Mendel University in Brno

MENDELU Working Papers
in Business and Economics
99/2025

Risk Without Reward? The Introduction of Bitcoin spot ETFs

Daniel Pastorek, Peter Albrecht

MENDELU Working Papers in Business and Economics
Research Centre
Faculty of Business and Economics
Mendel University in Brno
Zemědělská 1, 613 00 Brno
Czech Republic
http://vyzc.pef.mendelu.cz/en
+420 545 132 605

Citation

Pastorek, D., & Albrecht, P. (2025). *Risk Without Reward? The Introduction of Bitcoin Spot ETFs.* MENDELU Working Papers in Business and Economics, 99/2025. Mendel University in Brno. Mendel University in Brno. Cited from: XXX

Abstract

Daniel Pastorek and Peter Albrecht: Risk Without Reward? The Introduction of Bitcoin spot ETFs

Our study examines to what extent the introduction of Bitcoin spot exchange-traded funds (ETFs) affected Bitcoin's properties, including market dynamics, volatility, returns, return distribution, and tracking errors. Using block bootstrap simulations, OLS regression, EGARCH modeling, and non-parametric tests, we find that Bitcoin ETFs increase volatility and downside risk while leaving average returns unchanged. Return distribution shifts, including reduced skewness and kurtosis, suggest partial normalization, typically linked to greater liquidity and market participation. However, unlike traditional ETFs, Bitcoin ETFs introduce fail-to-deliver (FTD) occurrences—previously absent in Bitcoin markets—which mitigate extreme price movements through delayed settlement. Tracking error analysis confirms that spot ETFs more accurately track Bitcoin's price than futures-based ETFs. These findings offer critical insights into Bitcoin ETFs' market effects, particularly regarding stability and investor behavior.

Key words

Bitcoin, ETFs, volatility, market dynamics, FTDs

JEL: G11, G12, G14, G23, C58

Contacts

Mendel University in Brno, Faculty of Business and Economics, Zemědělská 1, 613 00 Brno, Czech Republic, email: daniel.pastorek@mendelu.cz; peter.albrecht@mendelu.cz.

Acknowledgements

Introduction

The launch of a traditional stock ETF typically meets well-established expectations based on decades of market behavior. Financial theory and empirical evidence indicate that these ETFs generally improve market stability and efficiency (Ben-David et al., 2017; Ben-David et al., 2018; Laborda et al., 2024). In contrast, the introduction of Bitcoin (BTC) spot ETFs in early 2024 presents a unique challenge. Bitcoin differs significantly from traditional stocks due to its high volatility, minimal regulation, and fundamentally distinct characteristics (Katsiampa, 2017; Corbet et al., 2018). Operating on a decentralized, peer-to-peer network, Bitcoin's price is influenced by factors such as speculation (Ciaian et al., 2018), regulatory changes (Auer and Claessens, 2018), and macroeconomic trends that are distinct from those affecting traditional stock markets (Bouri et al., 2017).

The introduction of the Bitcoin spot ETF raises several critical questions for academic discussion. Primarily, has the launch of these ETFs made Bitcoin a less risky asset? The question is rooted in the expectation that ETFs generally stabilize prices and reduce volatility across various asset classes (Todorov, 2021). By enhancing liquidity and increasing market participation, especially from institutional investors, Bitcoin spot ETFs could potentially lower volatility. Theoretically, greater liquidity enables more market participants to absorb price fluctuations, thereby mitigating extreme price movements. Research on ETFs in other markets, such as commodities and stock indices, supports this hypothesis, showing that ETFs can indeed moderate volatility (Todorov, 2024). However, no studies have yet examined this effect in the context of Bitcoin spot ETFs. To address this gap, we empirically assess volatility and risk-return dynamics before and after the ETF introduction, utilizing block bootstrap analysis to evaluate distributional shifts and a regression model to quantify changes in return volatility.

Furthermore, research on traditional financial markets suggests that ETF adoption can impact return dynamics by improving market efficiency (Israeli et al., 2017; Glosten et al., 2021). Increased liquidity and faster price corrections tend to reduce arbitrage opportunities, potentially leading to lower returns (Ben-David et al., 2018). However, it remains unclear whether this effect extends to Bitcoin, given its distinct market structure and the persistence of speculative activity. While some studies suggest that arbitrage opportunities in cryptocurrency markets have declined over time (Borri, 2019), inefficiencies remain due to price fragmentation across exchanges (Makarov and Schoar, 2020). We empirically assess whether the introduction of Bitcoin spot ETFs has influenced average returns, testing if increased market efficiency leads to a measurable change in Bitcoin's return levels through regression analysis.

Moreover, previous experiences in other markets indicate that the launch of ETFs tends to alter other statistical properties as well. By reducing extreme price fluctuations, particularly negative shocks, an ETF might theoretically decrease both kurtosis and skewness in Bitcoin's return distribution. Research on derivative products and ETFs suggest that greater liquidity and market maturity often lead to return distributions that approximate normality, exhibiting reduced skewness and kurtosis (Poterba & Shoven, 2002). However, this effect may not straightforwardly apply to Bitcoin, which retains a high degree of speculative trading, as noted by Bouri et al. (2017). Using block bootstrap analysis and non-parametric tests, we confirm a decrease in kurtosis and skewness following the introduction of Bitcoin ETFs. To explain these shifts, we examine the role of fail-to-deliver (FTD) occurrences, which, unlike in equity markets, were previously absent in Bitcoin trading. Leveraging statistical clustering techniques, we investigate whether FTDs contribute to reduced tail risk by smoothing extreme fluctuations through delayed settlement, introducing a novel mechanism to Bitcoin's market structure.

Lastly, while Bitcoin spot ETFs aim to closely track Bitcoin's spot price through direct asset holdings, some tracking error remains inevitable, as observed in traditional stock ETFs (Aber et al., 2009). Broad-market stock ETFs typically exhibit minimal tracking errors (0.05%–0.50% annually), while more volatile sector-specific ETFs, such as those tracking emerging markets, can experience tracking errors up to 1.5% (Cremers et al., 2013; Elton et al., 2002). Given Bitcoin's inherent volatility, we extend our analysis to quantify tracking errors in Bitcoin ETFs, comparing spot-based and futures-based structures to assess their relative tracking accuracy.

