

# Article A Two-Step Estimation Method for a Time-Varying INAR Model

Yuxin Pang <sup>1,2</sup>, Dehui Wang <sup>3,\*</sup> and Mark Goh <sup>2</sup>

- <sup>1</sup> School of Mathematics, Jilin University, Changchun 130012, China; pangyx19@mails.jlu.edu.cn
- <sup>2</sup> NUS Business School & The Logistics Institute-Asia Pacific, National University of Singapore, Singapore 119613, Singapore; mark.goh@nus.edu.sg
- <sup>3</sup> School of Mathematics and Statistics, Liaoning University, Shenyang 110031, China
- Correspondence: wangdehui@lnu.edu.cn

Abstract: This paper proposes a new time-varying integer-valued autoregressive (TV-INAR) model with a state vector following a logistic regression structure. Since the autoregressive coefficient in the model is time-dependent, the Kalman-smoothed method is applicable. Some statistical properties of the model are established. To estimate the parameters of the model, a two-step estimation method is proposed. In the first step, the Kalman-smoothed estimation method, which is suitable for handling time-dependent systems and nonstationary stochastic processes, is utilized to estimate the time-varying parameters. In the second step, conditional least squares is used to estimate the parameter in the error term. This proposed method allows estimating the parameters in the nonlinear model and deriving the analytical solutions. The performance of the estimation method is evaluated through simulation studies. The model is then validated using actual time series data.

**Keywords:** time-varying integer-valued autoregressive model; parameter estimation; Kalman-smoother; logistic regression; conditional least squares

MSC: 62M10

# 1. Introduction

Time series models with an integer-valued structure are prevalent in fields such as economics, insurance, medicine and crime. Recent reviews on count time series models based on thinning operators, including modeling and numerous examples, can be found in refs. [1–3]. The most classic one is the first-order integer-valued autoregressive INAR(1) model introduced by ref. [4] and ref. [5] using the properties of binomial sparse operators (ref. [6]). The class of INAR models is typically based on the assumption of observations that follow a Poisson distribution, which facilitates subsequent computations due to the equidispersion feature of the Poisson distribution (mean and variance being equal). Parameter estimation for these models is usually achieved by Yule–Walker, conditional least squares and conditional maximum likelihood methods, as discussed in refs. [7–9], among others. Recently, there has been a growing interest in research regarding the autoregressive coefficients in these models. Generally, the autoregressive coefficients can be treated as random variables, with rules such as a stationary process or specified mean and variance; see refs. [10–12]. Unlike the above models, this paper proposes a time-varying integer-valued autoregressive (TV-INAR) model in which the state equation is a nonstationary form.

In our work, the model is characterized by a state equation and time-varying parameters. The concept of state-space time series analysis was first introduced by ref. [13] in the field of engineering. Over time, the term "state space" became entrenched in statistics and econometrics, as the state-space model provides an effective method to deal with a wide range of problems in time series analysis. Ref. [14] presented a comprehensive treatment of the state-space approach to time series analysis. Recently, some progress has been made on the use of state space in integer autoregressive and count time series



Citation: Pang, Y.; Wang, D.; Goh, M. A Two-Step Estimation Method for a Time-Varying INAR Model. *Axioms* 2024, *13*, 19. https://doi.org/ 10.3390/axioms13010019

Academic Editor: Hari Mohan Srivastava

Received: 28 November 2023 Revised: 21 December 2023 Accepted: 22 December 2023 Published: 27 December 2023



**Copyright:** © 2023 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/).



models. Ref. [15] proposed a first-order random coefficient integer-valued autoregressive model by introducing a Markov chain with a finite state space and derived some statistical properties of the model in a random environment. Ref. [16] introduced a parameter-driven state-space model to analyze integer-valued time series data. And, to accommodate the features of overdispersion, zero-inflation and temporal correlation of count time series, ref. [17] proposed a flexible class of dynamic models in the state-space framework. As noted in ref. [18], the popularity of these models stems in large part from the development of the Kalman recursion, proving a quick updating scheme for predicting, filtering and smoothing a time series. The traditional understanding of Kalman filtering can be found in the works of refs. [14,19,20]. Nevertheless, most of this research has focused on the use of continuous time series models in an economic context, such as the analysis of stocks, trade and currency. For example, ref. [21] proposed a class of time-varying parameter autoregressive models and proved the equivalence of the Kalman-smoothed estimate and generalized least squares estimate. Following this lead, [22] developed a trade growth relationship model with time-varying parameters and estimated the transition of changing parameters with a Kalman filter. Both of them all involved cases of Gaussian and linear parameters. Ref. [14] attested that the results obtained by Kalman smoothing are equivalent when the model is non-Gaussian or nonlinear. Therefore, integer-valued time series models in the non-Gaussian case are also worth investigating.

Time series models with state-space and time-varying parameters are common in economics. Many macroeconomists believe that time-varying parameter models can better predict and adapt to the data than fixed-parameter models. Early research mostly focused on time-varying continuous time series models, such as the autoregressive (AR) models, vector autoregressive (VAR) models and autoregressive moving average (ARMA) models (see refs. [21,23,24]). In recent years, research on INAR models with time-varying parameters has attracted more attention and has been applied to natural disasters and medical treatment. Ref. [25] introduced a multivariate integer-valued autoregressive model of order one with periodic time-varying parameters and adopted a composite likelihoodbased approach. Ref. [26] considered a change-point analysis of counting time series data through a Poisson INAR(1) process with time-varying smoothing covariates. However, the time-varying parameters in the above models are not controlled by a state equation. Additionally, it is difficult to deal with time-varying parameter models when there are unobserved variables that need to be estimated. There are only a few effective methods available. Ref. [27] proposed a Bayesian estimation method for time-varying parameters and claimed that the Bayesian method is superior to the maximum likelihood method. Both the Bayesian and maximum likelihood methods require Kalman filtering to estimate state vectors that contain time-varying parameters. The application of the Kalman filter would be made possible once the model is put into a state-space form. It is worth exploring some new methods to deal with TV-INAR models.

Motivated by ref. [21], a new TV-INAR model with a state equation is presented in this paper. Since the model is in a state-space form, the Kalman smoothing method uses only known observation variables to estimate the parameters. Unlike traditional INAR models, which are limited to modeling stationary time series data, our TV-INAR model is equipped to handle nonstationary structures and time-varying dynamics, resulting in improved model fit and more accurate predictions. The Kalman-smoothed estimates of the time-varying unobserved state variables are derived. Furthermore, the mean of the Poisson error term is estimated through the estimates obtained in the previous step and conditional least squares methods.

The rest of this paper is organized as follows. A new TV-INAR model is presented and its basic properties are constructed in Section 2. In Section 3, the Kalman smoother is utilized to derive an estimate of the time-varying parameter. Then, incorporating conditional least squares (CLS) methods, the estimation of the mean of the Poisson error term is established. Numerical simulations and results are discussed in Section 4. In Section 5, the proposed model is applied to the offense data sets in Rochester. The conclusion is given in Section 6.

#### 3 of 17

#### 2. Poisson TV-INAR Model

In this section, the INAR(1) model is reviewed. A new TV-INAR model incorporating time-varying and nonstationary features is introduced. Then, some basic statistical properties are derived.

Suppose *Y* is a non-negative integer-valued random variable and  $\beta \in (0, 1)$ . The binomial thinning  $\circ$ , introduced in ref. [6], is defined as  $\beta \circ Y = \sum_{i=1}^{Y} B_i$ , where  $\{B_i\}$  is a sequence of independent and identically distributed (i.i.d.) Bernoulli random variables, independent of *Y*, and satisfying  $P(B_i = 1) = 1 - P(B_i = 0) = \beta$ . Then, the INAR(1) model is given by

$$Y_t = \beta \circ Y_{t-1} + \varepsilon_t, \qquad t = 1, 2, \cdots, \tag{1}$$

where  $\{\varepsilon_t\}$  is a sequence of uncorrelated non-negative integer-valued random variables, with mean  $\mu_{\varepsilon}$  and finite variance  $\sigma_{\varepsilon}^2$ , and  $\varepsilon_t$  is independent of all  $\{Y_{t-i}\}, i = 1, 2, \cdots$ .

#### 2.1. Definition of TV-INAR Process

It is very common to extend the autoregressive coefficient to a random parameter in time series models. However, it differs from the time-varying parameter introduced in this paper. In random coefficient (RC) models, the variable is usually assigned a definite distribution or given its expectation and variance. Although it is a random variable, its expectation and variance do not change with time. In contrast, in time-varying parameter models, the parameter does not have a fixed distribution, and its expectation and variance often change over time. This is also one of the challenges of such models compared with the ordinary RC models. Thus, based on the above INAR(1) process, we define the time-varying integer-valued autoregressive (TV-INAR) process as follows.

**Definition 1.** Let  $\{y_t\}_{t \in \mathbb{N}_0}$  be an integer-valued regressive process. It is a time-varying integer-valued regressive model if

$$y_t = g(\alpha_t) \circ Z_t + \varepsilon_t,$$
  

$$\alpha_t = \alpha_{t-1} + \eta_t,$$
(2)

where  $g(\cdot) \in (0,1)$  is a function of  $\alpha_t$ ;  $\{\varepsilon_t\}$  is a sequence of i.i.d. Poisson-distributed random variables with mean  $\lambda$ ;  $\{\eta_t\}$  is a sequence of i.i.d. standard normally distributed random variables; and  $\varepsilon_t$  is independent of  $Z_t$  and  $\eta_t$  when  $t \ge 1$ .

In the model (2),  $y_t$  and  $Z_t$  are observation variables, and  $\alpha_t$  is an unobserved variable, often called a time-varying parameter. Our model allows the class of TV-INAR models. For example,  $Z_t = y_{t-1}$  yields a TV-INAR(1) model. Model (2) becomes a TV-INAR(p) model when  $Z_t = (y_{t-1}, \dots, y_{t-p})'$  and the autocorrelation coefficient expands to  $(g(\alpha_t), \dots, g(\alpha_{t-p+1}))$ . Furthermore,  $Z_t$  can also be considered as a covariate of  $y_t$ .

In this paper, we devote the case of  $g(\alpha_t) = \frac{e^{\alpha_t}}{1+e^{\alpha_t}}$ . The idea of this model is that the development of the system over time is determined by  $\alpha_t$  in the second equation of (2). However, since  $\alpha_t$  cannot be observed directly, we need to analyze it based on the observations  $y_t$ . The first equation of (2) is called the observation equation; the second is called the state equation [14]. We assume initially that  $\alpha_0 \sim N(\mu_0, \sigma_0^2)$  where  $\mu_0$  and  $\sigma_0^2$  are known, because the state equation is nonstationary. Variances of the error terms  $\varepsilon_t$  and  $\eta_t$  in (2) are time-invariant.

