What Influences Bitcoin Implied Volatility?

Hai Lin* **Mike Zhou†** **Yanlin Jiang‡**

September 19, 2024

*School of Economics and Finance, Victoria University of Wellington, New Zealand. Email: hai.lin@vuw.ac.nz.

†School of Economics and Finance, Victoria University of Wellington, New Zealand. Email: mike.zhou@vuw.ac.nz.

‡Risk Management Institute, National University of Singapore, Singapore. Email: E1204741@u.nus.edu.

What Influences Bitcoin Implied Volatility?

Abstract

This study uses Bitcoin options data to create a model-free Bitcoin implied volatility index (BVIX) and investigates the factors that influence it. The research shows that investor attention is the most significant predictor of BVIX, while investor sentiment, interest rates, and trading volume also have predictive power. These findings are both statistically and economically significant. Additionally, the study suggests that these predictors affect call and put options differently, and non-fundamental sentiment is more important in influencing Bitcoin implied volatility than fundamental sentiment.

JEL classification: G12; G13; G17;

Keywords: Bitcoin; cryptocurrency markets; implied volatility; options.

1 Introduction

Cryptocurrency has become increasingly popular worldwide in recent years, sparking two main opposing perspectives. Some argue that it is a fraudulent scheme disguised as blockchain technology, essentially a bubble that will eventually burst and become worthless. However, many believe that cryptocurrency represents a technological innovation and will be a significant form of decentralized finance (DeFi). Regardless of which view is correct, it is undeniable that cryptocurrency, especially Bitcoin as the flagship, has been playing an increasingly important role in the global economy. In February 2021, Bitcoin surpassed the $1 trillion mark in market capitalization for the first time, [1] just over a decade since its invention in 2008 by Nakamoto (2008). Such rapid growth in cryptocurrency brings significant market volatility, which has led to short-term high returns and substantial losses.[2] This relationship between high volatility and extreme positive expected returns is described as lottery features in Lee and Wang (2024).

Due to the significant growth of the cryptocurrency market, there has been growing research on analyzing their returns. See, for example, Makarov and Schoar (2020), Liu and Tsyvinski (2021), Liu et al. (2022), Biais et al. (2023), Sockin and Xiong (2023), and Lee and Wang (2024). Meanwhile, the research on cryptocurrency market volatility is limited. Griffin and Shams (2020) investigate the driving forces behind the growth of Tether, the most commonly-used stable cryptocurrency pegged to the U.S. dollar, during the 2017 boom of cryptocurrency markets. They provide evidence that Tether is issued to purchase Bitcoin when its price falls, thus inflating the market. This finding reveals a significant deviation from fundamental prices and raises concerns about potential substantial negative impacts on prices when these distortions are unwound in the cryptocurrency markets. Pagnotta (2022)

[1] https://markets.businessinsider.com/currencies/news/bitcoin-market-value-1-trillion-for-the-first-time-cryptocurrency-2021-2-1030102808

[2] For instance, a man from Korea became the nation's youngest Rolls-Royce owner at 29 years old through Bitcoin trading. However, he lost most of the money shortly thereafter.https://www.nytimes.com/2019/02/10/business/south-korea-bitcoin-cryptocurrencies.html

and Biais et al. (2023) establish equilibrium models of Bitcoin and find that cryptocurrency prices can fluctuate dramatically even when no change occurs in the fundamentals. Instead of analyzing the equilibrium states, Makarov and Schoar (2021) analyze Bitcoin network in detail and document that large and concentrated participants still dominate the Bitcoin eco-system and this concentration increases systematic risk in the market. Consequently, these findings highlight the critical role of understanding and predicting market volatility, as it directly impacts the effectiveness of risk management, investment strategies, and hedging with cryptocurrency.

In this paper, we construct a model-free Bitcoin implied volatility index and study its influencing factors. We construct an implied volatility index of the Bitcoin market, which we call BVIX for short, to measure the level of expected volatility. Compared with the realized variance that sums squared intraday returns, implied volatility derived from options prices tends to subsume the information contained in past realized volatility and has superior forecasting performance in future volatility (Christensen and Prabhala, 1998; Jiang and Tian, 2005). In the Bitcoin market, participants value more about what happens in the future than in the past. Thus, implied volatility is a better measure than realized volatility. We apply a model-free methodology similar to that adopted by the CBOE for the VIX index instead of using implied volatility derived from the Black-Scholes formula (Black and Scholes, 1973) because the former tends to encapsulate all the information included in the latter (Jiang and Tian, 2005). We use historical data of Bitcoin options from LedgerX, a leading cryptocurrency derivatives exchange regulated by the U.S. Commodities and Futures Trading Commission (CFTC). The data used is different from Alexander and Imeraj (2023), who obtain options data from Deribit. LedgerX and Deribit are prominent cryptocurrency derivative exchanges that serve different users. LedgerX focuses more on regulatory compliance and institutional investors, while Deribit provides a wider range of products with a more global reach. The trading volume of LedgerX is typically lower than Deribit due to its focus on the U.S. market and regulatory compliance. However, its concentration among institutional investors and sophisticated individual traders makes its prices more stable and reflective of available information. Fig. 1 plots the BVIX from 1 January 2020 to 30 April 2024. The Bitcoin market exhibits significant volatility, particularly from early 2021 to 2022 when the Delta and Omicron variants were prevalent (Christensen et al., 2022; Elliott et al., 2022).

We next examine the factors influencing Bitcoin implied volatility. Liu and Tsyvinski (2021) and Liu et al. (2022) identify investor attention as a significant factor in capturing cryptocurrency returns. We extend their studies by investigating the predictive power of investor attention in the volatility of Bitcoin. Investor sentiment is another potential predictor as Pagnotta (2022) documents that sentiment shifts amplify Bitcoin volatility. Lopez-Lira and Tang (2023) and Chen et al. (2023) demonstrate ChatGPT's superior performance compared to traditional sentiment analysis methods. We follow their studies to utilize ChatGPT to analyze relevant news and construct a Bitcoin market sentiment measure. Since both Griffin and Shams (2020) and Biais et al. (2023) argue that Bitcoin volatility is primarily influenced by non-fundamental news, we also separate daily news into fundamental news and non-fundamental news based on the criterion in Biais et al. (2023) to examine which part has a more significant influence.

Additionally, we include several predictors from different dimensions. Mixon (2002) states that short-term interest rates can influence implied volatility in the stock market due to a leverage effect. Vähämaa and Äijö (2011) and Amengual and Xiu (2018) show that Federal Open Market Committee (FOMC) meetings significantly influence the behavior of the implied volatility of index options through the uncertainty channel. Because the FOMC is closely linked to and responsible for setting interest rates, it suggests a potential influence of interest rates on implied volatility. Trading volume is another predictor we use. Literature has documented that it influences stock market volatility (Harris and Raviv, 1993; Chan and Fong, 2000; Avramov et al., 2006). To account for the information spillover from the stock and commodity market, we include the VIX index, and gold returns to represent stock market implied volatility and commodity market performance, respectively. Finally, we incorporate the first lag of BVIX as a control variable to account for autocorrelation. We assess how these variables influence Bitcoin implied volatility. Lee and Wang (2024) emphasize the importance of frequently capturing the dynamic nature of cryptocurrency market volatility. Accordingly, we collect and calculate data daily rather than weekly or monthly.

We find that investor attention, investor sentiment, interest rate, and trading volume have significant predictive power on next-day implied volatility, while the VIX and gold returns are insignificant. The multiple regression incorporating all the predictors exhibits the best performance with the highest adjusted R^2 of 53.31%. We then employ the method proposed by Rapach et al. (2010) to compare the out-of-sample performance of each predictive model relative to a benchmark model. Since the implied volatility index is highly autocorrelated, we use the AR(1) model as the benchmark. The results are consistent with the in-sample regression models, indicating significant predictive power in both in-sample and out-of-sample contexts. In addition to the multiple regression, we calculate various combination forecasts following Rapach et al. (2010). The results show that all combination forecasts achieve significantly better forecasting performance than the benchmark.

To test whether the predictability of Bitcoin implied volatility has economic significance, we follow Bali et al. (2023) to construct option trading strategies based on the forecasts of each predictive model and compare the option portfolios' performance with that of the benchmark AR(1) model. The trading strategies involve grouping options based on expected returns and creating a zero-cost long-short portfolio weighted by open interest. Effective spreads are used to reflect transaction costs because Ofek et al. (2004) illustrate that transaction costs can considerably reduce economic profits in options markets. Using average annualized returns as the performance measure, we find that the predictive models incorporating investor attention or interest rate, along with the multiple regression and combination forecasts, generate significant excess returns over the benchmark when no transaction costs are assumed. The multiple regression achieves the highest annualized excess return and daily standard deviation of 36.50% and 1.11%, respectively, with a t-statistic of 3.43. When transaction costs are included, the annualized excess returns increase as the costs widen in most cases, suggesting better performance relative to the benchmark. For instance, when the effective spread increases from 0 to 100%, the annualized excess returns from the predictive model incorporating interest rate rise from 22.48% to 55.97%, with a t-statistic from 1.99 to 3.16. The predictive models incorporating investor sentiment, the VIX, or trading volume can outperform the benchmark, but their excess returns remain insignificant across all transaction cost levels.

To examine whether risk factors can explain the excess returns, we conduct time-series regressions of excess returns on the cryptocurrency risk factors identified in Liu et al. (2022). We employ a one-factor model incorporating the market factor and a three-factor model including the market, size, and momentum factors. The trading strategies yield positive alphas in most cases. However, the corresponding significance decreases under the three-factor model, suggesting that a proportion of the excess returns can be attributed to the influence of size and momentum factors.

We further study whether the forecast power of predictors is different between calls and puts. We assess the performance of portfolios constructed by trading exclusively in either call or put options. The results suggest that investor sentiment and attention exhibit more substantial predictive power for call options, while interest rates are more effective for put options. Additionally, when we categorize Bitcoin news into fundamental and non-fundamental based on the criteria outlined by Biais et al. (2023), we observe that fundamental news generally generates positive sentiment. In contrast, non-fundamental news shows a more significant fluctuation. Furthermore, by conducting bivariate and multiple regressions, we find that the influence of investor sentiment on implied volatility is more significantly driven by factors associated with non-fundamental news rather than fundamental news.

Our study contributes to the literature in several ways. Makarov and Schoar (2020), Liu and Tsyvinski (2021), Liu et al. (2022), Biais et al. (2023), Sockin and Xiong (2023), and Lee and Wang (2024) analyze cryptocurrency returns. We extend these studies to volatility, another important aspect of asset prices. Our finding underscores the importance of investor attention in cryptocurrency markets as Liu and Tsyvinski (2021) find that it is also an important factor in predicting returns.

Pagnotta (2022) document that the blockchain design leads to Bitcoin's extreme volatility, and the price can exhibit high fluctuations even when fundamentals are constant. Biais et al. (2023) show that the fundamental price of Bitcoin is based on its future expected net transaction benefits. These fundamentals are priced in equilibrium, but the price variation is largely attributed to extrinsic volatility driven by sunspots. Our work complements these papers by empirically examining factors influencing Bitcoin volatility. We find investor sentiment presents a significant predictive power when controlled for investor attention, suggesting a different channel of influence on volatility. Moreover, when we separate the sentiment into fundamental sentiment and non-fundamental sentiment, we find non-fundamental sentiment is more significant than fundamental sentiment, providing empirical support to Pagnotta (2022) and Biais et al. (2023).

Liu and Tsyvinski (2021) show that macroeconomic factors do not have a significant relationship with cryptocurrency returns. We document a different finding in volatility and find that interest rate is an effective predictor for cryptocurrency market implied volatility. This finding extends our understanding of how macroeconomic variables affect the cryptocurrency market. Consistent with the findings in stock market (Harris and Raviv, 1993; Chan and Fong, 2000; Avramov et al., 2006), we find that trading volume also plays a role in influencing cryptocurrency market volatility.

The rest of the paper is organized as follows. Section 2 discusses our empirical methodology, including the calculation of BVIX, the construction of predictive models, the out-of-sample performance measurement, and the option trading strategies used to examine economic significance. Section 3 explains the data used in our empirical study. Section 4 reports the primary empirical results. Section 5 provides additional tests. Section 6 concludes this paper.

2 Methodology

This section outlines the methodology for calculating BVIX, constructing predictive models for Bitcoin implied volatility, measuring the out-of-sample performance, and building option trading strategies to assess economic significance.

2.1 BVIX

To construct the implied volatility in the Bitcoin market, we follow the methodology employed by Cboe Global Markets (2022). They calculate the VIX, which represents the market's expectation of 30-day volatility implied by S&P 500® index options prices and has become the premier benchmark for the U.S. stock market volatility, often referred to as the "fear gauge". According to Cboe Global Markets (2022), a vital feature of the Cboe VIX index is that constituent options are weighted inversely proportional to the square of their strike (K^2). This weighting scheme aligns with replicating variance swap payoffs with option portfolios (Carr and Wu, 2009).

Like VIX, we construct an index representing the Bitcoin market's expectation of 30-day volatility implied by Bitcoin option prices. The process involves four steps. The first step is to define the near- and next-term Bitcoin options. In the stock options market, near-term options are the options with time-to-maturity less than and closest to 30 days, and the next-term options are those with maturity longer than and closest to 30 days. Compared with the stock market, options in the Bitcoin market are relatively limited. To best use Bitcoin options data, we categorize near-term options as those with a time-to-maturity of less than or equal to 30 days. Similarly, next-term options are defined as those with a time-to-maturity greater than 30 days.

The second step is to obtain the interest rate. We follow Gürkaynak et al. (2007) to obtain the values:

$$r_T = \beta_0 + \beta_1 \frac{1-\exp\left(-\frac{T}{\tau_1}\right)}{\frac{T}{\tau_1}} + \beta_2 \left[\frac{1-\exp\left(-\frac{T}{\tau_1}\right)}{\frac{T}{\tau_1}} - \exp\left(-\frac{T}{\tau_1}\right)\right] + \beta_3 \left[\frac{1-\exp\left(-\frac{T}{\tau_2}\right)}{\frac{T}{\tau_2}} - \exp\left(-\frac{T}{\tau_2}\right)\right], \quad (1)$$

where $\beta_0, \beta_1, \beta_2, \beta_3$ and τ_1, τ_2 are the corresponding parameters obtained from the U.S. Treasury Yield Curve. T represents the time to maturity.

