A Nonparametric Unit Root Test under Nonstationary Volatility

Burak Alparslan Eroğlu\textsuperscript{a}, Taner Yiğit\textsuperscript{a}\textsuperscript{*}

\textsuperscript{a}Department of Economics, Bilkent University, Ankara, 06800 Turkey;
email:tyigit@bilkent.edu.tr; tel: +90 312 2901898

September 17, 2015

\textbf{Abstract:} We develop a new nonparametric unit root testing method that is robust to permanent shifts in the innovation variance. Unlike other methods in the literature, our test does not require a parametric specification or the selection of lag length and bandwidth to adjust for serial correlation.

\textit{Keywords:} Nonstationary Volatility, Fractionally Integrated Time Series, Variance Ratio Statistic, Unit Root Testing

\textit{JEL classification:} C22, C32.
1. Introduction

Recent body of empirical evidence indicates that variance shifts (nonstationary volatility) is a common occurrence in macroeconomic and financial data; see Busetti and Taylor (2003), McConnell and Perez-Quiros (1998) and Sensier and Van Dijk (2004). This finding coupled with nonstationarity in the levels of these types of data led the researchers to investigate the impact of variance shifts on unit root tests. In one of these studies, Cavaliere and Taylor (2007), henceforth CT, document that under nonstationary volatility, the asymptotic distributions of standard unit root tests are altered by the inclusion of a new nuisance parameter called the "variance profile", leading to size distortions in these tests. In order to achieve correct inference, CT suggest first consistently estimating this nuisance parameter and then updating the asymptotic distribution of Phillips and Perron’s (1988) tests with this estimate. While their inclusion of the new nuisance parameter generates significant gains in size over classical unit root tests, they still rely on the methodologies used in earlier studies to correct for other nuisance parameters such as serial correlation in errors. CT adjust their test statistic via the estimation of the long run variance, obtained by a semi-parametric kernel or a parametric ADF based regression estimation. The success of these methods highly depends on lag length, bandwidth and Kernel selection in terms of finite sample properties. In this paper, we propose a nonparametric unit root test that is robust to nonstationary volatility problem yet does not require a long run variance estimation.

We derive our test statistic by modifying Nielsen’s (2009) nonparametric variance ratio statistic with the nonparametric variance profile estimator of CT. Computation of the proposed test statistic involves a fractional transformation of observed series, but it does not require any parametric regression or the choice of any tuning parameters like lag length and bandwidth. Therefore, we not only modify Nielsen’s test to be robust against nonstationary volatility, but also improve on the finite sample properties of CT statistic for various types of serial correlation, especially where short run cycles are strong. Derivation of the limiting distribution of fractionally integrated processes with nonstationary volatility and the proofs are placed in the Appendix.\(^1\)

2. Model and Variance Ratio Test

2.1. Model

Let \( \{x_t\}_{t=0}^T \) be generated by:

\[
\begin{align*}
   x_t &= y_t + \theta \delta_t \quad (1) \\
   y_t &= \rho y_{t-1} + u_t \quad (2) \\
   u_t &= C(L)e_t \quad (3) \\
   e_t &= \sigma_t e_i \quad (4)
\end{align*}
\]

where \( e_t \sim iid(0,1) \) and \( \theta \delta_t \) is the deterministic term and \( C(L) \) is the lag polynomial. From CT, we have following assumptions:

**Assumption. A.1** The lag polynomial \( C(L) \neq 0 \) for all \( |L| \leq 1 \), and \( \sum_{j=0}^{\infty} |c_j| < \infty \).

\[ \mathbb{E} |e_t| < K < \infty \] for some \( r \geq 4 \).

---

\(^1\) The notation in the paper follows Cavaliere and Taylor (2007).

*Corresponding author. Email: tyigit@bilkent.edu.tr*
A.2 \( \rho \) satisfies \(|\rho| \leq 1\).

A.3 \( \sigma_t \) satisfies \( \sigma_{[t,x]} := \omega(s) \) for all \( s \in [0,1] \), where \( \omega(.) \in \mathcal{D} \) is non-stochastic and strictly positive with for \( t < 0 \), \( \sigma_t \leq \sigma^* < \infty \).

