Multi-Horizon Test for Market Frictions*

Z. Merrick Li† Xiye Yang‡

September 30, 2025

Abstract

We test for the presence of market frictions that induce transitory deviations of observed asset prices from the underlying efficient prices. Our test is based on the joint inference of return covariances across multiple horizons. We demonstrate that a small set of horizons suffices to identify a broad spectrum of frictions, both theoretically and practically. Our method works for high- and low-frequency data under different asymptotic regimes. Extensive simulations show our method outperforms widely used state-of-the-art tests. Our empirical studies indicate that intraday transaction prices from recent years can be considered effectively friction-free at significantly higher frequencies.

1 Introduction

The majority of studies within the field of financial economics are founded upon the well-known proposition that capital markets exhibit efficiency (Fama, 1970, 1991). The Efficient Market Hypothesis has various forms, and assessing the hypothesis is contingent on certain equilibrium models that specify the expected stock returns. However, over short intervals, such as daily or intraday periods, expected returns are negligible in efficient markets. In such cases, market efficiency implies that the autocovariances of returns are close to zero. Thus,

*Our research has significantly benefited from valuable conversations with Yacine Aït-Sahalia, Torben Andersen, Federico Bandi, Peter Boswijk, Peter R. Hansen, Arthur Lewbel, Jia Li, Ming Li, Yingying Li, Whitney Newey, Alexei Onatski, Andrew Patton, Roberto Renò, Shuping Shi, Liangjun Su, Dacheng Xiu, and Jun Yu. We are especially grateful to Oliver Linton for illuminating discussions during the early stages of this paper. We thank the participants from the China Meeting of the Econometric Society (July 2021) and the Fifteenth Annual SoFiE Conference (Seoul, 2023) for their invaluable feedback. Additionally, we are grateful for the insights received from scholars at The Chinese University of Hong Kong, University of Cambridge, University of Hong Kong, University of Macau, Nanyang Technological University, National University of Singapore, and Singapore Management University. Li gratefully acknowledges the financial support from the Research Grants Council of Hong Kong (grant number 24500923). Yunxiao Zhai provided excellent research assistance.

†Department of Economics, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China. Email: merrickli@cuhk.edu.hk.

‡Department of Economics, Rutgers University, 75 Hamilton Street, New Jersey Hall, New Brunswick, NJ 08901, USA. Email: xiyeyang@economics.rutgers.edu.

efficiency can be assessed by analyzing the autocovariances of returns (Lehmann, 1990; Fama, 2014).

Therefore, nonzero return autocovariances at finer time scales indicates the presence of market frictions, which induce temporary deviations of asset prices from their efficient or true values (Hasbrouck, 2002). Such frictions can emerge from liquidity provision (Hasbrouck, 2007), behavioral fads (Lehmann, 1990), or herding in trading activity (Cipriani and Guarino, 2014). Detecting market frictions by examining return autocovariances appears straightforward. The absence of such frictions implies that autocovariances of returns should equal zero across *all* horizons, i.e., $\mathbf{Cov}(Y_i - Y_{i-k}, Y_{i+k} - Y_i) = 0$ for all $k \geq 1$, where $\{Y_i\}_i$ represents a sequence of observed logarithmic prices. Thus, detecting a single horizon k with nonzero autocovariances is sufficient to confirm the presence of market frictions. However, empirical studies have documented nonzero covariances over various horizons (k values), suggesting that examining returns over any *single* horizon may often be inadequate. Examining a wider range of horizons seems essential. This is a common practice in empirical research studies.[1]

We propose a simple method to detect market frictions using autocovariances of returns across *multiple* horizons. Our testing method is motivated by classic asset pricing literature, which shows that various forms of market frictions lead to return *reversals*, and these reversals are expected to produce *negative* autocovariances in returns over specific horizons. These patterns have been empirically observed for decades as evidence against market efficiency, supported by studies such as Roll (1984); Summers (1986); Fama and French (1988); Jegadeesh (1990); Lehmann (1990); Fama (1991); Huang et al. (2010); Bogousslavsky (2016). However, the existing literature falls short in addressing the selection of horizons and offers even less discussion on the aggregation of information across various horizons.

We demonstrate that analyzing the covariance between $Y_i - Y_{i-k}$ and $Y_{i+k} - Y_i$ for a small set of k values, such as $k \in \{1, 2, 3, 4\}$, is sufficient to detect a broad class of market frictions. To capture the most significant return reversals within the selected small set of horizons, we propose a *minimum* statistic that identifies the most negative autocovariances. The minimal statistic effectively identifies inefficiencies driven by the complex behavior of the market frictions, which can involve unknown serial dependencies, time-varying and stochastic patterns, and endogeneity with respect to the efficient prices. Furthermore, we show that our testing method can be applied to both intraday high-frequency data and interday low-frequency data, each employing different asymptotic regimes. Such robustness greatly broadens the applicability of our approach to diverse empirical scenarios.

Therefore, our method contributes to empirical asset pricing by offering a clear and robust solution to the enduring empirical challenge of return reversals or predictability observed across different horizons. This variability may result in conflicting interpretations of market efficiency when disparate horizons are employed. We demonstrate that analyzing and aggregating information from a small set of horizons is sufficient, thus obviating the need for an exhaustive exploration of return autocovariances across excessive horizons.

[1]For example, using monthly returns, Poterba and Summers (1988) examines return covariances over 96 horizons.

We also contribute to the classic literature on testing for serial correlation. The primary strength of our method is its ability to identify alternative directions with minimal tuning of parameters, alongside its robustness in detecting weak market frictions—those that are asymptotically dominated by efficient returns. In simulations, the test performs well in finite samples. We compare our tests against several widely used state-of-the-art tests, including modern updates of the classic portmanteau test and various versions of variance ratio tests. Across a comprehensive range of alternative models, our method outperforms *all* other tests in terms of both power and size control.

Another major contribution is our provision of an effective tool for conducting econometric inferences using intraday data, which has gained popularity over the past two decades. Intraday transaction prices deviate from efficient prices due to various microstructural effects. Econometric methods for intraday data differ significantly depending on the presence or absence of such microstructure frictions (Aït-Sahalia and Jacod, 2014). Therefore, a statistical test to evaluate the conformity of observed prices to the underlying efficient price is crucial for selecting the proper analytical approach.[2] While conventional practice subsamples intraday returns at 5- to 15-minute intervals to reduce microstructure frictions, our analysis shows that recent high-frequency data can be treated as effectively friction-free at much higher sampling frequencies. This allows empirical studies to use finer-sampled transaction prices, yielding larger datasets for more accurate estimation of fundamental quantities like volatility (Andersen et al., 2003), market betas (Bollerslev et al., 2024), and other key financial metrics.[3]

Our paper also contributes to a burgeoning body of literature that examines the intricate dynamics of asset prices that the standard jump-diffusion models cannot capture. Recent studies provide increasing evidence on temporary explosive drifts and gradual jumps that lead to persistent returns (Christensen et al., 2022; Andersen et al., 2023). Related studies in this area include Laurent et al. (2024); Shi and Phillips (2024), who employ return autocovariances to measure ”realized drift.” Additionally, Kolokolov et al. (2025) develop block BUMVU theory and propose a drift test emphasizing covariance across overlapping blocks. Chong and Todorov (2025) employs short-lag autocovariances, applying them to volatility increments to test for rough volatility.

The structure of the paper is as follows. Section 2 outlines the setup of the efficient prices and market frictions. Section 3 introduces the testing method and discusses the asymptotic properties of the testing statistics. Section 4 discusses several extensions of the setup and the testing method. Section 5 presents extensive numerical studies that illustrate and compare the finite sample performance of our method against other testing methods. Section 6 demonstrates the application of our method using empirical data. Section 7 concludes the pa-

[2] A graphic tool to examine the presences of deviations, or microstructure frictions in high frequency data is the well-known *volatility signature plot* (Andersen et al., 2000). Aït-Sahalia and Xiu (2019) proposed several formal tests for the presence of i.i.d. deviations based on the Hausman principle (Hausman, 1978) in volatility estimation of the efficient returns.

[3] For example, our empirical findings suggest that one can focus on intraday returns over intervals of 1 minute or even shorter to estimate volatility. The advantage of utilizing these larger subsamples, in contrast to sampling every 5 or 15 minutes, is the potential for negligible bias from weak microstructure frictions.

per. The Appendix explains some technical details and provides additional theoretical results. Supplementary materials Li and Yang (2025) include mathematical proofs for various technical lemmas and additional simulation results. The code to implement the tests is available online.[4]

2 Model Setup

We consider a fixed time period $[0, t]$, which could be a trading day or a year. Let $Y_i^n, i = 1, 2, \ldots, n_t$ be the observed logarithmic prices collected within the period $[0, t]$, where n_t is the number of observations upon time t.

We will adopt a *continuous-time* framework for the underlying price processes. Let X denote the logarithmic efficient price process, modeled as a standard Brownian semimartingale,[5] which can be expressed as

$$X_t = X_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s. \tag{1}$$

The drift b and volatility σ processes are locally bounded. We will denote the efficient price at time $i\Delta_n$ as $X_i^n := X_{i\Delta_n}$, where $\Delta_n := t/n_t$ is the spacing of observations. In the *infill* limit, as $n_t \to \infty$ with t fixed, we have $\Delta_n \to 0$.

Price inefficiencies are captured by temporary deviations from the efficient price, which are modeled as latent *pricing errors* in the observed prices, reflecting various *market frictions* that arise in the trading process. The pricing errors are commonly attributed to two main factors: (1) behavioral aspects, such as cognitive biases, heuristics, and other irrational behaviors (Lehmann, 1990; De Long et al., 1990; Shiller, 2003), and (2) market frictions arising from elements like inventory controls, bid-ask bounce, transaction costs, and other trading-related impediments (Hasbrouck, 2004).

Let ε_i^n be the pricing error associated with the i-th transaction, defined as

$$\varepsilon_i^n := \Delta_n^\theta \cdot \chi_i, \quad \theta \in [0, 3/4), \tag{2}$$

where χ is a stationary ρ-mixing sequence with mean 0, variance γ^2, independent of the efficient price process X, and has an autocorrelation function (ACF) $r_{(\cdot)}$ of unknown form. We impose no restrictions on the ACF, except that the ρ-mixing coefficients decay at a polynomial rate controlled by a parameter $v > 0$.[6] The ACF can exhibit an arbitrarily slow decay rate. Thus, it can capture the gradually diminishing deviations from efficient prices as modeled

[4] `R`, `Matlab`, and `Python` packages can be found at `https://github.com/merrickli/minimac`.

[5] Jumps will be discussed in Section 4.

[6] Specifically, we can define two discrete filtrations $\mathcal{G}_p := \sigma\{\chi_j : p \geq j\}$, $\mathcal{G}^q := \sigma\{\chi_j : q \leq j\}$. For any positive integer j, we assume the *mixing coefficients* satisfy $\rho_j \leq Cj^{-v}$ for some $C > 0, v > 0$, where the coefficients are defined by

$$\rho_j := \sup\left\{|\mathbb{E}(V_h V_{j+h})| : \mathbb{E}(V_j) = \mathbb{E}(V_{j+h}) = 0, \mathbf{Var}(V_h) \leq 1, \mathbf{Var}(V_{j+h}) \leq 1, V_h \in \mathcal{G}_h, V_{j+h} \in \mathcal{G}^{j+h}\right\}.$$

by Summers (1986); Poterba and Summers (1988). In the presence of pricing errors, the i-th observed price is the sum of two latent components: $Y_i^n = X_i^n + \varepsilon_i^n$.

Note that in (2), we introduce a factor Δ_n^θ to control the relative scale of pricing errors compared to the efficient returns, which are of order $\sqrt{\Delta_n}$. Given that we assume $\Delta_n \to 0$ as the sample size increases, the errors are asymptotically dominating (or dominated by) the efficient returns when $\theta \in [0, 1/2)$ (or $\theta \in (1/2, 3/4)$, respectively). In this paper, we adopt the following classification:

Definition 2.1 (Weak and strong pricing errors)**.** *Pricing errors defined in* (2) *are considered* weak *when $\theta > \frac{1}{2}$ and* strong *when $\theta \leq \frac{1}{2}$.*

The econometrics literature (Aït-Sahalia and Jacod, 2014) on intraday data often sets $\theta = 0$, reflecting the rationale that microstructural effects, such as bid-ask bounces, dominate ultra-high-frequency returns. Conversely, when $\theta = 1/2$, the errors and efficient returns share the same asymptotic scale. In this scenario, we can rescale the prices using $\sqrt{\Delta_n}$. The rescaled prices and the errors resemble the *permanent-transitory* decompositions frequently used in low-frequency settings in asset pricing and macroeconomics. We establish that our statistics remain invariant under this rescaling, allowing their application to both high-frequency and low-frequency data. Further details can be found in Section 3.4.

