The Marginal Distribution Function of Threshold-type Processes with Central Symmetric Innovations

de Gooijer, J.G.

DOI
10.1080/02331888.2022.2029862

Publication date
2022

Document Version
Final published version

Published in
Statistics

License
CC BY-NC-ND

Statistics
A Journal of Theoretical and Applied Statistics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gsta20

The marginal distribution function of threshold-type processes with central symmetric innovations

Jan G. De Gooijer

To cite this article: Jan G. De Gooijer (2022) The marginal distribution function of threshold-type processes with central symmetric innovations, Statistics, 56:1, 1-33, DOI: 10.1080/02331888.2022.2029862

To link to this article: https://doi.org/10.1080/02331888.2022.2029862

Published online: 14 Feb 2022.

STATISTICS
2022, VOL. 56, NO. 1, 1–33
https://doi.org/10.1080/02331888.2022.2029862

Taylor & Francis
Taylor & Francis Group

The marginal distribution function of threshold-type processes with central symmetric innovations

Jan G. De Gooijer

Amsterdam School of Economics, University of Amsterdam, Amsterdam, the Netherlands

ABSTRACT
This paper addresses the problem of finding exact and explicit (closed-form) expressions for the stationary marginal distribution of threshold-type time series processes, their associated moments, autocovariance and autocorrelation coefficients. The innovation process of the models under consideration follows three central symmetric distribution functions: Gaussian, Laplace, and Cauchy. Theoretical results for both two- and three-regime threshold-type models are derived. Various examples give rise to a deeper understanding of certain features of the stationary process structure. Exact results for the stationary density, central moments, and autocorrelations of threshold-type processes are compared with approximate density and moment results obtained through an existing numerical methods.

ARTICLE HISTORY
Received 18 April 2021
Accepted 10 January 2022

KEYWORDS
Autocovariance coefficient; Markov chain; moments; numerical method; probability density function

2020 MATHEMATICS SUBJECT CLASSIFICATIONS
62E15; 62E17; 62M10

1. Introduction

The impact of univariate threshold-type time series processes is enormous across many areas of science. Particular forms of these processes, such as the subclass of self-exciting threshold autoregressive (SETAR) processes have been defined and explored in detail; see, e.g. Chen et al. [1] and Hansen [2]. Indeed, much is known about the estimation and testing performance of threshold-type processes. By contrast, and somewhat surprising, little is known about the exact stationary marginal distribution function of the data generating process (DGP) underlying these processes. Knowledge of the exact marginal distribution of a threshold-type process is very useful in understanding the structure of the DGP, and in formulating or selecting models appropriate to given situations. For instance, it is often desirable to investigate such statistics as process mean, variance, skewness, kurtosis and autocorrelation. By obtaining an explicit exact expression for the stationary marginal distribution function, these statistics can readily be obtained. The distribution may also shed light on many other process characteristics, including multi-modality. The stationary marginal distribution function may serve as a basis for generating a typical value of a threshold-type process to start off a simulation study. Furthermore, the exact calculation of the likelihood function requires knowledge of the stationary marginal distribution since it enters the likelihood through the first observation.

CONTACT Jan G. De Gooijer j.g.degooijer@uva.nl Amsterdam School of Economics, University of Amsterdam, PO Box 15867, Amsterdam 1001NJ, The Netherlands

In this paper, we focus on the exact stationary marginal distribution function of low-order threshold-type processes and their moment and (auto)correlation properties. In particular, special attention is given to two- and three regime self-exciting threshold autoregressive (SETAR) processes; see Section 2 for a definition. We assume that these processes have error terms (innovations) following a central symmetric distribution function, and hence, we go beyond the usual assumption of Gaussian distributed innovations. The structure of low-order SETAR processes is simple, useful and interpretable in a wide range of contexts. Our aim is to derive explicit exact expressions of the stationary marginal distribution function of threshold-type models. This will allow us to better understand the structure of these models. For completeness, we also summarize several studies on this topic. To some extent, this part of the paper complements the limited review of stationary distributions provided by Ranocha [3].

Closely linked to some of the threshold models under consideration, is the two-regime threshold AR(1) process with an exogenous trigger and Gaussian innovations of Knight and Satchell [4]. In the case, the trigger follows a Bernoulli distribution, these authors obtained closed-form expressions for the stationary marginal distribution of the DGP, and the associated mean, variance and autocovariance function; see also Grynkiv and Stentoft [5] for a two-regime threshold vector AR model with an exogenous trigger. Guay and Scaillet [6] derived exact expressions for the conditional skewness and the conditional kurtosis of a time series following a contemporaneous first-order asymmetric process with Gaussian innovations, a special case of the threshold asymmetric moving average model briefly discussed in Remark 6.3. It appears that for higher-order processes with non-Gaussian innovations, extensions of the results in the above studies get very complicated. Therefore, these threshold-type specifications are outside the scope of this paper.

The rest of the paper is organized as follows. In Section 2, we give a general formulation of the stationary distribution of a Markov chain. Next, we introduce the multiple-regime SETAR process with central symmetric innovations. Also, we present some moment expressions to be evaluated explicitly later. In Section 3.1, we derive an explicit and exact expression for the stationary marginal distribution function, its moments and covariance function of a two-regime piecewise constant model (PCM), a subclass of the SETAR process introduced in Section 2. The PCM has Gaussian innovations. Section 3.2 provides a similar explicit expression but now for PCMs with central symmetric innovations. The section also exemplifies some features of the marginal density function for central symmetric tick- and thin-tailed innovation distribution functions. Section 4 contains an explicit formula for the stationary density of a three-regime PCM with central symmetric innovations. Section 5 considers a multiplicative PCM with general innovations. In Section 6, the focus is on two-regime first-order SETAR process with Gaussian, Laplace, and Cauchy innovations. In Section 7, we discuss a numerical procedure to approximate the stationary marginal distribution function. In addition, we compare the quality of the procedure with the exact pdf given earlier in Section 6. Finally, in Section 8, we investigate the stationary marginal density of a simple three-regime SETAR model for a modified random walk process. Some concluding remarks are given in Section 9. The proofs of new theoretical results are presented in the Appendix.

2. Preliminaries

Consider a strictly stationary and ergodic univariate time series process $\{Y_t, t \in \mathbb{Z}\}$. We wish to evaluate its invariant probability distribution π. Let $\{Y_t\}$ be an ergodic Markov chain in $\mathbb{R}$ with a given initial distribution. Then π satisfies the integral equation

$$\pi(A) = \int_{-\infty}^{\infty} \mathbb{P}(A|y)\pi(\mathrm{d}y), \tag{1}$$

where A is a Borel set of $\mathbb{R}$, and $\mathbb{P}(\cdot|\cdot)$ is the usual conditional probability (transition kernel). Finding an explicit expression of (1) for linear and nonlinear time series processes is a non-trivial problem. In this paper, we mainly focus on special cases of a k-regime SETAR model of order $(k; p_1, \ldots, p_k)$ with delay $d = 1$. This model is defined as

$$Y_t = \alpha_0^{(j)} + \sum_{u=1}^{p_j} \alpha_u^{(j)} Y_{t-u}\mathbb{I}_{(r_{j-1}, r_j]}(Y_{t-1}) + \varepsilon_t, \quad (j = 1, \ldots, k), \tag{2}$$

where $-\infty = r_0 < r_1 < \cdots < r_k = \infty$ are the threshold values, $\{\varepsilon_t\}$ is a sequence of independent and identically distributed (i.i.d.) central symmetric innovations having an absolutely continuous probability density function (pdf) on $\mathbb{R}$ with density $\mathfrak{g}(\varepsilon; \cdot)$ and distribution function $\mathcal{G}_{(\cdot)}$, and $\mathbb{I}_A(\cdot)$ denotes the indicator of a set A, i.e. $\mathbb{I}_A(x) = 1$ if $x \in A$ and $\mathbb{I}_A(x) = 0$ otherwise.

One possible extension of (2) is to allow the ε_t's to be different in each regime with different variances, assuming that they are independent across regimes. This extension, while not pursued in this paper, can easily be accommodated in the derivation of the stationary marginal densities. Another extension is to allow for moving average (MA) lag polynomials in (2); see, e.g. Remark 6.4, and Remark 7.1.

In the next sections, we give explicit expressions (often in closed form) of the exact stationary marginal pdf $f(y)$ for various special cases of (2). Using these expressions, we provide explicit expressions for the sth non-central moment $\nu_s = \int_{-\infty}^{\infty} y^s f(y)\,\mathrm{d}y$, assuming it exists. Based on the non-central moments, we also show exact and explicit expressions for the mean, variance, skewness and kurtosis which are, respectively, defined as

$$\mu = \nu_1, \quad \sigma^2 = \nu_2 - \mu^2, \quad \mathcal{S} = \left(\nu_3 - 3\nu_2\nu_1 + 2\nu_1^3\right)/\sigma^3,$$
$$\mathcal{K} = \left(\nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4\right)/\sigma^4.$$

In addition, we provide exact expressions for the lag ℓ covariance function γ_ℓ and the lag ℓ autocorrelation function ρ_ℓ which for a stationary process $\{Y_t, t \in \mathbb{Z}\}$ are, respectively defined as

$$\gamma_\ell = \mathbb{E}(Y_t Y_{t-\ell}), \quad \rho_\ell = \left(\gamma_\ell - \mu^2\right)/\sigma^2, \quad (\ell = 1, 2, \ldots).$$

Knowledge of these statistics makes it possible to obtain a deeper understanding of the statistical properties of the threshold-type models without resorting to extensive Monte Carlo simulations.

3. A two-regime piecewise constant model

3.1. Gaussian innovations

We start with a simple subclass of the SETAR model in (2) in which the autoregressive function has the form of a piecewise constant function of its arguments. The resulting zero-order SETAR process, called two-regime piecewise constant model (PCM), is defined by

$$Y_t = \alpha_0^- \mathbb{I}_{(-\infty,r]}(Y_{t-1}) + \alpha_0^+ \mathbb{I}_{(r,\infty)}(Y_{t-1}) + \varepsilon_t, \tag{3}$$

where $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)$ with $\sigma_\varepsilon^2 < \infty$. Without loss of generality, we take $\sigma_\varepsilon^2 = 1$. Note that the process $\{\mathbb{I}_{(-\infty,r]}(Y_{t-1})\}$ is a Markov chain. As we will see below, the simple structure of (3) makes a full investigation of features of nonlinearity possible.

The conditional distribution of Y_t given $Y_{t-1} = y_{t-1}$ is

$$h(y_t|y_{t-1}) = \begin{cases} \varphi(y_t; \alpha_0^-) & \text{if } y_{t-1} \le r, \\ \varphi(y_t; \alpha_0^+) & \text{if } y_{t-1} > r, \end{cases}$$

where $\varphi(y;\mu) \equiv \varphi(y;\mu,1)$ is the pdf of a Gaussian distributed random variable X with mean μ and variance 1. Denoting the stationary density of Y_{t-1} by $f(y_{t-1})$, the joint density of Y_t and Y_{t-1} is given by

$$g(y_t, y_{t-1}) = h(y_t|y_{t-1})f(y_{t-1}) = \begin{cases} \varphi(y_t; \alpha_0^-)f(y_{t-1}) & \text{if } y_{t-1} \le r, \\ \varphi(y_t; \alpha_0^+)f(y_{t-1}) & \text{if } y_{t-1} > r. \end{cases} \tag{4}$$

The density $f(y_t)$ is stationary if and only if it is also the marginal density of $\{Y_t, t \in \mathbb{Z}\}$. This gives the condition

$$f(y_t) = \int_{-\infty}^{r} \varphi(y_t; \alpha_0^-)f(y_{t-1})\,\mathrm{d}y_{t-1} + \int_{r}^{\infty} \varphi(y_t; \alpha_0^+)f(y_{t-1})\,\mathrm{d}y_{t-1}. \tag{5}$$

Using (5), the stationary marginal density $f_{\alpha_0^-,\alpha_0^+}(y)$ is given in Proposition 3.1.

Proposition 3.1: *The stationary marginal density of the* PCM (3) *is given by*

$$f_{\alpha_0^-,\alpha_0^+}(y) = w\varphi(y;\alpha_0^-) + (1-w)\varphi(y;\alpha_0^+), \tag{6}$$

where

$$w = \mathbb{P}(Y_t \le r) = \frac{\Phi_{(r-\alpha_0^+)}}{1 - \Phi_{(r-\alpha_0^-)} + \Phi_{(r-\alpha_0^+)}} \in [0,1), \tag{7}$$

with $\Phi_{(x)} = \int_{-\infty}^{x} \varphi(z;0)\mathrm{d}z$ *the distribution function of* $\mathcal{N}(0,1)$.

Next, we derive the first four non-central moments of $\{Y_t\}$. Making use of the result $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x\exp(-\frac{1}{2}(x-a)^2)\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(x+a)\exp(-\frac{1}{2}x^2)\mathrm{d}x = a$ $(a \in \mathbb{R})$, it is easy to see that the mean of $\{Y_t\}$ is given by

$$\nu_1 = \int_{-\infty}^{\infty} yf(y)\,\mathrm{d}y = w\alpha_0^- + (1-w)\alpha_0^+.$$

To obtain the second-, third- and fourth-order non-central moments, we first introduce the following integral equations:

$$\int_{-\infty}^{\infty} x^2\varphi(x;a)\,\mathrm{d}x = (1+a^2), \quad \int_{-\infty}^{\infty} x^3\varphi(x;a)\,\mathrm{d}x = (3+3a^3),$$
$$\int_{-\infty}^{\infty} x^4\varphi(x;a)\,\mathrm{d}x = (3+6a^2+a^4).$$

Then, on expanding $(y+\alpha_0^{\pm})^s$ in a binomial form for each value of s, we have

$$\nu_2 = w\big(1+(\alpha_0^-)^2\big) + (1-w)\big(1+(\alpha_0^+)^2\big),$$
$$\nu_3 = w\big(3\alpha_0^- + (\alpha_0^-)^3\big) + (1-w)\big(3\alpha_0^+ + (\alpha_0^+)^3\big),$$
$$\nu_4 = w\big(3+6(\alpha_0^-)^2 + (\alpha_0^-)^4\big) + (1-w)\big(3+6(\alpha_0^+)^2 + (\alpha_0^+)^4\big).$$

For general central symmetric innovations, explicit and exact expressions for $\mathcal{S}$ and $\mathcal{K}$ are given in Example 3.1.

Corollary 3.1: *The covariance coefficient γ_ℓ at lag ℓ ($\ell \geq 1$) is given by*

$$\gamma_\ell = A_1\gamma_{1,\ell} + A_2\gamma_{2,\ell}, \tag{8}$$

where the lag ℓ, regime-specific covariance coefficients $\gamma_{1,\ell} = \mathbb{E}(Y_tY_{t-\ell}\mathbb{I}_{(-\infty,r]}(Y_{t-1}))$ and $\gamma_{2,\ell} = \mathbb{E}(Y_tY_{t-\ell}\mathbb{I}_{(r,\infty)}(Y_{t-1}))$ ($\ell \geq 2$) can be expressed recursively as

$$\gamma_{1,\ell} = \big(1-\Phi_{(r-\alpha_0^-)}\big)\gamma_{2,\ell-1} + \Phi_{(r-\alpha_0^-)}\gamma_{1,\ell-1} \quad \text{with } \gamma_{1,1} = \alpha_0^-,$$
$$\gamma_{2,\ell} = \big(1-\Phi_{(r-\alpha_0^+)}\big)\gamma_{2,\ell-1} + \Phi_{(r-\alpha_0^+)}\gamma_{1,\ell-1} \quad \text{with } \gamma_{2,1} = \alpha_0^+,$$

and where A_1 and A_2 are the constants

$$A_1 = (1-w)\big\{-\varphi(r-\alpha_0^+) + \alpha_0^+\Phi_{(r-\alpha_0^+)}\big\} + w\big\{-\varphi(r-\alpha_0^-) + \alpha_0^-\Phi_{(r-\alpha_0^-)}\big\},$$
$$A_2 = (1-w)\big\{\varphi(r-\alpha_0^+) + \alpha_0^+\big(1-\Phi_{(r-\alpha_0^+)}\big)\big\} + w\big\{\varphi(r-\alpha_0^-) + \alpha_0^-\big(1-\Phi_{(r-\alpha_0^-)}\big)\big\},$$

with $\varphi(r-\alpha_0^{\pm}) = (1/\sqrt{2\pi})\exp(-(r-\alpha_0^{\pm})^2/2)$.