1 Data and Methods

1.1 Data

To assess the Bitcoin ETF's impact on Bitcoin's risk-return profile, we employ a multi-method approach, combining block bootstrap, OLS regression, EGARCH modelling, and non-parametric Kolmogorov-Smirnov (KS) and Mann-Whitney U tests. Such suite of methodologies enables us to capture potential shifts in returns, volatility clustering, and distributional changes, offering a comprehensive view of the ETF's effect on Bitcoin's risk dynamics. In these exercises, we use daily data from Bloomberg covering the period from the beginning of January 2023 till the end of December 2024, nearly a year after the introduction of Bitcoin spot ETFs and encompassing Bitcoin's all-time price high. The data covers a recent and comparable pre-ETF period, ensuring that the analysis captures Bitcoin's behavior under similar market conditions immediately preceding the introduction of spot ETFs.

Additionally, we incorporate daily data on settlement failures, specifically fails-to-Deliver (FTD) reports from the U.S. Securities and Exchange Commission (SEC) for Bitcoin spot ETFs. These ETFs include IBIT,

GBTC, FBTC, ARKB, BITB, HOLD, BRRR, BTCO, EZBC, and BTCW. The regulatory filings document instances where ETF shares were not successfully delivered at settlement. While FTDs have been studied in traditional equity markets as indicators of market inefficiencies and short-selling constraints, their emergence in Bitcoin spot ETFs represents a novel development in cryptocurrency markets. The inclusion of FTD data enables a structured examination of potential liquidity constraints and market structure effects arising from these settlement failures.

1.2 *Risk and distribution analysis*

We employ the block bootstrap technique in order to simulate return distributions pre- and post-ETF, crucially preserving time-series autocorrelation by sampling overlapping blocks. Such simulation allows us to estimate key risk and distributional metrics: VaR and CVaR provides insights into tail risk, while skewness and kurtosis describes distributional asymmetry and tail characteristics, highlighting shifts in return distributions due to the ETF's introduction.

Moreover, we apply the Kolmogorov-Smirnov and Mann-Whitney U tests to evaluate distributional changes in Bitcoin returns by comparing cumulative and rank distributions for pre- and post-ETF periods. These non-parametric tests identify significant shifts in distribution shape and scale, and offer insights beyond mean change, particularly important for assets like Bitcoin that often deviate from normality.

1.3 Volatility and return analysis

A To analyse changes in average returns and volatility levels, we employ two simple OLS regression models incorporating a dummy variable, *Period*, to differentiate between pre-ETF (Period = 0) and post-ETF (Period = 1) periods. The *Returns Model* estimates the effect of the ETF launch on daily returns, where the coefficient β_1 represents the average pre-ETF return and captures the incremental return effect following the ETF's introduction:

$$Returns_t = \beta_0 + \beta_1 Period_t + \varepsilon_t. \tag{1}$$

For volatility, the *Absolute Returns Model* used the absolute value of returns as a proxy, where α_0 denotes average pre-ETF volatility, and α_1 measures post-ETF changes:

$$|Returns_t| = \alpha_0 + \alpha_1 Period_t + \varepsilon_t. \tag{2}$$

To capture time-varying volatility and volatility clustering, we employ separate Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) models for the pre- and post-ETF periods. The EGARCH(1,1) specification is suitable as it allows for asymmetric responses to shocks, capturing the distinctive behaviour of volatility in response to positive and negative returns (Yildirim and Bekun, 2023):

$$\ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \alpha \frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}}. \tag{3}$$

where ω is the constant term in the variance equation; σ_t^2 is the conditional variance at time t. ε_t is the residual error term at time t; γ is the coefficient for the sign of the lagged standardized residuals, capturing the asymmetric effect of shocks on volatility; α represents the coefficient for the magnitude of the lagged standardized residuals, capturing the symmetric effect of shocks on volatility, and β is the coefficient for the lagged conditional variance, indicating volatility persistence.

By comparing the estimated coefficients - specifically, ω (the baseline volatility component), α_1 (shock sensitivity), and β_1 (volatility persistence) - across periods, we analyse potential shifts in volatility dynamics that may be associated with the ETF's introduction. The approach enables us to observe changes in both the persistence and asymmetry of volatility, providing a more nuanced view of how market conditions and investor sentiment might evolve in response to the availability of Bitcoin ETFs.

1.4 Fail-to-Deliver Occurrences and Market Stability

A To analyze the role of fail-to-deliver (FTD) occurrences in Bitcoin ETFs, we examine daily FTD quantities alongside Bitcoin price and volume data. Given the T+2 settlement cycle, FTD data are shifted backward by two business days to align with the original trading activity that led to settlement failures.

Extreme price movements are defined dynamically based on Bitcoin's historical volatility rather than a fixed return threshold. Specifically, we compute a 30-day rolling mean and standard deviation of Bitcoin's daily returns and classify extreme price movements as days when absolute returns exceed two standard deviations from the rolling mean. Such an approach provides a more robust identification of large price swings, as it adjusts for changing market volatility and avoids the limitations of an arbitrary threshold. The methodology is consistent with prior financial econometric research (e.g., Pastorek et al., 2023) that defines extreme events relative to recent volatility regimes (Albrecht and Kočenda, 2024). To classify periods of heightened volatility, we construct a 30-day rolling volatility measure and define high volatility periods as the top 20% of observations. High trading volume days are similarly identified as those in the top 20% of daily Bitcoin trading volume. These percentile-based thresholds balance sample size considerations while capturing periods of significant market activity.

To test whether FTD occurrences are associated with specific market conditions, we conduct independent two-sample t-tests. First, FTD quantities on days with extreme price movements are compared against standard trading days to assess whether settlement failures are more frequent during large price fluctuations. Second, we examine FTD levels across high- and low-volatility periods to determine whether settlement failures correspond to shifts in market uncertainty. The method provides statistical evidence on whether FTDs are more likely to occur under specific market stress conditions.

