Unit Root, Mean Reversion and Nonstationarity in Financial Time Series*

Jihyun Kim†
Toulouse School of Economics
University of Toulouse Capitole

Joon Y. Park‡
Department of Economics
Indiana University
and Sungkyunkwan University

April 8, 2018

Abstract

This paper reveals how interrelated are the notions of unit root, mean reversion and nonstationarity in financial time series, which are typically regarded as samples from diffusion-type processes. Mean reversion is shown to be non-synonymous with stationarity for general diffusion models: Nonstationary diffusions may also be mean-reverting if they have drift terms dominating diffusion terms. Furthermore, we find that the unit root test is a consistent test of no mean reversion for general diffusion models, and the test is also consistent as a test of nonstationarity if it is applied to the appropriately transformed samples. To test for no mean reversion and nonstationarity, we introduce a subsample bootstrap test and develop its asymptotics. Our empirical study illustrates that the examples of nonstationary mean-reverting financial time series are not rare, and pairs of nonstationary non-mean-reverting financial time series often yield nonstationary mean-reverting errors and define more general cointegrating relationships.

JEL Classification: C12, C22, C58
Key words and phrases: unit root, mean reversion, nonstationarity, financial time series, diffusion model.

---

*We are grateful to Yacine Aït-Sahalia, Yoosoon Chang, Christian Gouriéroux, Oliver Linton, Nour Meddahi, Morten Ørregaard Nielsen, Olivier Scaillet and seminar participants at Indiana University, Toulouse School of Economics, University of Cambridge, University of Maastricht, Singapore Management University, Rutgers University, University of Illinois at Urbana-Champaign, University of York, University of Liverpool, CREST, Aarhus University, University of Glasgow, 2013 Midwest Econometrics Group Meetings, 2013 North American Econometric Society Meeting, 2013 SETA, 2014 Toulouse Financial Econometrics Conference, 7th French Econometrics Conference, 9th International Conference on Computational and Financial Econometrics, and 2017 Southampton Finance & Econometrics Workshop for useful comments and discussions. Various versions of the paper have been circulated under the titles "Mean Reversion and Stationarity: A New Perspective from Asymptotics of Diffusion Models" and "Mean Reversion and Unit Root Properties of Diffusion Models".

†31000 Toulouse, France. E-mail: jihyun.kim@tse-fr.eu.
‡Corresponding author. Bloomington, IN 47405-7104, USA. E-mail: joon@indiana.edu.

1 Introduction

The unit root test was originally developed to test for the unity of the largest root in an autoregressive time series model. However, the existence of a unit root also entails other important consequences on more general stochastic characteristics of a time series. If a unit root is present, shocks have permanent effects and they are accumulated to build up a long run stochastic trend, which makes the autoregressive process nonstationary and non-mean-reverting. On the other hand, an autoregressive process becomes stationary and mean-reverting if it has the largest root less than unity in modulus, since shocks have only transitory effects and their effects eventually vanish in the long run. Accordingly, the unit root test may also be regarded generally as a test for nonstationarity and no mean reversion of a time series. In fact, the unit root test has commonly been used to test for nonstationarity and no mean reversion in a much broader class of time series models than autoregressive models.

The unit root test has been routinely applied also to financial time series that are thought to be collected discretely at relatively high frequencies from diffusion type continuous time processes. See, e.g., Chan et al. (1992), Aït-Sahalia (1996a,b) and Jones (2003). However, the meaning of the presence or absence of a unit root in discrete samples from a general diffusion model has never been clearly understood. The previous studies only consider an unrealistically simple case, where the underlying diffusion models are driven by Brownian motion and stationary Ornstein-Uhlenbeck processes respectively for the null and alternative hypotheses, as in Shiller and Perron (1985), Perron (1989, 1991) and Chambers (2004, 2008), among others. In particular, they do not allow for state-dependent volatility, which is widely regarded as one of the most conspicuous characteristics of financial time series. It is completely unknown whether the unit root test has any discriminatory power in distinguishing general stationary and nonstationary, or mean-reverting and non-mean-reverting, diffusion models. The effects of time span and sampling frequency on the size and power of the unit root test are also largely undiscovered for general diffusion models.

In this paper, we investigate the long run behaviors of general diffusion models including their unit root, mean reversion and nonstationarity properties. Our investigation is comprehensive and thorough. We consider the entire class of recurrent diffusions covering all positive and null recurrent diffusions, which includes in particular stationary diffusions having no proper moments as well as general nonstationary diffusions.[1] For the sample from

[1] To obtain more explicit asymptotic results, we only consider pure diffusions without jumps. We believe that the presence of jumps is not important in our study. The unit root property is determined entirely by the asymptotic properties of the underlying model, which can be tested effectively by samples collected at the daily or lower frequencies as long as the time span is long enough. It is well known that jumps are not

such a general class of diffusions, we show that the Dickey-Fuller unit root test has a well defined limit distribution if and only if the underlying diffusion does not have mean reversion, and it diverges to minus infinity in probability if and only if the underlying diffusion has mean reversion. The unit root test therefore has perfect discriminatory power, if used to discriminate non-mean-reverting diffusions against mean-reverting diffusions. On the other hand, the test cannot be used to test for nonstationarity of the underlying diffusion. Although all stationary diffusions, including those without finite mean, are mean-reverting, not all nonstationary diffusions are non-mean-reverting. Nonstationary diffusions may also be mean-reverting if they have drift terms dominating diffusion terms.

The existence of mean-reverting nonstationary diffusions has some important and far-reaching implications. First, it implies that nonstationary financial time series may not be necessarily non-mean-reverting. This opens up a possibility, for instance, that stock prices, which are widely believed to be nonstationary, are still mean-reverting. See, e.g., Fama and French (1988), Poterba and Summers (1988) and Kim et al. (1991). Second, in any cointegrating relationship, we may allow for the disequilibrium error process to be nonstationary as long as it is mean-reverting. As a result, we may extend the notion of cointegration in financial time series studied in, e.g., Baillie and Bollerslev (1989, 1994) and Diebold et al. (1994). The extended notion of cointegration also has some important consequences in actual financial investments. In fact, one popular short-term speculation strategy known as pairs trading utilizes co-movements in the prices of two or more stocks, where we may define co-movements more generally using the extended notion of cointegration.[2] For more details on pairs trading, the reader is referred to Bossaerts (1988), Bossaerts and Green (1989) and Gatev et al. (2006).

A diffusion is non-mean-reverting if and only if it is nonstationary, when there is no interaction between its drift and diffusion terms. On the other hand, its stationarity and nonstationarity are both preserved under transformations such as the scale transformation and Lamperti transformation, which annihilate the interaction of drift and diffusion terms. Consequently, if applied to the transformed data using one of these transformations, we may use the unit root test to test for nonstationarity of the underlying diffusion. Clearly, the presence of a unit root in the transformed data implies no mean reversion, and therefore, nonstationarity in the transformed underlying diffusion, since it has no drift-diffusion

frequently observed at the daily or lower frequencies in most financial time series. Moreover, as shown by, e.g., Jeong and Park (2016), the jump diffusions that are reducible to martingales by their scale functions yield essentially the same asymptotics as pure diffusions, and therefore, the main asymptotic results in this paper are expected to be also applicable for a large class of, though not all, jump diffusion models.

[2]For a different approach in modeling longrun co-movements of multiple time series, the reader is referred to Müller and Watson (2017).

interactions. This, however, holds if and only if there is nonstationarity in the original underlying diffusion. Therefore, the unit root test can also be used to test for nonstationarity consistently, if we know the scale or Lamperti transformation. Of course, the required transformation has to be estimated in practical implementation of the test. We show in the paper that the unit root test based on the Lamperti transformation, if it is estimated appropriately, is generally consistent at high frequency as a test of nonstationarity.

As a test of no mean reversion or nonstationarity in our general setup, the limit distribution of the Dickey-Fuller test becomes heavily model-dependent and relies on the underlying diffusion model in a complicated manner. Therefore, the usual Dickey-Fuller critical values cannot be used. The limit distribution of the test is generally represented as a functional of a skew Bessel process, and reduces to the Dickey-Fuller distribution only if the underlying diffusion becomes a Brownian motion in the limit. In the paper, we develop a subsample bootstrap test based on the nonparametric estimation of the underlying diffusion model, and show that it is valid for, and consistent against, general diffusion models. Our simulation shows that our test has reasonably good size and power in finite samples. As an illustration, we use the test to examine the presence of mean reversion and nonstationarity in some major financial time series. Most nonstationary time series are non-mean-reverting. However, the examples of nonstationary mean-reverting time series are not rare either. The exchange rates between Australian Dollars and New Zealand Dollars yield some evidence of mean-reverting nonstationarity. Moreover, in the spreads between the prices of Gold and Silver, and those between the equity indexes of developed markets in the World and Europe, we see strong evidence of mean-reverting nonstationarity. This shows the practical usefulness of our new notion of cointegration.

The rest of the paper is organized as follows. Section 2 presents the background and preliminaries that are necessary to understand subsequent development of asymptotic theory developed in the paper. The diffusion model and various notions to investigate its long run behaviors, and some important regularity and integrability conditions are introduced. Section 3 considers the Dickey-Fuller unit root test and develops its asymptotics. The asymptotics are two-dimensional, relying on the sampling interval δ as well as the sample span T. Section 4 reveals how the notions of unit root, mean reversion and nonstationarity are interrelated. It introduces the precise notion of mean reversion, and shows that the unit root test is indeed a consistent test of no mean reversion, which reduces to a consistent test of nonstationarity only when it is applied to the appropriately transformed samples. Section 5 introduces a subsample bootstrap test and develops its asymptotic theory, and provides a set of simulation results and illustrative empirical applications. Section 6 concludes the paper, and all mathematical proofs are in Appendix.

Finally, a word on notation. We write "$P_T \sim Q_T$" to denote $P_T/Q_T \to 1$. Similarly, "$P_T \sim_p Q_T$" means $P_T/Q_T \to_p 1$, and $P_T \sim_d Q_T$ implies that P_T and Q_T have the same asymptotic distributions. Moreover, we let "$P_T \prec Q_T$", "$P_T \prec_p Q_T$" and "$P_T \precsim_p Q_T$" signify $P_T = o(Q_T)$, $P_T = o_p(Q_T)$ and $P_T = O_p(Q_T)$, respectively. These notations, as well as other standard notations used in asymptotics, will be used frequently throughout the paper without further references.

2 Background and Preliminaries

In this section, we present the diffusion model with some of its basic properties determining long run behaviors, and introduce some important regularity and integrability conditions.

2.1 Diffusion Model

We consider the diffusion process X given by the time-homogeneous stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \tag{1}$$

where μ and σ are, respectively, the drift and diffusion functions, and W is the standard Brownian motion. We denote by $\mathcal{D} = (\underline{x}, \overline{x})$ the domain of the diffusion process X, where we set $\underline{x} = -\infty$ or 0 with $\overline{x} = \infty$. This causes no loss in generality, since we may simply consider $X - \underline{x}$ or $-X$ to allow for a more general case. In what follows, we denote by $x_B = \underline{x}$ or $\overline{x}$ the boundary of $\mathcal{D}$. Throughout the paper, we assume

Assumption 2.1. *We assume that (a) $\sigma^2(x) > 0$ for all $x \in \mathcal{D}$, and (b) $\mu(x)/\sigma^2(x)$ and $1/\sigma^2(x)$ are locally integrable at every $x \in \mathcal{D}$.*

Assumption 2.1 provides a simple sufficient set of conditions to ensure that a weak solution to the stochastic differential equation (1) exists uniquely up to an explosion time. See, e.g., Theorem 5.5.15 in Karatzas and Shreve (1991).

The scale function of the diffusion process X in (1) is defined as

$$s(x) = \int_w^x \exp\left(-\int_w^y \frac{2\mu(z)}{\sigma^2(z)}dz\right)dy, \tag{2}$$

where the lower limits of the integrals can be arbitrarily chosen to be any point $w \in \mathcal{D}$. Defined as such, the scale function s is uniquely identified up to any increasing affine transformation, i.e., if s is a scale function, then so is $as + b$ for any constants $a > 0$ and

$-\infty < b < \infty$. We also define the speed density

$$m(x) = \frac{1}{(\sigma^2 s')(x)} \tag{3}$$

on $\mathcal{D}$, where s' is the derivative of s, often called the scale density, which is assumed to exist. The speed measure is defined to be the measure on $\mathcal{D}$ given by the speed density with respect to the Lebesgue measure. Note, under Assumption 2.1, that both the scale function and speed density are well defined, and that the scale function is strictly increasing, on $\mathcal{D}$.

Our asymptotic theory depends crucially on the recurrence property of the diffusion process X. To define the recurrence property, we let ρ_y be the hitting time of a point y in $\mathcal{D}$ that is given by $\rho_y = \inf\{t \geq 0 | X_t = y\}$. We say that the diffusion X is recurrent if $\mathbb{P}\{\rho_y < \infty | X_0 = x\} = 1$ for all $x, y \in \mathcal{D}$. The recurrent diffusion X is said to be positive recurrent if $\mathbb{E}[\rho_y < \infty | X_0 = x] < \infty$ for all $x, y \in \mathcal{D}$, and null recurrent if $\mathbb{E}[\rho_y < \infty | X_0 = x] = \infty$ for all $x, y \in \mathcal{D}$. Under Assumption 2.1, the diffusion X is recurrent if and only if the scale function s in (2) is unbounded at both boundaries $\underline{x}$ and $\overline{x}$, i.e.,

$$s(\underline{x}) = -\infty \quad \text{and} \quad s(\overline{x}) = \infty.$$

Throughout the paper, we assume that this condition holds and X is recurrent. The recurrent diffusion X becomes positive recurrent or null recurrent, depending upon

$$m(\mathcal{D}) < \infty \quad \text{or} \quad m(\mathcal{D}) = \infty,$$

where m is the speed measure defined in (3).[3] A diffusion which is not recurrent is said to be transient.

Positive recurrent diffusions are stationary. More precisely, they have time invariant distributions, and if they are started from the time invariant distributions they become stationary. The time invariant density of the positive recurrent diffusion X is given by

$$\pi(x) = \frac{m(x)}{m(\mathcal{D})}.$$

Null recurrent and transient diffusions are nonstationary. They do not have time invariant distributions, and their marginal distributions change over time. Out of these two different types of nonstationary processes, we mainly consider null recurrent diffusions in the paper.

[3] Throughout the paper, we follow the notational convention in the Markov process literature and use the same notation for both a measure and its density with respect to the Lebesgue measure. As an example, for a given measure or a density m and a function f on $\mathcal{D} \subset \mathbb{R}$, we write $m(\mathcal{D})$ and $m(f)$ interchangeably with $\int_{\mathcal{D}} m(x)dx$ and $\int_{\mathcal{D}} m(x)f(x)dx$ respectively.

Brownian motion is the prime example of null recurrent diffusions. Typically, transient processes have upward or downward trends, in which case we may eliminate their trends using appropriate detrending methods so that they behave like recurrent processes. Like unit root processes in discrete time, null recurrent processes have stochastic trends and the standard law of large numbers and central limit theory in continuous time are not applicable. See, e.g., Jeong and Park (2013) and Kim and Park (2017) for more details on the statistical properties of null recurrent diffusions.

Let $X^s = s(X)$ be the scale transformation of X, which may be defined as $dX_t^s = m_s^{-1/2}(X_t^s)dW_t$ with speed measure m_s given by

$$m_s = \frac{1}{(s'\sigma)^2 \circ s^{-1}}.$$

Both recurrence and stationarity are preserved under scale transformation. First, X is recurrent on $\mathcal{D}$ if and only if X^s is recurrent on $\mathbb{R}$. Trivially, the scale function of X^s is identity, since it is already in natural scale, and therefore, X^s is recurrent if and only if its domain is given by the entire real line $\mathbb{R}$. However, the domain of X^s becomes $\mathbb{R}$ if and only if X is recurrent, i.e., $s(\underline{x}) = -\infty$ and $s(\overline{x}) = \infty$. Second, X is stationary on $\mathcal{D}$ if and only if X^s is stationary on $\mathbb{R}$, since $m_s(\mathbb{R}) = m(\mathcal{D})$.

Example 2.1. For an illustration, we consider the generalized Höpfner and Kutoyants (GHK) diffusion defined as

$$dX_t = \frac{aX_t}{(c + X_t^2)^{1-b}}dt + (c + X_t^2)^{b/2}dW_t \tag{4}$$

on $\mathbb{R}$ for $a, b \in \mathbb{R}$ and $c > 0$. The GHK model encompasses several diffusion models that are used earlier for illustrative purposes. If, for instance, $a = 0$ or $b = 0$, the GHK diffusion reduces to the diffusion considered by Chen et al. (2010) or Höpfner and Kutoyants (2003), respectively. Moreover, the speed density and speed measure of the GHK model are given respectively by

$$s'(x) = (x^2 + c)^{-a} \quad \text{and} \quad m(x) = (x^2 + c)^{a-b}.$$

The GHK process becomes recurrent if $a \leq 1/2$. Moreover, it becomes positive recurrent if $a - b < -1/2$.

Recall that m_s is defined to be the speed measure of X^s. Concurrently, we define $f_s = f \circ s^{-1}$ for any function f (other than m) on $\mathcal{D}$. Under this notational convention, we have $m_s(f_s) = m(f)$, which follows immediately from a change of variables. This convention will be made throughout the paper. Moreover, for locally integrable f on $\mathbb{R}$, we define $[f]$

as

$$[f](\lambda) = \int_{|x|<\lambda} f(x)dx.$$

This notation will also be used without further reference in what follows.

2.2 Regular Variation and Integrability Condition

We say that $f : (0, \infty) \to \mathbb{R}$ is regularly varying at infinity with index κ, and write as $f \in RV_\kappa$, if $f(\lambda x)/f(\lambda) \to x^\kappa$ as $\lambda \to \infty$ for all $x > 0$ with some $\kappa \in (-\infty, \infty)$. In particular, if $\kappa = 0$ and $f \in RV_0$, then f is said to be slowly varying at infinity.[4] See Bingham et al. (1993) for more discussions on regularly varying functions, as well as their alternative concepts and definitions. For our asymptotics, it is necessary to deal with functions defined on $\mathbb{R}$ and consider both boundaries $x_B = \pm\infty$. The required extension is straightforward and may easily be done as shown in Kim and Park (2017). In particular, for $f \in RV_\kappa$ on $\mathbb{R}$ for some $\kappa \in (-\infty, \infty)$, we have $f(\lambda x)/f(\lambda) \to \overline{f}(x)$ as $\lambda \to \infty$ or $\lambda \to -\infty$ for all $x \neq 0$, where $\overline{f}$, called the *limit homogeneous function*, is given by

$$\overline{f}(x) = |x|^\kappa \big(a1\{x > 0\} + b1\{x < 0\}\big)$$

for some constants a and b such that $|a| + |b| \neq 0$. On the other hand, $f : \mathcal{D} \to \mathbb{R}$ is said to be rapidly varying at boundary x_B with index ∞ or $-\infty$ if $\overline{\kappa} = \underline{\kappa} = \infty$ or $-\infty$ with $\overline{\kappa}$ and $\underline{\kappa}$ defined as $\overline{\kappa} = \sup_\kappa\{x^{-\kappa} f(x) \sim f_\kappa(x)$ at x_B for some nondecreasing $f_\kappa\}$ and $\underline{\kappa} = \inf_\kappa\{x^{-\kappa} f(x) \sim f_\kappa(x)$ at x_B for some nonincreasing $f_\kappa\}$. We write $f \in RV_\infty$ or $f \in RV_{-\infty}$ at x_B for the rapidly varying f of index ∞ and $-\infty$ at x_B, respectively.

Throughout the paper, we assume

Assumption 2.2. *We assume that (a) s' is regularly or rapidly varying with index $\kappa \neq -1$, (b) σ^2 is regularly varying and (c) m is either integrable or regularly varying.*

Note that $s' \in RV_{-1}$ if and only if $s \in RV_0$. In this case, X may either be recurrent or transient, since a slowly varying function may converge or diverge. We exclude this boundary case in our asymptotic analysis. We may easily see that this case arises if and only if $x\mu(x)/\sigma^2(x) \to 1/2$ as $x \to x_B$ at $x_B = \pm\infty$ by the Karamata representation of regularly varying functions (see Bingham et al. (1993)).[5] Furthermore, in this case, s^{-1} becomes rapidly varying.

