Online Supplement: Complementary Results and Proofs

for "Mixed Causal-Noncausal AR Processes and the Modelling of Explosive Bubbles"

Sébastien Fries and Jean-Michel Zakoian

This Appendix consists of six sections of additional results: K) asymptotic prediction of the MAR$(1,q)$ when $\alpha \in (0,1)$ and an explicit example in the MAR$(1,1)$ case; L) expectation of MAR(p,q) processes conditionally on a linear combination of past values and proof of the unit root property; M) conditional correlation structure of noncausal AR(1) processes, proofs of Proposition 3.3 and of the conditional variance of the MAR(1,1); N) proof of Lemma E.1; O) recursion over polynomials P_h and Q_h; P) Cluster size distribution, an illustration with the noncausal AR(1); Q) complementary results on the empirical study and details about the estimation of the excess clustering term structure; R) Complementary estimation of the financial series using the R package 'MARX'.

K A complement to Corollary 3.1 in the case $\alpha \in (0,1)$ and $q > 1$

Corollary K.1 *Under the conditions of Proposition 3.2, when $\alpha \in (0,1)$, we have almost surely*

$$\left|\mathbb{E}\left[X_{t+h} \middle| \mathscr{F}_{t-1}\right]\right| \underset{h\to+\infty}{\longrightarrow} \begin{cases} 0 & if \quad \psi^{<\alpha-1>} + \sum_{i=0}^{+\infty} a_{0,i}(\psi^{<1-\alpha>})^i = 0, \\ +\infty & else, \end{cases}$$

where the $a_{0,i}$'s are defined in Lemma E.1.

Proof.

To complete the proof of Corollary 3.1 in this case, we will derive the limit of $Q_h(\psi^{<\alpha-1>}) = \left(\psi^{<\alpha-1>}\right)^{h-1}\left[\psi^{<\alpha-1>} + \sum_{i=0}^{h-1} a_{0,i}(\psi^{<1-\alpha>})^i\right]$ when $\alpha < 1$. Recall that we have shown $a_{0,h} \underset{h\to+\infty}{\sim} Ch^{m-1}\lambda^h$.

In this case, we have $|\lambda||\psi|^{1-\alpha} < 1$, thus $\left|\psi^{<\alpha-1>} + \sum_{i=0}^{h-1} a_{0,i}(\psi^{<1-\alpha>})^i\right| \underset{i\to+\infty}{\longrightarrow} D$, where D is a nonnegative constant.

- Assume $D > 0$. Then $|Q_h(\psi^{<\alpha-1>})| \to +\infty$ as h tends to infinity, since $|\psi|^{(\alpha-1)(h-1)} \to +\infty$.
- Assume $D = 0$. We will show that $|Q_h(\psi^{<\alpha-1>})| \longrightarrow 0$.

  Indeed, we have

$$\psi^{<\alpha-1>} + \sum_{i=0}^{+\infty} a_{0,i}(\psi^{<1-\alpha>})^i = 0$$

$$\psi^{<\alpha-1>} + \sum_{i=0}^{h-1} a_{0,i}(\psi^{<1-\alpha>})^i = -\sum_{i=h}^{+\infty} a_{0,i}(\psi^{<1-\alpha>})^i.$$

Thus,

$$\begin{aligned}|Q_h(\psi^{<\alpha-1>})| &= |\psi|^{(\alpha-1)(h-1)}\Big|\psi^{<\alpha-1>} + \sum_{i=0}^{h-1} a_{0,i}(\psi^{<1-\alpha>})^i\Big| \\ &= |\psi|^{(\alpha-1)(h-1)}\Big|\sum_{i=h}^{+\infty} a_{0,i}(\psi^{<1-\alpha>})^i\Big| \\ &\leq |\psi|^{(\alpha-1)(h-1)}\sum_{i=h}^{+\infty} |a_{0,i}||\psi|^{i(1-\alpha)},\end{aligned}$$

and

$$\sum_{i=h}^{+\infty} |a_{0,i}||\psi|^{i(1-\alpha)} \underset{h\to+\infty}{\sim} |C|\sum_{i=h}^{+\infty} i^{m-1}(|\lambda||\psi|^{1-\alpha})^i.$$

We will show that for any $x \in (0,1)$, and any integer $r \geq 0$,

$$\sum_{i=h}^{+\infty} i^r x^i \underset{h\to+\infty}{\sim} h^r x^h (1-x)^{-1}, \tag{K.1}$$

which will imply

$$|\psi|^{(\alpha-1)(h-1)}\sum_{i=h}^{+\infty} |a_{0,i}||\psi|^{i(1-\alpha)} \underset{h\to+\infty}{=} O(h^{m-1}|\lambda|^h),$$

and thus $|Q_h(\psi^{<\alpha-1>})| \longrightarrow 0$, yielding the conclusion.

Let us now prove Equation (K.1). Notice that for $x \in (0,1)$, the sequences $\Big(i^r x^i\Big)_i$ and $\Big(i(i-1)\dots(i-r+1)x^i\Big)_i$ are equivalent as i tends to infinity and are both general terms of absolutely convergent series. Thus,

$$\sum_{i=h}^{+\infty} i^r x^i \underset{h\to+\infty}{\sim} \sum_{i=h}^{+\infty} i(i-1)\dots(i-r+1)x^i = x^r g^{(r)}(x),$$

where $g(x) := \sum_{i=h}^{+\infty} x^i = x^h(1-x)^{-1}$.

By the general Leibniz rule for r-times differentiable functions, we obtain

$$g^{(r)}(x) = \sum_{j=0}^{r} \frac{h!(r-j)!}{(h-j)!}\frac{x^{h-j}}{(1-x)^{r-j+1}} \underset{h\to+\infty}{\sim} \frac{h^r x^{h-r}}{1-x},$$

and thus,

$$\sum_{i=h}^{+\infty} i^r x^i \underset{h\to+\infty}{\sim} x^r\frac{h^r x^{h-r}}{1-x} = \frac{h^r x^h}{1-x}.$$

Substituting x by $|\lambda||\psi|^{1-\alpha}$ concludes the proof.

In the case $\alpha \in (0,1)$, i.e. for the heavier tails within the stable family, the absolute conditional expectation tends to $+\infty$ in modulus whenever the quantity $\psi^{<\alpha-1>} + \sum_{i=0}^{+\infty} a_{0,i}(\psi^{<1-\alpha>})^i$ does not vanish. This divergence is coherent with the fact that the unconditional expectation of (X_t) does not exist when $\alpha < 1$. It would be striking to have a case for which the above quantity is exactly zero, which would imply that the conditional expectation vanishes even for this class of particularly extreme processes. However, as the following example shows, all MAR(1,1) feature diverging conditional expectation when $\alpha < 1$.