The third step is to calculate the volatility of near- and next-term using the options data defined in the first step. Specifically, we calculate the implied volatility of near- or next-term using the following formula, as outlined by Demeterfi et al. (1999):

$$\sigma_T^2 = \frac{2}{T} \sum_i \frac{\Delta K_i}{K_i^2} e^{r_T T} Q(K_i) - \frac{1}{T}\left[\frac{F}{K_0} - 1\right]^2, \quad (2)$$

where σ_T is the implied volatility of maturity T (near-term or next-term),[3] r_T is the interest rate of maturity T derived from Eq. (1), F is the option-implied forward price derived from at-the-money (ATM) option prices, and K_0 is the first strike price equal to or immediately below the F. K_i is the strike price for the i_{th} out-of-the-money (OTM) option, and ΔK_i is the strike interval defined as $K_i - K_{i-1}$ for the highest strike, $K_{i+1} - K_i$ for the lowest strike, and $(K_{i+1} - K_{i-1})/2$ otherwise. $Q(K_i)$ in the formula represents the midpoint of the bid and ask prices for the option with strike K_i. We calculate the implied volatilities for both near-term and next-term.

In the final step, we linearly interpolate the volatilities of near- and next-term to calculate the implied volatility of a constant maturity of 30 days,

$$IV_{30} = \sqrt{\left\{T_1 \sigma_{T_1}^2 \frac{T_2 - 30}{T_2 - T_1} + T_2 \sigma_{T_2}^2 \frac{30 - T_1}{T_2 - T_1}\right\} \times \frac{365}{30}}, \quad (3)$$

where T_1 and T_2 are the near- and next-term maturity (in days), respectively. We multiply

[3]When calculating σ_T^2, we assume all near-term options have the same maturity as the one less than but closest to 30 days. Similarly, we assume all next-term options have the same maturity as the one longer than but closest to 30 days.

IV_{30} by 100 to yield the BVIX index.

2.2 Predictive models

2.2.1 In-sample predictive regressions

We focus on a daily horizon to better capture dynamic nature of Bitcoin market volatility. Specifically, we regress the BVIX of given day t on the predictors from $t-1$. We use the AR(1) model as the benchmark because it only utilizes the information from its own past.

$$BVIX_t = \alpha_0 + \beta_0 BVIX_{t-1} + \varepsilon_{i,t}. \tag{4}$$

To examine whether incorporating another predictor brings incremental predictive power, we run the extended regression incorporating the AR(1) component and one additional predictor (X^i_{t-1}),

$$BVIX_t = \alpha_i + \beta_0 BVIX_{t-1} + \beta_i X^i_{t-1} + \varepsilon_{i,t}, i = 1, ..., N, \tag{5}$$

where X^i_{t-1} is the i-th predictor and N is the number of predictors used in the analysis. We then run a multiple regression including all predictors:

$$BVIX_t = \alpha_{all} + \beta_0 BVIX_{t-1} + \sum_{i=1}^{N} \beta_i X^i_{t-1} + \varepsilon_{all,t}. \tag{6}$$

2.2.2 Out-of-sample forecast

Literature has documented that in-sample predictability can be due to overfitting, and the out-of-sample test is more robust (Welch and Goyal, 2008; Lin et al., 2018). To run the out-of-sample test, we employ an extensive time window and use all available data until day t to run the predictive regressions and generate the forecast for $t+1$. Since we only use the information until day t, the forecast for $t+1$ is out-of-sample. Different predictive regressions will have different forecasts of BVIX. If one predictor contains valuable information, the predictive regressions incorporating it will forecast BVIX better than the benchmark AR(1) model.

In addition to the multiple predictive regression incorporating all predictors into one regression, we also use a combination forecast for the out-of-sample analysis. Rapach et al. (2010) document that combination provides a better forecast than multiple predictive regression. The combination forecast at time $t+1$ is weighted averages of the N individual forecasts generated from the extended models in Eq. (5):

$$B\hat{V}IX_{c,t+1} = \sum_{i=1}^{N} \omega_{i,t} B\hat{V}IX_{i,t+1}, \tag{7}$$

where $B\hat{V}IX_{i,t+1}$ is the individual forecast using AR(1) and the i-th predictor, and $\{\omega_{j,t}\}_{j=1}^{N}$ are the weights used to combine the individual forecasts.

We use different ways of combination by choosing different $\{\omega_{i,t}\}_{j=1}^{N}$. The first class of methods implements mean, median, and trimmed mean as averaging schemes. The mean or average combination forecast is calculated by assigning equal weights to each forecast, that is, $\omega_{i,t} = 1/N$. The median forecast is the median value of $\left\{B\hat{V}IX_{i,t+1}\right\}_{i=1}^{N}$, and the trimmed mean forecast is the average of the forecast series excluding the smallest and largest values, with the weights assigned to the remaining forecasts being equal to $1/(N-2)$. The second class uses the discount mean squared error (DMSE) over the holdout out-of-sample period (q_0 until $t-1$) to calculate the weights:

$$\omega_{i,t} = \frac{\phi_{i,t}^{-1}}{\sum \phi_{i,t}^{-1}}, i = 1, ..., N, \tag{8}$$

where $\phi_{i,t}^{-1} = \sum_{s=q_0}^{t-1} \theta^{t-1-s}(BVIX_{s+1} - B\hat{V}IX_{i,s+1})^2$, and θ is a discount factor. Intuitively, the DMSE method puts more weight on the forecast with a smaller error over the holdout period. When θ equals 1.0, there is no discounting. We consider setting θ to 1.0, 0.95, and

0.9, respectively.

2.3 Out-of-sample evaluation

To evaluate the the out-of-sample performance of predictive models, we calculate the following out-of-sample R^2 statistic, R^2_{OS}, developed by Campbell and Thompson (2008):

$$R^2_{OS} = 1 - \frac{\sum_{k=1}^{k=M}(BVIX_{t+k} - B\hat{V}IX_{t+k})^2}{\sum_{k=1}^{k=M}(BVIX_{t+k} - B\bar{V}IX_{t+k})^2}, \tag{9}$$

where $BVIX_{t+k}$ is the observed implied volatility, $B\hat{V}IX_{t+k}$ is the forecasted implied volatility from a predictive model, and $B\bar{V}IX_{t+k}$ is from the benchmark AR(1) model. M is the number of observations in the out-of-sample period. The R^2_{OS} statistic measures the reduction in mean squared error for the predictive regressions or combination forecasts relative to the benchmark. Thus, a positive R^2_{OS} indicates a better forecasting performance than the benchmark.

To test the significance of R^2_{OS}, we employ the MSPE-adjusted statistic developed by Clark and West (2007), which is a one-sided test for the null hypothesis that the squared forecasting errors generated by either the predictive regressions or the combination forecasts and the benchmark are equal, against the alternative hypothesis that they have lower squared forecasting errors than the benchmark. The MSPE-adjusted statistic is defined as follows:

$$f_{t+1} = (BVIX_{t+1} - B\bar{V}IX_{t+1})^2 - \left[(BVIX_{t+1} - B\hat{V}IX_{t+1})^2 - (B\bar{V}IX_{t+1} - B\hat{V}IX_{t+1})^2\right]. \tag{10}$$

By regressing $\{f_{t+k}\}_{k=1}^{k=M}$ on a constant, we use the p-value for the one-sided (upper tail) test to evaluate the significance of R^2_{OS} statistic.

We also apply encompassing test to examine whether incorporating an additional factor brings additional information and hence improves the forecast power. We follow the method proposed by Harvey et al. (1998), which is a one-sided test for the null hypothesis that the forecasts from model i encompass the forecasts from model j against the alternative hypothesis that the forecasts from model i do not encompass the forecasts from model j. Denote $d_{t+1} = (\hat{u}_{i,t+1} - \hat{u}_{j,t+1})\hat{u}_{i,t+1}$, where $\hat{u}_{i,t+1} = BVIX_{t+1} - B\hat{V}IX_{i,t+1}$ and $\hat{u}_{j,t+1} = BVIX_{t+1} - B\hat{V}IX_{j,t+1}$. The MHLN statistic is expressed as:

$$MHLN = \left[\frac{M-1}{M}\right]\left[\hat{V}(\bar{d})^{-1/2}\right]\bar{d} \tag{11}$$

where $\bar{d} = (1/M)\sum_{k=1}^{k=M} d_{t+k}$, $\hat{V}(\bar{d}) = M^{-1}\hat{\phi}_0$ and $\hat{\phi}_0 = M^{-1}\sum_{k=1}^{k=M}(d_{t+k} - \bar{d})^2$. The MHLN statistic follows a t_{M-1} distribution.

2.4 Economic significance

Statistical significance can differ from economic significance (Welch and Goyal, 2008; Thornton and Valente, 2012). We follow Bali et al. (2023) to develop option trading strategies based on out-of-sample forecasts to test whether the predictability of Bitcoin implied volatility has economic value. Specifically, on a given trading day t, we utilize a predictive model to calculate the implied volatility at day $t+1$ and compute the theoretical prices of each available option in the option pool by applying the Black-Scholes option pricing formula (Black and Scholes, 1973). Since the underlying asset price and interest rate at $t+1$ are unknown as of time t, we use their values at time t as tomorrow's forecasts (Goncalves and Guidolin, 2006).

We then sort the options into five groups based on the percentage deviation of the theoretical price at $t+1$ from the mid-price on day t. We construct a zero-cost long-short portfolio by buying the most undervalued options (top group) and selling the most overvalued options (bottom group). We calculate the option returns defined by Cao and Han (2013). Assume a call option is hedged discretely N times during a period $[t, t+\Delta t]$, where the hedged is rebalanced at the end of each trading day. The discrete delta-hedged return of a call option is then calculated as follows:

$$\pi_{t,t+\Delta t} = C_{t+\Delta t} - C_t - \sum_{n=0}^{N-1}\delta_{c,t_n}(S(t_{n+1}) - S(t_n)) - \sum_{n=0}^{N-1}\frac{a_n r_{t_n}}{365}(C(t_n) - \delta_{c,t_n}S(t_n)), \tag{12}$$

and,

$$r_{t,t+\Delta_t} = \frac{\pi_{t,t+\Delta t}}{|\delta_{c,t}S_t - C_t|}, \tag{13}$$

where $\pi_{t,t+\Delta t}$ is the dollar return and $r_{t,t+\Delta_t}$ is the delta-hedged call option return; C_t and $S(t)$

denote the call option price and the underlying asset price at time t; r_{t_n} is the risk-free rate at time t_n; a_n is the number of calendar days and equals to 1; δ_{c,t_n} is the observed delta of a call option at t_n calculated from the Black-Scholes formula. Similarly, we calculate the delta-hedged put option return by replacing call option price and delta with put option price and delta in Eq. (12) and Eq. (13). The portfolio's return is the average of options returns, weighted by their open interests on day t. We construct the portfolio each trading day and hold it for one week. The return from an AR(1) model is used as the benchmark for comparison.

3 Data

3.1 LedgerX

Founded in 2014, LedgerX is a leading cryptocurrency derivatives exchange regulated by the US Commodities and Futures Trading Commission (CFTC). The CFTC granted them the initial licenses in 2017. As one of the first cryptocurrency derivatives exchanges to receive CFTC approval, LedgerX holds three CFTC registrations: Designated Contract Market (DCM), Derivatives Clearing Organization (DCO), and Swap Execution Facility (SEF). It is available to individual and institutional investors and offers options and swaps on Bitcoin with minimum increments of 0.01 units and Ethereum with minimum increments of 0.1 units. According to its official website, LedgerX has recorded trading of over 12 million contracts since its launch, representing more than 120,000 Bitcoins and having a notional value exceeding $1.1 billion.

We obtain historical options data from 1 January 2020 to 30 Apr 2024, via the LedgerX database that contains all the options records on trading days. We download 172,166 records in total. Each record includes a contract name, contract type, last bid price, last ask price, and open interest. For convenience, we standardize each contract's name as *"Underlying-Maturity-Strike-Type"*. For example, "BTC-28Jun2024-110000-c" refers to a Bitcoin call option that expires on 28 June 2024, with a strike price of $110,000. We present a sample of the option records in Appendix A. Table 1 reports the absolute numbers and percentages of call and put records. The upper part of the table indicates that call options slightly outnumber put options across all categories of time to maturity.

The lower part reveals that put options exceed call options when the moneyness exceeds 1.15.

[Insert Table 1 about here]

We use the historical Bitcoin options data downloaded from LedgerX to calculate the BVIX index. Fig. 1 plots the BVIX from 1 January 2020 to 30 Apr 2024. The BVIX exhibits significant volatility during specific periods, particularly from early 2021 to early 2022 when the Delta and Omicron variants were prevalent (Christensen et al., 2022; Elliott et al., 2022).

[Insert Figure 1 about here]

3.2 Predictors

We study various predictors that can affect the implied volatility of Bitcoin. The predictors include:

1. Investor attention. Liu and Tsyvinski (2021) and Liu et al. (2022) demonstrate that investor attention plays a vital role in capturing cryptocurrency market returns. We examine its influence on implied volatility. We use the logarithmic value of the daily search volume of Bitcoin on Wikishark as a proxy for investor attention, denoted as $Attn_t$.
2. Investor sentiment. Pagnotta (2022) shows that sentiment shifts amplify Bitcoin volatility. Following Lopez-Lira and Tang (2023) and Chen et al. (2023), we apply ChatGPT to analyze Bitcoin-related news and construct a GPT sentiment index ($Sent_t$), given its demonstrated superiority over traditional sentiment analysis methods. We download daily news of Bitcoin from 1 January 2020 to 30 April 2024, via *Dow Jone Factiva*, a global news platform with over 33,000 sources operated by Dow Jones. We focus on the top five most influential outlets in the cryptocurrency field as they provide the most relevant and impactful news. The keywords used to search news are $Bitcoin, Bitcoins, BTC$, and $BTCs$. Each news item selected contains at least one keyword in the headline. We download a total of 40,261 pieces of news from *Factiva*. After dropping duplicates, we obtain 39,611 unique news. Table 2 describes the outlets and the proportions of news articles sourced from each outlet in our news database. Notably, CoinDesk and Cointelegraph account for approximately eighty percent of the total news, as they are currently the leading outlets in the field, producing significantly more information than their counterparts.

In addition to using the headline of each news article (Lopez-Lira and Tang, 2023; Chen et al., 2023), we also incorporate the first sentence of each article for analysis to enhance analytical accuracy. We then utilize ChatGPT through an application programming interface (API) to classify each news item into five categories, namely, *VERY NEGATIVE*, *NEGATIVE*, *UNKNOWN*, *POSITIVE*, and *VERY POSITIVE* by using the following prompt:

> Please disregard any previous instructions. As a professional financial analyst, your task is to analyze Bitcoin/BTC news and classify it into five categories: VERY NEGATIVE, NEGATIVE, UNKNOWN, POSITIVE, VERY POSITIVE. Here, VERY NEGATIVE indicates a major negative impact, while VERY POSITIVE indicates a major positive impact. NEGATIVE and POSITIVE encompass news that has either negative or positive influence on the market, albeit not to the extremes. Meanwhile, UNKNOWN indicates neutral news without significant impact. Provide only the classification result without any additional explanations or content. Now, analyze the news:"[The headline and first sentence of a given news are input here]".

