

Variance Decomposition and Cryptocurrency Return Prediction^{*}

Suzanne S. Lee[†] Minho Wang[‡]

January 2024

Abstract

This paper examines how realized variances predict cryptocurrency returns in the cross-section using intraday data. We find that cryptocurrencies with higher variances exhibit lower returns in subsequent weeks. Decomposing total variances into signed jump and jump-robust variances reveals that the negative predictability is attributable to positive jump and jump-robust variances. The negative pricing effect is more pronounced for smaller cryptocurrencies with lower prices, less liquidity, more retail trading activities, and more positive sentiment. Our results suggest that cryptocurrency markets are unique because retail investors and preferences for lottery-like payoffs play important roles in the partial variance effects.

JEL classification: G02, G10, G12

Key words: cryptocurrency, high-frequency data, partial variances, jumps, return prediction

^{*}The authors thank Dexin Zhou, Olivier Scaillet, Fabio Trojani, Narayan Jayaraman, Sudheer Chava, and seminar participants at the University of Geneva, the University of Georgia, Baruch College, Australian Finance and Banking Conference, Midwest Finance Association annual meeting, SKKU International Conference, and Joint Conference with the Allied Korea Finance Associations for their helpful discussions, comments, and encouragement. The authors particularly thank the editor and referee for their constructive suggestions and comments.

[†]Address: Scheller College of Business at the Georgia Institute of Technology, 800 W. Peachtree St. NW, Atlanta, GA 30308, e-mail: suzanne.lee@scheller.gatech.edu (Corresponding author)

[‡]Address: College of Business at Florida International University, 11200 SW 8th St, Miami, FL 33174, e-mail: minwang@fiu.edu

1. Introduction

Cryptocurrencies have recently emerged as a nonnegligible asset class. The aggregate market capitalization of cryptocurrencies exceeded two trillion U.S. dollars (USDs) in 2021.¹ This remarkable growth has been accompanied by unusually large price fluctuations with extreme returns.² Our analyses corroborate these findings, revealing that the annualized weekly volatility of cryptocurrencies frequently exceeds 100%. High volatility with extreme returns can arise because the equilibrium prices of cryptocurrencies reflect sunspots that drive high extrinsic volatility even when fundamentals are constant (Biais et al., 2023). This high volatility raises a fundamental question in finance: how do realized volatilities (or variances) play a role in cryptocurrency return prediction in the cross-section? Addressing this critical question is the primary objective of our study because it has substantial implications for risk assessments, investment strategies, and portfolio management in cryptocurrencies.

In light of the exceptionally wide ranges and heavy tails of cryptocurrency return distributions, we take a comprehensive approach by considering not only the traditional realized variance measure but also decomposed partial variances that separately account for normal and nonnormal returns, including signed jumps. This decomposition is motivated because cryptocurrency market participants may perceive and evaluate uncertainty differently when they face extreme positive or negative returns (Barberis and Xiong, 2009, 2012). This consideration is particularly relevant for cryptocurrency markets because of the lack of observable

¹Market capitalization data are from www.coinmarketcap.com. Participation in cryptocurrency markets has become widespread. Surveys indicate that 22% of institutional investors already own cryptocurrencies and that 11% of the American population holds Bitcoin (see <https://cointelegraph.com/news/hbus-survey-almost-12-of-us-cryptocurrency-holders-are-long-term-investors>).

²Liu and Tsyvinski (2021) and Scaillet et al. (2020) document high kurtosis (e.g., 15 to 100 for daily data) and frequent jumps. In fact, Bitcoin prices collapsed, with a nearly 50% drop on March 12, 2020. On April 30, 2020, Bitcoin prices increased by 18% within 24 hours.

fundamental information and the active participation of retail investors compared to established financial markets. Furthermore, the directions of jumps have differential impacts on risk and returns (Patton and Sheppard, 2015).³ By employing the decomposed variances, our analysis elucidates how realized partial variances affect cryptocurrency return prediction.

Notably, in cryptocurrency markets, variances are dramatically time-varying. Figure 1 illustrates this feature by depicting the 10th, 50th, and 90th percentiles for the weekly realized variances of individual cryptocurrencies and Bitcoin. For example, the weekly variances of Bitcoin range from 0.01 to 0.08 in 2017, with three discernible peaks in 2018. This evidence emphasizes the critical importance of frequently updating cryptocurrency return variances when identifying the effect of variances on future returns. Therefore, we use intra-day cryptocurrency data to estimate variances. Leveraging recent advances in the financial econometrics literature, we sum the squared high-frequency returns from the different segments of individual cryptocurrencies' return distributions. Our study extends the literature by addressing such distinctive characteristics of individual cryptocurrency markets and uncovering novel evidence regarding the return predictability of decomposed variances.⁴

We document that cryptocurrencies with high realized variances tend to provide substantially lower excess returns in subsequent weeks than those with low variances. In Figure 2, cryptocurrencies with the lowest (highest) variances have an average excess return of 0.1% (-3.6%) in subsequent weeks. The weekly return differential amounts to 3.7% (193% per

³Using parts of the return distribution for measuring risks is similar to using the lower partial moment measures of Price et al. (1982), although our decomposition is driven by extreme returns.

⁴Borri and Santucci de Magistris (2022) use high-frequency data on Bitcoin to find that jumps account for a large portion of the daily variation in Bitcoin returns and adopt a parametric estimation method to show that the conditional skewness and kurtosis of Bitcoin returns are priced using daily data. Our study differs in that we rely on high-frequency data of 100 cryptocurrencies to decompose individual cryptocurrencies' realized variances over time and discover the significant return predictability of decomposed variances.

annum), which is significant at the 1% level.⁵ This figure provides equally weighted (EW) average returns, but value-weighted (VW) averages yield consistent results.

To investigate the underlying drivers of this negative return predictability, we decompose the total variance into variances associated with positive jumps, negative jumps, and non-jump returns. By using detected jumps and separated variances, we discover that positive jump and jump-robust variances are significantly and negatively related to excess returns in subsequent weeks. The return prediction of (total) variances stems from return components that are not associated with negative jumps. These variance effects are robust to model specifications, the common risk factors of Liu et al. (2022), business cycles, and overall market conditions and are not totally attributable to the skewness effect.

The negative relationship between variances and future returns contradicts the traditional risk and return tradeoffs, which typically imply a positive relation in classical asset pricing theories with rational investors. Instead, our finding is in line with behavioral finance studies on speculative retail trading because more individual and retail investors participate in cryptocurrency markets than in other well-established financial markets (Kogan et al., 2023).⁶ The substantial participation of retail investors allows asset prices to deviate from fundamental values (De Long et al., 1990) and can increase volatility, as shown by Xiong and Yu (2011), Foucault et al. (2011), and Pedersen (2022). Retail investors prefer holding and

⁵This significant spread remains intact after controlling for the cryptocurrency pricing factors of Liu et al. (2022). Our variance estimation differs from that of Liu et al. (2022), who employ daily returns. Estimation using lower-frequency return data requires an assumption that volatility is stable over a longer estimation horizon. These authors indicate that return volatility is an insignificant pricing factor (we obtain the consistent results). Generally, the empirical relation between equity returns and volatility has not been strong. However, Bollerslev et al. (2020) use high-frequency data and identify significant relations.

⁶According to the financial statements of Coinbase, a cryptocurrency exchange platform company, 95% of the total transaction revenues come from retail investors' trading in 2020. Many news articles also support this idea. In addition, see, e.g., <https://www.wsj.com/articles/bitcoin-prices-pass-50-000-for-first-time-since-may-11629729934> and <https://www.bloomberg.com/news/articles/2021-12-21/crypto-funds-explode-in-2021-led-by-proshares-bitcoin-strategy-etf-bito>.

trading highly volatile securities and are willing to undertake risk that may yield low returns (Han and Kumar, 2013). In fact, the key drivers of high volatility and extreme positive returns are the important lottery features favored by investors.

Specifically, cryptocurrencies with high total and positive jump variances tend to have smaller sizes, lower prices, and wider bid-ask spreads than those with low variances. Such cryptocurrencies have significantly larger trading volumes than those with low variances, which suggests a significant disagreement about future prices of high-variance cryptocurrencies. The retail trading proportion (RTP) and positive investor sentiment tend to be greater for high-variance cryptocurrencies than for low-variance cryptocurrencies. Furthermore, our realized variance measures effectively forecast the lottery properties of individual cryptocurrencies, which makes our findings consistent with the cumulative prospect theory elaborated by Barberis and Huang (2008).⁷ Considering Xiong and Yu (2011), we test whether our results arise because of short-selling constraints. However, we do not find supporting evidence.

Our paper contributes to the literature on the relation between volatility and future returns. The literature indicates that the negative relationship between volatility and future returns can result from the preference for lottery features and that many investors are not fully diversified (e.g., Fama and MacBeth, 1973; Hou and Moskowitz, 2005; Ang et al., 2006; Fu, 2009; Huang et al., 2010; and Hou and Loh, 2016). Specifically, our paper demonstrates that the negative return predictability is associated with both jump-robust and positive jump variances but not with negative jump variances in cryptocurrency markets. Unlike our study, Kilic and Shaliastovich (2019) examine the role of implied and realized semivariances

⁷Investors who prefer lottery-like returns are willing to pay higher prices for assets with lottery features. Positive jumps represent lottery-like returns in that they are characterized by extremely large payoffs and a low probability. High volatility allows investors to anticipate a high probability of large returns.

in aggregate stock and bond returns. In addition, this paper supports that our inference methods can identify the nonlinear effect of unusually high uncertainty on returns.

Our paper extends the growing literature that examines cryptocurrencies as an asset class.⁸ Liu et al. (2022) show that only the standard deviation of price volume predicts future returns. Borri et al. (2022) find that volatility risk is positively priced, while Bianchi and Babiak (2021) show that their realized or idiosyncratic volatility generates a significantly negative return. More broadly, our study is related to studies that characterize return distributions and factor structures in cryptocurrency markets. Our findings of high volatility and large jumps echo those of Yermack (2015) and Scaillet et al. (2020), who study Bitcoin returns. Jia et al. (2021) and Borri and Santucci de Magistris (2022) investigate the effects of higher moments on cryptocurrency returns. Cong et al. (2022) provide a five-factor model to consider additional value and network adoption premiums.⁹ Sockin and Xiong (2023) present a model that supports the empirical results in the literature.

Our study builds on the literature on realized return moments and jump risk measures. Andersen et al. (2001a) support the approach of estimating realized variances with intraday data. Amaya et al. (2015) investigate the relation between firm-level realized return moments and subsequent returns. Bollerslev et al. (2020) examine the stock return predictability of realized jump variance components. Lee and Wang (2019) demonstrate the pricing of negative jumps in sovereign currency markets.¹⁰ Our study is the first to utilize cryptocurrency

⁸Our paper is broadly related to the literature on cryptocurrency markets, the economics of cryptocurrencies and blockchain technology, or valuation models for digital currencies. However, we focus on discovering novel empirical evidence on risk and returns and discussing relevant pricing models.

⁹As additional references, Liu and Tsyvinski (2021) analyze the time-series features of return distributions by using daily data on three cryptocurrencies. Shams (2020) studies return correlations. Borri (2019) uses *CoVar* to measure the conditional tail risk of four cryptocurrencies by employing daily data. Another strand of studies provides evidence of manipulation or dispersion in cryptocurrency prices across exchanges (e.g., Griffin and Shams, 2020; Makarov and Schoar, 2020; Li et al., 2021; Borri and Shakhnov, 2022).

¹⁰Many studies document the important role of jumps in pricing equities, bonds, options, or sovereign

markets as a unique laboratory, documenting cryptocurrencies’ special characteristics.

Finally, our work contributes to the behavioral finance literature on speculative trading in financial markets. For cryptocurrency markets, characterized by the active participation of retail investors, we show that investor sentiment is important and that investors’ lottery (or gambling) preference exists in these highly uncertain markets. For other financial markets, Baker and Wurgler (2006) show that highly volatile stocks are prone to fluctuations in sentiment. Han and Kumar (2013) show low returns for volatile assets with lottery features such as low prices, high variances, and positive skewness. Boyer et al. (2010) and Bali et al. (2011) use expected idiosyncratic skewness and maximum returns and find that these features are associated with low expected returns.

The remainder of this paper is organized as follows: Section 2 explains the variance decomposition with signed jumps and the estimation approaches. Section 3 introduces the high-frequency data used for this study and the estimation results. Section 4 investigates how the decomposed variances predict future cryptocurrency returns. Section 5 discusses potential explanations for our findings. Section 6 concludes the paper.

2. Inference methods

In this section, we describe our model of cryptocurrency price processes and explain our inference methods for the total and decomposed variances of individual cryptocurrencies.

As cryptocurrency markets operate 24 hours a day in real time, we assume that cryptocurrency prices follow a continuous-time model. We employ a general asset pricing framework

currencies (e.g., Merton, 1976; Piazzesi, 2005; and Chernov et al., 2018).

with diffusion and jump components to accommodate various forms of nonnormality, such as unusual volatility and heavy tails, often observed in cryptocurrency return data. Specifically, the i -th cryptocurrency price is set to follow a generic jump diffusion model that accommodates the potential intraday volatility and jump patterns:

$$dc_{i,t} = \mu_{i,t}dt + \sigma_{i,t}dB_{i,t} + Y_{i,t}dJ_{i,t}, \quad (1)$$

where $dc_{i,t}$ is the instantaneous change in the natural logarithmic price $c_{i,t(j)}$ of the i -th cryptocurrency at time t . The drift $\mu_{i,t}$ and diffusion $\sigma_{i,t}$ are bounded processes, and $B_{i,t}$ is a standard Brownian motion. $Y_{i,t}$ and $dJ_{i,t}$ are the jump size and arrival indicator, respectively.¹¹ We denote the intraday logarithmic return between discrete times $t(j-1)$ and $t(j)$ by $r_{i,t(j)} = c_{i,t(j)} - c_{i,t(j-1)}$ for cryptocurrency i .

We relate realized total variances and decomposed variances to returns in subsequent weeks. Similar analyses can be performed at different frequencies and return horizons. The total variance is defined as the sum of squared intraday returns in week w ,

$$\text{Total variance: } TV_{i,w} = \sum_{t(j) \in W_w} r_{i,t(j)}^2. \quad (2)$$

Andersen et al. (2001b) indicate that as the sampling frequency goes to infinity, this realized total variance converges to the quadratic variation composed of the integrated diffusive

¹¹Our statistical inference is based on discrete samples over a time horizon of $[0, T]$. We assume that there are n discrete observations for each cryptocurrency over the time horizon. In particular, we observe the i -th cryptocurrency price $c_{i,t}$ only at discrete times $0 \leq t(0) < t(1) < \dots < t(n) \leq T$, and for simplicity, we assume that $t(j+1) - t(j) = \Delta t$ for all j s. The total number of weeks within $[0, T]$ is set to be $\tilde{w}$, so that $[0, T] = \bigcup_{w=1}^{\tilde{w}} W_w$ with the weekly time interval W_w for week w .

variance and jump components as follows:

$$TV_{i,w} = \sum_{t(j) \in W_w} r_{i,t(j)}^2 \longrightarrow \int_{s \in W_w} \sigma_{i,s}^2 ds + \sum_{\tau \in W_w} Y_{i,\tau}^2. \quad (3)$$

This total variance measure does not differentiate signed jump variances from diffusive variances. In this paper, we recommend decomposing realized total variances into jump robust and signed jump variances by using jump detection tests that identify the arrival times of individual cryptocurrency jumps. In particular, we estimate the diffusive variance term $\int_{s \in W_w} \sigma_{i,s}^2 ds$ by using the jump-robust variance estimator, while the jump variance term $\sum_{\tau \in W_w} Y_{i,\tau}^2$ (which can be decomposed into $\sum_{\tau \in W_w, Y_{i,\tau} > 0} Y_{i,\tau}^2$ and $\sum_{\tau \in W_w, Y_{i,\tau} < 0} Y_{i,\tau}^2$) by using our signed jump variance estimators $JV_{i,w}^{(+)}$ and $JV_{i,w}^{(-)}$ as follows:

$$\text{Jump robust variance: } JRV_{i,w} = \sum_{t(j) \in W_w} r_{i,t(j)}^2 I(|T_{i,t(j)}| < \zeta),$$

$$\text{Positive jump variance: } JV_{i,w}^{(+)} = \sum_{t(j) \in W_w} r_{i,t(j)}^2 I(|T_{i,t(j)}| > \zeta) \times I(r_{i,t(j)} > 0),$$

$$\text{Negative jump variance: } JV_{i,w}^{(-)} = \sum_{t(j) \in W_w} r_{i,t(j)}^2 I(|T_{i,t(j)}| > \zeta) \times I(r_{i,t(j)} < 0),$$

$$\text{Jump variance: } JV_{i,w} = JV_{i,w}^{(+)} + JV_{i,w}^{(-)} = \sum_{t(j) \in W_w} r_{i,t(j)}^2 I(|T_{i,t(j)}| > \zeta), \quad (4)$$

where $I(a)$ is an indicator function that equals one if a is true. ζ is the rejection criterion for the Lee and Mykland (2008) jump test statistic $T_{i,t(j)}$ for cryptocurrency i at time $t(j)$.¹² We

¹²We use the standard Gumbel distribution for the rejection criteria, following Lee and Mykland (2008). As cryptocurrency markets may exhibit time-of-day patterns in volatility, following Lee and Wang (2020), we control for the intraday volatility pattern to mitigate concerns regarding the misclassification of jumps.

use all intraday return data within a time horizon to approximate the true latent diffusive and jump variances in that period.¹³ Our jump variance measures are separated according to the signs of jumps to identify signed jumps’ exclusive impact on future returns.¹⁴

We apply the Lee and Mykland (2008) method or its variant because it is important for us to distinguish individual jumps with different signs within a testing interval. The basic intuition behind this method is to discriminate between diffusive and jump returns by comparing instantaneous returns with the local volatility estimated over the preestimation window of size K . If absolute returns are significantly larger than the estimated local volatility, they are identified as jumps.¹⁵ To support our application for the purpose of this study, we assess the finite sample performance of our jump variance estimators with simulation studies.¹⁶ We find that the estimation error of the proposed jump variance estimators decreases as we use higher-frequency data. Moreover, jumps in volatility do not significantly affect the power to detect extremely large jumps, which is our main interest for this study.

3. Data

We obtain intraday data from Kaiko, which has collected high-quality tick-by-tick quotes and prices from liquid cryptocurrency exchanges since 2014.¹⁷ Many studies (e.g., Makarov

¹³Because our main analyses are performed at the weekly level, our notation is written with a weekly interval. However, the notation can be generalized to other fixed time intervals.

¹⁴In this study, we do not separate systematic jumps from idiosyncratic jumps.

¹⁵Most other jump tests depend on the integrated quantities over an interval during which the jump presence can be recognized but do not indicate the direction, arrival time, or size of each jump or the number of jumps within the interval. The design of the tests proposed by Bollerslev et al. (2007) is similar to ours except for the rejection criteria. Therefore, the results of the two tests are not expected to differ if the rejection regions are chosen similarly. See Barndorff-Nielsen and Shephard (2006), Bollerslev et al. (2007), Jiang and Oomen (2008), and Aït-Sahalia and Jacod (2009) for alternative approaches.

¹⁶The details of the simulation studies are reported in Online Appendix A. A theoretical justification for our variance decomposition based on asymptotic properties is available upon request.

