國立臺灣科技大學

財務金融研究所碩士班

碩士學位論文

學號：M11218014

比特幣選擇權隱含風險中立機率密度之平滑尾部導取

Extracting Smooth Tails of Option Implied Risk-neutral Densities in the Bitcoin Market

研究生：王士誠

指導教授：薛博今 博士

中　華　民　國　一　一　四　年　五　月

摘要

本研究旨在探討風險中立機率密度（Risk-neutral Density, RND）尾部配適方法對比特幣選擇權市場報酬率預測能力之影響，並提出創新的單點加斜率配適法，與 Birru 與 Figlewski (2012) 進行實證比較。

本研究以比特幣選擇權交易平台 Deribit 於 2021 年 1 月至 2024 年 4 月的交易數據為基礎，結合廣義柏拉圖分布（Generalized Pareto Distribution, GPD），建立完整的尾部延伸模型。透過比較距到期日 1 天與 7 天之選擇權樣本，檢驗 RND 統計特徵對現貨報酬率的預測能力。實證結果顯示，在計算效率方面，單點加斜率配適法較雙點配適法平均提升 10.95%。在日報酬率預測中，單點配適法建構之模型以偏態（Skewness）、中位數（Median）及前期報酬率為顯著預測變數（R-squared = 0.0130），表現優於雙點配適法（R-squared = 0.0098），此結果與 Conrad 等人 (2013) 及 Liu 與 Tsyvinski (2021) 關於偏態與動量效應的研究相呼應；在週報酬率預測方面，兩種方法表現接近，最佳預測變數組合為超額峰態（Excess Kurtosis）、中位數（Median）及加密貨幣恐懼與貪婪指數。超額峰態的預測力呼應 Amaya 等人 (2015) 研究結果，而恐懼與貪婪指數的顯著影響則與 He 等人 (2023) 的發現一致。樣本外預測結果顯示，無論在日報酬率或週報酬率預測中，兩種方法均產生正值的樣本外 R 平方，且單點配適法表現優於雙點配適法，顯示我們的預測模型具實質經濟價值 (Campbell & Thompson, 2008)。

本研究主要貢獻在於提出單點加斜率配適法，此方法提升 RND 尾部配適的效率與減少曲折，更驗證了 RND 統計特徵在高波動性市場中的應用價值。

關鍵字：加密貨幣、比特幣、選擇權、隱含波動率、風險中立機率密度、廣義柏拉圖分布

Abstract

This study proposes a methodological innovation for extracting smooth tails of risk-neutral density (RND) functions in Bitcoin options markets. We propose one single point with one slope method that addresses the kinks and computational inefficiencies in Birru and Figlewski (2012) method. While their approach requires matching conditions at two joining points per tail, the proposed method imposes cumulative probability matching and density slope continuity at a single joining point, enhancing computational efficiency by 10.95% while eliminating kinks.

Employing Deribit trading data (January 2021-April 2024) and integrating the Generalized Pareto Distribution (GPD), we demonstrate that RND moments exhibit significant predictive power for Bitcoin returns across different time horizons. For daily returns, the proposed method identifies skewness, median, and lagged returns as significant predictors (R-squared = 0.0130), outperforming Birru and Figlewski (2012) approach (R-squared = 0.0098). For weekly returns, excess kurtosis, median, and the Cryptocurrency Fear and Greed Index emerge as key predictors. These findings align with established research on skewness premium (Conrad et al., 2013), momentum effects (Liu & Tsyvinski, 2021), and sentiment indicators (M. He et al., 2023) in cryptocurrency markets.

Out-of-sample tests confirm the economic value of our approach with positive R-squared values across prediction horizons (Campbell & Thompson, 2008), providing valuable insights for risk management and investment strategies in cryptocurrency markets characterized by extreme volatility.

Keywords: Cryptocurrency, Bitcoin, Option, Implied Volatility, Risk-neutral Density, Generalized Pareto Distribution

Acknowledgement

Contents

List of Figures

List of Tables

1. Introduction

1.1 Research Background and Motivation

Risk assessment has remained a core focus in the financial industry. In particular, with the rapid expansion of the cryptocurrency ecosystem, Bitcoin has witnessed explosive growth in its derivatives market (Akyildirim et al., 2020). According to The Block (*The Block,* 2025), Bitcoin futures and options trading volume exceeded $21 trillion in 2024, with options trading growing at 130% annually, far outpacing traditional derivatives markets. This reflects increasing demand for cryptocurrency risk management tools while providing researchers a unique perspective on price discovery in emerging markets (Zulfiqar & Gulzar, 2021).

Cryptocurrency markets differ significantly from traditional financial markets in several aspects such as 24/7 trading, extreme volatility, decentralized structure, and unique investors' composition. Bitcoin's historical annualized volatility frequently exceeds 100%, substantially higher than the 15-20% volatility of traditional stock indices (Liu & Tsyvinski, 2021). This extreme volatility makes risk management crucial while exhibiting specific challenges for interpreting option price information.

As a derivative financial instrument, options contain rich market information in their prices. Call options grant holders the right to purchase the underlying asset at a predetermined price at a specific future time while put options give holders the right to sell. Their non-linear payoff structure reflects market expectations of future trends and volatility risk assessments (Hull, 2021). Option prices reveal market consensus on risk expectations and provide insights into market microstructure and investors' behavior (Bakshi et al., 2003). While the Black-Scholes model (1973) provides theoretical foundation for option pricing, the normal distribution assumption fails to capture the fat-

tailed distribution and negative skewness commonly observed in financial markets, particularly in high-volatility cryptocurrency markets (Chordia et al., 2021).

In the option pricing theory, the Black-Scholes model (1973) assumes geometric Brownian motion and applies no-arbitrage principles. Nevertheless, observed option prices often deviate from theoretical values, with implied volatility exhibiting systematic differences across strike prices, forming "volatility smile" or "volatility skew" (Rubinstein, 1994). This indicates that market expectations differ from the log-normal distribution assumed by Black-Scholes, particularly in the tail regions.

Extracting the risk-neutral density (RND) from option prices captures the market's complete expectation of future price distributions. According to Breeden and Litzenberger (1978), the RND can be derived by taking the second derivative of option prices with respect to strike prices. To extract the RND, implied volatility is converted from a finite number of observed market option prices through the Black-Scholes model. To obtain the continuum of implied volatility curve, Hagan and West (2006) indicated that quadratic splines can avoid the risk of overfitting while maintaining curve smoothness to ensure model robustness and reliability, particularly in the case of high market volatility. Haslip and Kaishev (2014) showed that when dealing with complex derivative financial products such as lookback options, quadratic splines combined with Fourier transforms can provide efficient and accurate pricing results, achieving a good balance between computational efficiency and precision, though with the risk of discontinuous first derivatives. Bliss and Panigirtzoglou (2004), Figlewski (2008) and Monteiro, Tütüncü, and Vicente (2008) demonstrated the interpolation of implied volatility using cubic spline functions. The cubic spline method has the advantages of high computational efficiency and relatively simple implementation, providing reasonable fitting results under most market conditions. Nevertheless, cubic splines only

guarantee the continuity of the first derivative, while the second derivative may reveal discontinuous at certain nodes, resulting in non-smooth fitting when dealing with extreme volatility. Since cubic splines may lead to insufficiently smooth RNDs, quartic spline functions with a single knot have been extensively adopted for implied volatility curve fitting (Figlewski, 2008; Birru & Figlewski, 2012; Reinke, 2020).

Due to limited market option prices, especially in deep out-of-the-money regions, RND extraction often lacks sufficient tail information. Previous research has developed two major streams of the RND tail-fitting methods: non-parametric and parametric approaches. Non-parametric methods make minimal assumptions about the underlying asset. Bondarenko (2000) proposed a non-parametric method for deriving the RND from option prices, finding daily RND changes correlate with index performance. Grith, Härdle, and Schienle (2012) explored kernel smoothing and spline functions in the RND estimation, highlighting their flexibility in capturing complex distributional features such as skewness and multimodality. Monteiro and Santos (2022) addressed local constraint limitations in kernel-based estimation by imposing broader no-arbitrage constraints using the Heston model. Dong, Xu, and Cui (2024) introduced the Implied Willow Tree method, reconstructing complete risk-neutral processes directly from cross-maturity option data without preset parametric models.

Parametric methods commonly employ Extreme Value Theory (EVT) to extend RND tails. Figlewski (2008) introduced Generalized Extreme Value Distribution (GEV) fitting requiring three continuity conditions per tail: matching cumulative probabilities at the first joining point and equal density values at both first and second joining points. Birru and Figlewski (2012) further replaced GEV with Generalized Pareto Distribution (GPD), discovering significant left skewness in the RND regardless of market volatility. Their findings confirmed that even during extreme market turbulence, the mean of RND

remains close to the futures price, indicating effective no-arbitrage relationships. Other approaches include mixed distribution methods, which decompose the RND into core and tail components. Glatzer and Scheicher (2005) combined log-normal distributions for the core with GPD for the tails, demonstrating effectiveness in Eurozone bond markets. Markose and Alentorn (2011) parameterized the RND tails using Generalized Extreme Value Family to capture market expectations of extreme events. Monteiro et al. (2008) proposed density function extrapolation using cubic spline functions with non-negativity constraints and exponential functions for the tail extrapolation, though computationally simple but potentially limited in capturing complex tail features. Recent developments include novel parametric methods that do not rely on extreme value theory. Orosi (2015) established appropriate functional forms with parameter constraints to produce well-behaved risk-neutral density estimates. Uberti (2023) developed a semi-parametric estimation method combining parametric stability with non-parametric flexibility. Y. Li, Nolte, and Pham (2024) introduced a Lognormal-Weibull mixture model offering improved performance when measuring skewness and analyzing multi-peak RNDs, demonstrating the continuous refinement in estimation methodology.

On top of implied volatility, RND contains vastly useful information in higher-order moments (i.e. skewness and excess kurtosis), which have been broadly used to predict asset returns. Bali and Murray (2013) and Conrad, Dittmar, and Ghysels (2013) found a significant negative relationship between risk-neutral skewness and future stock returns. Similar findings also appear in commodity futures (Fuertes et al., 2022), oil markets (Cortés et al., 2020), and foreign exchange (Chen et al., 2018). Recently, Böök, Imbet, and Reinke (2025) demonstrated that volatility, skewness, and kurtosis derived from options data outperform historical return-based indicators in predicting stock risk premiums. While these relationships are established in traditional markets, research in

cryptocurrency markets remains limited.

This study extends the framework of Birru and Figlewski (2012) to derive the entire RND since GPD is particularly adept at capturing the behavior of the extreme events. Nonetheless, certain constraints are observed in Birru and Figlewski (2012) when applied to cryptocurrency markets. First, Birru and Figlewski (2012) can produce discontinuities or kinks at the connection points. Also, it increases computational complexity, potentially affecting estimation stability.

To mitigate the limitations in Birru and Figlewski (2012), we propose a method which fits each tail of the RND using one single point with one slope to ensure the continuity without kinks for both the cumulative distribution function and density function. The proposed method imposes two essential conditions. First, at the joining point, the cumulative probability of GPD must equal that of the empirical RND. Second, the slope of the GPD density function must match the slope of the empirical RND at the same point. To further validate the effectiveness of the proposed method, this study examines the predictive power of the RND moments, such as skewness and excess kurtosis, for Bitcoin returns, in comparison with Birru and Figlewski (2012). Based on Bitcoin options trading data from the Deribit platform from January 2020 to April 2024, the proposed method significantly outperforms Birru and Figlewski (2012) in terms of daily returns and slightly surpasses in weekly returns. Further, the robustness tests using rolling window estimation (Campbell & Thompson, 2008) also confirms the proposed method perform superior, compared with Birru and Figlewski (2012).

In the daily return prediction, skewness, median, and lagged returns are identified as significant predictors, with superior explanatory power (R-squared = 0.0130), compared to Birru and Figlewski (2012) (R-squared = 0.0098). The skewness shows a negative effect on daily returns, in alignment with Conrad et al. (2013), indicating a phenomenon

where market participants assign greater premiums (positive returns) to downside risk (negative skewness). The significant predictive power of lagged returns supports Liu and Tsyvinski's (2021) argument about momentum effects in cryptocurrency markets. For weekly returns, both methods identify excess kurtosis, median, and the Crypto Fear and Greed Index as optimal predictors. Excess kurtosis exhibits a significant negative effect on returns, consistent with Amaya, Christoffersen, Jacobs, and Vasquez (2015), indicating that heightened tail risk leads to subsequent return reversals. The Crypto Fear and Greed Index shows a significant negative relationship with future returns, supporting M. He, Shen, Yaojie Zhang, and Yi Zhang (2023) that market sentiment serves as a contrarian indicator in cryptocurrency markets.

This research makes two main contributions to the literature as follows:

1. The proposed method ensures continuity in RND tail fitting through first-derivative continuity conditions while reducing computational complexity, achieving 10.95% greater computational efficiency compared to Birru and Figlewski (2012).

2. We pioneer the application of RND characteristics to multi-time scale prediction in cryptocurrency markets. While the predictive power of higher-order RND moments has been established in traditional markets (Chang et al., 2013; Conrad et al., 2013), their application in highly volatile cryptocurrency markets remains underexplored. We find that daily return prediction primarily relies on skewness, median, and lagged returns, whereas weekly prediction is influenced by excess kurtosis and market sentiment indicators. These differential findings provide practical tools for cryptocurrency risk management.

1.2 Thesis Structure

This dissertation comprises six chapters. Chapter 1 introduces the research background, motivation and findings. Chapter 2 reviews theoretical foundations of risk-neutral density and relevant empirical studies. Chapter 3 describes data sources and market context. Chapter 4 details the methodological framework, including the proposed tail-fitting method and empirical model specifications. Chapter 5 presents empirical results and comparative analysis of different methods. Chapter 6 concludes with key findings and implications for future research.

2. Literature Review

2.1 Fundamental Theory of the Risk-neutral Density (RND)

The risk-neutral density (RND) concept, pioneered by Breeden and Litzenberger (1978), demonstrated that market expectations about future price distributions could be extracted from option prices. Shimko (1993) enhanced this methodology by converting option prices into implied volatility space for interpolation, leveraging the smoother characteristics of volatility curves to improve RND estimation accuracy.

Christoffersen, Jacobs, and Chang (2013) established that risk-neutral skewness effectively predicts future return direction and magnitude, while Chang, Christoffersen, and Jacobs (2013) found higher risk-neutral kurtosis often precedes greater market volatility. For extreme market conditions, Birru and Figlewski (2012) documented significant RND shape transformations during the 2008 financial crisis, implementing the Generalized Pareto Distribution (GPD) for more accurate tail estimation. Jackwerth (2020) provided evidence that markets require time to fully incorporate major events, contributing valuable insights on market information efficiency.

2.2 Methods for Deriving the Risk-neutral Density

Figlewski (2008) developed a comprehensive framework categorizing RND estimation into parametric and non-parametric approaches. McNeil and Frey (2000) combined GARCH models with Extreme Value Theory (EVT) for financial time series tail risk estimation, highlighting GPD's advantages in modeling extreme events.

Methodological advancements include Orosi's (2015) parametric estimation

technique utilizing constrained functional forms, He, Peng, Zhang, and Zhao's (2022) GPD application for tail estimation, and Uberti's (2023) semi-parametric approach combining parametric stability with non-parametric flexibility. Ammann and Feser (2019) developed robust estimation methods for risk-neutral moments that reduce bias under market noise and liquidity constraints, while Hayashi (2020) proposed a method for analyzing RNDs from volatility smiles that eliminates numerical approximation errors.

Recent innovation comes from Dong, Xu, and Cui (2024) with their Implied Willow Tree (IWT) method, which reconstructs complete risk-neutral processes from cross-maturity option data without preset parametric models, demonstrating effectiveness in pricing complex options and handling noisy data.

2.3 Moments of the RND

2.3.1 Higher-Order Moments

Higher-order moments, particularly skewness and kurtosis, play critical roles in financial research. Bali and Murray (2013) and Conrad et al. (2013) documented risk-neutral skewness's significant negative relationship with future stock returns, aligning with investor preference for positive skewness. Kim and Park (2018) confirmed this relationship persists after controlling for firm characteristics.

Mei, Liu, Ma, and Chen (2017) found realized skewness and kurtosis negatively affect future volatility, with skewness outperforming kurtosis in medium to long-term predictions. Fuertes, Liu, and Tang (2022) demonstrated risk-neutral skewness's importance in commodity futures pricing, with strategies based on RNSK values generating significant excess returns, particularly in contango markets.

Cortés, Mora-Valencia, and Perote (2020) showed log-SNP distributions more

accurately capture oil price RNDs than traditional log-normal distributions, with skewness and kurtosis containing valuable market expectation information. Recent work by Böök et al. (2025) derived robust conditional volatility, skewness, and kurtosis indicators from S&P 500 options that outperform historical return-based indicators in predicting equity risk premiums.

2.3.2 Tail Risk and Extreme Value Theory

Balkema and de Haan's (1974) threshold exceedance model established that samples exceeding sufficiently high thresholds converge to the Generalized Pareto Distribution, foundational for financial market tail risk estimation. Wang and Yen (2018) found option-implied tail risk indicators effectively predict underlying asset movements, particularly during recessions.

Chen, Hsieh, and Huang (2018) documented higher-order RND moments' explanatory power for crash risk and risk premiums, with skewness positively correlating with risk premiums and kurtosis with foreign exchange swap spreads. Lehnert (2022) challenged traditional views by showing short-selling in options markets creates a negative relationship between risk-neutral market skewness and returns.

Conrad et al. (2013) extended methods combining GARCH with mixed normal distributions to capture asymmetric volatility, while Neumann and Skiadopoulos (2013) found market expectations regarding volatility, skewness, and kurtosis exhibit significant predictability, especially during heightened market volatility.