Note that $A_1 + A_2 = w\alpha_0^- + (1-w)\alpha_0^+$. By substituting this result in (8) and rearranging terms, we obtain the following relationship between γ_ℓ and $\gamma_{\ell-1}$ for $\ell \geq 2$,

$$\gamma_\ell = \big(\Phi_{(r-\alpha_0^-)} - \Phi_{(r-\alpha_0^+)}\big)\gamma_{\ell-1} + \Big((1-\Phi_{(r-\alpha^-)})\gamma_{2,\ell-1} + \Phi_{(r-\alpha_0^+)}\gamma_{1,\ell-1}\Big)(A_1+A_2), \tag{9}$$

where $\gamma_1 = A_1\alpha_0^- + A_2\alpha_0^+$, $\gamma_{1,1} = \alpha_0^-$ and $\gamma_{2,1} = \alpha_0^+$. The corresponding lag ℓ autocorrelation function of $\{Y_t, t \in \mathbb{Z}\}$ takes the form $\rho_\ell = (\gamma_\ell - \mu^2)/\sigma^2$ ($\ell \geq 1$) where $\sigma^2 = 1 + w(1-w)(\alpha_0^+ - \alpha_0^-)^2$. Consider the special case $\alpha_0^+ = -\alpha_0^- \equiv \alpha$, with $r = 0$. Then, $A_1 + A_2 = 0$, so that (9) reduces to $(\Phi_{(\alpha)} - \Phi_{(-\alpha)})^{\ell-1}\gamma_1$ ($\ell \geq 2$), where $\gamma_1 = 2\alpha A_2$ with $A_2 = \varphi(\alpha) + (\alpha/2)(\Phi_{(\alpha)} - \Phi_{(-\alpha)})$.

3.2. Central symmetric innovations

Assume that the random variable ε_1 follows a strictly and continuously positive pdf on $\mathbb{R}$ with a central symmetric distribution function $\mathcal{G}_{(\cdot)}$. This setup includes, for instance, the Gaussian distribution, the Student t-distribution, or the classical symmetric Laplace (standard double exponential) distribution. Then, in analogy with Proposition 3.1, the stationary marginal density $f_{\alpha_0^-,\alpha_0^+}(y)$ is given in Proposition 3.2.

Proposition 3.2: *The stationary marginal density $f_{\alpha_0^-,\alpha_0^+}(y)$ of the* PCM (3) *with central symmetric innovations is given by*

$$f_{\alpha_0^-,\alpha_0^+}(y) = \omega f_{\alpha_0^-}(y) + (1-\omega) f_{\alpha_0^+}(y), \tag{10}$$

where

$$\omega = \mathbb{P}(Y_t \le r) = \frac{\mathcal{G}_{(r-\alpha_0^+)}}{1 - \mathcal{G}_{(r-\alpha_0^-)} + \mathcal{G}_{(r-\alpha_0^+)}} \in [0,1), \tag{11}$$

and $f_{\alpha_0^-}(y)$ and $f_{\alpha_0^+}(y)$ are the densities of $\{Y_t, t \in \mathbb{Z}\}$ in the lower and upper regime, respectively.

Remark 3.1: The proof of Proposition 3.2 has been omitted from the paper since it can be obtained along the same lines as the proof of Proposition 3.1 given in the Appendix.

Example 3.1 (Li et al. [7]): The lag ℓ autocorrelation function ρ_ℓ, skewness $\mathcal{S}$, and kurtosis $\mathcal{K}$ of the PCM (3) with central symmetric innovations are given by

$$\rho_\ell = \frac{\gamma_\ell}{\gamma_0} = \frac{\lambda_\ell(\alpha_0^- - \alpha_0^+) + \omega(1-\omega)\big(\mathcal{G}_{(r-\alpha_0^-)} - \mathcal{G}_{(r-\alpha_0^+)}\big)^\ell (\alpha_0^- - \alpha_0^+)^2}{\sigma_\varepsilon^2 + \omega(1-\omega)(\alpha_0^- - \alpha_0^+)^2} \quad (\ell \ge 1), \tag{12}$$

$$\mathcal{S} = \frac{\mathbb{E}(\varepsilon_1^3) + (\omega - 3\omega^2 + 2\omega^3)(\alpha_0^- - \alpha_0^+)^3}{[\sigma_\varepsilon^2 + \omega(1-\omega)(\alpha_0^- - \alpha_0^+)^2]^{3/2}}, \tag{13}$$

$$\mathcal{K} = \frac{\mathbb{E}(\varepsilon_1^4) + 6\sigma_\varepsilon^2\omega(1-\omega)(\alpha_0^- - \alpha_0^+)^2 + (\omega - 4\omega^2 + 6\omega^3 - 3\omega^4)(\alpha_0^- - \alpha_0^+)^4}{[\sigma_\varepsilon^2 + \omega(1-\omega)(\alpha_0^- - \alpha_0^+)^2]^2}, \tag{14}$$

where $\lambda_\ell = \mathbb{E}[\varepsilon_{t-\ell}\mathbb{I}_{(-\infty,r]}(Y_{t-1})]$.

Remark 3.2: Consider the numerator of (12) with $\ell = 1$. Let $\alpha_0^+ = -\alpha_0^- \equiv \alpha$, $r = 0$, and $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Then $\gamma_1 = [2\alpha\varphi(\alpha) + (\Phi_{(\alpha)} - \Phi_{(-\alpha)})\alpha^2]$, which is identical to the expression $\gamma_1 = 2\alpha A_2$ given in Section 3.1.

Example 3.2: To provide some insight into the relation between r and the shape of the marginal stationary density, we consider the case $(\alpha_0^-, \alpha_0^+) = (-1.25, 1.25)$. Figure 1 shows $f_{\alpha_0^-,\alpha_0^+}(y)$ as a function of r for a PCM with $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ and for a PCM with $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} t_5$. For both innovation distributions the pdf for $r = 0$ is bimodal around the points α_0^- and

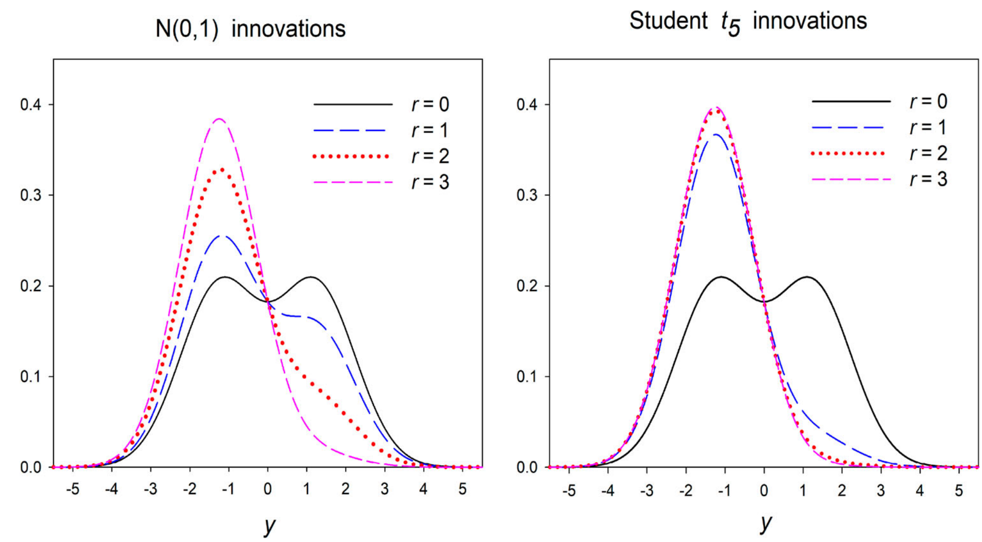

Figure 1. Stationary marginal densities $f_{\alpha_0^-,\alpha_0^+}(y)$ of the PCM (3) with $\mathcal{N}(0,1)$ and Student t_5 innovations, and with $(\alpha_0^-,\alpha_0^+) = (-1.25, 1.25)$.

α_0^+ which follows from the fact that $\omega = 1/2$. So, both DGPs alternate between these two points. As r increases the pdfs become unimodal around the point -1.25. Note that for the PCM with t_5 innovations the curves of the pdfs for $r = 1$, 2, and 3 are hard to distinguish at this scale.

Example 3.3: To explore the shape of the marginal stationary density of a PCM with central symmetric short-tailed (or thin-tailed) innovations, we consider a pdf proposed by Akkaya and Tiku [8]. In particular, for a random variable X the pdf is defined as $f(x;h) = (C/\sqrt{2\pi})\{1 + (x^2/2h)\}^2 \exp(-x^2/2)$ $(-\infty < x < \infty)$, where $C = 1/(1 + 1/h + 3/(4h^2))$ with $h = 2-c$, $c < 2$ a constant. All odd moments of X are zero. The second and fourth central moments of X are given by $C(1 + 3/h + 15/(4h^2))$ and $3 + [48(2 + h)/(3 + 4h(1 + h))]$, respectively. Adopting $f(\varepsilon;h)$ as the pdf of $\{\varepsilon_t, t \in \mathbb{Z}\}$, Figure 2 shows the skewness and kurtosis of the PCM (3) as a function of r for five values of c when $(\alpha_0^-,\alpha_0^+) = (4, -1)$. Note that as $|r|$ increases the kurtosis becomes larger with increasing values of c.

4. Multiple-regime PCM with central symmetric innovations

Underlying model (3) is a Markov chain, $\{M_t\}$, with $k = 2$ states or regimes. This setup can be extended to a PCM with $k \geq 2$ multiple states. Assuming $M_t = j$ if and only if $Y_{t-1} \in (r_{j-1}, r_j]$, the resulting k-regime PCM can be written as

$$Y_t = \alpha_0^{(j)} \mathbb{I}_{(r_{j-1},r_j]}(Y_{t-1}) + \varepsilon_t, \quad (j = 1,\ldots,k), \tag{15}$$

where $\{\varepsilon_t, t \in \mathbb{Z}\}$ are i.i.d. central symmetric innovations. Clearly, (15) can be analysed using Markov chain techniques. To this end, let $\mathbf{P} = (p_{ij})$ denote a $k \times k$ matrix of transition probabilities, where $p_{ij} = \mathbb{P}(M_t = j|M_{t-1} = i) = \mathcal{G}_{(r_j-\alpha_i)} - \mathcal{G}_{(r_{j-1}-\alpha_i)}$. Under fairly

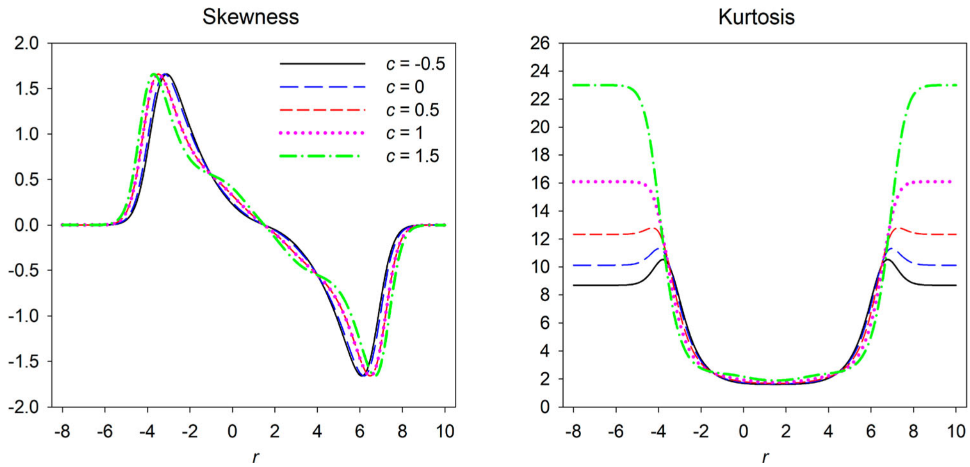

Figure 2. Skewness ($\mathcal{S}$) and kurtosis ($\mathcal{K}$) of the PCM (3) as a function of the threshold parameter r, for five values of c representing different innovation processes each with $\mathcal{K} < 3$.

weak conditions [9] on $\alpha_0^{(j)}$ and $\{\varepsilon_t\}$, $\{M_t\}$ will be ergodic and will possess a stationary distribution denoted by the row vector of probabilities $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_k)$. This is the unique solution to the system $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$ with $\sum_{j=1}^{k} \pi_j = 1$.

Let $f_{n,j}(y)$ denote the conditional distribution of Y_{t+n} given $Y_{t-1} \in (r_{j-1}, r_j]$ $(n \in \mathbb{Z}^+)$. Applying the law of total probability, the $k \times 1$ vector $\mathbf{f}_n(y) = (f_{n,1}(y), \ldots, f_{n,k}(y))^{\mathrm{T}}$ satisfies the recursive relationship $\mathbf{f}_n(y) = \mathbf{P}\mathbf{f}_{n-1}(y)$ $(n \geq 2)$, where $\mathbf{f}_1(y) = (\mathfrak{g}(y; \alpha_0^{(1)}), \ldots, \mathfrak{g}(y; \alpha_0^{(k)}))^{\mathrm{T}}$. Using this result, the stationary marginal density $f_{\alpha_0^{(1)}, \ldots, \alpha_0^{(k)}}(y)$ for the PCM (15) is easily derived to be

$$f_{\alpha_0^{(1)}, \ldots, \alpha_0^{(k)}}(y) = \sum_{j=1}^{k} \pi_j \mathfrak{g}(y; \alpha_0^{(j)}). \tag{16}$$

We see that (16) is a mixture of densities $\mathfrak{g}(y; \cdot)$ with weights given by k transition probabilities. From (16) it follows that the first and second non-central moments of (15) are given by $\nu_1 = \sum_{j=1}^{k} \pi_j \alpha_0^{(j)}$ and $\nu_2 = 1 + \sum_{j=1}^{k} \pi_j (\alpha_0^{(j)})^2$, respectively.

To obtain an expression for the covariance function γ_ℓ, we assume that $\mathbf{P}$ has distinct eigenvalues $1 = \lambda_1, \lambda_2, \ldots, \lambda_k$, where $|\lambda_j| < 1$ for $j \geq 2$. Let the corresponding right and left eigenvectors be $\mathbf{v}_1, \ldots, \mathbf{v}_k$ and $\mathbf{w}_1, \ldots, \mathbf{w}_k$, normalized so that $\mathbf{w}_i^{\mathrm{T}} \mathbf{v}_j = \delta_{ij}$, Kronecker's delta. Also, we introduce the regime-specific mean $\mu_n^{(j)}$ defined as

$$\mu_n^{(j)} = \mathbb{E}[Y_{t+n} \mathbb{I}_{(r_{j-1}, r_j]}(Y_{t-1})], \quad (n \geq 1).$$

Then the covariance function $\gamma_\ell = \mathbb{E}(Y_t Y_{t-\ell})$ $(\ell \geq 1)$ is given by

$$\begin{aligned} \gamma_\ell &= \sum_{j=1}^{k} \pi_j \mathbb{E}[Y_{t-\ell} \mathbb{I}_{(r_{j-1}, r_j]}(Y_{t-1})] \mathbb{E}[Y_t \mathbb{I}_{(r_{j-1}, r_j]}(Y_{t-1})], \\ &= \boldsymbol{\mu}_\ell^{\mathrm{T}} \boldsymbol{\mu}, \end{aligned} \tag{17}$$

where the $k \times 1$ vectors $\boldsymbol{\mu}_\ell$ and $\boldsymbol{\mu}$ are given by

$$\boldsymbol{\mu}_\ell = (\mu_\ell^{(1)}, \ldots, \mu_\ell^{(k)})^{\mathrm{T}} \quad \text{and} \quad \boldsymbol{\mu} = (\pi_1 \mu_0^{(1)}, \ldots, \pi_k \mu_0^{(k)})^{\mathrm{T}},$$

with $\pi_j \mu_0^{(j)} = \mathbb{E}[Y_t \mathbb{I}_{(r_{j-1}, r_j]}(Y_{t-1})] = \int_{r_{j-1}}^{r_j} y f_{\alpha_0^{(1)}, \ldots, \alpha_0^{(k)}}(y)\, \mathrm{d}y$ $(j = 1, \ldots, k)$. Since $\mathbf{f}_n(y) = \mathbf{P}^{n-1}\mathbf{f}_1(y)$ $(n \geq 2)$, it follows that

$$\boldsymbol{\mu}_\ell = \mathbf{P}^{\ell-1}\boldsymbol{\mu}_1 = \mathbf{P}^{\ell-1}\boldsymbol{\alpha}, \tag{18}$$

where $\boldsymbol{\alpha} = (\alpha_0^{(1)}, \ldots, \alpha_0^{(k)})^{\mathrm{T}}$, i.e. a $k \times 1$ parameter vector. On substituting (18) in (17) and using the spectral decomposition of $\mathbf{P}$, the covariance function at lag ℓ can be written as

$$\gamma_\ell = \sum_{j=2}^{k} \lambda_j^{\ell-1} \mathbf{v}_j \mathbf{w}_j^{\mathrm{T}} \boldsymbol{\alpha}, \quad (\ell \geq 1). \tag{19}$$

It is easy to see that if $\mathbb{I}_{(r_{j-1}, r_j]}(Y_{t-d})$ $(d \geq 1)$ is used instead of $\mathbb{I}_{(r_{j-1}, r_j]}(Y_{t-1})$ in (15) then the resulting model describes d independent processes, each of which has the same covariance structure as the equivalent model with delay one.

