

Spot Regressions with Candlesticks*

Yasin Simsek

Department of Economics, Duke University

April 30, 2026

Abstract

Betas from spot regressions are central to asset pricing and risk management, as measures of systematic risk. This paper develops a new estimation and inference framework for spot regressions by leveraging high-frequency candlesticks, extending conventional (open-to-close) returns with intra-period high/low prices. Specifically, I construct candlestick-based estimators of regression parameters, including spot beta, by minimizing a quadratic risk under a fixed-k asymptotic framework. I then develop a feasible hypothesis testing procedure for spot betas with correct asymptotic size. Simulation results show that the proposed estimator reduces estimation risk relative to return-based estimators, especially in small samples, and the test achieves notably higher power. I apply the framework to assess the market neutrality of Bitcoin using 1-minute data on IBIT and SPY, finding deviations from neutrality, particularly in high-volatility periods.

Keywords: High-frequency data, spot beta, spot covariance, candlesticks.

JEL Codes: C13, C32, C58, G11.

*This paper is built on the third chapter of my PhD dissertation at Duke University. I am very grateful to Andrew J. Patton, Tim Bollerslev and Anna Bykhovskaya for their guidance and support. I would also like to thank Anna Cieslak, Mehmet Caner, Peter R. Hansen, Ilze Kalnina, Raul Guarini Riva, Adam Rosen, George Tauchen, Christopher Walker and seminar participants at Duke University, 2025 Triangle Econometrics Conference, and 2025 Annual Meeting of SoFiE for their helpful comments and suggestions. Email: yasin.simsek@duke.edu.

1 Introduction

Time-series return regressions are widely employed in empirical finance to measure an asset’s exposure to risk factors, most prominently the market portfolio. The slope coefficients of these regressions are commonly referred to as betas and they are central to asset pricing and risk management research. The conventional approach for running these regressions relies on low-frequency returns (e.g., daily or weekly) over long horizons (e.g., months or years), thereby treating betas as constants throughout the estimation window. However, a substantial body of evidence shows that risk exposures are time-varying; see for example [Engle \(2004\)](#). High-frequency intraday returns provide an effective means for capturing this variation and have enabled more precise estimation over shorter horizons, such as a day or a month; see for example [Aït-Sahalia and Jacod \(2014\)](#) and references therein¹. More recently, a growing literature documents that betas vary significantly even within a single trading day (e.g., [Bibinger et al. \(2019\)](#); [Andersen et al. \(2021, 2023\)](#); [Liao and Todorov \(2024\)](#); [Patton and Simsek \(2024\)](#)). These findings motivate the estimation of risk exposures over very short intraday windows, commonly known as spot regressions. Since such spot estimates are constructed from returns observed over narrow time intervals, they are subject to limited local information, giving rise to a fundamental bias-variance trade-off between statistical precision and localization.

Building on this background, I develop a novel estimation framework for spot regressions using a new source of information to mitigate the localization-precision trade-off. Specifically, my approach leverages the richer information in high-frequency candlesticks that contain the open, high, low, and close prices within each sampling interval, moving beyond the conventional methods that rely solely on (open-to-close) returns. Importantly, this work provides a practical way to obtain more precise estimates and reliable inference, while preserving the local structure of the estimation.

I first construct a spot covariance estimator that exploits the second moments of candlestick variables and then derive estimates for regression parameters, including spot beta, implied by this estimator. The functional form of this estimator is obtained by minimizing a well-defined quadratic risk function. The resulting expression is analytically tractable and resembles the least squares formula. These features render the estimator readily applicable across settings without requiring additional economet-

¹Modeling time variation in betas is not limited to high-frequency methods and researchers have developed various approaches over time. One strand of the literature, for example, captures the time variation by modelling betas as parametric functions of conditioning variables (e.g., [Connor et al. \(2012\)](#); [Gagliardini et al. \(2016\)](#)) while another strand adopts nonparametric methods (e.g., [Ang and Kristensen \(2012\)](#)).

ric procedures. Importantly, my approach treats the size of the estimation window (equivalently, the number of observations in the estimation sample) as a fixed, possibly small, constant, thereby accounting for the scarcity of local information. I further show that the proposed estimator substantially reduces risk relative to standard estimators that rely solely on return observations, particularly in small samples. These properties make the approach especially well suited to high-frequency event studies, which require highly local estimates for identification (e.g., [Bollerslev et al. \(2018\)](#); [Nakamura and Steinsson \(2018\)](#)).²

Having established the candlestick-based estimator, I further develop a formal inference procedure on spot betas using the proposed estimator. While recent studies have investigated candlestick-based inference on spot volatility functionals (e.g., [Li et al. \(2024\)](#); [Bollerslev et al. \(2024a\)](#)), inference methods for spot betas remain underexplored. To address this gap, I propose a hypothesis test and associated test statistic. I show that this test statistic can be approximated by a limiting variable whose distribution under the null hypothesis can be characterized via simulations. Accordingly, I compute the critical values by simulating the quantiles of that limiting variable. The resulting procedure delivers a test with asymptotically correct size. A finite-sample simulation study further demonstrates that the test exhibits reasonable size control and achieves substantially higher power than its return-based counterpart ([Bollerslev et al. \(2024b\)](#)).

Finally, I demonstrate the practical value of the proposed framework through an empirical application studying the market neutrality of Bitcoin. Cryptocurrencies are often framed as “digital gold” by their advocates, suggesting potential hedging properties. Similarly, [Liu and Tsyvinski \(2021\)](#) find limited evidence of systematic exposure of crypto assets to market risk. Motivated by these observations, I test the null hypothesis of market neutrality (or zero market beta) for Bitcoin using my candlestick-based estimator and test. Specifically, I use high-frequency candlestick observations of the iShares Bitcoin ETF (IBIT) and the SPDR S&P 500 ETF (SPY). The analysis uncovers striking patterns with potential implications for risk management. I reject the null approximately 35% of the time, revealing pricing dynamics that differ markedly from those documented in prior work, such as [Liu and Tsyvinski \(2021\)](#). Notably, rejection rates peak around August and September 2024, coinciding with elevated financial market volatility. This is precisely when the diversification and hedging benefits of Bitcoin

²Specifically, [Nakamura and Steinsson \(2018\)](#) exploit high-frequency bond returns within a short window around FOMC announcements to identify the effects of monetary policy shocks, while [Bollerslev et al. \(2018\)](#) estimate investor disagreement using local jump regressions around news announcements.

would be most valuable.

In an attempt to improve the estimation, one could alternatively employ tick-level data which provides the most granular information, as it records every single transaction in the market usually at ultra-high frequencies (e.g., milliseconds). However, using tick-level data faces important limitations. First, this data is accessible only to well-resourced researchers since it requires costly subscriptions through commercial providers such as NYSE Trade and Quotes (TAQ) or TickData. Furthermore, prices sampled at such ultra-high frequencies are invariably contaminated by market microstructure noise, which necessitates imposing additional modeling assumptions on the noise structure to obtain enhanced estimates, see for example [Diebold and Strasser \(2013\)](#). By contrast, candlestick data is widely accessible through many different public sources at “not-too-fine” frequencies (e.g., 1-minute or 5-minute), which naturally guards against the impact of microstructure noise. From this perspective, my paper adopts a more practical and accessible approach for improving spot regression estimation.³

This paper contributes to multiple strands of the literature. From a technical perspective, this paper is closely related to the high-frequency econometrics literature on beta estimation, see for example [Barndorff-Nielsen and Shephard \(2004a\)](#); [Mykland and Zhang \(2009\)](#); [Li et al. \(2017\)](#); [Aït-Sahalia et al. \(2020\)](#); [Bollerslev et al. \(2024b\)](#) among many others. These papers develop estimators primarily based on high-frequency returns from open and close prices. My work complements this literature by introducing a new class of estimators that leverage candlestick observations, which expands the information set of high-frequency intervals with high and low prices.

My work is inspired by the range-based volatility estimation literature. Starting from seminal papers by [Garman and Klass \(1980\)](#) and [Parkinson \(1980\)](#), this literature highlights remarkable efficiency gains in estimating variances by extracting more information from candlestick prices. Notably, [Christensen and Podolskij \(2007\)](#) introduces the realized-range estimator for integrated variance, constructed by high-frequency ranges. Later, it is extended to be robust to microstructure noise ([Martens and Van Dijk \(2007\)](#); [Christensen et al. \(2009\)](#)); jumps ([Christensen and Podolskij \(2012\)](#)) and drifts ([Li et al. \(2025\)](#)). More recently, several papers study spot volatility estimation using candlesticks, see for example [Li et al. \(2024\)](#), [Bollerslev et al. \(2024a\)](#)

³Within the high-frequency econometrics literature, researchers have developed models for market microstructure noise; see, for example, [Hayashi and Yoshida \(2005\)](#); [Aït-Sahalia et al. \(2010\)](#); [Christensen et al. \(2010\)](#); [Barndorff-Nielsen et al. \(2011\)](#). While such approaches may be relevant in this context, their incorporation into the present framework is not straightforward. I leave a careful investigation of these issues for future research.

and [Bollerslev et al. \(2025\)](#). The main focus of these papers is to derive optimal estimation and inference frameworks for spot volatility and its functionals. This paper complements this literature by extending these ideas to multivariate settings which allows us to study dependence structure between assets with this data.

This paper is perhaps most closely related to [Bollerslev et al. \(2024a\)](#) which develops a decision-theoretic framework to construct optimal estimators for volatility functionals. In their approach, the risk function is asymptotically approximated using pivotal random variables, and the asymptotic risk is minimized through Monte Carlo simulations. However, in multivariate settings, the limiting distributions are generally non-pivotal, which makes direct minimization of a traditional risk function infeasible. To overcome this challenge, I integrate the asymptotic risk to obtain an average risk function which is pivotal, allowing me to determine the optimal estimators by minimizing this objective. Importantly, this approach delivers a feasible solution to an otherwise intractable problem while delivering superior efficiency in both estimation and inference.

Prior to this study, [Brandt and Diebold \(2006\)](#) and [Bannouh et al. \(2009\)](#) introduced range-based covariance estimators that exploit triangular no-arbitrage conditions in currency markets. While effective in that setting, this approach is generally not applicable to equity markets. Similarly, [Rogers and Zhou \(2008\)](#) estimated the correlation coefficient of a bivariate Brownian motion using open, high, low and close price observations. However, their analysis is conducted under a constant volatility framework. In contrast, this paper accommodates a more general Itô semimartingale environment and provides a feasible inference procedure with candlesticks which is not considered in the abovementioned papers.

The remainder of the paper is organized as follows. [Section 2](#) introduces the theoretical environment and notation. [Section 3](#) develops candlestick-based estimation framework. [Section 4](#) discusses the inference procedure. [Section 5](#) presents simulation results, and [Section 6](#) provides the empirical application. Finally, [Section 7](#) concludes. Additional technical details and all proofs are provided in the appendices.

2 Setup and Notation

2.1 The Price Process

Consider the bivariate log-price process $\mathbf{X}_t = [X_{1,t}, X_{2,t}]^\top$, observed at time t . Suppose that \mathbf{X}_t evolves as an Itô semimartingale on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$,

according to

$$d\mathbf{X}_t = \mathbf{b}_t dt + \boldsymbol{\sigma}_t d\mathbf{W}_t, \quad (1)$$

where \mathbf{b}_t is the drift process, $\boldsymbol{\sigma}_t$ is stochastic volatility matrix and $\mathbf{W}_t \equiv [W_{1,t}, W_{2,t}]^\top$ is a bivariate standard Brownian motion.

The spot covariance matrix, \mathbf{c}_t , and associated correlation matrix, $\boldsymbol{\rho}_t$, at time t are defined as:

$$\mathbf{c}_t \equiv \boldsymbol{\sigma}_t \boldsymbol{\sigma}_t^\top = \begin{pmatrix} c_{11,t} & c_{12,t} \\ c_{21,t} & c_{22,t} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\rho}_t \equiv \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix}, \quad (2)$$

where $\rho_t = c_{12,t} / \sqrt{c_{11,t} c_{22,t}}$ is the spot correlation between $X_{1,t}$ and $X_{2,t}$.

The Itô-semimartingale representation in Equation (1) can be motivated by no-arbitrage conditions and therefore serves as a workhorse framework in the continuous-time finance literature, see for example [Back \(2010\)](#). Consequently, it has become the fundamental model for analyzing high-frequency asset prices; see [Jacod and Protter \(2012\)](#) and [Aït-Sahalia and Jacod \(2014\)](#) for further discussions. For simplicity, I exclude jumps from the price process.⁴

I further tailor the price process \mathbf{X}_t to the following regression representation:

$$\begin{aligned} dX_{1,t} &= b_{1,t} + \nu_t^{1/2} dW_{1,t} \\ dX_{2,t} &= b_{2,t} + \beta_t dX_{1,t} + \varsigma_t^{1/2} dW_{2,t}. \end{aligned} \quad (3)$$

This is also equivalent to imposing the following spot covariance structure:

$$\mathbf{c}_t = \begin{pmatrix} \nu_t & \beta_t \nu_t \\ \beta_t \nu_t & \beta_t^2 \nu_t + \varsigma_t \end{pmatrix} \quad \text{and} \quad \boldsymbol{\sigma}_t = \begin{pmatrix} \nu_t^{1/2} & 0 \\ \beta_t \nu_t^{1/2} & \varsigma_t^{1/2} \end{pmatrix}. \quad (4)$$

Through the lens of factor models in asset pricing literature (e.g., [Sharpe \(1964\)](#); [Lintner \(1965\)](#); [Fama and French \(1992\)](#)), one can consider the first asset as the market portfolio and the second asset as a risky asset. Then, ν_t and ς_t refer to the market and idiosyncratic variances, respectively. Finally, β_t denotes the market beta of the risky asset which is the main object of interest in this paper.

The above representation implies that the spot beta, idiosyncratic variance and market variance can be identified directly from the spot covariance matrix \mathbf{c}_t as follows:

$$\beta_t = \frac{c_{12,t}}{c_{11,t}}, \quad \varsigma_t = c_{22,t} - \frac{c_{12,t}^2}{c_{11,t}}, \quad \text{and} \quad \nu_t = c_{11,t}. \quad (5)$$

⁴Standard jump-robust techniques such as truncation ([Mancini \(2009\)](#)) or bi-power variation ([Barndorff-Nielsen and Shephard \(2004b\)](#)) methods can in principle be adapted to this paper to explicitly account for price discontinuities.