For example, consider the following headline and first sentence from a randomly chosen news article:

> Headline: NYSE Files to List Shares of Valkyrie's Bitcoin ETF.
>
> First sentence: Valkyrie Digital Assets is getting ready to launch its bitcoin exchange-traded fund (ETF).The New York Stock Exchange (NYSE) filed a 19B-4 Form on behalf of the investment firm for its bitcoin ETF late on Friday.

The final prompt in this case is:

> Please disregard any previous instructions. As a professional financial analyst, your task is to analyze Bitcoin/BTC news and classify it into five categories: VERY NEGATIVE, NEGATIVE, UNKNOWN, POSITIVE, VERY POSITIVE. Here, VERY NEGATIVE indicates a major negative impact, while VERY POSITIVE indicates a major positive impact. NEGATIVE and POSITIVE encompass news that has either negative or positive influence on the market, albeit not to the extremes. Meanwhile, UNKNOWN indicates neutral news without significant im-

pact. Provide only the classification result without any additional explanations or content. Now, analyze the news: "NYSE Files to List Shares of Valkyrie's Bitcoin ETF. Valkyrie Digital Assets is getting ready to launch its bitcoin exchange-traded fund (ETF).The New York Stock Exchange (NYSE) filed a 19B-4 Form on behalf of the investment firm for its bitcoin ETF late on Friday."

Then ChatGPT replies:

POSITIVE

We use *gpt-3.5-turbo-1106* model to generate classification outputs. As such, we count the number of news articles in each category each day and assign different scores accordingly. Specifically, we assign a score of -2 to *VERY NEGATIVE*, -1 to *NEGATIVE*, 0 to *UNKNOWN*, +1 to *POSITIVE*, and +2 to *VERY POSITIVE* categories. Then, we calculate the daily GPT sentiment by:

$$Sent_t = \ln\left(\frac{1 + \Sigma POSITIVE_t}{1 + |\Sigma NEGATIVE|_t}\right), \tag{14}$$

where $\Sigma POSITIVE_t$ denotes the sum of all positive scores, and $|\Sigma NEGATIVE|_t$ represents the sum of all negative scores, taking their absolute values.

[Insert Table 2 about here]

3. Interest rate. Existing literature demonstrates that interest rates influence implied volatility in the stock market (Mixon, 2002; Vähämaa and Äijö, 2011; Amengual and Xiu, 2018). Following these studies, we use three-month U.S. T-bill rates downloaded from the Federal Reserve Bank of St. Louis as one predictor.

4. Trading volume. Harris and Raviv (1993), Chan and Fong (2000), Avramov et al. (2006) show that trading volume influences stock market volatility. We download the trading volume of Bitcoin acquired from `Coinmarketcap.com` and use its logarithmic value as one predictor.

   To account for influences from other markets, we incorporate the following predictors:

5. VIX. We download the VIX data from *CBOE Global Markets*.

6. Gold returns. We retrieve the gold return data from `Investing.com`.

Table 3 reports the summary statistics of these variables. Over the sample period, the average value of the BVIX is more than four times higher than the VIX, indicating that the Bitcoin market is much more volatile than the stock market. Additionally, the median, standard deviation, maximum, and minimum values of the BVIX are significantly higher than those of the VIX. However, the lower kurtosis and skewness of the BVIX suggest a flatter and more symmetric distribution than the VIX. We also plot the time series of these variables in Fig. 2.

[Insert Table 3 about here]

[Insert Figure 2 about here]

4 Empirical results

4.1 In-sample results

We first conduct in-sample predictive regressions of BVIX on the predictors. Table 4 reports the coefficients, t-statistics, and adjusted R^2 of the benchmark AR(1) model, extended bivariate regressions, and multiple regression. The results of bivariate regressions show that, except for the VIX (VIX) and gold return ($Gold$), all remaining predictors have significant in-sample predictive power, indicating that the stock market implied volatility and the commodity market's performance has little influence on Bitcoin market volatility. The first lag of BVIX demonstrates the highest predictive power in terms of t-statistic, aligning with our assumption to use the AR(1) model as our benchmark model. Investor attention ($Attn$) exhibits the second highest predictive power, followed by the three-month U.S. T-bill rate ($Tbill$) and trading volume (Vol). The GPT sentiment index ($Sent$) also has significant predictive power but is weaker than investor attention, indicating that attention is more important than sentiment in the Bitcoin market. The coefficients of AR(1), $Sent$, $Attn$, and Vol are significantly greater than zero, suggesting a positive relationship with implied volatility, while $Tbill$ has a negative influence. Higher investor sentiment, attention, and trading volume encourage market activity and speculation, increasing volatility. However, higher interest rates indicate a shift towards safer investments, reducing demand for speculative trading in Bitcoin and decreasing its implied volatility.

[Insert Table 4 about here]

When it comes to multiple regression, the right panel of Table 4 shows that the AR(1) component still demonstrates the highest predictive power according to t-statistics, followed by $Attn$, $Tbill$, Vol and $Sent$. In contrast, VIX and $Gold$ remain insignificant. Notably, $Sent$ appears to possess predictive power even after controlling for $Attn$, suggesting a different channel of influence on volatility. The adjusted R^2 reaches its highest level of 0.5331, indicating that incorporating all the predictors achieves the most substantial predictive power. The signs of each predictor remain consistent with those in the bivariate models.

4.2 Out-of-sample forecast

4.2.1 Statistical significance

We set 1 January 2021 as the beginning of the out-of-sample period. We use an extensive time window to run the predictive models and generate the forecast. Table 5 reports the out-of-sample R^2_{OS} for the six extended models, along with combination forecasts and the multiple regression(All). In the cluster of extended models, the models incorporating $Sent$, $Attn$, Vol, or $Tbill$ generate positive and significant R^2_{OS}. The model including $Attn$ achieves the highest R^2_{os} of 0.2252, indicating better out-of-sample performance than the other extended models. In contrast, the models incorporating VIX or $Gold$ display no out-of-sample predictive power with negative values of R^2_{os}. The results are consistent with the in-sample results reported in Table 4. The combination forecasts all have positive R^2_{os} and significantly outperform the benchmark, with the $Mean$ method performing best with an R^2_{os} equal to 0.1317. The multiple regression, which includes all the predictors, exhibits the strongest out-of-sample predictive power among all the models, with the R^2_{os} peaking at 0.2556.

Additionally, to study whether the predictive power concentrates on a particular period, we plot the cumulative squared forecasting errors, which are calculated by subtracting the cumulative squared forecasting errors of one predictive model from that of the benchmark AR(1) model in Fig. 3. Specifically, the the cumulative squared forecasting error up to day n is calculated as follows:

$$Diff_{i,n} = \sum_{k=1}^{k=n} (B\bar{V}IX_{t+k})^2 - BVIX_{t+k})^2 - \sum_{k=1}^{k=n} (B\hat{V}IX_{i,t+k} - BVIX_{t+k})^2, \tag{15}$$

where n denotes the time of forecast, taking values from 1 to M.

An increase in the graph indicates a smaller cumulative squared error than the benchmark, hence a better forecasting performance. Except for the extended models that include VIX and *Gold*, all the remaining models show an upward trend. The combination forecasts exhibit relatively flatter trends than other predictive models. The results indicate that the improvement of predictive models over the AR(1) model happens across the whole sample period.

[Insert Table 5 about here]

[Insert Figure 3 about here]

Table 6 reports the MHLN statistics developed by Harvey et al. (1998) over the out-of-sample period. Each item in the table corresponds to the null hypothesis that the forecasts generated by the column heading encompass the forecasts generated by the row heading against the alternative hypothesis that the forecasts generated by the column heading do not encompass the forecasts generated by the row heading.

The first column of Table 6 shows that we cannot reject the null hypothesis that the AR(1) forecasts encompass the forecasts for the VIX and *Gold* at a conventional significance level, suggesting including these two factors into the AR(1) model does not bring additional information to improve forecasting performance. The statistics of other predictive models are significant, indicating that other predictors or combination methods contain helpful information not included in the AR(1) component. The results are consistent with the findings from the R^2_{os} values. Moreover, according to the MHLN statistics in the second and fourth column, the *Attn* forecasts encompass the forecasts from *Sent* but not vice versa, demonstrating that investor attention is more critical in forming Bitcoin implied volatility than sentiment.

[Insert Table 6 about here]

4.3 Economic significance

To assess the economic significance of implied volatility predictability, we follow Bali et al. (2023) to construct zero-cost long-short options portfolios based on the forecasts from each predictive model and compare the portfolio performance with the benchmark. In order to avoid microstructure-related bias, we remove options with bid prices equal to zero and then select those with a positive bid-ask spread, a time-to-maturity of at least seven days, an open interest of at least 100, and a relative bid-ask spread (defined as the bid-ask spread divided by the mid-price) of less than 100%. As explained in Section 2, we sort the options into five groups based on the percentage deviation of the theoretical price at $t+1$ using one predictive model from the mid-price on the day t. We construct a zero-cost long-short portfolio by buying the most undervalued options (top group) and selling the most overvalued options (bottom group). The portfolios are weighted by the options' open interests and held for one week. We subtract the portfolio returns of the benchmark AR(1) from those of each predictive model to yield excess returns; thus, any model with economically significant predictability generates positive average excess returns. We apply a 1% winsorization to remove extreme returns values and calculate the t-statistics adjusted by Newey and West (1987) with lags set to five to account for autocorrelation . According to Ofek et al. (2004), transaction costs can considerably influence economic returns in options markets. Thus, we include effective spread to measure transaction cost level and follow Bali et al. (2023) to set it equal to 25%, 50%, and 100%, respectively.

Table 7 presents the summary statistics for the options of five ranked groups constructed based on the forecasts from each predictive model. We report the moneyness, time to maturity (TtM), proportion of call options ($\%Call$), proportion of out-of-the-money options ($\%OTM$), relative bid-ask spread ($Spread$), open interest (OI), and Delta. While there is slight variation among the predictive models within each group, the differences between the groups are significant. Compared to the bottom group, the top group contains a more significant fraction of out-of-the-money options with higher moneyness, longer time-to-maturity, and wider bid-ask spreads. In contrast, the proportion of call options is slightly lower. Additionally, the average open interest for options in the top group is considerably higher than that in the bottom group, suggesting a liquidity preference. The average Delta calculated from the Black-Scholes formula for the top group is only one-third of the bottom group.

[Insert Table 7 about here]

Table 8 reports the average annualized return of each model over the benchmark AR(1) model.[4] Panel A corresponds to the results when using the mid-price between the bid and ask as the trading price without considering transaction costs, while Panels B to D consider an increasing level of effective spreads to reflect transaction costs. The extended model using *Tbill*, the multiple regression, and all combination forecasts can achieve consistent significant positive excess returns across effective spreads, demonstrating their economic values. The investor attention, *Attn*, outperforms the benchmark significantly but only when transaction costs are low. In most cases, *Sent*, VIX, and Vol generate positive excess returns but are all insignificant. Gold returns *Gold* beat the benchmark to some extent when including transaction costs. The results indicate the economic significance of the predictive models. In addition, the results are slightly different from those of R^2_{OS} reported in Table 5, echoing the findings in Rapach et al. (2010) and Thornton and Valente (2012).

Some excess returns reported in Table 8 increase as the effective spread widens. For example, the excess return of the *Tbill* increases from 22.48% to 55.97% as the effective spread rises from 0 to 100%. This result is because we calculate the excess returns by deducting the returns of the benchmark AR(1) model, and the effective spreads affect the benchmark model more than the extended model using *Tbill*. An increase in excess return as transaction costs widen indicates a better performance than the benchmark.

[Insert Table 8 about here]

Fig. 4 and Fig. 5 plot the cumulative excess returns from predictive regressions (bivariate and multiple regressions) and combination forecasts, respectively. The cumulative excess return until day n is calculated by $CER_{i,n} = \sum_{k=1}^{k=n}(R_{i,t+k} - R_{bmk,t+k})$, where $R_{i,t+k}$ and $R_{bmk,t+k}$ are portfolio returns for the i_{th} predictive model and the benchmark AR(1) model at time $t+k$, respectively. The multiple regression yields the highest excess returns among all predictive models over the out-of-sample period except when the effective spread peaks at 100%. In contrast, the trimmed-mean and median methods produce the lowest returns. In summary, the predictive regressions exhibit greater volatility. At the same time, the combination forecasts display a more stable and staircase-

[4] We multiply the daily return by 252 to get the annualized one.

like pattern, demonstrating their capabilities in reducing the noise of individual forecasts (Rapach et al., 2010).

[Insert Figure 4 about here]

[Insert Figure 5 about here]

Furthermore, we also follow the approach of Goncalves and Guidolin (2006) to build value-weighted options portfolios consisting of up to ten options by purchasing the five most undervalued options and selling the five most overvalued options, with an initial fund of $1,000 on each trading day. For simplicity, we do not report the results. The portfolios returns using this strategy are much more volatile. Nevertheless, we receive significant results when assuming zero transaction costs. The excess returns reduce rapidly, all becoming insignificant when transaction costs are considered.

4.4 Gain on alpha

We document predictive models' positive and significant excess returns in Section 4.3. However, systematic risk might explain these excess returns, resulting in a zero alpha for them. We utilize the cryptocurrency risk factors data in Liu et al. (2022), downloaded from Liu's personal website (`https://www.yukunliu.com/research/`), to investigate whether the strategies can produce positive alpha. We use the one-factor model, which incorporates the market factor (MKT), and the three-factor model, which includes market, size (SMB), and momentum (MOM) factors. We run a time series regression of excess returns on these factors and test whether the alphas are positive. Specifically, we run the following regressions:

$$R_{i,t} - R_{bmk,t} = \alpha_i + \beta_{i,MKT} MKT_t + \epsilon_{i,t}, \tag{16}$$

$$R_{i,t} - R_{bmk,t} = \alpha_i + \beta_{i,MKT} MKT_t + \beta_{i,SMB} SMB_t + \beta_{i,MOM} MOM_t + \epsilon_{i,t}. \tag{17}$$

Eq. (16) and Eq. (17) represent the one-factor model and three-factor model respectively; MKT, SMB and MOM denote the market, size, and momentum factors, respectively. Since the return series is in the daily horizon, we convert the weekly factors data to daily data, assuming there is no change in the factors within one week. We focus on the signs and significance levels of alpha coefficients.