Studies on the class of TV-AR models have been extensive, especially in the field of economics. Such models are flexible enough to capture the complexity of macroeconomic systems and fit the data better than models with constant parameters. Retaining the advantages of the above models, we propose the TV-INAR model, which can deal with integer-valued discrete data and exists in real life. And, it is of interest that  $\alpha_t$  is always nonstationary, as  $\alpha_t$  is a random walk. Therefore, the existence of the above state equation is justified. And the binomial thinning operator  $\circ$  portrays the probability of the event, so  $g(\alpha_t)$  is given the form  $\frac{e^{\alpha_t}}{1+e^{\alpha_t}}$  to ensure that the probability is between (0,1).

#### 2.2. Statistical Properties

To have a better understanding of the model and to apply it directly to parameter estimation, some statistical properties of the model are provided. The mean, conditional mean, second-order conditional origin moment, variance and conditional variance of the time-varying integer-valued regressive process are given in the following proposition. The case of TV-INAR(1) is given in the corollary.

**Proposition 1.** Suppose  $\{Y_t\}$  is a process defined by (2), and  $\mathcal{F}_{Z,t}$  is a  $\sigma$ -field generated by  $\{Z_1, \cdots, Z_t, \eta_1, \cdots, \eta_t, \alpha_0\}$ . Then, when  $t \ge 1$ , we have

(1)  $E[Y_t] = E[Y_t | \mathcal{F}_{Z,t}] = \frac{e^{\alpha_t}}{1 + e^{\alpha_t}} Z_t + \mu_{\varepsilon} = \frac{e^{\alpha_t}}{1 + e^{\alpha_t}} Z_t + \lambda;$ 

$$\begin{aligned} (2) \quad & E[Y_t^2] = E[Y_t^2 | \mathcal{F}_{Z,t}] \\ &= E[(\frac{e^{\alpha_t}}{1 + e^{\alpha_t}} \circ Z_t)^2 | \mathcal{F}_{Z,t}] + E[\varepsilon_t^2 | \mathcal{F}_{Z,t}] + 2E[\frac{e^{\alpha_t}}{1 + e^{\alpha_t}} \circ Z_t | \mathcal{F}_{Z,t}] E[\varepsilon_t | \mathcal{F}_{Z,t}] \\ &= \frac{e^{\alpha_t}}{1 + e^{\alpha_t}} (1 - \frac{e^{\alpha_t}}{1 + e^{\alpha_t}}) Z_t + (\frac{e^{\alpha_t}}{1 + e^{\alpha_t}} Z_t)^2 + \lambda + \lambda^2 + 2(\frac{e^{\alpha_t}}{1 + e^{\alpha_t}} Z_t + \lambda)\lambda \\ &= \frac{e^{\alpha_t}}{1 + e^{\alpha_t}} Z_t (1 - \frac{e^{\alpha_t}}{1 + e^{\alpha_t}} + \frac{e^{\alpha_t}}{1 + e^{\alpha_t}} Z_t + 2\lambda) + 3\lambda^2 + \lambda; \end{aligned}$$

$$(3) \quad Var[Y_t] = Var[Y_t | \mathcal{F}_{Z,t}] \\ &= E[Y_t^2 | \mathcal{F}_{Z,t}] - E^2[Y_t | \mathcal{F}_{Z,t}] \\ &= \frac{e^{\alpha_t}}{1 + e^{\alpha_t}} (1 - \frac{e^{\alpha_t}}{1 + e^{\alpha_t}}) Z_t + 2\lambda^2 + \lambda. \end{aligned}$$

**Corollary 1.** Suppose  $\{Y_t\}$  satisfies the TV-INAR(1) model, i.e.,  $Z_t = Y_{t-1}$ , and  $\mathcal{F}_{Y,t-1} =$  $\sigma(Y_1, \cdots, Y_{t-1}, \eta_1, \cdots, \eta_t, \alpha_0)$ , then for  $t \geq 1$ ,

- (1)  $E[Y_t|\mathcal{F}_{Y,t-1}] = \frac{e^{\alpha_t}}{1+e^{\alpha_t}}Y_{t-1} + \lambda;$ (2)  $E[Y_t] = E[E(Y_t|\mathcal{F}_{Y,t-1})] = \frac{\lambda}{1-\frac{e^{\alpha_t}}{1+e^{\alpha_t}}};$

(3) 
$$Var[Y_t | \mathcal{F}_{Y,t-1}] = \frac{e^{\alpha_t}}{1 + e^{\alpha_t}} (1 - \frac{e^{\alpha_t}}{1 + e^{\alpha_t}}) Y_{t-1} + \lambda;$$
  
(4) 
$$Var[Y_t] = \frac{\frac{e^{\alpha_t}}{1 + e^{\alpha_t}} \lambda + \lambda}{1 + e^{\alpha_t} \lambda^2} = \lambda (1 + e^{\alpha_t}).$$

 $Var[Y_t] = \frac{1+e^{\alpha_t} \lambda^{(1+\alpha)}}{1-(\frac{e^{\alpha_t}}{1+e^{\alpha_t}})^2} = \lambda(1+e^{\alpha_t}).$ Clearly, when  $Z_t = Y_{t-1}$ ,  $\{Y_t, \alpha_t\}_{t \in \mathbb{N}_0}$  is a bivariate Markov chain with the following

$$P(Y_{t} = j; \alpha_{t} = n | Y_{t-1} = i, \alpha_{t-1} = m)$$

$$= P(\frac{e^{\alpha_{t}}}{1 + e^{\alpha_{t}}} \circ Y_{t-1} + \varepsilon_{t} = j; \alpha_{t-1} + \eta_{t} = n | Y_{t-1} = i, \alpha_{t-1} = m)$$

$$= \sum_{l=0}^{\min(i,j)} P(\frac{e^{\alpha_{t}}}{1 + e^{\alpha_{t}}} \circ Y_{t-1} = l | Y_{t-1} = i, \alpha_{t-1} = m) P(\varepsilon_{t} = j - l) P(\eta_{t} = n - m)$$

$$= \sum_{l=0}^{\min(i,j)} {i \choose l} (\frac{e^{n}}{1 + e^{n}})^{l} (\frac{1}{1 + e^{n}})^{i-l} \frac{\lambda^{j-l}e^{-\lambda}}{(j-l)!} \frac{1}{\sqrt{2\pi}} e^{-\frac{(n-m)^{2}}{2}}.$$

#### 3. Estimation Procedure

transition probabilities:

There are two interesting parameters in the model, one is the time-varying parameter  $\alpha_t$ , and the other is the mean  $\lambda$  of  $\varepsilon_t$ . The main goal of this section is to estimate these two parameters using a new two-step estimation method. The first step gives the estimate  $\hat{\alpha}_t$ by Kalman-smoothed state. In the second step, the estimate  $\hat{\lambda}$  is obtained by  $\hat{\alpha}_t$  and the conditional least squares (CLS) method.

#### 3.1. Kalman-Smoothed Estimate of State Vector $\alpha$

In this subsection, the Kalman-smoothed method is used to estimate the time-varying parameter. In fact, the essence of the Kalman filter is the minimum variance linear unbiased estimation (MVLUE). From ref. [14], we know that when the model is nonlinear and non-Gaussian, the results obtained from the standpoint of MVLUE are equivalent to the linear and Gaussian cases. To find the Kalman-smoothed estimate of the unobserved state vector, we employ the matrix formulation of Equation (2) following ref. [14]. The required moments are calculated and the analytic expression of this estimate is given at the end.

For convenience, set  $\beta_t = \frac{e^{\alpha_t}}{1+e^{\alpha_t}}$ . Then,  $\alpha_t = \log \frac{\beta_t}{1-\beta_t} \triangleq logit\beta_t$  is a logistic transformation. For  $t = 1, \dots, T$ , Equation (2) can be written in matrix form, similarly to ref. [21]:

$$Y = \beta \circ Z + \varepsilon,$$
  

$$\alpha = (\mu + \eta)C,$$
(3)

where

$$Y = (y_1, \cdots, y_T), \boldsymbol{\beta} = (\beta_1, \cdots, \beta_T), Z = \begin{pmatrix} Z_1 & 0 \\ & \ddots \\ 0 & Z_T \end{pmatrix}, \boldsymbol{\varepsilon} = (\varepsilon_1, \cdots, \varepsilon_T),$$
$$\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_T), \boldsymbol{\mu} = (\mu, 0, \cdots, 0), \boldsymbol{\eta} = (\eta_1, \cdots, \eta_T), \boldsymbol{C} = \begin{pmatrix} 1 & \cdots & 1 \\ & \ddots & \vdots \\ 0 & & 1 \end{pmatrix}.$$

Now, consider the problem of estimating the parameter. The following theorem adapted from ref. [14] provides the Kalman-smoothed estimate of the state vector  $\alpha$  in model (3).

