Perpetual Futures as Predictors of Bitcoin Volatility

Da-Hea Kim*

Abstract

This study examines whether information from Bitcoin perpetual futures improves the forecasting of Bitcoin spot volatility. Using high-frequency intraday data, we construct two types of predictors from perpetual futures: funding-rate-based measures and realized volatility measures. These predictors are incorporated into a Heterogeneous Autoregressive model to forecast daily, weekly, and monthly spot volatility. Both in-sample and out-of-sample results show significant improvements in volatility forecasts, particularly from funding rates and at longer forecast horizons. Forecast gains are concentrated in periods of high perpetual futures trading volume. These findings underscore the informational value of perpetual futures for volatility forecasting.

Keywords: Perpetual Futures; Funding Rates; Bitcoin Volatility; Realized Volatility; Volatility Forecasting

JEL Classification: G12, G13, G14, G17, C22

* Sungkyunkwan University Business School, 25-2 Sungkyunkwan-ro, Jongno-gu, Seoul (03063), Korea; Tel: +82-02-760-0889; Email: daheakim@skku.edu

1. Introduction

Volatility forecasting is central to financial decision-making and economic policy, as accurate predictions support effective risk management, portfolio optimization, and derivatives pricing, and enable regulators to monitor systemic risk and maintain financial stability (Poon and Granger, 2003). This role is especially critical in the Bitcoin market, where volatility often exceeds 80% on an annualized basis—substantially higher than that of major traditional asset classes such as equities (typically 15–20%) or government bonds (below 10%). Such elevated volatility underscores the need for robust risk management to protect investments. Moreover, as the Bitcoin options market remains in its early stages yet is expanding rapidly, accurate volatility forecasts are essential for fair pricing and efficient market functioning. Given Bitcoin's growing integration into the broader financial ecosystem,[2] monitoring its volatility is also crucial for safeguarding financial stability.

This study investigates whether information extracted from the Bitcoin perpetual futures market enhances the forecasting of Bitcoin spot market volatility. Perpetual futures—futures contracts that never expire—are the most actively traded derivatives in cryptocurrency markets.[3] Like traditional futures, perpetual futures offer leveraged exposure to price movements without requiring ownership of the underlying asset. However, unlike traditional futures, they allow for continuous exposure without the need for contract rollover.

We focus on the Bitcoin perpetual futures market for three main reasons. First, because perpetual futures aggregate investor expectations without requiring consensus on fundamental

[2] For example, an increasing number of public companies have begun allocating Bitcoin to their corporate treasuries as a strategic reserve asset, signaling growing institutional acceptance. See *Wall Street Journal*, "Businesses Are Bingeing on Crypto, Dialing Up the Market's Risks," April 28, 2025.

[3] In Q1 2025, the global crypto derivatives market recorded approximately $21 trillion in notional trading volume, with perpetual futures accounting for more than 90% of that activity (Source: tokeninsight.com).

value (Shiller, 1993), they are well-suited for price discovery of Bitcoin, which lacks income and valuation anchors. Second, the absence of contract expiration and rollover costs simplifies trading and concentrates liquidity in a single instrument, further enhancing price discovery.[4] Third, by enabling leveraged exposure without direct custody, perpetual futures mitigate custody, regulatory, and operational risks associated with Bitcoin ownership,[5] making them especially attractive to institutional and large-scale investors, who are more likely to be informed or influential. Collectively, these features suggest that the perpetual futures market may embed valuable information for forecasting spot market volatility.

However, early-stage markets often suffer from limited liquidity and a dominance of uninformed trading, resulting in inefficiencies and transitory noise (Shleifer and Summers, 1990; Makarov and Schoar, 2020). He et al. (2022) document frequent violations of theoretical pricing bounds in cryptocurrency perpetual futures, implying persistent arbitrage opportunities. While such mispricing has diminished over time, suggesting improving efficiency, the market may remain inefficient. Given the relative nascency of Bitcoin perpetual futures, their usefulness for volatility forecasting remains an open empirical question. These competing perspectives motivate our investigation into their informational value for predicting spot volatility.

We consider two sets of predictors derived from Bitcoin perpetual futures markets. The first is based on funding rates, an institutional feature unique to perpetual futures. Unlike

[4] Alexander et al. (2020) provide empirical evidence that perpetual futures are more liquid than traditional futures and lead spot prices.

[5] Bitcoin ownership entails custody risks—including private key loss, hacking, exchange failures, and physical threats. Recent "wrench attacks," where individuals with large cryptocurrency holdings are physically threatened to surrender private keys, highlight these risks (see *Wall Street Journal*, "Severed Fingers and 'Wrench Attacks' Rattle the Crypto Elite" May 17, 2025). In addition, institutions with fiduciary duties face regulatory and operational barriers to direct custody, while self-custody requires security infrastructure that many are unwilling or unable to implement (Fidelity Digital Assets, 2024).

traditional fixed-maturity futures, perpetual futures prices are not guaranteed to converge to the spot price due to the absence of a predetermined expiration date. To maintain alignment between perpetual futures and spot prices, a funding rate mechanism is employed: long position holders periodically pay short position holders a rate proportional to the price gap, thereby incentivizing trades that help close it (He et al., 2022). This mechanism suggests that funding rates serve as a powerful aggregator of market information, reflecting traders' willingness to pay, while also carrying implications for volatility dynamics: elevated or unstable funding rates indicate concentrated and leveraged positions that can trigger forced rebalancing, amplify order-flow shocks, and heighten fluctuations in the spot market.[6] We construct several funding-rate-based measures to capture both the level and the intraday variation of funding rates.

The second set of predictors consists of realized volatilities computed from intraday prices in the perpetual futures market. Unlike funding rates, which reflect traders' positioning incentives that drive price adjustments, realized volatility captures the ex-post market response to these adjustments. This provides complementary information by quantifying the actual magnitude of price fluctuation observed in the perpetual futures market. Moreover, divergences between spot and futures volatilities can be informative. When spot volatility exceeds futures volatility, it reflects stress or illiquidity in the spot market that is not immediately absorbed by derivatives trading. Such frictions in risk transmission across markets suggest that futures-based volatility measures contain predictive value for subsequent spot volatility. We construct daily, weekly, and monthly realized volatilities of perpetual futures to capture risk dynamics across different horizons.

We employ the Heterogenous Autoregressive (HAR) model of Corsi (2009), augmented with perpetual-futures-based predictors, to investigate whether these additional measures

[6] Similar mechanisms have been documented in other asset markets, where leverage constraints and order imbalances amplify volatility (e.g., Brunnermeier and Pedersen, 2009).

improve Bitcoin volatility forecasts. Both in-sample and out-of-sample analyses are conducted across multiple forecast horizons (daily, weekly, and monthly). The results indicate that signals from perpetual futures improve in-sample fit and, in out-of-sample tests, both Granger-cause spot volatility and enhance forecasting accuracy. Forecasting gains are more substantial at longer horizons, consistent with the idea that short-term volatility is obscured by temporary market frictions, while longer horizons allow the informational content of perpetual futures to manifest more clearly. Moreover, the gains are generally stronger for funding-rate-based predictors than for perpetual realized volatility measures, suggesting that funding rates embed more distinct information beyond lagged spot volatility, whereas the relatively limited independent information in perpetual volatility reflects the structural anchoring of perpetual futures prices to spot prices via the funding mechanism.

We further examine how the predictive value of perpetual-futures-based measures varies with trading activity. The results show that their contribution to volatility forecasting is primarily observed when perpetual futures trading volumes are relatively high, consistent with the view that more active markets incorporate information more efficiently.[7] Our findings are robust to a range of checks, including forecasting the level versus the log of volatility (linear versus log-linear specifications), employing alternative empirical proxies for return volatility, and using recursive versus rolling estimation schemes with different estimation window sizes.

This study contributes to two strands of literature. First, it extends the growing body of research on cryptocurrency perpetual futures markets. Prior studies have examined various aspects of these instruments, including their role in price discovery (Alexander et al., 2020), the effects of contract design and market microstructure on intraday pricing (De Blasis and Webb, 2022), their ability to enhance liquidity and reduce price dislocations under capital

[7] See, for example, Chordia, Roll, and Subrahmanyam (2008, 2011), Roll, Schwartz, and Subrahmanyam (2009), and Cao et al. (2024).

constraints (Gornall, Rinaldi, and Xiao, 2024), and the causal impact of their introduction or removal on spot market quality (Ruan and Streltsov, 2024). Other research has developed theoretical pricing models (Ackerer, Hugonnier, and Jermann, 2024), identified arbitrage opportunities via pricing bounds (He et al., 2022), and documented substantial returns from crypto carry trades involving shorting perpetual futures against spot holdings (Christin et al., 2022). While the existing literature focuses primarily on price levels and directional predictability (first moment), this paper explores a previously overlooked dimension: the informational content of perpetual futures for future volatility (second moment). By doing so, we provide a more comprehensive understanding of market efficiency.

Second, this study contributes to the literature on volatility forecasting, particularly in Bitcoin markets. Prior research has examined the predictive power of option-implied volatility (Hoang and Baur, 2020), macroeconomic and technical variables (Wang et al., 2022), sentiment signals from news media (Sapkota, 2022), GARCH and HAR models (Bergsli et al., 2022), and a range of statistical and machine learning methods (Dudek et al., 2024). However, the predictive value of perpetual futures, a dominant instrument in cryptocurrency trading, remains largely unexplored. This paper fills this gap by assessing whether signals from perpetual futures Granger-cause spot volatility and enhance the accuracy of realized volatility forecasts.

The remainder of the paper is organized as follows. Section 2 describes the data, the measurement of Bitcoin return volatility, the predictors used in the analysis, and the models employed for volatility forecasting. Section 3 investigates the predictive value of funding rate-based measures through in-sample analysis, out-of-sample forecasting, and robustness checks. Section 4 examines whether volatility in the perpetual futures market contains predictive information for spot volatility, following a parallel structure of in-sample, out-of-sample, and robustness analyses. Section 5 concludes.

2. Data and Methodology

2.1. Data

We collect intraday price and volume data for both Bitcoin spot and perpetual futures markets from Binance, one of the largest and most liquid cryptocurrency exchanges, using the official Binance API. Specifically, we use the BTC-USDT trading pair, which is the most actively traded perpetual futures contract on the platform, along with its corresponding spot market series. Using BTC-USDT for both markets ensures consistency, as this pair is by far the most liquid globally and is widely regarded as the effective spot benchmark in cryptocurrency trading.[8] The data are sampled at 5-minute intervals in Coordinated Universal Time (UTC), starting at 00:00:00 on January 1, 2020, and ending at 23:55:00 on April 30, 2025.[9] Prices are based on last transaction prices, and all series are synchronized to a common 5-minute grid. We also collect trading volumes for both spot and perpetual markets at the same frequency, which we later use in robustness checks to account for liquidity conditions.

2.2. Measurement of Bitcoin Volatility

Following McAleer and Medeiros (2008) and Liu et al. (2015), we measure realized volatility (RV) as the square root of the sum of squared 5-minute log returns, which serves as a proxy for daily volatility:

$$RV_t = \sqrt{\sum_{i=1}^{N} r_{t-1+i/N}^2}, \tag{1}$$

[8] While USDT may deviate slightly from USD, such deviations are typically small and transitory, and have negligible impact on high-frequency realized volatility measures.

[9] While BTC-USDT perpetual futures on Binance were officially launched on September 9, 2019, historical data for the contract is only available staring on January 1, 2020, which determines the beginning of our sample period.

where N denotes the number of intraday intervals in a day, and $r_{t-1+i/N}$ is the log return over the ith interval.

To address microstructure noise inherent in high-frequency data, we also compute the realized kernel (RK) estimator, following Barndorff-Nielsen et al. (2008, 2009):

$$RK_t = \sqrt{\sum_{h=-H}^{H} k\left(\frac{h}{H+1}\right)\gamma_h}, \qquad where \;\; \gamma_h = \sum_{i=|h|+1}^{N} r_{t-1+i/N} r_{t-1+(i-|h|)/N}. \tag{2}$$

We use the Parzen kernel for the weight function $k(\cdot)$ and set the bandwidth parameter H to 5.[10]

[Insert Table 1 around here]

Panel A of Table 1 reports descriptive statistics for the RV and RK measures, computed separately for the spot and perpetual futures markets. We also report their squared values, as some studies defines "realized volatility" in terms of realized variance.[11] The distributional characteristics of these measures are broadly similar across the two markets. Both RV and RK exhibit substantial positive skewness and excess kurtosis, indicating departures from normality, as confirmed by the Jarque-Bera test. To mitigate the impact of non-normality and fat tails, we also consider logarithmic transformations of these measures. The Ljung-Box statistics indicate significant autocorrelation, consistent with the well-documented phenomenon of volatility clustering (Andersen et al., 2001; 2003). Figure 1 plots the time series of RV and RK for the spot and perpetual futures markets, highlighting the persistent nature of Bitcoin volatility.

[10] Results are similar when using H=10; not reported for brevity.

[11] See, for example, Andersen et al. (2003), Andersen, Bollerslev, and Diebold (2007), Busch et al. (2011), and Bonato et al. (2023).

[Insert Figure 1 around here]

2.3. Funding-Rate-Based Variables

We collect historical funding rate data for perpetual futures contracts from the Binance website.[12] On Binance, funding rate payments are determined and published by the exchange every eight hours, yielding three observations per day.[13] Although this frequency is limited,[14] these realized funding rates represent the actual cost of holding long versus short positions in the perpetual futures market and therefore convey economically meaningful signals about positioning imbalances and market sentiment. To explore the informational content of funding rates for volatility forecasting, we construct several explanatory variables that capture both their level and intraday variation:

- AvgFR: The daily average of the three funding rates, capturing the overall directional pressure from net long or short positions.
- AbsFR: The absolute value of the daily mean funding rate, reflecting the strength of market imbalance regardless of direction.
- StdFR: The standard deviation of the three intraday funding rates, measuring variation in funding costs throughout the day.
- TrendFR: A monotonicity indicator that equals +1 if funding rates increase across the three intervals, −1 if they decrease, and 0 otherwise, summarizing whether intraday funding rates follow a consistent directional pattern.

[12] See Binance: https://www.binance.com/en/futures/funding-history/perpetual/funding-fee-history.