Table 9 reports the alphas under different settings. For the one-factor model, *Attn* and the multiple regression can generate positive and significant alphas when effective spreads are low. In contrast, the significance of alphas from *Tbill* and *Gold* increases as effective spreads widen. The majority of combination forecasts produce significant and positive alphas across different spreads, while the alphas of *Sent*, VIX, and Vol are insignificant. When the three-factor model is introduced, the significance levels for most cases diminish, indicating that part of the excess returns from the predictive models are explained by size and momentum factors. For example, the t-statistic of the alpha using multiple regression (All) is 2.66 for the one-factor model and decreases to 2.11 for the three-factor model. On the other hand, *Attn*, the multiple regression, and the combination forecasts under the averaging scheme of DMSE can still achieve significant and positive alphas. Overall, the results suggest that the cryptocurrency risk factors of Liu et al. (2022) help explain the excess returns of predictive models. Nevertheless, significant and positive alphas still exist when the risk factors are accounted for.

[Insert Table 9 about here]

5 Additional tests

5.1 Call vs. put options

We have used both call and put options in our analysis. One question of interest is whether the power of predictive models is different among options. For example, Lemmon and Ni (2014) document that investor sentiment and past market return affect stock options, whereas the demand for index options is invariant to them. To study this question, we split options into calls and puts and evaluate the performance of trading strategies using call and put options, respectively. The filter conditions are the same as above.

Table 10 reports the annualized excess returns and alphas of trading call options only. The extended models incorporating *Sent* yield significant positive excess returns and alphas under all scenarios, while the extended model using $Attn$ generates significant and positive results when the effect spread is zero (mid-price) or 25%. Other extended models that use VIX, Vol, or $Tbill$ fail to outperform the benchmark significantly. These results suggest investor sentiment and attention exhibit significant predictive power for call options. Given Bitcoin's strong lottery-like characteristics (Lee and Wang, 2024), this result is consistent with Byun and Kim (2016) that investor sentiment plays a vital role in pricing call options written on lottery-like assets. As for the combination forecasts, the mean method and DMSE methods achieve significantly positive excess returns and alphas under high transaction cost levels while the trimmed mean method performs better under lower cost levels. In contrast, the median method fails to outperform the benchmark significantly in most cases.

Table 11 reports the annualized excess returns and alphas of trading put options only. The performance of predictive models shows a distinctly different pattern from call options reported in Table 10. The extended model using $Tbill$ generates positive values that are highly significant and robust to transaction costs, while other extended models fail to outperform the benchmark in most cases. One possible explanation for this superior performance is that interest rates heavily influence the demand for put options in the Bitcoin market. Barraclough and Whaley (2012) find that, in the stock market, many put options that should have been exercised are left unexercised, resulting in substantial foregone interest income. We find a different phenomenon in the Bitcoin option market. Additionally, the multiple regression model and combination forecasts adopting DMSE methods produce significantly positive excess returns and alphas under various transaction cost settings. The comparison between Table 10 and Table 11 indicates that predictors affect call and put options differently.

[Insert Table 10 about here]

[Insert Table 11 about here]

5.2 Fundamental vs. non-fundamental sentiment

Biais et al. (2023) demonstrate that sentiment can be classified into fundamental and non-fundamental, affecting Bitcoin price differently. Following Biais et al. (2023), we classify Bitcoin daily news into fundamental news and non-fundamental news to investigate the effectiveness of each group. According to Biais et al. (2023), fundamental news refers to information that facilitates the exchange of Bitcoin to other currencies or the use of Bitcoin to purchase goods and services. Conversely, news that does not fit these criteria is classified as non-fundamental. We use the following prompt:

> Forget all previous instructions. As a professional financial analyst, your task is to analyze Bitcoin/BTC news. Label news as FUNDAMENTAL if relevant to Bitcoin's fundamentals, and as NONFUNDAMENTAL if not. A news is defined as a fundamental one if it falls into one of the following two categories: The ease with which Bitcoins can be exchanged for currencies such as, for example, the euro, Japanese yen, or U.S. dollar; The ease with which Bitcoin can be used to buy goods or services and thus reap transactional benefits. Otherwise, it is a nonfundamental news. Provide only the label result without any additional explanations or content. Now, analyze the news: "[The headline and first sentence of a daily news are input here]".

Fig. 6 presents the graph of the GPT sentiment indices based on fundamental news and non-fundamental news separately. The graph reveals that, compared with the non-fundamental index, the values of the fundamental index are, in most cases, greater than zero, indicating that fundamental news tends to cause positive investor sentiment. In contrast, non-fundamental news shows a more significant fluctuation, ranging from negative to positive sentiment.

We then run the bivariate and multiple predictive regressions by replacing the GPT sentiment index ($Sent$) with the fundamental sentiment index ($FSent$) and non-fundamental sentiment index ($NFSent$). Table 12 reports the results. In bivariate regressions, both the fundamental and non-fundamental sentiment index coefficients are statistically significant, with t-statistics of 1.67 and 2.64, respectively. The non-fundamental sentiment index has a more significant influence on BVIX than the fundamental sentiment index. When all predictors are used in the multiple regression, the fundamental sentiment index becomes insignificant, while the non-fundamental index remains significant with a t-statistic of 1.83. These results suggest that the influence of investor sentiment on Bitcoin implied volatility primarily stems more from factors related to non-fundamental news than fundamental news, supporting the findings in Pagnotta (2022) and Biais et al. (2023).

[Insert Figure 6 about here]

[Insert Table 12 about here]

6 Conclusion

In this paper, we introduce a model-free methodology to construct a Bitcoin implied volatility index (BVIX) and investigate its influencing predictors, contributing to the growing body of research on cryptocurrency markets.

Our research underscores the significantly higher volatility of Bitcoin compared to the stock market. We identify investor attention as the most significant predictor, highlighting its crucial role in cryptocurrency markets. Additionally, we find that the interest rate, measured as a three-month U.S. Tbill rate, holds strong predictive power, while other factors such as investor sentiment and trading volume show limited predictive power.

We document the statistical significance of the BVIX predictability utilizing out-of-sample R^2_{os} and its economic significance by constructing option trading strategies based on forecasts from the predictive models. The option strategies generate significant and positive returns over the benchmark AR(1) model. The excess returns are also robust to the cryptocurrency market risk factors of Liu et al. (2022) and transaction costs.

We also find that predictors affect call and put options in different ways. $Attn$ and $Sent$ contain helpful information for call options, while put options are mainly influenced by $Tbill$. Additionally, we investigate the influences of fundamental and non-fundamental sentiment on Bitcoin's implied volatility. Our findings provide evidence that non-fundamental sentiment plays a more important role in shaping implied volatility than fundamental sentiment.

As Bitcoin and other prominent cryptocurrencies continue to grow in importance within the global economy, this paper provides new insights for academia and industry in measuring and forecasting their volatility, offering a deeper understanding for researchers and market participants to manage the risk of the cryptocurrency market. Future research can be extended to study how it affects the behavior of firms with positions in the cryptocurrency market due to its high volatility.

Future research can also be extended to use other machine learning methods to analyze the news.

References

Alexander, C., Imeraj, A., 2023. The Bitcoin VIX and its variance risk premium. Journal of Alternative Investments 23, 84–109.

Amengual, D., Xiu, D., 2018. Resolution of policy uncertainty and sudden declines in volatility. Journal of Econometrics 203, 297–315.

Avramov, D., Chordia, T., Goyal, A., 2006. The impact of trades on daily volatility. The Review of Financial Studies 19, 1241–1277.

Bali, T. G., Beckmeyer, H., Moerke, M., Weigert, F., 2023. Option return predictability with machine learning and big data. The Review of Financial Studies 36, 3548–3602.

Barraclough, K., Whaley, R. E., 2012. Early exercise of put options on stocks. The Journal of Finance 67, 1423–1456.

Biais, B., Bisiere, C., Bouvard, M., Casamatta, C., Menkveld, A. J., 2023. Equilibrium Bitcoin pricing. The Journal of Finance 78, 967–1014.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654.

Byun, S.-J., Kim, D.-H., 2016. Gambling preference and individual equity option returns. Journal of Financial Economics 122, 155–174.

Campbell, J. Y., Thompson, S. B., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? The Review of Financial Studies 21, 1509–1531.

Cao, J., Han, B., 2013. Cross section of option returns and idiosyncratic stock volatility. Journal of Financial Economics 108, 231–249.

Carr, P., Wu, L., 2009. Variance risk premiums. The Review of Financial Studies 22, 1311–1341.

Cboe Global Markets, L., 2022. Volatility index® methodology: Cboe volatility index. Available from CBOE website.

Chan, K., Fong, W.-M., 2000. Trade size, order imbalance, and the volatility–volume relation. Journal of Financial Economics 57, 247–273.

Chen, J., Tang, G., Zhou, G., Zhu, W., 2023. Chatgpt, stock market predictability and links to the macroeconomy. Available at SSRN 4660148 .

Christensen, B. J., Prabhala, N. R., 1998. The relation between implied and realized volatility. Journal of Financial Economics 50, 125–150.

Christensen, P. A., Olsen, R. J., Long, S. W., Subedi, S., Davis, J. J., Hodjat, P., Walley, D. R., Kinskey, J. C., Saavedra, M. O., Pruitt, L., et al., 2022. Delta variants of sars-cov-2 cause significantly increased vaccine breakthrough covid-19 cases in houston, texas. The American Journal of Pathology 192, 320–331.

Clark, T. E., West, K. D., 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138, 291–311.

Demeterfi, K., Derman, E., Kamal, M., Zou, J., 1999. More than you ever wanted to know about volatility swaps. Goldman Sachs Quantitative Strategies Research Notes 41, 1–56.

Elliott, P., Bodinier, B., Eales, O., Wang, H., Haw, D., Elliott, J., Whitaker, M., Jonnerby, J., Tang, D., Walters, C. E., et al., 2022. Rapid increase in omicron infections in england during december 2021: React-1 study. Science 375, 1406–1411.

Goncalves, S., Guidolin, M., 2006. Predictable dynamics in the S&P 500 index options implied volatility surface. The Journal of Business 79, 1591–1635.

Griffin, J. M., Shams, A., 2020. Is Bitcoin really untethered? The Journal of Finance 75, 1913–1964.

Gürkaynak, R. S., Sack, B., Wright, J. H., 2007. The US Treasury yield curve: 1961 to the present. Journal of Monetary Economics 54, 2291–2304.

Harris, M., Raviv, A., 1993. Differences of opinion make a horse race. The Review of Financial Studies 6, 473–506.

Harvey, D. I., Leybourne, S. J., Newbold, P., 1998. Tests for forecast encompassing. Journal of Business & Economic Statistics 16, 254–259.

Jiang, G. J., Tian, Y. S., 2005. The model-free implied volatility and its information content. The Review of Financial Studies 18, 1305–1342.

Lee, S. S., Wang, M., 2024. Variance decomposition and cryptocurrency return prediction. Journal of Financial and Quantitative Analysis pp. 1–60.

Lemmon, M., Ni, S. X., 2014. Differences in trading and pricing between stock and index options. Management Science 60, 1985–2001.

Lin, H., Wu, C., Zhou, G., 2018. Forecasting corporate bond returns with a large set of predictors: An iterated combination approach. Management Science 64, 4218–4238.

Liu, Y., Tsyvinski, A., 2021. Risks and returns of cryptocurrency. The Review of Financial Studies 34, 2689–2727.

Liu, Y., Tsyvinski, A., Wu, X., 2022. Common risk factors in cryptocurrency. The Journal of Finance 77, 1133–1177.

Lopez-Lira, A., Tang, Y., 2023. Can chatgpt forecast stock price movements? Return predictability and large language models. arXiv preprint arXiv:2304.07619 .

Makarov, I., Schoar, A., 2020. Trading and arbitrage in cryptocurrency markets. Journal of Financial Economics 135, 293–319.

Makarov, I., Schoar, A., 2021. Blockchain analysis of the Bitcoin market. Available at SSRN .

Mixon, S., 2002. Factors explaining movements in the implied volatility surface. Journal of Futures Markets 22, 915–937.

Nakamoto, S., 2008. Bitcoin: A peer-to-peer electronic cash system .

Newey, W., West, K., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Ofek, E., Richardson, M., Whitelaw, R. F., 2004. Limited arbitrage and short sales restrictions: Evidence from the options markets. Journal of Financial Economics 74, 305–342.

Pagnotta, E. S., 2022. Decentralizing money: Bitcoin prices and blockchain security. The Review of Financial Studies 35, 866–907.

Rapach, D. E., Strauss, J. K., Zhou, G., 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. The Review of Financial Studies 23, 821–862.

Sockin, M., Xiong, W., 2023. A model of cryptocurrencies. Management Science 69, 6684–6707.

Thornton, D. L., Valente, G., 2012. Out-of-sample predictions of bond excess returns and forward rates: An asset allocation perspective. The Review of Financial Studies 25, 3141–3168.

Vähämaa, S., Äijö, J., 2011. The Fed's policy decisions and implied volatility. Journal of Futures Markets 31, 995–1010.

Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. The Review of Financial Studies 21, 1455–1508.

Figure 1: The BVIX index

This graph plots the BVIX index from 1 January 2020 to 30 April 2024. The historical options data, which contain 172,166 records, were downloaded from LedgerX Derivatives Exchange.

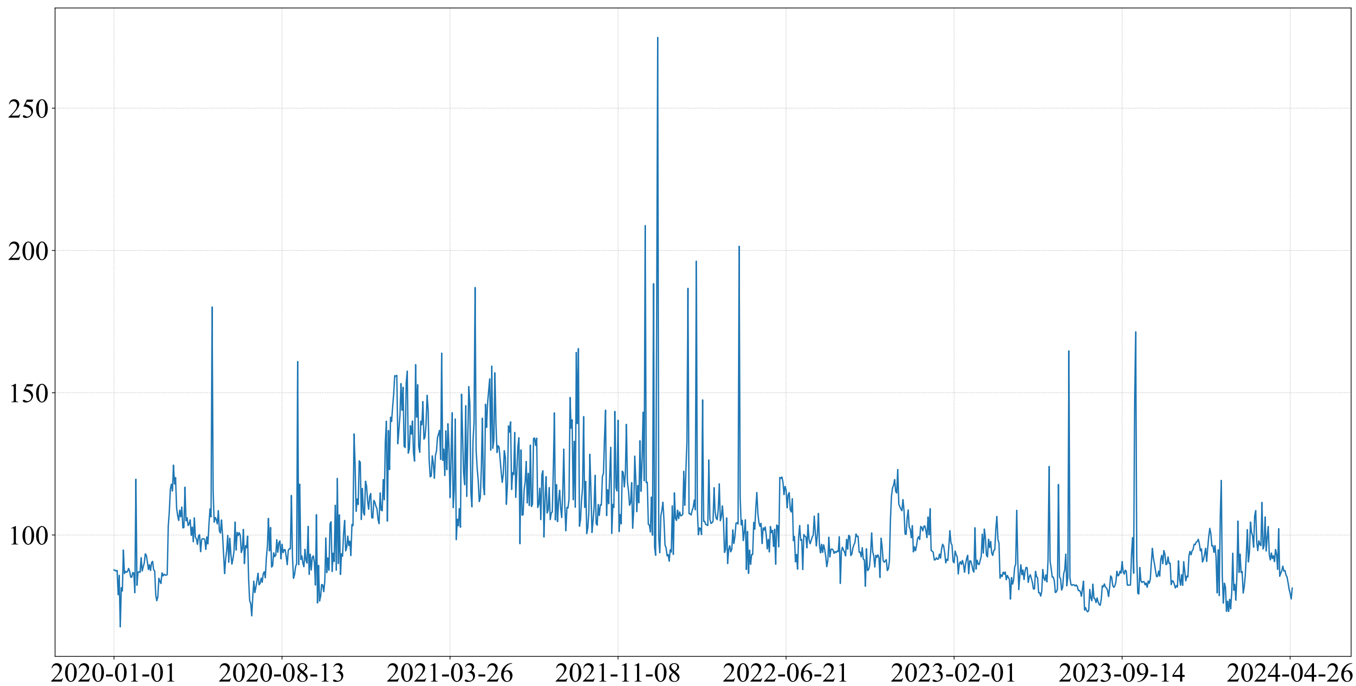

Figure 2: Predictors

This graph plots the time series of six predictors from 1 January 2020 to 30 April 2024. *Sent*, VIX, *Attn*, Vol, *Tbill*, and *Gold* represent the GPT sentiment index, the VIX index, investor attention, trading volume, three-month U.S. Tbill rate, and gold returns, respectively. *Attn* and Vol are measured in logarithmic values.