Example K.1 (Asymptotic predictions of the MAR(1,1) process) Let (X_t) be defined by Equation (3.3). From the explicit predictions formulated in Section 3.4, we deduce the asymptotic equivalents as the horizon h tends to infinity:

$$\mathbb{E}\Big[X_{t+h}\Big|\mathscr{F}_{t-1}\Big] \underset{h\to+\infty}{\overset{a.s.}{\sim}} \begin{cases} \dfrac{(\psi^{<\alpha-1>})^{h+1}}{1-\psi^{<1-\alpha>}\phi}\left(X_{t-1}-\phi X_{t-2}\right), & \text{if} \quad |\phi| \quad < |\psi|^{\alpha-1}, \\ \dfrac{\phi^{h+2}}{\phi-\psi^{<\alpha-1>}}\Big(X_{t-1}-\psi^{<\alpha-1>}X_{t-2}\Big), & \text{if} \quad |\phi| \quad > |\psi|^{\alpha-1}, \\ \phi^{h+1}\left(X_{t-1}-\dfrac{1+(-1)^h}{2}(X_{t-1}-\phi X_{t-2})\right), & \text{if} \quad \phi \quad = -\psi^{<\alpha-1>}, \\ (h+1)\phi^{h+1}(X_{t-1}-\phi X_{t-2}), & \text{if} \quad \phi \quad = \psi^{<\alpha-1>}. \end{cases}$$

Noticing that the condition $|\phi| < |\psi|^{\alpha-1}$ is equivalent to $\alpha < 1 + \frac{\ln|\phi|}{\ln|\psi|}$, with $\frac{\ln|\phi|}{\ln|\psi|} > 0$, it can be seen that the three asymptotic limits of Corollary 3.1 are consistent with these equivalents. In particular, when $\alpha = 1$, we always have $|\phi| < 1 = |\psi|^{\alpha-1}$ and we get that, almost surely,

$$\Big|\mathbb{E}\left[X_{t+h}\right| X_{t-1}, X_{t-2}]\Big| \underset{h\to+\infty}{\longrightarrow} \ell_{t-1} = \left|\frac{X_{t-1}-\phi X_{t-2}}{1-\operatorname{sign}(\psi)\phi}\right|.$$

L Unit root property and extension

The equality $\mathbb{E}\big[X_t\big|X_{t-1}\big] = X_{t-1}$ for the noncausal Cauchy AR(1) with positive AR coefficient shows the existence of a unit root. Indeed, we have $X_t = X_{t-1} + \eta_t$ where $\mathbb{E}\big[\eta_t\big|X_{t-1}\big] = 0$. We show in this section that this property actually extends to more general MAR processes. The next result provides the conditional expectation of X_t given X_{t-1}.

Proposition L.1 *Let X_t be the MAR(p,q) process solution of* (2.1) *with symmetric α-stable errors,* $1 < \alpha < 2$. *Denoting (d_k) the coefficients sequence of its MA(∞) representation, we have*

$$\mathbb{E}\left[X_t\right| X_{t-1}] = \frac{\sum_{k\in\mathbb{Z}} d_k\,(d_{k+1})^{<\alpha-1>}}{\sum_{k\in\mathbb{Z}} |d_{k+1}|^{\alpha}} X_{t-1}.$$

The condition for the existence of a unit root is now straightforward.

Corollary L.1 *Under the assumptions of Proposition L.1,*

$$\mathbb{E}\left[X_t\right| X_{t-1}] = X_{t-1} \qquad \Longleftrightarrow \qquad \sum_{k\in\mathbb{Z}} d_k\,(d_{k+1})^{<\alpha-1>} = \sum_{k\in\mathbb{Z}} |d_{k+1}|^{\alpha}.$$

The case $\alpha \leq 1$ is more intricate because the expectation on the left-hand side of (L.1) might not exist. However, the conditions for existence can be established using Theorem 2.13 of Samorodnistky and Taqqu (1994). This is left for further research. Proposition L.1 is a consequence of the more general conditional expectation of X_t given any linear combination of the past that we provide in the next result.

Proposition L.2 *Let X_t be the MAR(p,q) process solution of* (2.1) *with symmetric α-stable errors,* $1 < \alpha < 2$. *Denote (d_k) the coefficients sequence of its MA(∞) representation. Then for any $h \geq 0$, $k \geq 1$, and*

$a_1, \ldots, a_k$ *such that there exists* $\ell \in \mathbb{Z}$, $a_1 d_{\ell+1} + \ldots + a_k d_{\ell+k} \neq 0$, *we have*

$$\mathbb{E}\Big[X_{t+h}\Big|\sum_{j=1}^{k} a_j X_{t-j}\Big] = \frac{\sum_{\ell\in\mathbb{Z}} d_{\ell-h}\left(\sum_{j=1}^{k} a_j d_{\ell+j}\right)^{<\alpha-1>}}{\sum_{\ell\in\mathbb{Z}}\left|\sum_{j=1}^{k} a_j d_{\ell+j}\right|^{\alpha}}(a_1 X_{t-1} + \ldots + a_k X_{t-k}). \tag{L.1}$$

Proposition L.1 is obtained for $k = 1$, $a_1 = 1$.

Proof.

Let us introduce $Y_{t-1,k} = a_1 X_{t-1} + \ldots + a_k X_{t-k}$. Let $\varphi(u,v) = \mathbb{E}\left[e^{iuY_{t-1,k}+ivX_{t+h}}\right]$. For any $(u,v) \in \mathbb{R}^2$ we have,

$$\begin{aligned}
\varphi(u,v) &= \mathbb{E}\left[\exp\left\{iu\sum_{j=1}^{k} a_j \sum_{\ell\in\mathbb{Z}} d_\ell \varepsilon_{t+\ell-j} + v\sum_{\ell\in\mathbb{Z}} d_\ell \varepsilon_{t+\ell+h}\right\}\right] \\
&= \mathbb{E}\left[\exp\left\{i\sum_{\ell\in\mathbb{Z}}\left(u\sum_{j=1}^{k} a_j d_{\ell+j} + v d_{\ell-h}\right)\varepsilon_{t+\ell}\right\}\right] \\
&= \exp\left\{-\sigma^\alpha \sum_{\ell\in\mathbb{Z}}\left|u\sum_{j=1}^{k} a_j d_{\ell+j} + v d_{\ell-h}\right|^{\alpha}\right\}.
\end{aligned}$$

Thus,

$$\frac{\partial\varphi}{\partial u}(u,v) = -\alpha\sigma^\alpha\varphi(u,v)\sum_{\ell\in\mathbb{Z}}\left(\sum_{j=1}^{k} a_j d_{\ell+j}\right)\left(u\sum_{j=1}^{k} a_j d_{\ell+j} + v d_{\ell-h}\right)^{<\alpha-1>},$$

and

$$\left.\frac{\partial\varphi}{\partial u}\right|_{v=0} = -\alpha\sigma^\alpha u^{<\alpha-1>}\varphi(u,0)\sum_{\ell\in\mathbb{Z}}\left|\sum_{j=1}^{k} a_j d_{\ell+j}\right|^{\alpha}.$$

We also have

$$\begin{aligned}
\frac{\partial\varphi}{\partial v}(u,v) &= -\alpha\sigma^\alpha\varphi(u,v)\sum_{\ell\in\mathbb{Z}} d_{\ell-h}\left(u\sum_{j=1}^{k} a_j d_{\ell+j} + v d_{\ell-h}\right)^{<\alpha-1>}, \\
\left.\frac{\partial\varphi}{\partial v}\right|_{v=0} &= -\alpha\sigma^\alpha u^{<\alpha-1>}\varphi(u,0)\sum_{\ell\in\mathbb{Z}} d_{\ell-h}\left(\sum_{j=1}^{k} a_j d_{\ell+j}\right)^{<\alpha-1>}.
\end{aligned}$$

Therefore,

$$\left.\frac{\partial\varphi}{\partial v}\right|_{v=0} = \frac{\sum_{\ell\in\mathbb{Z}} d_{\ell-h}\left(\sum_{j=1}^{k} a_j d_{\ell+j}\right)^{<\alpha-1>}}{\sum_{\ell\in\mathbb{Z}}\left|\sum_{j=1}^{k} a_j d_{\ell+j}\right|^{\alpha}}\left.\frac{\partial\varphi}{\partial u}\right|_{v=0} \tag{L.2}$$