To further assess whether FTD occurrences are driven by liquidity constraints or arise as a response to volatility, we estimate an ordinary least squares (OLS) regression model with log-transformed FTD quantities as the dependent variable. The model specification is given as:

$$\begin{gathered}\log(FTD_t) = \beta_0 + \beta_1 High\,Volume_t + \beta_2 High\,Volatility_t \\ +\beta_3(High\,Volume\ x\ High\,Volatility)_t + \varepsilon_t\end{gathered} \tag{4}$$

where *High Volume* is a binary variable indicating whether Bitcoin's trading volume is in the top 20% of observations; *High Volatility* is a binary indicator for rolling volatility in the highest 20% of the sample, and the interaction term captures whether FTD occurrences intensify disproportionately under simultaneous high volume and high volatility conditions. The log transformation of FTD quantities ensures that the dependent variable exhibits a more normal distribution, improving the robustness of the regression estimates. By evaluating the statistical significance and magnitude of these coefficients, the model assesses whether settlement failures primarily reflect market liquidity constraints or function as a stabilizing mechanism in response to volatility shocks.

1.5 Identification of the tracking error

Additionaly, we utilize daily data spanning from January 2021 to December 2024 in order to analyze the tracking error of Bitcoin ETFs and further explore potential changes in correlation between Bitcoin and the broader financial market following the introduction of spot ETFs. The dataset captures Bitcoin's price and volume, S&P 500 index data, and two categories of Bitcoin ETFs: spot-based and futures-based. The extended period provides a robust pre- and post-ETF approval comparison, enabling a detailed investigation of how these ETFs perform relative to the underlying asset and their potential market impact.

We calculate tracking errors for each ETF as the rolling 30-day standard deviation of the difference in returns between each ETF and Bitcoin. Average tracking errors for spot and futures ETFs are computed separately and compared over time to assess whether spot ETFs exhibit lower tracking error, as suggested by literature on similar asset classes.

2 Results

In this section, we examine how the observed effects of Bitcoin ETFs align with expectations derived from research on traditional ETFs in the stock and commodity markets (see Todorov, 2024). Specifically, we assess the extent to which Bitcoin spot ETFs reflect the anticipated impacts commonly associated with ETF introduction, including: 1) reduction in volatility through enhanced market liquidity and participation, 2) reduced returns resulting from greater market efficiency, 3) alterations in the distributional properties of returns, particularly skewness and kurtosis, 4) risk of tracking error.

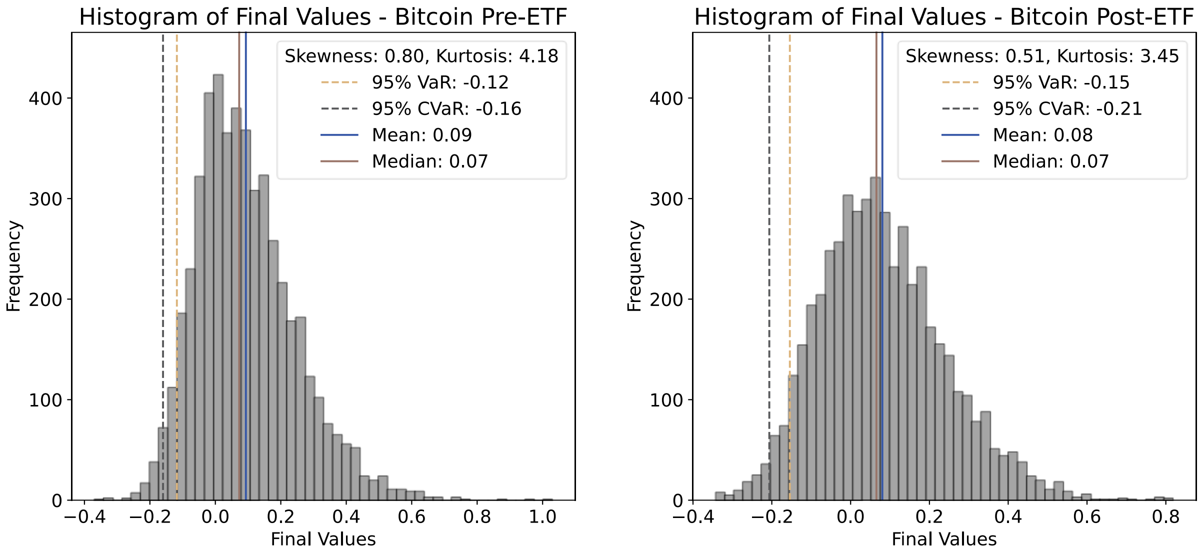
Table 1: Hypothesized and Observed Effects of Bitcoin Spot ETFs on Bitcoin

Effect Area	**Expectation**	**Estimated Resulting Effect**
Reduction in Volatility	▼ Decrease	▲ Increase
Reduced Returns (Market Efficiency)	▼ Decrease	↔ Stable
Changes in Return Distribution	↔ Normalization	↔ Partial normalization
Tracking Error	▼ Decrease	▼ Decrease

2.1 Risk-return properties

In Figure 1, we present histograms of final investment values using a block bootstrap method. The figure reveals a decrease in both skewness and kurtosis in the post-ETF period. Table 2 corroborates these findings, showing lower skewness and kurtosis metrics and higher downside risk, as indicated by more negative CVaR and VaR values. Consequently, we cannot confirm the initial hypothesis of reduced volatility, which has been observed in other markets like stocks and commodities (Todorov, 2024). However, both Figure 1 and Table 2 confirm a decrease in skewness and kurtosis. These statistical changes impact several financial factors. Reduced skewness and kurtosis enhance the predictability of returns (Poterba and Shoven, 2002) and extreme risks (Bouri et al., 2017). Additionally, the improved normality of the data indicates a more mature market (Ben-David et al., 2018) and enhances the efficiency of diversification strategies (Corbet et al., 2019).