The assumptions A.1 and A.2 are very standard in unit root testing literature. CT characterizes the dynamics of innovation variance in A.3, which should be bounded and display a countable number of jumps.

A fundamental object that is defined in CT is given below:

\[
\eta(s) := \left( \int_0^1 \omega(r)^2 \, dr \right)^{-1} \left( \int_0^1 \omega(r) \, dr \right)
\]  

(5)

This object is referred to as the variance profile of the process. Further, CT shows that

\[
\int_0^1 \omega(r)^2 \, dr = \tilde{\sigma}^2
\]

is the limit of \( T^{-1} \sum_{t=1}^T \sigma_t^2 \).

### 2.2. Variance Ratio test under nonstationary volatility

So as to modify the Variance Ratio test (Nielsen, 2009) statistic we first need the fractional partial sum operator for some \( d > 0 \):

\[
\bar{x}_t := \Delta_t^d x_t = (1-L)^d x_t = \sum_{k=0}^{t-1} \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)} x_{t-k} = \sum_{k=0}^{t-k} \tau_k(d)x_{t-k}
\]  

(6)

where \( \Gamma(.) \) is gamma function. Under the assumptions A, following Lemmas hold:

**Lemma 1.** Assume that \( \{u_{it}\}_{t=0}^T \) is generated by (3)-(4) and \( \rho = 1-c / T \) with \( c \geq 0 \).

i. \( y_T(t) = T^{-1/2} \sum_{k=1}^{[Tn]} e^{-c(n-k)} u_k \) \( \overset{w}{\longrightarrow} \) \( \tilde{\sigma} \mathcal{C}(1)J_{\alpha}^c(t) \), where \( J_{\alpha}^c(t) = \int_0^t \exp(-c(t-s))dB_{\alpha}(s) \) and

\[
B_{\alpha}(s) = \tilde{\sigma}^{-1} \int_0^s \omega(r)dB(r).
\]

ii. \( B_{\eta}(s) := B(\eta(s)) \) where \( B_{\eta}(s) \) variance transformed Brownian motion, \( \eta(s) \) is defined in (5). Thus, \( J_{\eta}^c(t) := J_{\eta}^c(t) = \int_0^t \exp(-c(t-s))dB_{\eta}(s) \).

iii. \( \tilde{y}_T(t) = T^{-d} \Delta_{t-d} y_T(t) \overset{w}{\longrightarrow} \tilde{\sigma} \mathcal{C}(1)J_{\alpha,d}^c(t) \), where \( J_{\alpha,d}^c(t) = \Gamma(d+1)^{-1} \int_0^t (t-s)^d dJ_{\alpha}^c(s) \). Further, we have \( J_{\eta,d}^c(t) = J_{\eta,d}^c(t) \).

**Remark 1.** Lemma 1.(i) and 1.(ii) are from Cavaliere (2005) and CT. Lemma 1.(iii) is new and establishes weak convergence for fractionally integrated processes with non-stationary volatility. Although Demetrescu and Sibbertsen (2014) models the fractional integrated process with non-stationary volatility, they do not establish weak convergence of this object.

**Remark 2.** Note that under the null hypothesis of \( \rho = 1 \) or \( c = 0 \) the above variance transformed Uhlenbeck-Orstein process becomes a variance transformed Brownian motion. For instance, under the null the partial sum process \( \tilde{y}_T(t) \) will converge to \( \tilde{\sigma} \mathcal{C}(1) \int_0^t (t-s)^d dB_{\eta}(s) \) where we
can define $B_{\eta,d}(t) := \int_0^t (t-s)^d dB_{\eta}(s)$. This limiting distribution resembles the type II fractional Brownian motions defined by Marinucci and Robinson (2000), since $B_{\eta,d}(t)$ does not contain any pre-historic influence (see also Wang et al. (2002)).