On the other hand, for $u \neq 0$:

$$\left.\frac{\partial\varphi}{\partial u}\right|_{v=0} = i\mathbb{E}\left[Y_{t-1,k}e^{iuY_{t-1,k}}\right], \quad \left.\frac{\partial\varphi}{\partial v}\right|_{v=0} = i\mathbb{E}\left[X_{t+h}e^{iuY_{t-1,k}}\right].$$

Therefore, for $u \in \mathbb{R}^*$:

$$\mathbb{E}\left[\left(X_{t+h} - \frac{\sum_{\ell\in\mathbb{Z}} d_{\ell-h}\left(\sum_{j=1}^k a_j d_{\ell+j}\right)^{<\alpha-1>}}{\sum_{\ell\in\mathbb{Z}}\left|\sum_{j=1}^k a_j d_{\ell+j}\right|^{\alpha}} Y_{t-1,k}\right) e^{iuY_{t-1,k}}\right] = 0. \tag{L.3}$$

Hence, from Bierens (Theorem 1, 1982): Thus

$$\mathbb{E}\left[X_{t+h}\middle| Y_{t-1,k}\right] = \frac{\sum_{\ell\in\mathbb{Z}} d_{\ell-h}\left(\sum_{j=1}^k a_j d_{\ell+j}\right)^{<\alpha-1>}}{\sum_{\ell\in\mathbb{Z}}\left|\sum_{j=1}^k a_j d_{\ell+j}\right|^{\alpha}} Y_{t-1,k}.$$

M Conditional heteroscedasticity of the MAR$(1, q)$ process

In order to prove Proposition 3.3, we need to show some preliminary results about the conditional covariance of noncausal AR(1) processes. We will then turn to the conditional covariance of a MAR$(1, q)$ process from which the conditional variance will be obtained.

M.1 Conditional correlation structure of the MAR$(1, q)$

Lemma M.1 *Let X_t be a noncausal AR(1) process satisfying $X_t = \psi X_{t+1} + \varepsilon_t$, with $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(1, 0, \sigma, 0)$. Then, for any nonnegative integers h and τ:*

$$\mathbb{E}\Big[X_{t+h}X_{t+h+\tau}\Big|X_{t-1}\Big] = (sign\ \psi)^{\tau}\left[|\psi|^{-h-1}\left(X_{t-1}^2 + \frac{\sigma^2}{(1-|\psi|)^2}\right) - \frac{\sigma^2}{(1-|\psi|)^2}\right].$$

Remark M.1 From the previous result, it is possible to derive the whole conditional correlation structure of (X_t). It can be shown that for any $t \in \mathbb{Z}$, and any positive integers h and τ:

$$\frac{\text{Cov}\Big(X_{t+h}, X_{t+h+\tau}\Big|X_{t-1}\Big)}{\sqrt{\mathbb{V}\Big(X_{t+h}\Big|X_{t-1}\Big)}\sqrt{\mathbb{V}\Big(X_{t+h+\tau}\Big|X_{t-1}\Big)}} = (\text{sign}\ \psi)^{\tau}\sqrt{\frac{|\psi|^{-h-1}-1}{|\psi|^{-h-\tau-1}-1}},$$

which, when $\tau \to +\infty$, is asymptotically equivalent to $(\text{sign}\ \psi)^{\tau}|\psi|^{\tau/2}\sqrt{1-|\psi|^{h+1}}$ for any $h \geq 0$, and to $(\text{sign}\ \psi)^{\tau}|\psi|^{\tau/2}$ when h becomes large. Although in our infinite variance framework, the unconditional correlation is not defined, empirical correlations can always be computed. We know from Davis and Resnick (1985,1986) that they converge in probability towards the theoretical autocorrelations that would prevail in the L^2 framework. Given n observations of process (X_t), we have for any $\tau \geq 0$,

$$\frac{\sum_{t=1}^{n-\tau+1} X_t X_{t+\tau}}{\sum_{t=1}^{n} X_t^2} \xrightarrow[n\to+\infty]{p} \psi^{\tau}.$$

Surprisingly, the "unconditional" autocorrelations of (X_t) do not converge to the conditional ones when $n \to +\infty$, and vanish at a much slower rate ($|\psi|^{\tau/2}$ instead of $|\psi|^{\tau}$).

We now turn to the MAR$(1, q)$ process.

Proposition M.1 *Let X_t be a MAR(1, q) process, $q \geq 0$, solution of Equation (2.1) with $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(1, 0, \sigma, 0)$. Then, for any positive integers h and τ, there exist polynomials P_h, $P_{h+\tau}$, both of degrees $q-1$, and Q_h, $Q_{h+\tau}$ of respective degrees h and $h+\tau$ such that*

$$\begin{aligned}\mathbb{E}\left[X_{t+h}X_{t+h+\tau} \,|\, \mathscr{F}_{t-1}\right] &= (P_h(B)X_{t-1})(P_{h+\tau}(B)X_{t-1}) \\ &+ sign(\psi)(\phi(B)X_{t-1})\Big[(P_h(B)X_{t-1})Q_{h+\tau}(sign\ \psi) + (P_{h+\tau}(B)X_{t-1})Q_h(sign\ \psi)\Big] \\ &+ c_{h,\tau}\left((\phi(B)X_{t-1})^2 + \frac{\sigma^2}{(1-|\psi|)^2}\right) - \frac{\sigma^2}{(1-|\psi|)^2}Q_h(sign\ \psi)Q_{h+\tau}(sign\ \psi),\end{aligned}$$

with $c_{h,\tau} = \sum_{i=0}^{h+\tau}\sum_{j=0}^{h} q_{i,h+\tau}q_{j,h}(sign\ \psi)^{i+j}|\psi|^{-\min(i,j)-1}$ and $Q_k(z) = \sum_{i=0}^{k} q_{i,k}z^i$, for any $k \geq 0$.

This result yields Proposition 3.2 by taking $h = \tau = 0$, with $P_0(B) = \phi_1 + \phi_2 B + \ldots + \phi_q B^q$ and $Q_0(B) = 1$.

M.2 Proof of Lemma M.1

Consider $\varphi(x,y,z) := \mathbb{E}\Big(e^{ixX_{t+k}+iyX_{t+\ell}+izX_{t-1}}\Big)$, with $0 \leq \ell \leq k$, $X_t = \psi X_{t+1} + \varepsilon_t$ and $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, 0, \sigma, 0)$. We have

$$\varphi(x,y,z) = \mathbb{E}\left(e^{i\sum_{n\in\mathbb{Z}}(xd_{n-k}+yd_{n-\ell}+zd_{n+1})\varepsilon_{t+n}}\right) = \exp\left\{-\sigma^\alpha \sum_{n\in\mathbb{Z}} |xd_{n-k} + yd_{n-\ell} + zd_{n+1}|^\alpha\right\}.$$

Thus, on the one hand,

$$\begin{aligned}\frac{\partial\varphi}{\partial z} &= -\alpha\sigma^\alpha \sum_{n\in\mathbb{Z}} d_{n+1}(xd_{n-k} + yd_{n-\ell} + zd_{n+1})^{<\alpha-1>}\varphi(x,y,z), \\ \frac{\partial^2\varphi}{\partial z^2} &= (\alpha\sigma^\alpha)^2\left(\sum_{n\in\mathbb{Z}} d_{n+1}(xd_{n-k} + yd_{n-\ell} + zd_{n+1})^{<\alpha-1>}\right)^2 \varphi(x,y,z) \\ &\quad - \alpha(\alpha-1)\sum_{n\in\mathbb{Z}} d_{n+1}^2|xd_{n-k} + yd_{n-\ell} + zd_{n+1}|^{\alpha-2}\varphi(x,y,z), \\ \left.\frac{\partial^2\varphi}{\partial z^2}\right|_{\substack{x=0\\y=0}} &= (\alpha\sigma^\alpha)^2|z|^{2(\alpha-1)}\left(\sum_{n\in\mathbb{Z}}|d_{n+1}|^\alpha\right)^2\varphi(0,0,z) - \alpha(\alpha-1)|z|^{\alpha-2}\sum_{n\in\mathbb{Z}}|d_{n+1}|^\alpha\varphi(0,0,z).\end{aligned}$$