**Theorem 1.** The Kalman-smoothed estimate of  $\alpha$  in model (3) is given by the conditional expectation of  $\alpha$ , which contains all the observations of  $y_t$ :

$$\hat{\boldsymbol{\alpha}} = E[\boldsymbol{\alpha}|\boldsymbol{Y}] = E(\boldsymbol{\alpha}) + (\boldsymbol{Y} - E(\boldsymbol{Y}))Var(\boldsymbol{Y})^{-1}Cov(\boldsymbol{Y}, \boldsymbol{\alpha}). \tag{4}$$

**Proof.** Let  $\hat{\alpha} = \mathbf{Y}A + B$ , where  $\mathbf{Y}$  is a vector  $(1 \times T)$ , B is a vector  $(1 \times T)$  and A is a matrix  $(T \times T)$ . Denote

$$I = E[(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})^T]$$
  
=  $E[(\boldsymbol{Y}A + B - \boldsymbol{\alpha})(\boldsymbol{Y}A + B - \boldsymbol{\alpha})^T]$   
=  $trE[(\boldsymbol{Y}A + B - \boldsymbol{\alpha})^T(\boldsymbol{Y}A + B - \boldsymbol{\alpha})].$ 

This is the error covariance matrix of state vector estimation, and it can be regarded as a function on *A* and *B*. To minimize *I*, we need

$$\begin{cases} \frac{\partial I}{\partial A} = 2E[\mathbf{Y}^T(\mathbf{Y}A + B - \boldsymbol{\alpha})] = 0, \\ \frac{\partial I}{\partial B} = 2E[\mathbf{Y}A + B - \boldsymbol{\alpha}] = 0. \end{cases}$$

We obtain

$$\begin{cases} A = \frac{E[\mathbf{Y}^T \boldsymbol{\alpha}] - E[\mathbf{Y}^T] E[\boldsymbol{\alpha}]}{E[\mathbf{Y}^T \mathbf{Y}] - E[\mathbf{Y}^T] E[\mathbf{Y}]} = Var(\mathbf{Y})^{-1}Cov(\mathbf{Y}, \boldsymbol{\alpha}), \\ B = E(\boldsymbol{\alpha}) - E(\mathbf{Y})A = E(\boldsymbol{\alpha}) - E(\mathbf{Y})Var(\mathbf{Y})^{-1}Cov(\mathbf{Y}, \boldsymbol{\alpha}). \end{cases}$$

This proves that  $\hat{\boldsymbol{\alpha}} = \boldsymbol{Y} Var(\boldsymbol{Y})^{-1} Cov(\boldsymbol{Y}, \boldsymbol{\alpha}) + E(\boldsymbol{\alpha}) - E(\boldsymbol{Y}) Var(\boldsymbol{Y})^{-1} Cov(\boldsymbol{Y}, \boldsymbol{\alpha}).$ 

Theorem 1 contains some numerical characteristics of the random variable. Following model (3), we denote  $\Lambda := E(\varepsilon) = (\lambda, \dots, \lambda)$ . Then, the mean of  $\alpha$  and Y, the variance of Y and the covariance between Y and  $\alpha$  are given in the following proposition.

**Proposition 2.** Suppose  $\{Y\}$  is a TV-INAR process defined by (3). Then, for any time T,

- (i)  $E(\boldsymbol{\alpha}) = \boldsymbol{\mu} \boldsymbol{C};$
- (*ii*)  $E(\mathbf{Y}) = E(\mathbf{\beta})\mathbf{Z} + \Lambda;$
- (iii)  $Var(\mathbf{Y}) = \mathbf{Z}Var(\boldsymbol{\beta})\mathbf{Z}' + E[diag(1 \beta_1, \cdots, 1 \beta_t)diag(\beta_1, \cdots, \beta_T)]\mathbf{Z} + diag(\lambda, \cdots, \lambda);$
- (*iv*)  $Cov(\mathbf{Y}, \boldsymbol{\alpha}) = E(\boldsymbol{\beta}'\boldsymbol{\alpha})\mathbf{Z} \mathbf{Z}E(\boldsymbol{\beta}')\boldsymbol{\mu}\mathbf{C}.$

The proofs of Proposition 2 are given in Appendix A. From Proposition 2, the Kalmansmoothed estimate of the time-varying parameter is

$$\hat{\boldsymbol{\alpha}} = E[\boldsymbol{\alpha}|\boldsymbol{Y}] = \boldsymbol{\mu}\boldsymbol{C} + [\boldsymbol{Y} - E(\boldsymbol{\beta})\boldsymbol{Z} - \boldsymbol{\Lambda}] \{\boldsymbol{Z} Var(\boldsymbol{\beta})\boldsymbol{Z}' + E[diag(1 - \beta_1, \cdots, 1 - \beta_T) \\ diag(\beta_1, \cdots, \beta_T)]\boldsymbol{Z} + diag(\lambda, \cdots, \lambda)\}^{-1} [E(\boldsymbol{\beta}'\boldsymbol{\alpha})\boldsymbol{Z} - \boldsymbol{Z}E(\boldsymbol{\beta}')\boldsymbol{\mu}\boldsymbol{C}],$$
(5)

where  $\alpha$  and  $\beta$  are random variables. Hence, in order to obtain the concrete result of the estimate, we need to compute the numerical characteristics of the random variables in Equation (5), i.e.,  $E(\beta_k)$ ,  $Var(\beta_k)$ ,  $E[\beta_k(1 - \beta_k)]$ ,  $E(\beta_i\alpha_j)$  and  $E(\beta_{k-l}\beta_k)$ , where k, i, j denote instances in time.

First, we set  $\alpha_0 \sim N(\mu_0, \sigma_0^2)$ . Then, we have

$$\alpha_k = \alpha_0 + \sum_{i=1}^k \eta_i \sim N(\mu_0, \sigma_0^2 + k),$$
$$e^{\alpha_k} \sim LN(\mu_0, \sigma_0^2 + k).$$

Here, *LN* stands for the logarithmic Normal distribution. Denote *X* as  $e^{\alpha_k}$  ( $k = 1, \dots, T$ ),  $\mu$  as its mean and  $\sigma^2$  as its variance, respectively. Then,  $X \sim LN(\mu, \sigma^2)$ , and the probability density function of *X* is

$$p(x) = \frac{1}{\sqrt{2\pi\sigma x}} \exp\left\{-\frac{(\ln x - \mu)^2}{2\sigma^2}\right\}.$$

Similarly, denote *Y* as  $e^{\sum_{l=j+1}^{i} \eta_l}$  and  $\tau^2$  as its variance, respectively. Then,  $Y \sim LN(0, \tau^2)$ . According to the logistic transformation,  $\beta_t = \frac{e^{\alpha_t}}{1+e^{\alpha_t}}$ . So  $\beta_k$  can be expressed in terms of *X*, i.e.,  $\beta_k = 1 - \frac{1}{1+X}$ . Thus, from the distributions of *X* and *Y*, the following proposition can be established.

**Proposition 3.** Suppose  $\beta_k$  and  $\beta_i$  are elements of  $\beta$  in Proposition 2, and  $\alpha_j$  is an element of  $\alpha$ . Denote  $ru_{\alpha+\frac{j^2}{2}(\sigma^2+k)}$ 

$$e_{r,k} = e^{r\mu_0 + \frac{1}{2}(v_0 + k)},$$
  

$$\Phi_{r,k} = \Phi\left(\frac{\mu_0}{\sqrt{\sigma_0^2 + k}} - r\sqrt{\sigma_0^2 + k}\right),$$
  

$$\Phi_{-1,r,k} = \Phi\left(-\frac{\mu_0}{\sqrt{\sigma_0^2 + k}} - r\sqrt{\sigma_0^2 + k}\right).$$

Then, for any  $k, i, j \in (1, \dots, T), l \in (1, \dots, k-1)$ ,

(1) 
$$E(\beta_k) = 1 - \sum_{r=0}^{\infty} \left[ (-1)^r e_{r,k} \Phi_{-1,r,k} + (-1)^{r+1} e_{-(r+1),k} \Phi_{r+1,k} \right];$$

(2) 
$$Var(\beta_k) = \sum_{r=0}^{\infty} (-1)^r (r+1) \left[ e_{r,k} \Phi_{-1,r,k} + e_{-(r+2),k} \Phi_{r+2,k} \right] - \left\{ \sum_{r=0}^{\infty} \left[ (-1)^r e_{r,k} \Phi_{-1,r,k} + (-1)^{r+1} e_{-(r+1),k} \Phi_{r+1,k} \right] \right\}^2;$$

(3) 
$$E[\beta_k(1-\beta_k)] = \sum_{r=1}^{\infty} (-1)^{r+1} r \Big[ e_{r,k} \Phi_{-1,r,k} + e_{-r,k} \Phi_{r,k} \Big];$$

(4) When 
$$j < i$$
,  

$$E(\beta_{i}\alpha_{j}) = \frac{1}{2}(\sqrt{\frac{\sigma_{0}^{2}+j}{2\pi}} + \mu_{0})\sum_{r=1}^{\infty}(-1)^{r+1}\left[(e^{r\mu_{0}} - e^{-r\mu_{0}})(\sqrt{\frac{\sigma_{0}^{2}+j}{2\pi}} + \mu_{0}) + (e^{r\mu_{0}} + e^{-r\mu_{0}})r(\sigma_{0}^{2} + j)\right]e^{\frac{r^{2}}{2}(\sigma_{0}^{2}+i)}\Phi(-r\sqrt{i-j});$$

$$When j > i,$$

$$E(\beta_{i}\alpha_{j}) = \sum_{r=1}^{\infty}(-1)^{r-1}\sqrt{2}\left[(\mu_{0} + r(\sigma_{0}^{2}+i))e_{r,i}\Phi_{-1,r,i} + (\mu_{0} - (r-1)(\sigma_{0}^{2}+i))e_{-(r-1),i} + \Phi_{r-1,i}\right];$$

(5) 
$$E(\beta_{k-l}\beta_k) = \frac{1}{\sqrt{2}} \sum_{r=1}^{\infty} (-1)^{r-1} \Big[ e_{r,k-l} \Phi_{-1,r,k-l} + e_{-(r-1),k-l} \Phi_{r-1,k-l} \Big] + \sum_{r=1}^{\infty} \sqrt{2} e^{\frac{r^2}{2}l} \Phi(-r\sqrt{l}) \Big[ \sum_{m=1}^{r} (-1)^{m-1} e_{m,k-l} + (-1)^m e_{-(m-1),k-l} \Big].$$

The proofs of Proposition 3 are given in Appendix B. And the distribution function  $\Phi(\cdot)$  of the standard normal distribution and Taylor's formula are used. Clearly, the exact solution  $\hat{\alpha}$  of the time-varying parameter  $\alpha$  is found by substituting (1)–(5) of Proposition 3 into Equation (5). Since  $\lim_{u\to+\infty} \Phi(u) = 1$ , if *r* tends to infinity, all of the  $\Phi(\cdot)$  in Proposition 3 converge to 0. The corresponding results are specified in Remark 1.

**Remark 1.** The items involving  $\Phi(\cdot)$  in Proposition 3 can be divided into two categories, namely  $\Phi(-a - rb)$  and  $\Phi(a - rb)$ , where a and b are constants greater than 0. According to the properties of the distribution function of the standard normal distribution, if r tends to infinity, the following limits apply:

$$\lim_{r \to +\infty} \Phi(-a - rb) = 1 - \lim_{r \to +\infty} \Phi(rb + a) = 0,$$
$$\lim_{r \to +\infty} \Phi(a - rb) = 1 - \lim_{r \to +\infty} \Phi(rb - a) = 0.$$

Therefore, (1)–(5) in Proposition 3 are convergent. Moreover, the Kalman-smoothed estimate  $\hat{\alpha}$  in (5) is also convergent.

