

Article

First-Order Random Coefficient Multinomial Autoregressive Model for Finite-Range Time Series of Counts

Jie Zhang ¹, Dehui Wang ², Kai Yang ¹ and Xiaogang Dong ^{1,*}

¹ School of Mathematics and Statistics, Changchun University of Technology, Changchun 130000, China; zhangjie@ccut.edu.cn (J.Z.); yangkai@ccut.edu.cn (K.Y.)

² School of Economics, Liaoning University, Shenyang 110000, China; wangdehui@lnu.edu.cn

* Correspondence: dongxiaogang@ccut.edu.cn

Abstract: In view of the complexity and asymmetry of finite range multi-state integer-valued time series data, we propose a first-order random coefficient multinomial autoregressive model in this paper. Basic probabilistic and statistical properties of the model are discussed. Conditional least squares (CLS) and weighted conditional least squares (WCLS) estimators of the model parameters are derived, and their asymptotic properties are established. In simulation studies, we compare these two methods with the conditional maximum likelihood (CML) method to verify the proposed procedure. A real example is applied to illustrate the advantages of our model.

Keywords: multinomial autoregressive process; random coefficient; binomial thinning; parameter estimation



Citation: Zhang, J.; Wang, D.; Yang, K.; Dong, X. First-Order Random Coefficient Multinomial Autoregressive Model for Finite-Range Time Series of Counts. *Symmetry* **2021**, *13*, 2271. <https://doi.org/10.3390/sym13122271>

Academic Editors: Fukang Zhu and Shishun Zhao

Received: 23 October 2021

Accepted: 21 November 2021

Published: 29 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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

Integer-valued time series data are fairly common in practice. For example, the number of major global earthquakes per year; the number of road accidents in successive months; the number of births at hospital per month, etc. The structure of this kind of data is complex and asymmetrical, and thus many statistical methods may not be suitable. In order to describe this kind of data more effectively, Ref. [1] proposed a binomial thinning operator “ \circ ” defined as:

$$\rho \circ X = \sum_{i=1}^X Y_i, \quad (1)$$

where $\{Y_i\}$ is a sequence of independent and identically distributed (i.i.d.) Bernoulli random variables with distribution $B(1, \rho)$.

Based on the thinning operator (1), Refs. [2,3] introduced a most widely used INAR(1) process; Ref. [4] extended INAR(1) process to the case of random coefficients and proposed the RCINAR(1) process; Ref. [5] discussed the generalized RCINAR(p) process with signed thinning operator; Ref. [6] introduced the threshold RCINAR(1) process and the related forecasting problems; Ref. [7] discussed the RCINAR(1) process with generalized negative binomial marginals.

Comparing the INAR(1) model and its extension, the studies of integer-valued time series with a finite range are relatively few. Ref. [3] attempted to replace INAR(1) process by binomial AR(1) process to describe this kind of data in 1985. Moreover, Ref. [8] studied the threshold autoregressive analysis for a finite-range time series of counts and applied it on measles data. Ref. [9] extended the BAR(1) model to some types of bivariate binomial autoregressive models. A new bivariate binomial time series model with two identical binomial distributions was introduced by Ref. [10].

The models discussed above are all aimed at the integer-valued time series data of up to two states, but there is little research on the integer-valued time series of multistates. Ref. [11] considered the specification tests for the multinomial logit model. Ref. [12]

introduced an INAR(1) negative multinomial regression model for longitudinal count data, and Ref. [13] used the autoregressive negative multinomial regression models to analyze multiple emotions over time. Ref. [14] studied a comparison of choice models for voting research by multinomial probit and multinomial logit. Recently, Ref. [15] extended the binomial AR(1) process to a multinomial autoregressive model, which can describe finite-range integer-valued data with three states. The definition of the F-MAR(1) process is as follows:

$$\begin{aligned} X_{1t} &= Z_{1t} + Z_{2t} + Z_{3t}, \\ X_{2t} &= \alpha_2 \circ (X_{1,t-1} - Z_{1t}) + \beta_2 \circ (X_{2,t-1} - Z_{2t}) \\ &\quad + \gamma_2 \circ (n - X_{1,t-1} - X_{2,t-1} - Z_{3t}), \end{aligned} \quad (2)$$

where α_i , β_i and $\gamma_i \in (0, 1)$ for $i = 1, 2$. $Z_{1t} = \alpha_1 \circ X_{1,t-1}$, $Z_{2t} = \beta_1 \circ X_{2,t-1}$, and $Z_{3t} = \gamma_1 \circ (n - X_{1,t-1} - X_{2,t-1})$.

The F-MAR(1) process is easy to interpret. For a closed group of Alzheimer's disease subjects with n mutually independent subjects, each of these subjects may be in one of three states: cognitively normal (CN), mild cognitive impairment (MCI) and diagnosed with Alzheimer's disease (AD). With drug treatment or other means, these three states can interchange mutually with a certain probability. X_{1t} and X_{2t} represent the number of subjects in states CN and MCI at time t , respectively. Clearly, the number of subjects in state AD can be obtained directly from $n - X_{1t} - X_{2t}$. Although the data is not symmetric, good properties and conclusions can still be obtained due to the stationarity and ergodicity of the model.

However, similar to the RCINAR(1) process, in most situations, the coefficients in F-MAR(1) process can also be affected by many factors. As in the above example, the transition probabilities among the three states should be affected by the patients' age, physical condition, economic conditions, and so on. In this paper, we attempt to extend the F-MAR(1) process to a case with random coefficients, where the fixed α_i , β_i and γ_i in Equation (2) are replaced by i.i.d. random variables α_{it} , β_{it} and γ_{it} . This type of extension will lead to more complex model structure and more difficult derivation of relevant probabilistic and statistical properties. Correspondingly, the extended model can be applied to more practical problems, and the data fitting results may also be better than the model with fixed coefficients.

The rest of the paper is organized as follows: In Section 2, we introduce the first order finite-range random coefficient multinomial autoregressive (F-RCMAR(1)) process, and investigate some basic properties. In Section 3, we discuss the estimations of model and establish the related theoretical properties. The related simulation studies are presented in Section 4. In Section 5, we apply the proposed model to a real example. The results demonstrate the availability and advantage of the proposed model in fitting finite-range data with three states. Some concluding remarks are provided in Section 6. All proof details and some figures are reported in Appendices A–D.

2. Definition and Basic Properties of the F-RCMAR(1) Process

In this section, we introduce a first order finite-range random coefficients multinomial autoregressive (F-RCMAR(1)) process, which is defined as the following recursive equation:

Definition 1. Let $X_t = [X_{1t}, X_{2t}]^T$ be a bivariate random variable, the F-RCMAR(1) process defined as:

$$\begin{aligned} X_{1t} &= Z_{1t} + Z_{2t} + Z_{3t}, \\ X_{2t} &= \alpha_{2t} \circ (X_{1,t-1} - Z_{1t}) + \beta_{2t} \circ (X_{2,t-1} - Z_{2t}) \\ &\quad + \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - Z_{3t}), \end{aligned} \quad (3)$$

where

- (1) $Z_{1t} = \alpha_{1t} \circ X_{1,t-1}$, $Z_{2t} = \beta_{1t} \circ X_{2,t-1}$, $Z_{3t} = \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1})$;
- (2) $n \in \mathbb{N}$ is a given number denoted the upper limit for the multinomial range;
- (3) $\{\alpha_{it}\}$, $\{\beta_{it}\}$ and $\{\gamma_{it}\}$ are i.i.d. random sequence with cumulative distribution functions (CDF) P_{α_i} , P_{β_i} and P_{γ_i} on $(0,1)$ for $i = 1, 2$, respectively, and they are independent of each other;
- (4) let $\mu_{\alpha_i} = E(\alpha_{it})$, $\mu_{\beta_i} = E(\beta_{it})$, $\mu_{\gamma_i} = E(\gamma_{it})$, $\sigma_{\alpha_i}^2 = \text{Var}(\alpha_{it})$, $\sigma_{\beta_i}^2 = \text{Var}(\beta_{it})$, and $\sigma_{\gamma_i}^2 = \text{Var}(\gamma_{it})$ for $i = 1, 2$, note that they are all assumed to exist and be finite;
- (5) all thinnings are performed independently of each other, and the thinnings at time t are independent of $(X_s)_{s < t}$ where $X_s = [X_{1s}, X_{2s}]^T$.

For a fixed population containing n individuals, each individual can only be in one of the three states of I, II and III. X_{1t} in Definition 1 represents the number of individuals in state I in the population at time t , where Z_{it} represents the number of individuals transiting from the i th state at time $t - 1$ to the state I at time t .

The corresponding transition probability will be affected by random coefficients. X_{2t} and $(n - X_{1t} - X_{2t})$ represent the number of individuals in state II and III at time t , respectively. X_{2t} is expressed as shown in Definition 1 to ensure that $\alpha_{1t} \circ X_{1,t-1} + \alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) \leq X_{1,t-1}$, $\beta_{1t} \circ X_{2,t-1} + \beta_{2t} \circ (X_{2,t-1} - \beta_{1t} \circ X_{2,t-1}) \leq X_{2,t-1}$, and $\gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}) + \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1})) \leq n - X_{1,t-1} - X_{2,t-1}$, so as to ensure that the condition of $X_{1t} + X_{2t} \leq n$ is established.

Proposition 1. If $\{X_t\}$ is generated from F-RCMAR(1) process presented in Equation (3), the related transition probability function is deduced as follows,

$$\begin{aligned}
 & P(X_{1t} = x_{1t}, X_{2t} = x_{2t} \mid X_{1,t-1} = x_{1,t-1}, X_{2,t-1} = x_{2,t-1}) \\
 &= \sum_{k=\max\{0, x_{1t}+x_{1,t-1}+x_{2,t-1}-n\}}^{\min\{x_{1t}, x_{1,t-1}+x_{2,t-1}\}} \sum_{j_1=\max\{0, k-x_{2,t-1}\}}^{\min\{x_{1,t-1}, k\}} \\
 & \quad \left\{ \left[\int_0^1 C_{x_{1,t-1}}^{j_1} \alpha_{1t}^{j_1} (1 - \alpha_{1t})^{x_{1,t-1}-j_1} dP_{\alpha_{1t}} \cdot \int_0^1 C_{x_{2,t-1}}^{k-j_1} \beta_{1t}^{k-j_1} (1 - \beta_{1t})^{x_{2,t-1}-k+j_1} dP_{\beta_{1t}} \right. \right. \\
 & \quad \cdot \int_0^1 C_{n-x_{1,t-1}-x_{2,t-1}}^{x_{1t}-k} \gamma_{1t}^{x_{1t}-k} (1 - \gamma_{1t})^{n-x_{1,t-1}-x_{2,t-1}-x_{1t}+k} dP_{\gamma_{1t}} \Big] \\
 & \quad \sum_{s=\max\{0, x_{1t}+x_{2t}+x_{1,t-1}+x_{2,t-1}-k-n\}}^{\min\{x_{2t}, x_{1,t-1}+x_{2,t-1}-k\}} \sum_{j_2=\max\{0, s+j_1-x_{1,t-1}\}}^{\min\{s, x_{2,t-1}-k+j_1\}} \\
 & \quad \left[\int_0^1 C_{x_{1,t-1}-j_1}^{s-j_2} \alpha_{2t}^{s-j_2} (1 - \alpha_{2t})^{x_{1,t-1}-j_1-s+j_2} dP_{\alpha_{2t}} \right. \\
 & \quad \cdot \int_0^1 C_{x_{2,t-1}-k+j_1}^{j_2} \beta_{2t}^{j_2} (1 - \beta_{2t})^{x_{2,t-1}-k+j_1-j_2} dP_{\beta_{2t}} \\
 & \quad \left. \cdot \int_0^1 C_{n-x_{1,t-1}-x_{2,t-1}-x_{1t}+k}^{x_{2t}-s} \gamma_{2t}^{x_{2t}-s} (1 - \gamma_{2t})^{n-x_{1,t-1}-x_{2,t-1}-x_{1t}+k-x_{2t}+s} dP_{\gamma_{2t}} \right] \Big\},
 \end{aligned} \tag{4}$$

where $C_{x_{1,t-1}}^{j_1} = \binom{x_{1,t-1}}{j_1} = \frac{x_{1,t-1}!}{j_1!(x_{1,t-1}-j_1)!}$.

Theorem 1. The process $\{X_t\}$ in Equation (3) is an irreducible, aperiodic and positive recurrent (i.e., ergodic) Markov chain. Hence, there exists a strictly stationary process satisfying the F-RCMAR(1) process.

The proof of Theorem 1 is similar to the proof of Theorem 1 in Ref. [15]. Furthermore, the following properties can be obtained.

Proposition 2. For the stationary F-RCMAR(1) process defined in Equation (3), we have

- (1) $E(X_{1t} | \mathbf{X}_{t-1}) = n\mu_{\gamma_1} + (\mu_{\alpha_1} - \mu_{\gamma_1})X_{1,t-1} + (\mu_{\beta_1} - \mu_{\gamma_1})X_{2,t-1},$
 $E(X_{2t} | \mathbf{X}_{t-1}) = n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) + [\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1})]X_{1,t-1} + [\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1})]X_{2,t-1};$
- (2) $Var(X_{1t} | \mathbf{X}_{t-1}) = (\mu_{\alpha_1} - \sigma_{\alpha_1}^2 - \mu_{\alpha_1}^2)X_{1,t-1} + (\mu_{\beta_1} - \sigma_{\beta_1}^2 - \mu_{\beta_1}^2)X_{2,t-1}$
 $+ (\mu_{\gamma_1} - \sigma_{\gamma_1}^2 - \mu_{\gamma_1}^2)(n - X_{1,t-1} - X_{2,t-1}) + \sigma_{\alpha_1}^2 X_{1,t-1}^2$
 $+ \sigma_{\beta_1}^2 X_{2,t-1}^2 + \sigma_{\gamma_1}^2(n - X_{1,t-1} - X_{2,t-1})^2,$
 $Var(X_{2t} | \mathbf{X}_{t-1}) = [\pi - (\mu_{\alpha_2}^2 + \sigma_{\alpha_2}^2)(1 - \mu_{\alpha_1})^2]X_{1,t-1}$
 $+ [\rho - (\mu_{\beta_2}^2 + \sigma_{\beta_2}^2)(1 - \mu_{\beta_1})^2]X_{2,t-1}$
 $+ [\lambda - (\mu_{\gamma_2}^2 + \sigma_{\gamma_2}^2)(1 - \mu_{\gamma_1})^2](n - X_{1,t-1} - X_{2,t-1})$
 $+ \sigma_{\alpha_2}^2(1 - \mu_{\alpha_1})^2 X_{1,t-1}^2 + \sigma_{\beta_2}^2(1 - \mu_{\beta_1})^2 X_{2,t-1}^2$
 $+ \sigma_{\gamma_2}^2(1 - \mu_{\gamma_1})^2(n - X_{1,t-1} - X_{2,t-1})^2;$
- (3) $E(X_{1t}) = \frac{n[\mu_{\gamma_1}(1-\rho)+\mu_{\beta_1}\lambda]}{1+\mu_{\gamma_1}-\mu_{\alpha_1}+(1-\mu_{\alpha_1}+\mu_{\beta_1})\lambda+(\mu_{\alpha_1}-\mu_{\gamma_1}-1)\rho+(\mu_{\gamma_1}-\mu_{\beta_1})\pi},$
 $E(X_{2t}) = \frac{n[\mu_{\gamma_1}\pi+(1-\mu_{\alpha_1})\lambda]}{1+\mu_{\gamma_1}-\mu_{\alpha_1}+(1-\mu_{\alpha_1}+\mu_{\beta_1})\lambda+(\mu_{\alpha_1}-\mu_{\gamma_1}-1)\rho+(\mu_{\gamma_1}-\mu_{\beta_1})\pi};$
- (4) $Var(X_{1t}) = \left\{ (\mu_{\alpha_1} - \sigma_{\alpha_1}^2 - \mu_{\alpha_1}^2 - \mu_{\gamma_1} + \sigma_{\gamma_1}^2 + \mu_{\gamma_1}^2 - 2n\sigma_{\gamma_1}^2) \cdot E(X_{1t}) + (\mu_{\beta_1} - \sigma_{\beta_1}^2 - \mu_{\beta_1}^2 - \mu_{\gamma_1} + \sigma_{\gamma_1}^2 + \mu_{\gamma_1}^2 - 2n\sigma_{\gamma_1}^2) \cdot E(X_{2t}) + (\sigma_{\alpha_1}^2 + \sigma_{\gamma_1}^2)E^2(X_{1t}) + (\sigma_{\beta_1}^2 + \sigma_{\gamma_1}^2)E^2(X_{2t}) + 2\sigma_{\gamma_1}^2 E(X_{1t})E(X_{2t}) + n(\mu_{\gamma_1} - \sigma_{\gamma_1}^2 - \mu_{\gamma_1}^2) + n^2\sigma_{\gamma_1}^2 + [\sigma_{\beta_1}^2 + \sigma_{\gamma_1}^2 + (\mu_{\beta_1} - \mu_{\gamma_1})^2 + 2\sigma_{\gamma_1}^2(\mu_{\beta_1} - \mu_{\gamma_1})(\rho - \lambda)/A] Var(X_{2t}) \right\} / \left\{ 1 - \sigma_{\alpha_1}^2 - \sigma_{\gamma_1}^2 - (\mu_{\alpha_1} - \mu_{\gamma_1})^2 - 2\sigma_{\gamma_1}^2(\mu_{\alpha_1} - \mu_{\gamma_1})(\rho - \lambda)/A \right\},$
 $Var(X_{2t}) = \left\{ [\pi - (\mu_{\alpha_2}^2 + \sigma_{\alpha_2}^2)(1 - \mu_{\alpha_1})^2 - \lambda + (\mu_{\gamma_2}^2 + \sigma_{\gamma_2}^2 - 2n\sigma_{\gamma_2}^2)(1 - \mu_{\gamma_1})^2] \cdot E(X_{1t}) + [\rho - (\mu_{\beta_2}^2 + \sigma_{\beta_2}^2)(1 - \mu_{\beta_1})^2 - \lambda + (\mu_{\gamma_2}^2 + \sigma_{\gamma_2}^2 - 2n\sigma_{\gamma_2}^2)(1 - \mu_{\gamma_1})^2] \cdot E(X_{2t}) + [\sigma_{\alpha_2}^2(1 - \mu_{\alpha_1})^2 + \sigma_{\gamma_2}^2(1 - \mu_{\gamma_1})^2] E^2(X_{1t}) + [\sigma_{\beta_2}^2(1 - \mu_{\beta_1})^2 + \sigma_{\gamma_2}^2(1 - \mu_{\gamma_1})^2] E^2(X_{2t}) + 2\sigma_{\gamma_2}^2(1 - \mu_{\gamma_1})^2 E(X_{1t})E(X_{2t}) + n\lambda + (n^2\sigma_{\gamma_2}^2 - n\sigma_{\gamma_2}^2 - n\mu_{\gamma_2}^2)(1 - \mu_{\gamma_1})^2 + \{\sigma_{\alpha_2}^2(1 - \mu_{\alpha_1})^2 + \sigma_{\gamma_2}^2(1 - \mu_{\gamma_1})^2 + (\pi - \lambda)^2 + [2\sigma_{\gamma_2}^2(1 - \mu_{\gamma_1})^2 + 2(\pi - \lambda)(\rho - \lambda)](\mu_{\alpha_1} - \mu_{\gamma_1})(\pi - \lambda)/A\} [\mu_{\alpha_1} - \sigma_{\alpha_1}^2 - \mu_{\alpha_1}^2 - \mu_{\gamma_1} + \sigma_{\gamma_1}^2 + \mu_{\gamma_1}^2 - 2n\sigma_{\gamma_1}^2] E(X_{1t}) + (\mu_{\beta_1} - \sigma_{\beta_1}^2 - \mu_{\beta_1}^2 - \mu_{\gamma_1} + \sigma_{\gamma_1}^2 + \mu_{\gamma_1}^2 - 2n\sigma_{\gamma_1}^2) E(X_{2t}) + (\sigma_{\alpha_1}^2 + \sigma_{\gamma_1}^2) E^2(X_{1t}) + (\sigma_{\beta_1}^2 + \sigma_{\gamma_1}^2) E^2(X_{2t}) + 2\sigma_{\gamma_1}^2 E(X_{1t})E(X_{2t}) + n(\mu_{\gamma_1} - \sigma_{\gamma_1}^2 - \mu_{\gamma_1}^2) + n^2\sigma_{\gamma_1}^2 + [\sigma_{\beta_1}^2 + \sigma_{\gamma_1}^2 + (\mu_{\beta_1} - \mu_{\gamma_1})^2 + 2\sigma_{\gamma_1}^2(\mu_{\beta_1} - \mu_{\gamma_1})(\rho - \lambda)/A] / \left[1 - \sigma_{\alpha_1}^2 - \sigma_{\gamma_1}^2 - (\mu_{\alpha_1} - \mu_{\gamma_1})^2 - 2\sigma_{\gamma_1}^2(\mu_{\alpha_1} - \mu_{\gamma_1})(\pi - \lambda)/A \right] \right\} / \left\{ 1 - \sigma_{\beta_2}^2(1 - \mu_{\beta_1})^2 - \sigma_{\gamma_2}^2(1 - \mu_{\gamma_1})^2 - (\rho - \lambda)^2 - [2\sigma_{\gamma_2}^2(1 - \mu_{\gamma_1})^2 + 2(\pi - \lambda)(\rho - \lambda)](\mu_{\beta_1} - \mu_{\gamma_1})(\rho - \lambda)/A - \{\sigma_{\alpha_2}^2(1 - \mu_{\alpha_1})^2 + \sigma_{\gamma_2}^2(1 - \mu_{\gamma_1})^2 + (\pi - \lambda)^2 + [2\sigma_{\gamma_2}^2(1 - \mu_{\gamma_1})^2 + 2(\pi - \lambda)(\rho - \lambda)](\mu_{\alpha_1} - \mu_{\gamma_1})(\pi - \lambda)/A\} [\sigma_{\beta_1}^2 + \sigma_{\gamma_1}^2 + (\mu_{\beta_1} - \mu_{\gamma_1})^2 + 2\sigma_{\gamma_1}^2(\mu_{\beta_1} - \mu_{\gamma_1})(\rho - \lambda)/A] / \left[1 - \sigma_{\alpha_1}^2 - \sigma_{\gamma_1}^2 - (\mu_{\alpha_1} - \mu_{\gamma_1})^2 - 2\sigma_{\gamma_1}^2(\mu_{\alpha_1} - \mu_{\gamma_1})(\pi - \lambda)/A \right] \right\};$
- (5) $Cov(X_{1t}, X_{2t}) = \frac{(\mu_{\alpha_1} - \mu_{\gamma_1})(\pi - \lambda)Var(X_{1t}) + (\mu_{\beta_1} - \mu_{\gamma_1})(\rho - \lambda)Var(X_{2t})}{A},$
 $Cov(X_{1t}, X_{2,t+1}) = (\pi - \lambda)Var(X_{1t}) + (\rho - \lambda)Cov(X_{1t}, X_{2t}),$
 $Cov(X_{1t}, X_{2,t+2}) = [(\mu_{\alpha_1} - \mu_{\gamma_1})(\pi - \lambda) + (\pi - \lambda)(\rho - \lambda)]Var(X_{1t}) + [(\mu_{\beta_1} - \mu_{\gamma_1})(\pi - \lambda) + (\rho - \lambda)^2]Cov(X_{1t}, X_{2t}),$