Using the above approach, Pemberton [10] noted that the covariance structure of (15) is the same as that of a linear autoregressive moving average (ARMA) process of order (p, p) with $p \leq k - 1$, and with i.i.d. central symmetric innovations. The following example provides a simple illustration.

Example 4.1 (Pemberton [11]): Consider the PCM (15) with $k = 2$, $\alpha_0^{(2)} = -\alpha_0^{(1)} \equiv \alpha$, and $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$, i.e.

$$Y_t = -\alpha \mathbb{I}_{(-\infty, 0]}(Y_{t-1}) + \alpha \mathbb{I}_{(0, \infty)}(Y_{t-1}) + \varepsilon_t. \tag{20}$$

Then the stationary marginal density of $\{Y_t, t \in \mathbb{Z}\}$ is given by

$$f_\alpha(y) = \sum_{j=1}^{2} \pi_j \varphi(y; \alpha_j) = \{\varphi(y; -\alpha) + \varphi(y; \alpha)\}/2. \tag{21}$$

Note that (21) is a special case of (6). In particular, for $r = 0$ and $\alpha_0^{+} = -\alpha_0^{-} \equiv \alpha$ the weight factor $w = 1/2$ so that (6) reduces to (21). Further, we see that $\mu = 0$ and $\sigma^2 \equiv \gamma_0 = (1 + \alpha^2)$.

At frequency $\tau \in [-\pi, \pi]$, the spectral density function $f(\tau)$ is given by

$$f(\tau) = \frac{1}{2\pi}\Big\{\gamma_0 - 2C(\alpha)\alpha \sum_{n=1}^{\infty} \beta^{n-1} \cos(n\tau)\Big\}, \tag{22}$$

where $C(\alpha) = 2\varphi(y; \alpha) - \alpha\beta$, and $\beta = 1 - 2\Phi_{(\alpha)}$. The function $f(\tau)$ defined by (22) is identical to the spectral density of a linear ARMA$(1, 1)$ process of the form

$$Y_t - \beta Y_{t-1} = \varepsilon_t - \theta \varepsilon_{t-1}, \tag{23}$$

where θ is the root of $z^2 + A(\alpha)z + 1 = 0$ which lies inside the unit circle, $A(\alpha) = (\beta^2 - 2\beta\rho_1 + 1)/(\rho_1 - \beta)$ with $\rho_1 = -\alpha C(\alpha)/\gamma_0$ the lag one autocorrelation of $\{Y_t, t \in \mathbb{Z}\}$. It can be shown that $0 > \theta > \beta$ so that there is a peak at the frequency $\tau = \pi$. This corresponds to a limit cycle of period 2; recall the comment in Example 3.2 about the limit cycle of (3) for the case $(\alpha_0^-, \alpha_0^+) = (-1.25, 1.25)$.

Proposition 4.1: *Consider the* PCM (15) *with* $k = 3$ *regimes,* $\alpha_j \equiv \alpha_0^{(j)}$ $(j = 1, 2, 3)$*, and* $\{\varepsilon_t\} \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$*, i.e.*

$$Y_t = \alpha_1 \mathbb{I}_{(-\infty, r_1]}(Y_{t-1}) + \alpha_2 \mathbb{I}_{(r_1, r_2]}(Y_{t-1}) + \alpha_3 \mathbb{I}_{(r_2, \infty)}(Y_{t-1}) + \varepsilon_t. \tag{24}$$

Then the stationary marginal density of (24) *is given by*

$$f_{\alpha_1,\alpha_2,\alpha_3}(y) = \tau_1 f_{\alpha_1}(y) + \tau_2 f_{\alpha_2}(y) + \tau_3 f_{\alpha_3}(y), \tag{25}$$

where

$$\begin{aligned}
\tau_1 &= w_1 \Phi_{(r_1-\alpha_1)} + w_2 \Phi_{(r_1-\alpha_2)} + w_3 \Phi_{(r_1-\alpha_3)},\\
\tau_2 &= w_1(\Phi_{(r_2-\alpha_1)} - \Phi_{(r_1-\alpha_1)}) + w_2(\Phi_{(r_2-\alpha_2)} - \Phi_{(r_1-\alpha_2)}),\\
\tau_3 &= w_1(1 - \Phi_{(r_2-\alpha_1)}) + w_2(1 - \Phi_{(r_2-\alpha_2)} + w_3(1 - \Phi_{(r_2-\alpha_3)}),
\end{aligned}$$

with

$$\begin{aligned}
w_1 &= \mathbb{P}(Y_t \le r_1) = \frac{\Phi_{(r_2-\alpha_3)}\Phi_{(r_1-\alpha_2)} + \Phi_{(r_1-\alpha_3)}\Phi_{(\alpha_2-r_2)}}{\Delta(r_1, r_2)} \in [0, 1),\\
w_2 &= \mathbb{P}(r_1 < Y_t \le r_2) = \frac{\Phi_{(r_2-\alpha_3)}\Phi_{(\alpha_1-r_1)} - \Phi_{(r_1-\alpha_3)}\Phi_{(\alpha_1-r_2)}}{\Delta(r_1, r_2)} \in [0, 1),\\
w_3 &= 1 - w_1 - w_2,
\end{aligned}$$

and

$$\begin{aligned}
\Delta(r_1, r_2) = {} & [\Phi_{(r_1-\alpha_2)} - \Phi_{(r_1-\alpha_3)}][\Phi_{(\alpha_1-r_2)} + \Phi_{(r_2-\alpha_3)}]\\
& + [\Phi_{(\alpha_2-r_2)} + \Phi_{(r_2-\alpha_3)}][\Phi_{(\alpha_1-r_1)} + \Phi_{(r_1-\alpha_3)}].
\end{aligned}$$

Remark 4.1: The proof of the proposition can be obtained along similar lines as the proof of Proposition 3.2, using standard but lengthy algebra. The link between Propositions 3.2 and 4.1 can be made explicit by rewriting τ_i $(i = 1, 2, 3)$ as follows:

$$\begin{aligned}
\tau_1 = {} & \frac{1}{\Delta(r_1, r_2)} \Big\{ \Phi_{(r_1-\alpha_3)}\Phi_{(r_1-\alpha_1)} + [\Phi_{(r_2-\alpha_3)} - \Phi_{(r_1-\alpha_3)}]\Phi_{(r_1-\alpha_2)}\\
& - \Phi_{(r_1-\alpha_3)} \sum_{i=1}^{2} (-1)^i \Phi_{(r_2-\alpha_i)} \Phi_{(r_1-\alpha_{3-i})} + \Phi_{(r_2-\alpha_3)} \sum_{i=1}^{2} (-1)^i \Phi_{(r_1-\alpha_i)} \Phi_{(r_{3-i}-\alpha_{3-i})} \Big\}\\
& + w_3 \Phi_{(r_1-\alpha_1)},\\
\tau_2 = {} & \frac{1}{\Delta(r_1, r_2)} \Big\{ [\Phi_{(r_2-\alpha_3)} - \Phi_{(r_1-\alpha_3)}][\Phi_{(r_2-\alpha_2)} - \Phi_{(r_1-\alpha_2)}
\end{aligned}$$

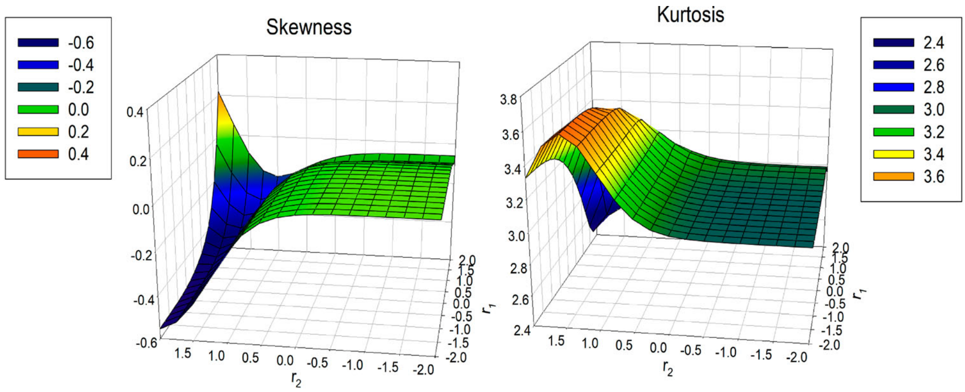

Figure 3. Skewness ($\mathcal{S}$) and kurtosis ($\mathcal{K}$) of the PCM (24) ($\mathcal{N}(0,1)$ innovations) as a function of the threshold parameters r_1 and r_2.

$$- \sum_{i=1}^{2} (-1)^i \Phi_{(r_1-\alpha_{3-i})} \Phi_{(r_2-\alpha_i)}] \Big\} + w_3 \{ \Phi_{(r_2-\alpha_3)} - \Phi_{(r_1-\alpha_3)} \},$$

$$\tau_3 = \frac{1}{\Delta(r_1, r_2)} \Big\{ \Phi_{(r_2-\alpha_3)} [1 - \Phi_{(r_1-\alpha_1)} - \Phi_{(r_2-\alpha_2)} + \Phi_{(r_1-\alpha_2)}$$

$$- \sum_{i=1}^{2} (-1)^i \Phi_{(r_1-\alpha_i)} \Phi_{(r_2-\alpha_{3-i})}] \Big\} + w_3 \{ 1 - \Phi_{(r_2-\alpha_3)} \}.$$

In the case $k = 2$, we have $\Phi_{(r_1-\alpha_3)} = 1$, $\Phi_{(r_2-\alpha_3)} = 1$, and $r_1 = r_2 = r$. Then, it is easy to see that $\Delta(r, r) = 1 - \Phi_{(r-\alpha_1)} + \Phi_{(r-\alpha_2)}$, $w_3 = 0$, $\tau_1 = \Phi_{(r-\alpha_1)}/\Delta(r, r)$, $\tau_2 = 0$, and $\tau_3 = (1 - \Phi_{(r-\alpha_1)})/\Delta(r, r)$. On setting $\alpha_1 = \alpha_0^-$, $\alpha_2 = 0$ and $\alpha_3 = \alpha_1^+$, and substituting the above results in (25), we obtain the stationary marginal density (6). Observe that the assumption of $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ can be relaxed to $\{\varepsilon_t\}$ is i.i.d. with mean zero and unit variance. The shape of $f_{\alpha_1,\alpha_2,\alpha_3}(y)$ is unimodal but not symmetric; see Example 4.2.

Corollary 4.1: *The first four non-central moments of the* PCM (24) *are given by*

$$\nu_1 = \sum_{i=1}^{3} \tau_i \alpha_i, \quad \nu_2 = \sum_{i=1}^{3} \tau_i \big(1 + \alpha_i^2\big),$$

$$\nu_3 = \sum_{i=1}^{3} \tau_i \big(3\alpha_i + \alpha_i^3\big), \quad \nu_4 = \sum_{i=1}^{3} \tau_i \big(3 + 6\alpha_i^2 + \alpha_i^4\big). \tag{26}$$

Example 4.2: Figure 3 shows the skewness $\mathcal{S}$ and kurtosis $\mathcal{K}$ for the PCM (24) with $\alpha_1 = 1$, $\alpha_2 = 2$, and $\alpha_3 = 3$. For $\mathcal{S}$ the minimum and maximum values are, respectively, -0.5547 ($r_1 = -2, r_2 = 2$) and 0.2180 ($r_1 = 2, r_2 = 2$). For $\mathcal{K}$ the minimum and maximum values are, respectively, 2.4596 ($r_1 = 1.75, r_2 = 2$) and 3.5936 ($r_1 = -2, r_2 = 1.5$).

5. Multiplicative PCM with general innovations

The PCMs in (3) and (15) are additive in structure. Chan et al. [12] studied multiplicative PCMs, called threshold heteroscedastic AR models or T-CHARM for short. In its simplest form the model is given by

$$Y_t = \sigma(Y_{t-1})\varepsilon_t, \tag{27}$$

where $\{\varepsilon_t, t \in \mathbb{Z}\}$ are i.i.d. (but not necessarily normal), and $\sigma(Y_{t-1}) = \sigma_j > 0$ for $Y_{t-1} \in \mathbb{R}^{(j)} = (r_{j-1}, r_j]$ where $\{\mathbb{R}^{(j)}; j = 1, \dots, k\}$ defines a partitioning of the real line $\mathbb{R}$ such that $\cup_{j=1}^{k}\mathbb{R}^{(j)} = \mathbb{R}$ and $\mathbb{R}^{(j)} \cap \mathbb{R}^{(j')} = \emptyset$ if $j \neq j'$. Similar to (15), the process underlying (27) is a Markov chain with k states or regimes.

If the density of ε_t is positive on $\mathbb{R}$ and the Lebesgue measure of each regime is positive, then it is straightforward to show that the volatility process $\{\sigma^2(Y_{t-1})\}$ has the following covariance structure

$$\text{Cov}\big(\sigma^2(Y_{t-1}), \sigma^2(Y_{t-\ell})\big) = \sum_{i=1}^{k}\sum_{j=1}^{k}\sigma_i^2\sigma_j^2\delta_{ij}(\ell), \quad (\ell \geq 0), \tag{28}$$

where $\delta_{ij}(\ell)$ satisfies the recursive equations

$$\delta_{ij}(\ell) = \sum_{s=1}^{k}\mathbb{P}(\sigma_s\varepsilon_t \in \mathbb{R}^{(j)})\delta_{si}(\ell-1),$$
$$\delta_{ij}(0) = \mathbb{P}(Y_t \in \mathbb{R}^{(i)} \cup \mathbb{R}^{(j)}) - \mathbb{P}(Y_t \in \mathbb{R}^{(i)})\mathbb{P}(Y_t \in \mathbb{R}^{(j)}),$$

with

$$\mathbb{P}(Y_t \in \mathbb{R}^{(j)}) = \sum_{i=1}^{k}\mathbb{P}(\sigma_i\varepsilon_t \in \mathbb{R}^{(j)})\mathbb{P}(Y_t \in \mathbb{R}^{(i)}) \quad \text{and} \quad \sum_{j=1}^{k}\mathbb{P}(Y_t \in \mathbb{R}^{(j)}) = 1.$$

It is easy to see that the autocovariance function γ_ℓ of $\{Y_t, t \in \mathbb{Z}\}$ satisfies the Yule–Walker equations of a stationary linear ARMA$(k-1, k-1)$ process, which also follows from generalizing the results in Example 4.1.

Example 5.1: Consider (27) for the case $k = 2$, i.e.

$$Y_t = \{\sigma_1\mathbb{I}_{(-\infty,r]}(Y_{t-1}) + \sigma_2\mathbb{I}_{(r,\infty)}(Y_{t-1})\}\varepsilon_t. \tag{29}$$

Using (28), the lag $\ell \geq 0$ covariance structure of (29) is given by

$$\text{Cov}\big(\sigma^2(Y_{t-1}), \sigma^2(Y_{t-\ell})\big) = (\sigma_2^2 - \sigma_1^2)^2\delta(1-\delta)\big\{\mathbb{P}(\sigma_2\varepsilon_t \in \mathbb{R}^{(2)}) - \mathbb{P}(\sigma_1\varepsilon_t \in \mathbb{R}^{(1)})\big\}^{\ell},$$

where

$$\delta \equiv \mathbb{P}(Y_t \in \mathbb{R}^{(1)}) = \frac{\mathbb{P}(\sigma_2\varepsilon_t \in \mathbb{R}^{(2)})}{1 - \mathbb{P}(\sigma_1\varepsilon_t \in \mathbb{R}^{(1)}) + \mathbb{P}(\sigma_2\varepsilon_t \in \mathbb{R}^{(2)})} \in [0, 1). \tag{30}$$

Note the equivalence in structure between (30) and, respectively, (7) and (11).

6. SETAR(2; 1, 1)

6.1. Gaussian innovations

A SETAR(2; 1, 1) process with parameters of opposite signs and the same absolute value, with $r = 0$, and with $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ is given by

$$\begin{aligned} Y_t &= \alpha_1^- Y_{t-1}\mathbb{I}_{(-\infty,0]}(Y_{t-1}) + \alpha_1^+ Y_{t-1}\mathbb{I}_{(0,\infty)}(Y_{t-1}) + \varepsilon_t, \\ \alpha_1^+ &= -\alpha_1^- \equiv \alpha, \quad (0 < \alpha < 1). \end{aligned} \tag{31}$$

Necessary and sufficient conditions for geometric ergodicity of a general SETAR(2; 1, 1) process are $\alpha_1^+ < 1$, $\alpha_1^- < 1$, and $\alpha_1^+\alpha_1^- < 1$. Model (31) has also been called an absolute AR(1) process, and its model formulation is given by $Y_t = \alpha|Y_{t-1}| + \varepsilon_t$.

