IDENTIFICATION AND ESTIMATION OF
TAR MODELS WITH t NOISE

HANWEN ZHANG

Departamento de Estadística

División de Ciencias Económicas y Administrativas

Universidad Santo Tomás

Colombia

e-mail: hanwenzhang@usantotomas.edu.co

Abstract

In this paper, we consider the problem of the identification and estimation of TAR models with one threshold variable, when the noise process follows a t distribution. A Bayesian procedure of three stages based on Gibbs samplings is proposed. A simulated example indicates that the procedure proposed has a good performance.

1. Introduction

The TAR models were proposed by Tong [6], these models assume that the values of a process (the threshold process) $\{Z_t\}$ determine not only the values of the process of interest $\{X_t\}$, but also its dynamic. When the threshold process is the same process of interest but lagged, the model is known as SETAR (self-exciting TAR) models. Tsay [8] extended the univariate SETAR models to the multivariate case. On the other side, Nieto [4] characterized the univariate TAR models in terms of its mean, conditional mean, variance, conditional variance, and also found the

expressions for the best predictor. It was found that the conditional variance of the TAR models is no constant, so the TAR models can be thought as an alternative to the GARCH models in order to describe the heteroscedasticity in the data. Also, the TAR models can be easily extended to the TARMA models, and its Bayesian modelling with two regimes has been investigated in Sáfadi & Morettin [5] and Xia et al. [10]; another extension of the TAR model is when there are two threshold variables instead of one, this class of models have been studied by Chen et al. [2] within the particular case when there are two regimes.

Despite the usefulness of the TAR models, these models are not easy to be identified due to the large number of parameters and the nesting structure between the parameters. For SETAR models, Tsay [7] provided a simple and widely applicable model-building procedure. However, generally speaking, most of the parameters as the thresholds and the number of regimes are assumed to be known; otherwise, they can be identified by using Tong's NAIC criterion together with some graphical techniques. Assuming that the noise process is Gaussian, Nieto [3] developed a Bayesian procedure in order to identify the number of regimes and estimate the other parameters once the thresholds are identified by using NAIC criterion for each possible number of regimes in presence of missing data in both process of interest and the threshold process.

In many cases, we can face with data which can not be appropriately described by the Gaussian distribution, for example, it is well-known that financial series data often have heavy tails, and we may think that the t distribution could be more appropriate for the noise process than the Gaussian distribution. However, on our actual knowledge, there is no theoretical development for the TAR modelling with t noise. For this reason, the focus of this paper is to propose a Bayesian methodology in order to identify and estimate a TAR model, when the noise process follows a standardized student-t distribution extending the methodology of Nieto [3]. We propose a three-stage methodology: In the first stage, we

estimate the number of regimes together with the thresholds; in the second stage, we estimate the autoregressive orders in the regimes; and finally, in the last stage, we estimate the autoregressive orders, the variance weights, and the degree of freedom of the noise process. This way, the identification of the model takes place in the first two stages, and the estimation of the model in the last stage.

In Section 2, we introduce the TAR model with t noise and its likelihood function; in Section 3, we present the way to estimate the model when the model has been identified; in Section 4, we present the identification of the model; in Section 5, we illustrate the performance of the proposed methodology with a simulated example; conclusion and further works are presented in Section 6.

2. TAR Model with t Noise

A TAR model with l regimes when the noise process follows a t distribution is given by

$$X_t = a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} X_{t-i} + h^{(j)} e_t, \tag{1}$$

when $Z_t \in R_j = (r_{j-1}, r_j]$ for some $j = 1, \cdots, l$, $r_0 = -\infty$, $r_l = \infty$. The values $r_1 < \cdots < r_{l-1}$ are denominated the thresholds and they define the regimes of the model. The process $\{Z_t\}$ is called the *threshold process*, and its stochastic behaviour is described by a Markov chain of order p. The values $k_1, \cdots, k_l$ are positive integer numbers representing the autoregressive orders in the l regimes. The noise process $\{e_t\} \sim iid \dfrac{t_n}{\sqrt{n/(n-2)}}$ and is mutually independent of the process $\{Z_t\}$. The $h^{(j)}$ with $j = 1, \cdots, l$ represent the variance weights in the l regimes. Additionally, we assume that the process $\{Z_t\}$ is exogeneous in the sense that there is no feedback of $\{X_t\}$ towards it.

The parameters of the model can be divided in two groups:

- *Structural parameters*: the number of regimes l, the $l-1$ thresholds $r_1, \cdots, r_{l-1}$, and the autoregressive orders of the l regimes $k_1, \cdots, k_l$.

- *Non structural parameters*: the autoregressive coefficients $a_i^{(j)}$ with $i = 0, \cdots, k_j$ and $j = 1, \cdots, l$, the variance weights $h^{(1)}, \cdots, h^{(l)}$ and the degree of freedom of the noise process n.

