ATHENS UNIVERSITY OF ECONOMICS

BIVARIATE INAR(1) MODELS

Xanthi Pedeli and Dimitris Karlis
Department of Statistics
Athens University of Economics and Business

Technical Report No 247, June, 2009

DEPARTMENT OF STATISTICS
TECHNICAL REPORT

Bivariate INAR(1) models

Xanthi Pedeli and Dimitris Karlis *
Department of Statistics
Athens University of Economics and Business

Abstract

The study of time series models for count data has become a topic of special interest during the last years. However, while research on univariate time series for counts now flourish, the literature on multivariate time series models for count data is notably more limited. In the present paper, a bivariate integer-valued autoregressive process of order 1 (BINAR(1)) is introduced. Emphasis is placed on the special cases of bivariate Poisson and bivariate negative binomial innovations. We discuss properties of the Poisson BINAR(1) model and propose estimation methods for its unknown parameters. A simulation experiment provides evidence for the superiority of maximum likelihood estimators. Issues of diagnostics and forecasting are considered and predictions are produced by means of the conditional forecast distribution. Estimation uncertainty is accomodated by taking advantage of the asymptotic normality of maximum likelihood estimators and constructing appropriate confidence intervals for the h-step-ahead conditional probability mass function. The proposed model is applied to a bivariate data series concerning syndromic surveillance during Athens 2004 Olympic Games.
Keywords: BINAR; count data; Poisson; negative binomial; bivariate time series.

1 Introduction

Multivariate count data occur in several different disciplines like epidemiology, marketing, criminology and engineering just to name a few. In many cases the data are observed across time leading to multivariate time series

*Corresponding author

data as, for example, when one studies the purchases of different products across time, or the occurrence of different diseases across time.

In the literature there are several models to fit univariate count time series models (see Davis et al., 1999). A commonly used class of such models consists of the so-called integer autoregressive time series models, introduced by McKenzie (1985) and Al-Osh and Alzaid (1987). The interested reader is referred to McKenzie (2003) and Jung and Tremayne (2006) for a brief but detailed review of such models. The literature on multivariate time series models for count data is less developed. Some interesting attempts have been made during the last decade but most of them do not arise in the context of INAR processes (see Gurmu and Elder, 2000; Chib and Winkelmann, 2001; Egan and Herriges, 2006). Among the models that have been built in the aforementioned setting are those of Latour (1997); Brännäs and Nordström (2000); Heinen and Rengifo (2007) and Quoreshi (2006).

The aim of this paper is to introduce and examine in detail a bivariate Poisson integer-valued autoregressive model of order 1 (BINAR(1)). To motivate the model consider the case of health surveillance systems and especially syndromic surveillance. In such systems the number of patients with some symptoms are counted in each hospital allowing for examining if there is a sudden increase of some symptoms implying perhaps some epidemic or some hazard for the public health. Syndromic surveillance is considered as more efficient than diagnostic surveillance since it is based on symptoms rather than diagnosis and thus is generally faster in creating alerts. Usually the data are small counts and since a number of different symptoms is counted we have multiple series of data. This leads to the need of creating appropriate time series models to handle multiple time series together as the interest is to examine if their concurrence implies a threat to public health.

The remaining of this paper is structured as follows. A general specification of the BINAR(1) process and alternative methods for the estimation of its unknown parameters are given in section 2. In sections 3 we concentrate on the special cases of bivariate Poisson and negative binomial innovations respectively. In section 4 we give a specification of the model residuals as a diagnostic tool while issues of forecasting are discussed in section 5. The advantages and drawbacks of the proposed estimators are presented in section 6 using simulation. An application to real data concerning cause-specific hospital admissions during Athens 2004 Olympic Games, follows in section 7. Some concluding remarks are presented in section 8.

2 The BINAR(1) Process

2.1 Model

Let $\mathbf{X}$ and $\mathbf{R}$ be non-negative integer-valued random 2-vectors. Let $\mathbf{A}$ be a 2×2 diagonal matrix with independent elements $\{\alpha_{jj}\}_{j=1,2}$. The bivariate integer-valued autoregressive process of order 1 can be defined as

$$\mathbf{X}_t = \mathbf{A} \circ \mathbf{X}_{t-1} + \mathbf{R}_t = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \circ \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} R_{1t} \\ R_{2t} \end{bmatrix}, \quad t \in \mathbb{Z} \tag{2.1}$$

where "$\circ$"is the binomial thinning operator defined as $\alpha \circ X = \sum_{i=1}^{X} Y_i = Y$, where $\{Y_i\}_{i=1}^{X}$ is a sequence of *iid* Bernoulli random variables such that $P(Y_i = 1) = \alpha = 1 - P(Y_i = 0)$ and $\alpha \in [0, 1]$ (Steutel and van Harn, 1979). In the bivariate case, the $\mathbf{A}\circ$ operation is a matricial operation which acts as the usual matrix multiplication keeping in the same time the properties of the binomial thinning operation. One can see that with the above definition the jth element, $j = 1, 2$ is given by $X_{jt} = \alpha_j \circ X_{j,t-1} + R_{jt}$. The elements $\mathbf{R}_t$ which entered the system in the interval $(t-1, t]$ are usually called as innovations.

Assuming independence between and within the thinning operations and $\{R_{jt}\}$ an *iid* sequence with mean λ_j and variance $\sigma_j^2 = \upsilon_j \lambda_j$, $\upsilon_j > 0, j = 1, 2,$ the unconditional first and second order moments based on second order stationarity conditions are:

$$E(X_{jt}) = \mu_{X_j} = \frac{\lambda_j}{1 - \alpha_j} \tag{2.2}$$

$$Var(X_{jt}) = \sigma_{X_j}^2 = \frac{(\alpha_j + \upsilon_j)\lambda_j}{1 - \alpha_j^2} \tag{2.3}$$

$$\text{Cov}\left(X_{jt}, X_{j,t+h}\right) = \gamma_{X_j}(h) = \alpha_j^h \sigma_{X_j}^2 \ ; \quad h = 1, 2, \ldots \tag{2.4}$$

$$\text{Corr}\left(X_{jt}, X_{j,t+h}\right) = \rho_{X_j}(h) = \alpha_j^h \ ; \quad h = 1, 2, \ldots \tag{2.5}$$

Obviously, the mean, variance and autocovariance functions can take only positive values, since λ_j,σ_j^2 and α_j are all positive. Depending on whether $\upsilon_j > 1$, $\upsilon_j \in (0, 1)$, or $\upsilon_j = 1$, the variance may be larger than the mean (overdispersion), smaller than the mean (underdispersion), or equal to the mean (equidispersion) respectively.

Dependence between the two series that comprise the BINAR(1) process is introduced by allowing for dependence between R_{1t} and R_{2t} while retaining all the previous assumptions fixed. Whatever the underlying joint distribution of $\{R_{1t}, R_{2t}\}$ is, it can be shown that the covariance between the innovations of the two series at time t, totally determines the covariance between the current value of the one process and the innovations of the other process at the same point in time t and vice versa (see Appendix I):

$$\mathrm{Cov}(X_{1t}, R_{2t}) = \mathrm{Cov}(R_{1t}, R_{2t}) \tag{2.6}$$

As expected, the covariance between the sequences $\{X_{1t}\}$ and $\{X_{2t}\}$ at time t is also affected by the corresponding "survival"parts of the two processes. More specifically it can be shown that,

$$\mathrm{Cov}(X_{1,t+h}, X_{2t}) = \frac{\alpha_1^h}{(1-\alpha_1\alpha_2)}\mathrm{Cov}(R_{1t}, R_{2t}) \ ; \quad h = 0, 1, \ldots \text{ and} \tag{2.7}$$

$$\mathrm{Corr}(X_{1,t+h}, X_{2t}) = \frac{\alpha_1^h\sqrt{(1-\alpha_1^2)(1-\alpha_2^2)}}{(1-\alpha_1\alpha_2)\sqrt{(\alpha_1+\upsilon_1)(\alpha_2+\upsilon_2)\lambda_1\lambda_2}}\mathrm{Cov}(R_{1t}, R_{2t}) \ ; \quad h = 0, 1, \ldots \tag{2.8}$$

Covariances and correlations between X_{1t} and $X_{2,t+h}$, $h = 0, 1, \ldots$, can be defined analogously (see Appendix I).

Note that (2.7) presumes that $\{\mathbf{X}_t\}$ is a strictly stationary process, i.e. that the joint distribution of $\begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix}$ is the same as that of $\begin{pmatrix} X_{1,t+h} \\ X_{2,t+h} \end{pmatrix}$, for all h. Using the analytical representations

$$\begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \circ \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} R_{1t} \\ R_{2t} \end{bmatrix} \tag{2.9}$$

and

$$\begin{pmatrix} X_{1,t+h} \\ X_{2,t+h} \end{pmatrix} = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \circ \begin{bmatrix} X_{1,t+h-1} \\ X_{2,t+h-1} \end{bmatrix} + \begin{bmatrix} R_{1,t+h} \\ R_{2,t+h} \end{bmatrix} \tag{2.10}$$

it is easy to see that strict stationarity does indeed hold for $\{\mathbf{X}_t\}$ since the variables involved in the right-hand sides of (2.9) and (2.10) have identical distributions (see also Latour, 1997).

2.2 Estimation

2.2.1 Yule-Walker estimation

As already noted, the structure of the BINAR(1) model implies that the two innovation series $\{R_{1t}, R_{2t}\}$ follow jointly a bivariate distribution. Let

$G_{R_1,R_2}(s_1, s_2)$ be the joint probability generating function (*jpgf*) of $\{R_{1t}, R_{2t}\}$. Then, the *jpgf* of $\mathbf{X}_t = \{X_{1t}, X_{2t}\}$ is given by

$$\begin{aligned} G_{\mathbf{X}_t}(\mathbf{s}) = G_{X_{1t},X_{2t}}(s_1, s_2) &= G_{X_{1,0}}(1 - \alpha_1^t + \alpha_1^t s_1) G_{X_{2,0}}(1 - \alpha_2^t + \alpha_2^t s_2) \\ &\times \prod_{i=0}^{t-1} G_{R_1,R_2}((1 - \alpha_1^i + \alpha_1^i s_1), (1 - \alpha_2^i + \alpha_2^i s_2)) \end{aligned} \tag{2.11}$$

which reduces to

$$G_{\mathbf{X}_t}(\mathbf{s}) = G_{X_{1t},X_{2t}}(s_1, s_2) = \prod_{i=0}^{\infty} G_{R_1,R_2}((1-\alpha_1^i+\alpha_1^i s_1), (1-\alpha_2^i+\alpha_2^i s_2)) \tag{2.12}$$

The moment generating function $M_{\mathbf{X}_t}(\mathbf{s}) = G_{\mathbf{X}_t}(e^{\mathbf{s}})$ can then be used to obtain appropriate sample moments for the estimation of the unknown model parameters.

2.2.2 Maximum-Likelihood estimation

The conditional density for the BINAR(1) model can be expressed as the convolution of two binomials, namely

$$f_1(x_1) = \binom{X_{1,t-1}}{x_1} \alpha_1^{x_1} (1 - \alpha_1)^{X_{1,t-1}-x_1} \tag{2.13}$$

$$f_2(x_2) = \binom{X_{2,t-1}}{x_2} \alpha_2^{x_2} (1 - \alpha_2)^{X_{2,t-1}-x_2}, \tag{2.14}$$

and a bivariate distribution of the form $f_3(k, s) = P(R_{1t} = k, R_{2t} = s)$. Thus the conditional density becomes

$$f(\mathbf{x}_t|\mathbf{x}_{t-1}, \boldsymbol{\theta}) = \sum_k \sum_s f_1(x_{1t} - k) f_2(x_{2t} - s) f_3(k, s) \tag{2.15}$$

where $\boldsymbol{\theta}$ is the vector of unknown parameters.
The conditional likelihood function is then given by

$$L(\boldsymbol{\theta}|\mathbf{x}) = \prod_{t=1}^{T} f(\mathbf{x}_t|\mathbf{x}_{t-1}, \boldsymbol{\theta}) \tag{2.16}$$

for some initial value $\mathbf{x}_0$ and hence maximization provides with the ML estimates. Numerical maximization is straightforward with standard statistical packages.

3 Parametric Cases

In this section we discuss two specific BINAR(1) models. The first one comes from the assumption that the innovations of the two series follow jointly a bivariate Poisson distribution. The second model assumes a bivariate negative binomial distribution for the two innovation processes. The two representations can be viewed as appropriate tools for modeling equidispersed and overdispersed bivariate time series respectively. Some additional specifications suited for time series data with negative correlation are also briefly considered.

3.1 The Poisson BINAR(1) Process

3.1.1 Model

Let assume that the joint probability mass function (*jpmf*) of the two innovation processes $\{R_{1t}, R_{2t}\}$ is a bivariate Poisson distribution given by

$$P(R_{1t} = x, R_{2t} = y) = e^{-(\lambda_1+\lambda_2-\phi)} \frac{(\lambda_1 - \phi)^x}{x!} \frac{(\lambda_2 - \phi)^y}{y!} \sum_{i=0}^{s} \binom{x}{i} \binom{y}{i} i! \left(\frac{\phi}{(\lambda_1 - \phi)(\lambda_2 - \phi)}\right)^i \quad (3.1)$$

where $s = min(x, y)$, $\lambda_1, \lambda_2 > 0$ and $\phi \in [0, min(\lambda_1, \lambda_2))$. We will denote this distribution as $BP(\lambda_1, \lambda_2, \phi)$. The bivariate Poisson distribution defined in (3.1) allows for dependence between the two random variables. Marginally each random variable follows a Poisson distribution with parameters λ_1 and λ_2 respectively. Parameter ϕ is the covariance between the two random variables. If $\phi = 0$ then the two variables are independent and the bivariate Poisson distribution reduces to the product of two independent Poisson distributions. For a comprehensive treatment of the bivariate Poisson distribution and its multivariate extensions the reader can refer to the books of Kocherlakota and Kocherlakota (1992) and Johnson et al. (1997).

The above assumption leads to the equidispersion case, i.e. $\upsilon_j = 1$, or equivalently assume that R_{jt} are *iid* Poisson sequences with $\sigma_j^2 = \lambda_j$, $j = 1, 2$. Obviously, in this case the covariance function (2.7) remains unaffected while the correlation function (2.8) is simplified due to the simplification of the variances of the two processes. Hence, the Poisson BINAR(1) model is characterized by the vector of expectations $\boldsymbol{\mu}_{\mathbf{X}_t} = E(\mathbf{X}_t)$ with elements

$$\mu_{X_{jt}} = \frac{\lambda_j}{1 - \alpha_j} \ ; \quad j = 1, 2 \quad (3.2)$$

the variance-covariance matrix $\boldsymbol{\gamma}_{\mathbf{X}_t}(h)$ with diagonal elements

$$\text{Cov}(X_{j,t+h}, X_{jt}) = \frac{\alpha_j^h \lambda_j}{1-\alpha_j} \ ; \quad j=1,2, \quad h=0,1,\ldots \tag{3.3}$$

and off-diagonal elements

$$\text{Cov}(X_{j,t+h}, X_{it}) = \frac{\alpha_j^h \phi}{1-\alpha_1\alpha_2} \ ; \quad j \neq i, \quad h=0,1,\ldots \tag{3.4}$$

and the correlation matrix $\boldsymbol{\rho}_{\mathbf{X}_t}(h)$ with diagonal and off-diagonal elements equal to

$$\text{Corr}(X_{j,t+h}, X_{jt}) = \alpha_j^h \ ; \quad j=1,2, \quad h=0,1,\ldots \tag{3.5}$$

and

$$\text{Corr}(X_{j,t+h}, X_{it}) = \frac{\alpha_j^h \sqrt{(1-\alpha_1)(1-\alpha_2)}\phi}{(1-\alpha_1\alpha_2)\sqrt{\lambda_1\lambda_2}} \ ; \quad j \neq i, \quad h=0,1,\ldots \tag{3.6}$$

respectively. Note also that conditionally on the previous observations $\mathbf{X}_{t-1} = \{X_{1,t-1}, X_{2,t-1}\}$, the vector of conditional means $\boldsymbol{\mu}_{\mathbf{X}_{t|t-1}} = E(\mathbf{X}_{t|t-1})$ has elements $\mu_{X_{jt|t-1}} = \alpha_j X_{j,t-1} + \lambda_j$, $j=1,2$. For $h=0$, the conditional variance-covariance matrix $\boldsymbol{\gamma}_{\mathbf{X}_{t|t-1}}(h)$ has diagonal and off-diagonal elements equal to $\text{Cov}(X_{j,t+h}, X_{jt}|X_{j,t-1}) = \alpha_j(1-\alpha_j)X_{j,t-1} + \lambda_j$ and $\text{Cov}(X_{j,t+h}, X_{it}|X_{j,t-1}, X_{i,t-1}) = \phi$ respectively, while otherwise it is the zero matrix.