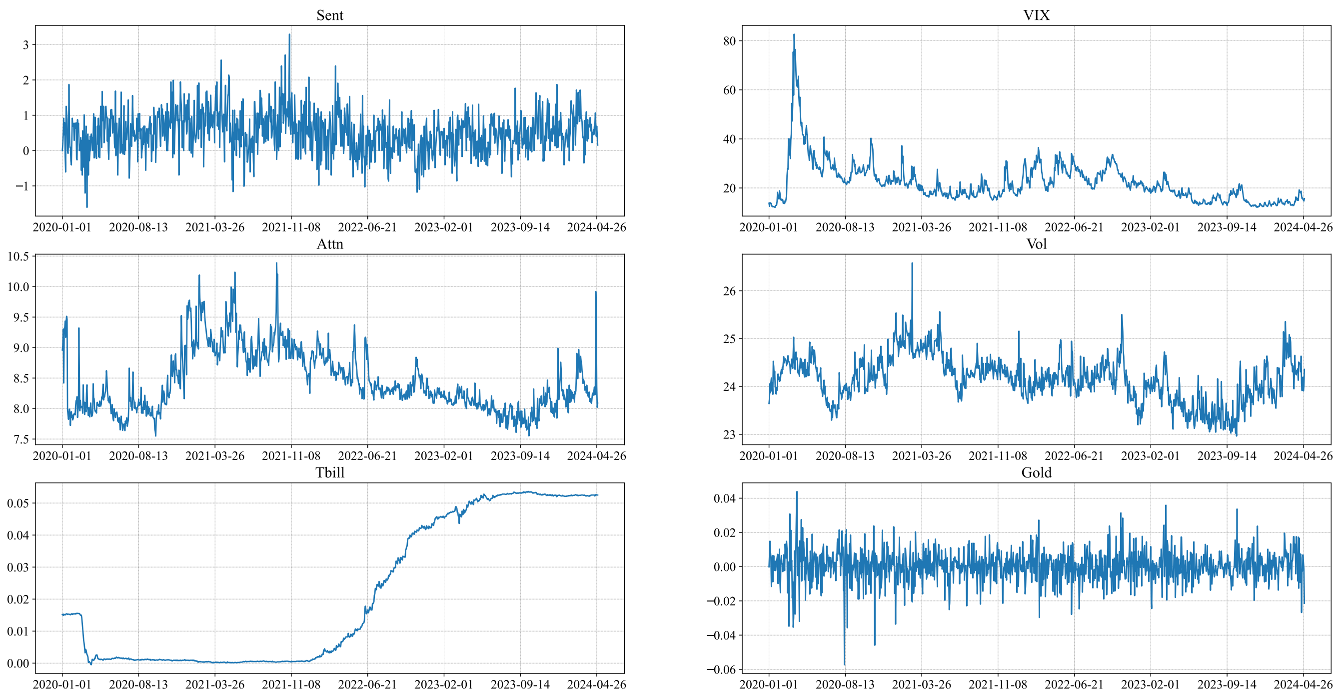

Figure 3: Cumulative squared forecasting errors

This graph illustrates the difference between cumulative squared forecasting errors of each predictive model and the benchmark AR(1) model, calculated as $Diff_{i,n} = \sum_{k=1}^{k=n}(B\bar{V}IX_{t+k} - BVIX_{t+k})^2 - \sum_{k=1}^{k=n}(B\hat{V}IX_{i,t+k} - BVIX_{t+k})^2$, where $B\bar{V}IX_{t+k}$, $B\hat{V}IX_{i,t+k}$, and $BVIX_{t+k}$ represent the forecasted value from the benchmark AR(1) model, the forecasted value from the i_{th} predictive model, and the actual value of Bitcoin implied volatility at time $t+k$, respectively; n denotes the day of forecasts, taking values from 1 to M, where M is the number of observations in the out-of-sample period. We set 1 January 2021 as the beginning of the out-of-sample period. An increase in the graph indicates a better performance than the benchmark. The graph uses the following abbreviations for the combination methods: Med. = Median; TM = Trimmed mean; D(0.9) = DMSE, θ = 0.9; D(0.95) = DMSE, θ = 0.95; D(1.0) = DMSE, θ = 1.0.

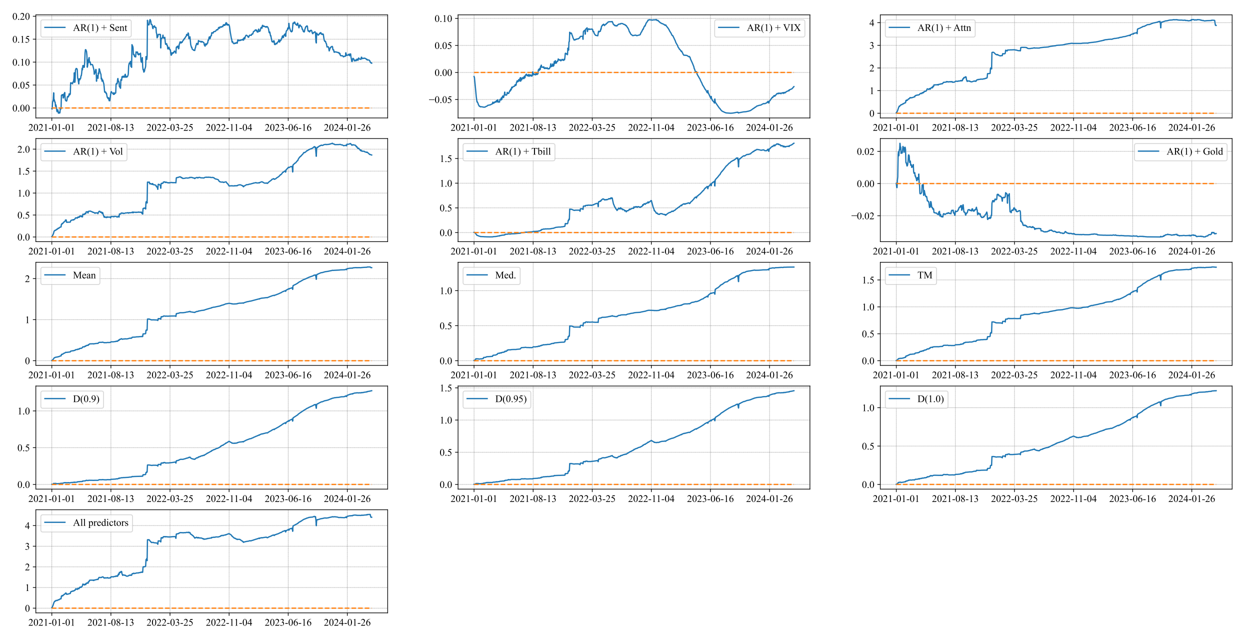

Figure 4: Cumulative excess returns for predictive models

This graph plots the cumulative excess returns of the extended predictive models across different transaction costs, computed as $CER_n^i = \sum_{k=1}^{k=n}(R_{i,t+k} - R_{bmk,t+k})$, where $R_{i,t+k}$ and $R_{bmk,t+k}$ are portfolio returns for the i_{th} predictive model and the benchmark AR(1) model at time $t+k$, respectively; n denotes the number of forecasts, taking values from 1 to M, where M is the number of observations in the out-of-sample period. We set 1 January 2021 as the beginning of the out-of-sample period.

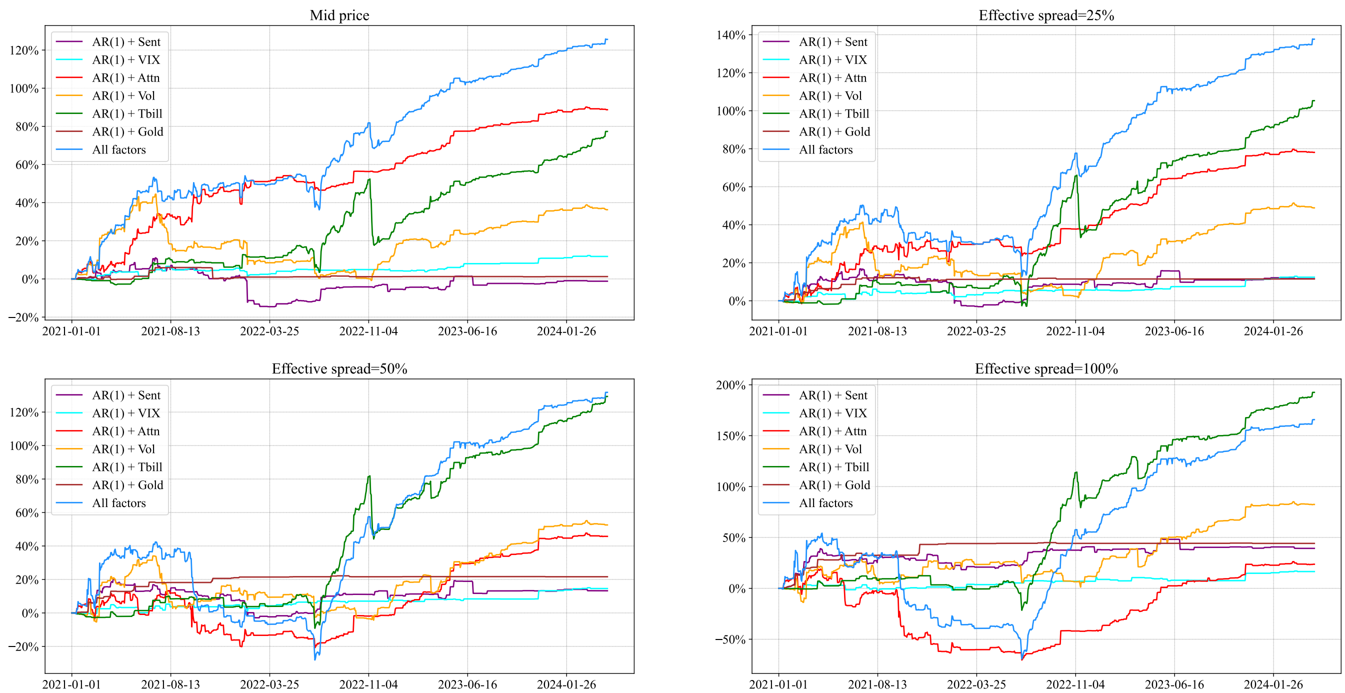

Figure 5: Cumulative excess returns for combination forecasts

This graph plots the cumulative excess returns of the combination forecasts across different transaction costs, computed as $CER_n^c = \sum_{k=1}^{k=n}(R_{t+k}^c - R_{t+k}^{bmk})$, where R_{t+k}^c and R_{t+k}^{bmk} are portfolio returns for the c_{th} combination forecasts and the benchmark AR(1) model at time $t+k$; n denotes the number of forecasts, taking values from 1 to M, where M is the number of observations in the out-of-sample period. We set 1 January 2021 as the beginning of the out-of-sample period.

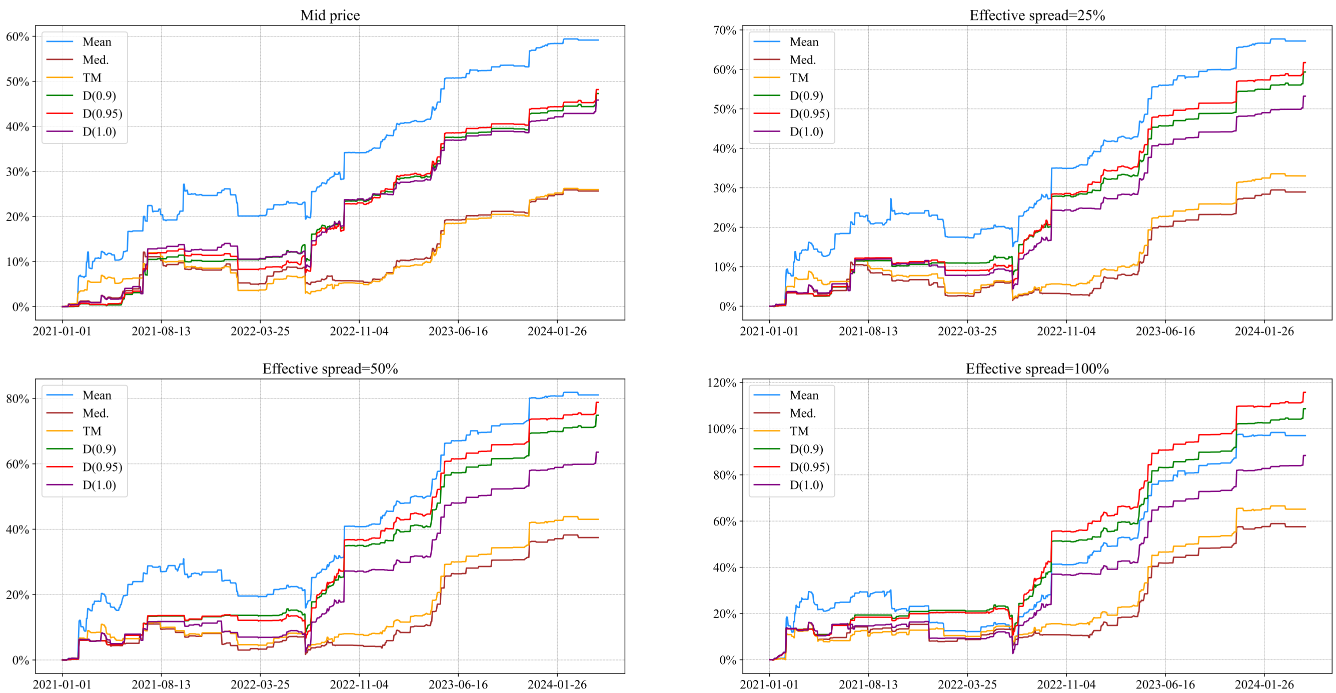

Figure 6: Fundamental and Non-fundamental GPT sentiment indices

This graph illustrates the patterns of GPT sentiment indices when decomposed into fundamental and non-fundamental indices. The former relates to news that facilitates Bitcoin's exchange into other currencies or its use for purchasing goods and services. At the same time, the latter covers news that does not meet these criteria. The period covers from 1 January 2020 to 30 April 2024.