Figure 1: Distribution of Final Values for Bitcoin Pre- and Post-ETF Periods



Note: The figure presents histograms of simulated final investment values over a 30-month horizon for Bitcoin pre- and post-ETF periods. Using a block bootstrap methodology with a 6-month block size, we generate 5000 resampled returns series to capture time-series dependence, critical for maintaining Bitcoin's volatility clustering and autocorrelation. The final values represent cumulative returns over the horizon, expressed as multiples of the initial investment. Value at Risk (VaR) and Conditional Value at Risk (CVaR) at the 95% confidence level capture tail risk, while mean, median, skewness, and kurtosis summarize the distribution's central tendency, asymmetry, and tail behaviour.

Furthermore, by employing a combination of block bootstraps, OLS regression, GARCH models, and non-parametric distributional tests, we examine shifts in Bitcoin's risk-return dynamics. The block bootstrap analysis (Figure 1 and Table 2) reveals higher values for Value at Risk (VaR) and Conditional Value at Risk (CVaR) in the post-ETF period, indicating an increased likelihood of significant losses. Such a rise in downside risk may could be linked to intensified speculative activity, a well-documented driver of Bitcoin's price dynamics (Ciaian et al., 2018).

Table 2: Distribution Metrics for Bitcoin Returns Pre- and Post-ETF

Metrics	**Pre-ETF Distribution**	**Post-ETF Distribution**
Mean	0.0933	0.0806
Median	0.0728	0.0655
VaR (5%)	-0.1175	-0.1542
CVaR (5%)	-0.1599	-0.2063

Skewness	0.7959	0.5100
Kurtosis	4.1770	3.4513

Note: The table presents key distributional metrics for Bitcoin returns, calculated using a block bootstrap approach with 5,000 resamples. The metrics include Value at Risk (VaR) and Conditional Value at Risk (CVaR) at the 5% level, mean, median, skewness, and kurtosis for the pre-ETF and post-ETF periods.

Moreover, we analyze the returns as the ETF adoption tends to be followed by reduced returns due to iincreased market efficiency. Such findings were confirmed for stock markets (Israeli et al., 2017), but the effect of BTC ETF introduction remains a question. The OLS regression results indicated no substantial change in average returns, as the post-ETF coefficient on the period dummy was not statistically significant ($\beta_1 = -0.0007$) (see Table 3). However, the regression model for volatility, proxied by absolute returns, showed a statistically significant increase in daily volatility ($\alpha_1 = 0.0047$) post-ETF, which suggests that the post-era of ETF launch has amplified price fluctuations rather than stabilized them (Table 3). Therefore, in contrast to one of pivotal studies on stocks by Todorov (2021), we find that the introduction of Bitcoin ETFs increased its risk profile while leaving average returns unchanged.

Table 3: Regression Results for Returns and Volatility Models Pre- and Post-ETF Periods

Model	**Coefficient**	**Estimate**	**Std. Error**	**t-Statistic**	**p-value**	**Observations**
Returns Model	Constant (β0)	0.0030	0.0013	2.31	0.02**	729
	Period Dummy (β1)	-0.0007	0.0019	-0.37	0.71	729
	R-squared	0.0002				729
Volatility Model	Constant (α0)	0.0156	0.0010	15.60	<0.01***	729
	Period Dummy (α1)	0.0047	0.0014	3.36	<0.01***	729
	R-squared	0.0159				729

Note: This table reports the results of two Ordinary Least Squares (OLS) regression models that examine the impact of the Bitcoin ETF on average returns and volatility. The Returns Model regresses daily returns on a constant term (β_0) and a period dummy variable (β_1) that distinguishes between pre-ETF and post-ETF periods. dSignificance levels are marked as ***p < 0.01, **p < 0.05, and *p < 0.1.

Following the previous results, we employ the EGARH[1] models to account for the persistence of volatility and sensitivity of previous shocks. These EGARCH(1,1) models for the pre and post-ETF periods (Table 4) further highlight a shift in volatility dynamics. The pre-ETF model exhibits limited volatility persistence, as evidenced by the statistically insignificant β_1 coefficient ($\beta_1 = 0.0848$, $p = 0.624$). Additionaly, it indicates that volatility shocks dissipated quickly. In contrast, the post-ETF EGARCH model revealed substantial volatility persistence ($\beta_1 = 0.8489$, $p < 0.001$), suggesting that price fluctuations became more prolonged following the ETF introduction. Moreover, the statistically significant α_1 coefficient ($\alpha_1 = 0.2290$, $p = 0.028$) in the post-ETF period indicates heightened sensitivity to past market shocks. These results underscore a structural change in Bitcoin's volatility profile, which could be attributed to the ETF's impact on market depth, potentially drawing institutional investors whose trading behaviors contribute to sustained volatility clustering. Such results are in contrast to prior expectations derived from other markets (Table 1) (Ben-David et al., 2017; Laborda et al., 2024).

Table 4: EGARCH Model Parameter Estimates for Pre- and Post-ETF Periods

Metric	**Parameter**	**Estimate**	**Std. Error**	**t-Statistic**	**p-value**	**Observations**
EGARCH Pre-ETF	μ	0.1706	0.115	1.485	0.138	374
	ω	1.5471	0.339	4.558	<0.01***	374
	α_1	0.3751	0.152	2.465	0.014*	374
	β_1	0.0848	0.173	0.490	0.624	374
EGARCH Post-ETF	μ	0.2225	0.137	1.625	0.104	355
	ω	0.3148	0.196	1.609	0.108^{*}	355
	α_1	0.2990	0.104	2.196	0.028^{*}	355

[1] We employ EGARCH due to its advantages in modeling the asymmetric effects (Yildirim and Bekun, 2023).

	β_1	0.8489	0.096	8.875	<0.01***	355

Note: The table presents the estimated parameters from the Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) model, fitted separately for the pre-ETF and post-ETF periods. The parameters ω (omega), α_1 (alpha), and β_1 (beta) represent the model's conditional volatility dynamics, with ω capturing the constant component of variance, α_1 representing the impact of past shocks on volatility, and β_1 indicating volatility persistence. Significance levels are marked as $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.1$.