¹⁷Trading volumes are also provided by Kaiko. See Makarov and Schoar (2020) for the details of Kaiko.

and Schoar, 2020; Li et al., 2021) also use Kaiko for cryptocurrency prices.

To construct an unbiased sample with the largest cross-section, we examine all cryptocurrencies that have intraday data longer than nine months and are traded on Coinbase, which is ranked as the top exchange in Kaiko’s overall evaluation (e.g., data quality and popularity).¹⁸ The minimum sample period of nine months is chosen because our inference requires a sufficient estimation horizon for detecting jumps and computing decomposed realized variances. Kaiko’s order book data provide intraday bid and ask quotes (and volumes) for 198 cryptocurrencies. We exclude stable coins (e.g., Tether). Adopting simple coin selection criteria, our sample comprises 100 cryptocurrencies with various characteristics and includes a delisted coin.¹⁹ Accordingly, survivorship bias is not critical. We also confirm the minimal effect of delisting on our results by following Liu et al. (2022). Our results are confirmed with data from other cryptocurrency exchanges, such as Bitfinex and Bittrex.

The sample period is from October 2015 to June 2023 and includes the failure of large financial institutions such as the FTX or Terra-Luna crash and the crypto winter.²⁰ From these data, the first return appears in October 2015 for BTC, which is the starting time of our sample period.²¹ These cryptocurrencies account for approximately 80% of the total market capitalization of all cryptocurrencies as of 2023.²²

¹⁸As also indicated by Makarov and Schoar (2020), many nonintegrated cryptocurrency exchanges exist in parallel across countries. The majority of these exchanges function like regulated equity markets, but they lack provisions to ensure the best price for trading. Because this unusual feature increases price deviations across cryptocurrency exchanges, the use of cryptocurrency price data from multiple exchanges can result in contamination with frictions from different exchanges and thus is undesirable for our study.

¹⁹The list of cryptocurrencies is provided in Online Appendix B.

²⁰The crypto winter refers to a bear market with significant declines in prices and market capitalization. Our results continue to hold with data from shorter sample periods and different exchanges.

²¹The order book data start in April 2015, but observations in the early period are extremely sparse.

²²We provide details about our sample construction and filtering procedure in Online Appendix C.

3.1. Intraday cryptocurrency returns and jumps

We choose the 15-minute interval to compute intraday returns because of the tradeoff between the following considerations. First, the accuracy of the price data can suffer from measurement errors that result from microstructure noise in data sampled too frequently. Second, a lower sampling frequency can hinder the consistent estimation of realized moments.²³ For each interval, we select the latest observation to construct evenly spaced data. We remove quotes that do not change for three consecutive intervals because these quotes might be inactive. We perform this filtering process for bid and ask quotes and construct mid quotes (i.e., $\text{mid} = 0.5 \times (\text{bid} + \text{ask})$).²⁴ Using the mid quotes, we compute log returns.

Table 1 summarizes the 15-minute returns of the 25 selected sample cryptocurrencies.²⁵ Bitcoin has the largest number of observations and the longest sample period. Notably, cryptocurrency markets are extremely volatile. Standard deviations range from 0.27% to 1.25%, which implies annualized standard deviations greater than 100%. Interestingly, the skewness ranges from -1.34 to 1.71. The dispersed skewness implies that some cryptocurrencies provide lottery-like returns, while others have crash-like returns. The kurtosis is higher than nine, indicating that it is important to analyze the tails of the return distributions.

We implement the variance decomposition and estimate weekly realized measures using detected jumps.²⁶ Table 2 summarizes the jump detection results for the 25 selected sample

²³Like sovereign currency markets, cryptocurrency markets allow 24-hour trading and have real-time trading features. Therefore, we follow Lee and Wang (2019, 2020) in terms of the sampling frequency.

²⁴We follow the realized variance literature by using mid quotes as the measures of true prices. Mid quotes are generally less noisy than transaction prices because they do not suffer from bid-ask bounce effects. See Bandi and Russell (2006) for details. Following Andersen et al. (2001a) and Lee (2012), who suggest the problems of bid-ask bounce effects, we filter out observations when large returns are canceled out. We confirm the reliability of our data by comparing our mid-quotes with daily transaction data on coinmarketcap.com.

²⁵The summary statistics for the 100 cryptocurrencies are provided in Online Appendix B.

²⁶The Lee and Mykland jump test requires setting a window size K to estimate the instantaneous volatility using the first $K - 1$. The window size K must be large enough to ensure that the jump effect disappears for a

cryptocurrencies.²⁷ The average jump frequency is 0.55%, which intuitively implies that approximately one jump occurs every two days. The jump frequency in cryptocurrency markets is similar to that in typical sovereign currency markets. Interestingly, the sizes of cryptocurrency jumps are relatively large. The average of the medians of positive (negative) jump sizes is 3.2% (-2.9%), which is more than 10 times larger than that in sovereign currency markets.²⁸ This comparison highlights the importance of jumps in cryptocurrency returns. This large jump size is consistent with the findings of Liu and Tsyvinski (2021), who show that cryptocurrencies tend to have extreme returns in their distributions.

The frequencies and sizes of jumps are considerably dispersed across individual cryptocurrencies. The jump frequencies range from 0.22% (SUSHI) to 2.06% (BTRST) of the available return observations, and the median positive jump sizes range from 0.1% (RAI) to 4.9% (RGT). The ranges of the jump sizes are wider than those of sovereign currency markets. These findings indicate that cross-sectional differences in jump measures could play a role in cryptocurrency pricing. Overall, the numbers of positive and negative jumps are similar, and the distributions of positive and negative jump sizes are not discernibly different. These symmetries are also observed in other financial markets.

3.2. Weekly realized variances

We use the intraday return and detected jump data summarized in the previous subsection to compute the weekly decomposed variances and return measures. To construct weekly spans,

consistent volatility estimation. In this study, we use $K = 156$ for our 15-minute frequency, as recommended by Lee and Mykland (2008). We identify jumps with a 5% significance level. Our results are robust when we use a 1% significance level and filter out small jumps by using a 10% false discovery rate.

²⁷The jump detection results for the 100 cryptocurrencies are provided in Online Appendix B.

²⁸Lee and Wang (2019) report that the median positive (negative) jump size ranges from $\pm 0.1\%$ to $\pm 0.3\%$, with an average of $\pm 0.24\%$.

we follow the approach of Liu et al. (2022). Our choice of weekly analyses is also consistent with that of Amaya et al. (2015) and Bollerslev et al. (2020), who aggregate intraday data to construct realized risk measures over longer horizons and examine how these measures are related to subsequent equity returns in the cross-section.

We describe our weekly realized variances, which are estimated as explained in Section 2. Panel A of Table 3 shows the summary statistics of the weekly realized variances. The mean weekly realized total variance is 0.146, which corresponds to an annualized standard deviation of 138%. To put this into perspective, this annualized realized volatility can be compared with that of other financial asset markets. For example, the annualized realized volatility in sovereign currency markets ranges from 5.3% to 19.2%, with a mean of 12.25% (Lee and Wang, 2019). In addition, regarding U.S. stock markets, Andersen et al. (2001a) indicate that the mean of the annualized standard deviation is 28%, and Amaya et al. (2015) show that the median of annualized realized volatility is approximately 45%. This comparison suggests unusually high volatility in cryptocurrency markets.

Because cryptocurrency returns have substantial jump components, the mean of the jump-robust variances drops to 0.127. The mean jump variances appear relatively small because jumps are rare events; thus, jump variances often take a value of zero. However, when jumps occur, the jump variances dominantly contribute to the total variances. Symmetric jumps allow positive and negative jump variances to have similar distributions.

Panel B of Table 3 shows the correlations of these weekly realized variances. The jump-robust variances are highly correlated with the total variances because jump returns are rare. The correlation coefficient of these two variances is 0.98, which implies that they carry similar information. The positive and negative jump variances are positively correlated, with

a correlation coefficient of 0.64, indicating that positive jumps are occasionally accompanied by negative jumps. Because jumps tend to be detected during volatile periods, total variances are positively correlated with jump variances (i.e., the correlation coefficients are greater than 0.33). However, the relatively low correlation coefficients imply that signed jump variances capture information that differs from that captured by total and jump robust variances.

To examine how our weekly realized variance measures change over time and to assess whether there are cross-sectional differences in these measures, we plot the 10th, 50th, and 90th percentiles of the weekly realized variances. We provide the time-series plots for the total, positive jump, negative jump, and jump-robust variances in Panels A, B, C, and D of Figure 3, respectively. The percentiles for realized decomposed variances are clearly time-varying within relatively short horizons compared to those observable in stock markets. The total variances peak in 2018 because of turbulence in cryptocurrency markets.²⁹ The positive and negative jump variances are relatively high in 2018 and after 2021, and multiple peaks with large cross-sectional variations occurred in 2016, as bullish cryptocurrency markets attracted major market players, setting the stage for their growth.³⁰ Bullish and bearish markets consecutively occurred in the 2020s, which widened the cross-sectional variations in jump variances. Because the total and jump-robust variances are estimated with almost the same observations except for rare jumps, they yield similar patterns (Panels A and D). Overall, all variance measures dynamically change over time with discernible cross-sectional dispersion. Therefore, it is important to frequently update these risk measures for cryptocurrency markets to capture the time-varying features and pricing effects.

²⁹After substantial attention was given to cryptocurrency markets in 2017, there were hacking events (Corbet et al., 2020), and many governments (e.g., Korea, Japan, and the U.S.) announced the strengthening of regulations on cryptocurrency markets.

³⁰For example, Standard Chartered initiated investments in cryptocurrency markets.

4. Return prediction with variances

In this section, we investigate cryptocurrency return predictability, showing significantly negative spreads in returns to variance-sorted portfolios. Sorting analyses enable us to assess return predictability with selected variance measures and show the economic magnitudes of predictable returns. Then, we further support our findings by using cross-sectional predictability regressions that simultaneously control for multiple variance measures.

4.1. Sorting analyses

In this subsection, we compare the returns of total variance-sorted portfolios. At the end of every week (week w), we sort individual cryptocurrencies on total variances and construct tercile portfolios. The total variance measures are estimated with intraday observations from week $w - 3$ to w . Then, we compute equal-weighted (EW) and value-weighted (VW) returns in the subsequent week for each portfolio. We consider a long-short portfolio that purchases cryptocurrencies in the top tercile and sells cryptocurrencies in the bottom tercile.³¹

In Panel A of Table 4, we report the weekly returns of the total variance-sorted portfolios. The rows labeled “Excess return” show the clear negative relation between the realized total variances and the average subsequent returns. The portfolios with the highest total variances (High portfolios) provide significantly lower excess returns in subsequent weeks than those with the lowest total variances (Low portfolios). Specifically, the differentials between the weekly excess returns of High and Low portfolios are -3.7% and -3.0% for EW and VW portfolios, respectively, which are statistically significant at the 1% level. Unlike excess

³¹Although we use long-short portfolios for comparison purposes, there is a potential concern that some cryptocurrencies are difficult to short sell. Therefore, following Liu et al. (2022), we assume that Bitcoin is shorted and confirm that our results are robust.

returns, the standard deviations of High and Low portfolios are similar.

To investigate whether these return differences result from exposure to systematic risk factors, we compute alphas by regressing weekly excess returns on the three factors of Liu et al. (2022). The differentials between the alphas of High and Low portfolios remain negative and significant for both the EW and VW portfolios. Therefore, the negative relation between the total variances and subsequent returns cannot be explained solely by systematic factors.

Our results are partly inconsistent with those of Liu et al. (2022) in that these authors show insignificantly negative return differentials in subsequent weeks for volatility-sorted portfolios. These different results arise because these authors use daily data for variance/volatility estimations, while we employ intraday data. In fact, using the daily data of our sample coins, we obtain consistent results. By using higher-frequency observations for variance estimations, we can measure variances more precisely and better capture variances' time-varying features because the variance estimates are more frequently updated with recent data over shorter horizons and are less vulnerable to the smoothing effect of older data. As noted in Section 3, variances in cryptocurrency markets are clearly time-varying, and their cross-sectional variations change dramatically over short horizons. We incorporate this fact with high-frequency data and clearly demonstrate the significant role of realized variances in cryptocurrency return prediction, which has not been studied in the literature.

To examine which types of variances contribute to the return prediction, we perform additional sorting analyses by using positive jump, negative jump, and jump-robust variances. This variance decomposition with signed jump and non-jump returns differentiates the impact of extreme variations from nonnormal return distributions.³² The sorting analysis

³²A recent study by Bollerslev et al. (2021) and the references therein propose a similar variance decom-

results using these decomposed variances are reported in Panels B, C, and D of Table 4.

In Panel B, High positive jump variance portfolios have 3.6% (2.3%) lower excess returns for EW (VW) portfolios than do Low positive jump variance portfolios. The alphas for High positive jump variance portfolios are also significantly lower than those for Low positive jump variance portfolios. These results indicate that positive jump variances significantly contribute to the negative return predictability of total variances.

Panel C shows the results for negative jump variance-sorted portfolios, which are similar to those for positive jump variance-sorted portfolios in Panel B. The similarity results from the positive correlation of positive and negative jump variances. Our regression analyses in the next subsection clarify the exclusive effects of these variables with simultaneous controls. Finally, Panel D presents the results for jump-robust variance-sorted portfolios. The overall results in Panel D are similar to those presented in Panel A because the total and jump robust variances are similar, except for occasional jump arrivals.

4.2. Cross-sectional regression analyses

The sorting analyses could ignore the potential confounding effects of various independent variables. In this subsection, we address this concern by conducting a series of standard Fama and MacBeth (FMB, 1973) cross-sectional regressions with individual cryptocurrencies. These analyses expand our findings and simultaneously control for multiple factors and cryptocurrency characteristics. Our choice of regression models is motivated by Amaya et al. (2015), who use FMB regressions to identify the relationship between weekly realized central moments and stock returns in subsequent weeks. Our regression models are also consistent with partial (co)variance measures, which enables the multiple decompositions of realized variances.

with those of Bali et al. (2011), who adopt FMB regressions to investigate the relationship between maximum daily returns and subsequent monthly returns.³³

We first test the effect of the total variance measure on subsequent returns. Specifically, for each week in our sample, we run the following cross-sectional regression:

$$rx_{i,w+1} = \gamma_{0,w} + \gamma_{1,w}TV_{i,w}^{(+)} + c'_w X_{i,w} + \epsilon_{i,w+1}, \quad (5)$$

where $rx_{i,w+1}$ is the excess return of cryptocurrency i over week $w + 1$. The first variable of interest $TV_{i,w}^{(+)}$ and the vector of control variables $X_{i,w}$ are measured at the end of week w . After estimating the slope coefficients for each week in the sample, we take the time-series averages of the coefficient estimates to check whether the independent variables can significantly predict excess returns in the subsequent week.

Table 5 shows the related estimates and corresponding t -statistics.³⁴ As shown in column (I), the coefficients of realized total variances are significantly negative. To compare our intraday data-based measure with the daily return-based volatility measure of Liu et al. (2022), we use the standard deviations of daily returns as independent variables in column (II). Consistent with the result of Liu et al. (2022), the volatilities estimated with daily data provide insignificantly negative coefficients. This comparison elucidates the importance of using intraday data for capturing volatility effects in cryptocurrency markets. Column (III) shows that the negative relationship between total variances and future returns is robust to cryptocurrency-specific control variables such as lagged returns and cryptocur-

³³Bali et al. (2011) employ daily data, while we use intraday data, similar to Amaya et al. (2015). The regression frequency of Bali et al. (2011) is monthly, while ours is weekly, as is that of Amaya et al. (2015).

³⁴I.e., we report $\hat{\gamma}_l = (1/\tilde{w}) \sum_{w=1}^{\tilde{w}} \gamma_{l,w}$ with $l = 0, 1$ or $l = 0, 1, 2, 3$, where $\tilde{w}$ is the total number of weeks. We use the Newey-West standard errors. Our results are robust to lag length selection.

rency sizes. These controls are employed because of the momentum and size factors of Liu et al. (2022).³⁵

Then, we use multiple independent variables to test the effect of the decomposed variances on subsequent returns. Specifically, we conduct the following cross-sectional regression:

$$rx_{i,w+1} = \gamma_{0,w} + \gamma_{1,w}JV_{i,w}^{(+)} + \gamma_{2,w}JV_{i,w}^{(-)} + \gamma_{3,w}JRV_{i,w} + c'_wX_{i,w} + \epsilon_{i,w+1}, \quad (6)$$

where $JV_{i,w}^{(+)}$, $JV_{i,w}^{(-)}$, and $JRV_{i,w}$ are the positive jump, negative jump, and jump robust variances, respectively, as defined in Section 2. These decomposed variances are estimated with one month of observations. We conduct weekly cross-sectional regressions and then report the time series means of the coefficient estimates and t -statistics.

From column (IV) of Table 5, we show the results using decomposed variances as independent variables to identify which variance component significantly contributes to the negative predictability of the realized total variance. To represent the role of jumps without considering signs, we decompose total variances into jump and jump-robust variances. Column (IV) shows that the coefficient of jump variances is -0.669 and that of jump-robust variances is -0.152. Both coefficients are significant at the 1% level. Column (V) indicates that the negative relation of jump variances to subsequent returns mainly results from positive jump variance effects. In particular, positive jump variances continue to have significantly negative coefficients, while negative jump variances have positive coefficients. The negative return prediction of positive jump and jump-robust variances is maintained after controlling for

³⁵Our results indicate that the lagged return effects differ across return horizons or sample periods. For example, columns (III) and (IV) of Table 6 show that one-month and one-quarter lagged returns are significant, while one-week lagged returns are insignificant in Table 7, which also shows that the lagged return effects are sensitive to sample periods.

lagged returns and market capitalization, as shown in column (VI). These results imply that cryptocurrencies with higher variances tend to provide lower returns in subsequent periods and that this predictive relationship is amplified if positive jump variances are high.

From these analyses, we propose two key takeaways. First, it is important to account for the unique characteristics of cryptocurrency markets with unusually high volatility. Our realized variance measures capture short-term volatility dynamics and accommodate frequent updates with high-frequency data, which allows for the identification of unique variance effects. Another important and interesting point is that the negative return prediction results from return distributions beyond the left tail. This relationship could contrast with the classical risk and return trade-off, which is typically represented by the positive relation between risk measures (e.g., variances) and returns. In the next section, we explore the possible connection of our findings with previous studies that present this negative relationship.

4.3. Robustness tests

In this subsection, we address the concern that the return predictability of positive jump and jump-robust variances is attributable to our selection of dependent variables or omitted variables in the regressions. We also investigate how our results are affected by aggregate cryptocurrency uncertainty, average liquidity in cryptocurrency markets, and business cycles. We prove that our results are robust to model specifications and market conditions.

First, we test whether our results are sensitive to the choice of subsequent return horizons for dependent variables. We consider the two alternatives of two-week and one-month excess returns as dependent variables by replacing $rx_{i,w+1}$ with $rx_{i,w+1:w+2}$ or $rx_{i,w+1:w+4}$ in Equa-

tion (6), where $rx_{i,w+1:w+k}$ represents the excess returns of cryptocurrency i from weeks $w+1$ to $w+k$. Columns (I) and (II) of Table 6 present the results of applying Equation (6) with the replaced dependent variables. We continue to find that positive jump and jump-robust variances yield negative coefficients, which are significant at the 1% level.³⁶

In the second set of robustness checks, we consider additional control variables of lagged returns with alternative horizons while keeping the same dependent variables of subsequent-week returns as in our main regression specification. This set of robustness tests is motivated by Jegadeesh (1990), Lehmann (1990), and Jegadeesh and Titman (1993), who document short-term return reversals at weekly and monthly horizons and return momentum at 6-month to 12-month horizons in stock markets. Given our relatively short sample period, we use lagged one-month and one-quarter returns, instead of lagged one-week returns, as control variables. We report the results in columns (III) and (IV) of Table 6. These alternative controls are intended to capture momentum or reversal effects that may exist in cryptocurrency markets. We conclude that our main findings remain consistent.