3.1.2 Estimation

Yule-Walker estimation

For any $h \in \mathbb{Z}$, the autocovariance function of each one of the stationary processes $\{X_{jt}\}_{t\in\mathbb{Z}}$ at lag h is given by the formula $\gamma_j(h) = \alpha_j^{|h|}\gamma_j(0)$, $j=1,2$. The Yule-Walker estimator for α_j can therefore be found from $\alpha_j = \gamma_j(1)/\gamma_j(0)$ where $\gamma_j(1) = \text{Cov}(X_{jt}, X_{j,t-1})$ and $\gamma_j(0) = Var(X_{jt})$. Thus, by replacing $\gamma_j(1)$ and $\gamma_j(0)$ with their sample equivalents, we obtain:

$$\tilde{\alpha}_j = \frac{\sum_{t=2}^T (x_{jt}-\overline{x}_j)(x_{j,t-1}-\overline{x}_j)}{\sum_{t=1}^T (x_{jt}-\overline{x}_j)^2} \ ; \quad j=1,2, \quad t=1,\ldots,T \tag{3.7}$$

Estimation of λ_j is based on the moment condition arising from the corresponding marginal distribution, $E(X_{jt}) = \lambda_j/(1-\alpha_j)$, $j=1,2$ and has the form

$$\tilde{\lambda}_j = (1-\tilde{\alpha}_j)\overline{x}_j \ ; \quad j=1,2, \quad t=1,\ldots,T \tag{3.8}$$

The parameter ϕ can be estimated regarding its involvement in the covariance function $\mathrm{Cov}(X_{1t}, X_{2t}) = \phi/(1-\alpha_1\alpha_2)$:

$$\tilde{\phi} = \frac{(1-\tilde{\alpha}_1\tilde{\alpha}_2)}{T}\sum_{t=1}^{T}(x_{1t}-\overline{x}_1)(x_{2t}-\overline{x}_2)\ ; \quad t=1,\ldots,T \tag{3.9}$$

Note that under the Poisson BINAR(1) model, the parameter α_j can alternatively be estimated as

$$\tilde{\alpha}_j = \frac{\sum_{t=2}^{T}(x_{jt}-\overline{x}_j)(x_{j,t-1}-\overline{x}_j)}{T\overline{x}_j}\ ; \quad j=1,2, \quad t=1,\ldots,T \tag{3.10}$$

The estimators $\tilde{\lambda}_j$ and $\tilde{\phi}$ are also modified due to the involvement of the new estimator (3.10) in the corresponding formulas. We refer to this group of estimators as *moment-based estimators*.

The asymptotic distribution of the Yule-Walker estimators can be found by taking advantage from their asymptotic equivalence to the conditional least squares estimators in case of a Poisson AR(1) process (see Appendix I and Freeland and McCabe (2005)).

PROPOSITION 1. Let $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$ be a Poisson BINAR(1) process. Then, $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$ is an AR(1) process that can be written as

$$\mathbf{X}_t - \boldsymbol{\mu}_{\mathbf{X}_t} = \mathbf{A}(\mathbf{X}_{t-1} - \boldsymbol{\mu}_{\mathbf{X}_t}) + \mathbf{R}_t^{\star}\ ; \quad t\in\mathbb{Z} \tag{3.11}$$

where $\{\mathbf{R}_t^{\star}\}_{t\in\mathbb{Z}}$ has mean vector $E(\mathbf{R}_t^{\star}) = \mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{R}^{\star}} = \mathbf{A}(\mathbf{I}-\mathbf{A})\boldsymbol{\mu}_{\mathbf{X}_t} + \boldsymbol{\Sigma}_{\boldsymbol{R}}$ with $\boldsymbol{\Sigma}_{\boldsymbol{R}}$ denoting the variance-covariance matrix of $\{R_{1t}, R_{2t}\}$.
Equation (3.11) can alternatively be written as

$$\mathbf{X}_t = \mathbf{A}\mathbf{X}_{t-1} + \boldsymbol{\mu}_{\boldsymbol{R}} + \mathbf{R}_t^{\star}\ ; \quad t\in\mathbb{Z} \tag{3.12}$$

where $\boldsymbol{\mu}_{\boldsymbol{R}}$ is the mean vector of $\mathbf{R}$ (Latour, 1997).

Under this notation, it can be seen that the error of the Poisson BINAR(1) model is a martingale difference sequence. Thus, our model can be regarded as member of the general class of multivariate time series models of the form $\mathbf{Y}_t = \mathbf{f}(\mathcal{I}_{t-1};\boldsymbol{\theta}_0) + \mathbf{a}_t$ (Chabot-Hallé and Duchesne, 2008) where $\mathbf{Y} = \{\mathbf{Y}_t, t\in\mathbb{Z}\}$ is a stationary and ergodic multivariate stochastic process, $\mathcal{I}_t$ is the sigma-field generated by $\{Y_t, Y_{t-1}, ...\}$, $\boldsymbol{f} = (f_1, ..., fd)^T$ is a known real-valued function with values in $\mathbb{R}^d$, $\boldsymbol{\theta}_0$ corresponds to a $k\times 1$ vector of unknown parameters and $\mathbf{a} = \{\mathbf{a}_t, t \in \mathbb{Z}\}$ is the error term. For such models, the conditional least squares estimators can be found by minimizing the criterion:

$$\mathbf{S}_n(\boldsymbol{\theta}) = \sum_{t=\tau+1}^{n} \{\mathbf{Y}_{t-1} - \mathbf{f}(\mathcal{I}_{t-1};\boldsymbol{\theta})\}^T \boldsymbol{\Sigma}_{\mathbf{a}}^{-1} \{\mathbf{Y}_{t-1} - \mathbf{f}(\mathcal{I}_{t-1};\boldsymbol{\theta})\} \tag{3.13}$$

where τ is an appropriate integer.
Under certain regularity conditions it can be shown that $\tilde{\boldsymbol{\theta}}_n$ is asymptotically normal:

$$\tilde{\boldsymbol{\theta}}_n \sim AN\left(\boldsymbol{\theta}_0, \frac{\mathbf{U}^{-1}\mathbf{W}\mathbf{U}^{-1}}{n}\right) \tag{3.14}$$

where,

$$\mathbf{U} = E\left\{\frac{\partial \mathbf{f}_{t-1}^T}{\partial \boldsymbol{\theta}} \boldsymbol{\Sigma}_{\mathbf{a}}^{-1} \frac{\partial \mathbf{f}_{t-1}}{\partial \boldsymbol{\theta}^T}\right\} \tag{3.15}$$

$$\mathbf{W} = E\left\{\frac{\partial \mathbf{f}_{t-1}^T}{\partial \boldsymbol{\theta}} \boldsymbol{\Sigma}_{\mathbf{a}}^{-1} (\mathbf{Y}_t - \mathbf{f}_{t-1})(\mathbf{Y}_t - \mathbf{f}_{t-1})^T \boldsymbol{\Sigma}_{\mathbf{a}}^{-1} \frac{\partial \mathbf{f}_{t-1}}{\partial \boldsymbol{\theta}^T}\right\} \tag{3.16}$$

with $\mathbf{f}_{t-1} = \mathbf{f}(\mathcal{I}_{t-1};\boldsymbol{\theta}_0)$.

For the Poisson BINAR(1) model, let $\boldsymbol{\theta}_0 = (\alpha_1, \alpha_2, \lambda_1, \lambda_2, \phi)^T$ be the vector of unknown parameters and let $\tilde{\boldsymbol{\theta}}_0 = (\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\phi})^T$ be the vector of the respective Yule-Walker estimators. Then, their asymptotic distribution can be found by equations (3.14)-(3.16) by replacing $\mathbf{Y}_t$ by $\mathbf{X}_t$, $\mathbf{f}_{t-1}$ by $\mathbf{A}\mathbf{X}_{t-1} + \boldsymbol{\mu}_{\mathbf{R}}$ and $\boldsymbol{\Sigma}_{\mathbf{a}}$ by $\boldsymbol{\Sigma}_{\mathbf{R}^\star}$.

Maximum-Likelihood estimation

The conditional density for the Poisson BINAR(1) model can be obtained by substituting

$$f_3(k,s) = \frac{e^{-(\lambda_1+\lambda_2-\phi)} \sum_{m=0}^{min(k,s)} (\lambda_1-\phi)^{k-m}(\lambda_2-\phi)^{s-m}\phi^m}{(k-m)!(s-m)!m!} \tag{3.17}$$

in eq.(2.15). Then we get

$$f(\mathbf{x}_t|\mathbf{x}_{t-1}, \alpha_1, \alpha_2, \lambda_1, \lambda_2, \phi) = e^{-(\lambda_1+\lambda_2-\phi)} \sum_{k=0}^{g_1}\sum_{s=0}^{g_2} \frac{\sum_{m=0}^{min(k,s)} (\lambda_1-\phi)^{k-m}(\lambda_2-\phi)^{s-m}\phi^m}{(k-m)!(s-m)!m!}$$
$$\times \binom{x_{1,t-1}}{x_{1t}-k} \alpha_1^{x_{1t}-k}(1-\alpha_1)^{x_{1,t-1}-x_{1t}+k} \binom{x_{2,t-1}}{x_{2t}-s} \alpha_2^{x_{2t}-s}(1-\alpha_2)^{x_{2,t-1}-x_{2t}+s} \tag{3.18}$$

where $g_1 = min(x_{1t}, x_{1,t-1})$ and $g_2 = min(x_{2t}, x_{2,t-1})$.

Remark: An alternative way to obtain the conditional density of the Poisson BINAR(1) model is by considering the construction of the two innovation processes $\{R_{1t}, R_{2t}\}$ by means of the trivariate reduction technique (see Kocherlakota and Kocherlakota, 1992). Under this approach, we let W_1, W_2, W_3 be independent Poisson random variables with parameters $\lambda_1^\star$, $\lambda_2^\star$ and ϕ respectively. We also let $R_1 = W_1 + W_3$ and $R_2 = W_2 + W_3$. It can be shown that the joint distribution of $\{R_1, R_2\}$ is bivariate Poisson with parameters $\lambda_1 = \lambda_1^\star + \phi$, $\lambda_2 = \lambda_2^\star + \phi$ and ϕ. Then, the two series of counts can be written as $X_{1t} = Y_{1t} + W_{3t}$ and $X_{2t} = Y_{2t} + W_{3t}$ respectively, where $Y_{1t} = \alpha_1 \circ X_{1,t-1} + W_{1t}$, $Y_{2t} = \alpha_2 \circ X_{2,t-1} + W_{2t}$ and W_{3t} are all independently distributed. Hence, the joint conditional probability distribution of $\{X_{1t}, X_{2t}\}$ is $P(X_{1t} = r, X_{2t} = s) = \sum_{k=0}^{\min(r,s)} f_1(r-k) f_2(s-k) f_3(k)$ where $f_1(r-k) = P(r-k|X_{1,t-1}, \alpha_1, \lambda_1^\star)$ is the convolution of a Bin$(X_{1,t-1}, \alpha_1)$ and a Poisson$(\lambda_1^\star)$ distribution, $f_2(s-k) = P(s-k|X_{2,t-1}, \alpha_2, \lambda_2^\star)$ is the convolution of a Bin$(X_{2,t-1}, \alpha_2)$ and a Poisson$(\lambda_2^\star)$ distribution and $f_3(k)$ is Poisson with parameter ϕ.

3.2 A BINAR(1) Process with BVNB Innovations

3.2.1 Model

Assume that the *jpmf* of the innovations $\{R_{1t}, R_{2t}\}$ is a bivariate negative binomial distribution of the following form (Marshall and Olkin, 1990; Boucher et al., 2008; Cheon et al., 2009):

$$\begin{aligned} P(R_{1t} = x, R_{2t} = y) &= \frac{\Gamma(\beta^{-1} + x + y)}{\Gamma(\beta^{-1})\Gamma(x+1)\Gamma(y+1)} \\ &\times \left(\frac{\lambda_1}{\lambda_1 + \lambda_2 + \beta^{-1}}\right)^x \left(\frac{\lambda_2}{\lambda_1 + \lambda_2 + \beta^{-1}}\right)^y \left(\frac{\beta^{-1}}{\lambda_1 + \lambda_2 + \beta^{-1}}\right)^{\beta^{-1}} \end{aligned} \tag{3.19}$$

where $\lambda_1, \lambda_2, \beta > 0$. We will denote this distribution as BVNB$(\lambda_1, \lambda_2, \beta)$. Note that the marginal distribution of R_{jt} is univariate negative binomial with parameters β^{-1} and $p_j = \beta^{-1}/(\lambda_j + \beta^{-1})$, $j = 1, 2$ and that the correlation between the two count variables R_{1t} and R_{2t}

$$\text{Corr}(x, y) = \sqrt{\frac{\lambda_1 \lambda_2 \beta^2}{(1 + \lambda_1 \beta)(1 + \lambda_2 \beta)}} \tag{3.20}$$

must be positive. This assumption allows for more flexibility than the Poisson BINAR(1) model does, due to the involvement of the overdispersion parameter β in the model's specification.

Recall that in section 2.1, $\{R_{jt}\}$ was generally defined as an *iid* sequence with mean λ_j and variance $\sigma_j^2 = \upsilon_j \lambda_j$, $\upsilon_j > 0$, $j = 1, 2$. For the BVNB model, $\sigma_j^2 = \lambda_j(1 + \beta\lambda_j)$ implying that $\upsilon_j = 1 + \beta\lambda_j, \quad \lambda_j, \beta > 0$. Consequently $\upsilon_j > 1$ which indicates the overdispersion case. However, the resulting model is not a BINAR model with negative binomial marginals but a model that effectively accounts for overdispersion. In specific, the statistical properties of the BINAR(1) model with BVNB innovations are encompassed in the vector of expectations $\boldsymbol{\mu}_{\mathbf{X}_t} = E(\mathbf{X}_t)$ with elements

$$\mu_{X_{jt}} = \frac{\lambda_j}{1 - \alpha_j} \; ; \quad j = 1, 2 \tag{3.21}$$

the variance-covariance matrix $\boldsymbol{\gamma}_{\mathbf{X}_t}(h)$ with diagonal and off-diagonal elements equal to

$$\mathrm{Cov}(X_{j,t+h}, X_{jt}) = \frac{\alpha_j^h \lambda_j (1 + \beta\lambda_j + \alpha_j)}{1 - \alpha_j^2} \; ; \quad j = 1, 2, \quad h = 0, 1, \ldots \tag{3.22}$$

and

$$\mathrm{Cov}(X_{j,t+h}, X_{it}) = \frac{\alpha_j^h \beta \lambda_1 \lambda_2}{1 - \alpha_1 \alpha_2} \; ; \quad j \neq i, \quad h = 0, 1, \ldots \tag{3.23}$$

respectively, and the correlation matrix $\boldsymbol{\rho}_{\mathbf{X}_t}(h)$ with diagonal elements

$$\mathrm{Corr}(X_{j,t+h}, X_{jt}) = \alpha_j^h \; ; \quad j = 1, 2, \quad h = 0, 1, \ldots \tag{3.24}$$

and off-diagonal elements

$$\mathrm{Corr}(X_{j,t+h}, X_{it}) = \frac{\alpha_j^h \beta}{(1 - \alpha_1 \alpha_2)} \sqrt{\frac{(1 - \alpha_1^2)(1 - \alpha_2^2)\lambda_1 \lambda_2}{(1 + \beta\lambda_1 + \alpha_1)(1 + \beta\lambda_2 + \alpha_2)}} \; ; \quad j \neq i, \quad h = 0, 1, \ldots \tag{3.25}$$

Conditionally on the previous observations $\mathbf{X}_{t-1} = \{X_{1,t-1}, X_{2,t-1}\}$, the vector of conditional means $\boldsymbol{\mu}_{\mathbf{X}_{t|t-1}} = E(\mathbf{X}_{t|t-1})$ has elements $\mu_{X_{jt|t-1}} = \alpha_j X_{j,t-1} + \lambda_j$, $j = 1, 2$. For $h = 0$, the conditional variance-covariance matrix $\boldsymbol{\gamma}_{\mathbf{X}_{t|t-1}}(h)$ has diagonal and off-diagonal elements equal to $\mathrm{Cov}(X_{j,t+h}, X_{jt} | X_{j,t-1}) = \alpha_j(1 - \alpha_j)X_{j,t-1} + \lambda_j(1 + \beta\lambda_j)$ and $\mathrm{Cov}(X_{j,t+h}, X_{it} | X_{j,t-1}, X_{i,t-1}) = \beta\lambda_1\lambda_2$ respectively, while otherwise it is the zero matrix.

