A Nonparametric Unit Root Test under Nonstationary Volatility

Burak Alparslan Eroğlu[a], Taner Yiğit[a,*]

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

September 17, 2015

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.*

Keywords: Nonstationary Volatility, Fractionally Integrated Time Series, Variance Ratio Statistic, Unit Root Testing
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:

$$x_t = y_t + \theta'\delta_t \tag{1}$$

$$y_t = \rho y_{t-1} + u_t \tag{2}$$

$$u_t = C(L)\varepsilon_t \tag{3}$$

$$\varepsilon_t = \sigma_t e_t \tag{4}$$

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. $\mathcal{A}.1$ *The lag polynomial* $C(L) \neq 0$ *for all* $|L| \leq 1$, *and* $\sum_{j=0}^{\infty} j|c_j| < \infty$. $\mathbb{E}|e_t|^r < K < \infty$ *for some* $r \geq 4$.

[1] The notation in the paper follows Cavaliere and Taylor (2007).

*Corresponding author. Email: tyigit@bilkent.edu.tr

$\mathcal{A}.2$ ρ satisfies $|\rho|\leq 1$.

$\mathcal{A}.3$ σ_t satisfies $\sigma_{\lfloor Ts \rfloor} := \omega(s)$ 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 $\mathcal{A}.1$ and $\mathcal{A}.2$ are very standard in unit root testing literature. CT characterizes the dynamics of innovation variance in $\mathcal{A}.3$, which should be bounded and display a countable number of jumps.

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

$$\eta(s) := \left(\int_0^1 \omega(r)^2 dr \right)^{-1} \left(\int_0^s \omega(r)^2 dr \right) \tag{5}$$

This object is referred to as the *variance profile* of the process. Further, CT shows that $\int_0^1 \omega(r)^2 dr = \bar{\omega}^2$ 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 $d > 0$:

$$\tilde{x}_t := \Delta_+^{-d} x_t = (1-L)_+^{-d} x_t = \sum_{k=0}^{t-1} \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)} x_{t-k} = \sum_{k=0}^{t-k} \pi_k(d) x_{t-k} \tag{6}$$

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

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

i. $y_T(t) = T^{-1/2}\sum_{k=1}^{\lfloor Tt \rfloor} e^{-c(\lfloor Tt \rfloor - k)} u_k \xrightarrow{w} \bar{\omega} C(1) J_\omega^c(t)$, where $J_\omega^c(t) = \int_0^t exp(-c(t-s)) dB_\omega(s)$ and $B_\omega(s) = \bar{\omega}^{-1}\int_0^s \omega(r) dB(r)$.

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

iii. $\tilde{y}_T(t) = T^{-d}\Delta_+^{-d} y_T(t) \xrightarrow{w} \bar{\omega} C(1) J_{\omega,d}^c(t)$, where $J_{\omega,d}^c(t) = \Gamma(d+1)^{-1}\int_0^t (t-s)^d dJ_\omega^c(s)$. Further, we have $J_{\omega,d}^c(t) = J_{\eta,d}^c(t)$.

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* $\bar{\omega} C(1)\int_0^t (t-s)^d dB_\eta(s)$ *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 Rabinson (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:

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

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* $\delta_t = 0$, $j = 1$ *when* $\delta_t = 1$ *and when* $\delta_t = [1,t]'$ *for* $d > 0$

i. $\hat{x}_T(t) \xrightarrow{w} J_{\eta,j}^c(t)$ where $J_{\eta,j}^c(t) = J_\eta^c(t) - \left(\int_0^1 J_\eta^c(s)D_j(s)'ds\right)\left(\int_0^1 D_j(s)D_j(s)'ds\right)^{-1} D_j(t)$ for $j = 1,2$, and $D_1(s) = 1$, $D_2(s) = [1,s]'$ and $J_{\eta,0}^c(t) = J_\eta^c(t)$.

ii. $\tilde{\hat{x}}_T(t) \xrightarrow{w} J_{\eta,d,j}^c(t)$ where $J_{\eta,d,j}^c(t) = J_{\eta,d}^c(t) - \left(\int_0^1 J_{\eta,d}^c(s)D_j(s)'ds\right)\left(\int_0^1 D_j(s)D_j(s)'ds\right)^{-1} \int_0^t \frac{(t-r)^{d-1}}{\Gamma(d)} D_j(r)dr$ for $j = 1,2$. Further $J_{\eta,d,0}^c(t) = J_{\eta,d}^c(t)$

iii. $\tau_\eta(d) = T^{2d}\frac{\sum_{t=1}^{T}\hat{x}_t^2}{\sum_{t=1}^{T}\tilde{\hat{x}}_t^2} \xrightarrow{w} U_{j,\eta}(d) = \frac{(\varpi C(1))^2\int_0^1 J_{\eta,j}^c(s)^2 ds}{(\varpi C(1))^2\int_0^1 J_{\eta,d,j}^c(s)^2 ds} = \frac{\int_0^1 J_{\eta,j}^c(s)^2 ds}{\int_0^1 J_{\eta,d,j}^c(s)^2 ds}$.