To effectively present our asymptotics, we introduce

---

[4] Throughout the paper, we use the generic notation ℓ to denote any slowly varying function. The precise definition of ℓ varies from a line to a line.

[5] Note that Assumption 2.2 (a) is implied by $-2x\mu(x)/\sigma^2(x) \to \kappa \in [-\infty, \infty]\backslash\{-1\}$ due to the Karamata representation of regularly varying functions.

Definition 2.1. Let f be a nonintegrable (or m-nonintegrable) regularly varying function on $\mathcal{D}$. We say that f is *strongly nonintegrable* (or m-strongly nonintegrable) if $f\ell$ is not integrable (or not m-integrable) for any slowly varying function ℓ. On the other hand, we say that f is *nearly integrable* (m-nearly integrable) if there exists some slowly varying function ℓ such that $f\ell$ is integrable (or m-integrable).

Following Definition 2.1, we say that a null recurrent diffusion X is *strongly nonstationary* if its speed density m is strongly nonintegrable, and *nearly stationary* if its speed density m is nearly integrable. We assume that m_s and s^{-1} have $+\infty$ as their dominating boundary, i.e., for $f = m_s, s^{-1}$ we have $f(-x)/f(x) = O(1)$ as $x \to \infty$. This assumption is not restrictive and made just for the convenience of exposition.

In the development of our asymptotics, we consider the following three conditions. They characterize our asymptotics in terms of s' and m, which are functions of infinitesimal parameters μ and σ^2 of X. Here and elsewhere in the paper, we denote by ι the identity function on $\mathcal{D}$, i.e., $\iota(x) = x$ for all $x \in \mathcal{D}$.

(ST) : m is either integrable or nearly integrable,

(DD) : $1/s'$ is either integrable or nearly integrable, and

(SI) : ι^2 is either m-integrable or m-nearly integrable.

If ST, DD and SI hold with m being integrable, $1/s'$ being integrable or ι^2 being m-integrable, respectively, we will say that they hold in *strong form*. Clearly, ST is a condition related to the stationarity of X, and it holds if and only if X is stationary or nearly stationary. It is easy to see that ST holds if and only if either m is integrable or $m \in RV_\kappa$ with $\kappa \leq -1$ at $x_B = \pm\infty$ and $\kappa \geq -1$ at $x_B = 0$. On the other hand, SI requires the m-square integrability or near integrability of the identity function.[6]

The implications of DD are more involved. Roughly, DD provides a condition that the drift term of XdX dominates its diffusion term asymptotically. In fact, if we set

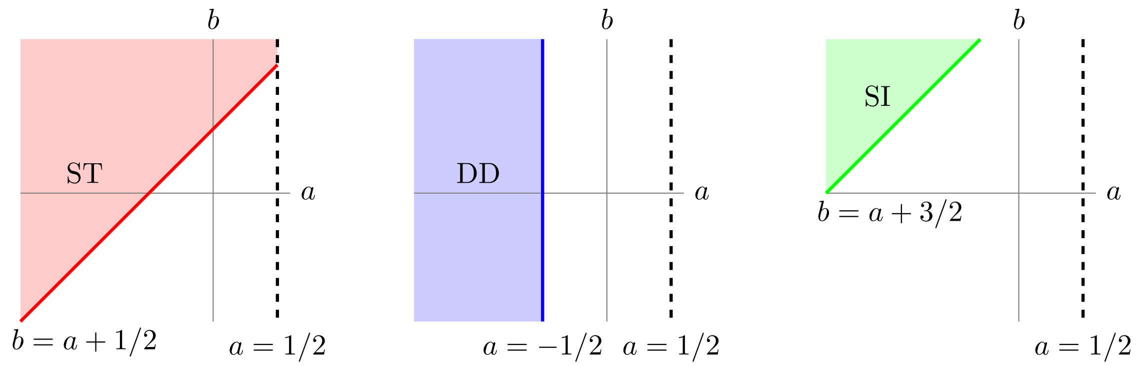
$$X_t dX_t = dN_t + dM_t$$

with $dN_t = X_t\mu(X_t)dt$ and $dM_t = X_t\sigma(X_t)dW_t$, then it follows that

Lemma 2.1. *Let Assumptions 2.1 and 2.2 hold. Then DD holds if and only if $M_T = o_p(N_T)$ as $T \to \infty$.*

[6] If $\mathcal{D} = \mathbb{R}$, SI is stronger than ST, and ST holds in strong form under SI. This, however, is not true if $\mathcal{D} = (0, \infty)$.

Figure 1: Asymptotic Characteristics of GHK Diffusions



Note that $1/s' = m\sigma^2$. Therefore, DD holds if and only if σ^2 is m-integrable or m-nearly integrable, i.e., either $m\sigma^2$ is integrable or $m\sigma^2 \in RV_\kappa$ with $\kappa \leq -1$ at $x_B = \pm\infty$. Due to the recurrence condition, σ^2 is always m-integrable at $x_B = 0$.

Lemma 2.2. *Let Assumption 2.1 hold. For any differentiable function ν on $\mathcal{D}$, we have*

$$m\left(\mu\nu + \frac{1}{2}\sigma^2\nu'\right) = 0$$

if and only if $(\nu/s')(x) = 0$ at $x = x_B$.

If DD holds, it follows from Lemma 2.2 with $\nu = 1$ that $m(\mu) = 0$. In case DD holds in strong form, we may also deduce from Lemma 2.2 with $\nu = \iota$ that $m(\iota\mu) = -(1/2)m(\sigma^2)$.

Example 2.2. The asymptotic characteristics of the GHK diffusion introduced in Example 2.1 are provided in Figure 1. DD and ST hold if and only if $a \leq -1/2$ and $a - b \leq -1/2$ respectively, and SI is satisfied if $a - b \leq -3/2$.

3 Asymptotic Theory of Unit Root Test

In the section, we develop the asymptotics for the Dickey-Fuller test for unit root, which is based on the discrete samples $(X_{i\delta})$, $i = 1, \ldots, n$, collected from the diffusion $X = (X_t)$ over the sample span $T = n\delta$. In what follows, we will simply write $x_i = X_{i\delta}$, $i = 1, \ldots, n$, with $x_0 = X_0$.

3.1 Primary Asymptotics of Unit Root Test

To test for a unit root in (x_i), we may use the first-order autoregression without any augmented lags, since X is a Markov process. In particular, we consider the regression

$$\Delta x_i = \alpha + \beta x_{i-1} + u_i, \tag{5}$$

where Δ is the usual difference operator, and test the null hypothesis $\beta = 0$ against the alternative hypothesis $\beta < 0$ using the least squares regression. The least square estimator and the t-statistic for β in (5), denoted respectively as $\hat{\beta}$ and $t(\hat{\beta})$, are given by

$$\hat{\beta} = \frac{\displaystyle\sum_{i=1}^{n}(x_{i-1} - \overline{x}_n)\Delta x_i}{\displaystyle\sum_{i=1}^{n}(x_{i-1} - \overline{x}_n)^2} \quad \text{and} \quad t(\hat{\beta}) = \frac{\hat{\beta}}{\hat{v}\left(\displaystyle\sum_{i=1}^{n}(x_{i-1} - \overline{x}_n)^2\right)^{-1/2}},$$

where $\overline{x}_n$ is the sample mean of (x_i) and $\hat{v}^2$ is the usual estimator for the variance of regression errors (u_i).

In the usual discrete time setup, the Dickey-Fuller test based on the t-statistic $t(\hat{\beta})$ from regression (5) is widely used to test for the null hypothesis of a unit root, i.e., $\beta = 0$, against the alternative hypothesis of stationarity, i.e., $\beta < 0$. Under the null hypothesis of a unit root, it has nonstandard, yet well-defined, limit distribution, as long as some mild regularity conditions are satisfied for the innovations (u_i). The limit distribution, which we call the Dickey-Fuller distribution, is usually represented as a functional of Brownian motion. On the other hand, under the alternative hypothesis of stationarity, it diverges to negative infinity in probability. Consequently, it provides a test for unit root nonstationarity that is consistent against a wide class of stationary time series. Therefore, we may say, loosely yet generally, that a unit root time series is nonstationary, and that a stationary time series does not have a unit root.

The Dickey-Fuller test has also routinely been applied to samples that are thought to be collected discretely from diffusion type continuous time processes. However, little is known about its asymptotic behavior for discrete samples obtained from general continuous time processes. In particular, except for simple models such as Brownian motion and Ornstein-Uhlenbeck process, it has been completely unknown whether the test can be used to effectively discriminate nonstationary diffusions from stationary diffusions. To facilitate our discussions on the asymptotic behavior of the Dickey-Fuller test, we will simply say in what follows that a diffusion has a unit root if and only if $t(\hat{\beta})$ is stochastically bounded and the test is expected to suggest the non-rejection of the unit root hypothesis at least with some positive probability.

In our asymptotics, we require that the sampling interval δ be sufficiently small relative to the extremal bounds of various functional transforms of X over time interval $[0,T]$. Following Aït-Sahalia and Park (2016), we define

$$T(f) = \max_{0\leq t\leq T} |f(X_t)|$$

for some function $f : \mathcal{D} \to \mathbb{R}$. For the identity function, we have $T(\iota) = \max_{0\leq t\leq T} |X_t|$ and $T(\iota)$ becomes the asymptotic order of extremal process of X. It is obvious that we have $[T(\iota)]^k = T(\iota^k)$ for any nonnegative k. More generally, for any regularly varying f, we may obtain the exact rate of $T(f)$ from the asymptotic behavior of extremal processes. For instance, the extremal processes of Ornstein-Uhlenbeck process and Feller's square root process are respectively of orders $O_p(\sqrt{\log T})$ and $O_p(\log T)$, and the extremal process of the general driftless diffusion process is of order $O_p(T)$. Thus if f is regularly varying and c_T is the order of the extremal process, then we have $T(f) = f(c_T)$. The order of the extremal process is known for a wide class of diffusions. For instance, under some regularity conditions on μ and σ^2, it is well known that the extremal processes of positive recurrent diffusions are of order $O_p(s^{-1}(T))$, to which the reader is referred to, e.g., Davis (1982). Moreover, asymptotic orders for the extremal processes of general null recurrent diffusions are obtained by Stone (1963), Jeong and Park (2013) and Kim and Park (2017).

Assumption 3.1. *We assume that ι, μ and σ are all majorized by $\omega : \mathcal{D} \to \mathbb{R}$ satisfying $\delta T(\omega^4) T \log(T/\delta) \to_p 0$.*

Assumption 3.1 makes it necessary to have $\delta \to 0$. If we fix T, $\delta \to 0$ is indeed the necessary and sufficient condition for Assumption 3.1. Clearly, we may also allow $T \to \infty$ as long as $\delta \to 0$ sufficiently fast. In this case, our asymptotic results will be more relevant for the case where δ is sufficiently small relative to T. Our asymptotics in the paper are derived under the condition $\delta \to 0$ and $T \to \infty$ jointly. For Assumption 3.1 to hold, it suffices to have $\delta = O(T^{-1-\epsilon})$ for any $\epsilon > 0$, if X is bounded so that $T(\omega^4)$ is a constant. The condition appears to be mild enough to yield asymptotics generally relevant for a very wide range of empirical analysis relying on samples collected from diffusion models. For daily observations over ten years, as an example, we have $\delta = 1/252$ and $T^{-1} = 1/10$. Our subsequent asymptotics hold jointly in δ and T as long as they satisfy Assumption 3.1 as $\delta \to 0$ and $T \to \infty$. In particular, we do not use sequential asymptotics, requiring $\delta \to 0$ and $T \to \infty$ sequentially.

The primary asymptotics for $\hat{\beta}$ and $t(\hat{\beta})$ are given by the following lemma. Here and elsewhere in the paper, we let $\overline{X} = (\overline{X}_t)$, $\overline{X}_t = t^{-1}\int_0^t X_s ds$, be the sample mean process of X.

Lemma 3.1. *Let Assumption 3.1 hold. The we have*

$$\hat{\beta} \sim_p \frac{\delta \int_0^T (X_t - \overline{X}_T) dX_t}{\int_0^T (X_t - \overline{X}_T)^2 dt} \quad \textit{and} \quad t(\hat{\beta}) \sim_p \frac{\sqrt{T} \int_0^T (X_t - \overline{X}_T) dX_t}{[X]_T^{1/2} \left(\int_0^T (X_t - \overline{X}_T)^2 dt \right)^{1/2}}$$

for all δ sufficiently small relative to T.

The limit theory in Lemma 3.1 holds for all small enough δ relative to T as long as δ and T satisfy Assumption 3.1. Therefore, we expect that they provide good approximations for finite sample distributions, whenever δ is relatively small compared with T. Note that we do not assume $T = \infty$ to obtain the asymptotics in Lemma 3.1. The asymptotics we have in Lemma 3.1 will be referred to in the paper as the primary asymptotics. The joint asymptotics for $\delta \to 0$ and $T \to \infty$, which are presented below, may be obtained simply by taking T-limits to our primary asymptotics.

Our primary asymptotics in Lemma 3.1 make it clear that we have $\hat{\beta} \to_p 0$ whenever $\delta \to 0$ fast enough relative to T. If, in particular, T is fixed, we have $\hat{\beta} \to_p 0$ for any diffusion X as long as $\delta \to 0$. For discrete samples from any diffusion, we will therefore always observe a root getting close to unity as we collect samples frequently enough and δ becomes sufficiently small. However, even in this case, the unit root test will not necessarily support the presence of a unit root. If we let $\delta \to 0$ with fixed T, $t(\hat{\beta})$ remains to be random and the unit root test will yield a completely random conclusion, regardless of the asymptotic properties of the underlying diffusion. The unit root test will be totally uninformative in this case. This was first observed in Shiller and Perron (1985), and further analyzed subsequently by Perron (1989, 1991), for the test of a unit root in samples from Brownian motion against those from Ornstein-Uhlenbeck process.

3.2 Asymptotics of Unit Root Test

Now we let $T \to \infty$ and establish large T asymptotics for the unit root test. To effectively present our asymptotics, we define (λ_T) to be the normalizing sequence given by

$$T = \lambda_T [m_s](\lambda_T) \quad \text{or} \quad \lambda_T^2 m_s(\lambda_T) \tag{6}$$

depending upon whether or not ST holds. Subsequently, we let

$$a_T = \begin{cases} \lambda_T[m_s\sigma_s^2](\lambda_T) & \text{if} \quad \text{DD holds} \\ \lambda_T^2(m_s\sigma_s^2)(\lambda_T) & \phantom{\text{if}} \quad \text{DD does not hold} \end{cases}$$

$$b_T = \begin{cases} \lambda_T[m_s\iota_s^2](\lambda_T) & \text{if} \quad \text{SI holds} \\ \lambda_T^2(m_s\iota_s^2)(\lambda_T) & \phantom{\text{if}} \quad \text{SI does not hold} \end{cases}$$

from (λ_T), and let

$$P = \begin{cases} L(\tau,0) & \text{if} \quad \text{DD holds} \\ \int_0^\tau \overline{m_s\sigma_s^2}(B_t)dt & \phantom{\text{if}} \quad \text{DD does not hold} \end{cases}$$

$$Q = \begin{cases} 1 - \big(m(\iota)\big)^2 / \big(m(\iota^2)m(\mathcal{D})\big) & \quad \text{SI holds and ST holds} \\ L(\tau,0) & \quad \text{SI holds and ST does not hold} \\ \int_0^\tau \overline{m_s\iota_s^2}(B_t)dt & \text{if} \quad \text{SI does not hold and ST holds} \\ \int_0^\tau \overline{m_s\iota_s^2}(B_t)dt - (\int_0^\tau \overline{m_s\iota_s}(B_t)dt)^2 & \quad \text{SI does not hold and ST does not hold} \end{cases}$$

where τ is a stopping time defined as

$$\tau = \inf\left\{t \,\middle|\, L(t,0) > 1\right\} \quad \text{or} \quad \inf\left\{t \,\middle|\, \int_{\mathbb{R}} L(t,x)\overline{m}_s(dx) > 1\right\}, \tag{7}$$

depending upon whether or not ST holds, from the local time L of Brownian motion B, and $\overline{f}$ denotes the limit homogeneous function of regularly varying f on $\mathbb{R}$. Note that $L(\tau,0) = 1$ a.s. under ST. Numerical sequences (a_T) and (b_T) and random variables P and Q introduced here will be used repeatedly in what follows.

Lemma 3.2. *Let Assumptions 2.1 and 2.2 hold. If either* ST *or* DD *holds, we have*

$$\frac{1}{a_T}[X]_T \to_d P, \quad \frac{1}{a_T}\int_0^T (X_t - \overline{X}_T)dX_t \to_d -\frac{P}{2}$$

and

$$\frac{1}{b_T}\int_0^T (X_t - \overline{X}_T)^2 dt \to_d Q$$

as $T \to \infty$, and $Ta_T/b_T \to \infty$ as $T \to \infty$.

If neither ST nor DD holds, we would have quite different asymptotics. Let $Y = s(X)$ be the scale transformation of X and define Y^T by $Y_t^T = \lambda_T^{-1}Y_{Tt}$ with the normalizing sequence (λ_T) in (6), we have $Y^T \to_d Y^\circ$ as $T \to \infty$ in the space $C[0,1]$ of continuous functions with uniform topology, where using Brownian motion B and its local time L we may represent the limit process Y° as

$$Y_t^\circ = B \circ \overline{A}_t \quad \text{with} \quad \overline{A}_t = \inf\left\{ s \,\middle|\, \int_{\mathbb{R}} L(s,x)\overline{m}_s(dx) > t \right\}. \tag{8}$$

To obtain the asymptotics for a general diffusion X, we write it as $X = s^{-1}(Y)$. Note that if X does not satisfy DD, then $s^{-1} \in RV_p$ with $p = 1/(q+1) > 1/2$ since $s' \in RV_q$ with $q < 1$. Therefore, if we define X^T as $X_t^T = X_{Tt}/s^{-1}(\lambda_T) = s^{-1}(Y_{Tt})/s^{-1}(\lambda_T)$, we may well expect that

$$X^T = \frac{s^{-1}(\lambda_T Y^T)}{s^{-1}(\lambda_T)} \to_d \overline{s^{-1}}(Y^\circ) = X^\circ,$$

in $C[0,1]$ as $T \to \infty$.

Lemma 3.3. *Let Assumptions 2.1 and 2.2 hold. If neither* ST *nor* DD *holds, we have* $\left(s^{-1}(\lambda_T)\right)^{-2}[X]_T \to_d [X^\circ]_1$ *and*

$$\frac{1}{\left(s^{-1}(\lambda_T)\right)^2}\int_0^T (X_t - \overline{X}_T)dX_t \to_d \int_0^1 (X_t^\circ - \overline{X}_1^\circ)dX_t^\circ$$
$$\frac{1}{T\left(s^{-1}(\lambda_T)\right)^2}\int_0^T (X_t - \overline{X}_T)^2 dt \to_d \int_0^1 (X_t^\circ - \overline{X}_1^\circ)^2 dt$$

with $\overline{X}_1^\circ = \int_0^1 X_t^\circ dt$*, as* $T \to \infty$*.*

The asymptotics for unit root test follow immediately from Lemmas 3.1, 3.2 and 3.3.