$$\begin{aligned}
Cov(X_{2t}, X_{1,t+1}) &= (\mu_{\alpha_1} - \mu_{\gamma_1})Cov(X_{1t}, X_{2t}) + (\mu_{\beta_1} - \mu_{\gamma_1})Var(X_{2t}), \\
Cov(X_{2t}, X_{1,t+2}) &= [(\mu_{\alpha_1} - \mu_{\gamma_1})^2 + (\mu_{\beta_1} - \mu_{\gamma_1})(\pi - \lambda)]Cov(X_{1t}, X_{2t}) \\
&\quad + [(\mu_{\alpha_1} - \mu_{\gamma_1})(\mu_{\beta_1} - \mu_{\gamma_1}) + (\mu_{\beta_1} - \mu_{\gamma_1})(\rho - \lambda)]Var(X_{2t}), \\
Cov(X_{1t}, X_{1,t+1}) &= (\mu_{\alpha_1} - \mu_{\gamma_1})Var(X_{1t}) + (\mu_{\beta_1} - \mu_{\gamma_1})Cov(X_{1t}, X_{2t}), \\
Cov(X_{1t}, X_{1,t+2}) &= [(\mu_{\alpha_1} - \mu_{\gamma_1})^2 + (\mu_{\beta_1} - \mu_{\gamma_1})(\pi - \lambda)]Var(X_{1t}) \\
&\quad + [(\mu_{\alpha_1} - \mu_{\gamma_1})(\mu_{\beta_1} - \mu_{\gamma_1}) + (\mu_{\beta_1} - \mu_{\gamma_1})(\pi - \lambda)]Cov(X_{1t}, X_{2t}), \\
Cov(X_{2t}, X_{2,t+1}) &= (\pi - \lambda)Cov(X_{1t}, X_{2t}) + (\rho - \lambda)Var(X_{2t}), \\
Cov(X_{2t}, X_{2,t+2}) &= [(\mu_{\alpha_1} - \mu_{\gamma_1})(\pi - \lambda) + (\pi - \lambda)(\rho - \lambda)]Cov(X_{1t}, X_{2t}) \\
&\quad + [(\mu_{\beta_1} - \mu_{\gamma_1})(\pi - \lambda) + (\rho - \lambda)^2]Var(X_{2t}).
\end{aligned}$$

Since the forms of variance and covariance are so complicated, the derivation procedure is simple and similar, we only present the covariance up to two steps. The corresponding correlations can be solved by the formula of covariance and variance. The rest steps can be deduced by analogy. Here, $\pi = \mu_{\alpha_2}(1 - \mu_{\alpha_1})$, $\rho = \mu_{\beta_2}(1 - \mu_{\beta_1})$, $\lambda = \mu_{\gamma_2}(1 - \mu_{\gamma_1})$, $A = 1 - (\mu_{\alpha_1} - \mu_{\gamma_1})(\rho - \lambda) - (\mu_{\beta_1} - \mu_{\gamma_1})(\pi - \lambda)$.

Remark 1. Based on the statistical properties in Proposition 2, the autocorrelation function $\{\text{Corr}(X_{it}, X_{i,t+k})\}_{i=1,2; k=1,2,\dots}$ of F-RCMAR(1) process can be negative or positive for different parameter combinations, which means that model (3) can describe both positively correlated and negatively correlated data.

3. Parameter Estimation

Suppose that $\{X_t\}_{t=1}^N$ is a strictly stationary and ergodic solution of the F-RCMAR(1) process. The parameter n is assumed to be known. Our major interest lies in the estimation of parameters $\boldsymbol{\theta} = (\mu_{\alpha_1}, \mu_{\alpha_2}, \mu_{\beta_1}, \mu_{\beta_2}, \mu_{\gamma_1}, \mu_{\gamma_2})^\top$. We mainly discuss three different methods to estimate the unknown parameters: conditional least squares (CLS), weighted conditional least squares (WCLS) and compared with conditional maximum likelihood (CML), in next section.

3.1. Conditional Least Squares Estimation

From Proposition 2, the CLS-estimators can be obtained by minimizing the following expression over $\boldsymbol{\theta} \in (0, 1)^6$:

$$\begin{aligned}
S(\boldsymbol{\theta}) &= \sum_{t=1}^N (X_t - E(X_t | \mathbf{X}_{t-1}))^\top (X_t - E(X_t | \mathbf{X}_{t-1})) \\
&= \sum_{t=1}^N [X_{1t} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{1,t-1} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{2,t-1}]^2 \\
&\quad + \sum_{t=1}^N [X_{2t} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{1,t-1} \\
&\quad - (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{2,t-1}]^2. \tag{5}
\end{aligned}$$

Take the partial derivative of Equation (5) with respect to parameters $\boldsymbol{\theta}$, the CLS estimators of $\boldsymbol{\theta}_1 = (\mu_{\alpha_1}, \mu_{\beta_1}, \mu_{\gamma_1})^\top$ can be obtained first as follows:

$$\begin{pmatrix} \hat{\mu}_{\alpha_1}^{CLS} \\ \hat{\mu}_{\beta_1}^{CLS} \\ \hat{\mu}_{\gamma_1}^{CLS} \end{pmatrix} = \mathbf{B}^{-1} \begin{pmatrix} \sum_{t=1}^N X_{1,t-1}X_{1t} \\ \sum_{t=1}^N X_{2,t-1}X_{1t} \\ \sum_{t=1}^N (n - X_{1,t-1} - X_{2,t-1})X_{1t} \end{pmatrix}, \tag{6}$$

where

$$\mathbf{B} = \begin{pmatrix} \sum_{t=1}^N X_{1,t-1}^2 & \sum_{t=1}^N X_{1,t-1}X_{2,t-1} & \sum_{t=1}^N X_{1,t-1}(n - X_{1,t-1} - X_{2,t-1}) \\ \sum_{t=1}^N X_{1,t-1}X_{2,t-1} & \sum_{t=1}^N X_{2,t-1}^2 & \sum_{t=1}^N X_{2,t-1}(n - X_{1,t-1} - X_{2,t-1}) \\ \sum_{t=1}^N X_{1,t-1}(n - X_{1,t-1} - X_{2,t-1}) & \sum_{t=1}^N X_{2,t-1}(n - X_{1,t-1} - X_{2,t-1}) & \sum_{t=1}^N (n - X_{1,t-1} - X_{2,t-1})^2 \end{pmatrix}.$$

With the representations of $\hat{\theta}_1^{CLS}$, based on the expressions of $-\frac{1}{2}\frac{\partial S(\theta)}{\partial \mu_{\alpha_2}}$, $-\frac{1}{2}\frac{\partial S(\theta)}{\partial \mu_{\beta_2}}$ and $-\frac{1}{2}\frac{\partial S(\theta)}{\partial \mu_{\gamma_2}}$, we can find the CLS estimators of $\theta_2 = (\mu_{\alpha_2}, \mu_{\beta_2}, \mu_{\gamma_2})^\top$ as:

$$\hat{\theta}_2^{CLS} = \begin{pmatrix} \hat{\mu}_{\alpha_2}^{CLS} \\ \hat{\mu}_{\beta_2}^{CLS} \\ \hat{\mu}_{\gamma_2}^{CLS} \end{pmatrix} = \mathbf{D}^{-1} \begin{pmatrix} \sum_{t=1}^N X_{1,t-1}X_{2,t} \\ \sum_{t=1}^N X_{2,t-1}X_{2,t} \\ \sum_{t=1}^N (n - X_{1,t-1} - X_{2,t-1})X_{2,t} \end{pmatrix}, \quad (7)$$

where

$$\mathbf{D} = \mathbf{B} \times \begin{pmatrix} 1 - \hat{\mu}_{\alpha_1}^{CLS} & 0 & 0 \\ 0 & 1 - \hat{\mu}_{\beta_1}^{CLS} & 0 \\ 0 & 0 & 1 - \hat{\mu}_{\gamma_1}^{CLS} \end{pmatrix}.$$

The following Theorem 2 and Theorem 3 show the asymptotic normality of $\hat{\theta}_2^{CLS}$.

Theorem 2. Considering $\{\mathbf{X}_t\}$ obtained from model (3), the CLS-estimators $\hat{\theta}_1^{CLS}$ given in Equation (6) are the consistent estimators of θ_1 and asymptotically normal as:

$$\sqrt{N}(\hat{\theta}_1^{CLS} - \theta_1) \xrightarrow{d} N(\mathbf{0}, \mathbf{V}_1^{-1} \mathbf{W}_1 \mathbf{V}_1^{-1}). \quad (8)$$

Theorem 3. Let $\{\mathbf{X}_t\}$ represent a process satisfying Equation (3), $\hat{\theta}_1^{CLS}$ are the consistent estimators of θ_1 . Then, the CLS-estimators $\hat{\theta}_2^{CLS}$ also satisfy the asymptotic normality as:

$$\sqrt{N}(\hat{\theta}_2^{CLS} - \theta_2) \xrightarrow{d} N(\mathbf{0}, \mathbf{V}_2^{-1} \mathbf{W}_2 \mathbf{V}_2^{-1}). \quad (9)$$

3.2. Weighted Conditional Least Squares Estimation

In this section, we consider adjusting function (5) by weights to improve the effect of CLS estimation. The related weighted conditional least squares (WCLS) estimators $\hat{\theta}_{WCLS} := (\hat{\mu}_{\alpha_1}^{WCLS}, \hat{\mu}_{\alpha_2}^{WCLS}, \hat{\mu}_{\beta_1}^{WCLS}, \hat{\mu}_{\beta_2}^{WCLS}, \hat{\mu}_{\gamma_1}^{WCLS}, \hat{\mu}_{\gamma_2}^{WCLS})^\top$ can be obtained by minimizing the following function:

$$Q(\theta) = \sum_{t=1}^N V_\tau^{-1} (\mathbf{X}_t - E(\mathbf{X}_t | \mathbf{X}_{t-1}))^\top (\mathbf{X}_t - E(\mathbf{X}_t | \mathbf{X}_{t-1})). \quad (10)$$

The weight V_τ in Equation (10) is expressed as follows:

$$\begin{aligned} V_\tau = & \{ [\hat{\mu}_{\alpha_1}(1 - \hat{\mu}_{\alpha_1})X_{1,t-1} + \hat{\mu}_{\beta_1}(1 - \hat{\mu}_{\beta_1})X_{2,t-1} + \hat{\mu}_{\gamma_1}(1 - \hat{\mu}_{\gamma_1}) \\ & (n - X_{1,t-1} - X_{2,t-1})]^2 + [\hat{\mu}_{\alpha_2}(1 - \hat{\mu}_{\alpha_2})(1 - \hat{\mu}_{\alpha_1})X_{1,t-1} \\ & + \hat{\mu}_{\beta_2}(1 - \hat{\mu}_{\beta_2})(1 - \hat{\mu}_{\beta_1})X_{2,t-1} + \\ & \hat{\mu}_{\gamma_2}(1 - \hat{\mu}_{\gamma_2})(1 - \hat{\mu}_{\gamma_1})(n - X_{1,t-1} - X_{2,t-1})]^2 \}^{\frac{1}{2}}. \end{aligned} \quad (11)$$

The expression of V_τ is based on the one-step conditional variance of F-MAR(1) process in Ref. [15], i.e., the related model with fixed coefficients, and the fixed coefficients in one-step conditional variance are replaced by the expectation of random coefficients. $\tau = (\hat{\mu}_{\alpha_1}, \hat{\mu}_{\alpha_2}, \hat{\mu}_{\beta_1}, \hat{\mu}_{\beta_2}, \hat{\mu}_{\gamma_1}, \hat{\mu}_{\gamma_2})^\top$ are the consistent estimators of θ . In the simulation, we use the CLS estimators to replace them.