Finally, we conduct the Kolmogorov-Smirnov (KS) and Mann-Whitney U tests to account for shifts in return distributions. As outlined by Poterba and Shoven (2002), return distribution provides information about the risk profile of the investment. Information about the distribution of returns can provide helpful insights for portfolio optimization. Following the results (Table 5), we confirm a significant shift in return distributions between the pre and post-ETF periods (KS statistic = 0.1060, p < 0.001). It indicates that the introduction of the Bitcoin ETF has notably altered Bitcoin's risk-return profile. However, the Mann-Whitney U test, which evaluates median differences, found no statistically significant change between the two periods (p = 0.7706). Together, these results suggest that while the ETF's introduction has led to distributional changes, particularly increased volatility and downside risk, these effects are not reflected in a shift in central tendencies, such as median returns (see Table 5).

Table 5: Non-Parametric Tests for Distributional Differences

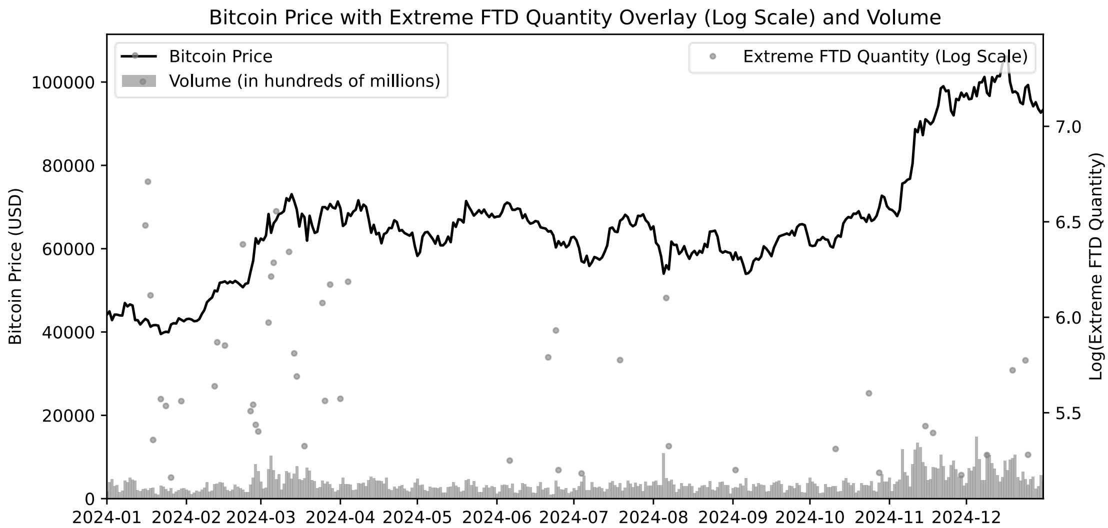
Test	**Statistic**	**p-Value**	**Observations**	**Interpretation**
Mann-Whitney U Test	67214	0.7706	Pre-ETF: 374, Post-ETF: 355	No significant difference in return distributions
Kolmogorov-Smirnov (KS)	0.1060	<0.01***	Pre-ETF: 374, Post-ETF: 355	Significant shift in return distributions

Note: The Mann-Whitney U test to evaluate median differences in return distributions, while the Kolmogorov-Smirnov test assess shifts in cumulative distributions, both comparing pre- and post-ETF periods.

2.2 Existence of fail-to-deliveries

The observed shifts in Bitcoin's statistical properties raise important questions about their underlying drivers. One of the aims of Bitcoin ETFs is to enhance market transparency and reduce manipulative practices, such as wash trading, by introducing higher standards of regulatory oversight and compliance. In traditional markets, ETFs are linked to increased transparency and reduced manipulation (Fotak et al., 2014). However, Bitcoin ETFs employing fail-to-deliver (FTD) mechanisms (or synthetic positions to meet demand without direct asset holdings) may introduce new risks, including counterparty exposure and potential market stress vulnerabilities (Pastorek et al., 2023). In this section, we examine the emergence of FTD occurrences in the Bitcoin market and their implications for market dynamics.

Figure 2: Bitcoin Price with Extreme FTD Quantity Overlay (Log Scale) and Volume



Notes: The figure presents Bitcoin's daily closing price alongside extreme fail-to-deliver (FTD) quantities, shown in a log scale, from January to December 2024. The black line represents the Bitcoin price, while the grey bar plot at the bottom indicates trading volume (scaled in hundreds of millions). The grey dots denote extreme FTD quantities, defined as values exceeding twice the standard deviation of typical FTD quantities after removing outliers. Additionally, FTD data are shifted backward by the standard settlement period to align with the trading conditions that precipitate settlement failures. This adjustment accounts for the lag between trading activity and FTD reporting, allowing for more accurate alignment of extreme FTD events with underlying market conditions, particularly in relation to Bitcoin's price trends and trading volume.

In Figure 2, we illustrate distinct clusters of extreme fail-to-deliver (FTD) occurrences that align with significant movements in Bitcoin's price, particularly during phases of upward trends or heightened volatility. Notably, the early and mid-2024 periods feature several events with extreme FTD occurrences coinciding with pronounced price increases and fluctuations. The clustering suggests that periods of rapid price movement may create conditions under which market participants struggle to settle trades, leading to elevated FTD levels. Conversely, during more stable price periods or moderate downward trends, the figure reveals fewer extreme FTDs. It implies that calmer markets with reduced price pressure and volatility are associated with lower FTD occurrences, likely due to steadier market dynamics and enhanced liquidity. These findings align with the results from Figure 1 and Table 2. In Figure 1 and Table 2, we found that despite heightened volatility after the ETF introduction, the normality of Bitcoin improved as the ETF led to lower skewness and kurtosis. Such results indicate that Bitcoin obtained less extreme values after introducing the ETF. The occurrence of FTDs further evaluates this finding – when Bitcoin price starts to decrease, FTDs occur and the settlement is spread over several days. It complements previous research stating that FTDs may lead to lower price extremes (Fotak et al., 2014).