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:

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

Theorem 2. *Under the conditions of Theorem 1*

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

ii. $B_{\hat{\eta},d,T}(s) := T^{-d}\Delta_+^{-d}B_{\hat{\eta},T}(s) \xrightarrow{w} B_{\eta,d}(s)$

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 $\hat{\eta}(\lfloor Ts \rfloor / T)$ using (8). By drawing e_t from $N(0,1)$, we obtain $B_{\hat{\eta},T}(s)$. Then applying fractional integration operator Δ_+^{-d} to this object and multiplying it by T^{-d}, we get $B_{\hat{\eta},d,T}(s)$. 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 $\tau_\eta(d)$ 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_t 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_t follows a simple AR(1) model: $u_t = 0.5u_{t-1} + \varepsilon_t$, third is an ARMA(2,2) process: $u_t = 0.1u_{t-1} + 0.07u_{t-2} - 0.4\varepsilon_{t-1} + 0.2\varepsilon_{t-2} + \varepsilon_t$. Last one follows a MA(2) process; $u_t = -0.2\varepsilon_{t-1} + 0.15\varepsilon_{t-2}$. 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_t^S test.[2]

Table 1: Empirical Size and Power with No serial correlation

		$\tau_\eta(d)$			MZ_t^s		
		$\rho = 1$	$\rho = 0.93$	$\rho = 0.86$	$\rho = 1$	$\rho = 0.93$	$\rho = 0.86$
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

[2] 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_t^S 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							
		$\tau_\eta(d)$			MZ_t^s		
		$\rho = 1$	$\rho = 0.93$	$\rho = 0.86$	$\rho = 1$	$\rho = 0.93$	$\rho = 0.86$
CV	T=100	0.028	0.209	0.505	0.041	0.343	0.654
	T=500	0.037	0.975	1.000	0.046	0.999	1.000
SBV	T=100	0.029	0.171	0.390	0.040	0.272	0.530
	T=500	0.044	0.952	0.999	0.051	0.997	0.999
TV	T=100	0.027	0.271	0.556	0.032	0.329	0.615
	T=500	0.045	0.977	1.000	0.048	0.999	1.000
EISV	T=100	0.028	0.114	0.311	0.032	0.133	0.261
	T=500	0.037	0.897	0.993	0.037	0.897	0.957

Table 3: Empirical Size and Power with ARMA(2,2)innovations							
		$\tau_\eta(d)$			MZ_t^s		
		$\rho = 1$	$\rho = 0.93$	$\rho = 0.86$	$\rho = 1$	$\rho = 0.93$	$\rho = 0.86$
CV	T=100	0.047	0.335	0.695	0.015	0.161	0.519
	T=500	0.050	0.989	1.000	0.043	0.999	1.000
SBV	T=100	0.039	0.286	0.619	0.014	0.116	0.367
	T=500	0.042	0.973	1.000	0.042	0.995	1.000
TV	T=100	0.046	0.365	0.747	0.015	0.193	0.513
	T=500	0.045	0.990	1.000	0.04	0.999	1.000
EISV	T=100	0.049	0.225	0.518	0.017	0.068	0.158
	T=500	0.049	0.933	0.998	0.04	0.888	0.954

Table 4: Empirical Size and Power with MA(2) innovations							
		$\tau_\eta(d)$			MZ_t^s		
		$\rho = 1$	$\rho = 0.93$	$\rho = 0.86$	$\rho = 1$	$\rho = 0.93$	$\rho = 0.86$
CV	T=100	0.044	0.348	0.733	0.023	0.272	0.6680
	T=500	0.050	0.990	1.000	0.037	1.000	1.000
SBV	T=100	0.047	0.305	0.618	0.021	0.193	0.533
	T=500	0.046	0.972	1.000	0.037	0.997	0.999
TV	T=100	0.049	0.383	0.765	0.018	0.245	0.661
	T=500	0.053	0.991	1.000	0.042	0.999	1.000
EISV	T=100	0.047	0.348	0.726	0.021	0.237	0.609
	T=500	0.054	0.937	0.999	0.041	0.900	0.949

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

Fabio Busetti and AM Robert Taylor. Testing against stochastic trend and seasonality in the presence of unattended breaks and unit roots. *Journal of Econometrics*, 117(1):21–53, 2003.