Through the similar algebraic operations, we can find the WCLS estimators $\hat{\theta}_{WCLS}$ as:

$$\hat{\theta}_1^{WCLS} = \begin{pmatrix} \hat{\mu}_{\alpha_1}^{WCLS} \\ \hat{\mu}_{\beta_1}^{WCLS} \\ \hat{\mu}_{\gamma_1}^{WCLS} \end{pmatrix} = \mathbf{B}_1^{-1} \begin{pmatrix} \sum_{t=1}^N \frac{1}{V_\tau} X_{1,t-1} X_{1t} \\ \sum_{t=1}^N \frac{1}{V_\tau} X_{2,t-1} X_{1t} \\ \sum_{t=1}^N \frac{1}{V_\tau} (n - X_{1,t-1} - X_{2,t-1}) X_{1t} \end{pmatrix}, \quad (12)$$

where $\mathbf{B}_1 = (\mathbf{B}_1^{(ij)})_{3 \times 3}$ with

$$\begin{aligned} \mathbf{B}_1^{(11)} &= \sum_{t=1}^N \frac{1}{V_\tau} X_{1,t-1}^2, & \mathbf{B}_1^{(12)} = \mathbf{B}_1^{(21)} &= \sum_{t=1}^N \frac{1}{V_\tau} X_{1,t-1} X_{2,t-1}, \\ \mathbf{B}_1^{(22)} &= \sum_{t=1}^N \frac{1}{V_\tau} X_{2,t-1}^2, & \mathbf{B}_1^{(13)} = \mathbf{B}_1^{(31)} &= \sum_{t=1}^N \frac{1}{V_\tau} X_{1,t-1} (n - X_{1,t-1} - X_{2,t-1}), \\ \mathbf{B}_1^{(33)} &= \sum_{t=1}^N \frac{1}{V_\tau} (n - X_{1,t-1} - X_{2,t-1})^2, \\ \mathbf{B}_1^{(23)} = \mathbf{B}_1^{(32)} &= \sum_{t=1}^N \frac{1}{V_\tau} X_{2,t-1} (n - X_{1,t-1} - X_{2,t-1}). \end{aligned}$$

$$\hat{\theta}_2^{WCLS} = \begin{pmatrix} \hat{\mu}_{\alpha_2}^{WCLS} \\ \hat{\mu}_{\beta_2}^{WCLS} \\ \hat{\mu}_{\gamma_2}^{WCLS} \end{pmatrix} = \mathbf{D}_1^{-1} \begin{pmatrix} \sum_{t=1}^N \frac{1}{V_\tau} X_{1,t-1} X_{2t} \\ \sum_{t=1}^N \frac{1}{V_\tau} X_{2,t-1} X_{2t} \\ \sum_{t=1}^N \frac{1}{V_\tau} (n - X_{1,t-1} - X_{2,t-1}) X_{2t} \end{pmatrix}, \quad (13)$$

where

$$\mathbf{D}_1 = \mathbf{B}_1 \times \begin{pmatrix} 1 - \hat{\mu}_{\alpha_1}^{WCLS} & 0 & 0 \\ 0 & 1 - \hat{\mu}_{\beta_1}^{WCLS} & 0 \\ 0 & 0 & 1 - \hat{\mu}_{\gamma_1}^{WCLS} \end{pmatrix}.$$

Theorem 4. For the WCLS estimators $\hat{\theta}_1^{WCLS} = (\hat{\mu}_{\alpha_1}^{WCLS}, \hat{\mu}_{\beta_1}^{WCLS}, \hat{\mu}_{\gamma_1}^{WCLS})^T$ and $\hat{\theta}_2^{WCLS} = (\hat{\mu}_{\alpha_2}^{WCLS}, \hat{\mu}_{\beta_2}^{WCLS}, \hat{\mu}_{\gamma_2}^{WCLS})^T$ given by Equations (12) and (13), we have

$$\sqrt{N}(\hat{\theta}_1^{WCLS} - \theta_1) \xrightarrow{d} N(\mathbf{0}, \mathbf{V}_3^{-1} \mathbf{W}_3 \mathbf{V}_3^{-1}), \quad (14)$$

and

$$\sqrt{N}(\hat{\theta}_2^{WCLS} - \theta_2) \xrightarrow{d} N(\mathbf{0}, \mathbf{V}_4^{-1} \mathbf{W}_4 \mathbf{V}_4^{-1}), \quad (15)$$

where

$$\begin{aligned} \mathbf{V}_3^{(11)} &= (1 - 2\mu_{\alpha_1}) E \left[\frac{X_{10}^2}{V_\tau} \right] + E \left[\frac{[X_{11} - \mu_{\gamma_1}(n - X_{10} - X_{20}) - \mu_{\beta_1}X_{20}]X_{10}}{V_\tau} \right], \\ \mathbf{V}_3^{(22)} &= (1 - 2\mu_{\beta_1}) E \left[\frac{X_{20}^2}{V_\tau} \right] + E \left[\frac{(X_{11} - \mu_{\gamma_1}(n - X_{10} - X_{20}) - \mu_{\alpha_1}X_{10})X_{20}}{V_\tau} \right], \\ \mathbf{V}_3^{(33)} &= (1 - 2\mu_{\gamma_1}) E \left[\frac{(n - X_{10} - X_{20})^2}{V_\tau} \right] + \\ &\quad E \left[\frac{(X_{11} - \mu_{\alpha_1}X_{10} - \mu_{\beta_1}X_{20})(n - X_{10} - X_{20})}{V_\tau} \right], \end{aligned}$$

$$\begin{aligned}
\mathbf{V}_3^{(12)} &= (1 - \mu_{\alpha_1}) E \left[\frac{X_{10} X_{20}}{V_\tau} \right], \quad \mathbf{V}_3^{(21)} = (1 - \mu_{\beta_1}) E \left[\frac{X_{10} X_{20}}{V_\tau} \right], \\
\mathbf{V}_3^{(13)} &= (1 - \mu_{\alpha_1}) E \left[\frac{X_{10}(n - X_{10} - X_{20})}{V_\tau} \right], \\
\mathbf{V}_3^{(23)} &= (1 - \mu_{\beta_1}) E \left[\frac{X_{20}(n - X_{10} - X_{20})}{V_\tau} \right], \\
\mathbf{V}_3^{(31)} &= (1 - \mu_{\gamma_1}) E \left[\frac{X_{10}(n - X_{10} - X_{20})}{V_\tau} \right], \\
\mathbf{V}_3^{(32)} &= (1 - \mu_{\gamma_1}) E \left[\frac{X_{20}(n - X_{10} - X_{20})}{V_\tau} \right];
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{W}_3^{(11)} &= E \left[\frac{1}{V_\tau^2} [X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 (1 - \mu_{\alpha_1})^2 X_{10}^2 \right], \\
\mathbf{W}_3^{(22)} &= E \left[\frac{1}{V_\tau^2} [X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 (1 - \mu_{\beta_1})^2 X_{20}^2 \right], \\
\mathbf{W}_3^{(33)} &= E \left[\frac{1}{V_\tau^2} [X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 \right. \\
&\quad \left. (1 - \mu_{\gamma_1})^2 (n - X_{10} - X_{20})^2 \right], \\
\mathbf{W}_3^{(12)} &= \mathbf{W}_3^{(21)} = E \left[\frac{1}{V_\tau^2} [X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 \right. \\
&\quad \left. (1 - \mu_{\alpha_1})(1 - \mu_{\beta_1})X_{10}X_{20} \right], \\
\mathbf{W}_3^{(13)} &= \mathbf{W}_3^{(31)} = E \left[\frac{1}{V_\tau^2} [X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 \right. \\
&\quad \left. (1 - \mu_{\alpha_1})(1 - \mu_{\gamma_1})X_{10}(n - X_{10} - X_{20}) \right], \\
\mathbf{W}_3^{(23)} &= \mathbf{W}_3^{(32)} = E \left[\frac{1}{V_\tau^2} [X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 \right. \\
&\quad \left. (1 - \mu_{\beta_1})(1 - \mu_{\gamma_1})X_{20}(n - X_{10} - X_{20}) \right].
\end{aligned}$$

$$\begin{aligned}
\mathbf{V}_4^{(11)} &= (1 - \mu_{\alpha_1})^2 E \left[\frac{X_{10}^2}{V_\tau} \right], \quad \mathbf{V}_4^{(22)} = (1 - \mu_{\beta_1})^2 E \left[\frac{X_{20}^2}{V_\tau} \right], \\
\mathbf{V}_4^{(33)} &= (1 - \mu_{\gamma_1})^2 E \left[\frac{(n - X_{10} - X_{20})^2}{V_\tau} \right], \\
\mathbf{V}_4^{(12)} &= \mathbf{V}_4^{(21)} = (1 - \mu_{\alpha_1})(1 - \mu_{\beta_1}) E \left[\frac{X_{10} X_{20}}{V_\tau} \right], \\
\mathbf{V}_4^{(13)} &= \mathbf{V}_4^{(31)} = (1 - \mu_{\alpha_1})(1 - \mu_{\gamma_1}) E \left[\frac{X_{10}(n - X_{10} - X_{20})}{V_\tau} \right], \\
\mathbf{V}_4^{(23)} &= \mathbf{V}_4^{(32)} = (1 - \mu_{\beta_1})(1 - \mu_{\gamma_1}) E \left[\frac{X_{20}(n - X_{10} - X_{20})}{V_\tau} \right];
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{W}_4^{(11)} &= E \left\{ \frac{1}{V_\tau^2} [X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10} \right. \\
&\quad \left. - (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2 (1 - \mu_{\alpha_1})^2 X_{10}^2 \right\};
\end{aligned}$$

$$\begin{aligned}
W_4^{(22)} &= E\left\{\frac{1}{V_\tau^2}[X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10}\right. \\
&\quad \left.- (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2(1 - \mu_{\beta_1})^2X_{20}^2\right\}; \\
W_4^{(33)} &= E\left\{\frac{1}{V_\tau^2}[X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10}\right. \\
&\quad \left.- (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2(1 - \mu_{\gamma_1})^2(n - X_{10} - X_{20})^2\right\}; \\
W_4^{(12)} = W_4^{(21)} &= E\left\{\frac{1}{V_\tau^2}[X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10}\right. \\
&\quad \left.- (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2(1 - \mu_{\alpha_1})(1 - \mu_{\beta_1})X_{10}X_{20}\right\}; \\
W_4^{(13)} = W_4^{(31)} &= E\left\{\frac{1}{V_\tau^2}[X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10}\right. \\
&\quad \left.- (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2(1 - \mu_{\alpha_1})(1 - \mu_{\gamma_1})X_{10}(n - X_{10} - X_{20})\right\}; \\
W_4^{(23)} = W_4^{(32)} &= E\left\{\frac{1}{V_\tau^2}[X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10}\right. \\
&\quad \left.- (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2(1 - \mu_{\beta_1})(1 - \mu_{\gamma_1})X_{20}(n - X_{10} - X_{20})\right\}.
\end{aligned}$$

Theorem 4 is similar in proof to the Theorem 2 and Theorem 3. Therefore, we will not prove this theorem any more.

4. Simulation

Consider the F-RCMAR(1) process (3):

$$\begin{aligned}
X_{1t} &= Z_{1t} + Z_{2t} + Z_{3t}, \\
X_{2t} &= \alpha_{2t} \circ (X_{1,t-1} - Z_{1t}) + \beta_{2t} \circ (X_{2,t-1} - Z_{2t}) \\
&\quad + \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - Z_{3t}),
\end{aligned}$$

where $Z_{1t} = \alpha_{1t} \circ X_{1,t-1}$, $Z_{2t} = \beta_{1t} \circ X_{2,t-1}$, $Z_{3t} = \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1})$.

In this section, we assume that α_{it} , β_{it} and γ_{it} follow the power function distributions $PF(1, \alpha_i)$, $PF(1, \beta_i)$, and $PF(1, \gamma_i)$ for $i = 1, 2$, respectively. Under these assumptions, we compare the performances of the CML estimators with those of the CLS estimators and WCLS estimators described above.

The probability density function of the power function distribution $PF(\theta, c)$ is:

$$p(x; \theta, c) = cx^{c-1}\theta^{-c}, \quad 0 < x < \theta, \theta > 0, c > 0.$$

We have $E(X) = \frac{c\theta}{c+1}$ and $Var(X) = \frac{c\theta^2}{(c+1)^2(c+2)}$. Here, we make $\theta = 1$ so that α_{it} , β_{it} and $\gamma_{it} \in (0, 1)$ for $i = 1, 2$. According to the expression of expectation and variance, we can also find the relationship between them as:

$$Var(X) = \frac{c\theta^2}{(c+1)^2(c+2)} = \frac{E(X)}{(c+1)(c+2)} = \frac{E(X)(1 - E(X))^2}{2 - E(X)}.$$

Therefore, when we find the estimators of $\{\mu_{\alpha_i}, \mu_{\beta_i}, \mu_{\gamma_i}, i = 1, 2\}$, the related estimators of $\{\sigma_{\alpha_i}^2, \sigma_{\beta_i}^2, \sigma_{\gamma_i}^2, i = 1, 2\}$ can also be obtained directly. The estimation of variance is not given in the simulation any more.

For a fixed value of x_0 , it is easy to derive the conditional likelihood function of process (3) as:

$$L(\eta) = \prod_{t=1}^N P(X_t = x_t | X_{t-1} = x_{t-1}).$$

The transition probability $P(X_t = x_t | X_{t-1} = x_{t-1})$ is:

$$\begin{aligned}
P(X_t | X_{t-1}) = & \sum_{k=\max\{0, X_{1,t}+X_{1,t-1}+X_{2,t-1}-n\}}^{\min\{X_{1,t}, X_{1,t-1}+X_{2,t-1}\}} \sum_{j_1=\max\{0, k-X_{2,t-1}\}}^{\min\{X_{1,t-1}, k\}} \\
& \left\{ \left[C_{X_{1,t-1}}^{j_1} \alpha_1 B(\alpha_1 + j_1, X_{1,t-1} - j_1 + 1) \right. \right. \\
& \cdot C_{X_{2,t-1}}^{k-j_1} \beta_1 B(\beta_1 + k - j_1, X_{2,t-1} - k + j_1 + 1) \\
& \cdot C_{n-X_{1,t-1}-X_{2,t-1}}^{X_{1t}-k} \gamma_1 \\
& \left. B(\gamma_1 + X_{1t} - k, n - X_{1,t-1} - X_{2,t-1} - X_{1t} + k + 1) \right] \cdot \\
& \sum_{s=\max\{0, X_{1t}+X_{2t}+X_{1,t-1}+X_{2,t-1}-k-n\}}^{\min\{X_{2t}, X_{1,t-1}+X_{2,t-1}-k\}} \sum_{j_2=\max\{0, s+j_1-X_{1,t-1}\}}^{\min\{s, X_{2,t-1}-k+j_1\}} \\
& \left[C_{X_{1,t-1}-j_1}^{s-j_2} \alpha_2 B(\alpha_2 + s - j_2, X_{1,t-1} - j_1 - s + j_2 + 1) \right. \\
& \cdot C_{X_{2,t-1}-k+j_1}^{j_2} \beta_2 B(\beta_2 + j_2, X_{2,t-1} - k + j_1 - j_2 + 1) \\
& \cdot C_{n-X_{1,t-1}-X_{2,t-1}-X_{1t}+k}^{X_{2t}-s} \gamma_2 \\
& \left. \left. \cdot B(\gamma_2 + X_{2t} - s, n - X_{1,t-1} - X_{2,t-1} - X_{1t} + k - X_{2t} + s + 1) \right] \right\},
\end{aligned}$$

where the Beta function $B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$, and $\boldsymbol{\eta} = (\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2)^\top$.

The CML-estimators

$$\hat{\boldsymbol{\eta}}_{CML} := (\hat{\alpha}_1^{CML}, \hat{\alpha}_2^{CML}, \hat{\beta}_1^{CML}, \hat{\beta}_2^{CML}, \hat{\gamma}_1^{CML}, \hat{\gamma}_2^{CML})^\top$$

can be easily obtained by solving the following equation:

$$\hat{\boldsymbol{\eta}}_{CML} = \arg \inf_{\boldsymbol{\eta} \in (0, +\infty)^6} \{-\log L(\boldsymbol{\eta})\}.$$

The consistency and the asymptotic normality of the CML-estimators $\hat{\boldsymbol{\eta}}_{CML}$ are established in Theorem 5.

Theorem 5. Let $\{X_t\}$ be a F-RCMAR(1) process satisfying Equation (3). Then, there exists consistent CML estimators of $\boldsymbol{\eta}$, which are also asymptotically normally distributed as

$$\sqrt{N}(\hat{\boldsymbol{\eta}}_{CML} - \boldsymbol{\eta}) \xrightarrow{d} N(\mathbf{0}, \mathbf{I}^{-1}(\boldsymbol{\eta})), \quad (16)$$

where $\mathbf{I}^{-1}(\boldsymbol{\eta})$ is the Fisher information matrix.

Furthermore, the CML estimators of $\boldsymbol{\theta} = (\mu_{\alpha_1}, \mu_{\alpha_2}, \mu_{\beta_1}, \mu_{\beta_2}, \mu_{\gamma_1}, \mu_{\gamma_2})^\top$ can be obtained by

$$\begin{aligned}
& \left(\hat{\mu}_{\alpha_1}^{CML}, \hat{\mu}_{\alpha_2}^{CML}, \hat{\mu}_{\beta_1}^{CML}, \hat{\mu}_{\beta_2}^{CML}, \hat{\mu}_{\gamma_1}^{CML}, \hat{\mu}_{\gamma_2}^{CML} \right)^\top \\
& = \left(\frac{\hat{\alpha}_1^{CML}}{1 + \hat{\alpha}_1^{CML}}, \frac{\hat{\alpha}_2^{CML}}{1 + \hat{\alpha}_2^{CML}}, \frac{\hat{\beta}_1^{CML}}{1 + \hat{\beta}_1^{CML}}, \frac{\hat{\beta}_2^{CML}}{1 + \hat{\beta}_2^{CML}}, \frac{\hat{\gamma}_1^{CML}}{1 + \hat{\gamma}_1^{CML}}, \frac{\hat{\gamma}_2^{CML}}{1 + \hat{\gamma}_2^{CML}} \right)^\top.
\end{aligned}$$

To report the performances of the proposed methods described above, we conduct simulation studies under the following various combinations of the F-RCMAR(1) model with sample size $N = 100, 200, 500, 1000$. The corresponding moments can be obtained

according to Proposition 2. The related sample paths and autocorrelation function (ACF) of X_{1t} and X_{2t} are provided in Figure 1.

Case A. $(n, \mu_{\alpha_1}, \mu_{\alpha_2}, \mu_{\beta_1}, \mu_{\beta_2}, \mu_{\gamma_1}, \mu_{\gamma_2}) = (50, 0.05, 0.05, 0.1, 0.1, 0.6, 0.9)$,

with $EX_{1t} = 16.0574$, $EX_{2t} = 10.2221$,

$VarX_{1t} = 122.7455$, $VarX_{2t} = 41.9010$, $Corr(X_{1t}, X_{2t}) = 0.5366$;

Case B. $(n, \mu_{\alpha_1}, \mu_{\alpha_2}, \mu_{\beta_1}, \mu_{\beta_2}, \mu_{\gamma_1}, \mu_{\gamma_2}) = (7, 0.4, 0.1, 0.3, 0.5, 0.6, 0.9)$,

with $EX_{1t} = 3.1070$, $EX_{2t} = 1.5722$,

$VarX_{1t} = 3.6837$, $VarX_{2t} = 1.6654$, $Corr(X_{1t}, X_{2t}) = 0.1005$;

Case C. $(n, \mu_{\alpha_1}, \mu_{\alpha_2}, \mu_{\beta_1}, \mu_{\beta_2}, \mu_{\gamma_1}, \mu_{\gamma_2}) = (7, 0.9, 0.2, 0.3, 0.5, 0.7, 0.8)$,

with $EX_{1t} = 5.9119$, $EX_{2t} = 0.4263$,

$VarX_{1t} = 1.2003$, $VarX_{2t} = 0.4269$, $Corr(X_{1t}, X_{2t}) = -0.1124$;

Case D. $(n, \mu_{\alpha_1}, \mu_{\alpha_2}, \mu_{\beta_1}, \mu_{\beta_2}, \mu_{\gamma_1}, \mu_{\gamma_2}) = (7, 0.9, 0.2, 0.4, 0.7, 0.1, 0.8)$,

with $EX_{1t} = 5.1532$, $EX_{2t} = 1.1021$,

$VarX_{1t} = 4.0181$, $VarX_{2t} = 1.8506$, $Corr(X_{1t}, X_{2t}) = -0.6112$;

The sample paths of X_{it} in Figure 1 fluctuate steadily around the mean of X_{it} for $i = 1, 2$. The ACFs of X_{it} display the correlation of data. In addition, the related simulation results of the empirical biased (bias) and mean squared errors (MSE) are summarized in Tables 1 and 2. These values are shown within parenthesis as the format $(\frac{bias}{MSE})$ in following tables. All simulations for parameter combinations above are calculated under MATLAB software based on 1000 replications.