Like Nielsen (2009), we apply OLS detrending to the observed series $x_t$ to clean out the deterministic terms. Let $\hat{x}_t$ be the OLS detrended residuals and defining $\tilde{\hat{x}}_t = \Delta^d_+ \hat{x}_t$, our test statistic is then given by:

$$\tau_\eta(d) = T^{2d} \sum_{t=1}^{T} \frac{\hat{x}_t^2}{\sum_{t=1}^{T} \tilde{\hat{x}}_t^2} \quad (7)$$

**Theorem 1.** Assume that the time series $\{x_t\}$ is generated by equations (1)-(4) and $\rho = 1 - c / T$ for $c \geq 0$. Let $j = 0$ when $\delta_j = 0$, $j = 1$ when $\delta_j = 1$ and when $\delta_j = [1, t]'$ for $d > 0$

i. $\hat{x}_t \Rightarrow J^c_{\eta,j}(t)$ where $J^c_{\eta,j}(t) = J^c_{\eta,j}(t) - \left(\int_0^t J^c_{\eta,j}(s) ds \right) \left(\int_0^t D_j(s) D_j(s)' ds \right) D_j(t)$ for $j = 1, 2$, and $D_1(s) = 1$, $D_2(s) = [1, s]'$ and $J^c_{\eta,0}(t) = J^c_{\eta,0}(t)$.

ii. $\tilde{\hat{x}}_t \Rightarrow J^{c, d}_{\eta,d,j}(t)$ where $J^{c, d}_{\eta,d,j}(t) = J^{c, d}_{\eta,d,j}(t) - \left(\int_0^t J^{c, d}_{\eta,d,j}(s) ds \right) \left(\int_0^t D_j(s) D_j(s)' ds \right) \left(\int_0^t (t-s)^{d-1} D_j(r) dr \right)$ for $j = 1, 2$. Further $J^{c, d}_{\eta,d,j}(t) = J^{c, d}_{\eta,d,j}(t)$.

iii. $\tau_\eta(d) = T^{2d} \sum_{t=1}^{T} \frac{\hat{x}_t^2}{\sum_{t=1}^{T} \tilde{\hat{x}}_t^2} U_{J,d}(d) = \frac{(\tilde{\eta} C(1)^2 \int_0^T J^c_{\eta,j}(s)^2 ds)}{(\tilde{\eta} C(1)^2 \int_0^T J^{c, d}_{\eta,d,j}(s)^2 ds)} \cdot \frac{\int_0^T J^c_{\eta,j}(s)^2 ds}{\int_0^T J^{c, d}_{\eta,d,j}(s)^2 ds}$

**Remark 3.** Note that short run dynamics cancel out in asymptotic distribution since the numerator and the denominator share the same long run variance component in part (iii).

**2.3. Simulated Asymptotic distribution**

The test statistic obtained in Theorem 1 involves $\eta(s)$ as nuisance parameter which can be consistently estimated by modifying the nonparametric estimator in CT:

$$\hat{\eta}(s) := \frac{\sum_{t=1}^{T} (\Delta \hat{x}_t)^2 + (Ts - \lceil Ts \rceil)(\Delta \hat{x}_t)_{\lceil Ts \rceil + 1}^2}{\sum_{t=1}^{T} (\Delta \hat{x}_t)^2} \quad (8)$$

**Theorem 2.** Under the conditions of Theorem 1

i. (CT shows) $B_{\eta,T}(s) := T^{-1/2} \sum_{t=1}^{\lfloor Ts \rfloor T} e_t \Rightarrow B_{\eta}(s)$

ii. $B^{d}_{\eta,d,T}(s) := T^{-d} \Delta^d_+ B_{\eta,T}(s) \Rightarrow B^{d}_{\eta,d}(s)$
After obtaining the consistent estimate for $\eta(s)$, we can simulate the asymptotic distribution and the critical value for the test statistic. First, we choose a step level $N$. For $s = j/N$ for $j = 1, 2, ..., N$, then we compute $\hat{\eta}(Ts)/T$ using (8). By drawing $e_i$ from $N(0, 1)$, we obtain $B_{\eta,j}(s)$. Then applying fractional integration operator $\Delta_{-d}^d$ to this object and multiplying it by $T^{-d}$, we get $B_{\eta,j,d,T}(s)$. This asymptotic distribution is then used to generate the critical values for the test. The proposed test rejects the null hypothesis for large values the test statistic, that is, we reject if $\tau_q(d)$ is greater than $(1 - \alpha)$ quantile of $U_{\eta,j}$.