The analysis of fail-to-deliver (FTD) occurrences in Bitcoin ETFs reveals important insights into their relationship with trading volume, volatility, and price dynamics. The volume data provides a clear indication that FTD events are strongly linked to high trading activity. Specifically, elevated FTD levels consistently coincide with spikes in trading volume, particularly during periods of heightened market activity, such as the early 2024 price surge, mid-year fluctuations, and the late 2024 rally. To further examine whether high volume is the primary driver of FTD occurrences or whether volatility also plays a role, we conduct a series of statistical tests.

The results of the t-tests (Table 7) show a significant increase in FTD levels on high trading volume days, confirming that trading intensity is a primary driver of settlement failures. In contrast, FTD levels do not differ significantly between high- and low-volatility periods, suggesting that market turbulence alone does not exacerbate settlement failures when decoupled from high trading activity. The finding implies that FTDs are more likely a response to liquidity constraints driven by trading demand rather than a reaction to broader market uncertainty.

Table 6: OLS Regression Results for FTDs

Variable	**Coefficient**	**Std Error**	**t-value**	**p-value**
Intercept	4.8568	0.344	14.116	0.000***
High Volume	3.2302	0.750	4.309	0.000***
High Volatility	0.9512	0.780	1.219	0.223
High Volume and Volatility	-2.9138	1.961	-1.486	0.138

Notes: This table presents the Ordinary Least Squares (OLS) regression results for fail-to-deliver (FTD) quantities, which are log-transformed to improve model fit and normalize the distribution. The model examines the effects of high trading volume, high volatility, and their interaction on FTD levels. Significance levels are marked as ***p < 0.01, **p < 0.05, and *p < 0.1.

To formally assess the impact of high trading volume and volatility on FTD occurrences, we estimate an OLS regression model (Table 6). The regression results reinforce the earlier findings, indicating that the coefficient for high trading volume is statistically significant and substantial, confirming a strong positive relationship between trading activity and FTD levels. By contrast, the coefficients for high volatility and the interaction between volume and volatility are statistically insignificant. It suggests that volatility alone or in combination with high trading volume does not meaningfully influence FTD occurrences. These results establish that high-volume trading days lead to increased FTD activity, while high volatility in isolation does not.

Table 7: T-Test Results for FTDs and Bitcoin Market Metrics

Metric	p-value
P-value for Significant Price Movements	<0.01***
P-value for High vs Low Volatility	0.6638

Notes: This table presents the results of two t-tests examining the relationship between fail-to-deliver (FTD) quantities and key Bitcoin market metrics. The P-value for Significant Price Movements tests whether FTD levels differ significantly on days with substantial Bitcoin price movements (returns exceeding 2%), compared to days without such movements. The P-value for High vs Low Volatility tests whether FTD levels vary between periods of high and low Bitcoin market volatility, defined by the top and bottom 20% of volatility values. Significance levels are marked as ***p < 0.01, **p < 0.05, and *p < 0.1.

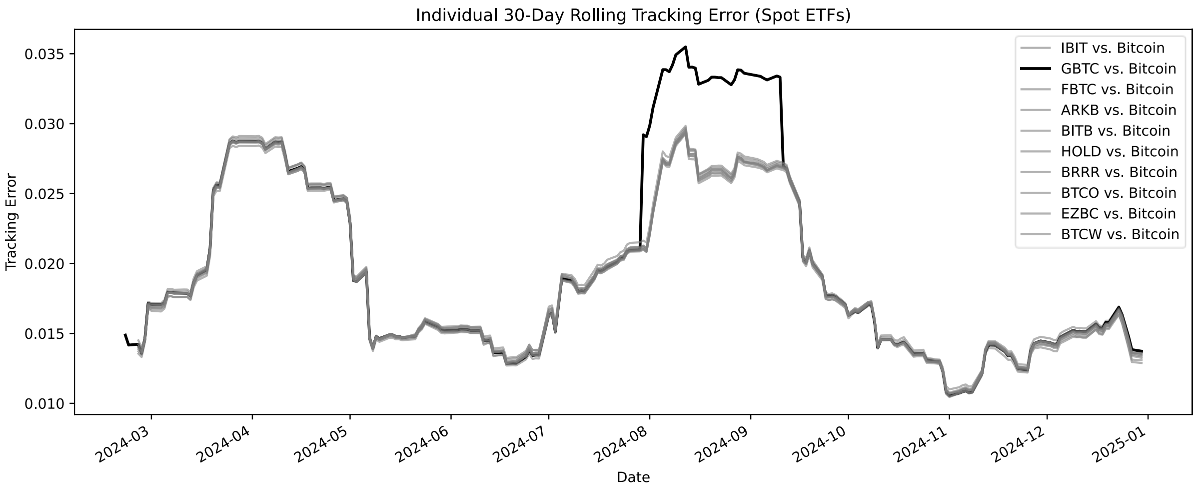
These results have significant implications for understanding the role of FTDs in Bitcoin ETF markets. The strong association between high trading volume and FTD occurrences highlights that periods of elevated market participation, possibly driven by speculative trading or shifts in sentiment, place strain on settlement systems, leading to an increase in FTD levels. Importantly, the lack of a significant relationship between FTDs and price volatility suggests that FTDs do not arise merely as a reaction to uncertainty or market instability but are more closely tied to transactional demand.

Moreover, the absence of FTD spikes during significant price movements—despite their strong link to trading volume—points to a potential stabilizing effect of FTD mechanisms. In markets without such mechanisms, one might expect sharp price movements to follow periods of intense trading due to liquidity imbalances. However, the findings suggest that FTDs help spread settlement obligations over time, mitigating the immediate impact of high trading activity on price dynamics. This indicates that FTDs may serve as a buffer, reducing the likelihood of extreme price fluctuations that might otherwise occur during periods of high volume.