Clearly, with the sample size N increasing, the bias and MSE of three kinds of estimators considered almost decrease as expected. That is to say, the three estimation methods discussed above are consistent for all parameters. Furthermore, the simulation results of WCLS estimation performed better than those of CLS estimation, which means that using Equation (11) as the weight function is a satisfactory choice. The WCLS estimation method can effectively improve the results of the CLS estimation method. On the whole, the CML estimators $\hat{\theta}_{CML}$ have the smallest MSE and the best robustness within all three methods. Thus, we can conclude that the CML estimators have the best performances when compared with other two methods.

Furthermore, Figures 2–4 show the boxplots of the biases for CLS, WCLS and CML estimations with Case B with sample sizes of $N = 100, 200, 500, 1000$. We can find that the estimates for parameters tend to be unbiased and consistent. Figures 5–7 display the QQ plots of three kinds of estimators for Case B with sample size $N = 500$, which imply that all the CLS, WCLS and CML estimators are asymptotically normal for all the parameters. The related histograms of the three kinds of estimators are shown in Appendix D. For other parameter combinations, the QQ plots and boxplots of simulation results should be similar, which will be omitted here.

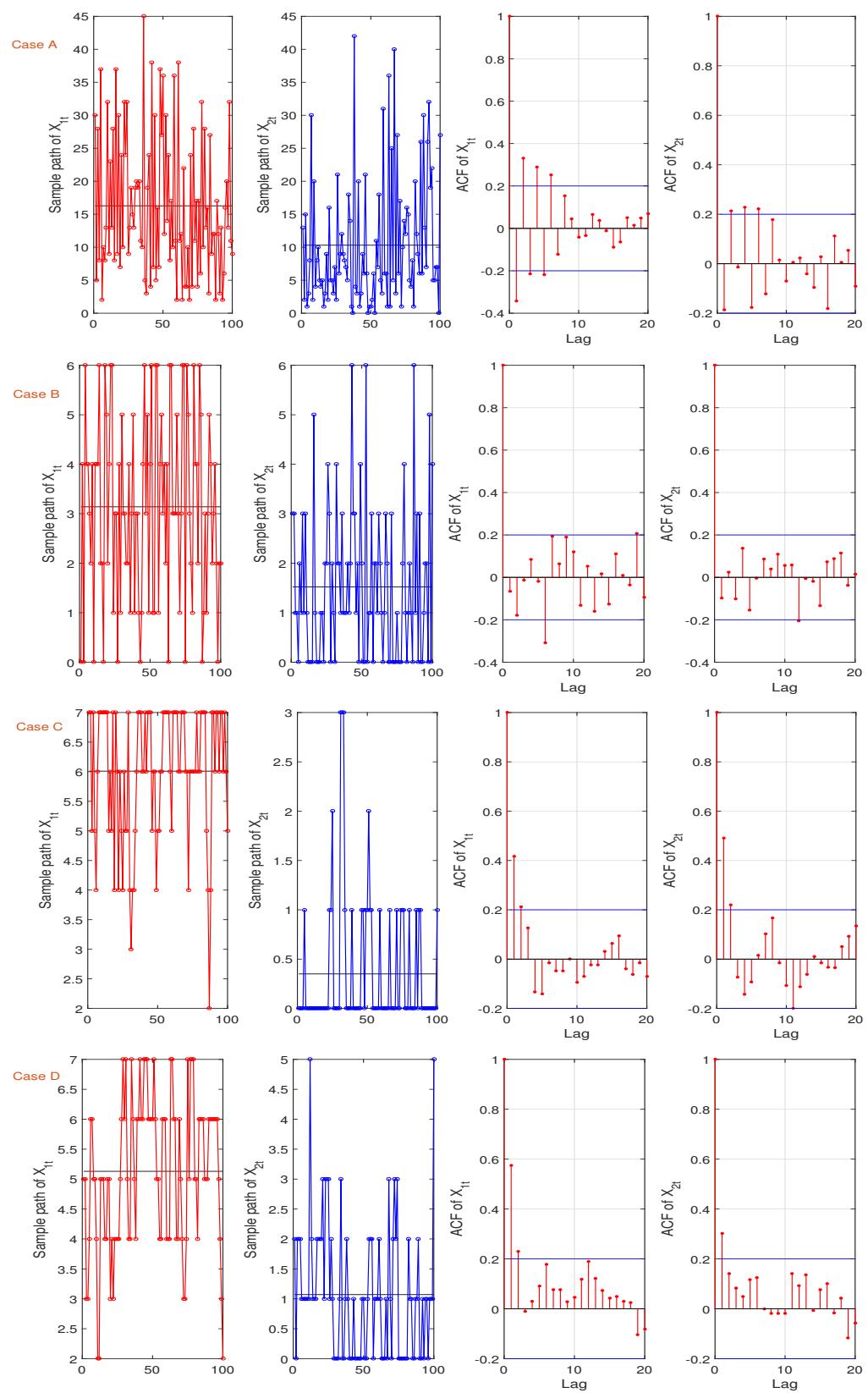


Figure 1. The sample paths and ACFs of X_{1t} and X_{2t} for Cases A, B, C and D with the sample size equal to 100. The horizontal lines in sample paths are the mean of X_{1t} and X_{2t} .

Table 1. Empirical bias and MSE of estimators of θ for cases A, B, C and D.

Case A	$n = 50$	$\mu_{\alpha_1} = 0.05$	$\mu_{\alpha_2} = 0.05$	$\mu_{\beta_1} = 0.1$	$\mu_{\beta_2} = 0.1$	$\mu_{\gamma_1} = 0.6$	$\mu_{\gamma_2} = 0.9$
$N = 100$	CLS	(0.0126) (0.0036)	(0.0058) (0.0028)	(0.0182) (0.0050)	(0.0009) (0.0048)	(−0.0183) (0.0033)	(−0.0672) (0.0091)
	WCLS	(0.0108) (0.0033)	(0.0009) (0.0026)	(0.0129) (0.0044)	(−0.0048) (0.0045)	(−0.0149) (0.0030)	(−0.0462) (0.0071)
	CML	(0.0026) (0.0008)	(0.0016) (0.0010)	(0.0009) (0.0018)	(−0.0024) (0.0022)	(−0.0036) (0.0010)	(−0.0247) (0.0012)
$N = 200$	CLS	(0.0123) (0.0015)	(0.0014) (0.0013)	(0.0189) (0.0040)	(−0.0030) (0.0032)	(−0.0187) (0.0020)	(−0.0450) (0.0048)
	WCLS	(0.0089) (0.0013)	(0.0004) (0.0013)	(0.0171) (0.0038)	(−0.0045) (0.0031)	(−0.0158) (0.0017)	(−0.0351) (0.0046)
	CML	(0.0035) (0.0005)	(0.0019) (0.0005)	(−0.0020) (0.0013)	(−0.0006) (0.0013)	(−0.0023) (0.0005)	(−0.0149) (0.0004)
$N = 500$	CLS	(0.0081) (0.0006)	(−0.0031) (0.0005)	(0.0090) (0.0012)	(0.0018) (0.0012)	(−0.0104) (0.0007)	(−0.0183) (0.0014)
	WCLS	(0.0068) (0.0006)	(−0.0039) (0.0005)	(0.0068) (0.0011)	(0.0014) (0.0012)	(−0.0086) (0.0006)	(−0.0126) (0.0016)
	CML	(0.0031) (0.0002)	(−0.0021) (0.0002)	(0.0021) (0.0005)	(0.0016) (0.0004)	(−0.0028) (0.0002)	(−0.0058) (0.0001)
$N = 1000$	CLS	(0.0009) (0.0002)	(0.0035) (0.0003)	(−0.0001) (0.0007)	(0.0014) (0.0005)	(−0.0021) (0.0003)	(−0.0120) (0.0006)
	WCLS	(0.0004) (0.0002)	(0.0039) (0.0004)	(−0.0010) (0.0006)	(0.0007) (0.0005)	(−0.0013) (0.0003)	(−0.0100) (0.0008)
	CML	(−0.0006) (−0.0001)	(0.0016) (0.0001)	(−0.0019) (0.0002)	(−0.0019) (0.0002)	(−0.0002) (0.0001)	(−0.0016) (0.0000)
Case B	$n = 7$	$\mu_{\alpha_1} = 0.4$	$\mu_{\alpha_2} = 0.1$	$\mu_{\beta_1} = 0.3$	$\mu_{\beta_2} = 0.5$	$\mu_{\gamma_1} = 0.6$	$\mu_{\gamma_2} = 0.9$
$N = 100$	CLS	(−0.0033) (−0.0044)	(0.0051) (0.0045)	(0.0195) (0.0098)	(0.0087) (0.0120)	(−0.0145) (0.0053)	(−0.0426) (0.0138)
	WCLS	(−0.0031) (−0.0042)	(0.0049) (0.0043)	(0.0192) (0.0097)	(0.0084) (0.0117)	(−0.0142) (0.0052)	(−0.0422) (0.0137)
	CML	(0.0016) (0.0039)	(0.0038) (0.0026)	(0.0147) (0.0071)	(0.0090) (0.0078)	(−0.0118) (0.0033)	(−0.0450) (0.0068)
$N = 200$	CLS	(0.0008) (0.0022)	(0.0010) (0.0021)	(0.0028) (0.0048)	(0.0020) (0.0050)	(−0.0060) (0.0030)	(−0.0197) (0.0071)
	WCLS	(0.0007) (0.0021)	(0.0009) (0.0020)	(0.0025) (0.0047)	(0.0018) (0.0048)	(−0.0059) (0.0029)	(−0.0195) (0.0069)
	CML	(0.0007) (0.0013)	(0.0004) (0.0012)	(0.0022) (0.0033)	(0.0021) (0.0033)	(−0.0058) (0.0015)	(−0.0216) (0.0028)
$N = 500$	CLS	(−0.0007) (−0.0009)	(0.0006) (0.0008)	(0.0026) (0.0020)	(0.0003) (0.0023)	(−0.0040) (0.0012)	(−0.0087) (0.0025)
	WCLS	(−0.0007) (−0.0009)	(0.0006) (0.0008)	(0.0025) (0.0020)	(0.0002) (0.0022)	(−0.0039) (0.0010)	(−0.0086) (0.0024)
	CML	(0.0006) (0.0005)	(−0.0002) (0.0004)	(0.0021) (0.0012)	(0.0007) (0.0013)	(−0.0038) (0.0006)	(−0.0094) (0.0009)
$N = 1000$	CLS	(−0.0005) (−0.0004)	(0.0006) (0.0004)	(0.0023) (0.0009)	(−0.0002) (0.0011)	(−0.0007) (0.0006)	(−0.0060) (0.0011)
	WCLS	(−0.0005) (−0.0004)	(0.0006) (0.0004)	(0.0022) (0.0009)	(−0.0001) (0.0010)	(−0.0007) (0.0006)	(−0.0059) (0.0009)
	CML	(0.0002) (0.0002)	(0.0002) (0.0002)	(0.0015) (0.0006)	(−0.0005) (−0.0006)	(−0.0001) (0.0004)	(−0.0065) (0.0004)

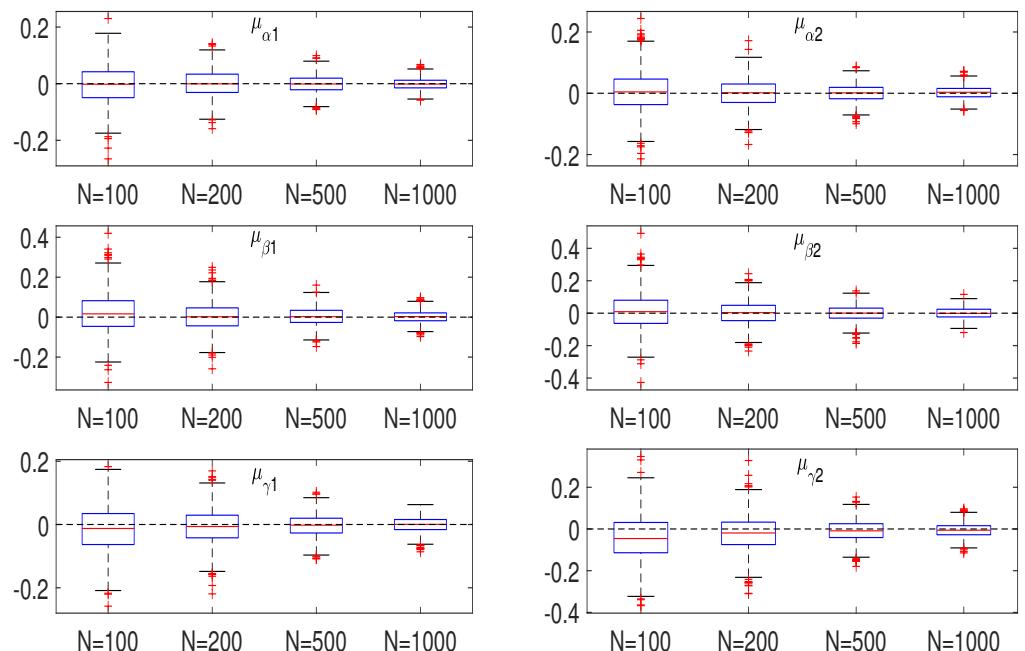
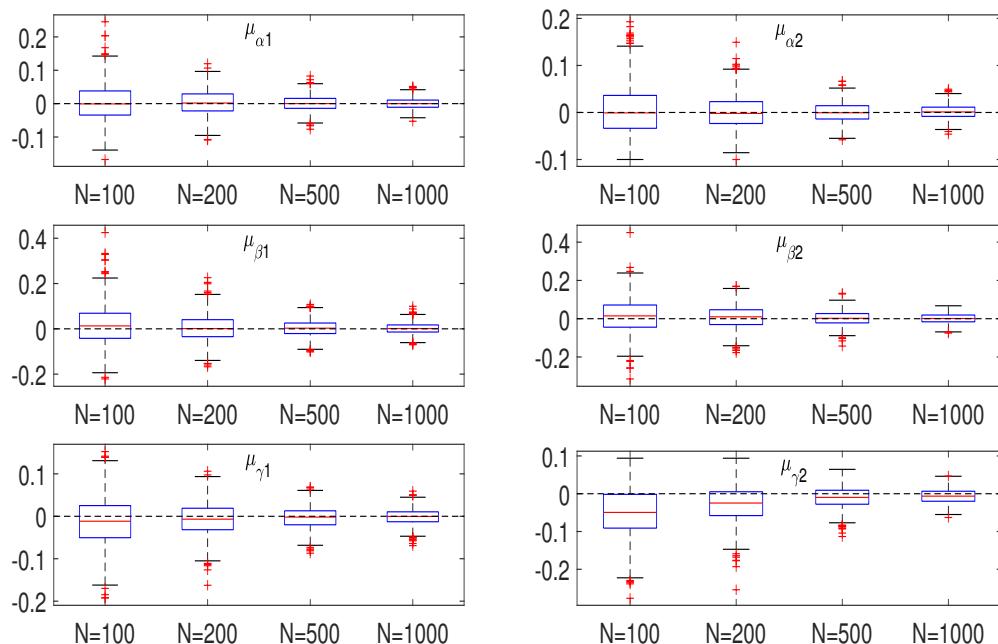
**Figure 2.** The bias of CLS estimators $(\hat{\mu}_{\alpha_1}^{CLS}, \hat{\mu}_{\alpha_2}^{CLS}, \hat{\mu}_{\beta_1}^{CLS}, \hat{\mu}_{\beta_2}^{CLS}, \hat{\mu}_{\gamma_1}^{CLS}, \hat{\mu}_{\gamma_2}^{CLS})^T$ for Case B with the sample size $N = 100, 200, 500$ and 1000 .

Table 2. Empirical bias and MSE of estimators of θ for cases A, B, C and D.

