APPENDIX for
Identifying trend nature in time series using autocorrelation functions and stationarity tests

M. Boutahar[a] and M. Royer-Carenzi[a]
[a]Aix Marseille Univ, CNRS, Centrale Marseille, I2M, UMR 7373, Marseille, France

Appendix A. Sample autocorrelation functions behavior for (Det,d) models – Proof of Theorem 2.2

We consider the polynomial case with degree $d \geq 1$. Let us define as $S_j(n)$ the sum of the j-th power of the first n integers:

$$S_j(n) \;=\; \sum_{k=1}^{n} k^j \,.$$

From Faulhaber's formula, detailed in [Conway and Guy, 1996], we know that

$$S_j(n) \;=\; \frac{n^{j+1}}{j+1} \,+\, \frac{1}{2}n^j \,+\, \frac{1}{j+1}\sum_{p=2}^{j} \mathcal{B}_p \binom{j+1}{p} n^{j-p+1}\,, \tag{A1}$$

where $\mathcal{B}_p$ are rational numbers called Bernoulli numbers.

Lemma A.1. *For any $j \in \mathbb{N}^*$, the sum of the j-th power of the first n integers, $S_j(n)$, is a $(j+1)$-degree polynomial with leading coefficient $\frac{1}{j+1}$. More precisely,*

$$S_j(n) \;=\; \frac{n^{j+1}}{j+1} \,+\, o(n^{j+1}), \tag{A2}$$

where $o(f)$ is one of the Landau symbols, as defined in [Hardy and Wright, 1975], p. 7-8. $o(f)$ is called "little-O of f", and expresses the convergence to 0 of a given function, when it is divided by f.

Then,

$$\overline{Z} \;=\; \frac{a_d}{d+1}n^d \,+\, o(n^d) \,+\, \overline{B}\,. \tag{A3}$$

In order to study the asymptotic behavior of variables $\Xi(h)$, estimators of the theoretical autocorrelation function for stationary square-integrable processes, we first compute variables $\Gamma(h)$, estimators of the theoretical autocovariance function $\mathrm{cov}(Z_t, Z_{t+h})$.

Definition A.2. *From random variables $(Z_1, \cdots, Z_n)$, we define the autocovariance function as*

$$\Gamma(h) \;=\; \frac{\sum\limits_{k=1}^{n-h} (Z_{k+h} - \overline{Z})(Z_k - \overline{Z})}{n}\,, \quad |h| \le n.$$

Using Equation (A3), we can express $\Gamma(h)$, by summing from $k = 1$ to $k = n - h$ the products of

$$Z_k - \overline{Z} \;=\; \sum_{j=0}^{d} a_j\, k^j \;-\; \frac{a_d}{d+1} n^d \;+\; o(n^d) \;+\; B_k - \overline{B}$$

with

$$Z_{k+h} - \overline{Z} \;=\; \sum_{j=0}^{d} a_j\, (k+h)^j \;-\; \frac{a_d}{d+1} n^d \;+\; o(n^d) \;+\; B_{k+h} - \overline{B}\,.$$

We have to study every term, and clarify its asymptotic behavior.

a) First, we consider all product terms involving either $\overline{B}$ or $\sum_{k=1}^{n-h} \frac{B_k}{n}$, or $\sum_{k=1}^{n-h} \frac{B_{k+h}}{n}$. Let us denote by T_a the sum of all these terms. Since $(B_t)_t$ is (**SN**) satisfying Hypotheses (H1) to (H3), then we can apply the weak law of large numbers for moving averages (see [Brockwell and Davis, 1991], Prop 6.3.10) and obtain

$$\overline{B} \;\xrightarrow[n\to+\infty]{\mathbb{P}}\; \mathbb{E}(B_1) \;=\; 0\,,$$

so do converge $\sum_{k=1}^{n-h} \frac{B_k}{n}$ and $\sum_{k=1}^{n-h} \frac{B_{k+h}}{n}$. Since $\overline{B}$, $\sum_{k=1}^{n-h} \frac{B_k}{n}$, or $\sum_{k=1}^{n-h} \frac{B_{k+h}}{n}$ do multiply either themselves or polynomials with degree $\le d$, then these product terms converge $\mathbb{P}$ to 0, as soon as they are divided by n^d.

