

Article

A Time-Varying Mixture Integer-Valued Threshold Autoregressive Process Driven by Explanatory Variables

Danshu Sheng¹, Dehui Wang¹, Jie Zhang^{2,*} , Xinyang Wang^{1,*} and Yiran Zhai³¹ School of Mathematics and Statistics, Liaoning University, Shenyang 110031, China² School of Mathematics and Statistics, Changchun University of Technology, Changchun 130012, China³ State Grid Jilin Electric Power Company Limited Information and Telecommunication Company, Changchun 132400, China

* Correspondence: zhangjie@ccut.edu.cn (J.Z.); wangxinyang@lnu.edu.cn (X.W.)

Abstract: In this paper, a time-varying first-order mixture integer-valued threshold autoregressive process driven by explanatory variables is introduced. The basic probabilistic and statistical properties of this model are studied in depth. We proceed to derive estimators using the conditional least squares (CLS) and conditional maximum likelihood (CML) methods, while also establishing the asymptotic properties of the CLS estimator. Furthermore, we employed the CLS and CML score functions to infer the threshold parameter. Additionally, three test statistics to detect the existence of the piecewise structure and explanatory variables were utilized. To support our findings, we conducted simulation studies and applied our model to two applications concerning the daily stock trading volumes of VOW.

Keywords: threshold integer-valued autoregressive models; mixture thinning operator; parameter estimation; Wald test; explanatory variables



Citation: Sheng, D.; Wang, D.; Zhang, J.; Wang, X.; Zhai, Y. A Time-Varying Mixture Integer-Valued Threshold Autoregressive Process Driven by Explanatory Variables. *Entropy* **2024**, *26*, 140. <https://doi.org/10.3390/e26020140>

Academic Editor: Boris Ryabko

Received: 4 January 2024

Revised: 1 February 2024

Accepted: 2 February 2024

Published: 4 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

An integer-valued time series represents count data reflecting the states of a particular phenomenon at different time points. It finds widespread applications across various real-world domains. For instance, in the field of economics, Ref. [1] employed the integer-valued moving average model to describe the number of transactions in intra-day stock data. In industrial contexts, Ref. [2] utilized the compound Poisson integer-valued autoregressive (INAR) model to characterize the count of workers in the heavy manufacturing industry receiving benefits due to burn-related injuries. Within the realm of insurance actuarial studies, Ref. [3] explored an extension of the classical discrete-time risk model, incorporating an INAR(1) process to capture temporal dependence among claim counts. A straightforward approach in the modeling and analyzing of count time series involves creating an integer-valued autoregressive model by using thinning operators. Ever since [4] introduced the first INAR(1) time series model, which relied on the binomial thinning operator [5], it has become a prevalent method to model INAR-type models utilizing various thinning operators (see [6–10], among others). This approach has been extensively applied across diverse fields such as epidemiology, social sciences, economics, life sciences, and more.

Even though INAR-type models are commonly used in practical applications, they often fall short when confronted with nonlinear phenomena. For instance, researchers in the field of epidemiology, as exemplified by [11], have detected temporal fluctuations in the incidence rate of epidemics. Therefore, trying to represent such data with an INAR-type model may not be the most-appropriate approach. Furthermore, time series data frequently undergo sudden changes that can either temporarily or permanently disrupt their dynamics. In such situations, the conventional INAR model may also prove inadequate in delivering an accurate fit. To address the nonlinear aspects of integer-valued time

series data, Ref. [12] introduced the integer-valued self-exciting threshold autoregressive (SETINAR(2,1)) process, which relies on the binomial thinning operator (“ \circ ”); Ref. [13] presented the self-excited threshold Poisson autoregressive (SETPAR) model and applied it to analyze major global earthquake data; Ref. [14] proposed a basic self-exciting threshold binomial AR(1) model (SETBAR(1)) with values across a finite range of counts; Ref. [15] investigated an integer-valued threshold autoregressive process (NBTINAR(1)) based on the negative binomial thinning operator (“ $*$ ”) and applied it to analyze the annual counts of major earthquakes with a magnitude of 7 or above from 1900 to 2015. In a comprehensive review, Ref. [16] surveyed threshold models for integer-valued time series with an infinite range and introduced two novel models tailored to cases with a finite range of values. In the latest research, Ref. [17] pointed out that employing different operators before and after the threshold can enhance the model’s ability to explain a wider range of phenomena. As a result, she has proposed the following threshold autoregressive model using the mixed operators.

$$X_t = \begin{cases} \alpha_1 \circ X_{t-1} + Z_{1,t}, & X_{t-1} \leq r \\ \alpha_2 * X_{t-1} + Z_{2,t}, & X_{t-1} > r; \end{cases}$$

where $\{Z_{1,t}\}$ and $\{Z_{2,t}\}$ are sequences of i.i.d. Poisson and Geometric distributed random variables, respectively. However, it is worth noting that the use of constant autoregressive coefficients in this model ignores the effect of exogenous variables on the observed data. For example, denote the daily trading volume of a specific stock as X_t . Clearly, in practice, its autoregressive coefficient is often not static and is often subject to some external factors related to change over time, such as: market factors: the overall volatility of the stock market can impact the volatility of individual stocks; industry factors: specific events or trends within an industry can also affect stock price fluctuations; interest rates and monetary policy: changes in interest rates and monetary policies can have a wide-ranging impact on the stock market; political and geopolitical factors: political events, elections, international relations, and geopolitical tensions can introduce uncertainty and volatility to the stock market. emotional and investor behavior: investor sentiment and behavior can significantly influence stock price movements, among other factors.

Inspired by the above discussion and learning the method of constructing models driven by explanatory variables (see [18,19], just to name a few), in this paper, we propose a first-order time-varying mixture thinning integer-valued threshold autoregressive (TVMTTINAR(1)) process driven by explanatory variables. For this, the definition of the TVMTTINAR(1) model is given, and the statistical inference for the proposed model is studied. Furthermore, considering that verifying the existence of a piecewise structure and explanatory variables is key to model construction, we propose three kinds of test statistics. Finally, from the simulation and two applications, we can also see that our proposed model is very competitive.

The organization of this paper is as follows. Section 2 gives the definition of the proposed model, and some properties are also investigated. In Section 3, the estimates of the model parameters are derived by using the conditional least squares (CLS) and conditional maximum likelihood (CML) methods. Three test statistics are also constructed to test the existence of the piecewise structure and explanatory variables, respectively. Some simulation studies are carried out to investigate the performances of the proposed estimates and test statistics in Section 4. Two real data examples are given in Section 5. Some concluding remarks are given in Section 6. All proofs are postponed to Appendix A.

2. The First-Order Time-Varying Mixture Thinning Integer-Valued Threshold Autoregressive Model

We first introduce the definition of the TVMTTINAR(1) process. Furthermore, we investigate the statistical properties of the proposed model.

Definition 1. The process $\{X_t\}$ is called the TVMTTINAR(1) process if X_t follows the recursion:

$$X_t = \begin{cases} \phi_{1,t} \circ X_{t-1} + \varepsilon_t, & X_{t-1} \leq r \\ \phi_{2,t} * X_{t-1} + \varepsilon_t, & X_{t-1} > r, \end{cases} \tag{1}$$

or

$$X_t = \begin{cases} \phi_{2,t} * X_{t-1} + \varepsilon_t, & X_{t-1} \leq r \\ \phi_{1,t} \circ X_{t-1} + \varepsilon_t, & X_{t-1} > r. \end{cases} \tag{2}$$

For convenience, we write the above two models by the symbol R as follows:

$$X_t = (\phi_{1,t} \circ X_{t-1})I_{1,t}^R + (\phi_{2,t} * X_{t-1})I_{2,t}^R + \varepsilon_t, \quad t \in \mathbb{Z}, \tag{3}$$

where:

1. $I_{1,t}^R = \begin{cases} I\{X_{t-1} \leq r\}, R = 0, \\ I\{X_{t-1} > r\}, R = 1, \end{cases}$ and $I_{2,t}^R = 1 - I_{1,t}^R = \begin{cases} I\{X_{t-1} > r\}, R = 0, \\ I\{X_{t-1} \leq r\}, R = 1; \end{cases}$ That is, $R = 0$ indicates that TVMTTINAR(1) represents the process (1).
2. For fixed $i \in 1, 2$, $\phi_{i,t} \in (0, 1)$

$$\log \left(\frac{\phi_{i,t}}{1 - \phi_{i,t}} \right) = \mathbf{Z}_t^\top \boldsymbol{\beta}_i,$$

where $\boldsymbol{\beta}_i = (\beta_{i,0}, \beta_{i,1}, \dots, \beta_{i,q})^\top$ are the regression coefficients, $\{\mathbf{Z}_t := (1, Z_{1,t}, \dots, Z_{q,t})^\top\}_{t \in \mathbb{Z}}$ is a sequence of stationary, weakly dependent, and observable explanatory variables with a constant mean vector and covariance matrix. For fixed t , \mathbf{Z}_t is assumed to be independent of $\{X_{t-l}\}_{l \geq 1}$.

3. The binomial thinning operator “ \circ ”, proposed by [5], is defined as $\phi_1 \circ X = \sum_{i=1}^X B_i$, where $\phi \in (0, 1)$, $\{B_i\}$ is a sequence of i.i.d. Bernoulli random variables satisfying $P(B_i = 1) = 1 - P(B_i = 0) = \phi$ and B_i is independent of X .
4. The negative binomial thinning operator “ $*$ ”, proposed by [20], is defined as $\phi * X = \sum_{i=1}^X W_i$, where $\phi \in (0, 1)$, $\{W_i\}$ is a sequence of i.i.d. Geometric random variables with parameter $\frac{\phi}{1+\phi}$ and W_i is independent of X .
5. $\{\varepsilon_t\}$ is a sequence of i.i.d. Poisson distributed random variables with mean λ . For fixed t , ε_t is assumed to be independent of $\phi \circ X_{t-1}$, $\phi * X_{t-1}$, and X_{t-l} for all $l \geq 1$.

In contrast to the common SETINAR-type model, the TVMTTINAR model does not require $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$. This is mainly due to the existence of mixed thinning operators. Even when $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$, there is a piecewise structure. However, this is a small probability event, and the model inference problem in this case is not specially considered in this paper. In addition, in practical applications, we can usually choose which of the two TVMTTINAR(1) models ($R = 0$ and $R = 1$) is more applicable based on criteria such as the AIC and BIC. We will conduct a specific analysis through data examples in Section 5.

Next, we are ready to state that there exists the strict stationarity and ergodicity of the TVMTTINAR(1) process. Note that, under the assumption that $\boldsymbol{\beta}_i$ satisfies $\sup_{\mathbf{Z} \in \mathbb{R}^{q+1}} |\boldsymbol{\beta}_i \mathbf{Z}| < \infty$ for $i = 1, 2$, there is

$$\phi_{i,t} = \frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)}{1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)} \in (0, 1), \forall t.$$

Thereby, similar to the method in [19], it is easy to verify that Theorem 3.1 in [21] holds, and the TVMTTINAR(1) process is a strictly stationary and ergodic Markov chain. Moreover, its transition probabilities are given by

$$\begin{aligned}
 P(\mathbf{z}_t, x_{t-1}, x_t) &= P(X_t = x_t | X_{t-1} = x_{t-1}, \mathbf{Z}_t = \mathbf{z}_t) \\
 &= P(\phi_{1,t} \circ X_{t-1} I_{1,t}^R + \phi_{2,t} * X_{t-1} I_{2,t}^R + \varepsilon_t = x_t | X_{t-1} = x_{t-1}, \mathbf{Z}_t = \mathbf{z}_t) \\
 &= p(x_{t-1}, x_t, \phi_{1,t}, \lambda) I_{1,t}^R + p(x_{t-1}, x_t, \phi_{2,t}, \lambda) I_{2,t}^R \\
 &= \sum_{i=1}^2 p(x_{t-1}, x_t, \phi_{i,t} I_{i,t}^R, \lambda),
 \end{aligned}
 \tag{4}$$

where

$$\begin{aligned}
 p(x_{t-1}, x_t, \phi_{1,t} I_{1,t}^R, \lambda) &= I_{1,t}^R \sum_{m=0}^{\min(x_{t-1}, x_t)} \binom{x_{t-1}}{m} e^{-\lambda} \frac{\lambda^{x_t-m}}{(x_t-m)!} \phi_{1,t}^m (1 - \phi_{1,t})^{x_{t-1}-m}, \\
 p(x_{t-1}, x_t, \phi_{2,t} I_{2,t}^R, \lambda) &= I_{2,t}^R \sum_{m=0}^{x_t} \frac{\Gamma(x_{t-1} + m)}{\Gamma(x_{t-1}) \Gamma(m + 1)} \frac{\phi_{2,t}^m}{(1 + \phi_{2,t})^{x_{t-1}+m}} e^{-\lambda} \frac{\lambda^{x_t-m}}{(x_t-m)!}.
 \end{aligned}
 \tag{5}$$

Since the existence of the first four moments under observable explanatory variables is a necessary condition for deriving the asymptotic properties of the parameter estimation in Section 3, we then give the following proposition.

Proposition 1. *Let $\{X_t\}$ be the process defined by Definition 1. Then, the first four conditional moments are bounded, that is $E(X_t^k | \mathbf{Z}_0, \dots, \mathbf{Z}_t) < \infty$ for $k = 1, 2, 3, 4$.*

Next, we consider the moments and conditional moments of TVMTTINAR(1). For the simplicity of the notation, we denote $E(I_{1,t}^R) = p_1, E(I_{2,t}^R) = p_2 = 1 - p_1, \mu_1 := E(X_t | X_t \leq r), \mu_2 := E(X_t | X_t > r), \mu_{\phi_1}^{(h)} := E(\frac{\exp(\mathbf{Z}_{t+h}^\top \beta_1)}{1 + \exp(\mathbf{Z}_{t+h}^\top \beta_1)} | X_{t+h-1} \leq r), \mu_{\phi_2}^{(h)} := E(\frac{\exp(\mathbf{Z}_{t+h}^\top \beta_2)}{1 + \exp(\mathbf{Z}_{t+h}^\top \beta_2)} | X_{t+h-1} > r), \phi_i := E(\frac{\exp(\mathbf{Z}_t^\top \beta_i)}{1 + \exp(\mathbf{Z}_t^\top \beta_i)}) (i = 1, 2), \sigma_{\phi_i}^2 := \text{Var}(\frac{\exp(\mathbf{Z}_t^\top \beta_i)}{1 + \exp(\mathbf{Z}_t^\top \beta_i)}) (i = 1, 2), \sigma_1^2 := \text{Var}(X_t | X_t \leq r), \sigma_2^2 := \text{Var}(X_t | X_t > r), \gamma_h^{(1)} := \text{Cov}(X_t, X_{t+h} | X_{t+h} \leq r), \gamma_h^{(2)} := \text{Cov}(X_t, X_{t+h} | X_{t+h} > r), where $\gamma_0^{(i)} = [(\sigma_i^2 + \mu_i^2) - \mu_i E(X_t)], i = 1, 2$.$

The conditional mean and variance for the TVMTTINAR(1) model can be given by

$$\begin{aligned}
 E(X_t | X_{t-1}, \mathbf{Z}_t) &= \sum_{i=1}^2 \frac{\exp(\mathbf{Z}_t^\top \beta_i)}{1 + \exp(\mathbf{Z}_t^\top \beta_i)} X_{t-1} I_{i,t}^R + \lambda, \\
 E(X_t | \mathbf{Z}_t) &= \sum_{i=1}^2 \frac{\exp(\mathbf{Z}_t^\top \beta_i)}{1 + \exp(\mathbf{Z}_t^\top \beta_i)} p_i \mu_i + \lambda, \\
 \text{Var}(X_t | X_{t-1}, \mathbf{Z}_t) &= \frac{\exp(\mathbf{Z}_t^\top \beta_1)}{[1 + \exp(\mathbf{Z}_t^\top \beta_1)]^2} X_{t-1} I_{1,t}^R + \frac{\exp(\mathbf{Z}_t^\top \beta_2)[1 + 2 \exp(\mathbf{Z}_t^\top \beta_2)]}{[1 + \exp(\mathbf{Z}_t^\top \beta_2)]^2} X_{t-1} I_{1,t}^R + \lambda, \\
 \text{Var}(X_t | \mathbf{Z}_t) &= \sum_{i=1}^2 \left\{ \left[\frac{\exp(\mathbf{Z}_t^\top \beta_i)}{1 + \exp(\mathbf{Z}_t^\top \beta_i)} \right]^2 [p_i(\sigma_i^2 + \mu_i^2) - p_i^2 \mu_i^2] \right\} + \frac{\exp(\mathbf{Z}_t^\top \beta_1)}{[1 + \exp(\mathbf{Z}_t^\top \beta_1)]^2} p_1 \mu_1 \\
 &\quad + \frac{\exp(\mathbf{Z}_t^\top \beta_2)[1 + 2 \exp(\mathbf{Z}_t^\top \beta_2)]}{[1 + \exp(\mathbf{Z}_t^\top \beta_2)]^2} p_2 \mu_2 - 2 \prod_{i=1}^2 \frac{\exp(\mathbf{Z}_t^\top \beta_i)}{1 + \exp(\mathbf{Z}_t^\top \beta_i)} p_i \mu_i + \lambda.
 \end{aligned}$$

The unconditional expressions for the marginal mean and variance of the TVMTTINAR(1) model are

$$E(X_t) = \sum_{i=1}^2 p_i \phi_i \mu_i + \lambda,$$

$$\text{Var}(X_t) = \sum_{i=1}^2 \left\{ \phi_i^2 [p_i(\sigma_i^2 + \mu_i^2) - p_i^2 \mu_i^2] + p_i \sigma_{\phi_i}^2 (\sigma_i^2 + \mu_i^2) \right\} + p_1 \mu_1 (\phi_1 - \sigma_{\phi_1}^2 - \phi_1^2) + p_2 \mu_2 (\phi_2 + \sigma_{\phi_2}^2 + \phi_2^2) - 2p_1 p_2 \phi_1 \phi_2 \mu_1 \mu_2 + \lambda.$$