[13] Funding rates on Binance are charged at 00:00, 8:00, and 16:00 Universal Time Coordinated (UTC). Only traders holding positions at these times are subject to the funding payments, which reflect actual, realized costs. Although Binance provides predicted funding rates in real time, historical data on these forecasts is not publicly available, limiting their use in empirical analysis.

[14] We acknowledge that with only three observations per day, their aggregates are not statistically robust in the conventional sense. Yet because funding rates are released at this frequency by design, our construction should be viewed as a pragmatic way to extract economic signals from the realized funding-rate data given the market's data structure.

Panel B of Table 1 presents summary statistics for these funding-rate-based variables. Since funding rates are typically expressed in very small decimals, often below 0.001%, all variables are multiplied by 10,000 for ease of reference. The mean of AvgFR is positive (1.254), indicating that, on average, long position holders pay funding fees to short position holders, reflecting persistent net demand for long positions over the sample period. This pattern aligns with the existence of a structural premium for leveraged long exposure, as documented by Christin et al. (2022), who attribute this premium to persistent investor demand for upside exposure in cryptocurrencies. AbsFR has a mean of 1.406, suggesting sizable directional imbalances in positions. StdFR averages 0.532, indicating moderate intraday variation in funding conditions. TrendFR frequently assumes a value of zero, indicating that flat or non-monotonic funding rate sequences are more common than strictly increasing or decreasing intraday trends. Figure 2 plots the time series of these funding-rate-based measures.

[Insert Figure 2 around here]

2.4. Volatility Forecasting Models: HAR and Extensions

To model and forecast volatility, we employ the Heterogeneous Autoregressive (HAR) model proposed by Corsi (2009), which is widely used in the volatility literature. Although it is not a formal long-memory model, it effectively captures the long-memory-like behavior commonly observed in realized volatility. We choose the HAR model for its empirical robustness and its flexibility in incorporating additional predictors. Empirical evidence supports this choice: Bergsli et al. (2022) find that HAR models outperform GARCH-type models in forecasting Bitcoin volatility, while Dudek et al. (2024) show that simple linear models such as HAR perform comparably to more complex machine learning approaches.

As a benchmark, we implement the standard HAR model in the following form:

$$RV_{t+1} = \alpha + \beta_d RV_{t,d} + \beta_w RV_{t,w} + \beta_m RV_{t,m} + \varepsilon_{t+1}, \quad (3)$$

where $RV_{t,d} = RV_t$, $RV_{t,w} = \sqrt{\frac{1}{7}\sum_{i=0}^{6} RV_{t-i}^2}$, and $RV_{t,m} = \sqrt{\frac{1}{30}\sum_{i=0}^{29} RV_{t-i}^2}$ represent the lagged daily, weekly, and monthly realized volatilities from the spot market, respectively.

We then extend the model by including additional predictors in a vector $\mathbf{X}_t$:

$$RV_{t+1} = \alpha + \beta_d RV_{t,d} + \beta_w RV_{t,w} + \beta_m RV_{t,m} + \gamma' \mathbf{X}_t + \varepsilon_{t+1}, \quad (4)$$

where $\mathbf{X}_t$ includes one or more of the following: funding-rate-based variables and realized volatility measures from the perpetual futures market, either individually or in combination. This allows us to assess the incremental predictive power of perpetual futures signals beyond the information already embedded in the lagged spot market volatility.

To evaluate forecast performance across different horizons, we also estimate the model using h-day aggregated volatilities as the dependent variable, defined as the square root of the average realized variance over the next h days:

$$RV_{t,t+h} = \sqrt{\frac{1}{h}\sum_{j=1}^{h} RV_{t+j}^2}, \quad (5)$$

where $h = 1,\ 7,\ 30$ correspond to daily, weekly, and monthly horizons, respectively. This construction provides a volatility measure that is comparable across horizons by capturing the average level of realized variance over the forecast window $[t+1, t+h]$, thereby enabling us to investigate whether the predictive content of perpetual market variables differs across short-, medium-, and long-term horizons.

While the above specifications are based on realized volatility levels, we also consider log-transformed versions to account for the heavy-tailed nature of realized volatility and

mitigate the impact of extreme observations. Specifically, the log-transformed specification is given by

$$\ln(RV_{t,t+h}) = \alpha + \beta_d \ln(RV_{t,d}) + \beta_w \ln(RV_{t,w}) + \beta_m \ln(RV_{t,m}) + \gamma' \mathbf{X}_t + \varepsilon_{t+1}, \quad (6)$$

where $\ln(RV_{t,t+h})$ denotes the log of the realized volatility over forecast horizon h, and $\ln(RV_{t,d})$, $\ln(RV_{t,w})$, and $\ln(RV_{t,m})$ are the logs of daily, weekly, and monthly realized volatilities, respectively.

3. Forecasting Bitcoin Spot Volatility with Funding-Rate-Based Predictors

3.1. In-Sample Analysis

We begin by assessing whether funding-rate-based measures provide in-sample predictive power for Bitcoin spot volatility. Table 2 presents estimation results from linear regressions where funding rate variables are added to the benchmark HAR model, either individually or jointly:

$$\begin{aligned} RV_{t,t+h} = \alpha + \beta_d RV_{t,d} + \beta_w RV_{t,w} + \beta_m RV_{t,m} + \gamma_{AvgFR} AvgFR_t + \gamma_{AbsFR} AbsFR_t \\ + \gamma_{StdFR} StdFR_t + \gamma_{TrendFR} TrendFR_t + \varepsilon_{t+1}, \end{aligned} \quad (7)$$

[Insert Table 2 around here]

Panel A (daily), Panel B (weekly), and Panel C (monthly) present the results across three forecast horizons. The benchmark HAR model (column (1)) delivers strong explanatory power at all horizons, with all lagged realized volatility terms entering positively and significantly. As expected, the importance of short-horizon volatility declines with longer forecast horizons: β_d falls from 0.427 (t-statistic=16.90) in Panel A to 0.134 (t-statistic=6.20)

in Panel C. Conversely, the weight on longer-horizon volatility increases, with β_m rising from 0.087 (t-statistic=2.54) in Panel A to 0.231 (t-statistic=7.86) in Panel C.

Adding funding-rate-based predictors improves model fit. The average funding rate (AvgFR) has consistently significant and positive coefficients across all horizons and specifications, suggesting that greater net long positioning, reflected in higher funding rates, is associated with increased future spot volatility, likely due to liquidation risk and position unwinding pressures. The coefficients on the absolute average (AbsFR) are positively significant without controls for AvgFR but become insignificant when AvgFR is included, suggesting that its predictive power is subsumed by AvgFR. Both variables capture position imbalances in the market; however, AvgFR not only reflects the magnitude of directional pressure (as AbsFR does) but also its direction. This indicates that AvgFR captures a more comprehensive measure of market positioning, rendering the incremental explanatory power of AbsFR negligible. The standard deviation (StdFR) shows some predictive power at longer horizons, but its significance and coefficient sign vary across specifications, indicating limited robustness. The directional trend measure (TrendFR) is mostly insignificant, suggesting little forecasting value.

The incremental explanatory power from funding-rate-based predictors is more pronounced at longer forecast horizons, as indicated by adjusted R^2 values. In Panel C (monthly horizon), the adjusted R^2 increases from 24.10% in the HAR-only model (column (1)) to 34.11% in the full model with all predictors (column (6)). In comparison, the daily horizon models show more modest gains, with increases of only 0.68 to 1.60 percentage points, while the weekly horizon models exhibit larger gains of up to 4.52 percentage points. This pattern likely reflects that shorter-term volatility is more contaminated by market microstructure frictions and transitory liquidity effects, whereas longer horizons allow the informational content of perpetual futures to materialize more clearly.

[Insert Table 3 around here]

Table 3 presents the results from a log-linear specification, where the dependent variable is the log of realized spot volatility and the lagged realized spot volatilities also enter in log form. The same set of funding-rate-based predictors—AvgFR, AbsFR, StdFR, and TrendFR—are added individually and jointly to the benchmark log-HAR model. The results are broadly consistent with those in Table 2. The coefficients on lagged realized spot volatilities remain positive and statistically significant across all horizons, with similar patterns in magnitude: the coefficient on lagged daily volatility (β_d) decreases with the forecast horizon, while that on lagged monthly volatility (β_m) increases. When funding-rate-based predictors are added individually, all except TrendFR display statistically significant incremental predictive power across horizons. In joint specifications, AvgFR and AbsFR retain their significance, while StdFR remains significant only at the monthly horizon. Adjusted R^2 improvements follow a similar pattern, with larger gains observed when either AvgFR or AbsFR is included and when forecasts are made at longer horizons. Overall, Table 3 confirms that funding-rate-based measures improve in-sample forecasting of Bitcoin spot volatility.

3.2. Out-of-Sample Forecasting Performance

The central question in this section is whether incorporating funding-rate-based predictors improves the out-of-sample performance of volatility forecasts relative to the benchmark HAR model. Following Paye (2012), we distinguish between two conceptually distinct notions of forecast improvement. The first adopts a structural perspective, examining the data-generating process: Do funding-rate-based variables Granger-cause spot volatility, such that spot volatility depends not only on its own lags but also on funding-rate-based

variables? The second takes a normative perspective, focusing on practical forecasting performance: Do the augmented models produce more accurate volatility forecasts than the benchmark?

While related, these two perspectives need not lead to the same empirical conclusion due to the classic bias-variance trade-off. Even if a funding-rate-based predictor is part of the true model and reduces conditional bias, including it may increase forecast variance through parameter estimation noise. As a result, a correctly specified model could underperform a simpler, mis-specified benchmark in terms of forecast error.

To evaluate both aspects of forecast performance, we adopt a two-pronged testing approach following Paye (2012). In Subsection 3.2.1, we test Granger causality using the Clark and West (CW, 2007) test, which accounts for overfitting in nested models and evaluates whether the inclusion of funding rate-based variables adds incremental information beyond past volatility. In Subsection 3.2.2, we assess forecast accuracy using the Giacomini and White (GW, 2006) test, which evaluates whether augmented models statistically outperform the benchmark in terms of predictive loss.

3.2.1. Testing for Granger Causality

We first test whether funding-rate-based predictors Granger-cause spot volatility by applying the CW test for nested models, following Paye (2012). The CW test evaluates whether the observed reduction in mean squared prediction error (MSPE) from an augmented model is sufficiently large to offset the penalty from estimating additional parameters. Unlike standard forecast comparison tests, which tend to favor simpler models in nested settings, the CW test introduces a correction term to account for this bias. Let $e_{t,0}\left(= y_t - \hat{y}_{t,0}\right)$ and $e_{t,1}\left(= y_t - \hat{y}_{t,1}\right)$ denote forecast errors from the benchmark and augmented models, respectively, where

y_t is the realized volatility at time t, and $\hat{y}_{t,0}$ and $\hat{y}_{t,1}$ are forecasts from the benchmark and augmented models. The CW test is based on the adjusted forecast error difference:

$$f_t = e_{t,0}^2 - e_{t,1}^2 + \left(\hat{y}_{t,0} - \hat{y}_{t,1}\right)^2 \tag{8}$$

The CW test statistic is then given by:

$$\text{CW} = \frac{\bar{f}}{\sqrt{\hat{\sigma}_f^2 / T}} \tag{9}$$

where $\bar{f}$ is the sample mean of f_t, T is the number of out-of-sample forecasts, and $\hat{\sigma}_f^2$ is the sample variance of f_t. A significantly positive CW statistic provides evidence of Granger causality, indicating that the funding-rate-based variable improves forecasts by offering incremental information not already captured by lagged volatility dynamics.

The last row of Table 4 reports CW statistics comparing the benchmark HAR model—based solely on lagged realized spot volatility—with four augmented HAR models: three that include individual funding-rate-based predictors (AvgFR, AbsFR, and StdFR), and a Kitchen Sink model that incorporates all three predictors.[15] Forecasting models are estimated recursively, starting with an initial window of 180 days and expanding daily as new observations become available. Panel A presents results from linear specifications. Across daily, weekly, and monthly horizons, the CW statistics are uniformly positive and highly significant at the 1% level. At the daily horizon, CW values range from 4.94 (StdFR) to 11.17 (AbsFR), while the Kitchen Sink model also yields a strong and significant statistic of 8.28. At the weekly horizon, the evidence is even stronger, with CW statistics exceeding 7.6 for all predictors and reaching 11.14 for AbsFR. At the monthly horizon, CW values remain consistently large,

[15] The directional measure (TrendFR) is excluded from the out-of-sample analysis, as it was found to be largely insignificant in the in-sample regressions (see Table 2, Column (5)).

ranging from 8.65 (Kitchen Sink) to 12.40 (StdFR), again confirming significance at the 1% level. These results demonstrate that each funding-rate-based variable contributes incremental predictive content beyond lagged realized spot volatility, and that the joint inclusion of all three predictors continues to deliver significant gains. Panel B reports results from log-linear specification and shows a similar pattern of statistically significant positive CW statistics across all forecast horizons. Taken together, these findings indicate that perpetual funding-rate-based variables Granger-cause spot volatility.

[Insert Table 4 around here]

3.2.2. Testing for Superior Predictive Ability

While the CW test results in Section 3.2.1 provide evidence that funding rate predictors contain unique information about future volatility, they do not necessarily imply that the augmented models with funding rate measures improve forecasting performance in a normative sense. To evaluate whether funding-rate-based predictors improve forecast accuracy, we use the GW test of equal predictive ability. This test assesses whether the average difference in forecast losses between the benchmark HAR model and the augmented model is statistically significant. The null hypothesis states that both models have equal predictive accuracy. The GW test statistic is given by:[16]

$$\text{GW} = \frac{\sqrt{T}\bar{d}_T}{\hat{\sigma}_T}, \text{ where } d_t = L(y_t, \hat{y}_{t,0}) - L(y_t, \hat{y}_{t,1}), \quad \bar{d}_T = \sum_{t=1}^{T} d_t / T \tag{10}$$

[16] The GW test statistic is equivalent to the Diebold and Mariano (2002) test statistic.

Here, y_t is the realized volatility at time t, $\hat{y}_{t,0}$ and $\hat{y}_{t,1}$ are forecasts from the benchmark and augmented models, $L(\cdot)$ is a loss function, and $\hat{\sigma}_T$ is a heteroscedasticity- and autocorrelation-consistent (HAC) estimator of the asymptotic standard deviation of $\sqrt{T}\bar{d}_T$.