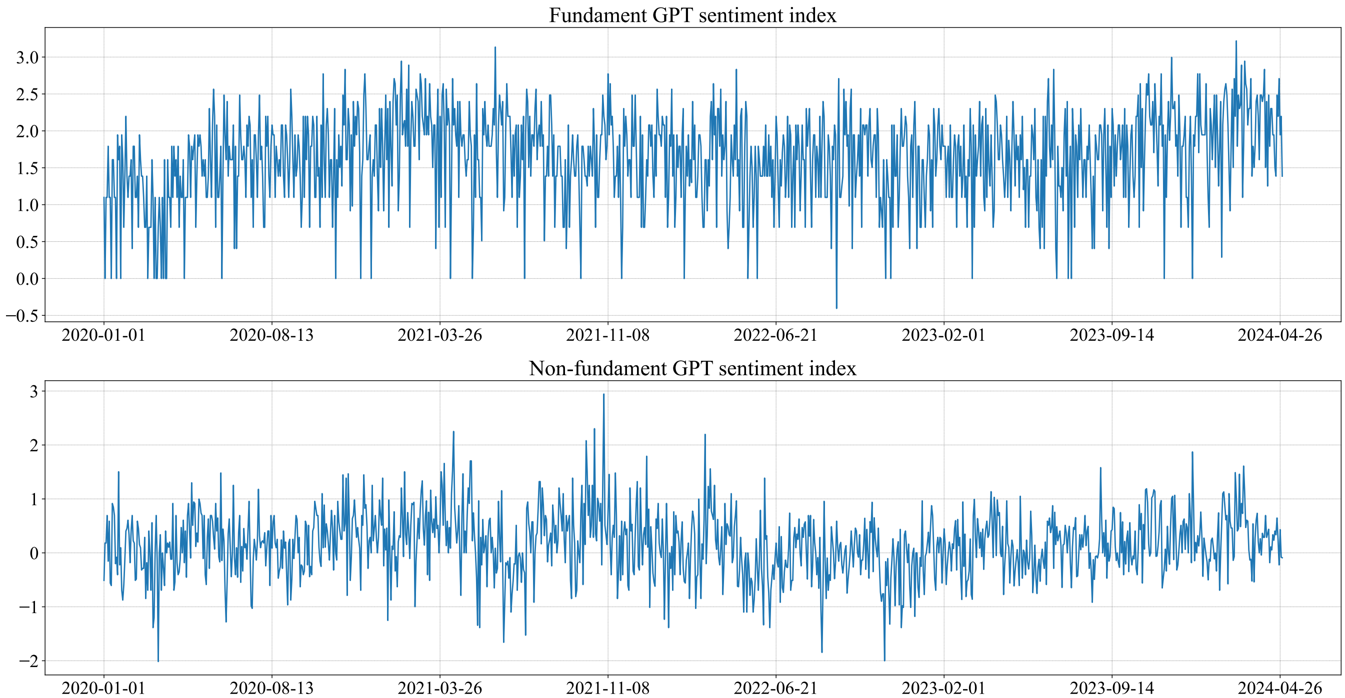

Table 1: Summary of option records.

This table reports the number of options records on the LedgerX exchange from 1 January 2020 to 30 April 2024, categorized by time to maturity and moneyness. It includes the total number of records and the distribution of call and put options, both in absolute numbers and percentages. Moneyness is defined as K/S, where K is the strike price of a certain option and S is the Bitcoin spot price on the recording day.

	Number of records			Percentage		
	All	Call	Put	All	Call	Put
Time to maturity						
Total	172,166	87,930	84,236	100.00%	51.07%	48.93%
$\leq$ 7 days	33,822	17,122	16,700	100.00%	50.62%	49.38%
8-90 days	72,930	37,379	35,551	100.00%	51.25%	48.75%
91-180 days	28,662	14,653	14,009	100.00%	51.12%	48.88%
$>$ 180 days	36,752	18,776	17,976	100.00%	51.09%	48.91%
Moneyness						
Total	172,166	87,930	84,236	100.00%	51.07%	48.93%
$\leq$ 0.85	48,353	26,992	21,361	100.00%	55.82%	44.18%
0.85-1.15	49,409	24,929	24,480	100.00%	50.45%	49.55%
$>$ 1.15	74,404	36,009	38,395	100.00%	48.40%	51.60%

Table 2: News Outlet Descriptions and Proportions

This table describes the top five outlets in the cryptocurrency field and shows the percentage of news articles sourced from each outlet in our news database.

Outlet	Description	Count	%
CoinDesk	CoinDesk is a leading media and data company in the global crypto economy, known for its trusted journalism, major events like Consensus, and expertise in digital asset indices.	13,586	34.30
Cointelegraph	Founded in 2013, Cointelegraph is an independent digital media resource providing comprehensive news on blockchain, crypto assets, and fintech trends.	17,603	44.44
CryptoNews	CryptoNews is a leading source for comprehensive and timely cryptocurrency and blockchain news, attracting over two million visitors monthly.	4,595	11.60
BTC Magazine	Bitcoin Magazine, established in 2012, is the oldest and most trusted source for news, information, and expert commentary on Bitcoin and blockchain technology.	1,891	4.77
Decrypt	Decrypt, founded in 2018, is a next-generation media company and creative studio at the intersection of emerging technology, alternative finance, and culture, powered by AI and Web3.	1,936	4.89
Total		39,611	100.00

Table 3: Summary Statistics

This table presents the summary statistics of the BVIX index and factors, including the GPT sentiment index (*Sent*), the VIX index (VIX), investor attention (*Attn*), trading volume (Vol), the three-month U.S. Tbill rate (*Tbill*), and gold return (*Gold*). The sample period contains 1,130 trading days from 1 January 2020 to 30 April 2024. Both Vol and *Attn* are expressed in logarithmic values.

	Mean	Median	Std.dev	Min.	Max.	Kurtosis	Skewness
BVIX	102.32	97.90	20.33	25.46	274.85	7.24	1.63
Sent	0.51	0.51	0.59	-1.61	3.30	0.71	0.20
VIX	22.14	20.69	8.42	12.07	82.69	10.74	2.47
Attention	8.44	8.29	0.51	7.55	10.39	0.12	0.80
Vol	24.15	24.17	0.47	22.96	26.58	0.38	0.09
Tbill (%)	2.13	0.89	2.25	-0.05	5.36	-1.65	0.41
Gold (%)	0.04	0.08	0.95	-5.73	4.39	3.07	-0.33

Table 4: In-sample predictive regression

This table reports the results of the predictive regressions. The predictors include the first lag of implied volatility ($BVIX_{t-1}$), GPT sentiment index ($Sent_{t-1}$), VIX index VIX_{t-1}, investor attention ($Attn_{t-1}$), trading volume (Vol_{t-1}), three-month U.S. Tbill rate ($Tbill_{t-1}$), and gold return ($Gold_{t-1}$). The t-statistics are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	$BVIX_t$	*Dependent variable:* $BVIX_t$						$BVIX_t$
Predictor	AR(1)	Bivariate regression						Multiple regression
IV_{t-1}	0.67***	0.66***	0.67***	0.47***	0.56***	0.56***	0.67***	0.41***
	(30.31)	(29.62)	(30.04)	(17.81)	(23.01)	(22.50)	(30.29)	(15.01)
$Sent_{t-1}$		2.70***						1.54**
		(3.50)						(2.07)
VIX_{t-1}			0.08					0.05
			(1.55)					(0.76)
$Attn_{t-1}$				12.91***				10.00***
				(12.22)				(7.68)
Vol_{t-1}					9.32***			2.87**
					(8.84)			(2.39)
$Tbill_{t-1}$						-2.04***		-1.17***
						(-9.09)		(-4.46)
$Gold_{t-1}$							-26.34	-19.26
							(-0.56)	(-0.44)
Constant	33.73***	33.51***	32.23***	-54.89***	-180.40***	49.69***	33.76***	-92.58***
	(14.62)	(14.59)	(12.89)	(-7.25)	(-7.42)	(17.52)	(14.62)	(-3.69)
Observations	1,129	1,129	1,129	1,129	1,129	1,129	1,129	1,129
Adj. R^2	0.449	0.454	0.449	0.513	0.484	0.486	0.448	0.533

Table 5: Out-of-sample forecast performance.

This table reports out-of-sample R^2_{os} values of the extended models, combination forecasts, and multiple regression. For combination forecasts, we report mean, median, trimmed mean and discount mean squared error (DMSE) forecasts with θ equal to 0.9, 0.95, and 1.0. Mean squared errors (MSE) are used to measure the out-of-sample forecasting errors. The statistical significance of R^2_{os} is evaluated by the p-value for the MSPE-adjusted statistic proposed by Clark and West (2007). ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	Error	R^2_{os}	t-stats	Obs.
Extended models				
AR(1) + Sent	197.36	0.0057***	2.31	867
AR(1) + VIX	198.79	-0.0015	0.21	867
AR(1) + Attn	153.80	0.2252***	5.70	867
AR(1) + Vol	176.97	0.1084***	5.33	867
AR(1) + Tbill	177.63	0.1051***	11.45	867
AR(1) + Gold	198.85	-0.0018	-0.67	867
Combination forecasts				
Mean	172.36	0.1317***	6.95	867
Med.	183.07	0.0777***	6.47	867
TM	178.47	0.1009***	6.66	867
$DMSE(\theta = 0.9)$	183.78	0.0741***	9.49	867
$DMSE(\theta = 0.95)$	181.69	0.0846***	9.45	867
$DMSE(\theta = 1.0)$	184.38	0.0711***	8.14	867
Multiple regression and benchmark				
Multiple regression	147.76	0.2556***	6.69	867
AR(1)	198.49			867

Table 6: Encompassing test results.

This table reports the MHLN statistics of Harvey et al. (1998). The statistic relates to a one-sided (upper-tail) test of the null hypothesis that the forecasts generated by the column heading encompass the forecasts generated by the row heading against the alternative hypothesis that the forecasts generated by the column heading do not encompass the forecasts generated by the row heading. *Sent*, VIX, *Attn*, Vol, *Tbill*, and *Gold* represent the extended models incorporating the corresponding predictor. The table uses the following abbreviations: Bmk=Benchmark; All=All predictors, which refers to the multiple regression; Sent=Sentiment; Attn=Attention; Med.=Median; TM=Trimmed mean; D(0.9)=DMSE, θ=0.9; D(0.95)=DMSE, θ=0.95; D(1.0)=DMSE, θ=1.0.

	AR(1)	Sent	VIX	Attn	Vol	Tbill	Gold	Mean	Med.	TM	D(0.9)	D(0.95)	D(1.0)	All
AR(1)		1.19	1.11	-1.54	0.53	3.59	1.29	-5.33	-5.28	-5.38	-7.00	-6.63	-6.61	2.64
Sent	2.31		2.66	-1.29	0.92	4.08	2.53	-5.64	-4.90	-5.44	-4.94	-5.45	-5.58	2.38
VIX	0.21	1.18		-1.56	0.82	3.72	0.83	-5.40	-5.39	-5.46	-6.90	-6.60	-6.57	2.71
Attn	5.70	6.29	5.84		8.90	7.49	5.75	4.31	4.99	4.77	4.89	4.80	4.99	4.80
Vol	5.33	5.95	5.70	0.67		8.03	5.40	2.23	3.85	3.18	3.93	3.73	3.98	2.56
Tbill	11.45	12.02	11.72	2.57	7.84		11.59	6.60	11.17	9.93	8.43	7.87	9.49	0.64
Gold	-0.67	1.05	0.67	-1.58	0.49	3.57		-5.46	-5.46	-5.53	-7.41	-6.95	-6.99	2.61
Mean	6.95	8.37	7.37	-0.44	4.39	9.39	7.11		6.29	6.35	4.71	4.35	5.35	2.54
Med.	6.47	7.89	6.97	-1.02	1.97	7.05	6.68	-5.31		-5.00	1.67	-0.01	2.14	2.55
TM	6.66	8.21	7.14	-0.78	2.72	8.42	6.84	-5.41	5.47		3.69	3.04	4.38	2.62
D(0.9)	9.50	9.72	9.96	-0.66	2.19	4.61	9.95	-3.98	0.38	-2.45		-6.58	2.26	2.26
D(0.95)	9.45	10.33	9.95	-0.50	2.53	4.98	9.82	-3.58	3.02	-1.35	6.92		8.47	2.20
D(1.0)	8.15	9.36	8.71	-0.84	1.96	5.31	8.54	-4.71	-0.78	-3.56	-1.29	-6.63		2.37
All	6.69	6.85	6.83	6.53	7.83	4.43	6.72	5.22	5.92	5.73	5.48	5.31	5.72	

Table 7: Portfolio decomposition

This table shows summary statistics for the options of five ranked groups constructed based on the forecasts from each predictive model. All characteristics are averaged over time. *Moneyness* denotes K/S, TtM is the time-to-maturity. $\%Calls$ is the proportion of call options in the portfolio. $\%OTMs$ is the share of out-of-the-money options in the portfolio. *Spread* is the relative bid-ask spread. OI the open interest. *Delta* is one option Greek that measures how much an option's price can be expected to move for every \$1 change in Bitcoin price.