Then, we have that the autocovariance function and autocorrelation function (ACF):

$$\text{Cov}(X_t, X_{t+h} | \mathbf{Z}_{t+1}, \dots, \mathbf{Z}_{t+h}) = \sum_{i=1}^2 \frac{\exp(\mathbf{Z}_{t+h}^\top \boldsymbol{\beta}_i)}{1 + \exp(\mathbf{Z}_{t+h}^\top \boldsymbol{\beta}_i)} p_i \gamma_{h-1}^{(i)},$$

$$\text{Cov}(X_t, X_{t+h}) = \sum_{i=1}^2 \mu_{\phi_i}^{(h)} p_i \gamma_{h-1}^{(i)},$$

$$\rho(h) := \text{Corr}(X_t, X_{t+h}) = \left[\sum_{i=1}^2 \mu_{\phi_i}^{(h)} p_i \gamma_{h-1}^{(i)} \right] \setminus \text{Var}(X_t)$$

3. Parameters' Estimation and Testing

Suppose we have a series of observations $\{X_t\}_{t=1}^n$ generated from the TVMTINAR(1) process. Denote $\boldsymbol{\theta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top, \lambda)^\top$ as the parameter vector under the known threshold parameter r , and $\boldsymbol{\eta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top, \lambda, r)^\top$ as the parameter vector under the unknown r case. Their parameter spaces are

$$\Theta_{\boldsymbol{\theta}} = \{\boldsymbol{\theta} \in \mathbb{R}^{q+1} \times \mathbb{R}^{q+1} \times (0, +\infty)\},$$

$$\Theta_{\boldsymbol{\eta}} = \{\boldsymbol{\eta} \in \mathbb{R}^{q+1} \times \mathbb{R}^{q+1} \times (0, +\infty) \times \mathbb{N}\}.$$

Furthermore, suppose the true values $\boldsymbol{\theta}_0 = (\boldsymbol{\beta}_{1,0}^\top, \boldsymbol{\beta}_{2,0}^\top, \lambda_0)^\top$ and $\boldsymbol{\eta}_0 = (\boldsymbol{\beta}_{1,0}^\top, \boldsymbol{\beta}_{2,0}^\top, \lambda_0, r_0)^\top$ of $\boldsymbol{\theta}_0$ and $\boldsymbol{\eta}_0$ are the interior points of $\Theta_{\boldsymbol{\theta}}$ and $\Theta_{\boldsymbol{\eta}}$, respectively. In this section, we mainly implement parameter estimation based on two different approaches, namely the conditional least squares (CLS) and conditional maximum likelihood (CML) methods. The objective function is not differentiable with respect to the threshold variable r since r is an integer. Therefore, we firstly propose solutions to estimate $\boldsymbol{\theta}$ under the assumption that the threshold variable r is known. Later, in Section 3.3, we turn to estimating the threshold variable r based on the estimation methods mentioned before. All the proofs are presented in Appendix A.

3.1. Conditional Least Squares Estimation

Denote

$$g(\boldsymbol{\theta}, X_{t-1}, \mathbf{Z}_t) = E(X_t | X_{t-1}, \mathbf{Z}_t) = \frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_1)}{1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_1)} X_{t-1} I_{1,t}^R + \frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_2)}{1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_2)} X_{t-1} I_{2,t}^R + \lambda,$$

$$Q(\boldsymbol{\theta}) := \sum_{t=1}^n (X_t - g(\boldsymbol{\theta}, X_{t-1}, \mathbf{Z}_t))^2 = \sum_{t=1}^n U_t^2(\boldsymbol{\theta}),$$

where

$$U_t(\boldsymbol{\theta}) = X_t - \frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_1)}{1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_1)} X_{t-1} I_{1,t}^R - \frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_2)}{1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_2)} X_{t-1} I_{2,t}^R - \lambda.$$

Then, the CLS estimator $\hat{\boldsymbol{\theta}}_{CLS} := (\hat{\boldsymbol{\beta}}_{1,CLS}^\top, \hat{\boldsymbol{\beta}}_{2,CLS}^\top, \hat{\lambda}_{CLS})^\top$ of $\boldsymbol{\theta}$ is obtained by minimizing the sum of the squared deviations, that is

$$\hat{\boldsymbol{\theta}}_{CLS} = \arg \min_{\boldsymbol{\theta} \in \Theta_{\boldsymbol{\theta}}} Q(\boldsymbol{\theta}). \tag{6}$$

Since the TVMTTINAR(1) model is stationary and ergodic and the first four moments are bounded, then using the Taylor expansion and the martingale central limit theorem, the following theorem about the consistency and asymptotic normality of the parameter estimator can be obtained. The detailed proof is presented in Appendix A.

Theorem 1. *Let $\{X_t\}$ be a TVMTTINAR(1) process. Then, the CLS estimator $\hat{\theta}_{CLS}$ is consistent and has the asymptotic distribution:*

$$\sqrt{n}(\hat{\theta}_{CLS} - \theta_0) \xrightarrow{d} N(\mathbf{0}, \mathbf{V}^{-1}\mathbf{W}\mathbf{V}^{-1}), \tag{7}$$

where \mathbf{V} and \mathbf{W} are square matrices of order $2q + 3$ with the jk th element given by

$$V_{jk} := E \left(\frac{\partial}{\partial \theta_j} g(\theta, X_{t-1}, \mathbf{Z}_t) \frac{\partial}{\partial \theta_k} g(\theta, X_{t-1}, \mathbf{Z}_t) \right)_{\theta_0},$$

$$W_{jk} := E \left(U_t^2(\theta) \frac{\partial}{\partial \theta_j} g(\theta, X_{t-1}, \mathbf{Z}_t) \frac{\partial}{\partial \theta_k} g(\theta, X_{t-1}, \mathbf{Z}_t) \right)_{\theta_0}.$$

3.2. Conditional Maximum Likelihood Estimation

For a fixed value of x_0 , the conditional log-likelihood function for the TVMTTINAR(1) model can be written as

$$L(\theta) := \sum_{t=1}^n \ell_t(\theta) = \sum_{t=1}^n \log P(z_t, X_{t-1}, X_t)$$

where $P(z_t, X_{t-1}, X_t)$ is the transition probabilities defined in (4). The CML estimator $\hat{\theta}_{CML} := (\hat{\beta}_{1,CML}^\top, \hat{\beta}_{2,CML}^\top, \hat{\lambda}_{CML})^\top$ of θ is obtained by maximizing the conditional log-likelihood function, that is

$$\hat{\theta}_{CML} = \arg \max_{\theta \in \Theta_\theta} L(\theta). \tag{8}$$

Since $\phi_{i,t}$ is nonlinear with respect to $\beta_{i,j}$ for arbitrary $i = 1, 2$ and $j = 0, 1, \dots, q$, there is no closed-form expressions for the CLS and CML estimators. The numerical solutions can be solved by the MATLAB(2021b) function “fmincon” or the R(4.2.1) function “optim”. The implementation details and performance are discussed in Section 4.

Theorem 2. *Let $\{X_t\}$ be a TVMTTINAR(1) process. Assume that the function $E[\ell_t(\theta)]$ has a unique maximizer in the compact parameter space Θ , $E\left(\frac{1}{n} \frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^\top}\right)_{\theta_0}$ is a nonsingular matrix, and for a neighborhood of θ_0 , say $N(\theta_0)$. For any $i, j, k = 1, \dots, 2q + 3$, there is*

$$\overline{\lim}_{n \rightarrow \infty} \sup_{\theta \in N(\theta_0)} \left| \frac{\partial^3 \ell_t(\theta)}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| < \infty.$$

Then, the CML estimator $\hat{\theta}_{CML}$ is consistent and has the asymptotic distribution:

$$\sqrt{n}(\hat{\theta}_{CML} - \theta_0) \xrightarrow{d} N(\mathbf{0}, \mathbf{J}^{-1}(\theta_0)\mathbf{I}(\theta_0)\mathbf{J}^{-1}(\theta_0)), \tag{9}$$

as $n \rightarrow \infty$, where $\mathbf{I}(\theta_0) = E\left[\frac{\partial \ell_t(\theta)}{\partial \theta} \frac{\partial \ell_t(\theta)}{\partial \theta^\top}\right]_{\theta_0}$, $\mathbf{J}(\theta_0) = E\left(\frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta^\top}\right)_{\theta_0}$.

3.3. Inference Methods for Threshold r

In this section, we concentrate on the estimation of threshold variable r . Since r is an integer, different from the continuous-type threshold models, the integer-type threshold models usually use a one-by-one search on a fixed interval $[\underline{r}, \bar{r}]$ to make the loss function optimal to obtain the threshold estimator. In applications, typically, the empirical 10th and 90th quantile value of the sample as \underline{r} and \bar{r} is used, respectively. The methods commonly used at present are the CLS and CML. For the CLS method, the estimation of the threshold variable r can be achieved based on the following steps:

Step 1. Denote \underline{r} and \bar{r} as the 10th and 90th quantile value of the observations $\{X_1, \dots, X_n\}$, for each $r \in [\underline{r}, \bar{r}] \cap \mathbb{N}$, and find \hat{r}_{CLS} such that

$$\hat{r}_{CLS} = \arg \min_{r \in [\underline{r}, \bar{r}] \cap \mathbb{N}} Q(\theta(r)).$$

Step 2. The parameter vector $\hat{\theta}_{CLS}(\hat{r}_{CLS})$ is estimated by (6) under the estimator \hat{r}_{CLS} , and all the parameters under r unknown cases are as follows:

$$\hat{\eta}_{CLS} = (\hat{\theta}_{CLS}^T(\hat{r}_{CLS}), \hat{r}_{CLS})^T.$$

Similarly, the CML estimates for the threshold variable r can also be achieved based on the following steps:

Step 1. Denote \underline{r} and \bar{r} as the 10th and 90th quantile value of the observations $\{X_1, \dots, X_n\}$, for each $r \in [\underline{r}, \bar{r}] \cap \mathbb{N}$, and find \hat{r}_{CML} such that

$$\hat{r}_{CML} = \arg \max_{r \in [\underline{r}, \bar{r}] \cap \mathbb{N}} L(\theta(r)).$$

Step 2. The parameter vector $\hat{\theta}_{CML}(\hat{r}_{CML})$ is estimated by (6) under the estimator \hat{r}_{CML} , and all the parameters under r unknown cases are as follows:

$$\hat{\eta}_{CML} = (\hat{\theta}_{CML}^T(\hat{r}_{CML}), \hat{r}_{CML})^T.$$

Similar procedures can be found in [13,14].

3.4. Testing the Existence of the Piecewise Structure

Threshold models are typically characterized by piecewise linearization, which divides a complex system into regimes using a specific threshold. Therefore, testing to detect the existence of segmented structures is very necessary. To date, many researchers have come up with different test statistics. A common high-performance method is to construct the likelihood ratio (LR) test based on the conditional likelihood function; see [14]. However, the LR test cannot be implemented because the TVMTTINAR(1) model is constructed by two operators. In this paper, we constructed the Wald test statistics to detect the existence of piecewise structures in the TVMTTINAR(1) model. The null hypothesis and the alternative hypothesis take the form:

$$H_0^{(1)} : \beta_1 = \beta_2 \quad \text{vs.} \quad H_1^{(1)} : \beta_1 \neq \beta_2. \quad (10)$$

Note that, although the TVMTTINAR(1) model does not degenerate to the INAR-type model when $\beta_1 = \beta_2$, the probability of this happening is extremely small and will not be considered here. That is to say, we only prove the existence of the piecewise structure by testing $\beta_1 \neq \beta_2$. A simple idea, learned from [22], is to use the test of the difference between two normal population means based on the asymptotic normality of some consistent estimators. Then, we construct the Wald test based on the asymptotic distribution (9) of the CLS estimator $\hat{\theta}_{CLS}$ and obtain the following result.

Let $\hat{\Sigma} = \hat{V}^{-1}\hat{W}\hat{V}^{-1}$, where \hat{V} and \hat{W} are square matrices of order $2q + 3$ with the jk th element given by

$$\hat{V}_{jk} := \left(\frac{1}{n} \sum_{t=1}^n \frac{\partial}{\partial \theta_j} g(\theta, X_{t-1}, \mathbf{Z}_t) \frac{\partial}{\partial \theta_k} g(\theta, X_{t-1}, \mathbf{Z}_t) \right)_{\theta=\hat{\theta}_{CLS}},$$

$$\hat{W}_{jk} := \left(\frac{1}{n} \sum_{t=1}^n U_t^2(\theta) \frac{\partial}{\partial \theta_j} g(\theta, X_{t-1}, \mathbf{Z}_t) \frac{\partial}{\partial \theta_k} g(\theta, X_{t-1}, \mathbf{Z}_t) \right)_{\theta=\hat{\theta}_{CLS}}.$$

Obviously, they are the consistent estimators of V and W (defined in Theorem 1). Then, the statistic for testing the problem (10) is defined by

$$T_n^{(1)} = \frac{\mathbf{I}(\hat{\beta}_{1,CLS} - \hat{\beta}_{2,CLS})}{\sqrt{\mathbf{A}\hat{\Sigma}\mathbf{A}^\top/n}},$$

where $\mathbf{I} = (1, \dots, 1)_{1 \times (q+1)}$, $\mathbf{A} = (\mathbf{I}, -\mathbf{I})_{1 \times 2(q+1)}$. Furthermore, under $H_0^{(1)}$,

$$T_n^{(1)} \xrightarrow{d} N(0, 1), \text{ as } n \rightarrow \infty.$$

3.5. Testing the Existence of Explanatory Variables

The existence of observable explanatory variables in the TVMTTINAR(1) model constructs time-varying characteristics. Once the explanatory variables are not present, the model degrades to a constant-coefficient mixture thinning operator threshold INAR model (MTTINAR(1)), i.e.,

$$X_t = (\alpha_1 \circ X_{t-1})I_{1,t}^R + (\alpha_2 * X_{t-1})I_{2,t}^R + \varepsilon_t, \quad t \in \mathbb{Z}, \tag{11}$$

where $\alpha_i \in (0, 1), i = 1, 2$. Therandom variables are similar to Definition 1. A more-general version of such a problem is to test whether the explanatory variable coefficients $\beta_{i,j} (i = 1, 2, j = 1, 2, \dots, q)$ in each regime are all zeros, i.e.,

$$H_0^{(2)} : \beta_{i,j} = 0, i = 1, 2, j = 1, \dots, q \text{ vs. } H_1^{(2)} : \text{At least one } \beta_{i,j} \neq 0, i \in \{1, 2\}, 1 \leq j \leq q. \tag{12}$$

For this, we construct the following two test statistics. The first method is to construct a test statistic using the asymptotic normality of the estimator $\hat{\theta}_{CLS}$. Let $\mathbf{0}_{j \times k}$ be a zero matrix with j rows and k columns, $\mathbf{B} = (\mathbf{0}_{q \times 1}, \mathbf{I}_{q \times q})$, $\mathbf{C} = \begin{pmatrix} \mathbf{B} & \mathbf{0}_{q \times (q+1)} \\ \mathbf{0}_{q \times (q+1)} & \mathbf{B} \end{pmatrix}$. We construct the test statistic as follows:

$$T_n^{(2)} = n\hat{\theta}_{CLS}^\top \mathbf{C}^\top (\mathbf{C}\hat{\Sigma}\mathbf{C}^\top)^{-1} \mathbf{C}\hat{\theta}_{CLS}.$$

Then, under $H_0^{(2)}$,

$$T_n^{(2)} \xrightarrow{d} \chi_{2q}^2, \text{ as } n \rightarrow \infty.$$

Another approach is to construct a classical likelihood ratio (LR) test statistic. Let $\tilde{\theta} := (\alpha_1, \alpha_2, \lambda)^\top$ be the parameter of the MTTINAR(1) model with the parameter set $\Theta_{\tilde{\theta}}$:

$$\Theta_{\tilde{\theta}} = \{\tilde{\theta} \in (0, 1) \times (0, 1) \times (0, +\infty)\}.$$

Then, the LR statistic for testing problem (12) is defined by

$$T_n^{(3)} = 2(\max_{\Theta_{\theta}} L(\theta) - \max_{\Theta_{\tilde{\theta}}} \tilde{L}(\tilde{\theta})),$$

where $\tilde{L}(\tilde{\theta})$ is the conditional log-likelihood function for the MTTINAR(1) model (11). Suppose we have a series of observations $\{x_t\}_{t=1}^n$ generated from the MTTINAR(1) process, then $\tilde{L}(\tilde{\theta})$ is given by

$$\begin{aligned} \tilde{L}(\tilde{\theta}) &:= \sum_{t=1}^n \tilde{\ell}_t(\tilde{\theta}) = \sum_{t=1}^n \log \tilde{P}(x_{t-1}, x_t), \\ \tilde{P}(x_{t-1}, x_t) &= I_{1,t}^R \left(\sum_{m=0}^{\min(x_{t-1}, x_t)} \binom{x_{t-1}}{m} e^{-\lambda} \frac{\lambda^{x_t-m}}{(x_t-m)!} \alpha_1^m (1-\alpha_1)^{x_{t-1}-m} \right) \\ &\quad + I_{2,t}^R \left(\sum_{m=0}^{x_t} \frac{\Gamma(x_{t-1}+m)}{\Gamma(x_{t-1})\Gamma(m+1)} \frac{\alpha_2^m}{(1+\alpha_2)^{x_{t-1}+m}} e^{-\lambda} \frac{\lambda^{x_t-m}}{(x_t-m)!} \right). \end{aligned}$$

Furthermore, under $H_0^{(2)}$,

$$T_n^{(3)} \xrightarrow{d} \chi_{2q}^2, \text{ as } n \rightarrow \infty.$$

4. Simulation Studies

To evaluate the finite-sample performance of the proposed inference methods and testing statistics, we conducted extensive simulation studies and split the simulation studies into the following four parts. In the first two parts, we considered the performance of the CLS and CML estimators in two cases where threshold r is known and unknown. In the third and fourth parts, we mainly focused on the performance of the proposed test statistics by empirical sizes and powers.