Throughout the paper, $(\beta_t, \varsigma_t, \nu_t)$ are referred to as the spot regression parameters.

2.2 Observation Scheme and Candlestick Returns

The price process \mathbf{X}_t is assumed to be sampled on a regular time grid, $\{i\Delta_n : i = 0, 1, \dots, n\}$ over a fixed time span $[0, T]$. Here, $\Delta_n = T/n$ refers to the sampling frequency and n is the number of observations, an integer. High-frequency intervals are denoted by $\mathcal{T}_i \equiv [(i-1)\Delta_n, i\Delta_n]$ for each $i \in \{1, \dots, n\}$. Following the standard practice in the high-frequency financial econometrics literature, I consider an in-fill asymptotic framework where $\Delta_n \rightarrow 0$ asymptotically, see for example [Aït-Sahalia and Jacod \(2014\)](#).⁵

A typical candlestick over the interval \mathcal{T}_i consists of four prices:

$$\mathbf{X}_{(i-1)\Delta_n}, \sup_{t \in \mathcal{T}_i} \mathbf{X}_t, \inf_{t \in \mathcal{T}_i} \mathbf{X}_t, \mathbf{X}_{i\Delta_n}$$

which are called the open, high, low, and close prices, respectively. From these prices, one can construct the following normalized returns:

$$\begin{aligned} \mathbf{r}_i &\equiv \frac{\mathbf{X}_{i\Delta_n} - \mathbf{X}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}}, \\ \mathbf{h}_i &\equiv \frac{\sup_{t \in \mathcal{T}_i} \mathbf{X}_t - \mathbf{X}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}}, \\ \mathbf{\ell}_i &\equiv \frac{\inf_{t \in \mathcal{T}_i} \mathbf{X}_t - \mathbf{X}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \end{aligned} \tag{6}$$

where sup and inf operators are applied element-wise. The first line defines the standard high-frequency (open-to-close) return, commonly employed in the literature. The second and third lines indicate the high-open and low-open returns, respectively. These returns stand as a new source of information in this framework. Indeed, the bundle $(\mathbf{r}_i, \mathbf{h}_i, \mathbf{\ell}_i)$ summarizes the price dynamics within the interval \mathcal{T}_i through the lens of candlesticks.

I also define two variables:

$$\begin{aligned} \mathbf{w}_i &\equiv \mathbf{h}_i - \mathbf{\ell}_i, \\ \mathbf{a}_i &\equiv \mathbf{h}_i + \mathbf{\ell}_i - \mathbf{r}_i. \end{aligned} \tag{7}$$

where \mathbf{w}_i and \mathbf{a}_i refer to the range and asymmetry of the candlestick returns, re-

⁵The choice of Δ_n is usually guided by volatility signature plots introduced in [Andersen et al. \(2000\)](#). Following the standard practice in the high-frequency econometrics literature, I adopt moderate sampling frequencies like $\Delta_n = 1, 5, 10$ -min in my practical implementations, thereby avoiding ultra high-frequency observations that could be contaminated by microstructure effects.

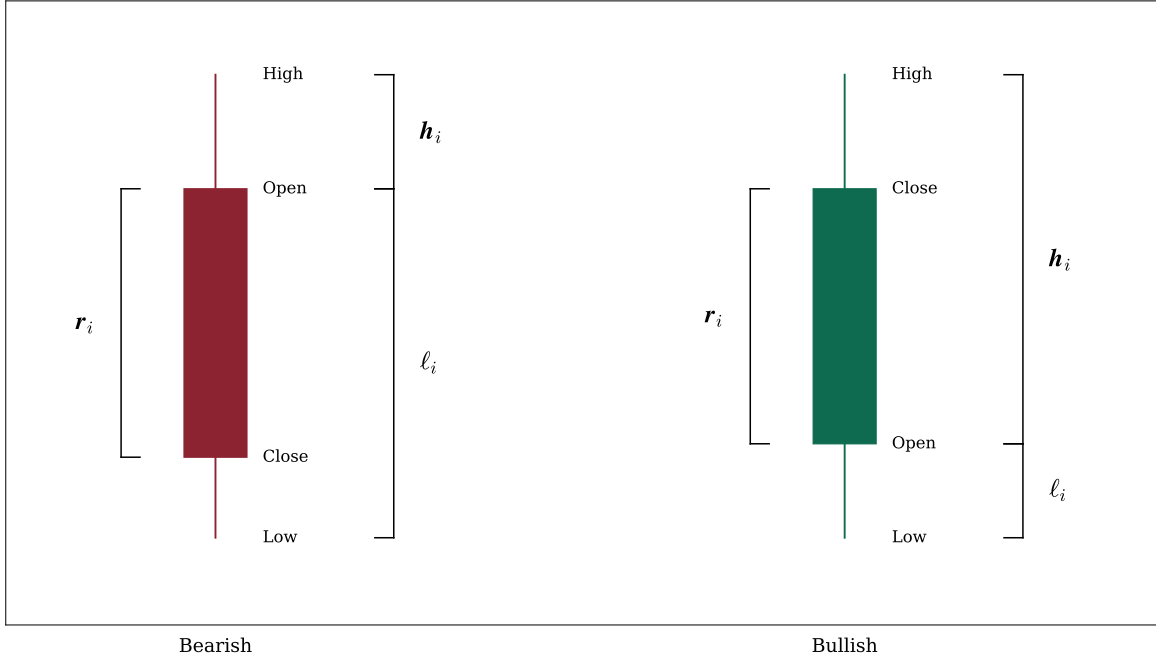


Figure 1: **Illustration of candlestick returns:** The figure shows examples of bearish (left) and bullish (right) candlesticks, illustrating the open, high, low, and close prices. Also it highlights the open-close return r_i , high return h_i , and low return ℓ_i as defined in Equation (6).

spectively. One can show that $(r_i, \mathbf{a}_i, \mathbf{w}_i)$ is an invertible linear transformation of $(r_i, \mathbf{h}_i, \ell_i)$, implying that both vectors carry the same information. Recent papers on candlestick-based volatility estimation (e.g., Li et al. (2024), Bollerslev et al. (2024a) and Bollerslev et al. (2025)) work with $(r_i, \mathbf{a}_i, \mathbf{w}_i)$. Thus, to be consistent with this literature and facilitate the analysis, I also work with $(r_i, \mathbf{a}_i, \mathbf{w}_i)$ for the rest of the paper. However, the proposed methodology can be easily adapted to the original set of returns $(r_i, \mathbf{h}_i, \ell_i)$ without changing any of the main results.

To build intuition, I present the graphical representation of the candlestick returns on a typical candlestick chart in Figure 1. As shown in this figure, the range \mathbf{w}_i is shown as the vertical distance between the high and low prices, while the return r_i reflects the length of the thick body. The asymmetry \mathbf{a}_i indicates the position of the returns (or the thick body) within that range. For example, if $\mathbf{a}_i = 0$ then the thick body of the candlestick is exactly centered between the high and low prices. On the other hand, if $\mathbf{a}_i > 0$ then the body is skewed towards the low price, and vice versa.

3 Candlestick-based Estimation

The primary objective of this paper is to construct estimators for spot regression parameters that efficiently exploit the information contained in candlestick features. To this end, I first show how to construct spot covariance estimators using candlestick returns, and then obtain spot regression estimates exploiting the relation in Equation (5).

3.1 A Class of Spot Covariance Estimators

A standard estimator for the spot covariance matrix \mathbf{c}_t is given by the local average of the outer products of high-frequency (open-to-close) returns. For any $t \in [0, T]$, this can be formally expressed as:

$$\hat{\mathbf{c}}_t = \frac{1}{k} \sum_{i \in \mathcal{I}_{n,t}} \mathbf{r}_i \mathbf{r}_i^\top. \quad (8)$$

where $\mathcal{I}_{n,t} = \{\lceil \frac{t}{\Delta_n} \rceil + 1, \dots, \lceil \frac{t}{\Delta_n} \rceil + k\}$ is the local estimation window and k stands for the size of that window.⁶ Correspondingly, the spot quantities in the regression representation in Equation (3) can be estimated as follows:

$$\hat{\beta}_{n,t} = \frac{\hat{c}_{12,t}}{\hat{c}_{11,t}}, \quad \hat{\varsigma}_{n,t} = \hat{c}_{22,t} - \frac{\hat{c}_{12,t}^2}{\hat{c}_{11,t}}, \quad \text{and} \quad \hat{\nu}_{n,t} = \hat{c}_{11,t}. \quad (9)$$

The estimator in Equation (8) is well-studied in the high-frequency econometrics literature; see, among others, [Fan and Wang \(2008\)](#). While the conventional estimators let $k = k_n$ grow with sample size n , analogous to bandwidth parameter in classic kernel nonparametrics, I treat k as a fixed constant, not dependent on n . Accordingly, throughout this paper I adopt fixed- k asymptotics developed by [Bollerslev et al. \(2021\)](#). This approach was originally proposed for inference on spot volatilities and subsequently extended to spot beta inference by [Bollerslev et al. \(2024b\)](#). The framework yields inference that remains valid for a fixed, potentially small, value of k , thereby directly capturing the local nature of spot estimates rather than relying on asymptotic approximations that require k to diverge.⁷

⁶Here, $\lceil \cdot \rceil$ denotes the ceiling function, which maps a real number to the smallest following integer. Thus, $\mathcal{I}_{n,t}$ can be interpreted as the right-sided local window that contains k observations after time t . Alternatively, left-sided or symmetric windows can be considered without changing the main results of the paper.

⁷Fixed- k asymptotics build on coupling (or strong approximation) arguments; see [Jacod et al. \(2021\)](#). The key idea is that, over short intervals, an Itô semimartingale can be locally approximated by a scaled Brownian motion. This feature allows for determining the (approximate) finite-sample distribution of the estimator under shrinking Δ_n (or diverging n). [Bollerslev et al. \(2021\)](#) show

By construction, Equation (8) relies solely on open-to-close returns, discarding potentially valuable information embedded in other candlestick returns. Thus, I introduce a flexible class of estimators for \mathbf{c}_t that incorporates all candlestick returns $(\mathbf{r}_i, \mathbf{a}_i, \mathbf{w}_i)$ and combines them through a weighted sum of quadratic forms over a local window.

Formally, the estimator is defined as:

$$\begin{aligned} \widehat{\mathbf{c}}_{n,t}(\lambda) = \frac{1}{k} \sum_{i \in \mathcal{I}_{n,t}} \left\{ \lambda_1 \mathbf{r}_i \mathbf{r}_i^\top + \lambda_2 \mathbf{a}_i \mathbf{a}_i^\top + \lambda_3 \mathbf{w}_i \mathbf{w}_i^\top \right. \\ \left. + \lambda_4 (\mathbf{r}_i \mathbf{a}_i^\top + \mathbf{a}_i \mathbf{r}_i^\top) + \lambda_5 (\mathbf{r}_i \mathbf{w}_i^\top + \mathbf{w}_i \mathbf{r}_i^\top) + \lambda_6 (\mathbf{a}_i \mathbf{w}_i^\top + \mathbf{w}_i \mathbf{a}_i^\top) \right\} \end{aligned} \quad (10)$$

where $\lambda \equiv (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) \in \mathbb{R}^6$ denote the weights assigned to each component. This estimator can be compactly expressed in matrix form as follows:

$$\widehat{\mathbf{c}}_{n,t}(\lambda) = \frac{1}{k} \sum_{i \in \mathcal{I}_{n,t}} \mathbf{z}_i \Lambda \mathbf{z}_i^\top, \quad \text{where } \mathbf{z}_i = \begin{pmatrix} \mathbf{r}_i & \mathbf{a}_i & \mathbf{w}_i \end{pmatrix} \quad \text{and} \quad \Lambda = \begin{pmatrix} \lambda_1 & \lambda_4 & \lambda_5 \\ \lambda_4 & \lambda_2 & \lambda_6 \\ \lambda_5 & \lambda_6 & \lambda_3 \end{pmatrix}. \quad (11)$$

The resulting estimator is always symmetric and positive semi-definite whenever Λ is positive semi-definite. The choice of λ will be discussed in subsequent sections in detail. Naturally, I propose the following estimates for the spot quantities in the regression representation:

$$\widehat{\beta}_{n,t}(\lambda) = \frac{\widehat{c}_{12,t}(\lambda)}{\widehat{c}_{11,t}(\lambda)}, \quad \widehat{\varsigma}_{n,t}(\lambda) = \widehat{c}_{22,t}(\lambda) - \frac{\widehat{c}_{12,t}^2(\lambda)}{\widehat{c}_{11,t}(\lambda)}, \quad \text{and} \quad \widehat{\nu}_{n,t}(\lambda) = \widehat{c}_{11,t}(\lambda). \quad (12)$$

The proposed class of estimators is highly flexible. For example, setting $\lambda = (1, 0, 0)$ recovers the classic spot covariance and beta estimator defined in Equation (8) and (9). Furthermore, the estimator can be viewed as a natural multivariate extension of the univariate candlestick-based volatility estimators, and hence my approach aims to generalize these efficiency gains to the multivariate setting.

Remark. *At a first glance, it might be surprising that moments of high and low returns are included in the covariance estimator just like the regular returns. In this regard, Rogers and Shepp (2006) show that the correlations of range and asymmetry variables are smooth and nonlinear functions of the correlation coefficient of the underlying price*

that fixed- k asymptotics deliver confidence intervals with more accurate coverage than conventional large- k approximations, especially in small samples, where the latter often suffer from nontrivial size distortions.

process. Thus this suggests that the cross moments of candlestick returns may contain valuable information about the underlying correlation structure and thus motivates my approach to incorporate them into the estimation process.