Our testing device is capable of detecting weak errors for θ values up to $\frac{3}{4}$. This feature sets our approach apart from many existing tests, such as the portmanteau test (Escanciano and Lobato, 2009), the variance ratio tests (Lo and MacKinlay, 1988) in low-frequency settings and the volatility methods applied in high-frequency frameworks (Andersen et al., 2000; Bandi and Russell, 2008; Aït-Sahalia and Xiu, 2019), where weak errors asymptotically vanish in variances or squared terms. Instead, we utilize a *covariance* approach in our method, and we illustrate that this approach remains effective for identifying weak pricing errors.[7] The broader range of detectable errors in theory is corroborated by our simulation experiments, where our tests prove more effective at identifying weak errors in finite samples.

Although strong and weak errors are largely indicative of asymptotic conditions, empirical scenarios often display pricing errors with varying magnitudes, rendering the concepts practically relevant. For example, subsampling is frequently used with high-frequency intraday data. In these coarser samples, the transitory deviations reflecting microstructural errors tend to diminish compared to efficient returns, illustrating a case of weak errors.

3 The Test Statistics and Large Sample Properties

In the absence of market frictions, the observed prices align with efficient prices, indicating no deviations whatsoever. Consequently, the variance of pricing error ε will equal zero. Mathe-

[7] Recent literature in high-frequency econometrics has revived the covariance approach for various applications, such as estimating microstructure noise (Li and Linton, 2022) and measuring integrated drift (Laurent et al., 2024; Shi and Phillips, 2024; Andersen et al., 2023).

matically, this hypothesis can be expressed as:

$$\mathbb{H}_0 : \gamma^2 = 0.$$

Under the alternative hypothesis, the observed prices swing away from the efficient prices. This is represented as:

$$\mathbb{H}_1 : \gamma^2 > 0.$$

This section presents the design of our test statistics aimed at distinguishing between the aforementioned hypotheses. Our testing method is inspired by classic asset pricing literature, which examines return autocovariances or autocorrelations to assess market efficiency. Given our emphasis on finer time scales—such as weekly, daily, and even intraday periods—expected returns remain nearly constant and, in fact, close to zero. Consequently, under the hypothesis of frictionless markets, observed returns show zero correlations across all horizons. However, the presence of pricing errors in asset prices induces negative autocorrelations in returns. This is particularly evident when the pricing errors are independently and identically distributed (i.i.d.). To see this, we note that the first-order autocovariance of the observed returns is given by[8]

$$\mathbf{Cov}\big(Y_{i+1}^n - Y_i^n, Y_i^n - Y_{i-1}^n\big) \approx \mathbf{Cov}\big(\varepsilon_{i+1}^n - \varepsilon_i^n, \varepsilon_i^n - \varepsilon_{i-1}^n\big) = -\mathbf{Var}(\varepsilon_i^n).$$

However, if the errors display strong persistence, negative autocovariances become more pronounced only over longer horizons as noted by Summers (1986) and Fama and French (1988). Recent studies (Chordia et al., 2005; Huang et al., 2010) also confirm that return reversals occur across various horizons. With this in mind, we will develop a multi-horizon testing approach in this section. This new method provides an effective tool for selecting return horizons, over which it examines return reversals to identify a broad spectrum of pricing errors. Surprisingly, the range of horizons to consider is quite small.

3.1 The testing statistics over a single horizon

We will analyze returns across different horizons. To begin, let's focus on a single horizon indexed by a positive integer k. Given a series of observed prices $\{Y_i^n\}_{i=1}^{n_t}$, we define the following statistics

$$F(Y;k)_t^n := \frac{1}{2k} \sum_{i=k+1}^{n_t-k} f(Y;k)_i^n,$$

[8]The first approximation follows from the martingale difference property of the efficient returns, and the fact that the expected returns—reflected in the drift terms—are negligible at finer time scales.

where $f(Y;k)_i^n := (Y_i^n - Y_{i-k}^n)(Y_{i+k}^n - Y_i^n)$.[9] The following theorem establish the limiting distribution under the $\mathbb{H}_0$.[10]

Theorem 3.1. *We have the following limiting distribution under $\mathbb{H}_0$ for a given horizon k:*

$$\frac{F(Y;k)_t^n}{\sqrt{\Delta_n}} \xrightarrow{\mathcal{L}_s} \mathcal{Z}_t,$$

where $\mathcal{Z}_t$ is a (conditional) Gaussian variable centered at 0, with conditional covariance given by $\Phi_k \int_0^t \sigma_s^4 \mathrm{d}s$, where $\Phi_k := \frac{1}{6}\left(k + \frac{1}{2k}\right)$.

Let $Q(Y)_t^n := \sum_{i=2}^{n_t-1} (f(Y;1)_i^n)^2$, we introduce the standardized testing statistic[11]

$$H(Y;k)_t^n := \frac{F(Y;k)_t^n}{\sqrt{\Phi_k\, Q(Y)_t^n}}. \tag{3}$$

Theorem 3.1 readily implies the following convergence result , which further gives the size of the single horizon statistic:

Corollary 3.1. *For any $t > 0$, we have the following limiting distribution:*

$$H(Y;k)_t^n \xrightarrow{\mathcal{L}_s} \mathcal{N}(0,1)\,. \tag{4}$$

Let z_α be the α-quantile of the standard normal distribution. The statistic $H(Y;k)_t^n$ has asymptotic size α under $\mathbb{H}_0$:

$$\mathbb{P}(H(Y;k)_t^n < z_\alpha \,|\, \mathbb{H}_0) \to \alpha.$$

3.2 The rationale behind power: return horizons and reversals

In the presence of pricing errors, it is expected that the covariances of returns will deviate from zero at specific lags or horizons. A natural approach to investigating such deviations involves comparing the squared autocovariances of returns across different lags or horizons. The comparison of squared autocovariances at individual lags aligns with the framework of portmanteau tests for zero autocorrelations (Box and Pierce, 1970; Ljung and Box, 1978; Escanciano and Lobato, 2009). On the other hand, examining autocovariances over broader horizons

[9]The division by $2k$ is for ease of calculation. Following the analysis in Lo and MacKinlay (1988), we can construct a sequence of non-overlapping statistics indexed by the starting observation index $q_k = 1, \ldots, 2k$: $F(Y;k,q_k)_t^n := \sum_{j=1}^{n(q_k)_t} f(Y;k)_{q_k+(2j-1)k}^n$, where $n(q_k)_t := \lfloor (n_t - q_k)/(2k) \rfloor$ and $\lfloor \cdot \rfloor$ is the floor function. The statistic $F(Y;k,q_k)_t^n$ is based on non-overlapping intervals, which simplifies the analysis using standard results on martingale differences. Consequently, our overlapping statistic $F(Y;k)_t^n$ becomes the average of the $2k$ non-overlapping statistics. That is, $F(Y;k)_t^n = \frac{1}{2k}\sum_{q_k=1}^{2k} F(Y;k,q_k)_t^n$.

[10]The notation $\xrightarrow{\mathcal{L}_s}$ denotes *stable convergence in law*, which is a slightly stronger notion than *convergence in law*. We require this stronger convergence here since the limiting variance process is stochastic. Stable convergence of the empirical process is preserved when it is normalized by a consistent estimator of the limiting variance. Such convergence is not guaranteed by standard convergence in law.

[11]Note that $H(Y;k)_t^n$ can also be rewritten as $H(Y;k)_t^n = \frac{F(Y;k)_t^n}{\sqrt{\Delta_n \Phi_k \widetilde{Q}(Y)_t^n}}$, where $\widetilde{Q}(Y)_t^n := Q(Y)_t^n/\Delta_n$ is a consistent estimator of the asymptotic variance in Theorem 3.1.

is consistent with the variance ratio tests for martingale difference sequences (Lo and MacKinlay, 1988; Cochrane, 1988; Deo and Richardson, 2003). These tests are inherently *two-sided*. We now show that a more powerful one-sided test can be conducted to detect pricing errors.

The economic intuition of the directional test is straightforward: pricing errors, as *transitory* components of observed prices, exert only temporary effects on asset prices. Consequently, returns will eventually revert. Thus, the statistic $F(Y;k)_t^n$, after some normalization, should converge to a probability limit that is negative for some k. We will show that the ACF of the pricing error process plays a crucial role in determining the "speed" of this reversion.

Let's revisit the special ACF where pricing errors follow an i.i.d. process. Then, we immediately observe $\mathbf{Cov}\big(Y_i^n - Y_{i-k}^n, Y_{i+k}^n - Y_i^n\big) = \mathbf{Cov}\big(\varepsilon_i^n - \varepsilon_{i-k}^n, \varepsilon_{i+k}^n - \varepsilon_i^n\big)$, as efficient returns are uncorrelated. The i.i.d. property of the pricing errors further simplifies the covariance to $-\mathbf{Var}(\varepsilon_i^n)$. Thus, the *covariance* of returns over any k horizon is the negative of the pricing errors' *variance*. That is, return reversals occur at *any* horizon.

For a general ACF $r_{(\cdot)}$, the covariance of the pricing error differences reduces to

$$\mathbf{Cov}\big(\varepsilon_i^n - \varepsilon_{i-k}^n, \varepsilon_{i+k}^n - \varepsilon_i^n\big) = -\mathbf{Var}(\varepsilon_i^n) \cdot g(k,r), \ \text{ where } g(k,r) := r_0 - 2r_k + r_{2k}. \tag{5}$$

Return reversals will be observed if $g(k,r) > 0$. The following theorem addresses the joint conditions on k and r that ensure a positive $g(k,r)$ value. Surprisingly, for a *broad* range of pricing errors (characterized by their ACFs $r_{(\cdot)}$), return reversals will emerge over a *small* set of horizons.

Theorem 3.2. *Let $\mathcal{R}(\ell) := \{r_{(\cdot)} : r_1 \leq 1 - 2^{-\ell}\}$ for some nonnegative integer ℓ. Then for any $r_{(\cdot)} \in \mathcal{R}(\ell)$, there exists some $k \in \mathbf{K}_\ell := \{1, \ldots, 2^\ell\}$ such that $g(k,r) > 0$.*

According to this theorem, return reversals (with $k = 1$) are immediately detected if $r_1 \leq 0$, and the i.i.d. case emerges as a specific example. If r_1 is allowed to be positive up to 0.5, checking returns over an additional horizon with $k = 2$ suffices. Extending the range for r_1 up to 0.75 requires checking returns over two more horizons with $k = 3, 4$; this class of ACFs encompasses most parametric models in empirical studies. Even considering extremely slowly decaying pricing errors as discussed in Summers (1986); Poterba and Summers (1988), Theorem 3.2 shows it is unnecessary to examine beyond a wide range of horizons. Checking for $k \in \{1, 2, \ldots, 8\}$ is sufficient to detect pricing errors with r_1 nearly 0.9!