2.4 Empirical Applications in Financial Markets

2.4.1 Traditional Financial Markets

Mohrschladt and Schneider (2021) revealed in-the-money options contain valuable market information through high-frequency trading data analysis. Li, Wu, H. Zhang, and L. Zhang (2024) demonstrated risk-neutral skewness's predictive power for future stock returns, particularly during recessions, consistent with Cujean and Hasler's (2017) theoretical predictions.

Feng, He, and Zhang (2024) established strong associations between market sentiment and option-implied volatility during uncertainty periods, while Köse et al. (2024) found institutional investor behavior significantly impacts RND shapes. Amaya et al. (2015) documented realized skewness's explanatory power for cross-sectional stock returns, persisting after controlling for risk factors.

Bali and Zhou (2016) identified significant associations between market uncertainty and expected returns during heightened macroeconomic uncertainty. Jondeau, Wang, Yan, and Zhang (2020) demonstrated individual stock skewness effectively predicts S&P 500 index futures returns, persisting after controlling for liquidity risk and economic cycles.

2.4.2 Cryptocurrency Market

Zulfiqar and Gulzar (2021) noted cryptocurrency exchange options provide diverse hedging instruments, while Baur and Smales (2022) found leveraged fund traders maintain key roles and net short positions in Bitcoin futures markets, accurately predicting major market fluctuations.

López-Cabarcos, Pérez-Pico, Piñeiro-Chousa, and Šević (2021) established social

media sentiment indicators' predictive power for short-term Bitcoin price movements. Chordia, Lin, and Xiang (2021) documented significant left skewness and excess kurtosis in Bitcoin options' RND, while Akyildirim, Corbet, Lucey, Sensoy, and Yarovaya (2020) found cryptocurrency volatility increases with investor fear sentiment, correlating with traditional market volatility indicators.

Liu and Tsyvinski (2021) identified significant momentum effects in cryptocurrency markets, with Li, Urquhart, Wang, and Zhang (2021) confirming particularly strong MAX momentum effects. Liu and Chen (2024) documented market capitalization-dependent skewness patterns, with asymmetric risk negatively correlating with future returns. Liu, Li, Nekhili, and Sultan (2023) employed machine learning to confirm lagged returns' strong predictive power for cryptocurrency returns.

3. Data

3.1 Data

This study utilizes historical trading data from Deribit exchange (*Deribit*, 2025), the dominant platform in the global cryptocurrency options market. The dataset spans January 2020 to April 2024, encompassing daily trading volume, closing prices, implied volatility, spot prices, futures prices, and related trading metrics.

According to The Block (*The Block*, 2025), Deribit commands over 80% market share in open interest among major trading platforms (Deribit, OKX, and Binance), attributable to its established operational history and strategic market development (as shown in Figure 3-1). Founded in 2016 in the Netherlands, Deribit pioneered professional cryptocurrency options trading, with its name reflecting the fusion of "Derivatives" and "Bitcoin."

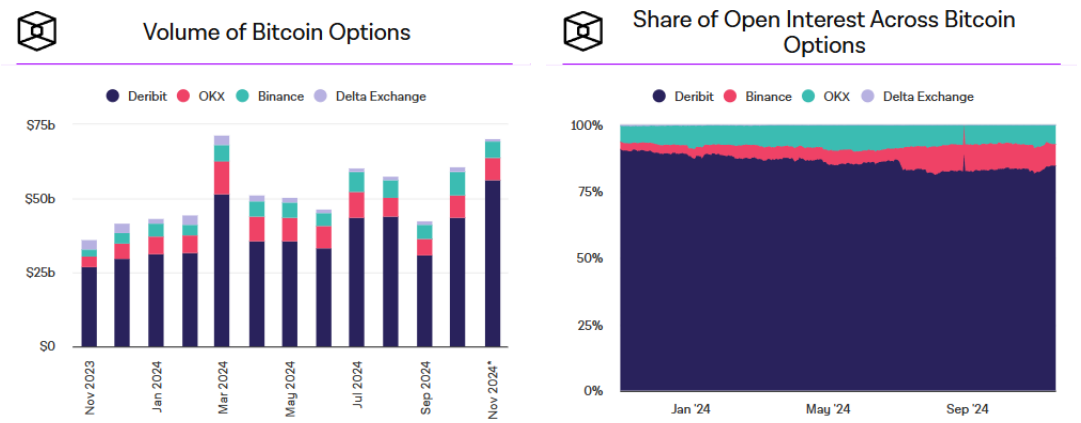

Figure 3-1: Statistics on Bitcoin Options Market Trading Volume and Open Interest (Data Source: The Block Official Website)

Deribit's Bitcoin European cash-settled options operate on a continuous 24/7 trading mechanism with uniform expiration at 08:00 UTC (*Deribit Options*, 2025). The exchange offers a comprehensive product matrix including short-term contracts (1-day, 2-day, 3-day), medium-term contracts (1-week, 2-week, 3-week), month-end expiration contracts

(January, February, April, May, July, August, October, November), and quarterly expiration contracts (March, June, September, December). This diversified product architecture satisfies varied investor requirements while enhancing market liquidity and price discovery efficiency, with particularly significant market participation growth following the October 2020 introduction of daily and weekly expiration products.

3.2 Overview of Bitcoin Options Trading Market

Transaction counts (Figure 3-2) and trading volumes (Figure 3-3) have grown substantially since October 2020, coinciding with Deribit's product diversification strategy that introduced daily and weekly expiration products. Market activity surged again in late 2023, with call option volume reaching historic highs in February 2024 amid rising Bitcoin prices, reflecting market optimism. The increasing trading volume since H2 2023 indicates improved market liquidity and depth, enhancing price discovery efficiency and attracting institutional participation.

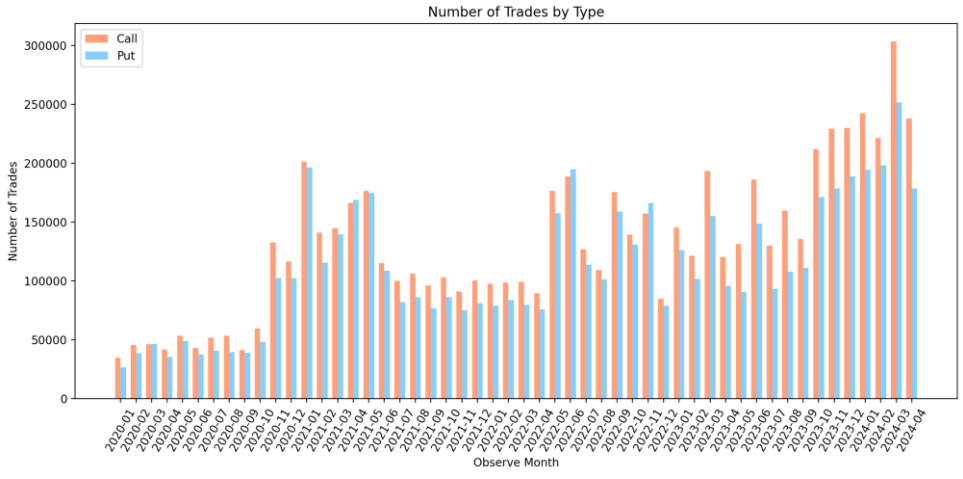

Figure 3-2: Monthly Transaction Counts of Bitcoin Call and Put Options from January 2020 to April 2024

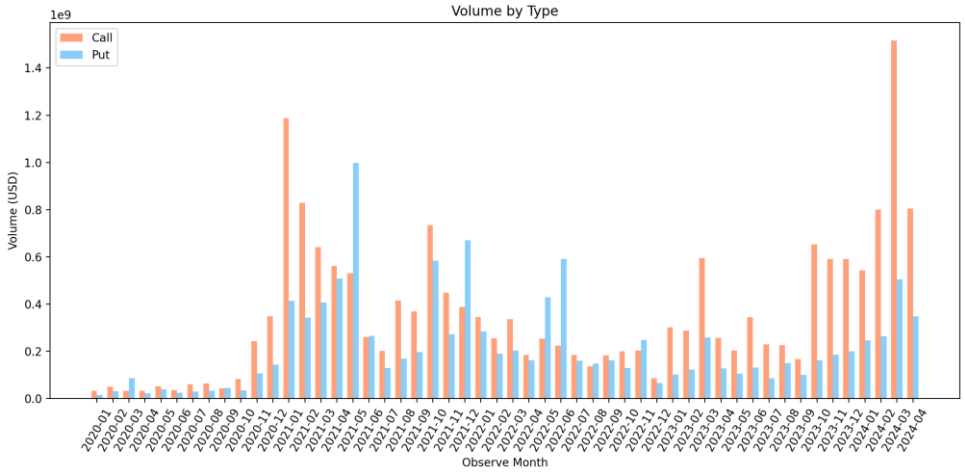

Figure 3-3: Monthly Trading Volume of Bitcoin Call and Put Options from January 2020 to April 2024

Trading activity heat maps reveal distinct patterns. Call options trading (Figure 3-4) concentrates around at-the-money positions (Moneyness ratio $\left(\frac{StrikePrice(K)}{SpotPrice(S)}\right)$ 0.9-1.1), with slightly out-of-the-money options (1.0-1.1) recording the highest volume ($2,072M), indicating investor preference for leveraged positions. Short-term call options (31-90 days) maintain substantial volume across various moneyness levels, while extremely short-term options (<14 days) show significant at-the-money activity, reflecting speculative demand. Long-term (>180 days) call options exhibit limited trading except for anomalous volume ($392M) in deep out-of-the-money positions (>2.0), likely reflecting institutional hedging strategies.

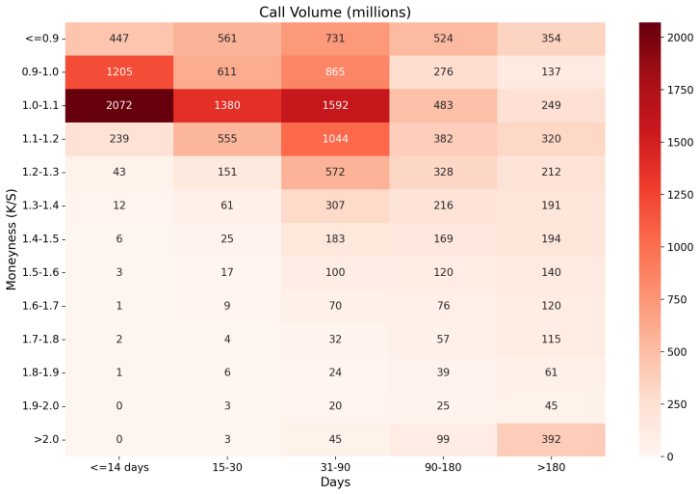

Figure 3-4: Heat Map of Total Bitcoin Call Options Trading Volume from January 2020 to April 2024

Put options (Figure 3-5) demonstrate different characteristics, with trading heavily concentrated near at-the-money (0.9-1.0) positions, peaking at $1,676M. Deep out-of-the-money puts (≤0.7) show limited activity, suggesting minimal demand for protection against significant price declines. Short-term puts are most active near at-the-money, while medium-term puts (90-180 days) show substantial volume ($821M) in deep in-the-money positions (>1.1). Long-term puts (>180 days) exhibit relatively uniform volume distribution across moneyness levels.

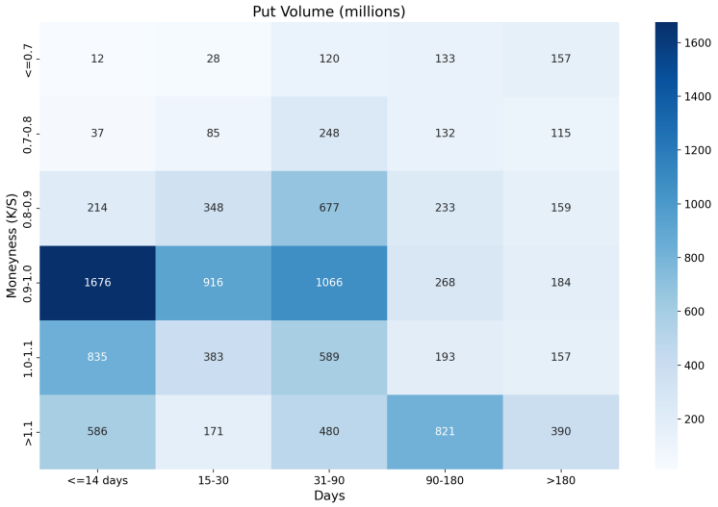

Figure 3-5: Heat Map of Total Bitcoin Put Options Trading Volume from January 2020 to April 2024

The market demonstrates considerable depth with trading concentrated around at-the-money positions and preference for short-term strategies. Higher call option volumes relative to puts reflect bullish market sentiment, while anomalous trading patterns in specific segments likely represent specialized institutional strategies. Despite its significant growth, the market remains characterized by predominantly short-term speculative trading behavior.

4. Research Methodology

4.1 Deriving the Risk-neutral Density (RND)

In the following text, symbols C, P, S, K, and T represent as follows: C is call option price, P is put option price, S is underlying asset current price, K is strike price, r is risk-free rate, T is days to option expiration. This research also uses $f(K)$ to represent the Risk-neutral Density Function (RND) and $F(K) = \int_{-\infty}^{K} f(z)dz$ to represent the Risk-neutral Distribution Function.

The call option price is the expected payoff before its expiration day T, discounted back to the present value. Under risk-neutral conditions, this expected price can be calculated based on risk-neutral probability and discounted using the risk-free rate, as follows:

$$C = \int_{K}^{\infty} e^{-rT} (S_T - K) f(S_T) dS_T \tag{1}$$

Next, by taking the first partial derivative of call option price with respect to strike price, we can derive the risk-neutral distribution function $F(K)$, as follows:

$$\begin{aligned}
\frac{\partial C}{\partial K} &= \frac{\partial}{\partial K}\left[\int_{K}^{\infty} e^{-rT} (S_T - K) f(S_T) dS_T\right] \\
&= e^{-rT}\left[-(K-K)f(K) + \int_{K}^{\infty} -f(S_T) dS_T\right] \\
&= -e^{-rT}\int_{K}^{\infty} f(S_T) dS_T \\
&= -e^{-rT}[1 - F(K)]
\end{aligned}$$

Rearranging terms, we obtain the risk-neutral distribution function $F(K)$:

$$F(K) = e^{rT}\frac{\partial C}{\partial K} + 1 \tag{2}$$

Then, by taking another partial derivative of equation (2) with respect to strike price, we can derive the RND at strike price:

$$f(K) = \frac{\partial}{\partial K}\left[e^{rT}\frac{\partial C}{\partial K} + 1\right] = e^{rT}\frac{\partial^2 C}{\partial K^2} \tag{3}$$

In actual options trading markets, since strike prices are in discrete form, we can use observed option prices and apply Finite Difference Methods (FDM) to obtain approximate solutions to equations (2) and (3). Assuming that at time T to expiration, there are N options with different strike prices in the market, where K_1 represents the lowest strike price and K_n represents the highest strike price. We uses three options with strike prices K_{n-1}, K_n and K_{n+1} to calculate the approximate value centered at K_n, as follows:

$$F(K_n) \approx e^{rT}\left[\frac{C_{n+1} - C_{n-1}}{X_{n+1} - X_{n-1}}\right] + 1 \tag{4}$$

$$f(K_n) \approx e^{rT}\frac{C_{n+1} - 2C_n + C_{n-1}}{(\Delta X)^2} \tag{5}$$

Equations (1) to (5) explain how to theoretically derive the RND between strike prices K_2 and K_{n-1} from a set of call option prices. Similar derivation methods can also be applied to extract the RND from put option prices. For put options, the equivalent expressions corresponding to equations (2) to (5) are as follows:

$$F(K) = e^{rT}\frac{\partial P}{\partial K} \tag{6}$$

$$f(K) = e^{rT} \frac{\partial^2 P}{\partial K^2} \tag{7}$$

$$F(K_n) \approx e^{rT} \left[\frac{P_{n+1} - P_{n-1}}{X_{n+1} - X_{n-1}} \right] \tag{8}$$

$$f(K_n) \approx e^{rT} \frac{P_{n+1} - 2P_n + P_{n-1}}{(\Delta X)^2} \tag{9}$$

In this research, ΔX is a fixed constant value used to construct artificially spaced option prices to fill gaps between discrete strike prices in the market. This approach addresses the problem of sparse or uneven trading data and ensures consistent spacing between strike prices, facilitating numerical calculations through finite difference methods and improving the accuracy of estimation.

4.2 Deriving the RND Using Bitcoin Options

4.2.1 Applying the Black-Scholes Model

In traditional financial markets, option pricing models (such as the Black-Scholes model) typically use risk-free interest rates as parameters, which are generally represented by the yields of low-risk assets such as government bonds. Nevertheless, in cryptocurrency markets such as Bitcoin, the applicability of risk-free interest rates is limited and therefore not widely used. This is because the Bitcoin market lacks a unified risk-free asset. Due to the decentralized nature of cryptocurrency markets, there are no widely accepted risk-free assets such as government bonds, making it difficult to determine a single universal risk-free rate, thus limiting its applicability in this market.

Additionally, Bitcoin price volatility is far higher than traditional assets. This highly volatile characteristic has a more significant impact on option prices than risk-free interest

rates, causing traders to focus more on changes in implied volatility rather than risk-free rates. Furthermore, in cryptocurrency markets, the interest rate environment may be influenced by exchange rules and market supply and demand, not necessarily related to traditional risk-free rates, making traditional interest rate indicators difficult to reflect the actual situation in cryptocurrency markets. Moreover, the cost of holding Bitcoin differs from the cost of holding traditional currencies or assets, including security aspects and technological risks, which are difficult to quantify through risk-free interest rates, further limiting the applicability of risk-free rates in Bitcoin option pricing.