Theorem 3.4. *Let Assumptions 2.1, 2.2 and 3.1 hold. If neither* ST *nor* DD *holds, we have*

$$n\hat{\beta} \to_d \frac{\int_0^1 (X_t^\circ - \overline{X}_1^\circ)dX_t^\circ}{\int_0^1 (X_t^\circ - \overline{X}_1^\circ)^2 dt} \quad \textit{and} \quad t(\hat{\beta}) \to_d \frac{\int_0^1 (X_t^\circ - \overline{X}_1^\circ)dX_t^\circ}{[X^\circ]_1^{1/2}\left(\int_0^1 (X_t^\circ - \overline{X}_1^\circ)^2 dt\right)^{1/2}}$$

as $\delta \to 0$ *and* $T \to \infty$*. On the other hand, if either* ST *or* DD *holds, we have*

$$\frac{b_T}{Ta_T} n\hat{\beta} \to_d -\frac{P}{2Q} \quad \textit{and} \quad \sqrt{\frac{b_T}{Ta_T}}\, t(\hat{\beta}) \to_d -\sqrt{\frac{P}{4Q}}$$

as $\delta \to 0$ *and* $T \to \infty$*, and* $Ta_T/b_T \to \infty$ *as* $T \to \infty$*.*

Theorem 3.4 shows that the tests based on $n\hat{\beta}$ and $t(\hat{\beta})$ have full asymptotic discriminatory powers for the null and alternative hypotheses, H_0: neither ST nor DD holds, and

H_1: either ST or DD holds. Under H_0, neither ST nor DD holds, and $n\hat{\beta}$ and $t(\hat{\beta})$ have nondegenerate limit distributions. On the other hand, under H_1, either ST or DD holds, $n\hat{\beta} \sim_d (Ta_T/b_T)(-P/2Q) \to_p -\infty$ and $t(\hat{\beta}) \sim_d (Ta_T/b_T)^{1/2}(-P/4Q)^{1/2} \to_p -\infty$, since $P, Q > 0$ with probability one and $Ta_T/b_T \to \infty$. Note that our limit theory in Theorem 3.4 is completely general and holds for a truly broad class of recurrent diffusions. In particular, we do not impose any assumptions on serial dependence or existence of moments. We only rely on some basic regularity conditions in Assumptions 2.1 and 2.2.

It is clear from Theorem 3.4 that the unit root test cannot be used to test for nonstationarity. Null recurrent diffusion may not have a unit root if it satisfies DD and has a dominating drift. As an illustrative example, we consider the GHK model in (4) with $a = -7/5$ and $b = -1$. With the given set of parameter values, X has a dominating drift though it is nonstationary, i.e., DD holds though ST is not satisfied. Therefore, the unit root test is expected to reject the unit root null hypothesis.

We should also note that the unit root test does not have any nontrivial power in discriminating diffusions with and without drift. A recurrent diffusion with linear drift is positive recurrent. Therefore, as long as it is recurrent, any diffusion is stationary and satisfies ST if it has a linear drift. This, in turn, implies that the unit root test rejects the unit root hypothesis for any recurrent diffusion with linear drift. However, in general, the unit root test does not have any discriminatory power for or against the presence of drift in diffusion. In fact, it is easy to see that ST is satisfied by a driftless diffusion, for which we have $m = 1/\sigma^2$, if $x/\sigma^2(x) = O(1)$ as $x \to x_B$ at $x_B = \pm\infty$. In this case, stationarity or near stationarity of X is induced by volatility, not by drift, and we may refer to it as volatility induced stationarity or near stationarity following Conley et al. (1997).

The asymptotics in Theorem 3.4 include the conventional unit root asymptotics as special cases. The conventional unit root asymptotics assume that X° is given by a Brownian motion under H_0, in which case our limit distribution reduces to the standard Dickey-Fuller distribution. To see how our asymptotics reduce the conventional asymptotics under H_1, we let ST, DD and SI all hold in strong form. In this case, we have $\lambda_T \sim Tm(\mathcal{D})$, $a_T \sim T\pi(\sigma^2)$ and $b_T \sim T\pi(\iota^2)$, and $P = 1$ and $Q = 1 - \big(\pi(\iota)\big)^2/\pi(\iota^2)$. Therefore, it follows that

$$\frac{1}{T}n\hat{\beta} \to_p -\frac{1}{2}\frac{\pi(\sigma^2)}{\pi(\iota^2) - \big(\pi(\iota)\big)^2}, \quad \sqrt{\frac{1}{T}}t(\hat{\beta}) \to_p -\sqrt{\frac{1}{4}\frac{\pi(\sigma^2)}{\pi(\iota^2) - \big(\pi(\iota)\big)^2}} \tag{9}$$

as $\delta \to 0$ and $T \to \infty$.

Example 3.1. We consider X with a linear drift function $\mu(x) = a(b - x)$ specified by some parameters a and b and a general diffusion function σ. Let

$$-\frac{2xa(b-x)}{\sigma^2(x)} \to \kappa$$

as $x \to x_B$, and assume $\kappa \in (1, \infty]$ and $[-\infty, 1)$ respectively at $x_B = \pm\infty$ and 0. Then we may readily deduce from the Karamata representation theorem that ST, DD and SI are all satisfied (see Bingham et al. (1993)). It follows from Lemma 2.2 that $\pi(\mu) = 0$ and $\pi(\sigma^2) = -2\pi(\iota\mu)$, from which we have $\pi(\iota) = b$ and $\pi(\sigma^2) = -2[ab\pi(\iota) - a\pi(\iota^2)] = 2a[\pi(\iota^2) - (\pi(\iota))^2]$. Therefore, (9) reduces to

$$\frac{1}{T} n\hat{\beta} \to_p -a, \quad \sqrt{\frac{1}{T}} t(\hat{\beta}) \to_p -\sqrt{\frac{a}{2}},$$

which generalizes the asymptotics of unit root test in Shiller and Perron (1985) and Perron (1989, 1991) obtained for Ornstein-Uhlenbeck process.

As shown, both statistics $n\hat{\beta}$ and $t(\hat{\beta})$ can be used to test for our null and alternative hypotheses introduced above. However, for the rest of the paper, we will exclusively focus on the test based on the t-statistic defined as $t(\hat{\beta})$, since it is used much more commonly in practice.

4 Unit Root, Mean Reversion and Nonstationarity

To fully understand what it means to have a unit root in financial time series, we need the notion of mean reversion, as well as stationarity. In this section, we introduce the notion of mean reversion, and how it is related to the unit root and nonstationarity.

4.1 Unit Root and Mean Reversion

We define

Definition 4.1. We say that X has *mean reversion* if and only if

$$\frac{1}{c_T} \int_0^T (X_t - \overline{X}_T) dX_t \to_d Z$$

as $T \to \infty$, for some normalizing sequence (c_T) and a random variable Z with support on a subset of $(-\infty, 0)$.

For X with mean reversion as defined in Definition 4.1, we say that it has *strong* mean reversion if Z has a point support, and *weak* mean reversion otherwise.

The motivation for our definition of mean reversion is clear. We define

$$\mathrm{MR}_T = \int_0^T (X_t - \overline{X}_T)dX_t \approx \sum_{i=1}^m (X_{t_{i-1}} - \overline{X}_T)(X_{t_i} - X_{t_{i-1}})$$

for $0 = t_0 < \cdots < t_m = T$ with $\max_{1\leq i\leq m} |t_i - t_{i-1}| \approx 0$. Roughly, negative MR_T implies that $(X_{t_{i-1}} - \overline{X}_T)$ has a negative sample correlation with $(X_{t_i} - X_{t_{i-1}})$, i.e., the deviation from sample mean in the current period is negatively correlated with the increment made in transition to the next period. This occurs if and only if X has tendency to increase whenever it is observed below its sample mean, and vice versa. In describing the mean-reverting behavior of diffusion, we may use the recursive mean and define $(X_t - \overline{X}_t)$ as the deviation from mean, in place of $(X_t - \overline{X}_T)$ relying on the sample mean $\overline{X}_T$ over the entire time span. Though we do not show it explicitly in the paper, all our subsequent theories are also applicable for this alternative definition of mean reversion only with some obvious minor changes.

Due to Lemmas 3.2 and 3.3, it follows straightforwardly from Definition 4.1 that

Corollary 4.1. *Let Assumptions 2.1 and 2.2 hold. Then X has mean reversion if and only if either* ST *or* DD *holds.*

In fact, we may readily deduce from Lemma 3.2 that

Corollary 4.2. *Let Assumptions 2.1 and 2.2 hold. Then* ST *or* DD *holds if and only if*

$$\int_0^T (X_t - \overline{X}_T)dX_t \sim_p -\frac{1}{2}\int_0^T \sigma^2(X_t)dt$$

as $T \to \infty$.

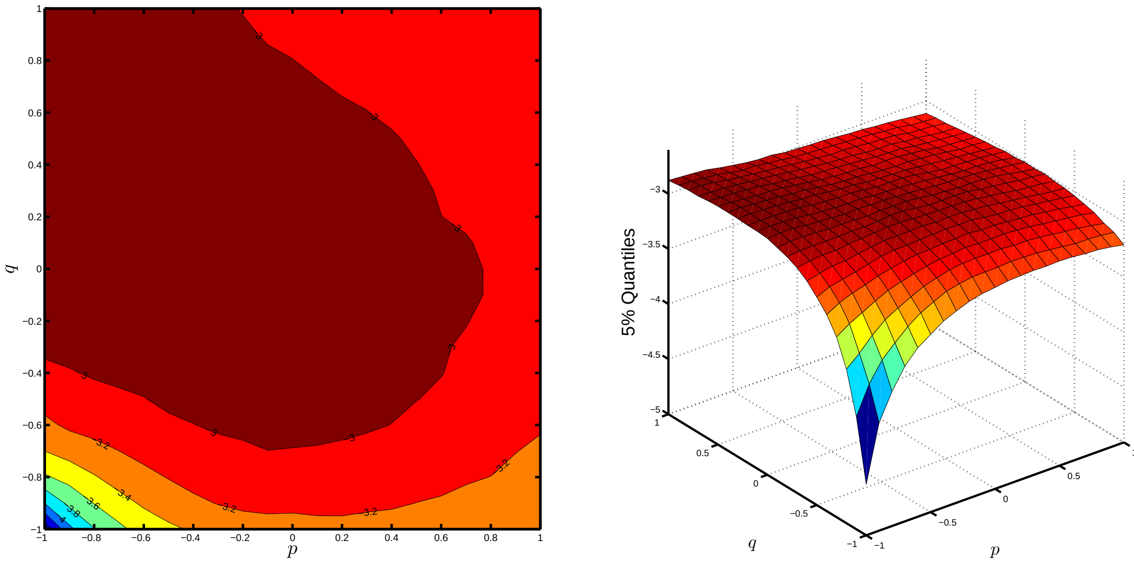
It can be seen from Corollary 4.2 why X has mean reversion if and only if either ST or DD holds. Note that we have $(X_t - \overline{X}_T)dX_t = (1/2)\big(d(X_t - \overline{X}_T)^2 - d[X]_t\big)$, due to Ito's formula, and $d[X]_t = \sigma^2(X_t)dt$. Corollary 4.2 shows that $(X_T - \overline{X}_T)^2$ and $(X_0 - \overline{X}_T)^2$ are asymptotically negligible if and only if either ST or DD holds. In case both ST and DD hold in strong form, we have

$$\frac{1}{T}\int_0^T (X_t - \overline{X}_T)dX_t \to_d -\frac{1}{2}\pi(\sigma^2)$$

as $T \to \infty$, and therefore, X has strong mean reversion. In all other cases, X has weak mean reversion.

Given Theorem 3.4, Corollary 4.1 implies that the unit root test is consistent as a test for the absence of mean reversion against the presence of mean reversion. In particular, it

Figure 2: Asymptotic Critical Values of Unit Root Test for GHK Diffusions



Notes: The data were generated according to the GHK model (4). The left plot presents 5% quantiles of t-statistics for $(p, q) \in [-1, 1] \times [-1, 1]$, and its corresponding Contour plot is on the right.

has asymptotically perfect power against all time series having non-mean-reverting diffusion limits. We may therefore use the unit root test to test for the absence of mean reversion against the presence of mean reversion. The test statistic diverges to minus infinity in the presence of mean reversion. Unfortunately, however, its limit null distribution is heavily model-dependent. In general, the null limit distribution of the unit root test is different from the Dickey-Fuller distribution, and therefore, the standard critical values are not applicable. Therefore, we will consider subsample inference for the test of the absence of mean reversion.

Example 4.1. For the GHK diffusion in Example 2.1, we have $s' \in RV_p$ and $m \in RV_q$ with $p = -2a$ and $q = 2a - 2b$. We conduct simulations for the GHK diffusion for various combinations of $(p, q) \in [-1, 1] \times [-1, 1]$ to examine the model dependency of the unit root test. For each experiment, we simulate 10,000 realizations with $\delta = 1/252$ and $T = 40$ which correspond to 40 years of daily observations. Figure 2 presents the 5% quantiles of the unit root test. It shows that the critical value of the unit root test is highly model dependent. In particular, if both p and q are close to, but not less than, -1, then 5% quantiles of the unit root test for the GHK diffusion are significantly smaller than the critical value from the Dickey-Fuller distribution. For example, if $p = -0.9$ and $q = -0.9$, then the 5% quantile of the test for the GHK diffusion is about -3.7, whereas the corresponding critical value from the Dickey-Fuller distribution is approximately -2.9.

Table 1: Divergence Rate of Unit Root Test

	ST and SI	ST and NSI	NST
DD	$T^{1/2}\ell(T)$	$T^{(p-q-2)/2(p+1)}\ell(T)$	$T^{(p-1)/2(p+q+2)}\ell(T)$
NDD	$T^{1/(p+1)}\ell(T)$	$T^{(-q-1)/2(p+1)}\ell(T)$	1

Notes: Presented is the divergence rate $\sqrt{Ta_T/b_T}$ of the unit root test for X having $s' \in RV_p$ and $m \in RV_q$ with $p \neq -1$. Here NST, NDD and NSI imply that ST, DD and SI do not hold, respectively.

Under the presence of mean reversion, the unit root test diverges up to $-\infty$ as $T \to \infty$ at the rate of $\sqrt{Ta_T/b_T}$. If $s' \in RV_p$ and $m \in RV_q$, then (i) ST holds if and only if $q \leq -1$ at $x_B = \pm\infty$ and $q \geq -1$ at $x_B = 0$, (ii) DD holds if and only if $p \geq 1$ at $x_B = \pm\infty$,[7] and (iii) SI holds if and only if $q \leq -3$ at $x_B = \pm\infty$ and $q \leq -3$ at $x_B = 0$. Moreover, if $p \neq -1$, then s^{-1} becomes regularly varying. In this case, it follows from Karamata's theorem that for some slowly varying function ℓ

$$a_T \sim \begin{cases} \lambda_T \ell(\lambda_T) & \text{if DD holds} \\ \lambda_T^{2/(p+1)}\ell(\lambda_T) & \text{otherwise,} \end{cases} \qquad b_T \sim \begin{cases} \lambda_T \ell(\lambda_T) & \text{if SI holds} \\ \lambda_T^{(p+q+4)/(p+1)}\ell(\lambda_T) & \text{otherwise,} \end{cases}$$

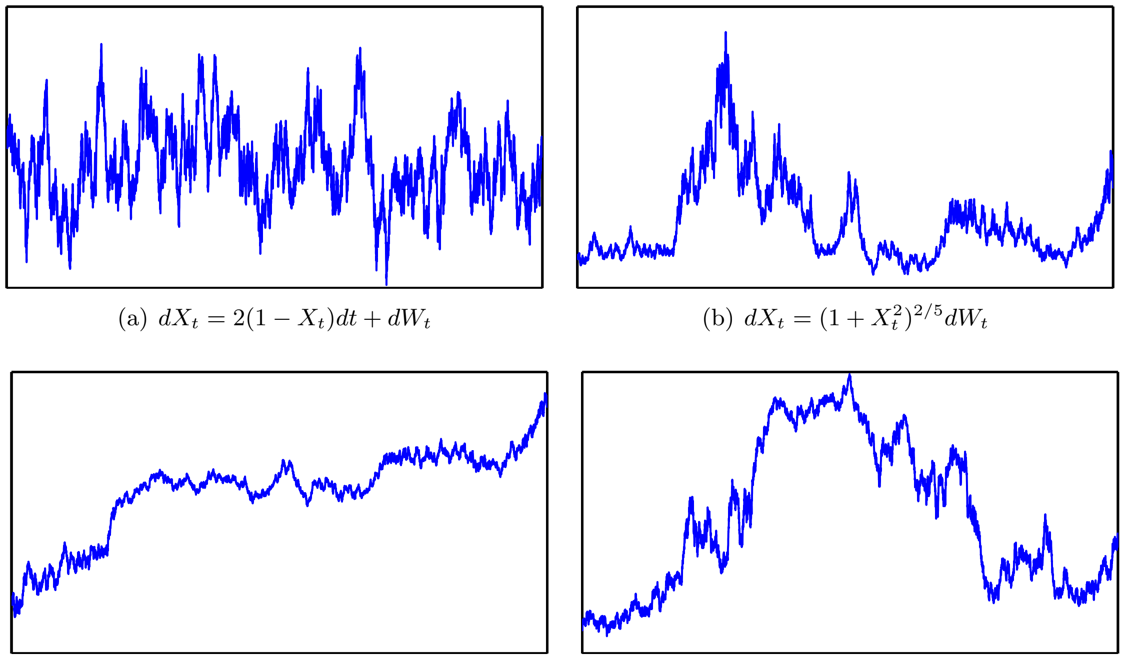
where $\lambda_T \sim T\ell(T)$ or $T^{(p+1)/(p+q+2)}\ell(T)$ depending upon whether or not ST holds, and therefore, we can easily obtain the rates of divergence as shown in Table 1. It is easy to see that $Ta_T/b_T \prec T$ for a nonstationary diffusion satisfying DD.

All stationary and nearly stationary diffusions are mean-reverting, regardless of the existence of mean and other moments. However, stationarity or near stationarity is necessary for mean reversion. Strongly nonstationary diffusions may also be mean-reverting, if their drift terms dominate diffusion terms. This is shown clearly in Corollary 4.1. The existence of mean-reverting nonstationary diffusions, especially strongly nonstationary diffusions, has some important and far-reaching implications on financial time series. It implies that stock prices clearly seen to be strongly nonstationary may still be mean-reverting. Moreover, it allows us to extend the notion of cointegration, and to meaningfully define long run relationships among multiple financial time series with nonstationary error terms, as long as they are mean-reverting. Such an extended notion of cointegration may be practically very useful, for instance, in asset managements relying on pairs trading.

Example 4.2. In Figure 3, we provide the simulated sample paths of four different diffusions with $\delta = 1/252$ and $T = 40$, which correspond to 40 years of daily observations. The sample path of a stationary Ornstein-Uhlenbeck process is in Part (a). It is a stationary

[7] As discussed in Section 2.2, DD always holds at $x_B = 0$.

Figure 3: Sample Paths of Diffusions with Distinctive Asymptotics



(a) $dX_t = 2(1 - X_t)dt + dW_t$

(b) $dX_t = (1 + X_t^2)^{2/5}dW_t$

(c) $dX_t = (1 + X_t^2)^{1/5}dW_t$

(d) $dX_t = -(7/5)X_t(1+X_t^2)^{-2}dt+(1+X_t^2)^{-1/2}dW_t$

process with strong mean reversion, satisfying both ST and DD. Parts (b) and (c) present the sample paths of driftless diffusions. The driftless diffusion in Part (b) satisfies ST but not DD, and becomes a stationary process with weak mean reversion. On the other hand, the driftless diffusion in Part (c) satisfies neither ST nor DD, which implies that it is a nonstationary process with no mean reversion. Finally, Part (d) presents the sample path of the GHK diffusion with parameter values $a = -1.4$ and $b = -1$. It satisfies DD but not ST, and provides an example of a nonstationary process with mean reversion.

4.2 Unit Root and Nonstationarity

As discussed, the unit root test tests the absence of mean reversion, not nonstationarity. There are, however, important special cases, where the former becomes identical to the latter.

Corollary 4.3. *Let Assumptions 2.1 and 2.2 hold. Then (a) if $\overline{\sigma^2}$ is a constant function on $\mathcal{D}$, DD holds if and only if ST holds, and (b) if $\overline{s'}$ is a constant function on $\mathcal{D}$, DD does not hold.*

The result in Corollary 4.3 is well expected from our definitions of ST and DD. If X has a constant diffusion function, we have $m = 1/s'$ up to a constant, and therefore, ST becomes identical to DD. On the other hand, if X has no drift term, DD cannot hold. Consequently, in any of these two cases, ST becomes the only relevant condition for mean reversion.