$$T_a \;=\; o_{\mathbb{P}}(n^d) \tag{A4}$$

b) Next, we study the behavior of the following terms :

$$T_{b,1} \;=\; \sum_{k=1}^{n-h} \frac{B_k\, B_{k+h}}{n}\,,$$

$$T_{b,2} \;=\; \sum_{k=1}^{n-h} \sum_{j=0}^{d} a_j \frac{k^j\, B_{k+h}}{n}\,,$$

$$T_{b,3} \;=\; \sum_{k=1}^{n-h} \sum_{j=0}^{d} a_j \frac{(k+h)^j\, B_k}{n}\,.$$

Applying again the weak law of large numbers for moving averages (see [Brockwell and Davis, 1991], Prop 7.3.5), we obtain the $\mathbb{P}$-convergence of term $T_{b,1}$

to $\gamma_B(h)$, and then

$$T_{b,1} \;=\; o_{\mathbb{P}}(n^d)$$

On the other hand, we need Cauchy-Schwarz's inequality to study terms $T_{b,2}$ and $T_{b,3}$ in the same way. We get

$$\begin{aligned}
T_{b,2} &= \sum_{j=0}^{d} a_j \sum_{k=1}^{n-h} \left(\frac{k^j}{\sqrt{n}} \times \frac{B_{k+h}}{\sqrt{n}} \right) \\
&\leq \sum_{j=0}^{d} a_j \left[\left(\sum_{k=1}^{n-h} \frac{k^{2j}}{n} \right)^{1/2} \times \left(\sum_{k=1}^{n-h} \frac{B_{k+h}^2}{n} \right)^{1/2} \right] \\
&\leq \left(\frac{n-h}{n} \sum_{k=1}^{n-h} \frac{B_{k+h}^2}{n-h} \right)^{1/2} \sum_{j=0}^{d} a_j \left(\frac{n^{2j}}{2j+1} + o(n^{2j}) \right)^{1/2} .
\end{aligned}$$

The weak law of large numbers for moving averages and Prop 7.3.5 in [Brockwell and Davis, 1991] imply the $\mathbb{P}$-convergence of the left hand term to $\gamma_B(0)^{1/2} = \sigma_B$. In addition, the right hand term is $o(n^{d+1})$. Consequently,

$$T_b \;=\; T_{b,1} + T_{b,2} + T_{b,3} \;=\; o_{\mathbb{P}}(n^{d+1}) \tag{A5}$$

c) Let us denote by T_c all the product terms involving $o(n^d)$, not studied yet. From Equation (A1), we get that $o(n^d)$ multiplies polynomials with degree $\leq d$. Then all product terms converge to 0, as soon as they are divided by n^d. Consequently, we have

$$T_c \;=\; o(n^{2d})\,. \tag{A6}$$

d) It remains to specify terms with polynomial products, in order to explicit the leading coefficient. Let us consider

$$\begin{aligned}
T_{d,1} &= \frac{a_d^2}{(d+1)^2} \, \frac{n-h}{n} \, n^{2d} , \\
T_{d,2} &= -\frac{a_d}{d+1} n^d \times \sum_{j=0}^{d} a_j \sum_{k=1}^{n-h} \frac{k^j}{n} , \\
T_{d,3} &= -\frac{a_d}{d+1} n^d \times \sum_{j=0}^{d} a_j \sum_{k=1}^{n-h} \frac{(k+h)^j}{n} , \\
T_{d,4} &= \sum_{k=1}^{n-h} \left(\sum_{i=0}^{d} a_i \frac{k^i}{\sqrt{n}} \times \sum_{j=0}^{d} a_j \frac{(k+h)^j}{\sqrt{n}} \right) .
\end{aligned}$$

All the terms $T_{d,1}$ to $T_{d,4}$ contain a leading term, associated to degree $2d$. There is nothing to do for $T_{d,1}$. Equation (A2) provides

$$\begin{aligned} T_{d,2} &= -\frac{a_d}{d+1} n^d \times \sum_{j=0}^{d} a_j \left(\frac{n^j}{j+1} + o(n^j) \right) \\ &= -\frac{a_d^2}{(d+1)^2} n^{2d} + o(n^{2d}) \,. \end{aligned}$$