3. Monte Carlo Experiments

In the Monte Carlo simulations, data is generated according to equations (1)-(4) with $T = \{100, 500\}$. We consider following specifications for error term variance:

i. **Constant volatility (CV)**: $\omega(s) = 1$ for $s \in [0, 1]$ and $\sigma_0 > 0$.

ii. **Single break in volatility (SBV)**: $\omega(s) = 1 + 2*1(s > 0.2*T)$ for $s \in [0, 1]$.

iii. **Trending volatility (TV)**: $\omega(s) = 1 + 2*s$ for $s \in [0, 1]$.

iv. **Exponential integrated Stochastic volatility (EISV)**: $\omega(s) = \sigma_0 \exp(4B(s))$ for $s \in [0, 1]$ where $B(s)$ is standard Brownian process.

The innovations $e_i$ are drawn from $N(0, 1)$. All simulations are conducted $MC = 10000$ times. We fix the step size $N$ to $T$ in simulating the variance shifted Brownian motions. We consider four scenarios for serial correlation in innovations. First one does not contain any serial correlation. In second, $u_i$ follows a simple AR(1) model: $u_i = 0.5u_{i-1} + e_i$, third is an ARMA(2,2) process: $u_i = 0.1u_{i-1} + 0.07u_{i-2} - 0.4e_{i-1} + 0.2e_{i-2} + e_i$. Last one follows a MA(2) process: $u_i = -0.2e_{i-1} + 0.15e_{i-2}$. We fix $\rho = \{1, 0.93, 0.86\}$. $\rho = 0$ indicates size and other values are for power evaluation. We also provide simulation for Cavaliere and Taylor (2007) $MZ^s_t$ test.  

| Table 1: Empirical Size and Power with No serial correlation |
|---------------|-------------|-------------|-------------|-------------|-------------|-------------|
|               | $\tau_q(d)$ | $MZ^s_t$   |             |             |             |
|               | $\rho = 1$  | $\rho = 0.93$ | $\rho = 0.86$ | $\rho = 1$  | $\rho = 0.93$ | $\rho = 0.86$ |
| **CV**        |             |             |             |             |             |
| $T=100$       | 0.043       | 0.353       | 0.726       | 0.003       | 0.035       | 0.076       |
| $T=500$       | 0.048       | 0.990       | 1.000       | 0.046       | 1.000       | 1.000       |
| **SBV**       |             |             |             |             |             |
| $T=100$       | 0.049       | 0.294       | 0.64        | 0.029       | 0.277       | 0.647       |
| $T=500$       | 0.052       | 0.976       | 1.000       | 0.043       | 0.997       | 0.999       |
| **TV**        |             |             |             |             |             |
| $T=100$       | 0.053       | 0.370       | 0.762       | 0.028       | 0.339       | 0.747       |
| $T=500$       | 0.050       | 0.992       | 1.000       | 0.042       | 0.999       | 1.000       |
| **EISV**      |             |             |             |             |             |
| $T=100$       | 0.045       | 0.233       | 0.534       | 0.021       | 0.110       | 0.298       |
| $T=500$       | 0.049       | 0.939       | 0.999       | 0.041       | 0.903       | 0.951       |

The confidence level is 0.05 and there is no trend and mean components. $d$ is fixed to 0.1 as recommended in Nielsen (2009). For formula and asymptotic distribution of $MZ^s_t$ test see CT. In fact, CT propose 3 different test statistic, but we only give the results of the best performing one from among these tests. For selection of lag length, we utilize MAIC proposed by Ng and Perron (2001). Simulation results for different serial correlation specifications will be provided by the authors upon request.
Table 2: Empirical Size and Power with AR(1) innovations