To get started, we first introduce the following models applied to Sections 4.1 and 4.2. The models are divided into the A-type and B-type models, which represent the $R = 0$ and $R = 1$ TVMTTINAR(1) models, respectively. The two types of models choose similar parameters. In order to save space, we introduce the A-type model, while the parentheses represent the similar B-type model:

Model A1 (B1): Generated from the TVMTTINAR(1) process (3) with $R = 0$ ($R = 1$), $\lambda = 5$, $r = 6$, $(\beta_{1,0}, \beta_{1,1}) = (0.1, 0.3)$, $(\beta_{2,0}, \beta_{2,1}) = (-0.5, -0.6)$. The explanatory variables $Z_{1,t}$ are generalized from the i.i.d. normal distribution $N(0, 1)$.

Model A2 (B2): Generated from the TVMTTINAR(1) process (3) with $R = 0$ ($R = 1$), $\lambda = 5$, $r = 6$, $(\beta_{1,0}, \beta_{1,1}) = (0.3, 0.3)$, $(\beta_{2,0}, \beta_{2,1}) = (-0.5, -0.6)$. The explanatory variables $Z_{1,t}$ are generalized from an AR(1) process, i.e., $Z_{1,t} = 0.5Z_{1,t-1} + \epsilon_t$ with $Z_{1,0} = 0$, $\epsilon_t \sim N(0, 1)$.

Model A3 (B3): Generated from the TVMTTINAR(1) process (3) with $R = 0$ ($R = 1$), $\lambda = 5$, $r = 6$, $(\beta_{1,0}, \beta_{1,1}, \beta_{1,2}) = (0.1, 0.5, 0.3)$, $(\beta_{2,0}, \beta_{2,1}, \beta_{2,2}) = (-0.3, -0.5, -0.6)$. The explanatory variables $Z_{1,t}$ are generalized from an AR(1) process, i.e., $Z_{1,t} = 0.5Z_{1,t-1} + \epsilon_t$ with $Z_{1,0} = 0$, $\epsilon_t \sim N(0, 1)$; $Z_{2,t}$ is generalized from a seasonal series, i.e., $Z_{2,t} = \sin(2\pi t/12) + \epsilon_t$ with $\epsilon_t \sim N(0, 0.25)$.

All simulations were implemented in MATLAB. The sample size considered in all simulations was $n = 200, 500, 1000$. For each model, the value of r was chosen such that the observations in each regime comprised at least 20% of the total sample size. The empirical results displayed in the tables and box plots, that is the empirical biases and mean square errors (MSEs), were computed over 10,000 replications.

4.1. Simulation Study When r Is Known

Tables 1 and 2 report the bias and MSE of the CLS and CML estimators for Models A1–B3 when r is known. It is easy to see that all the simulation results performed better as n increased, which implies that the two estimation methods can lead to good and consistent estimators when r is known. It is worth mentioning that, although it is not mentioned in the main conclusion, the simulation showed that the CML estimators are consistent. In addition, $\hat{\theta}_{CML}$ has smaller bias and MSE, which means that $\hat{\theta}_{CML}$ is better than $\hat{\theta}_{CLS}$.

For the sake of intuition, we also give the box plots and QQ plots of the CLS and CML estimators. Figure 1 plots the bias of 10,000 CLS and CML simulation estimators for Models A1 and B1. Note that the box plots are symmetric and centered on zero-bias; both the bias and MSE for the CML estimators are smaller than the CLS estimators, which is consistent with the previous conclusions. Figure 2 shows the QQ plots of the CLS and CML estimators for Models A1 and B1 with the sample size $n = 1000$. It is easy to see that the CLS and CML estimators are asymptotically normal for all parameters, especially for the CML estimator without the asymptotically normal theorem. Similar results were obtained for the remaining models, and the figures are omitted here to save space.

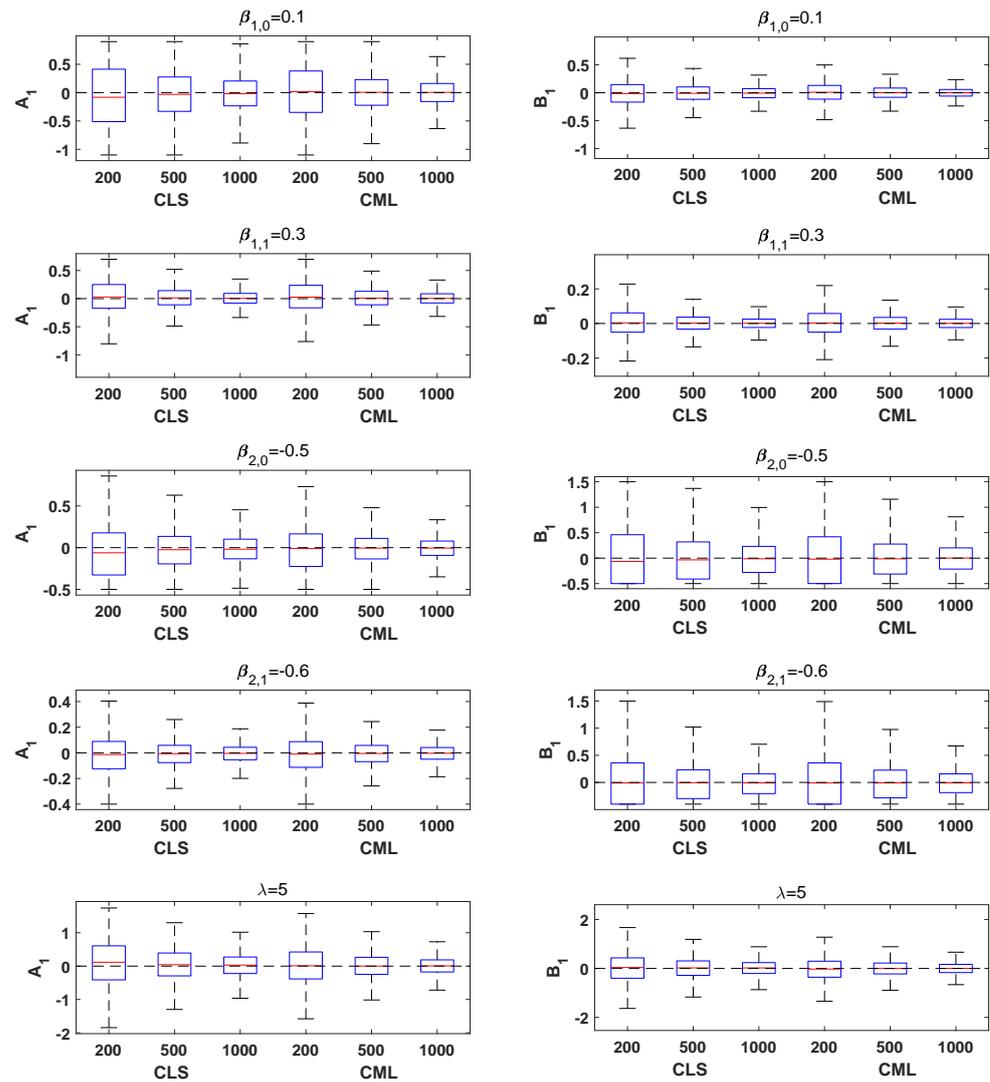


Figure 1. Box plots from 10,000 CLS and CML simulation estimators for Models A1 and B1, with the sample size $n = 200, 500, 1000$.

Table 1. Simulation results for Models A1–B2 when r is known.

Sample Size	Para.	A1				A2			
		CLS		CML		CLS		CML	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
$n = 200$	$\beta_{1,0}$	−0.0544	0.3579	0.0076	0.2669	−0.0856	0.3338	0.0071	0.2476
	$\beta_{1,1}$	0.0430	0.1017	0.0398	0.0955	0.0261	0.0915	0.0273	0.0882
	$\beta_{2,0}$	−0.0607	0.0991	−0.0302	0.0731	−0.0837	0.0931	−0.0419	0.0671
	$\beta_{2,1}$	−0.0216	0.0255	−0.0164	0.0229	−0.0186	0.0216	−0.0160	0.0189
	λ	0.0856	0.4181	0.0245	0.3047	0.1343	0.3908	0.0438	0.2752
$n = 500$	$\beta_{1,0}$	−0.0245	0.1947	0.0020	0.1150	−0.0472	0.1886	0.0032	0.1149
	$\beta_{1,1}$	0.0209	0.0374	0.0156	0.0337	0.0158	0.0323	0.0141	0.0297
	$\beta_{2,0}$	−0.0324	0.0547	−0.0156	0.0344	−0.0464	0.0515	−0.0194	0.0313
	$\beta_{2,1}$	−0.0114	0.0105	−0.0077	0.0090	−0.0098	0.0085	−0.0077	0.0068
	λ	0.0449	0.2357	0.0141	0.1417	0.0778	0.2224	0.0214	0.1307
$n = 1000$	$\beta_{1,0}$	−0.0157	0.1056	−0.0014	0.0564	−0.0190	0.1021	0.0046	0.0534
	$\beta_{1,1}$	0.0094	0.0169	0.0067	0.0154	0.0084	0.0152	0.0063	0.0135
	$\beta_{2,0}$	−0.0197	0.0300	−0.0101	0.0169	−0.0226	0.0283	−0.0075	0.0154
	$\beta_{2,1}$	−0.0070	0.0053	−0.0048	0.0045	−0.0058	0.0044	−0.0037	0.0035
	λ	0.0279	0.1324	0.0108	0.0717	0.0369	0.1239	0.0073	0.0645
Sample Size	Para.	B1				B2			
		CLS		CML		CLS		CML	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
$n = 200$	$\beta_{1,0}$	−0.0084	0.0553	0.0089	0.0338	−0.0304	0.0466	0.0031	0.0271
	$\beta_{1,1}$	0.0057	0.0070	0.0045	0.0066	0.0052	0.0039	0.0041	0.0036
	$\beta_{2,0}$	0.0718	0.3459	0.0684	0.2918	0.1148	0.4075	0.1362	0.3766
	$\beta_{2,1}$	0.0583	0.2060	0.0587	0.2037	0.0767	0.2385	0.0745	0.2307
	λ	0.0100	0.3801	−0.0315	0.2366	0.0794	0.3770	−0.0158	0.2228
$n = 500$	$\beta_{1,0}$	−0.0031	0.0257	0.0013	0.0147	−0.0135	0.0225	−0.0006	0.0121
	$\beta_{1,1}$	0.0024	0.0026	0.0018	0.0025	0.0026	0.0018	0.0019	0.0017
	$\beta_{2,0}$	0.0155	0.1914	0.0101	0.1454	0.0323	0.2214	0.0393	0.1878
	$\beta_{2,1}$	−0.0002	0.0982	0.0005	0.0951	0.0209	0.1220	0.0198	0.1179
	λ	0.0050	0.1838	−0.0043	0.1083	0.0370	0.1855	0.0003	0.1014
$n = 1000$	$\beta_{1,0}$	−0.0053	0.0140	−0.0007	0.0077	−0.0100	0.0110	−0.0032	0.0058
	$\beta_{1,1}$	0.0012	0.0013	0.0007	0.0012	0.0014	0.0008	0.0010	0.0008
	$\beta_{2,0}$	−0.0084	0.1157	−0.0018	0.0860	−0.0045	0.1376	0.0055	0.1119
	$\beta_{2,1}$	−0.0163	0.0590	−0.0135	0.0567	−0.0127	0.0662	−0.0114	0.0639
	λ	0.0121	0.1020	0.0007	0.0578	0.0278	0.0934	0.0083	0.0504

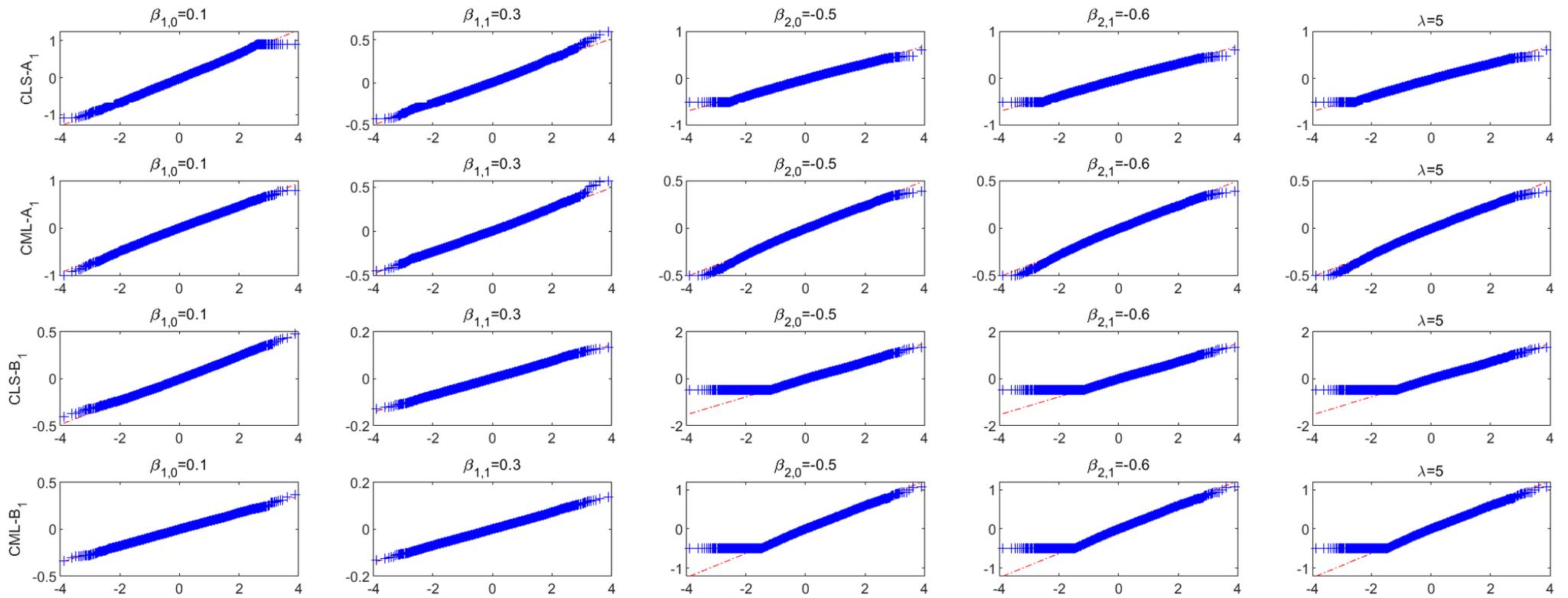


Figure 2. QQ plots of CLS and CML estimators for Models A1 and B1, with the sample size $n = 200$.

Table 2. Simulation results for Models A3–B3 when r is known.

Sample Size	Para.	A3				B3			
		CLS		CML		CLS		CML	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
$n = 200$	$\beta_{1,0}$	−0.0905	0.4153	0.0203	0.3111	−0.0285	0.0593	0.0005	0.0371
	$\beta_{1,1}$	0.0539	0.1056	0.0588	0.1019	0.0088	0.0070	0.0075	0.0065
	$\beta_{1,2}$	−0.0004	0.2013	0.0026	0.1929	0.0071	0.0126	0.0041	0.0118
	$\beta_{2,0}$	−0.0837	0.1029	−0.0360	0.0698	0.0269	0.3911	0.0432	0.3349
	$\beta_{2,1}$	−0.0179	0.0196	−0.0150	0.0172	−0.0159	0.1750	−0.0101	0.1707
	$\beta_{2,2}$	−0.0283	0.0371	−0.0203	0.0329	0.1030	0.2927	0.1025	0.2824
	λ	0.1419	0.4508	0.0369	0.3043	0.0598	0.3590	−0.0090	0.2323
$n = 500$	$\beta_{1,0}$	−0.0467	0.2223	0.0027	0.1373	−0.0137	0.0258	−0.0014	0.0149
	$\beta_{1,1}$	0.0394	0.0423	0.0351	0.0392	0.0043	0.0025	0.0040	0.0023
	$\beta_{1,2}$	0.0183	0.0768	0.0148	0.0708	0.0040	0.0052	0.0028	0.0050
	$\beta_{2,0}$	−0.0398	0.0490	−0.0166	0.0300	−0.0221	0.2194	−0.0052	0.1678
	$\beta_{2,1}$	−0.0082	0.0069	−0.0066	0.0059	−0.0422	0.0848	−0.0352	0.0803
	$\beta_{2,2}$	−0.0118	0.0142	−0.0068	0.0121	0.0182	0.1339	0.0204	0.1276
	λ	0.0716	0.2330	0.0216	0.1395	0.0312	0.1635	0.0011	0.0974
$n = 1000$	$\beta_{1,0}$	−0.0152	0.1161	0.0076	0.0638	−0.0089	0.0140	−0.0018	0.0077
	$\beta_{1,1}$	0.0222	0.0216	0.0172	0.0193	0.0016	0.0013	0.0013	0.0012
	$\beta_{1,2}$	0.0101	0.0356	0.0070	0.0329	0.0019	0.0025	0.0013	0.0023
	$\beta_{2,0}$	−0.0174	0.0252	−0.0053	0.0144	−0.0197	0.1215	−0.0047	0.0866
	$\beta_{2,1}$	−0.0037	0.0036	−0.0028	0.0032	−0.0282	0.0507	−0.0225	0.0470
	$\beta_{2,2}$	−0.0062	0.0071	−0.0028	0.0061	−0.0064	0.0781	−0.0025	0.0748
	λ	0.0305	0.1260	0.0049	0.0684	0.0202	0.0922	0.0025	0.0517

4.2. Simulation Study When r Is Unknown

Tables 3 and 4 report the performance of the proposed CLS and CML estimators in Section 3.3 for Models A1–B3 when r is known. It is easy to draw the following conclusions from the tabular results. For small sample sizes (such as $n = 200$), the bias and MSE of the estimator are still relatively large. However, with the increase of the sample size, this deviation decreases very quickly, mainly because the accuracy of threshold estimation is greatly improved with the increase of the sample size. Moreover, the CML estimator demonstrates a noticeable advantage over the CLS estimator. Nevertheless, this does not imply that CLS estimators lack any advantages. Table 5 reports the percentage of correctly identifying r (Frequency) and average time (s) across 10,000 replications. Without the closed-form solutions of the two methods, the CLS estimation method is still very advantageous in the calculation speed.