3.2.2 Estimation

Yule-Walker estimation

Using the moments $\mu_{X_{jt}}$, $\mathrm{Corr}(X_{jt}, X_{j,t-1})$ and $\mathrm{Cov}(X_{1t}, X_{2t})$, as defined in the previous section, and equating them to their sample equivalents, we get

$$\tilde{\alpha}_j = \frac{\sum_{t=2}^{T}(x_{jt} - \bar{x}_j)(x_{j,t-1} - \bar{x}_j)}{\sum_{t=2}^{T}(x_{jt} - \bar{x}_j)^2} \tag{3.26}$$

$$\tilde{\lambda}_j = (1 - \hat{\alpha}_j)\bar{x}_j \tag{3.27}$$

$$\tilde{\beta} = \frac{1 - \tilde{\alpha}_1\tilde{\alpha}_2}{\tilde{\lambda}_1\tilde{\lambda}_2}\left\{\frac{1}{T}\sum_{t=1}^{T}(x_{1t} - \bar{x}_1)(x_{2t} - \bar{x}_2)\right\} \tag{3.28}$$

Note that for a negative observed covariance there is a danger to obtain non-admissible estimates. This deficit concerns both the Poisson BINAR(1) model and the BINAR(1) model with BVNB innovations.

Maximum-Likelihood Estimation

For the BINAR(1) model with BVNB innovations it holds that

$$f_3(k, s) = \frac{\Gamma(\beta^{-1} + k + s)}{\Gamma(\beta^{-1})k!s!}\left(\frac{\lambda_1}{\lambda_1 + \lambda_2 + \beta^{-1}}\right)^k\left(\frac{\lambda_2}{\lambda_1 + \lambda_2 + \beta^{-1}}\right)^s\left(\frac{\beta^{-1}}{\lambda_1 + \lambda_2 + \beta^{-1}}\right)^{\beta^{-1}} \tag{3.29}$$

Thus the conditional density (2.15) becomes

$$\begin{aligned} f(\mathbf{x}_t|\mathbf{x}_{t-1}, \alpha_1, \alpha_2, \lambda_1, \lambda_2, \beta) = \\ \sum_{k=0}^{g_1}\sum_{s=0}^{g_2}\frac{\Gamma(\beta^{-1} + k + s)}{\Gamma(\beta^{-1})k!s!}\left(\frac{\lambda_1}{\lambda_1 + \lambda_2 + \beta^{-1}}\right)^k\left(\frac{\lambda_2}{\lambda_1 + \lambda_2 + \beta^{-1}}\right)^s\left(\frac{\beta^{-1}}{\lambda_1 + \lambda_2 + \beta^{-1}}\right)^{\beta^{-1}} \\ \times \binom{x_{1,t-1}}{x_{1t} - k}\alpha_1^{x_{1t}-k}(1 - \alpha_1)^{x_{1,t-1}-x_{1t}+k}\binom{x_{2,t-1}}{x_{2t} - s}\alpha_2^{x_{2t}-s}(1 - \alpha_2)^{x_{2,t-1}-x_{2t}+s} \end{aligned} \tag{3.30}$$

where $g_1 = min(x_{1t}, x_{1,t-1})$ and $g_2 = min(x_{2t}, x_{2,t-1})$.

3.3 Other distributional choices

As mentioned before, the choice of the joint distribution for R_{1t} and R_{2t} determines the properties of the underlying process. While the bivariate negative binomial provides overdisperison, it is interesting to note that a selection of a distribution with negative correlation can also produce negative correlation between the two series (see 2.7). The literature on bivariate count distributions with negative correlation is limited. One of the reasons is that negative correlation in bivariate counts occurs rather infrequently. However there are such models in the literature, as for example the bivariate Poisson-lognormal model of Aitchinson and Ho (1989) (see also Chib and Winkelmann, 2001), the finite mixture model developed in Karlis and Meligkotsidou (2007) and models based on copulas see, e.g. Nikoloulopoulos and Karlis (2009) and the references therein. Finally, note that while we used a certain bivariate negative binomial distribution, there are certain other alternatives in the literature which could have been used. We have selected this one mainly because of its relative simplicity.

4 Diagnostics

In this section we describe diagnostics for assessing the goodness of fit. Usually, in model fitting, this is accomplished by means of residual analysis. However, due to the structural distinctiveness of INAR-type models, the classical definition of residuals as differences between the observed and fitted values, may prove to be inadequate as a diagnostic tool. We follow Freeland and McCabe (2004a) by introducing a definition for residuals for count data that distinguishes between a set of residuals for the survival process $r_{1t} = \alpha \circ X_{t-1} - \alpha X_{t-1}$ and another for the arrival component $r_{2t} = R_t - \lambda$. In this section we attempt to extend the ideas of Freeland and McCabe (2004a) to the BINAR(1) model.

For each one of the two series $\{X_{1t}, X_{2t}\}$, we define two sets of residuals; one for each random component. So, for the survival components we let $r_{1t}^{(j)} = \alpha_j \circ X_{j,t-1} - \alpha_j X_{j,t-1}$ and for the arrival components we let $r_{2t}^{(j)} = R_{jt} - \lambda_j$, $j = 1, 2$. In order to arrive at a sensible and practical form of the above definitions, the unobservable quantities $\alpha_j \circ X_{j,t-1}$ and R_{jt} should be replaced with $E_t[\alpha_j \circ X_{j,t-1}]$ and $E_t[R_{jt}]$ respectively, i.e. with their conditional expectations given the observed values of X_{jt} and $X_{j,t-1}$.

PROPOSITION 2. Let $E_t[\cdot]$ denote the conditional expectation to the sigma field, $\Im_t = \sigma(X_{j0}, X_{j1}, ..., X_{jt})$, $j = 1, 2$. For the BINAR(1) model with bivariately distributed innovations the following equalities hold:

$$E_t[\alpha_1 \circ X_{1,t-1}] = \frac{\alpha_1 x_{1,t-1} P(x_{1t}-1|X_{1,t-1}-1, X_{2,t-1})}{P(x_{1t}|X_{1,t-1}, X_{2,t-1})} \tag{4.1}$$

$$E_t[\alpha_2 \circ X_{2,t-1}] = \frac{\alpha_2 x_{2,t-1} P(x_{2t}-1|X_{1,t-1}, X_{2,t-1}-1)}{P(x_{2t}|X_{1,t-1}, X_{2,t-1})} \tag{4.2}$$

$$E_t[R_{1t}] = \frac{\sum_{x_{2t}} \sum_{k=0}^{g_1} \sum_{s=0}^{g_2} k f_1(x_{1t}-k) f_2(x_{2t}-s) f_3(k,s)}{P(x_{1t}|X_{1,t-1}, X_{2,t-1})} \tag{4.3}$$

$$E_t[R_{2t}] = \frac{\sum_{x_{1t}} \sum_{k=0}^{g_1} \sum_{s=0}^{g_2} s f_1(x_{1t}-k) f_2(x_{2t}-s) f_3(k,s)}{P(x_{2t}|X_{1,t-1}, X_{2,t-1})} \tag{4.4}$$

where the densities $f_1(\cdot)$ and $f_2(\cdot)$ are given in (2.13) and (2.14), $f_3(k,s) = P(R_{1t}=k, R_{2t}=s)$, $g_1 = min(x_{1t}, x_{1,t-1})$ and $g_2 = min(x_{2t}, x_{2,t-1})$. Using Proposition 2 we can now define the residuals as

$$r_{1t}^{(j)\star} = E_t[r_{1t}^{(j)}] = E_t[\alpha_j \circ X_{j,t-1}] - \alpha_j X_{j,t-1}, \quad \text{and} \tag{4.5}$$

$$r_{2t}^{(j)\star} = E_t[r_{2t}^{(j)}] = E_t[R_{jt}] - \lambda_j, \quad \text{for } \ j=1,2 \tag{4.6}$$

Regarding separately each one of the two series that comprise the BINAR(1) model, it is noted that adding the components of the two new sets of residuals gives the usual definition of residuals, i.e.

$$\begin{aligned} r_{1t}^{(j)\star} + r_{2t}^{(j)\star} &= E_t[\alpha_j \circ X_{j,t-1}] - \alpha_j X_{j,t-1} + E_t[R_{jt}] - \lambda_j \\ &= E_t[\alpha_j \circ X_{j,t-1} + R_{jt}] - \alpha_j X_{j,t-1} - \lambda_j \\ &= X_{jt} - \alpha_j X_{j,t-1} - \lambda_j = r_t^{(j)}. \end{aligned} \tag{4.7}$$

Thus, the adequacy of each component of the model may by assessed by plotting the aformentioned sets of residuals.

5 Forecasting

The usual way to produce forecasts in time series models is via the conditional forecast distribution. Freeland and McCabe (2004b) established the h-step-ahead conditional distribution of the Poisson INAR(1) model, based on the remark of Al-Osh and Alzaid (1987) that

$$(X_t, X_{t-h}) \overset{d}{=} \left(\alpha^h \circ X_{t-h} + \sum_{i=0}^{h-1} \alpha^i \circ R_{t-i}, X_{t-h} \right) \tag{5.1}$$

where R_t is a sequence of uncorrelated non-negative integer-valued random variables with finite mean and variance.

The above result holds also for the marginal distribution of each one of the two series (X_{1t}, X_{2t}) that consist a BINAR(1) model. As in the univariate case, $\alpha_j^h \circ X_{j,t-h} \mid X_{j,t-h},\ j = 1, 2$, has a binomial distribution with parameters $(\alpha_j^h, X_{j,t-h})$. Moreover, the joint and marginal distributions of $\sum_{i=0}^{h-1} \alpha_1^i \circ R_{1,t-i}$ and $\sum_{i=0}^{h-1} \alpha_2^i \circ R_{2,t-i}$ are determined by the joint and marginal distributions of X_{1t} and X_{2t}. This relation can be described in terms of the *jpgf* of $\left\{\sum\limits_{i=0}^{h-1} \alpha_1^i \circ R_{1,t-i}, \sum\limits_{i=0}^{h-1} \alpha_2^i \circ R_{2,t-i}\right\}$. Denote by S_j the quantity $\sum\limits_{i=0}^{h-1} \alpha_j^i \circ R_{1,t-j},\ j = 1, 2$. Then,

$$G_{S_1,S_2}(s_1, s_2) = \prod_{i=0}^{h-1} G_{R_1,R_2}((1 - \alpha_1^i + \alpha_1^i s_1), (1 - \alpha_2^i + \alpha_2^i s_2)) \tag{5.2}$$

Hence, the joint distribution of $\{X_{1t}, X_{2t}\}$ given $\{X_{1,t-h}, X_{2,t-h}\}$ is a convolution of two binomial distributions with parameters $(\alpha_1^h, X_{1,t-h})$ and $(\alpha_2^h, X_{2,t-h})$ respectively, and a bivariate distribution with *jpgf* of the form (5.2). Obviously, if (5.2) has not a closed-form expression, then neither the h-step-ahead forecast distribution can be specified in closed-form. However, it is straightforward to evaluate it numerically.

For the Poisson BINAR(1) model, it can be proved that

$$\begin{aligned} G_{S_1,S_2}(s_1, s_2) &= \exp\left[\left(\frac{1-\alpha_1^h}{1-\alpha_1}\right)\lambda_1(s_1 - 1) + \left(\frac{1-\alpha_2^h}{1-\alpha_2}\right)\lambda_2(s_2 - 1)\right. \\ &+ \left.\left(\frac{1-\alpha_1^h\alpha_2^h}{1-\alpha_1\alpha_2}\right)\phi(s_1 - 1)(s_2 - 1)\right] \end{aligned} \tag{5.3}$$

while the corresponding *jpgf* for the BINAR(1) model with BVNB innovations is given by

$$G_{S_1,S_2}(s_1, s_2) = \prod_{i=0}^{h-1} \left[1 - \beta\lambda_1\alpha_1^i(s_1 - 1) - \beta\lambda_2\alpha_2^i(s_2 - 1)\right]^{-\beta^{-1}} \tag{5.4}$$

which is not of a convenient form.

Note however that irrespective of the *jpgf* of $\left\{\sum\limits_{i=0}^{h-1} \alpha_1^i \circ R_{1,t-i}, \sum\limits_{i=0}^{h-1} \alpha_2^i \circ R_{2,t-i}\right\}$, closed-form expressions are available for their conditional expectations and variances (conditional on X_{1t}, X_{2t}). More specifically, it can be proved that

$$E\left(\sum_{i=0}^{h-1} \alpha_j^i \circ R_{j,t-i}\right) = \left(\frac{1-\alpha_j^h}{1-\alpha_j}\right)\lambda_j \tag{5.5}$$

and

$$Var\left(\sum_{i=0}^{h-1} \alpha_j^i \circ R_{j,t-i}\right) = \left(\frac{1-\alpha_j^{2h}}{1-\alpha_j^2}\right)\upsilon_j\lambda_j + \left(\frac{1-\alpha_j^h}{1-\alpha_j} - \frac{1-\alpha_j^{2h}}{1-\alpha_j^2}\right)\lambda_j \tag{5.6}$$

Theorem 1 summarizes the h-step-ahead conditional *jpmf* and *jpgf* and the corresponding conditional means and variances for the BINAR(1) model. Specific expressions for the Poisson BINAR(1) and the BINAR(1) model with BVNB innovations follow in corollaries 1 and 2 respectively.

THEOREM 1. The *jpmf* of $\{X_{1,T+h}, X_{2,T+h}\}$ given $\{x_{1T}, x_{2T}\}$ is given by

$$\begin{aligned}
& P_h(X_{1,T+h} = x_1, X_{2,T+h} = x_2 | x_{1T}, x_{2T}) = \\
& \sum_{k=0}^{min(x_1,x_{1T})} \sum_{s=0}^{min(x_2,x_{2T})} \binom{x_{1T}}{x_1-k} (\alpha_1^h)^{x_1-k}(1-\alpha_1^h)^{x_{1T}-x_1+k} \\
\times\ & \binom{x_{2T}}{x_2-s} (\alpha_2^h)^{x_2-s}(1-\alpha_2^h)^{x_{2T}-x_2+s} \\
\times\ & f\left(\sum_{i=0}^{h-1} \alpha_1^i \circ R_{1,T+h-i} = k, \sum_{i=0}^{h-1} \alpha_2^i \circ R_{2,T+h-i} = s | x_{1T}, x_{2T}\right)
\end{aligned}$$

with means,

$$E(x_{j,T+h}|x_{2T}, x_{2T}) = \alpha_j^h x_{jT} + \left(\frac{1-\alpha_j^h}{1-\alpha_j}\right)E(R_{jt})\ ; \quad j = 1,2, \quad h = 1,2,\ldots \tag{5.7}$$

and variances,

$$\begin{aligned}
Var(x_{j,T+h}|x_{1T}, x_{2T}) &= \alpha_j^h(1-\alpha_j^h)x_{jT} + \left(\frac{1-\alpha_j^{2h}}{1-\alpha_j^2}\right)Var(R_{jt}) \\
&+ \left(\frac{1-\alpha_j^h}{1-\alpha_j} - \frac{1-\alpha_j^{2h}}{1-\alpha_j^2}\right)E(R_{jt})\ ; \quad j = 1,2, \quad h = 1,2,\ldots
\end{aligned} \tag{5.8}$$

The corresponding *jpgf* of $\{X_{1,T+h}, X_{2,T+h}\}$ given $\{x_{1T}, x_{2T}\}$ is of the form

$$G_{X_{1,T+h},X_{2,T+h}}(s_1, s_2|x_{1T}, x_{2T}) = (1-\alpha_1^h+\alpha_1^h s_1)^{X_{1T}}(1-\alpha_2^h+\alpha_2^h s_2)^{X_{2T}} G_{S_1,S_2}(s_1, s_2) \tag{5.9}$$

where $G_{S_1,S_2}(s_1, s_2)$ is given in (5.2).