3.2 Determining the Weights via Risk Minimization

The estimator in Equation (10), and by extension Equation (12), depends crucially on the choice of weights λ . Following a decision-theoretic approach, I determine the weighting scheme by minimizing a well-defined risk function. Specifically, the risk of an estimator $\widehat{\mathbf{c}}_{n,t}(\lambda)$ is defined as the conditional expectation of its loss:

$$R(\lambda; \mathbf{c}_t) \equiv \mathbb{E} [L(\lambda; \mathbf{c}_t) \mid \mathcal{F}_t]. \quad (13)$$

where $L(\lambda; \mathbf{c}_t)$ is a loss function that measures the estimation error of $\widehat{\mathbf{c}}_{n,t}(\lambda)$ relative to the true spot covariance \mathbf{c}_t . I employ the following quadratic loss function:

$$L(\lambda; \mathbf{c}_t) = \left\| \boldsymbol{\sigma}_t^{-1} \widehat{\mathbf{c}}_{n,t}(\lambda) \boldsymbol{\sigma}_t^{-1\top} - \mathbf{I} \right\|^2 \quad (14)$$

where $\| \cdot \|$ denotes the Frobenius norm and \mathbf{I} is the identity matrix. The term $\boldsymbol{\sigma}_t^{-1} \widehat{\mathbf{c}}_{n,t} \boldsymbol{\sigma}_t^{-1\top}$ indicates the multiplicative estimation error of the spot covariance matrix relative to the identity, naturally inducing a scale-invariant loss function. This property is particularly desirable, as covariance matrices represent scale parameters. In the univariate setting, the loss reduces to

$$\left(\frac{\widehat{c}_{n,t}}{c_t} - 1 \right)^2,$$

which corresponds to the relative squared error that has been previously analyzed in the context of volatility estimation with candlestick data by Li et al. (2024).

Remark. *The determination of the weights λ directly targets the full covariance matrix \mathbf{c}_t rather than the spot beta β_t which is the main focus in a regression context. A reader might therefore ask why covariance-optimal weights is considered here. I deliberately adopt this approach because targeting covariance matrix allows for unified treatment of all spot quantities in the regression representation in Equation (3) as they are nonlinear functions of the covariance matrix. In other words, this approach yields a single set of weights that can be used for estimating all spot quantities simultaneously. Furthermore, in Appendix A.1, I show that this loss function can be decomposed into three components that can be directly associated with the estimation errors of the spot re-*

gression parameters. This suggests that minimizing the overall loss effectively balances the trade-offs between the estimation errors of different parameters. Indeed, simulation results in Section 5 show that the proposed approach leads to more reliable inference for betas, serving as ex-post supportive evidence. In addition, from a practical perspective, this approach yields a closed-form solution for the weights, which greatly simplifies the computation.

3.2.1 Asymptotic Approximation for the Risk Function

$R(\lambda; \mathbf{c}_t)$ indicates the exact finite sample risk of an estimator. For a given λ , the risk depends on the joint distribution of $(\mathbf{r}_i, \mathbf{a}_i, \mathbf{w}_i)$ which in turn is determined by the law of $(\mathbf{b}, \boldsymbol{\sigma}, \mathbf{W})$. Since this law is unknown, the estimation of λ with direct minimization of $R(\lambda; \boldsymbol{\sigma}_t)$ is infeasible in practice. To circumvent this issue, I first establish an asymptotic approximation of the multiplicative estimation error $\boldsymbol{\sigma}_t^{-1} \widehat{\mathbf{c}}_{n,t} \boldsymbol{\sigma}_t^{-1\top}$ using coupling arguments (see, e.g., Jacod et al. (2021)). This characterization is subsequently extended to the loss function and, in turn, to the risk. The resulting asymptotic risk functional is then used to determine the optimal weights λ .

The assumption below gathers a set of regularity conditions, required for the asymptotic approximation results.

Assumption 1. *Suppose that \mathbf{X}_t has the form in Equation (1) and there exists a sequence $(T_m)_{m \geq 1}$ of stopping times increasing to infinity and the following conditions hold for each $m \geq 1$:*

$$(i) \quad \|\mathbf{b}_t\| + \|\boldsymbol{\sigma}_t\| + \|\boldsymbol{\sigma}_t^{-1}\| \leq K_m \text{ for some constant } K_m \text{ for all } t \in [0, T_m];$$

$$(ii) \quad \mathbb{E} [\|\boldsymbol{\sigma}_{t \wedge T_m} - \boldsymbol{\sigma}_{s \wedge T_m}\|^2] \leq K_m |t - s| \text{ for all } t, s \in [0, T_m].$$

These conditions are quite standard in the high-frequency econometrics literature; see, for example, Jacod and Protter (2012) for further details. The first part of the Assumption 1 implies local boundedness of the drift, volatility, and inverse volatility processes while the second part ensures a degree of smoothness in the volatility process, specifically called as locally 1/2-Hölder continuous. These assumptions are sufficiently general to accommodate a wide range of volatility dynamics, including volatility jumps, leverage effects, and intraday seasonality. For example, they are satisfied if the volatility process is itself an Itô semimartingale or long-memory process driven by fractional Brownian motion.

The following proposition describes the key approximation result for the estimation error $\boldsymbol{\sigma}_t^{-1} \widehat{\mathbf{c}}_{n,t} \boldsymbol{\sigma}_t^{-1\top}$. The proof is deferred to Appendix A.2.

Proposition 1. *Suppose that Assumption 1 holds. Fix any $t \in [0, T]$. For any $k \geq 1$ and λ , the following holds as $\Delta_n \rightarrow 0$:*

$$\|\boldsymbol{\sigma}_t^{-1} \widehat{\mathbf{c}}_{n,t}(\lambda) \boldsymbol{\sigma}_t^{-\top} - U_{n,t}(\lambda)\| = o_p(1) \quad (15)$$

where

$$U_{n,t}(\lambda) = \frac{1}{k} \sum_{i \in \mathcal{I}_{n,t}} \left\{ \lambda_1 \boldsymbol{\zeta}_{i,r} \boldsymbol{\zeta}_{i,r}^\top + \lambda_2 \boldsymbol{\zeta}_{i,a} \boldsymbol{\zeta}_{i,a}^\top + \lambda_3 \boldsymbol{\zeta}_{i,w} \boldsymbol{\zeta}_{i,w}^\top \right. \\ \left. + \lambda_4 (\boldsymbol{\zeta}_{i,r} \boldsymbol{\zeta}_{i,a}^\top + \boldsymbol{\zeta}_{i,a} \boldsymbol{\zeta}_{i,r}^\top) + \lambda_5 (\boldsymbol{\zeta}_{i,r} \boldsymbol{\zeta}_{i,w}^\top + \boldsymbol{\zeta}_{i,w} \boldsymbol{\zeta}_{i,r}^\top) + \lambda_6 (\boldsymbol{\zeta}_{i,a} \boldsymbol{\zeta}_{i,w}^\top + \boldsymbol{\zeta}_{i,w} \boldsymbol{\zeta}_{i,a}^\top) \right\} \quad (16)$$

and, for any $i \in \mathcal{I}_{n,t}$,

$$\begin{aligned} \boldsymbol{\zeta}_{i,r} &\equiv \frac{\mathbf{W}_{i\Delta_n} - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \\ \boldsymbol{\zeta}_{i,a} &\equiv \boldsymbol{\varrho}_t^{-1} \sup_{\tau \in \mathcal{T}_i} \boldsymbol{\varrho}_t \left(\frac{\mathbf{W}_{\tau} - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) + \boldsymbol{\varrho}_t^{-1} \inf_{\tau \in \mathcal{T}_i} \boldsymbol{\varrho}_t \left(\frac{\mathbf{W}_{\tau} - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) - \left(\frac{\mathbf{W}_{i\Delta_n} - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) \\ \boldsymbol{\zeta}_{i,w} &\equiv \boldsymbol{\varrho}_t^{-1} \sup_{\tau \in \mathcal{T}_i} \boldsymbol{\varrho}_t \left(\frac{\mathbf{W}_{\tau} - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) - \boldsymbol{\varrho}_t^{-1} \inf_{\tau \in \mathcal{T}_i} \boldsymbol{\varrho}_t \left(\frac{\mathbf{W}_{\tau} - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) \end{aligned} \quad (17)$$

with $\boldsymbol{\varrho}_t$ being the square root of spot correlation matrix $\boldsymbol{\rho}_t$, i.e., $\boldsymbol{\rho}_t = \boldsymbol{\varrho}_t \boldsymbol{\varrho}_t^\top$. All inf and sup operators are applied element-wise.

Proposition 1 establishes that the multiplicative estimation error of the spot covariance estimator $\widehat{\mathbf{c}}_{n,t}(\lambda)$ is asymptotically approximated by the random matrix $U_{n,t}$ in probability as $\Delta_n \rightarrow 0$. The structure of $U_{n,t}$ mirrors that of the original estimator, with the candlestick returns replaced by the variables $\boldsymbol{\zeta}_i$ s. These variables are functions of Brownian motion \mathbf{W} and square root of spot correlation matrix $\boldsymbol{\varrho}_t$. In addition, the approximation error ($o_p(1)$ term) captures nonparametric biases arising from stochastic volatility and drift. Put differently, in the limiting case of constant volatility and vanishing drift, the relationship in Proposition 1 holds exactly rather than approximately.

Using Proposition 1 and the conditions stated therein, analogous approximations apply to the loss and risk functions. Since the loss function is continuous, the continuous mapping theorem implies

$$L(\lambda; \mathbf{c}_t) = \|U_{n,t}(\lambda) - \mathbf{I}\|^2 + o_p(1). \quad (18)$$

This relation can similarly be extended to the risk function by taking the conditional

expectation of both sides given \mathcal{F}_t :

$$R(\lambda; \mathbf{c}_t) = \underbrace{\mathbb{E} [\|U_{n,t}(\lambda) - \mathbf{I}\|^2 \mid \mathcal{F}_t]}_{\tilde{R}(\lambda; \boldsymbol{\rho}_t)} + o_p(1) \quad (19)$$

where $\tilde{R}(\lambda; \boldsymbol{\rho}_t)$ is the asymptotic risk of the estimator $\hat{\mathbf{c}}_{n,t}$.⁸ Importantly, I switch to notation $\tilde{R}(\lambda; \boldsymbol{\rho}_t)$ to emphasize that the asymptotic risk depends on the spot correlation matrix $\boldsymbol{\rho}_t$ rather than the spot covariance matrix \mathbf{c}_t . This is because the risk formula involves a term, $U_{n,t}$, which depends on $\boldsymbol{\varrho}_t$, i.e., the square root of the spot correlation matrix. The use of the Frobenius norm in the final calculation effectively considers the product $\boldsymbol{\varrho}_t \boldsymbol{\varrho}_t^\top$, making the risk solely dependent on $\boldsymbol{\rho}_t$.

As a result, direct minimization of the asymptotic risk $\tilde{R}(\lambda; \boldsymbol{\rho}_t)$ with respect to λ to obtain the optimal weights is infeasible in practice, as $\boldsymbol{\rho}_t$ is unknown to the researcher. In the univariate case, by contrast, the term $U_{n,t}$ does not depend on any unknown parameters and thus the distribution of $U_{n,t}$ is pivotal. This allows for direct computation of the asymptotic risk and thus the derivation of optimal weights via direct risk minimization, see Theorem 1 in Li et al. (2024); Bollerslev et al. (2024a). This is because the scaling by $\boldsymbol{\varrho}_t$ and its inverse cancels, as they become scalars, in Equation (17). Hence, one can characterize the distribution of $U_{n,t}$ through simulations of Brownian motion functionals. This simplification does not hold in the multivariate setting, and thus the estimation problem studied in this paper requires more intricate treatment than the univariate case.

3.2.2 Average Risk

To address the complication discussed above, I propose an alternative approach for selecting the weights λ based on the average risk $\bar{R}(\lambda)$, defined as the integrated asymptotic risk $\tilde{R}(\lambda; \boldsymbol{\rho}_t)$ over the parameter space of $\boldsymbol{\rho}_t$. This approach effectively marginalizes over the unknown correlation structure, thereby enabling the derivation of “optimal” weights without requiring explicit knowledge of $\boldsymbol{\rho}_t$.⁹ This perspective is directly inspired by the statistical decision theory literature, see for example Lehmann and Casella (2006), which advocates the use of integrated risk functions to accommodate nuisance parameters.

Let \mathcal{P} denote the parameter space of positive semidefinite correlation matrices.

⁸This step requires an additional uniform integrability condition to interchange the limit and expectation operators. This condition is satisfied under the assumptions stated in Assumption 1.

⁹Here optimality is used in the context of minimizing the average risk which is specifically defined in Equation (20).

Define the average risk $\bar{R}(\lambda)$ as the integrated asymptotic risk

$$\bar{R}(\lambda) \equiv \int_{\mathcal{P}} \tilde{R}(\lambda; \boldsymbol{\rho}_t) d\boldsymbol{\rho}_t. \quad (20)$$

Put differently, $\bar{R}(\lambda)$ is averaging over the interval $(-1, 1)$ as the setup has two assets and single correlation parameter.

For any fixed λ , the mapping $\boldsymbol{\rho} \mapsto \tilde{R}(\lambda; \boldsymbol{\rho})$ can be computed via Monte Carlo simulation. Specifically, conditional on $\boldsymbol{\rho}_t$, the distribution of $U_{n,t}$ is fully characterized by functionals of Brownian motion defined. Thus, the risk $\tilde{R}(\lambda; \boldsymbol{\rho}_t)$ can be computed by simulating these functionals and evaluating the corresponding loss function. Consequently, the average risk $\bar{R}(\lambda)$ depends only on λ , and the optimal weights can be obtained by solving

$$\lambda^* = \arg \min_{\lambda} \bar{R}(\lambda). \quad (21)$$

The solution to this optimization problem is available in closed form and resembles a least-squares formula:

$$\lambda^* = \left(\int_{\mathcal{P}} \mathbb{E}[\boldsymbol{\Pi}^\top \boldsymbol{\Pi} | \mathcal{F}_t] d\boldsymbol{\rho}_t \right)^{-1} \left(\int_{\mathcal{P}} \mathbb{E}[\boldsymbol{\Pi}^\top \mathbf{y} | \mathcal{F}_t] d\boldsymbol{\rho}_t \right) \quad (22)$$

where $\boldsymbol{\Pi}$ is stacking the vectorized version of outer products of limiting variables and \mathbf{y} is the vectorized version of the identity matrix \mathbf{I} . More explicitly, $\mathbf{y} = \text{vech}(\mathbf{I})$ is 3×1 vector and

$$\begin{aligned} \boldsymbol{\Pi} \equiv & \left[\frac{1}{k} \sum_i \text{vech}(\boldsymbol{\zeta}_{i,r} \boldsymbol{\zeta}_{i,r}^\top) \quad \frac{1}{k} \sum_i \text{vech}(\boldsymbol{\zeta}_{i,a} \boldsymbol{\zeta}_{i,a}^\top) \quad \frac{1}{k} \sum_i \text{vech}(\boldsymbol{\zeta}_{i,w} \boldsymbol{\zeta}_{i,w}^\top) \cdots \right. \\ & \left. \frac{1}{k} \sum_i \text{vech}(\boldsymbol{\zeta}_{i,r} \boldsymbol{\zeta}_{i,a}^\top + \boldsymbol{\zeta}_{i,a} \boldsymbol{\zeta}_{i,r}^\top) \quad \frac{1}{k} \sum_i \text{vech}(\boldsymbol{\zeta}_{i,r} \boldsymbol{\zeta}_{i,w}^\top + \boldsymbol{\zeta}_{i,w} \boldsymbol{\zeta}_{i,r}^\top) \quad \frac{1}{k} \sum_i \text{vech}(\boldsymbol{\zeta}_{i,a} \boldsymbol{\zeta}_{i,w}^\top + \boldsymbol{\zeta}_{i,w} \boldsymbol{\zeta}_{i,a}^\top) \right] \end{aligned} \quad (23)$$

is a 3×6 matrix and each column corresponds to the half-vectorized version of the outer products of the limiting variables defined in Equation (17).