In this paper, we use the following notation: $\boldsymbol{\theta}'_j = (a_1^{(j)}, \cdots, a_{k_j}^{(j)})'$ for $j = 1, \cdots, l$, $\boldsymbol{\theta}' = (\boldsymbol{\theta}'_1, \cdots, \boldsymbol{\theta}'_l)'$, and $\mathbf{h}' = (h^{(1)}, \cdots, h^{(l)})'$.

2.1. Likelihood function of the model

Conditioned on the values of the structural parameters, the initial values $\mathbf{x}_k = (x_1, \cdots, x_k)'$, where $k = \max\{k_1, \cdots, k_l\}$ and the observed data of the threshold process $\mathbf{z} = (z_1, \cdots, z_T)'$, we have that

$$f(\mathbf{x}|\mathbf{z}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z) = f(x_{k+1}|\mathbf{x}_k, \mathbf{z}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z) \cdots f(x_T|x_{T-1}, \cdots, x_1, \mathbf{z}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z).$$

As $e_t \sim \dfrac{t_n}{\sqrt{n/(n-2)}}$, for $t = k+1, \cdots, T$, the variable $x_t|x_{t-1}, \cdots, x_1$ distributes as a t_n variable multiplied by $\dfrac{h^{(j_t)}}{\sqrt{n/(n-2)}}$ and plus $a_0^{(j_t)} + \sum_{i=1}^{k_{j_t}} a_i^{(j_t)} x_{t-i}$. This way

$$f(x_t|x_{t-1}, \cdots, x_1, \mathbf{z}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi(n-2)}\Gamma(\frac{n}{2})} \frac{1}{h^{(j_t)}}$$

$$\times \left(1 + \frac{\left[x_t - a_0^{(j_t)} - \sum_{i=1}^{k_{j_t}} a_i^{(j_t)} x_{t-i}\right]^2}{(h^{(j_t)})^2(n-2)}\right)^{-\frac{n+1}{2}},$$

where $j_t = j$ if $Z_t \in R_j = (r_{j-1}, r_j]$ for some $j = 1, \cdots, l$.

Consequently, the likelihood function is given by

$$f(\mathbf{x}|\mathbf{z}, \boldsymbol{\theta}_x, \boldsymbol{\theta}_z) = \left[\frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi(n-2)}\Gamma(\frac{n}{2})} \right]^{T-k} \prod_{t=k+1}^{T} \left[h^{(j_t)} \right]^{-1} \prod_{t=k+1}^{T} \left(1 + \frac{e_t^2}{n-2} \right)^{-\frac{n+1}{2}},$$

(2)

with $e_t = \dfrac{1}{h^{(j_t)}} \left(x_t - a_0^{(j_t)} - \sum_{i=1}^{k_{j_t}} a_i^{(j_t)} x_{t-i} \right)$ and $j_t = j$ when $Z_t \in R_j$.

3. Estimation of Non Structural Parameters

In this part of the paper, the structural parameters are assumed to be known, and we focus on finding the posterior conditional distributions of the autoregressive coefficients $\boldsymbol{\theta}_j$, the variance weights $h^{(j)}$, and the degree of freedom of the noise process n. Additionally, we assume prior independency between the parameters $\boldsymbol{\Theta}, \mathbf{h}$, and n, as well as prior independency between the non-structural parameters in the l regimes.

The prior distribution for the vector $\boldsymbol{\theta}_j$ is a multivariate normal distribution, denoted as $\boldsymbol{\theta}_j \sim N(\boldsymbol{\theta}_{0,j}, \mathbf{V}_{0,j}^{-1})$, and the posterior conditional distribution of $\boldsymbol{\theta}_j$ is given by the following result:

Result 1. For each $j = 1, \cdots, l$, the conditional distribution of $\boldsymbol{\theta}_j$ given the structural parameters, $\boldsymbol{\theta}_i$, with $i \neq j$, $\mathbf{h}$, and n is given by

$p(\boldsymbol{\theta}_j|\boldsymbol{\theta}_i, i \neq j, \mathbf{h}, \mathbf{x}, \mathbf{z}, n)$

$$\propto \prod_{\{t:j_t=j\}} \left(1 + \frac{\left[x_t - a_0^{(j)} - \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i} \right]^2}{(h^{(j)})^2 (n-2)} \right)^{-\frac{n+1}{2}}$$

$$\times \exp \left\{ -\frac{1}{2} (\boldsymbol{\theta}_j - \boldsymbol{\theta}_{0,j})' \mathbf{V}_{0,j} (\boldsymbol{\theta}_j - \boldsymbol{\theta}_{0,j}) \right\}.$$

(3)

Proof.