Theorem 3.2 implies that considering a small set of horizons suffices under weak conditions. We now provide further elaboration on this theorem. We rewrite $g(k,r) = 2\left(\frac{r_0 + r_{2k}}{2} - r_k\right)$. From this expression, it becomes clear that $g(k,r) \leq 0$ if the ACF maintains *concavity*. However, being bounded between -1 and 1, the concavity of the ACF cannot be sustained indefinitely because it will eventually intersect with the lower bound. The left panel of Figure 1 illustrates this phenomenon by plotting four different ACFs with different values of r_1, from which we observe that the concavity of the ACF will be violated after a few lags when it cuts the lower bound. The figure also demonstrates that a smaller r_1 (black dots in the left figure) diminishes the interval over which concavity–and thus negativity in $g(k,r)$–is maintained. This implies

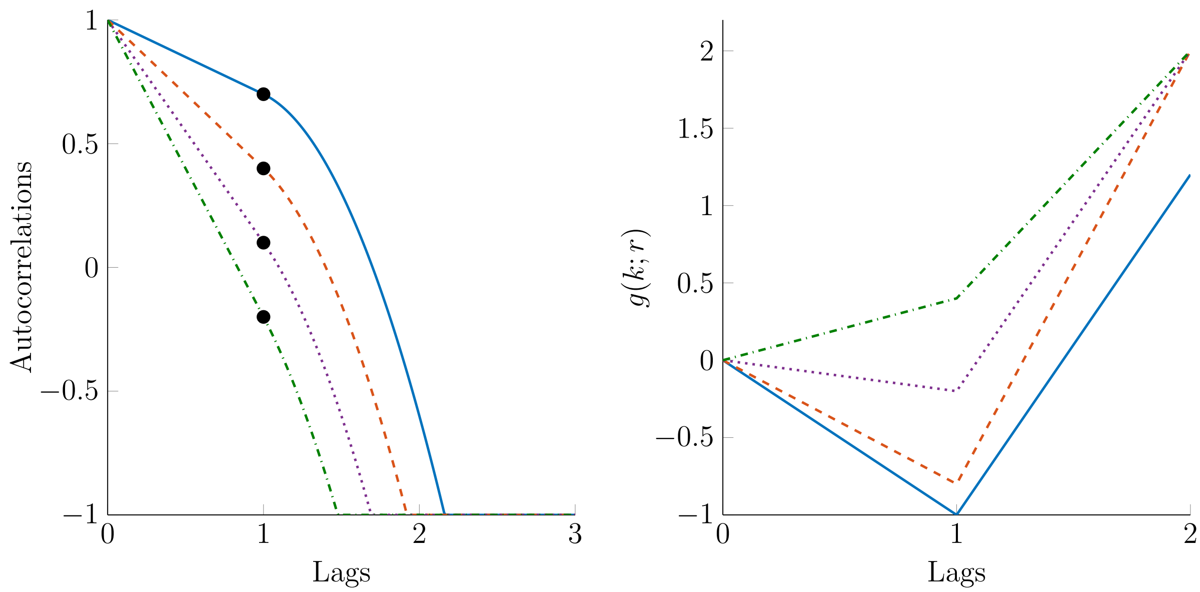

Figure 1: Graphic illustration of Theorem 3.2: ACF plots (left) and the $g(\cdot)$ function plots (right). The ACFs and their $g(\cdot)$ functions are matched by the line styles.

that as r_1 decreases, the interval for identifying a k such that $g(k, r) > 0$ becomes narrower, which aligns with the predictions of Theorem 3.2. Indeed, the right panel of Figure 1 shows that $g(k, r)$ is more likely to be positive when r_1 decreases. Moreover, given ACFs range in $(-1, 1)$, we have $g \in (-2, 4)$. This asymmetry also suggests greater potential power for left-sided tests (positive g) compared to right-sided tests (negative g).

The upcoming theorem demonstrates the consistency of our testing statistics $H(Y; k)_t^n$ that relies on having a positive g value.

Theorem 3.3. *Suppose $r_{(\cdot)} \in \mathcal{R}(\ell)$ and $k \in \mathbf{K}_\ell$ as stated in Theorem 3.2. Then, $H(Y; k)_t^n$ is consistent under $\mathbb{H}_1$:*

$$\mathbb{P}(H(Y; k)_t^n < z_\alpha \,|\, \mathbb{H}_1) \to 1.$$

Theorem 3.3 provides a theoretical basis for the consistency of the test statistics $H(Y; k)_t^n$ under the alternative hypothesis. However, implementing the test requires an explicit choice of the horizon k from a finite set, as per Theorem 3.2. Furthermore, one might consider selecting k from a larger set, since Theorem 3.2 suggests that the set of ACFs that can be detected expands as ℓ increases. We first provide a formal analysis that a large k is not favored.

We define the a function $\overline{g}$ for a horizon k and an ACF $r_{(\cdot)}$: $\overline{g}(k; r) := \frac{-g(k,r)}{\sqrt{k^3+k/2}}$. The function $\overline{g}$ captures all aspects related to k in the limit of $H(Y; k)_t^n$ under $\mathbb{H}_1$, as one can observe from the definition of the testing statistic $H(Y; k)_t^n$ in (3). Intuitively, a more negative $\overline{g}(k; r)$ will yield higher power for our left-sided test. On one hand, Theorem 3.2 ensures the existence of k within a small set such that $-4 < -g(k, r) < 0$ for a broad range of ACFs. On the other hand, the denominator of $\overline{g}$ increases at a rate of $k^{3/2}$. Therefore, $\overline{g}$ will reach its minimum (negative) value at a small k for most ACFs. This explains the preferred choice of a small set of

k's in practice, such as $k \in \{1, 2, 3, 4\}$. Further simulation evidence, provided in Section 5, will justify the effectiveness of using such a small set of k's.

Given the small set of horizons under consideration, finding the optimal one to boost power seems appealing. However, achieving this optimality requires additional parametric assumptions about the ACF and precise estimates of the underlying ACFs, which are often unavailable in practice.

We advocate an approach based on testing statistics over multiple horizons. This approach not only alleviates the burden of identifying a single optimal k, but also allows us to explore additional information from various time horizons. As shown in the simulation study, the multiple horizon method is simple to implement and offers robust power against diverse pricing errors.

3.3 The multi-horizon test

In this subsection, we will demonstrate the multi-horizon testing approach. Let k, k' be two positive integers that represent two horizons. We need the following quantities to describe the result (where $\mathbf{1}_{\{\cdot\}}$ is the indicator function):

$$\overline{\Phi}_{k,k'} := k^3 - \frac{1}{3}\left[(2k-k')^3 - (2k-k')\right]\mathbf{1}_{\{k' \le 2k\}}, \;\; \Phi_{k,k'} := \frac{\overline{\Phi}_{k,k'}}{4kk'}, \;\; \rho_{k,k'} := \frac{\Phi_{k,k'}}{\sqrt{\Phi_k \Phi_{k'}}}.$$

Note that when $k = k'$, $\Phi_{k,k'}$ reduces to Φ_k that is defined in Theorem 3.1.

Theorem 3.4. *Let $k_1, \ldots, k_m$ be a set of positive integers. We have the following joint limiting distribution under $\mathbb{H}_0$:*

$$\left(\frac{F(Y;k_1)_t^n}{\sqrt{\Delta_n}}, \ldots, \frac{F(Y;k_m)_t^n}{\sqrt{\Delta_n}}\right) \xrightarrow{\mathcal{L}_s} \left(\mathcal{Z}_t^1, \ldots, \mathcal{Z}_t^m\right),$$

where the limting variables $\left(\mathcal{Z}_t^1, \ldots, \mathcal{Z}_t^m\right)$ are (conditionally) joint Gaussian with mean zero and conditional covariances $\Phi_{k_j,k_{j'}} \int_0^t \sigma_s^4 \mathrm{d}s$ for $1 \le j \le j' \le m$.

Remark 3.1. *It is important to note that while the limiting distribution presented in Theorem 3.1 is not surprising, the joint distribution derived in Theorem 3.4 is, to the best of our knowledge, new to the literature. This is particularly true for the non-trivial correlations $\Phi_{k_j,k_{j'}}$.*

Corollary 3.2. *We have*

$$(H(Y;k_1)_t^n, \ldots, H(Y;k_m)_t^n) \xrightarrow{\mathcal{L}_s} (\Theta_1, \ldots, \Theta_m),$$

where $(\Theta_1, \ldots, \Theta_m)$ are standard normal variables with correlation $\mathbf{corr}\left(\Theta_j, \Theta_{j'}\right) = \rho_{k_j,k_{j'}}$ *for $1 \le j, j' \le m$.*

We propose a minimum statistic to effectively capture return reversals, which serves as evidence of pricing errors. Let $\mathbf{K} = (k_1, k_2, \ldots, k_m)$ be an m-tuple of positive integers. We define $\Sigma_{\mathbf{K}}$ as the law of a multivariate normal distribution with mean zero and a covariance matrix given by $\rho_{k_i,k_j}, 1 \leq i,j \leq m$, as defined in Corollary 3.2. The minimum statistic is defined as

$$H(Y;\mathbf{K})_t^n := \min_{k_i \in \mathbf{K}} H(Y;k_i)_t^n. \tag{6}$$

Note that when $\mathbf{K} = \{k\}$ is a singleton, the statistic $H(Y;\mathbf{K})_t^n$ is consistent with $H(Y;k)_t^n$ introduced in (3).

Based on Theorem 3.4 and Corollary 3.2, we can derive the limiting distribution of the minimum statistic for any arbitrary $\mathbf{K}$ under the null hypothesis. In contrast, under the alternative hypothesis, we can identify return reversals over a small set of horizons, as predicted by Theorem 3.2. Consequently, the minimum statistic will converge to the left tail when pricing errors are present. Recall $\mathbf{K}_\ell := \{1, 2, \ldots, 2^\ell\}$ in Theorem 3.2.

Theorem 3.5. *Given $\alpha \in (0,1)$, let c_α be the critical value that satisfies $\Sigma_{\mathbf{K}_\ell}(x_1 \geq c_\alpha, \ldots, x_m \geq c_\alpha) = 1 - \alpha$. The statistic $H(Y;\mathbf{K}_\ell)_t^n$ has asymptotic size α under $\mathbb{H}_0$:*

$$\mathbb{P}(H(Y;\mathbf{K}_\ell)_t^n < c_\alpha \,|\, \mathbb{H}_0) \to \alpha.$$

When $r_1 \leq 1 - 2^{-\ell}$, $H(Y;\mathbf{K}_\ell)_t^n$ is consistent under $\mathbb{H}_1$:

$$\mathbb{P}(H(Y;\mathbf{K}_\ell)_t^n < c_\alpha \,|\, \mathbb{H}_1) \to 1.$$

Theorem 3.5 states that it is adequate to examine a small number of horizons to detect pricing errors with a wide range of ACFs. Moreover, we can pinpoint the *exact* horizons to ensure an effective test of *targeted* alternatives. For instance, if the ACF of the pricing error at lag 1 is believed to be mildly positive and bounded by 0.5, checking two horizons, $k \in \{1, 2\}$, is sufficient.

Therefore, our approach is nearly independent of tuning parameter selection. This freedom differs from data-driven methods like the automatic portmanteau test (Escanciano and Lobato, 2009) or the automatic variance ratio test (Choi, 1999). We choose a small set of horizons, serving as our tuning parameters, that perform universally across a broad range of alternative hypotheses. This feature contributes to the robustness and accuracy of our testing method, as demonstrated in the simulation study in Section 5.

3.4 Beyond the infill framework: The scenario $\theta = \frac{1}{2}$

Up to this point, discussions have assumed an infill setting, which implies that as the sample size tends to infinity, the time spacing between observations, Δ_n, approaches zero in the limit. This might suggest our test method is restricted to intraday high-frequency data, where efficient returns of order $\sqrt{\Delta_n}$ disappear in the limit. We will now show that our method is applicable beyond this context.

The flexibility of our method arises from the adaptable modelling of the pricing errors through the parameter θ. Recall that the deviation of the i-th transaction price from the underlying efficient price is $\varepsilon_i^n = \Delta_n^{\theta} \cdot \chi_i$, with $\theta \in [0, 3/4)$. The high-frequency econometrics literature (Aït-Sahalia and Jacod, 2014) typically emphasizes strong deviations, i.e., $\theta \in [0, 1/2)$. We now explore the implications when $\theta = 1/2$.

When $\theta = \frac{1}{2}$, the pricing errors and the efficient returns have the same asymptotic orders. Denote the scaled prices by $\widetilde{V}_i := V_i^n / \sqrt{\Delta_n}$, $V = X$ or Y. Then the additive model $Y_i^n = X_i^n + \varepsilon_i^n$ can be expressed as:

$$\widetilde{Y}_i = \widetilde{X}_i + \chi_i, \tag{7}$$

where $\widetilde{X}$ is a random walk with time-varying drift and volatility. This formulation connects the classic *permanent-transitory decomposition* commonly used in macroeconomics and asset pricing literature. In this model, the permanent component follows a random walk, while the transitory component behaves as a stationary process, typically modelled as an i.i.d. or AR(1) process.[12] Thus, the model in (7) can be seen as the counterpart in a long-span, fixed-frequency setting, where the efficient returns and deviations are of the same order, with neither vanishing as the sample size grows.[13]

A key observation is the rescaling-invariance property for our test statistics: $H(V; \mathbf{K})_t^n = H(\widetilde{V}; \mathbf{K})_t^n$ for $V = X$ or Y. This property leads to the robustness to the asymptotic regimes and data frequencies, and it guarantees the validity and effectiveness of our test across a wide range of data settings, whether applied to intraday noisy high-frequency data with increasingly dominant errors ($\theta < 1/2$) or to lower-frequency data, such as daily returns, where the pricing errors do not necessarily dominate as the sample size grows ($\theta \geq 1/2$). The following corollary follows directly from this invariance property and Theorem 3.5. The assumed data generating process accommodates a broad class of permanent–transitory decomposition models commonly used in the low-frequency time series literature.