Table 3. Simulation results for Models A1–B2 when r is unknown.

Sample Size	Para.	A1				A2			
		CLS		CML		CLS		CML	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
$n = 200$	$\beta_{1,0}$	−0.1673	0.4649	−0.0095	0.3159	−0.2118	0.4925	−0.0109	0.2932
	$\beta_{1,1}$	−0.0600	0.2097	0.0236	0.1342	−0.0609	0.1886	0.0198	0.1131
	$\beta_{2,0}$	−0.0811	0.0973	−0.0319	0.0738	−0.1050	0.0948	−0.0450	0.0680
	$\beta_{2,1}$	−0.0372	0.0451	−0.0259	0.0283	−0.0250	0.0350	−0.0222	0.0212
	λ	0.1526	0.4610	0.0338	0.3177	0.2131	0.4566	0.0547	0.2883
	r	1.0142	5.3096	0.2577	1.2313	0.8968	5.2976	0.1642	0.8174
$n = 500$	$\beta_{1,0}$	−0.0566	0.2308	−0.0009	0.1222	−0.0682	0.2172	0.0014	0.1180
	$\beta_{1,1}$	0.0054	0.0536	0.0157	0.0367	0.0061	0.0416	0.0140	0.0307
	$\beta_{2,0}$	−0.0408	0.0566	−0.0166	0.0350	−0.0517	0.0534	−0.0199	0.0316
	$\beta_{2,1}$	−0.0187	0.0132	−0.0093	0.0093	−0.0123	0.0096	−0.0082	0.0069
	λ	0.0703	0.2566	0.0166	0.1451	0.0935	0.2394	0.0226	0.1321
	r	0.1719	0.8427	0.0181	0.0705	0.1054	0.6042	0.0067	0.0281

Table 3. Cont.

		B1				B2			
		CLS		CML		CLS		CML	
Sample Size	Para.	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
<i>n</i> = 1000	$\beta_{1,0}$	−0.0188	0.1093	−0.0015	0.0565	−0.0191	0.1023	0.0046	0.0534
	$\beta_{1,1}$	0.0080	0.0181	0.0067	0.0155	0.0083	0.0152	0.0064	0.0135
	$\beta_{2,0}$	−0.0207	0.0304	−0.0101	0.0169	−0.0226	0.0284	−0.0075	0.0154
	$\beta_{2,1}$	−0.0078	0.0055	−0.0048	0.0045	−0.0058	0.0044	−0.0037	0.0035
	λ	0.0304	0.1350	0.0109	0.0718	0.0370	0.1240	0.0073	0.0645
	<i>r</i>	0.0121	0.0411	0.0009	0.0059	0.0003	0.0003	−0.0001	0.0001
<i>n</i> = 200	$\beta_{1,0}$	0.0394	0.0738	0.0146	0.0382	0.0001	0.0673	−0.0008	0.0333
	$\beta_{1,1}$	0.0164	0.0232	0.0078	0.0074	0.0200	0.0155	0.0066	0.0040
	$\beta_{2,0}$	0.3713	0.6543	0.1386	0.3408	0.5624	0.8281	0.3572	0.4869
	$\beta_{2,1}$	0.3190	0.4057	0.1768	0.2646	0.5179	0.5136	0.3818	0.3591
	λ	−0.2125	0.7856	−0.0367	0.3023	−0.1115	0.8718	0.0031	0.3169
	<i>r</i>	2.5127	15.9965	0.4938	1.2152	3.2435	25.3375	0.1035	1.0955
<i>n</i> = 500	$\beta_{1,0}$	0.0188	0.0330	0.0021	0.0152	0.0361	0.0372	−0.0037	0.0137
	$\beta_{1,1}$	0.0077	0.0048	0.0025	0.0025	0.0114	0.0055	0.0025	0.0017
	$\beta_{2,0}$	0.1333	0.3180	0.0325	0.1566	0.5518	0.6754	0.3191	0.2900
	$\beta_{2,1}$	0.0926	0.1770	0.0390	0.1133	0.4852	0.4148	0.3708	0.2623
	λ	−0.0858	0.3009	−0.0053	0.1144	−0.1831	0.4647	0.0112	0.1245
	<i>r</i>	0.8616	5.2522	0.1170	0.1640	2.3496	18.1162	−0.1513	0.3999
<i>n</i> = 1000	$\beta_{1,0}$	−0.0019	0.0150	−0.0007	0.0077	0.0423	0.0197	−0.0004	0.0063
	$\beta_{1,1}$	0.0020	0.0015	0.0008	0.0012	0.0072	0.0018	0.0013	0.0008
	$\beta_{2,0}$	0.0105	0.1346	0.0023	0.0884	0.5346	0.5129	0.3481	0.2311
	$\beta_{2,1}$	0.0005	0.0723	−0.0054	0.0608	0.4726	0.3270	0.3973	0.2285
	λ	−0.0011	0.1162	0.0008	0.0582	−0.1783	0.2391	−0.0017	0.0570
	<i>r</i>	0.1303	0.7299	0.0207	0.0231	1.2562	9.8408	−0.1482	0.2074

Table 4. Simulation results for Models A3–B3 when *r* is unknown.

		A3				B3			
		CLS		CML		CLS		CML	
Sample Size	Para.	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
<i>n</i> = 200	$\beta_{1,0}$	−0.2088	0.4902	0.0025	0.3416	−0.0076	0.0699	0.0050	0.0394
	$\beta_{1,1}$	−0.1031	0.2460	0.0372	0.1274	0.0179	0.0152	0.0107	0.0069
	$\beta_{1,2}$	−0.1352	0.3164	−0.0169	0.2225	0.0144	0.0274	0.0055	0.0123
	$\beta_{2,0}$	−0.1069	0.1053	−0.0389	0.0711	0.1478	0.4684	0.0648	0.3516
	$\beta_{2,1}$	−0.0330	0.0369	−0.0219	0.0196	0.1450	0.3211	0.0455	0.2047
	$\beta_{2,2}$	−0.0315	0.0598	−0.0251	0.0365	0.2349	0.3862	0.1481	0.3138
	λ	0.2373	0.5380	0.0512	0.3236	−0.0311	0.5100	−0.0161	0.2623
	<i>r</i>	1.2538	7.7742	0.1827	0.8443	1.3943	9.2127	0.2116	0.5360
<i>n</i> = 500	$\beta_{1,0}$	−0.0612	0.2341	0.0012	0.1395	−0.0109	0.0266	−0.0009	0.0150
	$\beta_{1,1}$	0.0284	0.0513	0.0352	0.0402	0.0052	0.0027	0.0042	0.0023
	$\beta_{1,2}$	0.0084	0.0855	0.0148	0.0721	0.0045	0.0057	0.0029	0.0050
	$\beta_{2,0}$	−0.0433	0.0499	−0.0171	0.0302	−0.0082	0.2310	−0.0029	0.1692
	$\beta_{2,1}$	−0.0106	0.0078	−0.0069	0.0059	−0.0301	0.0983	−0.0304	0.0834
	$\beta_{2,2}$	−0.0137	0.0151	−0.0071	0.0122	0.0305	0.1447	0.0251	0.1307
	λ	0.0837	0.2443	0.0228	0.1406	0.0206	0.1745	0.0001	0.0982
	<i>r</i>	0.0853	0.4439	0.0044	0.0140	0.1002	0.5810	0.0163	0.0277
<i>n</i> = 1000	$\beta_{1,0}$	−0.0161	0.1170	0.0075	0.0639	−0.0088	0.0141	−0.0018	0.0077
	$\beta_{1,1}$	0.0216	0.0218	0.0172	0.0193	0.0017	0.0013	0.0013	0.0012
	$\beta_{1,2}$	0.0095	0.0359	0.0070	0.0330	0.0020	0.0025	0.0013	0.0023
	$\beta_{2,0}$	−0.0177	0.0253	−0.0053	0.0144	−0.0190	0.1224	−0.0045	0.0867
	$\beta_{2,1}$	−0.0039	0.0036	−0.0028	0.0032	−0.0280	0.0513	−0.0221	0.0473
	$\beta_{2,2}$	−0.0063	0.0071	−0.0028	0.0061	−0.0060	0.0785	−0.0020	0.0750
	λ	0.0313	0.1269	0.0049	0.0685	0.0198	0.0927	0.0024	0.0517
	<i>r</i>	0.0031	0.0103	0.0000	0.0006	0.0036	0.0228	0.0014	0.0020

Table 5. The performances of $\hat{\rho}$ for Models A1–B3.

Model	Sample Size	CLS		CML	
		Frequency	Average Time (s)	Frequency	Average Time (s)
A1	200	0.6330	0.5269	0.7841	2.3501
	500	0.9159	0.6343	0.9694	4.8531
	1000	0.9913	0.7666	0.9989	9.0141
A2	200	0.7164	0.4614	0.8564	2.2148
	500	0.9619	0.6495	0.9885	5.2616
	1000	0.9997	0.7582	0.9999	9.3752
A3	200	0.6700	0.6728	0.8479	3.8789
	500	0.9619	0.9298	0.9880	9.2877
	1000	0.9974	1.0712	0.9994	16.3961

Model	Sample Size	CLS		CML	
		Frequency	Average Time (s)	Frequency	Average Time (s)
B1	200	0.4341	0.3850	0.6511	2.0574
	500	0.7767	0.4581	0.8909	4.6666
	1000	0.9574	0.5514	0.9793	9.2278
B2	200	0.1633	0.4005	0.4463	2.4881
	500	0.2720	0.5062	0.6398	5.6456
	1000	0.5250	0.6105	0.7944	11.5242
B3	200	0.6513	0.4865	0.8069	3.1971
	500	0.9540	0.6448	0.9776	7.5834
	1000	0.9965	0.8010	0.9980	14.9456

4.3. Empirical Sizes and Powers of the Wald Test

Some simulations were conducted to investigate the performances of the Wald test $T_n^{(1)}$. We selected the significance level as $\alpha = 0.05$ (the associated critical value was 1.65). For analyzing the empirical size, we first introduced the following two time-varying integer-valued autoregressive models (TVINAR).

$$TVINAR(1)\text{-B: } X_t = \frac{\exp(Z_t^\top \beta_1)}{1 + \exp(Z_t^\top \beta_1)} \circ X_{t-1} + \epsilon_t, \tag{13}$$

$$TVINAR(1)\text{-G: } X_t = \frac{\exp(Z_t^\top \beta_2)}{1 + \exp(Z_t^\top \beta_1)} * X_{t-1} + \epsilon_t. \tag{14}$$

where the explanatory variables $Z_{1,t}$ are generalized from an AR(1) process, i.e., $Z_{1,t} = -0.5 * Z_{1,t-1} + \epsilon_t$ with $Z_{1,0} = 0$, $\epsilon_t \sim N(0, 1)$; $Z_{2,t}$ is generalized from a seasonal series, i.e., $Z_{2,t} = \sin(2\pi t/12) + \epsilon_t$ with $\epsilon_t \sim N(0, 0.25)$.

For analyzing the empirical size, Models T₁₁–T₂₂ were considered. For analyzing the empirical power, Models T₃₁–T₃₂ were considered:

Model T₁₁: Generated from the TVINAR(1)-B process (13) with $(\beta_1^\top, \lambda) = (\beta_{1,0}, \beta_{1,1}, \beta_{1,2}, \lambda) = (0.1, 0.5, 0.3, 5)$.

Model T₁₂: Generated from the TVINAR(1)-B process (13) with $(\beta_1^\top, \lambda) = (\beta_{1,0}, \beta_{1,1}, \beta_{1,2}, \lambda) = (-0.7, -0.8, 0.6, 2)$.

Model T₂₁: Generated from the TVINAR(1)-G process (14) with $(\beta_2^\top, \lambda) = (\beta_{2,0}, \beta_{2,1}, \beta_{2,2}, \lambda) = (-0.3, -0.3, -0.6, 2)$.

Model T₂₂: Generated from the TVINAR(1)-G process (14) with $(\beta_2^\top, \lambda) = (\beta_{2,0}, \beta_{2,1}, \beta_{2,2}, \lambda) = (-0.6, 0.8, -0.5, 7)$.

Model T₃₁: Generated from the TVMTINAR(1) process (3) with $R = 0$, $\lambda = 5$, $r = 6$, $\beta_1^\top = (\beta_{1,0}, \beta_{1,1}, \beta_{1,2}) = (0.1, 0.5, 0.3)$, $\beta_2^\top = (\beta_{2,0}, \beta_{2,1}, \beta_{2,2}) = (-0.3, -0.3, -0.6)$. The explanatory variables $Z_{1,t}$ are generalized from an AR(1) process, i.e., $Z_{1,t} = -0.5 * Z_{1,t-1} + \epsilon_t$ with $Z_{1,0} = 0$, $\epsilon_t \sim N(0, 1)$; $Z_{2,t}$ is generalized from a seasonal series, i.e., $Z_{2,t} = \sin(2\pi t/12) + \epsilon_t$ with $\epsilon_t \sim N(0, 0.25)$.

Model T₃₂: Generated from the TVMTINAR(1) process (3) with $R = 1$, $\lambda = 3$, $r = 6$, $\beta_1^\top = (\beta_{1,0}, \beta_{1,1}, \beta_{1,2}) = (-0.4, -0.8, 0.3)$, $\beta_2^\top = (\beta_{1,0}, \beta_{1,1}, \beta_{1,2}) = (-0.3, 0.7, -0.4)$.

The explanatory variables $Z_{1,t}$ are generalized from an AR(1) process, i.e., $Z_{1,t} = -0.5 * Z_{1,t-1} + \epsilon_t$ with $Z_{1,0} = 0, \epsilon_t \sim N(0, 1)$; $Z_{2,t}$ is generalized from a seasonal series, i.e., $Z_{2,t} = \sin(2\pi t/12) + \epsilon_t$ with $\epsilon_t \sim N(0, 0.25)$.

The results are reported in Table 6. As can be seen from Table 6, for the empirical sizes, the Wald test $T_n^{(1)}$ gave satisfactory performances and the empirical sizes for Models T₁₁–T₂₂ were closer to the significant level of $\alpha = 0.05$ as the sample size increased. For the empirical powers, Table 6 also indicates that the $T_n^{(1)}$ succeeded in showing high values in almost each case. The above discussion shows the success of the proposed Wald test $T_n^{(1)}$ to detect the existence of the piecewise structure.

Table 6. Empirical sizes and powers of $T_n^{(1)}$ at level 0.05.

Empirical Sizes, Significance Level $\alpha = 0.05$									
Model	Method	Sample Size			Model	Method	Sample Size		
		$n = 200$	$n = 500$	$n = 1000$			$n = 200$	$n = 500$	$n = 1000$
T ₁₁	$T_n^{(1)}$	0.0497	0.0515	0.0517	T ₁₂	$T_n^{(1)}$	0.0255	0.0517	0.0559
T ₂₁	$T_n^{(1)}$	0.0123	0.0448	0.0527	T ₂₂	$T_n^{(1)}$	0.0012	0.0127	0.0501

Empirical Powers, Significance Level $\alpha = 0.05$									
Model	Method	Sample Size			Model	Method	Sample Size		
		$n = 200$	$n = 500$	$n = 1000$			$n = 200$	$n = 500$	$n = 1000$
T ₃₁	$T_n^{(1)}$	0.8414	0.9507	0.9972	T ₃₂	$T_n^{(1)}$	0.5016	0.7979	0.967

4.4. Empirical Sizes and Powers of the Proposed Test in Section 3.5

Similarly, we investigated the performances of the test $T_n^{(2)}$ and the LR test $T_n^{(3)}$ with the following models:

Model T₄₁: Generated from the MTTINAR(1) process (11) with $(\alpha_1, \alpha_2, \lambda, r, R) = (0.6225, 0.4502, 5, 6, 0)$.

Model T₄₂: Generated from the MTTINAR(1) process (11) with $(\alpha_1, \alpha_2, \lambda, r, R) = (0.4502, 0.6225, 3, 4, 1)$.

Model T₅₁: Generated from the TVMTTINAR(1) process (3) with $R = 0, \lambda = 5, r = 6, \beta_1^\top = (\beta_{1,0}, \beta_{1,1}) = (0.5, 0.3), \beta_2^\top = (\beta_{2,0}, \beta_{2,1}) = (-0.2, -0.2)$. The explanatory variables $Z_{1,t}$ are generalized from the i.i.d. uniform distribution $U(-10, 1)$.

Model T₅₂: Generated from the TVMTTINAR(1) process (3) with $R = 1, \lambda = 5, r = 6, \beta_1^\top = (\beta_{1,0}, \beta_{1,1}) = (0.5, 0.3), (\beta_2^\top) = (\beta_{2,0}, \beta_{2,1}) = (-0.2, -0.2)$. The explanatory variables $Z_{1,t}$ are generalized from the i.i.d. uniform distribution $U(-10, 1)$.

Models T₄₁ and T₄₂ were used to analyze the empirical sizes; Models T₄₁ and T₄₂ were applied to analyze the empirical powers. We selected the significance level of $\alpha = 0.05$ (since $q = 1$, the associated critical value was 5.991). Table 7 shows the empirical sizes and powers result. It is easy to see that, for the empirical sizes, both tests gave satisfactory performances and the empirical sizes for Models T₄₁–T₄₂ were closer to the significant level of $\alpha = 0.05$ as the sample size increased. For the empirical power, both proposed tests were increasingly close to 1 as the sample size increased in each case. In addition, although both tests can successfully detect the existence of the explanatory variables, the LR test performed significantly better.