Proposition 6.1 (Anděl et al. [13]): *The stationary marginal density $f_\alpha(y)$ of the* SETAR(2; 1, 1) *process* (31) *is given by*

$$\begin{aligned} f_\alpha(y) &= 2\Phi_{(-\alpha y)}\varphi\left(y; 0, \frac{1}{1-\alpha^2}\right), \\ &= 2\left(\frac{1-\alpha^2}{2\pi}\right)^{1/2} \exp\{-(1-\alpha^2)y^2/2\}\Phi_{(-\alpha y)}. \end{aligned} \tag{32}$$

Corollary 6.1: *The first four non-central moments of the stationary SETAR* (2; 1, 1) *process* (31) *are given by*

$$\nu_s = (-1)^s\alpha^{-s-1}[2(1-\alpha^2)/\pi]^{1/2}J_s, \quad (s = 1, \ldots, 4), \tag{33}$$

where

$$\begin{aligned} J_1 &= -k^{-1}(k+1)^{-1/2}, \ J_2 = k^{-1}\big(\pi/(2k)\big)^{1/2}, \\ J_3 &= \big(2/k + 1/(k+1)\big)^{1/2}J_1, \ J_4 = (3/k)J_2, \end{aligned}$$

with $k = (1-\alpha^2)/\alpha^2$.

Corollary 6.2: *The lag one covariance coefficient, $\gamma_1 = \mathbb{E}(Y_tY_{t-1})$, is given by*

$$\gamma_1 = \frac{\alpha}{1-\alpha^2} + \frac{2\alpha^2}{\pi(1-\alpha^2)^{1/2}} - \frac{2\alpha}{\pi(1-\alpha^2)}\arctan\left(\sqrt{\frac{1-\alpha^2}{\alpha^2}}\right). \tag{34}$$

Remark 6.1: Anděl et al. [13] overlooked that the parameter α should be introduced in the numerator of the third term of (34). For $\ell \geq 2$, γ_ℓ has not been obtained.

Remark 6.2: Das and Genton [14] obtained an explicit expression for the stationary marginal pdf of a multivariate (m-dimensional) SETAR(2; 1, 1) process with multivariate Gaussian innovations. Also formulae for measuring multivariate skewness and kurtosis are derived. In addition, these authors characterize the stationary marginal pdf for a subclass of the multivariate SETAR(2; 1, 1) process with central symmetric multivariate innovations.

Remark 6.3: There exists an interesting duality between the stationary marginal density in Proposition 6.1 and the marginal density of a time series $\{Y_t, t \in \mathbb{Z}\}$ following an asymmetric moving average (asMA) of order $(1, 1)$. In particular, consider the following asMA$(1, 1)$ process

$$Y_t = \beta_1^- \varepsilon_{t-1}\mathbb{I}_{(-\infty,0]}(\varepsilon_{t-1}) + \beta_1^+ \varepsilon_{t-1}\mathbb{I}_{(0,\infty)}(\varepsilon_{t-1}) + \varepsilon_t, \quad |\beta_1^-|, |\beta_1^+| < 1, \quad \beta_1^+ = -\beta_1^- \equiv \beta, \tag{35}$$

where $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$. From Brännäs and De Gooijer [15, Appendix], it can be deduced that the marginal density of $\{Y_t, t \in \mathbb{Z}\}$ is given by

$$f_\beta(y) = \frac{2}{\{(1+\beta^2)2\pi\}^{1/2}} \exp\{-(1+\beta^2)^{-1}y^2/2\}\Phi_{(-\beta y/(1+\beta^2)^{1/2})}. \tag{36}$$

The form of $f_\beta(y)$ is unimodal and symmetric. Note that by setting $\alpha = \beta/(1+\beta)^{1/2}$ in model (31) the marginal density (32) is completely identical to the marginal density (36). Thus in this special case, exact moment expressions of a SETAR$(2; 1, 1)$ process can be obtained from (36) directly. Using $f_\beta(y)$ the first four non-central moments of (35) are given by

$$\nu_1 = \frac{2\beta}{\sqrt{2\pi}}, \quad \nu_2 = (1+\beta^2), \quad \nu_3 = \frac{3}{2\sqrt{2\pi}}\left(2 + \frac{4}{3}\beta^2\right), \quad \nu_4 = 3(1+\beta^2)^2, \tag{37}$$

and the lag $\ell \geq 1$ covariance function is given by $\gamma_\ell = 0$.

Wecker [16] provides exact expressions for the mean, variance and autocovariances of a general asMA(q, q) process. We refer to Brännäs and De Gooijer [15, Appendix] for explicit/closed-form expressions of the marginal pdf and non-central moments of $\{Y_t, t \in \mathbb{Z}\}$ following an ARasMA process with AR(1) and asMA$(1, 1)$ lag polynomials. De Gooijer [17] gives explicit expressions for the mean vector, variance-covariance matrix, and lag ℓ cross-covariance matrix for an m-dimensional asymmetric vector moving average of order (q, q) with multivariate Gaussian innovations.

Example 6.1: Table 1 contains exact values of the mean, variance, skewness, kurtosis, and lag one autocorrelation coefficient of the SETAR$(2; 1, 1)$ process (31) for various parameter values α. However, using the density (32) it is impossible to obtain explicit exact formulae for ρ_ℓ $(\ell \geq 2)$. This problem can be "solved" by obtaining an approximation $\widetilde{\rho}_\ell$ of ρ_ℓ $(\ell \geq 2)$ which depends on a constant $c(\alpha)$ such that $\widetilde{\rho}_\ell/\widetilde{\rho}_{\ell-1}$, with $\widetilde{\rho}_1 \equiv \rho_1$. Column 7 of Table 1 shows $c(\alpha)$ values obtained for a time series $\{Y_t\}_{t=1}^T$ of length $T = 20,000$, using 20,000 replications. Table 2 contains the "true" values of ρ_ℓ, based on the simulation study, and the approximate values $\widetilde{\rho}_\ell$ for various values of α with $\ell = 1, \ldots, 5$. It is clear that the approximation method is excellent since the values of $\widetilde{\rho}_\ell$ are identical to those of ρ_ℓ in almost all cases.

Remark 6.4: Closely related to a SETAR process is a self-exciting threshold MA (SETMA) process. De Gooijer [18] derived an approximate expression for the lag ℓ $(\ell \geq 1)$ covariance function γ_ℓ of a two-regime SETMA process of order (q_1, q_2) and with i.i.d. $\mathcal{N}(0, 1)$ innovations. Li [19] obtained an explicit/closed form expression for the autocorrelation

Table 1. Exact values of the mean, variance, skewness, kurtosis, and ρ_1 of the SETAR(2; 1, 1) process (31) ($\mathcal{N}(0,1)$ innovations). The factor $c(\alpha)$ can be used to obtain approximate values of ρ_ℓ ($\ell = 2, 3, \ldots$); see Section 6.1.

α	Mean (μ)	Variance (σ^2)	Skewness ($\mathcal{S}$)	Kurtosis ($\mathcal{K}$)	ρ_1	$c(\alpha)$
0.1	0.0802	1.0037	0.0002	3.0000	0.0064	–
0.2	0.1629	1.0151	0.0018	3.0002	0.0258	0.04
0.3	0.2509	1.0359	0.0064	3.0010	0.0589	0.08
0.4	0.3482	1.0692	0.0164	3.0036	0.1072	0.16
0.5	0.4607	1.1210	0.0353	3.0101	0.1729	0.24
0.6	0.5984	1.2044	0.0696	3.0250	0.2594	0.35
0.7	0.7821	1.3491	0.1310	3.0582	0.3726	0.48
0.8	1.0638	1.6460	0.2447	3.1339	0.5220	0.63
0.9	1.6474	2.5491	0.4715	3.3210	0.7241	0.80

Notes: (1) The values of $c(\alpha)$ are based on 20, 000 replications of the SETAR(2; 1, 1) process with series of length $T = 20,000$; (2) For $-1 < \alpha < 0$, the exact values of μ and $\mathcal{S}$ have negative signs.

Table 2. Comparing two computed autocorrelation coefficients of the SETAR(2; 1, 1) process (31) ($\mathcal{N}(0,1)$ innovations): 1) ρ_ℓ obtained via Monte Carlo simulation, and 2) $\widetilde{\rho}_\ell = c(\alpha)\widetilde{\rho}_{\ell-1}$ with $\widetilde{\rho}_1 = \rho_1$.

	$\alpha = 0.2$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.8$	
ℓ	ρ_ℓ	$\widetilde{\rho}_\ell$	ρ_ℓ	$\widetilde{\rho}_\ell$	ρ_ℓ	$\widetilde{\rho}_\ell$	ρ_ℓ	$\widetilde{\rho}_\ell$
1	0.03	0.03	0.11	0.11	0.26	0.26	0.52	0.52
2	0.00	0.00	0.02	0.02	0.09	0.09	0.32	0.33
3	0.00	0.00	0.00	0.00	0.03	0.03	0.20	0.21
4	−0.00	0.00	0.00	0.00	0.01	0.01	0.13	0.13
5	−0.00	−0.000	0.00	0.00	0.00	0.00	0.08	0.08

Note: Using (34), the exact values of ρ_1 are: 0.0258 ($\alpha = 0.2$), 0.1072 ($\alpha = 0.4$), 0.2594 ($\alpha = 0.6$), and 0.5220 ($\alpha = 0.8$).

function ρ_ℓ of a simple three-regime SETMA process of order $(1, 1)$. The model is defined as

$$Y_t = \beta_j \varepsilon_{t-1} \mathbb{I}_{(r_{j-1}, r_j]}(Y_{t-2}) + \varepsilon_t, \quad (j = 1, 2, 3), \tag{38}$$

where $\{\varepsilon_t\}$ is i.i.d. with mean zero and $\mathbb{E}(\varepsilon_t^2) < \infty$. Let $a_j = \beta_j \varepsilon_{t-1} + \varepsilon_t$ $(j = 1, 2, 3)$. Then the exact result for ρ_ℓ is given by

$$\rho_1 = \frac{\beta_2 + (\beta_1 - \beta_2)w_1 + (\beta_3 - \beta_2)w_2}{1 + \beta_2^2 + (\beta_1^2 - \beta_2^2)w_1 + (\beta_3^2 - \beta_2^2)w_2} \quad \text{and} \quad \rho_\ell = 0 \quad \text{for } \ell \geq 2, \tag{39}$$

where

$$w_1 = \frac{\mathbb{P}(a_2 \leq r_1)\mathbb{P}(a_3 \leq r_2) + \mathbb{P}(a_2 > r_2)\mathbb{P}(a_3 \leq r_1)}{\Delta(r_1, r_2)} \in [0, 1),$$

$$w_2 = \frac{\mathbb{P}(a_1 > r_1)\mathbb{P}(a_2 > r_2) + \mathbb{P}(a_1 > r_2)\mathbb{P}(a_2 \leq r_1)}{\Delta(r_1, r_2)} \in [0, 1),$$

with

$$\Delta(r_1, r_2) = \{\mathbb{P}(a_1 > r_1) + \mathbb{P}(a_2 \leq r_1)\}\{\mathbb{P}(a_2 > r_2) + \mathbb{P}(a_3 \leq r_2)\} - \{\mathbb{P}(a_2 \leq r_1) - \mathbb{P}(a_3 \leq r_1)\}\{\mathbb{P}(a_1 \leq r_2) - \mathbb{P}(a_2 \leq r_2)\}.$$

Furthermore, the distribution function $F_{\beta_1,\beta_2,\beta_3}(\cdot)$ of $\{Y_t, t \in \mathbb{Z}\}$ takes the form

$$F_{\beta_1,\beta_2,\beta_3}(x) = w_1 \mathbb{P}(a_1 \leq x) + (1 - w_1 - w_2)\mathbb{P}(a_2 \leq x) + w_2 \mathbb{P}(a_3 \leq x), \tag{40}$$

which is a weighted average of $\mathbb{P}(a_j \le x)$'s $(j = 1, 2, 3)$.

6.2. Laplace innovations

Loges [20] derived explicit expressions of the stationary marginal density of a SETAR(2; 1, 1) processes with Laplace innovations in three parameter cases. These expressions are not in closed form since they involve infinite sums of certain functions (see below). Now a random variable X has a symmetric Laplace distribution with location parameter μ and shape parameter b, denoted by $\mathcal{L}(\mu, b)$, if its pdf is $f_X(x; \mu, b) = (1/2b)\exp(-|x - \mu|/b)$. Without loss of generality, we consider SETAR(2; 1, 1) processes with $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{L}(0, 1)$. This distribution implies that $\mathbb{E}(\varepsilon_t^s) = 0$ when s odd, and $\mathbb{E}(\varepsilon_t^s) = s!$ when s is even. We summarize and expand on Loges [20] results. In particular, new results are given in Corollary 6.4, Corollary 6.6, and Corollary 6.8.

Case 6.1: A SETAR(2; 1, 1) process with positive parameters, with $r = 0$, and with $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{L}(0, 1)$ is given by

$$Y_t = \alpha_1^- Y_{t-1}\mathbb{I}_{(-\infty,0]}(Y_{t-1}) + \alpha_1^+ Y_{t-1}\mathbb{I}_{(0,\infty)}(Y_{t-1}) + \varepsilon_t, \quad 0 < \alpha_1^-, \quad \alpha_1^+ < 1. \tag{41}$$

To simplify notation, we introduce the following functions:

$$g_1(\lambda) \coloneqq \sum_{q=0}^{\infty} d_1(\lambda, q), \quad g_2(\lambda) \coloneqq 1 - \sum_{q=0}^{\infty}\left(\frac{1}{2(1-\lambda^{q+1})}\right) d_1(\lambda, q)$$

and

$$g_3(\lambda) \coloneqq g_1(\lambda)/g_2(\lambda), \tag{42}$$

where

$$d_1(\lambda, 0) \coloneqq 1, \quad \text{and} \quad d_1(\lambda, q) \coloneqq \prod_{j=0}^{q-1}\left(\frac{\lambda^{2j+2}}{\lambda^{2j+2} - 1}\right) \quad \text{for } q \ge 1, \quad 0 \le \lambda < 1. \tag{43}$$

Proposition 6.2: *The invariant marginal density $f_{\alpha_1^-,\alpha_1^+}(y)$ of the* SETAR(2; 1, 1) *process* (41) *is given by*

$$f_{\alpha_1^-,\alpha_1^+}(y) = 2g(\alpha_1^-, \alpha_1^+)\big[g_3(\alpha_1^-)f_{\alpha_1^-}(y)\mathbb{I}_{(-\infty,0]}(y) + g_3(\alpha_1^+)f_{\alpha_1^+}(y)\mathbb{I}_{(0,\infty)}(y)\big], \tag{44}$$

where

$$g(\alpha_1^-, \alpha_1^+) = \big(g_3(\alpha_1^-) + g_3(\alpha_1^+)\big)^{-1} \quad \textit{for all } 0 \le \alpha_1^-, \alpha_1^+ < 1,$$
$$f_{\alpha_1^\pm}(y) = \big(2g_1(\alpha_1^\pm)\big)^{-1}\sum_{q=0}^{\infty} d_1(\alpha_1^\pm, q)(\alpha_1^\pm)^{-q}\exp\big(-(\alpha_1^\pm)^{-q}|y|\big).$$

Remark 6.5: It can be proved that $f_{\alpha_1^\pm}(y)$ is the stationary marginal density of an AR(1) process with parameter $|\alpha_1^\pm| < 1$ and $\mathcal{L}(0, 1)$ innovations; Loges [20]. Clearly, $f_{\alpha_1^-,\alpha_1^+}(y)$ is a mixture of pdfs $f_{\alpha_1^-}(y)$ and $f_{\alpha_1^+}(y)$ with weights $2g(\alpha_1^-,\alpha_1^+)g_3(\alpha_1^-)$ and $2g(\alpha_1^-,\alpha_1^+)g_3(\alpha_1^+)$, respectively. The weights are non-negative and they do not add up to one. The shape of $f_{\alpha_1^-,\alpha_1^+}(y)$ is unimodal (not symmetric) with vertex at the origin.