$p(\mathbf{\theta}_j | \mathbf{\theta}_i, \ i \neq j, \ \mathbf{h}, \ \mathbf{x}, \ \mathbf{z}, \ n)$

$\propto p(\mathbf{x} | \mathbf{\theta}, \ \mathbf{z}, \ \mathbf{h}, \ n) p(\mathbf{\theta}_j | \mathbf{\theta}_i, \ i \neq j, \ \mathbf{h}, \ n)$

$\propto p(\mathbf{x} | \mathbf{\theta}, \ \mathbf{z}, \ \mathbf{h}, \ n) p(\mathbf{\theta}_j)$

$$\propto \prod_{t=1}^{T} \left(1 + \frac{\left[x_t - a_0^{(j_t)} - \sum_{i=1}^{k_{j_t}} a_i^{(j_t)} x_{t-i} \right]^2}{(h^{(j_t)})^2 (n-2)} \right)^{-\frac{n+1}{2}}$$

$$\times \exp \left\{ -\frac{1}{2} (\mathbf{\theta}_j - \mathbf{\theta}_{0,j})' \mathbf{V}_{0,j} (\mathbf{\theta}_j - \mathbf{\theta}_{0,j}) \right\}$$

$$\propto \prod_{\{t : j_t = j\}} \left(1 + \frac{\left[x_t - a_0^{(j)} - \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i} \right]^2}{(h^{(j)})^2 (n-2)} \right)^{-\frac{n+1}{2}}$$

$$\times \exp \left\{ -\frac{1}{2} (\mathbf{\theta}_j - \mathbf{\theta}_{0,j})' \mathbf{V}_{0,j} (\mathbf{\theta}_j - \mathbf{\theta}_{0,j}) \right\}.$$

Note that, the posterior conditional distribution of $\mathbf{\theta}_j$ is affected only by $h^{(j)}$, but not the other components of $\mathbf{h}$ in regimes different from j, so we have some class of posterior independence between regimes.

Now, with respect to the variance weights $h^{(j)}$, we follow the standard Bayesian methodology assigning an inverse Gamma distribution with shape parameter α and scale parameter β, (inverse-Gamma (α, β)) as the prior distribution of $(h^{(j)})^2$, that is,

$$p((h^{(j)})^2) \propto (h^{(j)})^{-2\alpha-2} \exp \{-\beta / (h^{(j)})^2\} I_{(0, \infty)}((h^{(j)})^2).$$

Combining this prior distribution of $(h^{(j)})^2$ and the likelihood function, we have the following posterior conditional distribution $(h^{(j)})^2$:

Result 2. For each $j = 1, \cdots, l$, the conditional distribution of $(h^{(j)})^2$ given the structural parameters, $\boldsymbol{\theta}_j$, $j = 1, \cdots, l$, $h^{(i)}$, with $i \neq j$, is given by

$$p((h^{(j)})^2|\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_l, h^{(i)}, i \neq j, \mathbf{x}, \mathbf{z}, n)$$

$$\propto \prod_{\{t:j_t=j\}} \left(1 + \frac{\left[x_t - a_0^{(j)} - \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i} \right]^2}{(h^{(j)})^2(n-2)} \right)^{-\frac{n+1}{2}} (h^{(j)})^{-2\alpha-2-n_j} \exp\{-\beta/(h^{(j)})^2\}.$$

$$(4)$$

Proof.

$$p((h^{(j)})^2|\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_l, \mathbf{x}, \mathbf{z}, n)$$

$$\propto p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{z}, \mathbf{h}, n)p((h^{(j)})^2)$$

$$\propto \prod_{t=1}^{T} (h^{(j_t)})^{-1} \left(1 + \frac{\left[x_t - a_0^{(j_t)} - \sum_{i=1}^{k_{j_t}} a_i^{(j_t)} x_{t-i} \right]^2}{(h^{(j_t)})^2(n-2)} \right)^{-\frac{n+1}{2}}$$

$$\times (h^{(j)})^{-2\alpha-2} \exp\{-\beta/(h^{(j)})^2\}$$

$$\propto \prod_{\{t:j_t=j\}} \left(1 + \frac{\left[x_t - a_0^{(j)} - \sum_{i=1}^{k_j} a_i^{(j)} x_{t-i} \right]^2}{(h^{(j)})^2(n-2)} \right)^{-\frac{n+1}{2}}$$

$$\times (h^{(j)})^{-2\alpha-2-n_j} \exp\{-\beta/(h^{(j)})^2\}.$$

$$\square$$

Note that, the posterior conditional distribution of $(h^{(j)})^2$ is affected only by $\boldsymbol{\theta}_j$, but not by $\boldsymbol{\theta}_i$ with $i \neq j$, so again, we have the posterior independency between $\boldsymbol{\theta}_j$, $h^{(j)}$ of different regimes.