Because (jump) variances can be high when maximum returns are high, one might raise the question of whether our findings are related to maximum return effects. To distinguish our findings from the maximum return effects, we control for the maximum 15-minute returns in week w and maximum daily returns in the previous month. As columns (V) and (VI) of Table 6 indicate, neither the 15-minute nor one-day maximum return weakens our results. This result suggests that our decomposed variance measures capture the aspects of historical return variations that are different from one extreme realized return.³⁷

³⁶We perform additional tests with other horizons and find consistent results.

³⁷Univariate regression with one-day maximum returns yields a significantly negative coefficient, which is consistent with the results of Bali et al. (2011). By adding positive jump variances, we find that the effect of the maximum returns becomes weaker than that of the univariate regression result.

We also investigate whether our results are robust to changes in aggregate cryptocurrency market uncertainty, average liquidity in cryptocurrency markets, and business cycles. We compute weekly cryptocurrency market volatility by following the approach of Menkhoff et al. (2012) that is used for foreign currency markets. For average illiquidity, we adopt an approach similar to that of Chordia et al. (2001) by taking the average of the illiquidity measures of Amihud (2002) across cryptocurrencies. Using the medians of these measures, we construct high- and low-volatility or illiquidity subsamples. To consider business cycles, we separate our sample into two subsamples, the latter of which starts in August 2021 (i.e., the subsample for the former period includes recession periods). We choose this separation because the National Bureau of Economic Research (NBER) indicates that the peak during our sample period occurs in February 2020 and because the two subsamples can be of similar sizes. Then, we apply the FMB regressions of Equation (6) to each subsample.

Table 7 shows the results. In the columns denoted Market volatility, the variance effects of this paper are significant for both the high- and low-volatility subsamples. As the columns under Market illiquidity indicate, the return predictability of positive variances is robust, while that of jump-robust variances is significant only for the period with high levels of market illiquidity. This result implies that the effect of variances from extreme price movements (i.e., positive jumps) can be dominant in less liquid markets. The last two columns for business cycles indicate that our findings are robust regardless of the business cycle or subsample period. Overall, we confirm that our results are robust to sample selection and market conditions such as market volatility and illiquidity.

4.4. Comparison with the realized skewness effect

The variance effects in this paper might be linked to the skewness effect, which also implies a negative relationship between skewness and subsequent returns (Amaya et al., 2015). Borri and Santucci de Magistris (2022) and Jia et al. (2021) document the effects of coskewness and realized skewness on the cross-section of cryptocurrency returns, respectively. To clarify how the variance effect in this paper is related to the skewness effect, we compare these two effects. We first check with our data that the skewness effect exists in cryptocurrency markets. Then, we compare our decomposed variance effects with the skewness effect.³⁸

Following Amaya et al. (2015), we use lagged skewness as the main explanatory variable and volatility and kurtosis as control variables. We run the following FMB regression:

$$rx_{i,w+1} = \lambda_{0,w} + \lambda_{1,w}LSkew_{i,w} + \lambda_{2,w}Vol_{i,w} + \lambda_{3,w}Kurt_{i,w} + c'_w X_{i,w} + \varepsilon_{i,w+1}, \quad (7)$$

where $LSkew_{i,w}$, $Vol_{i,w}$, and $Kurt_{i,w}$ are the weekly realized skewness, volatility, and kurtosis of returns, respectively.³⁹

We report the results in Table 8. In column (I), we run a univariate regression that uses only realized skewness and show that cryptocurrencies with high realized skewness tend to have low excess returns in subsequent weeks, which is consistent with the findings of Amaya et al. (2015). We include volatility and kurtosis in column (II) and add lagged returns and market capitalization in column (III). We continue to find that high realized skewness significantly predicts low future returns. As column (IV) shows, this negative relationship

³⁸In Online Appendix D, we also use alternative jump variance measures instead of skewness.

³⁹We use jump-robust volatilities instead of jump-robust variances to be consistent with the FMB regression specification of Amaya et al. (2015). When we use variances or jump-robust variances instead of volatilities or jump-robust volatilities, we find essentially the same results.

is robust to the additional controls of jump-robust volatility and kurtosis. Therefore, the skewness effect appears to exist in cryptocurrency markets, as it does in U.S. equity markets.

We compare the return predictability of skewness with that of positive jump variances by conducting horse-race regressions. As columns (V) and (VI) of Table 8 show, the coefficients of realized skewness become insignificant after controlling for the decomposed partial variances. Jump-robust and positive jump variances continue to exhibit negative and statistically significant coefficients, which indicates that the realized variances, including positive jump variances, are more important return predictors in cryptocurrency markets than is realized skewness. Interestingly, in highly uncertain cryptocurrency markets, high realized variances are more important and preferred characteristics than high realized skewness as long as they are not associated with extreme negative returns.

5. Exploration of mechanism for return predictability

The negative relation between variances and future cryptocurrency returns contrasts with the positive risk-return tradeoffs that traditional finance theories suggest (e.g., Merton, 1987). However, considering other asset markets that share similar characteristics and investors with cryptocurrency markets, our findings can be explained by behavioral finance studies on speculative retail trading with lottery preferences, among others. In this section, we explore several economic mechanisms and discuss related studies in the literature to enhance our understanding of the potential drivers of our findings.

5.1. Retail investor trading

We consider speculative retail trading as one explanation because Kogan et al. (2023) show that retail investors actively participate in cryptocurrency markets. According to Han and Kumar (2013), retail investors are drawn toward stocks with speculative features (high volatility and skewness) and are willing to undertake risk that may yield lower returns. Barberis and Huang (2008) demonstrate investors' preference for lottery-type returns. Barberis and Xiong (2012) present a model in which investors' risk-seeking behavior allows them to prefer holding and trading highly volatile securities because of a greater chance of realizing large gains. Our finding for cryptocurrency markets is closely related to these studies.

For formal tests, we gauge retail investors' trading activities by modifying the RTP of Han and Kumar (2013) to accommodate our weekly studies using cryptocurrency market data. Specifically, we acknowledge a notably wide dispersion in the right-skewed distribution of trading volumes. Given the unique nature of cryptocurrency markets, we establish the 90th percentile of volumes, which is 22 million dollars, across all sample cryptocurrencies as the threshold for characterizing retail investors' trading activities.⁴⁰ We also use characteristic variables such as market capitalization, prices, and percentage bid-ask spreads (BASs) because Han and Kumar (2013) document that retail investors tend to favor stocks with low market capitalization and low prices, which tend to be less liquid. Daily dollar trading volumes are employed to check for trading activities associated with high variances.

We examine the cross-sectional differences in these characteristics of variance-sorted portfolios. This approach is similar to that used in asset pricing studies that explain the negative

⁴⁰We utilize dollar volumes at 5-minute intervals. Considering that approximately 90-95% of Coinbase's revenues originate from retail investors, we use the 90th percentile of 5-minute trading volumes. We perform additional tests using the 50th and 95th percentiles as alternative thresholds and confirm the robustness.

risk and return relationship in stock markets by considering firm or portfolio characteristics (Fu, 2009; Brandt et al., 2010) and price pressures resulting from illiquidity (Avramov et al., 2006). To be consistent with our sorting analyses in Section 4, at the end of week w , we sort cryptocurrencies on one of decomposed variances and construct tercile portfolios. For each portfolio, we compute the equally weighted averages of the characteristics in week w .⁴¹

Table 9 shows the time-series averages of the characteristic variables for the sorted portfolios. Using the total variance-sorted portfolios, Panel A indicates that retail trading activities are significantly greater for cryptocurrencies with high total variances than for those with low total variances. Cryptocurrencies with high total variances tend to have smaller market sizes, lower prices, and wider BASs than those with low total variances. These results are consistent with those of Han and Kumar (2013). Interestingly, the largest trading volumes of High total variance portfolios result mainly from High positive jump variance portfolios, which show the most discernible differences in trading volumes (Panel B). High positive jump variance portfolios also have higher RTPs than other portfolios, which implies that cryptocurrencies with high positive jump variances attract retail transactions.

The literature documents similar return predictability of extreme volatility in other asset markets in which retail investors actively participate. Xiong and Yu (2011) study asset price bubbles in China’s warrant markets with the limited presence of institutional investors, finding that warrant bubbles are accompanied by trading frenzy and large volatility and highlighting the role of short-selling constraints and heterogeneous beliefs in explaining the price bubbles (i.e., resale option theory). Our evidence of large trading volumes for High (positive jump) variance portfolios indicates that cryptocurrency investors indeed tend to disagree

⁴¹We report the contemporaneous characteristics for simplicity; the predictive analyses are consistent.

about the future prices of cryptocurrencies with high variances. However, our variance effects are unlikely to result from overpricing because of short-selling constraints. For our tests, we measure short-selling availability by using the trading volumes of associated futures contracts to initiate short-selling positions.⁴² As Table 9 shows, High variance portfolios tend to exhibit larger futures trading volumes than Low variance portfolios, which implies that cryptocurrencies with high variances tend to have fewer short-selling constraints.

5.2. Investor sentiment

Baker and Wurgler (2006) show that investor sentiment affects the cross-section of stock returns. These authors find that when sentiment is high, “riskier” stocks with high volatility tend to earn lower returns and confirm their prediction that highly volatile stocks are difficult to value and to arbitrage, making these stocks especially prone to fluctuations in sentiment. This effect is stronger for small, young, unprofitable, non-dividend-paying, extreme growth, and distressed stocks. As Sockin and Xiong (2023) discuss the importance of sentiment, in this subsection, we check whether our variance effects are related to sentiment.

Considering Vosoughi et al. (2018) and Duz Tan and Tas (2021), we use daily coin-level sentiment measures from Twitter. We supplement our analyses with Telegram sentiment measures.⁴³ We employ the numbers and percentages of buy opinions and examine the relationship between these sentiment and our realized variance measures in the cross-section. As Table 9 shows, more positive sentiment is shared among investors for cryptocurrencies with higher variances, particularly with positive jump variances. This finding implies greater

⁴²We collected daily futures trading volume and the volume of futures buyers (those taking the counterparty positions for sellers at prevailing market prices). These futures trading volume data are from binance.com.

⁴³We collect daily investor sentiment data from intotheblock.com via cryptocompare.com, which provides the number of buy, sell, and neutral opinions. We appreciate that the referee informs us of the data source.

enthusiasm for cryptocurrencies with high positive jump variances and is consistent with the active participation of retail investors in the previous subsection. In addition, our cryptocurrency market results support the model of Pedersen (2022), in which speculative investors learn about market sentiment through social network platforms, further increasing prices.

5.3. Cryptocurrencies as lotteries

In this subsection, we support that the variance effect is consistent with cumulative prospect theory based on investors' preferences for lottery-like returns (e.g., Barberis and Huang, 2008). The key insight of this theory is that investors may favor assets with ex ante return distributions with rare but extremely high returns and thus would be willing to pay higher prices for such assets, which results in low subsequent returns. Accordingly, investors with lottery preferences are attracted by and seek out cryptocurrencies with observable measures that can help them predict lottery-type returns in the future. We hypothesize that positive jump and jump-robust variances can predict future lottery returns because positive jumps share common features with lottery-type returns (i.e., unusually large payoffs with low probability) and because high variances increase the likelihood of realizing lottery-type returns as long as they are not associated with extremely negative jumps.

A widely accepted measure for lottery-like returns in the literature is skewness (e.g., Boyer et al., 2010; Barberis et al., 2016). Therefore, to support our explanation, we test whether our key variance measures (i.e., positive jump and jump-robust variances) can predict future skewness. If this condition holds, these variance measures can be negatively related to subsequent returns, as documented in this paper. We perform an empirical test similar to

that of Boyer et al. (2010). Specifically, we estimate the following regression model:

$$Skew_{i,w+1} = \psi_0 + \psi_1 LSkew_{i,w} + \psi_2 JV_{i,w}^{(+)} + \psi_3 JV_{i,w}^{(-)} + \psi_4 JRV_{i,w} + c' X_{i,w} + e_{i,w+1}, \quad (8)$$

where $Skew_{i,w}$ is the weekly realized skewness of cryptocurrency i in week w . $LSkew_{i,w}$ is the realized skewness estimated from the previous month of observations (i.e., observations from week $w - 3$ to week w). $X_{i,w}$ is the vector of control variables, such as weekly kurtosis, lagged excess returns, natural logarithmic market capitalization, and fixed effects.

Table 10 shows that lagged skewness does not provide significant coefficients in a robust manner. This evidence is consistent with that of Boyer et al. (2010), who also show that lagged realized skewness is not a strong predictor of future skewness for stocks. However, we find that both positive jump and jump-robust variances significantly predict skewness in the subsequent week. Column (III) shows that both positive jump and jump-robust variances have significantly positive coefficients, which indicates that one can expect to observe greater skewness for cryptocurrencies with higher positive jump and jump-robust variances realized in the current period. In the remaining columns, we continue to find robust results. This finding has important implications in practice because investors can expect future skewness by using our decomposed realized variances, which can be estimated with available data.

5.4. Nondiversifiable factors

Our variance measures are computed without a specific assumption of a factor model and thus include systematic and idiosyncratic components. In this subsection, we discuss whether our finding of the negative return prediction of variances is related to nondiversifiable com-

ponents. This discussion is related to that of Ang et al. (2009), who document that stocks with recent past high idiosyncratic volatility have low future average returns and that not easily diversifiable factors underlie their idiosyncratic volatility effects.

In Section 4, for our cryptocurrency return prediction, we examine the excess returns and alphas of variance-sorted portfolios, accounting for the three factors of Liu et al. (2022). In Table 4, both the returns and alphas of High total and positive jump variance portfolios are significantly lower than those of Low portfolios. This implies that the negative relation between realized variances and subsequent returns cannot be fully explained by these existing systematic factors. Given this context, we further investigate whether idiosyncratic components mainly contribute to the return predictability of total and positive jump variances.

To this end, we compute the correlations among the three-factor alphas of cryptocurrencies constituting each variance-sorted portfolio.⁴⁴ We find that High total and positive jump variance portfolios exhibit average correlations of -0.3% and 0.1%, respectively. Moreover, the correlations observed in High total and positive jump variance portfolios are significantly lower than those in Low and Mid portfolios.⁴⁵ The significantly low correlations of High total and positive jump variance cryptocurrencies suggest that the returns of cryptocurrencies in High portfolios tend to be more idiosyncratic than those in Low portfolios. Combined with the significant alpha differentials reported in our sorting analyses, these findings reveal that idiosyncratic components play important roles in our main finding of the total and positive jump variance effects in cryptocurrency markets. These additional findings distinguish the dynamics of cryptocurrency markets from those of equity markets

⁴⁴We report the results obtained using the three-factor model alphas and confirm the consistency of our conclusions using market model alphas.

⁴⁵The differentials are statistically significant at the 1% level. For total (positive jump) variance-sorted portfolios, the correlations of Low and Mid portfolios are 2% and 1% (2% and 0.4%), respectively.

6. Conclusion

The remarkable growth in cryptocurrency markets has been accompanied by unusually large price fluctuations, generating return distributions with exceptionally wide ranges and heavy tails. We study how realized variances associated with different parts of return distributions affect future cryptocurrency returns. Using high-frequency returns, we decompose total variances into jump-robust, positive jump, and negative jump variances.

Our cross-sectional analyses reveal that cryptocurrencies with high total variances tend to exhibit low excess returns in subsequent weeks. The weekly return spread between cryptocurrencies in the lowest and highest tercile portfolios is 3.7% (193% per annum). This negative return prediction is attributable to jump-robust and positive jump variances. Interestingly, this result is more pronounced for cryptocurrencies with smaller sizes, lower prices, less liquidity, and more retail trading activities and is affected by investor sentiment. Our results can be explained by the overpricing of such cryptocurrencies that results from the risk-taking behavior of retail investors seeking highly volatile assets with the potential for large gains. In addition, our findings support the impact of retail investors' active participation on cryptocurrency markets.

Our paper contributes to the literature by adopting comprehensive intraday cryptocurrency data to precisely measure variances and identifying how decomposed partial variances can be used to distinguish their differential impact on future returns. We show that it is important to frequently capture the dynamic nature of cryptocurrency market volatility by using high-frequency data. Our study provides important implications for assessing risks in highly volatile asset markets.

References

- Amaya, D., Christoffersen, P., Jacobs, K., Vasquez, A., 2015. Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics* 118, 135–167.
- Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5, 31–56.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens, H., 2001a. The distribution of realized stock return volatility. *Journal of Financial Economics* 61, 43–76.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001b. The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* 96, 42–55.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. *Econometrica* 71, 579–625.
- Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61, 259–299.
- Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2009. High idiosyncratic volatility and low returns: International and further u.s. evidence. *Journal of Financial Economics* 91, 1–23.
- Avramov, D., Chordia, T., Goyal, A., 2006. Liquidity and autocorrelations in individual stock returns. *Journal of Finance* 61, 2365–2394.
- Aït-Sahalia, Y., Jacod, J., 2009. Testing for jumps in a discretely observed process. *Annals of Statistics* 37, 184–222.
- Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61, 1645–1680.
- Bali, T.G., Cakici, N., Whitelaw, R.F., 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics* 99, 427–446.
- Bandi, F.M., Russell, J.R., 2006. Separating microstructure noise from volatility. *Journal of Financial Economics* 79, 655–692.
- Barberis, N., Huang, M., 2008. Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review* 98, 2066–2100.
- Barberis, N., Mukherjee, A., Wang, B., 2016. Prospect theory and stock returns: An empirical test. *Review of Financial Studies* 29, 3068–3107.
- Barberis, N., Xiong, W., 2009. What drives the disposition effect? an analysis of a long-standing preference-based explanation. *Journal of Finance* 64, 751–784.
- Barberis, N., Xiong, W., 2012. Realization utility. *Journal of Financial Economics* 104, 251–271.