2.3 Tracking error analysis

Consequently, we examine the tracking errors between Bitcoin spot ETFs and Bitcoin futures-based ETFs (Figures 3 and 4). The results reveal significant differences, with spot ETFs consistently exhibiting lower tracking errors (Figure 4). Spot ETFs, which replicate the underlying Bitcoin price through direct asset holdings, achieve average 30-day rolling tracking errors below 0.03. It indicates a high degree of alignment with Bitcoin's spot price. The finding is consistent with prior research on ETF tracking fidelity, which shows that direct asset-holding structures generally have reduced tracking errors, especially in highly volatile asset classes (Elton et al., 2002; Cremers et al., 2013). The low tracking error observed here suggests that Bitcoin spot ETFs are effective at minimizing tracking deviations despite the inherent volatility and liquidity constraints associated with Bitcoin.

Figure 3: Individual 30-Day Rolling Tracking Error of Bitcoin Spot ETFs Compared to Bitcoin Spot Price

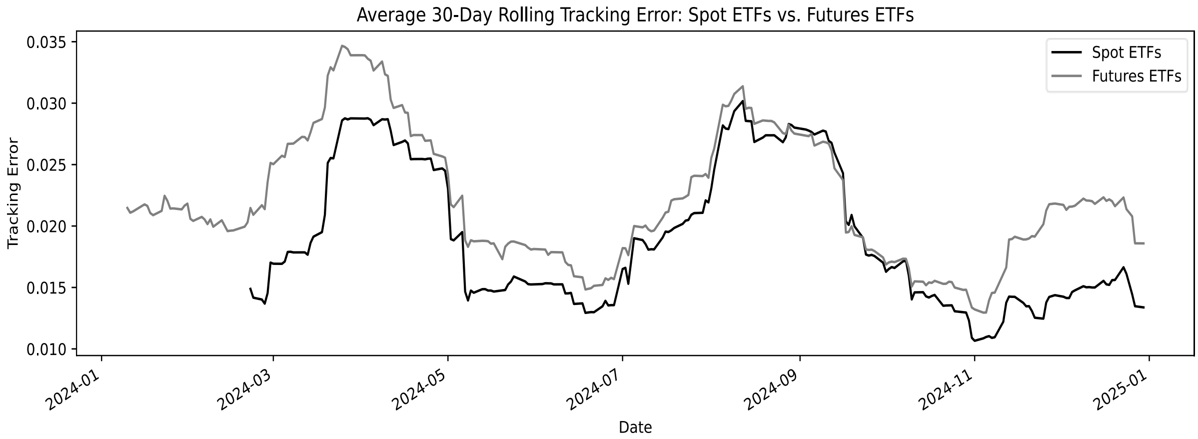


Note: The figure illustrates the 30-day rolling tracking error for individual Bitcoin spot ETFs relative to the Bitcoin spot price from January 2024 to December 2024. The tracking error, computed as the rolling standard deviation of the return differential between each ETF and Bitcoin, reflects deviations from the spot price, indicating potential inefficiencies in ETF replication of Bitcoin's price movements.

In contrast, futures-based ETFs show higher average tracking errors, frequently exceeding those of spot ETFs by 0.01 to 0.02, particularly during periods of elevated volatility. The average difference in tracking error (Futures ETFs - Spot ETFs) represents the value close to 0.004. Such divergence is likely attributable to structural complexities in futures-based ETFs, such as rolling costs, contango, and backwardation, which can impede accurate price replication as observed for different markets (Milonas and Henker, 2001). The findings align with broader evidence from commodities and other asset classes where futures-based ETFs often experience larger tracking discrepancies relative to spot ETFs, reflecting inefficiencies introduced by derivative exposures.

The results suggest that while both ETF types provide investors with Bitcoin exposure, spot ETFs offer a more precise representation of Bitcoin’s underlying price dynamics. The lower tracking error associated with spot ETFs underscores the advantage of direct asset holdings in achieving tracking accuracy.

Figure 4: Average 30-Day Rolling Tracking Error: Bitcoin Spot ETFs vs. Bitcoin Futures ETFs



Note: The figure compares the 30-day rolling average tracking error between Bitcoin spot ETFs and Bitcoin futures-based ETFs from January 2024 to December 2024. Tracking error is calculated as the rolling standard deviation of the return differential relative to the Bitcoin spot price, reflecting how closely each ETF type tracks Bitcoin's price movements.

Conclusions

Our study investigates the extent to which the introduction of Bitcoin ETFs has affected Bitcoin as an asset. We offer a thorough analysis of how Bitcoin spot ETFs influence Bitcoin's market dynamics, focusing on volatility, returns, return distribution, and tracking errors. Contrary to expectations based on traditional ETFs, the introduction of Bitcoin spot ETFs has been associated with increased volatility and heightened downside risk, as reflected in higher Value at Risk (VaR) and Conditional Value at Risk (CVaR) metrics. It suggests that, unlike in equity markets, where ETF adoption typically reduces volatility through enhanced liquidity and market depth, Bitcoin's market remains highly sensitive to speculative activity, even in the presence of institutional investments. As a result, the increase in volatility without a corresponding rise in returns suggests that the introduction of Bitcoin ETFs has deteriorated Bitcoin's long-term risk-return trade-off.

In terms of return distribution, our findings suggest a partial normalization of Bitcoin's statistical properties. The post-ETF period exhibits lower skewness and kurtosis, suggesting a reduction in extreme price fluctuations and tail risk. While Bitcoin continues to exhibit the characteristics of a speculative asset, the introduction of ETFs has contributed to a more structured trading environment. Our results indicate that fail-to-deliver (FTD) occurrences contributed in this normalization process. Unlike in traditional markets, where FTDs often signal liquidity constraints or settlement inefficiencies, their emergence in Bitcoin ETFs appears to function as a stabilizing mechanism. By deferring settlement over multiple days, spot FTDs prevent extreme price movements that might otherwise occur in a purely spot-driven market. It suggests that the presence of FTDs could mitigate the impact of large buy or sell orders, spreading settlement obligations over time and dampening immediate price distortions.