We get the same formula for T_{d3}. In addition, Equation (A2) also provides

$$\begin{aligned} T_{d,4} &= \sum_{i,j=0}^{d} \frac{a_i\, a_j}{n} \sum_{k=1}^{n-h} k^i\, (k+h)^j \,, \\ &= \frac{a_d^2}{2d+1} n^{2d} + o(n^{2d}) \,. \end{aligned}$$

Finally,

$$\begin{aligned} T_d &= T_{d,1} + T_{d,2} + T_{d,3} + T_{d,4} \\ &= \left(\frac{a_d^2}{(d+1)^2} \frac{n-h}{n} - 2 \frac{a_d^2}{(d+1)^2} + \frac{a_d^2}{2d+1} \right) n^{2d} + o(n^{2d}) \end{aligned} \tag{A7}$$

Adding Equations (A4) to (A7), we obtain

$$\Gamma(h) = \left(\frac{a_d^2}{(d+1)^2} \frac{n-h}{n} - 2 \frac{a_d^2}{(d+1)^2} + \frac{a_d^2}{2d+1} \right) n^{2d} + o_{\mathbb{P}}(n^{2d})$$

Finally since $a_d \neq 0$,

$$\begin{aligned} \Xi(h) &= \frac{\Gamma(h)}{\Gamma(0)} \\ &= \frac{\left(\frac{a_d^2}{(d+1)^2} \frac{n-h}{n} - 2 \frac{a_d^2}{(d+1)^2} + \frac{a_d^2}{2d+1} \right) n^{2d} + o_{\mathbb{P}}(n^{2d})}{\left(\frac{-a_d^2}{(d+1)^2} + \frac{a_d^2}{2d+1} \right) n^{2d} + o_{\mathbb{P}}(n^{2d})} \\ &\xrightarrow[n \to +\infty]{\mathbb{P}} 1 \,. \end{aligned}$$

Appendix B. Sample autocorrelation functions behavior for (Sto,d) models – Proof of Theorem 2.3

We just give the proof for $d = 1$, since the general case can be deduced using the decomposition technique suggested in [Chan and Wei, 1988].

We first differentiate the initial series at a given lag h :

$$V_{k,h} = Z_k - Z_{k-h} = \sum_{j=0}^{h-1} B_{k-j} \,.$$

Let L denotes de lag operator i.e. $LX_t = X_{t-1}$, hence $V_{k,h}$ can be written as

$$V_{t,h} = A(L)B_t,$$

where A is the following polynomial

$$A(z) = 1 + z + ... + z^{h-1}.$$

Since $B_t = \Psi(L)\mathcal{E}_t$, where $\Psi(z) = \sum_{j\in\mathbb{Z}} b_j z^j$, it follows that

$$V_{t,h} = A(L)\Psi(L)\mathcal{E}_t = V(L)\mathcal{E}_t,$$

where

$$V(z) = A(z)\Psi(z) = \sum_{j\in\mathbb{Z}} v_j z^j, \quad \text{with } v_j = \sum_{i=0}^{h-1} b_{j-i}.$$

Consequently the process $(V_{t,h})$ is also a moving average, by straightforward calculations we can show that $(V_{t,h})$ satisfies Hypotheses (H1) to (H3).
Let us set

$$\begin{aligned} B_n(t) &= \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} B_k\,, \\ V_{n,h}(t) &= \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} V_{k,h}, \text{ for all t } \in [0,1]\,, \end{aligned}$$

where $[x]$ is the integer part of x. Then using [Boutahar, 2009] and Theorem 2 in [Davydov, 1970], we get the weak convergence :

$$\begin{aligned} B_n(.) &\xrightarrow[n\to+\infty]{D[0,1]} \sqrt{2\pi f_B(0)}W. \\ V_{n,h}(.) &\xrightarrow[n\to+\infty]{D[0,1]} \sqrt{2\pi f_V(0)}W.\,, \end{aligned}$$