- Barndorff-Nielsen, O., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4, 1–30.
- Biais, B., Bisiere, C., Bouvard, M., Casamatta, C., Menkveld, A.J., 2023. Equilibrium bitcoin pricing. *Journal of Finance* 78, 967–1014.
- Bianchi, D., Babiak, M., 2021. A factor model for cryptocurrency returns. CERGE-EI Working Paper.
- Bollerslev, T., Andersen, T., Dobrev, D., 2007. No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. *Journal of Econometrics* 138, 125–180.
- Bollerslev, T., Li, S.Z., Zhao, B., 2020. Good volatility, bad volatility, and the cross section of stock returns. *Journal of Financial and Quantitative Analysis* 55, 751–781.
- Bollerslev, T., Medeiros, M.C., Patton, A.J., Quaadvlieg, R., 2021. From zero to hero: Realized partial (co) variances. *Journal of Econometrics* 231, 348–360.
- Borri, N., 2019. Conditional tail-risk in cryptocurrency markets. *Journal of Empirical Finance* 50, 1–19.
- Borri, N., Santucci de Magistris, P., 2022. Crypto premium, higher-order moments and tail risk. Available at SSRN.
- Borri, N., Massacci, D., Rubin, M., Ruzzi, D., 2022. Crypto risk premia. LUISS University Working Paper.
- Borri, N., Shakhnov, K., 2022. The cross-section of cryptocurrency returns. *Review of Asset Pricing Studies* 12, 667–705.
- Boyer, B., Mitton, T., Vorkink, K., 2010. Expected idiosyncratic skewness. *Review of Financial Studies* 23, 169–202.
- Brandt, M.W., Brav, A., Graham, J.R., Kumar, A., 2010. The idiosyncratic volatility puzzle: Time trend or speculative episodes? *Review of Financial Studies* 23, 863–899.
- Chernov, M., Graveline, J., Zviadadze, I., 2018. Crash risk in currency returns. *Journal of Financial and Quantitative Analysis* 53, 137–170.
- Chordia, T., Roll, R., Subrahmanyam, A., 2001. Market liquidity and trading activity. *Journal of Finance* 56, 501–530.
- Cong, L.W., Karolyi, G.A., Tang, K., Zhao, W., 2022. Value premium, network adoption, and factor pricing of crypto assets. Cornell University Working Paper.
- Corbet, S., Cumming, D.J., Lucey, B.M., Peat, M., Vigne, S.A., 2020. The destabilising effects of cryptocurrency cybercriminality. *Economics Letters* 191, 108741.

- De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990. Noise trader risk in financial markets. *Journal of Political Economy* 98, 703–738.
- Duz Tan, S., Tas, O., 2021. Social media sentiment in international stock returns and trading activity. *Journal of Behavioral Finance* 22, 221–234.
- Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81, 607–636.
- Foucault, T., Sraer, D., Thesmar, D.J., 2011. Individual investors and volatility. *Journal of Finance* 66, 1369–1406.
- Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics* 91, 24–37.
- Griffin, J.M., Shams, A., 2020. Is bitcoin really untethered? *Journal of Finance* 75, 1913–1964.
- Han, B., Kumar, A., 2013. Speculative retail trading and asset prices. *Journal of Financial and Quantitative Analysis* 48, 377–404.
- Hou, K., Loh, R.K., 2016. Have we solved the idiosyncratic volatility puzzle? *Journal of Financial Economics* 121, 167–194.
- Hou, K., Moskowitz, T.J., 2005. Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies* 18, 981–1020.
- Huang, W., Liu, Q., Rhee, S.G., Zhang, L., 2010. Return reversals, idiosyncratic risk, and expected returns. *Review of Financial Studies* 23, 147–168.
- Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. *Journal of Finance* 45, 881–898.
- Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of finance* 48, 65–91.
- Jia, Y., Liu, Y., Yan, S., 2021. Higher moments, extreme returns, and cross-section of cryptocurrency returns. *Finance Research Letters* 39, 101536.
- Jiang, G.J., Oomen, R.C., 2008. Testing for jumps when asset prices are observed with noise—a “swap variance” approach. *Journal of Econometrics* 144, 352–370.
- Kilic, M., Shaliastovich, I., 2019. Good and bad variance premia and expected returns. *Management Science* 65, 2522–2544.
- Kogan, S., Makarov, I., Niessner, M., Schoar, A., 2023. Are crypto different? evidence from retail trading. NBER Working Paper.
- Lee, S.S., 2012. Jumps and information flow in financial markets. *Review of Financial Studies* 25, 439–479.

- Lee, S.S., Mykland, P.A., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial Studies* 21, 2535–2563.
- Lee, S.S., Wang, M., 2019. The impact of jumps on carry trade returns. *Journal of Financial Economics* 131, 433–455.
- Lee, S.S., Wang, M., 2020. Tales of tails: Jumps in currency markets. *Journal of Financial Markets* 48, 100497.
- Lehmann, B.N., 1990. Fads, martingales, and market efficiency. *Quarterly Journal of Economics* 105, 1–28.
- Li, H., Wells, M.T., Yu, C.L., 2008. A bayesian analysis of return dynamics with lévy jumps. *Review of Financial Studies* 21, 2345–2378.
- Li, T., Shin, D., Wang, B., 2021. Cryptocurrency pump-and-dump schemes. University of Florida, Working paper.
- Liu, Y., Tsyvinski, A., 2021. Risks and returns of cryptocurrency. *Review of Financial Studies* 34, 2689–2727.
- Liu, Y., Tsyvinski, A., Wu, X., 2022. Common risk factors in cryptocurrency. *Journal of Finance* 77, 1133–1177.
- Makarov, I., Schoar, A., 2020. Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics* 135, 293–319.
- Menkhoff, L., Sarno, L., Schmeling, M., Schrimpf, A., 2012. Carry trades and global foreign exchange volatility. *Journal of Finance* 67, 681–718.
- Merton, R.C., 1976. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3, 125–144.
- Merton, R.C., 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42.
- Patton, A.J., Sheppard, K., 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics* 97, 683–697.
- Pedersen, L.H., 2022. Game on: Social networks and markets. *Journal of Financial Economics* 146, 1097–1119.
- Piazzesi, M., 2005. Bond yields and the federal reserve. *Journal of Political Economy* 113, 311–344.
- Price, K., Price, B., Nantell, T.J., 1982. Variance and lower partial moment measures of systematic risk: Some analytical and empirical results. *Journal of Finance* 37, 843–855.
- Scaillet, O., Treccani, A., Trevisan, C., 2020. High-frequency jump analysis of the bitcoin market. *Journal of Financial Econometrics* 18, 209–232.

- Shams, A., 2020. The structure of cryptocurrency returns. Charles A. Dice Center Working Paper.
- Sockin, M., Xiong, W., 2023. A model of cryptocurrencies. *Management Science* 69, 6684–6707.
- Vosoughi, S., Roy, D., Aral, S., 2018. The spread of true and false news online. *Science* 359, 1146–1151.
- Xiong, W., Yu, J., 2011. The chinese warrants bubble. *American Economic Review* 101, 2723–2753.
- Yermack, D., 2015. Is bitcoin a real currency? An economic appraisal. *Handbook of Digital Currency*, 31–43.

Table 1

Summary statistics.

This table shows the summary statistics of cryptocurrency data. Column Cryptocurrency lists the abbreviated names of the cryptocurrencies in the sample. Column # of obs shows the number of 15-minute return observations after data filtering. Column Start indicates the months in which the earliest observations appear in the sample. Columns under 15-min return provide the mean, standard deviation, skewness, and kurtosis of the 15-minute returns for the 25 selected cryptocurrencies. Columns Market cap., Volume, Price, and BAS show the unconditional averages of market capitalization, daily trading volumes in billion U.S. dollars (USD), prices per coin in USD, and percentage bid-ask spreads (BASs), respectively. Row Avg.100 shows the averages across the 100 sample cryptocurrencies.

Crypto currency	# obs	Start	15-min return				Market cap (\$B)	Volume (\$B)	Price (\$)	BAS (%)
			Mean (%)	Std (%)	Skew	Kurt				
BTC	263,892	Oct. 2015	0.0023	0.3983	-0.188	19.076	270.99	17.82	15,517.00	0.011
ETH	229,770	Jul. 2016	0.0022	0.5168	-0.068	15.162	100.31	8.96	1,003.27	0.030
LTC	216,697	Oct. 2016	0.0012	0.5936	0.016	15.027	4.67	1.55	91.74	0.067
BCH	190,108	Dec. 2017	-0.0003	0.5673	0.053	15.242	8.00	1.92	416.19	0.043
ETC	161,484	Aug. 2018	-0.0001	0.6249	0.069	16.790	2.10	0.80	19.88	0.092
ZRX	160,121	Oct. 2018	-0.0010	0.6995	0.071	24.108	0.39	0.06	0.50	0.136
XRP	64,229	Feb. 2019	0.0005	0.5197	-0.109	20.925	15.87	1.79	0.28	0.064
XLM	149,494	Mar. 2019	-0.0002	0.5553	-0.001	16.177	2.75	0.28	0.16	0.072
EOS	135,076	Apr. 2019	0.0014	0.5601	-0.095	14.771	3.33	1.43	3.09	0.138
REP	119,864	Apr. 2019	-0.0064	0.7078	0.188	20.966	0.20	0.02	16.54	0.292
LINK	139,040	Jun. 2019	0.0010	0.6488	0.015	11.856	3.74	0.68	12.20	0.079
XTZ	129,846	Aug. 2019	0.0003	0.6640	0.023	11.500	1.80	0.14	2.60	0.146
ALGO	132,270	Aug. 2019	-0.0007	0.7436	0.002	23.743	2.54	0.20	0.60	0.090
DASH	131,336	Sep. 2019	-0.0010	0.5870	-0.016	13.990	1.14	0.26	102.48	0.120
OXT	111,485	Dec. 2019	-0.0065	0.8280	0.489	17.496	0.12	0.03	0.26	0.146
ATOM	119,789	Jan. 2020	0.0004	0.6997	0.022	11.126	2.97	0.40	13.87	0.083
KNC	116,027	Feb. 2020	-0.0003	0.7623	0.103	13.128	0.17	0.03	1.46	0.115
OMG	94,855	May. 2020	0.0002	0.7718	0.182	11.982	0.56	0.19	4.30	0.142
MKR	106,360	Jun. 2020	-0.0007	0.6037	0.048	12.437	1.06	0.08	1,528.35	0.071
COMP	104,874	Jun. 2020	-0.0012	0.6993	-0.107	11.129	0.98	0.13	187.46	0.076
LRC	96,365	Sep. 2020	-0.0024	0.8149	0.276	14.184	0.40	0.08	0.56	0.087
ZEC	88,913	Dec. 2020	0.0000	0.6522	0.002	12.724	0.84	0.26	102.28	0.076
ADA	78,975	Mar. 2021	-0.0014	0.5473	0.036	12.444	15.03	1.05	0.95	0.031
DOGE	71,844	Jun. 2021	-0.0041	0.5537	0.071	15.383	5.67	0.67	0.13	0.048
ZEN	59,632	Sep. 2021	-0.0049	0.5999	-0.160	12.542	0.23	0.02	30.28	0.147
Avg.100	81,626		-0.0052	0.7370	0.187	16.443	5.44	0.49	773.57	0.192

Table 2

Summary statistics of cryptocurrency jumps.

This table summarizes the results of jump detection tests in cryptocurrency markets. To identify jumps, we apply the approach of Lee and Mykland (2008) and adjust the intraday volatility patterns of individual cryptocurrencies, following Lee and Wang (2020). Column Cryptocurrency lists the currency codes of the 25 sample cryptocurrencies. For each cryptocurrency, we provide the number of filtered jumps and the jump frequencies relative to the available observations (% jp). For signed jumps, we report the 25th, 50th, and 75th percentiles of positive and negative jump sizes in the last six columns. The numbers in columns % jp, positive jump size, and negative jump size are in percentage terms. Row Avg.100 shows the averages across the 100 sample cryptocurrencies.

Crypto currency	Jump frequency			% jp	Positive jump size (%)			Negative jump size (%)		
	Total	Positive	Negative		25p	50p	75p	25p	50p	75p
BTC	2,180	1,037	1,143	0.826	0.0083	0.0129	0.0192	-0.0212	-0.0141	-0.0094
ETH	1,318	639	679	0.574	0.0125	0.0185	0.0262	-0.0263	-0.0197	-0.0138
LTC	1,091	526	565	0.503	0.0171	0.0235	0.0325	-0.0327	-0.0235	-0.0176
BCH	1,025	517	508	0.539	0.0170	0.0229	0.0320	-0.0306	-0.0221	-0.0164
ETC	968	466	502	0.599	0.0175	0.0252	0.0342	-0.0341	-0.0244	-0.0165
ZRX	840	364	476	0.525	0.0188	0.0278	0.0431	-0.0392	-0.0255	-0.0183
XRP	380	203	177	0.592	0.0133	0.0186	0.0261	-0.0295	-0.0208	-0.0130
XLM	594	248	346	0.397	0.0156	0.0207	0.0276	-0.0286	-0.0191	-0.0144
EOS	675	307	368	0.500	0.0175	0.0228	0.0312	-0.0314	-0.0231	-0.0162
REP	1,048	532	516	0.874	0.0224	0.0320	0.0441	-0.0399	-0.0283	-0.0204
LINK	366	151	215	0.263	0.0191	0.0258	0.0351	-0.0366	-0.0259	-0.0176
XTZ	444	197	247	0.342	0.0205	0.0283	0.0399	-0.0333	-0.0260	-0.0180
ALGO	451	183	268	0.341	0.0232	0.0362	0.0535	-0.0431	-0.0295	-0.0211
DASH	462	205	257	0.352	0.0180	0.0264	0.0365	-0.0338	-0.0249	-0.0177
OXT	843	493	350	0.756	0.0220	0.0331	0.0476	-0.0426	-0.0281	-0.0198
ATOM	311	128	183	0.260	0.0217	0.0285	0.0409	-0.0373	-0.0263	-0.0193
KNC	416	190	226	0.359	0.0228	0.0335	0.0444	-0.0384	-0.0266	-0.0188
OMG	314	154	160	0.331	0.0235	0.0322	0.0460	-0.0393	-0.0275	-0.0175
MKR	387	180	207	0.364	0.0169	0.0247	0.0332	-0.0311	-0.0239	-0.0172
COMP	273	95	178	0.260	0.0253	0.0362	0.0438	-0.0407	-0.0305	-0.0235
LRC	365	178	187	0.379	0.0254	0.0383	0.0545	-0.0422	-0.0333	-0.0228
ZEC	261	92	169	0.294	0.0195	0.0270	0.0412	-0.0344	-0.0251	-0.0180
ADA	291	131	160	0.368	0.0171	0.0228	0.0310	-0.0280	-0.0204	-0.0139
DOGE	425	193	232	0.592	0.0184	0.0241	0.0338	-0.0288	-0.0218	-0.0164
ZEN	221	85	136	0.371	0.0197	0.0264	0.0367	-0.0365	-0.0264	-0.0202
Avg.100	420	206	214	0.550	0.0228	0.0320	0.0446	-0.0395	-0.0285	-0.0205

Table 3

Summary statistics of weekly realized variances.

This table reports the summary statistics of the decomposed variances estimated with intraday data. The variance is defined as the sum of squared 15-minute returns and is estimated from the previous month of observations. The positive (negative) jump variance is defined as the sum of squared positive (negative) jump sizes and is estimated from the previous month of observations. The total jump variance is the sum of the positive and negative jump variances. The jump-robust variance is estimated from observations without a jump. Panel A reports the first four central moments (Mean, Stdev, Skew, and Kurt), 25th percentile (25p), median (Median), and 75th percentile (75p). Panel B shows the pairwise correlations between these measures.

Panel A. Distributional characteristics

	Variance Total	Positive	Jump variance Negative	Total	Jump-robust variance
Mean	0.146	0.010	0.008	0.018	0.127
Stdev	0.161	0.015	0.010	0.024	0.154
Skew	6.031	3.671	5.073	3.736	6.752
Kurt	91.738	25.338	67.674	24.953	110.573
25p	0.058	0.002	0.003	0.006	0.047
Median	0.103	0.005	0.005	0.010	0.087
75p	0.179	0.012	0.010	0.021	0.155

Panel B. Correlation matrix

	Variance Total	Positive	Jump variance Negative	Total
Positive jump variance	0.420			
Negative jump variance	0.339	0.640		
Total jump variance	0.425	0.934	0.873	
Jump-robust variance	0.981	0.248	0.175	0.239

Table 4

Excess returns of cryptocurrencies sorted by decomposed variances.

This table shows the relationship between the decomposed variances and subsequent excess returns. At the end of every week, we sort the 100 sample cryptocurrencies on the estimated (total) variance, positive and negative jump variances, or jump-robust variance and then construct tercile portfolios. We estimate the total and decomposed variance measures by using the previous month of observations. For each sorted portfolio, we compute excess returns in the subsequent week. The portfolios are constructed with equal weights (EW) and value (market capitalization) weights (VW). In Panels A, B, C, and D, we report the results for the portfolios sorted by the total, positive jump, negative jump, and jump-robust variances. In each panel, we report the average excess returns, standard deviation, and Sharpe ratio for the excess returns. In addition, we provide the alphas of the time-series regressions of the portfolio returns on the three-factor model of Liu, Tsyvinski and Wu (2022). Column Low (High) concerns portfolios with the lowest (highest) sorting measures. Column H - L shows the differences between the values of High and Low portfolios (the standard deviations and Sharpe ratios in this column are for longing High portfolios and shorting Low portfolios). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Total variance

	Low	Mid	High	H - L
Total variance	0.021	0.034	0.052	0.031 ***
Excess return (EW)	0.001	-0.007	-0.036	-0.037 ***
Standard deviation	0.108	0.130	0.128	0.063
Sharpe ratio	0.009	-0.052	-0.285	-0.589
Alpha	0.013	0.010	-0.012	-0.024 ***
Excess return (VW)	0.008	0.000	-0.022	-0.030 ***
Standard deviation	0.098	0.132	0.146	0.093
Sharpe ratio	0.084	-0.001	-0.147	-0.321
Alpha	0.021	0.003	-0.006	-0.027 ***

Panel B. Positive jump variance

	Low	Mid	High	H - L
Positive jump variance	0.001	0.002	0.004	0.003 ***
Excess return (EW)	-0.001	-0.004	-0.037	-0.036 ***
Standard deviation	0.119	0.117	0.127	0.054
Sharpe ratio	-0.007	-0.036	-0.291	-0.672
Alpha	0.010	0.011	-0.011	-0.021 ***
Excess return (VW)	0.007	0.007	-0.016	-0.023 ***
Standard deviation	0.103	0.107	0.137	0.078
Sharpe ratio	0.066	0.061	-0.117	-0.291
Alpha	0.017	0.016	-0.009	-0.026 ***

Table 4

Excess returns of cryptocurrencies sorted on decomposed variances (continued).

Panel C. Negative jump variance

	Low	Mid	High	H - L
Negative jump variance	0.001	0.002	0.003	0.002***
Excess return (EW)	-0.005	-0.004	-0.032	-0.027***
Standard deviation	0.118	0.124	0.121	0.051
Sharpe ratio	-0.044	-0.035	-0.267	-0.531
Alpha	0.006	0.013	-0.008	-0.014***
Excess return (VW)	0.011	0.000	-0.012	-0.023***
Standard deviation	0.103	0.116	0.130	0.079
Sharpe ratio	0.106	0.002	-0.093	-0.292
Alpha	0.017	0.016	-0.003	-0.020***

Panel D. Jump-robust variance

	Low	Mid	High	H - L
Jump-robust variance	0.019	0.030	0.046	0.027***
Excess return (EW)	-0.001	-0.008	-0.033	-0.032***
Standard deviation	0.110	0.127	0.129	0.063
Sharpe ratio	-0.010	-0.062	-0.258	-0.512
Alpha	0.011	0.007	-0.008	-0.020***
Excess return (VW)	0.008	-0.001	-0.017	-0.026***
Standard deviation	0.098	0.129	0.143	0.090
Sharpe ratio	0.086	-0.006	-0.121	-0.286
Alpha	0.021	0.004	0.001	-0.020***

Table 5

Return prediction with decomposed variances.