#### 3.2. CLS Estimation of Parameter $\lambda$

Based on the proposed TV-INAR model and its statistical properties, we consider the parameter estimation problem on  $\lambda$  in the model. Here, the estimate  $\hat{\alpha}_t$  obtained in the previous section is used directly. Due to the characteristics of the model, that is, the state equation is nonstationary, the autoregressive coefficient in the observation equation is time-varying and the distribution is unknown, many estimation methods are not suitable for this model. Considering that the CLS estimation method does not need the distribution hypothesis of the model but only the corresponding moment information, and the estimation accuracy is relatively high, we prefer using the CLS method to estimate  $\lambda$ .

Suppose there is a TV-INAR process  $\{Y_t\}$ . Let

$$Q(\lambda) = \sum_{t=1}^{T} [Y_t - E(Y_t | Z_t)]^2$$
(6)

be the CLS criterion function. Then, the CLS estimator of the parameter  $\lambda$  can be obtained by minimizing the CLS criterion function, i.e.,

$$\hat{\lambda}_{CLS} = \arg\min Q(\lambda).$$

Taking the derivative of  $Q(\lambda)$  with respect to  $\lambda$  and setting it to zero, the minimizer  $\hat{\lambda}_{CLS}$  can be found:

$$\hat{\lambda}_{CLS} = \frac{1}{T} \sum_{t=1}^{I} Y_t - \frac{1}{T} \sum_{t=1}^{I} Z_t \frac{e^{\alpha_t}}{1 + e^{\hat{\alpha}_t}}$$

This is the CLS estimate of  $\lambda$ .

#### 4. Simulation

In this section, we perform simulations using Matlab R2018b to assess the performance of the two-step estimation approach discussed in Section 3. The approach recovers the true time-varying parameters, as well as the performance of the CLS estimator.

For the data generating process, using Equation (3), pseudodata are randomly generated by changing time T ( $T = \{50, 100, 200\}$ ) and the parameter  $\lambda$  ( $\lambda = \{0.05, 0.1, 0.8, 1.3, 4\}$ ) of the error term in the observation equation. The number of replications was 100. Here, the choice of  $\lambda$  is based on the size of the signal-to-noise ratio (SNR), which is defined as the variance in  $\eta_t$  relative to that in  $\varepsilon_t$ , i.e., SNR  $= \sigma_{\eta_t}^2 / \sigma_{\varepsilon_t}^2$ . In our model, the error of the state equation is constant. So, we consider the SNRs for 1/0.05, 1/0.1, 1/0.8, 1/1.3 and 1/4 by changing the variance in the error ( $\sigma_{\varepsilon_t}^2$ ) in the observation equation. And these five sample paths of the TV-INAR(1) model are plotted in Figure 1, respectively. We can see that the sample path of  $y_t$  is unsteady, and the variation in parameter combinations results in a change in the sample dispersion of the samples.



**Figure 1.** Sample paths of the TV-INAR(1) model for  $\lambda = \{0.05, 0.1, 0.8, 1.3, 4\}$ .

Then, we should notice the choice of the parameter's initial value. The initial value of  $\alpha_t$  follows a known normal distribution, which is mentioned in Section 2.1. In practice, it is difficult to gauge the true value of its mean and the variance of the distribution. For simplicity, we assume  $\alpha_0$  is known and nonstochastic, i.e.,  $\alpha_0 = 0(\mu_0 = 0, \sigma_0^2 = 0)$ . This assumption brings great convenience in simulation studies and does not affect numerical results. This can be observed in Proposition 3 when the sample size is sufficiently large.

Our simulation concerns a first-order integer-valued autoregressive model with timevarying parameters (TV-INAR(1)). For the first step estimation, let  $\alpha_{t,n}$  denote the true value of the parameter in the data generation process and  $\hat{\alpha}_{t,n}$  denote the Kalman-smoothed estimate of the parameter. We compute the sample means and sample standard deviations of the true and estimated values, respectively:

$$ar{lpha}_n = rac{1}{T}\sum_{t=1}^T lpha_{t,n}, \qquad sd(lpha_n) = \sqrt{rac{1}{T-1}\sum_{t=1}^T (lpha_{t,n} - ar{lpha}_n)^2},$$

$$\bar{\hat{\alpha}}_n = \frac{1}{T} \sum_{t=1}^T \hat{\alpha}_{t,n}, \qquad sd(\hat{\alpha}_n) = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (\hat{\alpha}_{t,n} - \bar{\hat{\alpha}}_n)^2}.$$

After N = 100 replications, the averages of the above four indicators are found, respectively:

$$\bar{\alpha} = \frac{1}{N} \sum_{n=1}^{N} \bar{\alpha}_n, \qquad \bar{s} = \frac{1}{N} \sum_{n=1}^{N} sd(\alpha_n),$$
$$\bar{\hat{\alpha}} = \frac{1}{N} \sum_{n=1}^{N} \bar{\hat{\alpha}}_n, \qquad \bar{\hat{s}} = \frac{1}{N} \sum_{n=1}^{N} sd(\hat{\alpha}_n).$$

Let  $bias_{\alpha} = |\bar{\alpha} - \bar{\hat{\alpha}}|$ . We take the difference between  $\bar{\alpha}$  and  $\bar{\hat{\alpha}}$  directly to evaluate the effectiveness of the Kalman-smoothed estimate approach. In addition, denote

$$rat = \frac{1}{N} \sum_{n=1}^{N} \frac{sd(\hat{\alpha}_n)}{sd(\alpha_n)}$$

as the mean of the ratio of the standard deviation of  $\hat{\alpha}_n$  to the standard deviation of the real process  $\alpha_n$ . The quality of the estimator depends on whether *rat* is close to one. This is similar to the criterion in [21].

Next, we evaluate the performance of the second step estimation, which is to apply the  $\hat{\alpha}_t$  and the conditional least squares (CLS) method to estimate the parameter  $\lambda$  in the model (2). As we mentioned above, the true value of  $\lambda$  is considered  $\lambda = \{0.05, 0.1, 0.8, 1.3, 4\}$ . To evaluate the estimation performance, besides considering  $bias_{\lambda} = |\lambda - \hat{\lambda}|$ , two other indicators were selected, which are the mean absolute deviation (MAD) and mean square error (MSE), as defined below:

$$MAD = \frac{1}{N} \sum_{n=1}^{N} |\hat{\lambda}_n - \lambda|, \quad MSE = \frac{1}{N} \sum_{n=1}^{N} (\hat{\lambda}_n - \lambda)^2,$$

where  $\hat{\lambda}_n$  is the estimation result of  $\lambda$  at the *n*th replication. Then, considering various sample sizes and initial parameter values, the simulation results of the two-step estimation process are listed in Tables 1 and 2.

**Table 1.** Simulation results of  $\alpha$ .

λ	SNR	Т	ā	ā	bias <sub>α</sub>	rat
0.05	20	50	0.7132	0.7790	0.0658	0.4982
		100	0.5941	0.6414	0.0473	0.5305
		200	-0.8605	-0.8671	0.0066	0.7479
0.1	10	50	-0.2026	-0.1137	0.0889	0.6547
		100	-0.9860	-0.9333	0.0527	0.7899
		200	-0.6248	-0.6175	0.0073	0.9138
0.8	1.25	50	0.3801	0.4772	0.0971	0.8308
		100	0.4309	0.3666	0.0643	0.8965
		200	0.6438	0.6154	0.0284	0.9276
1.3	0.77	50	-0.5826	-0.4699	0.1127	0.8869
		100	0.3095	0.2246	0.0849	0.9304
		200	-0.2924	-0.3385	0.0461	1.0483
4	0.25	50	-0.3752	-0.2143	0.1609	0.9025
		100	0.7106	0.5856	0.1250	0.9514
		200	0.6634	0.5762	0.0872	1.1287

10	of	17

λ	SNR	Т	$\hat{\lambda}$	$bias_{\lambda}$	MAD	MSE
0.05	20	50	0.0621	0.0121	0.0256	0.0018
		100	0.0587	0.0087	0.0179	0.0015
		200	0.0441	0.0059	0.0103	0.0010
0.1	10	50	0.0803	0.0197	0.0407	0.0023
		100	0.0858	0.0142	0.0373	0.0021
		200	0.0906	0.0094	0.0211	0.0017
0.8	1.25	50	0.6632	0.1368	0.1014	0.0751
		100	0.6829	0.1171	0.0762	0.0657
		200	0.7157	0.0843	0.0399	0.0481
1.3	0.77	50	1.1347	0.1652	0.1729	0.1251
		100	1.1827	0.1173	0.1094	0.0727
		200	1.2090	0.0910	0.0529	0.0512
4	0.25	50	4.1880	0.1880	0.1931	0.1634
		100	3.8397	0.1603	0.1652	0.1252
		200	3.8903	0.1097	0.1140	0.0938

**Table 2.** Simulation results of  $\lambda$ .

It is shown that the smaller the variance ( $\lambda$ ) of the error term, the smaller the biases of estimates  $bias_{\alpha}$  and  $bias_{\lambda}$ . This implies that the Kalman smoothing and CLS approaches work better in the sense of *bias*. In the first-step estimation, the larger the variance  $(\lambda)$ , the closer *rat* is to one. In this sense, the Kalman smoothing method is the best if  $\lambda$  is equal to 1.3. This suggests that only using one criterion to measure the effectiveness of the estimation method is insufficient. In addition, by increasing the sample size,  $bais_{\alpha}$  is smaller and *rat* is closer to one. This shows that the Kalman-smoothed estimate is closer to the true process. From the perspective of SNR, the larger the SNR, the smaller the bias in the estimation. The smaller the SNR, the closer the estimated median sample variance is to the true process. In summary, when the SNR is around 1, the estimation is relatively good. In the second-step estimation, the values of MAD and MSE are small, suggesting a relatively acceptable estimation effect. The estimation results as a whole show a trend that the larger the SNR, the smaller the estimation error. Consequently, the CLS estimate method works better. Additionally, when the sample size T increases, the corresponding MAD and MSE gradually decrease, and the estimates of the parameter gradually converge to the true values. In conclusion, the proposed two-step estimation method is a credible approach.