where $D[0,1]$ is the set of càdlàg functions with Skorokhod topology, and where $(W_t)_t$ is a standard Brownian motion. Moreover f_B and f_V are the spectral densities associated to processes (B_t) and $(V_{t,h})$:

$$f_B(\lambda) = \frac{\sigma_{\mathcal{E}}^2}{2\pi} \left| \sum_{j\in\mathbb{Z}} b_j e^{ij\lambda} \right|^2 \tag{B1}$$

$$f_V(\lambda) = \frac{\sigma_{\mathcal{E}}^2}{2\pi} \left| A(e^{ij\lambda}) \sum_{j\in\mathbb{Z}} b_j e^{ij\lambda} \right|^2 \tag{B2}$$

We also define

$$Z_n(t) = Z_{[nt]} - \overline{Z}\,, \text{ such that } Z_n\left(\frac{k}{n}\right) = Z_k - \overline{Z}\,.$$

We recall that

$$\frac{1}{\sqrt{n}} Z_n(t) \;=\; B_n(t) - \frac{1}{n}\sum_{j=1}^{n} B_n\left(\frac{j}{n}\right)\,,$$

Then by weak convergence continuity, we obtain that

$$\frac{1}{\sqrt{n}} Z_n(.) \xrightarrow[n\to+\infty]{D[0,1]} \sqrt{2\pi f_B(0)}\; W_{1,.}\,,$$

with

$$W_{1,t} = \left(W_t - \int_0^1 W_s ds\right)\,.$$

Autocorrelation function definition was given in Equation (5). We deduce that

$$\Xi(h) \;=\; 1 + \frac{\sum\limits_{k=1}^{n-h}(Z_k - \overline{Z})\; V_{k+h,h}}{\sum\limits_{k=1}^{n}(Z_k - \overline{Z})^2} \;+\; O_{\mathbb{P}}\left(\frac{1}{n}\right)\,.$$

Then,

$$\begin{aligned}
n\;(\hat{\Xi}(h) - 1) \quad &= \quad n\;\frac{\sum\limits_{k=1}^{n-h}(Z_k - \overline{Z})\; V_{k+h,h}}{\sum\limits_{k=1}^{n}(Z_k - \overline{Z})^2} \;+\; O_{\mathbb{P}}(1) \\
&= \quad \frac{\sum\limits_{k=1}^{n-h}\frac{Z_n\left(\frac{k}{n}\right)}{\sqrt{n}}\;\left(V_n\left(\frac{k+h}{n}\right) - V_n\left(\frac{k+h-1}{n}\right)\right)}{\frac{1}{n}\sum\limits_{k=1}^{n}\left(\frac{Z_n\left(\frac{k}{n}\right)}{\sqrt{n}}\right)^2} \;+\; O_{\mathbb{P}}(1) \\
&\xrightarrow[n\to+\infty]{\mathcal{L}} \quad \frac{\sqrt{2\pi f_V(0)}}{\sqrt{2\pi f_B(0)}}\;\frac{\int_0^1 W_{1,s}\,dW_s}{\int_0^1 W_{1,s}^2\,ds} \;=\; |h|\;\frac{\int_0^1 W_{1,s}\,dW_s}{\int_0^1 W_{1,s}^2\,ds}\,.
\end{aligned}$$

Prokhorov theorem ([Prokhorov, 1956]) permits to deduce that $n\;(\hat{\Xi}(h) - 1) \;=\; O_{\mathbb{P}}(1)$.

Appendix C. Complements on KPSS statistic behavior in several particular cases – Complements to Section 2.2.4

C.1. Non-invertible MA(1) process

Let $(\mathcal{E}_t)_t$ be a white noise. We consider the following process

$$u_t \;=\; a + \mathcal{E}_t - \mathcal{E}_{t-1}\,,$$

where constant a can be equal to 0. Such process $(u_t)_t$ accounts for the differentiated series $(\Delta(Z_t))_t$ when $(Z_t)_t$ is either (**WN**) or ($\mathbf{Det}_W$,**1**). Let us note that $(u_t)_t$ is a non-invertible MA(1) process, that does not satisfy our condition (H3-b). Above all its spectral density is given by