Following Patton (2011), we consider three loss functions for evaluating volatility forecasts:

$$\text{Mean Squared Error (MSE): } L(RV_{t,t+h}, \widehat{RV}_{t,t+h}) = \left(RV_{t,t+h}^2 - \widehat{RV}_{t,t+h}^2\right)^2, \tag{11}$$

$$\text{Mean Absolute Percentage Error (MAPE): } L(RV_{t,t+h}, \widehat{RV}_{t,t+h}) = \left|\frac{RV_{t,t+h}^2 - \widehat{RV}_{t,t+h}^2}{RV_{t,t+h}^2}\right|, \tag{12}$$

$$\text{Quasi-Likelihood (QLIKE): } L\left(RV_{t,t+h}, \widehat{RV}_{t,t+h}\right) = \log\left(RV_{t,t+h}^2\right) + \frac{\widehat{RV}_{t,t+h}^2}{RV_{t,t+h}^2}, \tag{13}$$

where $RV_{t,t+h}$ denotes the ex-post realized volatility and $\widehat{RV}_{t,t+h}$ is the corresponding forecast.

Table 4 reports average out-of-sample forecast losses for five models: the HAR benchmark, three HAR models augmented with individual funding-rate-based predictors (AvgFR, AbsFR, and StdFR), and a Kitchen Sink model that includes all three predictors. The table also presents GW test statistics comparing each augmented model to the benchmark. Panel A presents results based on linear specifications. Across all forecast horizons, both AvgFR and AbsFR consistently improve forecast accuracy relative to the HAR benchmark. The GW statistics for these two predictors are uniformly positive and statistically significant across all three loss functions, suggesting robust forecast gains from the augmented models. For example, under the monthly forecast horizon, AvgFR yields GW statistics of 13.78 (MSE), 14.63 (MAPE), and 14.87 (QLIKE), all significant at the 1% level.

In contrast, StdFR exhibits more mixed results. While it improves predictive accuracy at weekly and monthly horizons, it performs poorly at the daily horizon: the GW statistics for

StdFR are significantly negative at the daily horizon. This pattern mirrors its weak in-sample performance, where StdFR fails to achieve statistical significance and even reduces the model's adjusted R^2 relative to the HAR benchmark (see Table 2, Panel A, Column (4)). The combination of significantly positive CW statistics and insignificant GW statistics for StdFR indicates that, although StdFR contains unique information about spot volatility, it does not translate into forecast accuracy improvements, likely due to large estimation noise. These findings suggest that including funding rate dispersion may increase estimation noise when forecasting short-term volatility, outweighing its potential informational benefits.

The Kitchen Sink model generally outperforms the HAR benchmark across all horizons, with the larger gains observed at longer horizons. However, it fails to outperform models that include AvgFR or AbsFR individually, often producing higher average losses and lower GW statistics. Although its improvements over the benchmark remain statistically significant, these patterns suggest that StdFR contributes little incremental value and that AvgFR and AbsFR are highly correlated, making models with AvgFR or AbsFR alone more effective.

A similar pattern emerges in Panel B, which presents results from the log-linear specifications. AvgFR and AbsFR continue to yield substantial forecast improvements across all horizons, with significantly positive GW statistics under all loss functions. Although StdFR performs better than in the linear case, its improvements remain modest at short horizons relative to models with AvgFR or AbsFR. Overall, the log-linear results reinforce the conclusion from the linear specification: funding-rate-based predictors, particularly AvgFR and AbsFR, consistently enhance out-of-sample forecast accuracy.

In sum, the findings in Table 4 suggest that funding-rate-based variables provide incremental predictive information beyond that captured by lagged realized volatility, and that this information translates into consistent gains in forecast accuracy rather than spurious improvements from added model complexity.

3.2.3. Variation with Perpetual Trading Volume

This subsection investigates whether the predictive power of funding-rate-based variables depends on trading activity in the perpetual futures market. Since funding rates are generated within this market, their informativeness should be greater when trading is more active and liquid. To test this, we replicate the out-of-sample analysis in Table 4 across periods of high and low perpetual trading volume. We classify high versus low volume periods based on the median of the 30-day moving average of the ratio of perpetuals to spot trading volume.[17]

[Insert Table 5 around here]

Table 5 presents the results separately for high- and low-volume periods. As expected, forecast improvements from funding-rate-based predictors are more pronounced during high-volume periods. In Panel A (linear specification), GW statistics are consistently positive and statistically significant in the high-volume subsample across most models and forecast horizons.[18] In contrast, the low-volume subsample exhibits weaker GW statistics, with more frequent insignificance and lower magnitudes when positive. CW statistics are significantly positive across all models and both subsamples but are markedly larger during high-volume periods, often nearly double, indicating stronger Granger causality when market activity is elevated.

Panel B (log-linear specification) shows similar patterns. The augmented models with funding-rate-based measures again perform notably better in high-volume periods, and CW

[17] The 30-day average aligns with the longest volatility lag in the HAR specification. Results are robust to alternative subsample definitions using daily volume (not tabulated).

[18] The one exception is the model using StdFR alone at the daily horizon, which yields a significantly negative GW statistic, echoing earlier in-sample and out-of-sample results on its limited short-term predictive power.

statistics support greater informational value from funding rates under these circumstances. Together, these results highlight that funding rates contribute more meaningfully to volatility forecasting when derived from a more active perpetual futures market.

3.3. Robustness Checks

To evaluate the robustness of our forecasting results, we conduct two complementary exercises. First, we examine the sensitivity of results to the choice of volatility proxy by using the realized kernel instead of realized volatility. Second, we consider alternative model estimation methods, including longer recursive windows and rolling estimations. This section reports results from the linear specification only, given the minimal differences from log-linear models, and focuses on subsamples split by high- and low-volume periods rather than the full sample.[19]

3.3.1. Realized Kernel as Alternative Volatility Proxy

We first assess whether our findings are sensitive to the choice of volatility proxy by replicating the out-of-sample forecasting analysis using the realized kernel (RK), as defined in Eq. (2) and proposed by Barndorff-Nielsen et al. (2008). The RK estimator is designed to address the limitations of realized volatility in high-frequency settings, where microstructure noise can introduce significant bias. Although RV converges to integrated volatility under certain idealized conditions, those assumptions are typically violated in practice. In particular, the presence of microstructure noise, stemming from bid-ask bounce, discrete pricing, and other market frictions, induces autocorrelations in returns, undermining the consistency of RV. The RK estimator mitigates this bias by accounting for such noise, and is widely considered a more robust volatility measure in high-frequency data environments.

[19] Results using log-linear specifications and the full sample are available upon request.

[Insert Table 6 around here]

Table 6 presents the out-of-sample results using RK as a volatility proxy. The overall patterns closely resemble those reported in Table 5 based on realized volatility. Funding-rate-based predictors, such as AvgFR and AbsFR, continue to enhance forecast accuracy relative to the HAR benchmark, with stronger effects during high-volume periods. GW statistics for models augmented with these predictors remain highly significantly positive in high-volume periods but tend to weaken in low-volume periods. CW statistics are significantly positive across all specifications, with larger magnitudes in high-volume periods. Thus, the predictive value of funding-rate-based variables persists when RK is used as the volatility proxy, reinforcing the robustness of our results to alternative volatility measures.

3.3.2. Robustness to Estimation Methods

We assess the robustness of our results to alternative model estimation strategies. While the baseline forecasts are generated using recursive estimation with a 180-day initial window, Table 7 considers two alternatives: recursive estimation with a longer 365-day initial window (Panel A) and rolling estimation with a fixed 180-day window (Panel B).

[Insert Table 7 around here]

Panel A confirms the robustness of our main findings to a longer estimation window. Funding-rate-based predictors, except for StdFR at the daily forecast horizon, improve forecast accuracy relative to the HAR benchmark, with stronger forecast gains in high-volume periods. Both GW and CW statistics remain significantly positive, with larger magnitudes in high-volume periods, closely matching the baseline results in Table 5, Panel A.

Panel B shows that the forecasting gains from funding-rate-based variables are somewhat attenuated under rolling estimation. This is likely due to its greater sensitivity to short-term noise and reduced ability to capture longer-term dynamics. Nonetheless, a clear contrast emerges across volume subsamples. During high-volume periods, augmented models significantly improve forecasts (except at the daily horizon), as reflected in strongly positive and statistically significant GW statistics. In contrast, GW statistics turn significantly negative during low-volume periods, indicating that the simple HAR benchmark outperforms the augmented models. CW statistics remain significantly positive in both subsamples, suggesting that funding-rate-based variables contain unique information about spot volatility but fail to improve forecast accuracy under low-volume conditions due to large estimation noise.

4. Forecasting Bitcoin Spot Volatility with Perpetual Futures Volatility

This section examines whether realized volatilities computed from intraday perpetual futures prices improve forecasts of Bitcoin spot volatility. While perpetual futures prices are structurally linked to funding rates, their predictive value for spot volatility may differ. Funding rates function as an adjustment tool to promote convergence between perpetual and spot prices by altering traders' incentives and correcting imbalances in long and short positions. In contrast, realized volatility measures based on price dynamics capture the ex-post market response to these adjustments. By incorporating these high-frequency price-based volatility signals, we assess whether intraday perpetual price dynamics provide incremental predictive information about the evolution of spot volatility.

We follow the same structure as in Section 3 but replace funding-rate-based predictors with perpetual volatility measures. Since the methodology has already been described in detail in the previous section, we focus here on interpreting the empirical results.

4.1. In-Sample Analysis

Table 8 reports the results from in-sample predictive regressions of Bitcoin spot volatility, incorporating volatility differences between the spot and perpetual futures markets as additional predictors. Rather than including perpetual market volatility levels directly, we use their differences from corresponding spot volatilities to mitigate concerns over multicollinearity in coefficient interpretation. Given that the HAR benchmark model already includes lagged spot volatilities, which are highly correlated with perpetual volatility, using the difference helps isolate the incremental predictive content of perpetual volatility. Notably, this modeling choice has no effect on out-of-sample forecasts, as both specifications yield identical predictions.

[Insert Table 8 around here]

Each panel of Table 8 is organized by forecast horizon: columns (1)–(5) report results for daily forecasts, (6)–(10) for weekly forecasts, and (11)–(15) for monthly forecasts. For each horizon, we report estimates from the HAR benchmark, models augmented with a single volatility difference (daily, weekly, or monthly), and a Kitchen Sink model that includes all three differences.

Panel A reports the results from the linear specification. The coefficients on the difference between spot and perpetual volatilities are generally positive and statistically significant, particularly when each is included individually. A positive coefficient indicates that when spot volatility exceeds perpetual volatility, interpreted as excess realized risk in the spot market, future spot volatility tends to remain high. This pattern is most pronounced for the weekly and monthly differences: δ_w and δ_m are consistently positive across all forecast horizons, while δ_d is negative and insignificant at the daily horizon. These results suggest that a high spot–perpetual volatility gap observed over a longer past period reflects persistent

market stress and supply-demand imbalances, signaling prolonged future risk in the spot market.

In the Kitchen Sink model, the monthly volatility gap remains a robust predictor even after controlling for the others, suggesting that longer-horizon volatility gaps contain more robust information about future risk. This is likely because long-term volatility captures persistent market dynamics and structural risk, whereas short-term volatility is more susceptible to transitory noise and temporary market fluctuations. Regarding model fit, all augmented models yield higher adjusted R^2 values than the HAR benchmark, except for the specification including only the daily difference (column 2). As expected, the largest gains are observed in models with the monthly difference (columns 4, 9, and 14), supporting the view that persistent volatility gaps between spot and perpetual markets are most informative about continued market turbulence.

Panel B presents results from the log-linear specification and largely mirrors the patterns in Panel A. The sign and significance of the δ coefficients remain consistent, and improvements in adjusted R^2 follow similar patterns. Overall, the explanatory power of perpetual volatility measures is weaker than that of the funding-rate-based predictors in Section 3 (Table 2). This can be attributed to the structural feature of perpetual futures prices, which are anchored to spot prices through the funding rate mechanism. As a result, perpetual volatility contains less independent information beyond lagged spot volatility compared to funding rates. Nevertheless, perpetual volatility provides consistent predictive signals for future spot volatility regardless of model specifications, indicating its incremental informational value beyond lagged spot volatility.

4.2. Out-of-Sample Forecasting Performance

Table 9 reports the out-of-sample performance of perpetual volatility measures in predicting Bitcoin spot volatility. Forecasting models include the HAR benchmark, models augmented with a single perpetual volatility measure (daily, weekly, or monthly), and a Kitchen Sink model including all three. All models are estimated recursively with an initial 180-day window.

[Insert Table 9 around here]

Panel A presents results from the linear specification. CW test statistics are significantly positive across all specifications, indicating robust Granger causality from perpetual to spot volatility. Adding perpetual volatility generally improves forecast accuracy relative to the HAR benchmark, with the largest gains observed in models including monthly volatility ($RV_{t,m}^{perp}$). GW statistics are significantly positive across all loss functions for these models. The main exception is the model with daily volatility ($RV_{t,d}^{perp}$), which yields significantly negative GW statistics at the daily forecast horizon, echoing its weak in-sample performance.

Panel B presents results from the log-linear specification, which closely mirrors the findings in Panel A. Both GW and CW statistics are significantly positive, with larger magnitudes for longer-term perpetual volatilities at longer forecast horizons. This indicates greater reductions in forecast loss and stronger Granger causality from perpetuals to spot markets. These findings suggest that longer-term perpetual volatilities convey forward-looking information beyond what is captured by lagged spot volatility, particularly at longer horizons. Given the similarity between linear and log-linear results, subsequent analyses focus on the linear specification for brevity.[20]

[20] Results from the log-linear models are available upon request and are qualitatively similar.

4.2.1. Variation with Perpetual Trading Volume

Table 10 evaluates out-of-sample forecasting performance across high- and low-volume periods, defined by the median of the 30-day relative trading volume between perpetual and spot markets. Forecast improvements are more pronounced during high-volume periods, with GW statistics consistently positive and significant across most models and horizons.[21] In contrast, low-volume periods show mixed GW results with weaker significance, indicating diminished predictive power when the perpetual futures market is less active. CW statistics remain positive and significant in both subsamples but are notably stronger during high-volume periods. Overall, the findings confirm the enhanced informational role of perpetual volatility when trading activity is high.

