

THE NON-STATIONARY BIVARIATE INAR UNDER AN UNCONSTRAINED INTER CORRELATION STRUCTURE

YUVRAJ SUNECHER*

Department of Accounting, Finance and Economics, University of Technology Mauritius, Mauritius

Email: ysunecher@umail.utm.ac.mu

NAUSHAD MAMODE KHAN

Department of Economics and Statistics, University of Mauritius, Mauritius

Email: n.mamodekhan@uom.ac.mu

VANDNA JOWAHEER

Department of Mathematics, University of Mauritius, Mauritius

Email: vandnaj@uom.ac.mu

SUMMARY

It is commonly observed in medical and financial studies that large volume of time series of count data are collected for several variates. The modelling of such time series and the estimation of parameters under such processes are rather challenging since these high dimensional time series are influenced by time-varying covariates that eventually render the data non-stationary. This paper considers the modelling of a bivariate integer-valued autoregressive (BINAR(1)) process where the innovation terms are distributed under non-stationary Poisson moments. Since the full and conditional likelihood approaches are cumbersome in this situation, a Generalized Quasi-likelihood (GQL) approach is proposed to estimate the regression effects while the serial and time-dependent cross correlation effects are handled by method of moments. This new technique is assessed over several simulation experiments and the results demonstrate that GQL yields consistent estimates and is computationally stable since few non-convergent simulations are reported.

Keywords and phrases: Bivariate, Moving Average, Poisson, GQL, Non-Stationarity

AMS Classification: 65C60, 62J12, 62H12, 62J20, 62J10

* Corresponding author

© Institute of Statistical Research and Training (ISRT), University of Dhaka, Dhaka 1000, Bangladesh.

1 Introduction

Most of the univariate time series models that have been developed so far are based on the observation-driven (OD) integer-valued autoregressive (INAR(1)) process (See McKenzie, 1985; Al Osh and Alzaid, 1987; Kim and Park, 2008). In these models, it is relatively parsimonious to formulate the conditional likelihood estimating equations and used the iterative Newton-Raphson or Fisher-scoring algorithm to obtain the parameter estimates of the regression and serial correlation parameters. However, in some important practical settings, the time series of counts are collected for two variables where apart from the serial correlation within each series, there exists as well a cross-correlation between the two series and this lead to a bivariate INAR(1) (BINAR(1)) time series set-up. In addition, these time series of counts are subject to some time-independent and time-varying covariates that induce non-stationary correlations. Thus, the modelling of the BINAR(1) time series is challenging.

In literature, Pedeli and Karlis (2009, 2011) developed the first BINAR(1) time series of counts (CBINAR1M1) which was a simple extension of two INAR(1) processes where the cross correlation between the two series was induced by the innovation terms of the two series. The innovation series were assumed to follow the bivariate Poisson or some bivariate Generalized Poisson distributions. Later, Pedeli and Karlis (2013a) extended the CBINAR1M1 to a full BINAR(1) process (FBINAR1M1) by considering a more complex cross correlation structure where apart from the bivariate distributed innovation terms, the cross correlation was also induced by relating the observation from the first variate with observation from the second variate via the binomial thinning operator. However, both CBINAR1M1 and FBINAR1M1 were developed under stationary distributional or time-independent covariate assumptions. Recently, Mamode Khan et al. (2016b) introduced the first constrained BINAR(1) (CBINAR1M2) Poisson time series model under non-stationary set-up. In their model, the mean and variance parameters of the bivariate innovation series were functions of time-dependent covariates, but so far, there is no full BINAR(1) process for non-stationary bivariate time series of counts such as the number of stock transactions of two companies operating in the same line of business, the number of daytime and nighttime accidents in an area, the number of day and night larceny acts and many others.

As for the inference procedures, Pedeli and Karlis (2009, 2011, 2013a) developed CMLE and method of moment approaches to estimate the mean and correlation parameters respectively in CBINAR1M1 and FBINAR1M1 models. Through simulation studies, CMLE was shown to provide estimates with lower bias than the moment based approach. Based on these results, Mamode Khan et al. (2016b) developed the CMLE under time-dependent covariates for the non-stationary CBINAR1M2 and remarked that its log-information Hessian matrix was computationally intensive. As an alternative, Mamode Khan et al. (2016b) proposed a generalized quasi-likelihood approach (GQL) that depends only on the notion of the marginal means, variances and joint covariances. This GQL equation was then solved iteratively using Newton-Raphson to obtain estimates of the mean parameters while a robust method of moment was implemented to estimate the non-stationary serial and cross correlations based on the findings of Sutradhar and Das (1999) and Jowaheer and Sutradhar (2005). Besides, through simulation studies, GQL yielded equally efficient estimates as CMLE but far more efficient estimates than GMM (Qu and Lindsay, 2003) and GEE (Liang and Zeger, 1986). Thus, in this paper, we propose to develop a GQL approach for the full non-stationary Poisson BINAR(1)

process (FBINAR1M2) as well as CMLE and GMM adaptive equations. The performance of these methodologies are tested through a simulation study.

The outline of the paper is as follows: In Section 2, we introduce the FBINAR1M2 Poisson time series model and the marginal moments and joint covariances are derived under non-stationary distributional assumptions. In Section 3, we develop CMLE, GMM adaptive and GQL estimating functions to estimate the regression, serial and cross correlation parameters under the non-stationary FBINAR1M2 model which have not yet been explored in time series literature so far. This paper also compares the performance of CMLE, GMM adaptive and GQL on their basis of their statistical and computational efficiencies through a simulation study. The paper is concluded in Section 5.

2 The Non-Stationary Poisson BINAR(1) Model

Following Mamode Khan et al. (2016b), we extend the CBINAR1M2 to account for the cross-correlation between the innovation series and the cross-correlation between the current observation at time t with the other variate previous lagged observations such that the FBINAR1M2 model is written as:

$$Y_t^{[1]} = \rho_1 * Y_{t-1}^{[1]} + \alpha_{1,t-1} * Y_{t-1}^{[2]} + d_t^{[1]} \quad (2.1)$$

$$Y_t^{[2]} = \rho_2 * Y_{t-1}^{[2]} + \alpha_{2,t-1} * Y_{t-1}^{[1]} + d_t^{[2]}, \quad (2.2)$$

where $0 < \alpha_{1,t-1}, \alpha_{2,t-1}, \rho_1, \rho_2 < 1$ and ρ_k represent the serial correlation and $\alpha_{k,t-1}$ are the cross-correlation that needs to be estimated. $Y_{t-1}^{[k]}$ and $d_t^{[k]}$ are independent. As for the thinning operation (Steutel and Van Harn, 1979), $\theta * Y_{t-1}^{[k]} = \sum_{j=1}^{Y_{t-1}^{[k]}} b_j(\theta)$ and $b_j(\theta)$ is a binary variable with $prob[b_j(\theta) = 1] = \theta$ and $prob[b_j(\theta) = 0] = 1 - \theta$ such that

$$\theta * Y_{t-1}^{[k]} | Y_{t-1}^{[k]}, \theta \sim \text{Binomial}(Y_{t-1}^{[k]}, \theta)$$

In the above model, the $Y_t^{[k]}$ are subject to a $p \times 1$ vector of covariates $x_t = [x_{t1}, x_{t2}, \dots, x_{tp}]$ and $\mu_t^{[k]} = \exp(x_t^T \beta^{[k]})$, where $\beta^{[k]}$ is the $p \times 1$ vector of regression coefficients for the k^{th} series. From Equations (2.1) and (2.2), $Y_{t-1}^{[k]} \sim \text{Poisson}(\mu_{t-1}^{[k]})$. $d_t^{[1]}$ and $d_t^{[2]}$ follow the bivariate Poisson distribution (BivPoisson) (Kocherlakota and Kocherlakota, 2001) such that

$$P(d_t^{[1]} = u, d_t^{[2]} = v) = e^{-(\lambda_t^{[1]} + \lambda_t^{[2]} - \phi_t)} \frac{(\lambda_t^{[1]} - \phi_t)^u (\lambda_t^{[2]} - \phi_t)^v}{u! v!} \times \sum_{i=0}^s \binom{u}{i} \binom{v}{i} \phi_t^i i! [(\lambda_t^{[1]} - \phi_t) (\lambda_t^{[2]} - \phi_t)]^{-i} \quad (2.3)$$

and following equation (2.3), $d_t^{[k]} \sim \text{Poisson}(\lambda_t^{[k]})$, where $\lambda_t^{[1]} = (\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]})$, $\lambda_t^{[2]} = (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]})$, $s = \min(u, v)$, $\lambda_t^{[1]}, \lambda_t^{[2]} > 0$, and $\phi_t \in [0, \min(\lambda_t^{[1]}, \lambda_t^{[2]})]$.

ϕ_t denotes the covariance between the two error terms, such that

$$Cov(d_t^{[1]}, d_{t'}^{[2]}) = \begin{cases} \phi_t & \text{if } t = t', \\ 0 & \text{if } t \neq t'. \end{cases}$$

and alternatively

$$\phi_t = \rho_{12,t} \sqrt{\lambda_t^{[1]}} \sqrt{\lambda_t^{[2]}},$$

where $\rho_{12,t}$ denotes the bivariate cross-correlation between the innovation series $d_t^{[1]}$ and $d_t^{[2]}$ and $0 \leq \rho_{12,t} \leq 1$. From the above assumption, it can be proved that

$$E(Y_t^{[k]}) = Var(Y_t^{[k]}) = \mu_t^{[k]}$$

Similarly, the lag- h correlation for the k^{th} series is

$$Corr(Y_t^{[k]}, Y_{t+h}^{[k]}) = \frac{\rho_k^h \sqrt{\mu_t^{[k]}}}{\sqrt{\mu_{t+h}^{[k]}}} \quad (2.4)$$

As for the joint covariance $\sigma_{t,12} = Cov(Y_t^{[1]}, Y_t^{[2]})$, from Equations (2.1) and (2.2)

$$\begin{aligned} E(Y_t^{[1]} Y_t^{[2]}) &= \rho_1 \rho_2 E[Y_{t-1}^{[1]} Y_{t-1}^{[2]}] + \alpha_{1,t-1} \alpha_{2,t-1} E[Y_{t-1}^{[1]} Y_{t-1}^{[2]}] + \alpha_{2,t-1} \rho_1 \mu_{t-1}^{[1]} + \alpha_{2,t-1} \rho_1 (\mu_{t-1}^{[1]})^2 \\ &\quad + \rho_1 \mu_{t-1}^{[1]} \mu_t^{[2]} - \rho_1 \rho_2 \mu_{t-1}^{[1]} \mu_{t-1}^{[2]} - \rho_1 \alpha_{2,t-1} (\mu_{t-1}^{[1]})^2 + \alpha_{1,t-1} \rho_2 \mu_{t-1}^{[2]} \\ &\quad + \alpha_{1,t-1} \rho_2 (\mu_{t-1}^{[2]})^2 + \alpha_{1,t-1} \mu_{t-1}^{[2]} \mu_t^{[2]} - \alpha_{1,t-1} \rho_2 (\mu_{t-1}^{[2]})^2 - \alpha_{1,t-1} \alpha_{2,t-1} \mu_{t-1}^{[1]} \mu_{t-1}^{[2]} \\ &\quad + \rho_2 \mu_t^{[1]} \mu_{t-1}^{[2]} - \rho_2 \rho_1 \mu_{t-1}^{[1]} \mu_{t-1}^{[2]} - \alpha_{1,t-1} \rho_2 (\mu_{t-1}^{[2]})^2 + \alpha_{2,t-1} \mu_{t-1}^{[1]} \mu_t^{[1]} \\ &\quad - \rho_1 \alpha_{2,t-1} (\mu_{t-1}^{[1]})^2 - \alpha_{1,t-1} \alpha_{2,t-1} \mu_{t-1}^{[1]} \mu_{t-1}^{[2]} + E(d_t^{[1]} d_t^{[2]}) \end{aligned}$$

and by simplifying $\sigma_{t,12}$, a first-order difference equation is obtained, where

$$\begin{aligned} \sigma_{t,12} &= (\rho_1 \rho_2 + \alpha_{1,t-1} \alpha_{2,t-1}) Cov(Y_{t-1}^{[1]}, Y_{t-1}^{[2]}) + \rho_{12,t} \sqrt{\lambda_t^{[1]}} \sqrt{\lambda_t^{[2]}} \\ &\quad + \rho_1 \alpha_{2,t-1} \mu_{t-1}^{[1]} + \alpha_{1,t-1} \rho_2 \mu_{t-1}^{[2]} \end{aligned}$$