Corollary 6.3: *The non-central moments of the stationary* SETAR(2; 1, 1) *process* (41) *can be expressed as*

$$\nu_{2s-1} = 2g(\alpha_1^-,\alpha_1^+)(2s-1)\left[\left\{\prod_{j=1}^{s}\left(1-(\alpha_1^+)^{2j-1}\right)\right\}^{-1} \left\{\prod_{j=1}^{s}\left(1-(\alpha_1^-)^{2j-1}\right)\right\}^{-1}\right],$$

$$\nu_{2s} = 2g(\alpha_1^-,\alpha_1^+)(2s)!\left[g_3(\alpha_1^+)\left\{\prod_{j=1}^{s}\left(1-(\alpha_1^+)^{2j}\right)\right\}^{-1} + g_3(\alpha_1^-)\left\{\prod_{j=1}^{s}\left(1-(\alpha_1^-)^{2j}\right)\right\}^{-1}\right],$$

for all $s \in \mathbb{N}$.

Remark 6.6: For all $s \in \mathbb{N}$, the following relations can be used to re-express the non-central moments in Corollary 6.3:

$$g_4(\lambda, 2s) = g_1(\lambda)\left\{\prod_{j=1}^{s}(1-\lambda^{2j})\right\}^{-1} \quad \text{and}$$

$$g_4(\lambda, 2s-1) = \frac{2g_1(\lambda)}{g_3(\lambda)}\left\{\prod_{j=1}^{s}(1-\lambda^{2j-1})\right\}^{-1},$$

where

$$g_4(\lambda, s) = \sum_{q=0}^{\infty} d_1(\lambda, q)\lambda^{qs}, \quad (0 \le \lambda < 1). \tag{45}$$

Corollary 6.4: *The lag one covariance coefficient,* $\gamma_1 = \mathbb{E}(Y_t Y_{t-1})$, *is given by*

$$\gamma_1 = 2g(\alpha_1^-,\alpha_1^+)\left\{\alpha_1^-\frac{g_3(\alpha_1^-)}{g_1(\alpha_1^-)}g_4(\alpha_1^-,2) + \alpha_1^+\frac{g_3(\alpha_1^+)}{g_1(\alpha_1^+)}g_4(\alpha_1^+,2)\right\}. \tag{46}$$

Case 6.2: A SETAR(2; 1, 1) process with parameters of opposite signs, $r = 0$, and with $\mathcal{L}(0, 1)$ innovations is given by

$$Y_t = \alpha_1^- Y_{t-1}\mathbb{I}_{(-\infty,0)}(Y_{t-1}) + \alpha_1^+ Y_{t-1}\mathbb{I}_{[0,\infty)}(Y_{t-1}) + \varepsilon_t, \quad -1 < \alpha_1^- < 0, \quad 0 < \alpha_1^+ < 1. \tag{47}$$

In addition to the notations introduced with Case 6.1, we define the following functions:

$$g_5(\alpha_1^-,\alpha_1^+) = \sum_{q=0}^{\infty} d_2(\alpha_1^-,\alpha_1^+,q), \quad g_6(\alpha_1^-,\alpha_1^+) = \sum_{q=0}^{\infty} \frac{1}{2\big(1-(\alpha_1^+)^q|\alpha_1^-|\big)} d_2(\alpha_1^-,\alpha_1^+,q),$$

$$g_7(\alpha_1^-,\alpha_1^+,p) = |\alpha_1^-|^p \sum_{q=1}^{\infty} d_2(\alpha_1^-,\alpha_1^+,q)(\alpha_1^+)^{(q-1)s}, \quad s\in\mathbb{N},$$

$$g_8(\alpha_1^-,\alpha_1^+) = g_2(\alpha_1^+)/g_6(\alpha_1^-,\alpha_1^+),$$

$$g_9(\alpha_1^-,\alpha_1^+) \coloneqq \big(g_3(\alpha^+)g_6(\alpha_1^-,\alpha_1^+) + g_5(\alpha_1^-,\alpha_1^+)\big)^{-1}, \tag{48}$$

where

$$d_2(\alpha_1^-,\alpha_1^+,0) = 1, \quad d_2(\alpha_1^-,\alpha_1^+,q) = \prod_{j=0}^{q-1} \frac{(\alpha_1^+)^{2j}(\alpha_1^-)^2}{(\alpha_1^+)^{2j}(\alpha_1^-)^2 - 1}.$$

Proposition 6.3: *The invariant marginal density $f_{\alpha_1^-,\alpha_1^+}(y)$ of the SETAR process* (47) *is given by*

$$f_{\alpha_1^-,\alpha_1^+}(y) = h(\alpha_1^-,\alpha_1^+)\Big(g_8(\alpha_1^-,\alpha_1^+)\exp(y)\mathbb{I}_{(-\infty,0]}(y) + \big(r_1(y,\alpha_1^+) + g_8(\alpha_1^-,\alpha_1^+)r_2(y,\alpha_1^-,\alpha_1^+)\big)\mathbb{I}_{(0,\infty)}(y)\Big), \tag{49}$$

where

$$h(\alpha_1^-,\alpha_1^+) = \big(g_1(\alpha_1^+) + g_5(\alpha_1^-,\alpha_1^+)g_8(\alpha_1^-,\alpha_1^+)\big)^{-1},$$

$$r_1(y,\alpha_1^+) = \sum_{q=0}^{\infty} d_1(\alpha_1^+,q)(\alpha_1^+)^{-q}\exp(-(\alpha_1^+)^{-q}y),$$

$$r_2(y,\alpha_1^-,\alpha_1^+) = \sum_{q=1}^{\infty} d_2(\alpha_1^-,\alpha_1^+,q)(\alpha_1^+)^{1-q}|\alpha_1^-|^{-1}\exp(-(\alpha_1^+)^{1-q}|\alpha_1^-|^{-1}y). \tag{50}$$

Remark 6.7: The shape of the function $f_{\alpha_1^-,\alpha_1^+}(y)$ is unimodal but not symmetric.

Corollary 6.5: *The non-central moments of the stationary* SETAR(2; 1, 1) *process* (47) *can be expressed as*

$$\nu_{2s-1} = g_9(\alpha_1^-,\alpha_1^+)(2s-1)!\left[2g_6(\alpha_1^-,\alpha_1^+)\left\{\prod_{j=1}^{s}\big(1-(\alpha_1^+)^{2j-1}\big)\right\}^{-1} - 1 + g_7(\alpha_1^-,\alpha_1^+,2s-1)\left\{\prod_{j=1}^{s}\big(1-(\alpha_1^+)^{2j-1}\big)\right\}^{-1}\right],$$

$$\nu_{2s} = g_9(\alpha_1^-, \alpha_1^+)(2s)! \left[g_3(\alpha_1^+) g_6(\alpha_1^-, \alpha_1^+) \left\{ \prod_{j=1}^{s} \left(1 - (\alpha_1^+)^{2j}\right) \right\}^{-1} + 1 + g_7(\alpha_1^-, \alpha_1^+, 2s) \right],$$

for all $s \in \mathbb{N}$.

Corollary 6.6: *The lag one covariance coefficient,* $\gamma_1 = \mathbb{E}(Y_t Y_{t-1})$*, is given by*

$$\gamma_1 = 2\alpha_1^- h(\alpha_1^-, \alpha_1^+) g_8(\alpha_1^-, \alpha_1^+) + 2\alpha_1^+ h(\alpha_1^-, \alpha_1^+) g_4(\alpha_1^+, 2) + 2\alpha_1^+ h(\alpha_1^-, \alpha_1^+) g_8(\alpha_1^-, \alpha_1^+) g_7(\alpha_1^+, 2). \tag{51}$$

Case 6.3: As a special case of model specification (47), we consider a SETAR(2; 1, 1) process with parameters of opposite signs and the same absolute value, $r = 0$, and with $\mathcal{L}(0, 1)$ innovations, i.e.,

$$Y_t = \alpha_1^- Y_{t-1} \mathbb{I}_{(-\infty,0]}(Y_{t-1}) + \alpha_1^+ Y_{t-1} \mathbb{I}_{(0,\infty)}(Y_{t-1}) + \varepsilon_t, \;\; 0 < \alpha_1^- < 1, \; \alpha_1^- = -\alpha_1^+ \equiv \alpha. \tag{52}$$

From the formulas in Case 6.1 and Case 6.2, it can be deduced that

$$h(\alpha, -\alpha) g_8(\alpha, -\alpha) \equiv \frac{1}{g_3(\alpha)}, \quad d_1(\alpha, q) \equiv d_3(\alpha, -\alpha, q), \tag{53a}$$

$$g_1(\alpha) \equiv g_5(\alpha, -\alpha), \tag{53b}$$

$$h(-\alpha, \alpha) + \frac{1}{g_3(\alpha)} \equiv \frac{1}{g_1(\alpha)}. \tag{53c}$$

Then the following proposition is a special case of Proposition 6.3.

Proposition 6.4: *The invariant marginal density* $f_\alpha(y)$ *of the SETAR*(2; 1, 1) *process* (52) *is given by*

$$f_\alpha(y) = \frac{\exp(y)}{g_3(\alpha)} \mathbb{I}_{(-\infty,0]}(y) + \left(\frac{r(y, \alpha)}{g_1(\alpha)} - \frac{\exp(-y)}{g_3(\alpha)} \right) \mathbb{I}_{(0,\infty)}(y), \tag{54}$$

where $r(y, \alpha) = \sum_{q=0}^{\infty} d_1(\alpha, q) \alpha^{-q} \exp(-\alpha^{-q} y)$.

Remark 6.8: For $0 \le \alpha < 1$, the infinite sum $r(y, \alpha)$ quickly goes to zero as q increases. In that case, $f_\alpha(y)$ is a mixture of a standard inverse exponential pdf and a standard exponential pdf with weights $1/g_3(\alpha)$ and $-1/g_3(\alpha)$, respectively. The form of $f_\alpha(y)$ is unimodal but not symmetric.

Corollary 6.7: *The non-central moments of the stationary* SETAR(2; 1, 1) *process* (52) *can be expressed as*

$$\nu_{2s-1} = \frac{2(2s-1)!}{g_3(\alpha)} \left[\left\{ \prod_{j=1}^{s} (1 - \alpha^{2j-1}) \right\}^{-1} - 1 \right], \quad \nu_{2s} = (2s)! \left\{ \prod_{j=1}^{s} \left(1 - \alpha^{2j}\right) \right\}^{-1}$$

for all $s \in \mathbb{N}$.

Table 3. Exact values of the mean, variance, skewness, kurtosis, and ρ_1 of the SETAR(2; 1, 1) process (52) ($\mathcal{L}(0,1)$ innovations).

α	Mean (μ)	Variance(σ^2)	Skewness ($\mathcal{S}$)	Kurtosis ($\mathcal{K}$)	ρ_1
0.1	0.1009	2.0100	0.0007	5.9703	0.0051
0.2	0.2069	2.0405	0.0055	5.8844	0.0209
0.3	0.3228	2.0936	0.0181	5.7489	0.0493
0.4	0.4541	2.1748	0.0418	5.5704	0.0926
0.5	0.6091	2.2956	0.0799	5.3526	0.1538
0.6	0.8025	2.4810	0.1372	5.0949	0.2374
0.7	1.0634	2.7908	0.2216	4.7916	0.3498
0.8	1.4662	3.4060	0.3502	4.4359	0.5016
0.9	2.3004	5.2345	0.5644	4.0477	0.7110

Note: see note 2) to Table 1.

Corollary 6.8: *The lag one covariance coefficient,* $\gamma_1 = \mathbb{E}(Y_t Y_{t-1})$, *is given by*

$$\gamma_1 = -\frac{4\alpha}{g_3(\alpha)} + \frac{2\alpha}{g_1(\alpha)} g_4(\alpha, 2). \tag{55}$$

Remark 6.9: Note that the formula for the even non-central moments, ν_{2s}, is the same as for the moments of a stationary *linear* AR(1) process with $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} (0,1)$ innovations.

Remark 6.10: In all Cases 6.1–6.3, the functions $d_1(\cdot, q)$ and $d_2(\cdot, q)$ converge quickly to zero as q increases and, hence, the infinite sums $g_5(\cdot), g_6(\cdot), g_7(\cdot)$ are equal zero for some low values of q. In fact, one can safely set $q = 11$, which gives the same moment values (agreement to four decimal places) when cutting the infinite sums off at $q = 1,001$, which we adopt for obtaining the numerical results of this paper.

Example 6.2: Table 3 shows exact values of the mean, variance, skewness, kurtosis, and the lag one autocorrelation coefficient (ρ_1) of the SETAR(2; 1, 1) process (52) for various parameter values α. We see that the distribution function is positively skewed for all values of α.

Remark 6.11: Recall the conditions for stationarity of a general SETAR(2; 1, 1) process with delay $d = 1$ given in Section 6.1. Given these conditions, Tables 1 and 3 show that as α approaches unity, the process $\{Y_t, t \in \mathbb{Z}\}$ has a higher variability (as indicated by its variance σ^2) than for a low value of α. Clearly, the variability is higher for $\mathcal{L}(0,1)$ innovations (Table 3) than for $\mathcal{N}(0,1)$ innovations (Table 1). There is also a lack of symmetry in both cases (as indicated by the skewness values $\mathcal{S}$) with stronger evidence for SETAR(2; 1, 1) processes with $\mathcal{L}(0,1)$ innovations.

Remark 6.12: For the asMA(1, 1) process (35) with $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{L}(0,1)$ it is easy to show that the first four non-central moments are given by

$$\nu_1 = \frac{1}{2}(\beta_1^+ - \beta_1^-), \quad \nu_2 = 1 + (\beta_1^+)^2 + (\beta_1^-)^2 + \frac{1}{2}\beta_1^+\beta_1^-,$$

$$\nu_3 = 3(\beta_1^+ - \beta_1^-) + \frac{3}{4}\big((\beta_1^-)^2 - (\beta_1^+)^2\big) + 3\big((\beta_1^+)^3 - (\beta_1^-)^3\big),$$

$$\nu_4 = 24 + 12\big((\beta_1^+)^2 + (\beta_1^-)^2\big) + 12\big((\beta_1^+)^4 + (\beta_1^-)^4\big).$$

6.3. Cauchy innovations

A random variable X has a Cauchy distribution with location parameter zero, and scale parameter λ ($\lambda > 0$), denoted by $\mathcal{C}(0,\lambda)$, if its pdf is given by $f_X(x;0,\lambda) = (\pi\lambda)^{-1}(1 + (x/\lambda)^2)$. For $\lambda = 1$, this distribution coincides with the Student t_1-distribution. Anděl and Bartoň [21] obtained an explicit expression for the stationary marginal density $f_\alpha(y)$ of the SETAR(2; 1, 1) process defined in (31) with parameter α ($|\alpha| < 1$) and assuming that $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{C}(0,1)$. However, Li and Tong [22] noted that the expression for $f_\alpha(y)$ contains an error, resulting in negative values of the density for some special cases of y and α. The correct version of the density is given in Proposition 6.5.

Proposition 6.5 (Li and Tong [22]): *Denote $A = \alpha/(1-|\alpha|)$ ($|\alpha| < 1$). Then the stationary marginal density $f_\alpha(y)$ of a* SETAR(2; 1, 1) *process with a single parameter α and $\mathcal{C}(0,1)$ innovations is given by*

$$f_\alpha(y) = \frac{2}{\pi^2[y^2+(A-1)^2][y^2+(A+1)^2]} \times \left\{ Ay\log\left(\frac{y^2+1}{A^2}\right) + A(y^2+A^2-1)\arctan(y) + \frac{|A|\pi}{2}(y^2+A^2-1) + \frac{\pi}{2}(y^2-A^2+1) \right\}. \quad (56)$$

Remark 6.13: The shape of the function $f_\alpha(y)$ is unimodal but not symmetric. For $\alpha > 0$ the distribution function is positively skewed, and for $\alpha < 0$ negatively skewed.

Remark 6.14: Consider the asMA(1,1) process in (36) but now with $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{C}(0,1)$. Let $\mathfrak{g}(\epsilon)$ be the density of $\{\varepsilon_t\}$. Then the marginal density of $\{Y_t\}$ is given by $f_\beta(y) = \int_{-\infty}^0 f(y|\epsilon)\mathfrak{g}(\epsilon)\mathrm{d}\epsilon + \int_0^\infty f(y|\epsilon)\mathfrak{g}(\epsilon)\mathrm{d}\epsilon$, where $f(y|\epsilon)$ is the regime-specific conditional density. After some algebra, it can be deduced that the marginal density is given by

$$f_\beta(y) = \frac{2}{\pi^2[y^2+(\beta-1)^2][y^2+(\beta+1)^2]} \times \left\{ \beta y\log\left(\frac{y^2+1}{\beta^2}\right) + \beta(y^2+\beta^2-1)\arctan(y) + \frac{|\beta|\pi}{2}(y^2+\beta^2-1) + \frac{\pi}{2}(y^2-\beta^2+1) \right\}. \quad (57)$$

Clearly, if $\beta = A = \alpha/(1-|\alpha|)$ ($|\alpha| < 1$) the density (57) is the same as the density (56). This is another example of the duality between the stationary marginal densities of a special form of an asMA(1,1) process and a special form of a SETAR(2; 1, 1) process; see also Remark 6.3.