### 5. Case Application

In this section, we apply the model and method of Section 3 to predict real-time series data. The data set is a count time series of offense data, obtainable from the NSW Bureau of Crime Statistics and Research covering January 2000 to December 2009. The observations represent monthly counts of Sexual offenses in Ballina, NSW, Australia, which comprise 120 monthly observations. Figure 2 shows the time plot and partial autocorrelation function (PACF). It shows that the data are nonstationary and are first-order autocorrelated, which is an indication that it could be reasonable to model this data set with our model. The descriptive statistics of the data are displayed in Table 3. Next, we compared our TV-INAR(1) model with the INAR(1) model to fit the data set. In general, it is expected that the better model to fit the data presents smaller values for -log-likelihood, AIC and BIC. From the results in Table 4, we can conclude that the proposed model fits the data better.

Table 3. Descriptive statistics for sexual offense data.

Sample Size	Minimum	Maximum	Median	Mean	Variance
120	0	32	4	5.1	14.225



Figure 2. Time plot and PACF of sexual offense data for the period 2000–2009.

Table 4.	Estimates of th	e parameters and	l goodness-of	f-fit statistics	for the o	offense data.
			()			

Model	$\bar{\hat{lpha}}(\hat{lpha})$	$\hat{\lambda}$	-log-likelihood	AIC	BIC	RMSE
TV- INAR(1)	0.1574	2.5928	315.96	635.92	641.49	2.5278
INAR(1)	0.2299	3.9236	317.51	639.01	644.59	3.2137

For the prediction, the predicted values of the offense data are given by

$$E(Y_{t+k}|\mathcal{F}_{Y,t}) = \prod_{h=1}^{k} \frac{e^{\hat{\alpha}_{t+h}}}{1+e^{\hat{\alpha}_{t+h}}} Y_t + \sum_{h=1}^{k-1} \prod_{m=0}^{h-1} \frac{e^{\hat{\alpha}_{t+k-m}}}{1+e^{\hat{\alpha}_{t+k-m}}} \hat{\lambda} + \hat{\lambda}.$$

Specifically, the one-step-ahead conditional expectation point predictor is given by

$$E(Y_t|\mathcal{F}_{Y,t-1}) = \frac{e^{\hat{\alpha}_t}}{1+e^{\hat{\alpha}_t}}Y_{t-1} + \hat{\lambda}.$$

We compute the root mean square of the prediction errors (RMSEs) of the data in the past 6 months, with the RMSE defined as

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2}.$$

The estimators and RMSE results are also shown in Table 4. To analyze the adequacy of the model, we analyze the standardized Pearson residuals, defined as

$$e_t = \frac{Y_t - E(Y_t | \mathcal{F}_{Y,t-1})}{\sqrt{Var(Y_t | \mathcal{F}_{Y,t-1})}},$$

with  $t \in (1, \dots, n)$ . For our model, the mean and variance of the Pearson residuals are 0.8022 and 1.2000, respectively. As discussed in ref. [28], for an adequately chosen model, the variance of the residuals should take a value close to 1. And 1.2 seems to be close to 1. Therefore, we conclude that the proposed model fits the data well.

#### 6. Conclusions

This paper proposes a time-varying integer-valued autoregressive model with autoregressive coefficients driven by a logistic regression structure. It can be more flexible and efficient to handle integer-valued discrete and even nonstationary data. Some statistical properties of the model are derived, such as mean, variance, covariance, conditional mean and conditional variance. A two-step estimation method was introduced. Since the model is in a state-space form, the Kalman smoothing method would instead be used to estimate the time-varying parameter. In the first step, the Kalman-smoothed estimate of the state vector is obtained using the information of the known observation variables. In the second step, the estimate from the previous step and the CLS method are used to obtain the estimate of the error term parameter. The analysis formulae of the estimates in the two steps are both derived. The solution challenge lies in calculating the covariance matrix and the correlation between the variables. The advantage of this method is that the approximate estimation of the unknown parameters can be obtained from all the observable variables, and the approach has superior performance in practical applications, even in the case of nonlinear and non-Gaussian errors. The proposed method can also estimate the time-varying parameters well. In addition, an application of forecasting a real data set is presented. The results suggest that the TV-INAR(1) model is more suitable for practical data sets. In model (3), if the variances of the error terms  $\varepsilon_t$  and  $\eta_t$  are allowed to be time-dependent, the model will be regarded as a stochastic volatility model. This is a topic for discussion. Moreover, extending the research outcomes to the *p*-order model TV-INAR(*p*) is one direction of future research.

**Author Contributions:** Every author contributed equally to the development of this paper. Conceptualization, D.W.; ethodology, Y.P.; software, Y.P.; investigation, Y.P.; writing—original draft preparation, Y.P.; writing—review and editing, D.W. and M.G.; supervision, D.W. and M.G.; funding acquisition, D.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (Nos. 12271231, 12001229, 11901053).

**Data Availability Statement:** Publicly available data sets were analyzed in this study. These data can be found here: [https://data.nsw.gov.au/data/dataset/nsw-criminal-court-statistics] (accessed on 3 July 2023).

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

#### Appendix A. Proof of Proposition 2

From models (3) and (2), we have

(i) 
$$E(\alpha) = E[(\mu + \eta)C] = \mu C$$

- (ii)  $E(\mathbf{Y}) = E(\boldsymbol{\beta} \circ \mathbf{Z} + \boldsymbol{\varepsilon}) = E(\boldsymbol{\beta})\mathbf{Z} + \boldsymbol{\Lambda};$
- (iii)  $Var(\boldsymbol{Y}) = Var(\boldsymbol{\beta} \circ \boldsymbol{Z}) + diag(\lambda, \cdots, \lambda);$

$$Var(\boldsymbol{\beta} \circ \boldsymbol{Z}) = E\{[\boldsymbol{\beta} \circ \boldsymbol{Z} - E(\boldsymbol{\beta} \circ \boldsymbol{Z})]'[\boldsymbol{\beta} \circ \boldsymbol{Z} - E(\boldsymbol{\beta} \circ \boldsymbol{Z})]\} \\ = \begin{pmatrix} Cov(\beta_1 \circ Z_1, \beta_1 \circ Z_1) & \cdots & Cov(\beta_1 \circ Z_1, \beta_T \circ Z_T) \\ \vdots & \ddots & \vdots \\ Cov(\beta_T \circ Z_T, \beta_1 \circ Z_1) & \cdots & Cov(\beta_T \circ Z_T, \beta_T \circ Z_T) \end{pmatrix}.$$

The main diagonal elements are

$$Var(\beta_k \circ Z_k) = Var(\beta_k)Z_k^2 + E[\beta_k(1-\beta_k)]Z_k, \ k = 1, \cdots, T,$$

and the rest of the elements are

$$\begin{aligned} & Cov(\beta_{k-l} \circ Z_{k-l}, \beta_k \circ Z_k) \\ &= E\{E[(\beta_{k-l} \circ Z_{k-l})(\beta_k \circ Z_k)|\beta_{k-l}, \beta_k]\} - E[E(\beta_{k-l} \circ Z_{k-l}|\beta_{k-l})]E[E(\beta_k \circ Z_k|\beta_k)] \\ &= E\{E[(B'_1 + \dots + B'_{Z_{k-l}})(B''_1 + \dots + B''_{Z_k})|\beta_{k-l}, \beta_k]\} - E(\beta_{k-l})Z_{k-l}E(\beta_k)Z_k \\ &= E[E(B'_1B''_1 + \dots + B'_1B''_{Z_k}) + \dots + E(B'_{Z_{k-l}}B''_1 + \dots + B'_{Z_{k-l}}B''_{Z_k})|\beta_{k-l}, \beta_k] - E(\beta_{k-l}) \\ &Z_{k-l}E(\beta_k)Z_k \\ &= E(\beta_{k-l}\beta_k)Z_{k-l}Z_k - E(\beta_{k-l})E(\beta_k)Z_{k-l}Z_k. \end{aligned}$$

### Therefore,

$$Var(\mathbf{Y}) = \mathbf{Z}Var(\boldsymbol{\beta})\mathbf{Z}' + E[diag(1 - \beta_1, \dots, 1 - \beta_t)diag(\beta_1, \dots, \beta_T)]\mathbf{Z} + diag(\lambda, \dots, \lambda).$$
  
(iv) 
$$Cov(\mathbf{Y}, \boldsymbol{\alpha}) = Cov(\boldsymbol{\beta} \circ \mathbf{Z}, \boldsymbol{\alpha}) = \begin{pmatrix} Cov(\beta_1 \circ Z_1, \alpha_1) & \cdots & Cov(\beta_1 \circ Z_1, \alpha_T) \\ \vdots & \ddots & \vdots \\ Cov(\beta_T \circ Z_T, \alpha_1) & \cdots & Cov(\beta_T \circ Z_T, \alpha_T) \end{pmatrix}.$$

For any  $i, j \in (1, \cdots, T)$ ,

$$Cov(\beta_i \circ Z_i, \alpha_j) = E\{E[(\beta_i \circ Z_i)\alpha_j | \beta_i, \beta_j]\} - E[E(\beta_i \circ Z_i | \beta_i)]E[E(\alpha_j | \beta_j)]$$
  
=  $E[(B_1 + \dots + B_{Z_i})\alpha_j | \beta_i, \beta_j] - E(\beta_i)Z_i\mu_0$   
=  $E[B_1\alpha_j + \dots + B_{Z_j}\alpha_j | \beta_i, \beta_j] - E(\beta_i)Z_i\mu_0$   
=  $E[\beta_i\alpha_j]Z_i - E(\beta_i)Z_i\mu_0.$ 

Thus,  $Cov(\mathbf{Y}, \boldsymbol{\alpha}) = E(\boldsymbol{\beta}'\boldsymbol{\alpha})\mathbf{Z} - \mathbf{Z}E(\boldsymbol{\beta}')\boldsymbol{\mu}\mathbf{C}.$ 