For a regularly varying function ω with limit homogeneous function $\overline{\omega}$, $\overline{\omega}$ is a constant function on $\mathbb{R}$ if and only if ω is slowly varying and $\omega(-\lambda)/\omega(\lambda) \to 1$ as $\lambda \to \infty$. These conditions are easy to check for σ^2. For s', we have

Lemma 4.4. *$\overline{s'}$ is a constant function on $\mathcal{D} = (-\infty, \infty)$, if and only if*

$$\int_{|x|\leq\lambda} \frac{\mu(x)}{\sigma^2(x)} dx \to 0 \quad \textit{and} \quad \frac{x\mu(x)}{\sigma^2(x)} \to 0$$

as $\lambda \to \infty$ and $x \to x_B = \pm\infty$, respectively.

Clearly, $\overline{s'}$ being a constant function means that X has no drift and is in natural scale asymptotically. If both $\overline{\sigma^2}$ and $\overline{s'}$ are constant functions, X reduces to a Brownian motion asymptotically. For instance, the diffusion defined as

$$dX_t = \frac{X_t}{1 + X_t^2} dt + \big(1 + \log(1 + |X_t|)\big) dW_t$$

becomes a Brownian motion in the limit. For such a diffusion, our limit distribution of the unit root test reduces to the Dickey-Fuller distribution.

Corollary 4.3 makes it clear that if either $\overline{\sigma^2}$ or $\overline{s'}$ is constant, the unit root test may be used to test for nonstationarity. Any of these two conditions, of course, does not hold for general diffusions. However, there are two transformations, which make one of these two conditions hold for all diffusions: the scale transformation and the Lamperti transformation. As discussed, the scale transformation $s(X)$ of X yields a martingale diffusion with no drift function. The Lamperti transformation is given by

$$r(x) = \int_w^x (1/\sigma)(y) dy \tag{10}$$

for some $w \in \mathcal{D}$, and we may easily see that $r(X)$ with r in (10) becomes a unit diffusion having unit diffusion function. Moreover, they preserve ST strongly, as shown below.

Lemma 4.5. *Let Assumptions 2.1 and 2.2 hold. Then (a) X satisfies ST in strong or weak form if and only if $s(X)$ satisfies ST in strong or weak form, and (b) X satisfies ST in strong or weak form if and only if $r(X)$ satisfies ST in strong or weak form if σ is continuously differentiable.*

An immediate consequence of Lemma 4.5 is that the unit root test can be used to test for nonstationarity if it is applied to X transformed by the scale transformation s or the Lamperti transformation r.[8] Of course, such a test is feasible only when s or r can be estimated precisely enough so that the estimated transformation does not affect the test in any essential way. The estimation of s involves the estimation of both the drift and diffusion functions, while the estimation of r only needs an estimate for the diffusion function. It is well known that the drift function is much harder to precisely estimate than the diffusion function. Consistent estimation of the drift term requires the sample span to increase up to infinity. In contrast, the diffusion term estimator is consistent if either the sampling interval shrinks to zero or the sample span diverges. The interest reader is referred to, e.g., Bandi and Phillips (2003) and Aït-Sahalia and Park (2016) for the details.

In the paper, we use the Lamperti transformation r rather than the scale transformation s. The effect of using an estimated s on the unit root test becomes asymptotically negligible only if $S \prec T$, where S and T are the spans of samples used to test for a unit root and to estimate s, respectively. However, the required condition implies that we may only use a subsample of span S for the unit root test with a given sample of span T. This will necessarily entail nontrivial power loss in the unit root test. In contrast, the unit root test based on an estimated r is subject to no such power loss, since we may use samples of the same span to estimate r and to test for a unit root. All that is required is to estimate r using a sample collected at a frequency higher than the frequency the sample used subsequently to test for a unit root is obtained at. This will be shown more precisely below.

We define

$$Y = r(X).$$

and $\hat{Y} = \hat{r}(X)$ correspondingly with an estimate $\hat{r}$ of r. Furthermore, we let $\hat{r}$ be given by $\int_w^x (1/\hat{\sigma})(y)dy$, where $\hat{\sigma}^2$ is an estimator of σ^2 obtained nonparametrically from a discrete sample of time span T and sampling interval ε. As before, we use a discrete sample of time span T and sampling interval δ to test for a unit root. We denote by $\mathcal{D}_T$ the set of values (X_t) takes for $0 \leq t \leq T$, and suppose that

$$\sup_{x \in \mathcal{D}_T} \left| \hat{\sigma}^2(x) - \sigma^2(x) \right| = O_p(\varepsilon^a \lambda_T^{-b} K_T) \tag{11}$$

for some $0 < a, b < 1/2$ and a random sequence (K_T) given by a function of the maximal

[8] Stationarity and nonstationarity of diffusion-type models can also be tested using the approach in either Bandi and Corradi (2014) or Kanaya (2011). Though more general, their tests rely crucially on the choice of test functions and require quite complicated procedures to implement. Moreover, our test proposed here appears to be generally much more powerful in discriminating stationary and nonstationary diffusions.

process of $|X|$ at time T. The standard nonparametric estimators of σ^2 considered in Bandi and Phillips (2003) and Aït-Sahalia and Park (2016) satisfy (11), as shown by, e.g., Bu et al. (2017).

Assumption 4.1. *We assume that (a) Y is a diffusion satisfying Assumptions 2.1 and 2.2 introduced for X, and (b) ι, r and $1/\sigma^3$ are all majorized by $\omega : \mathcal{D} \to \mathbb{R}$ satisfying*

$$\varepsilon^a K_T T(\omega^3) T/\lambda_T^b \to_p 0, \quad \varepsilon^a K_T T(\omega^2) T^{1/2}/(\delta^{1/2}\lambda_T^b) \to_p 0$$

as $\varepsilon \to 0, \delta \to 0$ and $T \to \infty$.

The conditions in Assumption 4.1 do not appear to be stringent. Part (a) is expected to hold widely for X with continuously differentiable σ. Part (b) is satisfied if we choose ε sufficiently small, relative to δ and T. If X is bounded, it holds as long as $\varepsilon^a T/\lambda_T^b \to 0$ and $\varepsilon^a T^{1/2}/(\delta^{1/2}\lambda_T^b) \to 0$.

Proposition 4.6. *Let Assumption 4.1 hold. Then the unit root test using a discrete sample from $\hat{Y}$ is asymptotically equivalent to the test using a discrete sample from Y.*

Proposition 4.6 implies that using an estimated r has no asymptotic effect on the unit root test. Therefore, in particular, the unit root test for Y can be used to test for nonstationarity of X.

As well expected, we may use any other transformation asymptotically equivalent to r. If X is a nonstationary process satisfying Assumptions 2.1 and 2.2, then $1/\sigma$ is nonintegrable at the boundary of m being nonintegrable. Therefore, it follows immediately that

$$r(\theta_T X_t^T) \sim_p \frac{1}{\kappa+1}(\iota/\sigma)(\theta_T X_t^T),$$

where $\theta_T = s^{-1}(\lambda_T)$ and $\kappa > -1$ is the regularly varying index of $1/\sigma$ (see Bingham et al. (1993)). Consequently, we may use ι/σ in place of r, as long as $(\iota/\sigma)(X)$ satisfies Assumptions 2.1 and 2.2.

5 Tests of No Mean Reversion and Nonstationarity

To test for the null hypotheses of no mean reversion and nonstationarity, we consider a subsample test and evaluate its finite sample performance through Monte Carlo simulation. Some empirical illustrations will follow.

5.1 Subsample Test and Monte Carlo Simulation

For the test of no mean reversion and nonstationarity, we may use a subsample test. The conventional bootstrap test seems impossible, since there is no obvious way to impose no mean reversion or nonstationarity even in a very simple class of diffusion models. In the paper, we use a subsample bootstrap method based on a diffusion model fitted using a full sample. For several reasons, we do not use the usual model-free subsample approach. In general, a subsample approach should be implemented with much caution to make it valid in our framework, which allows for general form of nonstationarity. The standard subsample approach does not work, because a nonstationary diffusion has a stochastic trend and subsamples collected in the usual manner have initial values stochastically increasing at a rate of full sample. Therefore, subsamples may not have even a common distribution. Though there are ways to avoid this problem, they would necessarily involve some additional tuning parameters, which is not very desirable.

Our subsample bootstrap method relies on consistent estimates of the drift and diffusion functions, μ_T and σ_T, using samples of time span T and sampling interval δ. We let $\hat{X}$ be generated as

$$d\hat{X}_t = \mu_T(\hat{X}_t)dt + \sigma_T(\hat{X}_t)dW_t$$

from the diffusion defined by μ_T and σ_T.[9] We let S be the time span of subsamples, and assume

$$\sup_{x\in\mathcal{D}_S} |\mu_T(x) - \mu(x)| = O_p(\lambda_T^{-b} K_S) \quad \text{and} \quad \sup_{x\in\mathcal{D}_S} \left|\sigma_T^2(x) - \sigma^2(x)\right| = O_p(\delta^a \lambda_T^{-b} K_S) \tag{12}$$

for $0 < a, b < 1/2$, where all notations are defined similarly as in (11). The conditions in (12) hold for the standard nonparametric estimators of μ and σ^2 considered in Bandi and Phillips (2003) and Aït-Sahalia and Park (2016). This is shown in, e.g., Bu et al. (2017).

Assumption 5.1. *We assume that $\iota, \mu, \sigma, 1/\sigma^2, s', 1/s', m/s'$, and s'/m are all majorized by $\omega : \mathcal{D} \to \mathbb{R}$ satisfying (a) $\delta^a S(\omega^2) \to_p 0$ and (b) $\lambda_T^{-b} K_S S(\omega^7) S \to_p 0$.*

Lemma 5.1. *Let Assumptions 2.1 and 5.1 hold. Then*

$$\sup_{0\leq t\leq S} |\hat{X}_t - X_t| = O_p(\zeta(S,T)),$$

[9]For simplicity, we assume that $\hat{X}$ is generated continuously in time. It is, of course, impossible to generate $\hat{X}$ in continuous time. However, we may use the Euler or Milstein method to generate $\hat{X}$ at an arbitrarily high frequency, so that the discretization effect in generating $\hat{X}$ is negligible compared to other errors. Moreover, we assume that $\hat{X}$ is defined on $\mathcal{D}$. In fact, for our simulations and empirical illustrations, we use the spline method to extend the domain of μ_T and σ_T^2 to the entire $\mathcal{D}$.

where

$$\zeta(S,T) = \sqrt{\lambda_T^{-b} K_S S(\omega^9) S \log\left(\lambda_S^2 \lambda_T^b / (K_S S(\omega^7) S)\right)}.$$

We note that the random sequence $\zeta(S,T)$ in Lemma 5.1 is decreasing in T and increasing in S. Consequently, we have $\zeta(S,T) \to_p 0$ as long as $T \to \infty$ sufficiently faster than $S \to \infty$.

In what follows, we let the numerical sequences (a_S) and (b_S) be defined similarly as in Section 3.2.

Assumption 5.2. *We assume that* $\zeta(S,T)S(\omega)S/a_S \to_p 0$ *and* $\zeta(S,T)S(\omega)S/b_S \to_p 0$.

Lemma 5.2. *Let Assumptions 2.1, 5.1 and 5.2 hold. Then*

$$\begin{aligned}
\frac{1}{a_S}\int_0^S \hat{\sigma}^2(\hat{X}_t)dt &= \frac{1}{a_S}\int_0^S \sigma^2(X_t)dt + o_p(1), \\
\frac{1}{a_S}\int_0^S (\hat{X}_t - \overline{\hat{X}}_S)d\hat{X}_t &= \frac{1}{a_S}\int_0^S (X_t - \overline{X}_S)dX_t + o_p(1), \\
\frac{1}{b_S}\int_0^S (\hat{X}_t - \overline{\hat{X}}_S)^2 dt &= \frac{1}{b_S}\int_0^S (X_t - \overline{X}_S)^2 dt + o_p(1),
\end{aligned}$$

jointly, as $S, T \to \infty$.

The validity of our subsample bootstrap may now be easily established.

Theorem 5.3. *Let Assumptions 2.1, 2.2, 3.1, 5.1 and 5.2 hold. Then the unit root test based on a subsample of time span* S *generated from a diffusion defined by* μ_T *and* σ_T *has the same asymptotics as the test based on the original sample as* $\delta \to 0$ *and* $S, T \to \infty$.

To implement our subsample bootstrap method, we need to determine the subsample horizon S. For the simulations and empirical illustrations in the paper, we choose S that minimizes the running standard deviation of sample quantiles similarly as in Romano and Wolf (2001). More precisely, we set S as described below.

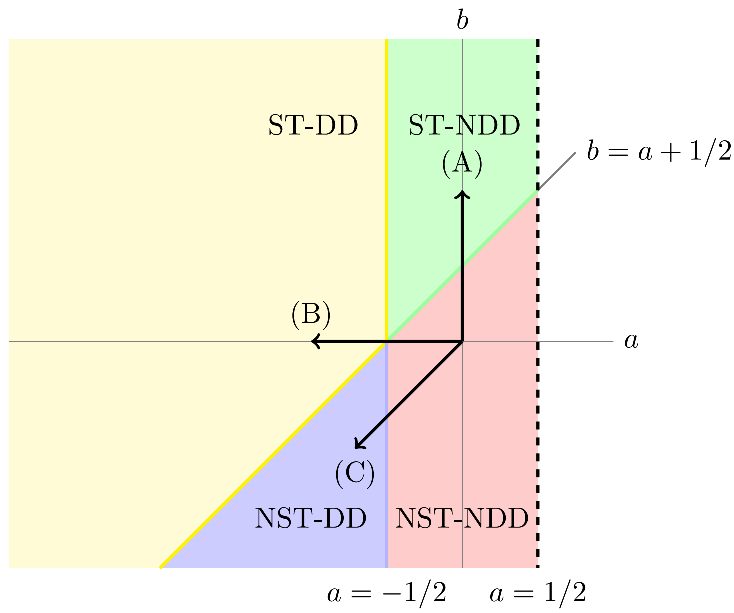
Step 1: For each $T_s \in [(1/2)T^{1/4}, 2T^{1/4}]$, we simulate $B = 5,000$ subsamples $(\hat{X}_{i\delta}^b)_{i=0}^{n_s}$ for $b = 1, ..., B$ from the estimated diffusion $\hat{X}$ with $\hat{X}_0 = X_0$ and $n_s\delta = T_s$, and obtain the empirical distribution of $t_b(T_s)$ for each T_s.[10]

Step 2: For each T_s, we compute the α% quantile, $Q(T_s)$, of the empirical distribution of $\big(t_b(T_s)\big)$ obtained in Step 1.

Step 3: We find $S = T_s$ that minimizes the running standard deviation of sample quantile $Q(T_s)$ over $T_s \pm k/D$ for $k = 0, 1, \ldots, K$ with $K = 4$ and $D = 12 \times 4.2$.

[10] Since the interval $[(1/2)T^{1/4}, 2T^{1/4}]$ is continuous, we discretize the interval and consider $T_s \in [(1/2)T^{1/4}, 2T^{1/4}]$ given by $T_s = (1/2)T^{1/4} + k/D$ for $k = 0, 1, \ldots$, where D is defined in Step 3.

Figure 4: Asymptotic Properties of Simulated GHK Models



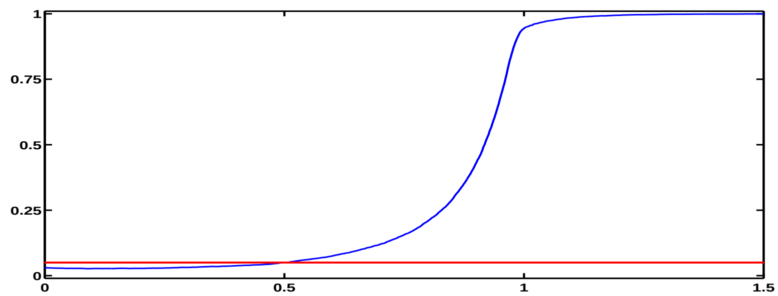
Notes: The areas labeled as ST and DD (NST and NDD) present the set of parameter values (a, b) of the GHK model (not) satisfying the conditions for mean reversion. For $(a, b) = (0, 0)$, the GHK model reduces to Brownian motion, which is nonstationary and non-mean-reverting. Three vectors originated from $(a, b) = (0, 0)$ indicate the directions we look for and obtain our simulation models under the alternative hypothesis of mean reversion and stationarity.

Our procedure is a natural extension of that in Romano and Wolf (2001) for subsamplings in continuous time models. In our procedure, we search for the subsample span, instead of the subsample size, which minimizes the running standard deviation of a sample quantile of the test statistic. Our choice of $K = 4$ and $D = 12 \times 4.2$ imply that the running standard deviation is computed from sample quantiles obtained for subsample sizes of plus and minus four weeks around any given subsample span in weekly increments.

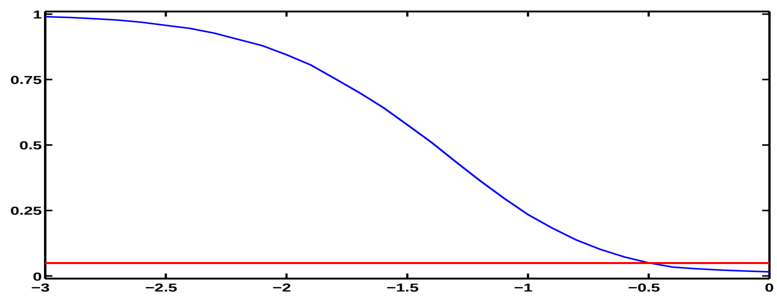
5.2 Simulations

For our simulations, we use the GHK model in (4). As discussed, the mean reversion and nonstationarity properties of the model are fully characterized by (a, b). In particular, ST holds if and only if $b \geq a + 1/2$, DD holds if and only if $a \leq -1/2$, and SI holds if and only if $b \geq a + 3/2$. This is described in Figure 1. To effectively compare the size and power performances of the test, we consider three different specifications: (A) $a = 0$, (B) $b = 0$, and (C) $a = b$ as illustrated in Figure 4. Recall that, under (A) and (B), the GHK model reduces respectively to the models considered by Chen et al. (2010) and Höpfner

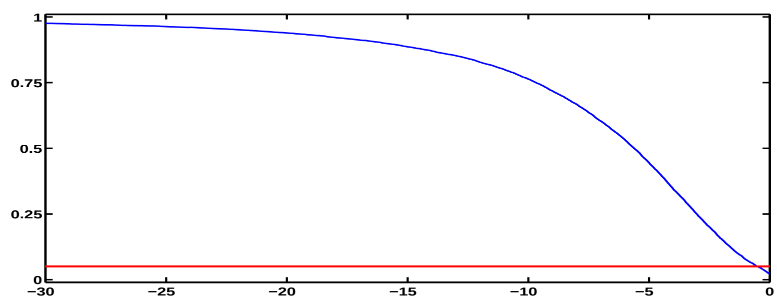
Figure 5: Power Functions for Simulated GHK Models



(a) Specification (A): $a = 0$ and $0 \leq b \leq 1.5$



(b) Specification (B): $-3 \leq a \leq 0$ and $b = 0$



(c) Specification (C): $-30 \leq a, b \leq 0$ and $a = b$

Notes: Presented are the power functions of the GHK model for each of our specifications (A), (B) and (C), respectively, in case of $T = 40$.

and Kutoyants (2003). On the other hand, under (C), the GHK model has speed measure given by the Lebesgue measure as for Brownian motion, and it is always nonstationary. The divergence rate of the unit root test is given by $T^{(2b-1)/2}\ell(T)$ and $T\ell(T)$ respectively for $1/2 \leq b < 3/2$ and $b \geq 3/2$ under (A), $T^{-(1+2a)/(1-2a)}\ell(T)$ and $T^{1/2}\ell(T)$ respectively for $-3/2 < a \leq -1/2$ and $a \leq -3/2$ under (B), and $T^{-(1+2a)/(4-4a)}$ under (C).