Case C	$n = 7$	$\mu_{\alpha_1} = 0.9$	$\mu_{\alpha_2} = 0.2$	$\mu_{\beta_1} = 0.3$	$\mu_{\beta_2} = 0.5$	$\mu_{\gamma_1} = 0.7$	$\mu_{\gamma_2} = 0.8$
$N = 100$	CLS	(-0.0066) (0.0005)	(0.0474) (0.0105)	(0.0220) (0.0253)	(-0.0057) (0.0261)	(0.0564) (0.0115)	(-0.0413) (0.1911)
	WCLS	(-0.0055) (0.0005)	(0.0342) (0.0088)	(0.0209) (0.0245)	(-0.0038) (0.0244)	(0.0474) (0.0111)	(-0.0256) (0.2247)
	CML	(-0.0004) (0.0003)	(0.0092) (0.0061)	(0.0123) (0.0124)	(-0.0045) (0.0166)	(0.0066) (0.0073)	(-0.0295) (0.0304)
$N = 200$	CLS	(-0.0039) (0.0002)	(0.0313) (0.0056)	(0.0094) (0.0128)	(-0.0028) (0.0143)	(0.0373) (0.0065)	(-0.0155) (0.0338)
	WCLS	(-0.0030) (0.0002)	(0.0219) (0.0047)	(0.0083) (0.0121)	(-0.0024) (0.0128)	(0.0298) (0.0063)	(0.0005) (0.0319)
	CML	(0.0001) (0.0001)	(0.0083) (0.0029)	(0.0039) (0.0057)	(-0.0016) (0.0079)	(0.0026) (0.0040)	(-0.0185) (0.0171)
$N = 500$	CLS	(-0.0013) (0.0001)	(0.0131) (0.0022)	(0.0018) (0.0048)	(-0.0042) (0.0048)	(0.0169) (0.0029)	(-0.0090) (0.0121)
	WCLS	(-0.0009) (0.0001)	(0.0082) (0.0018)	(0.0017) (0.0045)	(-0.0037) (0.0043)	(0.0129) (0.0027)	(-0.0007) (0.0115)
	CML	(0.0002) (0.0001)	(0.0017) (0.0011)	(0.0022) (0.0022)	(-0.0020) (0.0029)	(0.0017) (0.0018)	(-0.0045) (0.0073)
$N = 1000$	CLS	(-0.0007) (0.0000)	(0.0062) (0.0010)	(0.0028) (0.0025)	(-0.0001) (0.0024)	(0.0064) (0.0014)	(-0.0073) (0.0056)
	WCLS	(-0.0004) (0.0000)	(0.0038) (0.0008)	(0.0023) (0.0023)	(-0.0002) (0.0021)	(0.0037) (0.0013)	(-0.0041) (0.0056)
	CML	(0.0002) (0.0000)	(0.0006) (0.0006)	(0.0009) (0.0012)	(-0.0000) (0.0014)	(-0.0013) (0.0009)	(-0.0050) (0.0039)
Case D	$n = 7$	$\mu_{\alpha_1} = 0.9$	$\mu_{\alpha_2} = 0.2$	$\mu_{\beta_1} = 0.4$	$\mu_{\beta_2} = 0.7$	$\mu_{\gamma_1} = 0.1$	$\mu_{\gamma_2} = 0.8$
$N = 100$	CLS	(-0.0245) (0.0013)	(0.0901) (0.0257)	(-0.0028) (0.0110)	(0.0163) (0.0135)	(0.1689) (0.0588)	(-0.0099) (0.0106)
	WCLS	(-0.0181) (0.0011)	(0.0780) (0.0187)	(0.0024) (0.0101)	(0.0162) (0.0132)	(0.1385) (0.0376)	(-0.0026) (0.0095)
	CML	(-0.0026) (0.0004)	(0.0044) (0.0078)	(-0.0000) (0.0059)	(0.0055) (0.0088)	(0.0365) (0.0047)	(0.0035) (0.0055)
$N = 200$	CLS	(-0.0143) (0.0006)	(0.0694) (0.0133)	(-0.0006) (0.0059)	(0.0060) (0.0070)	(0.1265) (0.0213)	(-0.0035) (0.0050)
	WCLS	(-0.0103) (0.0005)	(0.0474) (0.0096)	(0.0010) (0.0053)	(0.0061) (0.0067)	(0.0958) (0.0138)	(0.0003) (0.0045)
	CML	(-0.0014) (0.0002)	(0.0005) (0.0038)	(-0.0011) (0.0030)	(0.0018) (0.0043)	(0.0191) (0.0021)	(0.0038) (0.0023)
$N = 500$	CLS	(-0.0052) (0.0002)	(0.0301) (0.0050)	(-0.0018) (0.0021)	(0.0018) (0.0025)	(0.0545) (0.0052)	(-0.0001) (0.0021)
	WCLS	(-0.0037) (0.0001)	(0.0202) (0.0037)	(-0.0004) (0.0019)	(0.0024) (0.0023)	(0.0412) (0.0037)	(0.0013) (0.0019)
	CML	(0.0001) (0.0001)	(0.0007) (0.0015)	(0.0001) (0.0011)	(-0.0004) (0.0016)	(0.0064) (0.0008)	(0.0019) (0.0010)
$N = 1000$	CLS	(-0.0030) (0.0001)	(0.0164) (0.0024)	(-0.0018) (0.0010)	(0.0016) (0.0012)	(0.0280) (0.0019)	(-0.0016) (0.0009)
	WCLS	(-0.0023) (0.0001)	(0.0112) (0.0017)	(-0.0010) (0.0009)	(0.0020) (0.0011)	(0.0209) (0.0015)	(-0.0010) (0.0009)
	CML	(-0.0003) (0.0000)	(0.0003) (0.0007)	(-0.0003) (0.0005)	(0.0007) (0.0007)	(0.0027) (0.0004)	(-0.0002) (0.0005)

**Figure 3.** The bias of WCLS estimators $(\hat{\mu}_{\alpha_1}^{WCLS}, \hat{\mu}_{\alpha_2}^{WCLS}, \hat{\mu}_{\beta_1}^{WCLS}, \hat{\mu}_{\beta_2}^{WCLS}, \hat{\mu}_{\gamma_1}^{WCLS}, \hat{\mu}_{\gamma_2}^{WCLS})^T$ for Case B with the sample size $N = 100, 200, 500$ and 1000 .

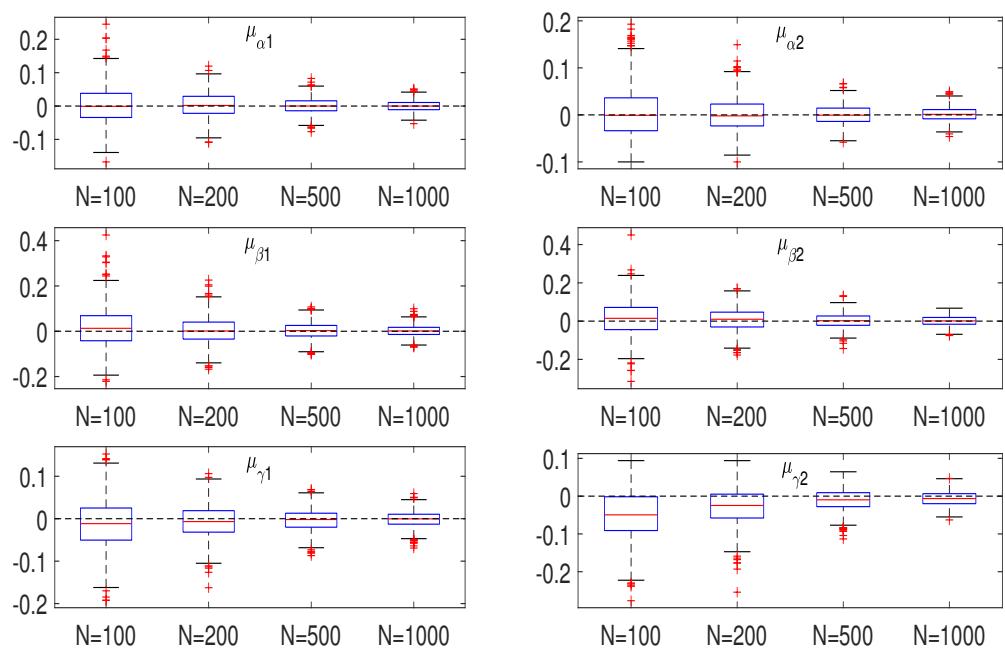


Figure 4. The bias of CML estimators $(\hat{\mu}_{\alpha_1}^{CML}, \hat{\mu}_{\alpha_2}^{CML}, \hat{\mu}_{\beta_1}^{CML}, \hat{\mu}_{\beta_2}^{CML}, \hat{\mu}_{\gamma_1}^{CML}, \hat{\mu}_{\gamma_2}^{CML})^T$ for Case B with the sample size $N = 100, 200, 500$ and 1000 .

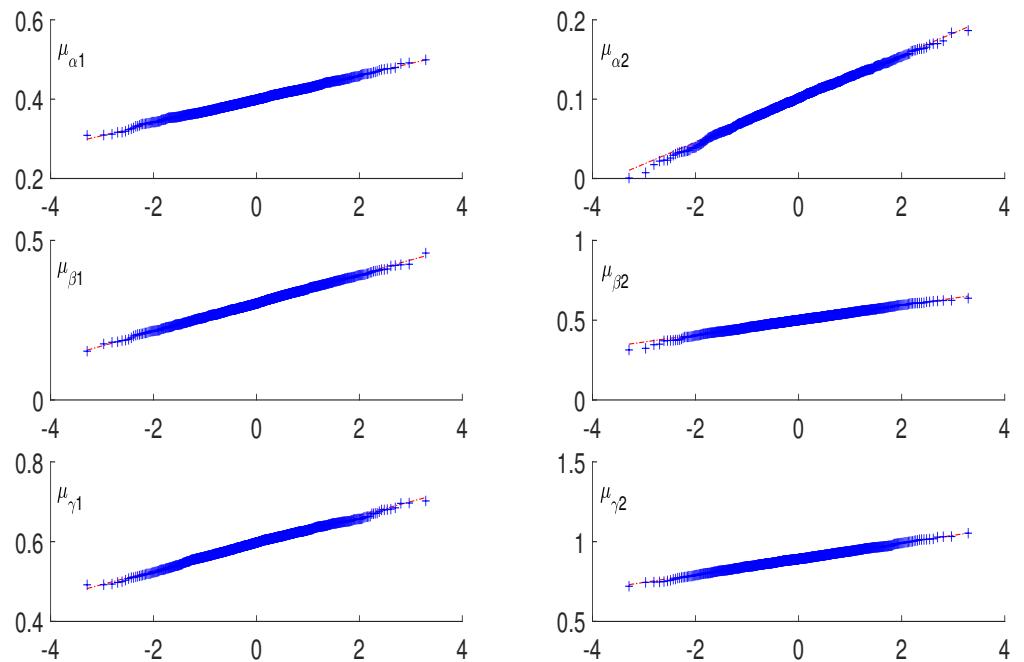


Figure 5. QQplots of CLS estimators $(\hat{\mu}_{\alpha_1}^{CLS}, \hat{\mu}_{\alpha_2}^{CLS}, \hat{\mu}_{\beta_1}^{CLS}, \hat{\mu}_{\beta_2}^{CLS}, \hat{\mu}_{\gamma_1}^{CLS}, \hat{\mu}_{\gamma_2}^{CLS})^T$ for Case B with the sample size equals to 500 .

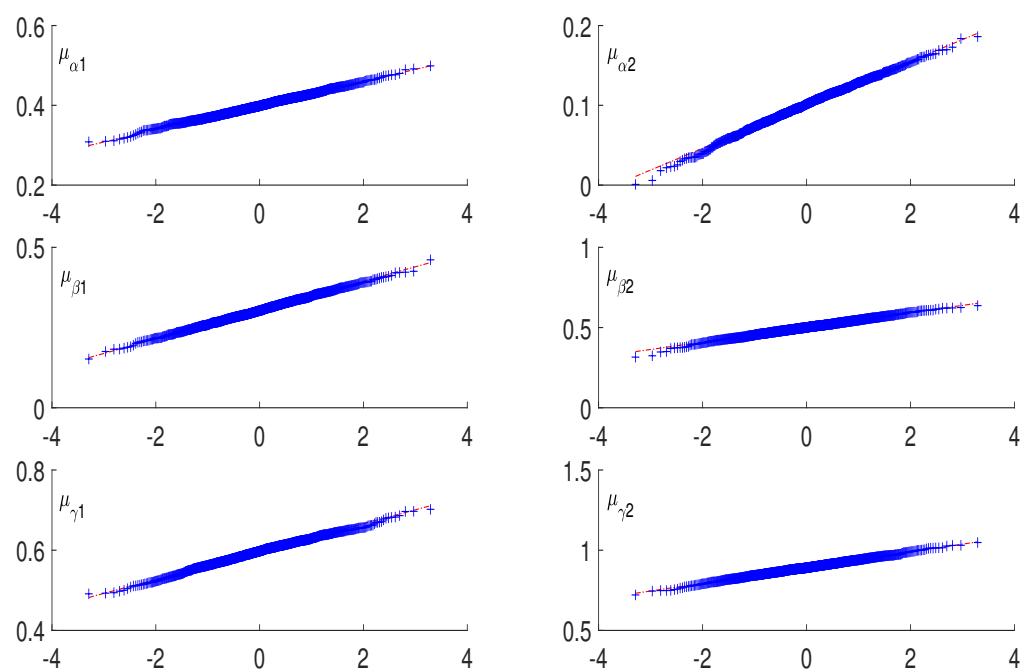


Figure 6. QQplots of WCLS estimators $(\hat{\mu}_{\alpha_1}^{WCLS}, \hat{\mu}_{\alpha_2}^{WCLS}, \hat{\mu}_{\beta_1}^{WCLS}, \hat{\mu}_{\beta_2}^{WCLS}, \hat{\mu}_{\gamma_1}^{WCLS}, \hat{\mu}_{\gamma_2}^{WCLS})^T$ for Case B with the sample size equals to 500.

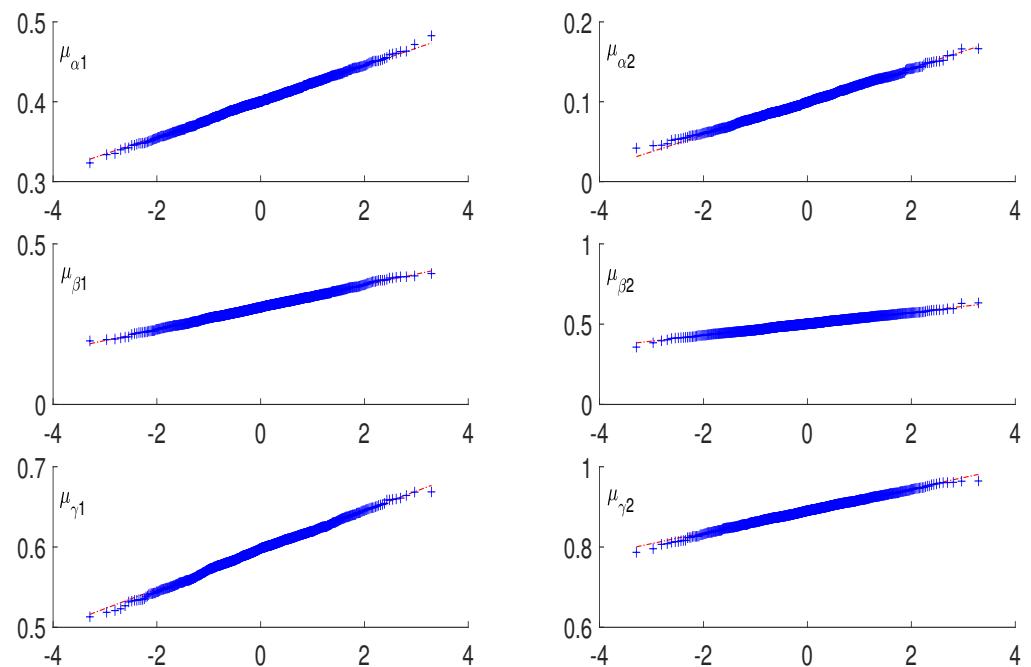


Figure 7. QQplots of CML estimators $(\hat{\mu}_{\alpha_1}^{CML}, \hat{\mu}_{\alpha_2}^{CML}, \hat{\mu}_{\beta_1}^{CML}, \hat{\mu}_{\beta_2}^{CML}, \hat{\mu}_{\gamma_1}^{CML}, \hat{\mu}_{\gamma_2}^{CML})^T$ for Case B with the sample size equals to 500.

5. Real Data Example

As an application, we use the F-RCMAR(1) model to fit a set of monthly incomes data. The data set represents the number of 34 employees with different levels of monthly incomes in a financial institution in China (Jan 2017–May 2021).

For each month t , X_{1t} represents the number of employees whose monthly incomes are between 5000 to 8000 (people at this level are required to pay 3% personal income tax), X_{2t} means the number of employees between 8000 to 17,000 (they need to pay 10%

personal income tax) and $n - x_{1t} - x_{2t}$ is the number of employees whose incomes are less than 5000 (they do not need to pay any tax). Therefore, the data set has finite range with fixed upper limit $n = 34$, and the series contains 53 observations. We plot the sample path and the ACF of the observations in Figure 8.

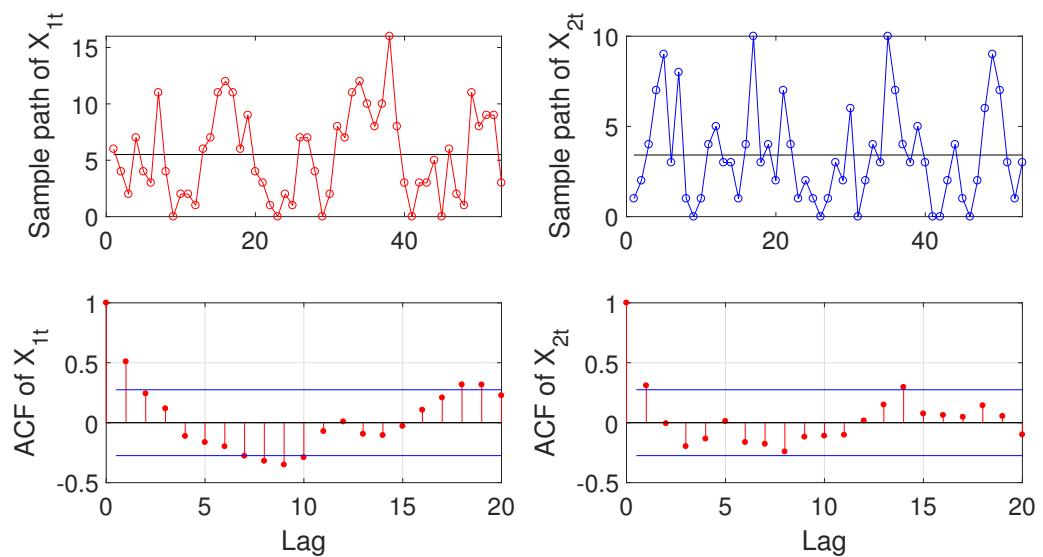


Figure 8. The sample paths and ACFs of the monthly incomes data from January 2017 to May 2021. The horizontal lines in sample paths are the mean of X_{1t} and X_{2t} . If the vertical lines in ACFs fall within two horizontal lines, it can be considered that the ACF is approximately 0.

According to the data set, the minimum of X_{1t} is 0, the median is 5, and the maximum is 16. The corresponding values of X_{2t} are 0, 3 and 10. In order to verify that the F-RCMAR(1) process can be better applied to the real data, we compare our model with F-MAR model in Ref. [15] and BVB_{II}-AR(1) model in Ref. [9]. We use the CML method to estimate the parameters of the fitted models.