## Appendix B. Proof of Proposition 3

(1)  $E(\beta_k) = E(\frac{e^{\alpha_k}}{1+e^{\alpha_k}}) = 1 - E(\frac{1}{1+e^{\alpha_k}}) = 1 - E(\frac{1}{1+X})$ . Denote  $X \sim LN(\mu, \sigma^2)$ . Using substitution and Taylor's expansion, we obtain

$$\begin{split} & E\Big(\frac{1}{1+\chi}\Big) \\ &= \int_{0}^{\infty} \frac{1}{1+x} \frac{1}{\sqrt{2\pi\sigma}x} \exp\Big\{-\frac{(\ln x - \mu)^{2}}{2\sigma^{2}}\Big\} dx \\ &= \int_{-\infty}^{\infty} \frac{1}{1+e^{\sqrt{2\sigma}s+\mu}} \frac{1}{\sqrt{2\pi\sigma}e^{\sqrt{2\sigma}s+\mu}} e^{-s^{2}} e^{\sqrt{2}\sigma s+\mu} \sqrt{2}\sigma ds \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-s^{2}} (1+e^{\sqrt{2}\sigma s+\mu})^{-1} ds \\ &= \frac{1}{\sqrt{\pi}} \left[\int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma}}} e^{-s^{2}} (1+e^{\sqrt{2}\sigma s+\mu})^{-1} ds + \int_{-\frac{\mu}{\sqrt{2\sigma}}}^{\infty} e^{-s^{2} - \sqrt{2}\sigma s-\mu} (1+e^{-\sqrt{2}\sigma s-\mu})^{-1} ds \right] \\ &= \frac{1}{\sqrt{\pi}} \left[\int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma}}} e^{-s^{2}} ds - e^{\mu + \frac{\sigma^{2}}{2}} \int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma}}} e^{-(s-\frac{\sigma}{\sqrt{2}})^{2}} ds + e^{2\mu + 2\sigma^{2}} \int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma}}} e^{-(s-\sqrt{2}\sigma)^{2}} ds - \cdots \right] \\ &+ e^{-\mu + \frac{\sigma^{2}}{2}} \int_{-\frac{\mu}{\sqrt{2\sigma}}}^{\infty} e^{-(s+\frac{\sigma}{\sqrt{2}})^{2}} ds - e^{-2\mu + 2\sigma^{2}} \int_{-\frac{\mu}{\sqrt{2\sigma}}}^{\infty} e^{-(s+\sqrt{2}\sigma)^{2}} ds + e^{-3\mu + \frac{9}{2}\sigma^{2}} \int_{-\frac{\mu}{\sqrt{2\sigma}}}^{\infty} e^{-(s+\frac{3}{\sqrt{2}}\sigma)^{2}} ds - \cdots \right] \\ &= \left[ \Phi\Big(-\frac{\mu}{\sigma}\Big) - e^{\mu + \frac{\sigma^{2}}{2}} \Phi\Big(-\frac{\mu}{\sigma} - \sigma\Big) + e^{2\mu + 2\sigma^{2}} \Phi\Big(-\frac{\mu}{\sigma} - 2\sigma\Big) - \cdots \right] + \left[ e^{-\mu + \frac{\sigma^{2}}{2}} \Phi\Big(\frac{\mu}{\sigma} - \sigma\Big) - e^{-2\mu + 2\sigma^{2}} \Phi\Big(\frac{\mu}{\sigma} - z\sigma\Big) + e^{-3\mu + \frac{9}{2}\sigma^{2}} \Phi\Big(\frac{\mu}{\sigma} - \sigma\Big) - \cdots \right] \right] \\ &= \sum_{r=0}^{\infty} \left[ (-1)^{r} e^{r\mu + \frac{r^{2}}{2}\sigma^{2}} \Phi\Big(-\frac{\mu}{\sigma} - r\sigma\Big) + (-1)^{r+1} e^{-(r+1)\mu + \frac{(r+1)^{2}}{2}\sigma^{2}}} \Phi\Big(\frac{\mu}{\sigma} - (r+1)\sigma\Big) \right]. \end{split}$$

Then,

$$\begin{split} E(\beta_k) = & 1 - \sum_{r=0}^{\infty} \left[ (-1)^r e^{r\mu_0 + \frac{r^2}{2} (\sigma_0^2 + k)} \Phi\left( - \frac{\mu_0}{\sqrt{\sigma_0^2 + k}} - r\sqrt{\sigma_0^2 + k} \right) + (-1)^{r+1} \\ & e^{-(r+1)\mu_0 + \frac{(r+1)^2}{2} (\sigma_0^2 + k)} \Phi\left( \frac{\mu_0}{\sqrt{\sigma_0^2 + k}} - (r+1)\sqrt{\sigma_0^2 + k} \right) \right]. \end{split}$$

$$\begin{aligned} &(2) \ Var(\beta_k) = Var(1 - \frac{1}{1 + e^{\alpha_k}}) = Var(\frac{1}{1 + e^{\alpha_k}}) = E\left(\frac{1}{1 + X}\right)^2 - \left[E\left(\frac{1}{1 + X}\right)\right]^2, \\ &\left[E\left(\frac{1}{1 + X}\right)\right]^2 \\ &= \int_0^\infty \frac{1}{(1 + x)^2} \frac{1}{\sqrt{2\pi\sigma x}} \exp\left\{-\frac{(\ln x - \mu)^2}{2\sigma^2}\right\} dx \\ &= \int_{-\infty}^\infty \frac{1}{(1 + e^{\sqrt{2\sigma s} + \mu})^2} \frac{1}{\sqrt{2\pi\sigma e^{\sqrt{2\sigma s} + \mu}}} e^{-s^2} e^{\sqrt{2\sigma s} + \mu} \sqrt{2\sigma} ds \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^\infty e^{-s^2} (1 + e^{\sqrt{2\sigma s} + \mu})^{-2} ds \\ &= \frac{1}{\sqrt{\pi}} \left[\int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma}}} e^{-s^2} (1 + e^{\sqrt{2\sigma s} + \mu})^{-2} ds + \int_{-\frac{\mu}{\sqrt{2\sigma}}}^\infty e^{-s^2 - 2\sqrt{2\sigma s} - 2\mu} (1 + e^{-\sqrt{2\sigma s} - \mu})^{-2} ds \right] \\ &= \frac{1}{\sqrt{\pi}} \left[\int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma}}} e^{-s^2} ds - 2e^{\mu + \frac{q^2}{2}} \int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma}}} e^{-(s - \frac{\pi}{\sqrt{2}})^2} ds + 3e^{2\mu + 2\sigma^2} \int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma}}} e^{-(s - \sqrt{2}\sigma)^2} ds - \cdots \\ &+ e^{-2\mu + 2\sigma^2} \int_{-\frac{\mu}{\sqrt{2\sigma}}}^\infty e^{-(s + \sqrt{2}\sigma)^2} ds - 2e^{-3\mu + \frac{9}{2}\sigma^2} \int_{-\frac{\mu}{\sqrt{2\sigma}}}^\infty e^{-(s + \frac{3}{\sqrt{2}}\sigma)^2} ds + 3e^{-4\mu + \frac{16}{2}\sigma^2} \int_{-\frac{\mu}{\sqrt{2\sigma}}}^\infty e^{-(s + \frac{4}{\sqrt{2}}\sigma)^2} ds - \cdots \right] \\ &= \left[\Phi\left(-\frac{\mu}{\sigma}\right) - 2e^{\mu + \frac{\sigma^2}{2}} \Phi\left(-\frac{\mu}{\sigma} - \sigma\right) + 3e^{2\mu + 2\sigma^2} \Phi\left(-\frac{\mu}{\sigma} - 2\sigma\right) - \cdots \right] + \left[e^{-2\mu + 2\sigma^2} \Phi\left(\frac{\mu}{\sigma} - 3\sigma\right) + 3e^{-4\mu + \frac{16}{2}\sigma^2} \Phi\left(\frac{\mu}{\sigma} - 4\sigma\right) - \cdots \right] \\ &= \sum_{r=0}^\infty (-1)^r (r+1) \left[e^{r\mu + \frac{r^2}{2}\sigma^2} \Phi\left(-\frac{\mu}{\sigma} - r\sigma\right) + e^{-(r+2)\mu + \frac{(r+2)^2}{2}\sigma^2} \Phi\left(\frac{\mu}{\sigma} - (r+2)\sigma\right)\right]. \end{aligned}$$

Then,

$$\begin{split} &Var(\beta_k) = \sum_{r=0}^{\infty} (-1)^r (r+1) \left[ e^{r\mu_0 + \frac{r^2}{2} (\sigma_0^2 + k)} \Phi \Big( -\frac{\mu_0}{\sqrt{\sigma_0^2 + k}} - r\sqrt{\sigma_0^2 + k} \Big) + e^{-(r+2)\mu_0 + \frac{(r+2)^2}{2} (\sigma_0^2 + k)} \\ &\Phi \Big( \frac{\mu_0}{\sqrt{\sigma_0^2 + k}} - (r+2)\sqrt{\sigma_0^2 + k} \Big) \right] - \left\{ \sum_{r=0}^{\infty} \left[ (-1)^r e^{r\mu_0 + \frac{r^2}{2} (\sigma_0^2 + k)} \Phi \Big( -\frac{\mu_0}{\sqrt{\sigma_0^2 + k}} - r\sqrt{\sigma_0^2 + k} \Big) \right. \\ &+ (-1)^{r+1} e^{-(r+1)\mu_0 + \frac{(r+1)^2}{2} (\sigma_0^2 + k)} \Phi \Big( \frac{\mu_0}{\sqrt{\sigma_0^2 + k}} - (r+1)\sqrt{\sigma_0^2 + k} \Big) \right] \right\}^2. \end{split}$$

$$(3) E[\beta_{k}(1-\beta_{k})] = E\left(\frac{1}{1+X}\right) - E\left(\frac{1}{1+X}\right)^{2} \\ = e^{\mu + \frac{\sigma^{2}}{2}} \Phi\left(-\frac{\mu}{\sigma} - \sigma\right) - 2e^{2\mu + 2\sigma^{2}} \Phi\left(-\frac{\mu}{\sigma} - 2\sigma\right) + 3e^{3\mu + \frac{9}{2}\sigma^{2}} \Phi\left(-\frac{\mu}{\sigma} - 3\sigma\right) - \cdots \\ + e^{-\mu + \frac{\sigma^{2}}{2}} \Phi\left(\frac{\mu}{\sigma} - \sigma\right) - 2e^{-2\mu + 2\sigma^{2}} \Phi\left(\frac{\mu}{\sigma} - 2\sigma\right) + 3e^{-3\mu + \frac{9}{2}\sigma^{2}} \Phi\left(\frac{\mu}{\sigma} - 3\sigma\right) - \cdots \\ = \sum_{r=1}^{\infty} (-1)^{r+1} r \left[ e^{r\mu + \frac{r^{2}}{2}\sigma^{2}} \Phi\left(-\frac{\mu}{\sigma} - r\sigma\right) + e^{-r\mu + \frac{r^{2}}{2}\sigma^{2}} \Phi\left(\frac{\mu}{\sigma} - r\sigma\right) \right] \\ = \sum_{r=1}^{\infty} (-1)^{r+1} r \left[ e^{r\mu_{0} + \frac{r^{2}}{2}(\sigma_{0}^{2} + k)} \Phi\left(-\frac{\mu_{0}}{\sqrt{\sigma_{0}^{2} + k}} - r\sqrt{\sigma_{0}^{2} + k}\right) + e^{-r\mu_{0} + \frac{r^{2}}{2}(\sigma_{0}^{2} + k)} \right].$$