Finally, we find the posterior conditional distribution of the degree of freedom of the noise process n. The prior distribution of n is a Gamma distribution, following the suggestion of Watanabe [9], since in the distribution Gamma (α', β'), the expectation and the variance are given by $\alpha'\beta'$ and $\alpha'\beta'^2$, respectively; and in the practice, α' and β' can be chosen according to the prior knowledge about n, and in case that there is no prior information about n, we can choose prior variance quite large to represent the high degree of uncertainty in the prior information of n. The prior distribution of n is given by

$$p(n) \propto n^{\alpha'-1} \exp\{-n/\beta'\}.$$

Using this prior distribution, we find the following posterior conditional distribution of n:

Result 3. The posterior conditional distribution of the degree of freedom of the noise process $\{e_t\}$ is given by

$$p(n|\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_l, \mathbf{h}, \mathbf{x}, \mathbf{z})$$

$$\propto \left[\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n-2}\,\Gamma(\frac{n}{2})}\right]^{T-k} \prod_{t=k+1}^{T}\left(1 + \frac{\left[x_t - a_0^{(j_t)} - \sum_{i=1}^{k_{j_t}} a_i^{(j_t)} x_{t-i}\right]^2}{(h^{(j_t)})^2(n-2)}\right)^{-\frac{n+1}{2}}$$

$$\times n^{\alpha'-1} \exp\{-n/\beta'\}. \tag{5}$$

Proof.

$$p(n|\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_l, \mathbf{h}, \mathbf{x}, \mathbf{z})$$

$$\propto p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{z}, \mathbf{h}, n)p(n)$$

$$
\propto \left[\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n-2}\,\Gamma(\frac{n}{2})} \right]^{T-k} \prod_{t=k+1}^{T} \left(1 + \frac{\left[x_t - a_0^{(j_t)} - \sum_{i=1}^{k_{j_t}} a_i^{(j_t)} x_{t-i} \right]^2}{(h^{(j_t)})^2 (n-2)} \right)^{-\frac{n+1}{2}}
$$

$$
\times\, n^{\alpha'-1} \exp\{-n/\beta'\}.
$$

$\square$

In conclusion, the estimation of the non-structural parameters can be carried out by means of a Gibbs sampler, using the conditional densities (3), (4), and (5).

4. Estimation of the Structural Parameters

In this section, we develop the results concerning the estimation of the structural parameters, i.e., the identification of a TAR model. Firstly, we assume that the number of regimes l and the $l-1$ thresholds are known, and we estimate the autoregressive orders in these regimes; and finally, we have the general case, where all the structural parameters are unknown.

4.1. Estimation of the autoregressive orders $k_1, \cdots, k_l$

As we assume that the number of the regimes and the thresholds are known, remaining parameters to be estimated are the autoregressive orders and the non-structural parameters.

We assume that the autoregressive orders $k_1, \cdots, k_l$ are realizations of discrete random variables $K_1, \cdots, K_l$, and each of theses variables ravkes value in the set $\{0, 1, \cdots, k_{\max}\}$. It is important to note that when the values of some autoregressive orders change, the specification of the TAR model changes and the dimension of the vector of the autoregressive coefficients Θ also changes. Carlin & Chib [1] developed a Bayesian methodology for the selection of models and Nieto [3] adapted this

methodology in order to identify the TAR model with Gaussian noise. In this paper, we adapt the same methodology to identify the TAR model with t noise. Suppose that M is a discrete random variable indexing the model, which takes values $1, \cdots, (k_{\max} + 1)^l$. For each possible model $M = m$, we define the vector of parameters $\boldsymbol{\Theta}_m$ as $\boldsymbol{\Theta}'_m = (\boldsymbol{\theta}'_1, \cdots, \boldsymbol{\theta}'_l, \mathbf{h}')$ for the model m with $m = 1, \cdots, (k_{\max} + 1)^l$. The degree of freedom n can be considered as a nuisance parameter, since its dimension is the same for all models as well as its interpretation, so we can sample values of n using the densities $p(n|\mathbf{k}^{(g)}, \boldsymbol{\theta}^{(g)}, \mathbf{h}^{(g)}, \mathbf{y})$ in the g-th iteration.

Carlin & Chib [1] found the following conditional densities:

$$p(M = m|\boldsymbol{\Theta}, \mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\Theta}_m, M = m)P(M = m)}{\sum_{m'} p(\mathbf{y}|\boldsymbol{\Theta}_m, M = m')P(M = m')}, \qquad (6)$$

where $\boldsymbol{\Theta} = \{\boldsymbol{\Theta}_1, \ldots, \boldsymbol{\Theta}_{(k_{\max}+1)^l}\}$;

and

$$p(\boldsymbol{\Theta}_m|\boldsymbol{\Theta}_{m' \neq m}, M, \mathbf{y}) \propto \begin{cases} p(\mathbf{y}|\boldsymbol{\Theta}_m, M = m)p(\boldsymbol{\Theta}_m|M = m), & \text{if } M = m, \\ p(\boldsymbol{\Theta}_m|M = m), & \text{if } M \neq m, \end{cases}$$

$$(7)$$

with $\mathbf{y} = (\mathbf{x}, \mathbf{z})$, and the densities $p(\boldsymbol{\Theta}_m|M = m)$ are denominated as the link functions, which can be taken as the prior distribution of $\boldsymbol{\Theta}_m$.