The resulting spot covariance estimator can be denoted as $\hat{\mathbf{c}}_{n,t}(\lambda^*)$ and the corresponding estimators for spot beta, idiosyncratic variance, and systematic variance can be obtained by substituting λ^* into Equation (12).

$$\hat{\beta}_{n,t}(\lambda^*) = \frac{\hat{c}_{12,t}(\lambda^*)}{\hat{c}_{11,t}(\lambda^*)}, \quad \hat{\varsigma}_{n,t}(\lambda^*) = \hat{c}_{22,t}(\lambda^*) - \frac{\hat{c}_{12,t}^2(\lambda^*)}{\hat{c}_{11,t}(\lambda^*)}, \quad \text{and} \quad \hat{\nu}_{n,t}(\lambda^*) = \hat{c}_{11,t}(\lambda^*). \quad (24)$$

In practice, the integrals in λ^* formula (Equation (22)) can be computed by simula-

tion. Specifically, one can draw many realizations of $\boldsymbol{\rho}$ uniformly from \mathcal{P} , calculate the corresponding conditional expectations via Monte Carlo simulation, and replace the integrals with sample averages. The estimator is then obtained by substituting these into Equation (22). Importantly, this numerical procedure takes estimation window size k as a given constant, consistent with this paper’s asymptotic framework.

Equation (25) presents the optimal weights λ^* for different values of k . There are several noteworthy observations regarding these results. First, the optimal weights for the first two components, λ_1 and λ_2 , are notably larger than those for the remaining components, suggesting that the open-close returns and asymmetry variables play a more dominant role in the estimates. Surprisingly, the optimal weight for the range variable, λ_3 , is near zero across all settings. The reason is that the range variable loses the sign information of the underlying correlation structure and is heavily biased for negative correlations. Consequently, the estimation process substantially downweights it. On the other hand, the cross terms mechanically receive exact zero weight since $(\mathbf{r}_i, \mathbf{a}_i, \mathbf{w}_i)$ are asymptotically orthogonal to each other.

$$\lambda^* = \begin{cases} (0.438, 1.523, \approx 0, 0, 0, 0)^\top & \text{if } k = 5, \\ (0.488, 1.648, \approx 0, 0, 0, 0)^\top & \text{if } k = 10, \\ (0.526, 1.692, \approx 0, 0, 0, 0)^\top & \text{if } k = 20. \end{cases} \quad (25)$$

Remark. *When defining the average risk, the integration over \mathcal{P} is performed uniformly which also induces a uniform marginal distribution for the correlation parameter. This may appear simplistic and a natural concern can arise regarding this specification. Indeed alternative choices including more flexible distributional assumptions or those informed by the historical data and expert judgment could be certainly considered. Nevertheless, the asymptotic risk comparisons presented in Section 5.1 reveals that this approach performs nearly as well as oracle estimators, leaving limited potential improvement through alternative specifications. Moreover, solving this optimization problem for λ does not inherently depend on uniform integration. As such, the optimal weights can be obtained for any given marginal distribution of $\boldsymbol{\rho}$ following the same computational steps.*

Remark. *An estimator of similar form, a linear combination of $\mathbf{r}_i \mathbf{r}_i^\top$ and $\mathbf{a}_i \mathbf{a}_i^\top$, was previously studied by Rogers and Zhou (2008) for the purpose of estimating the correlation between two Brownian motions. Their analysis assumes that the price process follows a scaled Brownian motion with constant volatility and determines the weights by minimizing the variance of the estimation error, subject to an unbiasedness constraint*

under the assumption of zero correlation. By contrast, my approach is formulated within a general Itô-semimartingale framework which accommodates stochastic volatility and selects the weights by minimizing the risk without imposing restrictions on the correlation level. In addition, the next section develops an inference procedure that enables hypothesis testing for spot betas, whereas [Rogers and Zhou \(2008\)](#) focus exclusively on point estimation. Beyond these methodological differences, the Monte Carlo experiments in [Section 5](#) and the empirical application in [Section 6](#) provide a comprehensive evaluation of candlestick-based estimators in a multivariate setting, which is also absent from the prior literature.

4 Inference on Spot Betas

Spot betas are fundamental objects of interest in asset pricing and risk management, yet inference on these quantities remains an underexplored area in the candlestick-based estimation literature. Existing work has focused on inference for spot volatility functionals (see, e.g., [Li et al. \(2024\)](#); [Bollerslev et al. \(2024a\)](#)). This section fills that gap. Specifically, I develop a hypothesis test for the spot beta, derive its limiting null distribution, and provide a feasible procedure for computing critical values.

4.1 Test Statistic

For a fixed time point $t \in [0, T]$, and hypothesized value $\beta_0 \in \mathbb{R}$, consider testing the null hypothesis

$$H_0 : \beta_t = \beta_0 \quad \text{against} \quad H_1 : \beta_t \neq \beta_0$$

Building on the candlestick-based estimator, I propose the following studentized test statistic:

$$\hat{T}_n(\lambda) = \frac{\sqrt{k-1} \left(\hat{\beta}_{n,t}(\lambda) - \beta_0 \right)}{\sqrt{\hat{\varsigma}_{n,t}(\lambda) / \hat{\nu}_{n,t}(\lambda)}}$$

where $\hat{\beta}_{n,t}(\lambda)$, $\hat{\nu}_{n,t}(\lambda)$, and $\hat{\varsigma}_{n,t}(\lambda)$ are candlestick-based estimates for some λ . Therefore, it encompasses both the return-based statistic of [Bollerslev et al. \(2024b\)](#), recovered at $\lambda = (1, 0, \dots, 0)^\top$, and the candlestick-based statistic associated with the risk-minimizing weights λ^* derived in [Section 3](#).

To study the asymptotic properties of the test statistic, it is useful to establish an asymptotic approximation to $\hat{T}_n(\lambda)$, which can be achieved by exploiting [Proposition 1](#) and the continuous mapping theorem. This result is summarized in the following

proposition. Proof is provided in Appendix A.3. Similar to the estimation problem, fixed- k asymptotics is adopted here.

Proposition 2. *Suppose that the conditions of Proposition 1 hold and \mathbf{X}_t follows Equation (3). Then, for any fixed $k \geq 2$ and λ , the following holds as $\Delta_n \rightarrow 0$:*

$$|\widehat{T}_n(\lambda) - \widetilde{T}_n(\lambda)| = o_p(1) \quad \text{where} \quad \widetilde{T}_n(\lambda) \equiv \frac{-\sqrt{k-1}[U_{n,t}^{-1}]_{12}}{\sqrt{[U_{n,t}^{-1}]_{11}[U_{n,t}^{-1}]_{22} - [U_{n,t}^{-1}]_{12}^2}}.$$

This proposition shows that $\widehat{T}_n(\lambda)$ can be asymptotically approximated by $\widetilde{T}_n(\lambda)$, which is a function of the matrix $U_{n,t}$ defined in Proposition 1. As such, the distributional properties of the limiting variable $\widetilde{T}_n(\lambda)$ can be exploited to compute the critical values for the test. Next, I turn to this problem.

4.2 Critical Values

For a given nominal level $\alpha \in (0, 1)$, critical values for the test can be described formally as follows:

$$\mathbb{P}\left(B_\alpha^- < \widetilde{T}_n(\lambda) < B_\alpha^+ \mid H_0\right) = 1 - \alpha. \quad (26)$$

where constants B_α^+ and B_α^- are the upper and lower critical values for the spot beta test, respectively. The test then rejects the null hypothesis H_0 if $\widehat{T}_n(\lambda)$ falls outside the interval (B_α^-, B_α^+) . Since infinitely many pairs can satisfy this condition, one can select the highest density interval (HDI), which is the narrowest interval containing the desired probability mass. By construction, this procedure delivers a test with asymptotically correct size.

The computation of the critical values (B_α^-, B_α^+) depends on whether the distribution of $\widetilde{T}_n(\lambda)$ under H_0 is pivotal or not. For instance, for the return-based statistic, i.e., when $\lambda = (1, 0, \dots, 0)^\top$, the distribution of $\widetilde{T}_n(\lambda)$ admits a closed-form Student's t distribution with $k - 1$ degrees of freedom as shown by Bollerslev et al. (2024b). For general λ , and in particular for the candlestick-based weights λ^* , $U_{n,t}(\lambda)$ depends on the unknown correlation parameter $\boldsymbol{\rho}_t$ in a non-trivial manner, and thus the distribution of $\widetilde{T}_n(\lambda)$ under H_0 is generally non-pivotal.

As a special case, the pivotal property is restored when $\beta_0 = 0$ in the null hypothesis, consequently its distribution can be simulated directly. This is because under $H_0 : \beta_t = 0$, the covariance matrix \mathbf{c}_t defined in Equation (4) simplifies to a diagonal matrix and consequently the $\boldsymbol{\rho}_t$ terms in the coupling returns provided in Equation (17) cancel out, leaving the limiting variable as a function of the Brownian motion functionals

only. Therefore, the quantiles of $\tilde{T}_n(\lambda)$ can be simulated which in turn allows for feasible inference on spot betas using the proposed estimator. Indeed, this case yields an asymptotically exact size test.

I now consider the general case, i.e. $H_0 : \beta_t = \beta_0$ with $\beta_0 \neq 0$. When the hypothesized value is non-zero, the distribution of $\tilde{T}_n(\lambda)$ is generally non-pivotal, and hence direct simulation is not feasible. To restore the feasibility of inference, I follow a sup-t test approach by constructing critical values that uniformly control size across the parameter space of $\boldsymbol{\rho}_t$.

Specifically, let the pair $(B_\alpha^-(\boldsymbol{\rho}), B_\alpha^+(\boldsymbol{\rho}))$ be the HDI of $\tilde{T}_n(\lambda)$ given some $\boldsymbol{\rho} \in \mathcal{P}$ and significance level α . Define

$$B_\alpha^- = \inf_{\boldsymbol{\rho} \in \mathcal{P}} B_\alpha^-(\boldsymbol{\rho}) \quad \text{and} \quad B_\alpha^+ = \sup_{\boldsymbol{\rho} \in \mathcal{P}} B_\alpha^+(\boldsymbol{\rho}). \quad (27)$$

By construction, (B_α^-, B_α^+) contains every conditional HDI and, consequently, this guarantees that

$$\inf_{\boldsymbol{\rho} \in \mathcal{P}} \mathbb{P}\left(B_\alpha^- < \tilde{T}_n(\lambda) < B_\alpha^+ \mid H_0\right) \geq 1 - \alpha. \quad (28)$$

The cutoffs B_α^- and B_α^+ are the envelopes of the conditional HDIs over \mathcal{P} and therefore deliver uniform asymptotic size control: the test rejects H_0 with probability at most α . This uniformity comes at the cost of conservativeness, but it requires no preliminary estimation step and critical values can be tabulated in a single simulation step. Simulation results in Section 5 indicate that the resulting test is not overly conservative and delivers good power properties.

Table 1 reports the critical values for $H_0 : \beta_t = 0$ and $H_0 : \beta_t = \beta_0$ with $\beta_0 \neq 0$, respectively, for range of k and α combinations, computed via 10,000 Monte Carlo simulations. For comparison, I also report the analogous critical values for the return-based beta estimator, which are obtained in closed form from the Student's t -distribution following Bollerslev et al. (2024b). To facilitate a direct assessment of efficiency, I further report the interval widths $B_\alpha^+ - B_\alpha^-$, for both estimators.

These results indicate that the critical values are tighter for the candlestick-based beta estimator compared to the return-based estimator, across all values of k and α . Specifically, the difference in interval widths becomes more pronounced as k decreases. This highlights the usefulness of my estimator, particularly in limited data scenarios.

Table 1: **Critical values for $H_0 : \beta_t = \beta_0$** : This table reports the critical values for the hypothesis test of $H_0 : \beta_t = \beta_0$ at significance levels $\alpha \in \{5\%, 10\%\}$ and for different local window sizes $k \in \{5, 10, 20\}$. The left panel shows the critical values for the return-based beta estimator derived from the Student’s t -distribution with $k - 1$ degrees of freedom. The right panel presents the critical values for the candlestick-based beta estimator computed via 10,000 Monte Carlo simulations. The interval width is defined as $B_\alpha^+ - B_\alpha^-$.

k	Return			Candlestick ($\beta_0 = 0$)			Candlestick ($\beta_0 \neq 0$)		
	B_α^-	B_α^+	Width	B_α^-	B_α^+	Width	B_α^-	B_α^+	Width
Panel A: $\alpha = 5\%$									
5	-2.776	2.776	5.552	-1.657	1.415	3.072	-1.695	1.636	3.332
10	-2.262	2.262	4.524	-1.430	1.460	2.890	-1.609	1.568	3.177
20	-2.093	2.093	4.186	-1.429	1.383	2.812	-1.567	1.563	3.130
Panel B: $\alpha = 10\%$									
5	-2.132	2.132	4.264	-1.204	1.262	2.466	-1.375	1.336	2.711
10	-1.383	1.383	2.766	-1.203	1.199	2.402	-1.338	1.337	2.674
20	-1.328	1.328	2.656	-1.103	1.244	2.347	-1.372	1.364	2.737

5 Monte Carlo Evidence

In this section, I examine the performance of the candlestick-based estimator and the test through Monte Carlo experiments. The analysis proceeds in two parts. First, I investigate the efficiency of the proposed estimator by comparing its asymptotic risk against an oracle estimator and a return-based estimator. Second, I evaluate the power of the hypothesis test for spot betas introduced in Section 4.