Here we only report our simulation results for the unit root test without the Lamperti transformation, i.e., the test of no mean reversion. As discussed, the test of nonstation-

Table 2: Parameter Values for Simulated GHK Models

Model	I	II	III	IV
A	$(0, 1/3)$	$(0, 1/2)$	$(0, 0.892)$	$(0, 0.967)$
B	$(-1/3, 0)$	$(-1/2, 0)$	$(-1.25, 0)$	$(-1.89, 0)$
C	$(-1/3, -1/3)$	$(-1/2, -1/2)$	$(-4.53, -4.53)$	$(-11.0, -11.0)$

Notes: Presented are the parameter values (a, b) of the GHK models used in our simulations, which are referred to as Models I, II, III and IV for each of the specifications (A), (B) and (C), respectively.

arity requires the Lamperti transformation based on the estimated diffusion function in a preliminary step. However, we find in our simulations that the required preliminary transformation virtually has no effect on the finite sample performance of the unit root test. The finite sample performance of the test of nonstationarity is therefore essentially the same as the test of no mean reversion for the diffusion models with a constant diffusion, as given by our specification (B).

To find the parameter values of the GHK model suitable for our simulations, we let $T = 40$ and compute the power functions of the 5% unit root test under our specifications (A), (B) and (C). They are presented in Figure 5. To obtain them, we simulate 10,000 sample paths at $\delta = 1/252$. The critical values of the test are set so that the GHK models with parameter values at the boundaries of the null and alternative hypotheses have 5% rejection probability. We consider four models in our simulations for each of our specifications (A), (B) and (C), which are referred to as Models I-IV. Models I and II are defined as the GHK model with parameter values in the non-mean-reverting region and at the boundaries of the mean-reverting region, respectively. On the other hand, Models III and IV are given by the GHK models with parameter values chosen to have 30% and 90% rejection probabilities, respectively, when Model II has 5% rejection probability, in case that $T = 40$. The simulations for our subsample test are based on 5,000 sample paths generated for $T = 20, 40, 60$ at $\delta = 1/252$. Our simulation models are summarized in Table 2.

In our simulations, we consider two versions of our subsample test: infeasible and feasible versions. To compute the critical values, the former uses subsamples generated from the true diffusion model, whereas the latter relies on subsamples obtained from the estimated diffusion model as discussed earlier. To estimate the underlying diffusion model required for the feasible subsample test, we use the standard nonparametric estimators of μ and σ^2 considered in Bandi and Phillips (2003) and Aït-Sahalia and Park (2016). As discussed, we use the spline method to extend the domain of our estimates μ_T and σ_T^2 of μ and σ^2 to facilitate generating subsamples. For both versions of the test, the initial values of the

Table 3: Rejection Probabilities

T	Model	Feasible Test I	II	III	IV	Infeasible Test I	II	III	IV
20	A	0.0344	0.0476	0.1494	0.2244	0.0548	0.0726	0.1920	0.2814
	B	0.0326	0.0516	0.2260	0.5192	0.0650	0.0838	0.2658	0.5306
	C	0.0236	0.0236	0.1502	0.4000	0.0582	0.0620	0.1758	0.3990
40	A	0.0540	0.0814	0.4950	0.8690	0.0620	0.0916	0.5038	0.8806
	B	0.0568	0.0918	0.4684	0.8260	0.0766	0.1078	0.4720	0.8108
	C	0.0360	0.0362	0.2882	0.6292	0.0636	0.0668	0.3064	0.6164
60	A	0.0530	0.0916	0.7798	0.9560	0.0566	0.1002	0.7708	0.9734
	B	0.0662	0.1148	0.6212	0.9176	0.0768	0.1266	0.6118	0.9032
	C	0.0364	0.0382	0.3914	0.7650	0.0558	0.0634	0.4240	0.7464

simulated subsamples are set to be the same as those of the original sample.

Our simulation results are summarized in Table 3. Overall finite sample performance of the feasible test is not much worse than that of the infeasible test. The rejection probabilities for the feasible and infeasible tests are largely comparable, and we do not see any obvious systematic pattern in their relative magnitudes. It seems that using the estimated model, instead of the true model, has no significant deleterious effect on the finite sample performance of our subsample test. Of course, we may only use the feasible test in practical application. Recall that Model I is non-mean-reverting, whereas Models II-IV are all mean-reverting. Under Model I, the test appears to slightly under-reject the null hypothesis when T is small. However, the problem disappears, or at least clearly improves, as T increases. As expected, the rejection probabilities of the test for Model II, the boundary case, are not noticeably larger than the actual size of the test. Nevertheless, they seem to steadily increase as T gets large. The finite sample powers of the test become larger for Models III and IV in all three specifications (A), (B) and (C). For large T, in particular, the test has reasonable powers against Models III and IV in most cases.

5.3 Empirical Illustrations

For the purpose of illustration, we apply our subsample test to a selected set of financial time series including daily observations of exchange rates, interest rates, commodity prices and stock indexes. The test is implemented exactly as in our simulations. Considered are two tests, one for the test of no mean reversion and the other for the test of nonstationarity, which are referred respectively to as MR and ST tests. As discussed, we need to estimate

Table 4: Exchange Rates: P-values and Test Results

	MR Test	ST Test	5%-Test	10%-Test
JPY/USD	0.2930	0.3578	NMR-NST	NMR-NST
CHF/USD	0.7356	0.6894	NMR-NST	NMR-NST
GBP/USD	0.0250	0.0228	MR-ST	MR-ST
ASD/NZD	0.0400	0.0802	MR-NST	MR-ST

the Lamperti transformation to implement the ST test. Following our theoretical results, we estimate the Lamperti transformation at higher frequency (daily), and apply the unit root test at lower frequency (weekly) after the transformation.

We test four exchange rates data, JPY/USD, CHF/USD, GBP/USD and ASD/NZD, respectively for Japan-US, Swiss-US, UK-US and Australia-New Zealand, from October 2, 1980 to August 5, 2016, which are obtained from FRED from St. Louis Fed. The tests are applied to the logarithms of exchange rates, and we report the results in Table 4. The evidence for no mean reversion and nonstationarity in JPY/USD and CHF/USD appears to be rather strong. In contrast, the evidence against no mean reversion and nonstationarity in GBP/USD is equally strong, suggesting unambiguously that it is stationary (and mean-reverting). Interestingly, we see some evidence of mean-reverting nonstationarity in ASD/NZD.

For the interest rates, we test 3 month T-bill and 10 year T-bill rates, and their spreads over the period from January 2, 1962 to August 11, 2016. The subperiod ranging from January 2, 1962 to August 14, 2008 is also considered to exclude the effect of recent near zero interest rates. The data are downloaded from FRED. We report the results in Table 5. For both rates, the tests strongly suggest no mean reversion and nonstationarity, regardless of whether we include or exclude the recent period of near zero interest rates. The p-values of the tests are slightly lower for the subperiod, compared to those for the entire period. The differences, however, are not significant. On the other hand, the spreads between 3 month T-bill and 10 year T-bill rates are clearly stationary (and mean-reverting) with p-values less than 1-2%. It seems clear that they are cointegrated, sharing one common stochastic trend. Our findings here are consistent with those of Stock and Watson (1988) based on monthly observations.

We also test some commodity prices. Considered are three different types of commodity prices, agricultural products (Corn and Wheat, December 8, 1988 - March 3, 2016), precious metals (Gold and Silver, July 2, 1973 - August 23, 2016) and oils (WTI and Gasoline, June 2, 1986 - August 15, 2016). Corn and Wheat prices are non-adjusted future prices based on

Table 5: Interest Rates: P-values and Test Results

Full Sample	MR Test	ST Test	5%-Test	10%-Test
3M	0.5340	0.8442	NMR-NST	NMR-NST
10Y	0.7964	0.8268	NMR-NST	NMR-NST
10Y−3M	0.0010	0.0092	MR-ST	MR-ST
Sub-Sample	**MR Test**	**ST Test**	**5%-Test**	**10%-Test**
3M	0.3268	0.6228	NMR-NST	NMR-NST
10Y	0.6822	0.6174	NMR-NST	NMR-NST
10Y−3M	0.0112	0.0188	MR-ST	MR-ST

spot-month continuous contract calculations from Chicago Mercantile Exchange, and Gold and Silver prices are London Fixing Price in London Bullion Market with fixing levels set per troy ounce. Corn, Wheat, Gold and Silver prices are downloaded from Quandl, and WTI and Gasoline prices are from FRED. All prices are taken logarithms, and their differences are also tested. The results are summarized in Table 6. All commodity prices considered here seem to be clearly non-mean-reverting and nonstationary. They fail to reject both hypotheses unambiguously with large p-values. The spreads in their prices, however, are all mean-reverting with p-values almost negligible. The price spreads in Wheat/Corn and WTI/Gasoline are also strongly supportive of stationarity with negligible p-values. Therefore, our test results suggest, quite convincingly, the presence of cointegration between these commodity prices in logarithms. On the contrary, the spreads in the prices of Gold/Silver shows some strong evidence of nonstationarity. This implies that the Gold/Silver prices in logarithms are cointegrated only in an extended sense.[11]

Finally, we test four international stock indexes from Morgan Stanley Capital International (MSCI), developed markets (DW), emerging markets (EM) and developed market in Europe (DE) from August 2, 1988 to June 9, 2016. The data are obtained from the Datastream Research Service. We apply the tests to the cumulative log returns, of each index and the differences of the cumulative returns, and report the results in Table 7. All stock indexes we consider are non-mean-reverting and nonstationary. The evidence is strong and unambiguous. The differences in log cumulative returns of DW/EM and DE/EM are also non-mean-reverting and nonstationary, implying the absence of cointegration. However,

[11] Our finding here sheds light on the previous test results for cointegration between the Gold/Silver prices. Escribano and Granger (1998), Lucey and Tully (2006) and Baur and Tran (2014) all find some evidence of a cointegrating relationship between the Gold/Silver prices, but they also observe quite clearly that the nature of the relationship is time varying and unstable.

Table 6: Commodity Prices: P-values and Test Results

	MR Test	ST Test	5%-Test	10%-Test
Wheat	0.2866	0.2918	NMR-NST	NMR-NST
Corn	0.4232	0.3954	NMR-NST	NMR-NST
Wheat/Corn	0.0010	0.0004	MR-ST	MR-ST
Gold	0.7508	0.6570	NMR-NST	NMR-NST
Silver	0.3792	0.3956	NMR-NST	NMR-NST
Gold/Silver	0.0096	0.1220	MR-NST	MR-NST
WTI	0.5390	0.4082	NMR-NST	NMR-NST
Gasoline	0.4568	0.3746	NMR-NST	NMR-NST
WTI/Gasoline	0.0000	0.0000	MR-ST	MR-ST

Table 7: Stock Indexes: P-values and Test Results

	MR Test	ST Test	5%-Test	10%-Test
DW	0.4886	0.4778	NMR-NST	NMR-NST
DE	0.2814	0.2942	NMR-NST	NMR-NST
EM	0.3266	0.2706	NMR-NST	NMR-NST
DW−EM	0.5888	0.5748	NMR-NST	NMR-NST
DE−EM	0.6706	0.6952	NMR-NST	NMR-NST
DW−DE	0.0284	0.2226	MR-NST	MR-NST

Notes: For each of DE, DW and EM indexes, the stock markets in the following countries are included: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and United Kingdom are in DE, Australia, Canada, Hong Kong, Japan, New Zealand, Singapore, and United States, as well as all countries in DE, are in DW, and Argentina, Brazil, Chile, China, Colombia, Czech Republic, Egypt, Hungary, India, Indonesia, Israel, Jordan, Korea, Malaysia, Mexico, Morocco, Pakistan, Peru, Philippines, Poland, Russia, South Africa, Taiwan, Thailand, Turkey, and Venezuela are in EM.

there is one exception. Though the differences in cumulative log returns of DW/DE are nonstationary, they are mean-reverting. This supports the presence of an extended notion of cointegration between DW/DE cumulative log returns.

6 Conclusion

In this paper, we clarify how the notions of unit root, mean reversion and nonstatonarity in financial time series generated from general diffusion models are related to each other.

In particular, we show that mean reversion is not synonymous with stationarity. While all stationary diffusions are mean-reverting, not all nonstationary diffusions are non-mean-reverting. Nonstationary diffusions may also be mean-reverting if they have time-varying stochastic volatilities and the drift terms dominating the diffusion terms. The existence of nonstationary mean-reverting processes makes it useful to extend the concept of cointegration. The unit root test can be used to test for no mean reversion, and also for nonstationarity, in general diffusion models. Both tests are fully consistent within our general framework allowing for all recurrent diffusions. The limit distributions of the tests are generally model-dependent, and the usual critical values of the unit root test are not applicable. To implement the tests, we propose a subsample bootstrap test and develop its asymptotics. In our illustrative applications of the tests, we demonstrate that many nonstationary time series are mean-reverting, and also that two non-mean-reverting nonstationary time series may well have errors that are mean-reverting though nonstationary.

Appendix A Useful Lemmas

A.1 Lemmas for Diffusion Asymptotics

Throughout, we assume that Assumptions 2.1 and 2.2 hold and X satisfies the recurrence condition $s(\underline{x}) = -\infty$ and $s(\overline{x}) = \infty$.

Lemma A1. *Let f be regularly varying on $\mathcal{D}$. (a) If f is m-nonintegrable at x_B, then $s^{-1} \in RV_\kappa$ at $s(x_B)$ with $0 < \kappa < \infty$ for $x_B = \pm\infty$, and $-\infty < \kappa < 0$ for $x_B = 0$. (b) f is m-nearly integrable/strongly nonintegrable at x_B if and only if f_s is m_s-nearly integrable/strongly nonintegrable at $s(x_B)$.*

Proof of Lemma A1. For the first part, note that $mf = f/(s'\sigma^2)$ with regularly varying σ and s' is nonintegrable at x_B. If s' is rapidly varying, it has to be rapidly increasing since it is nonintegrable at x_B. Therefore, for any regularly varying f, f becomes m-integrable. It follows that $s' \in RV_p$ with $-1 < p < \infty$ for $x_B = \pm\infty$, and $-\infty < p < -1$ for $x_B = 0$. By the property of regularly varying function (see, e.g., Proposition 1.5.7. in Bingham et al. (1993)), $s^{-1} \in RV_\kappa$ at $s(x_B)$, where $\kappa = 1/(p+1)$ with $0 < \kappa < \infty$ for $x_B = \pm\infty$, and $-\infty < \kappa < 0$ for $x_B = 0$. This completes the proof for the first part.

For the second part, let $s' \in RV_p$ at x_B. If $\sigma^2 \in RV_q$ and $f \in RV_r$ at x_B, $mf \in RV_a$ with $a = r - p - q$, whereas $m_s f_s \in RV_b$ at $s(x_B)$ with $b = (r - 2p - q)/(p+1)$ since $f_s m_s = (f/(s'\sigma)^2) \circ s^{-1}$. However, we have $m(f) = m_s(f_s)$, from which it follows that $b \geq -1$ if and only if $a \geq -1$ for $x_B = \pm\infty$, and $a \leq -1$ for $x_B = 0$. The stated result therefore follows immediately from Karamata's theorem. □

Lemma A2. *Let m be nonintegrable at x_B. (a) $1/\sigma$ is nonintegrable at x_B. (b) $r(x_B) = \infty$ for $x_B = \overline{x}$, and $r(x_B) = -\infty$ for $x_B = \underline{x}$. (c) m is nearly integrable/strongly nonintegrable at x_B if and only if $1/(s'\sigma) \circ r^{-1}$ is nearly integrable/strongly nonintegrable at $r(x_B)$.*

Proof of Lemma A2. Note that $m \in RV_a$ with $a \geq -1$ for $x_B = \pm\infty$, and $a \leq -1$ for $x_B = 0$. As shown in the proof of Lemma A1 (a), m being nonintegrable at x_B implies that $s' \in RV_p$ with $-1 < p < \infty$ for $x_B = \pm\infty$, and $-\infty < p < -1$ for $x_B = 0$. If $\sigma^2 \in RV_q$ at x_B, then $a = -p - q$ at x_B since $m = 1/(s'\sigma^2)$. Therefore, $q < 2$ at $x_B = \pm\infty$ and $2 < q$ at $x_B = 0$, and hence, $1/\sigma \in RV_{-q/2}$ is nonintegrable at x_B by Karamata's theorem, which completes the proof of the part (a). The part (b) follows immediately from the part (a).

For the part (c), note that m is nonintegrable at x_B if and only if $1/(s'\sigma) \circ r^{-1}$ is nonintegrable at $r(x_B)$ by a change of variables in integrals. Moreover, $r \in RV_{(2-q)/2}$ at $r(x_B)$ due to the proof of the part (a) with Karamata's theorem. Consequently, $1/(s'\sigma) \circ r^{-1} \in RV_b$ at $r(x_B)$, where $b = -(2p+q)/(2-q)$ with $-1 \leq b$ if and only if $-1 \leq a$ for

$x_B = \pm\infty$, and $a \leq -1$ for $x_B = 0$, from which, jointly with the part (b), we have the stated result in the part (c). $\square$

Lemma A3. *Let τ be the stopping time defined in* (7).

(a) If f and g^2 are m-integrable, then

$$\frac{1}{\lambda_T}\int_0^T f(X_t)dt \to_d L(\tau,0),$$
$$\frac{1}{\sqrt{\lambda_T}}\int_0^T g(X_t)dW_t \to_d \sqrt{L(\tau,0)}N,$$

where N is a standard normal random variate independent of $L(\tau,0)$.

(b) If f and g^2 are m-strongly nonintegrable, then

$$\frac{1}{\lambda_T^2(m_s f_s)(\lambda_T)}\int_0^T f(X_t)dt \to_d \int_0^\tau \overline{m_s f_s}(B_t)dt,$$
$$\frac{1}{\lambda_T(m_s^{1/2}g_s)(\lambda_T)}\int_0^T g(X_t)dW_t \to_d \int_0^\tau \overline{m_s^{1/2}g_s}(B_t)dB_t.$$

(c) If f and g^2 are m-nearly integrable, then

$$\int_0^T f(X_t)dt \sim_d \lambda_T[m_s f_s](\lambda_T)L(\tau,0) + \lambda_T^2(m_s f_s)(\lambda_T)\lim_{\varepsilon\to 0}\int_0^\tau \frac{1}{B_t}1\{|B_t| > \varepsilon\}dt,$$

or

$$\frac{1}{\lambda_T[m_s f_s](\lambda_T)}\int_0^T f(X_t)dt \to_d L(\tau,0),$$

depending upon whether $\overline{m_s f_s}(x) = 1/x$ or $\overline{m_s f_s}(x) \neq 1/x$, and

$$\frac{1}{\sqrt{\lambda_T[m_s g_s^2](\lambda_T)}}\int_0^T g(X_t)dW_t \to_d \sqrt{L(\tau,0)}N,$$

where N is a standard normal random variate independent of $L(\tau,0)$.

(d) If f is m-nearly integrable, then

$$\int_0^T f(X_t)dt = O_p(\lambda_T[m_s|f_s|](\lambda_T)).$$

(e) If ST holds, then $L(\tau,0) = 1$ with probability one.

Proof of Lemma A3. Due to Lemma A1 (b), the stated results in (a), (b) and (c) follow respectively from Theorems 3.3, 3.4 and 3.5 of Kim and Park (2017).

As for the part (d), note that if f is m-nearly integrable, then $|f|$ is m-nearly integrable and $\overline{m_s|f_s|} \neq 1/x$. It then follows from the part (c) of this lemma that

$$\frac{1}{\lambda_T[m_s|f_s|](\lambda_T)} \int_0^T f(X_t)dt \leq \frac{1}{\lambda_T[m_s|f_s|](\lambda_T)} \int_0^T |f|(X_t)dt \to_d L(\tau, 0)$$

which completes the proof of the part (d). Finally, the part (e) follows immediately from the construction of τ. $\square$

Lemma A4. *Let X be null recurrent, and define X^T by $X_t^T = \theta_T^{-1} X_{Tt}$ for $0 \leq t \leq 1$.*

(a) We have $X^T \to_d X^\circ$, where $X^\circ = \overline{s^{-1}}(Y^\circ)$ for a strongly nonstationary X, and $X^\circ = 0$ a.s. for a nearly stationary X. Here Y° is defined in (8).