<table>
<thead>
<tr>
<th></th>
<th>$\tau_e(d)$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho = 1$</td>
<td>$\rho = 0.93$</td>
<td>$\rho = 0.86$</td>
<td>$\rho = 1$</td>
<td>$\rho = 0.93$</td>
<td>$\rho = 0.86$</td>
</tr>
<tr>
<td>CV</td>
<td>T=100</td>
<td>0.028</td>
<td>0.209</td>
<td>0.505</td>
<td>0.041</td>
<td>0.343</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.037</td>
<td>0.975</td>
<td>1.000</td>
<td>0.046</td>
<td>0.999</td>
</tr>
<tr>
<td>SBV</td>
<td>T=100</td>
<td>0.029</td>
<td>0.171</td>
<td>0.390</td>
<td>0.040</td>
<td>0.272</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.044</td>
<td>0.952</td>
<td>0.999</td>
<td>0.051</td>
<td>0.997</td>
</tr>
<tr>
<td>TV</td>
<td>T=100</td>
<td>0.027</td>
<td>0.271</td>
<td>0.556</td>
<td>0.032</td>
<td>0.329</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.045</td>
<td>0.977</td>
<td>1.000</td>
<td>0.048</td>
<td>0.999</td>
</tr>
<tr>
<td>EISV</td>
<td>T=100</td>
<td>0.028</td>
<td>0.114</td>
<td>0.311</td>
<td>0.032</td>
<td>0.133</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.037</td>
<td>0.897</td>
<td>0.993</td>
<td>0.037</td>
<td>0.897</td>
</tr>
</tbody>
</table>

Table 3: Empirical Size and Power with ARMA(2,2) innovations

<table>
<thead>
<tr>
<th></th>
<th>$\tau_e(d)$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho = 1$</td>
<td>$\rho = 0.93$</td>
<td>$\rho = 0.86$</td>
<td>$\rho = 1$</td>
<td>$\rho = 0.93$</td>
<td>$\rho = 0.86$</td>
</tr>
<tr>
<td>CV</td>
<td>T=100</td>
<td>0.047</td>
<td>0.335</td>
<td>0.695</td>
<td>0.015</td>
<td>0.161</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.050</td>
<td>0.989</td>
<td>1.000</td>
<td>0.043</td>
<td>0.999</td>
</tr>
<tr>
<td>SBV</td>
<td>T=100</td>
<td>0.039</td>
<td>0.286</td>
<td>0.619</td>
<td>0.014</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.042</td>
<td>0.973</td>
<td>1.000</td>
<td>0.042</td>
<td>0.995</td>
</tr>
<tr>
<td>TV</td>
<td>T=100</td>
<td>0.046</td>
<td>0.365</td>
<td>0.747</td>
<td>0.015</td>
<td>0.193</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.045</td>
<td>0.990</td>
<td>1.000</td>
<td>0.04</td>
<td>0.999</td>
</tr>
<tr>
<td>EISV</td>
<td>T=100</td>
<td>0.049</td>
<td>0.225</td>
<td>0.518</td>
<td>0.017</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.049</td>
<td>0.933</td>
<td>0.998</td>
<td>0.04</td>
<td>0.888</td>
</tr>
</tbody>
</table>

Table 4: Empirical Size and Power with MA(2) innovations

<table>
<thead>
<tr>
<th></th>
<th>$\tau_e(d)$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho = 1$</td>
<td>$\rho = 0.93$</td>
<td>$\rho = 0.86$</td>
<td>$\rho = 1$</td>
<td>$\rho = 0.93$</td>
<td>$\rho = 0.86$</td>
</tr>
<tr>
<td>CV</td>
<td>T=100</td>
<td>0.044</td>
<td>0.348</td>
<td>0.733</td>
<td>0.023</td>
<td>0.272</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.050</td>
<td>0.990</td>
<td>1.000</td>
<td>0.037</td>
<td>1.000</td>
</tr>
<tr>
<td>SBV</td>
<td>T=100</td>
<td>0.047</td>
<td>0.305</td>
<td>0.618</td>
<td>0.021</td>
<td>0.193</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.046</td>
<td>0.972</td>
<td>1.000</td>
<td>0.037</td>
<td>0.997</td>
</tr>
<tr>
<td>TV</td>
<td>T=100</td>
<td>0.049</td>
<td>0.383</td>
<td>0.765</td>
<td>0.018</td>
<td>0.245</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.053</td>
<td>0.991</td>
<td>1.000</td>
<td>0.042</td>
<td>0.999</td>
</tr>
<tr>
<td>EISV</td>
<td>T=100</td>
<td>0.047</td>
<td>0.348</td>
<td>0.726</td>
<td>0.021</td>
<td>0.237</td>
</tr>
<tr>
<td></td>
<td>T=500</td>
<td>0.054</td>
<td>0.937</td>
<td>0.999</td>
<td>0.041</td>
<td>0.900</td>
</tr>
</tbody>
</table>