This table shows how decomposed variances are related to excess returns in the subsequent week. We estimate the coefficients of the following FMB regression:

$$rx_{i,w+1} = \gamma_{0,w} + \gamma_{1,w}TV_{i,w}^{(+)} + c'_w X_{i,w} + \epsilon_{i,w+1} \text{ or}$$

$$rx_{i,w+1} = \gamma_{0,w} + \gamma_{1,w}JV_{i,w}^{(+)} + \gamma_{2,w}JV_{i,w}^{(-)} + \gamma_{3,w}JRV_{i,w} + c'_w X_{i,w} + \epsilon_{i,w+1},$$

where $rx_{i,w}$ is the excess return of cryptocurrency i in week w . $TV_{i,w}^{(+)}$, $JV_{i,w}^{(+)}$, $JV_{i,w}^{(-)}$, and $JRV_{i,w}$ are the total, positive jump, negative jump, and jump-robust variances, respectively. These decomposed variances are estimated with the previous month of observations (i.e., observations from week $w - 3$ to week w). For comparison, we replace $TV_{i,w}^{(+)}$ with the volatility estimated from one-month daily return data (column (II)). $X_{i,w}$ is the vector of control variables such as lagged excess returns and natural logarithmic market capitalization. Then, we report the time-series averages of the estimated coefficients and the corresponding t -statistics. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(I)	(II)	(III)	(IV)	(V)	(VI)
Constant	0.012	-0.005	-0.054**	0.014*	0.120	-0.033
t -stat	1.57	-0.47	-2.10	1.81	1.53	-1.29
Total variance	-0.221***		-0.171***			
t -stat	-7.77		-5.62			
Jump variance				-0.669***		
t -stat				-4.71		
Positive jump variance					-1.470***	-1.492***
t -stat					-5.24	-4.69
Negative jump variance					0.626	0.791
t -stat					1.07	1.31
Jump-robust variance				-0.152***	-0.151***	-0.091**
t -stat				-4.83	-4.41	-2.00
Volatility (daily data)		-0.134				
t -stat		-0.86				
Lagged return			0.006			0.003
t -stat			0.44			0.16
Market capitalization			0.003***			0.002**
t -stat			3.13			2.11
Adj. R^2	0.115	0.092	0.186	0.163	0.200	0.266

Table 6

Robustness test.

This table shows the robustness of the decomposed variance effect. We use the following FMB regression:

$$rx_{i,w+1} = \gamma_{0,w} + \gamma_{1,w}JV_{i,w}^{(+)} + \gamma_{2,w}JV_{i,w}^{(-)} + \gamma_{3,w}JRV_{i,w} + c'_wX_{i,w} + \epsilon_{i,w+1},$$

where $rx_{i,w}$ is the excess return of cryptocurrency i in week w . $JV_{i,w}^{(+)}$, $JV_{i,w}^{(-)}$, and $JRV_{i,w}$ are the positive jump, negative jump, and jump-robust variances, respectively. These decomposed variances are estimated from the previous month of observations (i.e., observations from week $w - 3$ to week w). $X_{i,w}$ is the vector of control variables such as lagged excess returns and natural logarithmic market capitalization. For the part denoted “Dependent variable”, we replace the dependent variable with two-week excess returns (column (I)) and one-month excess returns (column (II)). For the part denoted “Control: Long-term”, we control for the potential return momentum/reversal effect and maximum return effect, employing the lagged excess returns over periods from week $w - 3$ to week w (i.e., one-month returns) for column (III) and those from week $w - 11$ to w (i.e., one-quarter returns) for column (IV). For the part denoted “Control: Maximum”, we use the maximum 15-minute excess returns in week w for column (V) and the maximum one-day returns during the period from week $w - 3$ to week w for column (VI). In this table, we report the time-series averages of the estimated coefficients and the corresponding t -statistics. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent variable		Control: Long-term		Control: Maximum	
	(I)	(II)	(III)	(IV)	(V)	(VI)
Constant	-0.090*	-0.192**	-0.023	-0.001	-0.021	-0.024
t -stat	-1.88	-2.24	-0.96	-0.04	-0.87	-0.98
Positive jump variance	-2.693***	-4.562***	-1.588***	-1.343***	-1.268***	-1.266***
t -stat	-5.05	-6.32	-5.20	-4.01	-4.02	-3.15
Negative jump variance	1.139	-0.333	0.960	0.889	0.740	0.908
t -stat	1.02	-0.25	1.53	1.44	1.22	1.49
Jump-robust variance	-0.169**	-0.234**	-0.104***	-0.126***	-0.117***	-0.159***
t -stat	-2.19	-2.13	-2.57	-3.17	-2.89	-3.91
Lagged return (or Max)	0.036	0.082**	0.059**	0.156***	-0.200***	0.051
t -stat	1.32	2.32	2.00	3.43	-3.27	1.13
Market capitalization	0.005***	0.011***	0.002*	0.001	0.002*	0.002*
t -stat	2.80	3.50	1.80	0.73	1.82	1.79
Adj. R^2	0.281	0.288	0.271	0.267	0.250	0.260

Table 7

Time series subsample analyses.

This table shows how the effects of decomposed variances on subsequent excess returns depend on market conditions. We construct time series subsamples depending on the overall volatility and illiquidity of cryptocurrency markets and business cycles. Cryptocurrency market volatility is defined as the means of the average absolute values of 15-minute log cryptocurrency returns across individual cryptocurrencies during the corresponding week, following Menkhoff et al. (2012). Cryptocurrency market illiquidity is captured by the averages of individual weekly Amihud (2002) illiquidity measures. The High (Low) period subsample is obtained by excluding the 20% of weeks with the lowest (highest) volatility or illiquidity measures. Business cycles are separated considering NBER business cycles (i.e., in our sample, the former (latter) period is before (after) August 2021, and the former period includes recession periods). We apply the following FMB regression to these time series subsamples:

$$rx_{i,w+1} = \gamma_{0,w} + \gamma_{1,w}JV_{i,w}^{(+)} + \gamma_{2,w}JV_{i,w}^{(-)} + \gamma_{3,w}JRV_{i,w} + c'_wX_{i,w} + \epsilon_{i,w+1},$$

where $rx_{i,w}$ is the excess return of cryptocurrency i in week w . $JV_{i,w}^{(+)}$, $JV_{i,w}^{(-)}$, and $JRV_{i,w}$ are the positive jump, negative jump, and jump-robust variances, respectively. These decomposed variances are estimated from the previous month of observations (i.e., observations from week $w - 3$ to week w). $X_{i,w}$ is the vector of control variables such as lagged excess returns and natural logarithmic market capitalization. Then, we report the time-series averages of the estimated coefficients and the corresponding t -statistics. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Market volatility		Market illiquidity		Sample period	
	Low	High	Low	High	Former	Latter
Constant	-0.032**	-0.034**	-0.035**	-0.036***	0.034*	-0.100***
t -stat	-2.37	-2.36	-2.26	-3.86	1.92	-11.61
Positive jump variance	-1.032***	-1.890***	-1.477***	-1.504***	-2.098***	-0.877***
t -stat	-8.38	-7.36	-7.96	-8.15	-7.52	-8.91
Negative jump variance	-0.443**	1.860***	0.457*	1.009***	2.199***	-0.638***
t -stat	-2.14	5.10	1.74	3.73	5.39	-4.67
Jump-robust variance	-0.076***	-0.105***	-0.028	-0.099***	-0.148***	-0.032*
t -stat	-4.77	-3.25	-1.49	-3.92	-4.50	-1.83
Lagged return	-0.003	0.007	-0.008	0.008	0.020	-0.015*
t -stat	-0.30	0.51	-0.76	0.80	1.38	-1.94
Market capitalization	0.002***	0.002***	0.002***	0.002***	-0.000	0.004***
t -stat	3.96	3.70	2.97	6.78	-0.25	13.34
Adj. R^2	0.228	0.301	0.238	0.281	0.361	0.170

Table 8

Comparison of positive jump variance effects with skewness effects.

This table shows the return predictability with skewness, which is compared with that with positive jump variances. We employ the following FMB regression:

$$rx_{i,w+1} = \lambda_{0,w} + \lambda_{1,w}LSkew_{i,w} + \lambda_{2,w}Vol_{i,w} + \lambda_{3,w}Kurt_{i,w} + c'_w X_{i,w} + \varepsilon_{i,w+1},$$

where $rx_{i,w}$ is the excess return of cryptocurrency i in week w . $LSkew_{i,w}$, $Vol_{i,w}$, and $Kurt_{i,w}$ are the realized weekly skewness, volatility (i.e., square root of the variance), and kurtosis of the returns, respectively. These realized moments are estimated from the previous month of observations. To compare the positive jump variance effects with the skewness effects, we replace $Vol_{i,w}$ with the jump-robust volatility (i.e., the square root of the jump-robust variance) and the positive and negative jump variances. $X_{i,w}$ is the vector of control variables such as lagged excess returns and natural logarithmic market capitalization. Then, we report the time-series averages of the estimated coefficients and the corresponding t -statistics. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(I)	(II)	(III)	(IV)	(V)	(VI)
Constant	-0.016	0.063***	0.025	0.065***	0.043***	-0.002
t -stat	-1.73	6.16	0.80	6.08	3.69	-0.08
Skewness	-0.286***	-0.007	-0.009*	-0.009*	0.005	0.002
t -stat	-7.52	-1.46	-1.86	-1.84	0.82	0.23
Volatility		-0.049***	-0.039***			
t -stat		-8.76	-5.88			
Jump-robust volatility				-0.052***	-0.035***	-0.018**
t -stat				-8.08	-4.80	-2.24
Kurtosis		-0.002**	-0.001**	-0.002***	-0.001	-0.000
t -stat		-2.21	-2.17	-2.85	-1.00	-0.20
Lagged return			0.005			0.001
t -stat			0.34			0.06
Market capitalization			0.001			0.001
t -stat			1.12			1.19
Positive jump variance					-1.632***	-1.672***
t -stat					-3.75	-3.57
Negative jump variance					1.019	0.715
t -stat					1.31	0.94
Adj. R^2	0.041	0.173	0.244	0.173	0.253	0.319

Table 9

Characteristics of decomposed variance-sorted portfolios.

This table shows the characteristics of decomposed variance-sorted portfolios. At the beginning of every week, we sort the 100 sample cryptocurrencies based on the estimated (total) variance, positive and negative jump variances, or jump-robust variance and construct tercile portfolios. We estimate the total and decomposed variance measures by using the previous month of observations. For each sorted portfolio, we provide the average market capitalization (in billions of USD), daily trading volume (in billions of USD), prices (in USD), bid-ask spread (BAS), and retail trading proportion (RTP) of Han and Kumar (2013) in the corresponding week. In addition, we report sentiments for individual cryptocurrencies. We use the number of buy opinions (relative to market capitalization) and the percentage of buy opinions relative to total opinions on Twitter and Telegram. To examine the possibility of short selling, we also consider the futures trading volume (in millions of USD) and volumes that buyers take the sellers' futures prices (in millions of USD). In each panel, column Low (High) presents portfolios with the lowest (highest) sorting measures. Column H - L shows the differences between the values of High and Low portfolios. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Total variance

	Low	Mid	High	H - L
Size/trade variable				
Market capitalization (\$B)	39.2	2.2	0.9	-38.3***
Daily trading volume (\$B)	11,704	21,263	23,957	12,253***
Price (\$)	3,124	227	261	-2,863***
Bid-ask spread	0.001	0.001	0.002	0.001***
RTP	0.913	0.958	0.952	0.039***
Sentiment				
# of Twitter buy /market cap	181.147	311.487	598.336	417.188***
% of Twitter buy	25.389	24.965	25.923	0.534
# of Telegram buy /market cap	36.410	109.980	280.846	244.437***
% of Telegram buy	23.347	18.089	16.142	-7.205
Short selling availability				
Futures volumes (\$M)	660	772	1,358	698***
Futures buyers' volume (\$M)	320	375	666	346***

Panel B. Positive jump variance

	Low	Mid	High	H - L
Size/trade variable				
Market capitalization (\$B)	26.9	15.2	1.5	-25.4***
Daily trading volume (\$B)	1,552	16,045	39,133	37,581***
Price (\$)	2,308	1,121	294	-2,014***
Bid-ask spread	0.001	0.001	0.002	0.001***
RTP	0.933	0.938	0.956	0.023***
Sentiment				
# of Twitter buy /market cap	169.808	278.230	639.958	470.151***
% of Twitter buy	25.484	25.524	25.416	-0.068
# of Telegram buy /market cap	52.182	78.820	278.874	226.692***
% of Telegram buy	21.929	19.907	15.941	-5.988
Short selling availability				
Futures volumes (\$M)	587	1,302	935	348***
Futures buyers' volume (\$M)	286	633	458	172***

Table 9

Characteristics of decomposed variance-sorted portfolios (continued).

Panel C. Negative jump variance

	Low	Mid	High	H - L
Size/trade variable				
Market capitalization (\$B)	28.2	12.0	2.7	-25.5***
Daily trading volume (\$B)	8,594	22,304	25,823	17,229***
Price (\$)	2,368	954	381	-1,986***
Bid-ask spread	0.001	0.001	0.002	0.001***
RTP	0.935	0.927	0.960	0.025***
Sentiment				
# of Twitter buy /market cap	239.442	278.201	568.234	328.792***
% of Twitter buy	25.182	25.586	25.556	0.374
# of Telegram buy /market cap	51.585	51.797	296.433	244.848***
% of Telegram buy	21.361	19.777	16.545	-4.816
Short selling availability				
Futures volumes (\$M)	736	1,114	960	225***
Futures buyers' volume (\$M)	358	543	469	112***

Panel D. Jump-robust variance

	Low	Mid	High	H - L
Size/trade variable				
Market capitalization (\$B)	38.9	2.4	1.0	-37.9***
Daily trading volume (\$B)	19,248	26,639	11,064	-8,184***
Price (\$)	3,114	195	299	-2,816***
Bid-ask spread	0.001	0.001	0.002	0.001***
RTP	0.920	0.951	0.951	0.031***
Sentiment				
# of Twitter buy /market cap	239.160	293.993	552.437	313.278***
% of Twitter buy	25.556	25.097	25.678	0.123
# of Telegram buy /market cap	47.538	100.089	277.792	230.255***
% of Telegram buy	22.797	18.068	16.582	-6.215
Short selling availability				
Futures volumes (\$M)	675	724	1,389	714***
Futures buyers' volume (\$M)	327	351	681	354***

Table 10

Prediction of weekly skewness.

This table shows whether weekly skewness can be predicted by lagged weekly skewness or decomposed variances. To investigate this possibility, we use the following panel regression:

$$Skew_{i,w+1} = \psi_0 + \psi_1 LSkew_{i,w} + \psi_2 JV_{i,w}^{(+)} + \psi_3 JV_{i,w}^{(-)} + \psi_4 JRV_{i,w} + c' X_{i,w} + e_{i,w+1},$$

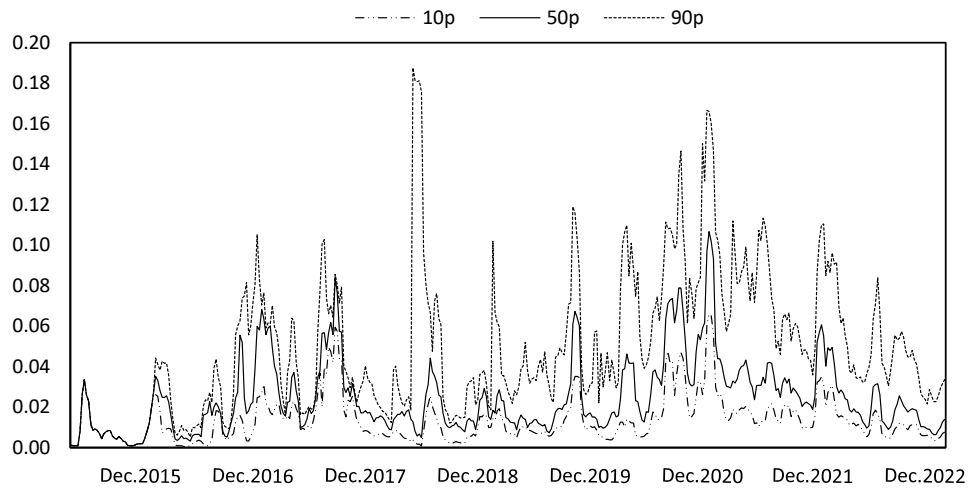
where $Skew_{i,w}$ is the realized weekly skewness of cryptocurrency i in week w and is estimated from the observations in week w . $LSkew_{i,w}$ is the realized skewness estimated from the previous month of observations (i.e., observations from week $w - 3$ to week w). $JV_{i,w}^{(+)}$, $JV_{i,w}^{(-)}$, and $JRV_{i,w}$ are the positive jump, negative jump, and jump-robust variances, respectively. These decomposed variances are estimated from the previous month of observations (i.e., observations from week $w - 3$ to week w). $X_{i,w}$ is the vector of control variables such as weekly kurtosis, lagged excess returns, natural logarithmic market capitalization, and fixed effects. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
Constant	-0.048***	-0.170***	-0.239***	-0.343***	-0.270***	-0.376***	-0.083***
t -stat	-4.96	-13.47	-15.02	-2.99	-10.21	-3.07	-0.25
Skewness	0.224***				0.136***	0.132***	-0.039
t -stat	11.18				5.25	4.96	-1.35
Positive jump variance		11.127***	9.043***	10.391***	4.322***	5.517***	2.842**
t -stat		11.80	9.32	10.15	3.46	4.20	1.99
Negative jump variance		-0.966	-0.196	-1.585	3.887**	2.466	1.089
t -stat		-0.67	-0.14	-1.03	2.42	1.44	0.61
Jump-robust variance			0.649***	0.647***	0.684***	0.692***	0.414***
t -stat			10.50	9.75	10.01	9.30	5.61
Kurtosis					0.004*	0.004**	0.004
t -stat					1.93	1.97	1.67
Lagged return				-0.196***		-0.213***	-0.116**
t -stat				-4.19		-4.57	-2.42
Market capitalization				0.005		0.005	0.002
t -stat				0.87		0.85	0.13
Fixed effects	No	No	No	No	No	No	Yes
Adj. R^2	0.014	0.024	0.031	0.032	0.034	0.035	0.069

Figure 1. Weekly variance over time

This figure illustrates how weekly variances change during the sample period from October 2015 to June 2023. Panel A shows the 10th, 50th, and 90th percentiles of individual cryptocurrency variances in the cross-section. Panel B provides weekly variances for bitcoin.

Panel A. All cryptocurrencies



Panel B. Bitcoin

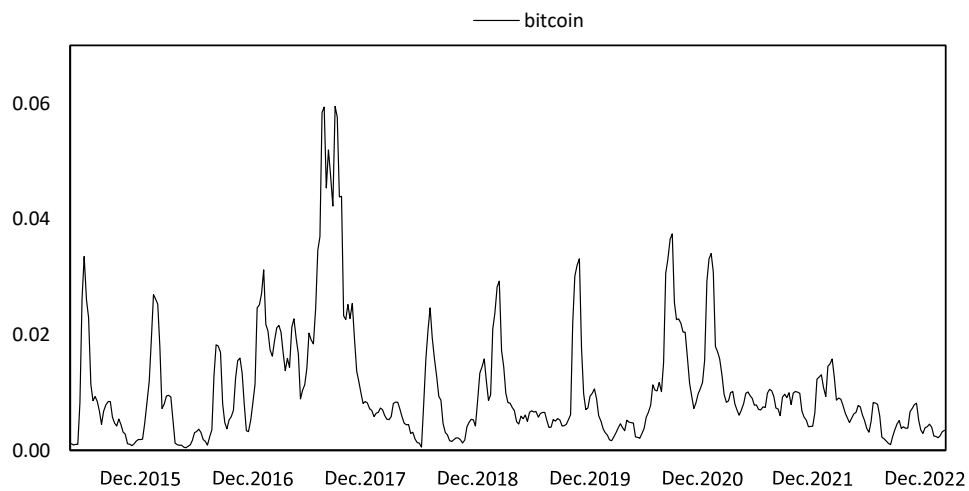


Figure 2. Variances and future excess returns

This figure shows the effect of variances on excess returns in the subsequent week. At the end of every week, the cryptocurrencies are sorted on their variances, and tercile portfolios are constructed. For each sorted portfolio, we compute the equally weighted average of excess returns in the subsequent weeks. We compare the average excess returns of low-variance portfolios with those of high-variance portfolios.