The related results are listed in Table 3. From Table 3, it is clear that the F-RCMAR(1) model has the smaller Akaike information criterion (AIC) and Bayesian information criterion (BIC) than other two models, which means that our model can be well adapted to the data set with three states and possesses better fitting performance compared to other two models. Therefore, it is necessary and significant to study the extension of random coefficient situation.

Table 3. Estimates of the parameters and goodness-of-fit statistics for the monthly income dataset.

Model	Estimators	AIC	BIC
F-RCMAR(1)	$\hat{\mu}_{\alpha_1} = 0.4777, \hat{\mu}_{\alpha_2} = 0.4337$ $\hat{\mu}_{\beta_1} = 0.1367, \hat{\mu}_{\beta_2} = 0.3328$ $\hat{\mu}_{\gamma_1} = 0.1609, \hat{\mu}_{\gamma_2} = 0.1211$	560.4151	572.2369
F-MAR(1)	$\hat{\alpha}_1 = 0.4141, \hat{\alpha}_2 = 0.2207$ $\hat{\beta}_1 = 0.1520, \hat{\beta}_2 = 0.3246$ $\hat{\gamma}_1 = 0.1070, \hat{\gamma}_2 = 0.0806$	603.8675	615.6892
BVB _{II} -AR(1)	$\hat{p} = 0.4347, \hat{q} = 0.3025$ $\hat{\beta}_1 = 0.2603, \hat{\beta}_2 = 0.1084$ $\hat{\phi}_\alpha = 0.0802, \hat{\phi}_\beta = 0.0175$	581.4334	593.2552

6. Conclusions

This article introduced a finite-range random coefficient multinomial autoregressive process. The stationarity and ergodicity of the process were established. Some probabilistic and statistical properties of the new model were explored. CLS-estimators and WCLS-

estimators of the model parameters were derived, and the related asymptotic properties were obtained.

In simulation studies, we compared these two methods with CML estimators. The results showed that it is better to use the CML method to estimate the parameters θ because of its great effect and robustness. Comparing with F-MAR model and BVB_{II}-AR(1) model in the literature, the brilliant performance of the new model was illustrated by an application to the number of employees at different levels of monthly incomes. We can conclude that the F-RCMAR(1) model is a significant process and a necessary extension of the F-MAR process.

Author Contributions: Conceptualization, J.Z.; methodology, D.W. and X.D.; software, J.Z. and K.Y.; validation, J.Z., D.W. and X.D.; formal analysis, J.Z.; investigation, J.Z.; resources, D.W., X.D. and K.Y.; data curation, K.Y.; writing—original draft preparation, J.Z.; writing—review and editing, D.W.; visualization, J.Z. and K.Y.; supervision, D.W. and X.D.; project administration, D.W., X.D. and K.Y.; funding acquisition, D.W., X.D. and K.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (No. 11871028, 11731015, 11901053), Program for Changbaishan Scholars of Jilin Province (2015010).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Proposition 1. Based on the definition of F-RCMAR(1) process, we have:

$$\begin{aligned}
 & P(X_{1t} = x_{1t}, X_{2t} = x_{2t} | X_{1,t-1} = x_{1,t-1}, X_{2,t-1} = x_{2,t-1}) \\
 &= P(\alpha_{1t} \circ X_{1,t-1} + \beta_{1t} \circ X_{2,t-1} + \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}) = x_{1t}, \\
 &\quad \alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) + \beta_{2t} \circ (X_{2,t-1} - \beta_{1t} \circ X_{2,t-1}) \\
 &\quad + \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1})) = x_{2t} | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \\
 &= \sum_{k=\max\{0, x_{1t}+x_{1,t-1}+x_{2,t-1}-n\}}^{\min\{x_{1t}, x_{1,t-1}+x_{2,t-1}\}} P(\alpha_{1t} \circ X_{1,t-1} + \beta_{1t} \circ X_{2,t-1} = k | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \\
 &\quad \cdot P(\gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}) = x_{1t} - k | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \\
 &\quad \cdot P(\alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) + \beta_{2t} \circ (X_{2,t-1} - \beta_{1t} \circ X_{2,t-1}) \\
 &\quad + \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - x_{1t} + k) = x_{2t} | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \\
 &= \sum_{k=\max\{0, x_{1t}+x_{1,t-1}+x_{2,t-1}-n\}}^{\min\{x_{1t}, x_{1,t-1}+x_{2,t-1}\}} \sum_{j_1=\max\{0, k-x_{2,t-1}\}}^{\min\{x_{1,t-1}, k\}} P(\alpha_{1t} \circ X_{1,t-1} = j_1 | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \\
 &\quad \cdot P(\beta_{1t} \circ X_{2,t-1} = k - j_1 | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \cdot P(\gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}) = \\
 &\quad x_{1t} - k | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \cdot P(\alpha_{2t} \circ (X_{1,t-1} - j_1) + \beta_{2t} \circ (X_{2,t-1} - k + j_1) + \\
 &\quad \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - x_{1t} + k) = x_{2t} | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \\
 &= \sum_{k=\max\{0, x_{1t}+x_{1,t-1}+x_{2,t-1}-n\}}^{\min\{x_{1t}, x_{1,t-1}+x_{2,t-1}\}} \sum_{j_1=\max\{0, k-x_{2,t-1}\}}^{\min\{x_{1,t-1}, k\}} \left\{ P(\alpha_{1t} \circ X_{1,t-1} = j_1 | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \right. \\
 &\quad \cdot P(\beta_{1t} \circ X_{2,t-1} = k - j_1 | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \cdot P(\gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}) = \\
 &\quad x_{1t} - k | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \cdot \sum_{s=\max\{0, x_{1t}+x_{2t}+x_{1,t-1}+x_{2,t-1}-k-n\}}^{\min\{x_{2t}, x_{1,t-1}+x_{2,t-1}-k\}} \left[P(\alpha_{2t} \circ (X_{1,t-1} - j_1) + \right. \\
 &\quad \left. \beta_{2t} \circ (X_{2,t-1} - k + j_1) = s | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \cdot P(\gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - x_{1t} + k) = \right. \\
 &\quad \left. x_{2t} - s | \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \right] \left. \right\}
 \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=\max\{0, x_{1t}+x_{1,t-1}+x_{2,t-1}-n\}}^{\min\{x_{1t}, x_{1,t-1}+x_{2,t-1}\}} \sum_{j_1=\max\{0, k-x_{2,t-1}\}}^{\min\{x_{1,t-1}, k\}} \left\{ P(\alpha_{1t} \circ X_{1,t-1} = j_1 \mid \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \right. \\
&\quad \cdot P(\beta_{1t} \circ X_{2,t-1} = k - j_1 \mid \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \cdot P(\gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}) = \\
&\quad x_{1t} - k \mid \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \cdot \sum_{s=\max\{0, x_{1t}+x_{2t}+x_{1,t-1}+x_{2,t-1}-k-n\}}^{\min\{x_{2t}, x_{1,t-1}+x_{2,t-1}-k\}} \sum_{j_2=\max\{0, s+j_1-x_{1,t-1}\}}^{\min\{s, x_{2,t-1}-k+j_1\}} \\
&\quad \left[P(\alpha_{2t} \circ (X_{1,t-1} - j_1) = s - j_2 \mid \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \cdot P(\beta_{2t} \circ (X_{2,t-1} - k + j_1) = \right. \\
&\quad j_2 \mid \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \cdot P(\gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - x_{1t} + k) = x_{2t} - s \mid \mathbf{X}_{t-1} = \mathbf{x}_{t-1}) \left. \right] \Bigg\} \\
&= \sum_{k=\max\{0, x_{1t}+x_{1,t-1}+x_{2,t-1}-n\}}^{\min\{x_{1t}, x_{1,t-1}+x_{2,t-1}\}} \sum_{j_1=\max\{0, k-x_{2,t-1}\}}^{\min\{x_{1,t-1}, k\}} \\
&\quad \left\{ \left[\int_0^1 C_{x_{1,t-1}}^{j_1} \alpha_{1t}^{j_1} (1 - \alpha_{1t})^{x_{1,t-1}-j_1} dP_{\alpha_{1t}} \cdot \int_0^1 C_{x_{2,t-1}}^{k-j_1} \beta_{1t}^{k-j_1} (1 - \beta_{1t})^{x_{2,t-1}-k+j_1} dP_{\beta_{1t}} \right. \right. \\
&\quad \cdot \int_0^1 C_{n-x_{1,t-1}-x_{2,t-1}}^{x_{1t}-k} \gamma_{1t}^{x_{1t}-k} (1 - \gamma_{1t})^{n-x_{1,t-1}-x_{2,t-1}-x_{1t}+k} dP_{\gamma_{1t}} \Big] \cdot \\
&\quad \sum_{s=\max\{0, x_{1t}+x_{2t}+x_{1,t-1}+x_{2,t-1}-k-n\}}^{\min\{x_{2t}, x_{1,t-1}+x_{2,t-1}-k\}} \sum_{j_2=\max\{0, s+j_1-x_{1,t-1}\}}^{\min\{s, x_{2,t-1}-k+j_1\}} \\
&\quad \left. \left[\int_0^1 C_{x_{1,t-1}-j_1}^{s-j_2} \alpha_{2t}^{s-j_2} (1 - \alpha_{2t})^{x_{1,t-1}-j_1-s+j_2} dP_{\alpha_{2t}} \right. \right. \\
&\quad \cdot \int_0^1 C_{x_{2,t-1}-k+j_1}^{j_2} \beta_{2t}^{j_2} (1 - \beta_{2t})^{x_{2,t-1}-k+j_1-j_2} dP_{\beta_{2t}} \\
&\quad \left. \left. \cdot \int_0^1 C_{n-x_{1,t-1}-x_{2,t-1}-x_{1t}+k}^{x_{2t}-s} \gamma_{2t}^{x_{2t}-s} (1 - \gamma_{2t})^{n-x_{1,t-1}-x_{2,t-1}-x_{1t}+k-x_{2t}+s} dP_{\gamma_{2t}} \right] \right\}.
\end{aligned}$$

Therefore, Proposition 1 is established. \square

Appendix B

Proof of Proposition 2. Based on Definition 1, we have

$$\begin{aligned}
E(X_{1t} \mid \mathbf{X}_{t-1}) &= E(\alpha_{1t} \circ X_{1,t-1} + \beta_{1t} \circ X_{2,t-1} + \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}) \mid \mathbf{X}_{t-1}) \\
&= E(\alpha_{1t} \circ X_{1,t-1} \mid \mathbf{X}_{t-1}) + E(\beta_{1t} \circ X_{2,t-1} \mid \mathbf{X}_{t-1}) + E(\gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}) \mid \mathbf{X}_{t-1}) \\
&= E\left(\sum_{i=1}^{X_{1,t-1}} Y_i^{(\alpha_{1t})} \mid \mathbf{X}_{t-1}\right) + E\left(\sum_{i=1}^{X_{2,t-1}} Y_i^{(\beta_{1t})} \mid \mathbf{X}_{t-1}\right) + E\left(\sum_{i=1}^{n-X_{1,t-1}-X_{2,t-1}} Y_i^{(\gamma_{1t})} \mid \mathbf{X}_{t-1}\right),
\end{aligned}$$

where $\{Y_i^{(\alpha_{1t})}\}$ is a sequence of i.i.d. Bernoulli random variables with distribution $B(1, \alpha_{1t})$. Other $\{Y_i\}$ follows the similar distributions. Therefore, the above equation is equal to:

$$\begin{aligned}
&= X_{1,t-1} E(Y_i^{(\alpha_{1t})}) + X_{2,t-1} E(Y_i^{(\beta_{1t})}) + (n - X_{1,t-1} - X_{2,t-1}) E(Y_i^{(\gamma_{1t})}) \\
&= X_{1,t-1} E[E(Y_i^{(\alpha_{1t})} \mid \alpha_{1t})] + X_{2,t-1} E[E(Y_i^{(\beta_{1t})} \mid \beta_{1t})] \\
&\quad + (n - X_{1,t-1} - X_{2,t-1}) E[E(Y_i^{(\gamma_{1t})} \mid \gamma_{1t})] \\
&= X_{1,t-1} \mu_{\alpha_1} + X_{2,t-1} \mu_{\beta_1} + (n - X_{1,t-1} - X_{2,t-1}) \mu_{\gamma_1} \\
&= n \mu_{\gamma_1} + (\mu_{\alpha_1} - \mu_{\gamma_1}) X_{1,t-1} + (\mu_{\beta_1} - \mu_{\gamma_1}) X_{2,t-1}.
\end{aligned}$$

Moreover, in order to solve $Var(X_{1t} | \mathbf{X}_{t-1})$, we need to calculate $Var(X_{1t} | \mathbf{X}_{t-1}, \alpha_{1t}, \beta_{1t}, \gamma_{1t})$ first.

$$\begin{aligned}
& Var(X_{1t} | \mathbf{X}_{t-1}, \alpha_{1t}, \beta_{1t}, \gamma_{1t}) \\
&= Var(\alpha_{1t} \circ X_{1,t-1} + \beta_{1t} \circ X_{2,t-1} + \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}) | \mathbf{X}_{t-1}, \alpha_{1t}, \beta_{1t}, \gamma_{1t}) \\
&= Var(\alpha_{1t} \circ X_{1,t-1} | \mathbf{X}_{t-1}, \alpha_{1t}, \beta_{1t}, \gamma_{1t}) + Var(\beta_{1t} \circ X_{2,t-1} | \mathbf{X}_{t-1}, \alpha_{1t}, \beta_{1t}, \gamma_{1t}) \\
&\quad + Var(\gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}) | \mathbf{X}_{t-1}, \alpha_{1t}, \beta_{1t}, \gamma_{1t}) \\
&= Var\left(\sum_{i=1}^{X_{1,t-1}} Y_i^{(\alpha_{1t})} | \mathbf{X}_{t-1}, \alpha_{1t}\right) + Var\left(\sum_{i=1}^{X_{2,t-1}} Y_i^{(\beta_{1t})} | \mathbf{X}_{t-1}, \beta_{1t}\right) \\
&\quad + Var\left(\sum_{i=1}^{n-X_{1,t-1}-X_{2,t-1}} Y_i^{(\gamma_{1t})} | \mathbf{X}_{t-1}, \gamma_{1t}\right) \\
&= X_{1,t-1}\alpha_{1t}(1 - \alpha_{1t}) + X_{2,t-1}\beta_{1t}(1 - \beta_{1t}) + (n - X_{1,t-1} - X_{2,t-1})\gamma_{1t}(1 - \gamma_{1t}),
\end{aligned}$$

Therefore, we have:

$$\begin{aligned}
Var(X_{1t} | \mathbf{X}_{t-1}) &= E[Var(X_{1t} | \mathbf{X}_{t-1}, \alpha_{1t}, \beta_{1t}, \gamma_{1t})] + Var[E(X_{1t} | \mathbf{X}_{t-1}, \alpha_{1t}, \beta_{1t}, \gamma_{1t})] \\
&= E_{\alpha_{1t}, \beta_{1t}, \gamma_{1t}}[X_{1,t-1}\alpha_{1t}(1 - \alpha_{1t}) + X_{2,t-1}\beta_{1t}(1 - \beta_{1t}) \\
&\quad + (n - X_{1,t-1} - X_{2,t-1})\gamma_{1t}(1 - \gamma_{1t})] \\
&\quad + Var_{\alpha_{1t}, \beta_{1t}, \gamma_{1t}}[\alpha_{1t}X_{1,t-1} + \beta_{1t}X_{2,t-1} + \gamma_{1t}(n - X_{1,t-1} - X_{2,t-1})] \\
&= X_{1,t-1}[E(\alpha_{1t}) - E(\alpha_{1t}^2)] + X_{2,t-1}[E(\beta_{1t}) - E(\beta_{1t}^2)] \\
&\quad + (n - X_{1,t-1} - X_{2,t-1})[E(\gamma_{1t}) - E(\gamma_{1t}^2)] + X_{1,t-1}^2 Var(\alpha_{1t}) \\
&\quad + X_{2,t-1}^2 Var(\beta_{1t}) + (n - X_{1,t-1} - X_{2,t-1})^2 Var(\gamma_{1t}) \\
&= X_{1,t-1}(\mu_{\alpha_1} - \sigma_{\alpha_1}^2 - \mu_{\alpha_1}^2) + X_{2,t-1}(\mu_{\beta_1} - \sigma_{\beta_1}^2 - \mu_{\beta_1}^2) \\
&\quad + (n - X_{1,t-1} - X_{2,t-1})(\mu_{\gamma_1} - \sigma_{\gamma_1}^2 - \mu_{\gamma_1}^2) + X_{1,t-1}^2 \sigma_{\alpha_1}^2 + X_{2,t-1}^2 \sigma_{\beta_1}^2 \\
&\quad + (n - X_{1,t-1} - X_{2,t-1})^2 \sigma_{\gamma_1}^2.
\end{aligned}$$

For $X_{2t} = \alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) + \beta_{2t} \circ (X_{2,t-1} - \beta_{1t} \circ X_{2,t-1}) + \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}))$, as shown in Ref. [16], we can draw a conclusion as: $\alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) \stackrel{d}{=} \alpha_{2t}(1 - \alpha_{1t}) \circ X_{1,t-1}$. Thus, we can find

$$\begin{aligned}
& E(\alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) | \mathbf{X}_{t-1}) \\
&= E(\alpha_{2t}(1 - \alpha_{1t}) \circ X_{1,t-1} | \mathbf{X}_{t-1}) \\
&= E\left(\sum_{i=1}^{X_{1,t-1}} Y_i^{(\alpha_{2t}(1-\alpha_{1t}))} | \mathbf{X}_{t-1}\right) = X_{1,t-1} E\left(Y_i^{(\alpha_{2t}(1-\alpha_{1t}))}\right) \\
&= X_{1,t-1} E\left[E\left(Y_i^{(\alpha_{2t}(1-\alpha_{1t}))} | \alpha_{1t}, \alpha_{2t}\right)\right] = X_{1,t-1} \mu_{\alpha_2}(1 - \mu_{\alpha_1});
\end{aligned}$$