This research uses Bitcoin option trading prices from the Deribit exchange, which has adopted a more suitable model (priced in Bitcoin) to compute option prices, adapting to the characteristics of the cryptocurrency market and meeting trading market needs. To meet research requirements, this study observes the traditional Black-Scholes model (Equation (10)) and compares it with the calculation formula provided by Deribit exchange (Equation (11)). It can be seen that multiplying the Deribit exchange quote $C_{Deribit}$ by the Bitcoin spot price S_0 yields the Bitcoin option price denominated in US dollars.

$$C_{BS} = S_0 \times N(d_1) - Ke^{-rT} \times N(d_2)$$

$$\Rightarrow C_{BS} = S_0 \times \left[N(d_1) - \frac{Ke^{-rT}}{S_0} \times N(d_2) \right] \; and \; F = S_0 e^{rT}$$

$$\Rightarrow \frac{C_{BS}}{S_0} = N(d_1) - \frac{K}{F} \times N(d_2) \tag{10}$$

$$C_{Deribit} = N(d_1) - \frac{K}{F} \times N(d_2) \tag{11}$$

Where $d_1 = \frac{\ln\left(\frac{F}{K}\right) + \left(\frac{\sigma^2}{2}\right) \times T}{\sigma \times \sqrt{T}}$, $d_2 = d_1 - \sigma \times \sqrt{T}$, C_{BS} is the Black-Scholes call

option price (denominated in USD), $C_{Deribit}$ is the Deribit exchange Bitcoin call option price (denominated in Bitcoin), S_0 is the Bitcoin spot price, $N(x)$ is the cumulative distribution function (CDF) of the normal distribution, K is the strike price, r is the risk-free interest rate, T is the option's time to expiration, F is the Bitcoin futures price, ln is the natural logarithm, σ is the annualized standard deviation.

4.2.2 Calculating Bitcoin Option Implied Volatility

In the Bitcoin options market, traders are predominantly risk-seeking and tend to operate with out-of-the-money (OTM) options, primarily due to their lower cost, high leverage effect, and particularly strong sensitivity to volatility. For buyers, given the extremely high volatility of the Bitcoin market itself, these options are highly attractive to speculators and high-risk-preferring investors. Despite the higher probability of these options expiring worthless, traders are still willing to take on such risks. For sellers, since Bitcoin price volatility is significantly higher than traditional financial markets, the premium levels for OTM options are usually higher, further enhancing the incentives for seller participation, making OTM options a core tool for many sellers to create stable cash flow. In summary, OTM options have better trading volume and liquidity, and their prices can more efficiently reflect market sentiment. This research uses OTM option trading data to compute implied volatility; nevertheless, to avoid anomalies caused by unreasonable trades in deeply out-of-the-money areas, options data with strike prices below $10 is excluded.

Shimko (1993) proposed converting market option prices to implied volatility (IV) before interpolation, as implied volatility curves are typically smoother and more continuous than price data, making them suitable for interpolation and smoothing processes. The interpolated curves are then converted back to call option prices to extract

the RND. This method does not rely on option prices conforming to Black-Scholes model assumptions, but rather uses the Black-Scholes formula merely as a calculation tool to convert data into a form more suitable for smoothing.

Figlewski (2008) proposed a method aimed at resolving abnormal fluctuations in implied volatility data near at-the-money options, particularly the jump phenomenon in call and put option prices near the at-the-money point. Such jumps may lead to non-smooth implied volatility curves, thereby affecting the stability of the RND extraction. Following this method, if a strike price is between 0.9 and 1.1 times the futures price, the average of call and put implied volatilities is taken as the data point. The formula is as follows:

$$IV_{mix}(K) = 0.5 \times IV_{call}(K) + 0.5 \times IV_{put}(K)$$

$$K \in [0.9 \times F, 1.1 \times F]$$

As shown in Figure 4-1 through Figure 4-3, this approach effectively reduces fluctuation amplitude near the at-the-money point while maintaining OTM option data outside this range.

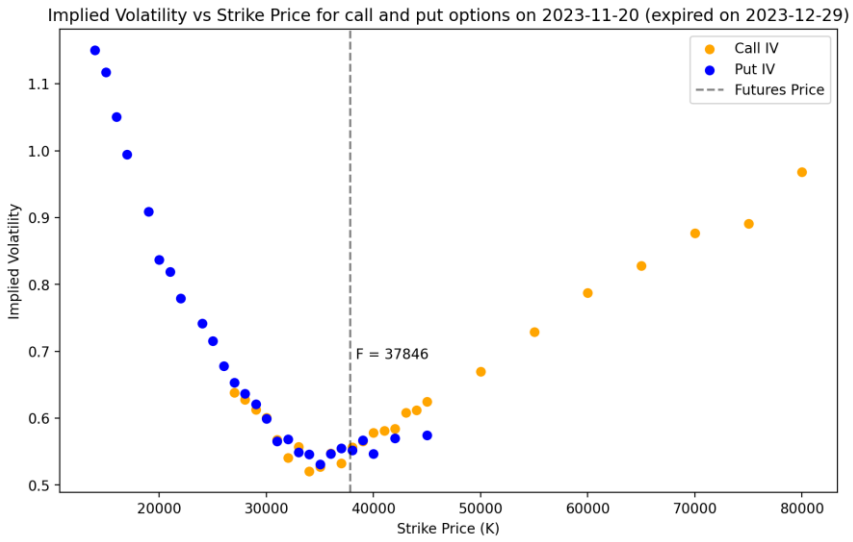

Figure 4-1: Bitcoin Option Implied Volatility Distribution on November 20, 2023 (Expiring December 29, 2023)

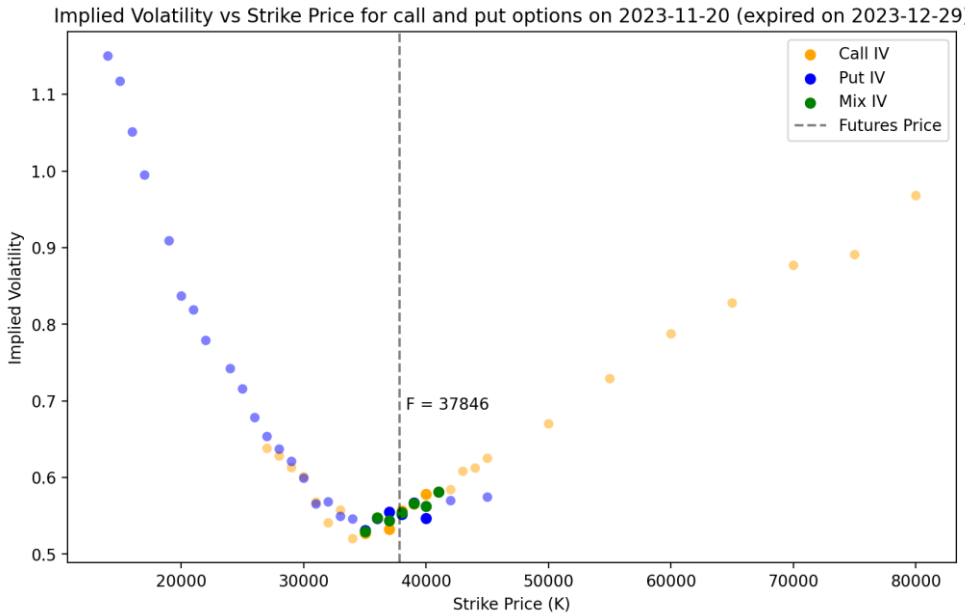

Figure 4-2: Bitcoin Option Implied Volatility Distribution on November 20, 2023 (Expiring December 29, 2023)

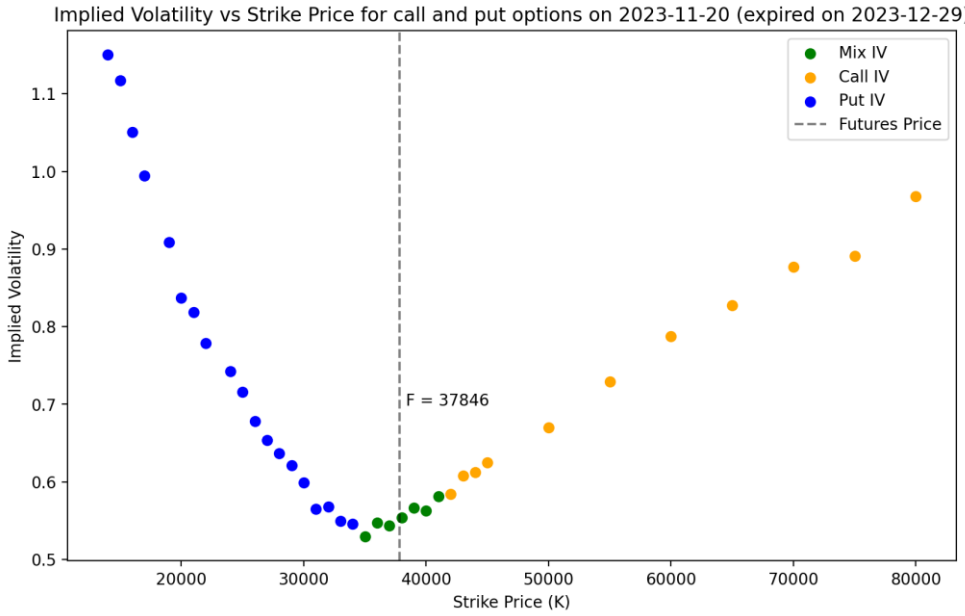

Figure 4-3: Bitcoin Option Implied Volatility Distribution on November 20, 2023 (Expiring December 29, 2023)

4.2.3 Fitting Bitcoin Option Implied Volatility Curve

To more precisely fit the processed implied volatility data, this research adopts a 4th-order spline function with a single knot for curve fitting. The knot is placed at the futures price, a design that allows greater flexibility at this key position while maintaining the overall continuity of the curve. Using a 4th-order spline function ensures that the fitted curve has third-order continuous differentiability (C^3 continuity), effectively capturing subtle changes in the implied volatility curve while avoiding over-fitting problems.

The mathematical representation of the 4th-order spline function is as follows:

$$S(x) = \begin{cases} \sum_{i=0}^{4} a_i (x - x_0)^i, & x < k \\ \sum_{i=0}^{4} b_i (x - x_0)^i, & x \geq k \end{cases}$$

Where k is the knot position (i.e., the futures price), x_0 is the reference point, and a_i and b_i are coefficients to be determined. At the knot, the function must satisfy the following continuity conditions:

$$\begin{cases} \sum_{i=0}^{4} a_i (k - x_0)^i = \sum_{i=0}^{4} b_i (k - x_0)^i \\ \sum_{i=1}^{4} i\, a_i (k - x_0)^{i-1} = \sum_{i=1}^{4} i\, b_i (k - x_0)^{i-1} \\ \sum_{i=2}^{4} i(i-1) a_i (k - x_0)^{i-2} = \sum_{i=2}^{4} i(i-1) b_i (k - x_0)^{i-2} \\ \sum_{i=3}^{4} i(i-1)(i-2) a_i (k - x_0)^{i-3} = \sum_{i=3}^{4} i(i-1)(i-2) b_i (k - x_0)^{i-3} \end{cases}$$

These conditions ensure the continuity of the function value and its first, second, and third derivatives at the knot.

Setting the sole knot at the futures price has important economic significance, as this position typically corresponds to at-the-money options. This knot placement divides the curve into two segments, corresponding to areas above and below the futures price, allowing the fitted curve to more accurately reflect volatility characteristics near the at-the-money point. This segmented fitting method is particularly suitable for handling the asymmetric features that may appear in option implied volatility before and after the at-the-money position.

At the implementation level, this research uses the LSQUnivariateSpline method from Python's SciPy package for curve fitting. This method employs the least squares approach for parameter estimation, effectively handling non-uniformly distributed data points and achieving segmented fitting through specified internal knots. Through the combination of the least squares method and knot placement, LSQUnivariateSpline provides a flexible and stable mathematical tool capable of fitting smooth and accurate implied volatility curves in non-uniformly distributed data (as shown in Figure 4-4), providing a solid foundation for subsequent risk-neutral density function (RND) extractions. The least squares method formula is as follows:

$$\min_{\{a_i\},\{b_i\}} \sum_{j=1}^{n} [y_j - S(x_j)]^2$$

Where y_j is the actual observed value, $S(x_j)$ is the observed value of the spline function at x_j, $\{a_i\},\{b_i\}$ is the set of parameters to be estimated for the spline function, and n is the total number of data points.

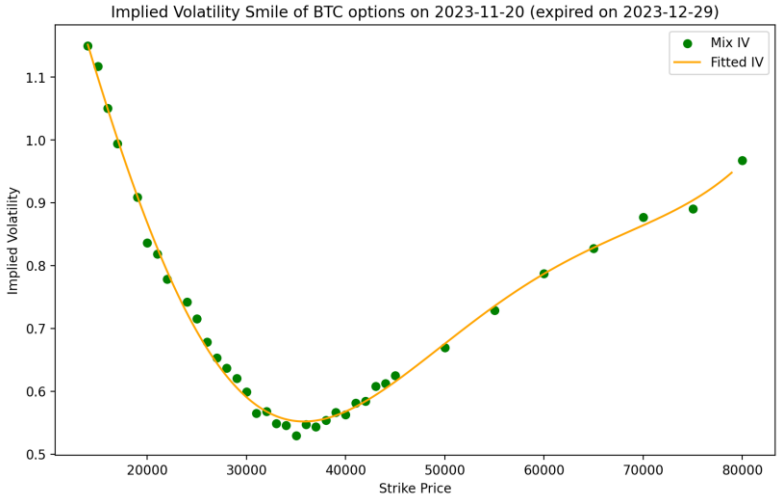

Figure 4-4: Bitcoin Option Implied Volatility Fitted Curve on November 20, 2023 (Expiring December 29, 2023)

4.2.4 Extracting the Empirical RND from Bitcoin Options

After completing the fitting of the implied volatility curve, the derivation of the risk-neutral probability density follows. First, this research uses the fitted implied volatility curve, combined with the pricing model adopted by the Deribit exchange (Formula (11)), to calculate theoretical call option prices at different strike prices, as shown in Figure 4-5.

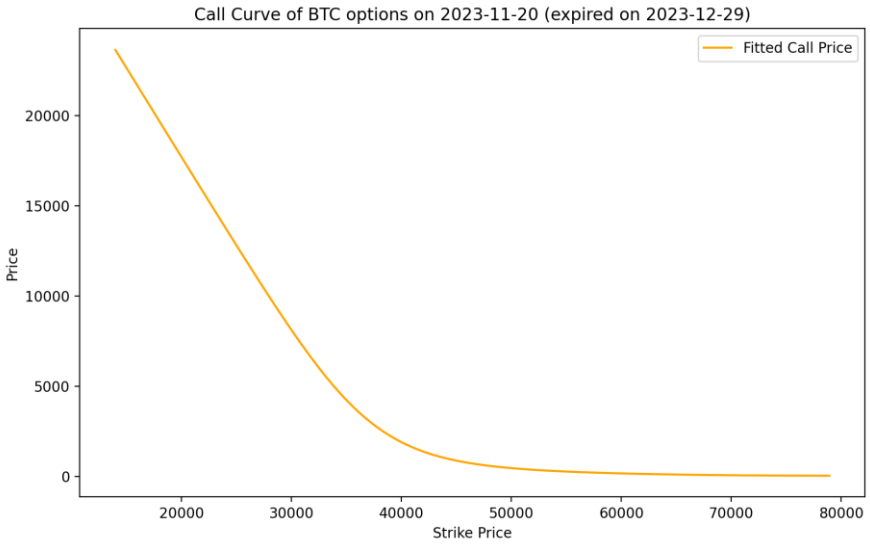

Figure 4-5: Bitcoin Option Theoretical Call Option Prices on November 20, 2023 (Expiring December 29, 2023)

After obtaining the theoretical call option prices, this research employs the finite difference method (central difference method) for discrete data differentiation to derive the empirical RND. Compared to forward or backward differences, the central difference method can effectively reduce truncation errors (Formulas (4) and (5)). To ensure the stability and accuracy of numerical calculations, this research adopts an equidistant partitioning approach in setting the price spacing, with ΔX set to 0.1. Choosing a smaller price spacing not only provides more refined density estimation, but the equidistant partitioning approach also helps improve the stability of numerical differentiation calculations.

After completing extracting the empirical RND, to ensure the reliability and reasonableness of the data, this research focuses on the empirical CDF for data validation. In terms of data integrity validation, we first ensure that empirical CDF value does not contain missing values, with observations containing missing values excluded from the analysis scope. Furthermore, to maintain theoretical consistency of extraction, this research further restricts these two probability values to be strictly between 0 and 1, while excluding boundary values equal to 0 or 1, to avoid extreme cases affecting subsequent analysis.

$$\begin{cases} F(K)\ exists\ (non-missing) \\ F_R(K)\ exists\ (non-missing) \\ 0 < F(K) < 1 \\ 0 < F_R(K) < 1 \end{cases}$$

Where $F(K)$ is the cumulative distribution function, and $F_R(K) = 1 - F(K)$。

The validation criteria are based on three aspects of consideration. First, from the perspective of theoretical consistency, these criteria ensure that the estimation conforms to the basic properties of CDF in probability theory, while also satisfying the basic requirements of PDF. Second, in terms of numerical stability, this filtering approach can

avoid computational problems in subsequent analyses due to extreme or abnormal values, effectively improving the reliability of the overall estimation. Finally, from a practical application perspective, removing abnormal values that might lead to misinterpretation better ensures that results accurately reflect market participants' true price expectations.