Table 7. Empirical sizes and powers of $T_n^{(2)}$ and $T_n^{(3)}$ at level 0.05.

Empirical Sizes, Significance Level $\alpha = 0.05$									
Model	Method	$n = 200$	Sample Size		Model	Method	$n = 200$	Sample Size	
			$n = 500$	$n = 1000$				$n = 500$	$n = 1000$
T ₄₁	$T_n^{(2)}$	0.0580	0.0466	0.0475	T ₄₂	$T_n^{(2)}$	0.0320	0.0413	0.0445
	$T_n^{(3)}$	0.0504	0.0502	0.0495		$T_n^{(3)}$	0.0456	0.0487	0.0509
Empirical Power, Significance Level $\alpha = 0.05$									
Model	Method	$n = 200$	Sample Size		Model	Method	$n = 200$	Sample Size	
			$n = 500$	$n = 1000$				$n = 500$	$n = 1000$
T ₅₁	$T_n^{(2)}$	0.9987	1.0000	1.0000	T ₅₂	$T_n^{(2)}$	0.9987	1.0000	1.0000
	$T_n^{(3)}$	1.0000	1.0000	1.0000		$T_n^{(3)}$	1.0000	1.0000	1.0000

5. Real Data Example

In this section, we will utilize the TVMTTINAR(1) model to match the daily stock trading volume dataset of an automotive company, Volkswagen Corporation (VOW). There are some explanations for the selection of the factors affecting the trading volume of stocks in the automotive industry. The volume of stock trading in the automotive industry can be affected by a number of factors, the most well-known of which are the state of the economy and the state of the oil market. After all, the state of the economy largely determines consumers’ ability and willingness to buy. On the one hand, the fluctuation of oil prices increases the production cost, and on the other hand, it affects the willingness of consumers to buy traditional fuel vehicles and directly or indirectly affects the automobile industry.

Economic conditions can be represented by some stock market indices. We selected the Dow Jones Industrial average indices (DJI) here. The DJI can reflect the overall performance of the stock market and is also used as an indicator of the health of the economy. Therefore, it is reasonable to choose the DJI’s stock data series as economic indicators. On the other hand, oil prices vary widely between countries and regions, making it difficult to find a uniform measure, the Crude oil (Co) stock data series was selected here to represent the oil market.

All datasets were originally downloaded online from the Yahoo finance web site (<https://hk.finance.yahoo.com/>, accessed on 7 December 2010). Both the DJI and Co stock data series include open, high, low, close, and adjusted (Adj) close prices, wherein the Adj close price datasets were selected as the explanatory variables for the analysis. In addition, the two explanatory variable datasets need to be differentiated to reflect fluctuations in the economy and the oil market.

5.1. Volkswagen Corporation Daily Stock Trading Volume Data

We first considered the VOW daily stock trading volume dataset, which consist of 281 observations starting in 7 December 2010 and ending in 11 January 2012. As the data are relatively large, considering the convenience of calculation, we analyzed the data by the unit of a “ 2×10^5 ” trading volume. Figure 3 shows the sample path and the sample autocorrelation (ACF) of the observations, where the first line shows the sample path and ACF of the VOW daily stock trading volume dataset and the second line shows the sample path after the difference in the DJI and Co’s Adj close prices.

Next, we used the TVMTTINAR(1) model and the following integer-valued threshold autoregressive models to fit the VOW corporation dataset and compare different models via the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC):

- SETINAR(2,1) model [12].
- NBTINAR(1) model [15].
- RCTINAR(1) model [11].
- BiNB-MTTINAR(1) ($R = 0$) model [17].
- BiNB-MTTINAR(1) ($R = 1$) model [17].

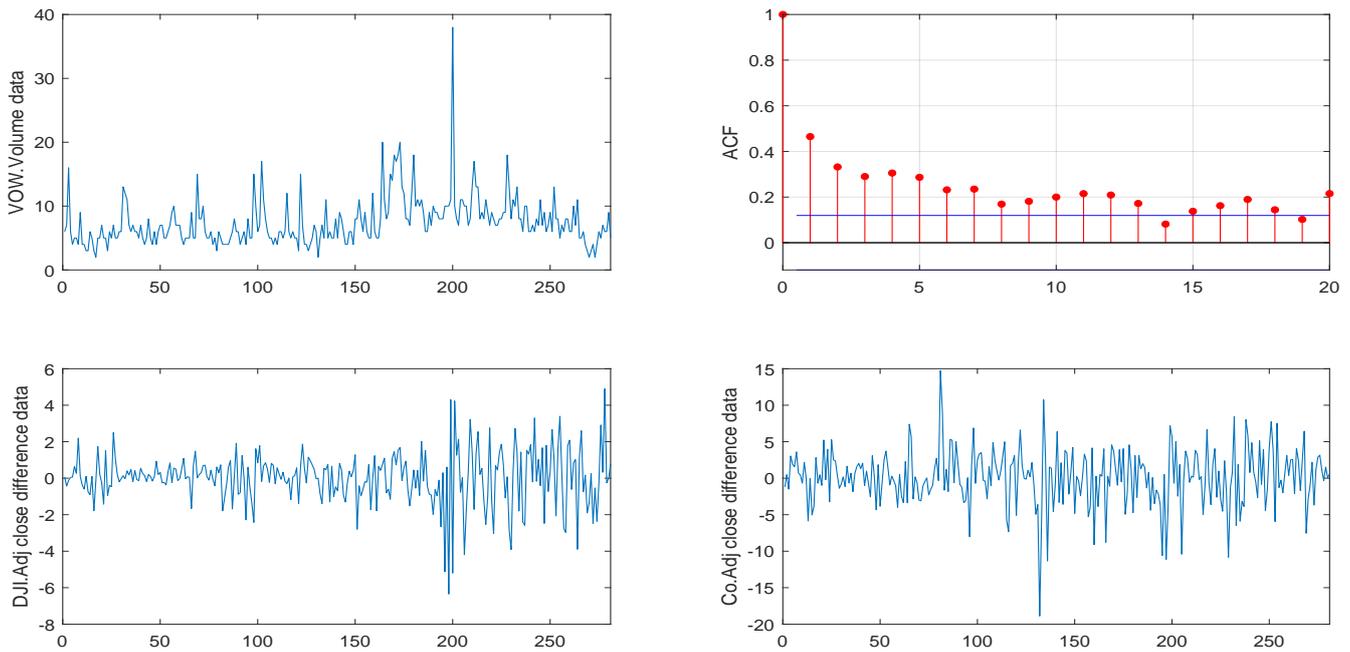


Figure 3. Sample path and ACF of the VOW daily stock trading volume dataset (from 7 December 2010 and ending in 11 January 2012).

For each of the above models, we estimated the CML of the parameters and the threshold r with the range $r \in \{4, 5, \dots, 12\}$, where 4 and 12 are the 10th and 90th quantiles of the data. Furthermore, the standard error (SE) of $\hat{\theta}_{CML}$, the root mean square of the differences between the observations and forecasts (RMS), and the AIC and BIC values are also given. Among them, the standard error for the CML estimator can be obtained as the square roots of the elements in the diagonal of the inverse of the negative Hessian of the log-likelihood calculated at the CML estimates. The RMS is defined as follows:

$$RMS = \sqrt{\frac{1}{n-1} \sum_{t=2}^n \left(X_t - \sum_{i=1}^2 \frac{\exp(\mathbf{Z}_t^\top \beta_i)}{1 + \exp(\mathbf{Z}_t^\top \beta_i)} X_{t-1} I_{i,t}^R - \lambda \right)^2}.$$

The fitting results are summarized in Table 8. As seen from the results presented in Table 8, the proposed TVMTTINAR(1) ($R = 0$) model outperformed the other SETINAR models when considering the AIC as an information criterion. However, due to the excessive number of parameters, when considering the BIC as an information criterion, the model appeared slightly less favorable. Additionally, the TVMTTINAR(1) ($R = 0$) model had the lowest RMS value. Taking all factors into account, the TVMTTINAR(1) ($R = 0$) model was highly competitive, making it reasonable to apply it for fitting this VOW dataset.

Then, we computed the (standardized) Pearson residuals $Pr_t(\hat{\theta})$ to check if the fit model was adequate for the data.

$$Pr_t(\hat{\theta}) = \frac{X_t - \sum_{i=1}^2 \frac{\exp(\mathbf{Z}_t^\top \beta_i)}{1 + \exp(\mathbf{Z}_t^\top \beta_i)} X_{t-1} I_{i,t}^R - \lambda}{\sqrt{\frac{\exp(\mathbf{Z}_t^\top \beta_1)}{[1 + \exp(\mathbf{Z}_t^\top \beta_1)]^2} X_{t-1} I_{1,t}^R + \frac{\exp(\mathbf{Z}_t^\top \beta_2)[1 + 2\exp(\mathbf{Z}_t^\top \beta_2)]}{[1 + \exp(\mathbf{Z}_t^\top \beta_2)]^2} X_{t-1} I_{1,t}^R + \lambda}}. \quad (15)$$

We proceeded to apply the TVMTTINAR(1) ($R = 0$) model to fit this dataset and computed some additional fitting-related information beyond Table 8. These details are summarized in Table 9, encompassing the proportion of samples below the threshold value relative to the total sample size (rate), the test statistic $T_n^{(1)}$ for testing the presence of the segmented

structure, the test statistic $T_n^{(3)}$ for testing the existence of explanatory variables, along with the mean and variance of the Pearson residuals.

Table 8. Fitting results of different models: CML, SE, \hat{r}_{CML} , AIC, BIC, and RMS based on the first VOW dataset.

Model	Para.	CML	SE	\hat{r}_{CML}	AIC	BIC	RMS
SETINAR(2,1)	α_1	0.0169	0.0221	6	1395.9685	1406.8836	3.3266
	α_2	0.3071	0.0092				
	λ	5.9064	0.0032				
NBTINAR(1)	α_1	0.2689	0.0034	6	1369.2936	1380.2087	12.3344
	α_2	0.3371	0.0041				
	v	17.0000	0.0004				
RCTINAR(1)	ϕ_1	0.0000	4.4123	6	1405.0371	1415.9521	3.3656
	ϕ_2	0.2295	0.0034				
	λ	6.3869	0.0006				
BiNB–MTTINAR(1) (R = 0)	ϕ_1	0.4852	0.0076	10	1358.6573	1369.5724	3.3655
	ϕ_2	0.5202	0.0027				
	λ	3.7781	0.0013				
BiNB–MTTINAR(1) (R = 1)	ϕ_1	0.3910	0.0038	4	1430.5887	1441.5038	3.7434
	ϕ_2	0.6003	0.0012				
	λ	4.6621	0.0061				
TVMTTINAR(1) (R = 0)	$\beta_{1,0}$	−0.5391	−0.0006	6	1352.1336	1377.6020	3.1897
	$\beta_{1,1}$	0.0203	0.0001				
	$\beta_{1,2}$	0.0215	0.0000				
	$\beta_{2,0}$	0.0300	0.0001				
	$\beta_{2,1}$	−0.2134	−0.0005				
	$\beta_{2,2}$	0.0573	0.0003				
	λ	4.0801	0.0030				
TVMTTINAR(1) (R = 1)	$\beta_{1,0}$	−0.1093	−0.0002	12	1366.3313	1391.7998	3.2179
	$\beta_{1,1}$	−0.1298	−0.0003				
	$\beta_{1,2}$	0.0793	0.0002				
	$\beta_{2,0}$	−0.0765	−0.0002				
	$\beta_{2,1}$	−0.2039	−0.0007				
	$\beta_{2,2}$	0.0367	0.0002				
	λ	3.8956	0.0030				

Table 9. The other fitting results of the TVMTTINAR(1) model.

Rate	$T_n^{(1)}$	$T_n^{(3)}$	Mean($\text{Pr}_t(\hat{\theta})$)	Var($\text{Pr}_t(\hat{\theta})$)
0.5018	2.2343	35.2307	0.0041	1.1116

After computing, the proportion of estimated values below the threshold was 0.5018, indicating reliable results on both sides of the threshold. $T_n^{(1)} = 2.2343$, surpassing the critical value of 1.65 at a 0.05 significance level, leading us to reject the null hypothesis $H_0^{(1)} : \beta_1 = \beta_2$, confirming the presence of a segmented structure. $T_n^{(3)} = 35.2307$, exceeding the critical value of 9.487 at a 0.05 significance level, compelling us to accept the alternative hypothesis $H_1^{(2)} : \text{At least one } \beta_{i,j} \neq 0, i \in \{1, 2\}, 1 \leq j \leq q$. Although the TVMTTINAR(1) ($R = 0$) model in Table 8 demonstrates superiority in terms of the AIC and RMS, its superiority is not immediately evident. However, it is important to highlight that these test results fully justify the introduction of mixture thinning operators and observable explanatory variables. This provides another level of evidence supporting the suitability of the TVMTTINAR(1) ($R = 0$) model for fitting this dataset, thus highlighting the model’s competitiveness. Additionally, the model’s fit Pearson residuals exhibited a mean of 0.0041 and a variance of 1.1116, indicating a well-balanced fit.

Finally, Figure 4 shows the diagnostic checking plots for our out fit model, including the (a) standardized residuals, (b) Histogram of the standardized residuals, (c) ACF plot of the residuals, and (d) PACF plot of the residuals. From Figure 4, it can be seen that the Pearson residual samples' ACF and PACF had values close to zero, which reveals that our fit model was more suitable.

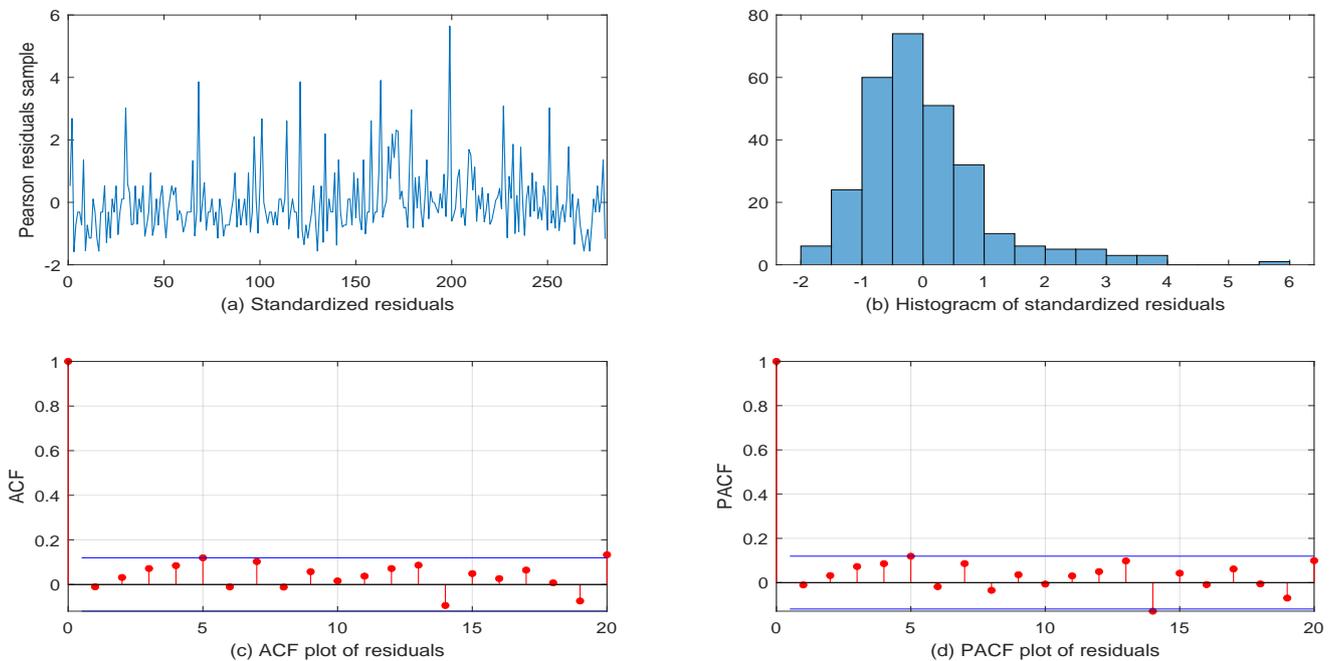


Figure 4. Diagnostic checking plots for the first VOW dataset.

5.2. Another VOW Daily Stock Trading Volume Dataset

Similar to the first data analysis, we considered another dataset of the VOW daily stock trading volume, which also consists of 281 observations starting in 4 May 2010 and ending in 6 June 2011. As the data are relatively large, considering the convenience of calculation, we analyzed the data by the unit of a “ 2×10^5 ” trading volume. Also, Figure 5 shows the sample path and the sample autocorrelation (ACF) of the observations.

Next, we compared the performances of SETINAR(2,1), NBTINAR(1), RCTINAR(1), and BiNB-MTTINAR(1) versus TVMTTINAR(1) for this series. The estimation results are shown in Table 10. Clearly, for this dataset, the TVMTTINAR(1) ($R = 1$) model demonstrated superior performance in terms of the AIC and RMS. Additionally, it is worth noting that, while both sets of data analysis belong to the VOW dataset, particularly with overlapping data (from 7 December 2010 to 6 June 2011), the optimal model selection changed. This indicates the presence of a change point in this period, suggesting a need for further discussion and analysis in subsequent research.