For the other off-diagonal covariance,

$$E(Y_t^{[1]} Y_{t+1}^{[2]}) = \rho_2 Cov(Y_t^{[1]}, Y_t^{[2]}) + \alpha_{2,t} Cov(Y_t^{[1]}, Y_t^{[1]}) + \mu_t^{[1]} \mu_{t+1}^{[2]} \quad (2.5)$$

$$E(Y_t^{[1]} Y_{t+2}^{[2]}) = \rho_2 Cov(Y_t^{[1]}, Y_{t+1}^{[2]}) + \alpha_{2,t+1} Cov(Y_t^{[1]}, Y_{t+1}^{[1]}) + \mu_t^{[1]} \mu_{t+2}^{[2]} \quad (2.6)$$

(See detailed derivation for Equations (2.5) and (2.6) in Appendix A).

$$E(Y_t^{[1]} Y_{t+h}^{[2]}) = \rho_2 Cov(Y_t^{[1]}, Y_{t+h-1}^{[2]}) + \alpha_{2,t+h-1} Cov(Y_t^{[1]}, Y_{t+h-1}^{[1]}) + \mu_t^{[1]} \mu_{t+h}^{[2]} \quad (2.7)$$

Using Equations(2.5) – (2.7), we have

$$\begin{aligned} Cov(Y_t^{[1]}, Y_{t+h}^{[2]}) &= \rho_2 Cov(Y_t^{[1]}, Y_{t+h-1}^{[2]}) + \alpha_{2,t+h-1} Cov(Y_t^{[1]}, Y_{t+h-1}^{[1]}) \\ &= \rho_2^h \left[(\rho_1 \rho_2 + \alpha_{1,t-1} \alpha_{2,t-1}) Cov(Y_{t-1}^{[1]}, Y_{t-1}^{[2]}) + \rho_{12,t} \sqrt{\lambda_t^{[1]}} \sqrt{\lambda_t^{[2]}} \right. \\ &\quad \left. + \rho_1 \alpha_{2,t-1} \mu_{t-1}^{[1]} + \alpha_{1,t-1} \rho_2 \mu_{t-1}^{[2]} \right] + \mu_t^{[1]} \sum_{j=0}^{h-1} \rho_2^j \rho_1^{h-j-1} \alpha_{2,t+h-1-j} \end{aligned}$$

and similarly,

$$\begin{aligned} Cov(Y_{t+h}^{[1]}, Y_t^{[2]}) &= \rho_1^h \left[(\rho_1 \rho_2 + \alpha_{1,t-1} \alpha_{2,t-1}) Cov(Y_{t-1}^{[1]}, Y_{t-1}^{[2]}) + \rho_{12,t} \sqrt{\lambda_t^{[1]}} \sqrt{\lambda_t^{[2]}} \right. \\ &\quad \left. + \rho_1 \alpha_{2,t-1} \mu_{t-1}^{[1]} + \alpha_{1,t-1} \rho_2 \mu_{t-1}^{[2]} \right] + \mu_t^{[2]} \sum_{j=0}^{h-1} \rho_1^j \rho_2^{h-j-1} \alpha_{1,t+h-1-j}. \end{aligned}$$

Under time-independent covariates assumption, $\sigma_{t,12}$ is reduced to

$$\begin{aligned} \sigma_{12} &= \frac{\rho_1 \alpha_{2,1} \mu_1^{[1]} + \alpha_{1,1} \rho_2 \mu_1^{[2]} + \rho_{12} \sqrt{\lambda_1^{[1]}} \sqrt{\lambda_1^{[2]}}}{(1 - \rho_1 \rho_2 - \alpha_{1,1} \alpha_{2,1})} \\ Cov(Y_t^{[1]}, Y_{t+h}^{[2]}) &= \rho_2 Cov(Y_t^{[1]}, Y_{t+h-1}^{[2]}) + \alpha_{2,1} Cov(Y_t^{[1]}, Y_{t+h-1}^{[1]}) \\ &= \rho_2^h \sigma_{12} + \alpha_{2,1} \mu_1^{[1]} \sum_{j=0}^{h-1} \rho_2^j \rho_1^{h-j-1} \\ Cov(Y_{t+h}^{[1]}, Y_t^{[2]}) &= \rho_1 Cov(Y_{t+h-1}^{[1]}, Y_t^{[2]}) + \alpha_{1,1} Cov(Y_{t+h-1}^{[1]}, Y_t^{[1]}) \\ &= \rho_1^h \sigma_{12} + \alpha_{1,1} \mu_1^{[2]} \sum_{j=0}^{h-1} \rho_1^j \rho_2^{h-j-1} \end{aligned}$$

where $\alpha_{1,1}$ and $\alpha_{2,1}$ are the stationary cross-dependence terms in Equations (2.1) and (2.2). Note that these stationary cross-covariance expressions have the same formulation as Pedeli and Karlis (2013a).

3 Estimation of Parameters

This section reviews the estimation methodologies employed in estimating the regression, serial- and cross-correlation parameters. For the CBINAR1M1 and FBINAR1M1, Pedeli and Karlis (2009, 2011, 2013a) developed the CMLE approach and method of moments approach but as mentioned, these two models were designed under stationary assumptions and yet the CMLE was computationally intensive. Mamode Khan et al. (2016b) recently developed the GQL estimation technique for the non-stationary CBINAR1M2 where through simulation studies, it was noted that CMLE and GQL yielded equally efficient estimates asymptotically while far more efficient estimates than

GMM. Besides, GQL is more parsimonious than CMLE since it depends only on the notion of the marginal and joint moments. However, CMLE, GMM and GQL have not yet been developed for the FBINAR1M2 model. In the subsections that follow, these estimation approaches are derived in details and their merits and drawbacks are also highlighted.

3.1 CMLE

The conditional density for the FBINAR1M2 model is expressed as (Pedeli and Karlis, 2013a):

$$f_1(k) = \sum_{j_1=0}^k \binom{y_{t-1}^{[1]}}{j_1} \binom{y_{t-1}^{[2]}}{k-j_1} \rho_1^{j_1} (1-\rho_1)^{y_{t-1}^{[1]}-j_1} \alpha_{1,t-1}^{k-j_1} (1-\alpha_{1,t-1})^{y_{t-1}^{[2]}-k+j_1}$$

$$f_2(s) = \sum_{j_2=0}^s \binom{y_{t-1}^{[2]}}{j_2} \binom{y_{t-1}^{[1]}}{s-j_2} \rho_2^{j_2} (1-\rho_2)^{y_{t-1}^{[2]}-j_2} \alpha_{2,t-1}^{s-j_2} (1-\alpha_{2,t-1})^{y_{t-1}^{[1]}-s+j_2}$$

and a bivariate distribution of the innovation terms of the form $f_3(u, v) = P(d_t^{[1]} = u, d_t^{[2]} = v)$. Then, the conditional density is written as (Pedeli and Karlis, 2013a)

$$\begin{aligned} & f(y_t^{[1]}, y_t^{[2]} | y_{t-1}^{[1]}, y_{t-1}^{[2]}, \theta) \\ &= \sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f_1(k) f_2(s) f_3(y_t^{[1]} - k, y_t^{[2]} - s) \\ &= \sum_{k=0}^{g_1} \sum_{s=0}^{g_2} \left[\sum_{j_1=0}^k \binom{y_{t-1}^{[1]}}{j_1} \binom{y_{t-1}^{[2]}}{k-j_1} \rho_1^{j_1} (1-\rho_1)^{y_{t-1}^{[1]}-j_1} \alpha_{1,t-1}^{k-j_1} (1-\alpha_{1,t-1})^{y_{t-1}^{[2]}-k+j_1} \right] \\ & \times \left[\sum_{j_2=0}^s \binom{y_{t-1}^{[2]}}{j_2} \binom{y_{t-1}^{[1]}}{s-j_2} \rho_2^{j_2} (1-\rho_2)^{y_{t-1}^{[2]}-j_2} \alpha_{2,t-1}^{s-j_2} (1-\alpha_{2,t-1})^{y_{t-1}^{[1]}-s+j_2} \right] \\ & \times e^{-(\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} + \mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t)} \\ & \times \sum_{m=0}^{\min(k,s)} \left[(\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} - \phi_t)^{y_t^{[1]}-k-m} \right. \\ & \times (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t)^{y_t^{[2]}-s-m} \phi_t^m \left. \right] \\ & \quad / [(y_t^{[1]} - k - m)! (y_t^{[2]} - s - m)! m!] \end{aligned}$$

where $\theta = (\rho_k, \alpha_{k,t-1}, \phi_t, \beta^{[k]})'$ is the vector of unknown parameters, $g_1 = \min(y_t^{[1]}, y_{t-1}^{[1]})$ and $g_2 = \min(y_t^{[2]}, y_{t-1}^{[2]})$.

The conditional likelihood function is then given by

$$L(\theta | y) = \prod_{t=1}^T f(y_t^{[k]} | y_{t-1}^{[k]}, \theta)$$

for some initial values of θ and by maximizing the conditional likelihood function

$$\log L(\theta | y) = \log \sum_{t=1}^T f(y_t^{[k]} | y_{t-1}^{[k]}, \theta)$$

the following expressions are obtained:

Differentiating with respect to ρ_1 , we have

$$\begin{aligned} \frac{\partial \log L(\theta | y)}{\partial \rho_1} = & \sum_{t=1}^T \left[\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} \left[f_2(s) f_3(y_t^{[1]} - k, y_t^{[2]} - s) \frac{\partial f_1(k)}{\partial \rho_1} \right. \right. \\ & \left. \left. + f_1(k) f_2(s) \frac{\partial}{\partial \rho_1} f_3(y_t^{[1]} - k, y_t^{[2]} - s) \right] \left[\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f(y_t^{[k]} | y_{t-1}^{[k]}, \theta) \right]^{-1} \right], \quad (3.1) \end{aligned}$$

where

$$\begin{aligned} \frac{\partial f_1(k)}{\partial \rho_1} = & \sum_{j_1=0}^k \binom{y_{t-1}^{[1]}}{j_1} \binom{y_{t-1}^{[2]}}{k-j_1} \alpha_{1,t-1}^{k-j_1} (1 - \alpha_{1,t-1})^{y_{t-1}^{[2]} - k + j_1} \\ & \times \left[j_1 \rho_1^{j_1-1} (1 - \rho_1)^{y_{t-1}^{[1]} - j_1} - \rho_1^{j_1} (y_{t-1}^{[1]} - j_1) (1 - \rho_1)^{y_{t-1}^{[1]} - j_1 - 1} \right] \end{aligned}$$

and

$$\begin{aligned} & \frac{\partial}{\partial \rho_1} f_3(y_t^{[1]} - k, y_t^{[2]} - s) \\ = & \sum_{m=0}^{\min(k,s)} \frac{e^{-(\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} + \mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t)}}{(y_t^{[1]} - k - m)! (y_t^{[2]} - s - m)! m!} \\ & \times \left[\mu_{t-1}^{[1]} (\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} - \phi_t) y_t^{[1] - k - m} \right. \\ & \times (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t) y_t^{[2] - s - m} \phi_t^m \\ & - \mu_{t-1}^{[1]} (y_t^{[1]} - k - m) (\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} - \phi_t) y_t^{[1] - k - m - 1} \\ & \left. \times (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t) y_t^{[2] - s - m} \phi_t^m \right] \end{aligned}$$