4. Conclusion

Simulation evidence suggests the proposed nonparametric unit root test has desirable size and power properties in all scenarios. Especially in no serial correlation case, our test almost dominates CT’s test in terms of size. Furthermore, small sample properties of our test are better than CT’s tests when complicated serial correlation structures are considered with a short run cycles (MA component).

5. References


6. Appendix

**Proof of Lemma 1:** Part (i) can be found in Theorem 1 of CT and Remark 3.1. Part (ii) is from Proposition 3 of Cavaliere (2005)

For part (iii), write the partial sum process for $\tilde{y}_T$ as:

$$\tilde{y}_T(t) = T^{-1/2-d} \sum_{k=1}^{[T]} \pi_{[T]-k} (d) y_k$$
where \( \pi_k(d) = \frac{\Gamma(k + d)}{\Gamma(d)\Gamma(k + 1)} \), and Wang et al. (2000) claims that \( \sum_{j=0}^{m} \pi_j(d) = \pi_a(d + 1) \), thus we have:

\[
\tilde{y}_T(t) = T^{-1/2-a} \sum_{k=1}^{[Tt]} \pi_{[Tt] - k}(d) \sum_{j=1}^{k} u_k \\
= T^{-1/2-a} \sum_{k=1}^{[Tt]} \sum_{j=1}^{k} \pi_{[Tt] - k}(d) u_k = T^{-1/2-a} \sum_{k=1}^{[Tt]} (Tt - k) \frac{d^d}{\Gamma(d + 1)} u_k \\
= T^{-1/2-a} \sum_{k=1}^{[Tt]} \left( \frac{Tt}{d} - k \right) \frac{d^d}{\Gamma(d + 1)} \frac{1}{T^{-1/2}} \Delta y_k
\]

here \( \Delta y_k \) can be written as \( \int_{(k-1)/T}^{k/T} dy_T(s) \) in the limit (see Phillips (1987). Then,

\[
\tilde{y}_T(t) = \sum_{k=1}^{[Tt]} \left( \frac{Tt}{d} - k \right) \frac{d^d}{\Gamma(d + 1)} \int_{(k-1)/T}^{k/T} dy_T(s) = \sum_{k=1}^{[Tt]} \sum_{j=1}^{k} \frac{d^d}{\Gamma(d + 1)} dy_T(s) \\
\xrightarrow{w} \int_{0}^{t} (t - s) \frac{d^d}{\Gamma(d + 1)} dy_T(s) \xrightarrow{w} C(1) \int_{0}^{t} (t - s) dJ_{s}^c(s)
\]

**Proof of Theorem 1:** To prove part (i), consider the residuals from the regression of \( \delta_{s} \) on \( x_{s} \) for \( t = 1, \ldots, T \), for \( s \in [0,1] \):

\[
\hat{\delta}_{s[Ts]} = y_{s[Ts]} - (\hat{\theta} - \theta) \delta_{s[Ts]} \\
T^{-1/2} \hat{\delta}_{s[Ts]} = T^{-1/2} y_{s[Ts]} - T^{-1/2} (\hat{\theta} - \theta) \delta_{s[Ts]} \tag{11}
\]