5.1 Efficiency of the Spot Covariance Estimator

To assess the efficiency of my approach, I compare the asymptotic risk of my estimator against two natural benchmarks.¹⁰ The first benchmark is an infeasible oracle estimator that minimizes asymptotic risk under knowledge of the true correlation structure. This estimator serves as a theoretical upper bound for the performance of any feasible estimator, as it uses population information. The second benchmark is the return-based estimator, which relies solely on (open-to-close) returns. All three estimators belong to the class defined in Equation (10), differing only in the choice of λ : the return-based

¹⁰It is important to note that the asymptotic risk $\tilde{R}(\lambda; \rho_t)$ can be computed via simulations for any λ and k if the correlation structure is known.

Table 2: **Asymptotic Risk of Estimators for Different ρ Values:** The table presents the asymptotic risk of three estimators: the return-based estimator, the candlestick-based estimator, and an oracle estimator that minimizes asymptotic risk with knowledge of the true correlation structure. The results are shown for various local window sizes $k \in \{5, 10, 20\}$ and correlation levels $\rho \in \{0, 0.2, 0.6\}$. The asymptotic risk is computed via Monte Carlo simulations.

k	Return	$\rho = 0$		$\rho = 0.2$		$\rho = 0.6$	
		Candlestick	Oracle	Candlestick	Oracle	Candlestick	Oracle
5	1.212	0.427	0.384	0.429	0.384	0.440	0.391
10	0.599	0.240	0.222	0.241	0.221	0.249	0.224
20	0.297	0.128	0.119	0.128	0.118	0.134	0.119

estimator sets $\lambda = (1, 0, \dots, 0)^\top$, the oracle sets $\lambda = \operatorname{argmin}_\lambda \tilde{R}(\lambda, \rho_t)$, and the proposed estimator sets $\lambda = \lambda^*$, minimizing the average risk across correlation structures.

Table 2 reports the asymptotic risk values for various configurations. I consider 3 different levels of correlation, $\rho \in \{0, 0.2, 0.6\}$.¹¹ Moreover, the size of the local estimation window is set to $k \in \{5, 10, 20\}$. All results are based on 10,000 Monte Carlo simulations.

These results provide several key insights. First, my estimator achieves asymptotic risk levels that are very close to those of the oracle estimator across all values of k and ρ , implying roughly 8 – 9% efficiency loss. This indicates that the proposed approach effectively addresses complications arising from the distribution of the estimator being non-pivotal, resulting in an estimator that performs nearly as well as the infeasible oracle estimator. Second, the proposed estimator consistently demonstrates a significant reduction in asymptotic risk compared to the return-based estimator, particularly for smaller values of k . That is, for $k = 5$, the candlestick-based estimator’s risk is roughly 65% lower than that of the return-based estimator, and this difference is about 55% for $k = 20$. This suggests that my approach effectively leverages the candlestick prices and yields more efficient estimators. With only $k = 5$ observations, the proposed estimator achieves a lower risk than the return-based estimator using $k = 10$ observations, demonstrating its superior efficiency in extracting information from narrow time windows. This advantage is particularly valuable in high-frequency event studies, where identification often relies on short time intervals and data are inherently limited.

¹¹These numbers represent the 10th, 50th, 90th percentiles of cross-section of pairwise correlations among the S&P 500 stocks.

Overall, this analysis reveals that incorporating candlestick information substantially improves efficiency, providing a powerful and practical alternative to traditional return-based methods. Importantly, my approach performs nearly as well as oracle estimators, leaving limited potential improvement through alternative specifications.

5.2 Size and Power of the Spot Beta Test

Next, I evaluate the properties of the hypothesis test for spot betas introduced in Section 4 with finite sample simulations.

To do this, I simulate high-frequency price data from a data generating process (DGP), previously studied in [Bollerslev et al. \(2024b\)](#).¹² Specifically, this DGP assumes a two-factor structure for the market variance process: $\nu_t = V_{1,t} + V_{2,t}$ where $V_{1,t}$ and $V_{2,t}$ follow the processes:

$$\begin{aligned} dV_{1,t} &= 0.0128 (0.4068 - V_{1,t}) dt + 0.0954 \sqrt{V_{1,t}} \left(\gamma dW_{1,t} + \sqrt{1 - \gamma^2} dB_{1,t} \right), \\ dV_{2,t} &= 0.6930 (0.4068 - V_{2,t}) dt + 0.7023 \sqrt{V_{2,t}} \left(\gamma dW_{1,t} + \sqrt{1 - \gamma^2} dB_{2,t} \right). \end{aligned} \quad (29)$$

Here, $B_{1,t}$ and $B_{2,t}$ are independent Brownian motions that are also independent of $W_{1,t}$ and $W_{2,t}$. The parameter γ is set to -0.7 , capturing the well-documented leverage effect in financial markets. The coefficients are set so that the first volatility factor is highly persistent, showing 2.5-month half-life, while the second factor reverts to the mean quickly, with a half-life of a day. This allows the DGP to capture both short-term fluctuations and long-term trends in volatility.

The idiosyncratic variance ς_t and the spot beta are assumed to follow:

$$\beta_t = 1 + 0.25 \sin(t)^2, \quad \varsigma_t = (1.5 + 0.25 \sin(t)^2) \nu_t \quad (30)$$

where β_t fluctuates between 1 and 1.25 over time and ς_t is set to be proportional to the market variance ν_t . Finally, the price processes are generated using the representation provided in Equation (3).

I simulate the price path using an Euler scheme on a one-second grid. Then I sample the prices at one-minute frequency to construct the candlestick data, implying $\Delta_n = 1/390$ and $n = 390$ intraday observations. This setup mimics the empirical applications involving high-frequency financial data. I consider three different values

¹²This DGP is built on the univariate setup originally proposed by [Bollerslev and Todorov \(2011\)](#). Later, it is implemented by a number of subsequent studies (e.g., [Bollerslev et al. \(2021\)](#); [Li et al. \(2024\)](#)).

Table 3: **Power of Tests Derived from Return and Candlestick Estimators (%)**: The table reports the rejection rates (in percent) for the hypothesis test of $H_0 : \beta_t = 0$ at significance levels $\alpha \in \{1\%, 5\%, 10\%\}$ and for different local window sizes $k \in \{5, 10, 20\}$ based on simulated data. The left panel shows the rejection rates for the return-based beta estimator, while the right panel presents the rejection rates for the candlestick-based beta estimator.

k	Return			Candlestick		
	1%	5%	10%	1%	5%	10%
5	8.84	28.02	42.11	31.49	58.82	65.67
10	32.53	59.98	72.66	60.46	84.81	91.22
20	73.99	90.32	94.80	94.95	99.15	99.44

for the local window size, $k \in \{5, 10, 20\}$.

For each configuration, I compute empirical rejection rates for the null hypothesis $H_0 : \beta_t = 0$ at significance levels of 1%, 5%, and 10% using the critical values in the middle block of Table 1. The results are summarized in Table 3. I also report the rejection rates for the return-based beta estimator for comparison.¹³ All results are based on 10,000 Monte Carlo simulations.

Simulation results demonstrate that, across all scenarios, the candlestick-based test consistently rejects false null hypotheses more frequently than the return-based test. This gap widens as the local window size decreases. For instance, when $k = 20$ and $\alpha = 5\%$, the difference in rejection rates is approximately 9%, increasing to 25% for $k = 10$ and $\alpha = 5\%$. These results demonstrate a substantial reduction in false negatives, highlighting the superior power of the candlestick-based procedure.

To further illustrate the power advantage of the proposed test, Figure 2 presents the power curves for both tests as a function of the true β_t value. In this analysis, I use the same DGP as before but set β_t to be constant, varying from 0 to 2. The null hypothesis is that $\beta_t = 0$, and the empirical rejection rates are computed at a 5% significance level using 10,000 Monte Carlo simulations. The local window size is fixed at $k = 10$.

Figure 2 shows the rejection rates for both return-based and candlestick-based tests. The x-axis shows how far β_t is from the null value of 0. Looking at the leftmost point where β_t is set to zero, both tests appear to correctly maintain the size at approximately

¹³Note that the return-based test is developed by [Bollerslev et al. \(2024b\)](#). As discussed in Section 4, the corresponding test statistic is shown to be Student's t distributed.

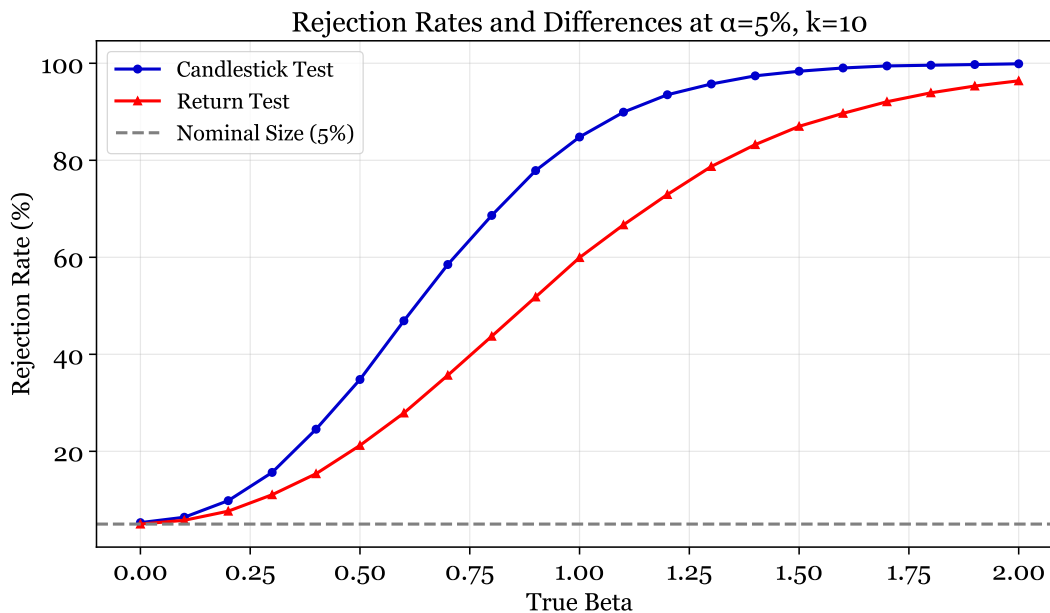


Figure 2: **Power Curves for $H_0 : \beta_t = 0$** : The figure plots the empirical power of the return-based and candlestick-based tests as a function of the true β_t value (assumed to be constant), with $k = 10$ and a significance level of 5%. The power is computed using 10,000 Monte Carlo simulations.

5%. At the right-end where β_t is far from the null, unsurprisingly, both tests achieve a power close to 100%. As β_t deviates from zero, the power of both tests increases, with the candlestick-based test exhibiting a notably steeper increase. For example, when β_t is around 1, the candlestick-based test achieves a power of approximately 85%, while the return-based test is around 60%, indicating a 25% difference.

So far, this analysis has focused on the case of a zero null hypothesis. I also evaluate the power of the tests for non-zero null hypotheses. Particularly, I consider $H_0 : \beta_t = 1$. For this case, the same simulated price data is employed as before, but the critical values are now based on rightmost block of Table 1. Figure 3 exhibits the power curves for both tests under this non-zero null hypothesis. A similar pattern emerges here: the candlestick-based test consistently outperforms the return-based test across almost all values of β_t , with the power gap increases up to 25% as β_t moving away from the null value of 1. Not surprisingly, at the null, the candlestick-based test slightly underrejects, around 4%, while the return-based test maintains the correct size.

Overall, these results demonstrate that candlestick-based inference substantially improves power for spot betas, particularly at moderate deviations from the null. Furthermore, I show that the advantage extends beyond the zero-null case to more general

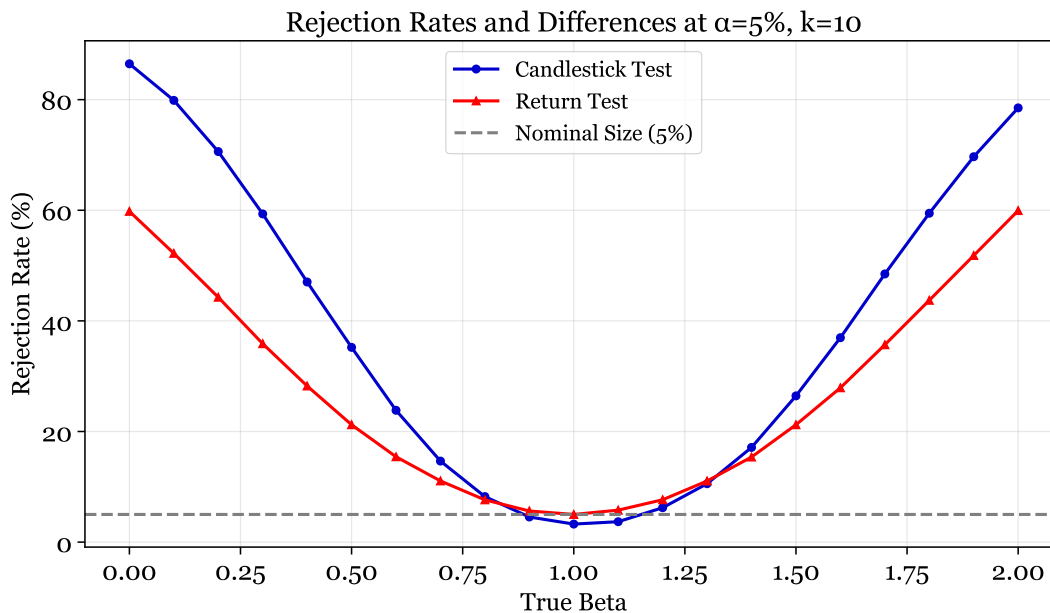


Figure 3: **Power Curves for $H_0 : \beta_t = 1$** : The figure plots the empirical power of the return-based and candlestick-based tests as a function of the true β_t value (assumed to be constant), with $k = 10$ and a significance level of 5%. The power is computed using 10,000 Monte Carlo simulations.

settings in which sup test approach is implemented, underscoring the practical value of the proposed test for inference on spot betas.

6 Empirical Application: Is Bitcoin Market Neutral?

To demonstrate the practical value of the proposed estimator and inference procedure, I analyze Bitcoin’s market exposure, i.e., its market beta. Given the growing role of crypto assets in institutional and retail portfolios, this empirical question can have important implications for risk management and portfolio selection. Specifically, recent years have seen several developments that facilitated investment in crypto assets. In 2017, the Chicago Mercantile Exchange (CME) launched Bitcoin futures contracts, followed by the introduction of Bitcoin options in 2020. More recently, in January 2024, the U.S. Securities and Exchange Commission (SEC) approved the first Bitcoin exchange-traded fund (ETF).