and the double derivative with respect to ρ_1 is

$$\begin{aligned} & \sum_{t=1}^T \frac{\partial^2 \log L(\theta | y)}{\partial \rho_1^2} \\ = & \sum_{t=1}^T \left[\frac{\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f_2 f_3 \frac{\partial^2 f_1}{\partial \rho_1^2} + f_2 \frac{\partial f_1}{\partial \rho_1} \frac{\partial f_3}{\partial \rho_1} + f_1 f_2 \frac{\partial^2 f_3}{\partial \rho_1^2} + f_2 \frac{\partial f_1}{\partial \rho_1} \frac{\partial f_3}{\partial \rho_1}}{\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f(y_t^{[k]} | y_{t-1}^{[k]}, \theta)} \right. \\ & \left. - \frac{\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} (f_2 f_3 \frac{\partial f_1}{\partial \rho_1} + f_1 f_2 \frac{\partial f_3}{\partial \rho_1})^2}{\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} \{f(y_t^{[k]} | y_{t-1}^{[k]}, \theta)\}^2} \right], \quad (3.2) \end{aligned}$$

where

$$\begin{aligned} \frac{\partial^2 f_1(k)}{\partial \rho_1^2} &= \sum_{j_1=0}^k \binom{y_{t-1}^{[1]}}{j_1} \binom{y_{t-1}^{[2]}}{k-j_1} \alpha_{1,t-1}^{k-j_1} (1 - \alpha_{1,t-1})^{y_{t-1}^{[2]} - k + j_1} \\ &\quad \times \left[j_1(j_1 - 1) \rho_1^{j_1-2} (1 - \rho_1)^{y_{t-1}^{[1]} - j_1} - j_1 \rho_1^{j_1-1} (y_{t-1}^{[1]} - j_1) (1 - \rho_1)^{y_{t-1}^{[1]} - j_1 - 1} \right. \\ &\quad \left. - j_1 \rho_1^{j_1-1} (y_{t-1}^{[1]} - j_1) (1 - \rho_1)^{y_{t-1}^{[1]} - j_1 - 1} \right. \\ &\quad \left. + \rho_1^{j_1} (y_{t-1}^{[1]} - j_1) (y_{t-1}^{[1]} - j_1 - 1) (1 - \rho_1)^{y_{t-1}^{[1]} - j_1 - 2} \right] \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2}{\partial \rho_1^2} f_3(y_t^{[1]} - k, y_t^{[2]} - s) &= \sum_{m=0}^{\min(k,s)} \frac{e^{-(\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} + \mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t)}}{(y_t^{[1]} - k - m)! (y_t^{[2]} - s - m)! m!} \\ &\quad \times [-\mu_{t-1}^2 \binom{[1]}{[1]} (\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} - \phi_t)^{y_t^{[1]} - k - m} \\ &\quad \times (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t)^{y_t^{[2]} - s - m} \phi_t^m \\ &\quad + \mu_{t-1}^2 \binom{[1]}{[1]} (y_t^{[1]} - k - m) (\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} - \phi_t)^{y_t^{[1]} - k - m - 1} \\ &\quad \times (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t)^{y_t^{[2]} - s - m} \phi_t^m \\ &\quad - \mu_{t-1}^2 \binom{[1]}{[1]} (y_t^{[1]} - k - m) (\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} - \phi_t)^{y_t^{[1]} - k - m - 1} \\ &\quad \times (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t)^{y_t^{[2]} - s - m} \phi_t^m \\ &\quad + \mu_{t-1}^2 \binom{[1]}{[1]} (y_t^{[1]} - k - m) (y_t^{[1]} - k - m - 1) \\ &\quad \times (\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} - \phi_t)^{y_t^{[1]} - k - m - 2} \\ &\quad \times (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t)^{y_t^{[2]} - s - m} \phi_t^m]. \end{aligned}$$

The above derivation confirms that the double derivative for CMLE is computationally intensive under non-stationary assumptions. Differentiating with respect to ρ_2 , we have

$$\begin{aligned} \sum_{t=1}^T \frac{\partial \log L(\theta | y)}{\partial \rho_2} &= \sum_{t=1}^T \left[\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f(y_t^{[k]} | y_{t-1}^{[k]}, \theta) \right]^{-1} \left[\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f_1(k) f_3(y_t^{[1]} - k, y_t^{[2]} - s) \frac{\partial f_2(s)}{\partial \rho_2} \right. \\ &\quad \left. + f_1(k) f_2(s) \frac{\partial}{\partial \rho_2} f_3(y_t^{[1]} - k, y_t^{[2]} - s) \right], \end{aligned}$$

where

$$\begin{aligned} \frac{\partial f_2(s)}{\partial \rho_2} &= \sum_{j_2=0}^s \binom{y_{t-1}^{[2]}}{j_2} \binom{y_{t-1}^{[1]}}{s-j_2} \alpha_{2,t-1}^{s-j_2} (1 - \alpha_{2,t-1})^{y_{t-1}^{[1]} - s + j_2} \\ &\quad \left[j_2 \rho_2^{j_2-1} (1 - \rho_2)^{y_{t-1}^{[2]} - j_2} - \rho_2^{j_2} (y_{t-1}^{[2]} - j_2) (1 - \rho_2)^{y_{t-1}^{[2]} - j_2 - 1} \right] \end{aligned}$$

and

$$\begin{aligned} \frac{\partial}{\partial \rho_2} f_3(y_t^{[1]} - k, y_t^{[2]} - s) &= \sum_{m=0}^{\min(k,s)} \frac{e^{-(\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} + \mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t)}}{(y_t^{[1]} - k - m)!(y_t^{[2]} - s - m)!m!} \\ &\times [\mu_{t-1}^{[2]} (\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} - \phi_t) y_t^{[1] - k - m} \\ &\times (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t) y_t^{[2] - s - m} \phi_t^m \\ &- \mu_{t-1}^{[2]} (y_t^{[2]} - s - m) (\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} - \phi_t) y_t^{[1] - k - m} \\ &\times (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t) y_t^{[2] - s - m - 1} \phi_t^m]. \end{aligned}$$

Differentiating with respect to $\alpha_{1,t-1}$, we have

$$\begin{aligned} \sum_{t=1}^T \frac{\partial \log L(\theta | y)}{\partial \alpha_{1,t-1}} &= \sum_{t=1}^T \left[\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f(y_t^{[k]} | y_{t-1}^{[k]}, \theta) \right]^{-1} \left[\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f_2(s) f_3(y_t^{[1]} - k, y_t^{[2]} - s) \frac{\partial f_1(k)}{\partial \alpha_{1,t-1}} \right. \\ &\quad \left. + f_1(k) f_2(s) \frac{\partial}{\partial \alpha_{1,t-1}} f_3(y_t^{[1]} - k, y_t^{[2]} - s) \right], \end{aligned}$$

where

$$\begin{aligned} \frac{\partial f_1(k)}{\partial \alpha_{1,t-1}} &= \sum_{j_1=0}^k \binom{y_{t-1}^{[1]}}{j_1} \binom{y_{t-1}^{[2]}}{k - j_1} \rho_1^{j_1} (1 - \rho_1)^{y_{t-1}^{[1]} - j_1} \left[(k - j_1) \alpha_{1,t-1}^{k - j_1 - 1} (1 - \alpha_{1,t-1})^{y_{t-1}^{[2]} - k + j_1} \right. \\ &\quad \left. - \alpha_{1,t-1}^{k - j_1} (y_{t-1}^{[2]} - k + j_1) (1 - \alpha_{1,t-1})^{y_{t-1}^{[2]} - k + j_1 - 1} \right] \end{aligned}$$

and

$$\begin{aligned} \frac{\partial}{\partial \alpha_{1,t-1}} f_3(y_t^{[1]} - k, y_t^{[2]} - s) &= \sum_{m=0}^{\min(k,s)} \frac{e^{-(\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} + \mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t)}}{(y_t^{[1]} - k - m)!(y_t^{[2]} - s - m)!m!} \\ &\times \left[\mu_{t-1}^{[2]} (\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} - \phi_t) y_t^{[1] - k - m} \right. \\ &\times (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t) y_t^{[2] - s - m} \phi_t^m \\ &- \mu_{t-1}^{[2]} (y_t^{[1]} - k - m) (\mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} - \phi_t) y_t^{[1] - k - m - 1} \\ &\left. \times (\mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]} - \phi_t) y_t^{[2] - s - m} \phi_t^m \right]. \end{aligned}$$

Differentiating with respect to $\alpha_{2,t-1}$, we have

$$\begin{aligned} \sum_{t=1}^T \frac{\partial \log L(\theta | y)}{\partial \alpha_{2,t-1}} &= \sum_{t=1}^T \left[\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f(y_t^{[k]} | y_{t-1}^{[k]}, \theta) \right]^{-1} \left[\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f_1(k) f_3(y_t^{[1]} - k, y_t^{[2]} - s) \frac{\partial f_2(s)}{\partial \alpha_{2,t-1}} \right. \\ &\quad \left. + f_1(k) f_2(s) \frac{\partial}{\partial \alpha_{2,t-1}} f_3(y_t^{[1]} - k, y_t^{[2]} - s) \right], \end{aligned}$$

where

$$\begin{aligned} \frac{\partial f_2(s)}{\partial \alpha_{2,t-1}} &= \sum_{j_2=0}^k \binom{y_{t-1}^{[2]}}{j_2} \binom{y_{t-1}^{[1]}}{s-j_2} \rho_2^{j_2} (1-\rho_2)^{y_{t-1}^{[2]}-j_2} \left[(s-j_2) \alpha_{2,t-1}^{s-j_2-1} (1-\alpha_{2,t-1})^{y_{t-1}^{[1]}-s+j_2} \right. \\ &\quad \left. - \alpha_{2,t-1}^{s-j_2} (y_{t-1}^{[1]}-s+j_2) (1-\alpha_{2,t-1})^{y_{t-1}^{[1]}-s+j_2-1} \right] \end{aligned}$$

and

$$\begin{aligned} &\frac{\partial}{\partial \alpha_{2,t-1}} f_3(y_t^{[1]}-k, y_t^{[2]}-s) \\ &= \sum_{m=0}^{\min(k,s)} \frac{e^{-(\mu_t^{[1]}-\rho_1\mu_{t-1}^{[1]}-\alpha_{1,t-1}\mu_{t-1}^{[2]}+\mu_t^{[2]}-\rho_2\mu_{t-1}^{[2]}-\alpha_{2,t-1}\mu_{t-1}^{[1]}-\phi_t)}}{(y_t^{[1]}-k-m)!(y_t^{[2]}-s-m)!m!} \\ &\quad \times \left[\mu_{t-1}^{[1]}(\mu_t^{[1]}-\rho_1\mu_{t-1}^{[1]}-\alpha_{1,t-1}\mu_{t-1}^{[2]}-\phi_t) y_t^{[1]-k-m} \right. \\ &\quad \times (\mu_t^{[2]}-\rho_2\mu_{t-1}^{[2]}-\alpha_{2,t-1}\mu_{t-1}^{[1]}-\phi_t) y_t^{[2]-s-m} \phi_t^m \\ &\quad - \mu_{t-1}^{[1]}(y_t^{[2]}-s-m)(\mu_t^{[1]}-\rho_1\mu_{t-1}^{[1]}-\alpha_{1,t-1}\mu_{t-1}^{[2]}-\phi_t) y_t^{[1]-k-m} \\ &\quad \left. \times (\mu_t^{[2]}-\rho_2\mu_{t-1}^{[2]}-\alpha_{2,t-1}\mu_{t-1}^{[1]}-\phi_t) y_t^{[2]-s-m-1} \phi_t^m \right]. \end{aligned}$$