Corollary 1. For the Poisson BINAR(1) model, the *jpgf* and *jpmf* of $\{X_{1,T+h}, X_{2,T+h}\}$ given $\{x_{1T}, x_{2T}\}$ are given by

$$\begin{aligned} &G_{X_{1,T+h},X_{2,T+h}}(s_1, s_2|x_{1T}, x_{2T}) = (1-\alpha_1^h+\alpha_1^h s_1)^{X_{1T}}(1-\alpha_2^h+\alpha_2^h s_2)^{X_{2T}} \\ \times\ &\exp\left\{\left(\frac{1-\alpha_1^h}{1-\alpha_1}\right)\lambda_1 s_1 + \left(\frac{1-\alpha_2^h}{1-\alpha_2}\right)\lambda_2 s_2 + \left(\frac{1-\alpha_1^h\alpha_2^h}{1-\alpha_1\alpha_2}\right)\phi(s_1-1)(s_2-1)\right\} \end{aligned}$$

and

$$\begin{aligned} &P_h(X_{1,T+h} = x_1, X_{2,T+h} = x_2|x_{1T}, x_{2T}) = \\ &\sum_{k=0}^{min(x_1,x_{1T})}\ \sum_{s=0}^{min(x_2,x_{2T})} \binom{x_{1T}}{x_1-k} (\alpha_1^h)^{x_1-k}(1-\alpha_1^h)^{x_{1T}-x_1+k} \\ \times\ &\binom{x_{2T}}{x_2-s} (\alpha_2^h)^{x_2-s}(1-\alpha_2^h)^{x_{2T}-x_2+s} \\ \times\ &\exp\left\{-\left[\left(\frac{1-\alpha_1^h}{1-\alpha_1}\right)\lambda_1 + \left(\frac{1-\alpha_2^h}{1-\alpha_2}\right)\lambda_2 - \left(\frac{1-\alpha_1^h\alpha_2^h}{1-\alpha_1\alpha_2}\right)\phi\right]\right\} \\ \times\ &\sum_{m=0}^{min(k,s)} \frac{\left[\left(\frac{1-\alpha_1^h}{1-\alpha_1}\right)\lambda_1 - \left(\frac{1-\alpha_1^h\alpha_2^h}{1-\alpha_1\alpha_2}\right)\phi\right]^{k-m}\left[\left(\frac{1-\alpha_2^h}{1-\alpha_2}\right)\lambda_2 - \left(\frac{1-\alpha_1^h\alpha_2^h}{1-\alpha_1\alpha_2}\right)\phi\right]^{s-m}\left[\left(\frac{1-\alpha_1^h\alpha_2^h}{1-\alpha_1\alpha_2}\right)\phi\right]^{m}}{(k-m)!(s-m)!m!} \end{aligned}$$

respectively, with means,

$$E(x_{j,T+h}|x_{1T}, x_{2T}) = \alpha_j^h x_{jT} + \frac{1-\alpha_j^h}{1-\alpha_j}\lambda_j\ ; \quad j = 1, 2, \quad h = 1, 2, \ldots \tag{5.10}$$

variances,

$$Var(x_{j,T+h}|x_{1T}, x_{2T}) = \alpha_j^h(1-\alpha_j^h)x_{jT} + \frac{1-\alpha_j^h}{1-\alpha_j}\lambda_j\ ; \quad j = 1, 2, \quad h = 1, 2, \ldots \tag{5.11}$$

and covariance,

$$\mathrm{Cov}(x_{1,T+h}, x_{2,T+h}|x_{1T}, x_{2T}) = \left(\frac{1-\alpha_1^h\alpha_2^h}{1-\alpha_1\alpha_2}\right)\phi\ ; \quad h = 1, 2, \ldots \tag{5.12}$$

Corollary 2. For the BINAR(1) model with BVNB innovations, the *jpgf* and the *jpmf* of $\{X_{1,T+h}, X_{2,T+h}\}$ given $\{x_{1T}, x_{2T}\}$ are given by

$$G_{X_{1,T+h},X_{2,T+h}}(s_1, s_2|x_{1T}, x_{2T}) = (1-\alpha_1^h+\alpha_1^h s_1)^{X_{1T}}(1-\alpha_2^h+\alpha_2^h s_2)^{X_{2T}}$$
$$\times \prod_{i=0}^{h-1}\left[1-\beta\lambda_1\alpha_1^i(s_1-1)-\beta\lambda_2\alpha_2^i(s_2-1)\right]^{-\beta^{-1}}$$

and

$$P_h(X_{1,T+h} = x_1, X_{2,T+h} = x_2|x_{1T}, x_{2T}) =$$
$$\sum_{k=0}^{min(x_1,x_{1T})}\sum_{s=0}^{min(x_2,x_{2T})}\begin{pmatrix} x_{1T} \\ x_1-k \end{pmatrix}(\alpha_1^h)^{x_1-k}(1-\alpha_1^h)^{x_{1T}-x_1+k}$$
$$\times \begin{pmatrix} x_{2T} \\ x_2-s \end{pmatrix}(\alpha_2^h)^{x_2-s}(1-\alpha_2^h)^{x_{2T}-x_2+s}$$
$$\times\ f\left(\sum_{i=0}^{h-1}\alpha_1^i\circ R_{1,T+h-i} = k, \sum_{i=0}^{h-1}\alpha_2^i\circ R_{2,T+h-i} = s|x_{1T}, x_{2T}\right)$$

respectively, where $f\left(\sum_{i=0}^{h-1}\alpha_1^i\circ R_{1,T+h-i} = k, \sum_{i=0}^{h-1}\alpha_2^i\circ R_{2,T+h-i} = s|x_{1T}, x_{2T}\right)$ can be numerically calculated.
The means and variances of this process are given by,

$$E(x_{j,T+h}|x_{1T}, x_{2T}) = \alpha_j^h x_{jT} + \frac{1-\alpha_j^h}{1-\alpha_j}\lambda_j\ ; \quad j = 1, 2, \quad h = 1, 2, \ldots \tag{5.13}$$

and

$$\begin{aligned} Var(x_{j,T+h}|x_{1T}, x_{2T}) &= \alpha_j^h(1-\alpha_j^h)x_{jT} + \left(\frac{1-\alpha_j^{2h}}{1-\alpha_j^2}\right)(1+\beta\lambda_j)\lambda_j \\ &+ \left(\frac{1-\alpha_j^h}{1-\alpha_j} - \frac{1-\alpha_j^{2h}}{1-\alpha_j^2}\right)\lambda_j\ ; \quad j = 1, 2, \quad h = 1, 2, \ldots \end{aligned} \tag{5.14}$$

whereas the covariance function is not of a closed-form.

The marginal probabilities $P_h(x_1|x_{1T}, x_{2T})$ and $P_h(x_2|x_{1T}, x_{2T})$ can be calculated directly as, $P_h(x_1|x_{1T}, x_{2T}) = \sum_{x_2} P_h(x_1, x_2|x_{1T}, x_{2T})$ and $P_h(x_2|x_{1T}, x_{2T}) = \sum_{x_1} P_h(x_1, x_2|x_{1T}, x_{2T})$ respectively.

Given the fact that the vector of parameters $\boldsymbol{\theta}$ is unknown, in practice we are only able to compute $P_h(x_1, x_2|x_{1T}, x_{2T}; \hat{\boldsymbol{\theta}})$ where $\hat{\boldsymbol{\theta}}$ are typically the maximum likelihood estimators introduced in section 2.2. Lack of knowledge about the true values of the model parameters and the need to estimate them introduce uncertainty in the estimation of the h-step-ahead *jpmf*'s. Estimation uncertainty, i.e. the error made in estimating these probabilities, can be assessed by taking advantage of the asymptotic normality of ML estimators. Under standard regularity conditions, the ML estimator $\boldsymbol{\theta}$, denoted by $\hat{\boldsymbol{\theta}}$, is asymptotically normally distributed around the true parameter value, i.e. $\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \overset{a}{\sim} N(0, \boldsymbol{i}^{-1})$, where $\boldsymbol{i}^{-1}$ is the inverse of the Fisher information matrix (Bu and McCabe, 2008). The δ-method can then be used for finding the asymptotic distribution of a random variable $g(\hat{\boldsymbol{\theta}})$. An application of the δ-method to $g(\hat{\boldsymbol{\theta}}) = P_h(\mathbf{x}|\mathbf{x}_T; \hat{\boldsymbol{\theta}})$ provides us with a confidence interval for the probability associated with any fixed value of $\mathbf{x} = (x_1, x_2)$ in the forecast distribution. Obviously, these intervals may be truncated outside [0,1] (Freeland and McCabe, 2004b).

THEOREM 2 (Freeland and McCabe, 2004b)**:** The quantity $P_h(\mathbf{x}|\mathbf{x}_T; \hat{\boldsymbol{\theta}})$ has an asymptotically normal distribution with mean $P_h(\mathbf{x}|\mathbf{x}_T; \boldsymbol{\theta}_\mathbf{0})$ and variance

$$\sigma_h^2(\mathbf{x}; \boldsymbol{\theta}_0) = T^{-1} \left\{ \left(\left. \frac{\partial P_h}{\partial \hat{\boldsymbol{\theta}}'} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right) \boldsymbol{i}^{-1} \left(\left. \frac{\partial P_h}{\partial \hat{\boldsymbol{\theta}}} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right) \right\} \tag{5.15}$$

It is apparent that analytical expressions for (5.15) are only available in cases where $P_h(\mathbf{x}|\mathbf{x}_T)$ has a closed-form expression as is the case for the Poisson BINAR(1) model (see Appendix I).

Closing this section, it is worth noting that summarizing the forecast distribution by means of conditional expectations, while ensures a minimum mean square error, it has drawbacks with respect to data coherency since the integer-valued property of the time series is not taken into account. Freeland and McCabe (2004b) suggest instead the use of the median of this distribution which always lies in the support of the series and is therefore coherent. Also, Pavlopoulos and Karlis (2008) propose a parametric bootstrap approach which guarantees both integer-valued predictions and prediction intervals with integer-valued ends. In our case, since the distributions are discrete, it is relatively easy to find the median to use as prediction instead of the mean, since the median will satisfy the discrete nature of the data.

6 Simulations

6.1 Simulation Design

The small sample properties of the Moment-based (MoM), Yule-Walker (YW) and Maximum Likelihood (ML) estimators under the Poisson BINAR(1) model, were assessed by conducting a series of simulation experiments. Count series were generated assuming that the two innovation processes $\{R_{1t}, R_{2t}\}$ follow jointly a bivariate Poisson distribution. For the various estimators outlined in section 3.1, we derived the biases and computed the ratios of MoM to ML and YW to ML standard deviations under different scenarios for the model parameters α_j; $j = 1, 2$ and $\lambda_1^\star$, $\lambda_2^\star$, ϕ (where $\lambda_1^\star + \phi = \lambda_1$ and $\lambda_2^\star + \phi = \lambda_2$). Simulations were carried out using `R`.

The sample sizes n considered were 50, 200, 500 and 1000. The parameters α_1, α_2 designating the marginal dependence structures of the series X_{1t} and X_{2t} respectively, were allowed to alternate between 0.3 and 0.5. The parameters $\lambda_1^\star$, $\lambda_2^\star$, ϕ which govern the correlation between the innovations R_{1t}, R_{2t} of the two processes at time t, were chosen as follows. The parameter ϕ was fixed at 0.5 or 1 while $\lambda_1^\star$ and $\lambda_2^\star$ were allowed to be either 1 or 3. Some combinations were omitted since the permutation between $\lambda_1^\star$ and $\lambda_2^\star$ makes in fact no difference. Thus, we finally considered the following configurations of the parameters: $(\lambda_1^\star, \lambda_2^\star, \phi) = \{(1, 1, 0.5); (1, 1, 1); (1, 3, 0.5); (1, 3, 1); (3, 3, 0.5); (3, 3, 1)\}$. The alternative combinations of the design parameters α_j; $j = 1, 2$ and $\lambda_1^\star$, $\lambda_2^\star$, ϕ resulted in a total number of 24 simulation experiments per sample size and for each experiment 500 iterations were conducted.

The simulated data sets that produced MoM or/and YW estimates in an inadmissible range were disregarded and iterations were continued till reaching the prespecified number of 500 per experiment. The number of invalid results that were disregarded has also been recorded. For the optimization of the likelihood function, we employed the `nlm` function in `R` adopting suitable parameter transformations.

The main discussion below focuses on the medium sample size of $n = 200$ since in this case, the parameter estimators show a representative behavior under the various simulation designs. For smaller sample sizes, higher biases and lower SD ratios are obtained but no differences are observed in terms of inference. For $n = 500$ and $n = 1000$ all estimators generally show minimal biases while the SD ratios clearly indicate the superiority of ML estimators. Full details regarding the simulation design are available from the authors upon request. Results are presented in detail in Appendix II.

6.2 Simulation Results

The degree of numerical instability problems encountered during the simulation procedure, varied depending on the sample size. More specifically, the tendency of MoM and/or YW methods to produce inadmissible estimates was markedly greater for smaller sample sizes. In the 24 simulation experiments of $n = 50$, the percentage of extra data sets needed to be sampled in order to gather an array of admissible estimates ranged from 19% to 55%. This percentage decreases rapidly as the sample size increases. For $n = 200$ it does not exceed the 15% while for $n = 500$ it hardly reaches the 3%. For $n = 1000$ the higher observed precentage was about 0.8% but in the vast majority of the simulation experiments no inadmissible estimates were generated.

The biases of the estimates $\hat{\alpha}_j$, $j = 1, 2$ and $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$, $\hat{\phi}$ of the Poisson BINAR(1) model and the ratios of MoM to ML and YW to ML standard deviations are presented in Table 1. All estimators perform well and exhibit a downward bias for α_1, α_2. Regarding the parameters $\lambda_1^\star$, $\lambda_2^\star$, ϕ, biases are mainly upward and clearly elevated in absolute value, especially in the case of MoM and YW estimates. Increasing n to 500 or 1000 dramatically decreases all biases while reducing the sample size to $n = 50$ has inverse consequences. However, the performance of all the three estimators remains poorer for $\lambda_1^\star$, $\lambda_2^\star$ and ϕ in comparison to α_1, α_2, independently of the sample size. Furthermore, while the superiority of ML estimators in terms of bias is clear for α_j's, $j = 1, 2$, the picture is more vague for $\lambda_1^\star$, $\lambda_2^\star$ and especially for ϕ. Comparison between the biases of MoM and YW estimators yields similar conclusions. In particular, the biases of the estimators $\hat{\alpha}_j$, $j = 1, 2$, are usually lower in absolute value when adopting the MoM but the same does not hold with clarity for the estimators $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and even more for $\hat{\phi}$.

Regarding the ratios of MoM to ML and YW to ML standard deviations for alternative parameter configurations and simulated data sets of $n = 200$, the advantages of ML estimators in terms of precision clearly emerge, since all ratios are greater than 1. It is also worth noting that the ratios of MoM to ML standard deviations are greater than the respective YW to ML ratios, indicating that the YW method usully leads to more precise estimates than those produced by MoM. However, this behavior is consistently reversed when estimating the parameter ϕ. Similar conclusions hold for shorter and larger count series.

The effect of sample size on the performance of the group of MoM, YW and ML estimators is depicted in Figure 1. This figure also provides a graphical inspection and comparison of the results obtained by employing different estimators. We present graphically only one representative case i.e., $(\alpha_1, \alpha_2, \lambda_1^\star, \lambda_2^\star, \phi) = (0.3, 0.5, 1, 3, 1)$. In view of a confrontation between the different methods of estimation and conditionally on a limited sample size (e.g. $n = 50$), the median estimates obtained by adopting the method of maximum likelihood are apparently closer (and indeed impressively close) to the real parameter values, than the median MoM and YW estimates. For larger samples, all estimators perform well in terms of location. Regarding dispersion issues, both the interquartile ranges and the overall range of the produced values are narrower for the ML than for YW and even more for MoM estimators. The only case in which dispersion differences are not definite, is in the estimation of the parameter ϕ but all its estimators should at least be considered as equivalent.

Summing up, according to our results the ML, YW and MoM methods are recommended in order of priority. The ML estimators are undoubtedly superior while the inferiority of MoM to YW estimators is less discernible. Finally, all the three methods of estimation yield more precise and less biased estimates for the parameters α_1, α_2 than for $\lambda_1^\star, \lambda_2^\star$ and ϕ.

7 Applications

The data are part of a large database related to syndromic surveillance during Athens 2004 Olympic Games and they refer to eleven different symptoms recorded during the period from March 2004 until end of September 2004 covering the period of Olympics and Paralympics games. For the purposes of the present application we selected a dataset of moderate length ($n = 94$) and considered two particular symptoms: “gastroenteritis (diarrhea, vomiting), without blood”and “other syndrome with potential interest for public health”. The latter was a general category including all symptoms that could not be classified in any of the following prespecified categories: respiratory infection with fever, bloody diarrhea, gastroenteritis (diarrhea, vomiting) without blood, febrile illness with rash, meningitis, encephalitis or unexplained acute encephalopathy / delirium, suspected viral hepatitis (acute), botulism like syndrome, lymphadenitis with fever, sepsis or unexplained shock and unexplained death with history of fever. From now on, we will use the brevities “gastroenteritis”and “other symptom”accordingly for the two symptoms under consideration. The series and their autocorrelations

Table 1: Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 200$.