(b) If X is strongly nonstationary, then X° becomes a semimartingale with quadratic variation $[X^\circ]$ is given by $[X^\circ]_t = \int_0^{\overline{A}_t} \left((\overline{s^{-1}})'_- \right)^2 (B_s) ds$, where $(\overline{s^{-1}})'_-$ is the left-hand derivative of $\overline{s^{-1}}$, and $\overline{A}_t = \inf \left\{ s | \int_{\mathbb{R}} \overline{m}_s(x) L(s, x) > t \right\}$.

Proof of Lemma A4 (a). Since m_s and s^{-1} have $+\infty$ as their dominating boundary, $s^{-1} \in RV_\kappa$ at $+\infty$ with $0 < \kappa < \infty$ by Lemma A1 (a) with $f = 1$. Moreover, if we let $Y = s(X)$,

$$X_t^T = \frac{s^{-1}(\lambda_T Y_{Tt}/\lambda_T)}{s^{-1}(\lambda_T)} = \overline{s^{-1}}(Y_{Tt}/\lambda_T)(1 + o_p(1)) \to_d \overline{s^{-1}}(Y_t^\circ)$$

by Proposition 3.2 of Kim and Park (2017), the continuous mapping theorem and the uniform convergence of regularly varying functions (see, e.g., pages 21-22 in Bingham et al. (1993)). This completes the proof. $\square$

Proof of Lemma A4 (b). Note that $\overline{s^{-1}}$ is nondecreasing and can be represented as the difference of two convex functions. As in Kim and Park (2017), we may apply Itô-Tanaka formula to $Y^\circ = B \circ \overline{A}$ to deduce

$$\begin{aligned} X_t^\circ - X_0^\circ &= \overline{s^{-1}}(B \circ \overline{A}_t) - \overline{s^{-1}}(B \circ \overline{A}_0) \\ &= \int_0^t (\overline{s^{-1}})'_-(B \circ \overline{A}_s) d(B \circ \overline{A}_s) + \frac{1}{2} \int_{\mathbb{R}} L(\overline{A}_t, x)(\overline{s^{-1}})''(dx), \end{aligned} \tag{A.1}$$

where $(\overline{s^{-1}})''$ is the second derivative of $\overline{s^{-1}}$ in the sense of distributions. In (A.1), the first and second terms represent respectively the martingale and bounded variation components of the semimartingale X°, for which we have $[X^\circ]_t = \int_0^{\overline{A}_t} \left((\overline{s^{-1}})'_- \right)^2 (B_s) ds$. This completes the proof. $\square$

Lemma A5. *If ST holds, we have (a) $X_T = o_p(\theta_T)$, and (b) $\overline{X}_T = o_p(\theta_T)$.*

Proof of Lemma A5. If X is stationary, $X_T = O_p(1) = o_p(\theta_T)$ since $\theta_T = s^{-1}(\lambda_T) \to \infty$ as $T \to \infty$. If X is nearly stationary, $X_T = o_p(\theta_T)$ as shown in Lemma A4 (a), which completes the proof for the part (a).

For the part (b), note that $[m_s]$ is slowly varying by Karamata's theorem, and

$$\overline{X}_T = \frac{1}{T}\int_0^T X_t dt = \begin{cases} O_p(\lambda_T/T), & \text{if } \iota \text{ is } m\text{-integrable} \\ O_p(\lambda_T[m_s|\iota_s|](\lambda_T)/T), & \text{if } \iota \text{ is } m\text{-nearly integrable} \\ O_p(\lambda_T^2(m_s\iota_s)(\lambda_T)/T), & \text{if } \iota \text{ is } m\text{-strongly nonintegrable} \end{cases} \tag{A.2}$$

by Lemma A3. If ι is m-integrable, then $\overline{X}_T = o_p(\theta_T)$ since $\lambda_T/T = O(1)$ and $\theta_T \to \infty$.

If ι is m-nearly integrable, then $|\iota_s|$ is m_s-nearly integrable, and hence, $[m_s|\iota_s|]$ becomes slowly varying due to Karamata's theorem. It follows that $[m_s|\iota_s|](\lambda_T)/[m_s](\lambda_T) \leq \ell(\lambda_T)$ for some slowly varying ℓ, and

$$\frac{\lambda_T[m_s|\iota_s|](\lambda_T)}{\theta_T T} = \frac{[m_s|\iota_s|](\lambda_T)}{s^{-1}(\lambda_T)[m_s](\lambda_T)} \leq \frac{\ell(\lambda_T)}{s^{-1}(\lambda_T)} \to 0 \tag{A.3}$$

since $T = \lambda_T[m_s](\lambda_T)$ and $s^{-1} \in RV_\kappa$ with $0 < \kappa < \infty$ at $+\infty$, as shown in Lemma A1 (a).

Finally, let ι be m-strongly nonintegrable. In this case, we have

$$\frac{\lambda_T^2(m_s\iota_s)(\lambda_T)}{\theta_T T} = \frac{\lambda_T(m_s\iota_s)(\lambda_T)}{s^{-1}(\lambda_T)[m_s](\lambda_T)} = \frac{\lambda_T m_s(\lambda_T)}{[m_s](\lambda_T)} \to 0, \tag{A.4}$$

where, in particular, the last convergence follows from Karamata's theorem since m_s is either integrable or nearly integrable. The part (b) follows from (A.2)-(A.4). $\square$

Lemma A6. *If X is defined on $\mathcal{D} = (0, \infty)$, then $(\iota/s')(x) \to 0$ and $1/s'(x) \to 0$ as $x \to 0$.*

Proof of Lemma A6. Since X is recurrent on $(0, \infty)$, $s(0) = \int_\omega^0 s'(x)dx = -\infty$ for any $\omega > 0$, and hence, $s' \in RV_p$ with $p < -1$ at $x_B = 0$. This completes the proof. $\square$

Lemma A7. *Let DD hold. (a) $1/s'(x) \to 0$ as $x \to \underline{x}, \overline{x}$, (b) $\iota/s' \circ s^{-1}(\lambda) \prec [m_s\sigma_s^2](\lambda)$ and $\iota/s' \circ s^{-1}(-\lambda) \prec [m_s\sigma_s^2](\lambda)$, (c) $[m_s\sigma_s^2](\lambda) \sim -2[m_s\iota_s\mu_s](\lambda)$, and (d) $\theta_T^2 = o(\lambda_T[m_s\sigma_s^2](\lambda_T))$.*

Proofs of Lemma A7 (a) and (b). Due to Lemma A6, it suffice to prove the statements at $x_B = \pm\infty$. Note that DD holds if and only if $1/s'$ is either integrable or nearly integrable since $m\sigma^2 = 1/s'$. It follows that $1/s'(x) \to 0$ as $x \to \pm\infty$ which completes the proof of the part (a).

For the part (b), we write

$$[m_s\sigma_s^2](\lambda) = \int_{-\lambda}^{\lambda} \frac{1}{(s')^2 \circ s^{-1}(x)} dx = \int_{s^{-1}(-\lambda)}^{s^{-1}(\lambda)} \frac{1}{s'(x)} dx. \tag{A.5}$$

By Karamata's theorem, we have

$$\frac{\iota/s' \circ s^{-1}(\lambda)}{[m_s\sigma_s^2](\lambda)},\ \frac{\iota/s' \circ s^{-1}(-\lambda)}{[m_s\sigma_s^2](\lambda)} \to 0$$

which completes the proof of the part (b). □

Proof of Lemma A7 (c). We have

$$\begin{aligned}-2[m_s\iota_s\mu_s](\lambda) &= -2\int_{-\lambda}^{\lambda} m_s\iota_s\mu_s(x)dx = -2\int_{s^{-1}(-\lambda)}^{s^{-1}(\lambda)} m\iota\mu(x)dx = \int_{s^{-1}(-\lambda)}^{s^{-1}(\lambda)} \frac{xs''(x)}{[s'(x)]^2}dx \\ &= \int_{s^{-1}(-\lambda)}^{s^{-1}(\lambda)} \frac{1}{s'(x)}dx - \left[\frac{\iota}{s'} \circ s^{-1}(\lambda) - \frac{\iota}{s'} \circ s^{-1}(-\lambda)\right] \end{aligned} \tag{A.6}$$

due to the integration by parts, from which, jointly with Lemma A7 (b) and (A.5), we have the stated result. □

Proof of Lemma A7 (d). Under DD, $s' \in RV_p$ with $-1 < p \le \infty$ at $+\infty$. If $s' \in RV_\infty$, then $s^{-1} \in RV_0$, and hence,

$$\frac{\theta_T}{\sqrt{\lambda_T[m_s\sigma_s^2](\lambda_T)}} = \frac{s^{-1}(\lambda_T)}{\sqrt{\lambda_T[m_s\sigma_s^2](\lambda_T)}} \to 0.$$

If $-1 < p < \infty$, then $\theta_T s'(\theta_T) \sim (p+1)s(\theta_T)$ by Karamata's theorem, and hence,

$$\frac{\theta_T}{\sqrt{\lambda_T[m_s\sigma_s^2](\lambda_T)}} \le \frac{\theta_T}{\sqrt{s(\theta_T)\int_0^{\theta_T} 1/s'(x)dx}} \sim \left(\frac{(p+1)\theta_T}{s'(\theta_T)\int_0^{\theta_T} 1/s'(x)dx}\right)^{1/2} \to 0 \tag{A.7}$$

since $\int_0^{\theta_T} 1/s'(x)dx \le [m_s\sigma_s^2](\lambda_T)$ by (A.5), and $\int_0^{\theta_T} 1/s'(x)dx$ is slowly varying and satisfies $(\theta_T/s'(\theta_T))/\int_0^{\theta_T} 1/s'(x)dx \to 0$ by Karamata's theorem. This completes the proof. □

Lemma A8. *If DD holds, then*

$$\int_0^T X_t\sigma(X_t)dW_t \prec_p \int_0^T X_t\mu(X_t)dt.$$

Proof of Lemma A8. For a brevity of notations, we write I, NI and SN instead of integrable, nearly integrable and strongly nonintegrable, respectively. It then follows from Lemma A3 that

$$\int_0^T X_t\sigma(X_t)dW_t = \begin{cases} O_p\big(\sqrt{\lambda_T[m_s\iota_s^2\sigma_s^2](\lambda_T)}\big), & \text{if } \iota^2\sigma^2 \text{ is } m\text{-I or } m\text{-NI} \\ O_p\big(\lambda_T(m_s^{1/2}\iota_s\sigma_s)(\lambda_T)\big), & \text{if } \iota^2\sigma^2 \text{ is } m\text{-SN.} \end{cases} \tag{A.8}$$

and

$$\frac{1}{\lambda_T[m_s\sigma_s^2](\lambda_T)}\int_0^T X_t\mu(X_t)dt \to_d -\frac{1}{2}L(\tau,0). \tag{A.9}$$

due, in particular, to Lemma A7 (c).

If $\iota^2\sigma^2$ is either m-I or m-NI, then $[m_s\iota_s^2\sigma_s^2]$ is slowly varying. In this case, the stated result follows from (A.8) and (A.9) since $\sqrt{\lambda_T[m_s\iota_s^2\sigma_s^2](\lambda_T)}/\lambda_T[m_s\sigma_s^2](\lambda_T) \to 0$.

Finally, let $\iota^2\sigma^2$ be m-SN. Then

$$\frac{\lambda_T(m_s^{1/2}\iota_s\sigma_s)(\lambda_T)}{\lambda_T[m_s\sigma_s^2](\lambda_T)} = \frac{(\iota/s')(\theta_T)}{[m_s\sigma_s^2](\lambda_T)} \leq \frac{(\iota/s')(\theta_T)}{\int_0^{\theta_T} 1/s'(x)dx} \to 0, \tag{A.10}$$

where the first equality is due to $m_s^{1/2}\iota_s\sigma_s = (\iota/s')\circ s^{-1}$, the second inequality follows from $\int_0^{\theta_T} 1/s'(x)dx \leq [m_s\sigma_s^2](\lambda_T)$ as in (A.7), and the convergence to zero holds by Karamata's theorem with the fact that $\int_0^{\theta_T} 1/s'(x)dx$ is slowly varying. The stated result for m-SN $\iota^2\sigma^2$ follows immediately from (A.8), (A.9) and (A.10), which completes the proof. □

Lemma A9. *If DD does not hold, then*

$$\int_0^T X_t\mu(X_t)dt \precsim_p \int_0^T X_t\sigma(X_t)dW_t.$$

Proof of Lemma A9. If DD does not hold, then $\iota^2\sigma^2$ is m-strongly nonintegrable and

$$\frac{1}{\lambda_T(m_s^{1/2}\iota_s\sigma_s)(\lambda_T)}\int_0^T X_t\sigma(X_t)dW_t \to_d \int_0^\tau \overline{m_s\iota_s\sigma_s}(B_t)dB_t \tag{A.11}$$

by Lemma A3. Following the notation in the proof of Lemma A8, we have

$$\int_0^T X_t\mu(X_t)dt = \begin{cases} O_p\big(\lambda_T[m_s\iota_s\mu_s](\lambda_T)\big), & \text{if } \iota\mu \text{ is } m\text{-I or } m\text{-NI} \\ O_p\big(\lambda_T^2(m_s\iota_s\mu_s)(\lambda_T)\big), & \text{if } \iota\mu \text{ is } m\text{-SN.} \end{cases} \tag{A.12}$$

We first let $\iota\mu$ be either m-I or m-NI. It then follows from (A.6) and $m_s^{1/2}\iota_s\sigma_s = (\iota/s')\circ$

s^{-1} that

$$\frac{\lambda_T[m_s\iota_s\mu_s](\lambda_T)}{\lambda_T(m_s^{1/2}\iota_s\sigma_s)(\lambda_T)} = -\frac{1}{2}\frac{\int_{-\theta_T}^{\theta_T} 1/s'(x)dx}{(\iota/s')(\theta_T)} + O(1) = O(1) \tag{A.13}$$

by Karamata's theorem with the fact that $1/s'$ is strongly nonintegrable.

Finally, let $\iota\mu$ be m-SN. Note that $s' \in RV_p$ at $x_B = +\infty$ for $-\infty \leq p \leq \infty$ if and only if $x\mu(x)/\sigma^2(x) \to -p/2$ as $x \to +\infty$ by the representation of regularly and rapidly varying functions. Moreover, if DD does not hold, $-1 < p < 1$ under Assumption 2.2. Therefore,

$$\frac{\lambda_T^2(m_s\iota_s\mu_s)(\lambda_T)}{\lambda_T(m_s^{1/2}\iota_s\sigma_s)(\lambda_T)} = \frac{\lambda_T(m_s^{1/2}\mu_s)(\lambda_T)}{\sigma_s(\lambda_T)} = \frac{s(\theta_T)\mu(\theta_T)}{s'(\theta_T)\sigma^2(\theta_T)} \to \frac{-p}{2+2p} = O(1) \tag{A.14}$$

since $s(\theta_T)/(\theta_T s'(\theta_T)) \to 1/(1+p)$ by Karamata's theorem. The stated result follows from (A.11)-(A.14). $\square$

Lemma A10. *If ST holds, then*

$$\int_0^T (X_t - \overline{X}_T)dX_t \sim_p -\frac{1}{2}\int_0^T \sigma^2(X_t)dt.$$

Proof of Lemma A10. We write

$$\int_0^T (X_t - \overline{X}_T)dX_t = \frac{1}{2}\left[(X_T^2 - X_0^2) - 2\overline{X}_T(X_T - X_0) - \int_0^T \sigma^2(X_t)dt\right]. \tag{A.15}$$

By applying Lemma A3 to the last term in (A.15), we have

$$\frac{1}{a_T}\int_0^T \sigma^2(X_t)dt \to_d P, \tag{A.16}$$

where a_T and P are defined in Section 3.2. Moreover,

$$(X_T^2 - X_0^2) - 2\overline{X}_T(X_T - X_0) = o_p(\theta_T^2) \tag{A.17}$$

due to Lemma A5. To complete the proof, it suffice to show $\theta_T^2 = O(a_T)$.

If DD holds, $a_T = \lambda_T[m_s\sigma_s^2](\lambda_T)$ and $\theta_T^2 = o\left(\lambda_T[m_s\sigma_s^2](\lambda_T)\right)$ as shown in Lemma A7 (d). On the other hand, if DD does not hold, then $a_T = \lambda_T^2(m_s\sigma_s^2)(\lambda_T)$ and $\theta_T^2 =$

$O(\lambda_T^2(m_s\sigma_s^2)(\lambda_T))$ because

$$\frac{\theta_T^2}{\lambda_T^2(m_s\sigma_s^2)(\lambda_T)} = \left(\frac{\theta_T s'(\theta_T)}{s(\theta_T)}\right)^2 \to (1+p)^2 = O(1) \qquad \text{(A.18)}$$

since $s(\lambda) \sim \lambda s'(\lambda)/(1+p)$ and $p \neq \pm\infty$ as discussed in the proof of Lemma A1 (a). In all cases, we have $\theta_T^2 = O(a_T)$, and hence, the stated result follows from (A.15)-(A.17). □

Lemma A11. *If DD holds, then*

$$\int_0^T (X_t - \overline{X}_T)dX_t \sim_p -\frac{1}{2}\int_0^T \sigma^2(X_t)dt.$$

Proof of Lemma A11. Due to Lemma A10, it suffice to prove the lemma for a strongly nonstationary X. If X is strongly nonstationary, then

$$\overline{X}_T = O_p(\lambda_T^2(m_s\iota_s)(\lambda_T)/T) = O_p(\theta_T)$$

by Lemma A3, from which, together with Lemma A4 (a), we have

$$(X_T^2 - X_0^2) - 2\overline{X}_T(X_T - X_0) = O_p(\theta_T^2). \qquad \text{(A.19)}$$

However, if DD holds, $\theta_T^2 = o\left(\lambda_T[m_s\sigma_s^2](\lambda_T)\right)$ as shown in Lemma A7 (d). Therefore, the stated result follows from (A.15), (A.16) and (A.19). □

Lemma A12. *We have*

$$\sum_{i=1}^n \left(\Delta x_i - \overline{\Delta x}_n\right)^2 = \sum_{i=1}^n (\Delta x_i)^2 + O_p(\delta T(\iota^2)/T),$$

where $\overline{\Delta x}_n = n^{-1}\sum_{i=1}^n \Delta x_i$. *Moreover, if* $\delta \to 0$, *then*

$$\begin{aligned}\sum_{i=1}^n (\Delta x_i)^2 = &\int_0^T \sigma^2(X_t)dt + O_p(\delta T(\mu^2)T) + O_p(\delta T(\mu\sigma)T^{1/2}) \\ &+ O_p\left(\delta^{1/2}T(\mu\sigma)T\sqrt{\log(T/\delta)}\right) + O_p\left(\delta^{1/2}T(\sigma^2)T^{1/2}\sqrt{\log(T/\delta)}\right).\end{aligned}$$

Proof of Lemma A12. We have

$$\sum_{i=1}^n \left(\Delta x_i - \overline{\Delta x}_n\right)^2 = \sum_{i=1}^n (\Delta x_i)^2 - n(\overline{\Delta x}_n)^2. = \sum_{i=1}^n (\Delta x_i)^2 + O_p(\delta T(\iota^2)/T) \qquad \text{(A.20)}$$

since $n(\overline{\Delta x}_n)^2 = n\left((X_T - X_0)/n\right)^2 = O_p(\delta T(\iota^2)/T)$. The leading term of (A.20) satisfies

$$\begin{aligned}\sum_{i=1}^n (x_i - x_{i-1})^2 - [X]_T &= \sum_{i=1}^n \Big(\big(X_{i\delta} - X_{(i-1)\delta}\big)^2 - \big([X]_{i\delta} - [X]_{(i-1)\delta}\big)\Big) \\ &= 2\sum_{i=1}^n \int_{(i-1)\delta}^{i\delta} \big(X_t - X_{(i-1)\delta}\big)\, dX_t \\ &= 2\left(P_T + Q_T + R_T + S_T\right)\end{aligned}$$

by Itô's formula, where

$$\begin{aligned}P_T &= \sum_{i=1}^n \int_{(i-1)\delta}^{i\delta} \left(\int_{(i-1)\delta}^t \mu(X_s)ds\right)\mu(X_t)dt = O_p(\delta T(\mu^2)T), \\ Q_T &= \sum_{i=1}^n \int_{(i-1)\delta}^{i\delta} \left(\int_{(i-1)\delta}^t \mu(X_s)ds\right)\sigma(X_t)dW_t = O_p(\delta T(\mu\sigma)T^{1/2}), \\ R_T &= \sum_{i=1}^n \int_{(i-1)\delta}^{i\delta} \left(\int_{(i-1)\delta}^t \sigma(X_s)dW_s\right)\mu(X_t)dt = O_p\left(\delta^{1/2}T(\mu\sigma)T\sqrt{\log(T/\delta)}\right), \\ S_T &= \sum_{i=1}^n \int_{(i-1)\delta}^{i\delta} \left(\int_{(i-1)\delta}^t \sigma(X_s)dW_s\right)\sigma(X_t)dW_t = O_p\left(\delta^{1/2}T(\sigma^2)T^{1/2}\sqrt{\log(T/\delta)}\right)\end{aligned}$$

due to Lemmas B4 and B5 of Kim and Park (2017). This completes the proof. □

A.2 Lemmas for Subsample Bootstrap

In what follows, we assume that Assumption 2.1 holds. The proofs of subsequential lemmas rely on the representation of a scale transformed diffusion $Y = s(X)$ as a time-changed Brownian motion, i.e., $Y = B \circ A$, where B is Brownian motion and A is time change defined as $A_t = \inf\left\{s \middle| \int_{\mathbb{R}} m_s(x)L(s,x)dx > t\right\}$ from the local time L of B. For the details of the representation, the reader is referred to, e.g., Rogers and Williams (2000).