Differentiating with respect to ϕ_t , we have

$$\sum_{t=1}^T \frac{\partial \log L(\theta|y)}{\partial \phi_t} = \sum_{t=1}^T \frac{\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f_1(k) f_2(s) \frac{\partial}{\partial \phi_t} f_3(y_t^{[1]}-k, y_t^{[2]}-s)}{\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f(y_t^{[k]} | y_{t-1}^{[k]}, \theta)},$$

where

$$\begin{aligned} \frac{\partial}{\partial \phi_t} f_3(y_t^{[1]}-k, y_t^{[2]}-s) &= \sum_{m=0}^{\min(k,s)} \frac{e^{-(\mu_t^{[1]}-\rho_1\mu_{t-1}^{[1]}-\alpha_{1,t-1}\mu_{t-1}^{[2]}+\mu_t^{[2]}-\rho_2\mu_{t-1}^{[2]}-\alpha_{2,t-1}\mu_{t-1}^{[1]}-\phi_t)}}{(y_t^{[1]}-k-m)!(y_t^{[2]}-s-m)!m!} \\ &\quad \times \left[(\mu_t^{[1]}-\rho_1\mu_{t-1}^{[1]}-\alpha_{1,t-1}\mu_{t-1}^{[2]}-\phi_t) y_t^{[1]-k-m} \right. \\ &\quad \times (\mu_t^{[2]}-\rho_2\mu_{t-1}^{[2]}-\alpha_{2,t-1}\mu_{t-1}^{[1]}-\phi_t) y_t^{[2]-s-m} \phi_t^m \\ &\quad + m(\mu_t^{[1]}-\rho_1\mu_{t-1}^{[1]}-\alpha_{1,t-1}\mu_{t-1}^{[2]}-\phi_t) y_t^{[1]-k-m} \\ &\quad \times (\mu_t^{[2]}-\rho_2\mu_{t-1}^{[2]}-\alpha_{2,t-1}\mu_{t-1}^{[1]}-\phi_t) y_t^{[2]-s-m} \phi_t^{m-1} \\ &\quad - (y_t^{[1]}-k-m)(\mu_t^{[1]}-\rho_1\mu_{t-1}^{[1]}-\alpha_{1,t-1}\mu_{t-1}^{[2]}-\phi_t) y_t^{[1]-k-m-1} \\ &\quad \times (\mu_t^{[2]}-\rho_2\mu_{t-1}^{[2]}-\alpha_{2,t-1}\mu_{t-1}^{[1]}-\phi_t) y_t^{[2]-s-m} \phi_t^m \\ &\quad - (y_t^{[2]}-s-m)(\mu_t^{[1]}-\rho_1\mu_{t-1}^{[1]}-\alpha_{1,t-1}\mu_{t-1}^{[2]}-\phi_t) y_t^{[1]-k-m} \\ &\quad \left. \times (\mu_t^{[2]}-\rho_2\mu_{t-1}^{[2]}-\alpha_{2,t-1}\mu_{t-1}^{[1]}-\phi_t) y_t^{[2]-s-m-1} \phi_t^m \right]. \end{aligned}$$

Differentiating with respect to $\beta^{[k]}$, we have

$$\sum_{t=1}^T \frac{\partial \log L(\theta | y)}{\partial \beta^{[k]}} = \sum_{t=1}^T \frac{\partial \log L(\theta | y)}{\partial \lambda_t^{[k]}} \times \frac{\partial \lambda_t^{[k]}}{\partial \beta^{[k]}},$$

where

$$\begin{aligned} \lambda_t^{[1]} &= \mu_t^{[1]} - \rho_1 \mu_{t-1}^{[1]} - \alpha_{1,t-1} \mu_{t-1}^{[2]} \\ \lambda_t^{[2]} &= \mu_t^{[2]} - \rho_2 \mu_{t-1}^{[2]} - \alpha_{2,t-1} \mu_{t-1}^{[1]}, \quad \frac{\partial \lambda_t^{[k]}}{\partial \beta^{[k]}} = \mu_t^{[k]} x_{t,j}^T - \rho_k \mu_{t-1}^{[k]} x_{t-1,j}^T \\ \sum_{t=1}^T \frac{\partial \log L(\theta | y)}{\partial \lambda_t^{[k]}} &= \sum_{t=1}^T \frac{\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f_1(k) f_2(s) \frac{\partial}{\partial \lambda_t^{[k]}} f_3(y_t^{[1]} - k, y_t^{[2]} - s)}{\sum_{k=0}^{g_1} \sum_{s=0}^{g_2} f(y_t^{[k]} | y_{t-1}^{[k]}, \theta)} \end{aligned}$$

with

$$\begin{aligned} \frac{\partial}{\partial \lambda_t^{[1]}} f_3(y_t^{[1]} - k, y_t^{[2]} - s) &= \sum_{m=0}^{\min(k,s)} \frac{e^{-(\lambda_t^{[1]} + \lambda_t^{[2]} - \phi_t)}}{(y_t^{[1]} - k - m)! (y_t^{[2]} - s - m)! m!} \\ &\quad \left[(y_t^{[1]} - k - m) (\lambda_t^{[1]} - \phi_t)^{y_t^{[1]} - k - m - 1} (\lambda_t^{[2]} - \phi_t)^{y_t^{[2]} - s - m} \phi_t^m \right. \\ &\quad \left. - (\lambda_t^{[1]} - \phi_t)^{y_t^{[1]} - k - m} (\lambda_t^{[2]} - \phi_t)^{y_t^{[2]} - s - m} \phi_t^m \right] \\ \frac{\partial}{\partial \lambda_t^{[2]}} f_3(y_t^{[1]} - k, y_t^{[2]} - s) &= \sum_{m=0}^{\min(u,v)} \frac{e^{-(\lambda_t^{[1]} + \lambda_t^{[2]} - \phi_t)}}{(y_t^{[1]} - k - m)! (y_t^{[2]} - s - m)! m!} \\ &\quad \left[(y_t^{[2]} - s - m) (\lambda_t^{[1]} - \phi_t)^{y_t^{[1]} - k - m} (\lambda_t^{[2]} - \phi_t)^{y_t^{[2]} - s - m - 1} \phi_t^m \right. \\ &\quad \left. - (\lambda_t^{[1]} - \phi_t)^{y_t^{[1]} - k - m} (\lambda_t^{[2]} - \phi_t)^{y_t^{[2]} - s - m} \phi_t^m \right]. \end{aligned}$$

From the above, it is clear that implementing CMLE under non-stationary assumptions for a full BINAR(1) is computationally intensive, in particular the double derivative Hessian.

3.2 GQL/GMM Adaptive

Further to Mamode Khan et al. (2016b), the GQL equation is given as follows:

$$D^T \Sigma^{-1} (\tilde{y} - \tilde{\mu}) = 0,$$

where

$$\begin{aligned} \tilde{y} &= \begin{pmatrix} y^{[1]} \\ y^{[2]} \end{pmatrix}, y^{[1]} = [y_1^{[1]}, y_2^{[1]}, \dots, y_T^{[1]}]^T, y^{[2]} = [y_1^{[2]}, y_2^{[2]}, \dots, y_T^{[2]}]^T \\ \tilde{\mu} &= \begin{pmatrix} \mu^{[1]} \\ \mu^{[2]} \end{pmatrix}, \mu^{[1]} = [\mu_1^{[1]}, \mu_2^{[1]}, \dots, \mu_T^{[1]}]^T, \mu^{[2]} = [\mu_1^{[2]}, \mu_2^{[2]}, \dots, \mu_T^{[2]}]^T. \end{aligned}$$

The derivative matrix is a block diagonal $2T \times 2p$ matrix with $[D^{[1]}, D^{[2]}]$ on the diagonals where

$$D^{[k]} = \begin{pmatrix} \frac{\partial \mu_1^{[k]}}{\partial \beta_1^{[k]}} & \frac{\partial \mu_1^k}{\partial \beta_2^{[k]}} & \cdots & \frac{\partial \mu_1^{[k]}}{\partial \beta_p^{[k]}} \\ \frac{\partial \mu_2^{[k]}}{\partial \beta_1^{[k]}} & \frac{\partial \mu_2^k}{\partial \beta_2^{[k]}} & \cdots & \frac{\partial \mu_2^{[k]}}{\partial \beta_p^{[k]}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mu_T^{[k]}}{\partial \beta_1^{[k]}} & \frac{\partial \mu_T^k}{\partial \beta_2^{[k]}} & \cdots & \frac{\partial \mu_T^{[k]}}{\partial \beta_p^{[k]}} \end{pmatrix}_{(T \times p)} \quad (3.3)$$

and $\frac{\partial \mu_t^{[k]}}{\partial \beta^{[k]}} = \mu_t^{[k]} x_t^T$. The covariance structure $\hat{\Sigma}$ is

$$\begin{pmatrix} V(Y_1^{[1]}) & \cdots & Cov(Y_1^{[1]}, Y_T^{[1]}) & \vdots & Cov(Y_1^{[1]}, Y_1^{[2]}) & \cdots & Cov(Y_1^{[1]}, Y_T^{[2]}) \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ Cov(Y_T^{[1]}, Y_1^{[1]}) & \cdots & V(Y_T^{[1]}) & \vdots & Cov(Y_T^{[1]}, Y_1^{[2]}) & \cdots & Cov(Y_T^{[1]}, Y_T^{[2]}) \\ \cdots & \cdots & \vdots & \cdots & \cdots & \cdots & \cdots \\ Cov(Y_1^{[2]}, Y_1^{[1]}) & \cdots & Cov(Y_1^{[2]}, Y_T^{[1]}) & \vdots & V(Y_1^{[2]}) & \cdots & Cov(Y_1^{[2]}, Y_T^{[2]}) \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ Cov(Y_T^{[2]}, Y_1^{[1]}) & \cdots & Cov(Y_T^{[2]}, Y_T^{[1]}) & \vdots & Cov(Y_T^{[2]}, Y_1^{[2]}) & \cdots & V(Y_T^{[2]}) \end{pmatrix}$$

The Newton-Raphson iterative technique is used to solve the GQL in Equation (3.3) and this yields

$$\begin{pmatrix} \hat{\beta}_{r+1}^{[1]} \\ \hat{\beta}_{r+1}^{[2]} \end{pmatrix} = \begin{pmatrix} \hat{\beta}_r^{[1]} \\ \hat{\beta}_r^{[2]} \end{pmatrix} + [D^T \Sigma^{-1} D]_r^{-1} [D^T \Sigma^{-1} (\tilde{y} - \tilde{\mu})]_r,$$

where $\hat{\beta}_r^{[k]}$ is the estimate at the r^{th} iteration and $[\cdot]_r$ is the value of the expression at the r^{th} iteration.

The moment equations to estimate the serial correlations in Equation (2.4) following Jowaheer and Sutradhar (2005) are given by

$$\frac{\sum_{t=1}^{T-1} \tilde{y}_t^{[k]} \tilde{y}_{(t+1)}^{[k]} / (T-1)}{\sum_{t=1}^T (\tilde{y}_t^{[k]})^2 / T} = \hat{\rho}_k \sum_{t=1}^{T-1} \left(\sqrt{\mu_t^{[k]} / \mu_{t+1}^{[k]}} \right) / (T-1)$$

which yields

$$\hat{\rho}_k = \frac{T \sum_{t=1}^{T-1} \tilde{y}_t^{[k]} \tilde{y}_{(t+1)}^{[k]}}{\sum_{t=1}^T (\tilde{y}_t^{[k]})^2 \sum_{t=1}^{T-1} \sqrt{\mu_t^{[k]} / \mu_{t+1}^{[k]}}}, \quad (3.4)$$

where $\tilde{y}_t^{[k]} = (y_t^{[k]} - \mu_t^{[k]}) / \sqrt{\mu_t^{[k]}}$. Under the assumption of time-independent covariates with $\mu_t^{[k]} = \mu_1^{[k]}$,

$$\hat{\rho}_{k,h} = \frac{\sum_{t=1}^{T-1} \tilde{y}_t^{[k]} \tilde{y}_{(t+h)}^{[k]} / (T-1)}{\sum_{t=1}^T (\tilde{y}_t^{[k]})^2 / T},$$

where $\hat{\rho}_{k,h}$ is the estimated serial correlation for the k^{th} series at lag h . The time dependent cross correlation parameter and correlation parameter between the error terms are obtained as follows:

$$\begin{aligned}\hat{\alpha}_{1,t} &= \frac{\tilde{\sigma}_{t+1,t} - \hat{\rho}_1 \tilde{\sigma}_{tt}}{\hat{\mu}_t^{[2]}}, & \hat{\alpha}_{2,t} &= \frac{\tilde{\sigma}_{t,t+1} - \hat{\rho}_2 \tilde{\sigma}_{tt}}{\hat{\mu}_t^{[1]}} \\ \hat{\rho}_{12,t} &= \left[\tilde{\sigma}_{tt} - \hat{\rho}_1 \hat{\rho}_2 \tilde{\sigma}_{t-1,t-1} - \frac{\tilde{\sigma}_{t-1,t-1}(\tilde{\sigma}_{t+1,t} - \hat{\rho}_1 \tilde{\sigma}_{tt})(\tilde{\sigma}_{t,t+1} - \hat{\rho}_2 \tilde{\sigma}_{tt})}{\hat{\mu}_{t-1}^{[1]} \hat{\mu}_{t-1}^{[2]}} \right. \\ &\quad \left. - \hat{\rho}_1(\tilde{\sigma}_{t,t+1} - \hat{\rho}_2 \tilde{\sigma}_{tt}) - \hat{\rho}_2(\tilde{\sigma}_{t+1,t} - \hat{\rho}_1 \tilde{\sigma}_{tt}) \right] \\ &\quad \times \left[\sqrt{\hat{\mu}_t^{[1]} - \hat{\rho}_1 \hat{\mu}_{t-1}^{[1]} - \tilde{\sigma}_{t+1,t} + \hat{\rho}_1 \tilde{\sigma}_{tt}} \sqrt{\hat{\mu}_t^{[2]} - \hat{\rho}_2 \hat{\mu}_{t-1}^{[2]} - \tilde{\sigma}_{t,t+1} + \hat{\rho}_2 \tilde{\sigma}_{tt}} \right]^{-1},\end{aligned}$$

where $\hat{\mu}_t^{[k]} = \exp(x_t^T \hat{\beta}^{[k]})$ and

$$\begin{aligned}\tilde{\sigma}_{t-j,t-j} &= Cov(y_{t-j}^{[1]}, y_{t-j}^{[2]}) = \sum_{t=j+1}^T (y_{t-j}^{[1]} - \hat{\mu}_{t-j}^{[1]})(y_{t-j}^{[2]} - \hat{\mu}_{t-j}^{[2]}) / (T - j) \\ \tilde{\sigma}_{t+1,t} &= Cov(y_{t+1}^{[1]}, y_t^{[2]}) = \sum_{t=1}^{T-1} (y_{t+1}^{[1]} - \hat{\mu}_{t+1}^{[1]})(y_t^{[2]} - \hat{\mu}_t^{[2]}) / (T - 1) \\ \tilde{\sigma}_{t,t+1} &= Cov(y_t^{[1]}, y_{t+1}^{[2]}) = \sum_{t=1}^T (y_t^{[1]} - \hat{\mu}_t^{[1]})(y_{t+1}^{[2]} - \hat{\mu}_{t+1}^{[2]}) / (T - 1).\end{aligned}$$

Under the assumption of time-independent covariates, the above correlation parameters reduce to Equations (3.5) and (3.6). Hence,

$$\hat{\alpha}_{1,1} = \frac{\tilde{\sigma}_{t+1,t} - \hat{\rho}_1 \tilde{\sigma}_{tt}}{\hat{\mu}_1^{[2]}}, \quad \hat{\alpha}_{2,1} = \frac{\tilde{\sigma}_{t,t+1} - \hat{\rho}_2 \tilde{\sigma}_{tt}}{\hat{\mu}_1^{[1]}}, \quad (3.5)$$

where $\hat{\alpha}_{k,1}$ is the cross-correlation between the time series and $\hat{\rho}_{12,t} = \hat{\rho}_{12}$ is given by

$$\begin{aligned}\hat{\rho}_{12} &= \left[\tilde{\sigma}_{tt} - \hat{\rho}_1 \hat{\rho}_2 \tilde{\sigma}_{tt} - \frac{\tilde{\sigma}_{tt}(\tilde{\sigma}_{t+1,t} - \hat{\rho}_1 \tilde{\sigma}_{tt})(\tilde{\sigma}_{t,t+1} - \hat{\rho}_2 \tilde{\sigma}_{tt})}{\hat{\mu}_1^{[1]} \hat{\mu}_1^{[2]}} \right. \\ &\quad \left. - \hat{\rho}_1(\tilde{\sigma}_{t,t+1} - \hat{\rho}_2 \tilde{\sigma}_{tt}) - \hat{\rho}_2(\tilde{\sigma}_{t+1,t} - \hat{\rho}_1 \tilde{\sigma}_{tt}) \right] \\ &\quad \times \left[\sqrt{\hat{\mu}_1^{[1]} - \tilde{\sigma}_{t+1,t} + \hat{\rho}_1 \tilde{\sigma}_{tt} - \hat{\rho}_1 \hat{\mu}_1^{[1]}} \sqrt{\hat{\mu}_1^{[2]} - \tilde{\sigma}_{t,t+1} + \hat{\rho}_2 \tilde{\sigma}_{tt} - \hat{\rho}_2 \hat{\mu}_1^{[2]}} \right]^{-1}. \quad (3.6)\end{aligned}$$

For an initial value of $\hat{\beta}^{[k]}$, $\hat{\alpha}_{1,t}$, $\hat{\alpha}_{2,t}$, $\hat{\rho}_1$, $\hat{\rho}_2$ and $\hat{\rho}_{12,t}$ are obtained using Equations (3.4) to (3.6). These estimates are then substituted into the covariance matrix and the iterative Equation (3.4) is solved to obtain an updated estimate of the regression parameters and these are, in turn, recycled into the correlation estimating function. These iterations continue until the difference between the updated and previous regression estimates is less than 10^{-5} . The standard errors of the regression estimates are calculated using the Hessian $[D^T \Sigma^{-1} D]^{-1}$. Under mild regularity conditions, the

regression estimates are consistent and $((\hat{\beta}^{[1]}, \hat{\beta}^{[2]}) - (\beta^{[1]}, \beta^{[2]}))^T$ follows the asymptotic normal distribution with mean 0 and covariance matrix

$$[D^T \Sigma^{-1} D]^{-1} [D^T \Sigma^{-1} (\tilde{y} - \tilde{\mu})(\tilde{y} - \tilde{\mu})^T \Sigma^{-1} D] [D^T \Sigma^{-1} D]^{-1}$$

as demonstrated by Sutradhar et al. (2014).

The parameters under the GMM adaptive approach are estimated as

$$g_I = \begin{bmatrix} D^T (\tilde{y} - \tilde{\mu}) \\ \hat{\alpha}^T D^T V (\tilde{y} - \tilde{\mu}) \end{bmatrix},$$

where $V = (\tilde{y} - \tilde{\mu})(\tilde{y} - \tilde{\mu})^T / (2T)$ and $\hat{\alpha}$ is the orthogonal vector. The GMM objective function is introduced because the dimension of the score vector is larger than the number of regression parameters to estimate, which is given by

$$Q_I(\beta) = g_I^T C_I^{-1} g_I \quad (3.7)$$

where C_I is the sample variance of g_I :

$$\begin{pmatrix} D^T V D & (D^T V^2 D) \hat{\alpha} \\ \hat{\alpha}^T (D^T V^2 D) & \hat{\alpha}^T (D^T V^3 D) \hat{\alpha} \end{pmatrix}, \quad (3.8)$$

where $\hat{\alpha}$ is the product of the inverse of the upper triangular matrix, G , obtained from the Cholesky decomposition of the matrix $(D^T V - L D^T) V (D^T V - L D^T)$, where $L = (D^T V^2 D) (D^T V D)^{-1}$ and the eigenvector corresponding to the largest eigenvalue of the matrix $(G^T)^{-1} \tilde{D} \tilde{D}^T G^{-1}$ where $\tilde{D} = D^T V D - L D^T D$. More detail is given by Qu and Lindsay (2003) and Mamode khan and Jowaheer (2013). Using the Newton-Raphson algorithm,

$$\begin{pmatrix} \hat{\beta}_{r+1}^{[1]} \\ \hat{\beta}_{r+1}^{[2]} \end{pmatrix} = \begin{pmatrix} \hat{\beta}_r^{[1]} \\ \hat{\beta}_r^{[2]} \end{pmatrix} - [\ddot{Q}(\hat{\beta}_r)]^{-1} \dot{Q}(\hat{\beta}_r)$$

where asymptotically, $\dot{Q}_I(\beta) = 2\dot{g}_I^T C_I^{-1} g_I$ and the hessian matrix $\ddot{Q}_I(\beta) = 2\dot{g}_I^T C_I^{-1} \dot{g}_I$ with $\dot{g}_I = \frac{\partial g_I}{\partial \beta^T}$ where \dot{g}_I is of dimension $(p+1) \times p$.

For an initial value of $\beta^{[k]}$, the iterative Equation (??) is used to compute $\dot{Q}(\beta^{[k]})$ and $\ddot{Q}(\beta^{[k]})$. These two components are then used to update the value $\beta^{[k]}$. The cycle of iteration continues in this way. Qu and Lindsay (2003) showed that the estimator of $\beta^{[k]}$ obtained by minimizing the objective function (??) is consistent and asymptotically normal with

$$\hat{\beta}^{[k]} - \beta^{[k]} \sim N\left(0, (E[\dot{g}_I^T] E[C_I^{-1}] E[\dot{g}_I])^{-1}\right).$$

Using the converged estimates, the forecasting equations are written as: For $t = 1 : T$, and $j = 2, 3, \dots, T-1$.

$$\left[\begin{pmatrix} y_{t-j+1}^{[1]} \\ y_{t-j+1}^{[2]} \end{pmatrix}, \begin{pmatrix} y_{t-j}^{[1]} \\ y_{t-j}^{[2]} \end{pmatrix} \right] = \left[\begin{pmatrix} \hat{\rho}_1 * y_{t-j}^{[1]} + \hat{\alpha}_{1,t-j} * y_{t-j}^{[2]} + d_{t-j+1}^{[1]} \\ \hat{\rho}_2 * y_{t-j}^{[2]} + \hat{\alpha}_{2,t-j} * y_{t-j}^{[1]} + d_{t-j+1}^{[2]} \end{pmatrix}, \begin{pmatrix} y_{t-j}^{[1]} \\ y_{t-j}^{[2]} \end{pmatrix} \right],$$

where $(d_{t-j+1}^{[1]}, d_{t-j+1}^{[2]})$ follows

$$\text{BIVPoisson}(\hat{\mu}_{t-j+1}^{[1]} - \hat{\rho}_1 \hat{\mu}_{t-j}^{[1]} - \hat{\alpha}_{1,t-j} \hat{\mu}_{t-j}^{[2]}, \hat{\mu}_{t-j+1}^{[2]} - \hat{\rho}_2 \hat{\mu}_{t-j}^{[2]} - \hat{\alpha}_{2,t-j} \hat{\mu}_{t-j}^{[1]}, \hat{\rho}_{12,t-j+1})$$

and for $t = T$ and $j = 1, 2, 3, \dots$

$$y_{T+j}^{[1]} = \rho_1^j * y_T^{[1]} + \sum_{i=0}^{j-1} \rho_1^i * \alpha_{1,T+j-i-1} * y_{T+j-i-1}^{[2]} + \sum_{i=0}^{j-1} \rho_1^i * d_{T+j-i}^{[1]} \quad (3.9)$$

$$y_{T+j}^{[2]} = \rho_2^j * y_T^{[2]} + \sum_{i=0}^{j-1} \rho_2^i * \alpha_{2,T+j-i-1} * y_{T+j-i-1}^{[1]} + \sum_{i=0}^{j-1} \rho_2^i * d_{T+j-i}^{[2]}, \quad (3.10)$$

where $[d_{T+j}^{[1]}, d_{T+j}^{[2]}]$ follows

$$\text{BIVPoisson}(\hat{\mu}_{T+j}^{[1]} - \hat{\rho}_1 \hat{\mu}_{T+j-1}^{[1]} - \hat{\alpha}_{1,T+j-1} \hat{\mu}_{T+j-1}^{[2]}, \hat{\mu}_{T+j}^{[2]} - \hat{\rho}_2 \hat{\mu}_{T+j-1}^{[2]} - \hat{\alpha}_{2,T+j-1} \hat{\mu}_{T+j-1}^{[1]}, \hat{\rho}_{12,T+j})$$

4 Simulation Study

In this section, we present a simulation study that consists of generating stationary and non-stationary bivariate time series of Poisson counts under the assumption of a 2×1 covariate vector of the form

$$x_{t1} = \begin{cases} -1 + t & (t = 1, \dots, T/4), \\ \text{rnorm}(1, 0, 1) & (t = (T/4) + 1, \dots, 3T/4), \\ 1 + t & (t = (3T/4) + 1, \dots, T), \end{cases}$$

where `rnorm` generates random numbers from normal distribution and the second element x_{t2} is fixed Poisson with parameter 2. The simulation study also consists of estimating the regression, serial- and cross-correlation parameters using the approaches mentioned in section 3. In this context, various values of the sample size are considered, $T = 60, 100$ and 500 , along with different values of the cross-correlation/serial-correlation parameters, $\rho_k = 0.2, 0.6$, with $\rho_{12} = 0.3, 0.9$ and $\alpha_{k,1} = 0.2, 0.6$, where $\beta_1^{[k]} = \beta_2^{[k]} = 1$ for $k = 1, 2$. These values were taken to assess performance of the estimating algorithm under low and high correlation values. For the stationary set-up, $x_{tk} = \bar{x}_k$ where $\bar{x}_k = \sum_{t=1}^T x_{tk}/T$.