[Insert Table 10 around here]

4.3. Robustness Checks

4.3.1. Realized Kernel as Alternative Volatility Proxy

Table 11 reports out-of-sample forecasting performance using RK as the volatility proxy. The results show stronger and more consistent forecast improvements from perpetual volatility when RK is employed. Compared to Table 10, GW statistics are more uniformly positive and significant during high-volume periods. Notably, even the model with daily perpetual volatility yields significantly positive GW statistics for daily forecasts in high-volume periods, contrasting with its negative or insignificant performance when using realized volatility. The divergence between high- and low-volume periods is more pronounced: high-volume periods show broadly positive and significant GW statistics across all models and horizons, whereas low-volume periods exhibit mixed signs and weaker significance. CW statistics also display

[21] The daily volatility model at the daily horizon is an exception.

larger positive values during high-volume periods, confirming stronger Granger causality from perpetual to spot volatility when trading activity is high. Overall, these results highlight the benefits of using RK, suggesting that mitigating microstructure noise enhances the predictive content of perpetual volatility.

[Insert Table 11 around here]

4.3.2. Robustness to Estimation Methods

Table 12 assesses the robustness of our findings to alternative estimation methods, focusing on the Kitchen Sink model relative to the HAR benchmark. Panel A reports results based on recursive estimation with a 365-day initial window. The results closely align with the main findings in Table 10 (180-day initial window), with GW and CW statistics remaining consistently positive and significant during high-volume periods, reaffirming the predictive power of perpetual market volatility when trading activity is high.

[Insert Table 12 around here]

Panel B uses rolling estimation with a fixed 180-day window. Compared to recursive estimation, rolling estimation generally yields weaker forecasting performance, reflected in smaller and less significant GW and CW statistics, likely due to its greater sensitivity to transitory noise. Nonetheless, the pattern of stronger forecast performance during high-volume periods persists. Overall, while the magnitude of improvement varies by estimation methods, these robustness checks confirm that perpetual market volatility enhances forecast accuracy under active market conditions.

5. Conclusion

Given the novelty of perpetual futures and their growing role in the cryptocurrency ecosystem, this paper investigates whether signals extracted from the Bitcoin perpetual futures market improve forecasts of Bitcoin spot volatility. We focus on two classes of predictors: funding-rate-based measures and realized volatility measures derived from perpetual futures prices.

Using an extended Heterogeneous Autoregressive model, we find that both types of predictors provide significant incremental information about future spot volatility beyond that contained in lagged spot volatility. They improve in-sample fit and, in out-of-sample tests, deliver robust evidence of Granger causality and enhanced forecasting accuracy, particularly at longer horizons. The predictive value of perpetual futures signals is most pronounced when trading volumes in the perpetual market are relatively high. These results remain robust to alternative model specifications, different volatility proxies, and both rolling and recursive estimation methods.

Our findings suggest that perpetual-futures-based signals are useful for forecasting Bitcoin volatility. These signals can aid investors in risk management, improve derivatives pricing, and assist regulators in monitoring market conditions. This study contributes to the literature by demonstrating that perpetual futures contain information about higher-order return moments, complementing prior research that has primarily focused on price discovery and return predictability. In doing so, it also highlights the role of perpetual futures markets in shaping spot market risk dynamics and enhancing market efficiency, thereby extending the understanding of volatility transmission and price discovery in cryptocurrency markets.

References

Ackerer, D., Hugonnier, J., & Jermann, U. (2024). Perpetual futures pricing (No. w32936). *National Bureau of Economic Research*.

Alexander, C., Choi, J., Park, H., & Sohn, S. (2020). BitMEX bitcoin derivatives: Price discovery, informational efficiency, and hedging effectiveness. *Journal of Futures Markets*, *40*(1), 23-43.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, *61*(1), 43-76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*(2), 579-625.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics*, *89*(4), 701-720.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2008). Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, *76*(6), 1481-1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2009). Realized kernels in practice: trades and quotes. *Econometrics Journal*, *12*(3), C1-C32.

Bergsli, L. Ø., Lind, A. F., Molnár, P., & Polasik, M. (2022). Forecasting volatility of Bitcoin. *Research in International Business and Finance*, *59*, 101540.

Bonato, M., Cepni, O., Gupta, R., & Pierdzioch, C. (2023). Climate risks and state-level stock market realized volatility. *Journal of Financial Markets*, *66*, 100854.

Brunnermeier, M. K., & Pedersen, L. H. (2009). Market liquidity and funding liquidity. *The Review of Financial Studies*, *22*(6), 2201-2238.

Busch, T., Christensen, B. J., & Nielsen, M. Ø. (2011). The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. *Journal of Econometrics*, *160*(1), 48-57.

Cao, J., Goyal, A., Ke, S., & Zhan, X. (2024). Options trading and stock price informativeness. *Journal of Financial and Quantitative Analysis*, *59*(4), 1516-1540.

Chordia, T., Roll, R., & Subrahmanyam, A. (2008). Liquidity and market efficiency. *Journal of Financial Economics*, *87*(2), 249-268.

Chordia, T., Roll, R., & Subrahmanyam, A. (2011). Recent trends in trading activity and market quality. *Journal of Financial Economics*, *101*(2), 243-263.

Christin, N., Routledge, B., Soska, K., & Zetlin-Jones, A. (2022). The crypto carry trade. *Preprint at http://gerbil. life/papers/CarryTrade. v1*, *2*.

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, *138*(1), 291-311.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, *7*(2), 174-196.

De Blasis, R., & Webb, A. (2022). Arbitrage, contract design, and market structure in Bitcoin futures markets. *Journal of Futures Markets*, *42*(3), 492-524.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *20*(1), 134-144.

Dudek, G., Fiszeder, P., Kobus, P., & Orzeszko, W. (2024). Forecasting cryptocurrencies volatility using statistical and machine learning methods: A comparative study. *Applied Soft Computing*, *151*, 111132.

Fidelity Digital Assets. (2024). Channels for exposure to Bitcoin revisited.

Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, *74*(6), 1545-1578.

Gornall, W., Rinaldi, M., & Xiao, Y. (2024). Perpetual Futures and Basis Risk: Evidence from Cryptocurrency SSRN: https://ssrn.com/abstract=5036933

He, S., Manela, A., Ross, O., & von Wachter, V. (2022). Fundamentals of perpetual futures. *arXiv preprint arXiv:2212.06888*.

Hoang, L. T., & Baur, D. G. (2020). Forecasting bitcoin volatility: Evidence from the options market. *Journal of Futures Markets*, *40*(10), 1584-1602.

Hung, J. C., Liu, H. C., & Yang, J. J. (2021). Trading activity and price discovery in Bitcoin futures markets. *Journal of Empirical Finance*, *62*, 107-120.

Liu, L. Y., Patton, A. J., & Sheppard, K. (2015). Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *Journal of Econometrics*, *187*(1), 293-311.

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

McAleer, M., & Medeiros, M. C. (2008). Realized volatility: A review. *Econometric Reviews*, *27*(1-3), 10-45.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, *160*(1), 246-256.

Paye, B. S. (2012). 'Déjà vol': Predictive regressions for aggregate stock market volatility using macroeconomic variables. *Journal of Financial Economics*, *106*(3), 527-546.

Poon, S. H., & Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, *41*(2), 478-539.

Roll, R., Schwartz, E., & Subrahmanyam, A. (2009). Options trading activity and firm valuation. *Journal of Financial Economics*, *94*(3), 345-360.

Ruan, Q., & Streltsov, A. (2024). Perpetual Futures Contracts and Cryptocurrency Market Quality. *Available at SSRN*.

Sapkota, N. (2022). News-based sentiment and bitcoin volatility. *International Review of Financial Analysis*, *82*, 102183.

Shiller, R. J. (1993). Measuring asset values for cash settlement in derivative markets: hedonic repeated measures indices and perpetual futures. *The Journal of Finance*, *48*(3), 911-931.

Shleifer, A., & Summers, L. H. (1990). The noise trader approach to finance. *Journal of Economic Perspectives*, *4*(2), 19-33.

Wang, J., Ma, F., Bouri, E., & Guo, Y. (2023). Which factors drive Bitcoin volatility: macroeconomic, technical, or both?. *Journal of Forecasting*, *42*(4), 970-988.

Figure 1. Time Series of RVs and RKs for Bitcoin Spot and Perpetual Futures

This figure presents the time series of realized volatilities (RV, top row), their logarithmic transformation (middle row), and realized kernels (RK, bottom row) for Bitcoin spot (left column) and Bitcoin perpetual futures (right column) traded on Binance from December 31, 2019 to April 30, 2025. All series are expressed in percentage points.

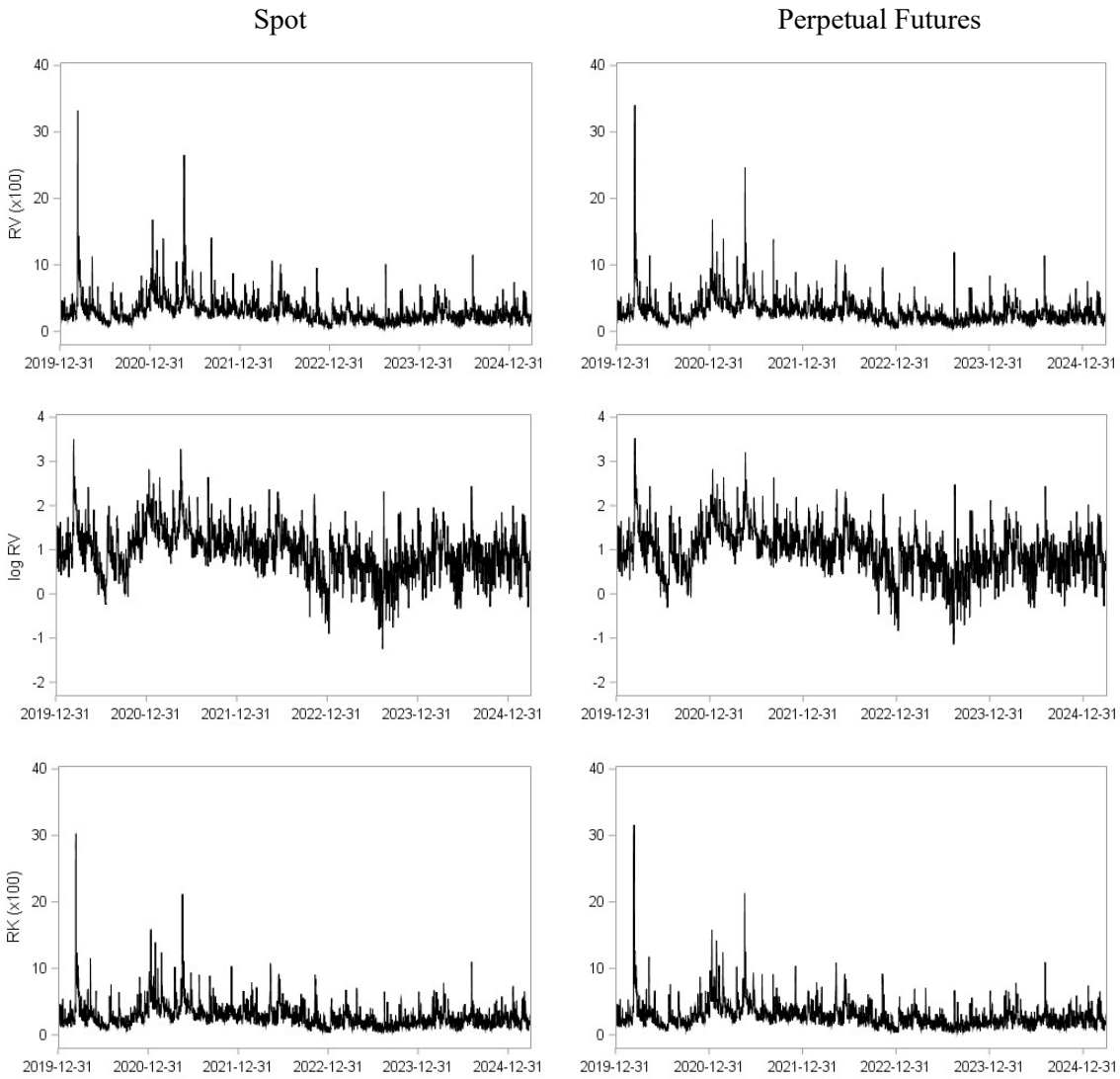

Figure 2. Time Series of Funding-Rate-Based Measures

This figure displays the time series of three funding-rate-based measures for Bitcoin perpetual futures contracts traded on Binance: the daily average of funding rates (*AvgFR*, top panel), the daily absolute average (*AbsFR*, middle panel), and the daily standard deviation (*StdFR*, bottom panel). The sample period spans from December 31, 2019 to April 30, 2025. All funding-rate-based measures are multiplied by 10,000 and expressed in percentage terms.