Corollary 3.3. *Suppose that $Y_i = X_i + \varepsilon_i$, where $\{\varepsilon_i\}_{i=1}^n$ is a stationary mixing sequence with variance γ^2 and ACF satisfying $r_1 \leq 1 - \frac{1}{2\ell}$ for some positive integer ℓ. The efficient price process X evolves as $X_i = X_{i-1} + e_i$, where $\{e_i\}_{i=1}^n$ is a sequence of i.i.d. innovations independent of $\{\varepsilon_i\}_i$.*

Let $\alpha \in (0, 1)$, and define c_α and $\Sigma_{\mathbf{K}_\ell}$ as in Theorem 3.5. Then, under the null hypothesis $\mathbb{H}_0$, the test statistic $H(Y; \mathbf{K}_\ell)_t^n$ has asymptotic size α as $n \to \infty$: $\mathbb{P}(H(Y; \mathbf{K}_\ell)_t^n < c_\alpha \mid \mathbb{H}_0) \to \alpha$. Moreover $H(Y; \mathbf{K}_\ell)_t^n$ is consistent under the alternative hypothesis $\mathbb{H}_1$: $\mathbb{P}(H(Y; \mathbf{K}_\ell)_t^n < c_\alpha \mid \mathbb{H}_1) \to 1$.

[12] This structure is well-documented in studies such as Summers (1986); Cochrane (1988); Fama and French (1988); Hasbrouck (1993); Cochrane (1994); Corwin and Schultz (2012); Hendershott and Menkveld (2014).

[13] While Δ_n approaches zero, it functions merely as a scalar and does not have any substantial impact beyond this role.

4 Extensions

In this section, we discuss various extensions of our framework. These extensions are both theoretically and empirically relevant, and they demonstrate the versatility of our testing method. The technical details and regularity conditions for these extensions are provided in Appendix A. The asymptotic properties of our testing method under these extensions are given in Appendix B, which encompasses the results from Section 3 as special cases.

4.1 Stochastic observation scheme

Our assumption that observed prices are equally spaced by Δ_n holds for interday prices, such as daily or weekly data, as well as for high-frequency data collected at regular intervals. However, this scheme may not adequately represent all intraday prices, where transaction times are often irregular and stochastic. We now broaden the regular observation scheme to accommodate a general stochastic scheme, as established in Jacod et al. (2017, 2019); Li and Linton (2022).

The randomness of the observation scheme is governed by a positive stochastic process α, which serves as the observation density process controlling the random observation spacing relative to the regular spacing Δ_n. In simple terms, around time s, observations occur more frequently than the regular scheme when $\alpha_s > 1$, and less frequently when $\alpha_s < 1$.

The random observation times $\{T_i^n\}_i$, random spacings $\{\Delta(n,i)\}_i$, and observation densities $\{\alpha_{T_i^n}\}_i$ are related as follows: Assume $T_0^n = 0$. For any $i \geq 1$, the random observation duration is approximately $\Delta(n,i) \approx \Delta_n / \alpha_{T_{i-1}^n}$. Hence, the observation times are determined by $T_i^n = \sum_{j \leq i} \Delta(n,j)$. The total number of observations by time t is given by $n_t = \sum_i \mathbf{1}_{\{T_i^n \leq t\}}$. We define a cumulative observation density process

$$A_t := \int_0^t \alpha_s \mathrm{d}s,$$

which will be used to describe the limiting processes under this random observation scheme. The detailed regularity conditions on α is presented in Appendix A.2.

4.2 Jumps

We can accommodate jumps in the efficient price process by imposing mild conditions to restrict jump activities. Employing the truncation method (Mancini, 2001; Li et al., 2013), we eliminate jumps by carefully setting a threshold value that depends on the moments of observed returns. In the absence of pricing errors, if the absolute efficient returns exceed this threshold, they are identified as jumps in the price level and are removed. As a consequence, the limiting distribution of our testing statistics will not be changed.[14] With the alternative hy-

[14]This paper's motivation differs from that of estimating volatility-related parameters, where jumps create bias. In our case, including the jump component in our statistic does not introduce any nonzero bias under the null hypothesis, nor does it cause size distortion. However, it does increase the asymptotic variance. Therefore, we

pothesis, the threshold is relatively larger than the pricing errors, allowing them to be retained. Consequently, consistency is also maintained. For the details of the jumps and the truncation method, one is referred to Appendix A.3.

4.3 Stochastic scales in pricing errors

The pricing error process is assumed to be stationary. We can also follow Jacod et al. (2017) to allow the pricing error process to exhibit stochastic variances, capturing time-varying scales in pricing errors. Instead of fixing the variance of the χ process at γ^2, we introduce a nonnegative process γ, so the i-th pricing error is defined as:

$$\varepsilon_i^n = \Delta_n^\theta \cdot \gamma_{T_i^n} \cdot \chi_i.$$

For identification purposes, we normalize the variance of χ to be 1.

The integrated variance of the pricing errors on $[0,t]$ is given by $\int_0^t \gamma_s^2 \mathrm{d}A_s$. Now the null hypothesis posits that $\int_0^t \gamma_s^2 \mathrm{d}A_s = 0$ almost surely, while the alternative hypothesis asserts that $\int_0^t \gamma_s^2 \mathrm{d}A_s > 0$ with positive probability.

4.4 Endogenous pricing errors

A common assumption in the literature is that pricing errors and efficient returns are independent. However, this assumption may not hold in practice and in many leading microstructure models. In the previous subsection, we introduced the random scale process γ to capture time-varying scales in pricing errors. In principle, the γ process can be endogenously determined by the efficient returns; for example, it could be a function of the volatility process of the efficient price. However, we still assume that the stationary pricing error component, χ, is independent of the efficient returns. Thus, while efficient returns and pricing errors may display higher-order dependence via γ, they remain uncorrelated.

We demonstrate that our testing statistics remain valid in a broad setting for endogenous pricing errors that are correlated with the efficient returns. The key element of our testing statistics can be rewritten as:

$$f(Y;k)_i^n = f(X;k)_i^n + f(\varepsilon;k)_i^n + (X_i^n - X_{i-k}^n)(\varepsilon_{i+k}^n - \varepsilon_i^n) + (X_{i+k}^n - X_i^n)(\varepsilon_i^n - \varepsilon_{i-k}^n). \tag{8}$$

Assuming a constant γ, under the alternative hypothesis, we expect $\mathbb{E}(f(X;k)_i^n) \approx 0$ due to the martingale property of X, and $\mathbb{E}(f(\varepsilon;k)_i^n) < 0$ for some k, as shown by Theorem 3.2. Thus, if the covariance between efficient returns and pricing errors at displacements is not overly positive, such that:

$$\mathbf{Cov}\big(X_i^n - X_{i-k}^n, \varepsilon_{i+k}^n - \varepsilon_i^n\big) + \mathbf{Cov}\big(X_{i+k}^n - X_i^n, \varepsilon_i^n - \varepsilon_{i-k}^n\big) < -\mathbf{Cov}\big(\varepsilon_i^n - \varepsilon_{i-k}^n, \varepsilon_{i+k}^n - \varepsilon_i^n\big),$$

employ the truncation method to deal with intraday data, as it reduces variance and enhance the statistical power of our tests.

our test statistics can still detect pricing errors. According to Theorem 3.2, this condition holds for some k if pricing errors have a large variance, e.g., large γ or asymptotically $\theta < 1/2$. The empirical implication is that for large pricing errors, cross-covariances will not negate the power of our test statistics.

However, with moderate or small pricing errors, positive cross-covariances between efficient returns and pricing error differences may offset the negative covariances between $\varepsilon_i^n - \varepsilon_{i-k}^n$ and $\varepsilon_{i+k}^n - \varepsilon_i^n$. Positive cross-covariances are suggested by many leading microstructure models, such as those by Kyle (1985) and Easley and O'hara (1992). These reflect market makers' learning effects on market dynamics, see Diebold and Strasser (2013); Andersen et al. (2022) for a detailed discussion. If these positive cross-covariances are strong, our left-sided test statistics may fail to detect pricing errors. In such cases, we recommend using two-sided test statistics discussed below.[15]

4.5 Right-sided tests, two-sided tests, and other alternative hypotheses

In this paper, we concentrate on returns over short periods, such as daily or intraday intervals, where the expected returns—represented by the drift term in model (1)—are nearly zero. Consequently, deviations of transaction prices from efficient price lead to return reversals. Our left-sided test statistics are specifically designed to effectively identify these reversals, thereby uncovering deviations and market inefficiencies.

However, recent literature on intraday prices, such as Christensen et al. (2022); Andersen et al. (2023), documents the presence of strong drift or gradual jump terms that lead to locally persistent returns, which invariably introduces short-term inefficiencies. Although such phenomena are less common compared to the temporary deviations examined in this paper, they can still be detected using our test statistics. In fact, it can be shown that our test statistics will exhibit strong positive values, allowing a *right*-sided test to effectively identify these deviations from the efficient price. Section 6 offers an empirical illustration of this application in recognizing extreme drift terms.

For those primarily interested in detecting deviations from efficient prices—regardless of their causes—a two-sided test is appealing. It is effective against a broader spectrum of alternatives. Specifically, for a given set of horizons $\mathbf{K}$ and a significance level α, we can numerically calculate a pair of critical values $L^*_{\mathbf{K},\alpha}$ and $U^*_{\mathbf{K},\alpha}$ that satisfies:

$$\mathbb{P}\big(H(Y,\mathbf{K})_t^n \in [L^*_{\mathbf{K},\alpha}, U^*_{\mathbf{K},\alpha}]^c | \mathbb{H}_0\big) \to \alpha.$$

Moreover, we can find the pair of critical values $L^*_{\mathbf{K},\alpha}$ and $U^*_{\mathbf{K},\alpha}$ such that the associated confidence intervals have the minimal length.[16]

In the supplementary materials Li and Yang (2025), particularly Section S.1.2, we present

[15] Additional simulation studies in the presence of endogenous pricing errors are presented in Appendix S.1 of the supplementary materials Li and Yang (2025).

[16] Note that the distribution of our test statistics under the null hypothesis is not symmetric. Therefore, the confidence intervals are not symmetric under the null hypothesis.

simulation evidence that confirms the robustness of our test statistics across a wider array of alternative models. This includes fractional Brownian motion with diverse Hurst parameters, ARMA processes with varying coefficients, and drift burst processes characterized by different degrees of local explosiveness.

5 Simulation Studies

In this section, we conduct numerical studies to evaluate the finite-sample performance of our test statistics and examine different models of pricing errors under the alternative hypotheses. We will compare our method with several other popular methods in the literature.

5.1 Other competing tests

Under the null hypothesis, returns will be uncorrelated at the finer time scales considered in this paper. Therefore, the classical portmanteau tests developed by Box and Pierce (1970); Ljung and Box (1978) can, in principle, serve as a benchmark for evaluating the performance of our tests in finite samples. These portmanteau tests rely on the (weighted) sum of the squared sample ACFs at lags less than or equal to a positive integer p. Under the null hypothesis, the statistic will following a χ_p^2 distribution. Recent studies, such as Escanciano and Lobato (2009), have developed data-driven methods to select p, which are found robust and more powerful than other tests in practice.

The variance-ratio test (Lo and MacKinlay, 1988; Poterba and Summers, 1988) is a popular approach in the finance and macroeconomics literature. It compares the variance of increments of a series aggregated over multiple periods to the variance of single-period increments. Thus, such test requires an explicit choice of multiple periods or horizons. Various versions of the test have been proposed in the literature. For instance, Choi (1999) introduced an automatic variance ratio test where the optimal value of horizons K is determined through a completely data-dependent procedure. Chow and Denning (1993) proposed testing multiple holding periods simultaneously.

Within the high-frequency econometrics literature, Aït-Sahalia and Xiu (2019) proposed to test for the presence of i.i.d. pricing errors based on the Hausman principle (Hausman, 1978) in volatility estimation. The tests are based on the well-known result that realized volatility estimators are consistent and efficient in the absence of pricing errors. However, these estimators become inconsistent when pricing errors are present. In contrast, the quasi-maximum likelihood estimator remains consistent regardless of the presence or absence of errors.

Compared to the aforementioned methods, our approach offers several advantages. First, although our method is essentially nonparametric, it largely avoids the common tuning parameter selection issues—specifically, the returns horizons—faced by nonparametric techniques. Indeed, we demonstrate that in most practical applications, focusing on a small set of horizons, such as $\{1, 2, 3, 4\}$, is often sufficient. This leads to significant robustness in testing a broad range of alternatives, as we will illustrate in the subsequent subsections. Secondly, our method can handle weak errors—errors that are asymptotically dominated by the efficient returns.[17] The manifestation of this theoretical advantage is that our tests exhibit exceptional sensitivity and power in detecting small-scale errors in practice. A third advantage is that our test naturally connects to market liquidity measurement, providing an economically intuitive interpretation. Additionally, our testing methods offer other benefits, such as robustness to random sampling, endogeneity and serial dependence in pricing errors, as well as computational efficiency—an important advantage for handling large datasets.