These estimates are obtained using small initial values. From the simulation tables, it is remarked that the GQL regression, serial and bivariate correlation estimates are consistent and close to the population values and GQL yields far more efficient estimates than GMM under both stationary and non-stationary set-ups, while almost equally efficient estimates as CMLE under the stationary set-ups. Note for the non-stationary set-ups, we implemented only GQL and GMM adaptive since CMLE was computationally not feasible. As the number of time points increases to $T = 100$ and $T = 500$, the standard errors decrease, with the standard errors under non-stationary set-up relatively lower than under stationary set-up, irrespective of the size of the serial- and cross-correlations.

However, a few non-convergent simulations were observed particularly under non-stationary set-up, as shown in the table below:

The non-convergent simulations were mainly due to either the covariance matrix was ill-conditioned or the Hessian matrix was near to singularity. To overcome this problem, we took the initial estimates for GQL in the non-stationary set-up as the converged estimates obtained under stationary set-up. This drastically reduced the number of non-convergent simulations under the non-stationary set-up. This indicates that the proper choice of initial values of the estimates to run the algorithm to obtain GQL estimates under non-stationary set-up is important. Note that, in these tables, the non-stationary serial- and cross-correlations for the different lags as computed in Equations (2.4) and (3.7) and the bivariate dependence $\rho_{12,t}$ are not all displayed due to their high dimensions but are reliable values in the interval (0,1). It is remarked that the GQL estimating equation performs very well for bivariate time series with cross-correlation up to 0.7.

Table 1: Estimates of the parameters (standard errors) under stationary set-up for $\rho_k = [\rho_1 = 0.6, \rho_2 = 0.6]$ and $\alpha_{k,1} = [\alpha_{1,1} = 0.2, \alpha_{2,1} = 0.2]$

$\rho_{12,1}$	T	Method	$\hat{\beta}_1^{[1]}$	$\hat{\beta}_2^{[1]}$	$\hat{\beta}_1^{[2]}$	$\hat{\beta}_2^{[2]}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{2,1}$	$\hat{\rho}_{12,1}$
0.3	60	CMLE	1.0297 (0.0025)	1.0115 (0.0091)	1.0225 (0.0013)	1.0541 (0.0047)	0.5940	0.6049	0.1917	0.2011	0.2942
		GMM	1.2990 (0.0849)	1.2426 (0.0834)	1.2402 (0.0857)	1.2954 (0.0825)					
		GQL	1.0294 (0.0022)	1.0091 (0.0087)	1.0221 (0.0012)	1.0877 (0.0046)	0.5921	0.6053	0.1957	0.2098	0.2998
	100	CMLE	0.9946 (0.00130)	1.0114 (0.00075)	0.9964 (0.00027)	0.9989 (0.00016)	0.6078	0.5993	0.2030	0.2029	0.2938
		GMM	0.9757 (0.06580)	1.2317 (0.06741)	0.9730 (0.06322)	0.9770 (0.06240)					
		GQL	0.9930 (0.0011)	1.0131 (0.0007)	0.9942 (0.00025)	0.9989 (0.00015)	0.6092	0.5931	0.2055	0.2078	0.2981
	500	CMLE	1.0258 (0.00068)	0.9921 (0.00015)	0.9930 (0.00016)	0.9902 (0.00009)	0.6016	0.6088	0.2073	0.2060	0.2917
		GMM	1.211 (0.03719)	0.9824 (0.03510)	0.9847 (0.03202)	0.9809 (0.03831)					
		GQL	1.0251 (0.00060)	0.9957 (0.00013)	0.9976 (0.00014)	0.9948 (0.00008)	0.6069	0.6067	0.2046	0.2072	0.2958
0.9	60	CMLE	0.9932 (0.0048)	0.9994 (0.0097)	0.9992 (0.0069)	0.9997 (0.0046)	0.5981	0.5921	0.1906	0.1930	0.8910
		GMM	0.9790 (0.0886)	0.9741 (0.0878)	0.9791 (0.0895)	0.9755 (0.0911)					
		GQL	0.9949 (0.0045)	0.9956 (0.0095)	0.9997 (0.0067)	0.9947 (0.0043)	0.5982	0.5903	0.1963	0.1979	0.8946
	100	CMLE	1.0195 (0.0015)	1.0199 (0.0026)	0.9943 (0.0010)	0.9959 (0.0008)	0.6083	0.6062	0.1932	0.1913	0.8998
		GMM	1.2951 (0.0617)	1.2343 (0.0613)	0.9783 (0.0611)	0.9763 (0.0634)					
		GQL	1.0118 (0.0012)	1.0199 (0.0022)	0.9969 (0.0009)	0.9993 (0.0007)	0.6083	0.6084	0.1999	0.1945	0.8957
	500	CMLE	1.0914 (0.00055)	1.0195 (0.00014)	0.9972 (0.00013)	0.9918 (0.00024)	0.6081	0.6060	0.2043	0.1988	0.8972
		GMM	1.2951 (0.02310)	1.2316 (0.02460)	0.9851 (0.02352)	0.9814 (0.02013)					
		GQL	1.0936 (0.00051)	1.0079 (0.00012)	0.9965 (0.00011)	0.9993 (0.00021)	0.6071	0.6095	0.2010	0.1953	0.8963

Table 2: Estimates of the parameters (standard errors) under non-stationary set-up for $\rho_k = [\rho_1 = 0.6, \rho_2 = 0.6]$ and $\alpha_{k,1} = [\alpha_{1,1} = 0.2, \alpha_{2,1} = 0.2]$

$\rho_{12,1}$	T	Method	$\hat{\beta}_1^{[1]}$	$\hat{\beta}_2^{[1]}$	$\hat{\beta}_1^{[2]}$	$\hat{\beta}_2^{[2]}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{2,1}$	$\hat{\rho}_{12,1}$	
0.3	60	GMM	1.3123 (0.0837)	1.3095 (0.0801)	1.3140 (0.0828)	0.9721 (0.0889)						
		GQL	1.0103 (0.0019)	1.0024 (0.0045)	1.0791 (0.0007)	0.9979 (0.0027)	0.6097	0.6062	0.1957	0.2059	0.2943	
	100	GMM	0.9791 (0.05950)	1.3690 (0.05758)	0.9720 (0.05726)	0.9766 (0.05406)						
		GQL	0.9979 (0.00080)	1.0159 (0.00012)	0.9988 (0.00015)	0.9980 (0.00012)	0.6056	0.5962	0.2022	0.2049	0.2951	
	500	GMM	0.9824 (0.02314)	0.9843 (0.02229)	0.9883 (0.02187)	0.9841 (0.02867)						
		GQL	0.9912 (0.00040)	0.9971 (0.00011)	0.9956 (0.00012)	0.9959 (0.00003)	0.6087	0.6039	0.2044	0.2054	0.2929	
0.9	60	GMM	1.3883 (0.0827)	0.9712 (0.0818)	0.9761 (0.0823)	1.3013 (0.0893)						
		GQL	1.0111 (0.0021)	0.9941 (0.0057)	0.9979 (0.0042)	1.0731 (0.0031)	0.5945	0.5964	0.2061	0.2059	0.9012	
	100	GMM	1.3616 (0.0433)	1.3133 (0.0466)	0.9909 (0.0409)	0.9908 (0.0445)						
		GQL	1.0151 (0.0011)	1.0188 (0.0014)	0.9954 (0.0007)	0.9960 (0.0005)	0.6039	0.6046	0.2021	0.1911	0.8978	
	500	GMM	1.3955 (0.02882)	1.3053 (0.02482)	0.9923 (0.02703)	0.9966 (0.02435)						
		GQL	1.0101 (0.00042)	1.0211 (0.00011)	0.9965 (0.00007)	0.9967 (0.00013)	0.6062	0.6022	0.2012	0.1911	0.8969	

Table 3: Estimates of the parameters (standard errors) under stationary set-up for $\rho_k = [\rho_1 = 0.2, \rho_2 = 0.2]$ and $\alpha_{k,1} = [\alpha_{1,1} = 0.2, \alpha_{2,1} = 0.6]$

$\rho_{12,1}$	T	Method	$\hat{\beta}_1^{[1]}$	$\hat{\beta}_2^{[1]}$	$\hat{\beta}_1^{[2]}$	$\hat{\beta}_2^{[2]}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{2,1}$	$\hat{\rho}_{12,1}$
0.3	60	CMLE	0.9972 (0.0044)	0.9991 (0.0059)	0.9914 (0.0035)	1.0939 (0.0079)	0.1930	0.2029	0.1947	0.5922	0.2943
		GMM	0.9759 (0.0880)	0.9767 (0.0841)	0.9735 (0.0864)	1.3895 (0.0803)					
		GQL	0.9953 (0.0041)	0.9957 (0.0055)	0.9971 (0.0034)	1.0828 (0.0076)	0.1949	0.2053	0.1986	0.5910	0.2921
	100	CMLE	0.9923 (0.0015)	0.9914 (0.0026)	1.0936 (0.0017)	1.0910 (0.0009)	0.1910	0.2018	0.2011	0.5996	0.2962
		GMM	0.9751 (0.0527)	0.9719 (0.0500)	1.3222 (0.0550)	1.3754 (0.0519)					
		GQL	0.9951 (0.0012)	0.9918 (0.0022)	1.0918 (0.0014)	1.0951 (0.0008)	0.1985	0.2008	0.2080	0.5930	0.2925
	500	CMLE	0.9958 (0.00048)	0.9901 (0.00099)	0.9995 (0.00088)	0.9954 (0.00045)	0.2042	0.1970	0.2011	0.5955	0.2915
		GMM	0.9873 (0.02501)	0.9899 (0.02874)	0.9881 (0.02707)	0.9849 (0.02921)					
		GQL	0.9925 (0.00046)	0.9992 (0.00098)	0.9959 (0.00087)	0.9961 (0.00042)	0.2072	0.1969	0.2059	0.5948	0.2923
0.9	60	CMLE	0.9915 (0.0015)	0.9993 (0.0035)	1.0963 (0.0058)	1.0994 (0.0066)	0.1938	0.1919	0.2093	0.5965	0.8983
		GMM	0.9710 (0.0835)	0.9774 (0.0818)	1.3822 (0.0856)	1.3663 (0.0839)					
		GQL	0.9913 (0.0011)	0.9911 (0.0034)	1.0891 (0.0057)	1.0899 (0.0062)	0.1905	0.1976	0.2091	0.5991	0.8992
	100	CMLE	0.9977 (0.00044)	0.9910 (0.00016)	0.9912 (0.00036)	0.9984 (0.00028)	0.2055	0.2085	0.2021	0.5964	0.8927
		GMM	0.9720 (0.05106)	0.9752 (0.05392)	0.9759 (0.05676)	0.9771 (0.05481)					
		GQL	0.9911 (0.00041)	0.9946 (0.00015)	0.9921 (0.00036)	0.9939 (0.00027)	0.2064	0.2073	0.2004	0.5924	0.8970
	500	CMLE	1.0281 (0.00025)	0.9989 (0.00009)	0.9969 (0.00016)	0.9997 (0.00010)	0.2064	0.2074	0.2078	0.5935	0.8982
		GMM	1.2408 (0.02172)	0.9840 (0.02830)	0.9885 (0.02108)	0.9822 (0.02986)					
		GQL	1.0108 (0.00022)	0.9905 (0.00008)	0.9998 (0.00013)	0.9915 (0.00009)	0.2034	0.2038	0.2005	0.5909	0.8981