7. Approximations

7.1. Markov chain

Anděl et al. [13] provided a Markov chain approximation as well as a numerical method to approximate the stationary pdf of the SETAR(2; 1, 1) process defined in (31) ($\mathcal{N}(0,1)$ innovations). However, the Markov chain approach gives only rough estimates of a number of distribution characteristics when compared to the exact solution. One reason is that its accuracy depends on the number of states and their location. The numerical method provides a more accurate solution, but its quality decreases for AR parameters approaching unity in absolute value. More importantly, the method induces cumulative errors due to repeated, recursive, iterations used for solving an integral equation. Other numerical integration methods for approximating the stationary pdf of nonlinear processes are provided by Tong [23, Section 4.2]. In general, these methods are quite complicated to handle, and they depend heavily on certain tuning parameters.

7.2. Riemann–Stieltjes integration

As an alternative to the numerical methods discussed by Tong [23], Li and Qiu [24] provide a simple numerical method to solve the following integral equation

$$H(x) = \int_{\mathbb{R}} K(x,y)H(\mathrm{d}y), \tag{58}$$

where $H(\cdot)$ is an arbitrary distribution function, and $K(\cdot,\cdot)$ is a bounded, continuous and positive function, i.e. a transition kernel density. Note the equivalence between (58) and (1).

For a positive integer m, let $-\infty = x_0 < x_1 < x_2 < \cdots < x_m < x_{m+1}$ be a partition of $\mathbb{R}$. Then, by the definition of the Riemann–Stieltjes integral, a discrete approximation of (58) is given by

$$\begin{aligned} H(x_k) &= \sum_{j=1}^{m+1} K(x_k, x_j^*)\big(H(x_j) - H(x_{j-1})\big) \\ &= \sum_{j=1}^{m} \Big[K(x_k, x_j^*) - K(x_k, x_{j+1}^*)\Big]H(x_j) + K(x_k, x_{m+1}^*), \end{aligned} \tag{59}$$

where $x_j^* \in [x_{j-1}, x_j]$ $(j = 1, \ldots, m+1)$. A convenient choice is to take $x_j^* = (x_{j-1} + x_j)/2$ $(j = 2, \ldots, m)$ and $x_1^* = x_1 - 1$ and $x_{m+1}^* = x_m + 1$. Define the $m \times m$ matrix $\mathbf{K} = (k_{ij})$ with elements $k_{ij} = K(x_i, x_j^*) - K(x_i, x_{j+1}^*)$, the $m \times m$ matrix $\mathbf{H} = (H(x_1), \ldots, H(x_m))^{\mathrm{T}}$, and the $m \times 1$ vector $\mathbf{a} = (K(x_1, x_{m+1}^*), \ldots, K(x_m, x_{m+1}^*))^{\mathrm{T}}$. Then $\mathbf{H} = \mathbf{K}\mathbf{H} + \mathbf{a}$, which yields $\mathbf{H} = (\mathbf{I}_m - \mathbf{K})^{-1}\mathbf{a}$.

Li and Qiu [24] apply the above approximation method to various first-order nonlinear processes, including a two-regime SETARMA process with a regime-dependent constant, and regime-dependent AR(1) and MA(1) terms. For this particular process, Example 7.1 shows results on the kernel $K(x,y)$ of the DGP and on the implied non-central moments.

Example 7.1: Consider a stationary time series process $\{Y_t, t \in \mathbb{Z}\}$ generated by the following two-regime SETARMA model

$$
\begin{aligned}
Y_t &= (\alpha_0^- + \alpha_1^- Y_{t-1} + \beta_1^- \varepsilon_{t-1})\mathbb{I}_{(-\infty,r]}(Y_{t-1}) \\
&\quad + (\alpha_0^+ + \alpha_1^+ Y_{t-1} + \beta_1^+ \varepsilon_{t-1})\mathbb{I}_{(r,\infty)}(Y_{t-1}) + \varepsilon_t, \\
&= V_t + \varepsilon_t,
\end{aligned} \tag{60}
$$

where $V_t = (\alpha_0^- + \alpha_1^- Y_{t-1} + \beta_1^- \varepsilon_{t-1})\mathbb{I}_{(-\infty,r]}(Y_{t-1}) + (\alpha_0^+ + \alpha_1^+ Y_{t-1} + \beta_1^+ \varepsilon_{t-1})\mathbb{I}_{(r,\infty)}$ $(V_{t-1} + \varepsilon_{t-1})$ (a Markov chain), and $\{\varepsilon_t, t \in \mathbb{Z}\}$ is a central symmetric innovation process having an absolutely continuous pdf on $\mathbb{R}$ with density $g_\varepsilon(\cdot)$. Model (60) will be denoted by SETARMA$(2; 1, 1, 1, 1)$. Clearly, the model includes the two-regime PCM of Section 3, and the two-regime SETAR$(2; 1, 1)$ process of Section 6 as special cases. Chan and Goracci [25] derived necessary and sufficient conditions for the ergodicity of invertible SETARMA$(2; 1, 1, 1, 1)$ processes.

Let $G^*(x)$ be the distribution function of the process $\{V_t\}$. Then

$$
G^*(x) = \int_{\mathbb{R}} K(x, y) G^*(\mathrm{d}y),
$$

where for $x \in \mathbb{R}$ the transition kernel $\mathbb{P}(V_{t+1} \le x | V_t = y) \equiv K(x, y)$ is given by

$$
\begin{aligned}
K(x, y) &= \mathbb{P}\left(\varepsilon_t \le \frac{x - \alpha_0^- - \alpha_1^- y}{\alpha_1^- + \beta_1^-},\ \varepsilon_t \le r - y\right) \\
&\quad + \mathbb{P}\left(\varepsilon_t \le \frac{x - \alpha_0^+ - \alpha_1^+ y}{\alpha_1^+ + \beta_1^+},\ \varepsilon_t > r - y\right).
\end{aligned} \tag{61}
$$

In other words, $G^*(x)$ satisfies the form (58) and consequently by the convolution property, it follows that an approximation of the stationary marginal density of the SETARMA process (60) is given by

$$
f^*_{\theta_1^-, \theta_1^+}(y) = \int_{\mathbb{R}} g_\varepsilon(x - u) G^*(\mathrm{d}u), \tag{62}
$$

where θ_1^- and θ_1^+ denote the set of parameters associated with the negative and positive regime, respectively. That is,

$$
\theta_1^- = \{\alpha_0^-, \alpha_1^-, \beta_1^-\}, \quad \theta_1^+ = \{\alpha_0^+, \alpha_1^+, \beta_1^+\}.
$$

To simplify notation, let $\phi_1^- = \alpha_1^- + \beta_1^-$, $\phi_1^+ = \alpha_1^+ + \beta_1^+$, and

$$
\xi_1^- = (x - \alpha_0^- - \alpha_1^- y), \quad \xi_1^+ = (x - \alpha_0^+ - \alpha_1^+ y), \quad \zeta_1^- = \xi_1^- / \phi_1^-, \quad \zeta_1^+ = \xi_1^+ / \phi_1^+.
$$

Now, depending on $\phi_1^+ \gtreqless 0$ and $\phi_1^- \gtreqless 0$, (61) can be expressed into nine distinct cases:

$$K(x,y) = \begin{cases} G^*(\zeta_1^+) - G^*(\min\{(\zeta_1^+, r-y\}) & \\ \quad +G^*(\min\{\zeta_1^-, r-y\}) & \text{if } \phi_1^+ > 0 \quad \text{and} \quad \phi_1^- > 0, \\ G^*(\zeta_1^+) - G^*(\min\{\zeta_1^+, r-y\}) & \\ \quad +G^*(r-y)\mathbb{I}_{[0,\infty)}(\xi_1^-) & \text{if } \phi_1^+ > 0 \quad \text{and} \quad \phi_1^- = 0, \\ G^*(\zeta_1^+) - G^*(\min\{\zeta^+, r-y\}) & \\ \quad +G^*(r-y) - G^*(\min\{\zeta_1^-, r-y\}) & \text{if } \phi_1^+ > 0 \quad \text{and} \quad \phi_1^- < 0, \\ \big(1 - G^*(r-y)\big)\mathbb{I}_{[0,\infty)}(\xi_1^+) & \\ \quad +G^*(\min\{\zeta_1^-, r-y\}) & \text{if } \phi_1^+ = 0 \quad \text{and} \quad \phi_1^- > 0, \\ (1 - G^*(r-y))\mathbb{I}_{[0,\infty)}(\xi_1^+) & \\ \quad +G^*(r-y)\mathbb{I}_{[0,\infty)}(\xi_1^+) & \text{if } \phi_1^+ = 0 \quad \text{and} \quad \phi_1^- = 0, \\ (1 - G^*(r-y))\mathbb{I}_{[0,\infty)}(\xi_1^+) & \\ \quad +G^*(r-y) - G^*(\min\{\zeta_1^-, r-y\}) & \text{if } \phi_1^+ = 0 \quad \text{and} \quad \phi_1^- < 0, \\ 1 - G^*(\max\{\zeta_1^+, r-y\}) & \\ \quad +G^*(\min\{\zeta_1^-, r-y\}) & \text{if } \phi_1^+ < 0 \quad \text{and} \quad \phi_1^- > 0, \\ 1 - G^*(\max\{\zeta_1^+, r-y\}) & \\ \quad +G^*(r-y)\mathbb{I}_{[0,\infty)}(\xi_1^-) & \text{if } \phi_1^+ < 0 \quad \text{and} \quad \phi_1^- = 0, \\ 1 - G^*(\max\{\zeta_1^+, r-y\}) & \\ \quad +G^*(r-y) - G^*(\min\{\zeta_1^-, r-y\}) & \text{if } \phi_1^+ < 0 \quad \text{and} \quad \phi_1^- < 0. \end{cases}$$

For a SETARMA(2; 1, 1, 1, 1) process with $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ the upper and lower limits of the support interval of the distribution of $\{V_t\}$ are $\pm 2.5758 \pm C(\theta_1^-, \theta_1^+)$, where ± 2.5758 are the 0.5th and 99.5th percentiles of an $\mathcal{N}(0,1)$ distribution, and where $C(\theta_1^-, \theta_1^+) \equiv \max\{|\alpha_0^-| + |\alpha_0^+|, 0\} \times \max\{|\alpha_1^-| + |\alpha_1^+|, |\beta_1^-| + |\beta_1^+|\}$. For a SETARMA process with Laplace innovations, the lower and upper limits of the support interval of $\{V_t\}$ are $\pm 9.2103 \pm C(\theta_1^-, \theta^+)$, where ± 9.2103 are the 0.5th and 99.5th percentiles of an $\mathcal{L}(0,1)$ distribution. Similarly, for a SETARMA process with Cauchy innovations the lower and upper limits of the support interval are $\pm 13.4673 \pm C(\theta_1^-, \theta^+)$, where ± 13.4673 are the 0.5th and 99.5th percentiles of a $\mathcal{C}(0,1)$ distribution.

Let $\{y_i\}_{i=-n}^n$ denote the set of $(2n+1)$ equally spaced values, covering the support interval of $\{V_t\}$, and let $\{f^*_{\theta_1^-,\theta_1^+}(y_i)\}_{i=-n}^n$ be the associated set of values of the density function. Then the implied non-central moments of $f^*_{\theta_1^-,\theta_1^+}(y)$ are given by

$$\nu_s = \left\{ \sum_{i=-n}^{n} y_i^s f^*_{\theta_1^-,\theta_1^+}(y_i) \right\} \Big/ (2n+1), \quad (s = 1, 2, \ldots).$$

Thus, we can readily compute the central moments of $f^*_{\theta_1^-,\theta_1^+}(y)$. Finding, however, an expression for the lag one covariance coefficient $\gamma_1 = \mathbb{E}(Y_t Y_{t-1})$ depends on the model specification and the distribution function of the innovation process. For instance, the implied lag one covariance coefficient γ_1 for a SETAR(2; 1, 1) process with $\mathcal{N}(0,1)$

innovations and threshold parameter r is given by

$$\begin{aligned}\gamma_1 = \Bigg\{ & \alpha_1^- \sum_{\substack{i=-n \\ y_i<r}}^{n} f^*_{\alpha_1^-,\alpha_1^+}(y_i) \Big[(1+(\alpha_1^- y_i)^2)\Phi_{(r-\alpha_1^- y_i)} \\ & - (2\pi)^{-1/2}(r+\alpha_1^- y_i)\exp\left(-\frac{1}{2}(r-\alpha_1^- y_i)^2\right)\Big] \\ & + \alpha_1^- \sum_{\substack{i=-n \\ y_i\geq r}}^{n} f^*_{\alpha_1^-,\alpha_1^+}(y_i) \Big[(1+(\alpha_1^+ y_i)^2)\Phi_{(r-\alpha_1^- y_i)} \\ & - (2\pi)^{-1/2}(r+\alpha_1^+ y_i)\exp\left(-\frac{1}{2}(r-\alpha_1^+ y_i)^2\right)\Big] \\ & + \alpha_1^+ \sum_{\substack{i=-n \\ y_i<r}}^{n} f^*_{\alpha_1^-,\alpha_1^+}(y_i) \Big[(1+(\alpha_1^- y_i)^2)\Phi_{(\alpha_1^- y_i-r)} \\ & + (2\pi)^{-1/2}(r+\alpha_1^- y_i)\exp\left(-\frac{1}{2}(r-\alpha_1^- y_i)^2\right)\Big] \\ & + \alpha_1^+ \sum_{\substack{i=-n \\ y_i\geq r}}^{n} f^*_{\alpha_1^-,\alpha_1^+}(y_i) \Big[(1+(\alpha_1^+ y_i)^2)\Phi_{(\alpha_1^+ y_i-r)} \\ & + (2\pi)^{-1/2}(r+\alpha_1^+ y_i)\exp\left(-\frac{1}{2}(r-\alpha_1^+ y_i)^2\right)\Big]\Bigg\} \Big/ (2n+1). \end{aligned} \tag{63}$$

The proof of (63) is relegated to the Appendix. In a similar way an approximate expression of γ_1 for a SETAR(2; 1, 1) process with $\mathcal{L}(0,1)$ innovations can be obtained. In the numerical computation, we use equidistant grid points of y_i on the interval $[-10, 10]$ with step size of 0.1.

Example 7.2: Table 4 shows the implied mean, variance, skewness, and kurtosis for six SETAR(2; 1, 1) processes with $\mathcal{L}(0,1)$ innovations, and $r = 0$. The numbers within parentheses are the corresponding approximate values based on the Riemann–Stieltjes integration method. All densities are unimodal. Note that the quality of the approximation method deteriorates slightly as the α^- parameter approaches unity. The approximation method is excellent for μ and σ^2, but is less satisfactory for $\mathcal{S}$ and $\mathcal{K}$ due to numerical inaccuracies in the computation of the third and fourth non-central moments.

The quality of the approximation method can be further assessed by computing a norm of the vector $\boldsymbol{\nu} = [\nu_i]_{i=1}^{2n+1} := ((f^*_{\alpha_1^-,\alpha_1^+}(y_{-n}) - f_{\alpha_1^-,\alpha_1^+}(y_{-n})), \ldots, (f^*_{\alpha_1^-,\alpha_1^+}(y_n) - f_{\alpha_1^-,\alpha_1^+}(y_n)))$. Table 5 contains values of the Frobenius norm $||\boldsymbol{\nu}||_F = \{\sum_i \nu_i^2\}^{1/2}$, and values of the supremum norm $||\boldsymbol{\nu}||_\infty = \max_i(\nu_i)$ of three SETAR(2; 1, 1) models with parameters of opposite signs, and with $\mathcal{N}(0,1)$, $\mathcal{L}(0,1)$, and $\mathcal{C}(0,1)$ innovations. The exact stationary marginal density function $f_{\alpha_1^-,\alpha_1^+}(\cdot)$ of the processes with Gaussian and Cauchy innovations are based on (32) and (56), respectively. The exact stationary marginal distribution results

Table 4. Exact and approximate (in parentheses) values of the mean, variance, skewness, and kurtosis of various SETAR(2; 1, 1) processes with $\mathcal{L}(0,1)$ innovations, and $r = 0$.