Giuseppe Cavaliere and AM Robert Taylor. Testing for unit roots in time series models with non-stationary volatility. *Journal of Econometrics*, 140(2):919–947, 2007.

Giuseppe Cavaliere. Unit root tests under time-varying variances. *Econometric Reviews*, 23(3):259–292, 2005.

Matei Demetrescu and Philipp Sibbertsen. Inference on the long-memory properties of time series with non-stationary volatility. Technical report, Discussion Paper, Wirtschaftswissenschaftliche Fakultät, Leibniz University of Hannover, 2014.

Domenico Marinucci and Peter M Robinson. Weak convergence of multivariate fractional processes. *Stochastic Processes and their applications*, 86(1):103–120, 2000.

Margaret Mary McConnell and Gabriel Perez-Quiros. Output fluctuations in the United States: what has changed since the early 1980s? *FRB of New York Staff Report*, 41, 1998.

Serena Ng and Pierre Perron. Lag length selection and the construction of unit root tests with good size and power. *Econometrica*, pages 1519–1554, 2001.

Morten Ørregaard Nielsen. A powerful test of the autoregressive unit root hypothesis based on a tuning parameter free statistic. *Econometric Theory*, 25(06):1515–1544, 2009.

Peter CB Phillips. Towards a unified asymptotic theory for autoregression. *Biometrika*, 74(3):535–547, 1987.

P. C. B. Phillips and P. Perron. Testing for a unit root in time series regression. *Biometrika*, 75(2):335â€“346, 1988.

Marianne Sensier and Dick Van Dijk. Testing for volatility changes in us macroeconomic time series. *Review of Economics and Statistics*, 86(3):833–839, 2004.

Q. Wang, Y.X. Lin, and C. Gulati. Asymptotics for general nonstationary fractionally integrated processes without prehistoric influence. *Journal of Applied Mathematics and Decision Sciences*, 6(4):255–269, 2002.

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:

$$\tilde{y}_T(t) = T^{-1/2-d} \sum_{k=1}^{\lfloor Tt \rfloor} \pi_{\lfloor Tt \rfloor - k}(d) y_k$$

where $\pi_k(d) = \dfrac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)}$, and Wang et al. (2000) claims that $\sum_{j=0}^{m}\pi_j(d) = \pi_m(d+1)$, thus we have:

$$\begin{aligned}
\tilde{y}_T(t) &= T^{-1/2-d}\sum_{k=1}^{\lfloor Tt\rfloor}\pi_{\lfloor Tt\rfloor-k}(d)\sum_{j}^{k}u_k \\
&= T^{-1/2-d}\sum_{k=1}^{\lfloor Tt\rfloor}\sum_{j=1}^{k}\pi_{\lfloor Tt\rfloor-k}(d)u_k \quad = \quad T^{-1/2-d}\sum_{k=1}^{\lfloor Tt\rfloor}\pi_{\lfloor Tt\rfloor-k}(d+1)u_k \\
&= T^{-1/2-d}\sum_{k=1}^{\lfloor Tt\rfloor}\frac{\left(\lfloor Tt\rfloor-k\right)^d}{\Gamma(d+1)}u_k \quad = \quad \sum_{k=1}^{\lfloor Tt\rfloor}\frac{\left(t-k/T\right)^d}{\Gamma(d+1)}T^{-1/2}\Delta y_k
\end{aligned}$$

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

$$\tilde{y}_T(t) = \sum_{k=1}^{\lfloor Tt\rfloor}\frac{\left(t-k/T\right)^d}{\Gamma(d+1)}\int_{(k-1)/T}^{k/T} dy_T(s) = \sum_{k=1}^{\lfloor Tt\rfloor}\int_{(k-1)/T}^{k/T}\frac{\left(t-k/T\right)^d}{\Gamma(d+1)}dy_T(s) \tag{9}$$

$$\xrightarrow{w}\int_0^t\frac{\left(t-s\right)^d}{\Gamma(d+1)}dy_T(s)\xrightarrow{w}\frac{C(1)}{\Gamma(d+1)}\int_0^t(t-s)^d\,dJ_\omega^c(s) \tag{10}$$