In the context of the problem of identification of the autoregressive orders, the model indicator M is determined jointly by the values of variables $K_1, \cdots, K_l$. This way, the density (6) is equivalent to the densities $p(k_j|\boldsymbol{\Theta}, k_{i, i \neq j}, \mathbf{y})$ with $j = 1, \cdots, l$. In order to compute these densities, Nieto [3] found that

$$p(k_j|\boldsymbol{\Theta}, k_{i, i \neq j}, l, \mathbf{x}, \mathbf{z}) = \frac{p(\mathbf{x}|\mathbf{z}, \boldsymbol{\Theta}, \mathbf{h}, \mathbf{k}, l)p(k_j)}{\sum_{k'_j=0}^{\bar{k}_j} p(\mathbf{x}|\mathbf{z}, \boldsymbol{\theta}, \mathbf{h}, \mathbf{k}', l)p(k'_j)}, \qquad (8)$$

where $\mathbf{k} = (k_1, \cdots, k_l)$, and $\mathbf{k}'$ is obtained by replacing the component k_j of the vector $\mathbf{k}$ by k'_j for all $j = 1, \cdots, l$.

Finally, we use the conditional density (5) to samples values of n.

In summary, using the densities (7), (8), and (5), a Gibbs sampler can be implemented in order to obtain the estimations of the probabilities of all the possible values for each K_j with $j = 1, \cdots, l$. Denoting these estimated probabilities as $\hat{p}_{0j}, \hat{p}_{1j}, \cdots, \hat{p}_{k_{\max}j}$, we can choose the value of K_j as the one with major probability associated.

4.2. Estimation of the number of the regimes *l*

In order to estimate the number of regimes, we use again the approach developed in Nieto [3] adapting the methodology of Carlin & Chib [1]. Suppose that the number of regimes l is realization of a discrete random variable L, which takes values in the set $\{2, \cdots, l_{\max}\}$, and the prior distribution of L is denoted by $p(l)$.

Clearly, when the value of l changes, the model specification also changes, this way, we have $l_{\max} - 1$ possible models. Suppose that M is the discrete random variable indexing the model, then M takes values $2, \cdots, l_{\max}$, and for each possible model $M = j$, $\boldsymbol{\Theta}_j$ denotes the vector of the parameters in this model, that is,

$$\boldsymbol{\Theta}'_j = (\boldsymbol{\theta}'_1, \cdots, \boldsymbol{\theta}'_j, \mathbf{h}'_j, \mathbf{k}'_j),$$

with $\mathbf{k}'_j = (k_{1j}, \cdots, k_{jj})'$, where k_{ij} denotes the autoregressive order in the i-th regime in the model $M = j$, $\mathbf{h}'_j = (h^{(1)}, \cdots, h^{(j)})'$. Finally, let $\boldsymbol{\Theta}' = (\boldsymbol{\Theta}'_2, ..., \boldsymbol{\Theta}'_{l_{\max}})$, the vector containing all the parameters for all the possible models.

Nieto [3] found the following conditional densities:

$$p(M = j|\boldsymbol{\Theta}, \mathbf{y}) = p(l|\boldsymbol{\Theta}, \mathbf{y}) \propto p(\mathbf{x}|\mathbf{z}, \boldsymbol{\Theta}_l, l)p(l) \quad \text{for } l = 2, \cdots, l_{\max}; \qquad (9)$$

$$p(k_{ij}|\boldsymbol{\Theta}_{-k_{ij}}, l, \mathbf{y}) = \begin{cases} \dfrac{p(\mathbf{x}|\mathbf{z}, \boldsymbol{\Theta}_l, l)p(k_{ij})}{\sum_{k'_{il}=0}^{k_{\max}} p(\mathbf{x}|\mathbf{z}, \boldsymbol{\Theta}_l, l)p(k'_{il})}, & \text{if } j = l, \\ p(k_{ij}), & \text{if } j \neq l, \end{cases} \qquad (10)$$

where $\boldsymbol{\Theta}_{-k_{ij}}$ denotes the vector $\boldsymbol{\Theta}$ without the element k_{ij}, and

$$p(\boldsymbol{\theta}_j, h^{(j)}|\boldsymbol{\Theta}_{-\boldsymbol{\theta}_j, h^{(j)}}, l, \mathbf{y}) \propto \begin{cases} p(\mathbf{y}|\boldsymbol{\Theta}_l, l)p(\boldsymbol{\Theta}_l), & \text{if } j = l, \\ p(\boldsymbol{\Theta}_j), & \text{if } j \neq l, \end{cases} \qquad (11)$$

where $\boldsymbol{\Theta}_{-\boldsymbol{\theta}_j, h^{(j)}}$ denotes the vector $\boldsymbol{\Theta}$ without the components $\boldsymbol{\theta}_j$ and $h^{(j)}$.