Table 4: Estimates of the parameters (standard errors) under non-stationary set-up for $\rho_k = [\rho_1 = 0.2, \rho_2 = 0.2]$ and $\alpha_{k,1} = [\alpha_{1,1} = 0.2, \alpha_{2,1} = 0.6]$

$\rho_{12,1}$	T	Method	$\hat{\beta}_1^{[1]}$	$\hat{\beta}_2^{[1]}$	$\hat{\beta}_1^{[2]}$	$\hat{\beta}_2^{[2]}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{2,1}$	$\hat{\rho}_{12,1}$		
0.3	60	GMM	0.9796 (0.0933)	0.9707 (0.0919)	0.9710 (0.0989)	1.3441 (0.0920)							
		GQL	0.9955 (0.0033)	0.9968 (0.0041)	0.9975 (0.0021)	1.0121 (0.0047)	0.1979	0.2015	0.1991	0.6012	0.2947		
	100	GMM	0.9745 (0.0667)	0.9710 (0.0611)	1.3641 (0.0679)	1.3507 (0.0624)							
		GQL	0.9954 (0.0019)	0.9946 (0.0017)	1.0523 (0.0012)	1.0475 (0.0006)	0.2011	0.2002	0.2042	0.5945	0.2963		
	500	GMM	0.9899 (0.02710)	0.9813 (0.02915)	1.2841 (0.02803)	0.9809 (0.02207)							
		GQL	0.9955 (0.00028)	0.9997 (0.00071)	1.0012 (0.00066)	0.9985 (0.00025)	0.2021	0.1980	0.2041	0.5957	0.2967		
	0.9	60	GMM	0.9708 (0.0889)	0.9748 (0.0831)	1.3615 (0.0811)	1.3363 (0.0827)						
			GQL	0.9935 (0.0010)	0.9947 (0.0022)	1.0147 (0.0034)	1.0218 (0.0046)	0.1956	0.1982	0.2043	0.5993	0.9001	
		100	GMM	0.9775 (0.05551)	0.9701 (0.05667)	0.9760 (0.05492)	0.9770 (0.05332)						
			GQL	0.9919 (0.00024)	0.9957 (0.00012)	0.9926 (0.00023)	0.9944 (0.00021)	0.2084	0.2087	0.2002	0.5967	0.8981	
		500	GMM	1.2214 (0.01425)	0.9817 (0.01332)	0.9732 (0.01630)	0.9790 (0.01361)						
			GQL	1.0022 (0.00021)	0.9923 (0.00004)	0.9901 (0.00011)	0.9929 (0.00007)	0.2021	0.2022	0.2003	0.5918	0.8989	

Table 5: Estimates of the parameters (standard errors) under stationary set-up for $\rho_k = [\rho_1 = 0.2, \rho_2 = 0.2]$ and $\alpha_{k,1} = [\alpha_{1,1} = 0.6, \alpha_{2,1} = 0.2]$

$\rho_{12,1}$	T	Method	$\hat{\beta}_1^{[1]}$	$\hat{\beta}_2^{[1]}$	$\hat{\beta}_1^{[2]}$	$\hat{\beta}_2^{[2]}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{2,1}$	$\hat{\rho}_{12,1}$	
0.3	60	CMLE	0.9925 (0.0051)	0.9953 (0.0098)	0.9900 (0.0077)	0.9918 (0.0048)	0.2021	0.2006	0.6018	0.1901	0.2943	
		GMM	0.9793 (0.0784)	0.9772 (0.0760)	0.9739 (0.0755)	0.9750 (0.0739)						
		GQL	0.9955 (0.0049)	0.9964 (0.0096)	0.9946 (0.0075)	0.9936 (0.0046)	0.2071	0.2002	0.6049	0.1977	0.2902	
	100	CMLE	1.0911 (0.0018)	1.0954 (0.0025)	0.9961 (0.0027)	0.9919 (0.0022)	0.2030	0.2002	0.6097	0.2017	0.2975	
		GMM	1.3211 (0.0460)	1.3233 (0.0453)	0.9732 (0.0490)	0.9934 (0.0511)						
		GQL	1.0912 (0.0015)	1.0917 (0.0022)	0.9973 (0.0024)	0.9956 (0.0019)	0.2068	0.2037	0.6008	0.2054	0.2977	
	500	CMLE	0.9973 (0.00065)	0.9998 (0.00098)	0.9957 (0.00040)	0.9937 (0.00024)	0.2057	0.2013	0.6022	0.2096	0.2919	
		GMM	0.9866 (0.01100)	0.9800 (0.01676)	0.9888 (0.01102)	0.9824 (0.01453)						
		GQL	0.9957 (0.00063)	0.9929 (0.00097)	0.9985 (0.00039)	0.9991 (0.00021)	0.2043	0.2036	0.6013	0.2058	0.2911	
	0.9	60	CMLE	1.0283 (0.0078)	0.9950 (0.0088)	0.9982 (0.0047)	0.9933 (0.0028)	0.2073	0.2082	0.6018	0.1978	0.8911
			GMM	1.3834 (0.0808)	0.9725 (0.0896)	0.9792 (0.0850)	0.9742 (0.0883)					
			GQL	1.0171 (0.0076)	0.9985 (0.0087)	0.9976 (0.0043)	0.9959 (0.0027)	0.2004	0.2010	0.6036	0.1982	0.8915
100		CMLE	1.0918 (0.0015)	1.0910 (0.0027)	1.0928 (0.0038)	0.9910 (0.0009)	0.2040	0.2018	0.6087	0.2082	0.8975	
		GMM	1.3920 (0.0532)	1.3215 (0.0518)	1.3686 (0.0560)	0.9746 (0.0514)						
		GQL	1.0974 (0.0013)	1.0996 (0.0024)	1.0973 (0.0037)	0.9918 (0.0008)	0.2008	0.2023	0.6026	0.2020	0.8974	
500		CMLE	0.9959 (0.00101)	0.9911 (0.00022)	0.9974 (0.00034)	0.9910 (0.00027)	0.2017	0.2098	0.6094	0.1960	0.8961	
		GMM	0.9845 (0.01501)	0.9998 (0.01969)	0.9931 (0.01051)	0.9978 (0.01760)						
		GQL	0.9919 (0.00090)	0.9978 (0.00021)	0.9971 (0.00032)	0.9945 (0.00025)	0.2021	0.2089	0.6045	0.1984	0.8906	

Table 6: Estimates of the parameters (standard errors) under non-stationary set-up for $\rho_k = [\rho_1 = 0.2, \rho_2 = 0.2]$ and $\alpha_{k,1} = [\alpha_{1,1} = 0.6, \alpha_{2,1} = 0.2]$

$\rho_{12,1}$	T	Method	$\hat{\beta}_1^{[1]}$	$\hat{\beta}_2^{[1]}$	$\hat{\beta}_1^{[2]}$	$\hat{\beta}_2^{[2]}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{2,1}$	$\hat{\rho}_{12,1}$		
0.3	60	GMM	0.9743 (0.0825)	0.9721 (0.0851)	0.9758 (0.0812)	0.9794 (0.0825)							
		GQL	0.9966 (0.0031)	0.9975 (0.0042)	0.9959 (0.0051)	0.9947 (0.0024)	0.2019	0.2001	0.6031	0.1985	0.2942		
	100	GMM	1.3225 (0.0522)	1.3749 (0.0542)	1.3447 (0.0533)	0.9952 (0.0525)							
		GQL	1.0147 (0.0012)	1.0213 (0.0016)	1.0013 (0.0018)	0.9976 (0.0013)	0.2064	0.2025	0.6004	0.2033	0.2982		
	500	GMM	0.9723 (0.03010)	0.9828 (0.03899)	0.9820 (0.03438)	0.9875 (0.03119)							
		GQL	0.9965 (0.00054)	0.9947 (0.00066)	0.9991 (0.00031)	0.9995 (0.00015)	0.2059	0.2061	0.6007	0.2041	0.2933		
	0.9	60	GMM	1.3466 (0.0768)	0.9791 (0.0747)	0.9719 (0.0719)	0.9735 (0.0778)						
			GQL	1.0081 (0.0055)	0.9988 (0.0062)	0.9981 (0.0033)	0.9969 (0.0022)	0.2002	0.2005	0.6021	0.1988	0.8945	
		100	GMM	1.3168 (0.0435)	1.3613 (0.0485)	1.3568 (0.0445)	0.9920 (0.0453)						
			GQL	1.0416 (0.0011)	1.0382 (0.0019)	1.0258 (0.0013)	0.9952 (0.0006)	0.2016	0.2012	0.6015	0.2017	0.8986	
		500	GMM	0.9823 (0.01124)	0.9880 (0.01720)	0.9828 (0.01422)	0.9831 (0.01112)						
			GQL	0.9931 (0.00070)	0.9983 (0.00017)	0.9979 (0.00025)	0.9958 (0.00022)	0.2012	0.2071	0.6032	0.1989	0.8929	

Table 7: Estimates of the parameters (standard errors) under stationary set-up for $\rho_k = [\rho_1 = 0.2, \rho_2 = 0.2]$ and $\alpha_{k,1} = [\alpha_{1,1} = 0.2, \alpha_{2,1} = 0.2]$

$\rho_{12,1}$	T	Method	$\hat{\beta}_1^{[1]}$	$\hat{\beta}_2^{[1]}$	$\hat{\beta}_1^{[2]}$	$\hat{\beta}_2^{[2]}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{2,1}$	$\hat{\rho}_{12,1}$	
0.3	60	CMLE	1.0904 (0.0098)	1.0962 (0.0085)	0.9931 (0.0076)	0.9916 (0.0056)	0.2118	0.1987	0.2040	0.2031	0.2910	
		GMM	1.3726 (0.0871)	1.3991 (0.0824)	0.9792 (0.0831)	0.9738 (0.0837)						
		GQL	1.0941 (0.0096)	1.0971 (0.0082)	0.9985 (0.0073)	0.9913 (0.0055)	0.2173	0.1939	0.2058	0.2098	0.2922	
	100	CMLE	1.0991 (0.0015)	0.9918 (0.0024)	0.9917 (0.0017)	1.0923 (0.0016)	0.2013	0.2083	0.2026	0.2021	0.2979	
		GMM	1.3084 (0.0539)	0.9741 (0.0550)	0.9740 (0.0514)	1.3451 (0.0556)						
		GQL	1.0916 (0.0012)	0.9947 (0.0022)	0.9956 (0.0015)	1.0992 (0.0014)	0.2070	0.2044	0.2005	0.2057	0.2952	
	500	CMLE	0.9934 (0.00025)	0.9976 (0.00019)	0.9913 (0.00034)	0.9922 (0.00043)	0.2029	0.2082	0.2093	0.2092	0.2942	
		GMM	0.9891 (0.02270)	0.9817 (0.02052)	0.9835 (0.02505)	0.9811 (0.02187)						
		GQL	0.9922 (0.00023)	0.9982 (0.00017)	0.9945 (0.00031)	0.9971 (0.00041)	0.2021	0.2029	0.2043	0.2056	0.2977	
	0.9	60	CMLE	0.9951 (0.0018)	0.9954 (0.0085)	1.0961 (0.0098)	0.9980 (0.0057)	0.2065	0.2035	0.2032	0.2049	0.8922
			GMM	0.9789 (0.0721)	0.9717 (0.0718)	1.3178 (0.0717)	0.9767 (0.0795)					
			GQL	0.9964 (0.0016)	0.9928 (0.0081)	1.0987 (0.0097)	0.9914 (0.0054)	0.2015	0.2014	0.2096	0.2069	0.8932
100		CMLE	0.9905 (0.0008)	1.0965 (0.0015)	0.9990 (0.0026)	1.0931 (0.0034)	0.2046	0.2096	0.2034	0.2003	0.8914	
		GMM	0.9712 (0.0546)	1.3853 (0.0511)	0.9730 (0.0569)	1.3491 (0.0585)						
		GQL	0.9932 (0.0007)	1.0985 (0.0013)	0.9948 (0.0023)	1.0917 (0.0031)	0.2013	0.2006	0.2071	0.2013	0.8971	
500		CMLE	0.9980 (0.00015)	0.9901 (0.00029)	0.9919 (0.00017)	0.9946 (0.00059)	0.2087	0.2052	0.2005	0.1941	0.8958	
		GMM	0.9823 (0.01066)	0.9801 (0.01137)	0.9802 (0.01536)	0.9788 (0.01429)						
		GQL	0.9916 (0.00014)	0.9973 (0.00028)	0.9941 (0.00014)	0.9905 (0.00057)	0.2068	0.2016	0.2094	0.1988	0.8996	