And on the other hand,

$$
\begin{aligned}
\frac{\partial \varphi}{\partial y} &= -\alpha\sigma^\alpha \sum_{n\in\mathbb{Z}} d_{n-\ell}(xd_{n-k}+yd_{n-\ell}+zd_{n+1})^{<\alpha-1>}\varphi(x,y,z),\\
\frac{\partial^2 \varphi}{\partial x\partial y} &= (\alpha\sigma^\alpha)^2 \left(\sum_{n\in\mathbb{Z}} d_{n-\ell}(xd_{n-k}+yd_{n-\ell}+zd_{n+1})^{<\alpha-1>}\right)\\
&\qquad\qquad \times \left(\sum_{n\in\mathbb{Z}} d_{n-k}(xd_{n-k}+yd_{n-\ell}+zd_{n+1})^{<\alpha-1>}\right)\varphi(x,y,z)\\
&\quad - \alpha(\alpha-1)\sum_{n\in\mathbb{Z}} d_{n-\ell}d_{n-k}|xd_{n-k}+yd_{n-\ell}+zd_{n+1}|^{\alpha-2}\varphi(x,y,z),\\
\left.\frac{\partial^2 \varphi}{\partial x\partial y}\right|_{\substack{x=0\\y=0}} &= (\alpha\sigma^\alpha)^2|z|^{2(\alpha-1)}\left(\sum_{n\in\mathbb{Z}} d_{n-\ell}(d_{n+1})^{<\alpha-1>}\right)\left(\sum_{n\in\mathbb{Z}} d_{n-k}(d_{n+1})^{<\alpha-1>}\right)\varphi(0,0,z)\\
&\quad - \alpha(\alpha-1)|z|^{\alpha-2}\sum_{n\in\mathbb{Z}} d_{n-\ell}d_{n-k}|d_{n+1}|^{\alpha-2}\varphi(0,0,z).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\frac{1}{A_2}\left[\left.\frac{\partial^2 \varphi}{\partial x\partial y}\right|_{\substack{x=0\\y=0}} - (\alpha\sigma^\alpha)^2A_1|z|^{2(\alpha-1)}\varphi(0,0,z)\right] &= -\alpha(\alpha-1)|z|^{\alpha-2}\varphi(0,0,z),\\
\frac{1}{A_3}\left[\frac{\partial^2 \varphi}{\partial z^2} - (\alpha\sigma^\alpha)^2A_3^2|z|^{2(\alpha-1)}\varphi(0,0,z)\right] &= -\alpha(\alpha-1)|z|^{\alpha-2}\varphi(0,0,z),
\end{aligned}
$$

with

$$
\begin{aligned}
A_1 &= \left(\sum_{n\in\mathbb{Z}} d_{n-\ell}(d_{n+1})^{<\alpha-1>}\right)\left(\sum_{n\in\mathbb{Z}} d_{n-k}(d_{n+1})^{<\alpha-1>}\right),\\
A_2 &= \sum_{n\in\mathbb{Z}} d_{n-\ell}d_{n-k}|d_{n+1}|^{\alpha-2},\\
A_3 &= \sum_{n\in\mathbb{Z}} |d_{n+1}|^{\alpha}.
\end{aligned}
$$

Therefore,

$$
\frac{1}{A_2}\left[\left.\frac{\partial^2 \varphi}{\partial x\partial y}\right|_{\substack{x=0\\y=0}} - (\alpha\sigma^\alpha)^2A_1|z|^{2(\alpha-1)}\varphi(0,0,z)\right] = \frac{1}{A_3}\left[\frac{\partial^2 \varphi}{\partial z^2} - (\alpha\sigma^\alpha)^2A_3^2|z|^{2(\alpha-1)}\varphi(0,0,z)\right],
$$

This yields for $\alpha = 1$,

$$
\frac{1}{A_2}\left[\left.\frac{\partial^2 \varphi}{\partial x\partial y}\right|_{\substack{x=0\\y=0}} - \sigma^2A_1\varphi(0,0,z)\right] = \frac{1}{A_3}\left[\frac{\partial^2 \varphi}{\partial z^2} - \sigma^2A_3^2\varphi(0,0,z)\right].
$$

Taking into account that $d_n = \psi^n \mathbb{1}_{\{n\geq 0\}}$ for the noncausal AR(1) and noticing that

$$
\begin{aligned}
\frac{\partial^2 \varphi}{\partial x\partial y} &= -\mathbb{E}\left[X_{t+k}X_{t+\ell}e^{izX_{t-1}}\right],\\
\frac{\partial^2 \varphi}{\partial z^2} &= -\mathbb{E}\left[X_{t-1}^2e^{izX_{t-1}}\right],
\end{aligned}
$$

we get for any $z \in \mathbb{R}^*$:

$$\mathbb{E}\left[\left\{X_{t+kX_{t+\ell}} - (\text{sign } \psi)^{k+\ell}\left(|\psi|^{-\ell-1}(X_{t-1}^2 + \tilde{\sigma}^2) - \tilde{\sigma}^2\right)\right\} e^{izX_{t-1}}\right] = 0,$$

with $\tilde{\sigma} = \dfrac{\sigma}{1-|\psi|}$. From Bierens (Theorem 1, 1982):

$$\mathbb{E}\left[X_{t+k}X_{t+\ell}\middle|X_{t-1}\right] = (\text{sign } \psi)^{k+\ell}\left(|\psi|^{-\ell-1}(X_{t-1}^2 + \tilde{\sigma}^2) - \tilde{\sigma}^2\right),$$

which concludes the proof.

M.3 Proof of Proposition M.1

Let k and ℓ be two positive integers such that $\ell \leq k$. From Lemma D.1, we know that for any $h \geq 0$, there exist two polynomials P_h and Q_h of respective degrees $q-1$ and h such that:

$$X_{t+h} = P_h(B)X_{t-1} + Q_h(F)u_t.$$

Thus, using the same device as in the Proof of Proposition 3.2,

$$\begin{aligned}
\mathbb{E}\Big[X_{t+k}X_{t+\ell}\Big|X_{t-1},\ldots,X_{t-q-1}\Big] &= \mathbb{E}\Big[\Big(P_k(B)X_{t-1} + Q_k(F)u_t\Big)\Big(P_\ell(B)X_{t-1} + Q_\ell(F)u_t\Big)\Big|X_{t-1},\ldots,X_{t-q-1}\Big],\\
&= \Big(P_k(B)X_{t-1}\Big)\Big(P_\ell(B)X_{t-1}\Big)\\
&\quad + \Big(P_k(B)X_{t-1}\Big)\mathbb{E}\Big[Q_\ell(F)u_t\Big|u_{t-1}\Big] + \Big(P_\ell(B)X_{t-1}\Big)\mathbb{E}\Big[Q_k(F)u_t\Big|u_{t-1}\Big]\\
&\quad + \sum_{i=0}^{k}\sum_{j=0}^{\ell} q_i q_j \mathbb{E}\Big[u_{t+i}u_{t+j}\Big|u_{t-1}\Big].
\end{aligned}$$