		Bias			**SD ratios**	
(α_1, α_2, $\lambda_1^\star$, $\lambda_2^\star$, ϕ)		MoM	YW	ML	MoM/ML	YW/ML
(0.3, 0.3, 1, 1, 1)	$\hat{\alpha}_1$	-0.006	-0.009	-0.004	1.499	1.172
	$\hat{\alpha}_2$	-0.003	-0.006	-0.001	1.493	1.159
	$\hat{\lambda}_1^\star$	0.011	0.017	0.005	1.796	1.541
	$\hat{\lambda}_2^\star$	0.007	0.009	0.000	1.771	1.522
	$\hat{\phi}$	0.004	0.009	0.008	1.333	1.396
(0.3, 0.3, 1, 3, 1)	$\hat{\alpha}_1$	-0.006	-0.007	-0.002	1.340	1.078
	$\hat{\alpha}_2$	-0.008	-0.009	-0.005	1.363	1.106
	$\hat{\lambda}_1^\star$	0.029	0.029	0.011	1.521	1.353
	$\hat{\lambda}_2^\star$	0.048	0.050	0.026	1.505	1.262
	$\hat{\phi}$	-0.012	-0.009	-0.001	1.162	1.191
(0.3, 0.3, 3, 3, 1)	$\hat{\alpha}_1$	-0.011	-0.011	-0.004	1.349	1.075
	$\hat{\alpha}_2$	-0.007	-0.009	-0.003	1.413	1.106
	$\hat{\lambda}_1^\star$	0.049	0.049	0.007	1.368	1.186
	$\hat{\lambda}_2^\star$	0.027	0.033	-0.002	1.432	1.236
	$\hat{\phi}$	0.004	0.007	0.018	1.130	1.147
(0.3, 0.5, 1, 1, 1)	$\hat{\alpha}_1$	-0.012	-0.013	-0.008	1.488	1.175
	$\hat{\alpha}_2$	-0.022	-0.017	-0.005	1.990	1.241
	$\hat{\lambda}_1^\star$	0.034	0.034	0.009	1.648	1.470
	$\hat{\lambda}_2^\star$	0.085	0.059	0.007	2.215	1.587
	$\hat{\phi}$	-0.006	-0.003	0.013	1.209	1.270
(0.3, 0.5, 1, 3, 1)	$\hat{\alpha}_1$	-0.015	-0.014	-0.007	1.292	1.075
	$\hat{\alpha}_2$	-0.013	-0.017	-0.007	2.345	1.323
	$\hat{\lambda}_1^\star$	0.041	0.033	0.019	1.313	1.227
	$\hat{\lambda}_2^\star$	0.065	0.089	0.031	2.337	1.464
	$\hat{\phi}$	0.001	0.004	0.005	1.161	1.193
(0.3, 0.5, 3, 3, 1)	$\hat{\alpha}_1$	-0.012	-0.014	-0.009	1.440	1.104
	$\hat{\alpha}_2$	-0.017	-0.020	-0.008	2.290	1.291
	$\hat{\lambda}_1^\star$	0.046	0.051	0.054	1.412	1.243
	$\hat{\lambda}_2^\star$	0.085	0.101	0.054	2.100	1.401
	$\hat{\phi}$	0.028	0.034	0.010	1.133	1.163
(0.5, 0.5, 1, 1, 1)	$\hat{\alpha}_1$	-0.016	-0.014	-0.002	2.138	1.243
	$\hat{\alpha}_2$	-0.023	-0.019	-0.008	2.092	1.272
	$\hat{\lambda}_1^\star$	0.069	0.053	0.004	2.270	1.662
	$\hat{\lambda}_2^\star$	0.091	0.070	0.023	2.223	1.678
	$\hat{\phi}$	-0.012	-0.005	0.006	1.192	1.289
(0.5, 0.5, 1, 3, 1)	$\hat{\alpha}_1$	-0.022	-0.020	-0.007	2.032	1.238
	$\hat{\alpha}_2$	-0.007	-0.011	-0.001	2.394	1.308
	$\hat{\lambda}_1^\star$	0.069	0.055	0.004	1.739	1.408
	$\hat{\lambda}_2^\star$	0.028	0.053	-0.015	2.256	1.448
	$\hat{\phi}$	-0.005	0.003	0.012	1.197	1.229
(0.5, 0.5, 3, 3, 1)	$\hat{\alpha}_1$	-0.014	-0.017	-0.005	2.270	1.263
	$\hat{\alpha}_2$	-0.012	-0.015	-0.003	2.179	1.263
	$\hat{\lambda}_1^\star$	0.025	0.037	0.020	1.995	1.396
	$\hat{\lambda}_2^\star$	0.017	0.027	0.014	1.942	1.391
	$\hat{\phi}$	0.058	0.069	0.012	1.198	1.240

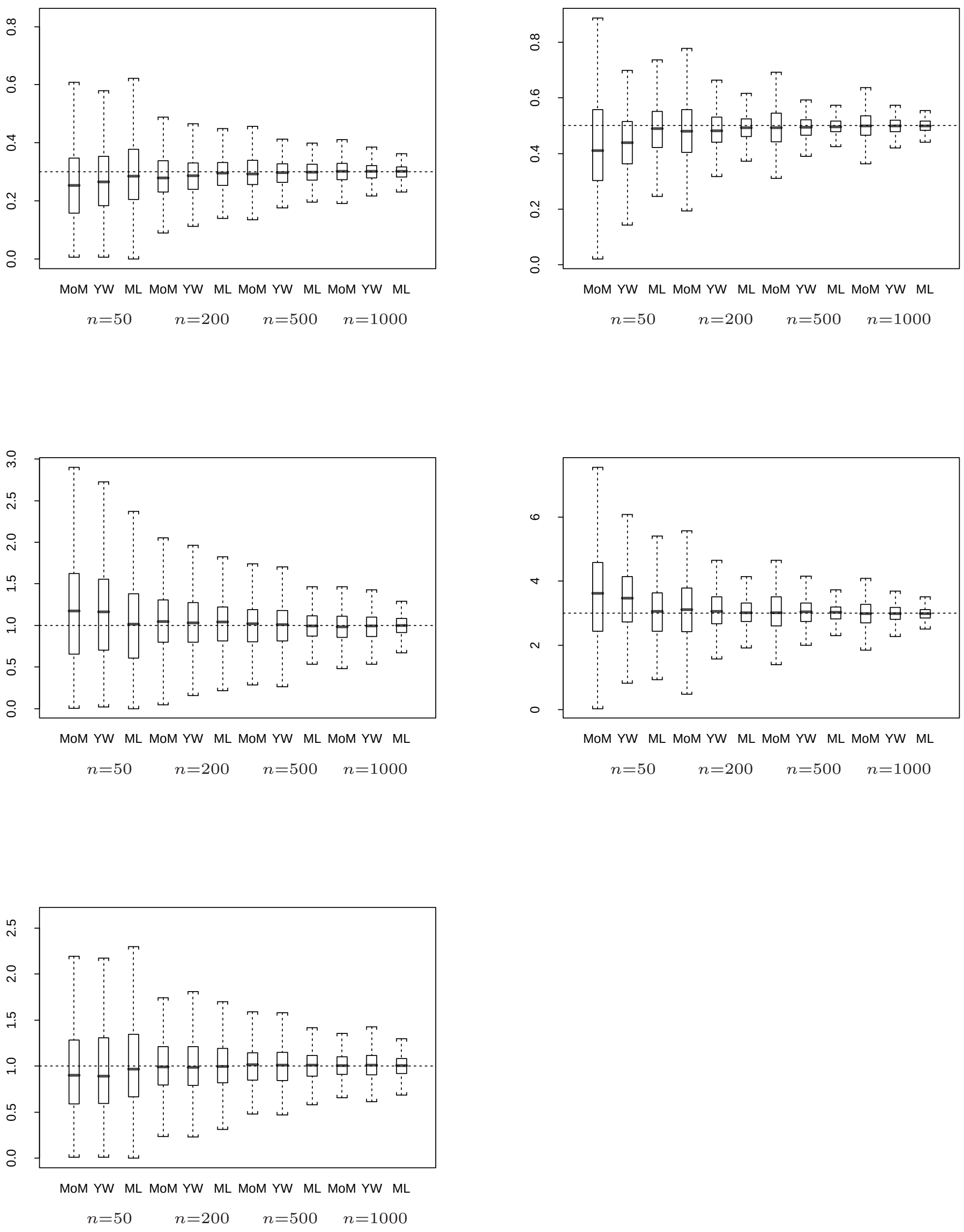

Figure 1: Boxplot with results from the simulation experiment with initial parameter values $(\alpha_1, \alpha_2, \lambda_1^\star, \lambda_2^\star, \phi) = (0.3, 0.5, 1, 3, 1)$.

can be seen in figure 2. The gastroenteritis and other symptom series have mean values (variances) equal to 1.29 (1.24) and 9.52 (16.53) respectively implying overdispersion. The frequent zero frequencies in the gastroenteritis series implies the appropriateness of fitting a count data model. The autocorrelation functions for both series present a rather exponential decay with a few exceptions. The first order autocorrelation coefficient is 0.24 and tends to decrease as passing through higher lags.

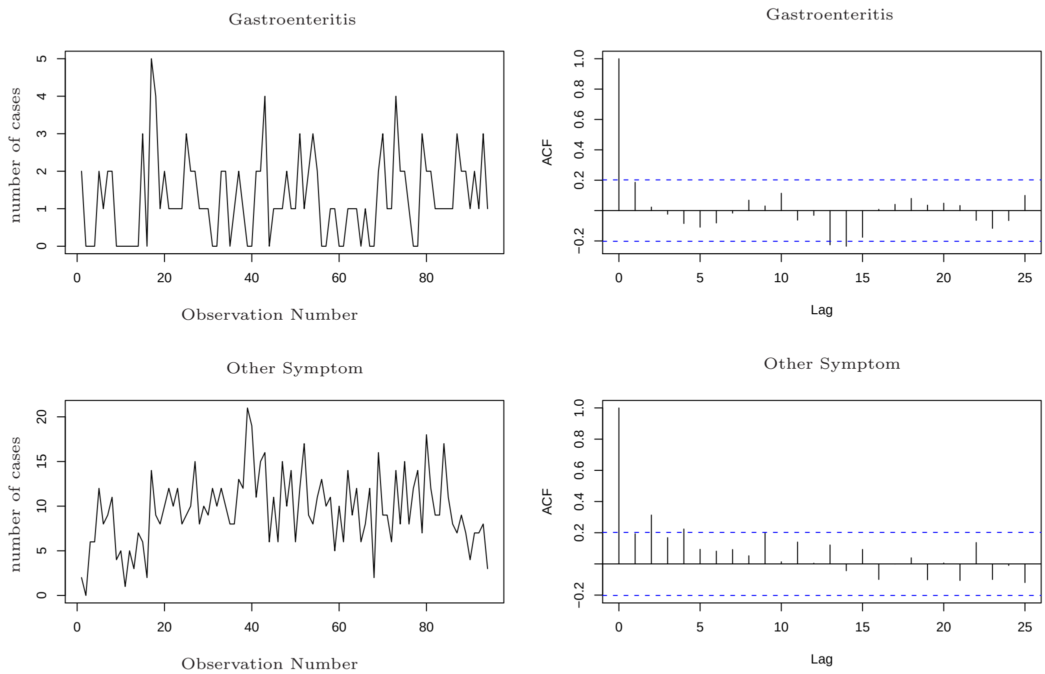

Figure 2: Time series plots for the series of gastroenteritis and other symptom.

In Table 2, ML estimates obtained by fitting a BINAR(1) model with BVNB innovations are contrasted to the corresponding estimates that are produced by fitting a Poisson BINAR(1) model, two independent Poisson INARs (ignoring correlation between the series) and a bivariate Poisson model (ignoring the time dependence). Compared with the independent INAR(1) estimates, the Poisson BINAR(1) model produces lower values for $\hat{\alpha}_1$, $\hat{\alpha}_2$ and higher values for $\hat{\lambda}_1$ and $\hat{\lambda}_2$. This seems to be a reasonable result since the correlation between the two series is ignored by the INAR approach. Thus, the explanatory ability of the parameters λ_1, λ_2 is underestimated by the two independent INAR(1) processes, a fact that naturally also affects the

Table 2: Maximum Likelihood Estimates from fitting alternatively a BINAR(1), two independent INAR(1) models and a simple bivariate Poisson model.

	BINAR(1)		Independent INAR(1)		Biv. Poisson		Neg.Bin BINAR(1)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
$\hat{\alpha}_1$	0.2172	0.1075	0.2423†	0.1121†			0.2760	0.1058
$\hat{\alpha}_2$	0.1539	0.0656	0.1544‡	0.0689‡			0.2650	0.0889
$\hat{\lambda}_1$	0.9993	0.1663	0.9669†	0.1656†	1.2796	0.1173	0.9234	0.1651
$\hat{\lambda}_2$	8.1259	0.6857	8.1211‡	0.7144‡	9.6021	0.3214	7.0602	0.9323
$\hat{\phi}$	0.9502	0.2878			1.1103	0.3147		
$\hat{\beta}$							0.1610	0.0673
Log-Lik	-394.2857		-399.2293		-398.6475		-387.3354	
AIC	798.5714		806.4586		803.2950		784.6709	

† *gastroenteritis*
‡ *other symptom*

estimation of the parameters α_1, α_2. Comparing the log-likelihoods, one can see that both the time series context and the correlation between the series are needed. The BINAR(1) model with BVNB innovations can also model the overdispersion and thus it provides the better fit.

It is clear that the time series models are better than the models that neglect this. In addition the BINAR(1) model with BVNB innovations is much better as it captures the overdispersion in the dataset together with the correlation between the two series but also the autocorrelation within each series. The standard errors of the estimates obtained by the two approaches (standard errors are derived numerically from the Hessian) show that fitting a BINAR(1) model to the data generally improves the precision of the produced estimates. On the other hand it is apparent that ignoring any form of the correlation (either within or between) or the overdispersion leads to incorrect standard errors and hence incorrect inferences.

In what follows we concern results obtained from fitting a BINAR(1) model with BVNB innovations. However, similar conclusions are derived regarding the Poisson BINAR(1) model. Figure 3 includes the plots of the residuals of the two series. Since these residuals have not been standardized, the survival and arrival residuals add up to the Pearson residuals. Moreover, a large Pearson residual is comprised by a large survival and arrival residual, while a small Pearson residual consists of a small survival and arrival residual. The signs of survival and arrival residuals may also differ in some cases. However, they still keep their similarity in pattern.

Another interesting point is the reflection of the model structure in the correlation between different pairs of residuals. More specifically, the sample correlations between the survival and arrival residuals of each series are very high: 0.51 for the gastroenteritis series and 0.82 for the series of other symptom. The arrival residuals of the two series are also significantly correlated at 0.30 while the survival residuals exhibit a weak correlation at 0.10. Those results are in accordance with the structural assumption underlying the BINAR(1) model that the correlation between the two series has been introduced by using correlated innovation terms.

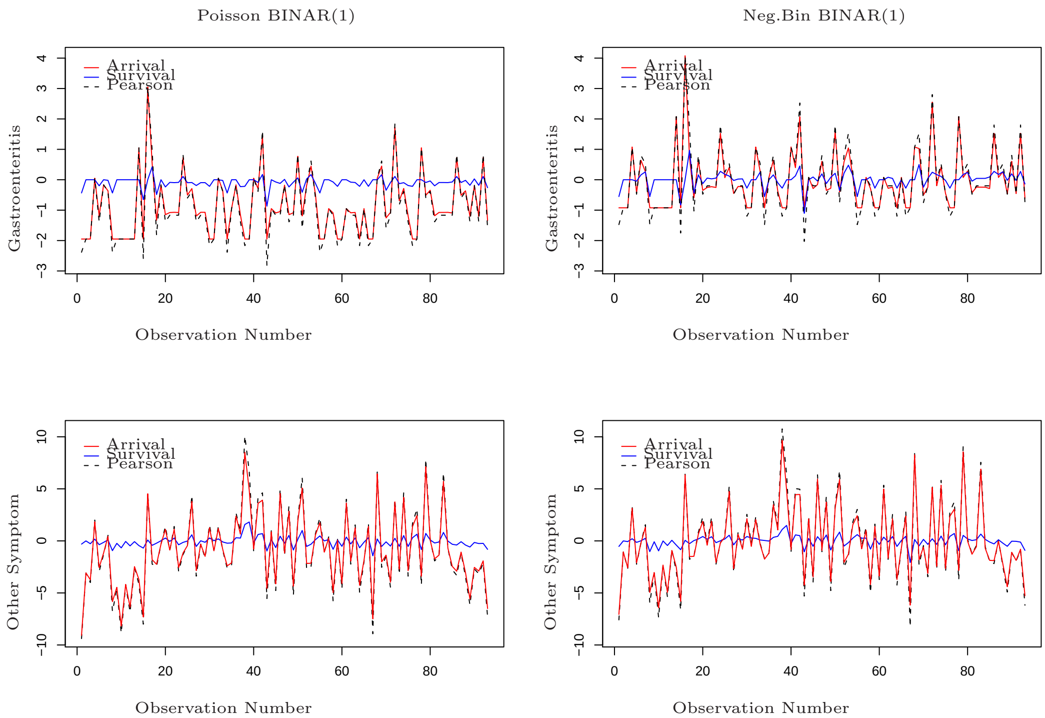

Figure 3: Non-standardized residuals of the gastroenteritis and other symptom series.