The derive empirical RND and its CDF are shown in Figure 4-6 and Figure 4-7:

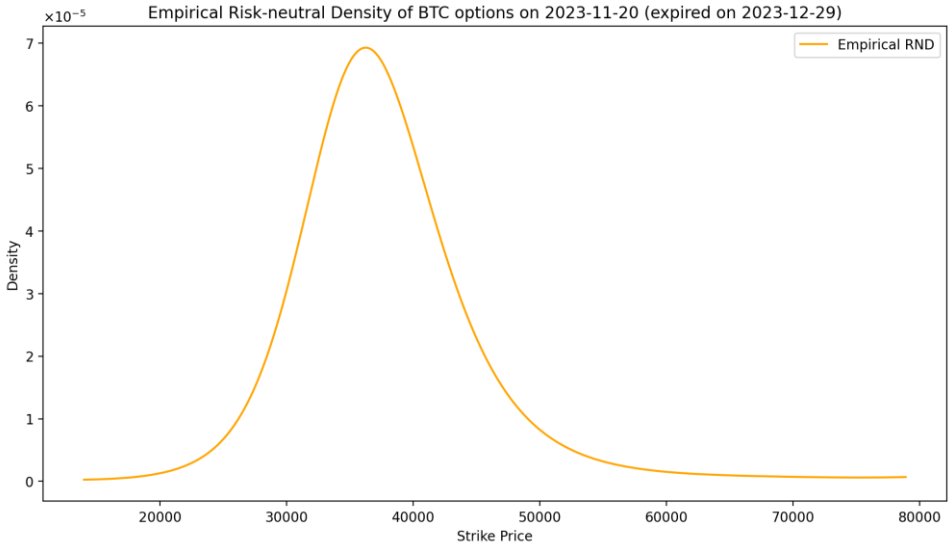

Figure 4-6: Bitcoin Option Empirical RND on November 20, 2023 (Expiring December 29, 2023)

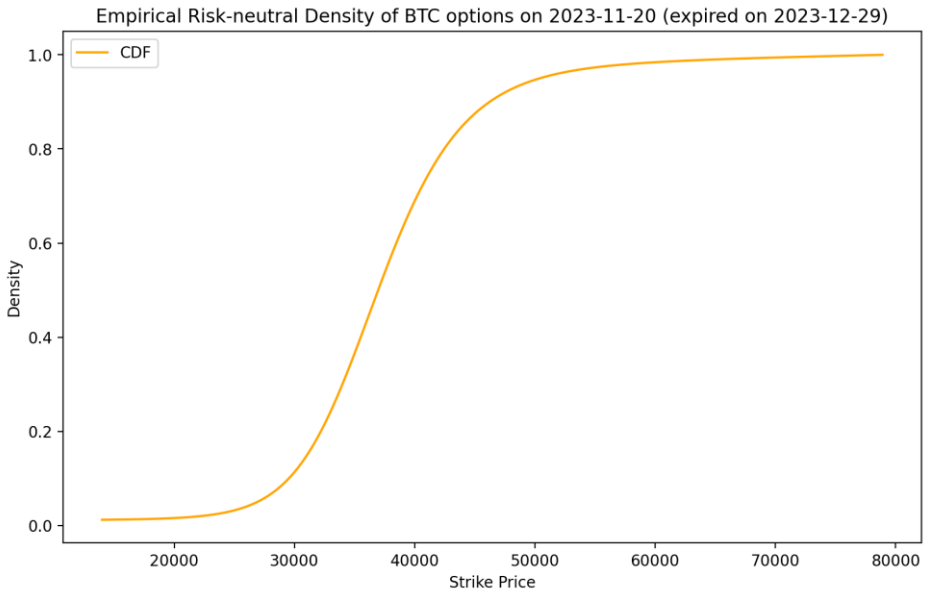

Figure 4-7: CDF of Bitcoin Option Empirical RND on November 20, 2023 (Expiring December 29, 2023)

4.3 Fitting the Tails of the Empirical RND

4.3.1 Fitting Method with GEV

The Empirical RND extracted from market option prices can only cover the range of effective trading strike prices. To fully describe market expectations, the tails of the empirical RND need to be extended. Figlewski (2008) proposed using the Generalized Extreme Value Distribution (GEV) to fit the tails of the empirical RND. This method requires setting three conditions for each tail to ensure the continuity of tail fitting:

Right tail conditions: $\begin{cases} F_{GEVR}\big(K(\alpha_{1R})\big) = \alpha_{1R} \\ f_{GEVR}\big(K(\alpha_{1R})\big) = f_{EMP}\big(K(\alpha_{1R})\big) \\ f_{GEVR}\big(K(\alpha_{2R})\big) = f_{EMP}\big(K(\alpha_{2R})\big) \end{cases}$

Left tail conditions: $\begin{cases} F_{GEVL}\big(-K(\alpha_{1L})\big) = 1 - \alpha_{1L} \\ f_{GEVR}\big(-K(\alpha_{1L})\big) = f_{EMP}\big(K(\alpha_{1L})\big) \\ f_{GEVR}\big(-K(\alpha_{2L})\big) = f_{EMP}\big(K(\alpha_{2L})\big) \end{cases}$

Where F_{GEVR} and f_{GEVR} are the CDF and PDF of the right tail GEV, respectively, F_{GEVL} is the CDF of the left tail GEV, f_{EMP} is the RND function derived in this research, and $K(\alpha)$ is the strike price corresponding to the α quantile of the Empirical RND. The fitting conditions from Figlewski (2008) can be summarized as:

(1) At the first joining point, the CDF of the GEV tail and the CDF of the empirical RND must be equal

(2) At the first joining point, the GEV density function value and the density function value of the empirical RND must be equal

(3) At the second joining point, the GEV density function value and the density function value of the empirical RND must be equal

4.3.2 GPD Distribution Theory

This research adopts the Generalized Pareto Distribution (GPD) for tail fitting, based on Balkema and Haan's (1974) proof that observations exceeding a high threshold asymptotically converge to GPD. GPD requires only two parameters (scale σ and shape ξ), enhancing computational efficiency and reducing overfitting risk compared to GEV's three parameters. Research by Hosking and Wallis (1987), McNeil and Frey (2000), and Birru and Figlewski (2012) confirmed GPD's advantages in financial applications.

The mathematical expression of the GPD's CDF is as follows:

$$F_{GPD}(x;\sigma,\xi) = \begin{cases} 1-(1+\xi\frac{x}{\sigma})^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1-exp\,(-\frac{x}{\sigma}), & \xi = 0 \end{cases}$$

The mathematical expression of the GPD's PDF is as follows:

$$f_{GPD}(x;\sigma,\xi) = \begin{cases} \frac{1}{\sigma}(1+\xi\frac{x}{\sigma})^{-\frac{1}{\xi}-1}, & \xi \neq 0 \\ \frac{1}{\sigma}exp\,(-\frac{x}{\sigma}), & \xi = 0 \end{cases}$$

Where $\sigma > 0$ is the scale parameter, used to control the degree of dispersion of the distribution, with larger σ values indicating greater variability in the data. The shape parameter ξ determines the type and tail characteristics of the distribution:

(1) When $\xi > 0$: The distribution is a Pareto Distribution with heavy-tailed characteristics; the distribution has infinite support, with domain $[0,\infty)$; the tail decays more slowly

(2) When $\xi = 0$: The distribution degenerates to an Exponential Distribution with a fixed decay rate; it is the simplest continuous memoryless distribution; the tail

decays at a moderate speed

(3) When $\xi < 0$: The distribution belongs to the Beta Family with finite support characteristics; the distribution function is only defined on the interval $\left[0, -\frac{\sigma}{\xi}\right)$; it is less commonly used in financial market applications because asset returns typically do not have a clear upper limit

4.3.3 Fitting Method with GPD

Birru and Figlewski (2012) proposed a fitting method for GPD tails, using two joining points to compare the density function values of GPD and the empirical RND. The tail fitting conditions are set as follows:

(1) At the first joining point, the GPD density function value and the empirical RND density function value must be equal

(2) At the second joining point, the GPD density function value and the empirical RND density function value must be equal

The mathematical expressions are as follows:

Right tail conditions:

$$\begin{cases} f_{GPD}(K(\alpha_{1R})) = f_{EMP}(K(\alpha_{1R})) & (PDF\ condition) \\ f_{GPD}(K(\alpha_{2R})) = f_{EMP}(K(\alpha_{2R})) & (PDF\ condition) \end{cases}$$

Left tail conditions:

$$\begin{cases} f_{GPD}(K(\alpha_{1L})) = f_{EMP}(K(\alpha_{1L})) & (PDF\ condition) \\ f_{GPD}(K(\alpha_{2L})) = f_{EMP}(K(\alpha_{2L})) & (PDF\ condition) \end{cases}$$

The scale parameter and shape parameter of GPD are solved by minimizing the following objective functions:

Right tail parameter minimization objective function:

$$\min_{\sigma,\xi}\{[f_{GPD}(K(\alpha_{1R})) - f_{EMP}(K(\alpha_{1R}))]^2 + [f_{GPD}(K(\alpha_{2R})) - f_{EMP}(K(\alpha_{2R}))]^2\}$$

Left tail parameter minimization objective function:

$$\min_{\sigma,\xi}\{[f_{GPD}(K(\alpha_{1L})) - f_{EMP}(K(\alpha_{1L}))]^2 + [f_{GPD}(K(\alpha_{2L})) - f_{EMP}(K(\alpha_{2L}))]^2\}$$

Where α_{1L} is preset to 0.05; α_{2L} is preset to 0.02; α_{1R} is preset to 0.95; α_{2R} is preset to 0.98. The empirical RND curve and its CDF completed using Birru and Figlewski (2012) method are shown in Figure 4-8 and Figure 4-9.

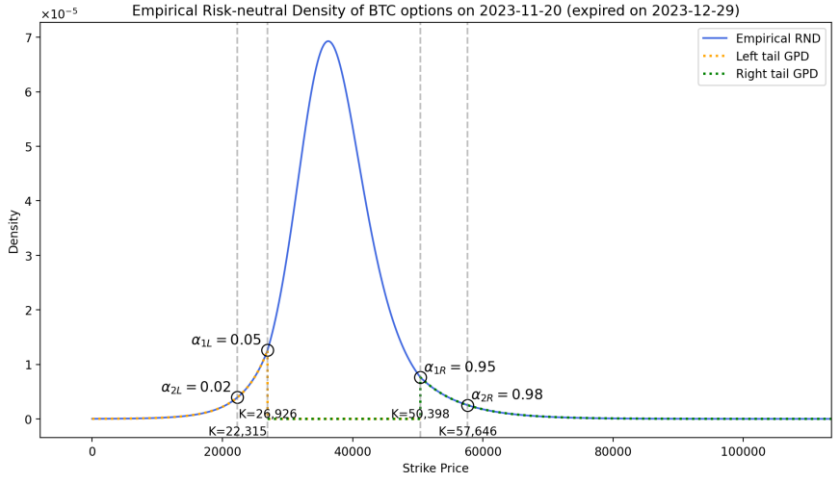

Figure 4-8: Bitcoin Option Empirical RND and GPD Tail Fitting on November 20, 2023 (Birru and Figlewski (2012))

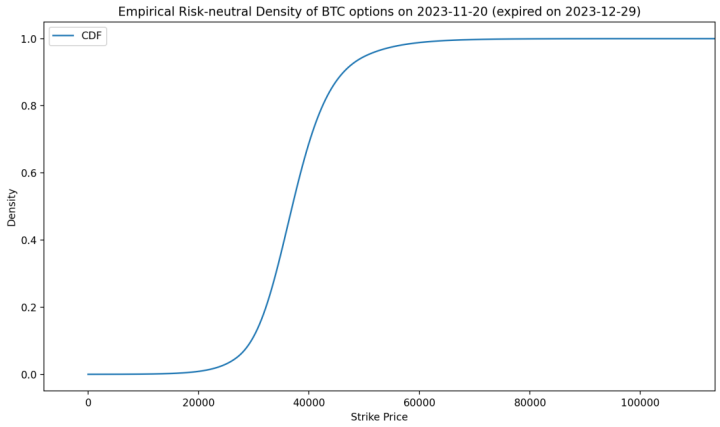

Figure 4-9: CDF of Bitcoin Option RND on November 20, 2023 (Birru and Figlewski (2012))

4.3.4 Fitting Method with GPD Using the Proposed Method

This research proposes fitting the empirical RND using the proposed method with GPD. The proposed method not only considers the fitting of CDF values but also adds continuity conditions for the slope of the density function. The main advantage of this method is that it can simultaneously ensure the continuity and smoothness of the density function while simplifying the fitting process and improving computational efficiency. The tail fitting conditions set in this research are as follows:

(1) At the joining point, the CDF of GPD and the CDF of the empirical RND must be equal.

(2) At the joining point, the slope of the GPD density function and the slope of the empirical RND density function must be equal.

The mathematical expressions are as follows:

Right tail conditions:

$$\begin{cases} F_{GPD}(K(\alpha_{1R})) = \alpha_{1R} & (CDF\ condition) \\ \dfrac{f_{GPD}(K(\alpha_{1R}) + \Delta x) - f_{GPD}(K(\alpha_{1R}))}{\Delta x} = \dfrac{f_{EMP}(K(\alpha_{1R}) + \Delta x) - f_{EMP}(K(\alpha_{1R}))}{\Delta x} & (Slope\ condition) \end{cases}$$

Left tail conditions:

$$\begin{cases} F_{GPD}(-K(\alpha_{1L})) = \alpha_{1L} & (CDF\ condition) \\ \dfrac{f_{GPD}(-K(\alpha_{1L}) + \Delta x) - f_{GPD}(-K(\alpha_{1L}))}{\Delta x} = \dfrac{f_{EMP}(K(\alpha_{1L}) + \Delta x) - f_{EMP}(K(\alpha_{1L}))}{\Delta x} & (Slope\ condition) \end{cases}$$

The scale parameter and shape parameter of GPD are solved by minimizing the following objective functions:

Right tail parameter minimization objective function:

$$\min_{\sigma,\xi}\Bigg\{\Big[F_{GPD}\big(K(\alpha_{1R})\big)-\alpha_{1R}\Big]^2 + \Bigg[\frac{f_{GPD}(K(\alpha_{1R})+\Delta x)-f_{GPD}(K(\alpha_{1R}))}{\Delta x} - \frac{f_{EMP}(K(\alpha_{1R})+\Delta x)-f_{EMP}(K(\alpha_{1R}))}{\Delta x}\Bigg]^2\Bigg\}$$

Left tail parameter minimization objective function:

$$\min_{\sigma,\xi}\Bigg\{[F_{GPD}(-K(\alpha_{1L}))-\alpha_{1L}]^2 + \Bigg[\frac{f_{GPD}(-K(\alpha_{1L})+\Delta x)-f_{GPD}(-K(\alpha_{1L}))}{\Delta x} - \frac{f_{EMP}(K(\alpha_{1L})+\Delta x)-f_{EMP}(K(\alpha_{1L}))}{\Delta x}\Bigg]^2\Bigg\}$$

Where Δx is a fixed constant value used to construct artificially spaced option prices, preset to 0.1; α_{1L} is preset to 0.05; α_{1R} is preset to 0.95.

This research performs fitting separately for the left and right tails. For the left tail, we select the point with a cumulative probability of 5% as the joining point; for the right tail, we select the point with a cumulative probability of 95% as the joining point. This design ensures the smoothness of tail fitting while maintaining the continuity of the overall distribution. To minimize fitting errors, this research adopts the least squares method for parameter estimation, solving for the scale parameter and shape parameter of GPD through numerical optimization methods. The completed RND curve and its CDF are shown in Figures 4-10 and 4-11.

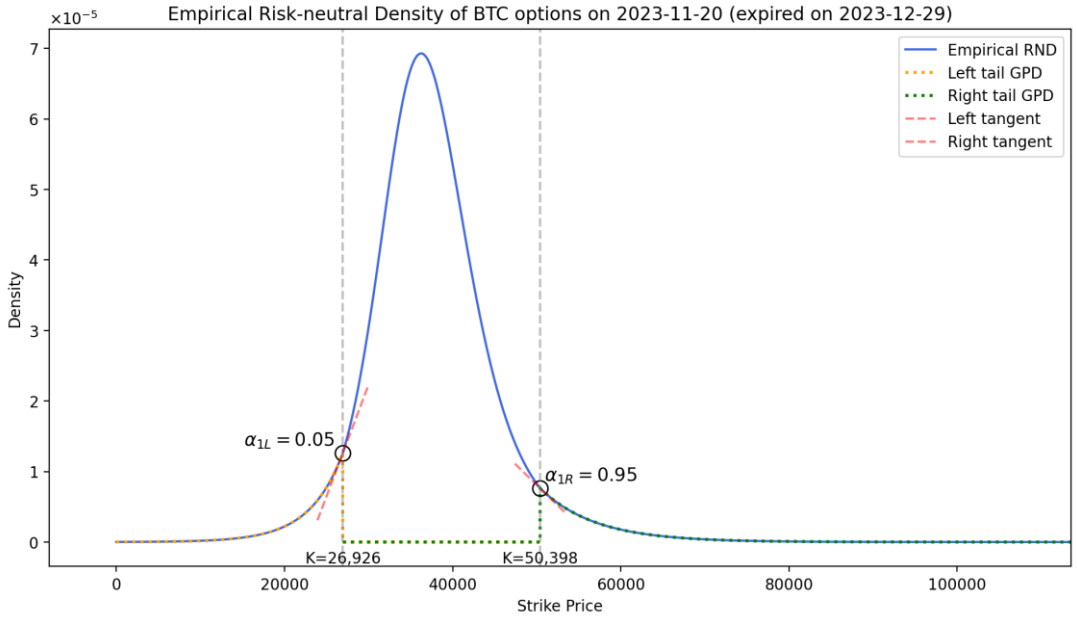

Figure 4-10: Bitcoin Option Empirical RND and GPD Tail Fitting on November 20, 2023 (The Proposed Method)

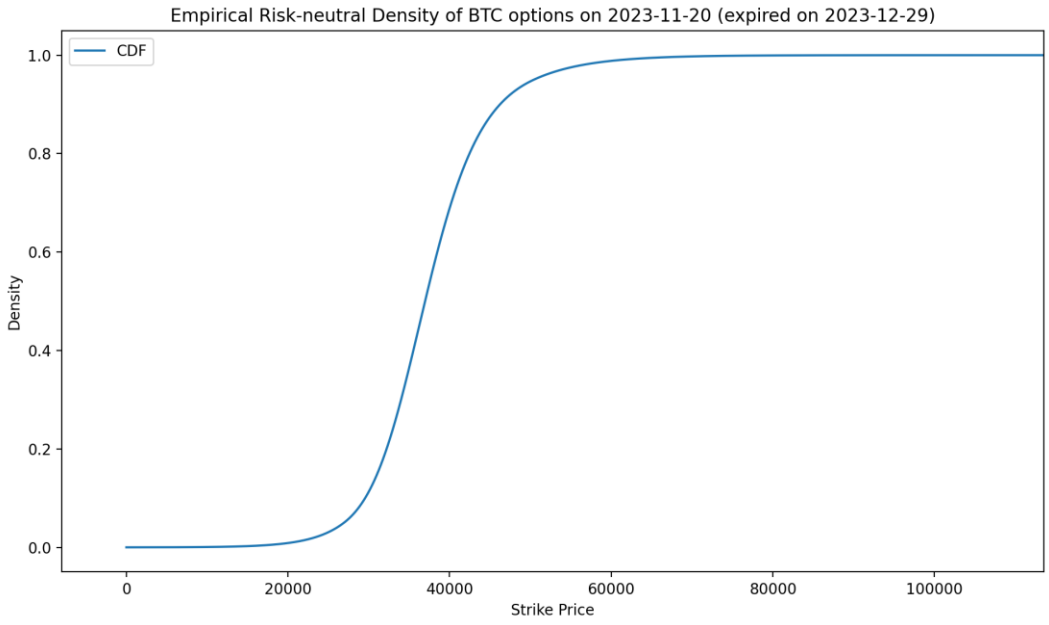

Figure 4-11: CDF of Bitcoin Option RND on November 20, 2023 (The Proposed Method)

4.4 Deriving the Moments of the RND

4.4.1 Definition and Implications of Moments

Moments are important statistical measures for describing the characteristics of probability distributions and can be divided into raw moments and central moments. For discrete strike prices, the nth-order raw moment is defined as:

$$m_n' = E[K^n] = \sum_{i=1}^{N} K_i^n f(K_i)\,\Delta K$$

And the nth-order central moment is defined as:

$$m_n = E[(K - \overline{K})^n] = \sum_{i=1}^{N} (K_i - \overline{K})^n f(K_i)\,\Delta K$$

Where $f(K_i)$ is the RND; ΔK is the price interval, set to 0.1; $\overline{K}$ is the expected value, i.e., the first-order raw moment m_1'; N is the number of observations.