The second and third terms can be expressed as:

$$\begin{aligned}
&\Big(P_k(B)X_{t-1}\Big)\mathbb{E}\Big[Q_\ell(F)u_t\Big|u_{t-1}\Big] + \Big(P_\ell(B)X_{t-1}\Big)\mathbb{E}\Big[Q_k(F)u_t\Big|u_{t-1}\Big] =\\
&\qquad \text{sign}(\psi)\Big(\phi(B)X_{t-1}\Big)\Big[Q_\ell(\text{sign } \psi)\Big(P_k(B)X_{t-1}\Big) + Q_k(\text{sign } \psi)\Big(P_\ell(B)X_{t-1}\Big)\Big],
\end{aligned}$$

whereas the fourth term can be rewritten using Lemma M.1:

$$\begin{aligned}
\sum_{i=0}^{k}\sum_{j=0}^{\ell} q_i q_j \mathbb{E}\Big[u_{t+i}u_{t+j}\Big|u_{t-1}\Big] &= \sum_{i=0}^{k}\sum_{j=0}^{\ell} q_i q_j (\text{sign } \psi)^{i+j}\left[|\psi|^{-\min(i,j)-1}\Big((\phi(B)X_{t-1})^2 + \tilde{\sigma}^2\Big) - \tilde{\sigma}^2\right],\\
&= -\tilde{\sigma}^2 Q_k(\text{sign } \psi)Q_\ell(\text{sign } \psi)\\
&\quad + \Big((\phi(B)X_{t-1})^2 + \tilde{\sigma}^2\Big)\sum_{i=0}^{k}\sum_{j=0}^{\ell} q_i q_j (\text{sign } \psi)^{i+j}|\psi|^{-\min(i,j)-1}.
\end{aligned}$$

M.4 Proof of Proposition 3.3

The result of Proposition 3.3 is obtained by substituting $\mathbb{E}\Big[X_{t+h}\Big|\mathscr{F}_{t-1}\Big]$ and $\mathbb{E}\Big[X_{t+h}^2\Big|\mathscr{F}_{t-1}\Big]$ in

$$\mathbb{V}\Big(X_{t+h}\Big|\mathscr{F}_{t-1}\Big) = \mathbb{E}\Big[X_{t+h}^2\Big|\mathscr{F}_{t-1}\Big] - \Big(\mathbb{E}\Big[X_{t+h}\Big|\mathscr{F}_{t-1}\Big]\Big)^2,$$

using the formulas of Propositions 3.2 and M.1.

M.5 Details on the conditional variance of the MAR(1,1) of Section 3.4

By Lemma E.1, the polynomial Q_h intervening in Proposition 3.3 reads in the case of the MAR(1,1)

$$Q_h(z) = \sum_{i=0}^{h} \phi^{h-i} z^i.$$

Applying Proposition 3.3, we know that

$$\mathbb{V}\Big(X_{t+h}\Big|\mathscr{F}_{t-1}\Big) = \left((X_{t-1} - \phi X_{t-2})^2 + \frac{\sigma^2}{(1-|\psi|)^2}\right)\left(c_h - \Big(Q_h(\text{sign } \psi)\Big)^2\right),$$

with $c_h = \sum_{i=0}^{h}\sum_{j=0}^{h} q_{i,h} q_{j,h} (\text{sign } \psi)^{i+j} |\psi|^{-\min(i,j)-1}$. Using the explicit form of the $q_{i,h}$'s, the coefficients of polynomial Q_h, we can deduce that for $\psi > 0$

$$Q_h(\text{sign}\,\psi) = \frac{1-\phi^{h+1}}{1-\phi},$$
$$c_h = \psi^{-h-1} \sum_{i=0}^{h}\sum_{j=0}^{h} \phi^i \phi^j \psi^{\max(i,j)},$$

which can be simplified by elementary calculations after splitting the sums according to whether $i \geq j$ or $j > i$.

N Proof of Lemma E.1

For $h = 0$, Equation (D.1) holds with $P_0(B) = \phi_1 + \phi_2 B^2 \ldots + \phi_q B^{q-1}$ and $Q_0(B) = 1$. We have

$$\begin{aligned}
X_{t+h} &= a_{0,h} X_{t-1} + \sum_{i=1}^{q-1} a_{i,h} X_{t-i-1} + \sum_{i=0}^{h} b_{i,h} u_{t+i} \\
&= a_{0,h} \left(\sum_{i=0}^{q-1} \phi_{i+1} X_{t-i-2} + u_{t-1} \right) + \sum_{i=1}^{q-1} a_{i,h} X_{t-i-1} + \sum_{i=0}^{h} b_{i,h} u_{t+i} \\
&= \sum_{i=0}^{q-2} \Big(a_{i+1,h} + a_{0,h}\phi_{i+1} \Big) X_{t-i-2} + a_{0,h}\phi_q X_{t-q-1} + a_{0,h} u_{t-1} + \sum_{i=0}^{h} b_{i,h} u_{t+i}.
\end{aligned}$$

Since this last formula holds at any $t \in \mathbb{Z}$, this last equation yields

$$X_{t+h+1} = \sum_{i=0}^{q-2} \Big(a_{i+1,h} + a_{0,h}\phi_{i+1} \Big) X_{t-i-1} + a_{0,h}\phi_q X_{t-q} + a_{0,h} u_t + \sum_{i=1}^{h+1} b_{i-1,h} u_{t+i}.$$

However, we also have by definition

$$X_{t+h+1} = P_{h+1}(B) X_{t-1} + Q_{h+1}(F) u_t = \sum_{i=0}^{q-1} a_{i,h+1} X_{t-i-1} + \sum_{i=0}^{h+1} b_{i,h+1} u_{t+i}.$$

Thus, by identification,

$$\begin{aligned}
a_{q-1,h+1} &= a_{0,h}\phi_q, \\
a_{i,h+1} &= a_{i+1,h} + a_{0,h}\phi_{i+1}, \quad \text{for} \quad 0 \leq i \leq q-2, \\
a_{0,h} &= b_{0,h+1}, \\
b_{i,h+1} &= b_{i-1,h}, \quad \text{for} \quad 1 \leq i \leq h+1.
\end{aligned}$$

We deduce from these equations that for any $h \geq 0$,

$$b_{i,h+1} = a_{0,h-i}, \quad \text{for} \quad 0 \leq i \leq h+1,$$
$$a_{i,h+1} = \sum_{j=0}^{\min(q-i-1,h)} a_{0,h-j}\phi_{i+1+j}, \quad \text{for} \quad 0 \leq i \leq q-1,$$

with the convention $a_{0,-1} = 1$. We obtain that $(a_{0,h})$ is the solution of the linear recurrent equation of order q

$$a_{0,h+q} = \phi_1 a_{0,h+q-1} + \ldots + \phi_q a_{0,h}, \quad \text{for} \quad h \geq 0, \tag{N.1}$$

with initial values $(a_{0,0}, \ldots, a_{0,q-1})$ that could be expressed as functions of $\phi_1, \ldots, \phi_q$. Denote $\lambda_1, \ldots, \lambda_s$ the distinct roots of the polynomial $F^q\phi(B)$ with respective multiplicities $m_1, \ldots, m_s$, with $s \leq q$, $m_1 + \ldots + m_s = q$. Since ϕ has all its roots outside the unit circle, we know that $|\lambda_i| < 1$ for all i. Therefore, there exist polynomials $C_1, \ldots, C_q$ of respective degrees $m_1, \ldots, m_s$ such that for any $h \geq q$,

$$a_{0,h} = C_1(h)\lambda_1^h + \ldots + C_s(h)\lambda_s^h.$$

O A recursive scheme for computing polynomials P_h and Q_h of Lemma D.1

Lemma O.1 *Polynomials P_h and Q_h of Lemma D.1 satisfy the following recursive equations:*

$$BP_{h+1}(B) = P_h(B) - P_h(0)\phi(B), \qquad Q_{h+1}(F) = FQ_h(F) + P_h(0), \tag{O.1}$$

with initial conditions $Q_0(B) = 1$, $P_0(B) = \phi_1 + \phi_2 B + \ldots + \phi_q B^{q-1}$.