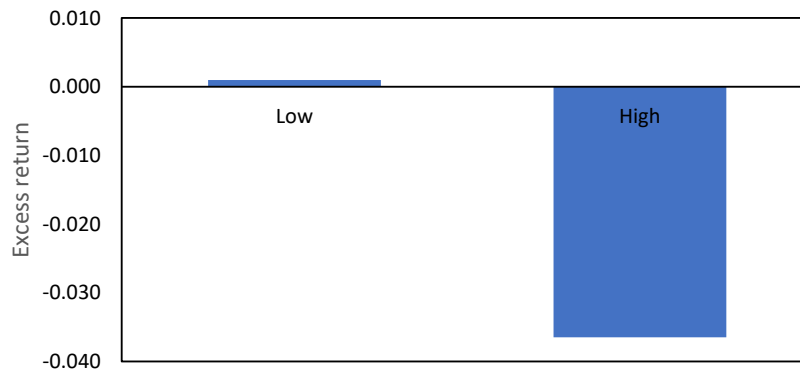
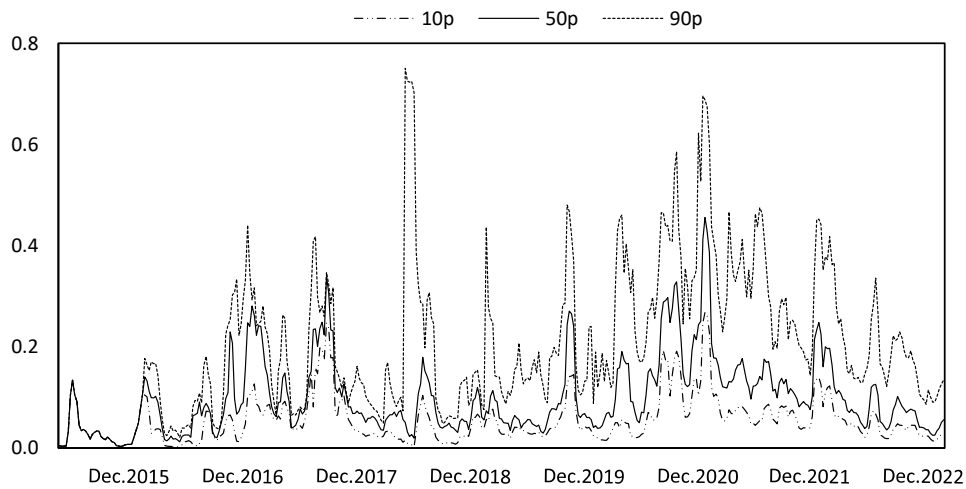


Figure 3. Time series of weekly realized variances

This figure illustrates how the weekly decomposed variances change during the sample period from October 2015 to June 2023. Panels A, B, C, and D are for the total, positive jump, negative jump, and jump-robust variances, respectively. These decomposed variance measures are estimated from the previous month of observations. In each panel, we provide the 10th, 50th, and 90th percentiles of decomposed variances in the cross-section.

Panel A. Total variance



Panel B. Positive jump variance

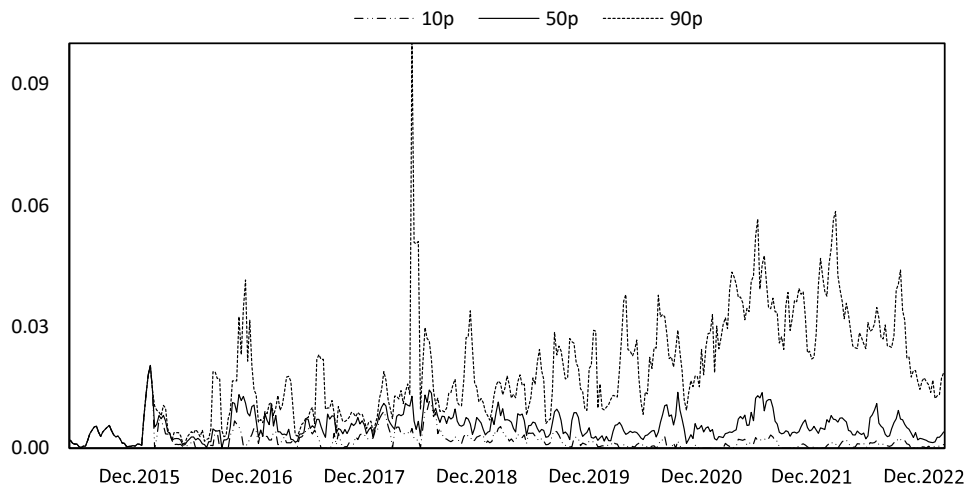
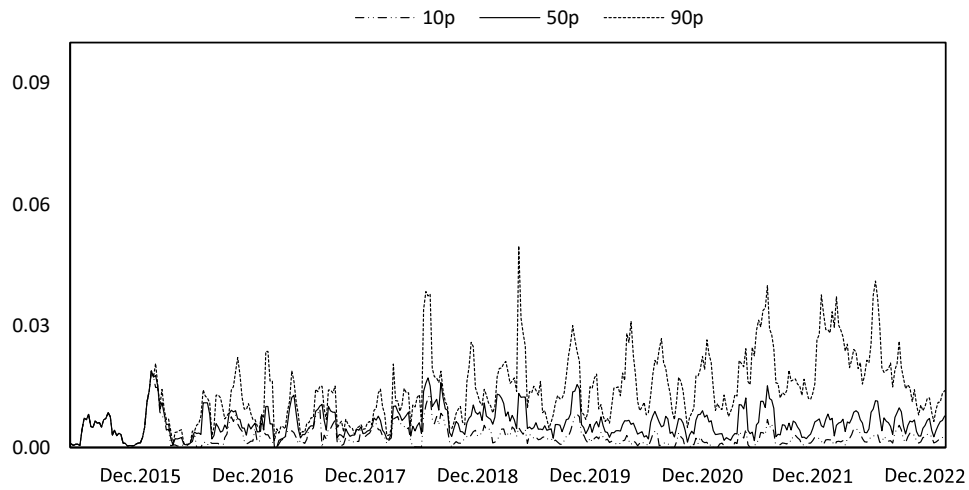
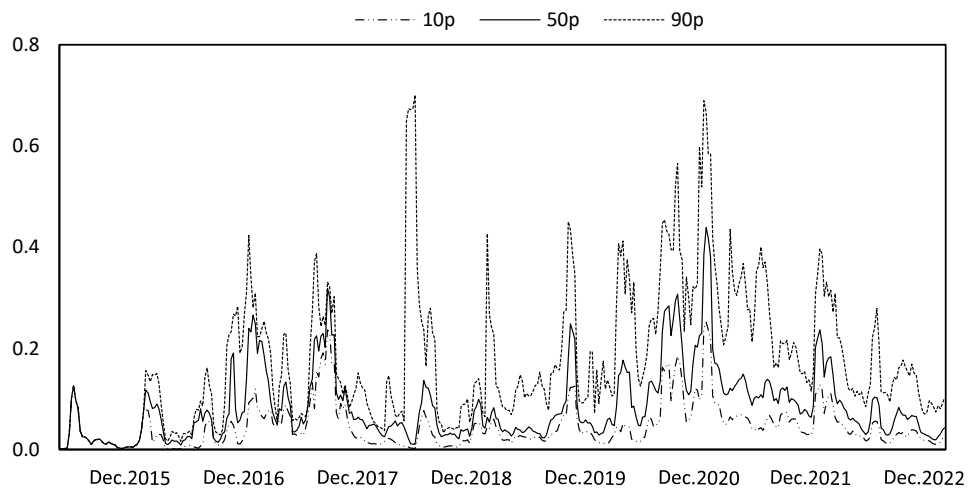


Figure 3. Time series of weekly realized variances (continued)

Panel C. Negative jump variance



Panel D. Jump-robust variance



Online Appendix for “Variance Decomposition and Cryptocurrency Return Prediction”

Suzanne Lee

Minho Wang

January 2024

To keep the main text brief, we provide this online appendix to further support our empirical results and to expand on the details of the summarized points in the main text.

This online appendix includes multiple appendixes, which are located in the order of the main text. Online Appendix A supports our empirical application of jump variance estimators by checking their finite sample performance. Online Appendix B provides the summary statistics and jump detection results for all 100 sample cryptocurrencies and the full names. Online Appendix C elaborates upon the approaches for constructing samples and filtering intraday data. Online Appendix D compares the results of using the measures that we introduce in this paper with those of employing alternative measures. Online Appendix E supplements the contemporaneous sorting analyses with characteristics by regressing price increases and partial variances on characteristic variables. Online Appendix F defines the variables used in this paper. Each appendix contains the corresponding figure and table.

O. Appendix A. Monte Carlo simulation

As indicated in the main text, taking the total realized variance as the estimator of integrated variance (volatility) is a common approach in the literature (see Andersen et al. (2003) and the references therein). In this paper, we suggest further decomposing total realized variances into jump-robust and (signed) jump variances by applying jump detection tests that identify the arrivals of individual cryptocurrency jumps (with different signs). To support the empirical application of our jump variance estimators for our study, in this appendix, we examine our measures' finite sample performance by using Monte Carlo simulation. The purpose of this simulation study is to prove that the proposed estimators converge to the true jump variances that we aim to identify in the model. By definition, the performance of our jump variance estimators depends on the performance of the jump tests because the estimators use cryptocurrency returns in the time intervals during which jump tests detect jump arrivals. Therefore, we design this simulation to study the effectiveness of our jump detection tests and the estimation errors that can be made by using our jump variance estimator(s). In sum, the overall results show that our jump variances perform well in estimating return variations associated with jumps under general market conditions.

For return generation, we use the Euler-Maruyama scheme, which is widely used to simulate data from continuous-time models. To avoid starting value effects, we discard the first five hundred observations from the burn-in period every time we generate a time series. We generate intraday data over one trading day from the general model in Section 2 as

$$dc_{i,t} = \mu_{i,t}dt + \sigma_{i,t}dB_{i,t} + Y_{i,t}dJ_{i,t}, \quad (9)$$

where the stochastic volatility model is specified as the following square root processes:

$$d\sigma_{i,t}^2 = \kappa(\theta - \sigma_{i,t}^2)dt + \omega\sigma_{i,t}dW_{i,t}. \quad (10)$$

The terms $dB_{i,t}$ and $dW_{i,t}$ are standard Brownian motion processes. $Y_{i,t}$ and $dJ_{i,t}$ denote the jump size and jump arrival indicator for the i -th cryptocurrency price at time t , respectively, as indicated in Subsection 2.1. The jump sizes are selected relative to the volatility level $\sigma_{i,t-}$ immediately before jump time t . Because we assume that volatility is stochastic, the jump sizes are also time varying and are set at $5\sigma_{i,t-}$, $3\sigma_{i,t-}$, $1\sigma_{i,t-}$, $0.5\sigma_{i,t-}$ and $0.25\sigma_{i,t-}$. We ignore the drift term in this simulation by setting $\mu_{i,t} = 0$ because the magnitudes of the drift terms, compared to those of the diffusion and jump terms, are negligible at intraday levels. We use the parameters for the volatility process from Li, Wells and Yu (2008): $\kappa = 0.0162$, $\theta = 0.8465$, and $\omega = 0.47$. We consider the full 24 trading hours per day and simulate 10,000 paths of the log price process $c_{i,t}$. We consider various sampling frequencies of 5 minutes, 15 minutes, 30 minutes, 1 hour, 2 hours, 3 hours, 6 hours, and 12 hours.

The results from this simulation study are reported in Table OA. Panel A shows that as we increase the sampling frequency, the likelihood of detecting true jump arrivals by our jump tests approaches 100%. As long as the goal of the application is to detect rare and extremely large jumps, our jump variance estimators are expected to perform well when we use high-frequency intraday data. Panel B confirms this expectation, indicating that the mean squared error (MSE) of the jump variance estimator significantly decreases as we increase the sampling frequency within a day. Because of this result, we use 15-minute cryptocurrency returns and confirm our results with different intraday return data.

Table OA

Finite sample performance of the jump variance estimator.

This table shows the results of simulation analyses to present the performance of the jump variance estimators used in our study. Considering that the performance of jump variance estimators depends on the performance of jump tests in detecting jump arrivals, we first examine how likely it is that our jump tests detect true jump arrivals and report the results in Panel A. Then, we investigate the estimator error that can be made by applying our jump variance estimators and report the results in Panel B. The number of simulations is set to 10,000. The jump sizes are set relative to the stochastic volatility level $\sigma_{i,t-}$ immediately before jump time t . We consider various intraday sampling frequencies. The estimation error is measured by the mean squared error (MSE).

Panel A. Likelihood of detecting true jump arrival

Sampling Frequency	Jump sizes relative to volatility level				
	$5 \times \sigma_{i,t-}$	$3 \times \sigma_{i,t-}$	$1 \times \sigma_{i,t-}$	$0.5 \times \sigma_{i,t-}$	$0.25 \times \sigma_{i,t-}$
12 hours	0.9656	0.9628	0.9566	0.9389	0.6846
6 hours	0.9878	0.9885	0.9867	0.9815	0.9440
3 hours	0.9961	0.9955	0.9943	0.9930	0.9862
2 hours	0.9969	0.9971	0.9955	0.9965	0.9911
1 hour	0.9989	0.9983	0.9980	0.9977	0.9972
30 minutes	0.9990	0.9993	0.9992	0.9990	0.9988
15 minutes	0.9996	1.0000	1.0000	0.9996	0.9994
5 minutes	1.0000	1.0000	1.0000	1.0000	1.0000

Panel B. Performance of the jump variance estimator measured by the MSE

Sampling Frequency	Jump sizes relative to volatility level				
	$5 \times \sigma_{i,t-}$	$3 \times \sigma_{i,t-}$	$1 \times \sigma_{i,t-}$	$0.5 \times \sigma_{i,t-}$	$0.25 \times \sigma_{i,t-}$
12 hours	0.000398	0.000405	0.000388	0.000377	0.000234
6 hours	0.000192	0.000191	0.000197	0.000185	0.000169
3 hours	0.000092	0.000094	0.000093	0.000093	0.000091
2 hours	0.000059	0.000061	0.000061	0.000061	0.000060
1 hour	0.000030	0.000030	0.000031	0.000031	0.000030
30 minutes	0.000015	0.000015	0.000015	0.000015	0.000015
15 minutes	0.000007	0.000007	0.000007	0.000008	0.000008
5 minutes	0.000003	0.000003	0.000002	0.000003	0.000003

O. Appendix B. Summary statistics for all sample cryptocurrencies

In this appendix, we provide the summary statistics for all 100 sample cryptocurrencies. Table OB1 shows the four central moments of 15-minute returns and the unconditional averages of market capitalization, daily trading volumes in billion U.S. dollars (USD), prices per coin in USD, and percentage bid-ask spreads (BASs). Table OB2 presents the number of filtered jumps, the jump frequencies relative to the available observations, and the distributions of jump sizes. In these tables, we provide the abbreviated and full names of the sample cryptocurrencies.

Table OB1

Summary statistics (all sample coins).

This table shows the summary statistics of the cryptocurrency data. Column Cryptocurrency lists the abbreviated and full names of the cryptocurrencies in the sample. Column # of obs shows the number of 15-minute return observations after data filtering. Column Start indicates the months in which the earliest observations appear in the sample. Columns under 15-min return provide the mean, standard deviation, skewness, and kurtosis of the 15-minute returns for each sample cryptocurrency. Columns Market cap., Volume, Price, and BAS show the unconditional averages of market capitalization, daily trading volumes in billion U.S. dollars (USD), prices per coin in USD, and percentage bid-ask spreads (BASs), respectively. Row Avg.100 shows the averages across the 100 sample cryptocurrencies.