We will conduct "horseraces" of our tests against these established methods discussed above: the automatic portmanteau test (AutoQ) proposed by Escanciano and Lobato (2009); variance ratio tests with different horizon choices ($\mathrm{VR}q, q = 2, 4, 10$); the automatic variance ratio test (AutoVR) studied by Choi (1999); the multiple periods version of the variance ratio test developed by Chow and Denning (1993) (CD); and the three tests recommended by Aït-Sahalia and Xiu (2019) (H_{3n}, T_n, AC_n).

5.2 Model specifications and numerical settings

The efficient price process is specified by the following stochastic differential equations:

$$\mathrm{d}X_t = \mu \mathrm{d}t + \sigma_t \mathrm{d}W_t + \xi_t^X \mathrm{d}J_t, \quad \mathrm{d}\sigma_t^2 = \kappa(\vartheta - \sigma_t^2)\mathrm{d}t + \eta\sigma_t \mathrm{d}B_t + \xi_t^\sigma \mathrm{d}J_t,$$

where W and B are two standard Brownian motions with correlation ρ, and J_t is a Poisson process with intensity λ. The log-price jump size ξ_t^X follows a double exponential distribution, with probability P of being negative. The means of positive and negative jump sizes are μ_+^X and $-\mu_-^X$, respectively. The volatility jump size ξ_t^σ follows an exponential distribution with mean μ^σ.

We set the time of unit to be one year and consider a time horizon of one trading day, represented by the interval $[0, T]$ with $T = 1/252$. The model parameters are given as follows:

$$\begin{aligned} &\mu = 0.05, \ \ \kappa = 5, \ \ \vartheta = 0.04, \ \ \eta = 0.05, \ \ \rho = -0.8, \\ &P = 0.65, \ \ \mu^\sigma = 0.003, \ \ \mu_+^X = 0.0075, \ \ \mu_-^X = 3\mu_+^X/2, \ \ \lambda = [1/T]. \end{aligned}$$

We use $n = 23,400$, which corresponds to the number of seconds in a typical trading day.[18] We

[17] Aït-Sahalia and Xiu (2019) show that under the simple i.i.d. setting with $\theta = 3/4$, their testing statistics follow a noncentral Chi-squared distribution with various noncentrality parameters, see their Corollaries 1, 2, 3 and 4. But those results do not yield consistency. The proofs on p.197 reveals that:

$$\begin{aligned} &\Delta_n^{-1}\sum_{i=1}^n (\Delta_i^n X)^4 = O_p(1), \quad \Delta_n^{-1}\sum_{i=1}^n a_n^4 (U_i - U_{i-1})^4 = O_p(\Delta_n^{4\theta-2}); \\ &\Delta_n^{-1}\sum_{i=1}^n a_n (\Delta_i^n X)^3 (U_i - U_{i-1}) = O_p(\Delta_n^{\theta}), \quad \Delta_n^{-1}\sum_{i=1}^n a_n^3 (\Delta_i^n X)(U_i - U_{i-1})^3 = O_p(\Delta_n^{3\theta-1}). \end{aligned}$$

Attaining consistency is possible if $\Delta_n^{-1}\sum_{i=1}^n a_n^4 (U_i - U_{i-1})^4$ supersedes the other three terms, thereby requires $\theta < 1/2$.

[18] In this numerical study, we focus on a single data frequency. While other frequencies could be explored, changing the frequency affects the noise-to-signal ratio (NSR), raising it when $\theta < 1/2$ and lowering it when

consider a model where the observation times T_i^n at stage n follow an inhomogeneous Poisson process with rate $n\alpha_t$, where $\alpha_t = 1 + \frac{1}{2}\Theta_t$, and $\Theta_t = \cos(2\pi \times 252 \times \mathrm{mod}(t, \frac{1}{252}))$, a function that models a trading indicator that peaks at the beginning and end of each trading day, with a trough midday, and repeats daily.

Under $\mathbb{H}_1$, the pricing error scale process γ satisfies

$$\gamma_t = K_\gamma \gamma'_t, \quad \mathrm{d}\gamma'_t = -\rho_\gamma(\gamma'_t - \mu_t)\mathrm{d}t + \sigma_\gamma \mathrm{d}W_t.$$

We set the parameters as follows: $\rho_\gamma = 10$, $\mu_t = 1+0.1\Theta_t$, and $\sigma_\gamma = 0.1$. After simulating the series $\{\gamma'_{T_i^n}\}_i$, we normalize it by dividing each element by the sample mean. This normalization ensures that the magnitude of the γ process is controlled by the parameter K_γ.[19]

For the stationary part of the pricing errors χ, we consider the class of ARMA processes. The error associated with the i-th trade is given by

$$\chi_i = e_i + \sum_{j=1}^{p} \varrho_j \chi_{i-j} + \sum_{j=1}^{q} \vartheta_j e_{i-j}.$$

We consider the pairs of $(p, q) \in \{(0,0), (1,0), (0,1), (2,0), (0,2), (1,1)\}$ that correspond to i.i.d., AR(1), MA(1), AR(2), MA(2), and ARMA(1,1) pricing errors. The innovations $\{e_i\}_i$ follow a Student's t-distribution with 5 degrees of freedom. The $\{\chi_i\}_i$ process is normalized to have unit variance. We analyze a total of 10 unique model specifications under $\mathbb{H}_1$. This set includes the 6 ARMA structures mentioned earlier, with some models employing multiple parameterizations to reflect various autocorrelation patterns of pricing errors found in empirical data.

5.3 Evaluating the size and power

Table 1 presents the rejection probabilities under the null hypothesis ($\mathbb{H}_0$) for our test statistics and various other tests. Overall, our statistics, the three variance ratio statistics, and the automatic portmanteau test statistics perform well, aligning closely with the nominal sizes. The three tests H_{3n}, T_n, and AC_n proposed by Aït-Sahalia and Xiu (2019) (hereafter the AX tests), together with the automatic and multiple-period variance ratio tests, exhibit more conservative behavior, particularly at lower levels.

We further examine the power of the testing statistics at significance level of 1%. Our findings are reported in Table 2. For each model, we examine small, medium, and large errors, as measured by noise-to-signal (NSR) ratios at three levels: 5%, 10%, and 15%. The NSR is defined as: $\mathrm{NSR} = \frac{2nK_\gamma^2(1-r_1)}{\mathrm{QV}(X)}$, where r_1 is the first-order autocorrelation of the χ process, and QV(X) represents the average realized variance of the efficient returns based on 1,000 simulations, serving as a proxy for the quadratic variation of the efficient price. Thus, NSR represents the true ratio of the realized variance of pricing errors to that of efficient returns. Consequently,

---

$\theta > 1/2$. Given that we are already evaluating different NSRs to assess test power, focusing on a single frequency prevents the presentation of redundant results.

[19] We set unit variance for the process χ. In the simulation setting, we set $\theta = 0$ in (2), as the value of θ does not affect the results in finite samples once n (or Δ_n) or γ are fixed.

Level \ Tests	$\mathbf{K}_0$	$\mathbf{K}_1$	$\mathbf{K}_2$	H_{3n}	T_n	AC_n
1%	0.8	0.8	1.1	0.6	0.6	0.6
5%	4.8	5.4	4.2	3.9	3.9	3.8
10%	10.7	10.5	10.0	8.0	7.9	7.8
Level \ Tests	VR_2	VR_4	VR_{10}	AutoVR	CD	AutoQ
1%	0.7	1.1	0.9	0.4	0.5	1.0
5%	4.3	4.7	3.9	1.2	3.1	4.5
10%	9.5	10.5	9.0	3.2	6.1	9.9

Table 1: Sizes of Various Tests. $\mathbf{K}_\ell$ represents the jump-truncated version of our testing statistics $H(Y;\mathbf{K}_\ell)_t^n$ as defined in (6), with $\mathbf{K}_\ell$ specified in Theorem 3.2. The statistic AutoQ refers to the automatic portmanteau test by Escanciano and Lobato (2009). $\text{VR}q, q = 2, 4, 10$ denote variance ratio tests with different horizon choices. AutoVR is the automatic variance ratio test studied by Choi (1999). CD represents the multiple periods version of the variance ratio test developed by Chow and Denning (1993). H_{3n}, T_n, AC_n are the three tests recommended by Aït-Sahalia and Xiu (2019). The sizes are calculated based on 1,000 simulations.

for each model, the three levels of K_γ^2 are determined by the formula $K_\gamma^2 = \frac{\text{NSR}\times\text{QV}(X)}{2n(1-r_1)}$, which are reported in the first rows of both the top and bottom panels of Table 2.

A quick observation is that our testing statistics win *all* 30 horseraces. Now, let's elaborate on the comparative performance of our tests versus others. When detecting *strong* errors with a 15% NSR, there is always one test among the competing nine that has rejection rates close to ours, often nearly 100%. However, the specific comparable test varies across models. For medium errors with a 10% NSR, our tests begin to demonstrate significantly greater power than others in several models, such as the AR(1) model with a coefficient of 0.9, the AR(2) model with coefficients of 0.7 and -0.2, and the ARMA(1,1) model. Smaller errors (5% NSR) are more challenging to detect, and the performance gap between the other tests and ours widens further in the presence of these small errors. Across all models, our testing statistics achieve greater rejection rates than the best of the alternative tests by a substantial margin. This consistent outperformance supports our theory that our tests are capable of detecting weak errors.

In summary, our method offers superior performance and robustness in detecting various types of pricing errors. Compared to several commonly used state-of-the-art tests for serial correlations, our test method outperforms all by a significant margin. We consider this an important contribution to the classic literature in time series econometrics.

The supplementary materials Li and Yang (2025) present additional simulation results that assess the performance of the two-sided tests under conditions featuring endogenous deviations. These results also encompass other alternative models, including drift burst and long memory processes.

6 Empirical Applications

In this section, we apply our method to demonstrate the applications using empirical data. First, we identify the frequencies at which the intraday transaction price are deemed friction-free. Next, we assess the use of mid-quote as a proxy of the efficient prices.

A widely accepted principle in econometric analysis of intraday data is that subsampling asset prices can greatly reduce the impact of pricing errors. Consequently, a prevalent approach in volatility estimation of the efficient price involves using subsamples at intervals over several minutes (Andersen et al., 2000, 2003; Corradi and Distaso, 2006; Bandi and Russell, 2008; Diebold and Strasser, 2013). This practice helps in analyzing how quickly and effectively market prices adjust to new information, thereby converging towards market efficiency.

Our empirical investigation is based on the transaction price of SPY, an S&P 500 ETF designed to track the performance of the S&P 500 Index. We consider subsampling at five distinct frequencies: every 1 second, 5 seconds, 10 seconds, 30 seconds, and 60 seconds for each trading day from 2014 to 2021. For each subsample, we perform both our tests based on various horizons and the other tests used in the simulation studies, namely, the three AX tests, the variance ratio tests with various horizons, the automatic variance ratio test, the multiple periods variance ratio test, and the automatic portmanteau test. We then calculate the rejection fractions for each year at 1% significance level, with the results reported in Table 3.

We have some quick observations. First, there is a clear trend of decreasing rejection rates as the sampling frequency decreases from 1 second to 60 seconds. This suggests that pricing errors are more likely to be detected at higher frequencies. Second, the rejection rates vary substantially across years. The data reveals that from 2014 to 2017, rejection rates are higher, especially at higher frequencies like 1-second and 5-second levels, compared to the period from 2018 to 2021. This trend suggests a significant improvement in market liquidity over time.

Comparing the test statistics, we observe results consistent with the simulation studies, particularly in higher-frequency samples: our tests often show higher rejection rates than the other tests. However, in 2017 to 2020, especially in 2020, our tests exhibit slightly lower rejection rates at the highest frequency (1-second) compared to the three AX tests.

One plausible explanation for the reduction in rejection rates of our test statistics is that, being left-sided tests, they may not capture persistent returns caused by drift bursts (Christensen et al., 2022) or gradual jumps (Andersen et al., 2023), which result in statistics falling on the right tails.[20] However, the AX tests, being two-sided, attribute all deviations from the efficient price such as drift bursts to pricing errors. While our two-sided tests provide a compelling basis for comparison with the AX tests, a right-sided test more directly identifies extreme drifts in the price process, which produce highly positive autocovariances in returns.