Similarly, $E(X_{2t} | \mathbf{X}_{t-1})$ can be obtained as:

$$\begin{aligned}
E(X_{2t} | \mathbf{X}_{t-1}) &= E[\alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) + \beta_{2t} \circ (X_{2,t-1} - \beta_{1t} \circ X_{2,t-1}) \\
&\quad + \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1})) | \mathbf{X}_{t-1}] \\
&= E[\alpha_{2t}(1 - \alpha_{1t}) \circ X_{1,t-1} + \beta_{2t}(1 - \beta_{1t}) \circ X_{2,t-1} | \mathbf{X}_{t-1}] \\
&\quad + E[\gamma_{2t}(1 - \gamma_{1t}) \circ (n - X_{1,t-1} - X_{2,t-1}) | \mathbf{X}_{t-1}] \\
&= X_{1,t-1}\mu_{\alpha_2}(1 - \mu_{\alpha_1}) + X_{2,t-1}\mu_{\beta_2}(1 - \mu_{\beta_1}) \\
&\quad + (n - X_{1,t-1} - X_{2,t-1})\mu_{\gamma_2}(1 - \mu_{\gamma_1}) = n\lambda + (\pi - \lambda)X_{1,t-1} + (\rho - \lambda)X_{2,t-1},
\end{aligned}$$

where $\pi = \mu_{\alpha_2}(1 - \mu_{\alpha_1})$, $\rho = \mu_{\beta_2}(1 - \mu_{\beta_1})$, $\lambda = \mu_{\gamma_2}(1 - \mu_{\gamma_1})$.

$$\begin{aligned} Var[\alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) \mid \mathbf{X}_{t-1}, \alpha_{2t}] &= Var[\alpha_{2t}(1 - \alpha_{1t}) \circ X_{1,t-1} \mid \mathbf{X}_{t-1}, \alpha_{2t}] \\ &= Var[\sum_{i=1}^{X_{1,t-1}} Y_i^{(\alpha_{2t}(1 - \alpha_{1t}))} \mid \mathbf{X}_{t-1}, \alpha_{2t}] = X_{1,t-1} Var[Y_i^{(\alpha_{2t}(1 - \alpha_{1t}))} \mid \alpha_{2t}] \\ &= X_{1,t-1} \{E_{\alpha_{1t}}[Var(Y_i^{(\alpha_{2t}(1 - \alpha_{1t}))} \mid \alpha_{2t}, \alpha_{1t})] + Var_{\alpha_{1t}}[E(Y_i^{(\alpha_{2t}(1 - \alpha_{1t}))} \mid \alpha_{2t}, \alpha_{1t})]\} \\ &= X_{1,t-1} \{E_{\alpha_{1t}}[(\alpha_{2t}(1 - \alpha_{1t})(1 - \alpha_{2t}(1 - \alpha_{1t})))] + Var_{\alpha_{1t}}[\alpha_{2t}(1 - \alpha_{1t})]\} \\ &= X_{1,t-1} \{\alpha_{2t} - \alpha_{2t}^2 + 2\mu_{\alpha_1}\alpha_{2t}^2 - \mu_{\alpha_1}\alpha_{2t} - (\sigma_{\alpha_1}^2 + \mu_{\alpha_1}^2)\alpha_{2t}^2 + \alpha_{2t}^2\sigma_{\alpha_1}^2\} \\ &= X_{1,t-1} [\alpha_{2t} - \alpha_{2t}^2 + 2\mu_{\alpha_1}\alpha_{2t}^2 - \mu_{\alpha_1}\alpha_{2t} - \mu_{\alpha_1}^2\alpha_{2t}^2], \end{aligned}$$

and then we have

$$\begin{aligned} Var[\alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) \mid \mathbf{X}_{t-1}] &= E_{\alpha_{2t}} \{Var[\alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) \mid \mathbf{X}_{t-1}, \alpha_{2t}]\} \\ &\quad + Var_{\alpha_{2t}} \{E[\alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) \mid \mathbf{X}_{t-1}, \alpha_{2t}]\} \\ &= E_{\alpha_{2t}} \{X_{1,t-1} [\alpha_{2t} - \alpha_{2t}^2 + 2\mu_{\alpha_1}\alpha_{2t}^2 - \mu_{\alpha_1}\alpha_{2t} - \mu_{\alpha_1}^2\alpha_{2t}^2]\} + Var_{\alpha_{2t}}[X_{1,t-1}(1 - \mu_{\alpha_1})\alpha_{2t}] \\ &= X_{1,t-1} [(1 - \mu_{\alpha_1})\mu_{\alpha_2} - (1 - \mu_{\alpha_1})^2(\mu_{\alpha_2}^2 + \sigma_{\alpha_2}^2)] + X_{1,t-1}^2(1 - \mu_{\alpha_1})^2\sigma_{\alpha_2}^2, \end{aligned}$$

and

$$\begin{aligned} Var(X_{2t} \mid \mathbf{X}_{t-1}) &= Var[\alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) \mid \mathbf{X}_{t-1}] \\ &\quad + Var[\beta_{2t} \circ (X_{2,t-1} - \beta_{1t} \circ X_{2,t-1}) \mid \mathbf{X}_{t-1}] \\ &\quad + Var[\gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1})) \mid \mathbf{X}_{t-1}] \\ &= \{X_{1,t-1} [(1 - \mu_{\alpha_1})\mu_{\alpha_2} - (1 - \mu_{\alpha_1})^2(\mu_{\alpha_2}^2 + \sigma_{\alpha_2}^2)] + X_{1,t-1}^2(1 - \mu_{\alpha_1})^2\sigma_{\alpha_2}^2\} \\ &\quad + \{X_{2,t-1} [(1 - \mu_{\beta_1})\mu_{\beta_2} - (1 - \mu_{\beta_1})^2(\mu_{\beta_2}^2 + \sigma_{\beta_2}^2)] + X_{2,t-1}^2(1 - \mu_{\beta_1})^2\sigma_{\beta_2}^2\} \\ &\quad + \{(n - X_{1,t-1} - X_{2,t-1}) [(1 - \mu_{\gamma_1})\mu_{\gamma_2} - (1 - \mu_{\gamma_1})^2(\mu_{\gamma_2}^2 + \sigma_{\gamma_2}^2)] \\ &\quad + (n - X_{1,t-1} - X_{2,t-1})^2(1 - \mu_{\gamma_1})^2\sigma_{\gamma_2}^2\}. \end{aligned}$$

Therefore, based on the equations: $E(X) = E_Y(E(X \mid Y))$ and $Var(X) = Var_Y(E(X \mid Y)) + E_Y(Var(X \mid Y))$, the properties (3) and (4) in Proposition 2 can be derived. $Cov(X_{1t}, X_{2t})$ can be solved by the following equations:

$$\begin{aligned} Cov(X_{1t}, X_{2t}) &= Cov \left(\alpha_{1t} \circ X_{1,t-1} + \beta_{1t} \circ X_{2,t-1} + \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}), \right. \\ &\quad \left. \alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1}) + \beta_{2t} \circ (X_{2,t-1} - \beta_{1t} \circ X_{2,t-1}) \right. \\ &\quad \left. + \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1})) \right) \\ &= Cov[\alpha_{1t} \circ X_{1,t-1}, \alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1})] \\ &\quad + Cov[\alpha_{1t} \circ X_{1,t-1}, \beta_{2t} \circ (X_{2,t-1} - \beta_{1t} \circ X_{2,t-1})] \\ &\quad + Cov[\alpha_{1t} \circ X_{1,t-1}, \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}))] \\ &\quad + Cov[\beta_{1t} \circ X_{2,t-1}, \alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1})] \\ &\quad + Cov[\beta_{1t} \circ X_{2,t-1}, \beta_{2t} \circ (X_{2,t-1} - \beta_{1t} \circ X_{2,t-1})] \\ &\quad + Cov[\beta_{1t} \circ X_{2,t-1}, \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}))] \\ &\quad + Cov[\gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}), \alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1})] \\ &\quad + Cov[\gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}), \beta_{2t} \circ (X_{2,t-1} - \beta_{1t} \circ X_{2,t-1})] \\ &\quad + Cov[\gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}), \gamma_{2t} \circ (n - X_{1,t-1} - X_{2,t-1} - \gamma_{1t} \circ (n - X_{1,t-1} - X_{2,t-1}))], \end{aligned}$$

in this place,

$$\begin{aligned} \text{Cov}[\alpha_{1t} \circ X_{1,t-1}, \alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1})] &= \text{Cov}[\alpha_{1t} \circ X_{1,t-1}, \alpha_{2t}(1 - \alpha_{1t}) \circ X_{1,t-1}] \\ &= E[\alpha_{1t} \circ X_{1,t-1} \cdot \alpha_{2t}(1 - \alpha_{1t}) \circ X_{1,t-1}] - E[\alpha_{1t} \circ X_{1,t-1}]E[\alpha_{2t}(1 - \alpha_{1t}) \circ X_{1,t-1}] \\ &= E\left[\left(\sum_{i=1}^{X_{1,t-1}} W_i^{(\alpha_{1t})}\right) \cdot \left(\sum_{j=1}^{X_{1,t-1}} W_j^{(\alpha_{2t}(1-\alpha_{1t}))}\right)\right] - E\left[\sum_{i=1}^{X_{1,t-1}} Y_i^{(\alpha_{1t})}\right]E\left[\sum_{j=1}^{X_{1,t-1}} Y_j^{(\alpha_{2t}(1-\alpha_{1t}))}\right]. \end{aligned}$$

One known conclusion is that, when $\{X_i\}$ is a sequence of i.i.d. random variables, random variable N can only take positive integer value and is independent of $\{X_i\}$. Then, we can find:

$$E\left(\sum_{i=1}^N X_i\right) = E(X_1)E(N).$$

Based on these, we have:

$$\begin{aligned} \text{Cov}[\alpha_{1t} \circ X_{1,t-1}, \alpha_{2t} \circ (X_{1,t-1} - \alpha_{1t} \circ X_{1,t-1})] \\ &= E(X_{1,t-1}^2)E(W_1^{(\alpha_{1t})} \cdot W_1^{(\alpha_{2t}(1-\alpha_{1t}))}) - E^2(X_{1,t-1})E(Y_1^{(\alpha_{1t})})E(Y_1^{(\alpha_{2t}(1-\alpha_{1t}))}) \\ &= \mu_{\alpha_1}\mu_{\alpha_2}(1 - \mu_{\alpha_1})\text{Var}(X_{1,t-1}), \end{aligned}$$

and the $\text{Cov}(X_{1t}, X_{2t})$ can be obtained as shown in property (5) in Proposition 2.

The deriving procedure of other properties is similar to these, which is not repeated any further. \square

Appendix C

Proof of Theorem 2 and Theorem 3. According to Equation (5), let

$$\begin{aligned} M_{N1} &= -\frac{1}{2} \frac{\partial S(\boldsymbol{\theta})}{\partial \mu_{\alpha_1}} \times (1 - \mu_{\alpha_1}) + \left(-\frac{1}{2}\right) \frac{\partial S(\boldsymbol{\theta})}{\partial \mu_{\alpha_2}} \times \mu_{\alpha_2} \\ &= \sum_{t=1}^N [X_{1t} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{1,t-1} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{2,t-1}] (1 - \mu_{\alpha_1})X_{1,t-1}, \quad (\text{A1}) \end{aligned}$$

$$\begin{aligned} M_{N2} &= -\frac{1}{2} \frac{\partial S(\boldsymbol{\theta})}{\partial \mu_{\beta_1}} \times (1 - \mu_{\beta_1}) + \left(-\frac{1}{2}\right) \frac{\partial S(\boldsymbol{\theta})}{\partial \mu_{\beta_2}} \times \mu_{\beta_2} \\ &= \sum_{t=1}^N [X_{1t} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{1,t-1} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{2,t-1}] (1 - \mu_{\beta_1})X_{2,t-1}, \quad (\text{A2}) \end{aligned}$$

$$\begin{aligned} M_{N3} &= -\frac{1}{2} \frac{\partial S(\boldsymbol{\theta})}{\partial \mu_{\gamma_1}} \times (1 - \mu_{\gamma_1}) + \left(-\frac{1}{2}\right) \frac{\partial S(\boldsymbol{\theta})}{\partial \mu_{\gamma_2}} \times \mu_{\gamma_2} \\ &= \sum_{t=1}^N [X_{1t} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{1,t-1} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{2,t-1}] \\ &\quad \cdot (1 - \mu_{\gamma_1})(n - X_{1,t-1} - X_{2,t-1}). \quad (\text{A3}) \end{aligned}$$

Solving the above functions we can find the CLS estimators of $\boldsymbol{\theta}_1 = (\mu_{\alpha_1}, \mu_{\beta_1}, \mu_{\gamma_1})^\top$ as Equation (6).

Let $\mathcal{F}_N = \sigma\{\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_N\}$ be the σ -filed generated by $\{\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_N\}$, Then, we have

$$\begin{aligned} E(M_{N1} | \mathcal{F}_{N-1}) &= E(M_{(N-1)1} + [X_{1N} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{1,N-1} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{2,N-1}] \\ &\quad (1 - \mu_{\alpha_1})X_{1,N-1} | \mathcal{F}_{N-1}) \\ &= M_{(N-1)1} + E([X_{1N} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{1,N-1} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{2,N-1}] \\ &\quad (1 - \mu_{\alpha_1})X_{1,N-1} | \mathcal{F}_{N-1}) = M_{(N-1)1}. \end{aligned}$$

As a consequence, $\{M_{N1}, \mathcal{F}_N, N \geq 1\}$ is a martingale. Via the Theorem 1.1 in Billingsley (1961) and the Corollary 3.2 from Hall and Heyde (1980), we can find

$$\begin{aligned} & \frac{1}{N} \sum_{t=1}^N [X_{1t} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{1,t-1} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{2,t-1}]^2 (1 - \mu_{\alpha_1})^2 X_{1,t-1}^2 \\ & \xrightarrow{a.s.} E[[X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 (1 - \mu_{\alpha_1})^2 X_{10}^2] \stackrel{\Delta}{=} \sigma_{11}, \end{aligned}$$

and $\frac{1}{\sqrt{N}} M_{N1} \xrightarrow{d} N(0, \sigma_{11})$ is established.

Similarly, we can prove that $\{M_{N2}, \mathcal{F}_N, N \geq 1\}$ and $\{M_{N3}, \mathcal{F}_N, N \geq 1\}$ are both martingales, as $N \rightarrow \infty$,

$$\begin{aligned} & \frac{1}{N} \sum_{t=1}^N [X_{1t} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{1,t-1} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{2,t-1}]^2 (1 - \mu_{\beta_1})^2 X_{2,t-1}^2 \\ & \xrightarrow{a.s.} E[[X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 (1 - \mu_{\beta_1})^2 X_{20}^2] \stackrel{\Delta}{=} \sigma_{22}, \\ & \frac{1}{\sqrt{N}} M_{N2} \xrightarrow{d} N(0, \sigma_{22}); \end{aligned}$$

and

$$\begin{aligned} & \frac{1}{N} \sum_{t=1}^N [X_{1t} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{1,t-1} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{2,t-1}]^2 (1 - \mu_{\gamma_1})^2 (n - X_{1,t-1} - X_{2,t-1})^2 \\ & \xrightarrow{a.s.} E[[X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 (1 - \mu_{\gamma_1})^2 (n - X_{10} - X_{20})^2] \stackrel{\Delta}{=} \sigma_{33}, \\ & \frac{1}{\sqrt{N}} M_{N3} \xrightarrow{d} N(0, \sigma_{33}). \end{aligned}$$

Thus, we have:

$$\frac{1}{\sqrt{N}} \begin{pmatrix} M_{N1} \\ M_{N2} \\ M_{N3} \end{pmatrix} \xrightarrow{d} N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix} \right) \stackrel{\Delta}{=} N(\mathbf{0}, \mathbf{W}_1),$$

where

$$\begin{aligned} \sigma_{12} &= E[[X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 (1 - \mu_{\alpha_1})(1 - \mu_{\beta_1})X_{10}X_{20}], \\ \sigma_{13} &= E[[X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 \\ &\quad (1 - \mu_{\alpha_1})(1 - \mu_{\gamma_1})X_{10}(n - X_{10} - X_{20})], \\ \sigma_{23} &= E[[X_{11} - n\mu_{\gamma_1} - (\mu_{\alpha_1} - \mu_{\gamma_1})X_{10} - (\mu_{\beta_1} - \mu_{\gamma_1})X_{20}]^2 (1 - \mu_{\beta_1}) \\ &\quad (1 - \mu_{\gamma_1})X_{20}(n - X_{10} - X_{20})]. \end{aligned}$$

Furthermore, by making the first-order Taylor expansion at $\hat{\theta}_1^{CLS} = \theta_1$ for the function:

$$\begin{pmatrix} \sum_{t=1}^N [X_{1t} - n\hat{\mu}_{\gamma_1}^{CLS} - (\hat{\mu}_{\alpha_1}^{CLS} - \hat{\mu}_{\gamma_1}^{CLS})X_{1,t-1} - (\hat{\mu}_{\beta_1}^{CLS} - \hat{\mu}_{\gamma_1}^{CLS})X_{2,t-1}] (1 - \hat{\mu}_{\alpha_1}^{CLS})X_{1,t-1} \\ \sum_{t=1}^N [X_{1t} - n\hat{\mu}_{\gamma_1}^{CLS} - (\hat{\mu}_{\alpha_1}^{CLS} - \hat{\mu}_{\gamma_1}^{CLS})X_{1,t-1} - (\hat{\mu}_{\beta_1}^{CLS} - \hat{\mu}_{\gamma_1}^{CLS})X_{2,t-1}] (1 - \hat{\mu}_{\beta_1}^{CLS})X_{2,t-1} \\ \sum_{t=1}^N [X_{1t} - n\hat{\mu}_{\gamma_1}^{CLS} - (\hat{\mu}_{\alpha_1}^{CLS} - \hat{\mu}_{\gamma_1}^{CLS})X_{1,t-1} - (\hat{\mu}_{\beta_1}^{CLS} - \hat{\mu}_{\gamma_1}^{CLS})X_{2,t-1}] (1 - \hat{\mu}_{\gamma_1}^{CLS})(n - X_{1,t-1} - X_{2,t-1}) \end{pmatrix} = \mathbf{0},$$

we have that

$$\begin{pmatrix} \hat{\mu}_{\alpha_1}^{CLS} - \mu_{\alpha_1} \\ \hat{\mu}_{\beta_1}^{CLS} - \mu_{\beta_1} \\ \hat{\mu}_{\gamma_1}^{CLS} - \mu_{\gamma_1} \end{pmatrix} = \frac{1}{N} V_1^{-1} \begin{pmatrix} M_{N1} \\ M_{N2} \\ M_{N3} \end{pmatrix},$$

where

$$\begin{aligned}
V_1^{(11)} &= (1 - 2\mu_{\alpha_1} + \mu_{\gamma_1})E(X_{10}^2) + E(X_{11}X_{10}) - n\mu_{\gamma_1}E(X_{10}) - (\mu_{\beta_1} - \mu_{\gamma_1})E(X_{20}X_{10}), \\
V_1^{(22)} &= (1 - 2\mu_{\beta_1} + \mu_{\gamma_1})E(X_{20}^2) + E(X_{11}X_{20}) - n\mu_{\gamma_1}E(X_{20}) - (\mu_{\alpha_1} - \mu_{\gamma_1})E(X_{10}X_{20}), \\
V_1^{(33)} &= (1 - 2\mu_{\gamma_1})E[(n - X_{10} - X_{20})^2] + E[(X_{11} - \mu_{\alpha_1}X_{10} - \mu_{\beta_1}X_{20})(n - X_{10} - X_{20})], \\
V_1^{(12)} &= (1 - \mu_{\alpha_1})E(X_{10}X_{20}), \quad V_1^{(21)} = (1 - \mu_{\beta_1})E(X_{10}X_{20}), \\
V_1^{(13)} &= (1 - \mu_{\alpha_1})E[X_{10}(n - X_{10} - X_{20})], \quad V_1^{(31)} = (1 - \mu_{\gamma_1})E[X_{10}(n - X_{10} - X_{20})], \\
V_1^{(23)} &= (1 - \mu_{\beta_1})E[X_{20}(n - X_{10} - X_{20})], \quad V_1^{(32)} = E(1 - \mu_{\gamma_1})E[X_{20}(n - X_{10} - X_{20})].
\end{aligned}$$