Meanwhile, the risk characteristics of crypto assets remain a subject of debate among academics and practitioners. Crypto advocates often describe these assets as

“digital gold”, suggesting potential hedging benefits against aggregate market risk. In a similar vein, Liu and Tsyvinski (2021) find limited evidence of systematic exposure of crypto assets to traditional factors including the equity market portfolio. Motivated by these developments, I estimate the spot beta and then test the null of market neutrality, i.e., zero beta, using the new candlestick-based framework proposed in Section 4, offering a more granular perspective to the same empirical question.

For this empirical analysis, I collect 1-minute candlestick observations from two prominent ETFs: SPY and IBIT. While the former is a well-known ETF that tracks the S&P 500 index and commonly used in empirical studies, the latter is a newly launched (as of January 2024) iShares Bitcoin Trust ETF designed to track the performance of Bitcoin.¹⁴ The data is sourced from NYSE Trade and Quote database through Wharton Research Data Services(WRDS). The sample spans the entire year 2024, covering 250 trading days and standard trading hours from 9 : 30 to 16 : 00.¹⁵

I estimate spot beta using non-overlapping local windows of $k = 10$ one-minute candlesticks, yielding 39 spot estimates per trading day. Then, I test the null hypothesis of zero beta $H_0 : \beta_t = 0$, or market neutrality, for each estimate using the test statistic developed in Section 4 with a significance level of 5%.

I start by assessing the monthly rejection rates (in terms of percentage) of the null hypothesis of market neutrality which is defined as the proportion of intraday spot beta estimates that are significantly different from zero at the 5% level. The results are presented in Figure 4. The figure shows that the rejection rates are around 10% in the first two months of 2024, later increasing to 40% in mid-2024 and ending the year with a similar rate. Notably, the rejection rates are more pronounced in August and September, reaching around 60%.

Interestingly, these months also coincide with a number of crucial economic events and heightened market volatility. For instance, in August 2024, Bank of Japan announced a rate hike and this was followed by global equity sell-offs. Moreover, the first week of September was marked by a series of weak production and labor market data releases, which raised concerns about a potential economic slowdown. Specifically, during that period, the ISM manufacturing index, the ADP jobs report and the non-farm

¹⁴Cryptocurrency ETFs are designed to provide investors with exposure to the price movements in crypto markets. The IBIT has been the most traded one since its launch and its net asset value exceeds \$70 billion as of 2025.

¹⁵One advantage of pairing SPY with IBIT is that both ETFs trade on regulated exchanges (NYSE Arca and NASDAQ) and thus share the same market microstructure like the same trading hours. This means that the co-movement between them resulted from genuine economic dynamics, not from differences in how or when each asset is traded. That makes the pair a clean setting for my empirical analysis.

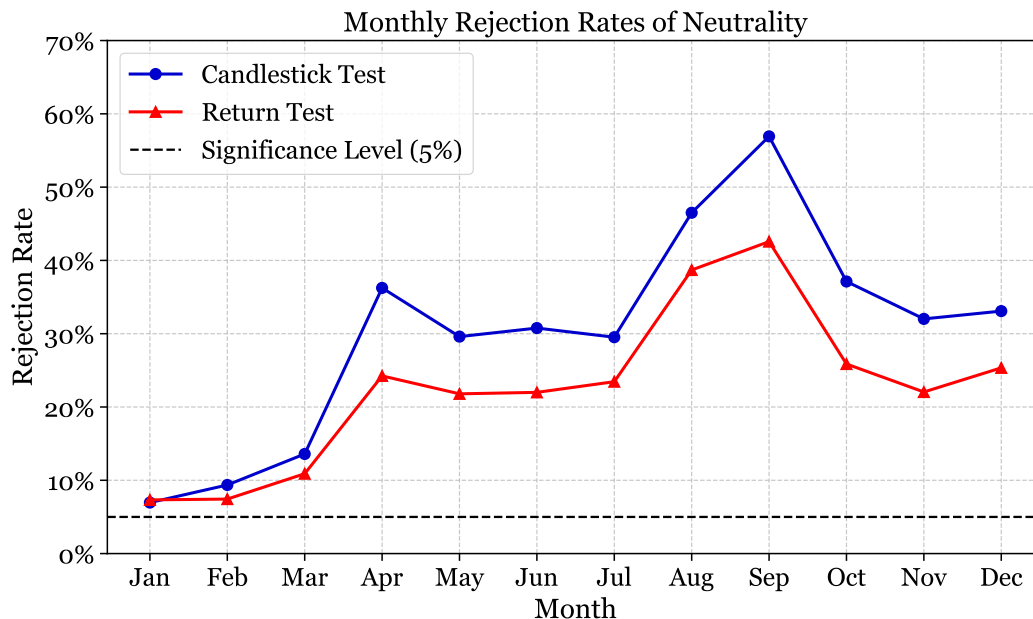


Figure 4: **Monthly Rejection Rates of the Null Hypothesis of Zero Beta for IBIT.** The figure shows the monthly rejection rates of the null hypothesis of zero beta for IBIT using candlestick-based inference framework. The sample covers the entire 2024 year, a total of 250 days, and usual trading hours from 9 : 30 to 16 : 00.

payrolls came in below expectations and consequently increased market volatility.

It may be instructive to look at these days in more detail. Figure 5 presents the spot beta estimates and corresponding confidence intervals for September 3–6. This figure reveals that the null hypothesis of zero beta is rejected in a substantial portion of the intraday intervals, approximately 65% of the time. For instance, looking at the bottom right panel, the null is rejected in 27 out of 39 intervals on September 6. On that day, the NFP report was released at 8:30 AM, prior to market open, and spot beta estimates were already significant and around 1.5 at the opening. Similar patterns are observed on the other days of that week.

As a result, this empirical analysis provides evidence that Bitcoin’s market exposure may not be negligible, and it can show significant positive beta with respect to the market portfolio, particularly during periods of heightened market volatility, precisely the periods when such instruments are supposed to be most valuable in terms of risk management purposes.

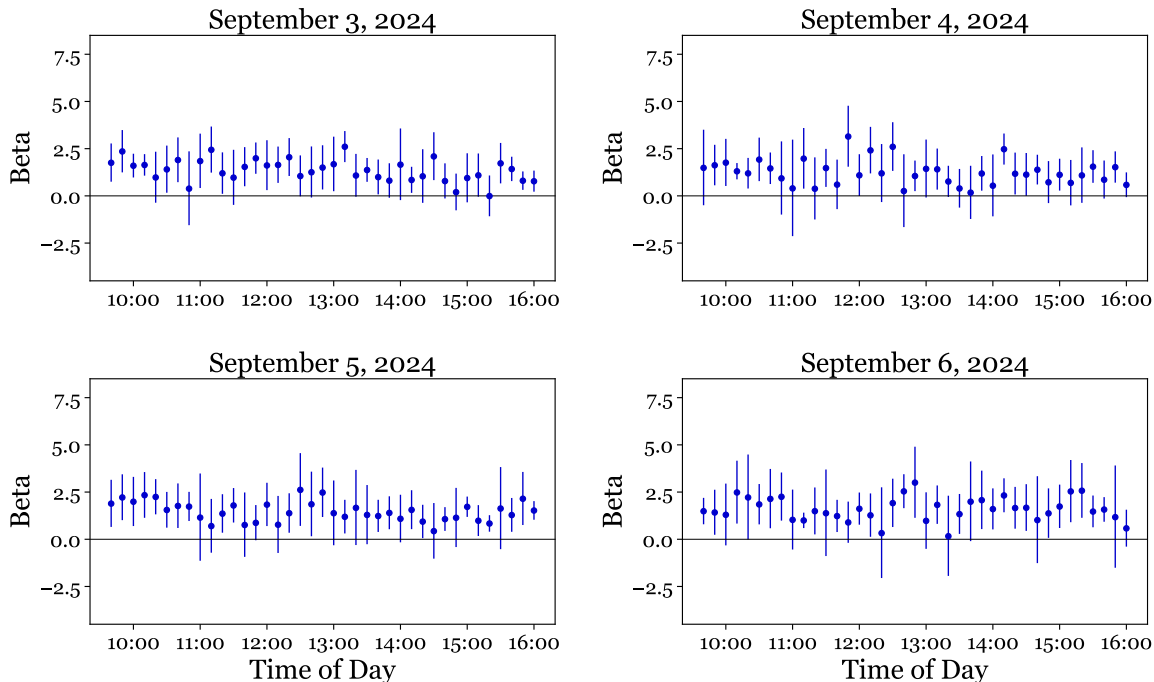


Figure 5: **Spot Beta of IBIT on September 3-6, 2024.** The figure shows the spot beta estimates of IBIT with respect to SPY using 10 1 – *min* candlestick observations in 10 – *min* frequency. The vertical lines indicate the 95% confidence intervals.

7 Conclusion

This paper develops a new framework for estimating and conducting inference on spot regressions, with a focus on spot betas, using high-frequency candlestick data, which contains high and low prices in addition to the open and close prices for each interval. The estimator is constructed by minimizing a quadratic risk function under fixed- k asymptotics, where k is the number of intraday observations used for local estimation. This approach yields a simple and implementable estimator that effectively combines information from all four prices. Furthermore, I develop a hypothesis testing procedure for spot betas based on the proposed estimator, where the critical values are computed by simulating the limiting distribution of the test statistic under the null hypothesis.

Monte Carlo experiments demonstrate that the candlestick-based estimator attains asymptotic risk levels close to those of an infeasible oracle benchmark, while substantially outperforming the return-based estimator. Moreover, the candlestick-based test exhibits markedly higher power (up to 25% improvement) and reduces false negatives in inference. These efficiency gains are especially valuable for high-frequency event studies, where identification relies on short time windows. In an empirical study with 1-minute candlestick data of SPY (S&P 500 ETF) and IBIT (iShares Bitcoin Trust

ETF), the method reveals significant positive market exposure of the Bitcoin ETF, especially during turbulent periods, challenging popular “digital gold” narrative for risk management.

Overall, this paper highlights the practical value of candlestick data for more precise and reliable estimation and inference in spot regressions, offering researchers and practitioners a readily implementable tool for analyzing time-varying risk exposures in financial markets.

Appendix A

A.1 Decomposition of the Loss Function

The remark in the Section 3.2 presents a decomposition of the loss function into three components plus cross terms: $L(\lambda; \mathbf{c}_t) = L_\nu + L_\zeta + L_\beta + \text{cross terms}$. This decomposition is crucial for understanding the estimation methodology of this paper. Therefore, I provide further details on how this decomposition is derived and its implications for the estimation procedure, in this section.

To derive the decomposition, I start by writing the estimator as Cholesky form:

$$\widehat{\mathbf{c}}_{n,t}(\lambda) = \widehat{\boldsymbol{\sigma}}_{n,t}(\lambda) \widehat{\boldsymbol{\sigma}}_{n,t}(\lambda)^\top \quad \text{where} \quad \widehat{\boldsymbol{\sigma}}_{n,t}(\lambda) = \begin{pmatrix} \widehat{\nu}_t^{1/2} & 0 \\ \widehat{\beta}_t \widehat{\nu}_t^{1/2} & \widehat{\zeta}_t^{1/2} \end{pmatrix}. \quad (\text{A.1})$$

Now, consider multiplying $\widehat{\boldsymbol{\sigma}}_{n,t}(\lambda)$ by $\boldsymbol{\sigma}_t^{-1}$:

$$\boldsymbol{\sigma}_t^{-1} \widehat{\boldsymbol{\sigma}}_{n,t}(\lambda) = \begin{pmatrix} \nu_t^{-1/2} & 0 \\ -\beta_t \nu_t^{-1/2} & \zeta_t^{-1/2} \end{pmatrix} \begin{pmatrix} \widehat{\nu}_t^{1/2} & 0 \\ \widehat{\beta}_t \widehat{\nu}_t^{1/2} & \widehat{\zeta}_t^{1/2} \end{pmatrix} = \begin{pmatrix} (\widehat{\nu}_t/\nu_t)^{1/2} & 0 \\ (\widehat{\beta}_t - \beta_t)(\widehat{\nu}_t/\nu_t)^{1/2} & (\widehat{\zeta}_t/\zeta_t)^{1/2} \end{pmatrix}.$$

To simplify the notation, define the following variables:

$$A \equiv (\widehat{\nu}_t/\nu_t)^{1/2} - 1, \quad B \equiv (\widehat{\zeta}_t/\zeta_t)^{1/2} - 1 \quad \text{and} \quad C \equiv (\widehat{\beta}_t - \beta_t)(\widehat{\nu}_t/\nu_t)^{1/2}$$

With these definitions, the multiplicative estimation error can be expressed as:

$$\boldsymbol{\sigma}_t^{-1} \widehat{\mathbf{c}}_{n,t}(\lambda) \boldsymbol{\sigma}_t^{-1\top} = \begin{pmatrix} A^2 & AB \\ AB & B^2 + C^2 \end{pmatrix}.$$

Then, the loss function can be written as:

$$\begin{aligned} L(\lambda; \mathbf{c}_t) &= \left\| \boldsymbol{\sigma}_t^{-1} \widehat{\mathbf{c}}_{n,t}(\lambda) \boldsymbol{\sigma}_t^{-1\top} - I \right\|^2 \\ &= (A^2 - 1)^2 + (B^2 + C^2 - 1)^2 + 2(AB)^2 \end{aligned}$$

After some algebraic manipulation, one can further express the loss function as:

$$L(\lambda; \mathbf{c}_t) = (A^2 - 1)^2 + (C^2 - 1)^2 + \frac{B^2}{A^2} (2A^4 + A^2 B^2 + 2A^2 (C^2 - 1))$$

Note that

$$(A^2 - 1)^2 = \left(\frac{\widehat{\nu}_t - \nu_t}{\nu_t} \right)^2, \quad (C^2 - 1)^2 = \left(\frac{\widehat{\varsigma}_t - \varsigma_t}{\varsigma_t} \right)^2 \quad \text{and} \quad \frac{B^2}{A^2} = \left(\frac{\widehat{\beta}_t - \beta_t}{\varsigma_t/\nu_t} \right)^2.$$

Finally, the loss function can be decomposed as:

$$L(\lambda; \mathbf{c}_t) = \underbrace{\left(\frac{\widehat{\nu}_t - \nu_t}{\nu_t} \right)^2}_{\equiv L_\nu} + \underbrace{\left(\frac{\widehat{\varsigma}_t - \varsigma_t}{\varsigma_t} \right)^2}_{\equiv L_\varsigma} + \underbrace{\left(\frac{\widehat{\beta}_t - \beta_t}{\varsigma_t/\nu_t} \right)^2}_{\equiv L_\beta} \cdot \psi$$

where L_ν , L_ς and L_β represent the squared relative errors of the market variance, idiosyncratic variance and spot beta, respectively, and $\psi = 2A^4 + A^2B^2 + 2A^2(C^2 - 1)$ is a scaling factor and function of the estimation errors of all three parameters.