In the sequel, for $f = s', s, s^{-1}, m, m_s, A$, we simply let $\hat{f}$ be defined similarly as f with μ and σ^2 replaced by μ_T and σ_T^2, respectively.

Lemma A13. *If $\delta^a\lambda_T^{-b}K_S S(\omega) \to_p 0$, then*

$$\sup_{x\in\mathcal{D}_S} \left|(1/\sigma_T^2)(x) - (1/\sigma^2)(x)\right| = O_p\left(\delta^a\lambda_T^{-b}K_S S(\omega^2)\right).$$

Proof of Lemma A13. We have

$$
\begin{aligned}
(1/\sigma_T^2)(x) - (1/\sigma^2)(x) &= \frac{(\sigma^2 - \sigma_T^2)(x)}{\sigma^4(x)\left(1 - ((\sigma^2 - \sigma_T^2)/\sigma^2)(x)\right)} \\
&= \frac{(\sigma^2 - \sigma_T^2)(x)}{\sigma^4(x)\left(1 + O_p\left(\delta^a \lambda_T^{-b} K_S S(1/\sigma^2)\right)\right)} = O_p\left(\delta^a \lambda_T^{-b} K_S S(1/\sigma^4)\right)
\end{aligned}
$$

uniformly in $x \in \mathcal{D}_S$, since $\sup_{x\in\mathcal{D}_S} \left|(\sigma^2 - \sigma_T^2)(x)\right| = O_p\left(\delta^a \lambda_T^{-b} K_S\right)$ by (12). □

Lemma A14. *If (a) $\delta^a \lambda_T^{-b} K_S S(\omega) \to_p 0$ and (b) $\delta^a S(\omega^2) \to_p 0$, then*

$$
\sup_{x\in\mathcal{D}_S} \left| \frac{\mu}{\sigma^2}(x) - \frac{\mu_T}{\sigma_T^2}(x) \right| = O_p\left(\lambda_T^{-b} K_S S(\omega)\right).
$$

Proof of Lemma A14. We have

$$
\frac{\mu_T}{\sigma_T^2}(x) - \frac{\mu}{\sigma^2}(x) = \frac{\mu_T - \mu}{\sigma_T^2}(x) + \mu(x)\left(\frac{1}{\sigma_T^2}(x) - \frac{1}{\sigma^2}(x)\right) = O_p\left(\lambda_T^{-b} K_S S(\omega)\right)
$$

since

$$
\frac{\mu_T - \mu}{\sigma_T^2}(x) = \frac{(\mu_T - \mu)(x)}{\sigma^2(x)\left(1 + O_p\left(\delta^a \lambda_T^{-b} K_S S(1/\sigma^2)\right)\right)} = O_p\left(\lambda_T^{-b} K_S S(1/\sigma^2)\right)
$$

due to (12) with the condition (a) in this lemma, and

$$
\sup_{x\in\mathcal{D}_S} \left| \mu(x)\left(\frac{1}{\sigma_T^2}(x) - \frac{1}{\sigma^2}(x)\right) \right| = O_p\left(\delta^a \lambda_T^{-b} K_S S(\omega^2) S(\mu)\right) = O_p\left(\lambda_T^{-b} K_S S(\omega)\right)
$$

due to Lemma A13 with the condition (b) in this lemma. □

Lemma A15. *Let the conditions in Lemma A14 hold. If $\lambda_T^{-b} K_S S(\omega^2) \to_p 0$, then*

$$
\begin{aligned}
\sup_{x\in\mathcal{D}_S} \left| s'^k(x) - \hat{s}'^k(x) \right| &= O_p\left(\lambda_T^{-b} K_S S(\omega^3)\right) \quad \textit{for} \quad k = 1, 2, \\
\sup_{x\in\mathcal{D}_S} \left| s(x) - \hat{s}(x) \right| &= O_p\left(\lambda_T^{-b} K_S S(\omega^4)\right).
\end{aligned}
$$

Moreover, if $\lambda_T^{-b} K_S S(\omega^4) \to_p 0$, then

$$
\sup_{x\in\mathcal{R}_S} \left| s^{-1}(x) - \hat{s}^{-1}(x) \right| = O_p(\lambda_T^{-b} K_S S(\omega^5)),
$$

where $\mathcal{R}_S = \{y|y = s(x), x \in \mathcal{D}_S\} \cap \{y|y = \hat{s}(x), x \in \mathcal{D}_S\}$.

Proof of Lemma A15. It follows from Lemma A14 that

$$\sup_{x\in\mathcal{D}_S}\left|\int_w^x \frac{\mu_T}{\sigma_T^2}(y)dy - \int_w^x \frac{\mu}{\sigma^2}(y)dy\right| = O_p\left(\lambda_T^{-b}K_S S(\omega)S(\iota)\right),$$

and therefore,

$$\begin{aligned}\hat{s}'(x) &= s'(x)\exp\left(-2\left(\int_w^x \frac{\mu_T}{\sigma_T^2}(y)dy - \int_w^x \frac{\mu}{\sigma^2}(y)dy\right)\right)\\ &= s'(x)\left(1 + O_p\left(\lambda_T^{-b}K_S S(\omega)S(\iota)\right)\right) \qquad (\text{A.21})\end{aligned}$$

uniformly in $x \in \mathcal{D}_S$. The stated result for s' follows immediately from (A.21). Similarly, we can show the result for s'^2.

The result for s can be deduced from (A.21), and we have

$$\sup_{x\in\mathcal{D}_S}|s(x) - \hat{s}(x)| = \sup_{x\in\mathcal{D}_S}\left|\int_w^x \left(s'(y) - \hat{s}'(y)\right)dy\right| = O_p\left(\lambda_T^{-b}K_S S(\omega^3)S(\iota)\right)$$

as desired.

As for s^{-1}, we let $x = \hat{s}^{-1}(y)$ for $y \in \mathcal{R}_S$. We then have

$$s^{-1}(y) - \hat{s}^{-1}(y) = s^{-1}(\hat{s}(x)) - x = s^{-1}(\hat{s}(x)) - s^{-1}(s(x)) \qquad (\text{A.22})$$

and

$$\begin{aligned}\sup_{x\in\mathcal{D}_S}|s^{-1}(\hat{s}(x)) - s^{-1}(s(x))| &\le \sup_{y\in\mathcal{R}_S}|(s^{-1})'(y)|\sup_{x\in\mathcal{D}_S}|\hat{s}(x) - s(x)|\\ &= O_p\left(\lambda_T^{-b}K_S S(\omega^4)S(1/s')\right) \qquad (\text{A.23})\end{aligned}$$

since s^{-1} is continuously differentiable with $(s^{-1})' = (1/s')\circ s^{-1}$. The stated result for s^{-1} follows immediately from (A.22) and (A.23). $\square$

Lemma A16. *Let the conditions in Lemma A15 hold. If* $\lambda_T^{-b}K_S S(\omega^6) \to_p 0$, *then*

$$\begin{aligned}\sup_{x\in\mathcal{D}_S}|(\hat{s}'\sigma_T)^2(x) - (s'\sigma)^2(x)| &= O_p\left(\lambda_T^{-b}K_S S(\omega^5)\right),\\ \sup_{x\in\mathcal{D}_S}\left|\frac{1}{(\hat{s}'\sigma_T)^2}(x) - \frac{1}{(s'\sigma)^2}(x)\right| &= O_p\left(\lambda_T^{-b}K_S S(\omega^7)\right).\end{aligned}$$

Proof of Lemma A16. It follows from Lemma A15 that

$$\begin{aligned}
\sup_{x\in\mathcal{D}_S} |(\hat{s}'\sigma_T)^2(x) - (s'\sigma)^2(x)| &= \sup_{x\in\mathcal{D}_S} |(\hat{s}'\sigma_T)^2(x) - (s'\sigma_T)^2(x)| + \sup_{x\in\mathcal{D}_S} |(s'\sigma_T)^2(x) - (s'\sigma)^2(x)| \\
&= O_p\left(\lambda_T^{-b} K_S S(\omega^3) S(\sigma^2)\right) + O_p\left(\delta^a \lambda_T^{-b} K_S S(s'^2)\right) \\
&= O_p\left(\lambda_T^{-b} K_S S(\omega^5)\right). \qquad \text{(A.24)}
\end{aligned}$$

Moreover, we can deduce from the proof of Lemma A13 with (A.24) that

$$\sup_{x\in\mathcal{D}_S} \left| \frac{1}{(\hat{s}'\sigma_T)^2}(x) - \frac{1}{(s'\sigma)^2}(x) \right| = O_p\left(\lambda_T^{-b} K_S S(\omega^5) S((m/s')^2)\right)$$

as desired. □

Lemma A17. *For any $\varepsilon > 0$, we have*

$$\sup_{0\leq s,t\leq S, |t-s|\leq\varepsilon} |A_t - A_s| = O_p(\varepsilon S(\omega)).$$

Proof of Lemma A17. Since $m_s(x) > 0$ for all $x \in \mathbb{R}$, A is invertible and

$$A^{-1}(t) = \int m_s(x) L(t,x) dx = \int_0^t m_s(B_u) du,$$

where the last equality follows from the occupation times formula. Moreover, A^{-1} is continuously differentiable and we have $(A^{-1})'(t) = m_s(B_t)$, and hence,

$$A'(t) = \frac{1}{(A^{-1})'(A(t))} = \frac{1}{m_s(B\circ A(t))} = (s'/m)(X_t).$$

Therefore, we have

$$\sup_{0\leq s,t\leq S, |t-s|\leq\varepsilon} |A_t - A_s| \leq \varepsilon \sup_{0\leq t\leq S} (s'/m)(X_t) = O_p(\varepsilon S(\omega))$$

as desired. □

Lemma A18. *Let the conditions in Lemma A16 hold. Then we have*

$$\sup_{0\leq t\leq S} |A_t - \hat{A}_t| = O_p\left(\lambda_T^{-b} K_S S(\omega^7) S\right).$$

Proof of Lemma A18. We have $A_t - \hat{A}_t = A\circ\hat{A}^{-1}(\hat{A}_t) - A\circ A^{-1}(\hat{A}_t)$, where $\hat{A}^{-1}(\hat{A}_t) = t$

and

$$A^{-1}(\hat{A}_t) = \int_0^{\hat{A}_t} m_s(B_u)du = \int_0^t (m_s/\hat{m}_s)(B_u)du$$

by the occupation times formula and change of variables. Moreover, we have

$$\begin{aligned}\sup_{x\in\mathcal{R}_S} |(m_s/\hat{m}_s)(x) - 1| &\leq \left(\sup_{x\in\mathcal{R}_S} m_s(x)\right)\left(\sup_{x\in\mathcal{R}_S} |(1/\hat{m}_s)(x) - (1/m_s)(x)|\right) \\ &\leq \left(\sup_{x\in\mathcal{D}_S} (m/s')(x)\right)\left(\sup_{x\in\mathcal{D}_S} |(\hat{s}'\sigma_T)^2(x) - (s'\sigma)^2(x)|\right) \\ &= O_p\left(\lambda_T^{-b} K_S S(\omega^5) S(m/s')\right)\end{aligned}$$

due to Lemma A16, and hence,

$$\sup_{0\leq t\leq S} |A^{-1}(\hat{A}_t) - \hat{A}^{-1}(\hat{A}_t)| = O_p\left(\lambda_T^{-b} K_S S(\omega^6) S\right)$$

from which, jointly with Lemma A17, we have the stated result. □

Appendix B Proofs for Main Results

Proof of Lemma 2.1. The stated result follows immediately from Lemmas A8 and A9. □

Proof of Lemma 2.2. It follows from the integration by parts with $s''(x) = -2(s'\mu/\sigma^2)(x)$ that for any $x_l, x_u \in \mathcal{D}$, we have

$$-2\int_{x_l}^{x_u} (m\mu\nu)(x)dx = \int_{x_l}^{x_u} (m\sigma^2\nu')(x)dx - \left[(\nu/s')(x_u) - (\nu/s')(x_l)\right],$$

from which we can obtain the stated result by letting $x_u \to \overline{x}$ and $x_l \to \underline{x}$. □

Proof of Lemma 3.1. We note that

$$\sum_{i=1}^n x_{i-1}\delta = \int_0^T X_t dt + O_p(\delta T(\mu)T) + O_p(\delta T(\sigma)T^{1/2}), \tag{B.1}$$

$$\sum_{i=1}^n x_{i-1}^2\delta = \int_0^T X_t^2 dt + O_p(\delta T(\iota\mu)T) + O_p(\delta T(\sigma^2)T) + O_p(\delta T(\iota\sigma)T^{1/2}) \tag{B.2}$$

due to Lemma B1 of Kim and Park (2017). Moreover, we may deduce from Lemma B3 of

Kim and Park (2017) that

$$\begin{aligned}\sum_{i=1}^{n} x_{i-1}\Delta x_i &= \sum_{i=1}^{n} X_{(i-1)\delta}\int_{(i-1)\delta}^{i\delta}\mu(X_t)dt + \sum_{i=1}^{n} X_{(i-1)\delta}\int_{(i-1)\delta}^{i\delta}\sigma(X_t)dW_t \\ &= \int_0^T X_t\mu(X_t)dt + \int_0^T X_t\sigma(X_t)dW_t + R_T + S_T \\ &= \int_0^T X_t dX_t + R_T + S_T, \end{aligned} \tag{B.3}$$

where

$$\begin{aligned} R_T &= O_p(\delta T(\mu^2)T) + O_p(\delta T(\mu\sigma)T^{1/2}), \\ S_T &= O_p(\delta T(\mu\sigma)T^{1/2}) + O_p\left(\delta^{1/2}T(\sigma^2)T^{1/2}\sqrt{\log(T/\delta)}\right). \end{aligned}$$

The stated result for $\hat{\beta}$ follows immediately from (B.1)-(B.3) given Assumption 3.1.

As for $t(\hat{\beta})$, we write

$$\begin{aligned} \hat{v}^2/\,\delta &= \frac{1}{n\delta}\sum_{i=1}^{n}\left[\Delta x_i - (\hat{\alpha} - \hat{\beta}x_{i-1})\delta\right]^2 \\ &= \frac{1}{n\delta}\sum_{i=1}^{n}\left(\Delta x_i - \overline{\Delta x}_n\right)^2 - \frac{1}{n\delta}\frac{\left[\sum_{i=1}^{n}\left(x_{i-1} - \overline{x}_n\right)\Delta x_i\right]^2}{\sum_{i=1}^{n}\left(x_{i-1} - \overline{x}_n\right)^2}. \end{aligned} \tag{B.4}$$

For the second term in (B.4), we may deduce from (B.1)-(B.3) with Assumption 3.1 that

$$\frac{1}{n\delta}\frac{\left[\sum_{i=1}^{n}\left(x_{i-1} - \overline{x}_n\right)\Delta x_i\right]^2}{\sum_{i=1}^{n}\left(x_{i-1} - \overline{x}_n\right)^2} = \frac{\delta}{T}\frac{\left(\int_0^T (X_t - \overline{X}_T)dX_t\right)^2}{\int_0^T (X_t - \overline{X}_T)^2 dt}(1 + o_p(1)), \tag{B.5}$$

where

$$\begin{aligned} \int_0^T (X_t - \overline{X}_T)dX_t &= \frac{1}{2}\left((X_T - \overline{X}_T)^2 - (X_0 - \overline{X}_T)^2 - \int_0^T \sigma^2(X_t)dt\right) \\ &= O_p(T(\iota^2)) + O_p(T(\sigma^2)T). \end{aligned}$$

The leading term of (B.5) satisfies

$$\frac{\delta}{T}\frac{\left(\int_0^T (X_t - \overline{X}_T)dX_t\right)^2}{\int_0^T (X_t - \overline{X}_T)^2 dt} = O_p\left(\delta T(\omega^4)T\right) = o_p(1), \tag{B.6}$$

where, in particular, the last equality is due to Assumption 3.1. It then follows from (B.4)-(B.6) with Lemma A12 that

$$\hat{v}^2/\delta = \frac{1}{T}\sum_{i=1}^{n}\left[\Delta x_i - (\hat{\alpha} - \hat{\beta}x_{i-1})\delta\right]^2 = \frac{1}{T}\int_0^T \sigma^2(X_t)dt(1+o_p(1)) \tag{B.7}$$

due to Assumption 3.1. The stated result for $t(\hat{\beta})$ follows immediately from (B.7) with the proof of the first part of this lemma. □

Proof of Lemma 3.2. By Lemma A3 with Lemmas A10 and A11, we have the desired convergences for $[X]_T$, $\int_0^T (X_t - \overline{X}_T)dX_t$ and $\int_0^T (X_t - \overline{X}_T)dt$. To complete the proof, it suffice to show that $Ta_T/b_T \to \infty$.

If SI holds, then $b_T = \lambda_T \ell(\lambda_T)$, and hence, $Ta_T/b_T = Ta_T/(\lambda_T\ell(\lambda_T)) \to \infty$ since $\lambda_T = O(T)$ and $a_T/\ell(\lambda_T) \to \infty$.

Now let SI do not hold. If both DD and ST hold, then

$$\begin{aligned}\frac{Ta_T}{b_T} = \frac{[m_s](\lambda_T)[m_s\sigma_s^2](\lambda_T)}{(m_s\iota_s^2)(\lambda_T)} &= \frac{[m_s](\lambda_T)}{\lambda_T m_s(\lambda_T)}\,\frac{[m_s\sigma_s^2](\lambda_T)}{\lambda_T(m_s\sigma_s^2)(\lambda_T)}\,\frac{\lambda_T^2(m_s\sigma_s^2)(\lambda_T)}{\iota_s^2(\lambda_T)}\\ &\equiv A_T B_T C_T \to \infty\end{aligned} \tag{B.8}$$

where $A_T, B_T \to \infty$ and $C_T = (s(\theta_T)/\theta_T s'(\theta_T))^2 \to 1/(p+1)^2$ by Karamata's theorem, since $s' \in RV_p$ with $-1 < p < \infty$ at $+\infty$ as shown in the proof of Lemma A1 (a).