Then, we also summarize the rate, the test statistic $T_n^{(1)}, T_n^{(3)}$, and the mean and variance of the Pearson residuals in Table 11. After computing, the proportion of estimated values below the threshold was 0.5267, indicating reliable results on both sides of the threshold. $T_n^{(1)} = 1.6682$, surpassing the critical value of 1.65 at a 0.05 significance level, leading us to reject the null hypothesis $H_0^{(1)} : \beta_1 = \beta_2$, confirming the presence of a segmented structure. $T_n^{(3)} = 63.9769$, exceeding the critical value of 9.487 at a 0.05 significance level, compelling us to accept the alternative hypothesis $H_1^{(2)} : \text{At least one } \beta_{i,j} \neq 0, i \in \{1, 2\}, 1 \leq j \leq q$. Also, it is worth noting that the TVMTTINAR(1) ($R = 1$) model in Table 10 exhibited some superiority in terms of the AIC and RMS. However, the degree of superiority was not clearly evident. It is important to mention that the aforementioned test results thoroughly justified the need to introduce mixture thinning operators and observable explanatory

variables. These findings provide further evidence that the TVMTTINAR(1) ($R = 1$) model is highly suitable for fitting the dataset, thereby highlighting its competitive nature. Additionally, the model's fit Pearson residuals exhibited a mean of 0.0015 and a variance of 1.0148, indicating a well-balanced fit.

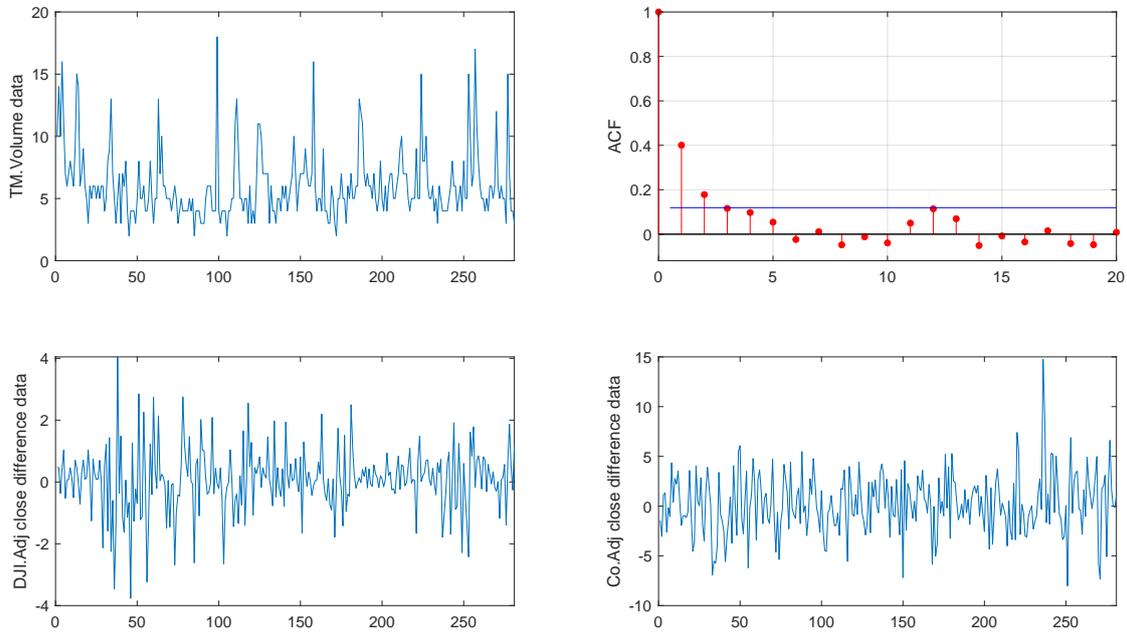


Figure 5. Sample path and ACF of the VOW daily stock trading volume dataset (from 7 December 2010 and ending in 6 June 2011).

Finally, Figure 6 shows the diagnostic checking plots for our out fit model, including the (a) standardized residuals, (b) Histogram of the standardized residuals, (c) ACF plot of the residuals, and (d) PACF plot of the residuals. From Figure 4, it can be seen that the Pearson residual samples' ACF and PACF had values close to zero, which reveals that our fitted model was suitable.

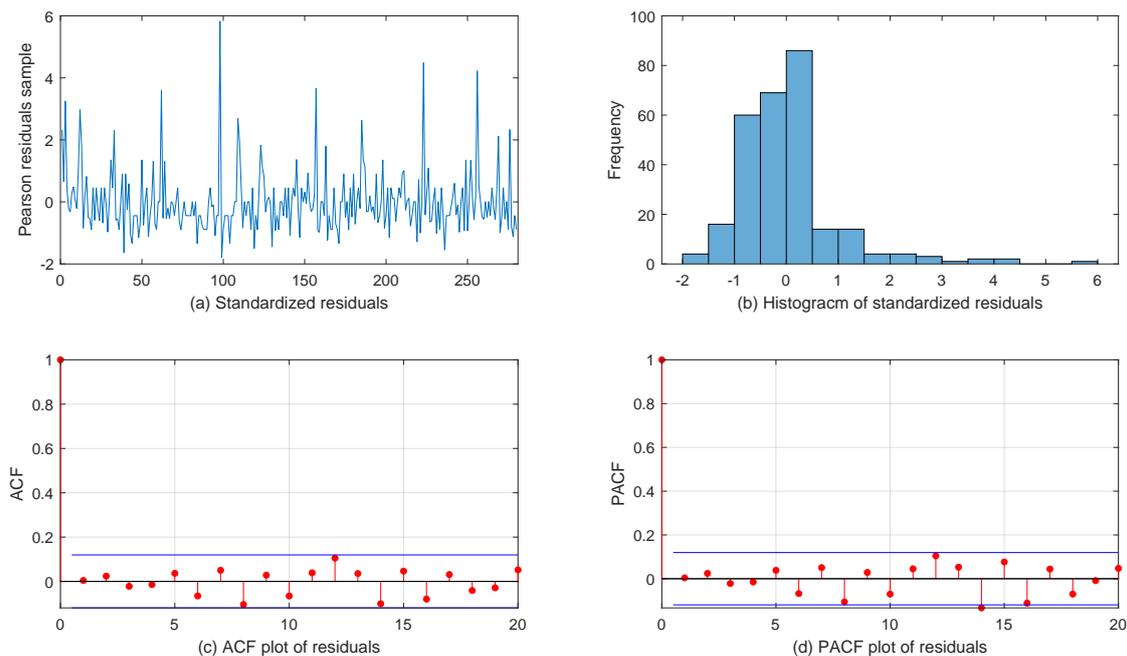


Figure 6. Diagnostic checking plots for another VOW dataset.

Table 10. Fitting results of different models: CML, SE, \hat{r}_{CML} , AIC, BIC, and RMS based on the second VOW dataset.

Model	Para.	CML	SE	\hat{r}_{CML}	AIC	BIC	RMS
SETINAR(2,1)	α_1	0.2470	0.0028	6	1257.6000	1268.5151	2.5162
	α_2	0.3511	0.0068				
	λ	4.2692	0.0053				
NBTINAR(1)	α_1	0.1245	0.0172	6	1259.0943	1270.0094	33.9008
	α_2	0.1605	0.0104				
	v	39.0000	0.0001				
RCTINAR(1)	ϕ_1	0.0228	0.0079	6	1268.9508	1279.8659	2.5308
	ϕ_2	0.2181	0.0017				
	λ	5.4232	0.0007				
BiNB–MTTINAR(1) (R = 0)	ϕ_1	0.4788	0.0073	8	1259.9032	1270.8183	2.5413
	ϕ_2	0.4955	0.0022				
	λ	3.1526	0.0010				
BiNB–MTTINAR(1) (R = 1)	ϕ_1	0.4441	0.0022	4	1295.9314	1306.8464	2.9048
	ϕ_2	0.6682	0.0018				
	λ	3.3134	0.0041				
TVMTTINAR(1) (R = 0)	$\beta_{1,0}$	−0.3448	−0.0003	6	1258.9090	1284.3775	2.5260
	$\beta_{1,1}$	0.0404	0.0000				
	$\beta_{1,2}$	0.0176	0.0000				
	$\beta_{2,0}$	−0.2200	−0.0002				
	$\beta_{2,1}$	0.0163	0.0000				
	$\beta_{2,2}$	−0.0080	0.0000				
	λ	3.4359	0.0063				
TVMTTINAR(1) (R = 1)	$\beta_{1,0}$	−1.1063	−0.0010	5	1254.5145	1279.9829	2.4856
	$\beta_{1,1}$	0.1417	0.0001				
	$\beta_{1,2}$	−0.0427	0.0000				
	$\beta_{2,0}$	−19.9998	−4.3550				
	$\beta_{2,1}$	−0.2643	−0.0002				
	$\beta_{2,2}$	3.3629	0.0030				
	λ	4.9869	0.0072				

Table 11. The other fitting results of the TVMTTINAR(1) model.

Rate	$T_n^{(1)}$	$T_n^{(3)}$	Mean($\text{Pr}_t(\hat{\theta})$)	Var($\text{Pr}_t(\hat{\theta})$)
0.5267	1.6682	63.9769	0.0015	1.0148

6. Conclusions

This article introduces a first-order time-varying coefficient mixture thinning threshold integer-valued autoregressive process. The process was proven to be stationary and ergodic. We investigated the CLS and CML techniques for parameter estimation, and the asymptotic properties of the estimators were demonstrated. Two methods were suggested for estimating the unknown threshold parameter r , based on the CLS and CML score functions. Additionally, we constructed the Wald test statistic to check for the existence of the piecewise structure and constructed two test statistics to test the existence of the explanatory variables. Finally, we successfully applied the TVMTINAR(1) model to Volkswagen Corporation’s daily stock trading volume datasets. From real data studies, potential problems for future research include extending the results to a mixture thinning threshold INAR model with random coefficients and studying a TVMTTINAR(1) model with change points. These will remain the subject of future research.

Author Contributions: Conceptualization, D.S. and J.Z.; methodology, D.S. and X.W.; software, D.S. and Y.Z.; validation, D.S., J.Z. and X.W.; formal analysis, D.S.; data curation, Y.Z.; writing—original draft preparation, D.S., J.Z. and X.W.; writing—review and editing, D.S., J.Z. and X.W.; visualization,

Y.Z.; supervision, D.S.; project administration, D.W.; funding acquisition, D.W., J.Z. and X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (Nos. 12271231, 12101417), the Natural Science Foundation of Jilin Province (No. YDZJ202301ZYTS384), and the Liaoning Provincial Social Science Foundation (No. L22ZD065).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Author Yiran Zhai was employed by the company State Grid Jilin Electric Power Company Limited Information and Telecommunication Company. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A

Proof of Proposition 1. Denote $\phi_{\max,t} = \max\{\phi_{1,t}, \phi_{2,t}\}$ for all t . To simplify the notation, denote $\mathcal{Z}_t = (\mathbf{Z}_0, \dots, \mathbf{Z}_t)$. Compute to see that, under the stationary distribution,

$$\begin{aligned} E(X_t|\mathcal{Z}_t) &= E[(\phi_{1,t} \circ X_{t-1})I_{1,t}^R|\mathcal{Z}_t] + E[(\phi_{2,t} * X_{t-1})I_{2,t}^R|\mathcal{Z}_t] + \lambda \\ &= E[E((\phi_{1,t} \circ X_{t-1})I_{1,t}^R|X_{t-1})|\mathcal{Z}_t] + E[E((\phi_{2,t} * X_{t-1})I_{2,t}^R|X_{t-1})|\mathcal{Z}_t] + \lambda \\ &\leq \phi_{\max,t}E(X_{t-1}|\mathcal{Z}_t) + \lambda \\ &\leq \dots \\ &\leq (\phi_{\max,t})^t E(X_0|\mathcal{Z}_t) + \lambda \sum_{i=0}^{t-1} (\phi_{\max,t})^i \leq \infty, \end{aligned} \tag{A1}$$

Similarly, we have

$$\begin{aligned} E(X_t^2|\mathcal{Z}_t) &= E\left[(\phi_{1,t} \circ X_{t-1})^2 I_{1,t}^R + (\phi_{2,t} * X_{t-1})^2 I_{2,t}^R + \varepsilon_t^2 \right. \\ &\quad \left. + 2(\phi_{1,t} \circ X_{t-1})I_{1,t}^R \varepsilon_t + 2\phi_{2,t} * X_{t-1} I_{2,t}^R \varepsilon_t \middle| \mathcal{Z}_t\right] \\ &\leq [(\phi_{1,t} - \phi_{1,t}^2 + 2\phi_{1,t}\lambda)E(X_{t-1}|\mathcal{Z}_t) + \phi_{1,t}^2 E(X_{t-1}^2|\mathcal{Z}_t) \\ &\quad + (\phi_{2,t} + \phi_{2,t}^2 + 2\phi_{2,t}\lambda)E(X_{t-1}|\mathcal{Z}_t) + \phi_{2,t}^2 E(X_{t-1}^2|\mathcal{Z}_t) + \lambda + \lambda^2 \\ &\leq u_{\max,t}E(X_{t-1}|\mathcal{Z}_t) + \phi_{\max,t}^2 E(X_{t-1}^2|\mathcal{Z}_t) + \lambda + \lambda^2 \end{aligned}$$

where $u_{\max,t} = \max\{(\phi_{1,t} - \phi_{1,t}^2 + 2\phi_{1,t}\lambda), (\phi_{2,t} + \phi_{2,t}^2 + 2\phi_{2,t}\lambda)\}$ for all t . If $t = 1$, $E(X_1^2|\mathcal{Z}_t) \leq u_{\max}E(X_0|\mathcal{Z}_t) + \phi_{\max}^2 E(X_0^2|\mathcal{Z}_t) + \lambda + \lambda^2 < \infty$. Else, if $t \geq 2$,

$$\begin{aligned} E(X_t^2|\mathcal{Z}_t) &\leq \sum_{i=0}^{t-1} u_{\max} \phi_{\max}^{t-1+i} E(X_0|\mathcal{Z}_t) + \lambda \sum_{i=0}^{t-2} u_{\max} \phi_{\max}^{t-2+i} + (\lambda + \lambda^2) \sum_{i=0}^{t-1} \phi_{\max}^{2i} \\ &< \infty. \end{aligned} \tag{A2}$$

A similar, but tedious calculation shows that $E(X_t^3|\mathcal{Z}_t) \leq \infty$ and $E(X_t^4|\mathcal{Z}_t) \leq \infty$. Combining (A1) and (A2), one can see that $E(X_t^k|\mathcal{Z}_t) < \infty$ for $k = 1, 2, 3, 4$. \square

Proof of Section 2. The results $E(X_t|X_{t-1}, \mathbf{Z}_t)$, $E(X_t|\mathbf{Z}_t)$ and $\text{Var}(X_t|X_{t-1}, \mathbf{Z}_t)$ are straightforward to verify. We prove the other results of the moments and conditional moments:

(1) The variance of X_t under the condition \mathbf{Z}_t is given by

$$\begin{aligned} \text{Var}(X_t|\mathbf{Z}_t) &= \text{Var}[I_{1,t}^R(\phi_{1,t} \circ X_{t-1})|\mathbf{Z}_t] + \text{Var}[I_{2,t}^R(\phi_{2,t} * X_{t-1})|\mathbf{Z}_t] \\ &\quad + 2\text{Cov}(I_{1,t}^R(\phi_{1,t} \circ X_{t-1}), I_{2,t}^R(\phi_{2,t} * X_{t-1})|\mathbf{Z}_t) + \lambda \\ &= I + II + III + \lambda. \end{aligned} \tag{A3}$$

A direct calculation shows

$$\begin{aligned}
 I &= \text{Var} \left\{ \mathbb{E} [I_{1,t}^R(\phi_1 \circ X_{t-1}) | X_{t-1}] | \mathbf{Z}_t \right\} + \mathbb{E} \left\{ \text{Var} [I_{1,t}^R(\phi_1 \circ X_{t-1}) | X_{t-1}] | \mathbf{Z}_t \right\} \\
 &= \phi_{1,t}^2 \text{Var}(I_{1,t}^R X_{t-1}) + (\phi_{1,t}(1 - \phi_{1,t}) \mathbb{E}(I_{1,t}^R X_{t-1})) \\
 &= \phi_{1,t}^2 \mathbb{E}(I_{1,t}^R X_{t-1}^2) - \phi_{1,t}^2 p_1^2 \mu_1^2 + p_1 \phi_{1,t}(1 - \phi_{1,t}) \mu_1 \\
 &= \left[\frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_1)}{1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_1)} \right]^2 [p_1(\sigma_1^2 + \mu_1^2) - p_1^2 \mu_1^2] + \frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_1)}{[1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_1)]^2} p_1 \mu_1. \tag{A4}
 \end{aligned}$$

Similarly, we have

$$\begin{aligned}
 II &= \text{Var} \left\{ \mathbb{E} [I_{2,t}^R(\phi_2 * X_{t-1}) | X_{t-1}] | \mathbf{Z}_t \right\} + \mathbb{E} \left\{ \text{Var} [I_{2,t}^R(\phi_2 * X_{t-1}) | X_{t-1}] | \mathbf{Z}_t \right\} \\
 &= \phi_{2,t}^2 \text{Var}(I_{2,t}^R X_{t-1}) + (\phi_{2,t}(1 - \phi_{2,t}) \mathbb{E}(I_{2,t}^R X_{t-1})) \\
 &= \phi_{2,t}^2 \mathbb{E}(I_{2,t}^R X_{t-1}^2) - \phi_{2,t}^2 p_2^2 \mu_2^2 + p_2 \phi_{2,t}(1 + \phi_{2,t}) \mu_2 \\
 &= \left[\frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_2)}{1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_2)} \right]^2 [p_2(\sigma_2^2 + \mu_2^2) - p_2^2 \mu_2^2] + \frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_2)(2 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_2))}{[1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_2)]^2} p_2 \mu_2. \tag{A5}
 \end{aligned}$$

and

$$\begin{aligned}
 III &= 2\text{Cov}(I_{1,t}^R(\phi_{1,t} \circ X_{t-1}), I_{2,t}^R(\phi_2 * X_{t-1}) | \mathbf{Z}_t) \\
 &= -2 \frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_1)}{1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_1)} \frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_2)}{1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_2)} p_1 p_2 \mu_1 \mu_2. \tag{A6}
 \end{aligned}$$

Then, $\text{Var}(X_t | \mathbf{Z}_t)$ follows by replacing (A4), (A5) and (A6) in (A3) and some algebra.