Figure 4 shows the one-step-ahead marginal predictive distributions $P(x_{1,T+1}|x_{1T}, x_{2T})$ and $P(x_{2,T+1}|x_{1T}, x_{2T})$ where X_1 corresponds to the series of gastroenteritis and X_2 corresponds to the series of other symptom. The last observation was equal to 1 for the series of gastroenteritis and equal to 3 for the series of other symptom. The usefulness of such predictive distributions is that they permit the recognition of an unexpectedly large observation and thus alarm for a more thorough examination of the situation under surveillance. As one can see in figure 4, both distributions are skewed to the right which is in accordance with the shapes of the Poisson and negative binomial distributions. According to the BINAR(1) model with BVNB innovations, the most probable one-step-ahead predictive value is equal to 1 for the gastroenteritis series and equal to 6 for the series of other symptom. The larger dispersion of the series of other symptom compared with the gastroenteritis series is also reflected in the plot of its predictive distribution.

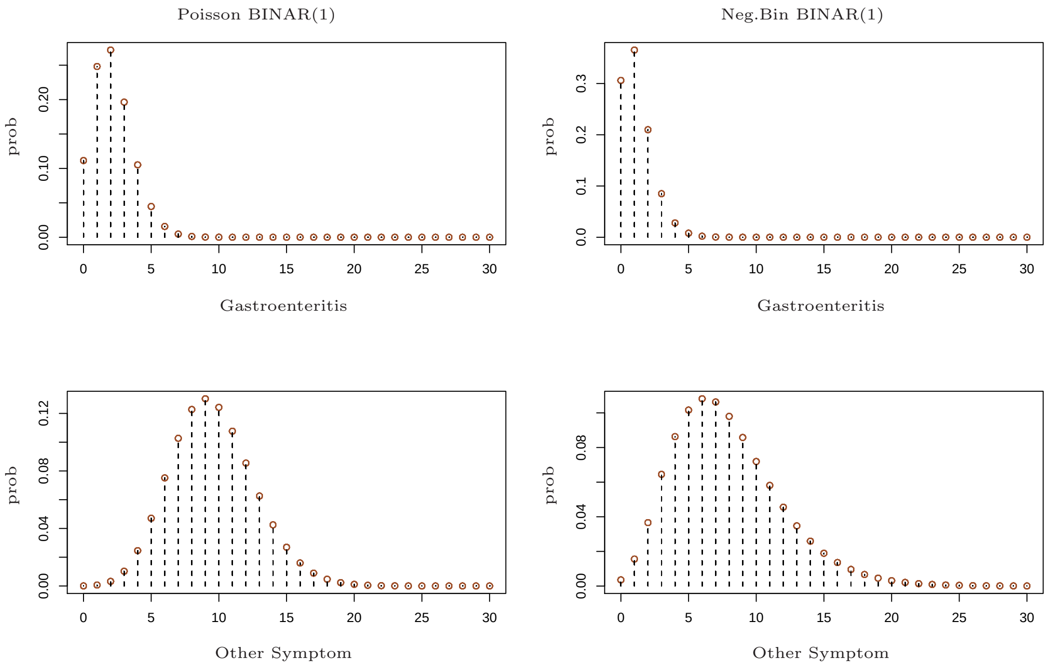

Figure 4: The one-step-ahead predictive distributions $P(x_{1,T+1}|x_{1T}, x_{2T})$ $P(x_{2,T+1}|x_{1T}, x_{2T})$ of the series of gastroenteritis and other symptom respectively. The last observed values ($T = 94$) are equal to 1 and 3 accordingly.

Application of Theorem 2 provides us with appropriate confidence intervals for the estimated one-step-ahead probability forecasts obtained from both BINAR(1) models under consideration. Some representative results are presented in Table 3 for both bivariate (lower panel) and univariate probability forecasts (upper panels). Most of the intervals are narrower when prediction is implemented through the BINAR(1) model with BVNB innovations confirming the superiority of the aformentioned model for our data.

Two examples of the contours of the joint probability density functions of the estimated one-step-ahead *jpmf*'s for pairs of possible counts are shown in figure 5. Since the δ-method has been used for the calculation of the conditional probability forecasts, their joint probability density functions are asymptotically bivariate normal (Bu and McCabe, 2008). It is obvious from

Table 3: 95% C.I.'s for some one-step-ahead probability forecasts.

	Poisson BINAR(1)		Neg.Bin BINAR(1)	
i	$\hat{P}(X_{T+1}=i)$	$\hat{P}(X_{T+1}\leq i)$	$\hat{P}(X_{T+1}=i)$	$\hat{P}(X_{T+1}\leq i)$
0	(0.053, 0.170)	(0.053, 0.170)	(0.205, 0.407)	(0.205, 0.407)
1	(0.183, 0.314)	(0.235, 0.484)	(0.360, 0.371)	(0.565, 0.778)
2	((0.265, 0.279)	(0.501, 0.762)	(0.169, 0.251)	(0.816, 0.947)
j	$\hat{P}(Y_{T+1}=j)$	$\hat{P}(Y_{T+1}\leq j)$	$\hat{P}(Y_{T+1}=j)$	$\hat{P}(Y_{T+1}\leq j)$
5	(0.028, 0.066)	(0.046, 0.125)	(0.089, 0.114)	(0.284, 0.333)
6	(0.052, 0.098)	(0.099, 0.224)	(0.086, 0.130)	(0.414, 0.419)
7	(0.081, 0.125)	(0.179, 0.348)	(0.081, 0.132)	(0.500, 0.546)
i, j	$\hat{P}(X_{T+1}=i, Y_{T+1}=j)$	$\hat{P}(X_{T+1}\leq i, Y_{T+1}\leq j)$	$\hat{P}(X_{T+1}=i, Y_{T+1}=j)$	$\hat{P}(X_{T+1}\leq i, Y_{T+1}\leq j)$
0, 5	(0.007, 0.009)	(0.002, 0.030)	(0.035, 0.037)	(0.106, 0.130)
1, 6	(0.021, 0.024)	(0.028, 0.127)	(0.040, 0.042)	(0.297, 0.329)
2, 7	(0.028, 0.030)	(0.105, 0.282)	(0.021, 0.023)	(0.447, 0.515)

the contours presented in figure 5 that the estimated forecast *jpmf*'s for different pairs of counts are more or less correlated. For example, the first contour suggest that $\hat{P}(X_{T+1}=1, Y_{T+1}=6)$ and $\hat{P}(X_{T+1}=1, Y_{T+1}=5)$ are highly correlated at 0.875 while the correlation between $\hat{P}(X_{T+1}=0, Y_{T+1}=5)$ and $\hat{P}(X_{T+1}=1, Y_{T+1}=8)$ is negative and weaker (-0.408).

Figure 6 shows the observed values of the series of gastroenteritis and other symptom together with the corresponding one-step-ahead predictions. The divergence between real data and forecasts is also portrayed. The horizontal lines correspond to the observed mean values of the two series. Obviously, divergence is larger for observations that lie far away from the mean. This seems to be expected since the one-step-ahead predictions have the same mean but are less dispersed than the original series. Note also that the correlation coefficient of the two series of forecasts is equal to the correlation coefficient of the real data series.

8 Concluding Remarks

The main focus in this paper is on bivariate time series for count data. Generally, the desired BINAR(1) model can be constructed in two different ways: The first approach prespecifies the form of the marginal distributions and subsequently identifies the required form of the distribution of the innovations in order for stationarity to hold. In the second approach it is the choice of the form of the innovations distribution that leads to the specification of the underlying marginal distributions. The models proposed in this paper

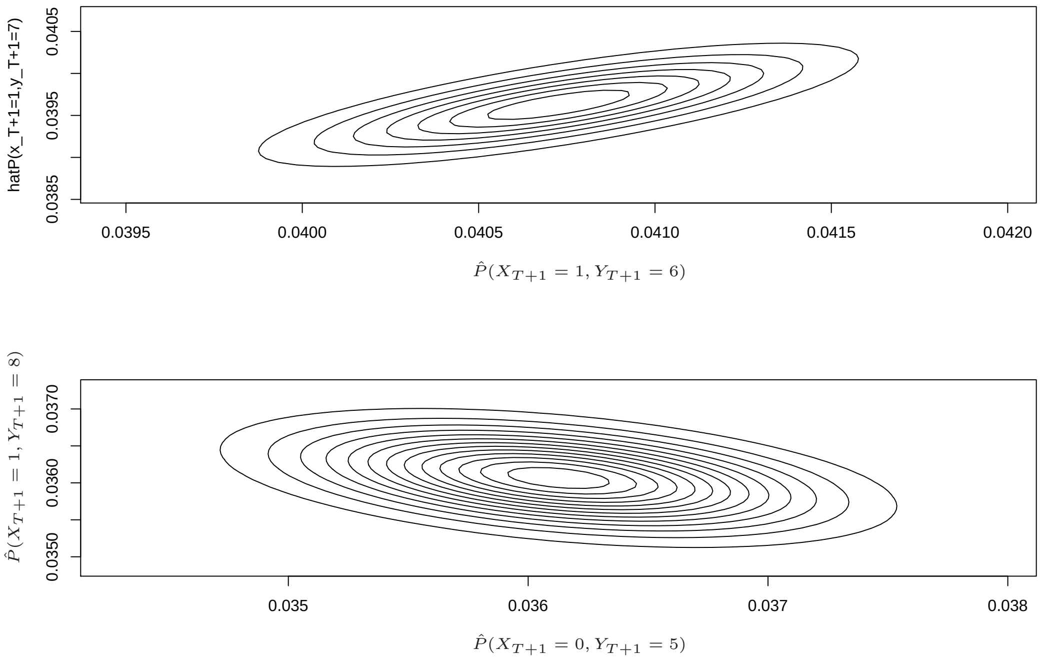

Figure 5: Examples of the joint density functions of the estimated one-step-ahead *jpmf*'s under the BINAR(1) model with BVNB innovations.

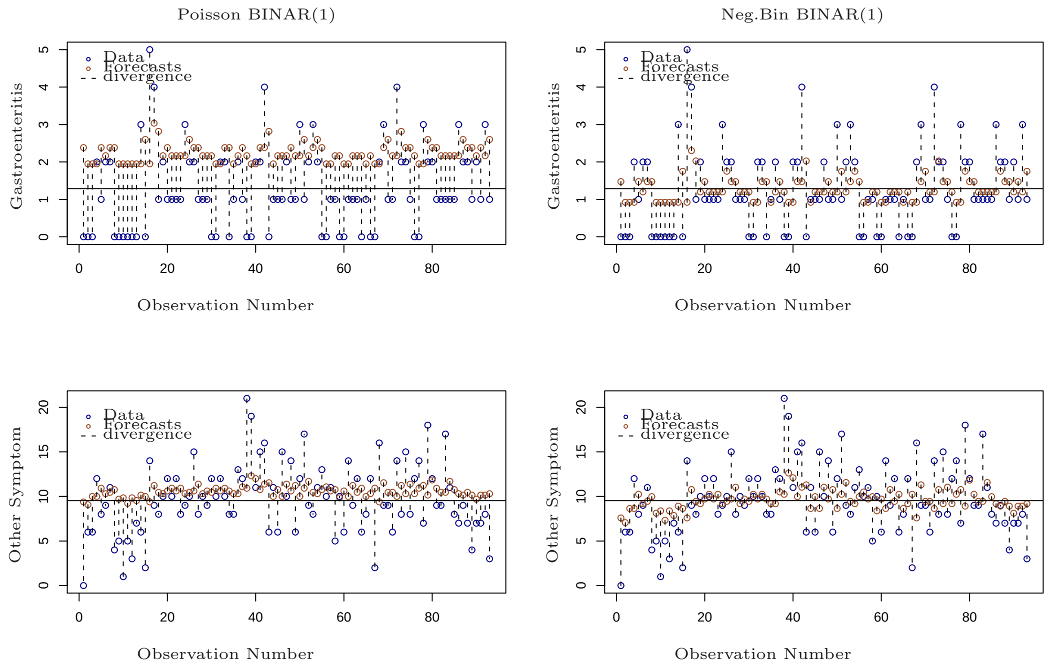

Figure 6: Observed values of the series of gastroenteritis and other symptom and the corresponding one-step-ahead predictions.The horizontal lines correspond to the observed mean values of the two series.

have been built following the last approach. In particular, we considered two different BINAR(1) models, one with bivariate Poisson innovations and another one with bivariate negative binomial innovations. The former specification has the facilitating property that the joint distribution of the two series under consideration is also bivariate Poisson. In the latter case, we don't end up with a bivariate negative binomial INAR(1) process but we obtain a BINAR(1) model that effectively accounts for overdispersion. Deviations from the equidispersion restriction could alternatively accounted for by assuming another distribution for the innovations, e.g. mixed Poisson, or by the inclusion of appropriate regressors. Results on such extensions will be reported elsewhere.

It is of course self-evident that the proposed model is not a panacea. For example, when significant correlation between the series under consideration is present at lags higher than 1, fitting a BINAR(1) model proves to be rather inadequate. Thus, extensions of the present model to higher orders would be a useful contribution to the improvement of its flexibility. Moreover, the structure of real-life data frequently implies need for the inclusion of both autoregressive and moving average components (when for example seasonal patterns are observed in time series of counts). So, extending the bivariate INAR model to a bivariate INARMA model seems to be another interesting challenge. Finally, generalization of the proposed process to the multivariate case would provide a great opportunity for modelling more than two time series of correlated count data. In this case, the definition of a multivariate discrete distribution for the innovation process is needed. The existing models have certain limitations and they do not lead to models with well specified marginals. Hence inference can be difficult with standard methods like maximum likelihood and some alternatives, like composite likelihood, should be considered.

Acknowledgements

The authors would like to thank the Hellenic Center for Infectious Diseases Control and Prof. Urania Dafni for making the data used in the paper available to us. We would also like to thank Prof. Klea Katsouyanni for her support.

References

J. Aitchinson and C. Ho. The multivariate Poisson-log normal distribution. *Biometrika*, 75:621–629, 1989.

M.A. Al-Osh and A.A. Alzaid. First-Order Integer-Valued Autoregressive Process. *Journal of Time Series Analysis*, 8(3):261–275, 1987.

J.P. Boucher, D. Michel, and G. Montserrat. Models of insurance claim counts with time dependence based on generalization of poisson and negative binomial distributions. *Variance*, 2(1):135–162, 2008.

K. Brännäs and J. Nordström. A Bivariate Integer Valued Allocation Model for Guest Nights in Hotels and Cottages. *Umeå Economic Studies*, 547, 2000.

R. Bu and B. McCabe. Model selection, estimation and forecasting in inar(p) models: A likelihood-based markov chain approach. *International Journal of Forecasting*, 24:151–162, 2008.

D. Chabot-Hallé and P. Duchesne. Diagnostic checking of multivariate nonlinear time series models with martingale difference errors. *Statistics and Probability Letters*, 78:997–1005, 2008.

S. Cheon, S.H. Song, and B.C. Jung. Tests for independence in a bivariate negative binomial model. *Journal of the Korean Statistical Society*, 38: 185–190, 2009.

S. Chib and R. Winkelmann. Markov Chain Monte Carlo Analysis of Correlated Count Data. *Journal of Business & Economic Statistics*, 19(4): 428–435, 2001.

R.A. Davis, W.T. Dunsmuir, and Y. Wang. *Modelling time series of count data*, pages 63–114. Ed. S. Ghosh, Marcel Dekker, 1999.

K. Egan and J. Herriges. Multivariate count data regression models with individual panel data from an on-site sample. *Journal of Environmental Economics and Management*, 52:567–581, 2006.

R.K. Freeland and B.P.M. McCabe. Analysis of low count time series data by Poisson autoregression. *Journal of Time Series Analysis*, 25(5):701–722, 2004a.

R.K. Freeland and B.P.M. McCabe. Forecasting discrete valued low count time series. *International Journal of Forecasting*, 20:427–434, 2004b.

R.K. Freeland and B.P.M. McCabe. Asymptotic properties of CLS estimators in the Poisson AR(1) model. *Statistics & Probability Letters*, 73:147–153, 2005.

S. Gurmu and J. Elder. Generalized bivariate count data regression models. *Econometric Letters*, 68:31–36, 2000.

A. Heinen and E. Rengifo. Multivariate autoregressive modeling of time series count data using copulas. *Journal of Empirical Finance*, 14:564–583, 2007.

N. Johnson, S. Kotz, and N. Balakrishnan. *Multivariate Discrete Distributions*. Wiley, New York, 1997.

R.C. Jung and A.R. Tremayne. Binomial thinning models for integer time series. *Statistical Modelling*, 6:21–96, 2006.

D. Karlis and L. Meligkotsidou. Finite multivariate Poisson mixtures with applications. *Journal of Statistical Planning and Inference*, 137:1942–1960, 2007.

S. Kocherlakota and K. Kocherlakota. *Bivariate Discrete Distributions, Statistics: textbooks and monographs*, volume 132. Markel Dekker, New York, 1992.

A. Latour. The Multivariate GINAR(p) Process. *Advances in Applied Probability*, 29(1):228–248, 1997.

A.W. Marshall and I. Olkin. *Multivarirate distribution generated from mixtures of convolution and product families*, pages 371–393. Topics in Statistical Dependence, Block,Sampson y Sanits (Eds), Institute of Mathematical Statistics, 1990.