(4)

$$\beta_{i}\alpha_{j} = \frac{e^{\alpha_{i}}}{1 + e^{\alpha_{i}}}\alpha_{j} = \begin{cases} \frac{\exp\left\{\alpha_{0} + \sum\limits_{l=1}^{j} \eta_{l} + \sum\limits_{l=j+1}^{i} \eta_{l}\right\}}{1 + \exp\left\{\alpha_{0} + \sum\limits_{l=1}^{j} \eta_{l} + \sum\limits_{l=j+1}^{i} \eta_{l}\right\}} (\alpha_{0} + \sum\limits_{l=1}^{j} \eta_{l}), & j < i, \\ \frac{\exp\left\{\alpha_{0} + \sum\limits_{l=1}^{i} \eta_{l}\right\}}{1 + \exp\left\{\alpha_{0} + \sum\limits_{l=1}^{i} \eta_{l}\right\}} (\alpha_{0} + \sum\limits_{l=1}^{i} \eta_{l} + \sum\limits_{l=i+1}^{j} \eta_{l}), & j \ge i, \end{cases}$$

where  $\alpha_0 + \sum_{l=1}^{j} \eta_l = \alpha_j \sim N(\mu_0, \sigma_0^2 + j), \sum_{l=j+1}^{i} \eta_l \sim N(0, i-j)$ , and they are independent of

each other. The same applies when  $j \ge i$ . Let Y denote  $e^{\sum_{l=j+1}^{i} \eta_l}$  and  $\tau^2$  denote its variance. Thus,  $Y \sim LN(0, \tau^2)$ .

When j < i,

$$\begin{split} E(\beta_i \alpha_j) &= \int_0^\infty \int_0^\infty \frac{xy \ln x}{1+xy} \frac{1}{\sqrt{2\pi\sigma_x}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi\tau_y}} e^{-\frac{(\ln y)^2}{2\tau^2}} dx dy \\ &= \int_0^\infty \frac{\ln x}{\sqrt{2\pi\sigma}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}} dx \int_0^\infty \frac{1}{\sqrt{2\pi\tau}(1+xy)} e^{-\frac{(\ln y)^2}{2\tau^2}} dy, \end{split}$$

where

$$\begin{aligned} &\frac{1}{\sqrt{2\pi\tau}} \int_{0}^{\infty} \frac{1}{1+xy} e^{-\frac{(\ln y)^2}{2\tau^2}} dy \\ &= \frac{1}{\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} \frac{1}{1+xe^{\sqrt{2\tau}s}} e^{-s^2} e^{\sqrt{2}\tau s} \sqrt{2}\tau ds \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-s^2 + \sqrt{2}\tau s} (1+xe^{\sqrt{2}\tau s})^{-1} ds \\ &= \frac{1}{\sqrt{\pi}} \left[ \int_{-\infty}^{0} e^{-s^2 + \sqrt{2}\tau s} (1+xe^{\sqrt{2}\tau s})^{-1} ds + \int_{0}^{\infty} x^{-1} e^{-s^2} (1+x^{-1}e^{-\sqrt{2}\tau s})^{-1} ds \right] \\ &= \frac{1}{\sqrt{\pi}} \left[ e^{\frac{\tau^2}{2}} \int_{-\infty}^{0} e^{-(s-\frac{\tau}{\sqrt{2}})^2} ds - xe^{2\tau^2} \int_{-\infty}^{0} e^{-(s-\sqrt{2}\tau)^2} ds + x^2 e^{\frac{9}{2}\tau^2} \int_{-\infty}^{0} e^{-(s-\frac{3}{\sqrt{2}}\tau)^2} ds - \cdots \\ &+ x^{-1} \int_{0}^{\infty} e^{-s^2} ds - x^{-2} e^{\frac{\tau^2}{2}} \int_{0}^{\infty} e^{-(s+\frac{\tau}{\sqrt{2}})^2} ds + x^{-3} e^{2\tau^2} \int_{0}^{\infty} e^{-(s+\sqrt{2}\tau)^2} ds - \cdots \\ &+ x^{-3} e^{2\tau^2} \Phi(-2\tau) - xe^{2\tau^2} \Phi(-2\tau) + x^2 e^{\frac{9}{2}\tau^2} \Phi(-3\tau) - \cdots + x^{-1} \Phi(0) - x^{-2} e^{\frac{\tau^2}{2}} \Phi(-\tau) \\ &+ x^{-3} e^{2\tau^2} \Phi(-2\tau) - \cdots . \end{aligned}$$

Thus,

$$\begin{split} E(\beta_{i}\alpha_{j}) &= \left(\frac{\sigma}{\sqrt{2\pi}} + \mu\right)\Phi(0) + \left[ (e^{\mu} - e^{-\mu})(\frac{\sigma}{\sqrt{2\pi}} + \mu) + (e^{\mu} + e^{-\mu})\sigma^{2} \right] e^{\frac{\sigma^{2} + \tau^{2}}{2}} \Phi(-\tau) - \left[ (e^{2\mu} - e^{-2\mu}) + (\frac{\sigma}{\sqrt{2\pi}} + \mu) + (e^{2\mu} + e^{-2\mu})2\sigma^{2} \right] e^{2\sigma^{2} + 2\tau^{2}} \Phi(-2\tau) + \left[ (e^{3\mu} - e^{-3\mu})(\frac{\sigma}{\sqrt{2\pi}} + \mu) + (e^{3\mu} + e^{-3\mu})3\sigma^{2} \right] e^{\frac{\sigma}{2}(\sigma^{2} + \tau^{2})} \Phi(-3\tau) - \cdots \\ &= \frac{1}{2} \left( \frac{\sigma}{\sqrt{2\pi}} + \mu \right) + \sum_{r=1}^{\infty} (-1)^{r+1} \left[ (e^{r\mu} - e^{-r\mu})(\frac{\sigma}{\sqrt{2\pi}} + \mu) + (e^{r\mu} + e^{-r\mu})r\sigma^{2} \right] \left( \frac{\sigma}{\sqrt{2\pi}} + \mu \right) \Phi(-r\tau) \\ &= \frac{1}{2} \left( \sqrt{\frac{\sigma_{0}^{2} + j}{2\pi}} + \mu_{0} \right) \sum_{r=1}^{\infty} (-1)^{r+1} \left[ (e^{r\mu_{0}} - e^{-r\mu_{0}})(\sqrt{\frac{\sigma_{0}^{2} + j}{2\pi}} + \mu_{0}) + (e^{r\mu_{0}} + e^{-r\mu_{0}})r(\sigma_{0}^{2} + j) \right] \\ &= e^{\frac{r^{2}}{2}(\sigma_{0}^{2} + i)} \Phi(-r\sqrt{i-j}). \end{split}$$

Similarly, when  $j \ge i$ ,

$$E(\beta_i \alpha_j) = E\left(\frac{X \ln X}{1+X} + \frac{XY}{1+X}\right),$$

X and Y are independent, and E(Y) = 0. Therefore,

$$\begin{split} & E(\beta_i \alpha_j) \\ &= E\left(\frac{X \ln X}{1 + X}\right) \\ &= \int_0^\infty \frac{x \ln x}{1 + x} \frac{1}{\sqrt{2\pi\sigma}x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^\infty \frac{\sqrt{2\sigma}s + \mu}{1 + e^{\sqrt{2\sigma}s + \mu}} e^{-s^2 + \sqrt{2\sigma}s + \mu} ds \\ &= \frac{\sqrt{2\sigma}}{\sqrt{\pi}} \int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma}}} se^{-s^2 + \sqrt{2\sigma}s + \mu} (1 + e^{\sqrt{2}\sigma s + \mu})^{-1} ds + \frac{\sqrt{2\sigma}}{\sqrt{\pi}} \int_{-\frac{\mu}{\sqrt{2\sigma}}}^\infty se^{-s^2} (1 + e^{-\sqrt{2}\sigma s - \mu})^{-1} ds \\ &+ \frac{\mu}{\sqrt{\pi}} \int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma}}} e^{-s^2 + \sqrt{2}\sigma s + \mu} (1 + e^{\sqrt{2}\sigma s + \mu})^{-1} ds + \frac{\mu}{\sqrt{\pi}} \int_{-\frac{\mu}{\sqrt{2\sigma}}}^\infty e^{-s^2} (1 + e^{-\sqrt{2}\sigma s - \mu})^{-1} ds \\ &= \sqrt{2}(\mu + \sigma^2)e^{\mu + \frac{\sigma^2}{2}} \Phi(-\frac{\mu}{\sigma} - \sigma) - \sqrt{2}(\mu + 2\sigma^2)e^{2\mu + 2\sigma^2} \Phi(-\frac{\mu}{\sigma} - 2\sigma) - \dots + \sqrt{2}\mu \Phi(\frac{\mu}{\sigma}) \\ &- \sqrt{2}(\mu - \sigma^2)e^{-\mu + \frac{\sigma^2}{2}} \Phi(\frac{\mu}{\sigma} - \sigma) + \sqrt{2}(\mu - 2\sigma^2)e^{-2\mu + 2\sigma^2} \Phi(\frac{\mu}{\sigma} - 2\sigma) - \dots \\ &= \sum_{r=1}^\infty (-1)^{r-1}\sqrt{2} \Big[ (\mu + r\sigma^2)e^{r\mu + \frac{r^2}{2}\sigma^2} \Phi(-\frac{\mu}{\sigma} - r\sigma) + (\mu - (r - 1)\sigma^2)e^{-(r - 1)\mu + \frac{(r - 1)^2}{2}\sigma^2} \\ &\Phi(\frac{\mu}{\sigma} - (r - 1)\sigma) \Big] \\ &= \sum_{r=1}^\infty (-1)^{r-1}\sqrt{2} \Big[ (\mu_0 + r(\sigma_0^2 + i))e^{r\mu_0 + \frac{r^2}{2}(\sigma_0^2 + i)} \Phi(-\frac{\mu_0}{\sqrt{\sigma_0^2 + i}} - r\sqrt{\sigma_0^2 + i}) + (\mu_0 - (r - 1) (\sigma_0^2 + i))e^{-(r - 1)\mu_0 + \frac{(r - 1)^2}{2}(\sigma_0^2 + i)} \Phi(\frac{\mu_0}{\sqrt{\sigma_0^2 + i}} - (r - 1)\sqrt{\sigma_0^2 + i}) \Big]. \end{split}$$