4.4.2 Moments of the RND

Recent research shows that option-implied moments not only effectively describe market expectations but also possess significant predictive power. Chang et al. (2013) found that risk-neutral skewness can effectively predict stock returns, especially during periods of high market volatility; while Neumann and Skiadopoulos (2013) pointed out that changes in risk-neutral kurtosis often lead market trends, providing important signals for investment decisions. This research computes the following four main moments:

1. Mean

The mean is the first-order raw moment, reflecting the overall market expectation

for the future price level of the underlying asset. Bali and Murray (2013) pointed out that under risk-neutral pricing theory, the mean of the RND should equal the forward price discounted by the risk-free rate, providing an important arbitrage constraint. The formula is as follows:

$$\bar{K} = m_1' = E[K] = \sum_{i=1}^{N} K_i f(K_i) \Delta K$$

2. Standard Deviation

The standard deviation is the square root of the second-order central moment, measuring the degree of price dispersion. Christoffersen et al. (2013) found that option-implied standard deviation has stronger predictive power than historical volatility, especially in emerging markets. Standard deviation reflects market expectations for price volatility, with higher values indicating greater uncertainty among market participants regarding future price movements. The formula is as follows:

$$\sigma = \sqrt{m_2} = \sqrt{E[(K - \bar{K})^2]} = \sqrt{\sum_{i=1}^{N} (K_i - \bar{K})^2 f(K_i) \Delta K}$$

3. Skewness

Skewness is the standardized third-order central moment, describing the asymmetry of the distribution:

$$\text{Skewness} = \frac{m_3}{\sigma^3} = \frac{E[(K - \bar{K})^3]}{\sigma^3} = \frac{\sum_{i=1}^{N} (K_i - \bar{K})^3 f(K_i) \Delta K}{\sigma^3}$$

The skewness coefficient has important implications in financial markets. When positive skewness is observed, it indicates that the price distribution has a longer right tail, implying that market participants expect a higher probability of significant upward

movements, reflecting overall optimistic market sentiment. Conversely, negative skewness indicates that the price distribution has a longer left tail, representing a greater perceived downside risk in the market, usually reflecting higher hedging demand among market participants. Research by Conrad et al. (2013) shows that risk-neutral skewness not only reflects market sentiment but also contains investors' expectations for extreme events. Li et al. (2024) shows that the dynamic changes in the skewness coefficient can often serve as a leading indicator of market sentiment shifts, with its changing trends providing important reference value for investment decisions.

4. Excess Kurtosis

Excess Kurtosis is the standardized fourth-order central moment minus 3 (the kurtosis value of the normal distribution), used to describe the tail characteristics of the distribution. It is calculated as:

$$\text{Excess Kurtosis} = \frac{m_4}{\sigma^4} - 3 = \frac{E[(K - \overline{K})^4]}{\sigma^4} - 3 = \frac{\sum_{i=1}^{N}(K_i - \overline{K})^4 f(K_i)\,\Delta K}{\sigma^4} - 3$$

In practical applications, excess kurtosis is an important indicator for assessing extreme market risk. Amaya et al. (2015) pointed out that excess kurtosis can effectively capture extreme market risk, with its predictive power being particularly significant during financial crises. When positive excess kurtosis is observed, it indicates that the distribution has more pronounced fat-tail characteristics compared to the normal distribution, meaning that the probability of extreme events occurring is higher than expected under a normal distribution.

4.4.3 Market Implications of the Moments

These moments provide rich market information beyond mathematical description, including price expectations, risk assessment, and market sentiment indicators. This

research employs regression analysis to empirically test their predictive power for spot returns.

4.5 Regression Analysis

4.5.1 Theoretical Background for Regression Models

To explore the predictive power of Bitcoin option-implied RND for underlying asset price movements, this research adopts multiple regression analysis for testing. The dependent variable (Y) is set as the logarithmic return of Bitcoin, while the independent variables (X) gradually incorporate moments of the RND such as Mean, Standard Deviation, Skewness, Excess Kurtosis, Median, as well as market sentiment indicators such as the Cryptocurrency Fear and Greed Index and the Chicago Board Options Exchange Volatility Index (VIX), along with historical returns from the previous 1 to 4 periods, to construct the most explanatory predictive model.

Bali and Zhou (2016) demonstrated that moments of the RND, particularly skewness and kurtosis, effectively predict cross-sectional asset return variations, reflecting market participants' risk preferences and containing crucial pricing information. Amaya et al. (2015) further established that the RND excess kurtosis specifically excels in predicting extreme market risk, a finding particularly relevant for high-volatility markets such as cryptocurrencies.

López-Cabarcos et al. (2021) examined relationships between Bitcoin volatility, stock market performance, and investor sentiment, indicating that during market stability periods, VIX returns and investor sentiment significantly impact Bitcoin volatility. Akyildirim et al. (2020), utilizing high-frequency data, identified a time-varying positive correlation between cryptocurrencies and market panic indicators (VIX, VSTOXX),

which intensifies during periods of financial market stress. M. He et al. (2023) employed the daily updated Cryptocurrency Fear and Greed Index as a predictor, demonstrating significant in-sample and out-of-sample predictive power for individual cryptocurrency and market index returns over one-day to one-week horizons.

Liu and Tsyvinski (2021) identified significant momentum effects in cryptocurrency markets, with Bitcoin's current period returns substantially predicting returns for the subsequent 1-6 days. Y. Li et al. (2021) documented a positive MAX momentum effect, where cryptocurrencies exhibiting higher extreme daily returns tend to generate superior future returns. Liu et al. (2023), applying machine learning methodologies to cryptocurrency return prediction, determined that previous 1-day returns possess the strongest predictive power, exceeding the combined effect of all other variables.

4.5.2 Regression Model Specification

Campbell and Thompson (2008) introduced the concept of "economic significance threshold," showing that by gradually introducing predictive variables and setting strict statistical significance standards, one can effectively distinguish variables with substantial predictive power. In Gu, Kelly, and Xiu (2020), a "staged variable introduction framework" was proposed specifically for high-dimensional data, which can alleviate overfitting problems compared to models that introduce all variables at once.

This research adopts a multi-level regression analysis, gradually expanding from univariate to four-variable models, to systematically explore the predictive power of various risk-neutral probability density characteristics for future returns. The following details the design of regression models at each level:

1. Univariate Regression Model

The univariate regression model is mainly used to test the explanatory power of individual variables for future returns. Its basic form is:

$$R_t = \beta_0 + \beta_1 Variable_{i,t-1} + \varepsilon_t, i \in \{1, 2, \ldots, 11\}$$

Where $R_t = ln(\frac{Close_t}{Close_{t-1}})$ is the return of Bitcoin price in the next period (T). If the option sample expires in 7 days, then the return is calculated using the closing price from the current day to the closing price 7 days later (expiration date).

$Variable_i$ is sequentially replaced with the following variables for univariate regression analysis: Mean, Standard Deviation (Std), Skewness, Excess Kurtosis, Median, Cryptocurrency Fear and Greed Index, Chicago Board Options Exchange Volatility Index (VIX), T-1 Return, T-2 Return, T-3 Return, T-4 Return.

2. Bivariate Regression Model

Considering the importance of Skewness in option pricing theory, this research designs bivariate models with Skewness as a fixed factor. The model is set as follows:

$$R_t = \beta_0 + \beta_1 Skewness_{t-1} + \beta_2 Variable_{i,t-1} + \varepsilon_t, i \in \{1, 2, \ldots, 10\}$$

$Variable_i$ is sequentially replaced with the following variables, paired with Skewness for bivariate regression analysis: Mean, Standard Deviation (Std), Excess Kurtosis, Median, Cryptocurrency Fear and Greed Index, Chicago Board Options Exchange Volatility Index (VIX), T-1 Return, T-2 Return, T-3 Return, T-4 Return.

3. Three-Variable Regression Model

The three-variable model further incorporates Excess Kurtosis as a fixed factor, forming:

$$R_t = \beta_0 + \beta_1 Skewness_{t-1} + \beta_2 ExcessKurtosis_{t-1} + \beta_3 Variable_{i,t-1} + \varepsilon_t, i \in \{1, 2, \ldots, 9\}$$

$Variable_i$ is sequentially replaced with the following variables, paired with Skewness and Excess Kurtosis for three-variable regression analysis: Mean, Standard Deviation (Std), Median, Cryptocurrency Fear and Greed Index, Chicago Board Options Exchange Volatility Index (VIX), T-1 Return, T-2 Return, T-3 Return, T-4 Return.

4. Four-Variable Regression Model

Building on the three-variable model, the four-variable model adds the Standard Deviation (Std) variable. The complete model is as follows:

$$R_t = \beta_0 + \beta_1 Skewness_{t-1} + \beta_2 ExcessKurtosis_{t-1} + \beta_3 Std_{t-1} + \beta_4 Variable_{i,t-1} + \varepsilon_t,$$

$i \in 1,2, \ldots ,8$

$Variable_i$ is sequentially replaced with the following variables, paired with Skewness, Excess Kurtosis, and Standard Deviation for four-variable regression analysis: Mean, Median, Cryptocurrency Fear and Greed Index, Chicago Board Options Exchange Volatility Index (VIX), T-1 Return, T-2 Return, T-3 Return, T-4 Return.

4.5.3 Validation of Model Effectiveness

The proposed method, compared to Birru and Figlewski (2012), has the advantage of computational efficiency in practical applications. To objectively validate the effectiveness of this method, this research adopts three quantitative indicators for evaluation: Mean Squared Error (MSE), Out-of-sample R-squared (R^2_{OS}) (Campbell & Thompson, 2008; Welch & Goyal, 2008), and computational efficiency, to comprehensively compare the accuracy and practicality of predictive models constructed

by the two methods.

1. Mean Squared Error (MSE) Calculation

Mean Squared Error is a commonly used indicator for evaluating the accuracy of predictive models (Orosi, 2015). Its calculation formula is as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Where y_i is the actual observed value (i.e., the actual Bitcoin return), $\hat{y}_i$ is the model′s predicted value, and n is the sample size. A smaller MSE value indicates smaller prediction errors and higher prediction accuracy.

2. Out-of-sample R-squared (R_{OS}^2) Calculation

Referring to–Campbell and Thompson (2008) and Welch and Goyal (2008), this research adopts out-of-sample R-squared (R_{OS}^2) to evaluate the predictive power of the model. R_{OS}^2 measures the improvement in prediction of the predictive model relative to the historical average benchmark model. Its calculation formula is as follows:

$$R_{OS}^2 = 1 - \frac{\sum_{t=s_0+1}^{T}(R_{t+1} - \hat{R}_{t+1})^2}{\sum_{t=s_0+1}^{T}(R_{t+1} - \bar{R}_{t+1})^2}$$

Where R_{t+1} is the actual return, $\hat{R}_{t+1}$ is the predicted value from the predictive model, $\bar{R}_{t+1}$ is the historical average return (benchmark model), s_0 is the initial in-sample period length, and T is the total sample size. When R_{OS}^2 is greater than zero, it indicates that the predictive model outperforms the historical average benchmark model; the larger the R_{OS}^2 value, the more significant the improvement in prediction.

This research adopts a rolling window approach for out-of-sample prediction, with the initial in-sample period set to 80% of the total sample size, and gradually advancing

forward for prediction. By comparing the R_{OS}^2 values of the proposed method and Birru and Figlewski (2012), the difference in out-of-sample predictive power between the two methods can be objectively evaluated.

3. Computational Efficiency Comparison

In addition to prediction accuracy, this research also values the practical applicability of the method, especially its computational efficiency when processing large amounts of data. To objectively evaluate the computational performance of the two methods, this research selects ten representative option expiration dates, derives the 7-day risk-neutral probability density for each expiration date, and uses both the proposed method and Birru and Figlewski (2012) to fit the tails with the Generalized Pareto Distribution (GPD). To ensure the reliability and stability of the results, this research performs ten repeated computations for each method, records their execution times, and computes the average, thereby comprehensively evaluating the differences in computational efficiency between the two methods in practical applications.

5. Empirical Results

This chapter presents empirical comparisons between the proposed method and Birru and Figlewski (2012) for RND tail fitting, analyzing fitting characteristics, computational efficiency, and predictive performance across different option expiration horizons.

Bitcoin's 24-hour trading environment, substantial liquidity, and efficient price discovery mechanism make it ideal for studying RND characteristics. The market's rapid information processing and technology-oriented trader base enable short-term products to effectively capture real-time risk assessments.

5.1 Analysis of Fitting Effects

5.1.1 Comparison between the Proposed Method and Birru and Figlewski (2012)

Comparing samples from July 10, 2022 (expiring July 11, 2022) and September 27, 2023 (expiring September 28, 2023), the proposed method demonstrates better stability, with smoother and more continuous fitting curves (Figure 5-1 and Figure 5-2).

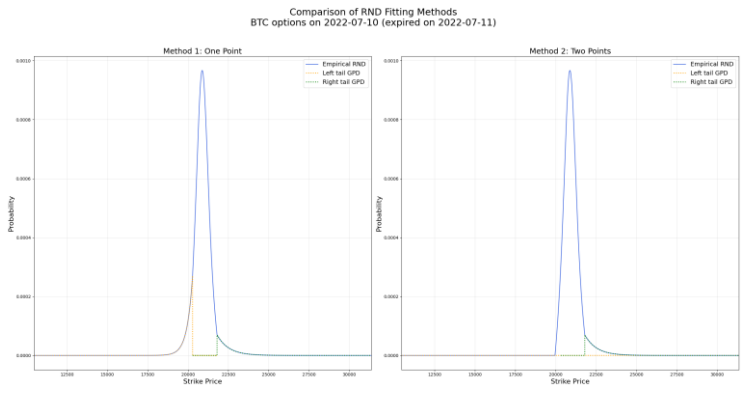

Figure 5-1: Comparison of Bitcoin Option GPD Tail Fitting on July 10, 2022
(Left: The proposed method; Right: Birru and Figlewski (2012))

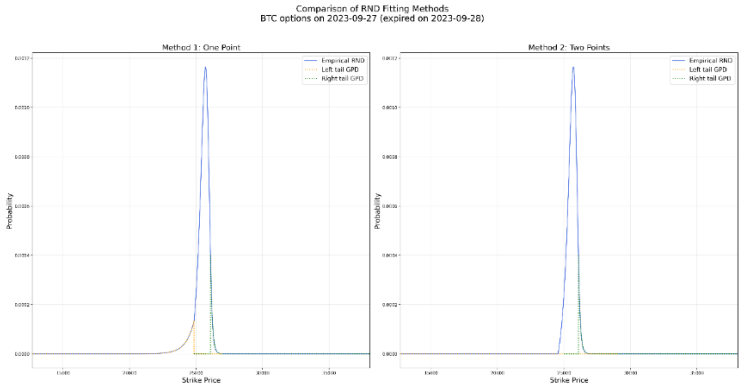

Figure 5-2: Comparison of Bitcoin Option GPD Tail Fitting on September 27, 2023
(Left: The proposed method; Right: Birru and Figlewski (2012))

The proposed method (left panels) reveals robust RND curve fitting across both samples, maintaining continuity at joining points and producing gradually decreasing tails that conform to probability density function properties. This demonstrates superior robustness in handling extreme values.

Conversely, while Birru and Figlewski (2012) (right panels) generally produces reasonable fitting curves, it occasionally exhibits fitting failures or discontinuities under certain market conditions. This limitation stems from simultaneous continuity requirements at two joining points, which becomes problematic during extreme market fluctuations or highly skewed price distributions.