Table 4 reports the rejection rates of our *right-sided* tests ($\mathbf{K}_\ell^+, \ell = 0, 1, 2$) at a 1% significance level. In the years 2014 to 2016, and 2021, the rejection rates are zero. From 2017 to 2020, the right-sided tests show positive rejection rates, which are especially high in 2018 and 2020. Indeed, the years 2018 and 2020 were marked by extreme financial turmoil that was driven by, e.g., trade tensions, interest rate hikes, the cryptocurrency crash, and the COVID-19 pandemic. As a result, prolonged returns were more often observed, as suggested by the high rejection rates of our right-sided tests. Therefore, the higher rejection rates by the three AX tests are

[20] In the presence of strong drifts, e.g., drifts of order $\sqrt{\Delta_n}$ (rather than Δ_n), it is trivial to show that our testing statistics will converge to positive infinity.

likely influenced by factors beyond the usual pricing errors, or microstructure noise discussed in the literature. Table 4 also highlights another advantage of our approach. By using left or right tests, we can identify various deviations from the efficient price processes. This ability is valuable for research in financial econometrics and empirical market microstructure.

In 2019 and 2020, AX tests show higher rejection rates than variance ratio tests and the portmanteau test, which is distinct from simulation findings where AX tests usually have lower rejection rates. Although empirical rejection rates may not fully reflect test power in simulations, a structural change in the serial correlation of SPY pricing errors could explain this difference from the simulation studies. In fact, simulating an AR(1) pricing error process with a negative coefficient will make AX tests have higher power than other tests, although their power remains lower than ours.

Despite we only focus on a single index for this empirical exercise, the results indicate that intraday transaction prices can be treated as "error-free" at much higher frequencies than the conventional 5 or 15 minutes. This finding is advantageous for empirical research using intraday asset prices, as researchers benefit from a more extensive set of intraday prices that can be considered as efficient.

Next, we assess the midquotes as proxies for efficient prices. The midquote, defined as the average of the best bid and ask prices, is a fundamental concept in financial markets. It often serves as a proxy for unobserved efficient prices (Hasbrouck, 2004). It is also crucial for calculating the effective bid-ask spread (Hasbrouck, 2009; Hendershott et al., 2011). However, Hagströmer (2021) recently argued that the midquote may not accurately represent the efficient price, which is a continuous process. This poor proxy can introduce significant biases in effective spread calculations.

We present a formal test to assess how closely the midquote approximates the efficient price. Similar to the analysis in the previous subsection, we examine the midquotes of SPY from 2014 to 2017, a period with fewer anomalies. The results are displayed in Table 5. The results indicate that midquotes sampled at high frequencies, such as the 1-second level, poorly approximate efficient prices. However, they are much closer to the efficient prices than the transaction prices, as shown in Table 3. The effective spread represents the gap between transaction prices and the midquote. Our findings indicate that this effective spread accounts for a substantial portion of the pricing errors, as midquotes converge to efficient prices more rapidly than transaction prices. However, it does not capture the entirety of the pricing errors; other components still contribute to serial covariances in the midquotes.

7 Conclusion

In efficient markets, asset returns over finer time scales exhibit no serial correlation across any horizons. However, transitory pricing errors can induce return reversals, resulting in nonzero autocovariances. Depending on the economic mechanisms and statistical properties of the pricing errors, such reversals may be more pronounced at certain horizons, as confirmed by extensive empirical studies in the finance literature.

We show that return reversals, driven by a broad range of pricing errors, actually occur over a small set of horizons. Using this surprising yet practical insight, we propose a minimum covariance approach to effectively capture these reversals within the identified horizons. Thus, our method directly addresses the behavior of asset prices, and offers insights into market efficiency, stock return predictability, and market liquidity. The test is easy to apply, and nearly free from the selection of tuning parameters. Extensive numerical experiments demonstrate the efficacy of our approach in detecting a wide range of pricing errors, and our test outperforms several leading tests in the literature.

When applied to intraday data, our method assesses how closely a given series aligns with martingale efficiency prices. It can identify subsampling frequencies that yield error-free returns and evaluate the reliability of midquotes as proxies for efficient prices . Furthermore, this method is instrumental in detecting various deviations from the semimartingales recently discussed in the literature, including drift bursts, gradual jumps, price staleness, and rough volatility.

Tests \ Model	i.i.d.			AR(1), $\varrho = 0.5$			AR(1), $\varrho = 0.9$			MA(1), $\vartheta = 0.5$			MA(1), $\vartheta = 0.9$		
	K_γ^2 =2.8e-10	5.6e-10	1.1e-09	5.6e-10	1.1e-09	2.2e-09	2.8e-09	5.6e-09	1.1e-08	4.7e-10	9.3e-10	1.9e-09	5.6e-10	1.1e-09	2.2e-09
$\mathbf{K}_0$	95.8	100.0	100.0	47.7	92.8	100.0	4.9	11.9	24.8	28.4	64.7	98.9	1.1	1.2	1.1
$\mathbf{K}_1$	93.5	100.0	100.0	51.3	96.1	100.0	8.2	18.2	42.6	65.0	99.3	100.0	78.8	100.0	100.0
$\mathbf{K}_2$	91.0	100.0	100.0	47.2	94.2	100.0	11.6	24.2	57.6	59.4	98.9	100.0	74.9	100.0	100.0
H_{3n}	77.9	100.0	100.0	19.4	73.3	100.0	1.1	2.3	8.0	8.0	39.6	95.2	0.4	0.6	1.5
T_n	77.9	100.0	100.0	19.4	73.1	100.0	1.1	2.3	8.0	8.0	39.6	95.2	0.4	0.6	1.5
AC_n	77.3	100.0	100.0	18.6	70.2	100.0	1.2	2.3	7.2	6.4	34.0	90.8	0.4	0.3	0.2
VR_2	84.1	99.9	100.0	27.8	79.2	99.9	1.3	4.2	12.2	11.7	44.1	94.2	0.6	0.4	0.2
VR_4	61.9	99.3	100.0	34.6	89.7	100.0	2.4	7.3	28.9	42.1	94.0	100.0	31.3	86.5	99.9
VR_{10}	28.0	86.7	100.0	21.2	76.5	100.0	3.0	13.4	54.3	23.3	81.1	100.0	22.0	76.6	100.0
AutoVR	47.9	99.5	100.0	9.4	70.1	100.0	0.4	1.2	4.1	9.6	60.4	99.7	2.1	5.4	10.0
CD	77.0	99.8	100.0	29.0	86.7	100.0	2.3	9.3	41.4	29.5	89.4	100.0	23.3	81.5	100.0
AutoQ	84.1	99.9	100.0	28.4	81.0	99.9	1.6	5.1	15.9	34.8	93.6	100.0	66.4	99.4	100.0

Tests \ Model	AR(2), $\boldsymbol{\varrho} = (0.7, -0.2)$			AR(2), $\boldsymbol{\varrho} = (0.5, 0.2)$			MA(2), $\boldsymbol{\vartheta} = (0.5, 0.4)$			MA(2), $\boldsymbol{\vartheta} = (0.5, -0.2)$			ARMA(1,1), $\varrho = 0.6, \vartheta = 0.4$		
	K_γ^2 =6.7e-10	1.3e-09	2.7e-09	7.5e-10	1.5e-09	3.0e-09	5.6e-10	1.1e-09	2.2e-09	4.1e-10	8.1e-10	1.6e-09	1.1e-09	2.3e-09	4.6e-09
$\mathbf{K}_0$	3.9	10.5	23.4	72.6	99.8	100.0	57.7	97.9	100.0	26.7	63.8	98.8	0.0	0.0	0.0
$\mathbf{K}_1$	46.2	93.8	100.0	66.9	99.8	100.0	55.0	97.3	100.0	76.1	99.8	100.0	33.5	78.0	99.9
$\mathbf{K}_2$	49.2	94.2	100.0	62.7	99.5	100.0	55.6	97.9	100.0	72.2	99.8	100.0	41.8	88.3	100.0
H_{3n}	1.2	2.7	10.2	41.6	96.4	100.0	25.4	85.0	100.0	7.4	39.5	95.6	3.6	15.9	68.5
T_n	1.2	2.7	10.2	41.5	96.3	100.0	25.2	85.0	100.0	7.3	39.3	95.6	3.6	16.0	68.4
AC_n	1.0	2.1	7.3	41.4	96.4	100.0	25.9	85.5	100.0	5.9	30.1	91.2	3.2	13.2	58.6
VR_2	1.6	4.0	12.0	52.4	97.4	100.0	37.6	89.2	100.0	12.0	41.9	93.0	6.0	23.8	68.9
VR_4	17.0	61.7	99.5	36.2	89.6	100.0	30.1	87.4	100.0	48.7	97.6	100.0	2.6	7.2	28.9
VR_{10}	19.3	71.9	99.9	18.5	69.7	99.8	20.7	79.0	100.0	25.4	82.6	100.0	10.8	46.3	97.2
AutoVR	0.8	4.6	21.9	13.0	81.4	99.9	7.2	57.3	99.9	13.0	70.4	99.9	0.2	0.1	0.0
CD	15.7	67.2	99.9	43.4	95.2	100.0	32.6	90.3	100.0	33.0	92.8	100.0	7.8	43.4	99.2
AutoQ	13.8	58.1	99.5	52.5	97.5	100.0	39.9	93.0	100.0	59.0	99.4	100.0	18.1	70.0	99.9

Table 2: Rejection percentage of $\mathbb{H}_0$ under $\mathbb{H}_1$ based on 1,000 simulations at 1% significance level.$\mathbf{K}_\ell$ represents the jump-truncated version of our testing statistics $H(Y;\mathbf{K}_\ell)_t^n$ as defined in (6), with $\mathbf{K}_\ell$ specified in Theorem 3.2. The statistic AutoQ refers to the automatic portmanteau test by Escanciano and Lobato (2009). $\mathrm{VR}q, q = 2, 4, 10$ denote variance ratio tests with different horizon choices. AutoVR is the automatic variance ratio test studied by Choi (1999). CD represents the multiple periods version of the variance ratio test developed by Chow and Denning (1993). H_{3n}, T_n, AC_n are the three tests recommended by Aït-Sahalia and Xiu (2019). NSR represents the noise-to-signal ratio for each model, expressed as a percentage. It is computed as the average ratio, over 1000 simulations, of the realized variance of pricing errors to the realized variance of efficient prices. For each model, three NSRs are considered: 5%, 10%, and 15%.

2014	1s	5s	10s	30s	60s
$\mathbf{K}_0$	99.6	65.1	36.5	7.9	7.9
$\mathbf{K}_1$	99.6	63.9	35.3	7.9	5.6
$\mathbf{K}_2$	100.0	63.1	34.5	6.3	4.0
H_{3n}	99.6	54.8	30.6	10.7	8.3
T_n	99.6	54.8	30.6	10.3	6.7
AC_n	99.6	52.0	30.6	9.1	9.1
VR_2	99.2	52.4	28.2	6.3	5.2
VR_4	98.8	46.4	25.8	6.3	2.0
VR_{10}	96.0	38.5	17.1	2.4	0.4
AutoVR	99.6	47.6	24.6	6.0	2.4
CD	99.2	52.4	28.2	6.3	5.2
AutoQ	99.2	52.8	27.8	7.5	4.4

2018	1s	5s	10s	30s	60s
$\mathbf{K}_0$	76.5	25.9	7.6	4.4	2.4
$\mathbf{K}_1$	78.5	25.5	6.8	2.0	1.6
$\mathbf{K}_2$	77.7	23.9	5.2	0.8	0.4
H_{3n}	76.9	22.3	8.8	4.0	1.2
T_n	76.9	21.5	8.0	4.4	1.6
AC_n	77.3	21.1	7.6	4.0	1.2
VR_2	60.2	15.1	3.6	1.6	0.8
VR_4	55.0	13.5	4.0	0.0	1.6
VR_{10}	45.4	8.8	1.6	0.0	0.0
AutoVR	71.7	19.9	4.0	0.4	0.4
CD	60.2	15.1	3.6	1.6	0.8
AutoQ	62.5	14.7	5.6	1.6	0.8