We have already establish limiting distribution for first factor on the left hand side of equation (12). This factor also corresponds to the case when \( \delta_{s} = 0 \). For second factor, define \( N(T) = 1 \) when \( \delta_{s} = 1 \) and \( N(T) = \left[ \begin{array}{cc} 1 & 0 \\ 0 & T^{-1} \end{array} \right] \) when \( \delta_{s} = [1, t]' \) and from Nielsen (2009):

\[
T^{-1/2}(\hat{\theta} - \theta) \delta_{s[Ts]} \\
= \left( T^{-1} \sum_{j=1}^{T} y_{s} D_{j}(s/T) \right) \left( T^{-1} \sum_{j=1}^{T} D_{j}(s/T) D_{j}(s/T)' \right)^{-1} N(T) D_{j}(\lfloor Ts \rfloor / T)
\]

by application of CMT and \( D(\lfloor Ts \rfloor / T) \rightarrow D_{j}(s) \) we have

\[
T^{-1/2}(\hat{\theta} - \theta) \delta_{s[Ts]} \xrightarrow{w} \bar{\sigma} C(1) \left( \int_{0}^{1} J_{s}^c(r) D_{j}(r) dr \right) \left( \int_{0}^{1} D_{j}(r) D_{j}(r)' dr \right)^{-1} D_{j}(s)
\]

Finally we have

\[
T^{-1/2} \hat{\delta}_{s[Ts]} \rightarrow \bar{\sigma} C(1) J_{s}^c(s) - \bar{\sigma} C(1) \left( \int_{0}^{1} J_{s}^c(r) D_{j}(r) dr \right) \left( \int_{0}^{1} D_{j}(r) D_{j}(r)' dr \right)^{-1} D_{j}(s) \tag{13}
\]

For part (ii) We can write the partial sum process to find the limits.
\[ T^{-1/2-d} \hat{X}_t = T^{-1/2-d} \sum_{k=0}^{T} \pi_k(d) y_{t-k} - T^{-1/2-d} (\hat{\theta} - \theta) \sum_{k=0}^{T} \pi_k(d) \delta_{t-k} \]

First factor converges by Lemma 1 part (iii). The second factor:

\[ T^{-1/2} (\hat{\theta} - \theta) T^{-d} \sum_{k=0}^{T} \pi_k(d) \delta_{t-k} \]

here \( T^{-1/2} (\hat{\theta} - \theta)' \rightarrow \delta C(1) \left( \int_{0}^{c} D_j(s) ds \right) \left( \int_{0}^{1} D_j(s) D_j(s)' ds \right)^{-1} \) by equation (13).

The convergence for \( T^{-d} \sum_{k=0}^{T} \pi_k(d) \delta_{t-k} \) is already proved by Nielsen (2009), that is

\[ T^{-d} \sum_{k=0}^{T} \pi_k(d) \delta_{t-k} \rightarrow \int_{0}^{1} \left( t - s \right)^{d-1} D_j(s) ds \]

Part (iii), is derived by application of CMT using the objects we found in parts (i)-(ii).

**Proof of Theorem 2.** Part (i) directly follows from Theorem 3 of CT.

For part (ii), define the partial sum \( S_T(t) = T^{-1/2} \sum_{r=1}^{\lfloor Tt \rfloor} e_r \) and \( \hat{S}_T(t) = T^{-d} \Delta^d S_T(t) \)

Note that from equation (10) we have:

\[ \hat{S}_T(t) \rightarrow \int_{0}^{t} \frac{(t-s)^d}{\Gamma(d+1)} dS_T(s) \]

Now, \( B_{\eta,T} := S_T(\hat{\eta}((\lfloor Ts \rfloor / T)T) \) and \( B_{\eta,d,T}(s) = T^{-d} \Delta^d S_T(\hat{\eta}((\lfloor Ts \rfloor / T)T) \)

, then we have

\[ B_{\eta,d,T}(s) \rightarrow \int_{0}^{t} \frac{(t-s)^d}{\Gamma(d+1)} dS_T(\hat{\eta}((\lfloor Ts \rfloor / T)T) \). But, from part (i), Theorem 3 of CT indicates that

\[ S_T(\hat{\eta}((\lfloor Ts \rfloor / T)T) \rightarrow B_{\eta}(s) \). From CMT we obtain the result.