$$f_u(\lambda) \;=\; \frac{\sigma_{\mathcal{E}}^2}{2\pi}|1-e^{i\lambda}|^2\,, \forall \lambda \in \mathbb{R}\,,$$

and hence

$$f_u(0) \;=\; 0\,. \tag{C1}$$

When developing KPSS test in [Kwiatkowski et al., 1992], the authors suppose conditions of Phillips and Perron [Phillips and Perron, 1988] on the error term u_t, among these conditions the following one on the long run variance:

$$\sigma^2 \;=\; \lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left(\left(\sum_{t=1}^{n} u_t\right)^2\right) \;=\; 2\pi\, f_u(0) > 0\,,$$

where n is the the sample size of time series u_t. But from Equation (C1) such condition is not satisfied for the non invertible MA(1) process. Thus the theoretical developments in [Kwiatkowski et al., 1992] on the asymptotic distribution of the test statistic LM_n cannot be applied here. Nevertheless doing computations in this case provides

$$LM_n \;\xrightarrow[n\to\infty]{\mathbb{P}}\; 0\,. \tag{C2}$$

As a direct consequence of Equation (C2) KPSS test will never reject the null hypothesis, as n tends to ∞, for series $(\Delta(Z_t))_t$ when $(Z_t)_t$ is either (**WN**) or ($\mathbf{Det}_W$,**1**).

C.2. ($\mathbf{Det}_W$,1) and ($\mathbf{Det}_W$,2) processes

We consider a processes $(Z_t)_t$ generated either under a ($\mathbf{Det}_W$,**1**) or a ($\mathbf{Det}_W$,**2**) model. Then the theoretical developments in [Kwiatkowski et al., 1992] on the asymptotic distribution of the test statistic LM_n cannot be applied, since $(Z_t)_t$ is generated under the alternative hypothesis. Nevertheless doing computations in these cases provides

$$LM_n \;\equiv\; n^3\,, \quad \text{as } n\to\infty\,,$$

where symbol $\equiv$ stands for the equivalence relation. Hence

$$LM_n \xrightarrow[n\to\infty]{\mathbb{P}} \infty \,. \tag{C3}$$

As a direct consequence of Equation (C3) KPSS test will always reject the null hypothesis, as n tends to ∞, for series $(Z_t)_t$ under either a (**Det**$_W$**,1**) or a (**Det**$_W$**,2**) model. Moreover this convergence remains valid if we consider the differentiated series $(\Delta(Z_t))_t$ when $(Z_t)_t$ is generated under a (**Det**$_W$**,2**) model, even if the associated noise becomes a non-invertible MA(1) process. This implies that KPSS test is powerful in all of these cases.

Appendix D. Diagram for high degree trends – Complement to Paragraph 2.2.5.3

We illustrate the strategy introduced in Paragraph 2.2.5.3, by a diagram. In addition, we indicate the R commands to use at the main steps. Note that the global approach is automated and implemented in our R-function *`trend.diag.high()`*, whose script can be found on our web page
www.i2m.univ-amu.fr/perso/manuela.royer-carenzi/AnnexesR.TrendTS/TrendTS.html

Appendix E. Boxplot of null-hypothesis rejection rate when $\sigma_\mathcal{E}$ varies – Complement to Table 2

In the main paper, Table 2 shows results for $\sigma_\mathcal{E} = 10$. Here, Figure E1 displays results when all the simulations with $\sigma_\mathcal{E}$ taking successive values in $\{0.5, 1, 3, 5, 10, 20, 30, 50, 100, 200, 300, 500\}$ are gathered. This illustrates the stability of Dickey-Fuller-testing response, as $\sigma_\mathcal{E}$ varies, for most data generating process, except for (**Det**$_W$**,1**) simulations, showing high variability.