Proof. By applying polynomial $\phi(B)$ to (D.1), we get by (B.1)

$$\phi(B)X_{t+h} = P_h(B)\phi(B)X_{t-1} + Q_h(F)\phi(B)u_t,$$
$$B^{-h}u_t = BP_h(B)u_t + Q_h(F)\phi(B)u_t,$$

which implies $B^{h+1}P_h(B) + B^h Q_h(F)\phi(B) = 1$. The same holds at rank $h+1$. Thus, denoting $Q_h(F) = \sum_{i=0}^h q_{i,h}F^i$ and $Q_h^*(B) := B^h Q_h(F) = \sum_{i=0}^h q_{h-i,h}B^i$, we also have: $B^{h+2}P_{h+1}(B) + Q_{h+1}(B)\psi^*(B)\phi(B) = 1$. Subtracting the expressions at ranks h and $h+1$ yields:

$$B^{h+1}\Big(BP_{h+1}(B) - P_h(B)\Big) + \phi(B)\Big(Q_{h+1}^*(B) - Q_h^*(B)\Big) = 0. \tag{O.2}$$

We can notice that the term of degree zero in this expression is: $\phi(0)\Big(Q_{h+1}^*(0) - Q_h^*(0)\Big) = 0$, hence $q_{h+1,h+1} = q_{h,h}$. Focusing on the next terms of degrees $i = 1, \ldots, h$, we can iteratively show that $q_{h+1-i,h+1} = q_{h-i,h}$. Finally, focusing on the term of degree $h+1$, we now deduce that $-P_h(0) + q_{1,h+1} - q_{0,h} = 0$. This leads us to the equality

$$Q_{h+1}^*(B) = Q_h^*(B) + B^{h+1}P_h(0), \tag{O.3}$$

or equivalently $Q_{h+1}(F) = FQ_h(F)+P_h(0)$, which establishes the right-hand side equation of (O.1). Finally, replacing (O.3) in (O.2) concludes the proof of Lemma O.1.

P Cluster size distribution: the noncausal AR(1) case

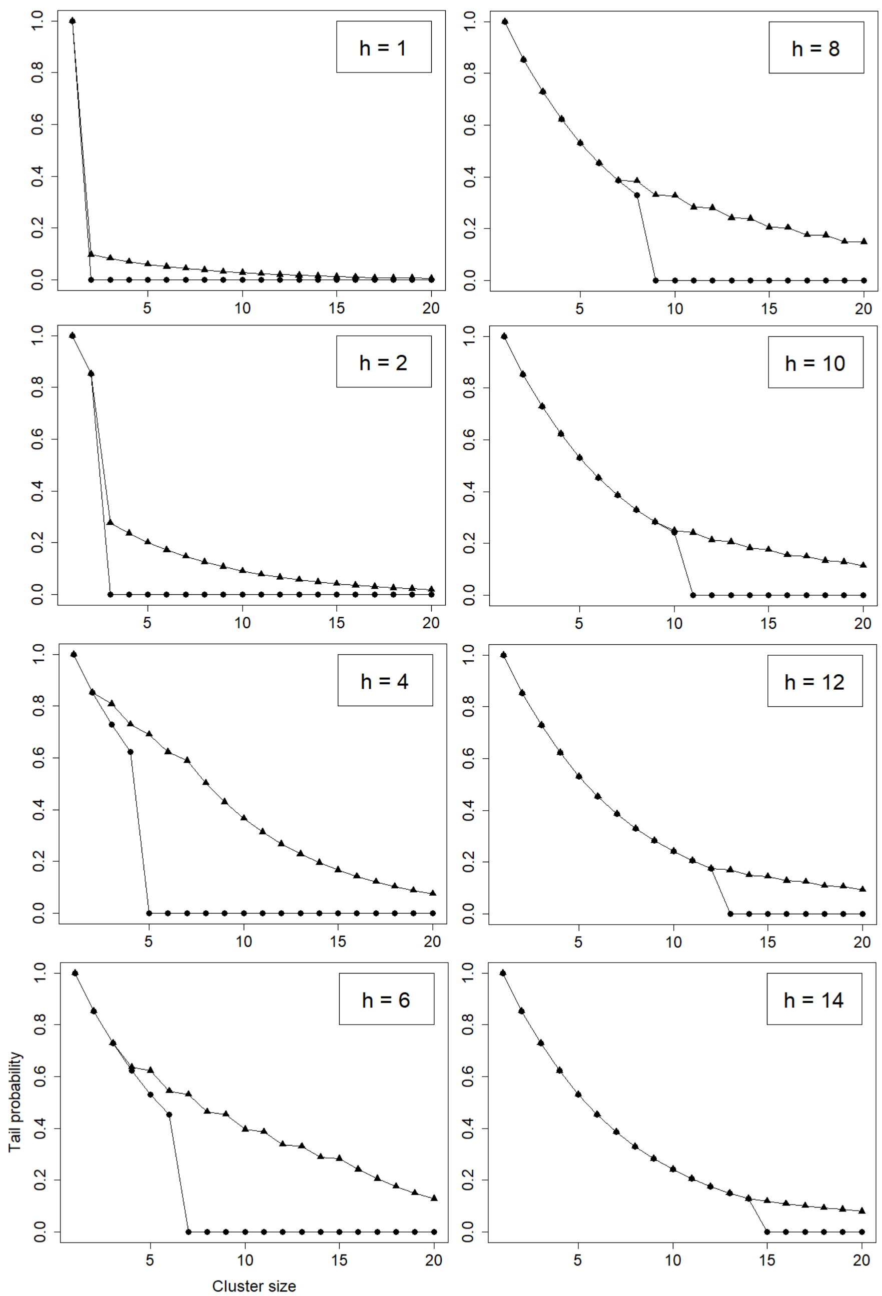

Figure 5: Theoretical tail probability given by Equation (4.16) of cluster sizes of extreme errors (4.18) (strong representation, points) and (4.19) (all-pass representation, triangles) for $\alpha = 1.5$, $\psi_0 = 0.9$ at different horizons h.

We illustrate the extreme clustering behaviors of the two error sequences (4.18) and (4.19) for various horizons and parameter values $\alpha = 1.5$, $\psi_0 = 0.9$. From equations (4.18) and (4.19), we deduce the sequence $(c_{(k)})$ and compute the tail probability distributions of the cluster size using (4.16). As depicted in Figure 5, the contrast between the errors of the all-pass representations and those of the strong representations is the highest for intermediate values of h.