Crypto currency	# obs	Start	15-min return				Market cap (\$B)	Volume (\$B)	Price (\$)	BAS (%)
			Mean (%)	Stdev (%)	Skew	Kurt				
1inch (1INCH)	68,474	Apr. 2021	-0.0043	0.6519	-0.214	12.320	0.52	0.13	1.75	0.19
aave (AAVE)	88,379	Dec. 2020	0.0000	0.6842	0.008	11.520	2.32	0.29	187.47	0.06
ach (Alchemy Pay)	66,572	Aug. 2021	-0.0059	0.9386	0.577	16.242	0.10	0.03	0.04	0.11
ada (Cardano)	78,975	Mar. 2021	-0.0014	0.5473	0.036	12.444	15.03	1.05	0.95	0.03
agld (Adventure Gold)	55,573	Sep. 2021	-0.0060	0.8321	0.314	13.859	0.07	0.03	0.90	0.34
algo (Algorand)	132,270	Aug. 2019	-0.0007	0.7436	0.002	23.743	2.54	0.20	0.60	0.09
amp (Amp)	58,889	Jun. 2021	-0.0115	0.7391	0.231	15.512	0.81	0.02	0.03	0.17
ankr (Ankr Network)	78,597	Mar. 2021	-0.0044	0.7041	0.121	14.203	0.30	0.06	0.06	0.08
arpa (ArpaCoin)	49,385	Oct. 2021	-0.0053	0.7322	0.283	15.443	0.05	0.03	0.06	0.28
asm (Assemble Protocol)	54,380	Oct. 2021	-0.0086	0.9920	0.439	21.955	0.03	0.01	0.04	0.56
atom (Cosmos)	119,789	Jan. 2020	0.0004	0.6997	0.022	11.126	2.97	0.40	13.87	0.08
avax (Avalanche)	60,069	Sep. 2021	-0.0023	0.5888	0.004	10.570	7.75	0.51	40.84	0.06
axs (Axie Infinity)	63,161	Aug. 2021	-0.0048	0.6271	0.250	13.614	1.93	0.30	39.41	0.12
badger (Badger DAO)	43,220	Oct. 2021	-0.0067	0.6605	0.283	11.017	0.12	0.02	8.30	0.32
bal (Balancer)	88,571	Oct. 2020	-0.0010	0.6357	-0.253	9.974	0.20	0.05	17.27	0.16
band (Band Protocol)	99,088	Aug. 2020	-0.0029	0.7843	0.017	11.578	0.14	0.05	5.57	0.12
bat (Basic Attention Token)	73,999	Apr. 2021	-0.0021	0.6091	-0.017	12.079	0.52	0.11	0.56	0.15
bch (Bitcoin Cash)	190,108	Dec. 2017	-0.0003	0.5673	0.053	15.242	8.00	1.92	416.19	0.04
bnt (Bancor)	77,769	Dec. 2020	-0.0013	0.5485	-0.173	12.038	0.25	0.03	2.38	0.21
bond (BarnBridge)	58,181	Jun. 2021	-0.0134	0.8376	0.658	19.046	0.06	0.02	11.82	0.27
btc (Bitcoin)	263,892	Oct. 2015	0.0023	0.3983	-0.188	19.076	270.99	17.82	15,517.00	0.01
btrst (Braintrust)	46,563	Sep. 2021	-0.0131	0.7803	0.235	31.846	0.23	0.00	2.98	0.31
chz (Chiliz)	69,606	Jun. 2021	-0.0023	0.6381	0.133	12.121	0.84	0.22	0.21	0.11
clv (Clover)	27,408	Jul. 2021	-0.0083	1.0971	0.723	18.628	0.07	0.02	0.60	0.55
comp (Compound Coin)	104,874	Jun. 2020	-0.0012	0.6993	-0.107	11.129	0.98	0.13	187.46	0.08
coti (COTI)	61,137	Aug. 2021	-0.0038	0.6872	-0.013	12.846	0.12	0.03	0.19	0.14
crv (Curve DAO Token)	78,420	Mar. 2021	-0.0003	0.7369	0.114	10.490	0.74	0.19	1.91	0.10
ctsi (Cartesi)	73,959	May. 2021	-0.0063	0.7874	0.282	14.966	0.13	0.03	0.37	0.14
dash (Digital Cash)	131,336	Sep. 2019	-0.0010	0.5870	-0.016	13.990	1.14	0.26	102.48	0.12
ddx (DerivaDAO)	37,028	Sep. 2021	-0.0225	1.0613	0.447	23.295	0.07	0.00	2.18	0.80
doge (DogeCoin)	71,844	Jun. 2021	-0.0041	0.5537	0.071	15.383	5.67	0.67	0.13	0.05
dot (Polkadot)	70,964	Jun. 2021	-0.0016	0.5304	-0.077	11.275	15.69	1.20	14.87	0.05
enj (Enjin Coin)	67,704	Apr. 2021	-0.0039	0.7402	-0.007	11.162	0.49	0.09	1.20	0.21
eos (Eos)	135,076	Apr. 2019	0.0014	0.5601	-0.095	14.771	3.33	1.43	3.09	0.14
etc (Ethereum Classic)	161,484	Aug. 2018	-0.0001	0.6249	0.069	16.790	2.10	0.80	19.88	0.09
eth (Ethereum)	229,770	Jul. 2016	0.0022	0.5168	-0.068	15.162	100.31	8.96	1,003.27	0.03
farm (Harvest Finance)	64,821	Jul. 2021	-0.0096	0.7227	0.632	21.175	0.06	0.01	81.15	0.17
fet (Fetch AI)	65,552	Jul. 2021	-0.0019	0.8350	0.406	12.278	0.16	0.03	0.33	0.14
fil (Filecoin)	88,894	Dec. 2020	-0.0010	0.6099	0.146	14.300	3.49	0.60	33.18	0.06
forth (Ampleforth Governance Token)	58,318	Apr. 2021	-0.0161	0.9642	0.719	20.215	0.08	0.02	9.60	0.32
grt (The Graph)	86,741	Dec. 2020	-0.0008	0.8287	0.076	12.083	1.70	0.17	0.52	0.09
gtc (Gitcoin)	55,492	Jun. 2020	-0.0042	0.8714	0.125	12.066	0.09	0.02	5.49	0.41
icp (Internet Computer)	74,552	May. 2021	-0.0053	0.7339	0.111	12.122	4.23	0.20	25.00	0.07
iotx (IoTeX)	64,761	Aug. 2021	-0.0056	0.7859	0.251	22.902	0.23	0.03	0.07	0.24

Table OB1

Summary statistics (all sample coins, continued).

Crypto currency	# obs	Start	15-min return				Market cap (\$B)	Volume (\$B)	Price (\$)	BAS (%)
			Mean (%)	Stdev (%)	Skew	Kurt				
jasmy (Jasmy)	56,425	Oct. 2021	-0.0067	0.9691	0.475	15.532	0.20	0.13	0.03	0.17
keep (Keep Token)	53,060	Jun. 2021	-0.0117	0.8218	0.242	19.801	0.20	0.02	0.37	0.26
knc (Kyber Network)	116,027	Feb. 2020	-0.0003	0.7623	0.103	13.128	0.17	0.03	1.46	0.12
link (ChainLink)	139,040	Jun. 2019	0.0010	0.6488	0.015	11.856	3.74	0.68	12.20	0.08
lpt (Livepeer)	65,917	Jun. 2021	-0.0078	0.7126	0.038	14.997	0.32	0.02	16.80	0.20
lrc (Loopring)	96,365	Sep. 2020	-0.0024	0.8149	0.276	14.184	0.40	0.08	0.56	0.09
ltc (Litecoin)	216,697	Oct. 2016	0.0012	0.5936	0.016	15.027	4.67	1.55	91.74	0.07
mana (Decentraland)	75,960	Apr. 2021	-0.0024	0.6967	0.192	13.467	0.93	0.19	1.23	0.08
mask (Mask Network)	57,353	Jul. 2021	-0.0038	0.9049	0.405	12.468	0.18	0.09	5.45	0.35
matic (Matic Network)	80,399	Mar. 2021	-0.0007	0.7557	-0.073	26.172	4.85	0.54	1.14	0.04
mir (Mirror Protocol)	64,287	May. 2021	-0.0144	0.8324	0.571	20.950	0.13	0.03	1.67	0.17
mkr (Maker)	106,360	Jun. 2020	-0.0007	0.6037	0.048	12.437	1.06	0.08	1,528.35	0.07
mln (Enzyme)	67,428	Jun. 2021	-0.0109	0.6735	0.468	25.940	0.05	0.00	52.34	0.17
nkn (NKN)	75,198	Apr. 2021	-0.0053	0.7894	0.089	13.601	0.08	0.01	0.23	0.14
nmr (Numeraire)	96,091	Aug. 2020	-0.0102	0.7907	0.631	19.595	0.09	0.01	28.69	0.16
nu (NuCypher)	73,003	Dec. 2020	-0.0060	0.9749	0.436	18.910	0.20	0.05	0.36	0.20
ogn (Origin Protocol)	73,659	Apr. 2021	-0.0058	0.7926	0.060	12.058	0.13	0.05	0.48	0.21
omg (OmiseGO)	94,855	May. 2020	0.0002	0.7718	0.182	11.982	0.56	0.19	4.30	0.14
orn (Orion Protocol)	59,227	Aug. 2021	-0.0085	0.6585	0.317	15.959	0.12	0.01	3.02	0.26
oxt (Orchid)	111,485	Dec. 2019	-0.0065	0.8280	0.489	17.496	0.12	0.03	0.26	0.15
perp (Perpetual Protocol)	54,088	Oct. 2021	-0.0094	0.7222	0.313	13.930	0.26	0.02	3.17	0.27
pla (PlayDapp Token)	63,188	Aug. 2021	-0.0080	0.7346	0.177	22.585	0.20	0.06	0.63	0.26
poly (Polymath)	62,822	Jul. 2021	-0.0073	0.7336	0.447	26.572	0.14	0.02	0.33	0.30
qnt (Quant)	70,421	Jun. 2021	-0.0013	0.7279	0.223	12.981	0.80	0.02	139.68	0.11
quick (QuickSwap)	64,169	Aug. 2021	-0.0113	0.7066	0.290	24.280	0.09	0.02	166.42	0.30
rad (Radicle)	40,183	Sep. 2021	-0.0197	0.9760	0.583	16.641	0.13	0.02	5.04	0.44
rai (Rai Reflex Index)	3,754	Aug. 2021	-0.0049	0.2697	-1.338	25.152	0.04	0.00	2.92	0.64
rari (Rarible)	33,261	Oct. 2021	-0.0195	1.1606	0.199	17.226	0.04	0.00	7.25	0.57
ren (Ren)	93,807	Oct. 2020	-0.0025	0.7998	-0.024	10.924	0.23	0.04	0.40	0.11
rep (Augur)	119,864	Apr. 2019	-0.0064	0.7078	0.188	20.966	0.20	0.02	16.54	0.29
req (Request Network)	59,605	Aug. 2021	-0.0113	0.8079	0.728	19.626	0.09	0.01	0.17	0.18
rgt (Rari Governance Token)	34,609	Sep. 2021	-0.0128	1.1810	0.611	21.721	0.11	0.00	16.17	0.57
rlc (iExec RLC)	69,368	May. 2021	-0.0042	0.7968	0.074	13.966	0.10	0.01	2.31	0.23
rly (Rally)	44,522	Jul. 2021	-0.0098	0.8733	1.706	50.526	0.22	0.01	0.25	0.32
shib (Shiba Inu)	59,288	Sep. 2021	0.0010	0.7420	0.628	32.800	8.44	1.10	0.00	0.12
skl (SKALE Network)	72,439	Mar. 2021	-0.0052	0.8148	0.022	11.673	0.31	0.04	0.19	0.18
snx (Synthetix Network Token)	87,599	Dec. 2020	-0.0003	0.7602	0.039	10.257	0.50	0.06	7.06	0.09
sol (Solana)	70,755	Jun. 2021	-0.0017	0.6460	0.084	12.256	15.20	1.08	68.14	0.04
storj (Storj)	69,457	Mar. 2021	-0.0053	0.7762	0.096	13.376	0.16	0.04	0.90	0.26
sushi (Sushi)	75,380	Mar. 2021	-0.0028	0.7009	-0.010	10.524	0.68	0.25	5.27	0.17
trb (Tellor Tributes)	73,914	May. 2021	-0.0030	0.7505	-0.023	13.102	0.05	0.03	27.77	0.16
tribe (Tribe)	40,963	Aug. 2021	-0.0079	0.4960	0.389	26.994	0.21	0.01	0.55	0.24
tru (TrueFi)	58,764	Aug. 2021	-0.0118	0.8603	0.376	15.483	0.07	0.01	0.19	0.26
uma (UMA)	95,853	Sep. 2020	-0.0089	0.7807	0.524	16.737	0.51	0.04	8.26	0.13
uni (Uniswap)	96,279	Sep. 2020	0.0001	0.6784	0.025	11.746	7.36	0.47	12.73	0.07
wbtc (Wrapped Bitcoin)	93,708	Oct. 2020	0.0008	0.3755	-0.059	13.031	4.82	0.15	34,096.31	0.10
wcfg (Wrapped Centrifuge)	23,013	Oct. 2021	-0.0385	1.2544	0.479	16.385			0.47	0.65
xlm (Stellar)	149,494	Mar. 2019	-0.0002	0.5553	-0.001	16.177	2.75	0.28	0.16	0.07
xrp (Ripple)	64,229	Feb. 2019	0.0005	0.5197	-0.109	20.925	15.87	1.79	0.28	0.06
xtz (Tezos)	129,846	Aug. 2019	0.0003	0.6640	0.023	11.500	1.80	0.14	2.60	0.15
xyo (XYO Network)	52,624	Sep. 2021	-0.0077	0.9156	0.566	16.526	0.08	0.00	0.02	0.25
yfi (DFI.money)	54,960	Sep. 2021	-0.0066	0.7134	-0.030	27.228	0.16	0.08	1,963.18	0.19
yfi (yearn.finance)	97,232	Sep. 2020	-0.0020	0.6745	0.014	11.974	0.66	0.24	21,157.01	0.07
zec (Zcash)	88,913	Dec. 2020	0.0000	0.6522	0.002	12.724	0.84	0.26	102.28	0.08
zen (Horizen)	59,632	Sep. 2021	-0.0049	0.5999	-0.160	12.542	0.23	0.02	30.28	0.15
zrx (0x)	160,121	Oct. 2018	-0.0010	0.6995	0.071	24.108	0.39	0.06	0.50	0.14
Avg.100	81,626		-0.0052	0.7370	0.187	16.443	5.44	0.49	773.57	0.19

Table OB2

Summary statistics of cryptocurrency jumps (all sample coins).

This table summarizes the results of jump detection tests in cryptocurrency markets. To identify jumps, we apply the approach of Lee and Mykland (2008) and adjust the intraday volatility patterns of individual cryptocurrencies, following Lee and Wang (2020). Column Cryptocurrency lists the currency codes of the 100 sample cryptocurrencies. For each cryptocurrency, we provide the number of filtered jumps and the jump frequencies relative to the available observations (% jp). For signed jumps, we report the 25th, 50th, and 75th percentiles of positive and negative jump sizes in the last six columns. The numbers in columns % jp, positive jump size, and negative jump size are in percentage terms. Row Avg.100 shows the averages across the 100 sample cryptocurrencies.

Crypto currency	Jump frequency				Positive jump size (%)			Negative jump size (%)		
	Total	Positive	Negative	% jp	25p	50p	75p	25p	50p	75p
1inch (1INCH)	213	85	128	0.311	0.0243	0.0342	0.0438	-0.0421	-0.0299	-0.0234
aave (AAVE)	235	98	137	0.266	0.0209	0.0319	0.0414	-0.0367	-0.0270	-0.0189
ach (Alchemy Pay)	407	242	165	0.611	0.0245	0.0344	0.0539	-0.0433	-0.0305	-0.0221
ada (Cardano)	291	131	160	0.368	0.0171	0.0228	0.0310	-0.0280	-0.0204	-0.0139
agld (Adventure Gold)	338	181	157	0.608	0.0243	0.0340	0.0494	-0.0441	-0.0307	-0.0232
algo (Algorand)	451	183	268	0.341	0.0232	0.0362	0.0535	-0.0431	-0.0295	-0.0211
amp (Amp)	335	193	142	0.569	0.0227	0.0300	0.0421	-0.0402	-0.0291	-0.0205
ankr (Ankr Network)	364	160	204	0.463	0.0229	0.0337	0.0439	-0.0365	-0.0261	-0.0204
arpa (ArpaCoin)	186	93	93	0.377	0.0311	0.0407	0.0576	-0.0439	-0.0364	-0.0277
asm (Assemble Protocol)	647	328	319	1.190	0.0283	0.0391	0.0588	-0.0514	-0.0349	-0.0245
atom (Cosmos)	311	128	183	0.260	0.0217	0.0285	0.0409	-0.0373	-0.0263	-0.0193
avax (Avalanche)	171	65	106	0.285	0.0173	0.0280	0.0366	-0.0318	-0.0227	-0.0178
axs (Axie Infinity)	276	132	144	0.437	0.0240	0.0324	0.0434	-0.0367	-0.0265	-0.0213
badger (Badger DAO)	150	87	63	0.347	0.0263	0.0346	0.0479	-0.0359	-0.0283	-0.0224
bal (Balancer)	224	73	151	0.253	0.0211	0.0267	0.0342	-0.0370	-0.0281	-0.0222
band (Band Protocol)	293	122	171	0.296	0.0294	0.0369	0.0476	-0.0451	-0.0330	-0.0245
bat (Basic Attention Token)	225	72	153	0.304	0.0209	0.0264	0.0363	-0.0316	-0.0235	-0.0159
bch (Bitcoin Cash)	1,025	517	508	0.539	0.0170	0.0229	0.0320	-0.0306	-0.0221	-0.0164
bnt (Bancor)	284	104	180	0.365	0.0166	0.0224	0.0306	-0.0286	-0.0212	-0.0147
bond (BarnBridge)	490	285	205	0.842	0.0280	0.0368	0.0560	-0.0443	-0.0328	-0.0247
btc (Bitcoin)	2180	1037	1143	0.826	0.0083	0.0129	0.0192	-0.0212	-0.0141	-0.0094
btrst (Braintrust)	960	483	477	2.062	0.0149	0.0226	0.0377	-0.0322	-0.0201	-0.0137
chz (Chiliz)	253	121	132	0.363	0.0223	0.0310	0.0411	-0.0362	-0.0274	-0.0197
clv (Clover)	120	81	39	0.438	0.0334	0.0458	0.0643	-0.0493	-0.0339	-0.0237
comp (Compound Coin)	273	95	178	0.260	0.0253	0.0362	0.0438	-0.0407	-0.0305	-0.0235
coti (COTI)	182	74	108	0.298	0.0245	0.0311	0.0416	-0.0430	-0.0326	-0.0210
crv (Curve DAO Token)	207	89	118	0.264	0.0236	0.0313	0.0428	-0.0392	-0.0297	-0.0222
ctsi (Cartesi)	293	167	126	0.396	0.0278	0.0350	0.0476	-0.0442	-0.0323	-0.0245
dash (Digital Cash)	462	205	257	0.352	0.0180	0.0264	0.0365	-0.0338	-0.0249	-0.0177
ddx (DerivaDAO)	544	288	256	1.469	0.0196	0.0338	0.0512	-0.0452	-0.0311	-0.0183
doge (Dogecoin)	425	193	232	0.592	0.0184	0.0241	0.0338	-0.0288	-0.0218	-0.0164
dot (Polkadot)	180	58	122	0.254	0.0178	0.0258	0.0311	-0.0326	-0.0212	-0.0141
enj (Enjin Coin)	171	57	114	0.253	0.0306	0.0389	0.0513	-0.0418	-0.0310	-0.0209
eos (Eos)	675	307	368	0.500	0.0175	0.0228	0.0312	-0.0314	-0.0231	-0.0162
etc (Ethereum Classic)	968	466	502	0.599	0.0175	0.0252	0.0342	-0.0341	-0.0244	-0.0165
eth (Ethereum)	1318	639	679	0.574	0.0125	0.0185	0.0262	-0.0263	-0.0197	-0.0138
farm (Harvest Finance)	595	334	261	0.918	0.0203	0.0296	0.0435	-0.0375	-0.0246	-0.0180
fet (Fetch AI)	262	159	103	0.400	0.0303	0.0381	0.0521	-0.0468	-0.0343	-0.0251
fil (Filecoin)	423	217	206	0.476	0.0219	0.0290	0.0374	-0.0365	-0.0268	-0.0184
forth (Ampleforth Governance Token)	491	302	189	0.842	0.0295	0.0416	0.0623	-0.0465	-0.0350	-0.0271
grt (The Graph)	227	93	134	0.262	0.0281	0.0334	0.0456	-0.0449	-0.0344	-0.0231
gtc (Gitcoin)	180	83	97	0.324	0.0333	0.0429	0.0590	-0.0520	-0.0406	-0.0317
icp (Internet Computer)	222	90	132	0.298	0.0242	0.0326	0.0440	-0.0363	-0.0287	-0.0201
iotx (IoTeX)	395	201	194	0.610	0.0210	0.0315	0.0456	-0.0423	-0.0295	-0.0202

Table OB2

Summary statistics of cryptocurrency jumps (all sample coins, continued).