The proposed method offers enhanced computational efficiency and reduced fitting failures, particularly valuable in large sample analyses.

5.1.2 Comparison of Computational Efficiency

Testing both methods across 10 option expiration dates (September-December 2023) using identical hardware configurations revealed significant performance differences (Table 5-1). The proposed method demonstrated superior performance, averaging 309.86 seconds execution time compared to Birru and Figlewski's (2012) 347.95 seconds, representing a 10.95% reduction in computation time.

This efficiency advantage stems from fundamental algorithmic differences. While Birru and Figlewski (2012) requires simultaneous satisfaction of continuity conditions at two joining points, introducing additional optimization constraints, the proposed method addresses only a single joining point, enabling more direct optimization with faster parameter convergence. The proposed method also exhibits greater execution time stability (lower standard deviation), indicating more consistent computational performance.

Table 5-1: Comparison of Computational Efficiency for Bitcoin Option GPD Tail Fitting
(Left: The proposed method; Right: Birru and Figlewski (2012))

Execution Time:	2025/2/5 00:48
Execution Conditions:	Each time generates 10 weekly return RNDs with GPD distribution tail fitting.
Option Expiration Dates:	2023/9/22, 2023/9/29, 2023/10/13, 2023/10/20, 2023/10/27, 2023/11/10, 2023/11/17, 2023/11/24, 2023/12/15, 2023/12/22

The proposed method		Birru and Figlewski (2012)	
1st Execution Time (sec)	297.58	1st Execution Time (sec)	346.49
2nd Execution Time (sec)	300.23	2nd Execution Time (sec)	346.95
3rd Execution Time (sec)	313.52	3rd Execution Time (sec)	347.78
4th Execution Time (sec)	313.51	4th Execution Time (sec)	347.68
5th Execution Time (sec)	312.25	5th Execution Time (sec)	347.81
6th Execution Time (sec)	311.37	6th Execution Time (sec)	347.94
7th Execution Time (sec)	313.52	7th Execution Time (sec)	348.39
8th Execution Time (sec)	311.86	8th Execution Time (sec)	349.10
9th Execution Time (sec)	312.36	9th Execution Time (sec)	348.59
10th Execution Time	312.42	10th Execution Time	348.80
Shortest Execution Time (sec)	297.58	Shortest Execution Time (sec)	346.49
Longest Execution Time (sec)	313.52	Longest Execution Time (sec)	349.10
Average Execution Time (sec)	309.86	Average Execution Time (sec)	347.95

5.2 Regression Analysis with 1 Day to Expiration

5.2.1 Fitting Tails with GPDs Based on the Proposed Method

This section uses option products expiring daily from January 10, 2021, to April 30,

2024, deriving the RND from the observation date one day before expiration, and constructs complete the RND functions using the proposed method. Moments such as mean, standard deviation, skewness, and kurtosis are then calculated as explanatory variables. Using the next period's Bitcoin spot return as the explained variable, multi-level regression analysis is conducted to observe whether the RND has predictive effects. The descriptive statistics of the variables are shown in Table 5-2, with a total of 832 samples. Observing Skewness and Excess Kurtosis, it can be found that the RND functions constructed using the proposed method have outliers in skewness and excess kurtosis, which in turn affect the mean and standard deviation of these variables.

Table 5-2: Descriptive Statistics of the RND Moments and Bitcoin Returns for Products with 1 Day to Expiration (The proposed method)

	Count	Mean	Std	Min	25%	Median	75%	Max
T Return (Y)	832	-0.0002	0.0336	-0.1670	-0.0155	-0.0003	0.0156	0.1353
Mean	832	35891.0208	14144.2251	0.0000	25047.4826	34274.1105	46035.6010	72614.4638
Std	832	1784.3452	1422.5234	0.0000	821.3941	1463.6223	2291.7552	12224.7221
Skewness	832	0.4054	7.2140	-12.5062	-0.7712	0.2451	1.0921	193.7010
Excess Kurtosis	832	56.6637	1392.5124	-140.8251	-0.2814	1.6822	4.2421	40084.7494
Median	832	36972.0948	14057.6155	7860.0000	25800.8750	35687.3500	46933.4000	72597.4000
Fear and Greed Index	832	46.9892	22.4770	6.0000	26.0000	49.0000	68.2500	95.0000
VIX	832	20.0917	5.2866	12.0700	16.1875	19.2250	23.0300	37.2100
T-1 Return	832	-0.0001	0.0337	-0.1670	-0.0155	-0.0003	0.0156	0.1353
T-2 Return	832	-0.0001	0.0337	-0.1670	-0.0153	-0.0003	0.0161	0.1353
T-3 Return	832	-0.0001	0.0337	-0.1670	-0.0153	-0.0003	0.0160	0.1353
T-4 Return	832	-0.0003	0.0337	-0.1670	-0.0157	-0.0004	0.0156	0.1353

Following the regression analysis set in Section 4.5, the univariate regression results are shown in Table 5-3. It can be observed that Mean, Skewness, Median, and T-4 Return are significant, mainly concentrated in the RND distribution characteristics. The predictive ability of market sentiment indicators such as the Cryptocurrency Fear and Greed Index and Volatility Index (VIX) is extremely low, indicating that technical

indicators may be more valuable than market sentiment indicators.

Table 5-3: Univariate Regression Results for Products with 1 Day to Expiration (The proposed method)

	Coefficient	p value	Significance	R-squared
Mean	-0.0866	0.0124	**	0.0075
Std	-0.0247	0.4761		0.0006
Skewness	0.0614	0.0770	*	0.0038
Excess Kurtosis	0.0515	0.1379		0.0027
Median	-0.0707	0.0416	**	0.0050
Fear and Greed Index	-0.0023	0.9462		0.0000
VIX	-0.0010	0.9777		0.0000
T-1 Return	-0.0402	0.2471		0.0016
T-2 Return	0.0276	0.4261		0.0008
T-3 Return	0.0093	0.7893		0.0001
T-4 Return	0.0617	0.0752	*	0.0038

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Table 5-4 presents bivariate regression results with Skewness as a fixed variable. The addition of Median and T-4 Return variables provides stable predictive capability with increased explanatory power (higher R-squared values), suggesting that RND median and short-term momentum effects contain significant predictive information for Bitcoin returns. Three-variable and four-variable regression analyses are documented in Appendix Table 1 and Appendix Table 2.

Table 5-4: Bivariate Regression Results for Products with 1 Day to Expiration (The proposed method)

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	R-squared
Mean	-0.0796	0.0229	**	0.0502	0.1510		0.0100
Std	-0.0230	0.5064		0.0607	0.0803	*	0.0043
Excess Kurtosis	-0.0631	0.5565		0.1211	0.2595		0.0042
Median	-0.0730	0.0352	**	0.0640	0.0647	*	0.0091
Fear and Greed Index	-0.0063	0.8551		0.0618	0.0758	*	0.0038
VIX	0.0009	0.9792		0.0614	0.0771	*	0.0038
T-1 Return	-0.0347	0.3194		0.0581	0.0956	*	0.0050
T-2 Return	0.0298	0.3906		0.0624	0.0724	*	0.0046
T-3 Return	0.0100	0.7722		0.0615	0.0765	*	0.0039
T-4 Return	0.0606	0.0802	*	0.0602	0.0821	*	0.0074

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

After multi-level regression analysis, this section finally adopts a regression model with three variables: Skewness, Median, and T-4 Return. The model results are shown in Table 5-5, indicating that Skewness and T-4 Return have positive effects on Bitcoin spot return prediction, while Median has a negative effect; the Mean Squared Error (MSE) shows the high volatility of Bitcoin. This section also uses this model as a basis to attempt to add a fourth variable to find a more explanatory model. Nevertheless, none of the added variables are statistically significant (Appendix Table 3), indicating that this model has already demonstrated relatively stable predictive ability.

Table 5-5: Regression Results for Products with 1 Day to Expiration (The proposed method)

	Coefficient	p value	Significance	R-squared	MSE
Skewness	0.0629	0.0690	*	0.0130	0.9858
Median	-0.0746	0.0311	**		
T-4 Return	0.0626	0.0705	*		

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

5.2.2 Fitting Tails with GPDs Based on Birru and Figlewski (2012)

This section uses option products expiring daily from January 10, 2021, to April 30, 2024, deriving the RND from the observation date one day before expiration, and constructs complete the RND functions using Birru and Figlewski (2012). Moments such as mean, standard deviation, skewness, and kurtosis are then calculated as explanatory variables. Using the next period's Bitcoin spot return as the explained variable, multi-level regression analysis is conducted to observe whether the RND has predictive effects. The descriptive statistics of the variables are shown in Table 5-6, with a total of 831 samples, one less than the proposed method. This is because Birru and Figlewski (2012) encountered fitting problems, indicating that the proposed method is more stable. Observing Skewness and Excess Kurtosis, it can be found that the statistical characteristic data calculated using Birru and Figlewski (2012) is less affected by outliers.

Table 5-6: Descriptive Statistics of the RND Moments and Bitcoin Returns for Products with 1 Day to Expiration (Birru and Figlewski (2012))

	Count	Mean	Std	Min	25%	Median	75%	Max
T Return (Y)	831	-0.0003	0.0336	-0.1670	-0.0156	-0.0004	0.0155	0.1353
Mean	831	36049.7626	13917.9793	771.3413	25159.9872	34158.9645	46118.6419	72614.4638
Std	831	1733.1035	1173.4002	179.8905	854.8143	1502.1723	2276.4792	10729.4925
Skewness	831	0.3140	1.7155	-12.5062	-0.5639	0.5011	1.1725	16.9925
Excess Kurtosis	831	2.9968	9.4980	-140.8251	-0.7554	1.3626	3.6404	61.3908
Median	831	37010.3730	14027.0369	15890.6000	25859.8500	35717.4000	46934.7000	72597.4000
Fear and Greed Index	831	47.0253	22.4806	6.0000	26.0000	49.0000	68.5000	95.0000
VIX	831	20.0818	5.2947	12.0700	16.1800	19.2000	23.0300	37.2100
T-1 Return	831	-0.0001	0.0337	-0.1670	-0.0156	-0.0003	0.0157	0.1353
T-2 Return	831	-0.0001	0.0337	-0.1670	-0.0153	-0.0003	0.0160	0.1353
T-3 Return	831	-0.0001	0.0337	-0.1670	-0.0153	-0.0003	0.0160	0.1353
T-4 Return	831	-0.0003	0.0337	-0.1670	-0.0157	-0.0004	0.0157	0.1353

Following the regression analysis set in Section 4.5, the univariate regression results are shown in Table 5-7. It can be observed that Mean, Median, and T-4 Return are significant.

Table 5-7: Univariate Regression Results for Products with 1 Day to Expiration (Birru and Figlewski (2012))

	Coefficient	p value	Significance	R-squared
Mean	-0.0707	0.0415	**	0.0050
Std	-0.0179	0.6054		0.0003
Skewness	-0.0250	0.4718		0.0006
Excess Kurtosis	0.0297	0.3922		0.0009
Median	-0.0703	0.0428	**	0.0049
Fear and Greed Index	-0.0022	0.9494		0.0000
VIX	-0.0009	0.9792		0.0000
T-1 Return	-0.0402	0.2476		0.0016
T-2 Return	0.0269	0.4383		0.0007
T-3 Return	0.0096	0.7830		0.0001
T-4 Return	0.0618	0.0748	*	0.0038

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

In the bivariate regression analysis, with Skewness as a fixed variable, the results are shown in Table 5-8. It can be observed that after adding a second variable, Skewness becomes insignificant in all cases, indicating that the statistical characteristic data of the RND functions constructed using Birru and Figlewski (2012) is less stable in predicting Bitcoin spot returns. The results of the three-variable and four-variable regression analyses are shown in Appendix Table 4 and Appendix Table 5.

Table 5-8: Bivariate Regression Results for Products with 1 Day to Expiration (Birru and Figlewski (2012))

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	R-squared
Mean	-0.0741	0.0336	**	-0.0327	0.3485		0.0061
Std	-0.0153	0.6629		-0.0232	0.5069		0.0009
Excess Kurtosis	0.0283	0.4168		-0.0232	0.5046		0.0014
Median	-0.0727	0.0367	**	-0.0307	0.3768		0.0059
Fear and Greed Index	-0.0008	0.9805		-0.0249	0.4735		0.0006
VIX	0.0015	0.9648		-0.0251	0.4715		0.0006
T-1 Return	-0.0466	0.1877		-0.0339	0.3380		0.0027
T-2 Return	0.0256	0.4613		-0.0236	0.4977		0.0013
T-3 Return	0.0091	0.7942		-0.0248	0.4754		0.0007
T-4 Return	0.0608	0.0804	*	-0.0221	0.5253		0.0043

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

To compare with the regression model constructed using the proposed method, this section also selects Skewness, Median, and T-4 Return as the three variables for the model. The model results are shown in Table 5-9, indicating that Skewness is not significant, and the regression model constructed using Birru and Figlewski (2012) does not have stable predictive ability.

Table 5-9: Regression Results for Products with 1 Day to Expiration (Birru and Figlewski (2012))

	Coefficient	p value	Significance	R-squared	MSE
Skewness	-0.0278	0.4233		0.0098	0.9890
Median	-0.0742	0.0330	**		
T-4 Return	0.0625	0.0718	*		

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

5.2.3 Comparison

Comparing the two methods for 1-day expiration options reveals significant differences in variable selection and predictive effects. The proposed method maintains one more effective sample than Birru and Figlewski (2012), reflecting better stability in practical applications.

In the three-variable model (Table 5-10), all variables reach statistical significance under the proposed method. Skewness shows a positive effect on returns, contrasting with findings by Bali and Murray (2013) and Conrad et al. (2013) in traditional markets. T-4 Return's significant predictive power aligns with Liu and Tsyvinski (2021) and Liu et al. (2023), indicating a significant momentum effect in cryptocurrency markets.

The proposed method demonstrates superior explanatory power (R-squared = 0.0130 vs. 0.0098) and lower Mean Squared Error (0.9858 vs. 0.9890). Out-of-sample R-squared values further confirm this advantage (0.0134 vs. 0.0121).

Table 5-10: Comparison of Regression Results for Products with 1 Day to Expiration
(Left: The proposed method; Right: Birru and Figlewski (2012))

	The proposed method						**Birru and Figlewski (2012)**					
	Coef.	p value	Sig.	R-squared	MSE	R^2_{OS}	Coef.	p value	Sig.	R-squared	MSE	R^2_{OS}
Skewness	0.0629	0.0690	*	0.0130	0.9858	0.0134	-0.0278	0.4233		0.0098	0.9890	0.0121
Median	-0.0746	0.0311	**				-0.0742	0.0330	**			
T-4 Return	0.0626	0.0705	*				0.0625	0.0718	*			

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

5.3 Regression Analysis with 7 Days to Expiration

5.3.1 Fitting Tails with GPDs Based on the Proposed Method

This section uses option products expiring daily from January 15, 2021, to April 19, 2024, deriving the RND from the observation date seven days before expiration, and constructs complete the RND functions using the proposed method. Moments such as mean, standard deviation, skewness, and kurtosis are then calculated as explanatory variables. Using the next period's Bitcoin spot return as the explained variable, multi-level regression analysis is conducted to observe whether the RND has predictive effects. The descriptive statistics of the variables are shown in Table 5-11, with a total of 119 samples. The means of Skewness and Excess Kurtosis are both greater than 0, and there are relatively few extreme values.

Table 5-11: Descriptive Statistics of the RND Moments and Bitcoin Returns for Products with 7 Days to Expiration (The proposed method)

	Count	Mean	Std	Min	25%	Median	75%	Max
T Return (Y)	119	-0.0070	0.0978	-0.3516	-0.0511	-0.0071	0.0423	0.3071
Mean	119	36529.4634	13781.4125	16591.1342	26070.6956	35496.4638	45453.7462	69393.9514
Std	119	3599.4222	2428.7791	710.3376	1727.0200	2868.8427	5085.2385	16566.9193
Skewness	119	0.0719	0.6877	-1.4832	-0.2484	0.0372	0.2935	4.2156
Excess Kurtosis	119	2.1573	2.9048	0.4791	1.2927	1.6335	1.9781	26.4692
Median	119	36497.0748	13738.8186	16678.0000	26066.9000	35379.1000	45497.1000	69137.5000
Fear and Greed Index	119	46.5378	22.4287	9.0000	25.0000	48.0000	69.0000	93.0000
VIX	119	20.0029	5.1284	12.2800	16.2950	18.8100	22.8100	32.0200
T-1 Return	119	-0.0060	0.0980	-0.3516	-0.0480	-0.0065	0.0463	0.3071
T-2 Return	119	-0.0061	0.0980	-0.3516	-0.0480	-0.0065	0.0463	0.3071
T-3 Return	119	-0.0055	0.0977	-0.3516	-0.0439	-0.0065	0.0463	0.3071
T-4 Return	119	-0.0055	0.0977	-0.3516	-0.0439	-0.0065	0.0463	0.3071

Following the regression analysis set in Section 4.5, the univariate regression results are shown in Table 5-12. It can be observed that Mean, Std, and Median are significant,

while the individual predictive ability of market sentiment indicators such as the Cryptocurrency Fear and Greed Index and Volatility Index (VIX) remains extremely low.

Table 5-12: Univariate Regression Results for Products with 7 Days to Expiration (The proposed method)

	Coefficient	p value	Significance	R-squared
Mean	-0.1559	0.0904	*	0.0243
Std	-0.1547	0.0931	*	0.0239
Skewness	-0.0607	0.5122		0.0037
Excess Kurtosis	-0.1145	0.2148		0.0131
Median	-0.1553	0.0917	*	0.0241
Fear and Greed Index	0.0277	0.7648		0.0008
VIX	-0.0505	0.5858		0.0025
T-1 Return	-0.0398	0.6677		0.0016
T-2 Return	0.0126	0.8920		0.0002
T-3 Return	0.0043	0.9626		0.0000
T-4 Return	-0.0849	0.3587		0.0072

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

In conducting multiple regression analyses with two, three, and four variables, we found that most explanatory variables did not exhibit significant predictive effects, as documented in Appendix Table 6 to Appendix Table 8. This phenomenon indicates that merely increasing the number of variables cannot effectively enhance the model's predictive capability and may instead lead to overfitting problems.