E. McKenzie. Some Simple Models for Discrete Variate Time Series. *Water Resources Bulletin*, 21(4):645–650, 1985.

E. McKenzie. *Discrete Variate Time Series*, volume 21. Shanbhag, D. and Rao, C. (eds.), Elsevier Science, 2003.

A.K. Nikoloulopoulos and D. Karlis. Finite normal mixture copulas for multivariate discrete data modeling. *Journal of Statistical Planning and Inference (In Press)*, 2009.

H. Pavlopoulos and D. Karlis. INAR(1) modeling of overdispersed count series with an environmental application. *Environmetrics*, 19(4):369–393, 2008.

A.M.M.S. Quoreshi. Bivariate Time Series Modeling of Financial Count Data. *Communications in Statistics - Theory and Methods*, 35:1343–1358, 2006.

F.W. Steutel and K. van Harn. Discrete Analogues of Self - Decomposability and Stability. *The Annals of Probability*, 7(5):893–899, 1979.

Appendix I.

PROOF OF EQUATION (2.6).

$$
\begin{aligned}
\mathrm{Cov}(X_{1t}, R_{2t}) &= E(X_{1t}R_{2t}) - E(X_{1t})E(R_{2t}) \\
&= E(X_{1t}R_{2t}) - \mu_1\lambda_2 \\
&= E\left[\left(\sum_{i=0}^{\infty} \alpha_1^i \circ R_{1t-i}\right) R_{2t}\right] - \mu_1\lambda_2 \\
&= \sum_{i=0}^{\infty} \alpha_1^i \left\{E(R_{1t-i}R_{2t})\right\} - \mu_1\lambda_2 \\
&= E(R_{1t}R_{2t}) + \sum_{i=1}^{\infty} \alpha_1^i \left\{E(R_{1t-1}R_{2t})\right\} - \mu_1\lambda_2 \\
&= \mathrm{Cov}(R_{1t}, R_{2t}) + \lambda_1\lambda_2 + \sum_{i=1}^{\infty} \alpha_1^i \left\{\lambda_1\lambda_2\right\} - \mu_1\lambda_2 \\
&= \mathrm{Cov}(R_{1t}, R_{2t}) + \sum_{i=0}^{\infty} \alpha_1^i \left\{\lambda_1\lambda_2\right\} - \mu_1\lambda_2 \\
&= \mathrm{Cov}(R_{1t}, R_{2t})
\end{aligned}
$$

PROOF OF EQUATION (2.7).

i. $h = 0$

$$
\begin{aligned}
\mathrm{Cov}(X_{1t}, X_{2t}) &= E(X_{1t}X_{2t}) - \mu_1\mu_2 \\
&= E\left\{X_{1t}(\alpha_2 \circ X_{2t-1} + R_{2t})\right\} - \mu_1\mu_2 \\
&= \alpha_2 E(X_{1t}X_{2t-1}) + E(X_{1t}R_{2t}) - \mu_1\mu_2 \\
&= \alpha_2 E(X_{1t}X_{2t-1}) + \mu_1\lambda_2 + \mathrm{Cov}(R_{1t}, R_{2t}) - \mu_1\mu_2 \\
&= \alpha_2 E(X_{1t}X_{2t-1}) - \alpha_2\mu_1\mu_2 + \mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_2 \mathrm{Cov}(X_{1t}, X_{2t-1}) + \mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_2 E\left\{(\alpha_1 \circ X_{1t-1} + R_{1t})X_{2t-1}\right\} - \alpha_2\mu_1\mu_2 + \mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_1\alpha_2 E(X_{1t-1}X_{2t-1}) + \alpha_2 E(R_{1t}X_{2t-1}) - \alpha_2\mu_1\mu_2 + \mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_1\alpha_2 E(X_{1t-1}X_{2t-1}) - \alpha_1\alpha_2\mu_1\mu_2 + \mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_1\alpha_2 \mathrm{Cov}(X_{1t-1}, X_{2t-1}) + \mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_1\alpha_2 E\left\{X_{1t-1}(\alpha_2 \circ X_{2t-2} + R_{2t-1})\right\} - \alpha_1\alpha_2\mu_1\mu_2 + \mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_1\alpha_2^2 E(X_{1t-1}X_{2t-2}) + \alpha_1\alpha_2 E(X_{1t-1}R_{2t-1}) - \alpha_1\alpha_2\mu_1\mu_2 + \mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_1\alpha_2^2 E(X_{1t-1}X_{2t-2}) - \alpha_1\alpha_2^2\mu_1\mu_2 + (1 + \alpha_1\alpha_2)\mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_1\alpha_2^2 \mathrm{Cov}(X_{1t-1}, X_{2t-2}) + (1 + \alpha_1\alpha_2)\mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_1^2\alpha_2^2 \mathrm{Cov}(X_{1t-2}, X_{2t-2}) + (1 + \alpha_1\alpha_2)\mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_1^2\alpha_2^3 \mathrm{Cov}(X_{1t-2}, X_{2t-3}) + (1 + \alpha_1\alpha_2 + \alpha_1^2\alpha_2^2)\mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_1^3\alpha_2^3 \mathrm{Cov}(X_{1t-3}, X_{2t-3}) + (1 + \alpha_1\alpha_2 + \alpha_1^2\alpha_2^2)\mathrm{Cov}(R_{1t}, R_{2t}) \\
&= \alpha_1^h\alpha_2^h \mathrm{Cov}(X_{1t-h}, X_{2t-h}) + \mathrm{Cov}(R_{1t}, R_{2t}) \sum_{j=0}^{h-1} (\alpha_1\alpha_2)^j
\end{aligned}
$$

Assuming that $\{X_{1t}, X_{2t}\}$ is a stationary bivariate process, it holds that

$$
\mathrm{Cov}(X_{1t+h}, X_{2t+h}) = \mathrm{Cov}(X_{1t-h}, X_{2t-h})
$$

$$\Rightarrow \quad \alpha_1^h\alpha_2^h\mathrm{Cov}(X_{1t},X_{2t}) + \mathrm{Cov}(R_{1t},R_{2t})\sum_{j=0}^{h-1}(\alpha_1\alpha_2)^j = \mathrm{Cov}(X_{1t-h},X_{2t-h})$$

$$\Rightarrow \quad \alpha_1^h\alpha_2^h\left\{\alpha_1^h\alpha_2^h\mathrm{Cov}(X_{1t-h},X_{2t-h}) + \mathrm{Cov}(R_{1t},R_{2t})\sum_{j=0}^{h-1}(\alpha_1\alpha_2)^j\right\} + \mathrm{Cov}(R_{1t},R_{2t})\sum_{j=0}^{h-1}(\alpha_1\alpha_2)^j$$
$$= \mathrm{Cov}(X_{1t-h},X_{2t-h})$$

$$\Rightarrow \quad \alpha_1^{2h}\alpha_2^{2h}\mathrm{Cov}(X_{1t-h},X_{2t-h}) + \mathrm{Cov}(R_{1t},R_{2t})(1+\alpha_1^h\alpha_2^h)\sum_{j=0}^{h-1}(\alpha_1\alpha_2)^j = \mathrm{Cov}(X_{1t-h},X_{2t-h})$$

$$\Rightarrow \quad \mathrm{Cov}(X_{1t-h},X_{2t-h}) = \frac{\mathrm{Cov}(R_{1t},R_{2t})(1+\alpha_1^h\alpha_2^h)\sum_{j=0}^{h-1}(\alpha_1\alpha_2)^j}{1-(\alpha_1^h\alpha_2^h)^2}$$

$$\Rightarrow \quad \mathrm{Cov}(X_{1t-h},X_{2t-h}) = \frac{\mathrm{Cov}(R_{1t},R_{2t})\sum_{j=0}^{h-1}(\alpha_1\alpha_2)^j}{1-(\alpha_1\alpha_2)^h}$$

$$\Rightarrow \quad \mathrm{Cov}(X_{1t},X_{2t}) = \frac{\mathrm{Cov}(R_{1t},R_{2t})}{1-\alpha_1\alpha_2}$$

ii. $h = 1, 2, \ldots$

$$\begin{aligned}
\mathrm{Cov}(X_{1t+h},X_{2t}) &= E(X_{1t+h}X_{2t}) - \mu_1\mu_2 \\
&= E\left\{\alpha_1 \circ X_{1t+h-1} + R_{1t+h})X_{2t}\right\} - \mu_1\mu_2 \\
&= \alpha_1 E(X_{1t+h-1}X_{2t}) + E(R_{1t+h}X_{2t}) - \mu_1\mu_2 \\
&= \alpha_1 E(X_{1t+h-1}X_{2t}) + \lambda_1\mu_2 - \mu_1\mu_2 \\
&= \alpha_1 E(X_{1t+h-1}X_{2t}) - \alpha_1\mu_1\mu_2 \\
&= \alpha_1 \mathrm{Cov}(X_{1t+h-1},X_{2t}) \\
&= \alpha_1 E\left\{\alpha_1 \circ X_{1t+h-2} + R_{1t+h-1})X_{2t}\right\} - \alpha_1\mu_1\mu_2 \\
&= \alpha_1^2 E(X_{1t+h-2}X_{2t}) + \alpha_1 E(R_{1t+h-1}X_{2t}) - \alpha_1\mu_1\mu_2 \\
&= \alpha_1^2 E(X_{1t+h-2}X_{2t}) + \alpha_1\lambda_1\mu_2 - \alpha_1\mu_1\mu_2 \\
&= \alpha_1^2 E(X_{1t+h-2}X_{2t} - \alpha_1^2\mu_1\mu_2 \\
&= \alpha_1^2 \mathrm{Cov}(X_{1t+h-2},X_{2t}) \\
&= \alpha_1^h \mathrm{Cov}(X_{1t},X_{2t}) \\
&= \frac{\alpha_1^h}{1-\alpha_1\alpha_2}\mathrm{Cov}(R_{1t},R_{2t})
\end{aligned}$$

PROOF OF THE EQUIVALENCE OF CLS AND YW ESTIMATORS FOR THE POISSON BINAR(1) MODEL

The CLS estimators of the parameters $\alpha_j, \lambda_j,\ j = 1, 2$ can be found by minimizing the criterion

$$\begin{aligned} Q(\alpha_j, \lambda_j) &= \sum_{t=2} T \left(X_{jt} - E(X_{jt}|X_{j,t-1})\right)^2 \\ &= \sum_{t=2}^{T} \left(X_{jt} - (\alpha_j X_{j,t-1} + \lambda_j)\right)^2 \end{aligned}$$

So, in evaluating the derivatives of Q with respect to α_j and λ_j, $j = 1, 2$, we obtain:

$$\tilde{\alpha}_j^{CLS} = \frac{\sum_{t=2}^{T} \left(X_{jt} - \bar{X}_j\right)\left(X_{j,t-1} - \bar{X}_j\right)}{\sum_{t=2}^{T} \left(X_{j,t-1} - \bar{X}_j\right)^2}$$

$$\tilde{\lambda}_j^{CLS} = \frac{1}{T-1}\left\{\sum_{t=2}^{T} X_{jt} - \tilde{\alpha}_j^{CLS} \sum_{t=2}^{T} X_{j,t-1}\right\}$$

These estimators coincide with the corresponding ones for the Poisson INAR(1) model. Freeland and McCabe (2005) prove the asymptotic equivalence between the YW and CLS estimators for the Poisson INAR(1) model and thus this equivalence also holds for the BINAR(1) model. In order to accomplish the proof we have to show that the equivalence also holds for the parameter ϕ of the BINAR(1) model:

For the evaluation of the CLS estimator for the parameter ϕ, we define the criterion

$$\begin{aligned} S &= \sum_{t=2}^{T} \left(Q(\alpha_1, \lambda_1) Q(\alpha_2, \lambda_2) - \gamma_{0|t-1}\right)^2 \\ &= \sum_{t=2}^{T} \left\{\left[X_{1t} - (\alpha_1 X_{1,t-1} + \lambda_1)\right]\left[X_{2t} - (\alpha_2 X_{2,t-1} + \lambda_2)\right] - \phi\right\}^2 \end{aligned}$$

Minimization of S with respect to ϕ yields

$$\begin{aligned}
\tilde{\phi}^{CLS} &= \frac{1}{T-1}\left\{\sum_{t=2}^{T}\left(X_{1t}-\tilde{\alpha}_1^{CLS}X_{1,t-1}\right)\left(X_{2t}-\tilde{\alpha}_2^{CLS}X_{2,t-1}\right)\right. \\
&- \left.\sum_{t=2}^{T}\left(X_{1t}-\tilde{\alpha}_1^{CLS}X_{1,t-1}\right)\sum_{t=2}^{T}\left(X_{2t}-\tilde{\alpha}_2^{CLS}X_{2,t-1}\right)\right\}
\end{aligned}$$

The YW estimator of ϕ is given by

$$\tilde{\phi}^{YW} = \frac{(1-\tilde{\alpha}_1^{YW}\tilde{\alpha}_2^{YW})}{T}\left\{\sum_{t=1}^{T}X_{1t}X_{2t}-\sum_{t=1}^{T}X_{1t}\sum_{t=1}^{T}X_{2t}\right\}$$

Obviously,

$$\begin{aligned}
T^{1/2}(\tilde{\phi}^{YW}-\tilde{\phi}^{CLS}) &= \frac{(1-\tilde{\alpha}_1^{YW}\tilde{\alpha}_2^{YW})}{T^{1/2}}\left\{\sum_{t=1}^{T}X_{1t}X_{2t}-\sum_{t=1}^{T}X_{1t}\sum_{t=1}^{T}X_{2t}\right\} \\
&- \frac{T^{1/2}}{T-1}\left\{\sum_{t=2}^{T}\left(X_{1t}-\tilde{\alpha}_1^{CLS}X_{1,t-1}\right)\left(X_{2t}-\tilde{\alpha}_2^{CLS}X_{2,t-1}\right)\right. \\
&- \left.\sum_{t=2}^{T}\left(X_{1t}-\tilde{\alpha}_1^{CLS}X_{1,t-1}\right)\sum_{t=2}^{T}\left(X_{2t}-\tilde{\alpha}_2^{CLS}X_{2,t-1}\right)\right\}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow T^{1/2}(\tilde{\phi}^{YW}-\tilde{\phi}^{CLS}) &= \frac{1}{T^{1/2}}\left\{\sum_{t=1}^{T}X_{1t}X_{2t}-\sum_{t=1}^{T}X_{1t}\sum_{t=1}^{T}X_{2t}\right\} \\
&- \frac{\tilde{\alpha}_1^{YW}\tilde{\alpha}_2^{YW}}{T^{1/2}}\left\{\sum_{t=1}^{T}X_{1t}X_{2t}-\sum_{t=1}^{T}X_{1t}\sum_{t=1}^{T}X_{2t}\right\} \\
&- \frac{T^{1/2}}{T-1}\left\{\sum_{t=2}^{T}X_{1t}X_{2t}-\sum_{t=2}^{T}X_{1t}\sum_{t=2}^{T}X_{2t}\right\} \\
&- \frac{\tilde{\alpha}_1^{CLS}\tilde{\alpha}_2^{CLS}T^{1/2}}{T-1}\left\{\sum_{t=2}^{T}X_{1,t-1}X_{2,t-1}-\sum_{t=2}^{T}X_{1,t-1}\sum_{t=2}^{T}X_{2,t-1}\right\} \\
&+ \frac{\tilde{\alpha}_1^{CLS}T^{1/2}}{T-1}\left\{\sum_{t=2}^{T}X_{1,t-1}X_{2t}-\sum_{t=2}^{T}X_{1,t-1}\sum_{t=2}^{T}X_{2t}\right\} \\
&+ \frac{\tilde{\alpha}_2^{CLS}T^{1/2}}{T-1}\left\{\sum_{t=2}^{T}X_{1t}X_{2,t-1}-\sum_{t=2}^{T}X_{1t}\sum_{t=2}^{T}X_{2,t-1}\right\}
\end{aligned}$$

$$\Rightarrow T^{1/2}(\tilde{\phi}^{YW} - \tilde{\phi}^{CLS}) = o_p(1)$$

and since,

$$T^{1/2}(\tilde{\phi}^{YW} - \phi) - T^{1/2}(\tilde{\phi}^{CLS} - \phi) = T^{1/2}(\tilde{\phi}^{YW} - \tilde{\phi}^{CLS}) \rightarrow^p 0$$

it follows that both estimators have asymptotically the same distribution.