(2) The variance of X_t is given by

$$\begin{aligned}
 \text{Var}(X_t) &= \text{Var}[I_{1,t}^R(\phi_{1,t} \circ X_{t-1})] + \text{Var}[I_{2,t}^R(\phi_2 * X_{t-1})] \\
 &\quad + 2\text{Cov}(I_{1,t}^R(\phi_{1,t} \circ X_{t-1}), I_{2,t}^R(\phi_2 * X_{t-1})) + \lambda \\
 &= I + II + III + \lambda. \tag{A7}
 \end{aligned}$$

Similar to the derivation of $\text{Var}(X_t | \mathbf{Z}_t)$,

$$\begin{aligned}
 I &= \text{Var} \left\{ \mathbb{E} [I_{1,t}^R(\phi_1 \circ X_{t-1}) | X_{t-1}, \mathbf{Z}_t] \right\} + \mathbb{E} \left\{ \text{Var} [I_{1,t}^R(\phi_1 \circ X_{t-1}) | X_{t-1}, \mathbf{Z}_t] \right\} \\
 &= \text{Var} \left[I_{1,t}^R \int f(\phi_{1,t}) \phi_{1,t} X_{t-1} d\phi_{1,t} \right] + \mathbb{E} \left\{ \mathbb{E} [I_{1,t}^R \phi_{1,t}(1 - \phi_{1,t}) X_{t-1}] + \text{Var}(I_{1,t}^R \phi_{1,t} X_{t-1}) \right\} \\
 &= \text{Var} [I_{1,t}^R X_{t-1} \mathbb{E}(\phi_{1,t})] + \mathbb{E} \left\{ I_{1,t}^R X_{t-1} \mathbb{E}[\phi_{1,t}(1 - \phi_{1,t})] + I_{1,t}^R X_{t-1} \text{Var}(\phi_{1,t}) \right\} \\
 &= \phi_1^2 [p_1(\sigma_1^2 + \mu_1^2) - p_1^2 \mu_1^2] + p_1 \mu_1 (\phi_1 - \sigma_{\phi_1}^2 - \phi_1^2) + p_1 \sigma_{\phi_1}^2 (\sigma_1^2 + \mu_1^2). \tag{A8}
 \end{aligned}$$

Similarly, we have

$$\begin{aligned}
 II &= \text{Var} \left\{ \mathbb{E} [I_{2,t}^R(\phi_2 \circ X_{t-1}) | X_{t-1}, \mathbf{Z}_t] \right\} + \mathbb{E} \left\{ \text{Var} [I_{2,t}^R(\phi_2 \circ X_{t-1}) | X_{t-1}, \mathbf{Z}_t] \right\} \\
 &= \text{Var} \left[I_{2,t}^R \int f(\phi_{2,t}) \phi_{2,t} X_{t-1} d\phi_{2,t} \right] + \mathbb{E} \left\{ \mathbb{E} [I_{2,t}^R \phi_{2,t}(1 + \phi_{2,t}) X_{t-1}] + \text{Var}(I_{2,t}^R \phi_{2,t} X_{t-1}) \right\} \\
 &= \text{Var} [I_{2,t}^R X_{t-1} \mathbb{E}(\phi_{2,t})] + \mathbb{E} \left\{ I_{2,t}^R X_{t-1} \mathbb{E}[\phi_{2,t}(1 + \phi_{2,t})] + I_{2,t}^R X_{t-1} \text{Var}(\phi_{2,t}) \right\} \\
 &= \phi_2^2 [p_2(\sigma_2^2 + \mu_2^2) - p_2^2 \mu_2^2] + p_2 \mu_2 (\phi_2 + \sigma_{\phi_2}^2 + \phi_2^2) + p_2 \sigma_{\phi_2}^2 (\sigma_2^2 + \mu_2^2). \tag{A9}
 \end{aligned}$$

and

$$\begin{aligned}
 III &= 2\text{Cov}(I_{1,t}^R(\phi_{1,t} \circ X_{t-1}), I_{2,t}^R(\phi_2 * X_{t-1})) \\
 &= -2\phi_1 \phi_2 p_1 p_2 \mu_1 \mu_2. \tag{A10}
 \end{aligned}$$

Then, $\text{Var}(X_t)$ follows by replacing (A8)–(A10) in (A7) and some algebra.

(3) For the conditional autocovariance $\text{Cov}(X_t, X_{t+h} | \mathbf{Z}_{t+1}, \dots, \mathbf{Z}_{t+h})$, when $h = 1$,

$$\begin{aligned} \text{Cov}(X_t, X_{t+1} | \mathbf{Z}_{t+1}) &= \text{Cov}[X_t, E(X_{t+1} | X_t) | \mathbf{Z}_{t+1}] \\ &= \text{Cov}\{X_t, [\phi_{1,t+1}X_t + \lambda]I_{1,t+1}^R + [\phi_{2,t+1}X_t + \lambda]I_{2,t+1}^R | \mathbf{Z}_{t+1}\} \\ &= \sum_{i=1}^2 [\phi_{i,t+1} \text{Cov}(X_t, I_{i,t+1}^R X_t)], \end{aligned}$$

where

$$\text{Cov}(X_t, I_{i,t+1}^R X_t) = E(I_{i,t+1}^R X_t X_t) - E(I_{i,t+1}^R X_t)E(X_t) = p_i(\sigma_i^2 + \mu_i^2) - p_i \mu_i E(X_t),$$

Then,

$$\text{Cov}(X_t, X_{t+1}) = \sum_{i=1}^2 \left\{ \phi_i p_i [(\sigma_i^2 + \mu_i^2) - \mu_i E(X_t)] \right\} = \sum_{i=1}^2 \phi_i p_i \gamma_0^{(i)}.$$

When $h > 1$, there is

$$\begin{aligned} \text{Cov}(X_t, X_{t+h} | \mathbf{Z}_{t+h}) &= \text{Cov}[X_t, E(X_{t+h} | X_{t+h-1}) | \mathbf{Z}_{t+h}] \\ &= \text{Cov}\{X_t, (\phi_{1,t+h}X_{t+h-1})I_{1,t+h}^R + (\phi_{2,t+h}X_{t+h-1})I_{2,t+h}^R + \lambda | \mathbf{Z}_{t+h}\} \\ &= \sum_{i=1}^2 \phi_{i,t+h} [\text{Cov}(X_t, I_{i,t+h}^R X_{t+h-1})], \\ &= \sum_{i=1}^2 \frac{\exp(\mathbf{Z}_{t+h}^\top \boldsymbol{\beta}_i)}{1 + \exp(\mathbf{Z}_{t+h}^\top \boldsymbol{\beta}_i)} \text{Cov}(X_t, I_{i,t+h}^R X_{t+h-1}), \end{aligned}$$

where

$$\text{Cov}(X_t, I_{i,t+h}^R X_{t+h-1}) = p_i \text{Cov}(X_t, X_{t+h-1} | X_{t+h-1} \leq r) = p_i \gamma_{h-1}^{(i)},$$

Then, we obtain

$$\text{Cov}(X_t, X_{t+h}) = \sum_{i=1}^2 \frac{\exp(\mathbf{Z}_{t+h}^\top \boldsymbol{\beta}_i)}{1 + \exp(\mathbf{Z}_{t+h}^\top \boldsymbol{\beta}_i)} p_i \gamma_{h-1}^{(i)}.$$

(4) For the unconditional autocovariance $\text{Cov}(X_t, X_{t+h})$, when $h = 1$,

$$\begin{aligned} \text{Cov}(X_t, X_{t+1} | \mathbf{Z}_{t+1}) &= \text{Cov}[X_t, E(X_{t+1} | X_t)] \\ &= \text{Cov}\{X_t, (\phi_{1,t+1}X_t)I_{1,t+1}^R + (\phi_{2,t+1}X_t)I_{2,t+1}^R + \lambda\} \\ &= \sum_{i=1}^2 \text{Cov}(X_t, [\phi_{i,t+1}I_{i,t+1}^R X_t]), \end{aligned}$$

where

$$\begin{aligned} \text{Cov}(X_t, \phi_{i,t+1}I_{i,t+1}^R X_t) &= E(\phi_{i,t+1}I_{i,t+1}^R X_t X_t) - E(\phi_{i,t+1}I_{i,t+1}^R X_t)E(X_t) \\ &= \mu_{\phi_i}^{(1)} p_i (\sigma_i^2 + \mu_i^2) - \mu_{\phi_i}^{(1)} p_i \mu_i E(X_t), \end{aligned}$$

Then,

$$\text{Cov}(X_t, X_{t+1}) = \sum_{i=1}^2 \left\{ \mu_{\phi_i}^{(1)} p_i [(\sigma_i^2 + \mu_i^2) - \mu_i E(X_t)] \right\} = \sum_{i=1}^2 \mu_{\phi_i}^{(1)} p_i \gamma_0^{(i)}.$$

When $h > 1$, there is

$$\begin{aligned} \text{Cov}(X_t, X_{t+h}) &= \text{Cov}[X_t, E(X_{t+h} | \mathbf{X}_{t+h-1})] \\ &= \text{Cov}\{X_t, (\phi_{1,t+h} X_{t+h-1}) I_{1,t+h}^R + (\phi_{2,t+h} X_{t+h-1}) I_{2,t+h}^R + \lambda | \mathbf{Z}_{t+h}\} \\ &= \sum_{i=1}^2 [\text{Cov}(X_t, \phi_{i,t+h} I_{i,t+h}^R X_{t+h-1})], \end{aligned}$$

where

$$\text{Cov}(X_t, \phi_{i,t+h} I_{i,t+h}^R X_{t+h-1}) = \mu_{\phi_i}^{(h)} p_i \text{Cov}(X_t, X_{t+h-1} | X_{t+h-1} \leq r) = \mu_{\phi_i}^{(h)} p_i \gamma_{h-1}^{(i)}$$

Then, we obtain

$$\text{Cov}(X_t, X_{t+h}) = \sum_{i=1}^2 \mu_{\phi_i}^{(h)} p_i \gamma_{h-1}^{(i)}.$$

Thus, the autocorrelation function $\rho(h) = [\sum_{i=1}^2 \mu_{\phi_i}^{(h)} \phi_i p_i \gamma_{h-1}^{(i)}] \setminus \text{Var}(X_t)$. \square

Proof of Theorem 1. By Taylor’s expansion, there is

$$\begin{aligned} \mathbf{0} &= \frac{1}{\sqrt{n}} \sum_{t=1}^n -\frac{1}{2} \frac{\partial U_t^2(\hat{\boldsymbol{\theta}}_{CLS})}{\partial \boldsymbol{\theta}} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n -\frac{1}{2} \frac{\partial U_t^2(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} + \left(\frac{1}{n} \sum_{t=1}^n -\frac{1}{2} \frac{\partial^2 U_t^2(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right) \sqrt{n}(\hat{\boldsymbol{\theta}}_{CLS} - \boldsymbol{\theta}_0) + o_p(n^{-1/2}). \end{aligned}$$

We first prove

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n -\frac{1}{2} \frac{\partial U_t^2(\hat{\boldsymbol{\theta}}_{CLS})}{\partial \boldsymbol{\theta}} \xrightarrow{d} \mathbf{N}(\mathbf{0}, \mathbf{W}).$$

Now, let $\mathcal{F}_n = \sigma\{X_0, X_1, \dots, X_n, \mathbf{Z}_1, \dots, \mathbf{Z}_{n+1}\}$, and for $i = 1, 2$,

$$M_n^{(i,0)} = \sum_{t=1}^n -\frac{1}{2} \frac{\partial U_t^2(\boldsymbol{\theta}_0)}{\partial \beta_{i,0}} = \sum_{t=1}^n \left[U_t(\boldsymbol{\theta}_0) X_{t-1} I_{i,t}^R \frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)]^2} \right].$$

Then, we have

$$\begin{aligned} E(M_n^{(i,0)} | \mathcal{F}_{n-1}) &= E\{M_{n-1}^{(i,0)} + U_n(\boldsymbol{\theta}_0) X_{n-1} I_{i,n}^R \frac{\exp(\mathbf{Z}_n^\top \boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{Z}_n^\top \boldsymbol{\beta}_i)]^2} \mid \mathcal{F}_{n-1}\} \\ &= M_{n-1}^{(i,0)} + E\{U_n(\boldsymbol{\theta}_0) X_{n-1} I_{i,n}^R \frac{\exp(\mathbf{Z}_n^\top \boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{Z}_n^\top \boldsymbol{\beta}_i)]^2} \mid \mathcal{F}_{n-1}\} = M_{n-1}^{(i,0)}, \end{aligned}$$

i.e., $\{M_n^{(i,0)}, \mathcal{F}_n, n \geq 0\}$ is a martingale. By Proposition (1), $E(X_t^4 | \mathcal{Z}_t) < \infty$, there is

$$E\left((U_t(\boldsymbol{\theta}_0))^2 X_{t-1}^2 I_{i,n}^R \left(\frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)]^2} \right)^2 \mid \mathcal{F}_{t-1} \right) < \infty.$$

Then, by the ergodic theorem, we have

$$\begin{aligned} \frac{1}{n} \sum_{t=2}^n \left((U_t(\boldsymbol{\theta}_0))^2 X_{t-1}^2 I_{i,n}^R \left(\frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)]^2} \right)^2 \right)^2 &\xrightarrow{\text{a.s.}} \mathbb{E} \left[U_t(\boldsymbol{\theta}_0) X_{t-1} I_{i,n}^R \frac{\exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)]^2} \right]^2 \\ &= \mathbb{E} \left[U_t^2(\boldsymbol{\theta}_0) \left(\frac{\partial}{\partial \beta_{i,0}} g(\boldsymbol{\theta}_0, X_{t-1}, \mathbf{Z}_t) \right)^2 \right] \\ &= \mathbf{W}_{(i-1)q+i, (i-1)q+i} \end{aligned}$$

Thus, by Corollary 3.2 from [23], the martingale central limit theorem applies, and we obtain $M_n^{(i,0)} / \sqrt{n} \xrightarrow{d} \mathcal{N}(0, \mathbf{W}_{(i-1)q+i, (i-1)q+i})$. Similarly, we can prove that, for any $i = 1, 2, s = 1, \dots, q$,

$$M_n^{(i,s)} = \sum_{t=1}^n -\frac{1}{2} \frac{\partial U_t^2(\boldsymbol{\theta}_0)}{\partial \beta_{i,s}} = \sum_{t=1}^n \left[U_t(\boldsymbol{\theta}_0) X_{t-1} I_{i,t}^R \frac{Z_{s,t} \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)]^2} \right].$$

is a martingale, and we have that

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \left[U_t(\boldsymbol{\theta}_0) X_{t-1} I_{i,t}^R \frac{Z_{s,t} \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)]^2} \right]^2 &\xrightarrow{\text{a.s.}} \mathbb{E} \left[U_t(\boldsymbol{\theta}_0) X_{t-1} I_{i,t}^R \frac{Z_{s,t} \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)]^2} \right]^2 \\ &= \mathbb{E} \left[U_t^2(\boldsymbol{\theta}_0) \left(\frac{\partial}{\partial \beta_{i,s}} g(\boldsymbol{\theta}_0, X_{t-1}, \mathbf{Z}_t) \right)^2 \right] \\ &= \mathbf{W}_{(i-1)q+i+s, (i-1)q+i+s} \end{aligned}$$

that is $M_n^{(i,s)} / \sqrt{n} \xrightarrow{d} \mathcal{N}(0, \mathbf{W}_{(i-1)q+i+s, (i-1)q+i+s})$. Furthermore, we can also prove that

$$M_n^{(2q+3)} = \sum_{t=1}^n -\frac{1}{2} \frac{\partial U_t^2(\boldsymbol{\theta}_0)}{\partial \lambda} = \sum_{t=1}^n U_t(\boldsymbol{\theta}_0).$$

is a martingale and

$$\frac{1}{n} \sum_{t=1}^n [U_t(\boldsymbol{\theta}_0)]^2 \xrightarrow{\text{a.s.}} \mathbb{E} U_t^2(\boldsymbol{\theta}_0) = \mathbf{W}_{2q+3, 2q+3},$$

that is $M_n^{(2q+3)} / \sqrt{n} \xrightarrow{d} \mathcal{N}(0, \mathbf{W}_{2q+3, 2q+3})$.