	AR(1)	Sent	VIX	Attn	Vol	Tbill	Gold	Mean	Med.	TM	D(0.9)	D(0.95)	D(1.0)	All
Group 1 (top group)														
Moneyness	1.54	1.54	1.54	1.54	1.53	1.49	1.54	1.53	1.54	1.54	1.53	1.53	1.53	1.49
TtM	128.18	128.32	128.22	128.40	128.64	127.90	128.25	128.63	128.42	128.62	128.32	128.40	128.42	128.12
% Calls	72.35%	72.28%	72.52%	72.24%	71.84%	71.32%	72.38%	72.24%	72.34%	72.28%	72.17%	72.19%	72.23%	71.49%
% OTMs	93.40%	93.41%	93.45%	94.51%	93.34%	93.16%	93.37%	93.66%	93.54%	93.56%	93.37%	93.41%	93.42%	94.28%
Spread	50.49%	50.36%	50.51%	50.27%	50.08%	49.48%	50.48%	50.31%	50.35%	50.32%	50.21%	50.19%	50.33%	49.43%
OI	4378.91	4369.12	4379.46	4428.63	4360.80	4277.21	4372.05	4390.35	4375.46	4379.55	4351.91	4342.95	4370.07	4331.99
Delta	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
Group 2														
Moneyness	1.25	1.25	1.25	1.27	1.26	1.22	1.25	1.25	1.25	1.25	1.25	1.24	1.25	1.23
TtM	132.81	133.21	133.05	134.39	132.71	135.68	132.75	133.07	133.05	132.94	133.19	133.34	133.03	137.25
% Calls	64.95%	64.85%	64.78%	65.06%	64.83%	64.52%	64.89%	64.75%	64.80%	64.85%	64.84%	64.75%	64.79%	64.14%
% OTMs	79.40%	79.31%	79.46%	80.46%	79.46%	76.78%	79.46%	79.18%	79.27%	79.31%	79.02%	78.93%	79.13%	78.02%
Spread	39.61%	39.67%	39.67%	40.22%	39.76%	38.55%	39.63%	39.49%	39.51%	39.57%	39.44%	39.32%	39.41%	39.25%
OI	4602.53	4593.73	4627.08	4645.54	4533.65	4455.66	4614.77	4543.86	4580.10	4577.59	4557.61	4578.25	4540.29	4478.68
Delta	0.20	0.20	0.20	0.19	0.19	0.20	0.20	0.19	0.19	0.19	0.20	0.20	0.20	0.19
Group 3														
Moneyness	1.05	1.05	1.06	1.06	1.06	1.06	1.05	1.06	1.05	1.05	1.05	1.05	1.05	1.07
TtM	132.77	132.42	132.59	134.69	134.68	130.96	133.01	133.43	133.34	133.60	132.99	132.77	133.06	132.73
% Calls	60.06%	59.89%	60.10%	60.08%	60.41%	61.03%	60.10%	60.28%	60.15%	60.23%	60.21%	60.21%	60.19%	60.79%
% OTMs	44.27%	43.82%	44.37%	43.83%	43.51%	41.37%	44.34%	43.80%	43.90%	43.91%	43.68%	43.57%	43.81%	41.31%
Spread	31.08%	30.90%	31.08%	31.61%	31.24%	31.10%	31.10%	31.22%	31.16%	31.17%	31.20%	31.29%	31.21%	31.55%
OI	4793.33	4756.61	4816.30	4937.11	4857.45	4787.82	4823.94	4910.83	4868.83	4894.36	4839.54	4805.91	4875.61	4906.55
Delta	0.23	0.23	0.23	0.22	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23
Group 4														
Moneyness	0.94	0.93	0.93	0.92	0.93	0.92	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.92
TtM	109.20	108.86	109.14	106.50	106.97	105.98	108.62	107.68	108.05	107.72	108.08	108.07	108.16	103.87
% Calls	72.02%	71.93%	71.77%	71.78%	72.25%	72.52%	71.95%	72.03%	72.08%	72.00%	72.06%	72.11%	72.06%	72.74%
% OTMs	10.87%	10.89%	10.83%	9.72%	10.69%	10.70%	10.79%	10.31%	10.63%	10.53%	10.53%	10.50%	10.58%	10.10%
Spread	21.55%	21.65%	21.60%	21.09%	21.20%	21.03%	21.48%	21.31%	21.47%	21.42%	21.25%	21.19%	21.29%	20.83%
OT	3607.67	3595.08	3555.61	3420.96	3552.14	3503.85	3571.60	3472.74	3520.20	3493.16	3536.90	3543.22	3509.81	3423.76
Delta	0.41	0.41	0.41	0.41	0.42	0.42	0.41	0.42	0.41	0.41	0.42	0.42	0.42	0.43
Group 5 (bottom group)														
Moneyness	1.38	1.39	1.38	1.37	1.40	1.48	1.38	1.39	1.39	1.39	1.40	1.40	1.40	1.45
TtM	102.22	102.37	102.15	101.14	102.21	104.67	102.48	102.31	102.28	102.27	102.60	102.60	102.48	103.10
%Calls	79.84%	80.28%	80.04%	80.09%	79.97%	79.90%	79.90%	79.95%	79.89%	79.88%	79.94%	79.96%	79.94%	80.12%
% OTMs	28.14%	28.69%	28.01%	27.59%	29.11%	34.13%	28.13%	29.12%	28.75%	28.78%	29.53%	29.70%	29.17%	32.47%
Spread	30.49%	30.67%	30.40%	30.07%	31.01%	33.12%	30.52%	30.91%	30.77%	30.78%	31.17%	31.27%	31.01%	32.24%
OI	3493.45	3568.15	3498.01	3444.39	3570.86	3863.20	3492.75	3562.54	3535.08	3536.68	3590.63	3604.67	3579.30	3742.44
Delta	0.53	0.53	0.53	0.54	0.53	0.51	0.53	0.53	0.53	0.53	0.53	0.53	0.53	0.52

Table 8: Annualized excess returns

This table presents the average annualized returns for each predictive model over the benchmark AR(1) model. Annualized excess returns are obtained by multiplying daily excess returns by 252. Each panel corresponds to the results using different levels of transaction costs. $NW-t$ is the Newey-West adjusted t-statistic with lags set to five to account for autocorrelation. $Sent$, VIX, $Attn$, Vol, $Tbill$, and $Gold$ represent the extended models incorporating the corresponding predictor and AR(1) component. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	Sent	VIX	Attn	Vol	Tbill	Gold	Mean	Med.	TM	D(0.9)	D(0.95)	D(1.0)	All
Panel A. *Mid Price*													
Annualized return	-0.35%	3.43%	25.79%***	10.55%	22.48%**	0.36%	17.20%***	7.46%**	7.55%**	13.75%***	14.00%***	13.32%***	36.50%***
NW-t	-0.07	1.60	3.63	1.21	1.99	0.18	3.63	2.56	2.53	3.76	3.56	3.48	3.43
Daily max.	4.00%	3.12%	7.99%	11.68%	7.29%	3.12%	5.97%	5.04%	3.56%	5.82%	5.82%	5.82%	11.79%
Daily min.	-11.37%	-2.88%	-6.80%	-11.37%	-11.09%	-5.43%	-3.77%	-3.77%	-3.77%	-3.77%	-3.77%	-3.77%	-7.54%
Daily std.	0.61%	0.27%	0.85%	0.99%	0.94%	0.23%	0.55%	0.35%	0.36%	0.43%	0.45%	0.44%	1.11%
Panel B. *25% Effective Spread*													
Annualized return	3.37%	3.61%	22.70%***	14.21%	30.61%***	3.34%**	19.53%***	8.41%**	9.59%***	17.25%***	17.94%***	15.47%***	40.01%***
NW-t	0.63	1.51	3.18	1.48	2.61	2.15	3.60	2.50	2.67	3.92	3.90	3.50	3.28
Daily max.	4.37%	3.60%	8.95%	15.39%	7.95%	3.13%	7.64%	4.32%	5.22%	7.64%	7.64%	7.64%	15.55%
Daily min.	-8.22%	-2.71%	-6.28%	-8.22%	-10.51%	-1.93%	-4.67%	-4.67%	-4.67%	-4.67%	-4.67%	-4.67%	-6.94%
Daily std.	0.60%	0.29%	0.91%	1.06%	1.02%	0.18%	0.62%	0.40%	0.42%	0.52%	0.53%	0.51%	1.27%
Panel C. *50% Effective Spread*													
Annualized return	3.85%	4.24%	13.29%	15.28%	37.58%***	6.28%**	23.57%***	10.89%***	12.51%***	21.76%***	22.91%***	18.48%***	38.30%**
NW-t	0.64	1.48	1.47	1.37	2.69	2.57	3.56	2.58	2.83	4.01	4.10	3.43	2.47
Daily max.	5.85%	4.78%	12.39%	18.75%	9.51%	5.85%	9.46%	5.85%	6.88%	9.46%	9.46%	9.46%	18.95%
Daily min.	-6.75%	-3.25%	-10.74%	-10.27%	-12.26%	-0.49%	-5.58%	-5.58%	-5.58%	-5.58%	-5.58%	-5.58%	-11.01%
Daily std.	0.66%	0.35%	1.15%	1.22%	1.22%	0.29%	0.73%	0.50%	0.52%	0.63%	0.65%	0.63%	1.60%
Panel D. *100% Effective Spread*													
Annualized return	11.41%	4.79%	6.88%	23.96%	55.97%***	12.83%**	28.19%***	16.73%***	18.94%***	31.57%***	33.63%***	25.69%***	48.19%**
NW-t	1.33	1.19	0.44	1.53	3.16	2.42	2.94	2.63	2.94	4.05	4.22	3.30	2.17
Daily max.	10.60%	7.13%	13.10%	25.90%	13.15%	10.60%	13.10%	10.60%	10.60%	13.10%	13.10%	13.10%	25.48%
Daily min.	-9.67%	-4.33%	-19.66%	-17.72%	-17.57%	-1.98%	-7.39%	-7.39%	-7.39%	-7.39%	-7.39%	-7.39%	-20.44%
Daily std.	0.96%	0.48%	1.67%	1.74%	1.61%	0.60%	1.06%	0.75%	0.77%	0.90%	0.92%	0.91%	2.36%

Table 9: Alpha

This table reports the alphas from the one-factor and three-factor models: $R_{i,t} - R_{bmk,t} = \alpha_i + \beta_{MKT} MKT_t + \epsilon_{i,t}$, and $R_{i,t} - R_{bmk,t} = \alpha_i + \beta_{MKT} MKT_t + \beta_{SMB} SMB_t + \beta_{MOM} MOM_t + \epsilon_{i,t}$, where $R_{i,t}$ is the return from the i-th predictive model and $R_{bmk,t}$ is from the benchmark AR(1) model; MKT, SMB and MOM denote the market, size, and momentum factors of Liu et al. (2022). In the table, *Sent*, VIX, *Attn*, Vol, *Tbill*, and *Gold* represent the extended models incorporating the corresponding predictor and AR(1) component. The t-statistics adjusted by Newey-West standard errors with lags set to five are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	Sent	VIX	Attn	Vol	Tbill	Gold	Mean	Med.	TM	D(0.9)	D(0.95)	D(1.0)	All
						Mid Price							
One-factor alpha	-2.00%	2.03%	29.00%***	7.19%	17.22%	0.42%	17.49%***	4.88%	4.51%	12.54%***	12.75%**	12.25%**	40.23%***
	(-0.28)	(0.73)	(2.92)	(0.58)	(1.04)	(0.14)	(2.79)	(1.35)	(1.26)	(2.73)	(2.52)	(2.44)	(2.66)
Three-factor alpha	-7.03%	1.32%	30.88%***	-0.31%	13.98%	1.52%	14.32%**	3.85%	2.50%	10.47%**	10.51%**	9.38%*	33.20%**
	(-0.87)	(0.41)	(2.84)	(-0.02)	(0.74)	(0.56)	(2.23)	(0.89)	(0.63)	(2.31)	(2.04)	(1.85)	(2.11)
						25% Effective Spread							
One-factor alpha	3.64%	2.39%	22.06%**	8.48%	23.93%	4.65%**	18.07%***	3.61%	4.75%	14.21%***	15.06%***	12.14%**	41.14%**
	(0.52)	(0.79)	(2.21)	(0.64)	(1.42)	(2.16)	(2.62)	(0.96)	(1.23)	(2.72)	(2.69)	(2.22)	(2.42)
Three-factor alpha	-3.13%	1.98%	20.60%**	1.09%	19.66%	3.67%*	13.04%**	1.59%	2.02%	11.03%**	11.87%**	7.97%	30.67%*
	(-0.42)	(0.59)	(2.15)	(0.08)	(1.05)	(1.73)	(2.08)	(0.37)	(0.49)	(2.32)	(2.23)	(1.57)	(1.75)
						50% Effective Spread							
One-factor alpha	4.30%	3.39%	5.26%	5.71%	28.91%	8.85%***	21.48%***	4.59%	6.26%	17.42%***	18.95%***	13.37%**	34.30%
	(0.58)	(0.96)	(0.42)	(0.38)	(1.45)	(2.69)	(2.59)	(1.01)	(1.38)	(2.77)	(2.87)	(2.05)	(1.61)
Three-factor alpha	-2.75%	3.40%	4.38%	0.73%	23.81%	5.81%**	13.03%*	2.00%	3.18%	13.29%**	14.93%**	7.96%	22.81%
	(-0.36)	(0.86)	(0.36)	(0.05)	(1.04)	(2.31)	(1.87)	(0.42)	(0.70)	(2.47)	(2.53)	(1.42)	(1.01)
						100% Effective Spread							
One-factor alpha	15.61%	4.37%	-9.49%	9.46%	45.51%*	18.14%**	22.33%*	7.85%	10.12%	24.96%***	27.85%***	17.53%*	38.68%
	(1.48)	(0.90)	(-0.43)	(0.45)	(1.85)	(2.47)	(1.90)	(1.13)	(1.57)	(2.79)	(3.00)	(1.87)	(1.26)
Three-factor alpha	3.98%	5.12%	-16.00%	6.95%	38.21%	10.55%**	11.38%	3.20%	5.62%	18.45%**	21.68%***	8.86%	26.36%
	(0.39)	(0.90)	(-0.69)	(0.33)	(1.34)	(2.18)	(1.15)	(0.49)	(0.97)	(2.53)	(2.75)	(1.14)	(0.83)