This decomposition reveals that the loss function can be expressed as a sum of weighted squared errors of the three parameters of interest: the market variance ν_t , the idiosyncratic variance ς_t and the spot beta β_t . Thus, minimizing the overall loss function $L(\lambda; \mathbf{c}_t)$ effectively balances tradeoffs between these different components.

A.2 Proof of Proposition 1

In this section, I provide the proof of Proposition 1 which builds on the coupling techniques developed in Jacod et al. (2021) and Bollerslev et al. (2021). The former studies the approximation of the estimation error for the spot covariance estimator when k increases with n , whereas the latter focuses on coupling the spot volatility estimator in a fixed- k framework. As noted in the main text, I consider a fixed- k setup and thus my work is in the same spirit as Bollerslev et al. (2021). To ensure consistency with the existing literature, I closely follow the notation introduced in the aforementioned papers.

I rewrite the Assumption 1 and Proposition 1 here for convenience:

Assumption 1. *Suppose that \mathbf{X}_t has the form in Equation (1) and there exists a sequence $(T_m)_{m \geq 1}$ of stopping times increasing to infinity and the following conditions hold for each $m \geq 1$:*

- (i) $\|\mathbf{b}_t\| + \|\boldsymbol{\sigma}_t\| + \|\boldsymbol{\sigma}_t^{-1}\| \leq K_m$ for some constant K_m for all $t \in [0, T_m]$;
- (ii) $\mathbb{E}[\|\boldsymbol{\sigma}_{t \wedge T_m} - \boldsymbol{\sigma}_{s \wedge T_m}\|^2] \leq K_m|t - s|$ for all $t, s \in [0, T_m]$.

Proposition 1. *Suppose that Assumption 1 holds. Fix any $t \in [0, T]$. For any $k \geq 1$ and Λ , the following holds as $\Delta_n \rightarrow 0$:*

$$\left\| \boldsymbol{\sigma}_t^{-1} \widehat{\mathbf{c}}_{n,t}(\lambda) \boldsymbol{\sigma}_t^{-\top} - U_{n,t} \right\| = o_p(1) \quad (\text{A.2})$$

where $U_{n,t} = \frac{1}{k} \sum_{i \in \mathcal{I}_{n,t}} \left\{ \lambda_1 \boldsymbol{\zeta}_{i,r} \boldsymbol{\zeta}_{i,r}^\top + \lambda_2 \boldsymbol{\zeta}_{i,a} \boldsymbol{\zeta}_{i,a}^\top + \lambda_3 \boldsymbol{\zeta}_{i,w} \boldsymbol{\zeta}_{i,w}^\top \right\}$ and, for any $i \in \mathcal{I}_{n,t}$,

$$\begin{aligned} \boldsymbol{\zeta}_{i,r} &\equiv \frac{\mathbf{W}_{i\Delta_n} - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \\ \boldsymbol{\zeta}_{i,a} &\equiv \boldsymbol{\varrho}_t^{-1} \sup_{\tau \in \mathcal{T}_i} \boldsymbol{\varrho}_t \left(\frac{\mathbf{W}_\tau - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) + \boldsymbol{\varrho}_t^{-1} \inf_{\tau \in \mathcal{T}_i} \boldsymbol{\varrho}_t \left(\frac{\mathbf{W}_\tau - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) - \left(\frac{\mathbf{W}_{i\Delta_n} - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) \\ \boldsymbol{\zeta}_{i,w} &\equiv \boldsymbol{\varrho}_t^{-1} \sup_{\tau \in \mathcal{T}_i} \boldsymbol{\varrho}_t \left(\frac{\mathbf{W}_\tau - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) - \boldsymbol{\varrho}_t^{-1} \inf_{\tau \in \mathcal{T}_i} \boldsymbol{\varrho}_t \left(\frac{\mathbf{W}_\tau - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) \end{aligned} \quad (\text{A.3})$$

with $\boldsymbol{\varrho}_t$ being the square root of spot correlation matrix $\boldsymbol{\rho}_t$, i.e., $\boldsymbol{\rho}_t = \boldsymbol{\varrho}_t \boldsymbol{\varrho}_t^\top$. In explicit terms, $\boldsymbol{\rho}_t = \text{diag}(\mathbf{c}_t)^{-\frac{1}{2}} \mathbf{c}_t \text{diag}(\mathbf{c}_t)^{-\frac{1}{2}}$ and $\boldsymbol{\varrho}_t = \text{diag}(\mathbf{c}_t)^{-\frac{1}{2}} \boldsymbol{\sigma}_t$ where $\text{diag}(\mathbf{c}_t)$ is a diagonal matrix with the same diagonal elements as \mathbf{c}_t .

Proof. Fix $k \geq 1$ and λ . Let K denote a generic positive constant. As is common in the spot estimation literature, one can strengthen Assumption 1 by assuming the conditions hold with $T_m = \infty$, which can be justified by a standard localization argument (see Jacod and Protter (2012) for details).

I begin by writing out the explicit expressions for the candlestick returns:

$$\begin{aligned} \mathbf{r}_i &\equiv \Delta_n^{-\frac{1}{2}} \left(\int_{(i-1)\Delta_n}^{i\Delta_n} \mathbf{b}_s ds + \int_{(i-1)\Delta_n}^{i\Delta_n} \boldsymbol{\sigma}_s d\mathbf{W}_s \right) \\ \mathbf{h}_i &\equiv \Delta_n^{-\frac{1}{2}} \left(\sup_{t \in \mathcal{T}_{n,i}} \left(\int_{(i-1)\Delta_n}^t \mathbf{b}_s ds + \int_{(i-1)\Delta_n}^t \boldsymbol{\sigma}_s d\mathbf{W}_s \right) \right) \\ \boldsymbol{\ell}_i &\equiv \Delta_n^{-\frac{1}{2}} \left(\inf_{t \in \mathcal{T}_{n,i}} \left(\int_{(i-1)\Delta_n}^t \mathbf{b}_s ds + \int_{(i-1)\Delta_n}^t \boldsymbol{\sigma}_s d\mathbf{W}_s \right) \right) \end{aligned} \quad (\text{A.4})$$

I also introduce the following definitions:

$$\begin{aligned} \mathbf{r}'_i &\equiv \boldsymbol{\sigma}_{(i-1)\Delta_n} \left(\frac{\mathbf{W}_{i\Delta_n} - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) \\ \mathbf{h}'_i &\equiv \sup_{t \in \mathcal{T}_{n,i}} \boldsymbol{\sigma}_{(i-1)\Delta_n} \left(\frac{\mathbf{W}_t - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) \\ \boldsymbol{\ell}'_i &\equiv \inf_{t \in \mathcal{T}_{n,i}} \boldsymbol{\sigma}_{(i-1)\Delta_n} \left(\frac{\mathbf{W}_t - \mathbf{W}_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \right) \end{aligned} \quad (\text{A.5})$$

which serve as the coupling variables for the candlestick returns. Similar variables can

be defined for the range and asymmetry variables:

$$\begin{aligned}\mathbf{a}'_i &\equiv \mathbf{h}'_i + \boldsymbol{\ell}'_i - \mathbf{r}'_i \\ \mathbf{w}'_i &\equiv \mathbf{h}'_i - \boldsymbol{\ell}'_i\end{aligned}\tag{A.6}$$

The proof consists of two steps. The first step controls how well the coupling returns in Equation (A.5) approximate the candlestick returns in Equation (A.4). The second step combines these results with the continuous mapping theorem to reach the desired conclusion. Before proceeding to the first step, I derive useful intermediate results.

By Assumption 1, it is easy to see that:

$$\left\| \int_{(i-1)\Delta_n}^{i\Delta_n} \mathbf{b}_s ds \right\| \leq \int_{(i-1)\Delta_n}^{i\Delta_n} \|\mathbf{b}_s\| ds = O_p(\Delta_n).\tag{A.7}$$

Moreover, the Burkholder-Davis-Gundy inequality and Assumption 1 imply that:

$$\begin{aligned}\mathbb{E} \left[\left\| \int_{(i-1)\Delta_n}^{i\Delta_n} (\boldsymbol{\sigma}_s - \boldsymbol{\sigma}_{(i-1)\Delta_n}) d\mathbf{W}_s \right\|^2 \right] &\leq K\Delta_n \mathbb{E} \left[\int_{(i-1)\Delta_n}^{i\Delta_n} \|\boldsymbol{\sigma}_s - \boldsymbol{\sigma}_{(i-1)\Delta_n}\|^2 ds \right] \\ &\leq K\Delta_n^2.\end{aligned}\tag{A.8}$$

Further, we can deduce that:

$$\sup_{t \in \mathcal{T}_{n,i}} \left\| \int_{(i-1)\Delta_n}^t (\boldsymbol{\sigma}_s - \boldsymbol{\sigma}_{(i-1)\Delta_n}) d\mathbf{W}_s \right\| = O_p(\Delta_n).\tag{A.9}$$

Step 1: I now establish approximation results for \mathbf{r}_i , \mathbf{h}_i and $\boldsymbol{\ell}_i$ separately. I start with the return:

$$\begin{aligned}\|\mathbf{r}_i - \mathbf{r}'_i\| &\leq \left\| \Delta_n^{-\frac{1}{2}} \int_{(i-1)\Delta_n}^{i\Delta_n} \mathbf{b}_s ds \right\| + \left\| \Delta_n^{-\frac{1}{2}} \int_{(i-1)\Delta_n}^{i\Delta_n} (\boldsymbol{\sigma}_s - \boldsymbol{\sigma}_{(i-1)\Delta_n}) d\mathbf{W}_s \right\| \\ &= O_p(\Delta_n^{\frac{1}{2}}).\end{aligned}\tag{A.10}$$

where the first line directly follows from the triangle inequality and the second line uses above intermediate results in Equation (A.7) and (A.9). For the high return, we

have:

$$\begin{aligned}
\|\mathbf{h}_i - \mathbf{h}'_i\| &\equiv \Delta_n^{-\frac{1}{2}} \left\| \sup_{t \in \mathcal{T}_{n,i}} \left(\int_{(i-1)\Delta_n}^t \mathbf{b}_s ds + \int_{(i-1)\Delta_n}^t \boldsymbol{\sigma}_s d\mathbf{W}_s \right) \right. \\
&\quad \left. - \sup_{t \in \mathcal{T}_{n,i}} \boldsymbol{\sigma}_{(i-1)\Delta_n} (\mathbf{W}_t - \mathbf{W}_{(i-1)\Delta_n}) \right\| \\
&\leq \Delta_n^{-\frac{1}{2}} \sup_{t \in \mathcal{T}_{n,i}} \left\| \int_{(i-1)\Delta_n}^t \mathbf{b}_s ds + \int_{(i-1)\Delta_n}^t (\boldsymbol{\sigma}_s - \boldsymbol{\sigma}_{(i-1)\Delta_n}) d\mathbf{W}_s \right\| \\
&\leq \Delta_n^{-\frac{1}{2}} \int_{(i-1)\Delta_n}^{i\Delta_n} \|\mathbf{b}_s\| ds + \sup_{t \in \mathcal{T}_{n,i}} \left\| \int_{(i-1)\Delta_n}^t (\boldsymbol{\sigma}_s - \boldsymbol{\sigma}_{(i-1)\Delta_n}) d\mathbf{W}_s \right\| \\
&= O_p(\Delta_n^{\frac{1}{2}}).
\end{aligned} \tag{A.11}$$

The first two lines are obviously implications of sup definition. Similarly, the last line follows from Equation A.7 and A.9. Finally, one can deduce the same inequality for the low return:

$$\|\boldsymbol{\ell}_i - \boldsymbol{\ell}'_i\| = O_p(\Delta_n^{\frac{1}{2}}), \tag{A.12}$$

and also for the asymmetry and range variables:

$$\begin{aligned}
\|\mathbf{a}_i - \mathbf{a}'_i\| &\leq \|\mathbf{h}_i - \mathbf{h}'_i\| + \|\boldsymbol{\ell}_i - \boldsymbol{\ell}'_i\| + \|\mathbf{r}_i - \mathbf{r}'_i\| = O_p(\Delta_n^{\frac{1}{2}}) \\
\|\mathbf{w}_i - \mathbf{w}'_i\| &\leq \|\mathbf{h}_i - \mathbf{h}'_i\| + \|\boldsymbol{\ell}_i - \boldsymbol{\ell}'_i\| = O_p(\Delta_n^{\frac{1}{2}}).
\end{aligned} \tag{A.13}$$

Furthermore, I claim that

$$\begin{aligned}
\left\| \boldsymbol{\sigma}_t^{-1} \mathbf{r}'_i - \boldsymbol{\zeta}_{i,r} \right\| &= O_p(\Delta_n^{1/2}) \\
\left\| \boldsymbol{\sigma}_t^{-1} \mathbf{a}'_i - \boldsymbol{\zeta}_{i,a} \right\| &= O_p(\Delta_n^{1/2}) \\
\left\| \boldsymbol{\sigma}_t^{-1} \mathbf{w}'_i - \boldsymbol{\zeta}_{i,w} \right\| &= O_p(\Delta_n^{1/2}).
\end{aligned} \tag{A.14}$$

The first line directly follows from Assumption 1. Note that $|i\Delta_n - t| \rightarrow 0$ for any $i \in \mathcal{I}_{n,t}$ as $\Delta_n \rightarrow 0$ and this implies $\|\boldsymbol{\sigma}_t - \boldsymbol{\sigma}_{(i-1)\Delta_n}\| = O_p(\Delta_n^{1/2})$. For notational convenience, I only consider the third line and the same steps can be adopted for the

second line as well. Specifically, one can write:

$$\begin{aligned}
\|\boldsymbol{\sigma}_t^{-1} \mathbf{w}'_i - \boldsymbol{\zeta}_{i,w}\| &\leq \|\boldsymbol{\sigma}_t^{-1}\| \cdot \|\mathbf{w}'_i - \boldsymbol{\sigma}_t \boldsymbol{\zeta}_{i,w}\| \\
&= \|\boldsymbol{\sigma}_t^{-1}\| \cdot \left\| \mathbf{w}'_i - \text{diag}(\mathbf{c}_t)^{1/2} \sup_{\tau,s \in \mathcal{T}_{n,i}} \boldsymbol{\varrho}_t \left(\frac{\mathbf{w}_\tau - \mathbf{w}_s}{\sqrt{\Delta_n}} \right) \right\| \\
&= \|\boldsymbol{\sigma}_t^{-1}\| \cdot \left\| \mathbf{w}'_i - \sup_{\tau,s \in \mathcal{T}_{n,i}} \boldsymbol{\sigma}_t \left(\frac{\mathbf{w}_\tau - \mathbf{w}_s}{\sqrt{\Delta_n}} \right) \right\| \tag{A.15} \\
&= \|\boldsymbol{\sigma}_t^{-1}\| \cdot \left\| \sup_{\tau,s \in \mathcal{T}_{n,i}} (\boldsymbol{\sigma}_t - \boldsymbol{\sigma}_{(i-1)\Delta_n}) \left(\frac{\mathbf{w}_\tau - \mathbf{w}_s}{\sqrt{\Delta_n}} \right) \right\| \\
&= O_p(\Delta_n^{1/2})
\end{aligned}$$

where the first line follows from sub-multiplicative property of matrix norm, the second and third lines use the definition of $\boldsymbol{\zeta}_{i,w}$ and $\boldsymbol{\varrho}_t$, the fourth line is a direct implication of sup definition and the last line uses Assumption 1 and properties of Brownian motion.