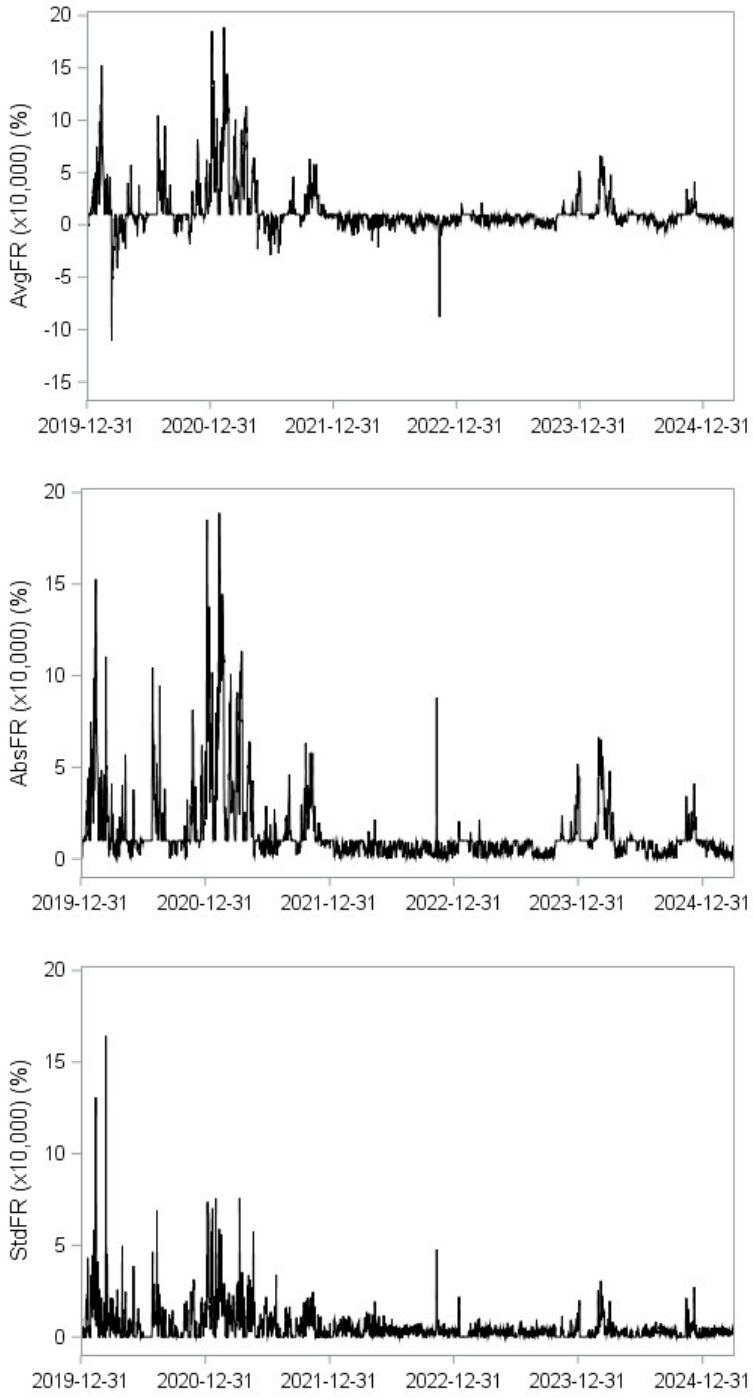

Table 1. Descriptive Statistics

Panel A reports summary statistics for realized volatilities (RVs), realized kernels (RKs), their squares, and log transformation for Bitcoin spot (Panel A.1) and BTCUSDT perpetual futures (Panel A.2) traded on Binance. RV and RK are in percentage points; their squares are in squared percentage points. The statistics include mean, standard deviation, skewness, kurtosis, minimum, and maximum, along with p-values from the Jarque-Bera test (normality) and the Ljung-Box test (autocorrelation, up to 10 lags, heteroscedasticity-adjusted). Panel B presents summary statistics for daily funding-rate based measures from the perpetual futures market, calculated from three intraday observations. *AvgFR* and *AbsFR* are the daily average and absolute average across the three funding rates, *StdFR* is their daily standard deviation, and *TrendFR* is a direction indicator: +1 if rates increase monotonically, -1 if they decrease, and 0 otherwise. *AvgFR*, *AbsFR*, and *TrendFR* are multiplied by 10,000 and expressed in percentage terms. The sample period spans January 1, 2020 to April 30, 2025.

Panel A: Volatility measures

	Mean	Std. Dev.	Skewness	Kurtosis	Min	Max	Jarque-Bera	Ljung-Box
				Panel A.1: Spot market				
RV^2	13.132	36.397	19.341	500.981	0.083	1,105.561	0.00	0.00
RV	3.023	1.999	4.520	44.658	0.288	33.250	0.00	0.00
Log RV	0.957	0.536	0.064	1.029	-1.245	3.504	0.00	0.00
RK^2	11.851	31.532	18.119	444.300	0.098	918.346	0.00	0.00
RK	2.853	1.926	4.077	37.405	0.313	30.304	0.00	0.00
Log RK	0.882	0.575	-0.099	0.765	-1.163	3.411	0.00	0.00
				Panel A.2: Perpetual futures market				
RV^2	13.208	36.753	19.922	541.275	0.101	1,157.907	0.00	0.00
RV	3.031	2.006	4.531	45.102	0.318	34.028	0.00	0.00
Log RV	0.960	0.534	0.096	0.989	-1.145	3.527	0.00	0.00
RK^2	11.912	33.085	19.194	493.587	0.100	998.246	0.00	0.00
RK	2.851	1.946	4.258	40.983	0.316	31.595	0.00	0.00
Log RK	0.880	0.577	-0.088	0.765	-1.151	3.453	0.00	0.00

Panel B: Funding-rate-based measures

Variable	Mean	Std. Dev.	Min	P25	P50	P75	Max
AvgFR (×10,000)	1.254	2.123	-11.041	0.360	0.954	1.000	18.890
AbsFR (×10,000)	1.406	2.026	0.000	0.455	0.991	1.000	18.890
StdFR (×10,000)	0.532	0.942	0.000	0.000	0.264	0.592	16.419
TrendFR	-0.012	0.491	-1.000	0.000	0.000	0.000	1.000

Table 2. In-Sample Forecasting with Funding-Rate-Based Measures: Linear Specification

The table presents results from in-sample predictive regressions for Bitcoin spot volatility using funding-rate based variables. The following regression model is estimated:

$$RV_{t,t+h} = \alpha + \beta_d RV_{t,d} + \beta_w RV_{t,w} + \beta_m RV_{t,m} + \gamma_{AvgFR} AvgFR_t + \gamma_{AbsFR} AbsFR_t + \gamma_{StdFR} StdFR_t + \gamma_{TrendFR} TrendFR_t + \varepsilon_{t+1},$$

where $RV_{t,t+h}$ is the realized volatility over forecast horizon h, and $RV_{t,d}$, $RV_{t,w}$, and $RV_{t,m}$ are daily, weekly, and monthly past realized volatilities, respectively. *AvgFR*, *AbsFR*, *StdFR*, and *TrendFR* are funding-rate-based measures that summarize the level, dispersion, and direction of intraday funding rates, as defined in Table 1. The table reports estimated coefficients and heteroscedasticity- and autocorrelation-robust t-statistics (in parentheses). Forecasting horizons correspond to h =1 (Panel A: daily), $h = 7$ (Panel B: weekly), and $h = 30$ (Panel C: monthly). The last two rows report the adjusted $R^2$and its increase relative to a benchmark HAR model (i.e., the same regression with $\gamma_k = 0$ for all k). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Daily forecasting horizon ($h = 1$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α	0.519***	0.456***	0.510***	0.520***	0.522***	0.394***	0.405***	0.462***
	(6.06)	(5.35)	(5.98)	(6.03)	(6.08)	(4.54)	(4.70)	(5.38)
β_d	0.427***	0.410***	0.405***	0.426***	0.426***	0.437***	0.436***	0.429***
	(16.90)	(16.38)	(15.93)	(16.18)	(16.84)	(16.78)	(16.77)	(16.46)
β_w	0.289***	0.302***	0.288***	0.289***	0.289***	0.311***	0.304***	0.285***
	(7.92)	(8.37)	(7.93)	(7.92)	(7.91)	(8.53)	(8.44)	(7.88)
β_m	0.087**	0.067**	0.074**	0.087**	0.087**	0.071**	0.072**	0.078**
	(2.54)	(1.98)	(2.15)	(2.53)	(2.55)	(2.09)	(2.12)	(2.29)
γ_{AvgFR}		0.109***				0.187***	0.139***	
		(6.84)				(4.50)	(7.73)	
γ_{AbsFR}			0.088***			-0.066		0.142***
			(5.02)			(-1.29)		(6.38)
γ_{StdFR}				0.002		-0.122**	-0.159***	-0.199***
				(0.04)		(-2.29)	(-3.55)	(-3.92)
$\gamma_{TrendFR}$					-0.050	-0.039		
					(-0.72)	(-0.57)		
$Adj.R^2$	45.47	46.74	46.15	45.44	45.45	47.06	47.06	46.55
$\Delta\, Adj.R^2$		1.27	0.68	-0.03	-0.01	1.60	1.60	1.08

Table 2 —*Continued*

Panel B: Weekly forecasting horizon ($h = 7$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α	1.080***	0.980***	1.061***	1.118***	1.078***	0.975***	0.969***	1.032***
	(13.73)	(12.85)	(13.92)	(14.21)	(13.69)	(12.51)	(12.53)	(13.42)
β_d	0.273***	0.247***	0.230***	0.241***	0.274***	0.252***	0.253***	0.245***
	(11.79)	(11.03)	(10.12)	(10.01)	(11.81)	(10.81)	(10.85)	(10.48)
β_w	0.222***	0.243***	0.219***	0.224***	0.222***	0.239***	0.243***	0.218***
	(6.62)	(7.52)	(6.76)	(6.72)	(6.63)	(7.30)	(7.53)	(6.72)
β_m	0.174***	0.142***	0.147***	0.163***	0.174***	0.144***	0.143***	0.150***
	(5.55)	(4.69)	(4.82)	(5.21)	(5.54)	(4.73)	(4.72)	(4.91)
γ_{AvgFR}		0.175***				0.147***	0.181***	
		(12.24)				(3.95)	(11.24)	
γ_{AbsFR}			0.177***			0.047		0.210***
			(11.29)			(1.02)		(10.54)
γ_{StdFR}				0.175***		-0.061	-0.035	-0.122***
				(4.77)		(-1.27)	(-0.87)	(-2.67)
$\gamma_{TrendFR}$					0.043	0.063		
					(0.68)	(1.02)		
$Adj.R^2$	37.44	41.95	41.32	38.14	37.42	41.95	41.95	41.51
$\Delta\, Adj.R^2$		4.52	3.88	0.70	-0.02	4.51	4.51	4.07

Panel C: Monthly forecasting horizon ($h = 30$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α	1.754***	1.629***	1.731***	1.822***	1.758***	1.652***	1.655***	1.732***
	(23.90)	(23.65)	(25.02)	(25.26)	(23.94)	(23.50)	(23.72)	(24.78)
β_d	0.134***	0.102***	0.080***	0.075***	0.133***	0.089***	0.089***	0.079***
	(6.20)	(5.03)	(3.89)	(3.40)	(6.13)	(4.20)	(4.22)	(3.74)
β_w	0.123***	0.150***	0.120***	0.127***	0.123***	0.152***	0.149***	0.120***
	(3.95)	(5.13)	(4.09)	(4.15)	(3.94)	(5.15)	(5.10)	(4.09)
β_m	0.231***	0.191***	0.196***	0.210***	0.231***	0.188***	0.188***	0.196***
	(7.86)	(6.94)	(7.08)	(7.32)	(7.88)	(6.84)	(6.85)	(7.07)
γ_{AvgFR}		0.218***				0.226***	0.203***	
		(16.95)				(6.72)	(13.96)	
γ_{AbsFR}			0.222***			-0.032		0.219***
			(15.58)			(-0.78)		(12.12)
γ_{StdFR}				0.318***		0.100**	0.082**	0.008
				(9.43)		(2.33)	(2.27)	(0.18)
$\gamma_{TrendFR}$					-0.084	-0.059		
					(-1.42)	(-1.06)		
$Adj.R^2$	24.10	33.98	32.62	27.44	24.14	34.11	34.12	32.58
$\Delta\, Adj.R^2$		9.88	8.51	3.34	0.04	10.01	10.02	8.48

Table 3. In-Sample Forecasting with Funding Rates: Log-Linear Specification

The table presents results from in-sample predictive regressions for Bitcoin spot log volatility using funding-rate based variables. The following log-linear regression model is estimated:

$$\ln(RV_{t,t+h}) = \alpha + \beta_d \ln(RV_{t,d}) + \beta_w \ln(RV_{t,w}) + \beta_m \ln(RV_{t,m}) + \gamma_{AvgFR} AvgFR_t + \gamma_{AbsFR} AbsFR_t + \gamma_{StdFR} StdFR_t + \gamma_{TrendFR} TrendFR_t + \varepsilon_{t+1},$$

where $\ln(RV_{t,t+h})$ is the log of realized volatility over forecast horizon h, and $\ln(RV_{t,d})$, $\ln(RV_{t,w})$, and $\ln(RV_{t,m})$ are the logs of daily, weekly, and monthly past realized volatilities, respectively. *AvgFR*, *AbsFR*, *StdFR*, and *TrendFR* are funding-rate-based measures that summarize the level, dispersion, and direction of intraday funding rates, as defined in Table 1. The table reports estimated coefficients and heteroscedasticity- and autocorrelation-robust t-statistics (in parentheses). Forecasting horizons correspond to h =1 (Panel A: daily), h =7 (Panel B: weekly), and h =30 (Panel C: monthly). The last two rows report the adjusted $R^2$and its increase relative to a benchmark log HAR model (i.e., the same regression with $\gamma_k = 0$ for all k). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Daily forecasting horizon ($h = 1$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α	0.036	0.030	0.038	0.041	0.035	0.030	0.030	0.035
	(1.35)	(1.14)	(1.45)	(1.55)	(1.33)	(1.11)	(1.11)	(1.33)
β_d	0.431***	0.417***	0.415***	0.422***	0.431***	0.418***	0.418***	0.418***
	(17.60)	(17.05)	(16.89)	(17.03)	(17.60)	(16.91)	(16.91)	(16.90)
β_w	0.367***	0.369***	0.362***	0.366***	0.368***	0.368***	0.369***	0.362***
	(9.54)	(9.63)	(9.44)	(9.49)	(9.54)	(9.57)	(9.62)	(9.44)
β_m	0.106***	0.099***	0.098***	0.101***	0.106***	0.099***	0.100***	0.099***
	(2.93)	(2.77)	(2.72)	(2.79)	(2.92)	(2.76)	(2.77)	(2.76)
γ_{AvgFR}		0.019***				0.017*	0.020***	
		(4.93)				(1.73)	(4.48)	
γ_{AbsFR}			0.019***			0.004		0.023***
			(4.52)			(0.29)		(4.14)
γ_{StdFR}				0.019**		-0.003	-0.001	-0.012
				(2.05)		(-0.27)	(-0.13)	(-0.97)
$\gamma_{TrendFR}$					0.007	0.009		
					(0.44)	(0.55)		
$Adj.R^2$	55.13	55.67	55.58	55.20	55.11	55.61	55.65	55.58
$\Delta\, Adj.R^2$		0.54	0.45	0.08	-0.02	0.48	0.52	0.45