Proof of Theorem 1: To prove part (i), consider the residuals from the regression of δ_t on x_t for $t=1\ldots T$, for $s\in[0,1]$:

$$\hat{x}_{\lfloor Ts\rfloor} = y_{\lfloor Ts\rfloor} - (\hat{\theta}-\theta)'\delta_{\lfloor Ts\rfloor} \tag{11}$$

$$T^{-1/2}\hat{x}_{\lfloor Ts\rfloor} = T^{-1/2}y_{\lfloor Ts\rfloor} - T^{-1/2}(\hat{\theta}-\theta)'\delta_{\lfloor Ts\rfloor} \tag{12}$$

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 $\delta_t = 0$. For second factor, define $N(T)=1$ when $\delta_t = 1$ and $N(T) = \begin{bmatrix} 1 & 0 \\ 0 & T^{-1}\end{bmatrix}$ when $\delta_t = [1,t]'$ and from Nielsen (2009):

$$\begin{aligned}
&T^{-1/2}(\hat{\theta}-\theta)'\delta_{\lfloor Ts\rfloor} \\
&= \left(T^{-1}\sum_{t=1}^{T}T^{-1/2}y_s D_j(s/T)\right)\left(T^{-1}\sum_{t=1}^{T}D_j(s/T)D_j(s/T)'\right)^{-1}N(T)D_j(\lfloor Ts\rfloor/T)
\end{aligned}$$

by application of CMT and $D(\lfloor Ts\rfloor/T)\to D_j(s)$ we have

$$T^{-1/2}(\hat{\theta}-\theta)\delta_{\lfloor Ts\rfloor}\xrightarrow{w}\bar{\omega}C(1)\left(\int_0^1 J_\eta^c(r)D_j(r)dr\right)\left(\int_0^1 D_j(r)D_j(r)'dr\right)^{-1}D_j(s)$$

Finally we have

$$T^{-1/2}\hat{x}_{\lfloor Ts\rfloor}\to\bar{\omega}C(1)J_\eta^c(s)-\bar{\omega}C(1)\left(\int_0^1 J_\eta^c(r)D_j(r)dr\right)\left(\int_0^1 D_j(r)D_j(r)'dr\right)^{-1}D_j(s) \tag{13}$$

For part (ii) We can write the partial sum process to find the limits

$$T^{-1/2-d}\tilde{\hat{x}}_t = T^{-1/2-d}\sum_{k=0}^{T}\pi_k(d)y_{t-k} - T^{-1/2-d}(\hat{\theta}-\theta)\sum_{k=0}^{T}\pi_k(d)\delta_{t-k}$$

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 $T^{-1/2}(\hat{\theta}-\theta)' \to \bar{\varpi}C(1)\left(\int_0^1 J_\eta^c(s)D_j(s)ds\right)\left(\int_0^1 D_j(s)D_j(s)'ds\right)^{-1}$ 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

$$T^{-d}\sum_{k=0}^{\lfloor Tt \rfloor}\pi_k(d)\delta_{\lfloor Tt \rfloor-k} \to \int_0^t \frac{(t-s)^{d-1}}{\Gamma(d)}D_j(s)ds$$

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 $S_T(t) = T^{-1/2}\sum_{t=1}^{\lfloor Tt \rfloor}e_t$ and $\tilde{S}_T(t) = T^{-d}\Delta_+^{-d}S_T(t)$

Note that from equation (10) we have: $\tilde{S}_T(t) \xrightarrow{w} \int_0^t \frac{(t-s)^d}{\Gamma(d+1)}dS_T(s)$

Now, $B_{\hat{\eta}T}(s) := S_T(\hat{\eta}((\lfloor Ts \rfloor/T)T)$ and $B_{\hat{\eta},d,T}(s) = T^{-d}\Delta_+^{-d}S_T(\hat{\eta}((\lfloor Ts \rfloor/T)T)$, then we have $B_{\hat{\eta},d,T}(s) = \xrightarrow{w} \int_0^t \frac{(t-s)^d}{\Gamma(d+1)}dS_T(\hat{\eta}((\lfloor Ts \rfloor/T)T)$. But, from part (i), Theorem 3 of CT indicates that $S_T(\hat{\eta}((\lfloor Ts \rfloor/T)T) \xrightarrow{w} B_\eta(s)$. From CMT we obtain the result.