	$\alpha_1^+ = -\alpha_1^- = \alpha$		$\alpha_1^+ = 0.5$		$\alpha_1^+ = 0.5$	
	$\alpha = 0.25$	$\alpha = 0.75$	$\alpha_1^- = 0.1$	$\alpha_1^- = 0.85$	$\alpha_1^- = -0.75$	$\alpha_1^- = -0.9$
μ	0.263 (0.263)	1.239 (1.236)	0.324 (0.324)	−0.969 (−0.962)	0.711 (0.710)	0.769 (0.769)
σ^2	2.064 (2.059)	3.038 (3.017)	2.302 (2.295)	4.719 (4.649)	2.399 (2.390)	2.486 (2.475)
$\mathcal{S}$	0.011 (0.007)	0.279 (0.250)	0.084 (0.079)	−0.395 (−0.348)	0.126 (0.113)	0.177 (0.193)
$\mathcal{K}$	5.822 (5.681)	4.620 (4.433)	5.344 (5.213)	3.937 (3.701)	5.243 (5.098)	5.222 (5.071)

Table 5. Values of the Frobenius norm $||\boldsymbol{\nu}||_F$ and the supremum norm $||\boldsymbol{\nu}||_\infty$ for a SETAR(2; 1, 1) model with $\alpha_1^+ = -\alpha_1^- \equiv \alpha$, and with $\boldsymbol{\nu}$ defined in Section 7.2.

	$\alpha = 0.25$		$\alpha = 0.5$		$\alpha = 0.75$	
Innovations	$\|\|\boldsymbol{\nu}\|\|_F$	$\|\|\boldsymbol{\nu}\|\|_\infty$	$\|\|\boldsymbol{\nu}\|\|_F$	$\|\|\boldsymbol{\nu}\|\|_\infty$	$\|\|\boldsymbol{\nu}\|\|_F$	$\|\|\boldsymbol{\nu}\|\|_\infty$
$\mathcal{N}(0,1)$	0.990	0.2625	0.866	5.913	0.686	0.175
$\mathcal{L}(0,1)$	2.039	1.228	1.217	0.655	0.680	0.277
$\mathcal{C}(0,1)$	7.493	1.249	22.870	4.272	64.707	9.583

All values are multiplied by 10^3.

for the SETAR(2; 1, 1) processes with $\mathcal{L}(0,1)$ innovations are based on formulas given in Propositions (6.2)–(6.4). As can be seen, the quality of the approximation by this particular numerical method is excellent for $\mathcal{N}(0,1)$ and $\mathcal{L}(0,1)$ innovations and irrespective of the value of α. Note, however, that the values for the Cauchy distribution are higher than those reported for $\mathcal{N}(0,1)$ and $\mathcal{L}(0,1)$. This result can be due to the fact that the Cauchy distribution does not possess finite non-central moments of order greater than or equal to 1. Obviously, this is a matter for further research.

Remark 7.1: The results in Example 7.1 can be readily extended to first-order $k \geq 3$ regime SETARMA processes. For instance, in the case $k = 3$ the model is given by

$$\begin{aligned} Y_t &= \sum_{j=1}^{3} \big(\alpha_0^{(j)} + \alpha_1^{(j)} Y_{t-1} + \beta_1^{(j)} \varepsilon_{t-1}\big)\mathbb{I}_{(r_{j-1},r_j]}(Y_{t-1}) + \varepsilon_t \\ &= V_t + \varepsilon_t, \end{aligned} \tag{64}$$

where $V_t = \sum_{j=1}^{3}(\alpha_0^{(j)} + \alpha_1^{(j)} Y_{t-1} + \beta_1^{(j)}\varepsilon_{t-1})\mathbb{I}_{(r_{j-1},r_j]}(V_{t-1} + \varepsilon_{t-1})$, and $\{\varepsilon_t\}$ is defined as in (60). Then, for $x \in \mathbb{R}$ the transition kernel $K(\cdot,\cdot)$ of the process $\{V_t\}$ is given by

$$\begin{aligned} K(x,y) = &\,\mathbb{P}\Big(\varepsilon_t \leq \frac{x - \alpha_0^{(1)} - \alpha_1^{(1)} y}{\alpha_1^{(1)} + \beta_1^{(1)}}, \varepsilon_t \leq r_1 - y\Big) \\ &+ \mathbb{P}\Big(\varepsilon_t \leq \frac{x - \alpha_0^{(2)} - \alpha_1^{(2)} y}{\alpha_1^{(2)} + \beta_1^{(2)}}, r_1 - y < \varepsilon_t \leq r_2 - y\Big) \\ &+ \mathbb{P}\Big(\varepsilon_t \leq \frac{x - \alpha_0^{(3)} - \alpha_1^{(3)} y}{\alpha_1^{(3)} + \beta_1^{(3)}}, \varepsilon_t > r_2 - y\Big). \end{aligned}$$

Section 8 contains an example of the corresponding pdf.

8. Illustration

A well-known DGP is the random walk without drift, which says that the next value of the process $\{Y_t, t \in \mathbb{Z}\}$ is the current value plus an "error" that is not correlated with the previous history. A simple modification of this process is given by

$$Y_t = -r\mathbb{I}_{(-\infty,-r]}(Y_{t-1}) + Y_{t-1}\mathbb{I}_{(-r,r]}(Y_{t-1}) + r\mathbb{I}_{(r,\infty)}(Y_{t-1}) + \varepsilon_t, \tag{65}$$

where $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. The constant term r in the outer regimes trivially makes the process stationary. The stationary marginal density $f_r(y_t)$ of (65) is given by

$$f_r(y_t) = \int_{-\infty}^{-r} \varphi(y_t; -r) f(y_{t-1})\,\mathrm{d}y_{t-1} + \int_{-r}^{r} \varphi(y_t; 1) f(y_{t-1})\,\mathrm{d}y_{t-1} + \int_{r}^{\infty} \varphi(y_t; r) f(y_{t-1})\,\mathrm{d}y_{t-1}. \tag{66}$$

This equation does not have a closed-form solution, but the approximation method of Section 7.2 can readily be used. Specifically, note that (65) is a special case of the SETAR(3; 1, 1, 1) process

$$Y_t = (\alpha_0^{(1)} + \alpha_1^{(1)} Y_{t-1})\mathbb{I}_{(-\infty,r_1]}(Y_{t-1}) + (\alpha_0^{(2)} + \alpha_1^{(2)} Y_{t-1})\mathbb{I}_{(r_1,r_2]}(Y_{t-1}) + (\alpha_0^{(3)} + \alpha_1^{(3)} Y_{t-1})\mathbb{I}_{(r_2,\infty)}(Y_{t-1}) + \varepsilon_t, \tag{67}$$

where $\{\varepsilon_t\} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Now, depending on $\alpha_1^{(1)} \gtreqless 0$, $\alpha_1^{(2)} \gtreqless 0$, and $\alpha_1^{(3)} \gtreqless 0$, there are 27 distinct cases to consider for the evaluation of the transition kernel (61).

Figure 4(a) shows the density $f_r^*(y)$ obtained for $r = 1, 2, 4$, and 6, using the approximation method of Section 7.2. We note here the general pattern of the near-uniformity between the thresholds as r increases, with slight 'shoulder' sections beyond each threshold to accommodate the few observations that are expected to overshoot them. This is intuitively in agreement with the idea that, for large r, the modified random walk results in series values being contained in an interval, symmetric around zero, in which no one value has a greater probability than any other.

Figure 4(b) displays the first 10 lags of the autocorrelation function ρ_ℓ^* associated with the approximation $f_r^*(y)$. We are led to the conclusion that a realization of model (65) will tend to exhibit the same properties as a linear AR(1) process with a parameter value near to unity. As r increases the model approaches non-stationarity, but for lower values of r, the autocorrelations are more moderate. Figure 4(a ,b) together suggests that the simple class of models given by (65) offer a wide range of patterns, seemingly useful additions to an armoury for dealing with non-stationarity.

9. Some concluding remarks

We have obtained explicit and exact expressions of the stationary marginal distribution density of several low-order threshold-type processes with central symmetric innovations.

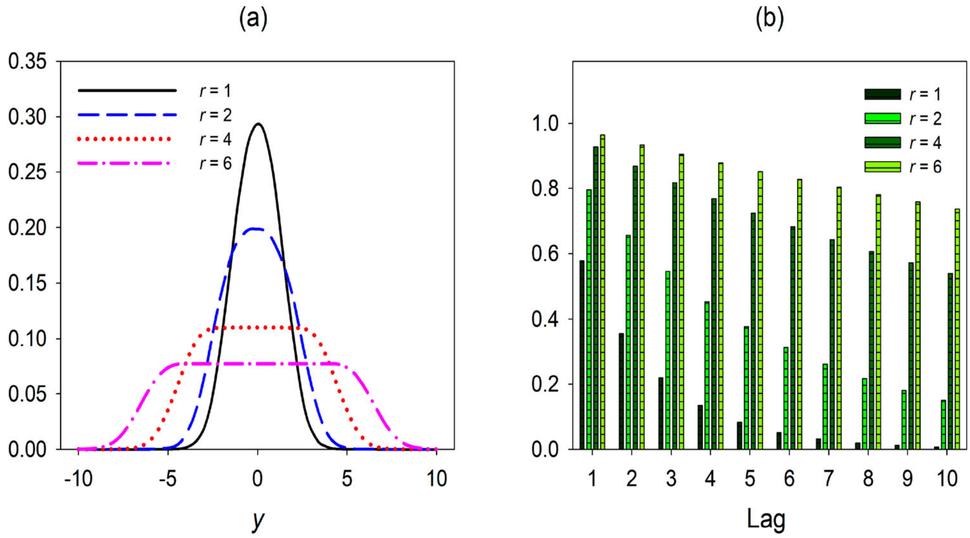

Figure 4. (a) Stationary marginal densities $f_r^*(y)$ and (b) autocorrelations ρ_ℓ^* ($\ell = 1, \ldots, 10$) of the threshold random walk model (65), using the approximation method of Section 7.2.

Also, closed-form and exact expressions for the associated moments, autocovariance and autocorrelation functions have been derived. The marginal distribution can be used to compute the exact likelihood function, and hence improve statistical inference. In empirical studies, comparing the empirical and the exact marginal distributions offers insight about the accuracy of the model fit. This may further be enhanced by a comparison of the sample- and exact autocorrelation function. In addition, the exact lag one autocorrelations may be used to generate pseudo-random low-order threshold-type processes.

For high-order and multiple-regime processes, exact analysis gets very complicated due to correlations among model parameters within and across regimes. Nonetheless, for two- and three-regime threshold-type models one may approximate the stationary marginal distribution by the Riemann-Stieltjes integration method discussed in Section 7.2. Through a comparison with the exact marginal distribution and its moments, we assessed the excellent accuracy of this approximation method.

Acknowledgments

We would like to thank Dong Li for making available his computer code. Also, we express our sincere thanks to an Associate Editor and two anonymous referees for their helpful comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Jan G. De Gooijer http://orcid.org/0000-0002-5434-8550

References

[1] Chen CWS, So MKP, Liu FC. A review of threshold time series models in finance. Stat Interface. 2011;4:167–181.
[2] Hansen BE. Threshold autoregression in economics. Stat Interface. 2011;4:123–127.
[3] Ranocha P. Stationary distribution of time series [doctoral dissertation, PhD thesis]. Prague: Charles University; 2006.
[4] Knight J, Satchell S. Some new results for threshold AR(1) models. J Time Ser Econom. 2011;3(2):1
[5] Grynkiv G, Stentoft L. Stationary threshold vector autoregressive models. J Risk Finance. 2018;11:45.
[6] Guay A, Scaillet O. Indirect inference, nuisance parameter, and threshold moving average models. J Bus Econ Stat. 2003;21:122–132.
[7] Li D, Ling S, Tong H. On moving-average models with feedback. Bernoulli. 2012;18:735–745.
[8] Akkaya AD, Tiku ML. Autoregressive models with short-tailed distributions. Statistics. 2008;42:207–221.
[9] Meyn SP, Tweedie RL. Markov chains and stochastic stability. New York (NY): Springer-Verlag; 1993.
[10] Pemberton J. Piecewise constant autoregressive models for univariate time series. Salford: Department of Mathematics, University of Salford; 1990. (Rep. MCS-90-04).
[11] Pemberton J. Measuring nonlinearity in time series. In: Subba Rao TS, editor, Developments in time series analysis. London: Chapman and Hall; 1993. p. 230–240.
[12] Chan KS, Li D, Ling S, et al. On conditionally heteroskedastic AR models with thresholds. Stat Sin. 2014;24:625–652.
[13] Anděl J, Netuka I, Zvára K. On threshold autoregressive processes. Kybernetika. 1984;20: 89–106.
[14] Das S, Genton MG. On the stationary marginal distribution of subclasses of multivariate SETAR processes of order one. J Time Ser Anal. 2020;41:406–420.
[15] Brännäs K, De Gooijer JG. Autoregressive-asymmetric moving average models for business cycle data. J Forecast. 1994;13:529–544.
[16] Wecker WE. Asymmetric time series. J Am Stat Assoc. 1981;6:16–21. Corrigenda: p. 954.
[17] De Gooijer JG. Asymmetric vector moving average models: estimation and testing. Comput Stat. 2021;36(2):1437–1460.
[18] De Gooijer JG. On threshold moving-average models. J Time Ser Anal. 1998;19(1):1–18.
[19] Li D. A note on moving-average models with feedback. J Time Ser Anal. 2012;33:873–879.
[20] Loges W. The stationary marginal distribution of a threshold AR(1) process. J Time Ser Anal. 2004;25:103–125.
[21] Anděl J, Bartoň T. A note on the threshold AR(1) model with Cauchy innovations. J Time Ser Anal. 1986;7:1–5.
[22] Li D, Tong H. On an absolute autoregressive model and skew symmetric distributions. Statistica. 2020;80:177–198.
[23] Tong H. Non-linear time series: a dynamical system approach. Oxford: Oxford University Press; 1990.
[24] Li D, Qiu J. The marginal density of a TMA(1) process. J Time Ser Anal. 2020;41:476–484.
[25] Chan KS, Goracci G. On the ergodicity of first-order threshold autoregressive moving-average processes. J Time Ser Anal. 2019;40:256–264.

Appendix. Proofs of new results

Proof of Proposition 3.1: The first term on the right-hand side (RHS) of (5) is given by

$$\text{RHS1} = \left((1-w) \int_{-\infty}^{r} \varphi(y_{t-1}; \alpha_0^-) \, \mathrm{d}y_{t-1} + w \int_{-\infty}^{r} \varphi(y_{t-1}; \alpha_0^+) \, \mathrm{d}y_{t-1} \right) \varphi(y; \alpha_0^-)$$

$$= \Big((1-w)\Phi_{(r-\alpha_0^+)} + w\Phi_{(r-\alpha_0^-)}\Big)\varphi(y;\alpha_0^-).$$

Similarly, the second term on the RHS of (5) is

$$\mathrm{RHS2} = \Big((1-w)\big(1-\Phi_{(r-\alpha_0^+)}\big) + w\big(1-\Phi_{(r-\alpha_0^-)}\big)\Big)\varphi(y;\alpha_0^+).$$

From the definition of w in (7), it can be readily verified that

$$w\{1-\Phi_{(r-\alpha_0^-)}\}+(1-w)\{1-\Phi_{(r-\alpha_0^+)}\} = 1-w, \quad \text{and} \quad w\Phi_{(r-\alpha_0^-)}+(1-w)\Phi_{(r-\alpha_0^+)} = w.$$

Using these results, and substituting RHS1 and RHS2 in (5), expression (6) follows directly. ■

Proof of Corollary 3.2: We may write

$$\gamma_\ell = \int_r^\infty \int_{-\infty}^\infty y_t y_{t-\ell} g(y_t, y_{t-\ell})\, \mathrm{d}y_t\, \mathrm{d}y_{t-\ell} + \int_{-\infty}^r \int_{-\infty}^\infty y_t y_{t-\ell} g(y_t, y_{t-\ell})\, \mathrm{d}y_t\, \mathrm{d}y_{t-\ell} \qquad \text{(A1)}$$

$$= \int_{-\infty}^\infty y_t k(y_t|y_{t-\ell+1})\varphi(y_{t-\ell+1};\alpha_0^+)\,\mathrm{d}y_t \int_r^\infty y_{t-\ell} f(y_{t-\ell})\,\mathrm{d}y_{t-\ell}$$

$$+ \int_{-\infty}^\infty y_t k(y_t|y_{t-\ell+1})\varphi(y_{t-\ell+1};\alpha_0^-)\,\mathrm{d}y_t \int_{-\infty}^r y_{t-\ell} f(y_{t-\ell})\,\mathrm{d}y_{t-\ell}$$

$$= A_2\gamma_{2,\ell} + A_1\gamma_{1,\ell}, \qquad \text{(A2)}$$

where $A_1 = \int_{-\infty}^r y_{t-\ell} f(y_{t-\ell})\,\mathrm{d}y_{t-\ell}$, $A_2 = \int_r^\infty y_{t-\ell} f(y_{t-\ell})\,\mathrm{d}y_{t-\ell}$, $k(y_t|y_{t-\ell+1}) = \prod_{j=0}^{\ell-2} h(y_{t-j}|y_{t-j-1})$ $(\ell \geq 2)$, and where

$$\gamma_{1,\ell} \equiv \int_{-\infty}^\infty y_t k(y_t|y_{t-\ell+1})\varphi(y_{t-\ell+1};\alpha_0^-)\,\mathrm{d}y_t, \quad \gamma_{2,\ell} \equiv \int_{-\infty}^\infty y_t k(y_t|y_{t-\ell+1})\varphi(y_{t-\ell+1};\alpha_0^+)\,\mathrm{d}y_t.$$

The constants A_1 and A_2 can easily be calculated using the identities

$$\int_{-\infty}^r y\varphi(y;\delta)\,\mathrm{d}y = -\varphi(r-\delta) + \delta\Phi_{(r-\delta)}, \qquad \int_r^\infty y\varphi(y;\delta)\,\mathrm{d}y = \varphi(r-\delta) + \delta\big(1-\Phi_{(r-\delta)}\big),$$

where $\varphi(r-\delta) = (1/\sqrt{2\pi})\exp(-(r-\delta)^2/2)$. By straightforward calculation, it can be shown that $\gamma_1 = A_1\alpha_0^- + A_2\alpha_0^+$. So, $\gamma_{1,1} = \alpha_0^-$ and $\gamma_{2,1} = \alpha_0^+$.