Using jointly the conditional densities (9), (10), and (11), we can obtain the posterior probabilities for all possible values of L, and choose the value with major probability as the estimation of the number of regimes l, or choose the mode of the value of L in the iterations of the Gibbs sampler as the estimation of l.

4.3. Estimation of the number of regimes l and the thresholds

Finally, we assume that the $l - 1$ threshold are also unknown, and they need to be estimated jointly with the number of regimes l. Following the approach of Carlin & Chib [1], the model is indexed by a discrete variable M, this variable takes values $2, \cdots, l_{\max}$ according to the value of the variable L. For each possible model $M = j$, the thresholds are denoted as $\mathbf{r}_j = (r_1, \cdots, r_{j-1})$ with $j = 2, \cdots, l_{\max}$.

It is straightforward to obtain the posterior conditional density of $r_1, \cdots, r_{j-1}$ given the values of other structural and non-structural parameters, given by

$$p(\mathbf{r}_j | l, \boldsymbol{\Theta}_{-\mathbf{r}_j}, \mathbf{y}) \propto \begin{cases} \left(1 + \dfrac{\left[x_t - a_0^{(j_t)} - \sum_{i=1}^{k_{j_t}} a_i^{(j_t)} x_{t-i} \right]^2}{(h^{(j_t)})^2 (n-2)} \right)^{-\frac{n+1}{2}}, & \text{if } j = l, \\[1em] p(\mathbf{r}_j), & \text{if } j \neq l, \end{cases}$$

$$\tag{12}$$

where $j_t = k$ if $Z_t \in R_j = (r_{j-1}, r_j]$ for some $j = 1, \cdots, l$. The posterior conditional density of l is given by (9). This way, using the posterior conditional density of $\mathbf{r}_j$, l, and $\boldsymbol{\Theta}_j$, we can implement a Gibbs sampler and obtain the estimation of the number of regimes and thresholds.

In order to define the prior density of $\mathbf{r}_j$, we recall that the values of thresholds are based on the values of the process $\{Z_t\}$, so we can assume that the thresholds take values in a interval (a, b), correctly specified, furthermore, we assume a uniform distribution for the thresholds $r_1, \cdots, r_{j-1}$, that is,

$$p(\mathbf{r}_j) = p(r_1, \cdots, r_{j-1}) \propto k \quad \text{if } a < r_1 < \cdots < r_{j-1} < b,$$

for $j = 2, \cdots, l_{\max}$.

4.4. Proposed algorithm

In conclusion, a three-stage process is proposed for the identification and estimation of TAR models with t noise as follows:

(1) In the first stage, the number of regimes and thresholds are estimated by using a Gibbs sampler based on the densities (9), (10), (11), and (12).

(2) In the next stage, the number of regimes and thresholds are fixed and the autoregressive orders are estimated by using a Gibbs sampler based on the densities (7) and (8).

(3) Finally, conditioned on the estimated structural parameters, we estimate the non-structural parameters by using a Gibbs sampler with densities (3), (4), and (5).

5. Simulated Example

We simulated a series $\{x_t\}$ of 100 observations

$$X_t = \begin{cases} 1 + 0.5X_{t-1} - 0.3X_{t-2} + e_t, & \text{if } Z_t \leq 0, \\ -0.5 - 0.7X_{t-1} + 1.5e_t, & \text{if } Z_t > 0, \end{cases}$$

with $e_t \sim \dfrac{t_5}{\sqrt{5/3}}$, $Z_t = 0.5Z_{t-1} + \epsilon_t$, and $\epsilon_t \sim RB(0, 1)$. The simulated series are shown in the Figure 1.
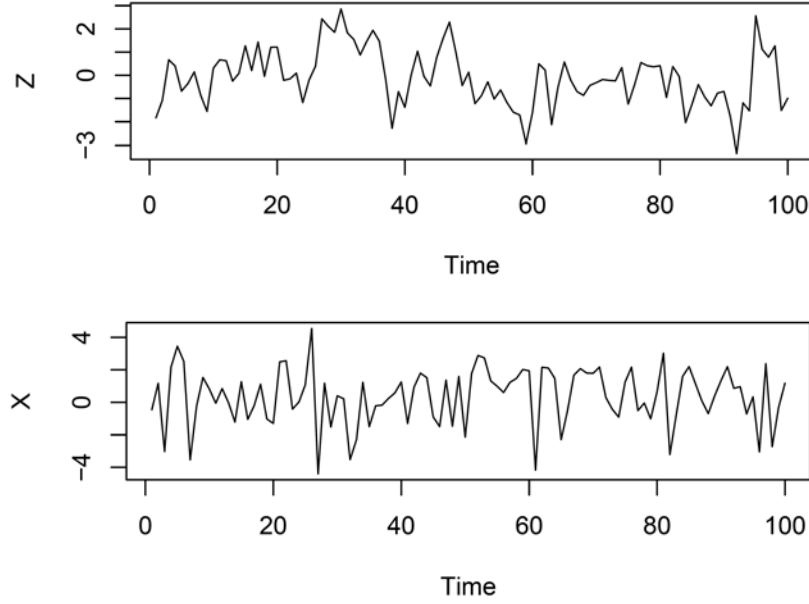

Figure 1. Simulated data.