(5) According to the state equation in (2),

$$\beta_{k-l}\beta_k = \frac{e^{\alpha_{k-l}}}{1+e^{\alpha_{k-l}}} \cdot \frac{e^{\alpha_k}}{1+e^{\alpha_k}} = \frac{e^{\alpha_{k-l}}}{1+e^{\alpha_{k-l}}} \cdot \frac{e^{\alpha_{k-l}+\sum_{i=k-l+1}^k \eta_i}}{1+e^{\alpha_{k-l}+\sum_{i=k-l+1}^k \eta_i}}.$$

Let *X* denote  $e^{\alpha_{k-l}}$  and *Y* denote  $e^{\sum_{i=k-l+1}^{k} \eta_i}$ . Then,  $X \sim LN(\mu_0, \sigma_0^2 + k - l), Y \sim LN(0, l)$ . For convenience, let  $\mu$  and  $\sigma$  represent the mean and variance of X in the computation. Thus,

$$\begin{split} & E(\beta_{k-l}\beta_k) \\ &= E\left(\frac{X}{1+X} \cdot \frac{XY}{1+XY}\right) \\ &= \int_0^\infty \int_0^\infty \frac{x}{(1+x)(1+xy)} \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(\ln x - \mu)^2}{2\sigma^2}\right\} \frac{1}{\sqrt{2\pi l}} \exp\left\{-\frac{(\ln y)^2}{2l}\right\} dy dx \\ &= \int_0^\infty \frac{x}{\sqrt{2\pi\sigma}(1+x)} \exp\left\{-\frac{(\ln x - \mu)^2}{2\sigma^2}\right\} dx \int_0^\infty \frac{1}{\sqrt{2\pi l}(1+xy)} \exp\left\{-\frac{(\ln y)^2}{2l}\right\} dy, \end{split}$$

where  $\int_0^\infty \frac{1}{\sqrt{2\pi l}(1+xy)} \exp\left\{-\frac{(\ln y)^2}{2l}\right\} dy$  can be obtained from (A1). Therefore,

$$\begin{split} E(\beta_{k-l}\beta_k) \\ &= \frac{1}{\sqrt{2\pi\sigma}} \int_0^\infty \frac{x}{1+x} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}} \left[ e^{\frac{l}{2}} \Phi(-\sqrt{l}) - x e^{2l} \Phi(-2\sqrt{l}) + x^2 e^{\frac{9}{2}l} \Phi(-3\sqrt{l}) - \dots + x^{-1} \right] \\ \Phi(0) - x^{-2} e^{\frac{l}{2}} \Phi(-\sqrt{l}) + x^{-3} e^{2l} \Phi(-2\sqrt{l}) - \dots \right] dx \\ &= \sqrt{2} \Phi(0) \left[ e^{\mu + \frac{\sigma^2}{2}} \Phi(-\frac{\mu}{\sigma} - \sigma) - e^{2\mu + 2\sigma^2} \Phi(-\frac{\mu}{\sigma} - 2\sigma) + e^{3\mu + \frac{9}{2}\sigma^2} \Phi(-\frac{\mu}{\sigma} - 3\sigma) - \dots + \right. \\ \Phi\left(\frac{\mu}{\sigma}\right) - e^{-\mu + \frac{\sigma^2}{2}} \Phi\left(\frac{\mu}{\sigma} - \sigma\right) + e^{-2\mu + 2\sigma^2} \Phi\left(\frac{\mu}{\sigma} - 2\sigma\right) - \dots \right] + \sqrt{2} e^{\frac{l}{2}} \Phi\left(-\sqrt{l}\right) \left[ e^{\mu + \frac{\sigma^2}{2}} - 1 \right] \\ &+ \sqrt{2} e^{2l} \Phi(-2\sqrt{l}) \left[ -e^{2\mu + 2\sigma^2} + e^{\mu + \frac{\sigma^2}{2}} - 1 + e^{-\mu + \frac{\sigma^2}{2}} \right] + \dots \\ &= \frac{1}{\sqrt{2}} \sum_{r=1}^\infty (-1)^{r-1} \left[ e^{r\mu_0 + \frac{r^2}{2}(\sigma_0^2 + k - l)} \Phi\left(-\frac{\mu_0}{\sqrt{\sigma_0^2 + k - l}} - r\sqrt{\sigma_0^2 + k - l}\right) + e^{-(r-1)\mu_0 + \frac{(r-1)^2}{2}(\sigma_0^2 + k - l)} \Phi\left(\frac{\mu_0}{\sqrt{\sigma_0^2 + k - l}} - (r-1)\sqrt{\sigma_0^2 + k - l}\right) \right] + \sum_{r=1}^\infty \sqrt{2} e^{\frac{r^2}{2}l} \Phi\left(-r\sqrt{l}\right) \\ &\left[ \sum_{m=1}^r (-1)^{m-1} e^{m\mu_0 + \frac{m^2}{2}(\sigma_0^2 + k - l)} + (-1)^m e^{-(m-1)\mu_0 + \frac{(m-1)^2}{2}(\sigma_0^2 + k - l)} \right]. \end{split}$$

#### References

- 1. Karlis, D.; Khan, N.M. Models for integer data. Annu. Rev. Stat. Appl. 2023, 10, 297–323. [CrossRef]
- Scotto, M.G.; Weiß, C.H.; Gouveia, S. Thinning-based models in the analysis of integer-valued time series: A review. *Stat. Model.* 2015, *15*, 590–618. [CrossRef]
- 3. Weiß, C.H. An Introduction to Discrete-Valued Time Series; John Wiley and Sons, Inc.: Chichester, UK, 2018.
- 4. McKenzie, E. Some simple models for discrete variate time series. *Water Resour. Bull.* **1985**, *21*, 645–650. [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]
- 6. Steutel, F.W.; Van, H.K. Discrete analogues of self-decomposability and stability. Ann. Probab. 1979, 7, 893–899. [CrossRef]
- Freeland, R.K.; McCabe, B. Analysis of low count time series data by Poisson autoregression. J. Time Ser. Anal. 2004, 25, 701–722. [CrossRef]
- Freeland, R.K.; McCabe, B. Asymptotic properties of CLS estimators in the Poisson AR(1) model. *Stat. Probab. Lett.* 2005, 73, 147–153. [CrossRef]
- 9. Yu, M.; Wang, D.; Yang, K.; Liu, Y. Bivariate first-order random coefficient integer-valued autoregressive processes. J. Stat. Plan. Inference 2020, 204, 153–176. [CrossRef]
- 10. Deng, Z.L.; Gao, Y.; Mao, L.; Li, Y.; Hao, G. New approach to information fusion steady-state Kalman filtering. *Automatica* 2005, 41, 1695–1707. [CrossRef]
- 11. Sbrana, G.; Silvestrini, A. Random coefficient state-space model: Estimation and performance in M3-M4 competitions. *Int. J. Forecast.* **2022**, *38*, 352–366. [CrossRef]
- 12. You, J.; Chen, G. Parameter estimation in a partly linear regression model with random coefficient autoregressive errors. *Commun. Stat. Theory Methods* **2022**, *31*, 1137–1158. [CrossRef]
- 13. Kalman, R.E. A new approach to linear filtering and prediction problems. J. Fluids Eng. 1960, 82, 35–45. [CrossRef]
- 14. Durbin, J.; Koopman, S.J. Time Series Analysis by State Space Methods. In *Oxford Statistical Science Series*, 2nd ed.; Oxford University Press: Oxford, UK, 2012.
- 15. Tang, M.T.; Wang, Y.Y. Asymptotic behavior of random coefficient INAR model under random environment defined by difference equation. *Adv. Differ. Equ.* **2014**, *1*, 99. [CrossRef]
- 16. Koh, Y.B.; Bukhari, N.A.; Mohamed, I. Parameter-driven state-space model for integer-valued time series with application. *J. Stat. Comput. Simul.* **2019**, *89*, 1394–1409. [CrossRef]
- 17. Yang, M.; Cavanaugh, J.E.; Zamba, G.K.D. State-space models for count time series with excess zeros. *Stat. Model.* **2015**, *15*, 70–90. [CrossRef]
- 18. Davis, R.A.; Dunsmuir, W.T.M. State space models for count time series. In *Handbook of Discrete-Valued Time Series*; Davis, R., Holan, S., Lund, R., Ravishanker, N., Eds.; CRC Press: Boca Raton, FL, USA, 2016; pp. 121–144.
- 19. Duncan, D.B.; Horn, S.D. Linear dynamic recursive estimation from the viewpoint of regression analysis. *J. Am. Stat. Assoc.* **1972**, 67, 815–821. [CrossRef]
- 20. Maddala, G.S.; Kim, I. Unit Roots, Cointegration, and Structural Change; Cambridge University Press: Cambridge, UK, 1998.
- 21. Ito, M.; Noda, A.; Wada, T. An Alternative Estimation Method for Time-Varying Parameter Models. *Econometrics* 2022, 10, 23. [CrossRef]
- 22. Omorogbe, J.A. Kalman filter and structural change revisited: An application to foreign trade-economic growth nexus. *Struct. Chang. Econom. Model.* **2019**, *808*, 49–62.
- 23. Bernanke, B.S.; Mihov, I. Measuring monetary policy. Q. J. Econ. 1998, 113, 869–902. [CrossRef]
- Ito, M.; Noda, A.; Wada, T. International stock market efficiency: A non-Bayesian time-varying model approach. *Appl. Econ.* 2014, 46, 2744–2754. [CrossRef]
- 25. Santos, C.; Pereir, I.; Scotto, M.G. On the theory of periodic multivariate INAR processes. Stat. Pap. 2021, 62, 1291–1348. [CrossRef]
- 26. Chattopadhyay, S.; Maiti, R.; Das, S.; Biswas, A. Change-point analysis through INAR process with application to some COVID-19 data. *Stat. Neerl.* **2022**, *76*, 4–34. [CrossRef] [PubMed]
- 27. Primiceri, G.E. Time varying structural vector autoregressions and monetary policy. Rev. Econ. Stud. 2005, 72, 821–852. [CrossRef]
- Aleksandrov, B.; Weiß, C.H. Testing the dispersion structure of count time series using Pearson residuals. AStA Adv. Stat. Anal. 2019, 104, 325–361. [CrossRef]

**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.