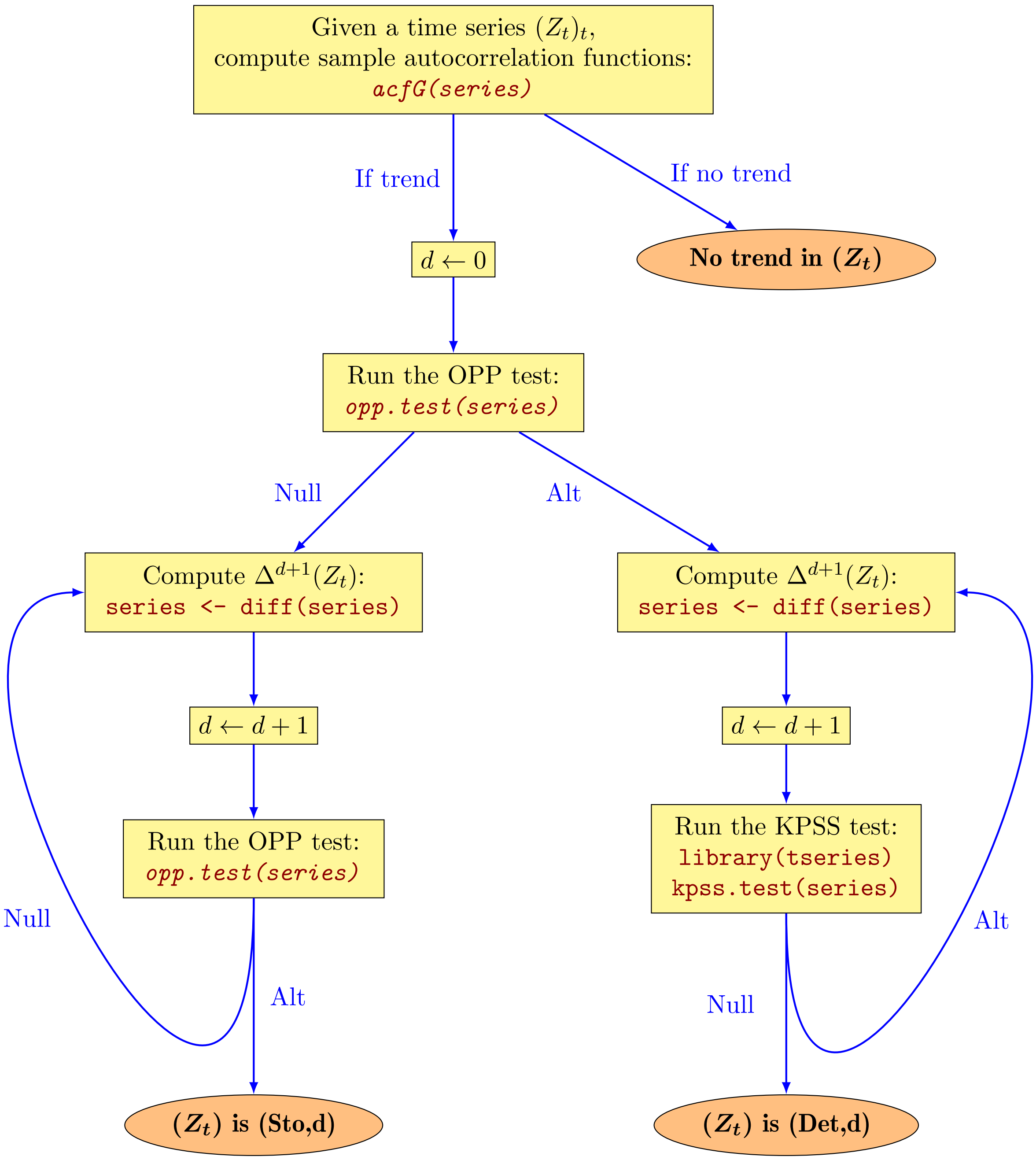

Figure D1.: TDT Strategy for high degree trends. Italic R-functions are available on our web page.