In the first stage, we identify the number of regimes and the thresholds. The prior distribution for l is the Poisson distribution

truncated in the set {2, 3, 4} with parameter 3, and the prior distribution of the thresholds is as described before. We run a Gibbs sampler of 1000 iterations and the estimation of the number of regimes is $\hat{l} = 2$ with probability 1, i.e., all the simulated values of l are 2 in all the last 800 iterations of the Gibbs sampler. The estimation of the threshold is -0.0561[1]. In the following figure, we present the histogram of the simulated values for the threshold. The 95% interval of credibility for the threshold is given by $(-0.8596, 0.1088)$ containing the real threshold 0.
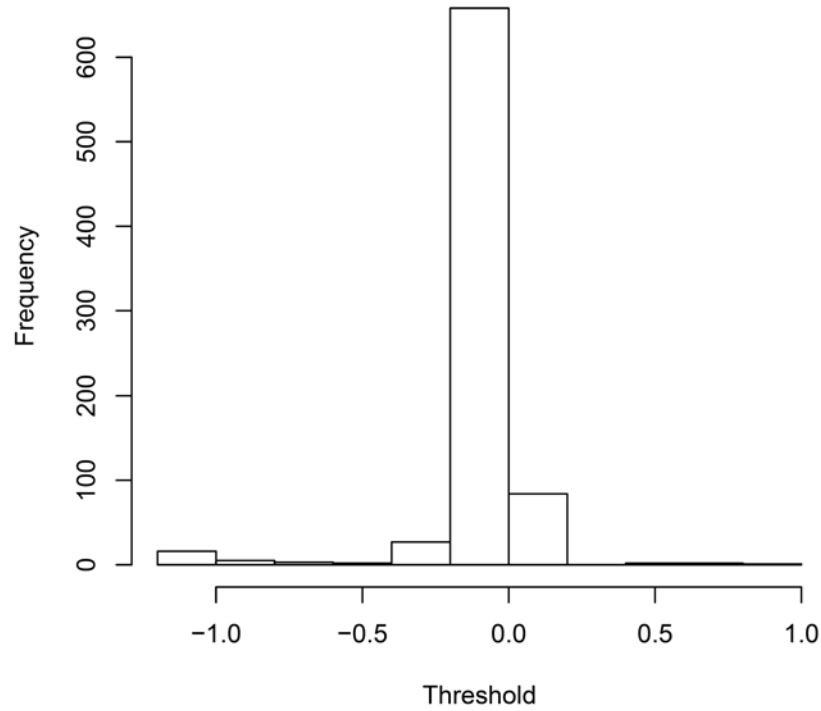

Figure 2. Histogram of the simulated values of the threshold for two regimes.

---

[1] For a certain model modelo $l = j$, the possible values of the thresholds $\mathbf{r}_j$ are the quantiles of the process $\{Z_t\}$, after removing the thresholds that induce regimes with too few data, in this case, we eliminate the thresholds that induce any regime with less than 20 data.

In the second stage, we estimate the autoregressive order in each of the two regimes. The prior distribution for k_j is the truncated Poisson distribution with parameter 2 in the set {0, 1, 2, 3} for each $j = 1, 2$. The maximum autoregressive order is chosed to be the autoregressive order p of the linear model $AR(p)$ which fitted best to the data, which is 3. We run a Gibbs sampler of 1000 iterations, and obtain the posterior probabilities for k_1 and k_2, displayed in the Table 1, we can see that the identified autoregressive orders are $\hat{k}_1 = 2$ and $\hat{k}_2 = 1$, corresponding to the real autoregressive orders.

Table 1. Posterior probabilities of the variables K_1 and K_2

	Regimes	
	1	2
0	0.0887	0.05125
1	0.09375	0.9075
2	0.8175	0.04125
3	0	0

Finally, we estimate the non-structural parameters: autoregressive coefficients, the variance weights, and the degree of freedom of the process of error. The prior distribution for these parameters are: $N(0, 10)$ for the autoregressive coefficients a_i^j with $i = 1, \cdots, k_j$ and $j = 1, 2$; distribution inverse-Gamma $(2, 3)$ for the variance weights $h^{(1)}$ and $h^{(2)}$; distribution Gamma $(1, 0.1)$ for the degree of freedom n, this way, the prior mean of n is 10 and the prior variance is 100, which can be considered as a non-informative prior distribution.