Finally, the Equation (8) is established. Similar to the proof of Theorem 2, according to Equation (5), let

$$\begin{aligned}
M_{N4} &= -\frac{1}{2} \frac{\partial S(\theta)}{\partial \mu_{\alpha_2}} = \sum_{t=1}^N [X_{2t} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) \\
&\quad - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{1,t-1} \\
&\quad - (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{2,t-1}](1 - \mu_{\alpha_1})X_{1,t-1},
\end{aligned} \tag{A4}$$

$$\begin{aligned}
M_{N5} &= -\frac{1}{2} \frac{\partial S(\theta)}{\partial \mu_{\beta_2}} = \sum_{t=1}^N [X_{2t} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) \\
&\quad - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{1,t-1} - (\mu_{\beta_2}(1 - \mu_{\beta_1}) \\
&\quad - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{2,t-1}](1 - \mu_{\beta_1})X_{2,t-1},
\end{aligned} \tag{A5}$$

$$\begin{aligned}
M_{N6} &= -\frac{1}{2} \frac{\partial S(\theta)}{\partial \mu_{\gamma_2}} = \sum_{t=1}^N [X_{2t} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) \\
&\quad - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{1,t-1} \\
&\quad - (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{2,t-1}](1 - \mu_{\gamma_1})(n - X_{1,t-1} - X_{2,t-1}).
\end{aligned} \tag{A6}$$

By replacing the parameters θ_1 in the above functions with their consistent estimators in Equation (6), the CLS estimators of $\theta_2 = (\mu_{\alpha_2}, \mu_{\beta_2}, \mu_{\gamma_2})^\top$ can be obtained as Equation (7).

Therefore, $\{M_{N4}, \mathcal{F}_N, N \geq 1\}$, $\{M_{N5}, \mathcal{F}_N, N \geq 1\}$ and $\{M_{N6}, \mathcal{F}_N, N \geq 1\}$ are all martingales and

$$\frac{1}{\sqrt{N}}M_{N4} \xrightarrow{d} N(0, W_2^{(11)}), \quad \frac{1}{\sqrt{N}}M_{N5} \xrightarrow{d} N(0, W_2^{(22)}), \quad \frac{1}{\sqrt{N}}M_{N6} \xrightarrow{d} N(0, W_2^{(33)}).$$

Thus, we have:

$$\frac{1}{\sqrt{N}} \begin{pmatrix} M_{N4} \\ M_{N5} \\ M_{N6} \end{pmatrix} \xrightarrow{d} N(\mathbf{0}, W_2),$$

where

$$\begin{aligned}
W_2^{(11)} &= E\{[X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10} \\
&\quad - (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2(1 - \mu_{\alpha_1})^2X_{10}^2\}, \\
W_2^{(22)} &= E\{[X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10} \\
&\quad - (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2(1 - \mu_{\beta_1})^2X_{20}^2\}, \\
W_2^{(33)} &= E\{[X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10} \\
&\quad - (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2(1 - \mu_{\gamma_1})^2(n - X_{10} - X_{20})^2\}, \\
W_2^{(12)} &= W_2^{(21)} = E\{[X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10} \\
&\quad - (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2(1 - \mu_{\alpha_1})(1 - \mu_{\beta_1})X_{10}X_{20}\}, \\
W_2^{(13)} &= W_2^{(31)} = E\{[X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10} \\
&\quad - (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2 \\
&\quad (1 - \mu_{\alpha_1})(1 - \mu_{\gamma_1})X_{10}(n - X_{10} - X_{20})\},
\end{aligned}$$

$$\begin{aligned} \mathbf{W}_2^{(23)} = \mathbf{W}_2^{(32)} &= E\{[X_{21} - n\mu_{\gamma_2}(1 - \mu_{\gamma_1}) - (\mu_{\alpha_2}(1 - \mu_{\alpha_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{10} \\ &\quad - (\mu_{\beta_2}(1 - \mu_{\beta_1}) - \mu_{\gamma_2}(1 - \mu_{\gamma_1}))X_{20}]^2 \\ &\quad (1 - \mu_{\beta_1})(1 - \mu_{\gamma_1})X_{20}(n - X_{10} - X_{20})\}. \end{aligned}$$

Furthermore, through the first-order Taylor expansion at $\hat{\theta}_2 = \theta_2$, we have:

$$\begin{pmatrix} \hat{\mu}_{\alpha_2}^{CLS} - \mu_{\alpha_2} \\ \hat{\mu}_{\beta_2}^{CLS} - \mu_{\beta_2} \\ \hat{\mu}_{\gamma_2}^{CLS} - \mu_{\gamma_2} \end{pmatrix} = \frac{1}{N} \mathbf{V}_2^{-1} \begin{pmatrix} M_{N4} \\ M_{N5} \\ M_{N6} \end{pmatrix},$$

where

$$\begin{aligned} \mathbf{V}_2^{(11)} &= (1 - \mu_{\alpha_1})^2 E(X_{10}^2), \quad \mathbf{V}_2^{(22)} = (1 - \mu_{\beta_1})^2 E(X_{20}^2), \\ \mathbf{V}_2^{(33)} &= (1 - \mu_{\gamma_1})^2 E((n - X_{10} - X_{20})^2), \\ \mathbf{V}_2^{(12)} = \mathbf{V}_2^{(21)} &= (1 - \mu_{\alpha_1})(1 - \mu_{\beta_1})E(X_{10}X_{20}), \\ \mathbf{V}_2^{(13)} = \mathbf{V}_2^{(31)} &= (1 - \mu_{\alpha_1})(1 - \mu_{\gamma_1})E[X_{10}(n - X_{10} - X_{20})], \\ \mathbf{V}_2^{(23)} = \mathbf{V}_2^{(32)} &= (1 - \mu_{\beta_1})(1 - \mu_{\gamma_1})E[X_{20}(n - X_{10} - X_{20})]; \end{aligned}$$

As a consequence, the proof of Theorem 3 is completed. \square

Appendix D. The Histograms of Estimation

In this section, we give the histograms of three estimators of $\boldsymbol{\theta}$ for Case B with $N = 500$. The abscissas of the vertical lines are the true values of the parameters. It is not difficult to see that the CLS, WCLS and CML estimation can obtain great simulation results. The distributions of the estimators are similar to the normal distributions, where the mean of the normal distribution is the true value of parameter.

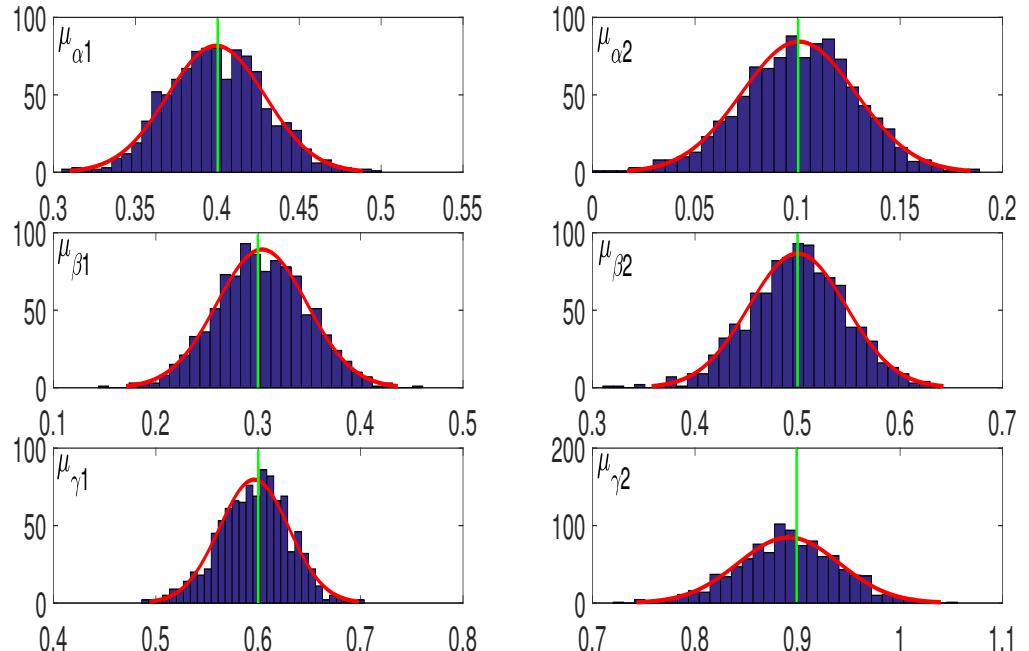


Figure A1. Histograms of CLS estimators $(\hat{\mu}_{\alpha_1}^{CLS}, \hat{\mu}_{\alpha_2}^{CLS}, \hat{\mu}_{\beta_1}^{CLS}, \hat{\mu}_{\beta_2}^{CLS}, \hat{\mu}_{\gamma_1}^{CLS}, \hat{\mu}_{\gamma_2}^{CLS})^T$ for Case B with the sample size equals to 500. The abscissas of the vertical lines are the true values of the parameters.

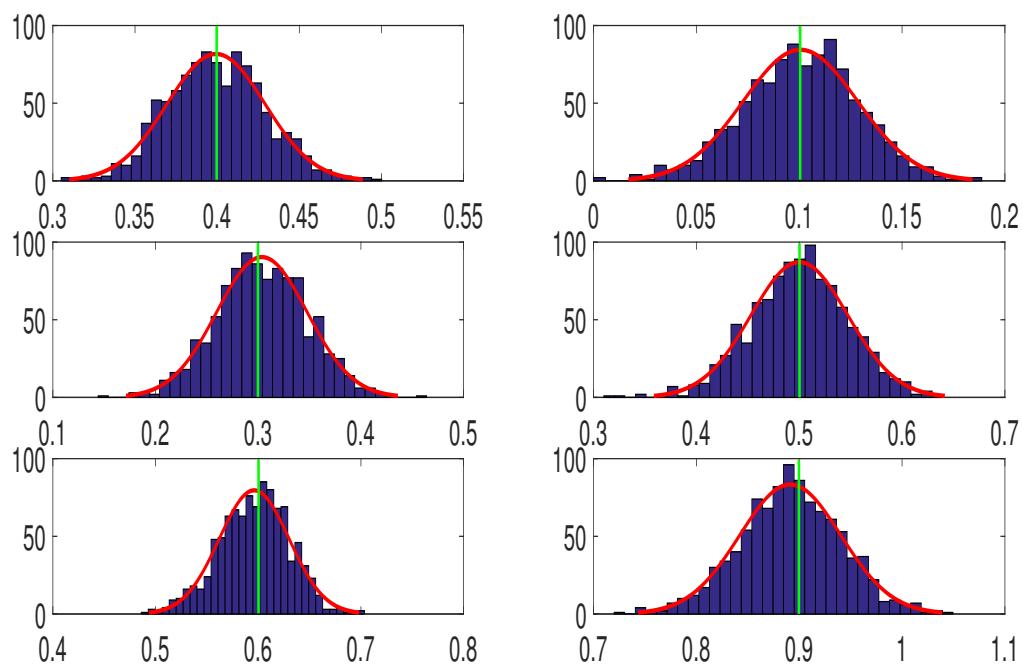


Figure A2. Histograms of WCLS estimators $(\hat{\mu}_{\alpha_1}^{WCLS}, \hat{\mu}_{\alpha_2}^{WCLS}, \hat{\mu}_{\beta_1}^{WCLS}, \hat{\mu}_{\beta_2}^{WCLS}, \hat{\mu}_{\gamma_1}^{WCLS}, \hat{\mu}_{\gamma_2}^{WCLS})^T$ for Case B with the sample size equals to 500. The abscissas of the vertical lines are the true values of the parameters.

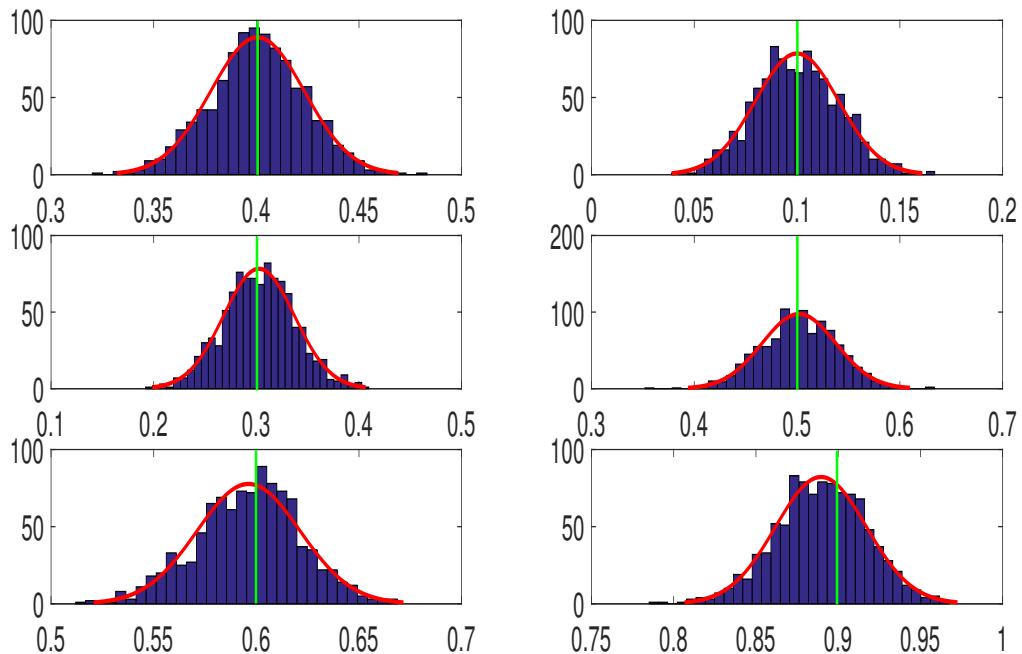


Figure A3. Histograms of CML estimators $(\hat{\mu}_{\alpha_1}^{CML}, \hat{\mu}_{\alpha_2}^{CML}, \hat{\mu}_{\beta_1}^{CML}, \hat{\mu}_{\beta_2}^{CML}, \hat{\mu}_{\gamma_1}^{CML}, \hat{\mu}_{\gamma_2}^{CML})^T$ for Case B with the sample size equals to 500. The abscissas of the vertical lines are the true values of the parameters.

References

- Steutel, F.W.; Van Harn, K. Discrete analogues of self-decomposability and stability. *Ann. Probab.* **1979**, *7*, 893–899. [[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](#)]
- McKenzie, E. Some simple models for discrete variate time series. *Water Resour. Bull.* **1985**, *21*, 645–650. [[CrossRef](#)]

4. Zheng, H.; Basawa, I.V.; Datta, S. First-order random coefficient integer-valued autoregressive process. *J. Stat. Plan. Inference* **2007**, *137*, 212–229. [[CrossRef](#)]
5. Wang, D.; Zhang, H. Generalized RCINAR(p) process with signed thinning operator. *Commun. Stat.—Simul. Comput.* **2010**, *40*, 13–44. [[CrossRef](#)]
6. 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](#)]
7. Zhang, J.; Wang, D.; Yang, K. A study of RCINAR(1) process with generalized negative binomial marginals. *Commun. Stat.—Simul. Comput.* **2019**, *49*, 1487–1510. [[CrossRef](#)]
8. Yang, K.; Wang, D.; Li, H. Threshold autoregression analysis for finite-range time series of counts with an application on measles data. *J. Stat. Comput. Simul.* **2018**, *88*, 597–614. [[CrossRef](#)]
9. Scotto, M.G.; Weiβ, C.H.; Silva, M.E.; Pereira, I. Bivariate binomial autoregressive models. *J. Multivar. Anal.* **2014**, *125*, 233–251. [[CrossRef](#)]
10. Ristić, M.M.; Popović, B.V. A new bivariate binomial time series model. *Markov Process. Relat. Fields* **2019**, *25*, 1–26.
11. Hausman, J.; McFadden, D. Specification tests for the multinomial logit model. *Econom. J. Econom. Soc.* **1984**, *52*, 1219–1240. [[CrossRef](#)]
12. Böckenholt, U. An INAR(1) negative multinomial regression model for longitudinal count data. *Psychometrika* **1999**, *64*, 53–67. [[CrossRef](#)]
13. Böckenholt, U. Analyzing multiple emotions over time by autoregressive negative multinomial regression models. *J. Am. Stat. Assoc.* **1999**, *94*, 757–765.
14. Dow, J.K.; Endersby, J.W. Multinomial probit and multinomial logit: A comparison of choice models for voting research. *Elect. Stud.* **2004**, *23*, 107–122. [[CrossRef](#)]
15. Zhang, J.; Wang, D.; Yang, K.; Yanju, X. A multinomial autoregressive model for finite-range time series of counts. *J. Stat. Plan. Inference* **2020**, *207*, 320–343. [[CrossRef](#)]
16. Scotto, M.G.; Weiβ, C.H.; Gouveia, S. Thinning-based models in the analysis of integer-valued time series: A review. *Stat. Model.* **2015**, *15*, 590–618. [[CrossRef](#)]