2015	1s	5s	10s	30s	60s
$\mathbf{K}_0$	98.0	43.7	22.6	6.3	3.6
$\mathbf{K}_1$	98.4	43.3	20.6	6.7	4.4
$\mathbf{K}_2$	98.0	42.1	17.5	4.4	2.8
H_{3n}	97.6	34.1	13.9	7.5	5.6
T_n	97.6	34.5	13.5	7.1	4.4
AC_n	97.6	34.1	12.3	5.6	4.4
VR_2	94.0	29.8	8.3	4.0	3.2
VR_4	92.9	26.2	8.3	3.6	2.8
VR_{10}	84.5	17.9	7.1	0.8	0.0
AutoVR	96.4	29.8	9.1	3.2	1.2
CD	94.0	29.8	8.3	4.0	3.2
AutoQ	94.0	29.0	8.7	4.0	2.4

2019	1s	5s	10s	30s	60s
$\mathbf{K}_0$	92.9	23.8	11.5	1.2	2.0
$\mathbf{K}_1$	91.3	21.8	9.5	1.6	1.2
$\mathbf{K}_2$	90.1	20.2	8.3	1.6	0.8
H_{3n}	90.1	29.4	17.1	5.6	4.0
T_n	89.7	29.8	17.5	4.8	4.4
AC_n	89.7	29.0	17.5	5.6	5.6
VR_2	62.7	15.9	4.8	1.6	1.2
VR_4	59.5	14.7	7.5	0.8	0.8
VR_{10}	52.8	10.7	2.4	0.4	0.0
AutoVR	87.3	23.0	13.9	2.0	1.2
CD	62.7	15.9	4.8	1.6	1.2
AutoQ	62.7	15.1	4.0	0.4	1.2

2016	1s	5s	10s	30s	60s
$\mathbf{K}_0$	98.4	52.4	27.8	9.1	8.3
$\mathbf{K}_1$	98.0	54.0	27.4	8.7	5.6
$\mathbf{K}_2$	97.6	53.2	25.8	7.1	3.2
H_{3n}	98.4	44.4	23.4	7.5	7.1
T_n	98.4	44.0	23.4	7.1	6.0
AC_n	98.4	43.7	23.4	7.9	4.8
VR_2	91.7	38.1	17.5	4.8	5.6
VR_4	88.5	33.3	16.3	5.6	3.6
VR_{10}	81.7	27.8	12.3	2.8	0.8
AutoVR	96.4	38.1	15.9	3.6	1.2
CD	91.7	38.1	17.5	4.8	5.6
AutoQ	91.7	38.5	17.1	6.0	5.2

2020	1s	5s	10s	30s	60s
$\mathbf{K}_0$	71.1	15.4	4.0	1.6	0.0
$\mathbf{K}_1$	72.3	15.8	4.0	0.8	0.4
$\mathbf{K}_2$	71.9	17.0	3.2	0.4	0.8
H_{3n}	76.3	21.7	8.7	3.6	7.1
T_n	76.3	22.5	9.1	4.0	5.5
AC_n	76.3	21.3	7.9	4.3	4.0
VR_2	47.4	11.1	3.2	0.8	0.8
VR_4	44.3	7.9	4.0	0.0	0.4
VR_{10}	37.2	4.3	1.2	0.4	0.0
AutoVR	70.8	13.4	6.3	0.4	1.2
CD	47.4	11.1	3.2	0.8	0.8
AutoQ	49.4	10.3	3.2	1.2	2.4

2017	1s	5s	10s	30s	60s
$\mathbf{K}_0$	99.6	72.1	36.7	14.3	8.8
$\mathbf{K}_1$	99.6	71.3	37.5	14.3	7.2
$\mathbf{K}_2$	99.6	71.3	34.7	10.8	6.0
H_{3n}	100.0	59.0	29.9	12.0	8.8
T_n	100.0	59.0	29.5	10.8	8.8
AC_n	100.0	59.4	27.1	8.8	7.2
VR_2	94.0	53.0	23.1	7.6	9.2
VR_4	93.6	51.4	22.3	10.0	3.6
VR_{10}	91.6	45.8	18.7	2.8	1.2
AutoVR	99.2	53.8	23.5	5.2	3.6
CD	94.0	53.0	23.1	7.6	9.2
AutoQ	94.4	51.4	21.5	7.2	5.2

2021	1s	5s	10s	30s	60s
$\mathbf{K}_0$	88.5	12.3	0.4	3.2	0.8
$\mathbf{K}_1$	86.9	9.5	2.4	1.6	0.4
$\mathbf{K}_2$	85.7	10.3	3.6	0.8	0.4
H_{3n}	84.9	15.5	12.7	5.2	4.8
T_n	84.9	15.5	13.5	6.0	3.2
AC_n	84.9	15.5	11.5	4.8	2.4
VR_2	68.7	6.3	6.3	3.6	0.8
VR_4	61.5	6.7	7.1	1.6	0.8
VR_{10}	49.2	5.2	4.0	2.0	0.0
AutoVR	79.0	11.9	8.7	3.2	0.8
CD	68.7	6.3	6.3	3.6	0.8
AutoQ	69.0	6.3	8.3	6.0	1.6

Table 3: Annual rejection rates of $\mathbb{H}_0$ for SPY from 2014 to 2021, across different subsampling frequencies. $\mathbf{K}_0, \mathbf{K}_1, \mathbf{K}_2$ denote our tests with various horizons. AutoQ is the automatic portmanteau test proposed by Escanciano and Lobato (2009). The VR$q, q = 2, 4, 10$ tests are variance ratio tests with chosen horizons. AutoVR refers to the automatic variance ratio test studied by Choi (1999). CD represents the multiple periods variance ratio test developed by Chow and Denning (1993). The tests H_{3n}, T_n, and AC_n are recommended by Aït-Sahalia and Xiu (2019). All evaluations are at the 1% significance level.

Tests \ Year	2014	2015	2016	2017	2018	2019	2020	2021
$\mathbf{K}_0^+$	0.0	0.0	0.0	0.4	4.4	0.8	6.7	0.0
$\mathbf{K}_1^+$	0.0	0.0	0.0	0.4	3.6	0.4	3.6	0.0
$\mathbf{K}_2^+$	0.0	0.0	0.0	0.4	2.0	0.8	3.6	0.0

Table 4: Rejection rates by right-sided tests ($\mathbf{K}_0^+, \mathbf{K}_1^+, \mathbf{K}_2^+$) at the 1% significance level, using samples at 1 second frequency. Right-sided tests examine return persistence, potentially driven by drift bursts or gradual jumps.

Year	**Tests**	1s	5s	10s	30s	60s
2014	$\mathbf{K}_0$	83.3	25.4	10.7	3.6	6.3
	$\mathbf{K}_1$	85.3	25.0	12.7	5.2	4.8
	$\mathbf{K}_2$	84.1	26.2	11.9	3.2	3.6
2015	$\mathbf{K}_0$	67.5	15.9	8.7	3.2	2.8
	$\mathbf{K}_1$	65.9	17.1	6.3	2.4	4.0
	$\mathbf{K}_2$	65.9	13.5	4.8	2.0	2.4
2016	$\mathbf{K}_0$	46.0	18.3	10.3	5.2	6.3
	$\mathbf{K}_1$	50.8	23.0	11.1	4.8	4.0
	$\mathbf{K}_2$	51.2	21.8	10.7	2.8	2.8
2017	$\mathbf{K}_0$	63.3	16.7	12.0	8.0	7.6
	$\mathbf{K}_1$	59.8	21.5	15.9	8.8	5.2
	$\mathbf{K}_2$	57.8	21.1	17.1	7.6	5.2

Table 5: Annual rejection rates of $\mathbb{H}_0$ for SPY midquotes (2014-2017) using tests $\mathbf{K}_\ell$ ($\ell = 0, 1, 2$) at 1% significance.

Appendices

Appendix A Technical Conditions

In this section, we present and discuss the regularity conditions for the extensions outlined in Section 4. First, we consider the Itô semimartingale, a comprehensive class of stochastic processes that encompass nearly all continuous-time models used in finance and economics.

A.1 Itô semimartingale

For any Itô semimartingale Z considered in this paper, we assume it is defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, and it can be represented as (Jacod and Protter, 2011):

$$Z_t = Z_0 + \int_0^t b_s^Z \mathrm{d}s + \int_0^t \sigma_s^Z \mathrm{d}W_s^Z + \left(\delta^Z \mathbf{1}_{\{|\delta^Z|\leq 1\}}\right) \star (\mu - \nu)_t + \left(\delta^Z \mathbf{1}_{\{|\delta^Z|>1\}}\right) \star \mu_t, \tag{9}$$

where W^Z is a Brownian motion, the Poisson random measure μ is defined on $\mathbb{R}_+ \times \mathbb{R}$, and its compensator is $\nu(\mathrm{d}s, \mathrm{d}z) = \mathrm{d}s \otimes \lambda(\mathrm{d}z)$ for some given σ-finite measure λ. The processes b^Z, σ^Z are optional and locally bounded. The function δ^Z on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$ is predictable. We further impose a mild regularity condition below for all Itô semimartingale considered in this paper, which is not significantly stronger than the property of being an Itô semimartingale. In fact, in virtually all models using Itô semimartingales, this condition is indeed satisfied, see a discussion on p.170 of Aït-Sahalia and Jacod (2014).

Assumption (H). *Let Z be a semimartingale represented by (9). There exists a sequence of stopping times $\{\tau_n\}_n$, a sequence of real numbers $\{w_n\}_n$, and for each n a deterministic nonnegative function Γ_n^Z on E so that the following hold for all (ω, t, z) with $t \leq \tau_n(\omega)$:*

$$\left|b_t^Z(\omega)\right| \leq w_n, \quad \left|\sigma_t^Z(\omega)\right| < w_n, \quad \left|\delta^Z(\omega, t, z)\right| \leq \Gamma_n^Z(z), \quad \int_E \left(\left(\Gamma_n^Z(z)\right)^2 \wedge 1\right) \lambda(\mathrm{d}z) < \infty.$$

A.2 Random observation scheme

Denote $\mathcal{F}_i^n := \mathcal{F}_{T_i^n}$. The observation density process α satisfies

Assumption (O-ρ). *Let α be an Itô semimartingales defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ satisfying $\alpha_t > 0$ and $\alpha_{t-} > 0$ for all $t > 0$. We further assume*

(i) $\Delta_n n_t \xrightarrow{\mathbb{P}} A_t$.

(ii) For any $\kappa \geq 2$, there exists a sequence of stopping times $\{\tau(\kappa)_m\}_m$ that increases to ∞, a sequence of real numbers $\{w(\kappa)_m\}_m$, and some $\rho > 0$, satisfying:

$$\left|\mathbb{E}\left(\alpha_{i-1}^n \Delta(n,i)\middle|\, \mathcal{F}_{i-1}^n\right) - \Delta_n\right| \leq w(\kappa)_m \Delta_n^{1+\rho}, \quad \mathbb{E}\left(\left|\alpha_{i-1}^n \Delta(n,i)\right|^\kappa \middle|\, \mathcal{F}_{i-1}^n\right) \leq w(\kappa)_m \Delta_n^\kappa,$$

when $T_{i-1}^n \leq \tau(\kappa)_m$.

A.3 Efficient price with jumps

We can introduce jumps into the efficient price process. We assume

Assumption (H-X)**.** *The efficient price process is represented as*

$$X_t = X_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + J_t, \tag{10}$$

where J_t is a jump process driven by a homogeneous Poisson process summable over each finite time interval. We further assume that the drift and volatility processes, b and σ, are also Itô semimartingales.[21]

Now we discuss the truncation procedure to eliminate jumps. Let $u_n = u\Delta_n^{\varpi}$, where $\varpi \in (0, 1/2)$, be a truncation level. The choice of u is well-established in the literature (Li et al., 2013). It may be a constant or vary over time, depending on the underlying processes, such as the average, or spot volatility of the (noisy) observed returns $\{\Delta_i^n Y\}_i$. In this paper, we adopt the following truncation levels: $u_j^n := K\sqrt{\hat{c}_j^n}\Delta_n^{\varpi}$, where $\hat{c}_j^n := \frac{1}{d_n\Delta_n}\sum_{i=j-d_n+1}^{j}(\Delta_i^n Y)^2$ with $d_n \to \infty$, or the average realized variance. Under the null hypothesis, this threshold will asymptotically eliminate jumps. Under the alternative hypothesis, it will retain most of the pricing errors that we aim to detect. We will use $\overline{F}(Y;k)_t^n, \overline{H}(Y;k)_t^n$ to denote the truncated version of our statistics $F(Y;k)_t^n$ and $H(Y;k)_t^n$.