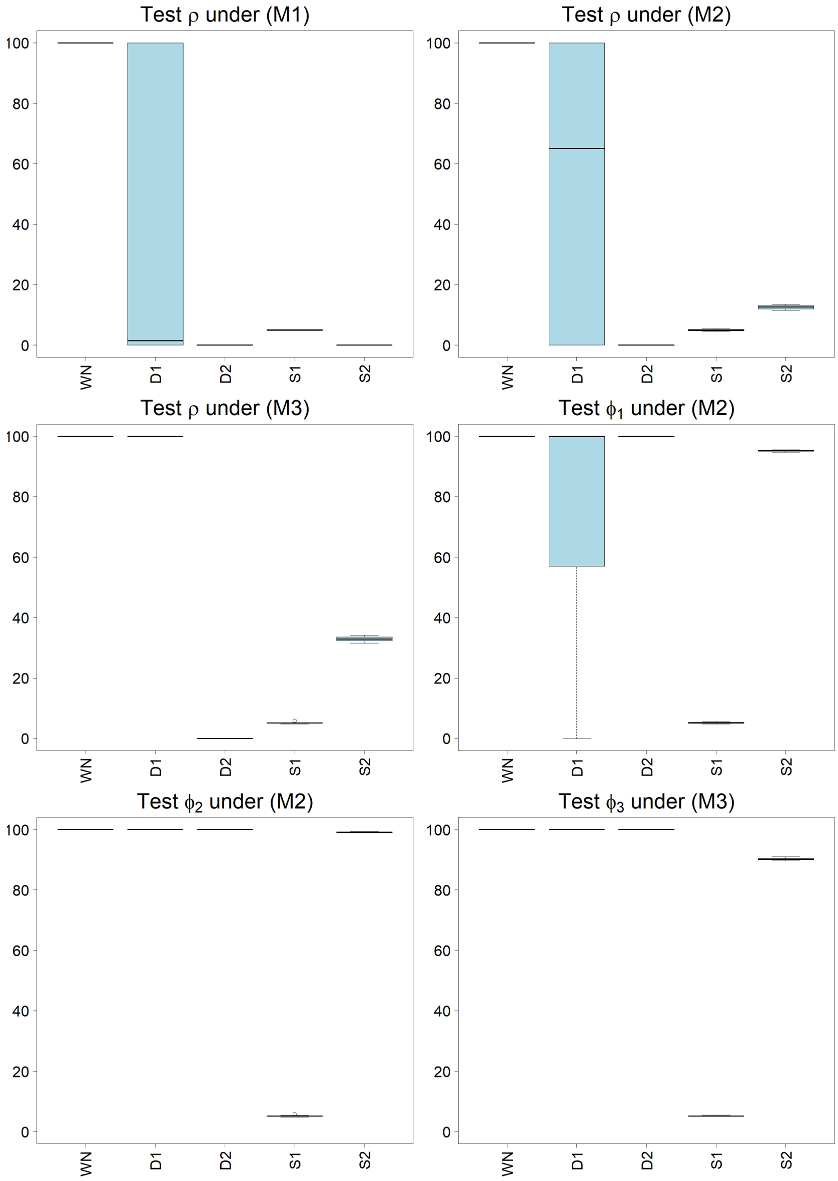

Figure E1.: Null hypothesis rejection rate for Dickey-Fuller tests, with respect to the underlying generating process used for simulations. All the simulations with $\sigma_{\mathcal{E}}$ taking successive values in $\{0.5, 1, 3,\ 5,\ 10,\ 20,\ 30,\ 50,\ 100,\ 200,\ 300,\ 500\}$ are gathered.

Appendix F. Boxplot of null-hypothesis rejection rate when $\sigma_{\mathcal{E}}$ varies, and when noise is (WN) – Complement to Table 3

In the main paper, Table 3 shows that KPSS and OPP tests perform accurately on (**WN**), (**Det**$_W$**,1**), (**Det**$_W$**,2**), (**Sto**$_W$**,1**) and (**Sto**$_W$**,2**) simulations. In Table, 3, $\sigma_{\mathcal{E}}$ successively takes values in the set $\{0.5, 1, 3,\ 5,\ 10,\ 20,\ 30,\ 50,\ 100,\ 200,\ 300,\ 500\}$ and the final rejection rate is computed by gathering all the simulations obtained for each $\sigma_{\mathcal{E}}$. Here, Figure F1 illustrates the stability of testing procedure for every data generating process as $\sigma_{\mathcal{E}}$ varies. Note that an outlier is observed when applying KPSS test to (**Det,1**) simulations. This means that KPSS test generally rejects the null hypothesis, as expected, in almost all cases. Actually, KPSS test sometimes fails to reject the null for several (**Det,1**) simulations with $\sigma_{\mathcal{E}} = 500$, that is to say when noise intensity is too high in relation to the linear coefficient a_1, so that the trend becomes imperceptible. Thus $\sigma_{\mathcal{E}} = 500$ is above the high-limit for noise intensity.