Q Monte Carlo study: complementary results and methodology

Q.1 Asymptotic distribution of the LS estimator

			$\alpha = 1.5$	$\psi = 0.7$	$\phi = 0.9$			$\alpha = 1$	$\psi = 0.7$	$\phi = 0.9$	
n		$q_{0.1}$	$q_{0.25}$	Median	$q_{0.75}$	$q_{0.9}$	$q_{0.1}$	$q_{0.25}$	Median	$q_{0.75}$	$q_{0.9}$
500	$\hat{\delta}_1$	-2.759	-1.338	-0.527	-0.061	0.231	-12.69	-3.569	-0.731	0.012	0.691
	$\hat{\delta}_2$	-0.265	0.038	0.495	1.284	2.653	-0.873	-0.049	0.694	3.430	12.13
2000	$\hat{\delta}_1$	-1.558	-0.746	-0.226	0.086	0.417	-6.321	-1.732	-0.221	0.247	1.382
	$\hat{\delta}_2$	-0.448	-0.105	0.214	0.730	1.521	-0.662	-0.320	0.001	0.322	0.655
5000	$\hat{\delta}_1$	-1.188	-0.565	-0.132	0.156	0.513	-4.564	-1.269	-0.097	0.387	1.824
	$\hat{\delta}_2$	-0.536	-0.172	0.125	0.561	1.177	-2.098	-0.469	0.096	1.357	4.749
∞	$\hat{\delta}_1$	-0.726	-0.252	0.000	0.246	0.719	-5.470	-0.856	0.000	0.954	5.686
	$\hat{\delta}_2$	-0.762	-0.264	0.000	0.268	0.768	-6.687	-1.110	0.000	1.006	6.503
			$\alpha = 0.5$	$\psi = 0.7$	$\phi = 0.9$			$\alpha = 1.7$	$\psi = 0.3$	$\phi = 0.4$	
500	$\hat{\delta}_1$	-1307	-114.6	-5.247	0.157	14.06	-1.003	-0.513	-0.042	0.408	0.870
	$\hat{\delta}_2$	-21.31	-0.412	5.176	114.8	1239	-0.958	-0.484	-0.008	0.466	0.956
2000	$\hat{\delta}_1$	-524.3	-40.97	-0.493	2.804	54.63	-0.662	-0.328	-0.016	0.290	0.618
	$\hat{\delta}_2$	-74.37	-4.171	0.506	46.28	563.9	-0.662	-0.320	0.001	0.322	0.655
5000	$\hat{\delta}_1$	-385.3	-28.11	-0.109	5.402	96.34	-0.641	-0.313	-0.008	0.292	0.608
	$\hat{\delta}_2$	-127.1	-7.493	0.111	33.07	445.0	-0.647	-0.318	-0.001	0.316	0.648
∞	$\hat{\delta}_1$	-1546	-31.43	0.000	32.34	1614	-0.555	-0.235	0.000	0.231	0.554
	$\hat{\delta}_2$	-2129	-42.88	0.000	41.63	2068	-0.614	-0.257	0.001	0.261	0.621

Table Q.1: Characteristics of the empirical distribution of $\hat{\delta}_i = \left(\frac{n}{\ln n}\right)^{1/\alpha}(\hat{\boldsymbol{\eta}}_i - \boldsymbol{\eta}_{0i})$, for $i = 1, 2$ over 100,000 simulated paths of α-stable MAR(1,1) processes (X_t) solution of $(1-\psi F)(1-\phi B)X_t = \varepsilon_t$ with four different parametrizations $(\alpha, \psi_0, \phi_0) \in \{(1.5, 0.7, 0.9), (1, 0.7, 0.9), (0.5, 0.7, 0.9), (1.7, 0.3, 0.4)\}$. The empirical a-quantile is denoted q_a. The results for $n = \infty$ are obtained by simulations of the asymptotic distribution in (4.10). [See Example 4.1]

Q.2 Direct implementation of the Portmanteau test

We conducted an experiment to assess the direct implementation of the portmanteau test (without Monte Carlo) and focused on $\alpha = 1.5$. We computed the residuals of the 100,000 simulated paths based on the all-pass causal AR(2) fits, evaluate the statistic (4.13) for $h = 1, \ldots, 10$ and simulate its asymptotic distribution. For each path, we performed the test at three different nominal sizes 1%, 5% and 10% by comparing the statistics to the appropriate quantile of the asymptotic distribution. The empirical sizes are reported in

Table Q.2. The test suffers heavy distortions, especially in smaller samples, which was expected from the results by Lin and McLeod (2008) in the pure causal AR framework. It is generally oversized for small lags and progressively becomes undersized as more lags are included. The empirical sizes slowly approach the nominal sizes as the number of observations increases and the discrepancy between few and more lags also gets smaller.

	$n = 500$			$n = 2000$			$n = 5000$		
H	1%	5%	10%	1%	5%	10%	1%	5%	10%
1	6.69	21.2	31.7	3.08	9.42	17.0	1.92	6.28	12.5
2	4.54	16.4	27.1	2.40	7.80	14.7	1.60	5.77	11.6
3	3.40	13.4	22.8	1.96	6.41	12.4	1.36	4.84	10.1
4	2.65	10.7	19.0	1.64	5.38	10.3	1.17	4.17	8.74
5	2.11	8.96	16.2	1.37	4.58	8.96	1.04	3.59	7.61
6	1.61	7.58	13.8	1.16	3.93	7.94	0.91	3.20	6.84
7	1.24	6.49	12.1	1.01	3.51	7.17	0.80	2.86	6.22
8	0.96	5.66	10.6	0.89	3.19	6.58	0.70	2.62	5.73
9	0.74	5.08	9.62	0.81	2.94	5.99	0.64	2.42	5.30
10	0.57	4.55	8.74	0.75	2.70	5.50	0.60	2.26	5.00

Table Q.2: Empirical sizes of portmanteau tests with nominal sizes 1%, 5% and 10% using the first H lags, $H = 1, \dots, 10$ of the residuals' autocorrelations of 100,000 simulated paths of process (X_t) solution of $(1 - 0.7F)(1 - 0.9B)X_t = \varepsilon_t$, with 1.5-stable noise.

Q.3 Extreme residuals clustering

Q.3.1 Estimating the term structure of excess clustering

In practice, for one simulated path of the MAR(1,1) process (X_t) and one horizon h, we have six series of residuals $(\hat{\zeta}^i_{t+h|t})_t$, $i = 1, \dots, 6$, one each for the pure causal and noncausal AR(2) competitors, and two each for the two MAR(1,1) competitors (one for the causal component, one for the noncausal component). To compute the cluster sequences $(\hat{\xi}^i_{k,h}(x))_k$ as defined in Section 5.2 for each residuals series, we need to choose a threshold $x > 0$. It would be desirable to use thresholds such that we can harmoniously compare the clustering behaviors of the six series of residuals. For the experiment detailed below, we worked with the autostandardised series of residuals

$$\hat{v}^i_{t+h|t} := \left(\frac{\hat{\zeta}^i_{t+h|t}}{\max\limits_s |\hat{\zeta}^i_{s+h|s}|} \right)_t, \tag{Q.1}$$

which lie between 0 and 1, and for each horizon h, we used the threshold

$$x_h := \max_{i=1,\dots,6} q_a\left(|\hat{v}^i_{t+h|t}|\right), \tag{Q.2}$$

where $q_a(\cdot)$ the a-percent quantile. In our experiments, $a = 0.9$ was used.

Outline of the experiment

For a given parameterization (α, ψ_0, ϕ_0) and path length n, we simulate 10000 paths of process (X_t) solution of $(1 - \psi_0 F)(1 - \phi_0 B)X_t = \varepsilon_t$ and conducted the experiment as follows. For each simulated path of (X_t) and a given horizon $h \geq 1$:

ι) Estimate the regression $X_t = \hat{\eta}_1 X_{t-1} + \hat{\eta}_2 X_{t-2} + \hat{\zeta}_t$.

$\iota\iota$) Obtain the set of (inverted) roots $\{\hat{\psi}, \hat{\phi}\}$ by solving for the zeros of $\hat{\eta}(z) = 1 - \hat{\eta}_1 z - \hat{\eta}_2 z^2$.