Moreover, the tracking error analysis highlights a key advantage of Bitcoin spot ETFs over their futures-based counterparts. Spot ETFs demonstrate superior tracking accuracy, maintaining a closer alignment with Bitcoin's price compared to futures ETFs, which suffer from inefficiencies related to roll costs, contango, and backwardation. The result aligns with prior research on ETF replication efficiency, confirming that direct asset holdings offer a more precise reflection of underlying price dynamics in highly volatile asset classes.

While these findings provide valuable insights into the role of Bitcoin spot ETFs in market evolution, this study has certain limitations. The relatively short post-ETF observation period may not fully capture long-term adjustments in trading behavior and market structure. Future research should extend the analysis to assess whether the observed volatility increase persists over time or diminishes as the market further integrates institutional participation. Additionally, incorporating microstructural data, such as order book dynamics and high-frequency trading behavior, could offer a deeper understanding of how Bitcoin ETFs influence market liquidity, price discovery, and arbitrage efficiency.

Overall, this study contributes to the growing literature on the financialization of cryptocurrencies by offering novel insights into the effects of Bitcoin ETFs. The findings underscore the unique challenges and opportunities introduced by ETFs in digital asset markets, emphasizing the need for continued research into their broader implications for price stability, market maturity, and investor behavior.

References

Aber, J. W., Li, D., & Can, L. (2009). Price volatility and tracking ability of ETFs. *Journal of Asset Management, 10*(4), 210–221.

Albrecht, P., Kočenda, E. (2024). Event-Driven Changes in Return Connectedness among Cryptocurrencies. SSRN 4908645. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4908645

Auer, R., & Claessens, S. (2018). Regulating cryptocurrencies: Assessing market reactions. *BIS Quarterly Review, September*, 51–65. https://www.bis.org/publ/qtrpdf/r_qt1809f.htm

Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking & Finance, 34*(8), 1886-1898.

Baur, D. G., Hong, K., & Lee, A. D. (2018). Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money, 54*, 177-189.

Ben-David, I., Franzoni, F. A., & Moussawi, R. (2017). Exchange-traded funds. *Annual Review of Financial Economics*, 9, 169–189.

Ben-David, I., Franzoni, F., & Moussawi, R. (2018). Do ETFs increase volatility? *The Journal of Finance, 73*(6), 2471-2535.

Bouri, E., Molnár, P., Azzi, G., Roubaud, D., & Hagfors, L. I. (2017). On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Finance Research Letters, 20*, 192-198.

Ciaian, P., Kancs, d'A., & Rajcaniova, M. (2018). The price of Bitcoin: GARCH evidence from high-frequency data. *Applied Economics*, 50(60), 6590–6610.

Chang, E. C., Sandnes, M., & Chiao, C. (2013). How closely do commodity ETF returns track the underlying commodity futures' returns? *Journal of Asset Management, 14*(1), 18-29.

Chang, E. C., Sandnes, A., & Chiao, C. (2013). The performance of commodity ETFs: tracking error and pricing deviation. Journal of Asset Management, 14(1), 47-59.

Corbet, S., Lucey, B., & Yarovaya, L. (2018). Datestamping the Bitcoin and Ethereum bubbles. *Finance Research Letters, 26*, 81-88.

Corbet, S., Lucey, B., Urquhart, A., & Yarovaya, L. (2019). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis, 62*, 182-199.

Cremers, M., Petajisto, A., & Zitzewitz, E. (2013). Should benchmark indices have alpha? Revisiting performance evaluation. Critical Finance Review, 2(1), 1-48.

Elton, E. J., Gruber, M. J., Comer, G., & Li, K. (2002). Spiders: Where are the bugs? Journal of Business, 75(3), 453-472.

Fotak, V., V. Roman, and P. K. Yadav. (2014). Fails-to-deliver, short selling, and market quality. Journal of Financial Economics, vol. 114, Issue 3, Pp. 493-516..

Glosten, L. R., Nallareddy, S., & Zou, Y. (2021). ETF Activity and Informational Efficiency of Underlying Securities. *Management Science, 67*(4), 2114–2136.

Israeli, D., Lee, C. M. C., & Sridharan, S. A. (2017). Is There a Dark Side to Exchange Traded Funds? An Information Perspective. *The Review of Accounting Studies, 22*(3), 1048–1083.

Katsiampa, P. (2017). Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics Letters, 158*, 3-6.

Laborda, J., Laborda, R., & de la Cruz, J. (2024). Can ETFs affect U.S. financial stability? A quantile cointegration analysis. *Financial Innovation, 10*(1), 64.

Madhavan, A. (2012). Exchange-traded funds, market structure, and the flash crash. *Financial Analysts Journal, 68*(4), 20-35.

Makarov, I., & Schoar, A. (2020). Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics, 135*(2), 293-319.

Pastorek, D., M. Drábek, P. Albrecht. (2023). Confirmation of T+35 Failures-To-Deliver Cycles: Evidence from GameStop Corp. Czech Journal of Economics and Finance, vol. 73, No. 1, pp. 56-80. Available at: https://doi.org/10.32065/CJEF.2023.01.03

Poterba, J. M., & Shoven, J. B. (2002). Exchange-traded funds: A new investment option for taxable investors. *American Economic Review, 92*(2), 422-427.

Shkilko, A., Van Ness, B., & Van Ness, R. (2012). Predatory trading. *The Journal of Finance, 67*(5), 2283-2312.

Todorov, K. (2021). *Launch of the first US bitcoin ETF: mechanics, impact, and risks*. BIS Quarterly Review. Retrieved from https://www.bis.org/publ/qtrpdf/r_qt2112t.htm

Todorov, K. (2024). When Passive Funds Affect Prices: Evidence from Volatility and Commodity ETFs. *Review of Finance, 28*(3), 831–863.

Yıldırım, H., Bekun, F.V. (2023). Predicting volatility of bitcoin returns with ARCH, GARCH and EGARCH models. Futur Bus J, vol. 9, pp. 75. https://doi.org/10.1186/s43093-023-00255-8