Table 3 —*Continued*

Panel B: Weekly forecasting horizon ($h = 7$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α	0.304***	0.294***	0.309***	0.317***	0.303***	0.302***	0.298***	0.306***
	(13.14)	(13.05)	(13.76)	(13.76)	(13.10)	(13.26)	(13.12)	(13.54)
β_d	0.231***	0.207***	0.200***	0.210***	0.232***	0.203***	0.203***	0.202***
	(10.78)	(9.85)	(9.49)	(9.76)	(10.81)	(9.59)	(9.59)	(9.56)
β_w	0.325***	0.327***	0.314***	0.321***	0.326***	0.319***	0.326***	0.314***
	(9.63)	(9.97)	(9.58)	(9.58)	(9.65)	(9.69)	(9.92)	(9.58)
β_m	0.168***	0.157***	0.152***	0.156***	0.168***	0.153***	0.154***	0.154***
	(5.32)	(5.09)	(4.94)	(4.96)	(5.31)	(4.98)	(5.00)	(4.99)
γ_{AvgFR}		0.035***				0.012	0.033***	
		(10.39)				(1.43)	(8.73)	
γ_{AbsFR}			0.039***			0.029***		0.042***
			(10.69)			(2.76)		(9.05)
γ_{StdFR}				0.047***		-0.005	0.012	-0.011
				(5.68)		(-0.49)	(1.36)	(-1.11)
$\gamma_{TrendFR}$					0.015	0.019		
					(1.05)	(1.34)		
$Adj.R^2$	47.52	50.29	50.45	48.36	47.52	50.51	50.32	50.46
$\Delta\, Adj.R^2$		2.78	2.94	0.84	0.00	2.99	2.80	2.94

Panel C: Monthly forecasting horizon ($h = 30$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α	0.562***	0.549***	0.569***	0.582***	0.563***	0.560***	0.557***	0.569***
	(24.62)	(25.28)	(26.19)	(25.87)	(24.63)	(25.53)	(25.51)	(26.04)
β_d	0.144***	0.111***	0.103***	0.113***	0.143***	0.102***	0.103***	0.103***
	(6.79)	(5.50)	(5.09)	(5.36)	(6.76)	(5.02)	(5.06)	(5.04)
β_w	0.173***	0.176***	0.159***	0.166***	0.172***	0.170***	0.173***	0.159***
	(5.19)	(5.56)	(5.01)	(5.09)	(5.17)	(5.35)	(5.49)	(5.01)
β_m	0.214***	0.199***	0.193***	0.196***	0.214***	0.193***	0.194***	0.193***
	(6.86)	(6.70)	(6.50)	(6.38)	(6.87)	(6.53)	(6.53)	(6.48)
γ_{AvgFR}		0.047***				0.034***	0.042***	
		(14.50)				(4.16)	(11.72)	
γ_{AbsFR}			0.050***			0.012		0.049***
			(14.24)			(1.18)		(10.97)
γ_{StdFR}				0.070***		0.019*	0.026***	0.003
				(8.81)		(1.75)	(3.02)	(0.25)
$\gamma_{TrendFR}$					-0.013	-0.007		
					(-0.87)	(-0.54)		
$Adj.R^2$	31.21	37.99	37.77	33.85	31.20	38.24	38.26	37.74
$\Delta\, Adj.R^2$		6.79	6.57	2.65	-0.01	7.04	7.05	6.54

Table 4. Out-of-Sample Forecasting with Funding Rates

The table presents out-of-sample forecasting performance for Bitcoin spot volatility using funding-rate based variables. Panel A reports results from linear forecasting regressions, and Panel B reports results from log-linear regressions. For each panel, we consider five models: a benchmark HAR model using lagged realized volatility at daily, weekly, and monthly frequencies; three models that augment the HAR specification with a single funding-rate-based measure (*AvgFR*, *AbsFR*, or *StdFR*); and a Kitchen Sink model including all three measures. Forecasting models are estimated using a recursive procedure with an initial sample of 180 days of data. The table reports average losses based on MSE, MAPE, and QLIKE. It also presents the Giacomini-White (GW) test statistics for equal predictive accuracy, comparing each model to the HAR benchmark. Positive and statistically significant GW values indicate superior forecasting performance relative to the HAR model. Clark and West (CW) test statistics assess Granger causality from each funding-rate-based measure to volatility; significance indicates rejection of the null of no causality. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Linear Specification

Horizon	Daily					Weekly					Monthly				
Model	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink
Average Loss															
MSE	0.585	0.561	0.572	0.601	0.579	0.429	0.379	0.385	0.422	0.383	0.430	0.363	0.371	0.407	0.363
MAPE	0.877	0.825	0.847	0.894	0.844	0.709	0.625	0.634	0.695	0.631	0.772	0.671	0.677	0.733	0.673
QLIKE	-5.673	-5.742	-5.714	-5.661	-5.724	-5.598	-5.688	-5.680	-5.614	-5.681	-5.353	-5.461	-5.456	-5.396	-5.459
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		5.71^{***}	5.07^{***}	-4.39^{***}	1.00		11.58^{***}	11.72^{***}	5.18^{***}	9.72^{***}		13.78^{***}	14.13^{***}	11.28^{***}	13.42^{***}
MAPE		9.90^{***}	9.41^{***}	-7.06^{***}	4.60^{***}		17.54^{***}	18.58^{***}	8.00^{***}	14.46^{***}		14.63^{***}	16.18^{***}	13.19^{***}	13.93^{***}
QLIKE		12.99^{***}	13.10^{***}	-4.80^{***}	6.88^{***}		18.77^{***}	20.32^{***}	9.55^{***}	15.26^{***}		14.87^{***}	16.74^{***}	14.46^{***}	14.23^{***}
Testing for Granger causality															
CW tests		9.00^{***}	11.17^{***}	4.94^{***}	8.28^{***}		10.01^{***}	11.14^{***}	7.62^{***}	9.00^{***}		8.90^{***}	9.62^{***}	12.40^{***}	8.65^{***}

Table 4 —*Continued*

Panel B: Log-Linear Specification

Horizon	Daily					Weekly					Monthly				
Model	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink
Average Loss															
MSE	0.523	0.517	0.517	0.522	0.518	0.366	0.343	0.342	0.359	0.343	0.342	0.306	0.307	0.326	0.306
MAPE	0.670	0.657	0.656	0.667	0.660	0.520	0.487	0.482	0.508	0.485	0.605	0.554	0.548	0.579	0.553
QLIKE	-5.969	-5.987	-5.989	-5.975	-5.984	-5.856	-5.893	-5.901	-5.871	-5.896	-5.602	-5.654	-5.664	-5.632	-5.658
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		2.71^{***}	3.02^{***}	1.60	1.94^{*}		7.71^{***}	7.75^{***}	7.58^{***}	6.66^{***}		9.96^{***}	10.00^{***}	9.76^{***}	8.96^{***}
MAPE		5.71^{***}	7.17^{***}	5.90^{***}	4.14^{***}		11.95^{***}	12.88^{***}	11.88^{***}	10.56^{***}		10.39^{***}	12.07^{***}	11.82^{***}	9.24^{***}
QLIKE		7.78^{***}	10.55^{***}	10.65^{***}	5.98^{***}		12.91^{***}	15.17^{***}	14.09^{***}	11.89^{***}		9.86^{***}	12.21^{***}	13.43^{***}	9.21^{***}
Testing for Granger causality															
CW test		6.01^{***}	5.99^{***}	3.88^{***}	4.98^{***}		10.79^{***}	11.96^{***}	12.41^{***}	9.71^{***}		11.32^{***}	12.29^{***}	16.62^{***}	11.21^{***}

Table 5. Out-of-Sample Forecasting with Funding Rates: High vs. Low Perpetual Trading Volume

The table compares out-of-sample volatility forecasting performance of funding-rate-based models across periods of high and low perpetual trading volume. Panel A presents results from linear regressions, while Panel B corresponds to log-linear specifications. The models follow the same structure as in Table 4, including a benchmark HAR model, three HAR extensions using AvgFR, AbsFR, or StdFR individually, and a Kitchen Sink model combining all three. Perpetual trading volume is measured relative to spot trading volume as the ratio of 30-day aggregated perpetual futures volume to spot market volume. The sample is split into high- and low-volume periods based on the median value of this ratio. Forecasting models are estimated recursively with an initial window of 180 days. The table reports average losses (MSE, MAPE, and QLIKE), GW test statistics comparing each model to the HAR benchmark, and CW test statistics for Granger causality from funding-rate-based measures to volatility. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

Panel A: Linear Specification

Horizon	Daily					Weekly					Monthly				
Model	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink
High-Volume Periods															
Average Loss															
MSE	0.660	0.636	0.644	0.662	0.643	0.385	0.338	0.337	0.372	0.341	0.383	0.309	0.313	0.356	0.311
MAPE	0.989	0.925	0.950	0.992	0.931	0.684	0.592	0.598	0.663	0.595	0.762	0.645	0.651	0.719	0.646
QLIKE	-5.784	-5.870	-5.838	-5.779	-5.864	-5.801	-5.911	-5.900	-5.826	-5.908	-5.540	-5.676	-5.664	-5.590	-5.676
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		4.89***	5.24***	-3.38***	2.26**		10.73***	12.39***	8.83***	9.28***		18.85***	19.06***	10.40***	18.68***
MAPE		9.94***	9.06***	-4.86***	6.88***		18.24***	18.07***	10.83***	16.53***		18.92***	19.06***	10.36***	18.91***
QLIKE		14.09***	13.11***	-7.30***	9.46***		22.92***	21.76***	12.38***	20.67***		21.98***	21.11***	11.88***	22.41***
CW test		14.38***	15.77***	8.37***	6.36***		14.71***	16.02***	10.80***	12.77***		14.74***	15.60***	11.43***	11.61***
Low-Volume Periods															
Average Loss															
MSE	0.509	0.486	0.501	0.541	0.514	0.472	0.420	0.433	0.472	0.424	0.477	0.417	0.430	0.458	0.416
MAPE	0.765	0.724	0.744	0.796	0.758	0.733	0.658	0.669	0.727	0.666	0.781	0.698	0.704	0.747	0.700
QLIKE	-5.563	-5.614	-5.591	-5.542	-5.584	-5.395	-5.465	-5.460	-5.403	-5.455	-5.166	-5.247	-5.248	-5.203	-5.242
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		3.50***	2.20**	-4.21***	-0.44		7.08***	6.19***	0.14	5.85***		6.77***	6.39***	6.02***	6.66***
MAPE		4.89***	4.46***	-6.54***	0.60		9.26***	9.82***	2.03**	7.16***		6.84***	7.67***	8.30***	6.39***
QLIKE		5.96***	5.99***	-3.96***	1.75*		8.49***	9.80***	2.93***	6.26***		6.17***	7.63***	8.64***	5.62***
CW test		7.73***	8.58***	4.95***	5.43***		8.02***	8.61***	5.28***	7.39***		7.56***	8.03***	9.04***	7.31***

Table 5 —*Continued*

Panel B: Log-Linear Specification

Horizon	Daily					Weekly					Monthly				
Model	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink
High-Volume Periods															
Average Loss															
MSE	0.594	0.588	0.588	0.592	0.589	0.337	0.315	0.310	0.327	0.312	0.284	0.244	0.240	0.264	0.241
MAPE	0.754	0.736	0.737	0.749	0.737	0.510	0.470	0.464	0.494	0.465	0.593	0.530	0.523	0.563	0.527
QLIKE	-6.117	-6.145	-6.144	-6.124	-6.144	-6.049	-6.099	-6.106	-6.068	-6.105	-5.797	-5.871	-5.875	-5.831	-5.876
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE_LOG		2.81^{***}	3.09^{***}	3.40^{***}	2.42^{**}		7.97^{***}	8.96^{***}	7.73^{***}	8.43^{***}		14.70^{***}	15.45^{***}	10.60^{***}	14.71^{***}
MAPE		6.83^{***}	6.71^{***}	6.96^{***}	6.45^{***}		13.18^{***}	13.47^{***}	10.65^{***}	13.33^{***}		14.30^{***}	15.36^{***}	10.59^{***}	14.40^{***}
QLIKE		11.05^{***}	11.27^{***}	10.25^{***}	10.46^{***}		17.26^{***}	17.31^{***}	12.64^{***}	17.53^{***}		17.46^{***}	17.59^{***}	12.25^{***}	17.69^{***}
CW test		11.32^{***}	11.47^{***}	4.87^{***}	10.84^{***}		14.17^{***}	16.17^{***}	12.29^{***}	15.34^{***}		14.76^{***}	16.93^{***}	14.90^{***}	14.86^{***}
Low-Volume Periods															
Average Loss															
MSE	0.451	0.445	0.446	0.451	0.447	0.394	0.371	0.373	0.391	0.373	0.399	0.369	0.374	0.387	0.371
MAPE	0.586	0.579	0.575	0.585	0.584	0.531	0.504	0.500	0.522	0.504	0.618	0.578	0.572	0.595	0.579
QLIKE	-5.822	-5.829	-5.835	-5.825	-5.824	-5.664	-5.686	-5.697	-5.674	-5.688	-5.408	-5.438	-5.453	-5.434	-5.440
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		1.47	1.56	-1.11	0.90		4.42^{***}	3.81^{***}	2.77^{***}	3.32^{***}		4.65^{***}	3.96^{***}	4.53^{***}	3.83^{***}
MAPE		2.07^{**}	3.73^{***}	1.23	0.70		5.70^{***}	6.31^{***}	6.08^{***}	4.50^{***}		4.56^{***}	5.45^{***}	6.67^{***}	3.75^{***}
QLIKE		1.94^{*}	4.59^{***}	4.74^{***}	0.65		4.71^{***}	6.79^{***}	7.28^{***}	4.13^{***}		3.10^{***}	4.98^{***}	7.43^{***}	2.88^{***}
CW test		4.99^{***}	4.76^{***}	2.75^{***}	4.23^{***}		8.08^{***}	8.69^{***}	7.03^{***}	7.47^{***}		8.93^{***}	9.50^{***}	11.04^{***}	8.88^{***}

Table 6. Out-of-Sample Forecasting with Funding Rates: Realized Kernel as Volatility Proxy