After iteratively testing various variable combinations, our research discovered that when predicting Bitcoin weekly returns, the pairing of Excess Kurtosis and Median demonstrated superior predictive performance. Building on this foundation, we further incorporated market sentiment indicators by adding the Cryptocurrency Fear and Greed Index to the model, which exhibited significant predictive power. Through careful selection of variable combinations, rather than indiscriminately increasing the number of variables, our research ultimately identified a prediction model with both statistical significance and economic meaning. The regression results are presented in Table 5-13.

Table 5-13: The RND Regression Results for Options with 7 Days to Expiration (The proposed method)

	Coefficient	p value	Significance	R-squared	MSE
Excess Kurtosis	-0.1620	0.0874	*	0.0666	0.9256
Median	-0.2706	0.0144	**		
Fear and Greed Index	0.2171	0.0561	*		

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

5.3.2 Fitting Tails with GPDs Based on Birru and Figlewski (2012)

This section employs options contracts expiring daily between January 15, 2021, and April 19, 2024, using observations 7 days prior to expiration to derive the RND. We construct the complete RND function using Birru and Figlewski (2012), then calculate statistics including mean, standard deviation, skewness, and kurtosis as explanatory variables. Using subsequent Bitcoin spot returns as the dependent variable, we conduct multi-level regression analyses to examine whether the RND possesses predictive power. The descriptive statistics of the variables are presented in Table 5-14, with a total sample size of 119. Both Skewness and Excess Kurtosis have means greater than 0, and the descriptive statistics are highly similar to those of the proposed method, indicating that for weekly returns, the two methods do not exhibit substantial differences.

Table 5-14: Descriptive Statistics of the RND Characteristics and Bitcoin Returns for Options with 7 Days to Expiration (Birru and Figlewski (2012))

	Count	Mean	Std	Min	25%	Median	75%	Max
T Return (Y)	119	-0.0070	0.0978	-0.3516	-0.0511	-0.0071	0.0423	0.3071
Mean	119	36529.5600	13781.3082	16592.5497	26071.3029	35496.4638	45453.7462	69393.9514
Std	119	3598.8321	2429.3086	708.7932	1725.0112	2868.8427	5085.2385	16566.9193
Skewness	119	0.0723	0.6879	-1.4832	-0.2469	0.0372	0.2912	4.2156
Excess Kurtosis	119	2.1404	2.9086	0.4327	1.2819	1.6078	1.9556	26.4692
Median	119	36497.0748	13738.8186	16678.0000	26066.9000	35379.1000	45497.1000	69137.5000
Fear and Greed Index	119	46.5378	22.4287	9.0000	25.0000	48.0000	69.0000	93.0000
VIX	119	20.0029	5.1284	12.2800	16.2950	18.8100	22.8100	32.0200
T-1 Return	119	-0.0060	0.0980	-0.3516	-0.0480	-0.0065	0.0463	0.3071
T-2 Return	119	-0.0061	0.0980	-0.3516	-0.0480	-0.0065	0.0463	0.3071
T-3 Return	119	-0.0055	0.0977	-0.3516	-0.0439	-0.0065	0.0463	0.3071
T-4 Return	119	-0.0055	0.0977	-0.3516	-0.0439	-0.0065	0.0463	0.3071

Following the regression analysis specified in Section 4.5, the univariate regression results are presented in Table 5-15. We observe that Mean, Standard Deviation, and Median exhibit statistical significance.

Table 5-15: Univariate Regression Results for Options with 7 Days to Expiration (The proposed method)

	Coefficient	p value	Significance	R-squared
Mean	-0.1559	0.0904	*	0.0243
Std	-0.1547	0.0930	*	0.0239
Skewness	-0.0608	0.5115		0.0037
Excess Kurtosis	-0.1157	0.2100		0.0134
Median	-0.1553	0.0917	*	0.0241
Fear and Greed Index	0.0277	0.7648		0.0008
VIX	-0.0505	0.5858		0.0025
T-1 Return	-0.0398	0.6677		0.0016
T-2 Return	0.0126	0.8920		0.0002
T-3 Return	0.0043	0.9626		0.0000
T-4 Return	-0.0849	0.3587		0.0072

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Birru and Figlewski (2012), when conducting multiple regression analyses with two,

three, and four variables, encounters the same issues as the proposed method, with most explanatory variables lacking significant predictive effects. These results are documented in Appendix Table 9 to Appendix Table 11.

To facilitate comparison with the regression model constructed using the proposed method, this section also selects Excess Kurtosis, Median, and the Cryptocurrency Fear and Greed Index as the three variables for inclusion in the model. The model results are presented in Table 5-16, with variable significance and model predictive capability closely resembling those of the proposed method model.

Table 5-16: The RND Regression Results for Options with 7 Days to Expiration (Birru and Figlewski (2012))

	Coefficient	p value	Significance	R-squared	MSE
Excess Kurtosis	-0.1620	0.0874	*	0.0666	0.9256
Median	-0.2706	0.0144	**		
Fear and Greed Index	0.2171	0.0561	*		

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

5.3.3 Comparison

For 7-day expiration options, both methods identify statistically significant variable combinations with remarkably similar RND characteristics. Both methods select Excess Kurtosis, Median, and the Cryptocurrency Fear and Greed Index as optimal predictors, with identical explanatory power (R-squared = 0.0666) and prediction error (MSE = 0.9256), as summarized in Table 5-17

The significant predictive power of Excess Kurtosis aligns with Amaya et al. (2015), who noted that excess kurtosis effectively captures extreme market risks. The Cryptocurrency Fear and Greed Index's significant predictive power for weekly returns is consistent with He et al. (2023) and López-Cabarcos et al. (2021).

Out-of-sample R-squared values are positive and substantial for both methods (0.3342 for the proposed method vs. 0.3335 for Birru and Figlewski (2012)), indicating economic value compared to historical average models.

Table 5-17: Comparison of the RND Regression Results for Options with 7 Days to Expiration
(Left: The proposed method; Right: Birru and Figlewski (2012))

	The proposed method						**Birru and Figlewski (2012)**					
	Coef.	p value	Sig.	R-squared	MSE	R^2_{OS}	Coef.	p value	Sig.	R-squared	MSE	R^2_{OS}
Excess Kurtosis	-0.1620	0.0874	*	0.0666	0.9256	0.3342	-0.1620	0.0874	*	0.0666	0.9256	0.3335
Median	-0.2706	0.0144	**				-0.2706	0.0144	**			
Fear and Greed Index	0.2171	0.0561	*				0.2171	0.0561	*			

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

5.4 Summary

The two methods exhibit different characteristics across prediction horizons. For daily return prediction, the proposed method clearly outperforms Birru and Figlewski (2012) in sample completeness, variable significance, and explanatory power. For weekly return prediction, performance is more comparable, possibly reflecting diminished impact of tail fitting methods over longer horizons.

Out-of-sample prediction results demonstrate positive R-squared values for both methods across both timeframes, with the proposed method consistently yielding higher values. Regarding computational efficiency, the proposed method demonstrates clear advantages with 10.95% faster execution time and more stable performance.

Different moments of risk-neutral probability density functions exhibit varying predictive capabilities across timeframes. For short-term (daily) prediction, skewness and historical returns demonstrate stronger predictive power, while for medium-term (weekly) prediction, excess kurtosis and market sentiment indicators play more important roles.

In conclusion, the proposed method not only performs better in daily return prediction but also offers significant advantages in computational efficiency, making it particularly valuable for large-scale analyses and time-sensitive applications.

6. Conclusions

6.1 Summary

This study introduces a methodological innovation for extracting smooth tails of risk-neutral density (RND) functions in Bitcoin options markets, addressing limitations in Birru and Figlewski (2012) approach. The proposed method ensures no kinks at connection points while reducing computational complexity, achieving 10.95% greater efficiency compared to the conventional approach.

Empirical analysis using Deribit trading data (January 2021-April 2024) demonstrates distinct predictive patterns across time horizons. For daily returns, the proposed method identifies skewness, median, and lagged returns as significant predictors (R-squared = 0.0130), outperforming Birru and Figlewski (2012) (R-squared = 0.0098). The negative relationship between skewness and returns aligns with Conrad et al. (2013), while the significance of lagged returns supports Liu and Tsyvinski's (2021) findings on cryptocurrency momentum effects. For weekly returns, both methods identify excess kurtosis, median, and the Cryptocurrency Fear and Greed Index as key predictors (R-squared = 0.0666), suggesting diminished impact of tail fitting methodology over longer horizons.

Out-of-sample tests confirm the economic value of the proposed approach with positive R-squared values (0.3342 versus 0.3335), substantially outperforming historical average models (Campbell & Thompson, 2008). These results validate that option-implied RND moments contain valuable forward-looking information, particularly in cryptocurrency markets characterized by extreme volatility. Our findings extend the application of RND characteristics to multi-time scale prediction in cryptocurrency markets, providing practical tools for risk management in these emerging markets.

6.2 Recommendations for Future Research

Based on our research findings, we propose the following recommendations for future research and practical applications:

1. Expansion of Research Scope

While our investigation centers on Bitcoin options, subsequent research should extend to additional cryptocurrencies and traditional financial markets to evaluate the comparative performance of the proposed method versus Birru and Figlewski (2012) across diverse market structures. Examining Ethereum options or analyzing stock index options markets would provide valuable insights into the cross-market predictive capabilities of RND moments.

2. Integration with Market Microstructure Factors

Our current analysis emphasizes options price-implied information. Future research should incorporate market microstructure factors (e.g., trading volume, bid-ask spread) to assess their relationship with cryptocurrency returns. Examining how these factors interact with RND moments could reveal whether microstructure effects complement or subsume the predictive power of option-implied information, potentially enhancing return forecasting models in these emerging markets.

References

Akyildirim, E., Corbet, S., Lucey, B., Sensoy, A., & Yarovaya, L. (2020). The relationship between implied volatility and cryptocurrency returns. *Finance Research Letters*, *33*, 101212. https://doi.org/10.1016/j.frl.2019.06.010

Amaya, D., Christoffersen, P., Jacobs, K., & Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, *118*(1), 135–167. https://doi.org/10.1016/j.jfineco.2015.02.009

Ammann, M., & Feser, A. (2019). Robust estimation of risk-neutral moments. *Journal of Futures Markets*, *39*(9), 1137–1166. https://doi.org/10.1002/fut.22020

Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. *The Review of Financial Studies*, *16*(1), 101–143. https://doi.org/10.1093/rfs/16.1.0101

Bali, T. G., & Murray, S. (2013). Does Risk-Neutral Skewness Predict the Cross Section of Equity Option Portfolio Returns? *The Journal of Financial and Quantitative Analysis*, *48*(4), 1145–1171.

Bali, T. G., & Zhou, H. (2016). Risk, Uncertainty, and Expected Returns. *The Journal of Financial and Quantitative Analysis*, *51*(3), 707–735.

Balkema, A. A., & Haan, L. de. (1974). Residual Life Time at Great Age. *The Annals of Probability*, *2*(5), 792–804. https://doi.org/10.1214/aop/1176996548

Baur, D. G., & Smales, L. A. (2022). Trading behavior in bitcoin futures: Following the "smart money." *Journal of Futures Markets*, *42*(7), 1304–1323. https://doi.org/10.1002/fut.22332

Birru, J., & Figlewski, S. (2012). Anatomy of a meltdown: The risk neutral density for the S&P 500 in the fall of 2008. *Journal of Financial Markets*, *15*(2), 151–180. https://doi.org/10.1016/j.finmar.2011.09.001

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, *81*(3), 637–654.

Bliss, R. R., & Panigirtzoglou, N. (2004). Option-Implied Risk Aversion Estimates. *The Journal of Finance*, *59*(1), 407–446. https://doi.org/10.1111/j.1540-6261.2004.00637.x

Bondarenko, O. (2000). *Recovering Risk-Neutral Densities: A New Nonparametric Approach* (SSRN Scholarly Paper No. 246063). Social Science Research Network. https://doi.org/10.2139/ssrn.246063

Böök, A., Imbet, J. F., Reinke, M., & Sala, C. (2025). The Forecasting Power of Short-Term Options. *The Journal of Derivatives*, *32*(3), 80–116. https://doi.org/10.3905/jod.2025.1.221

Breeden, D. T., & Litzenberger, R. H. (1978). Prices of State-Contingent Claims Implicit in Option Prices. *The Journal of Business*, *51*(4), 621–651.

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*(4), 1509–1531. https://doi.org/10.1093/rfs/hhm055

Chang, B. Y., Christoffersen, P., & Jacobs, K. (2013). Market skewness risk and the cross section of stock returns. *Journal of Financial Economics*, *107*(1), 46–68. https://doi.org/10.1016/j.jfineco.2012.07.002

Chen, R.-R., Hsieh, P., & Huang, J. (2018). Crash risk and risk neutral densities. *Journal of Empirical Finance*, *47*, 162–189. https://doi.org/10.1016/j.jempfin.2018.03.006

Chordia, T., Lin, T.-C., & Xiang, V. (2021). Risk-Neutral Skewness, Informed Trading, and the Cross Section of Stock Returns. *Journal of Financial and Quantitative Analysis*, *56*(5), 1713–1737. https://doi.org/10.1017/S0022109020000551

Christoffersen, P., Jacobs, K., & Chang, B. Y. (2013). Chapter 10—Forecasting with Option-Implied Information. In G. Elliott & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (Vol. 2, pp. 581–656). Elsevier. https://doi.org/10.1016/B978-0-444-53683-9.00010-4

Conrad, J., Dittmar, R. F., & Ghysels, E. (2013). Ex Ante Skewness and Expected Stock Returns. *The Journal of Finance*, *68*(1), 85–124. https://doi.org/10.1111/j.1540-6261.2012.01795.x

Cortés, L. M., Mora-Valencia, A., & Perote, J. (2020). Retrieving the implicit risk neutral density of WTI options with a semi-nonparametric approach. *The North American Journal of Economics and Finance*, *54*, 100862. https://doi.org/10.1016/j.najef.2018.10.010

Cujean, J., & Hasler, M. (2017). Why Does Return Predictability Concentrate in Bad Times? *The Journal of Finance*, *72*(6), 2717–2758. https://doi.org/10.1111/jofi.12544

Deribit. (2025). https://www.deribit.com/

Deribit Options. (2025). https://www.deribit.com/

Dong, B., Xu, W., & Cui, Z. (2024). Implied Willow Tree. *The Journal of Derivatives*, *31*(4), 44–74. https://doi.org/10.3905/jod.2024.1.200

Feng, Y., He, M., & Zhang, Y. (2024). Market Skewness and Stock Return Predictability: New Evidence from China. *Emerging Markets Finance and Trade*, *60*(2), 233–244. https://doi.org/10.1080/1540496X.2023.2217327

Figlewski, S. (2008). *Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio* (SSRN Scholarly Paper No. 1256783). Social Science Research Network. https://papers.ssrn.com/abstract=1256783

Fuertes, A.-M., Liu, Z., & Tang, W. (2022). Risk-neutral skewness and commodity futures pricing. *Journal of Futures Markets*, *42*(4), 751–785. https://doi.org/10.1002/fut.22308

Glatzer, E., & Scheicher, M. (2005). What moves the tail? The determinants of the option-implied probability density function of the DAX index. *Journal of Futures Markets*, *25*(6), 515–536. https://doi.org/10.1002/fut.20157

Grith, M., Härdle, W. K., & Schienle, M. (2012). Nonparametric Estimation of Risk-Neutral Densities. In J.-C. Duan, W. K. Härdle, & J. E. Gentle (Eds.), *Handbook of Computational Finance* (pp. 277–305). Springer. https://doi.org/10.1007/978-3-642-17254-0_11

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, *33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Hagan, P. S., & and West, G. (2006). Interpolation Methods for Curve Construction. *Applied Mathematical Finance*, *13*(2), 89–129. https://doi.org/10.1080/13504860500396032

Haslip, G. G., & and Kaishev, V. K. (2014). Lookback option pricing using the Fourier transform B-spline method. *Quantitative Finance*, *14*(5), 789–803. https://doi.org/10.1080/14697688.2014.882010

Hayashi, F. (2020). Analytically Deriving Risk-Neutral Densities from Volatility Smiles in Delta. *The Journal of Derivatives*, *27*(4), 6–12. https://doi.org/10.3905/jod.2020.1.099

He, M., Shen, L., Zhang, Y., & Zhang, Y. (2023). Predicting cryptocurrency returns for real-world investments: A daily updated and accessible predictor. *Finance Research Letters*, *58*, 104406. https://doi.org/10.1016/j.frl.2023.104406

He, Y., Peng, L., Zhang, D., & Zhao, Z. (2022). Risk Analysis via Generalized Pareto Distributions. *Journal of Business & Economic Statistics*, *40*(2), 852–867. https://doi.org/10.1080/07350015.2021.1874390

Hosking, J. R. M., & Wallis, J. R. (1987). Parameter and Quantile Estimation for the Generalized Pareto Distribution. *Technometrics*, *29*(3), 339–349. https://doi.org/10.2307/1269343

Hull, J. (2021). *Options, Futures, and Other Derivatives: Global Edition*. Pearson Deutschland. https://elibrary.pearson.de/book/99.150005/9781292410623

Jackwerth, J. (2020). What Do Index Options Teach Us About COVID-19? *The Review of Asset Pricing Studies*, *10*(4), 618–634. https://doi.org/10.1093/rapstu/raaa012