Table 8: Estimates of the parameters (standard errors) under non-stationary set-up for $\rho_k = [\rho_1 = 0.2, \rho_2 = 0.2]$ and $\alpha_{k,1} = [\alpha_{1,1} = 0.2, \alpha_{2,1} = 0.2]$

$\rho_{12,1}$	T	Method	$\hat{\beta}_1^{[1]}$	$\hat{\beta}_2^{[1]}$	$\hat{\beta}_1^{[2]}$	$\hat{\beta}_2^{[2]}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{2,1}$	$\hat{\rho}_{12,1}$	
0.3	60	GMM	1.3034 (0.0898)	1.3649 (0.0827)	0.9791 (0.0874)	0.9710 (0.0825)						
		GQL	1.0105 (0.0065)	1.0207 (0.0042)	0.9988 (0.0051)	0.9935 (0.0037)	0.2089	0.2098	0.2045	0.2064	0.2939	
	100	GMM	1.3467 (0.0655)	0.9753 (0.0644)	0.9759 (0.0616)	1.3866 (0.0649)						
		GQL	1.0527 (0.0011)	0.9971 (0.0013)	0.9964 (0.0012)	1.0613 (0.0011)	0.2021	0.2023	0.2002	0.2041	0.3013	
	500	GMM	0.9796 (0.02892)	0.9711 (0.02004)	0.9770 (0.02779)	0.9737 (0.02425)						
		GQL	0.9937 (0.00019)	0.9988 (0.00012)	0.9958 (0.00022)	0.9978 (0.00033)	0.2012	0.2017	0.2027	0.2024	0.2986	
0.9	60	GMM	0.9781 (0.0913)	0.9768 (0.0941)	1.3244 (0.0990)	0.9721 (0.0970)						
		GQL	0.9975 (0.0012)	0.9947 (0.0062)	1.0432 (0.0071)	0.9951 (0.0042)	0.2007	0.2008	0.2041	0.2052	0.8961	
	100	GMM	0.9739 (0.0646)	1.3335 (0.0676)	0.9935 (0.0632)	1.3017 (0.0610)						
		GQL	0.9941 (0.0006)	1.0991 (0.0011)	0.9963 (0.0021)	1.0456 (0.0024)	0.2009	0.2003	0.2024	0.2011	0.8982	
	500	GMM	0.9832 (0.03041)	0.9828 (0.03204)	0.9861 (0.0335)	0.9830 (0.03706)						
		GQL	0.9929 (0.00011)	0.9984 (0.00022)	0.9955 (0.00012)	0.9956 (0.00028)	0.2011	0.2012	0.2051	0.2005	0.9001	
			(0.00011)	(0.00022)	(0.00012)	(0.00028)						

5 Conclusion

In this paper, a new full BINAR(1) Poisson model (FBINAR1M2) is designed under the assumption that the cross-correlation is formed via the non-stationary bivariate innovative series and the cross-correlation between observations at the current time point with observation of the second variate at the previous time point. As a substitute to an extremely complicated conditional likelihood estimation approach, a consistent and equally efficient GQL approach is developed. Through simulation studies, it is shown that the GQL approach performs remarkably well and yields consistent estimates of the regression, serial- and cross-correlation parameters that are more efficient than those obtained under GMM and CMLE.

Table 9: Number of non-convergent simulations under GQL, CMLE and GMM for the different combinations under stationary set-up and under GQL and GMM for the different combinations under non-stationary set-up.

$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{2,1}$	$\hat{\rho}_{12,1}$	T	Stationary			Non-stationary	
						GQL	CMLE	GMM	GQL	GMM
0.6	0.6	0.2	0.2	0.3	60	140	270	160	1800	1850
					100	125	250	150	1700	1740
					500	100	215	120	1000	1100
				0.9	60	150	280	175	1830	1875
					100	130	260	160	1750	1790
					500	115	200	130	1100	1180
0.2	0.2	0.6	0.2	0.3	60	155	270	160	1750	1800
					100	125	240	130	1680	1725
					500	100	200	110	925	990
				0.9	60	140	280	175	1810	1860
					100	125	250	135	1710	1770
					500	105	195	120	990	1040
0.2	0.2	0.2	0.6	0.3	60	150	275	165	1740	1810
					100	120	240	130	1660	1720
					500	105	210	110	920	980
				0.9	60	140	270	170	1800	1850
					100	125	250	135	1700	1750
					500	110	190	125	970	1010
0.2	0.2	0.2	0.2	0.3	60	140	270	160	1630	1700
					100	110	230	120	1590	1630
					500	100	200	100	870	910
				0.9	60	135	260	150	1720	1780
					100	120	240	130	1625	1650
					500	105	170	110	890	925

References

- Al Osh, M. and Alzaid, A. (1987), "First-order integer-valued autoregressive process," *Journal of Time Series Analysis*, 8, 261–275.
- Jowaheer, V. and Sutradhar, B. (2005), "Fitting lower order non-stationary autocorrelation models to the time series of Poisson counts," *Wseas Transaction on Mathematics*, 4, 1109–2769.
- Jowaheer, V. and Sutradhar, B. (2005), "Fitting lower order nonstationary autocorrelation models to the time series of Poisson counts," *Transactions on Mathematics*, 4, 427–434.
- Kim, H. and Park, Y. (2008), "A non-stationary integer-valued autoregressive model," *Statistical Papers*, 49, 485–502.
- Kocherlakota, S. and Kocherlakota, K. (2001), "Regression in the Bivariate Poisson Distribution," *Communications in Statistics-Theory and Methods*, 30(5), 815–825.
- Liang, K. and Zeger, S. (1986), "Longitudinal data analysis using generalized linear models," *Biometrika*, 73, 13–22.
- Mamode khan, N. and Jowaheer, V. (2013), "Comparing joint GQL estimation and GMM adaptive estimation in COM-Poisson longitudinal regression model," *Commun Stat-Simul C.*, 42(4), 755–770.
- Mamode Khan, N., Sunecher, Y., and Jowaheer, V. (2016b), "Modelling a Non-Stationary BINAR(1) Poisson Process," *Journal of Statistical Computation and Simulation*, 86, 3106–3126.
- McKenzie, E. (1985), "Some simple models for discrete variate time series," *Water Resources Bulletin*, 21(4), 645–650.
- Pedeli, X. and Karlis, D. (2009), "Bivariate INAR(1) Models," Tech. rep., Athens University of Economics.
- (2011), "A bivariate INAR(1) process with application," *Statistical Modelling: An International Journal*, 11, 325–349.
- (2013a), "Some properties of multivariate INAR(1) processes." *Computational Statistics and Data Analysis*, 67, 213–225.
- Qu, A. and Lindsay, B. (2003), "Building adaptive estimating equations when inverse of covariance estimation is difficult," *Journal of Royal Statistical Society*, 65, 127–142.
- Steutel, F. and Van Harn, K. (1979), "Discrete analogues of self-decomposability and stability," *The Annals of Probability*, 7, 3893–899.
- Sutradhar, B. and Das, K. (1999), "On the efficiency of regression estimators in generalised linear models for longitudinal data," *Biometrika*, 86, 459–65.

Sutradhar, B., Jowaheer, V., and Rao, P. (2014), “Remarks on asymptotic efficient estimation for regression effects in stationary and non-stationary models for panel count data,” *Brazilian Journal of Probability and Statistics*, 28(2), 241–254.

A Derivations

$$\begin{aligned}
E(Y_t^{[1]}Y_{t+1}^{[2]}) &= E\left[Y_t^{[1]}(\rho_2 * Y_t^{[2]} + \alpha_{2,t} * Y_t^{[1]} + d_{t+1}^{[2]})\right] \\
&= E_{Y_t^{[2]}}E[Y_t^{[1]}(\rho_2 * Y_t^{[2]}) | Y_t^{[2]}] + E_{Y_t^{[1]}}E[Y_t^{[1]}(\alpha_{2,t} * Y_t^{[1]}) | Y_t^{[1]}] + E[Y_t^{[1]}d_{t+1}^{[2]}] \\
&= \rho_2 E[Y_t^{[1]}Y_t^{[2]}] + \alpha_{2,t}E[(Y_t^{[1]})^2] + \mu_t^{[1]}(\mu_{t+1}^{[2]} - \rho_2\mu_t^{[2]} - \alpha_{2,t}\mu_t^{[1]}) \\
&= \rho_2 E[Y_t^{[1]}Y_t^{[2]}] + \alpha_{2,t}[\mu_t^{[1]} + (\mu_t^{[1]})^2] + \mu_t^{[1]}(\mu_{t+1}^{[2]} - \rho_2\mu_t^{[2]} - \alpha_{2,t}\mu_t^{[1]}) \\
&= \rho_2 E[Y_t^{[1]}Y_t^{[2]}] - \rho_2\mu_t^{[1]}\mu_t^{[2]} + \alpha_{2,t}\mu_t^{[1]} + \mu_t^{[1]}\mu_{t+1}^{[2]} \\
&= \rho_2 Cov(Y_t^{[1]}, Y_t^{[2]}) + \alpha_{2,t}Cov(Y_t^{[1]}, Y_t^{[1]}) + \mu_t^{[1]}\mu_{t+1}^{[2]} \tag{A.1}
\end{aligned}$$

$$\begin{aligned}
E(Y_t^{[1]}Y_{t+2}^{[2]}) &= E\left[Y_t^{[1]}(\rho_2 * Y_{t+1}^{[2]} + \alpha_{2,t+1} * Y_{t+1}^{[1]} + d_{t+2}^{[2]})\right] \\
&= E_{Y_{t+1}^{[2]}}E[Y_t^{[1]}(\rho_2 * Y_{t+1}^{[2]}) | Y_{t+1}^{[2]}] + E_{Y_{t+1}^{[1]}}E[Y_t^{[1]}(\alpha_{2,t+1} * Y_{t+1}^{[1]}) | Y_{t+1}^{[1]}] + E[Y_t^{[1]}d_{t+2}^{[2]}] \\
&= \rho_2 E[Y_t^{[1]}Y_{t+1}^{[2]}] + \alpha_{2,t+1}E[Y_t^{[1]}Y_{t+1}^{[1]}] + \mu_t^{[1]}(\mu_{t+2}^{[2]} - \rho_2\mu_{t+1}^{[2]} - \alpha_{2,t+1}\mu_{t+1}^{[1]}) \\
&= \rho_2 E[Y_t^{[1]}Y_{t+1}^{[2]}] - \rho_2\mu_t^{[1]}\mu_{t+1}^{[2]} + \alpha_{2,t+1}E[Y_t^{[1]}Y_{t+1}^{[1]}] - \alpha_{2,t+1}\mu_t^{[1]}\mu_{t+1}^{[1]} + \mu_t^{[1]}\mu_{t+2}^{[2]} \\
&= \rho_2 Cov(Y_t^{[1]}, Y_{t+1}^{[2]}) + \alpha_{2,t+1}Cov(Y_t^{[1]}, Y_{t+1}^{[1]}) + \mu_t^{[1]}\mu_{t+2}^{[2]} \tag{A.2}
\end{aligned}$$

Received: February 10, 2018

Accepted: June 6, 2018