This table replicates the out-of-sample forecasting analysis in Table 5 using **realized kernel** as the volatility proxy instead of realized volatility. Forecasting models include the benchmark HAR, single-variable HAR extensions using AvgFR, AbsFR, or StdFR, and a Kitchen Sink model combining all three funding-rate-based measures. All models are estimated recursively with an initial window of 180 days. The table reports average losses (MSE, MAPE, QLIKE), Giacomini-White (GW) test statistics comparing each model to the HAR benchmark, and CW test statistics for Granger causality. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Horizon	Daily					Weekly					Monthly				
Model	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink
High-Volume Periods															
Average Loss															
MSE	0.860	0.835	0.844	0.865	0.845	0.375	0.336	0.337	0.365	0.340	0.387	0.318	0.324	0.362	0.319
MAPE	1.313	1.237	1.269	1.322	1.245	0.688	0.609	0.616	0.670	0.613	0.768	0.660	0.669	0.729	0.662
QLIKE	-5.638	-5.733	-5.694	-5.626	-5.724	-5.882	-5.983	-5.970	-5.903	-5.978	-5.628	-5.759	-5.745	-5.676	-5.759
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		4.77^{***}	5.06^{***}	-4.23^{***}	1.78^{*}		9.39^{***}	11.12^{***}	8.41^{***}	7.55^{***}		18.32^{***}	18.37^{***}	10.55^{***}	17.97^{***}
MAPE		8.80^{***}	8.52^{***}	-5.51^{***}	5.38^{***}		16.26^{***}	16.48^{***}	11.22^{***}	14.18^{***}		17.63^{***}	17.76^{***}	10.47^{***}	17.66^{***}
QLIKE		11.39^{***}	11.08^{***}	-7.25^{***}	6.81^{***}		22.32^{***}	21.77^{***}	13.34^{***}	19.42^{***}		22.14^{***}	21.54^{***}	12.89^{***}	22.62^{***}
CW test		14.45^{***}	16.69^{***}	11.22^{***}	5.92^{***}		14.73^{***}	16.90^{***}	12.01^{***}	11.96^{***}		14.87^{***}	16.29^{***}	13.37^{***}	10.99^{***}
Low-Volume Periods															
Average Loss															
MSE	0.695	0.672	0.689	0.737	0.703	0.478	0.434	0.446	0.479	0.437	0.444	0.393	0.404	0.427	0.393
MAPE	1.027	0.978	1.009	1.067	1.010	0.736	0.674	0.684	0.733	0.680	0.737	0.669	0.673	0.705	0.670
QLIKE	-5.483	-5.542	-5.507	-5.450	-5.507	-5.489	-5.548	-5.543	-5.494	-5.538	-5.300	-5.365	-5.369	-5.335	-5.360
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		2.94^{***}	1.26	-3.24^{***}	-0.72		6.09^{***}	5.17^{***}	-0.69	5.19^{***}		5.61^{***}	5.10^{***}	5.22^{***}	5.48^{***}
MAPE		4.60^{***}	2.96^{***}	-5.95^{***}	1.18		7.91^{***}	8.18^{***}	1.11	6.21^{***}		5.43^{***}	6.07^{***}	7.46^{***}	5.05^{***}
QLIKE		5.53^{***}	4.05^{***}	-4.53^{***}	1.64		7.37^{***}	8.51^{***}	1.86^{*}	5.30^{***}		4.87^{***}	6.21^{***}	8.10^{***}	4.37^{***}
CW test		7.48^{***}	8.55^{***}	4.93^{***}	7.17^{***}		7.69^{***}	8.18^{***}	4.41^{***}	7.25^{***}		7.41^{***}	7.74^{***}	9.16^{***}	7.19^{***}

Table 7. Out-of-Sample Forecasting with Funding Rates: Robustness to Estimation Methods

This table evaluates the robustness of the out-of-sample forecasting results in Table 5 to alternative estimation methods. Panel A reports results based on recursive estimation with an initial window of 365 days**,** while Panel B uses rolling estimation with a window size of 180 days. The forecasting models are estimated in linear specifications, and the model structure and evaluation metrics follow the same structure as in Table 5. The table reports average forecast losses (MSE, MAPE, and QLIKE), GW test statistics comparing each model to the benchmark HAR, and CW test statistics for Granger causality from funding-rate-based measures to volatility. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Recursive estimation with an initial sample of 365 days

Horizon	Daily					Weekly					Monthly				
Model	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink
High-Volume Periods															
Average Loss															
MSE	0.682	0.657	0.665	0.684	0.663	0.402	0.353	0.352	0.390	0.356	0.404	0.324	0.329	0.376	0.325
MAPE	1.020	0.953	0.980	1.023	0.956	0.706	0.609	0.617	0.686	0.612	0.795	0.667	0.675	0.751	0.669
QLIKE	-5.782	-5.873	-5.839	-5.778	-5.871	-5.806	-5.923	-5.910	-5.830	-5.921	-5.535	-5.680	-5.665	-5.584	-5.680
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		4.75***	5.09***	-3.16***	2.25***		10.24***	11.88***	7.54***	8.83***		19.03***	19.03***	9.45***	18.78***
MAPE		9.59***	8.64***	-4.21***	6.88***		17.70***	17.46***	9.32***	16.12***		19.68***	19.55***	9.68***	19.69***
QLIKE		13.76***	12.67***	-6.46***	9.53***		22.82***	21.40***	11.15***	20.66***		22.56***	21.40***	10.99***	23.00***
CW test		13.55***	14.82***	8.20***	6.14***		13.99***	15.21***	10.25***	11.98***		14.09***	14.85***	10.83***	10.86***
Low-Volume Periods															
Average Loss															
MSE	0.522	0.501	0.507	0.531	0.510	0.452	0.406	0.413	0.443	0.407	0.408	0.356	0.362	0.385	0.353
MAPE	0.795	0.750	0.761	0.801	0.763	0.713	0.633	0.640	0.696	0.637	0.717	0.633	0.637	0.678	0.634
QLIKE	-5.571	-5.627	-5.614	-5.569	-5.611	-5.436	-5.519	-5.515	-5.456	-5.514	-5.244	-5.338	-5.334	-5.286	-5.335
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		3.36***	4.03***	-2.77***	1.61		6.22***	5.73***	6.40***	5.64***		8.20***	8.22***	7.70***	8.68***
MAPE		5.70***	7.27***	-3.45***	3.39***		9.55***	10.05***	9.30***	8.46***		9.20***	10.08***	9.23***	9.14***
QLIKE		7.10***	9.32***	-1.04	4.14***		9.88***	10.82***	11.10***	8.63***		9.54***	10.41***	9.80***	9.25***
CW test		6.44***	7.82***	3.22***	5.39***		7.54***	8.37***	11.88***	6.96***		6.89***	7.21***	8.45***	6.77***

Table 7 —*Continued*

Panel B: Rolling estimation with a estimation window of 180 days

Horizon	Daily					Weekly					Monthly				
Model	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink	HAR	AvgFR	AbsFR	StdFR	Sink
High-Volume Periods															
Average Loss															
MSE	0.630	0.625	0.625	0.627	0.649	0.315	0.278	0.277	0.303	0.292	0.222	0.190	0.192	0.213	0.191
MAPE	0.866	0.853	0.852	0.852	0.877	0.501	0.449	0.449	0.482	0.464	0.469	0.421	0.424	0.453	0.423
QLIKE	-5.968	-5.987	-5.990	-5.989	-5.972	-6.074	-6.123	-6.120	-6.093	-6.109	-6.000	-6.030	-6.025	-6.012	-6.029
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		0.85	0.89	0.56	-1.98*		5.74***	5.95***	2.89***	2.61***		10.20***	9.74***	5.63***	8.17***
MAPE		1.68*	2.01**	2.69***	-0.83		8.25***	8.45***	4.81***	3.35***		12.94***	12.66***	7.44***	9.30***
QLIKE		2.44**	3.01***	3.95***	0.28		7.49***	6.96***	4.83***	3.17***		6.33***	5.37***	5.06***	4.63***
CW test		9.30***	9.09***	6.43***	5.83***		9.47***	8.88***	4.45***	9.35***		6.86***	6.33***	5.25***	5.20***
Low-Volume Periods															
Average Loss															
MSE	0.491	0.510	0.514	0.531	0.550	0.419	0.438	0.434	0.424	0.447	0.339	0.368	0.370	0.343	0.376
MAPE	0.713	0.746	0.740	0.748	0.787	0.624	0.655	0.647	0.627	0.666	0.563	0.601	0.602	0.569	0.612
QLIKE	-5.640	-5.609	-5.620	-5.616	-5.580	-5.547	-5.514	-5.522	-5.545	-5.504	-5.482	-5.436	-5.437	-5.478	-5.428
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		-2.67***	-3.18***	-3.15***	-5.21***		-2.64***	-1.96*	-1.21	-3.16***		-4.09***	-4.41***	-1.36	-4.47***
MAPE		-3.56***	-3.32***	-4.43***	-5.62***		-3.14***	-2.37**	-0.62	-3.50***		-3.74***	-3.88***	-1.55	-4.11***
QLIKE		-3.27***	-2.37**	-2.90***	-4.48***		-3.36***	-2.53**	-0.31	-3.47***		-4.42***	-4.39***	-1.10	-4.44***
CW test		8.22***	7.71***	5.30***	9.23***		8.93***	8.04***	7.78***	8.64***		5.28***	4.71***	6.20***	5.28***

Table 8. In-Sample Forecasting with Perpetual Market Volatility

The table presents results from in-sample predictive regressions for Bitcoin spot volatility using volatility differentials between the spot and perpetual futures markets. Panel A reports results from the linear specification:

$$RV_{t,t+h} = \alpha + \beta_d RV_{t,d} + \beta_w RV_{t,w} + \beta_m RV_{t,m} + \delta_d\left(RV_{t,d} - RV_{t,d}^{perp}\right) + \delta_w\left(RV_{t,w} - RV_{t,w}^{perp}\right) + \delta_m\left(RV_{t,m} - RV_{t,m}^{perp}\right) + \varepsilon_{t+1}.$$

Panel B reports results from the log-linear specification:

$$\ln\left(RV_{t,t+h}\right) = \alpha + \beta_d \ln\left(RV_{t,d}\right) + \beta_w \ln\left(RV_{t,w}\right) + \beta_m \ln\left(RV_{t,m}\right) + \delta_d \ln\left(RV_{t,d}/RV_{t,d}^{perp}\right) + \delta_w \ln\left(RV_{t,w}/RV_{t,w}^{perp}\right) + \delta_m \ln\left(RV_{t,m}/RV_{t,m}^{perp}\right) + \varepsilon_{t+1}.$$

RV_{t+h} denotes the realized spot volatility at horizon h, while $RV_{t,d}$, $RV_{t,w}$, $RV_{t,m}$ and $RV_{t,d}^{perp}$, $RV_{t,w}^{perp}$, $RV_{t,m}^{perp}$ denote past realized volatilities from the spot and perpetual markets, respectively, at daily, weekly, and monthly frequencies. Forecast horizons correspond to h =1 (daily), h =7 (weekly), and h =30 (monthly). The table reports estimated coefficients and robust t-statistics (in parentheses). The last two rows report the adjusted R^2 and its increase over the benchmark HAR model (i.e., excluding the volatility differential terms). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Linear Specification

Horizon	Daily					Weekly					Monthly				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
α	0.519***	0.519***	0.551***	0.591***	0.589***	1.080***	1.104***	1.135***	1.193***	1.197***	1.754***	1.789***	1.834***	1.899***	1.906***
	(6.06)	(6.01)	(6.31)	(6.63)	(6.60)	(13.73)	(13.95)	(14.20)	(14.65)	(14.68)	(23.90)	(24.31)	(24.72)	(25.13)	(25.23)
β_d	0.427***	0.427***	0.425***	0.423***	0.419***	0.273***	0.280***	0.270***	0.268***	0.271***	0.134***	0.145***	0.130***	0.128***	0.132***
	(16.90)	(16.76)	(16.83)	(16.78)	(16.40)	(11.79)	(12.03)	(11.70)	(11.63)	(11.63)	(6.21)	(6.68)	(6.05)	(5.97)	(6.13)
β_w	0.289***	0.289***	0.296***	0.291***	0.296***	0.222***	0.218***	0.234***	0.225***	0.224***	0.123***	0.118***	0.141***	0.127***	0.129***
	(7.92)	(7.91)	(8.08)	(7.99)	(8.02)	(6.62)	(6.51)	(6.98)	(6.75)	(6.66)	(3.95)	(3.78)	(4.54)	(4.12)	(4.13)
β_m	0.087**	0.087**	0.076**	0.075**	0.074**	0.174***	0.166***	0.155***	0.155***	0.151***	0.231***	0.219***	0.202***	0.205***	0.197***
	(2.54)	(2.53)	(2.18)	(2.17)	(2.14)	(5.55)	(5.27)	(4.86)	(4.91)	(4.76)	(7.86)	(7.45)	(6.82)	(7.03)	(6.71)
δ_d		-0.004			-0.405		0.803**			0.344		1.187***			0.522
		(-0.01)			(-1.04)		(2.54)			(0.96)		(4.04)			(1.58)
δ_w			0.856*		0.382			1.485***		0.199			2.167***		0.614
			(1.88)		(0.62)			(3.56)		(0.35)			(5.60)		(1.18)
δ_m				1.635***	1.542**				2.573***	2.265***				3.289***	2.589***
				(2.87)	(2.19)				(4.95)	(3.51)				(6.82)	(4.34)
$Adj.R^2$	45.47	45.44	45.54	45.67	45.65	37.44	37.62	37.82	38.20	38.18	24.10	24.71	25.29	25.87	26.04
$\Delta\, Adj.R^2$		-0.03	0.07	0.21	0.18		0.18	0.38	0.76	0.74		0.60	1.19	1.76	1.93