Jondeau, E., Wang, X., Yan, Z., & Zhang, Q. (2020). Skewness and index futures return. *Journal of Futures Markets*, *40*(11), 1648–1664. https://doi.org/10.1002/fut.22112

Kim, T. S., & Park, H. (2018). Is stock return predictability of option-implied skewness affected by the market state? *Journal of Futures Markets*, *38*(9), 1024–1042. https://doi.org/10.1002/fut.21921

Köse, N., Yildirim, H., Ünal, E., & Lin, B. (2024). The Bitcoin price and Bitcoin price uncertainty: Evidence of Bitcoin price volatility. *Journal of Futures Markets*, *44*(4), 673–695. https://doi.org/10.1002/fut.22487

Lehnert, T. (2022). Is Risk-Neutral Skewness an Indicator of Downside Risk? Evidence from Tail Risk Taking of Hedge Funds. *The Journal of Derivatives*, *29*(3), 65–84. https://doi.org/10.3905/jod.2022.1.148

Li, X., Wu, Z., Zhang, H., & Zhang, L. (2024). Risk-neutral skewness and stock market returns: A time-series analysis. *The North American Journal of Economics and Finance*, *70*, 102040. https://doi.org/10.1016/j.najef.2023.102040

Li, Y., Nolte, I., & Pham, M. C. (2024). Parametric risk-neutral density estimation via finite lognormal-Weibull mixtures. *Journal of Econometrics*, *241*(2), 105748. https://doi.org/10.1016/j.jeconom.2024.105748

Li, Y., Urquhart, A., Wang, P., & Zhang, W. (2021). MAX momentum in cryptocurrency markets. *International Review of Financial Analysis*, *77*, 101829. https://doi.org/10.1016/j.irfa.2021.101829

Liu, Y., & Chen, Y. (2024). Skewness risk and the cross-section of cryptocurrency returns. *International Review of Financial Analysis*, *96*, 103626. https://doi.org/10.1016/j.irfa.2024.103626

Liu, Y., Li, Z., Nekhili, R., & Sultan, J. (2023). Forecasting cryptocurrency returns with machine learning. *Research in International Business and Finance*, *64*, 101905. https://doi.org/10.1016/j.ribaf.2023.101905

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

López-Cabarcos, M. Á., Pérez-Pico, A. M., Piñeiro-Chousa, J., & Šević, A. (2021). Bitcoin volatility, stock market and investor sentiment. Are they connected? *Finance Research Letters*, *38*, 101399. https://doi.org/10.1016/j.frl.2019.101399

Markose, S., & Alentorn, A. (2011). The Generalized Extreme Value Distribution, Implied Tail Index, and Option Pricing. *The Journal of Derivatives*, *18*(3), 35–60. https://doi.org/10.3905/jod.2011.18.3.035

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, *7*(3), 271–300. https://doi.org/10.1016/S0927-5398(00)00012-8

Mei, D., Liu, J., Ma, F., & Chen, W. (2017). Forecasting stock market volatility: Do realized skewness and kurtosis help? *Physica A: Statistical Mechanics and Its Applications*, *481*, 153–159. https://doi.org/10.1016/j.physa.2017.04.020

Mohrschladt, H., & Schneider, J. C. (2021). Option-implied skewness: Insights from ITM-options. *Journal of Economic Dynamics and Control*, *131*, 104227. https://doi.org/10.1016/j.jedc.2021.104227

Monteiro, A. M., & Santos, A. A. F. (2022). Option prices for risk-neutral density estimation using nonparametric methods through big data and large-scale problems. *Journal of Futures Markets*, *42*(1), 152–171. https://doi.org/10.1002/fut.22258

Monteiro, A. M., Tütüncü, R. H., & Vicente, L. N. (2008). Recovering risk-neutral probability density functions from options prices using cubic splines and ensuring nonnegativity. *European Journal of Operational Research*, *187*(2), 525–542. https://doi.org/10.1016/j.ejor.2007.02.041

Neumann, M., & Skiadopoulos, G. (2013). Predictable Dynamics in Higher-Order Risk-Neutral Moments: Evidence from the S&P 500 Options. *The Journal of Financial and Quantitative Analysis*, *48*(3), 947–977.

Orosi, G. (2015). Estimating Option-Implied Risk-Neutral Densities: A Novel Parametric Approach. *The Journal of Derivatives*, *23*(1), 41–61. https://doi.org/10.3905/jod.2015.23.1.041

Reinke, M. (2020). Risk-Neutral Density Estimation: Looking at the Tails. *The Journal of Derivatives*, *27*(3), 99–125. https://doi.org/10.3905/jod.2019.1.090

Rubinstein, M. (1994). Implied Binomial Trees. *The Journal of Finance*, *49*(3), 771–818. https://doi.org/10.2307/2329207

Shimko, D. (1993). Bounds of probability. *Risk*, *6*(4), 33–37.

The Block. (2025). The Block. https://www.theblock.co/data/crypto-markets/options

Uberti, P. (2023). A theoretical generalization of the Markowitz model incorporating skewness and kurtosis. *Quantitative Finance*, *23*(5), 877–886. https://doi.org/10.1080/14697688.2023.2176250

Wang, Y.-H., & Yen, K.-C. (2018). The information content of option-implied tail risk on the future returns of the underlying asset. *Journal of Futures Markets*, *38*(4), 493–510. https://doi.org/10.1002/fut.21887

Welch, I., & Goyal, A. (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *The Review of Financial Studies*, *21*(4), 1455–1508. https://doi.org/10.1093/rfs/hhm014

Zulfiqar, N., & Gulzar, S. (2021). Implied volatility estimation of bitcoin options and the stylized facts of option pricing. *Financial Innovation*, *7*(1), 67. https://doi.org/10.1186/s40854-021-00280-y

Appendix

Appendix Table 1: Three-Variable Regression Results for Products with 1 Day to Expiration (The proposed method)

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	R-squared
Mean	-0.0783	0.0263	**	0.0844	0.4362		-0.0360	0.7387		0.0101
Std	-0.0239	0.4914		0.1232	0.2516		-0.0661	0.5387		0.0047
Median	-0.0718	0.0392	**	0.1055	0.3262		-0.0439	0.6833		0.0093
Fear and Greed Index	-0.0068	0.8444		0.1220	0.2566		-0.0636	0.5538		0.0042
VIX	0.0002	0.9961		0.1211	0.2600		-0.0631	0.5571		0.0042
T-1 Return	-0.0333	0.3398		0.1107	0.3049		-0.0555	0.6060		0.0053
T-2 Return	0.0300	0.3866		0.1233	0.2513		-0.0643	0.5492		0.0051
T-3 Return	0.0093	0.7899		0.1201	0.2638		-0.0620	0.5641		0.0043
T-4 Return	0.0610	0.0786	*	0.1229	0.2518		-0.0662	0.5369		0.0079

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 2: Four-Variable Regression Results for Products with 1 Day to Expiration (The proposed method)

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Std_Coef	Std_p	Std_Sig	R-squared
Mean	-0.0807	0.0341	**	0.0827	0.4478		-0.0343	0.7513		0.0063	0.8655		0.0101
Median	-0.0769	0.0496	**	0.1034	0.3373		-0.0411	0.7034		0.0111	0.7750		0.0094
Fear and Greed Index	-0.0016	0.9644		0.1234	0.2515		-0.0661	0.5386		-0.0235	0.5088		0.0048
VIX	-0.0012	0.9716		0.1233	0.2516		-0.0662	0.5383		-0.0240	0.4910		0.0048
T-1 Return	-0.0343	0.3269		0.1127	0.2968		-0.0584	0.5878		-0.0252	0.4691		0.0059
T-2 Return	0.0292	0.4005		0.1252	0.2442		-0.0671	0.5326		-0.0228	0.5113		0.0056
T-3 Return	0.0072	0.8356		0.1224	0.2552		-0.0651	0.5452		-0.0233	0.5046		0.0048
T-4 Return	0.0600	0.0837	*	0.1248	0.2451		-0.0688	0.5217		-0.0213	0.5401		0.0083

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 3: Four-Variable Regression Results Based on the Three-Variable Model (Daily Return The proposed method)

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	Median_Coef	Median_p	Median_Sig	T-4 Return_Coef	T-4 Return_p	T-4 Return_Sig	R-squared
Mean	-0.2567	0.2202		0.0177	0.7255		0.1761	0.3957		0.0633	0.0672	*	0.0148
Std	0.0169	0.6627		0.0637	0.0663	*	-0.0823	0.0342	**	0.0635	0.0670	*	0.0132
Excess Kurtosis	-0.0467	0.6641		0.1070	0.3185		-0.0733	0.0350	**	0.0628	0.0697	*	0.0132
Fear and Greed Index	0.0397	0.3479		0.0612	0.0771	*	-0.0966	0.0209	**	0.0576	0.0998	*	0.0140
VIX	-0.0237	0.5215		0.0625	0.0709	*	-0.0830	0.0250	**	0.0610	0.0788	*	0.0135
T-1 Return	-0.0329	0.3433		0.0598	0.0855	*	-0.0736	0.0335	**	0.0628	0.0695	*	0.0141
T-2 Return	0.0293	0.3970		0.0640	0.0647	*	-0.0751	0.0302	**	0.0618	0.0743	*	0.0139
T-3 Return	0.0141	0.6835		0.0631	0.0684	*	-0.0750	0.0305	**	0.0631	0.0685	*	0.0132

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 4: Three-Variable Regression Results for Products with 1 Day to Expiration (Birru and Figlewski (2012))

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	R-squared
Mean	-0.0730	0.0366	**	-0.0310	0.3751		0.0250	0.4717		0.0067
Std	-0.0157	0.6545		-0.0214	0.5418		0.0285	0.4133		0.0017
Median	-0.0712	0.0412	**	-0.0291	0.4034		0.0238	0.4941		0.0064
Fear and Greed Index	0.0005	0.9896		-0.0233	0.5050		0.0283	0.4172		0.0014
VIX	0.0001	0.9988		-0.0232	0.5069		0.0283	0.4177		0.0014
T-1 Return	-0.0447	0.2081		-0.0320	0.3677		0.0249	0.4755		0.0033
T-2 Return	0.0254	0.4661		-0.0219	0.5307		0.0280	0.4209		0.0021
T-3 Return	0.0102	0.7691		-0.0230	0.5092		0.0287	0.4107		0.0015
T-4 Return	0.0607	0.0809	*	-0.0203	0.5595		0.0281	0.4198		0.0051

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 5: Four-Variable Regression Results for Products with 1 Day to Expiration (Birru and Figlewski (2012))

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Std_Coef	Std_p	Std_Sig	R-squared
Mean	-0.0863	0.0318	**	-0.0355	0.3182		0.0241	0.4896		0.0269	0.5032		0.0072
Median	-0.0914	0.0305	**	-0.0350	0.3249		0.0220	0.5277		0.0358	0.3963		0.0073
Fear and Greed Index	0.0050	0.8909		-0.0215	0.5402		0.0287	0.4102		-0.0170	0.6400		0.0017
VIX	-0.0008	0.9813		-0.0213	0.5460		0.0285	0.4136		-0.0157	0.6542		0.0017
T-1 Return	-0.0451	0.2040		-0.0301	0.3996		0.0251	0.4721		-0.0168	0.6307		0.0036
T-2 Return	0.0251	0.4703		-0.0201	0.5675		0.0282	0.4175		-0.0153	0.6623		0.0023
T-3 Return	0.0088	0.8023		-0.0213	0.5439		0.0288	0.4084		-0.0148	0.6742		0.0017
T-4 Return	0.0605	0.0823	*	-0.0186	0.5957		0.0283	0.4165		-0.0147	0.6740		0.0053

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 6: Two-Variable Regression-Results for Products with 7 Days to Expiration (The proposed method)

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	R-squared
Mean	-0.1505	0.1066		-0.0401	0.6654		0.0259
Std	-0.1636	0.1212		0.0185	0.8601		0.0242
Excess Kurtosis	-0.1390	0.2782		0.0353	0.7822		0.0138
Median	-0.1502	0.1062		-0.0429	0.6430		0.0259
Fear and Greed Index	0.0803	0.4581		-0.1021	0.3462		0.0084
VIX	-0.0667	0.4830		-0.0750	0.4298		0.0079
T-1 Return	-0.0249	0.7963		-0.0538	0.5779		0.0043
T-2 Return	0.0248	0.7932		-0.0653	0.4901		0.0043
T-3 Return	0.0115	0.9021		-0.0620	0.5077		0.0038
T-4 Return	-0.0789	0.3982		-0.0515	0.5811		0.0098

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 7: Three-Variable Regression Results for Products with 7 Days to Expiration (The proposed method)

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	R-squared
Mean	-0.1591	0.0884	*	0.0685	0.5935		-0.1555	0.2228		0.0385
Std	-0.1410	0.2082		0.0644	0.6194		-0.0822	0.5435		0.0273
Median	-0.1591	0.0876	*	0.0660	0.6065		-0.1560	0.2212		0.0386
Fear and Greed Index	0.0621	0.5716		-0.0054	0.9709		-0.1264	0.3322		0.0165
VIX	-0.0355	0.7259		0.0163	0.9069		-0.1225	0.3708		0.0148
T-1 Return	-0.0750	0.4713		0.0820	0.5680		-0.1766	0.2032		0.0182
T-2 Return	0.0150	0.8744		0.0312	0.8120		-0.1370	0.2893		0.0140
T-3 Return	-0.0017	0.9859		0.0357	0.7840		-0.1393	0.2833		0.0138
T-4 Return	-0.0908	0.3330		0.0552	0.6700		-0.1524	0.2374		0.0218

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 8: Four-Variable Regression Results for Products with 7 Days to Expiration (The proposed method)

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Std_Coef	Std_p	Std_Sig	R-squared
Mean	-0.2019	0.2372		0.0648	0.6165		-0.1846	0.2507		0.0612	0.7640		0.0392
Median	-0.2031	0.2339		0.0615	0.6348		-0.1861	0.2477		0.0631	0.7574		0.0394
Fear and Greed Index	0.1350	0.2543		-0.0131	0.9283		-0.0336	0.8126		-0.1941	0.1101		0.0384
VIX	-0.0413	0.6830		0.0427	0.7614		-0.0622	0.6664		-0.1431	0.2038		0.0288
T-1 Return	-0.0743	0.4741		0.1105	0.4462		-0.1196	0.4106		-0.1406	0.2105		0.0317
T-2 Return	0.0145	0.8779		0.0603	0.6497		-0.0803	0.5561		-0.1410	0.2103		0.0275
T-3 Return	-0.0007	0.9940		0.0645	0.6251		-0.0823	0.5478		-0.1410	0.2102		0.0273
T-4 Return	-0.0904	0.3338		0.0841	0.5219		-0.0957	0.4819		-0.1407	0.2094		0.0353

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 9: Two-Variable Regression Results for Products with 7 Days to Expiration (Birru and Figlewski (2012))

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	R-squared
Mean	-0.1505	0.1066		-0.0401	0.6654		0.0259
Std	-0.1636	0.1212		0.0185	0.8601		0.0242
Excess Kurtosis	-0.1390	0.2782		0.0353	0.7822		0.0138
Median	-0.1502	0.1062		-0.0429	0.6430		0.0259
Fear and Greed Index	0.0803	0.4581		-0.1021	0.3462		0.0084
VIX	-0.0667	0.4830		-0.0750	0.4298		0.0079
T-1 Return	-0.0249	0.7963		-0.0538	0.5779		0.0043
T-2 Return	0.0248	0.7932		-0.0653	0.4901		0.0043
T-3 Return	0.0115	0.9021		-0.0620	0.5077		0.0038
T-4 Return	-0.0789	0.3982		-0.0515	0.5811		0.0098

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 10: Three-Variable Regression Results for Products with 7 Days to Expiration (Birru and Figlewski (2012))

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	R-squared
Mean	-0.1591	0.0884	*	0.0685	0.5935		-0.1555	0.2228		0.0385
Std	-0.1410	0.2082		0.0644	0.6194		-0.0822	0.5435		0.0273
Median	-0.1591	0.0876	*	0.0660	0.6065		-0.1560	0.2212		0.0386
Fear and Greed Index	0.0621	0.5716		-0.0054	0.9709		-0.1264	0.3322		0.0165
VIX	-0.0355	0.7259		0.0163	0.9069		-0.1225	0.3708		0.0148
T-1 Return	-0.0750	0.4713		0.0820	0.5680		-0.1766	0.2032		0.0182
T-2 Return	0.0150	0.8744		0.0312	0.8120		-0.1370	0.2893		0.0140
T-3 Return	-0.0017	0.9859		0.0357	0.7840		-0.1393	0.2833		0.0138
T-4 Return	-0.0908	0.3330		0.0552	0.6700		-0.1524	0.2374		0.0218

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 11: Four-Variable Regression Results for Products with 7 Days to Expiration (Birru and Figlewski (2012))

	Coef	p value	Sig	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Std_Coef	Std_p	Std_Sig	R-squared
Mean	-0.2019	0.2372		0.0648	0.6165		-0.1846	0.2507		0.0612	0.7640		0.0392
Median	-0.2031	0.2339		0.0615	0.6348		-0.1861	0.2477		0.0631	0.7574		0.0394
Fear and Greed Index	0.1350	0.2543		-0.0131	0.9283		-0.0336	0.8126		-0.1941	0.1101		0.0384
VIX	-0.0413	0.6830		0.0427	0.7614		-0.0622	0.6664		-0.1431	0.2038		0.0288
T-1 Return	-0.0743	0.4741		0.1105	0.4462		-0.1196	0.4106		-0.1406	0.2105		0.0317
T-2 Return	0.0145	0.8779		0.0603	0.6497		-0.0803	0.5561		-0.1410	0.2103		0.0275
T-3 Return	-0.0007	0.9940		0.0645	0.6251		-0.0823	0.5478		-0.1410	0.2102		0.0273
T-4 Return	-0.0904	0.3338		0.0841	0.5219		-0.0957	0.4819		-0.1407	0.2094		0.0353

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level