Next, we derive a recursive relationship between $\gamma_{i,\ell}$ and $\gamma_{i,\ell-1}$ for $i = 1, 2$. From (5) we have $g(y_t, y_{t-\ell}) = \int_{-\infty}^\infty k(y_t|y_{t-\ell+1})h(y_{t-\ell+1}|y_{t-\ell})f(y_{t-\ell})\,\mathrm{d}y_{t-\ell+1}$ Thus, the first term on the RHS of (A1) can be written as

$$\int_r^\infty \int_{-\infty}^\infty y_t y_{t-\ell} g(y_t, y_{t-\ell})\,\mathrm{d}y_t\,\mathrm{d}y_{t-\ell}$$

$$= \int_r^\infty y_{t-\ell} f(y_{t-\ell})\,\mathrm{d}y_{t-\ell}\left\{\int_{-\infty}^\infty \int_{-\infty}^\infty y_t k(y_t|y_{t-\ell+1})\varphi(y_{t-\ell+1};\alpha_0^-)\,\mathrm{d}y_{t-\ell+1}\,\mathrm{d}y_t\right\}. \qquad \text{(A3)}$$

Note that

$$k(y_t|y_{t-\ell+1}) = \begin{cases} k(y_t|y_{t-\ell+2})\varphi(y_{t-\ell+2};\alpha_0^-) & \text{if } y_{t-\ell+1} < r, \\ k(y_t|y_{t-\ell+2})\varphi(y_{t-\ell+2};\alpha_0^+) & \text{if } y_{t-\ell+1} \geq r. \end{cases}$$

Upon substituting $k(y_t|y_{t-\ell+1})$ into (A3), we get

$$\int_r^\infty y_{t-\ell} f(y_{t-\ell})\,\mathrm{d}y_{t-\ell}\left\{\int_{-\infty}^\infty \int_{-\infty}^r y_t k(y_t|y_{t-\ell+2})\varphi(y_{t-\ell+2};\alpha_0^-)\varphi(y_{t-\ell+1};\alpha_0^-)\,\mathrm{d}y_{t-\ell+1}\,\mathrm{d}y_t\right.$$

$$\left. + \int_{-\infty}^\infty \int_r^\infty y_t k(y_t|y_{t-\ell+2})\varphi(y_{t-\ell+2};\alpha_0^+)\varphi(y_{t-\ell+1};\alpha_0^-)\,\mathrm{d}y_{t-\ell+1}\,\mathrm{d}y_t\right\}$$

$$= A_2\Big\{\Phi_{(r-\alpha_0^+)}\gamma_{2,\ell-1} + \big(1-\Phi_{(r-\alpha_0^+)}\big)\gamma_{1,\ell-1}\Big\}.$$

The second term on the RHS of (A1) can be evaluated similarly. Then combining these results, we obtain

$$\gamma_\ell = A_2\Big\{\Phi_{(r-\alpha_0^+)}\gamma_{2,\ell-1}+\big(1-\Phi_{(r-\alpha_0^+)}\big)\gamma_{1,\ell-1}\Big\} + A_1\Big\{\Phi_{(r-\alpha_0^-)}\gamma_{1,\ell-1} + \big(1-\Phi_{(r-\alpha_0^-)}\big)\gamma_{2,\ell-1}\Big\}, \tag{A4}$$

which together with (A2) completes the proof. ■

Proof of Corollary 6.4: Let $Y_t = X$ and $Y_{t-1} = Y$. The joint density of Y_t and Y_{t-1} is given by

$$g_{\alpha_1^-,\alpha_1^+}(x,y) = \begin{cases} \frac{1}{2}\exp\{-|x-\alpha_1^- y|\}f_{\alpha_1^-,\alpha_1^+}(y) & \text{if } y \le 0, \\ \frac{1}{2}\exp\{-|x-\alpha_1^+ y|\}f_{\alpha_1^-,\alpha_1^+}(y) & \text{if } y > 0, \end{cases} \tag{A5}$$

where $f_{\alpha_1^-,\alpha_1^+}(y)$ is given by (44). Then

$$\begin{aligned}
\mathbb{E}(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, g_{\alpha_1^-,\alpha_1^+}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= g(\alpha_1^-,\alpha_1^+)\Bigg\{\frac{g_3(\alpha_1^-)}{g_1(\alpha_1^-)}\int_{-\infty}^{0} y\sum_{q=0}^{\infty} d_1(\alpha_1^-,q)(\alpha_1^-)^{-q}\exp(-(\alpha_1^-)^q y) \\
&\quad \times \int_{-\infty}^{\infty} x\exp(-|x-\alpha_1^- y|)\,\mathrm{d}x\,\mathrm{d}y \\
&\quad + \frac{g_3(\alpha_1^+)}{g_1(\alpha_1^+)}\int_{0}^{\infty} y\sum_{q=0}^{\infty} d_1(\alpha_1^+,q)(\alpha_1^+)^{-q}\exp(-(\alpha_1^+)^q y) \\
&\quad \times \int_{-\infty}^{\infty} x\exp(-|x-\alpha_1^+ y|)\,\mathrm{d}x\,\mathrm{d}y\Bigg\} \\
&= g(\alpha_1^-,\alpha_1^+)\Bigg\{\frac{g_3(\alpha_1^-)}{g_1(\alpha_1^-)}\int_{-\infty}^{0} y^2\sum_{q=0}^{\infty} d_1(\alpha_1^-,q)(\alpha_1^-)^{-q+1}\exp\big(-(\alpha_1^-)^q y\big)\,\mathrm{d}y \\
&\quad + \frac{g_3(\alpha_1^+)}{g_1(\alpha_1^+)}\int_{0}^{\infty} y^2\sum_{q=0}^{\infty} d_1(\alpha_1^+,q)(\alpha_1^+)^{-q+1}\exp\big(-(\alpha_1^+)^{-q} y\big)\,\mathrm{d}y\Bigg\}.
\end{aligned} \tag{A6}$$

With $g_4(\lambda,2)$ defined by (45) and since $\int_0^\infty x^2\exp(-\lambda^{-q}x)\mathrm{d}x = 2\lambda^{3q}$ ($\lambda > 0$), (A6) can be written as

$$\mathbb{E}(XY) = 2g(\alpha_1^-,\alpha_1^+)\left\{\frac{g_3(\alpha_1^-)}{g_1(\alpha_1^-)}g_4(\alpha_{1,2}^-) + \frac{g_3(\alpha_1^+)}{g_1(\alpha_1^+)}g_4(\alpha_{1,2}^+)\right\},$$

which completes the proof. ■

Proof of Corollary 6.6: Similar to the proof of Corollary 6.4, we have

$$\begin{aligned}
\mathbb{E}(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, g_{\alpha_1^-,\alpha_1^+}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= h(\alpha_1^-,\alpha_1^+)g_8(\alpha_1^-,\alpha_1^+)\int_{-\infty}^{0} y\exp(y)\int_{-\infty}^{\infty} x\left\{\frac{1}{2}\exp\{-|x-\alpha_1^- y|\}\mathrm{d}x\right\}\mathrm{d}y \\
&\quad + \int_0^{\infty} h(\alpha_1^-,\alpha_1^+)y\Big(r_1(y,\alpha_1^+) + g_8(\alpha_1^-,\alpha_1^+)r_2(y,\alpha_1^-,\alpha_1^+)\Big)
\end{aligned}$$

$$\begin{aligned}
&\times \int_{-\infty}^{\infty} x \left\{ \frac{1}{2} \exp\{-|x-\alpha_1^+ y|\}\, dx \right\} dy \\
&= \alpha_1^- h(\alpha_1^-, \alpha_1^+) g_8(\alpha_1^-, \alpha_1^+) \int_{-\infty}^{0} y^2 \exp(y)\, dy + \alpha_1^+ h(\alpha_1^-, \alpha_1^+) \\
&\times \int_0^{\infty} y^2 r_1(y, \alpha_1^+)\, dy \\
&+ \alpha_1^+ h(\alpha_1^-, \alpha_1^+) g_8(\alpha_1^-, \alpha_1^+) \int_0^{\infty} y^2 r_2(y, \alpha_1^-, \alpha_1^+)\, dy \\
&= 2\alpha_1^- h(\alpha_1^-, \alpha_1^+) g_8(\alpha_1^-, \alpha_1^+) + 2\alpha_1^+ h(\alpha_1^-, \alpha_1^+) \sum_{q=0}^{\infty} d_1(\alpha_1^+, q)(\alpha_1^+)^{2q} \\
&+ 2\alpha_1^+ h(\alpha_1^-, \alpha_1^+) g_8(\alpha_1^-, \alpha_1^+) |\alpha_1^-|^2 \sum_{q=0}^{\infty} d_2(\alpha_1^-, \alpha_1^+, q)(\alpha^+)^{2(q-1)} \\
&= 2\alpha_1^- h(\alpha_1^-, \alpha_1^+) g_8(\alpha_1^-, \alpha_1^+) + 2\alpha_1^+ h(\alpha_1^-, \alpha_1^+) g_4(\alpha_1^+, 2) \\
&+ 2\alpha_1^+ h(\alpha_1^-, \alpha_1^+) g_8(\alpha_1^-, \alpha_1^+) g_7(\alpha_1^+, 2),
\end{aligned}$$

where $g_4(\alpha_1^+, 2)$ is defined by (45), $g_7(\alpha_1^-, \alpha_1^+, 2)$ is defined by (48), and where we note that $\int_0^\infty y^2 \exp(-\lambda_1^{1-q} |\lambda_2|^{-1} y) dy = 2\lambda_1^{-3+3q} |\lambda_2|^3$ $(0 \le \lambda_1, \lambda_2 < 1)$. ■

Proof of Corollary 6.8: With $\alpha_1^+ = -\alpha_1^- \equiv \alpha$, the third term on the RHS of (51) can be written as

$$\begin{aligned}
\text{RHS3} &= 2\alpha h(-\alpha, \alpha) g_8(-\alpha, \alpha) \left(\alpha^2 \sum_{q=0}^{\infty} d_2(-\alpha, \alpha, q) \alpha^{2(q-1)} - \alpha^2 d_2(-\alpha, \alpha, 0) \alpha^{-2} \right) \\
&= 2\alpha h(-\alpha, \alpha) g_8(-\alpha, \alpha) \left(\sum_{q=0}^{\infty} d_2(-\alpha, \alpha, q) \alpha^{2q} - 1 \right) \\
&= 2\alpha h(-\alpha, \alpha) g_8(-\alpha, \alpha) \big(g_4(\alpha, 2) - 1 \big), \qquad \text{(A7)}
\end{aligned}$$

where $d_2(-\alpha, \alpha, 0) = 1$, $d_2(-\alpha, \alpha, q) = d_1(\alpha, q)$, and $g_4(\alpha, 2) = \sum_{q=0}^{\infty} d_1(\alpha, q) \alpha^{2q}$. Next, substituting (A7) into (51), and using (53a) and (53c) gives

$$\begin{aligned}
\gamma_1 &= 2\alpha \left(h(-\alpha, \alpha) + \frac{1}{g_3(\alpha)} \right) g_4(\alpha, 2) - \frac{4\alpha}{g_3(\alpha)} \\
&= \frac{2\alpha}{g_1(\alpha)} g_4(\alpha, 2) - \frac{4\alpha}{g_3(\alpha)},
\end{aligned}$$

which completes the proof. ■

Proof of (63)***:*** Let $Y_t = X$ and $Y_{t-1} = Y$. Then

$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{i=-n}^{n} f^*_{\alpha_1^-, \alpha_1^+}(y_i) \int_{-\infty}^{r} y q_{\alpha_1^-}(y|y_i) \int_{-\infty}^{\infty} x (2\pi)^{-1/2} \exp\left\{ -\frac{1}{2}(x - \alpha_1^- y)^2 \right\} dx\, dy \\
&+ \sum_{i=-n}^{n} f^*_{\alpha_1^-, \alpha_1^+}(y_i) \int_{r}^{\infty} y q_{\alpha_1^+}(y|y_i) \int_{-\infty}^{\infty} x (2\pi)^{-1/2} \exp\left\{ -\frac{1}{2}(x - \alpha_1^+ y)^2 \right\} dx\, dy
\end{aligned}$$

$$= \alpha_1^- \sum_{i=-n}^{n} f^*_{\alpha_1^-,\alpha_1^+}(y_i) \int_{-\infty}^{r} y^2 q_{\alpha_1^-}(y|y_i)\,\mathrm{d}y$$

$$+ \alpha_1^+ \sum_{i=-n}^{n} f^*_{\alpha_1^-,\alpha_1^+}(y_i) \int_{r}^{\infty} y^2 q_{\alpha_1^+}(y|y_i)\,\mathrm{d}y,$$

where $q_{\alpha_1^-}(y|y_i)$ is the density of $\mathcal{N}(\alpha_1^- y_i, 1)$ if $y_i \le r$, and $q_{\alpha_1^+}(y|y_i)$ is the density of $\mathcal{N}(\alpha_1^+ y_i, 1)$ if $y_i > r$. Consequently,

$$\mathbb{E}(XY) = \alpha_1^- \sum_{\substack{i=-n \\ y_i \le r}}^{n} f^*_{\alpha_1^-,\alpha_1^+}(y_i) \int_{-\infty}^{r} y^2 \left\{(2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(y - \alpha_1^- y_i)^2\right\}\right\} \mathrm{d}y$$

$$+ \alpha_1^- \sum_{\substack{i=-n \\ y_i > r}}^{n} f^*_{\alpha_1^-,\alpha_1^+}(y_i) \int_{-\infty}^{r} y^2 \left\{(2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(y - \alpha_1^+ y_i)^2\right\}\right\} \mathrm{d}y$$

$$+ \alpha_1^+ \sum_{\substack{i=-n \\ y_i \le r}}^{n} f^*_{\alpha_1^-,\alpha_1^+}(y_i) \int_{r}^{\infty} y^2 \left\{(2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(y - \alpha_1^- y_i)^2\right\}\right\} \mathrm{d}y$$

$$+ \alpha_1^+ \sum_{\substack{i=-n \\ y_i > r}}^{n} f^*_{\alpha_1^-,\alpha_1^+}(y_i) \int_{r}^{\infty} y^2 \left\{(2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(y - \alpha_1^+ y_i)^2\right\}\right\} \mathrm{d}y. \tag{A8}$$

Let $a \in \mathbb{R}$, $b \in \mathbb{R}$. From standard computation, we have the following integral equations

$$(2\pi)^{-1/2} \int_{-\infty}^{b} x^2 \exp\left\{-\frac{1}{2}(x-a)^2\right\} \mathrm{d}x = (1+a^2)\Phi_{(b-a)} - (2\pi)^{-1/2}(b+a) \exp\left\{-\frac{1}{2}(b-a)^2\right\}, \tag{A9}$$

$$(2\pi)^{-1/2} \int_{b}^{\infty} x^2 \exp\left\{-\frac{1}{2}(x-a)^2\right\} \mathrm{d}x = (1+a^2)\Phi_{(a-b)} + (2\pi)^{-1/2}(b+a) \exp\left\{-\frac{1}{2}(b-a)^2\right\}. \tag{A10}$$

Inserting (A9) and (A10) in (A8) completes the proof. ■