Table 8 —*Continued*

Panel B: Log-Linear Specification

Horizon	Daily					Weekly					Monthly				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
α	0.036	0.032	0.053*	0.077***	0.070**	0.304***	0.300***	0.329***	0.362***	0.352***	0.562***	0.584***	0.623***	0.657***	0.663***
	(1.35)	(1.17)	(1.88)	(2.72)	(2.41)	(13.14)	(12.44)	(13.36)	(14.61)	(13.96)	(24.62)	(24.57)	(25.96)	(27.37)	(27.09)
β_d	0.431***	0.430***	0.430***	0.424***	0.422***	0.231***	0.231***	0.230***	0.221***	0.219***	0.144***	0.146***	0.140***	0.128***	0.128***
	(17.60)	(17.58)	(17.57)	(17.33)	(17.22	(10.78)	(10.75)	(10.73)	(10.39)	(10.26)	(6.79)	(6.91)	(6.72)	(6.18)	(6.21)
β_w	0.367***	0.367***	0.379***	0.377***	0.379***	0.325***	0.325***	0.342***	0.338***	0.341***	0.173***	0.173***	0.216***	0.195***	0.205***
	(9.54)	(9.54)	(9.70)	(9.80)	(9.66)	(9.63)	(9.63)	(10.01)	(10.10)	(10.00)	(5.19)	(5.19)	(6.46)	(5.99)	(6.20)
β_m	0.106***	0.108***	0.084**	0.077**	0.083**	0.168***	0.171***	0.137***	0.128***	0.135***	0.214***	0.196***	0.136***	0.148***	0.133***
	(2.93)	(2.96)	(2.19)	(2.09)	(2.17)	(5.32)	(5.33)	(4.09)	(4.00)	(4.08)	(6.86)	(6.21)	(4.18)	(4.78)	(4.14)
δ_d		-0.167			-0.634		-0.196			-0.856**		1.138***			-0.058
		(-0.41)			(-1.37)		(-0.55)			(-2.12)		(3.23)			(-0.15)
δ_w			0.887*		0.059			1.273***		0.135			3.146***		0.974*
			(1.77)		(0.09)			(2.90)		(0.23)			(7.35)		(1.73)
δ_m				1.839***	2.000***				2.547***	2.733***				4.189***	3.677***
				(3.87)	(3.45)				(6.15)	(5.42)				(10.43)	(7.52)
$Adj.R^2$	55.13	55.11	55.18	55.45	55.46	47.52	47.50	47.72	48.51	48.59	31.21	31.54	33.06	34.87	34.92
$\Delta\ Adj.R^2$		-0.02	0.05	0.33	0.33		-0.02	0.20	0.99	1.07		0.34	1.86	3.67	3.71

Table 9. Out-of-Sample Forecasting with Perpetual Market Volatility

This table reports out-of-sample forecasting performance for Bitcoin spot volatility using perpetual market volatility as explanatory variables. Panel A presents results from linear forecasting regressions, and Panel B reports results from log-linear regressions. Forecasting models include a benchmark HAR model, three HAR extensions that incorporate perpetual market volatility at each horizon (daily, weekly, and monthly), and a Kitchen Sink model including all three. All models are estimated recursively with an initial window of 180 days. The table reports average forecast losses (MSE, MAPE, and QLIKE), GW test statistics comparing each model to the HAR benchmark, and CW test statistics for Granger causality from perpetual to spot market volatility. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Linear Specification

Horizon	Daily					Weekly					Monthly				
Model	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink
Average Loss															
MSE	0.585	0.592	0.582	0.574	0.582	0.429	0.426	0.421	0.413	0.413	0.430	0.425	0.417	0.406	0.410
MAPE	0.877	0.883	0.870	0.857	0.868	0.709	0.704	0.700	0.691	0.690	0.772	0.765	0.757	0.731	0.740
QLIKE	-5.673	-5.669	-5.682	-5.696	-5.684	-5.598	-5.603	-5.605	-5.613	-5.614	-5.353	-5.361	-5.363	-5.386	-5.384
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		-3.05***	1.41	2.82***	0.68		2.09**	2.22**	4.93***	4.43***		2.36**	3.79***	5.22***	3.65***
MAPE		-2.72***	2.66***	3.49***	1.52		2.16**	1.86*	3.87***	3.62***		2.06**	2.98***	6.49***	4.32***
QLIKE		-1.82*	3.69***	4.05***	1.85*		2.49**	1.56	3.08***	3.04***		2.09**	1.92*	5.21***	4.14***
CW test		4.58***	7.00***	11.66***	7.73***		1.41*	4.19***	10.72***	5.05***		1.93**	5.89***	14.95***	6.27***

Panel B: Log-Linear Specification

Horizon	Daily					Weekly					Monthly				
Model	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink
Average Loss															
MSE	0.523	0.524	0.524	0.520	0.522	0.366	0.367	0.364	0.358	0.358	0.342	0.340	0.331	0.323	0.327
MAPE	0.670	0.672	0.669	0.658	0.662	0.520	0.522	0.516	0.502	0.505	0.605	0.600	0.588	0.568	0.575
QLIKE	-5.969	-5.965	-5.972	-5.989	-5.982	-5.856	-5.852	-5.863	-5.880	-5.875	-5.602	-5.611	-5.625	-5.648	-5.648
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		-1.51	-0.79	0.90	0.25		-0.98	0.71	2.46**	2.26**		1.08	3.61***	4.19***	3.26***
MAPE		-1.91*	0.84	3.29***	2.07***		-2.22**	1.79*	4.65***	3.61***		2.94***	3.92***	5.72***	4.52***
QLIKE		-3.14***	1.86*	5.29***	3.07***		-4.09***	3.02***	5.98***	4.37***		5.43***	4.98***	6.82***	6.80***
CW test		2.33***	3.14***	6.27***	7.23***		4.29***	3.56***	6.88***	7.04***		5.44***	5.50***	8.95***	9.76***

Table 10. Out-of-Sample Forecasting with Perpetual Market Volatility: High vs. Low Perpetual Trading Volume

This table evaluates the out-of-sample forecasting performance of models using perpetual market volatility during periods of high and low perpetual trading volume. The analysis follows the linear specification in Panel A of Table 9. The sample is split into high- and low-volume subsamples based on the median of the 30-day relative trading volume ratio between perpetual and spot markets. Forecasting models include the HAR benchmark, three HAR extensions incorporating perpetual volatility at daily, weekly, and monthly horizons, and a Kitchen Sink model including all three. All models are estimated recursively with an initial window of 180 days. The table reports average forecast losses (MSE, MAPE, and QLIKE), GW test statistics comparing each model to the HAR benchmark, and CW test statistics for Granger causality. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Horizon	Daily					Weekly					Monthly				
Model	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink
High-Volume Periods															
Average Loss															
MSE	0.660	0.664	0.655	0.645	0.649	0.385	0.382	0.378	0.364	0.364	0.383	0.377	0.366	0.341	0.345
MAPE	0.989	0.994	0.978	0.958	0.965	0.684	0.678	0.666	0.652	0.650	0.762	0.751	0.734	0.701	0.706
QLIKE	-5.784	-5.779	-5.796	-5.821	-5.812	-5.801	-5.809	-5.824	-5.840	-5.839	-5.540	-5.553	-5.569	-5.595	-5.593
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		-2.44**	1.80*	2.16**	1.69*		1.54	1.19	3.86***	3.50***		1.69*	2.84***	6.08***	5.30***
MAPE		-2.27**	2.49**	2.91***	2.28**		1.74*	2.22**	4.13***	3.95***		1.84*	3.05***	6.14***	5.36***
QLIKE		-2.29**	2.70***	3.44***	2.65***		2.11**	2.74***	4.72***	4.33***		2.10**	3.12***	5.48***	4.98***
CW test		4.39***	6.03***	9.20***	6.72***		1.71**	3.69***	8.89***	5.42***		1.62*	4.68***	12.81***	5.17***
Low-Volume Periods															
Average Loss															
MSE	0.509	0.520	0.510	0.504	0.515	0.472	0.470	0.464	0.461	0.461	0.477	0.473	0.468	0.471	0.476
MAPE	0.765	0.772	0.763	0.757	0.771	0.733	0.731	0.734	0.730	0.730	0.781	0.778	0.779	0.760	0.773
QLIKE	-5.563	-5.559	-5.569	-5.572	-5.557	-5.395	-5.398	-5.387	-5.386	-5.389	-5.166	-5.168	-5.157	-5.178	-5.175
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		-2.41**	-0.30	2.33**	-1.19		1.54	2.39**	3.11***	2.72***		1.74*	2.61***	1.05	0.15
MAPE		-1.86*	0.97	2.43**	-1.12		1.44	-0.34	0.53	0.44		0.93	0.52	2.72 ***	0.75
QLIKE		-0.86	2.81***	2.64***	-1.07		1.44	-2.09**	-2.17**	-1.25		0.53	-2.03**	1.50	0.86
CW test		1.58*	3.91***	7.38***	5.79***		1.12	3.49***	7.45***	4.36***		1.45***	3.83***	8.01***	4.47***

Table 11. Out-of-Sample Forecasting with Perpetual Market Volatility: Realized Kernel as Volatility Proxy

This table replicates the out-of-sample forecasting analysis from Table 9 using realized kernel as the target volatility measure instead of realized volatility. The forecasting models are identical to those in Panel A of Table 9: a benchmark HAR model, three HAR extensions incorporating perpetual market volatility at daily, weekly, or monthly horizons, and a Kitchen Sink model including all three, estimated in linear specifications. All models are estimated recursively with an initial window of 180 days. The table reports average forecast losses (MSE, MAPE, and QLIKE), GW test statistics relative to the HAR benchmark, and CW test statistics for Granger causality. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Horizon	Daily					Weekly					Monthly				
Model	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink	HAR	$RV_{t,d}^{perp}$	$RV_{t,w}^{perp}$	$RV_{t,m}^{perp}$	Sink
High-Volume Periods															
Average Loss															
MSE	0.860	0.854	0.848	0.837	0.843	0.375	0.370	0.368	0.358	0.358	0.387	0.380	0.373	0.356	0.356
MAPE	1.313	1.282	1.280	1.252	1.251	0.688	0.679	0.676	0.657	0.656	0.768	0.756	0.746	0.722	0.722
QLIKE	-5.638	-5.675	-5.674	-5.704	-5.708	-5.882	-5.895	-5.899	-5.922	-5.923	-5.628	-5.644	-5.655	-5.680	-5.680
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		1.47	4.02***	4.08***	2.65***		2.73***	3.44***	3.82***	3.68***		4.32***	4.45***	6.51***	6.36***
MAPE		5.27***	6.18***	5.62***	5.63***		4.00***	4.92***	5.09***	5.04***		4.50***	4.60***	6.36***	6.25***
QLIKE		6.32***	6.66***	6.02***	6.42***		6.00***	6.93***	6.68***	6.56***		6.06***	5.59***	7.19***	7.15***
CW test		10.64***	14.43***	14.24***	11.56***		10.66***	14.08***	13.50***	13.39***		10.40***	13.88***	14.79***	14.46***
Low-Volume Periods															
Average Loss															
MSE	0.695	0.692	0.689	0.684	0.691	0.478	0.477	0.478	0.457	0.465	0.444	0.443	0.437	0.410	0.425
MAPE	1.027	1.012	1.015	1.011	1.014	0.736	0.733	0.742	0.730	0.742	0.737	0.739	0.751	0.727	0.751
QLIKE	-5.483	-5.503	-5.491	-5.487	-5.492	-5.489	-5.491	-5.472	-5.470	-5.461	-5.300	-5.296	-5.272	-5.265	-5.247
GW tests on Loss Differences Relative to the Benchmark HAR model															
MSE		0.40	1.36	1.91*	0.51		0.25	-0.04	3.52***	2.12**		0.38	1.15	6.29***	2.94***
MAPE		1.61	1.98**	1.84*	1.09		0.77	-1.45	0.71	-0.72		-0.67	-2.12**	1.50	-1.68*
QLIKE		2.10	1.30	0.52	0.75		0.49	-3.66***	-2.47***	-3.46***		-0.99	-3.84***	-4.80***	-6.10***
CW test		6.11***	8.49***	12.18***	7.73***		6.16***	6.80***	11.41***	10.41***		5.88***	7.28***	11.26***	9.72***

Table 12. Out-of-Sample Forecasting with Perpetual Market Volatility: Robustness to Estimation Methods

This table assesses the robustness of out-of-sample forecast performance to alternative estimation methods, using perpetual market volatilities as predictive variables. Only the benchmark HAR model and the Kitchen Sink model (which includes perpetual volatilities at all horizons) are considered in linear specifications. Panel A reports results based on recursive estimation with an initial sample of 365 days, while Panel B uses rolling estimation with a fixed window of 180 days. For each forecast horizon, the table presents average forecast losses (MSE, MAPE, and QLIKE), along with GW test statistics comparing each model against the HAR benchmark, and CW test statistics for Granger causality. The analysis is conducted for the full sample, and separately for high and low perpetual trading volume periods. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Recursive estimation with an initial sample of 365 days

Horizon	Daily		Weekly		Monthly	
	HAR	Sink	HAR	Sink	HAR	Sink
High-Volume Periods						
Average Loss						
MSE	0.682	0.667	0.402	0.383	0.404	0.367
MAPE	1.020	0.989	0.706	0.675	0.795	0.742
QLIKE	-5.782	-5.817	-5.806	-5.841	-5.535	-5.583
GW tests on Loss Differences Relative to the Benchmark HAR model						
MSE		1.97^{**}		2.90^{***}		4.54^{***}
MAPE		2.62^{***}		3.26^{***}		4.53^{***}
QLIKE		2.97^{***}		3.57^{***}		4.15^{***}
CW test		6.72^{***}		5.03^{***}		4.15^{***}
Low-Volume Periods						
Average Loss						
MSE	0.522	0.516	0.452	0.443	0.408	0.399
MAPE	0.795	0.787	0.713	0.706	0.717	0.711
QLIKE	-5.571	-5.580	-5.436	-5.442	-5.244	-5.251
GW tests on Loss Differences Relative to the Benchmark HAR model						
MSE		2.33^{**}		2.42^{**}		2.10^{**}
MAPE		2.29^{**}		1.32		1.06
QLIKE		2.65^{***}		1.30		1.31
CW test		6.94^{***}		3.28^{***}		3.06^{***}

Table 12 —*Continued*

Panel B: Rolling estimation with an estimation window of 180 days

Horizon	Daily		Weekly		Monthly	
	HAR	Sink	HAR	Sink	HAR	Sink
High-Volume Periods						
Average Loss						
MSE	0.491	0.491	0.419	0.409	0.423	0.339
MAPE	0.713	0.694	0.624	0.581	0.563	0.527
QLIKE	-5.640	-5.677	-5.547	-5.637	-5.482	-5.592
GW tests on Loss Differences Relative to the Benchmark HAR model						
MSE		-0.04		1.27		2.24**
MAPE		3.71**		4.67***		2.75***
QLIKE		7.47**		8.94***		7.85***
CW test		5.01**		7.14***		6.50***
Low-Volume Periods						
Average Loss						
MSE	0.630	0.660	0.315	0.342	0.222	0.218
MAPE	0.866	0.851	0.501	0.551	0.469	0.487
QLIKE	-5.968	-5.996	-6.074	-6.034	-6.000	-5.963
GW tests on Loss Differences Relative to the Benchmark HAR model						
MSE		-1.39		-1.73*		0.21
MAPE		0.95		-1.63		-0.40
QLIKE		1.62		-1.26		-0.78
CW test		4.08**		2.34**		2.88***