$\iota\iota\iota$) For each of the four competing models (5.1)-(5.4), decompose the process into pure causal and noncausal components and compute $(\hat{v}^i_{t+h|t})$, the series of autostandardised errors at horizons h as in (Q.1).

$\iota\nu$) Compute x_h as in (Q.2) and obtain the cluster sizes sequences $(\hat{\xi}^i_{k,h}(x_h))_k$ for each series $(\hat{v}^i_{t+h|t})$, $i = 1, \ldots, 6$.

ν) Compute the Excess Clustering at horizon h of each residuals series as in (5.5).

$\nu\iota$) For the two MAR(1,1) competitors, average the Excess Clustering indicators obtained from the residuals of the causal and noncausal components.

For a given simulated path (X_t), we repeat the above steps for horizons $h = 1, \ldots, H$ and obtain four estimators of the term structure of Excess Clustering, one for each of the competing models (5.1)-(5.4). Across the 10000 simulated paths of (X_t), one can then either:

(i) average model-wise across the obtained term structure estimators to gauge the typical excess clustering behavior of each competing model (as in Figures 3 and Q.1), or

(ii) for each of the simulated paths (X_t), compute the area under the four estimated term structures, select the least clustering model and evaluate the rate of correct selections.

Q.3.2 Excess clustering for additional parameterizations

We evaluated the residuals excess clustering behaviors of the four alternatives (5.1)-(5.4) for additional parameterizations and sample sizes of the MAR(1,1) data generating process. Excess clustering in all-pass residuals is apparent even for small sample sizes. The contrast between the residuals of the strong representation and those of the all-pass increases as the sample size grows (see the left panel of Figure 3 and the two upper panels of Figure Q.1). Also, even with a much smaller noncausal parameter $\psi = 0.2$ (lower right panel of Figure Q.1), the strong representation still clearly displays the least excess clustering compared to the three other competitors. We can nevertheless notice in this case that the pure causal AR(2) alternative is not far from the strong representation (points). This is coherent with the fact that the noncausal parameter ψ is relatively small, especially compared to the causal parameter ϕ, yielding much weaker dependence across the residuals of the misspecified pure causal AR(2).

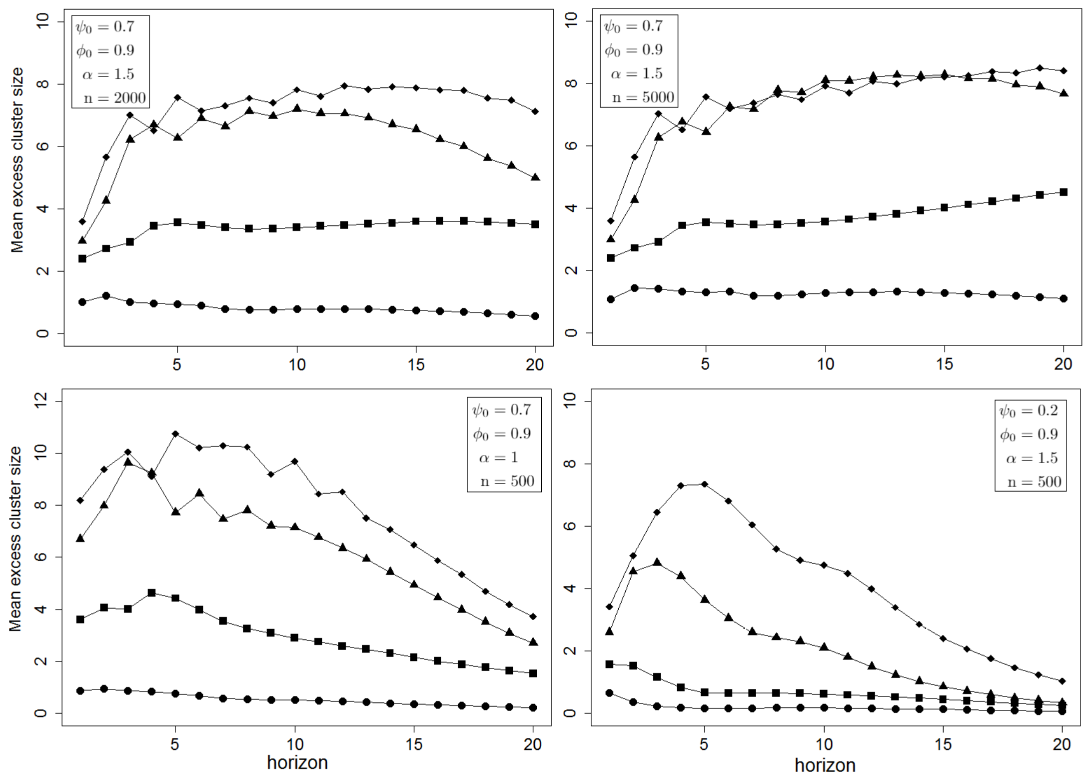

Figure Q.1: Across 10,000 simulations of the α-stable MAR(1,1) process (X_t) solution of $(1 - \psi_0 F)(1 - \phi_0 B)X_t = \varepsilon_t$, average of the term structure of excess clustering of the linear residuals of the four competing models (5.1) (squares), the strong representation (5.2) (points), (5.3) (triangles) and (5.4) (diamond). The parameterizations and path lengths are indicated on each panel.

R Real data: complementary results using the R package 'MARX'

R.1 Total AR orders selection by Information Criterion

The portmanteau procedure of Section 6.1 allowed to discard non-admissible low order models for the six financial and economic time series considered. Portmanteau tests are however not designed to select an «optimal» model. To go further, we report in Table Q.3 the orders that minimise Akaike's information criterion (AIC) using the R package 'MARX' available on CRAN (see Hecq, Telg and Lieb (2017b)). The validity of such AIC's for innovations in the domain of attraction of a stable law has been studied by Knight (1989). Except for the HSI, the results of the two procedures are compatible, the AIC criterion tending to select higher optimal orders.

	Cotton	Soybean	Sugar	Coffee	HSI	Shiller P/E
Selected total AR order	3	8	7	9	1	8

Table Q.3: Optimal order minimising the AIC criterion.

R.2 Identification of causal and noncausal roots

Given the lowest total AR orders validated by the portmanteau procedure (see Table 6.1), we used the routine marx.t of the 'MARX' package to fit MAR models on the six series by t-Student ML. The results are presented in Table Q.4. Except for the HSI and the sugar series, the causal/non-causal orders obtained are equal to those of the final specifications in Table 5. The estimated roots are also similar, but we note some discrepancies in their causal/non-causal allocations.

Series	Final specification	Noncausal (inverted) roots	Causal (inverted) roots
Cotton	MAR(2,0)	0.93, 0.11	–
Soybean	MAR(2,3)	$0.16 \pm 0.42i$	$0.94, -0.55, 0.30$
Sugar	MAR(2,2)	$0.29 \pm 0.41i$	$0.96, -0.43$
Coffee	MAR(1,3)	0.41	$0.95, -0.23 \pm 0.20$
HSI	MAR(3,0)	$0.92, 0.28, -0.21$	–
Shiller P/E	MAR(2,4)	$0.95, -0.48, 0.50 \pm 0.23i$	$-0.21 \pm 0.43i$

Table Q.4: Estimation of the MAR(p, q) specification for each financial series by t-Student ML using the routine marx.t of the 'MARX' package. This routine requires as input the total AR order $p + q$, for which we used the validated orders given by Table 4.

References

Bierens, H. (1982): Consistent Model Specification Tests. *Journal of Econometrics*, 20, 105-134.

Knight, K. (1989): Consistency of Akaike's information criterion for infinite variance autoregressive processes. *The Annals of Statistics*, 824-840.