A.4 Pricing errors with stochastic scales

We can allow for a stochastic scale process for the pricing errors:

Assumption (N-θ-v)**.** *The pricing error process satisfies the following factorization:*

$$\varepsilon_i^n := \Delta_n^{\theta} \cdot \gamma_i^n \cdot \chi_i, \quad \theta \in [0, 3/4),$$

where γ is a nonnegative Itô semimartingale; the process χ is stationary ρ-mixing sequence, and independent of the σ-filed $\mathcal{F}_\infty := \bigvee_{t>0}\mathcal{F}_t$ with mean 0 and variance 1, with $\mathbb{E}\left(|\chi_i|^{4+\epsilon}\right) < \infty$ for arbitrary $\epsilon > 0$. The ρ-mixing coefficient of χ satisfies $\rho_k \leq Kk^{-v}$ for some constants $K > 0, v > 0$.

Moreover, we define the null and alternative hypotheses as $\mathbb{H}_0 : \omega \in \Omega_t^0$ and $\mathbb{H}_1 : \omega \in \Omega_t^1$, where

$$\Omega_t^0 := \left\{\omega : \int_0^t \gamma_s(\omega)^2 \mathrm{d}A_s(\omega) = 0\right\} \text{ and } \Omega_t^1 := \left\{\omega : \int_0^t \gamma_s(\omega)^2 \mathrm{d}A_s(\omega) > 0\right\}. \tag{11}$$

Appendix B Additional Technical Results

In this section, we present the asymptotic results under the general setups. Recall that the process A is the cumulative observation density function $A_t := \int_0^t \alpha_s \mathrm{d}s$.

[21] Note that the assumption that b and σ are also Itô semimartingales is slightly stronger than local boundedness, is necessary to handle the random observation scheme discussed in the previous section.

Theorem B.1. *Suppose that Assumptions (H-X) and (O-ρ) hold. For a given set of positive integers $\mathbf{K} = \{k_1, \dots, k_m\}$, we have the following joint stable convergence on Ω_t^0:*

$$\left(\frac{\overline{F}(Y;k_1)_t^n}{\sqrt{\Delta_n}}, \dots, \frac{\overline{F}(Y;k_m)_t^n}{\sqrt{\Delta_n}}\right) \xrightarrow{\mathcal{L}_s} \left(\overline{\mathcal{Z}}_t^1, \dots, \overline{\mathcal{Z}}_t^m\right), \tag{12}$$

where conditional on $\mathcal{F}$, $(\overline{\mathcal{Z}}_t^j)_{j=1}^m$ are Gaussian variables with conditional covariances given by

$$\tilde{\mathbb{E}}\big(\overline{\mathcal{Z}}_t^j \overline{\mathcal{Z}}_t^{j'} \mid \mathcal{F}\big) = \Phi_{k_j,k_{j'}} \int_0^t \frac{\sigma_s^4}{\alpha_s^2} \mathrm{d}A_s.$$

Consequently, we also have the following stable convergence in law

$$(\overline{H}(Y;k_j)_t^n, \overline{H}(Y;k_{j'})_t^n) \xrightarrow{\mathcal{L}_s} (\overline{\Theta}_j, \overline{\Theta}_{j'}), \tag{13}$$

where $(\overline{\Theta}_j, \overline{\Theta}_{j'})$ are a pair of standard normal variables with correlation $\rho_{k_j,k_{j'}}$, defined on an extension of $(\Omega, \mathcal{F}, \mathbb{P})$ and is independent of $\mathcal{F}$.

Theorem B.2. *Let $\mathbf{K}_\ell := \{1, 2, \dots, 2^\ell\}$. Given $\alpha \in (0,1)$, let c_α be the critical value that satisfies $\Sigma_{\mathbf{K}_\ell}(x_1 \geq c_\alpha, \dots, x_m \geq c_\alpha) = 1 - \alpha$. The statistic $\overline{H}(Y;\mathbf{K}_\ell)_t^n$ has asymptotic size α in restriction to the set Ω_t^0:*

$$\mathbb{P}\big(\overline{H}(Y;\mathbf{K}_\ell)_t^n < c_\alpha \,|\, \Omega_t^0\big) \to \alpha.$$

When $r_1 \leq 1 - 2^{-\ell}$, $\overline{H}(Y;\mathbf{K}_\ell)_t^n$ is consistent on the set Ω_t^1:

$$\mathbb{P}\big(\overline{H}(Y;\mathbf{K}_\ell)_t^n < c_\alpha \,|\, \Omega_t^1\big) \to 1.$$

Appendix C Proof of Main Theorems

Proof of Theorem B.1. The convergence in (12) follows directly from Theorem S.11, and Lemma S.7 in Li and Yang (2025). The convergence in (13) follows from (12) and Lemma S.9 in Li and Yang (2025). □

Proof of Theomre 3.2. Suppose this is not true: we have $g(k,r) \leq 0$, or

$$r_k \geq \frac{1}{2}(1 + r_{2k}), \tag{14}$$

for all $k \in \{1, 2, \dots, 2^\ell\}$. We can use (14) to show recursively that

$$r_1 \geq 1 - \frac{1}{2^{\ell+1}} + \frac{r_{2^{\ell+1}}}{2^{\ell+1}} > 1 - \frac{1}{2^\ell},$$

which contradicts the fact that $r_1 \leq 1 - \frac{1}{2^\ell}$. □

Proof of Theorem B.2. The first part of the Theorem follows directly from Theorem B.1. We now show the consistency. In view of Theomre 3.2, Lemma S.8 and Lemma S.5 in Li and Yang (2025), it suffices to show that

$$\Delta_n^{2-4\theta} Q(Y)_t^n \xrightarrow{\mathbb{P}} 0, \tag{15}$$

for $\theta \in [0, 3/4)$. First of all, we observe that for $\theta \geq 1/2$, $Q(Y)_t^n = O_p(\Delta_n)$. Thus, (15) holds since $\theta < \frac{3}{4}$. Now consider the scenario $\theta \in [0, 1/2)$, thus the dominating part in $\Delta_i^n Y$ is the pricing error, which is of order Δ_n^{θ}. Then, it is trivial to get $\Delta_n^{2-4\theta} Q(Y)_t^n = O_p(\Delta_n) \xrightarrow{\mathbb{P}} 0$, which also implies (15). □

Proofs of Theorem 3.1, Corollary 3.1, Theorem 3.4, Corollary 3.2. These results follow from Theorem B.1. □

Proofs of Theorem 3.3 and Theorem 3.5. The two theorems are special cases of Theorem B.2. □

References

Aït-Sahalia, Y. and Jacod, J. (2014). *High-Frequency Financial Econometrics*. Princeton University Press.

Aït-Sahalia, Y. and Xiu, D. (2019). A hausman test for the presence of market microstructure noise in high frequency data. *Journal of Econometrics*, 211(1):176–205.

Andersen, T. G., Archakov, I., Cebiroglu, G., and Hautsch, N. (2022). Local mispricing and microstructural noise: A parametric perspective. *Journal of Econometrics*, 230(2):510–534.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2000). Great realizations. *Risk*, 13:105–108.

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

Andersen, T. G., Li, Y., Todorov, V., and Zhou, B. (2023). Volatility measurement with pockets of extreme return persistence. *Journal of Econometrics*, 237(2):105048.

Bandi, F. M. and Russell, J. R. (2008). Microstructure noise, realized variance, and optimal sampling. *Review of Economic Studies*, 75(2):339–369.

Bogousslavsky, V. (2016). Infrequent rebalancing, return autocorrelation, and seasonality. *Journal of Finance*, 71(6):2967–3006.

Bollerslev, T., Li, J., and Ren, Y. (2024). Optimal inference for spot regressions. *American Economic Review*, 114(3):678–708.

Box, G. E. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, 65(332):1509–1526.

Choi, I. (1999). Testing the random walk hypothesis for real exchange rates. *Journal of Applied Econometrics*, 14(3):293–308.

Chong, C. H. and Todorov, V. (2025). A nonparametric test for rough volatility. *Journal of the American Statistical Association*.

Chordia, T., Roll, R., and Subrahmanyam, A. (2005). Evidence on the speed of convergence to market efficiency. *Journal of Financial Economics*, 76(2):271–292.

Chow, K. V. and Denning, K. C. (1993). A simple multiple variance ratio test. *Journal of Econometrics*, 58(3):385–401.

Christensen, K., Oomen, R., and Renò, R. (2022). The drift burst hypothesis. *Journal of Econometrics*, 227(2):461–497.

Cipriani, M. and Guarino, A. (2014). Estimating a structural model of herd behavior in financial markets. *American Economic Review*, 104(1):224–251.

Cochrane, J. H. (1988). How big is the random walk in gnp? *Journal of Political Economy*, 96(5):893–920.

Cochrane, J. H. (1994). Permanent and transitory components of gnp and stock prices. *Quarterly Journal of Economics*, 109(1):241–265.

Corradi, V. and Distaso, W. (2006). Semi-parametric comparison of stochastic volatility models using realized measures. *Review of Economic Studies*, 73(3):635–667.

Corwin, S. A. and Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *Journal of Finance*, 67(2):719–760.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, 98(4):703–738.

Deo, R. S. and Richardson, M. (2003). On the asymptotic power of the variance ratio test. *Econometric Theory*, 19(2):231–239.

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.

Easley, D. and O'hara, M. (1992). Time and the process of security price adjustment. *Journal of Finance*, 47(2):577–605.

Escanciano, J. C. and Lobato, I. N. (2009). An automatic portmanteau test for serial correlation. *Journal of Econometrics*, 151(2):140–149.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417.

Fama, E. F. (1991). Efficient capital markets: Ii. *Journal of Finance*, 46(5):1575–1617.

Fama, E. F. (2014). Two pillars of asset pricing. *American Economic Review*, 104(6):1467–1485.

Fama, E. F. and French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy*, 96(2):246–273.

Hagströmer, B. (2021). Bias in the effective bid-ask spread. *Journal of Financial Economics*, 142(1):314–337.

Hasbrouck, J. (1993). Assessing the quality of a security market: A new approach to transaction-cost measurement. *Review of Financial Studies*, 6(1):191–212.

Hasbrouck, J. (2002). Stalking the "efficient price" in market microstructure specifications: an overview. *Journal of Financial Markets*, 5(3):329–339.

Hasbrouck, J. (2004). Liquidity in the futures pits: Inferring market dynamics from incomplete data. *Journal of Financial and Quantitative Analysis*, 39(2):305–326.

Hasbrouck, J. (2007). *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading*. Oxford University Press.

Hasbrouck, J. (2009). Trading costs and returns for us equities: Estimating effective costs from daily data. *Journal of Finance*, 64(3):1445–1477.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6):1251–1271.

Hendershott, T., Jones, C. M., and Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *Journal of Finance*, 66(1):1–33.

Hendershott, T. and Menkveld, A. J. (2014). Price pressures. *Journal of Financial Economics*, 114(3):405–423.

Huang, W., Liu, Q., Rhee, S. G., and Zhang, L. (2010). Return reversals, idiosyncratic risk, and expected returns. *Review of Financial Studies*, 23(1):147–168.

Jacod, J., Li, Y., and Zheng, X. (2017). Statistical properties of microstructure noise. *Econometrica*, 85(4):1133–1174.

Jacod, J., Li, Y., and Zheng, X. (2019). Estimating the integrated volatility with tick observations. *Journal of Econometrics*, 208(1):80–100.

Jacod, J. and Protter, P. E. (2011). *Discretization of Processes*, volume 67. Springer Science & Business Media.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance*, 45(3):881–898.

Kolokolov, A., Renò, R., and Zoi, P. (2025). Bumvu estimators. *Journal of Econometrics*, page 105942.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, pages 1315–1335.

Laurent, S., Renò, R., and Shi, S. (2024). Realized drift. *Journal of Econometrics*, page 105813.

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *Quarterly Journal of Economics*, 105(1):1–28.

Li, J., Todorov, V., and Tauchen, G. (2013). Volatility occupation times. *Annals of Statistics*, 41(1):1865–1891.

Li, Z. M. and Linton, O. (2022). A remedi for microstructure noise. *Econometrica*, 90(1):367–389.

Li, Z. M. and Yang, X. (2025). Supplementary materials for "multi-horizon test for market frictions". *working paper*.

Ljung, G. M. and Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

Lo, A. W. and MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1(1):41–66.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping brownian motion. *Giornale dell'Istituto Italiano degli Attuari*, 64:19–47.

Poterba, J. M. and Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics*, 22(1):27–59.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance*, 39(4):1127–1139.

Shi, S. and Phillips, P. C. (2024). Uncovering mild drift in asset prices with intraday high-frequency data. *Available at SSRN 4786774*.

Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *Journal of Economic Perspectives*, 17(1):83–104.

Summers, L. H. (1986). Does the stock market rationally reflect fundamental values? *Journal of Finance*, 41(3):591–601.