Table 10: Annualized excess returns and alphas: Call options only

This table presents the average annualized returns for each predictive model over the benchmark AR(1) model and their alphas from the one-factor and three-factor models. We only use call options to construct the portfolios. Annualized excess returns are obtained by multiplying daily excess returns by 252. Each panel corresponds to the results using different levels of transaction costs. *Sent*, VIX, *Attn*, Vol, *Tbill* and *Gold* represent the extended models incorporating the corresponding predictor and AR(1) component. The t-statistics adjusted by Newey-West standard errors with lags set to five are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	Sent	VIX	Attn	Vol	Tbill	Gold	Mean	Med.	TM	D(0.9)	D(0.95)	D(1.0)	All
						Panel A. *Mid Price*							
Annualized return	9.19%*	0.29%	23.67%***	11.76%	-5.08%	0.17%	8.87%*	4.97%	7.26%**	4.81%	4.39%	9.14%**	29.17%*
	(1.76)	(0.08)	(3.01)	(0.90)	(-0.37)	(0.08)	(1.82)	(1.61)	(2.14)	(1.14)	(0.99)	(2.14)	(1.82)
One-factor alpha	11.75%	0.55%	27.80%**	12.27%	-16.95%	0.40%	10.03%	4.27%	7.81%*	2.93%	2.41%	9.24%	33.55%
	(1.58)	(0.11)	(2.44)	(0.67)	(-0.84)	(0.13)	(1.47)	(1.00)	(1.62)	(0.49)	(0.38)	(1.55)	(1.46)
Three-factor alpha	14.60%*	2.26%	35.16%***	10.65%	-20.16%	0.92%	8.82%	3.34%	6.53%	3.78%	2.87%	11.03%	29.80%
	(1.69)	(0.39)	(2.92)	(0.52)	(-0.84)	(0.48)	(1.11)	(0.72)	(1.34)	(0.55)	(0.40)	(1.59)	(1.13)
						Panel B. *25% Effective Spread*							
Annualized return	16.90%**	-2.47%	18.60%**	15.75%	5.40%	0.34%	14.56%**	6.59%*	9.35%**	10.16%*	11.53%**	18.62%***	32.24%*
	(2.41)	(-0.69)	(2.05)	(1.09)	(0.39)	(0.15)	(2.26)	(1.82)	(2.32)	(1.96)	(2.10)	(3.06)	(1.81)
One-factor alpha	22.33%**	-3.72%	21.26%	18.72%	-1.25%	0.62%	18.12%**	6.12%	10.70%*	9.94%	12.09%	22.07%***	40.07%
	(2.32)	(-0.71)	(1.61)	(0.93)	(-0.06)	(0.20)	(2.03)	(1.23)	(1.87)	(1.39)	(1.58)	(2.65)	(1.59)
Three-factor alpha	27.30%**	-2.13%	26.13%**	18.98%	-1.88%	0.57%	16.45%	4.69%	9.06%*	12.27%	14.18%	27.06%***	35.09%
	(2.32)	(-0.36)	(1.98)	(0.85)	(-0.08)	(0.35)	(1.52)	(0.95)	(1.66)	(1.41)	(1.56)	(2.64)	(1.23)
						Panel C. *50% Effective Spread*							
Annualized return	18.78%**	-5.32%	12.28%	16.21%	6.48%	-0.46%	17.65%**	5.61%	8.87%*	15.37%**	18.98%**	21.24%***	22.71%
	(2.02)	(-1.24)	(1.07)	(0.91)	(0.38)	(-0.19)	(1.97)	(1.30)	(1.87)	(2.14)	(2.42)	(2.60)	(1.02)
One-factor alpha	24.11%*	-8.14%	12.26%	20.37%	0.20%	-0.57%	22.22%*	4.01%	9.65%	16.75%*	22.22%**	24.54%**	27.24%
	(1.92)	(-1.29)	(0.73)	(0.83)	(0.01)	(-0.17)	(1.79)	(0.68)	(1.44)	(1.69)	(2.04)	(2.24)	(0.86)
Three-factor alpha	29.44%*	-6.39%	16.21%	21.33%	0.28%	-1.02%	19.79%	1.85%	7.40%	20.54%*	25.90%**	30.86%**	21.09%
	(1.83)	(-0.90)	(0.98)	(0.79)	(0.01)	(-0.59)	(1.30)	(0.33)	(1.20)	(1.70)	(2.00)	(2.23)	(0.60)
						Panel D. *100% Effective Spread*							
Annualized return	32.04%**	-8.83%	4.00%	22.85%	14.01%	1.39%	26.74%*	7.72%	11.81%*	25.96%**	32.83%**	35.87%**	25.05%
	(2.04)	(-1.43)	(0.20)	(0.94)	(0.53)	(0.46)	(1.77)	(1.21)	(1.73)	(2.14)	(2.40)	(2.55)	(0.70)
One-factor alpha	41.85%**	-13.80%	1.80%	31.35%	10.90%	1.89%	34.86%*	5.94%	13.49%	30.60%*	40.92%**	43.53%**	34.05%
	(1.97)	(-1.51)	(0.06)	(0.93)	(0.29)	(0.45)	(1.65)	(0.67)	(1.39)	(1.82)	(2.15)	(2.31)	(0.66)
Three-factor alpha	51.10%*	-11.87%	1.09%	33.56%	11.88%	-0.44%	32.03%	2.51%	10.32%	37.19%*	47.73%**	55.76%**	17.86%
	(1.85)	(-1.16)	(0.04)	(0.93)	(0.26)	(-0.17)	(1.26)	(0.30)	(1.19)	(1.83)	(2.13)	(2.32)	(0.32)

Table 11: Annualized excess returns and alphas: Put options only

This table presents the average annualized returns for each predictive model over the benchmark AR(1) model and their alphas from the one-factor and three-factor models. We only use put options to construct the portfolios. Annualized excess returns are obtained by multiplying daily excess returns by 252. Each panel corresponds to the results using different levels of transaction costs. *Sent*, VIX, *Attn*, Vol, *Tbill* and *Gold* represent the extended models incorporating the corresponding predictor and AR(1) component. The t-statistics adjusted by Newey-West standard errors with lags set to five are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	Sent	VIX	Attn	Vol	Tbill	Gold	Mean	Med.	TM	D(0.9)	D(0.95)	D(1.0)	All
Panel A. *Mid Price*													
Annualized return	-0.23%	-0.53%	2.22%	11.08%*	16.87%***	0.34%	1.60%	1.45%	6.55%	8.73%**	7.74%*	1.79%	21.00%**
	(-0.09)	(-0.39)	(0.63)	(1.87)	(2.93)	(0.73)	(0.70)	(0.64)	(1.54)	(2.19)	(1.90)	(0.97)	(2.52)
One-factor alpha	-1.56%	-0.29%	0.74%	13.96%	20.95%**	0.61%	0.74%	0.62%	8.11%	11.33%**	9.28%*	0.44%	27.94%**
	(-0.51)	(-0.14)	(0.16)	(1.64)	(2.56)	(0.87)	(0.28)	(0.24)	(1.41)	(2.10)	(1.68)	(0.24)	(2.24)
Three-factor alpha	-1.59%	-0.10%	0.64%	18.43%*	19.79%**	0.42%	0.73%	1.16%	8.96%	11.25%*	9.82%	0.23%	29.12%**
	(-0.51)	(-0.05)	(0.13)	(1.69)	(2.36)	(0.86)	(0.27)	(0.42)	(1.43)	(1.91)	(1.64)	(0.12)	(2.06)
Panel B. *25% Effective Spread*													
Annualized return	-1.08%	-0.68%	1.86%	8.11%	18.92%***	0.50%	1.11%	1.66%	5.49%	7.62%**	7.36%**	2.53%	20.66%**
	(-0.42)	(-0.43)	(0.47)	(1.63)	(3.30)	(0.74)	(0.43)	(0.65)	(1.43)	(2.22)	(2.06)	(1.32)	(2.56)
One-factor alpha	-1.91%	-0.55%	0.49%	9.86%	23.57%***	0.94%	0.21%	1.25%	6.66%	9.63%**	8.51%*	1.57%	26.97%**
	(-0.56)	(-0.22)	(0.09)	(1.40)	(2.89)	(0.91)	(0.06)	(0.38)	(1.27)	(2.07)	(1.78)	(0.70)	(2.24)
Three-factor alpha	-1.60%	-0.46%	1.00%	13.50%	22.51%***	0.64%	0.53%	1.84%	7.51%	9.37%*	8.52%*	1.03%	27.43%**
	(-0.47)	(-0.18)	(0.18)	(1.56)	(2.73)	(0.92)	(0.16)	(0.54)	(1.34)	(1.86)	(1.66)	(0.45)	(2.13)
Panel C. *50% Effective Spread*													
Annualized return	-2.22%	-0.83%	0.56%	3.91%	21.06%***	0.66%	0.48%	2.06%	3.23%	6.48%*	6.23%*	3.45%	18.70%**
	(-0.77)	(-0.44)	(0.11)	(0.83)	(3.36)	(0.73)	(0.15)	(0.68)	(0.81)	(1.92)	(1.75)	(1.55)	(2.14)
One-factor alpha	-2.69%	-0.81%	-1.13%	3.97%	26.38%***	1.28%	-0.53%	2.16%	3.46%	7.94%*	6.68%	2.97%	23.65%*
	(-0.67)	(-0.28)	(-0.16)	(0.60)	(2.98)	(0.92)	(-0.12)	(0.53)	(0.63)	(1.73)	(1.41)	(1.05)	(1.83)
Three-factor alpha	-1.97%	-0.82%	-0.02%	6.93%	25.41%***	0.85%	0.19%	2.94%	4.48%	7.34%	6.38%	2.24%	23.35%*
	(-0.51)	(-0.26)	(0.00)	(0.94)	(2.85)	(0.94)	(0.04)	(0.70)	(0.77)	(1.49)	(1.26)	(0.78)	(1.78)
Panel D. *100% Effective Spread*													
Annualized return	-3.12%	-1.13%	1.79%	-0.10%	26.04%***	0.98%	-0.01%	2.52%	0.95%	5.58%*	5.03%	4.96%	20.31%*
	(-0.83)	(-0.44)	(0.24)	(-0.02)	(3.19)	(0.73)	(0.00)	(0.59)	(0.22)	(1.72)	(1.41)	(1.53)	(1.81)
One-factor alpha	-2.19%	-1.33%	1.34%	-1.14%	33.20%***	1.94%	-0.86%	3.49%	0.52%	6.75%	4.71%	5.28%	25.34%
	(-0.40)	(-0.34)	(0.13)	(-0.16)	(2.93)	(0.93)	(-0.14)	(0.59)	(0.09)	(1.60)	(1.08)	(1.24)	(1.56)
Three-factor alpha	-0.29%	-1.53%	3.50%	1.00%	32.66%***	1.29%	1.13%	4.92%	2.26%	5.39%	3.98%	4.44%	23.43%
	(-0.06)	(-0.36)	(0.33)	(0.14)	(2.85)	(0.96)	(0.18)	(0.84)	(0.39)	(1.24)	(0.90)	(1.03)	(1.45)

Table 12: Fundamental vs. non-fundamental sentiment

This table reports the results of the bivariate and multiple regressions after decomposing the GPT sentiment index into a fundamental sentiment index and a non-fundamental sentiment index. The former relates to news that facilitates Bitcoin's exchange into other currencies or its use for purchasing goods and services, while the latter covers news that does not meet these criteria. The predictors include the first lag of implied volatility ($BVIX_{t-1}$), fundamental GPT index ($FSent_{t-1}$), non-fundamental GPT index($NFSent_{t-1}$), VIX index (VIX_{t-1}), investmor attention ($Attn_{t-1}$), trading volume (Vol_{t-1}), three-month U.S. Tbill rate ($Tbill_{t-1}$), and gold return ($Gold_{t-1}$). The t-statistics are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	$BVIX_t$	*Dependent variable:* $BVIX_t$							$BVIX_t$
Predictor	AR(1)	Bivariate regression							Multiple regression
IV_{t-1}	0.67***	0.67***	0.66***	0.67***	0.47***	0.56***	0.56***	0.67***	0.41***
	(30.31)	(30.20)	(29.98)	(30.04)	(17.81)	(23.01)	(22.50)	(30.29)	(14.98)
$FSent_{t-1}$		1.22*							0.67
		(1.67)							(0.95)
$NFSent_{t-1}$			1.99***						1.33*
			(2.64)						(1.83)
VIX_{t-1}				0.08					0.05
				(1.55)					(0.80)
$Attn_{t-1}$					12.91***				10.00***
					(12.22)				(7.67)
Vol_{t-1}						9.32***			2.81**
						(8.84)			(2.33)
$Tbill_{t-1}$							-2.04***		-1.22***
							(-9.09)		(-4.66)
$Gold_{t-1}$								-26.34	-19.05
								(-0.56)	(-0.44)
Constant	33.73***	31.94***	33.99***	32.23***	-54.89***	-180.40***	49.69***	33.76***	-91.54***
	(14.62)	(12.56)	(14.76)	(12.89)	(-7.25)	(-7.42)	(17.52)	(14.62)	(-3.64)
Observations	1,129	1,129	1,129	1,129	1,129	1,129	1,129	1,129	1,129
Adj. R^2	0.449	0.450	0.452	0.450	0.513	0.484	0.486	0.448	0.533

Appendix A

Table A1: Historical options data from LedgerX.

Date	Standardized_name	Last_bid	Last_ask	Open_interest
2020-01-01	BTC-25Jun2021-50000-p	409	427	100
2020-01-01	BTC-25Jun2021-10000-p	41.4	48.3	0
2020-01-01	BTC-25Jun2021-25000-p	171	184	0
2020-01-01	BTC-25Jun2021-10000-c	24.1	50.7	20487
2020-01-01	BTC-25Jun2021-25000-c	12	40.9	16050
2020-01-01	BTC-25Jun2021-50000-c	4.87	12.2	550
2020-01-01	BTC-18Dec2020-50000-p	415	427	100
2020-01-01	BTC-18Dec2020-2000-c	52.2	57.8	487
2020-01-01	BTC-18Dec2020-50000-c	0.14	4.8	11765
2020-01-01	BTC-18Dec2020-5000-p	5.81	9.34	2700
2020-01-01	BTC-18Dec2020-7500-c	27.3	37.4	11070
2020-01-01	BTC-18Dec2020-25000-p	173	182	100
2020-01-01	BTC-18Dec2020-10000-p	38.1	42.5	1700
2020-01-01	BTC-18Dec2020-20000-p	127	134	0
2020-01-01	BTC-18Dec2020-7500-p	20.5	24.2	1811
2020-01-01	BTC-18Dec2020-100000-p	902	923	0
2020-01-01	BTC-18Dec2020-40000-p	317	329	0
2020-01-01	BTC-18Dec2020-15000-p	79.4	85.5	525
2020-01-01	BTC-18Dec2020-25000-c	5.69	13.2	7950
2020-01-01	BTC-18Dec2020-40000-c	1.44	6.62	1602
2020-01-01	BTC-18Dec2020-5000-c	34.1	41.7	17300
2020-01-01	BTC-18Dec2020-15000-c	12.9	26.3	24430
2020-01-01	BTC-18Dec2020-10000-c	17.3	26.1	28051
2020-01-01	BTC-18Dec2020-20000-c	8.52	21.7	38505
2020-01-01	BTC-26Jun2020-5000-p	2.8	6.07	1900
2020-01-01	BTC-26Jun2020-10000-p	31.1	34.6	600
2020-01-01	BTC-26Jun2020-25000-c	2.55	3.43	40305
...	...	...	...	...