If DD holds and ST does not hold, then

$$\frac{Ta_T}{b_T} = \frac{\lambda_T[m_s\sigma_s^2](\lambda_T)}{\iota_s^2(\lambda_T)} = B_T C_T \to \infty$$

by the same argument in (B.8).

If DD does not hold, then ST should be satisfied. In this case, we have

$$\frac{Ta_T}{b_T} = \frac{\lambda_T[m_s](\lambda_T)\sigma_s^2(\lambda_T)}{\iota_s^2(\lambda_T)} = A_T C_T \to \infty$$

by the same argument in (B.8). In all cases, we have $Ta_T/b_T \to \infty$, which completes the proof. □

Proof of Lemma 3.3. We have

$$\frac{1}{\lambda_T^2(m_s\sigma_s^2)(\lambda_T)}[X]_T \to_d \int_0^\tau \overline{m_s\sigma_s^2}(B_t)dt$$

by Lemma A3, and

$$\frac{1}{\theta_T^2}(X_T^2 - X_0^2) - 2\overline{X}_T(X_T - X_0) \to_d (X_1^{\circ 2} - X_0^{\circ 2}) - 2\overline{X}_1^{\circ}(X_1^{\circ} - X_0^{\circ}),$$
$$\frac{1}{T\theta_T^2}\int_0^T (X_t - \overline{X}_T)dt \to_d \int_0^1 (X_t^{\circ} - \overline{X}_1^{\circ})dt$$

by Lemma A4 (a). Moreover, we have

$$\int_0^1 (X_t^{\circ} - \overline{X}_1^{\circ})dX_t^{\circ} = \frac{1}{2}\left[(X_1^{\circ 2} - X_0^{\circ 2}) - 2\overline{X}_1^{\circ}(X_1^{\circ} - X_0^{\circ}) - [X^{\circ}]_t\right],$$

where $[X^{\circ}]_t = \int_0^{\overline{A}_t}\left(\overline{(s^{-1})'_-}\right)^2 (B_s)ds$, by Itô's formula with Lemma A4 (b). Since s' inf RV_p with $p \neq -1, \pm\infty$ as discussed in (A.18), we have $\theta_T^2 \sim (1+p)^2\lambda_T^2(m_s\sigma_s^2)(\lambda_T)$ by Karamata's theorem. Therefore, the stated result follows immediately if we show

$$\left(\overline{(s^{-1})'_-}\right)^2 = \frac{1}{(1+p)^2}\overline{m_s\sigma_s^2}. \tag{B.9}$$

To show (B.9), note that $m_s\sigma_s^2 = \left((s^{-1})'\right)^2$ since $m_s\sigma_s^2 = (1/s' \circ s^{-1})^2$ and $(s^{-1})' = 1/s' \circ s^{-1}$. It then follows from the property of regularly varying function (see, e.g., Proposition 1.5.7. in Bingham et al. (1993)) and Assumption 2.2 that $(s^{-1})' \in RV_\kappa$ with $\kappa = -p/(p+1) > -1$, from which, jointly with Karamata's theorem, we can show that $\overline{(s^{-1})'} = (1+p)\overline{(s^{-1})'_-}$. Therefore, we have (B.9), which completes the proof. □

Proof of Theorem 3.4. The stated results follow from Lemmas 3.1, 3.2 and 3.3. □

Proofs of Corollaries 4.1 and 4.2. The stated results follow from Lemmas 3.2 and 3.3. □

Proof of Corollary 4.3. For the part (a), we let $\overline{\sigma^2}$ be a constant function on $\mathcal{D}$, or equivalently σ^2 be a slowly varying function on $\mathcal{D}$. If ST holds, then $m\ell$ is integrable for some slowly varying function ℓ. We define $\ell_{\sigma^2} = \ell/\sigma^2$ from the slowly varying function ℓ. It then can easy to see that ℓ_{σ^2} is slowly varying since both ℓ and σ^2 are slowly varying. Then $m\sigma^2\ell_{\sigma^2}$ is integrable, and hence, DD holds. The converse can be shown analogously.

As for the part (b), we let $\overline{s'}$ be a constant function on $\mathcal{D}$, or equivalently s' be a slowly varying function on $\mathcal{D}$. Then $1/s'$ is also slowly varying, and is neither integrable nor nearly integrable by Karamata's theorem. Therefore, DD does not hold in this case. □

Proof of Lemma 4.4. It is easy to see that $\overline{s'}$ is a constant function on $\mathbb{R}$ if and only if s' is slowly varying such that $s'(\lambda)/s'(-\lambda) \to 1$ as $\lambda \to \infty$. Due to the representation of regularly varying functions, s' is slowly varying if and only if $x\mu(x)/\sigma^2(x) \to 0$ as $x \to x_B = \pm\infty$.

Moreover, $s'(\lambda)/s'(-\lambda) \to 1$ as $\lambda \to \infty$ if and only if $\int_{|x|\leq\lambda} \mu(x)/\sigma^2(x)dx \to 0$ as $\lambda \to \infty$ since $s'(\lambda)/s'(-\lambda) = \exp\left(-2\int_{|x|\leq\lambda} \mu(x)/\sigma^2(x)dx\right)$. This completes the proof. □

Proof of Lemma 4.5. The part (a) follows immediately from Lamma A1 (b) with $f = 1$. For the part (b), we let $Y = r(X)$, and define $\mathcal{D}_r = (r(\underline{x}), r(\overline{x}))$. Then, by Itô's lemma, we have

$$dY_t = (\mu/\sigma - \sigma'/2) \circ r^{-1}(Y_t)dt + dW_t$$

and its scale density s'_r and speed measure m_r are given, respectively, as

$$s'_r = (s'\sigma) \circ r^{-1} \quad \text{and} \quad m_r = 1/s'_r.$$

By a change of variables in integrals, $m_r(\mathcal{D}_r) = m(\mathcal{D})$, and hence, Y is stationary on $\mathcal{D}_r$ if and only if X is stationary on $\mathcal{D}$. Moreover, it follows from Lemma A2 (b) that m_r is nearly integrable/strongly nonintegrable on $\mathcal{D}_r$ if and only if m is nearly integrable/strongly nonintegrable on $\mathcal{D}$. This completes the proof of the lemma. □

Proof of Proposition 4.6. In the following, we simply write $\hat{y}_i = \hat{r}(X_{i\delta})$ and $y_i = r(X_{i\delta})$ with $\hat{y}_0 = \hat{r}(X_0)$ and $y_0 = r(X_0)$. If $\varepsilon^a\lambda_T^{-b}K_T T(1/\sigma^3) \to_p 0$, we can deduce from (11) that

$$\begin{aligned}
\sup_{0\leq i\leq n} |\hat{y}_i - y_i| &= \sup_{x\in\mathcal{D}_T} \left| \int_w^x \left(\frac{1}{\hat{\sigma}(y)} - \frac{1}{\sigma(y)} \right) dy \right| \\
&\leq \sup_{x\in\mathcal{D}_T} \left| \left(\frac{\sigma^2 - \hat{\sigma}^2}{\hat{\sigma}\sigma(\sigma + \hat{\sigma})} \right)(x) \right| T(\iota) \\
&\leq \sup_{x\in\mathcal{D}_T} |\sigma^2(x) - \hat{\sigma}^2(x)| T(\iota)T(1/\sigma^3) = O_p\left(\varepsilon^a\lambda_T^{-b}K'_T\right),
\end{aligned}$$

where $K' = K_T T(\iota)T(1/\sigma^3)$. Moreover, if $\varepsilon^a\lambda_T^{-b}K'_T \to_p 0$, then

$$\sup_{0\leq i\leq n} |\hat{y}_i^2 - y_i^2| = O_p\left(\varepsilon^a\lambda_T^{-b}K_T T(r)\right).$$

Therefore, we have

$$\begin{aligned}
\sum_{i=1}^n \hat{y}_{i-1}\delta &= \sum_{i=1}^n y_{i-1}\delta + O_p\left(\varepsilon^a\lambda_T^{-b}K'_T T\right), \\
\sum_{i=1}^n \hat{y}_{i-1}^2\delta &= \sum_{i=1}^n y_{i-1}^2\delta + O_p\left(\varepsilon^a\lambda_T^{-b}K'_T T(r)T\right), \\
\hat{y}_{n-1}^2 - \hat{y}_0^2 - \overline{y}_n(\hat{y}_{n-1} - \hat{y}_0) &= y_{n-1}^2 - y_0^2 - \overline{y}_n(y_{n-1} - y_0) + O_p\left(\varepsilon^a\lambda_T^{-b}K'_T T(r)\right),
\end{aligned}$$

and

$$\sum_{i=1}^{n}(\hat{y}_i - \hat{y}_{i-1})^2 = \sum_{i=1}^{n}(y_i - y_{i-1})^2 + P_T + Q_T,$$

where

$$P_T = \sum_{i=1}^{n}((\hat{y}_i - \hat{y}_{i-1}) - (y_i - y_{i-1}))^2 = O_p\left(\varepsilon^{2a}\lambda_T^{-2b}K_T'^2 T/\delta\right),$$
$$Q_T = 2\sum_{i=1}^{n}(y_i - y_{i-1})((\hat{y}_i - \hat{y}_{i-1}) - (y_i - y_{i-1})) = O_p\left(\varepsilon^{a}\lambda_T^{-b}K_T' T(r)\right).$$

Now we let $\hat{\beta}_{\hat{r}}$ and $\hat{\beta}_r$ be the least square estimators of the first-order autoregressions using, respectively, $(\hat{y}_i)$ and (y_i). If $\varepsilon^a\lambda_T^{-b}K_T'T(r)T \to_p 0$ and $\varepsilon^a\lambda_T^{-b}K_T'(T/\delta)^{1/2} \to_p 0$, then $\hat{\beta}_{\hat{r}} \sim_p \hat{\beta}_r$ since

$$\sum_{i=1}^{n}(z_{i-1} - \overline{z}_n)\Delta z_i = \frac{1}{2}\left(z_{n-1}^2 - z_0^2 - \overline{z}_n(z_{n-1} - z_0) - \sum_{i=1}^{n}(z_i - z_{i-1})^2\right)$$

for $z = \hat{y}, y$. We can prove $t(\hat{\beta}_{\hat{r}}) \sim_p t(\hat{\beta}_r)$ analogosly. □

Proof of Lemma 5.1. We let Y and $\hat{Y}$ be the scale transformed processes of X and $\hat{X}$, respectively. Due to the representation of martingale diffusion, we can write

$$X_t = s^{-1}(Y_t) = s^{-1}(B \circ A_t), \quad \hat{X}_t = \hat{s}^{-1}(\hat{Y}_t) = \hat{s}^{-1}(B \circ \hat{A}_t).$$

It then follows from the global modulus of continuity for Brownian motion (see Kanaya et al. (2017)) with Lemma A18 that

$$\sup_{0\le t\le S}|\hat{Y}_t - Y_t| = \sup_{0\le t\le S}|B \circ \hat{A}_t - B \circ A_t| = O_p\left(\sqrt{\lambda_T^{-b}K_S S(\omega^7)S\log\left(\lambda_S^2\lambda_T^b/(K_S S(\omega^7)S)\right)}\right)$$

since $A_S = O_p(\lambda_S^2)$ which is shown in Proposition 3.2 in Kim and Park (2017). Moreover, we have

$$\begin{aligned}\sup_{0\le t\le S}|\hat{X}_t - X_t| &= \sup_{0\le t\le S}|\hat{s}^{-1}(\hat{Y}_t) - s^{-1}(Y_t)| \\ &\le \sup_{0\le t\le S}|\hat{s}^{-1}(\hat{Y}_t) - s^{-1}(\hat{Y}_t)| + \sup_{0\le t\le S}|s^{-1}(\hat{Y}_t) - s^{-1}(Y_t)|,\end{aligned}$$

where

$$\sup_{0\le t\le S} |\hat{s}^{-1}(\hat{Y}_t) - s^{-1}(\hat{Y}_t)| = O_p\left(\lambda_T^{-b} K_S S(\omega^5)\right)$$

by Lemma A15, and

$$\sup_{0\le t\le S} |s^{-1}(\hat{Y}_t) - s^{-1}(Y_t)| \le \sup_{x\in\mathcal{D}_S} |(1/s')(x)| \sup_{0\le t\le S} |\hat{Y}_t - Y_t| = O_p\left(\zeta(S,T)\right)$$

since $(s^{-1})' = (1/s') \circ s^{-1}$, from which, together with Assumption 5.1, we have the stated result. □

Proof of Lemma 5.2. Note that

$$\int_0^S (X_t - \overline{X}_S) dX_t = \frac{1}{2}\left((X_S - \overline{X}_S)^2 - (X_0 - \overline{X}_S)^2 - \int_0^S \sigma^2(X_t)dt\right),$$
$$\int_0^S (X_t - \overline{X}_S)^2 dt = \int_0^S X_t^2 dt - S\overline{X}_S^2.$$

It then follows from Lemma 5.1 that

$$\int_0^S \hat{X}_t dt = \int_0^S X_t dt + O_p\left(\zeta(S,T)S\right),$$
$$\int_0^S \hat{X}_t^2 dt = \int_0^S X_t^2 dt + O_p\left(\zeta(S,T)S(\iota)S\right),$$
$$\int_0^S \sigma_T^2(\hat{X}_t)dt = \int_0^S \sigma^2(X_t)dt + O_p\left(\delta^a \lambda_T^{-b} K_S S\right) + O_p\left(\zeta(S,T)S((\sigma^2)')S\right)$$

from which we have the stated result. □

Proof of Theorem 5.3. The stated result follows immediately from Lemma 5.2. □

References

Aït-Sahalia, Y. (1996a): "Nonparametric Pricing of Interest Rate Derivative Securities," *Econometrica*, 64, 527–560.

——— (1996b): "Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9, 385–426.

Aït-Sahalia, Y. and J. Y. Park (2016): "Bandwidth Selection and Asymptotic Properties of Local Nonparametric Estimators in Possibly Nonstationary Continuous-Time Models," *Journal of Econometrics*, 192, 119–138.

Baillie, R. T. and T. Bollerslev (1989): "Common Stochastic Trends in a System of Exchange Rates," *Journal of Finance*, 44, 167–181.

——— (1994): "Cointegration, Fractional Cointegration, and Exchange Rate Dynamics," *Journal of Finance*, 49, 737–745.

Bandi, F. M. and V. Corradi (2014): "Nonparametric Nonstationarity Tests," *Econometric Theory*, 30, 127–149.

Bandi, F. M. and P. C. B. Phillips (2003): "Fully nonparametric estimation of scalar diffusion models," *Econometrica*, 71, 241–283.

Baur, D. G. and D. T. Tran (2014): "The Long-Run Relationship of Gold and Silver and the Influence of Bubbles and Financial Crises," *Empirical Economics*, 47, 1525–1541.

Bingham, N. H., C. M. Goldie, and J. L. Teugels (1993): *Regular Variation*, Cambridge, UK: Cambridge University Press.

Bossaerts, P. (1988): "Common Nonstationary Components of Asset Prices," *Journal of Economic Dynamics and Control*, 12, 347–364.

Bossaerts, P. and R. Green (1989): "A General Equilibrium Model of Changing Risk Premia: Theory and Evidence," *Review of Financial Studies*, 2, 467–493.

Bu, R., J. Kim, and B. Wang (2017): "Uniform and L_p Convergences of Nonparametric Estimators for Diffusion Processes," Working Paper, Toulouse School of Economics.

Chambers, M. J. (2004): "Testing for Unit Roots with Flow Data and Varying Sampling Frequency," *Journal of Econometrics*, 119, 1–18.

——— (2008): "Corrigendum to: Testing for Unit Roots with Flow Data and Varying Sampling Frequency," *Journal of Econometrics*, 144, 524–525.

Chan, K. C., A. Karolyi, F. Longstaff, and A. Sanders (1992): "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 47, 1209–27.

Chen, X., L. P. Hansen, and M. Carrasco (2010): "Nonlinearity and Temporal Dependence," *Journal of Econometrics*, 155, 155–169.

Conley, T. G., L. P. Hansen, E. G. J. Luttmer, and J. A. Scheinkman (1997): "Short-Term Interest Rates as Subordinated Diffusions," *Review of Financial Studies*, 10, 525–77.

Davis, R. A. (1982): "Maximum and Minimum of One-Dimensional Diffusions," *Stochastic Processes and their Applications*, 13, 1–9.

Diebold, F. X., J. Gardeazabal, and K. Yilmaz (1994): "On Cointegration and Exchange Rate Dynamics," *Journal of Finance*, 49, 727–735.

Escribano, A. and C. W. Granger (1998): "Investigating the Relationship between Gold and Silver Prices," *Journal of Forecasting*, 17, 81–107.

Fama, E. F. and K. R. French (1988): "Permanent and Transitory Component of Stock Prices," *Journal of Political Economy*, 96, 246–273.

Gatev, E., W. N. Goetzmann, and K. G. Rouwenhorst (2006): "Pairs Trading: Performance of a Relative-Value Arbitrage Rule," *Review of Financial Studies*, 19, 797–827.

Höpfner, R. and Y. Kutoyants (2003): "On a Problem of Statistical Inference in Null Recurrent Diffusions," *Statistical Inference for Stochastic Processes*, 6, 25–42.

Jeong, M. and J. Y. Park (2013): "Asymptotic Theory of Maximum Likelihood Estimator for Diffusion Model," Working Paper, Indiana University.

——— (2016): "An Asymptotic Theory of Jump Diffusion Model," Working Paper, Indiana University.

Jones, C. S. (2003): "Nonlinear Mean Reversion in the Short-Term Interest Rate," *Review of Financial Studies*, 16, 793–843.

KANAYA, S. (2011): "A Nonparametric Test for Stationarity in Continuous-Time Markov Processes," Working Paper, University of Aarhus.

KANAYA, S., J. KIM, AND J. Y. PARK (2017): "Moduli of Continuity of Brownian Motion and Related Processes over an Expanding Time Interval," Working Paper, Indiana University.

KARATZAS, I. AND S. E. SHREVE (1991): *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York.

KIM, J. AND J. Y. PARK (2017): "Asymptotics for Recurrent Diffusions with Application to High Frequency Regression," *Journal of Econometrics*, 196, 27–54.

KIM, M. J., C. R. NELSON, AND R. STARTZ (1991): "Mean Reversion in Stock Prices: A Reappraisal of the Empirical Evidence," *Review of Economic Studies*, 58, 515–528.

LUCEY, B. AND E. TULLY (2006): "The Evolving Relationship between Gold and Silver 1978-2002: Evidence from a Dynamic Cointegration Analysis: A Note," *Applied Financial Economics Letters*, 2, 47–53.

MÜLLER, U. K. AND M. W. WATSON (2017): "Long-Run Covariability," Forthcoming from *Econometrica*.

PERRON, P. (1989): "Testing for a Random Walk: A Simulation Experiment When the Sampling Interval Is Varied. In B. Raj (ed.)," in *Advances in Econometrics and Modeling*, Dordrecht: Kluwer Academic Publisher, 47–68.

——— (1991): "Test Consistency with Varying Sampling Frequency," *Econometric Theory*, 7, 341–368.

POTERBA, J. M. AND L. H. SUMMERS (1988): "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics*, 22, 27–59.

ROGERS, L. C. G. AND D. WILLIAMS (2000): *Diffusions, Markov Processes, and Martingales*, Cambridge University Press.

ROMANO, J. P. AND M. WOLF (2001): "Subsampling Intervals in Autoregressive Models with Linear Time Trend," *Econometrica*, 69, 1283–1314.

SHILLER, R. J. AND P. PERRON (1985): "Testing the Random Walk Hypothesis: Power versus Frequency of Observation," *Economics Letters*, 18, 381 – 386.

Stock, J. and M. Watson (1988): "Testing for Common Trends," *Journal of the American Statistical Association*, 83, 1097–1107.

Stone, C. (1963): "Limit Theorems for Random Walks, Birth and Death Processes, and Diffusion Processes," *Illinois Journal of Mathematics*, 7, 638–660.