Crypto currency	Jump frequency				Positive jump size (%)			Negative jump size (%)		
	Total	Positive	Negative	% jp	25p	50p	75p	25p	50p	75p
jasmy (Jasmy)	228	136	92	0.404	0.0311	0.0445	0.0613	-0.0464	-0.0336	-0.0270
keep (Keep Token)	574	291	283	1.082	0.0238	0.0348	0.0499	-0.0454	-0.0335	-0.0222
knc (Kyber Network)	416	190	226	0.359	0.0228	0.0335	0.0444	-0.0384	-0.0266	-0.0188
link (ChainLink)	366	151	215	0.263	0.0191	0.0258	0.0351	-0.0366	-0.0259	-0.0176
lpt (Livepeer)	283	127	156	0.429	0.0274	0.0382	0.0487	-0.0419	-0.0314	-0.0226
lrc (Loopring)	365	178	187	0.379	0.0254	0.0383	0.0545	-0.0422	-0.0333	-0.0228
ltc (Litecoin)	1091	526	565	0.503	0.0171	0.0235	0.0325	-0.0327	-0.0235	-0.0176
mana (Decentraland)	235	92	143	0.309	0.0200	0.0302	0.0425	-0.0349	-0.0242	-0.0169
mask (Mask Network)	174	82	92	0.303	0.0360	0.0479	0.0631	-0.0531	-0.0432	-0.0334
matic (Matic Network)	215	88	127	0.267	0.0181	0.0268	0.0374	-0.0371	-0.0254	-0.0171
mir (Mirror Protocol)	483	292	191	0.751	0.0244	0.0342	0.0503	-0.0426	-0.0308	-0.0229
mkr (Maker)	387	180	207	0.364	0.0169	0.0247	0.0332	-0.0311	-0.0239	-0.0172
mln (Enzyme)	807	396	411	1.197	0.0176	0.0291	0.0417	-0.0361	-0.0233	-0.0158
nkn (NKN)	309	124	185	0.411	0.0306	0.0376	0.0495	-0.0451	-0.0340	-0.0254
nmr (Numeraire)	965	536	429	1.004	0.0204	0.0326	0.0472	-0.0366	-0.0256	-0.0174
nu (NuCypher)	588	320	268	0.805	0.0291	0.0436	0.0621	-0.0543	-0.0376	-0.0239
ogn (Origin Protocol)	243	101	142	0.330	0.0230	0.0365	0.0506	-0.0427	-0.0310	-0.0216
omg (OmiseGO)	314	154	160	0.331	0.0235	0.0322	0.0460	-0.0393	-0.0275	-0.0175
orn (Orion Protocol)	397	236	161	0.670	0.0229	0.0321	0.0432	-0.0392	-0.0317	-0.0227
oxt (Orchid)	843	493	350	0.756	0.0220	0.0331	0.0476	-0.0426	-0.0281	-0.0198
perp (Perpetual Protocol)	316	173	143	0.584	0.0224	0.0297	0.0399	-0.0359	-0.0263	-0.0190
pla (PlayDapp Token)	474	214	260	0.750	0.0170	0.0294	0.0446	-0.0424	-0.0267	-0.0177
poly (Polymath)	690	343	347	1.098	0.0212	0.0304	0.0465	-0.0418	-0.0301	-0.0202
qnt (Quant)	225	116	109	0.320	0.0239	0.0313	0.0428	-0.0382	-0.0284	-0.0209
quick (QuickSwap)	664	336	328	1.035	0.0218	0.0315	0.0439	-0.0389	-0.0253	-0.0189
rad (Radicle)	380	223	157	0.946	0.0278	0.0402	0.0520	-0.0506	-0.0325	-0.0237
rai (Rai Reflex Index)	28	17	11	0.746	0.0005	0.0007	0.0018	-0.0024	-0.0013	-0.0005
rari (Rarible)	271	145	126	0.815	0.0328	0.0455	0.0664	-0.0598	-0.0442	-0.0309
ren (Ren)	281	120	161	0.300	0.0270	0.0386	0.0530	-0.0486	-0.0348	-0.0267
rep (Augur)	1048	532	516	0.874	0.0224	0.0320	0.0441	-0.0399	-0.0283	-0.0204
req (Request Network)	509	312	197	0.854	0.0227	0.0333	0.0514	-0.0408	-0.0271	-0.0198
rgt (Rari Governance Token)	399	208	191	1.153	0.0357	0.0493	0.0684	-0.0561	-0.0397	-0.0300
rlc (iExec RLC)	289	142	147	0.417	0.0258	0.0331	0.0441	-0.0419	-0.0313	-0.0208
rly (Rally)	337	185	152	0.757	0.0243	0.0410	0.0627	-0.0453	-0.0310	-0.0224
shib (Shiba Inu)	354	189	165	0.597	0.0200	0.0305	0.0456	-0.0354	-0.0252	-0.0182
skl (SKALE Network)	206	82	124	0.284	0.0277	0.0382	0.0498	-0.0451	-0.0350	-0.0262
snx (Synthetix Network Token)	212	97	115	0.242	0.0244	0.0338	0.0463	-0.0414	-0.0324	-0.0241
sol (Solana)	205	102	103	0.290	0.0200	0.0268	0.0338	-0.0390	-0.0288	-0.0203
storj (Storj)	257	106	151	0.370	0.0225	0.0349	0.0486	-0.0410	-0.0282	-0.0199
sushi (Sushi)	164	67	97	0.218	0.0242	0.0339	0.0481	-0.0417	-0.0303	-0.0246
trb (Tellor Tributes)	241	102	139	0.326	0.0287	0.0404	0.0505	-0.0455	-0.0351	-0.0237
tribe (Tribe)	455	235	220	1.111	0.0140	0.0204	0.0320	-0.0286	-0.0206	-0.0132
tru (TrueFi)	360	202	158	0.613	0.0304	0.0425	0.0551	-0.0515	-0.0380	-0.0275
uma (UMA)	694	400	294	0.724	0.0236	0.0348	0.0490	-0.0398	-0.0281	-0.0201
uni (Uniswap)	244	99	145	0.253	0.0217	0.0316	0.0424	-0.0383	-0.0290	-0.0214
wbtc (Wrapped Bitcoin)	509	237	272	0.543	0.0101	0.0141	0.0192	-0.0199	-0.0147	-0.0101
wcfcg (Wrapped Centrifuge)	249	144	105	1.082	0.0348	0.0478	0.0680	-0.0656	-0.0446	-0.0336
xlm (Stellar)	594	248	346	0.397	0.0156	0.0207	0.0276	-0.0286	-0.0191	-0.0144
xrp (Ripple)	380	203	177	0.592	0.0133	0.0186	0.0261	-0.0295	-0.0208	-0.0130
xtz (Tezos)	444	197	247	0.342	0.0205	0.0283	0.0399	-0.0333	-0.0260	-0.0180
xyo (XYO Network)	330	184	146	0.627	0.0273	0.0387	0.0523	-0.0416	-0.0307	-0.0219
yfi (DFI.money)	377	170	207	0.686	0.0220	0.0322	0.0492	-0.0448	-0.0312	-0.0211
yfi (yearn.finance)	253	95	158	0.260	0.0241	0.0331	0.0451	-0.0391	-0.0261	-0.0205
zec (Zcash)	261	92	169	0.294	0.0195	0.0270	0.0412	-0.0344	-0.0251	-0.0180
zen (Horizen)	221	85	136	0.371	0.0197	0.0264	0.0367	-0.0365	-0.0264	-0.0202
zrx (0x)	840	364	476	0.525	0.0188	0.0278	0.0431	-0.0392	-0.0255	-0.0183
Avg.100	420	206	214	0.550	0.0228	0.0320	0.0446	-0.0395	-0.0285	-0.0205

O. Appendix C. Sample construction and data filtering procedure

As noted in the introduction, cryptocurrency markets are relatively new and rapidly growing. We rely on intraday data that are long enough across multiple cryptocurrencies to meet our goal of analyzing return predictability. Because Kaiko provides reliable intraday data and because existing studies such as Makarov and Schoar (2020) employ intraday data from Kaiko, we collect data from Kaiko.

To construct an unbiased sample with the largest cross-section, we examine all cryptocurrencies that have available data longer than nine months and are traded on Coinbase, which is ranked as the top exchange in Kaiko’s overall evaluation (e.g., quality and popularity).⁴⁶ The minimum sample period is chosen because our variance measures require estimation horizons. Kaiko’s order book data provide intraday bid and ask quotes (and volumes) for 198 cryptocurrencies (as of 2022). We exclude stable coins (e.g., Tether) and cryptocurrencies with only a short sample period. Adopting simple coin selection criteria, our sample comprises 100 cryptocurrencies with various characteristics and includes a delisted coin.

To filter out 15-minute interval data, we follow two papers that study sovereign currency markets (i.e., Lee and Wang, 2019, 2020) because sovereign currency markets share many similar characteristics with cryptocurrency markets. For example, both cryptocurrency and sovereign currency markets trade continuously throughout the day. Specifically, we remove quotes that do not change for three consecutive time intervals because these quotes might be inactive. We perform this filtering process for bid and ask quotes and then construct mid

⁴⁶ An early version of this paper used the data of cryptocurrencies on Bitfinex and Bittrex and shows results consistent with those in this paper. To avoid concerns resulting from the different operating mechanisms of various exchanges, this paper employs data only from Coinbase because this exchange provides a wider cross section than the other two exchanges.

quotes (i.e., $\text{mid} = 0.5 \times (\text{bid} + \text{ask})$).

Using the mid quotes, we compute 15-minute log returns (i.e., $r_{i,t(j)} = c_{i,t(j)} - c_{i,t(j-1)}$), where $c_{i,t(j)}$ is the logarithm of the mid quote of the i -th cryptocurrency at time $t(j)$. Following Andersen et al. (2001a) and Lee (2012), who suggest potential problems resulting from bid-ask bounce effects, we filter out observations when large returns are canceled out by subsequent large returns with the opposite sign (i.e., the sum of two consecutive returns is close to zero, and the magnitude of each return is large). To mitigate the concern that small numbers of extreme observations drive the results, we remove extreme intraday returns if the absolute z -statistics are greater than seven. In addition, we exclude observations belonging to days that have fewer than 30 observations.

O. Appendix D. Comparison with alternative measures

Our positive jump variance measure is comparable to the relative signed jump (RSJ) variation measure of Bollerslev, Li and Zhao (2020), which is defined as the difference between the up and down semivariances divided by the total return variation. These authors find a clear negative relation between the RSJ measure and the average future returns in stock markets. Therefore, the return predictability of positive jump variances may appear consistent with that of the RSJ measure. However, using the RSJ measure, we have difficulty distinguishing whether the significant negative relation results from positive jump or negative jump variation (because the RSJ measure combines these two jump variations). Using our separated positive and negative jump variance measures, we can make the distinction, which can provide the additional benefit of identifying what drives the effects.

Instead of considering the relative differences in semivariances, one can still consider signed jump variations based on signed semivariances by subtracting half of the bipower variation. In this appendix, we confirm the role of positive jump variances in predicting cryptocurrency returns using this alternative measure and compare our findings. To this end, in Equation (6), we replace positive and negative jump variances with positive and negative semivariances minus half of the bipower variation.

We report the results in Table OD. As column (I) shows, the alternative positive jump variances also provide significantly negative coefficients with the control of jump-robust and alternative negative jump variances. In column (II), the return predictability of the alternative positive jump variance measures is robust to controlling for lagged returns and market capitalization. To compare our positive jump variance measure with the alternative mea-

asures, we perform a horse race regression in columns (III) and (IV). Interestingly, our positive jump variances have significantly negative coefficients, while the alternative measures lose significance. This finding can occur because our measures are estimated with only positive jumps and capture large price changes more effectively than do the alternative measures.

Table OD

Comparison of positive jump variances with alternative jump variances.

This table shows the return predictability with alternative jump variance measures, which is compared with that with positive jump variances. We employ the following FMB regression:

$$rx_{i,w+1} = \gamma_{0,w} + \gamma_{1,w}JV_{i,w}^{(+)} + \gamma_{2,w}JV_{i,w}^{(-)} + \gamma_{3,w}JRV_{i,w} + c'_wX_{i,w} + \epsilon_{i,w+1},$$

where $rx_{i,w}$ is the excess return of cryptocurrency i in week w . $JV_{i,w}^{(+)}$, $JV_{i,w}^{(-)}$, and $JRV_{i,w}$ are the positive jump, negative jump, and jump-robust variances, respectively. These decomposed variances are estimated from the previous month of observations (i.e., observations from week $w - 3$ to week w). The alternative positive (negative) jump variance is defined as the positive (negative) semivariance minus half of the bipower variation (Bollerslev, Li and Zhao, 2020). We first use the alternative jump variances in the regressions and then employ our jump variance measures as defined in Section 2. $X_{i,w}$ is the vector of control variables, such as lagged excess returns and natural logarithmic market capitalization. In this table, we report the time-series averages of the estimated coefficients and the corresponding t -statistics. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(I)	(II)	(III)	(IV)
Constant	0.013	-0.045*	0.009	-0.039
t -stat	1.60	-1.77	1.07	-1.57
Positive variance - 0.5BV	-1.271**	-1.526**	0.690	0.660
t -stat	-4.15	-4.15	0.95	0.92
Negative variance - 0.5BV	-0.322	0.118	0.502	1.402**
t -stat	-0.74	0.26	0.82	2.28
Jump-robust variance	0.131	0.209*	-0.372*	-0.463**
t -stat	1.15	1.71	-1.85	-2.10
Positive jump variance			-1.741***	-1.744***
t -stat			-2.83	-2.98
Negative jump variance			0.612	0.002
t -stat			0.87	0.00
Lagged return		0.018		0.010
t -stat		1.04		0.60
Market capitalization		0.002***		0.002**
t -stat		2.69		2.32
Adj. R^2	0.197	0.262	0.289	0.337

O. Appendix E. Relationship between characteristics and prices

Cryptocurrencies with higher total and positive jump variances have lower returns in the subsequent week. The characteristics that drive this result can be associated with contemporaneous price increases. To investigate this possibility, we contemporaneously regress excess returns or our variance measures on characteristic variables. To be consistent with Section 5, we use (log) market capitalization, (log) daily trading volumes, (log) futures volumes, bid-ask spreads (BASs), numbers of buy opinions on Twitter, and retail trading proportions (RTPs). These characteristic variables are estimated with the previous month of observations at the end of every week.

As Table OE shows, the prices of cryptocurrencies with low market capitalization and trading volumes tend to increase before they decrease in subsequent weeks. Because positive jumps capture large price increases, regressions that use positive jump variances as dependent variables provide results similar to those using excess returns. In addition, total and positive jump variances provide similar relationships. Interestingly, total and positive jump variances are positively related to futures volumes, which confirms that such cryptocurrencies do not relatively suffer from short-selling constraints.

Table OE

Contemporaneous relationship between characteristics and price increases.

This table shows the results of regressing excess returns, total variances, or positive jump variances on contemporaneous market capitalization, daily trading volumes, futures volumes, bid-ask spreads (BASs), numbers of buy opinions on Twitter, and retail trading proportions (RTPs). Characteristic variables are estimated with the previous month of observations at the end of every week. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Excess returns		Total variances		Pos. jump variances	
	(I)	(II)	(III)	(IV)	(V)	(VI)
Constant	-0.143***	0.092	-0.092***	-0.795***	0.003	-0.014**
<i>t</i> -stat	-5.51	1.20	-3.59	-11.23	1.11	-2.48
Log market capitalization	-0.011***	-0.037***	-0.048***	-0.035***	-0.002***	-0.003***
<i>t</i> -stat	-6.37	-10.79	-20.00	-9.25	-14.19	-8.89
Log trading volumes	0.016***	0.021***	0.049***	0.076***	0.002***	0.003***
<i>t</i> -stat	9.09	6.85	19.38	17.83	13.14	12.90
Log futures volumes	-0.001	0.009***	0.007***	0.001***	0.0005***	0.0003*
<i>t</i> -stat	-1.15	5.03	11.79	0.66	9.99	1.92
BAS	-8.969***	-9.201***	17.907***	22.421***	-0.231**	-0.254***
<i>t</i> -stat	-6.17	-5.18	15.15	21.15	-2.40	-2.88
Log Twitter buy	0.005***	0.000	-0.007***	-0.001	0.000	0.001***
<i>t</i> -stat	2.87	0.17	-4.30	-0.59	0.01	4.59
RTP	0.006	0.016*	0.004	-0.010	0.000	0.000
<i>t</i> -stat	1.00	1.86	0.44	-1.00	0.42	0.32
Year/coin fixed effects	No	Yes	No	Yes	No	Yes
Adj. R^2	0.082	0.191	0.289	0.592	0.129	0.347

O. Appendix F. Variable definition

In this appendix, we provide the definitions of the variables considered in our analyses.

Given the notation established in Section 2, we follow Amaya et al. (2015) for the definitions of realized moments that can be estimated with intraday return data. We obtain the realized variance as Equation (5), which defines the total variance. The realized volatility is computed as the square root of the realized variance as follows:

$$Vol_{i,w} = \sqrt{Var_{i,w}}. \quad (11)$$

We also consider ex post realized higher moments. The realized skewness is again based on cubed intraday returns standardized by the realized variance as follows:

$$Skew_{i,w} = \frac{\sqrt{N} \sum_{t(j) \in W_w} r_{i,t(j)}^3}{Var_{i,w}^{3/2}}, \quad (12)$$

where N is the total number of observations over the estimation window for week w . When this realized skewness measure is negative (positive), we can interpret the asset return distribution as having a left (right) tail that is fatter than the right (left) tail. Similarly, we consider the realized kurtosis defined as follows:

$$Kurt_{i,w} = \frac{N \sum_{t(j) \in W_w} r_{i,t(j)}^4}{Var_{i,w}^2}. \quad (13)$$

This measure is expected to capture the extremes of the return distributions.

For individual cryptocurrencies, using the notation in Sections 2 and 3, we compute

(percentage) BASs defined as follows:

$$BAS_{i,w} = \frac{1}{N} \sum_{t(j) \in W_w} \frac{bid_{i,t(j)} - ask_{i,t(j)}}{mid_{i,t(j)}}, \quad (14)$$

where $bid_{i,t(j)}$ ($ask_{i,t(j)}$) represents the bid (ask) quote of cryptocurrency i at time $t(j)$. $mid_{i,t(j)}$ is the mid quote (i.e., $0.5 \times (bid_{i,t(j)} + ask_{i,t(j)})$) and N is the total number of observations over the estimation window for week w .

To consider market conditions for robustness checks, we employ the overall volatility and illiquidity measures in cryptocurrency markets. We compute the cryptocurrency market volatility ($CVol$) by following the approach of calculating the global foreign exchange volatility of Menkhoff et al. (2012). Specifically, the cryptocurrency market volatility in week w is defined as

$$CVol_w = \frac{1}{N} \sum_{t(j) \in W_w} \sum_i^{\kappa_j} \frac{|r_{i,t(j)}|}{\kappa_j}, \quad (15)$$

where κ_j is the number of available cryptocurrencies at time $t(j)$. For the illiquidity measure, we modify the approach of Amihud (2002) because we use intraday data for weekly analyses, while the author uses daily data for monthly analyses. We first estimate the weekly illiquidity for individual cryptocurrencies as follows:

$$ILLIQ_{i,w} = \frac{1}{N} \sum_{t(j) \in W_w} \frac{|r_{i,t(j)}|}{V_{i,t(j)}}, \quad (16)$$

where $V_{i,t(j)}$ is the dollar volume of cryptocurrency i from time $t(j-1)$ to $t(j)$. Then, we take the average of individual illiquidity measures across available cryptocurrencies.

To examine retail investors' transactions, we use the retail trading proportion (RTP) of

Han and Kumar (2013). Using 5-minute trading volumes, we define the threshold of dollar transaction volumes for retail trading activities. Specifically, we use the 90th percentile of all cryptocurrencies' 5-minute trading volumes, which is 22 million dollars. If transaction volumes are lower than the bar, the transaction is classified as retail trading. Our threshold allows approximately 90% of total trading volume to be retail trading volume because 90-95% of Coinbase revenues originate from retail investors. We employ the previous month of trading volume data to compute weekly RTP measures.