We run another Gibbs sampler of 1000 iterations, the estimation of the autoregressive coefficients and the variance weights are given in the Table 2. These estimations are close to the true parameters and all the 95% credible intervals contain the true parameters.

Table 2. Estimation and 95% credible intervals for the non-structural parameters in the simulated example

Regimes	$a_i^{(j)}$			$h^{(j)}$
1	0.97	0.39	− 0.21	0.92
	(0.73, 1.18)	(0.25, 0.53)	(− 0.32, − 0.09)	(0.71, 1.30)
2	− 0.46	− 0.59		1.52
	(− 0.80, − 0.13)	(− 0.79, − 0.39)		(1.14, 2.26)

With respect to the degree of freedom n, the results obtained from the Gibbs sampler is displayed in the Figure 3, where we note that the values of n with large posterior probability is around the true parameter 5. The posterior mean of n is given by 5.6, and a 95% credible interval of n is given by (2.6, 22.8).
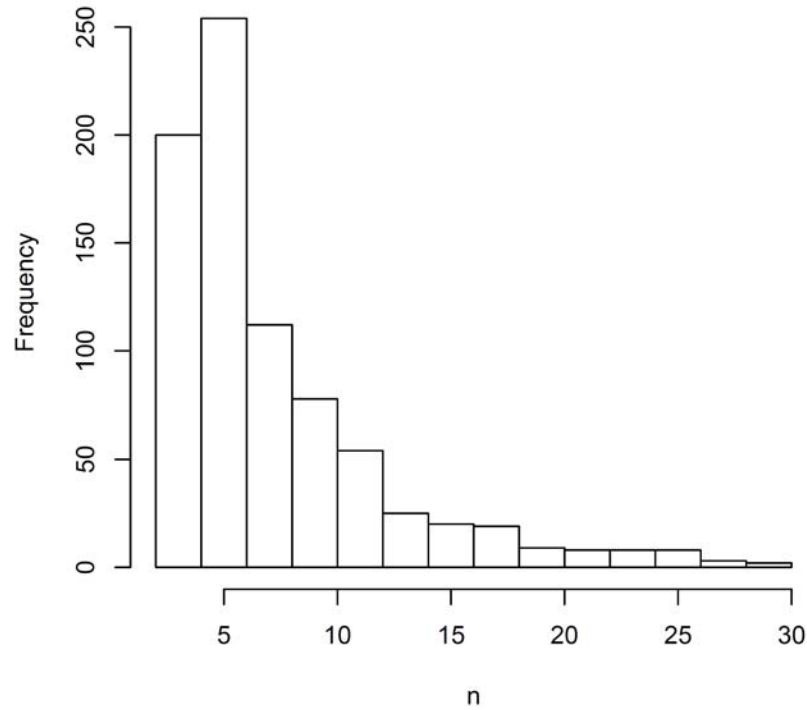

Figure 3. Histogram of the simulated values of the degree of freedom n.

6. Conclusion

In this paper, we presented a Bayesian methodology in order to identify and estimate the TAR models with t noise. The simulated example show that the results obtained from the methodology are acceptable. For future research, we will develop the forecasting procedure for these models, as well as the problem of estimating missing data in both processes $\{X_t\}$ and $\{Z_t\}$.

References

[1]   B. P. Carlin and S. Chib, Bayesian model choice via Markov chain Monte Carlo methods, Journal of the Royal Statistical Society, Serie B 37(3) (1995), 473-484.

[2]   H. Chen, T. T. Chong and J. Bai, Theory and applications of tar model with two threshold variables, Econometric Reviews 31 (2012), 142-170.

[3]   F. H. Nieto, Modelling bivariate threshold autoregressive processes in the presence of missing data, Communications in Statistics, Theory and Methods 34 (2005), 905-930.

[4]   F. H. Nieto, Forecasting with univariate tar models, Statistical Methodology 5 (2008), 263-276.

[5]   T. Sáfadi and P. Morettin, Bayesian analysis of thresholds autoregressive moving average models, The Indian Journal of Statistics 62 (2000), 353-371.

[6]   H. Tong, On a Threshold Model in Pattern Recognition and Signal Processing, C. H. Chen (Editor), Sijhoff & Noord-hoff, Amsterdam, 1978.

[7]   R. S. Tsay, Testing and modelling threshold autoregressive processes, Journal of American Statistical Association 84 (1989), 231-240.

[8]   R. S. Tsay, Testing and modelling multivariate threshold models, Journal of American Statistical Association 93 (1998), 1188-1202.

[9]   T. Watanabe, On sampling the degree-of-freedom of student's-t disturbances, Statistics & Probability Letters 52 (2001), 177-181.

[10] Q. Xia, L. Liu, J. Pan and R. Liang, Bayesian analysis of two-regime threshold autoregressive moving average model with exogeneous inputs, Communications in Statistics-Theory and Methods 41 (2012), 1089-1104.

∎