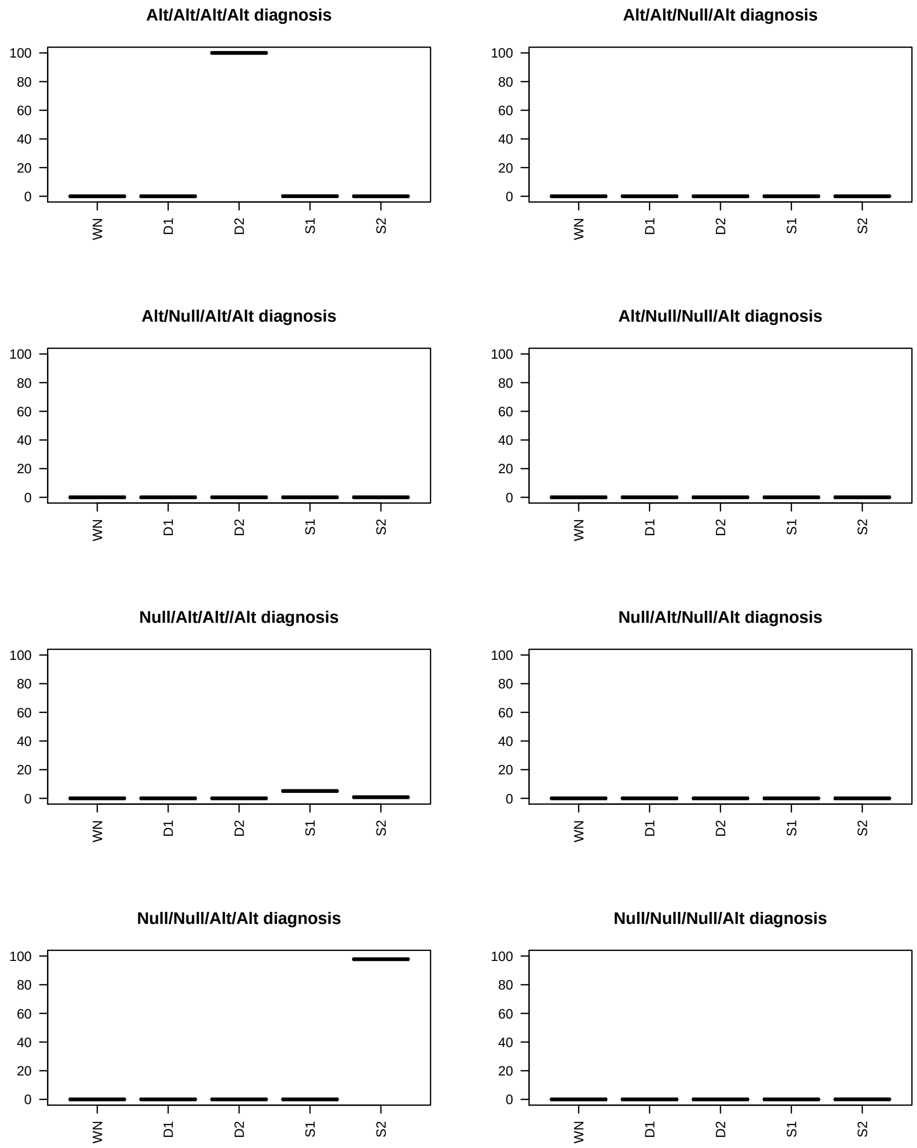

Figure F1.: Null hypothesis rejection rate for either KPSS or OPP stationarity tests applied upon either the initial or the differentiated series, with respect to the underlying generating process used for simulations. All the simulations, driven with a (**WN**), where $\sigma_{\mathcal{E}}$ takes successive values in $\{0.5, 1, 3,\ 5,\ 10,\ 20,\ 30,\ 50,\ 100,\ 200,\ 300,\ 500\}$, are gathered.