In the same way, for any $\mathbf{c} = (c_1, \dots, c_{2q+3})^\top \in \mathbb{R}^{2q+3} \setminus \mathbf{0}_{2q+3,1}$, we have

$$\begin{aligned} &\frac{1}{\sqrt{n}} \mathbf{c}^\top (M_n^{(1,0)}, \dots, M_n^{(1,q)}, M_n^{(2,0)}, \dots, M_n^{(2,q)}, M_n^{(2q+3)})^\top \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n U_t(\boldsymbol{\theta}_0) \left\{ \sum_{i=1}^2 \left[\sum_{s=0}^q c_{(i-1)q+s+i} X_{t-1} \frac{Z_{s,t} \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)]^2} \right] I_{i,t}^R + c_{2q+3} \right\} \\ &\xrightarrow{d} \mathcal{N} \left(0, \mathbb{E} \left\{ U_t(\boldsymbol{\theta}_0) \left\{ \sum_{i=1}^2 \left[\sum_{s=0}^q c_{(i-1)q+s+i} X_{t-1} \frac{Z_{s,t} \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{Z}_t^\top \boldsymbol{\beta}_i)]^2} \right] I_{i,t}^R + c_{2q+3} \right\} \right\}^2 \right). \end{aligned}$$

Thus, by the Cramér–Wold device,

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n -\frac{1}{2} \frac{\partial U_t^2(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} = \frac{1}{\sqrt{n}} (M_n^{(1,0)}, \dots, M_n^{(1,q)}, M_n^{(2,0)}, \dots, M_n^{(2,q)}, M_n^{(2q+3)})^\top \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{W}),$$

Consider the second item of Taylor’s expansion, $\frac{1}{n} \sum_{t=1}^n -\frac{1}{2} \frac{\partial^2 U_t^2(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}$:

$$\begin{aligned} & \frac{1}{n} \sum_{t=1}^n -\frac{1}{2} \frac{\partial^2 U_t^2(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \\ &= \frac{1}{n} \sum_{t=1}^n \left(-\frac{\partial g(\boldsymbol{\theta}_0, X_{t-1}, \mathbf{Z}_t)}{\partial \boldsymbol{\theta}^\top} \frac{\partial g(\boldsymbol{\theta}_0, X_{t-1}, \mathbf{Z}_t)}{\partial \boldsymbol{\theta}} + \frac{\partial^2 g(\boldsymbol{\theta}_0, X_{t-1}, \mathbf{Z}_t)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} (X_t - g(\boldsymbol{\theta}_0, X_{t-1}, \mathbf{Z}_t)) \right). \end{aligned}$$

Note that

$$\mathbb{E} \left(\frac{\partial^2 g(\boldsymbol{\theta}_0, X_{t-1}, \mathbf{Z}_t)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} (X_t - g(\boldsymbol{\theta}_0, X_{t-1}, \mathbf{Z}_t)) \right) = 0$$

and then, by the ergodic theorem, we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n -\frac{1}{2} \frac{\partial^2 U_t^2(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = -V.$$

Hence, we have that

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{CLS} - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, V^{-1}WV^{-1}),$$

and the proof has been completed. \square

Proof of Theorem 2. Considering that the case $R = 0$ is similar to $R = 1$, we only prove the case $R = 0$. We first prove the consistency of $\hat{\boldsymbol{\theta}}_{CML}$. See [24] for a similar technique. Clearly, $L(\boldsymbol{\theta})$ is a measurable function of x_t for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}_\theta$; it is continuous in an open and convex neighborhood $N(\boldsymbol{\theta}_0)$. From the assumption in Theorem 2, it ensures that $\mathbb{E}[\ell_t(\boldsymbol{\theta})]$ has a unique maximizer in the compact set $\boldsymbol{\Theta}_\theta$, say $\check{\boldsymbol{\theta}}$. We can assume $\check{\boldsymbol{\theta}} \in N(\boldsymbol{\theta}_0)$. Meanwhile, for arbitrary point $\boldsymbol{\theta} \in N(\boldsymbol{\theta}_0)$, by Jensen’s inequality, we have

$$\mathbb{E}[\ell_t(\boldsymbol{\theta}) - \ell_t(\boldsymbol{\theta}_0)] = \mathbb{E} \left[\log \left(\frac{P_\theta(\mathbf{z}_t, X_{t-1}, X_t)}{P_{\boldsymbol{\theta}_0}(\mathbf{z}_t, X_{t-1}, X_t)} \right) \right] \leq \log \left[\mathbb{E} \left(\frac{P_\theta(\mathbf{z}_t, X_{t-1}, X_t)}{P_{\boldsymbol{\theta}_0}(\mathbf{z}_t, X_{t-1}, X_t)} \right) \right] = 0. \tag{A11}$$

That is, $\mathbb{E}[\ell_t(\boldsymbol{\theta})]$ is a strict local maximum at $\boldsymbol{\theta}_0$.

In the following, we will show that the log-likelihood function $\sum_{t=1}^n \ell_t(\boldsymbol{\theta})/n$ converges almost surely and uniformly in $\boldsymbol{\Theta}_\theta$ to $\mathbb{E}[\ell_t(\boldsymbol{\theta})]$, that is

$$\left\| \frac{1}{n} \sum_{t=1}^n \ell_t(\boldsymbol{\theta}) - \mathbb{E}[\ell_t(\boldsymbol{\theta})] \right\|_{\boldsymbol{\Theta}_\theta} \xrightarrow{\text{a.s.}} 0,$$

as $n \rightarrow \infty$. Then, the almost sure convergence of $\hat{\boldsymbol{\theta}}_{CML}$ such that $\hat{\boldsymbol{\theta}}_{CML} \xrightarrow{\text{a.s.}} \boldsymbol{\theta}_0$ follows by standard arguments due to [25]. Note that $\{X_t\}$ is stationary and ergodic, then $\ell_t(\boldsymbol{\theta})$ is a stationary and ergodic sequence of random elements that take values in the space of continuous functions $\mathcal{C}(\boldsymbol{\Theta}_\theta, \mathbb{R})$ equipped with the uniform norm $\|\cdot\|_{\boldsymbol{\Theta}_\theta}$. Therefore, the desired convergence result follows by an application of the ergodic theorem of [26], if the uniform integrability condition $\mathbb{E}[\|\ell_t(\boldsymbol{\theta})\|_{\boldsymbol{\Theta}_\theta}] < \infty$ is satisfied. Note that

$$\begin{aligned} |\ell_t(\boldsymbol{\theta})| &= -\log \left(I_{1,t}^R \sum_{m=0}^{\min(x_{t-1}, x_t)} \binom{x_{t-1}}{m} e^{-\lambda} \frac{\lambda^{x_t-m}}{(x_t-m)!} \phi_{1,t}^m (1 - \phi_{1,t})^{x_{t-1}-m} \right. \\ &\quad \left. + I_{2,t}^R \sum_{m=0}^{x_t} \frac{\Gamma(x_{t-1} + m)}{\Gamma(x_{t-1})\Gamma(m+1)} \frac{\phi_{2,t}^m}{(1 + \phi_{2,t})^{x_{t-1}+m}} e^{-\lambda} \frac{\lambda^{x_t-m}}{(x_t-m)!} \right). \end{aligned}$$

For convenience, we first assume that $x_{t-1} \leq r$ and $x_{t-1} < x_t$, i.e., $(x_t - x_{t-1}) \geq 1$, then together with $\log(x) < x$, there is

$$\begin{aligned} |\ell_t(\boldsymbol{\theta})| &\leq -\log\left(e^{-\lambda} \frac{\lambda^{x_t-x_{t-1}}}{(x_t-x_{t-1})!} \phi_{1,t}^{x_{t-1}}\right) \\ &\leq -(-\lambda + (x_t - x_{t-1}) \log(\lambda) - \log[(x_t - x_{t-1})] + x_{t-1} \log(\phi_{1,t})) \\ &\leq (\lambda - (x_t - x_{t-1}) \log(\lambda) + (x_t - x_{t-1}) - x_{t-1} \log(\phi_{1,t})) \\ &\leq x_t[1 - \log(\lambda)] - x_{t-1}[\log(\phi_{1,t}) - \log(\lambda) + 1] + \lambda \end{aligned}$$

almost surely for any $\boldsymbol{\theta} \in \Theta_{\theta}$. Similarly, if $x_{t-1} = x_t$ and we assume $x_{t-1} \leq r$,

$$|\ell_t(\boldsymbol{\theta})| \leq \lambda - x_t \log \phi_{1,t}$$

almost surely for any $\boldsymbol{\theta} \in \Theta_{\theta}$. If $x_{t-1} > x_t$ and we assume $x_{t-1} \leq r$,

$$\begin{aligned} |\ell_t(\boldsymbol{\theta})| &\leq -\log\left(\binom{x_{t-1}}{x_t} e^{-\lambda} \phi_{1,t}^{x_t} (1 - \phi_{1,t})^{x_{t-1}-x_t}\right) \\ &\leq \left(\sum_{j=1}^{x_{t-1}-x_t} \log j - \sum_{j=1}^{x_{t-1}} \log j\right) + \lambda + x_t\{\log[(1 - \phi_{1,t})] - \log[\phi_{1,t}]\} - x_{t-1} \log[(1 - \phi_{1,t})] \\ &\leq \lambda + x_t\{\log[(1 - \phi_{1,t})] - \log[\phi_{1,t}]\} - x_{t-1} \log[(1 - \phi_{1,t})] \end{aligned}$$

almost surely for any $\boldsymbol{\theta} \in \Theta_{\theta}$. Analogously, in terms of $x_{t-1} > r$, there is

$$|\ell_t(\boldsymbol{\theta})| \leq \lambda + x_t\{\log[(1 - \phi_{2,t})] - \log[\phi_{2,t}]\} + x_{t-1} \log[(1 + \phi_{1,t})]$$

almost surely for any $\boldsymbol{\theta} \in \Theta_{\theta}$. Clearly, $E(X_t) < \infty$. We have the conclusion $E\|\ell_t(\boldsymbol{\theta})\|_{\Theta_{\theta}} < \infty$, and the strong consistency has been proven.

To prove this asymptotically, we study the method in [15] and perform the Taylor expansion of the score vector around $\boldsymbol{\theta}_0$:

$$\mathbf{0} = \frac{1}{\sqrt{n}} \frac{\partial L(\hat{\boldsymbol{\theta}}_{CML})}{\partial \boldsymbol{\theta}} = \frac{1}{\sqrt{n}} \frac{\partial L(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} + \left(\frac{1}{n} \frac{\partial^2 L(\boldsymbol{\theta}_n^*)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}\right) \sqrt{n}(\hat{\boldsymbol{\theta}}_{CML} - \boldsymbol{\theta}_0),$$

where $\boldsymbol{\theta}_n^*$ lies in between $\hat{\boldsymbol{\theta}}_{CML}$ and $\boldsymbol{\theta}_0$. It is easy to see $E\left(\frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)_{\boldsymbol{\theta}_0} = \mathbf{0}$, which implies that $\{\ell_t(\boldsymbol{\theta}_0)/\partial \boldsymbol{\theta}\}$ is a martingale difference, and

$$\text{Cov}\left(\frac{\partial \ell_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}}\right) = E\left[\left(\frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right) \left(\frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^\top\right]_{\boldsymbol{\theta}_0}.$$

By the Cramér–Wold device and the central limit theorem in Theorem 18.3 of Billingsley [27], it follows that

$$\frac{1}{\sqrt{n}} \frac{\partial L(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \xrightarrow{d} N(\mathbf{0}, \mathbf{I}(\boldsymbol{\theta}_0)) \text{ with } \mathbf{I}(\boldsymbol{\theta}_0) = E\left[\frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^\top}\right]_{\boldsymbol{\theta}_0}.$$

Next, we aim to show that $\frac{1}{n} \frac{\partial^2 \ell(\boldsymbol{\theta}_n^*)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}$ converges to a finite nonsingular matrix:

$$\mathbf{J}(\boldsymbol{\theta}_0) = E\left(\frac{\partial^2 \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}\right)_{\boldsymbol{\theta}_0}$$

for any $\theta_n^* \xrightarrow{\text{a.s.}} \theta_0$, where θ_n^* lies in between θ_0 and $\hat{\theta}_{CML}$. Note that, for any $i, j = 1, \dots, 2q + 3$, the Taylor's expansion of the second-order derivatives of $L(\theta_n^*)$ at θ_0 is

$$\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t(\theta_n^*)}{\partial \theta_i \partial \theta_j} = \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t(\theta_0)}{\partial \theta_i \partial \theta_j} + \left(\frac{1}{n} \sum_{t=1}^n \frac{\partial^3 l_t(\tilde{\theta})}{\partial \theta_i \partial \theta_j \partial \theta^\top} \right) (\tilde{\theta} - \theta_0),$$

where $\tilde{\theta}$ is between θ_0 and θ_n^* . According to the assumption in Theorem 2, there is

$$\overline{\lim}_{n \rightarrow \infty} \sup_{\theta \in N(\theta_0)} \left| \frac{\partial^3 \ell_t(\theta)}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| < \infty,$$

for any $i, j, k = 1, \dots, 2q + 3$. Thus, combining the assumption $E\left(\frac{1}{n} \frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^\top}\right)_{\theta_0}$ is a nonsingular matrix, the strong consistency of the estimates, and the ergodic theorem in [28], it is easy to obtain a conclusion:

$$\frac{1}{n} \frac{\partial^2 \ell(\theta_n^*)}{\partial \theta \partial \theta^\top} \xrightarrow{\text{a.s.}} \mathbf{J}(\theta_0),$$

as $n \rightarrow \infty$. To sum up, there is

$$\sqrt{n}(\hat{\theta}_{CML} - \theta_0) \xrightarrow{d} N(\mathbf{0}, \mathbf{J}^{-1}(\theta_0) \mathbf{I}(\theta_0) \mathbf{J}^{-1}(\theta_0)).$$

as $n \rightarrow \infty$, and the proof has been completed. \square

References

- Brännäs, K.; Shahiduzzaman Quoreshi, A.M.M. Integer-valued moving average modelling of the number of transactions in stocks. *Appl. Financ. Econ.* **2010**, *20*, 1429–1440. [\[CrossRef\]](#)
- Schweer, S.; Weiß, C.H. Compound Poisson INAR (1) processes: Stochastic properties and testing for overdispersion. *Comput. Stat. Data Anal.* **2014**, *77*, 267–284. [\[CrossRef\]](#)
- Guan, G.; Hu, X. On the analysis of a discrete-time risk model with INAR (1) processes. *Scand. Actuar. J.* **2022**, *2022*, 115–138. [\[CrossRef\]](#)
- Al-Osh, M.A.; Alzaid, A.A. First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.* **1987**, *8*, 261–275. [\[CrossRef\]](#)
- Steutel, F.W.; Van Harn, K. Discrete analogues of self-decomposability and stability. *Ann. Probab.* **1979**, *7*, 893–899. [\[CrossRef\]](#)
- Weiß, C.H. Thinning operations for modeling time series of counts—A survey. *Asta Adv. Stat. Anal.* **2008**, *92*, 319. [\[CrossRef\]](#)
- Freeland, R.K. True integer value time series. *Asta Adv. Stat. Anal.* **2010**, *94*, 217–229. [\[CrossRef\]](#)
- Kachour, M.; Truquet, L. A p -order signed integer-valued autoregressive (SINAR(p)) model. *J. Time Ser. Anal.* **2011**, *32*, 223–236. [\[CrossRef\]](#)
- Nasti, A.S.; Risti, M.M.; Bakouch, H.S. A combined geometric INAR(p) model based on negative binomial thinning. *Math. Comput. Model.* **2012**, *55*, 1665–1672. [\[CrossRef\]](#)
- Khoo, W.C.; Ong, S.H.; Biswas, A. Modeling time series of counts with a new class of INAR(1) model. *Stat. Pap.* **2017**, *58*, 1–24. [\[CrossRef\]](#)
- Li, H.; Yang, K.; Zhao, S.; Wang, D. First-order random coefficients integer-valued threshold autoregressive processes. *Asta Adv. Stat. Anal.* **2018**, *102*, 305–331. [\[CrossRef\]](#)
- Monteiro, M.; Scotto, M.G.; Pereira, I. Integer-valued self-exciting threshold autoregressive processes. *Commun.-Stat.-Theory Methods* **2012**, *41*, 2717–2737. [\[CrossRef\]](#)
- Wang, C.; Liu, H.; Yao, J.F.; Davis, R.A.; Li, W.K. Self-excited threshold Poisson autoregression. *J. Am. Stat. Assoc.* **2014**, *109*, 776–787. [\[CrossRef\]](#)
- Möller, T.A.; Silva, M.E.; Weiß, C.H.; Scotto, M.G.; Pereira, I. Self-exciting threshold binomial autoregressive processes. *Asta Adv. Stat. Anal.* **2016**, *100*, 369–400. [\[CrossRef\]](#)
- Yang, K.; Wang, D.; Jia, B.; Li, H. An integer-valued threshold autoregressive process based on negative binomial thinning. *Stat. Pap.* **2018**, *59*, 1131–1160. [\[CrossRef\]](#)
- Möller, T.A.; Weiß, C.H. Threshold models for integer-valued time series with infinite or finite range. *Stoch. Model. Stat. Their Appl.* **2015**, *122*, 327–334.
- Sheng, D.; Wang, D.; Sun, L.Q. A new First-Order mixture integer-valued threshold autoregressive process based on binomial thinning and negative binomial thinning. *J. Stat. Plan. Inference* **2024**, *231*, 106143. [\[CrossRef\]](#)

18. Yang, K.; Li, H.; Wang, D.; Zhang, C. Random coefficients integer-valued threshold autoregressive processes driven by logistic regression. *Asta Adv. Stat. Anal.* **2021**, *105*, 533–557. [[CrossRef](#)]
19. Sheng, D.; Wang, D.; Kang, Y. A new RCAR(1) model based on explanatory variables and observations. *Commun. Stat.-Theory Methods* **2022**, 1–22. [[CrossRef](#)]
20. Ristić, M.M.; Bakouch, H.S.; Nastixcx, A.S. A new geometric first-order integer-valued autoregressive (NGINAR(1)) process. *J. Stat. Plan. Inference* **2009**, *139*, 2218–2226. [[CrossRef](#)]
21. Tweedie, R.L. Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. *Stoch. Processes Appl.* **1975**, *3*, 385–403. [[CrossRef](#)]
22. Yang, K.; Li, H.; Wang, D. Estimation of parameters in the self-exciting threshold autoregressive processes for nonlinear time series of counts. *Appl. Math. Model.* **2018**, *57*, 226–247. [[CrossRef](#)]
23. Hall, P.; Heyde, C.C. *Martingale Limit Theory and Its Application*; Academic Press: New York, NY, USA, 1980.
24. Gorgi, P. Integer-valued autoregressive models with survival probability driven by a stochastic recurrence equation. *J. Time Ser. Anal.* **2018**, *39*, 150–171. [[CrossRef](#)]
25. Wald, A. Note on the consistency of the maximum likelihood estimate. *Ann. Math. Stat.* **1949**, *20*, 595–601. [[CrossRef](#)]
26. Rao, R.R. Relations between weak and uniform convergence of measures with applications. *Ann. Math. Stat.* **1962**, *33*, 659–680. [[CrossRef](#)]
27. Billingsley, P. *Convergence of Probability Measures*, 2nd ed.; John Wiley & Sons, Inc.: New York, NY, USA, 1999.
28. Durrett, R. *Probability: Theory and Examples*; Cambridge University Press: Cambridge, UK, 2019.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.