Step 2: Rewrite the main statement of the proposition as:

$$\begin{aligned}
\|\boldsymbol{\sigma}_t^{-1} \widehat{\mathbf{c}}_{n,t}(\lambda) \boldsymbol{\sigma}_t^{-\top} - U_{n,t}\| &= \left\| \frac{1}{k} \sum_{i \in \mathcal{I}_{n,t}} \{ \lambda_1 (\boldsymbol{\sigma}_t^{-1} \mathbf{r}_i) (\boldsymbol{\sigma}_t^{-1} \mathbf{r}_i)^\top - \boldsymbol{\zeta}_{i,r} \boldsymbol{\zeta}_{i,r}^\top \} \right. \\
&\quad + \frac{1}{k} \sum_{i \in \mathcal{I}_{n,t}} \{ \lambda_2 (\boldsymbol{\sigma}_t^{-1} \mathbf{a}_i) (\boldsymbol{\sigma}_t^{-1} \mathbf{a}_i)^\top - \boldsymbol{\zeta}_{i,a} \boldsymbol{\zeta}_{i,a}^\top \} \\
&\quad \left. + \frac{1}{k} \sum_{i \in \mathcal{I}_{n,t}} \{ \lambda_3 (\boldsymbol{\sigma}_t^{-1} \mathbf{w}_i) (\boldsymbol{\sigma}_t^{-1} \mathbf{w}_i)^\top - \boldsymbol{\zeta}_{i,w} \boldsymbol{\zeta}_{i,w}^\top \} \right\| \tag{A.16}
\end{aligned}$$

Using the results from Step 1, it follows that the terms in curly brackets are $O_p(\Delta_n^{1/2})$ for any $i \in \mathcal{I}_{n,t}$ and fixed $\lambda = (\lambda_1, \lambda_2, \lambda_3)^\top$. Therefore, the entire expression is $O_p(\Delta_n^{1/2}) = o_p(1)$. This completes the proof. \square

A.3 Proof of Proposition 2

Now, I provide the proof of Proposition 2, which establishes the coupling result for the test statistic defined as:

$$\widehat{T}_n = \frac{\sqrt{k-1} (\widehat{\beta}_{n,t} - \beta_t)}{\sqrt{\widehat{\varsigma}_{n,t} / \widehat{\nu}_{n,t}}}.$$

The proof is based on the algebraic manipulations of the previous proposition. For convenience, I restate the proposition here:

Proposition 2. *Under the conditions of Proposition 1, for any fixed $k \geq 2$ and λ , the following holds as $\Delta_n \rightarrow 0$:*

$$|\widehat{T}_n - \widetilde{T}_n| = o_p(1) \quad \text{where} \quad \widetilde{T}_n \equiv \frac{\sqrt{k-1}[U_{n,t}^{-1}]_{12}}{\sqrt{[U_{n,t}^{-1}]_{11}[U_{n,t}^{-1}]_{22} - [U_{n,t}^{-1}]_{12}^2}}.$$

Proof. Fix $k \geq 2$ and λ . By Proposition 1, we have:

$$\|\boldsymbol{\sigma}_t^{-1}\widehat{\mathbf{c}}_{n,t}(\lambda)\boldsymbol{\sigma}_t^{-\top} - U_{n,t}\| = o_p(1). \quad (\text{A.17})$$

Note that both $\boldsymbol{\sigma}_t^{-1}\widehat{\mathbf{c}}_{n,t}(\lambda)\boldsymbol{\sigma}_t^{-\top}$ and $U_{n,t}$ are positive definite matrices. Therefore, by the continuity of matrix inversion operator on the set of positive definite matrices, we have:

$$\|(\boldsymbol{\sigma}_t^{-1}\widehat{\mathbf{c}}_{n,t}(\lambda)\boldsymbol{\sigma}_t^{-\top})^{-1} - U_{n,t}^{-1}\| = o_p(1). \quad (\text{A.18})$$

This implies that each element of the matrix $(\boldsymbol{\sigma}_t^{-1}\widehat{\mathbf{c}}_{n,t}(\lambda)\boldsymbol{\sigma}_t^{-\top})^{-1}$ converges to the corresponding element of $U_{n,t}^{-1}$ in probability. These relations can be written in explicit forms as follows:

$$\begin{aligned} \left| \left(\frac{\nu_t}{\widehat{\nu}_{n,t}} + \frac{\nu_t(\beta_t - \widehat{\beta}_{n,t})^2}{\widehat{\varsigma}_{n,t}} \right) - [U_{n,t}^{-1}]_{11} \right| &= o_p(1) \\ \left| \left(\frac{\nu_t^{1/2}\varsigma_t^{1/2}(\beta_t - \widehat{\beta}_{n,t})}{\widehat{\varsigma}_{n,t}} \right) - [U_{n,t}^{-1}]_{12} \right| &= o_p(1) \\ \left| \frac{\varsigma_t}{\widehat{\varsigma}_{n,t}} - [U_{n,t}^{-1}]_{22} \right| &= o_p(1). \end{aligned}$$

Using the second and third lines, one can write:

$$\left| \frac{(\widehat{\beta}_{n,t} - \beta_t)}{\sqrt{\varsigma_{n,t}/\nu_{n,t}}} - \frac{[U_{n,t}^{-1}]_{12}}{[U_{n,t}^{-1}]_{22}} \right| = o_p(1). \quad (\text{A.19})$$

Moreover, from all three lines, one can deduce that:

$$\left| \frac{\nu_t}{\widehat{\nu}_{n,t}} - \left([U_{n,t}^{-1}]_{11} - \frac{[U_{n,t}^{-1}]_{12}^2}{[U_{n,t}^{-1}]_{22}} \right) \right| = o_p(1). \quad (\text{A.20})$$

Finally, combining the above equations yields:

$$\left| \frac{\sqrt{k-1}(\widehat{\beta}_{n,t} - \beta_t)}{\sqrt{\widehat{\varsigma}_{n,t}/\widehat{\nu}_{n,t}}} - \frac{\sqrt{k-1}[U_{n,t}^{-1}]_{12}}{\sqrt{[U_{n,t}^{-1}]_{11}[U_{n,t}^{-1}]_{22} - [U_{n,t}^{-1}]_{12}^2}} \right| = o_p(1). \quad (\text{A.21})$$

as required. □

References

- Aït-Sahalia, Y., Fan, J., and Xiu, D. (2010). High-frequency covariance estimates with noisy and asynchronous financial data. *Journal of the American Statistical Association*, 105(492):1504–1517.
- Aït-Sahalia, Y. and Jacod, J. (2014). *High-frequency financial econometrics*. Princeton University Press.
- Aït-Sahalia, Y., Kalnina, I., and Xiu, D. (2020). High-frequency factor models and regressions. *Journal of Econometrics*, 216(1):86–105.
- Andersen, T., Bollerslev, T., Diebold, F., and Labys, P. (2000). Great realizations. *RISK*, 13:105–108.
- Andersen, T. G., Riva, R., Thyrgaard, M., and Todorov, V. (2023). Intraday cross-sectional distributions of systematic risk. *Journal of Econometrics*, 235(2):1394–1418.
- Andersen, T. G., Thyrgaard, M., and Todorov, V. (2021). Recalcitrant betas: Intraday variation in the cross-sectional dispersion of systematic risk. *Quantitative Economics*, 12(2):647–682.
- Ang, A. and Kristensen, D. (2012). Testing conditional factor models. *Journal of Financial Economics*, 106(1):132–156.
- Back, K. (2010). *Asset pricing and portfolio choice theory*. Oxford University Press.
- Bannouh, K., Van Dijk, D., and Martens, M. (2009). Range-based covariance estimation using high-frequency data: The realized co-range. *Journal of Financial Econometrics*, 7(4):341–372.
- Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2011). Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, 162(2):149–169.
- Barndorff-Nielsen, O. E. and Shephard, N. (2004a). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, 72(3):885–925.

- Barndorff-Nielsen, O. E. and Shephard, N. (2004b). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37.
- Bibinger, M., Hautsch, N., Malec, P., and Reiss, M. (2019). Estimating the spot covariation of asset prices—statistical theory and empirical evidence. *Journal of Business & Economic Statistics*, 37(3):419–435.
- Bollerslev, T., Li, J., and Li, Q. (2024a). Optimal nonparametric range-based volatility estimation. *Journal of Econometrics*, 238(1):105548.
- Bollerslev, T., Li, J., Li, Q., and Li, Y. (2025). Optimal candlestick-based spot volatility estimation: New tricks and feasible inference procedures. *Available at SSRN 5046917*.
- Bollerslev, T., Li, J., and Liao, Z. (2021). Fixed-k inference for volatility. *Quantitative Economics*, 12(4):1053–1084.
- Bollerslev, T., Li, J., and Ren, Y. (2024b). Optimal inference for spot regressions. *American Economic Review*, 114(3):678–708.
- Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2018). Modeling and forecasting (un) reliable realized covariances for more reliable financial decisions. *Journal of Econometrics*, 207(1):71–91.
- Bollerslev, T. and Todorov, V. (2011). Estimation of jump tails. *Econometrica*, 79(6):1727–1783.
- Brandt, M. W. and Diebold, F. X. (2006). A no-arbitrage approach to range-based estimation of return covariances and correlations. *Journal of Business*, 79(1):61–74.
- Christensen, K., Kinnebrock, S., and Podolskij, M. (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *Journal of Econometrics*, 159(1):116–133.
- Christensen, K. and Podolskij, M. (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics*, 141(2):323–349.
- Christensen, K. and Podolskij, M. (2012). Asymptotic theory of range-based multi-power variation. *Journal of Financial Econometrics*, 10(3):417–456.
- Christensen, K., Podolskij, M., and Vetter, M. (2009). Bias-correcting the realized range-based variance in the presence of market microstructure noise. *Finance and Stochastics*, 13:239–268.

- Connor, G., Hagmann, M., and Linton, O. (2012). Efficient semiparametric estimation of the fama–french model and extensions. *Econometrica*, 80(2):713–754.
- Diebold, F. X. and Strasser, G. (2013). On the correlation structure of microstructure noise: A financial economic approach. *Review of Economic Studies*, 80(4):1304–1337.
- Engle, R. (2004). Risk and volatility: Econometric models and financial practice. *American economic review*, 94(3):405–420.
- Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.
- Fan, J. and Wang, Y. (2008). Spot volatility estimation for high-frequency data. *Statistics and its Interface*, 1(2):279–288.
- Gagliardini, P., Ossola, E., and Scaillet, O. (2016). Time-varying risk premium in large cross-sectional equity data sets. *Econometrica*, 84(3):985–1046.
- Garman, M. B. and Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, pages 67–78.
- Hayashi, T. and Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli*, 11(2):359–379.
- Jacod, J., Li, J., and Liao, Z. (2021). Volatility coupling. *The Annals of Statistics*, 49(4):1982–1998.
- Jacod, J. and Protter, P. (2012). *Discretization of Processes*. Springer.
- Lehmann, E. L. and Casella, G. (2006). *Theory of point estimation*. Springer Science & Business Media.
- Li, J., Todorov, V., and Tauchen, G. (2017). Adaptive estimation of continuous-time regression models using high-frequency data. *Journal of Econometrics*, 200(1):36–47.
- Li, J., Wang, D., and Zhang, Q. (2024). Reading the candlesticks: An ok estimator for volatility. *Review of Economics and Statistics*, 106(4):1114–1128.
- Li, Y., Nolte, I., Nolte, S., and Yu, S. (2025). Realized candlestick wicks. *Journal of Econometrics*, 250:106014.
- Liao, Y. and Todorov, V. (2024). Changes in the span of systematic risk exposures. *Quantitative Economics*, 15(3):817–847.

- Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, 20(4):587–615.
- Liu, Y. and Tsyvinski, A. (2021). Risks and returns of cryptocurrency. *The Review of Financial Studies*, 34(6):2689–2727.
- Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36(2):270–296.
- Martens, M. and Van Dijk, D. (2007). Measuring volatility with the realized range. *Journal of Econometrics*, 138(1):181–207.
- Mykland, P. A. and Zhang, L. (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica*, 77(5):1403–1445.
- Nakamura, E. and Steinsson, J. (2018). High-frequency identification of monetary non-neutrality: the information effect. *The Quarterly Journal of Economics*, 133(3):1283–1330.
- Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, pages 61–65.
- Patton, A. J. and Simsek, Y. (2024). Intraday variation in systematic risks and information flows. *Available at SSRN 5006587*.
- Rogers, L. and Shepp, L. (2006). The correlation of the maxima of correlated brownian motions. *Journal of Applied Probability*, 43(3):880–883.
- Rogers, L. C. and Zhou, F. (2008). Estimating correlation from high, low, opening and closing prices. *Annals of Applied Probability*, 18(2):813–823.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442.