PROOF OF EQUATION (4.1).

$$E_t[\alpha_1 \circ X_{1t-1}] = \frac{1}{P(x_{1t}|X_{1t-1}, X_{2t-1})} \sum_{x_{2t}} \sum_{k=0}^{min(x_{1t}, x_{1t-1})} \sum_{s=0}^{min(x_{2t}, x_{2t-1})} k \begin{pmatrix} x_{1t-1} \\ k \end{pmatrix}$$

$$\times \alpha_1^k (1-\alpha_1)^{x_{1t-1}-k} \begin{pmatrix} x_{2t-1} \\ s \end{pmatrix} \alpha_2^s (1-\alpha_2)^{x_{2t-1}-s}$$

$$\times f(R_{1t} = x_{1t} - k, R_{2t} = x_{2t} - s)$$

$$= \frac{1}{P(x_{1t}|X_{1t-1}, X_{2t-1})} \sum_{x_{2t}} \sum_{k=1}^{min(x_{1t}, x_{1t-1})} \sum_{s=0}^{min(x_{2t}, x_{2t-1})} x_{1t-1} \begin{pmatrix} x_{1t-1} - 1 \\ k-1 \end{pmatrix}$$

$$\times \alpha_1^k (1-\alpha_1)^{x_{1t-1}-k} \begin{pmatrix} x_{2t-1} \\ s \end{pmatrix} \alpha_2^s (1-\alpha_2)^{x_{2t-1}-s}$$

$$\times f(R_{1t} = x_{1t} - k, R_{2t} = x_{2t} - s)$$

$$= \frac{\alpha_1 x_{1t-1}}{P(x_{1t}|X_{1t-1}, X_{2t-1})} \sum_{x_{2t}} \sum_{k=1}^{min(x_{1t}, x_{1t-1})} \sum_{s=0}^{min(x_{2t}, x_{2t-1})} \begin{pmatrix} x_{1t-1} - 1 \\ k-1 \end{pmatrix}$$

$$\times \alpha_1^{k-1} (1-\alpha_1)^{x_{1t-1}-1-(k-1)} \begin{pmatrix} x_{2t-1} \\ s \end{pmatrix} \alpha_2^s (1-\alpha_2)^{x_{2t-1}-s}$$

$$\times f(R_{1t} = x_{1t} - k, R_{2t} = x_{2t} - s)$$

$$= \frac{\alpha_1 x_{1t-1}}{P(x_{1t}|X_{1t-1}, X_{2t-1})} \sum_{x_{2t}} \sum_{k=0}^{min(x_{1t}-1, x_{1t-1}-1)} \sum_{s=0}^{min(x_{2t}, x_{2t-1})} \begin{pmatrix} x_{1t-1} - 1 \\ k \end{pmatrix}$$

$$\times\alpha_1^k(1-\alpha_1)^{x_{1t-1}-1-k}\begin{pmatrix} x_{2t-1} \\ s \end{pmatrix}\alpha_2^s(1-\alpha_2)^{x_{2t-1}-s}$$
$$\times f(R_{1t}=x_{1t}-1-k, R_{2t}=x_{2t}-s)$$

$$=\frac{\alpha_1 x_{1t-1} P(x_{1t}-1|X_{1t-1}-1, X_{2t-1})}{P(x_{1t}|X_{1t-1}, X_{2t-1})}$$

Equation 4.2 can be proved similarly.

PROOF OF EQUATION (5.6).

For the proof of eq.(5.6) we 'll make use of the following property of thinning operation:

$$\begin{aligned} & E[(\alpha\circ X)^2]=\alpha^2E(X^2)+\alpha(1-\alpha)E(X) \\ \Rightarrow\quad & Var[(\alpha\circ X)^2]=\alpha^2Var(X)+\alpha(1-\alpha)E(X) \end{aligned}$$

Thus,

$$\begin{aligned} Var\left(\sum_{i=0}^{h-1}\alpha_j^i\circ R_{j,t-i}\right) &= \sum_{i=0}^{h-1}(\alpha_j^2)^iVar(R_{jt})+\sum_{i=0}^{h-1}\alpha_j^i(1-\alpha_j^i)E(R_{jt}) \\ &= \left(\frac{1-\alpha_j^{2h}}{1-\alpha_j^2}\right)Var(R_{jt})+\left(\frac{1-\alpha_j^h}{1-\alpha_j}-\frac{1-\alpha_j^{2h}}{1-\alpha_j^2}\right)E(R_{jt}) \end{aligned}$$

ESTIMATION UNCERTAINTY OF THE POISSON BINAR(1) MODEL

For the Poisson BINAR(1) model, we let $\hat{\boldsymbol{\theta}}_T=(\hat\alpha_1,\hat\alpha_2,\hat\lambda_1,\hat\lambda_2,\hat\phi)$ be the ML estimators of $\boldsymbol{\theta}=(\alpha_1,\alpha_2,\lambda_1,\lambda_2,\phi)$ based on a sample of size T. Then, eq.(5.15) can be written as

$$
\begin{aligned}
\sigma_h^2(\mathbf{x};\boldsymbol{\theta}_0) &= T^{-1}\left\{\left(\frac{\partial P_h}{\partial \alpha_1}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right)^2 i_{1,1}^{-1} + \left(\frac{\partial P_h}{\partial \alpha_2}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right)^2 i_{2,2}^{-1} + \left(\frac{\partial P_h}{\partial \lambda_1}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right)^2 i_{3,3}^{-1}\right.\\
&+ \left(\frac{\partial P_h}{\partial \lambda_2}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right)^2 i_{4,4}^{-1} + \left(\frac{\partial P_h}{\partial \phi}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right)^2 i_{5,5}^{-1} + 2\left(\frac{\partial P_h}{\partial \alpha_1}\frac{\partial P_h}{\partial \alpha_2}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right) i_{1,2}^{-1}\\
&+ 2\left(\frac{\partial P_h}{\partial \alpha_1}\frac{\partial P_h}{\partial \lambda_1}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right) i_{1,3}^{-1} + 2\left(\frac{\partial P_h}{\partial \alpha_1}\frac{\partial P_h}{\partial \lambda_2}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right) i_{1,4}^{-1} + 2\left(\frac{\partial P_h}{\partial \alpha_1}\frac{\partial P_h}{\partial \phi}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right) i_{1,5}^{-1}\\
&+ 2\left(\frac{\partial P_h}{\partial \alpha_2}\frac{\partial P_h}{\partial \lambda_1}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right) i_{2,3}^{-1} + 2\left(\frac{\partial P_h}{\partial \alpha_2}\frac{\partial P_h}{\partial \lambda_2}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right) i_{2,4}^{-1} + 2\left(\frac{\partial P_h}{\partial \alpha_2}\frac{\partial P_h}{\partial \phi}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right) i_{2,5}^{-1}\\
&+ \left. 2\left(\frac{\partial P_h}{\partial \lambda_1}\frac{\partial P_h}{\partial \lambda_2}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right) i_{3,4}^{-1} + 2\left(\frac{\partial P_h}{\partial \lambda_1}\frac{\partial P_h}{\partial \phi}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right) i_{3,5}^{-1} + 2\left(\frac{\partial P_h}{\partial \lambda_2}\frac{\partial P_h}{\partial \phi}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right) i_{4,5}^{-1}\right\}
\end{aligned}
$$

where $i_{k,j}^{-1}$ is the k,j-element of the matrix i^{-1}, $k,j = 1,2,...,5$ and

$$
\begin{aligned}
\frac{\partial P_h}{\partial \alpha_1} &= \frac{h}{1-\alpha_1^h} x_{1T}\left\{P_h(x_1-1,x_2|x_{1T}-1,x_{2T}) - \alpha_1^{h-1}P_h(x_1,x_2|x_{1T},x_{2T})\right\}\\
&- \frac{\lambda_1(1-h\alpha_1^{h-1}-(1-h)\alpha_1^h)}{(1-\alpha_1^2)}\{P_h(x_1,x_2|x_{1T},x_{2T}) - P_h(x_1-1,x_2|x_{1T},x_{2T})\}\\
&+ \frac{\alpha_2\phi(1-h\alpha_1^{h-1}\alpha_2^{h-1}-(1-h)\alpha_1^h\alpha_2^h)}{(1-\alpha_1\alpha_2)^2}\{P_h(x_1,x_2|x_{1T},x_{2T})\\
&- P_h(x_1-1,x_2|x_{1T},x_{2T}) - P_h(x_1,x_2-1|x_{1T},x_{2T}) + P_h(x_1-1,x_2-1|x_{1T},x_{2T})\}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial P_h}{\partial \alpha_2} &= \frac{h}{1-\alpha_2^h} x_{2T}\left\{P_h(x_1,x_2-1|x_{1T},x_{2T}-1) - \alpha_2^{h-1}P_h(x_1,x_2|x_{1T},x_{2T})\right\}\\
&- \frac{\lambda_2(1-h\alpha_2^{h-1}-(1-h)\alpha_2^h)}{(1-\alpha_2^2)}\{P_h(x_1,x_2|x_{1T},x_{2T}) - P_h(x_1,x_2-1|x_{1T},x_{2T})\}\\
&+ \frac{\alpha_1\phi(1-h\alpha_1^{h-1}\alpha_2^{h-1}-(1-h)\alpha_1^h\alpha_2^h)}{(1-\alpha_1\alpha_2)^2}\{P_h(x_1,x_2|x_{1T},x_{2T})\\
&- P_h(x_1-1,x_2|x_{1T},x_{2T}) - P_h(x_1,x_2-1|x_{1T},x_{2T}) + P_h(x_1-1,x_2-1|x_{1T},x_{2T})\}
\end{aligned}
$$

$$
\frac{\partial P_h}{\partial \lambda_1} = \left(\frac{1-\alpha_1^h}{1-\alpha_1}\right)\{P_h(x_1,x_2|x_{1T},x_{2T}) + P_h(x_1-1,x_2|x_{1T},x_{2T})\}
$$

$$\frac{\partial P_h}{\partial \lambda_2} = \left(\frac{1-\alpha_2^h}{1-\alpha_2}\right)\{P_h(x_1,x_2|x_{1T},x_{2T}) + P_h(x_1,x_2-1|x_{1T},x_{2T})\}$$

$$\begin{aligned}\frac{\partial P_h}{\partial \phi} &= \left(\frac{1-\alpha_1^h\alpha_2^h}{1-\alpha_1\alpha_2}\right)\{P_h(x_1,x_2|x_{1T},x_{2T}) - P_h(x_1-1,x_2|x_{1T},x_{2T}) \\ &- P_h(x_1,x_2-1|x_{1T},x_{2T}) + P_h(x_1-1,x_2-1|x_{1T},x_{2T})\}\end{aligned}$$

In order to obtain analytical expressions for the elements that comprise the Fisher information matrix i^{-1} we follow the notation of Freeland and McCabe (2004b) and denote by $\ddot{\ell}_{\boldsymbol{\theta}}$ the second derivatives of the log-likelihood of the Poisson BINAR(1) model with respect to $\boldsymbol{\theta} = [\alpha_1, \alpha_2, \lambda_1, \lambda_2, \phi]'$:

$$\ddot{\ell}_{\boldsymbol{\theta}} = \begin{bmatrix} \ddot{\ell}_{\alpha_1\alpha_1} & \ddot{\ell}_{\alpha_1\alpha_2} & \ddot{\ell}_{\alpha_1\lambda_1} & \ddot{\ell}_{\alpha_1\lambda_2} & \ddot{\ell}_{\alpha_1\phi} \\ & \ddot{\ell}_{\alpha_2\alpha_2} & \ddot{\ell}_{\alpha_2\lambda_1} & \ddot{\ell}_{\alpha_2\lambda_2} & \ddot{\ell}_{\alpha_2\phi} \\ & & \ddot{\ell}_{\lambda_1\lambda_1} & \ddot{\ell}_{\lambda_1\lambda_2} & \ddot{\ell}_{\lambda_1\phi} \\ & & & \ddot{\ell}_{\lambda_2\lambda_2} & \ddot{\ell}_{\lambda_2\phi} \\ & & & & \ddot{\ell}_{\phi\phi} \end{bmatrix}$$

Through ordinary algebra it can be shown that

$$\begin{aligned}\ddot{\ell}_{\alpha_1\alpha_1} &= \frac{1}{(1-\alpha_1)^2}\sum_{t=1}^{T}\left\{\frac{2x_{1,t-1}P(x_{1t}-1,x_{2t}|x_{1,t-1}-1,x_{2,t-1})}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})} - x_{1,t-1}\right. \\ &+ \frac{x_{1,t-1}(x_{1,t-1}-1)P(x_{1t}-2,x_{2t}|x_{1,t-1}-2,x_{2,t-1})}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})} \\ &- \left.\left(\frac{x_{1,t-1}P(x_{1t}-1,x_{2t}|x_{1,t-1}-1,x_{2,t-1})}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\right)^2\right\}\end{aligned}$$

$$\begin{aligned}\ddot{\ell}_{\alpha_2\alpha_2} &= \frac{1}{(1-\alpha_2)^2}\sum_{t=1}^{T}\left\{\frac{2x_{2,t-1}P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1}-1)}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})} - x_{2,t-1}\right. \\ &+ \frac{x_{2,t-1}(x_{2,t-1}-1)P(x_{1t},x_{2t}-2|x_{1,t-1},x_{2,t-1}-2)}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})} \\ &- \left.\left(\frac{x_{2,t-1}P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1}-1)}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\right)^2\right\}\end{aligned}$$

$$\ddot{\ell}_{\lambda_1\lambda_1} = \sum_{t=1}^{T}\left\{\frac{P(x_{1t}-2,x_{2t}|x_{1,t-1},x_{2,t-1})}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})} - \left(\frac{P(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\right)^2\right\}$$

$$\ddot{\ell}_{\lambda_2\lambda_2} = \sum_{t=1}^{T}\left\{\frac{P(x_{1t},x_{2t}-2|x_{1,t-1},x_{2,t-1})}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})} - \left(\frac{P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\right)^2\right\}$$

$$\begin{aligned}
\ddot{\ell}_{\phi\phi} &= \sum_{t=1}^{T}\left\{\frac{1}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\times\right.\\
&\times \{2P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1}) - 2P(x_{1t}-2,x_{2t}-1|x_{1,t-1},x_{2,t-1})\\
&- 2P(x_{1t}-1,x_{2t}-2|x_{1,t-1},x_{2,t-1}) + P(x_{1t}-2,x_{2t}-2|x_{1,t-1},x_{2,t-1})\\
&+ P(x_{1t}-2,x_{2t}|x_{1,t-1},x_{2,t-1}) + P(x_{1t},x_{2t}-2|x_{1,t-1},x_{2,t-1})\}\\
&+ \frac{1}{P^2(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\times\\
&\times \{2P(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1})\\
&+ 2P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1})\\
&- 2P(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})\\
&- P^2(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1}) - P^2(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})\\
&\left.- P^2(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})\}\right\}
\end{aligned}$$

$$\begin{aligned}
\ddot{\ell}_{\alpha_1\alpha_2} &= \sum_{t=1}^{T}\left\{\frac{x_{1,t-1}x_{2,t-1}}{(1-\alpha_1)(1-\alpha_2)P^2(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\times\right.\\
&\times \{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})P(x_{1t}-1,x_{2t}-1|x_{1,t-1}-1,x_{2,t-1}-1)\\
&\left.- P(x_{1t}-1,x_{2t}|x_{1,t-1}-1,x_{2,t-1})\ P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1}-1)\}\right\}
\end{aligned}$$

$$\begin{aligned}
\ddot{\ell}_{\alpha_1\lambda_1} &= \sum_{t=1}^{T}\left\{\frac{x_{1,t-1}}{(1-\alpha_1)P^2(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\times\right.\\
&\times \{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})P(x_{1t}-2,x_{2t}|x_{1,t-1}-1,x_{2,t-1})\\
&\left.- P(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})P(x_{1t}-1,x_{2t}|x_{1,t-1}-1,x_{2,t-1})\}\right\}
\end{aligned}$$

$$\begin{aligned}
\ddot{\ell}_{\alpha_2\lambda_2} &= \sum_{t=1}^{T}\left\{\frac{x_{2,t-1}}{(1-\alpha_2)P^2(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\times\right.\\
&\times\ \{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})P(x_{1t},x_{2t}-2|x_{1,t-1},x_{2,t-1}-1)\\
&-\ \left.P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1}-1)\}\right\}
\end{aligned}$$

$$\begin{aligned}
\ddot{\ell}_{\alpha_1\lambda_2} &= \sum_{t=1}^{T}\left\{\frac{x_{1,t-1}}{(1-\alpha_1)P^2(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\times\right.\\
&\times\ \{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})P(x_{1t}-1,x_{2t}-1|x_{1,t-1}-1,x_{2,t-1})\\
&-\ \left.P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})P(x_{1t}-1,x_{2t}|x_{1,t-1}-1,x_{2,t-1})\}\right\}
\end{aligned}$$

$$\begin{aligned}
\ddot{\ell}_{\alpha_2\lambda_1} &= \sum_{t=1}^{T}\left\{\frac{x_{2,t-1}}{(1-\alpha_2)P^2(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\times\right.\\
&\times\ \{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1}-1)\\
&-\ \left.P(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1}-1)\}\right\}
\end{aligned}$$

$$\begin{aligned}
\ddot{\ell}_{\alpha_1\phi} &= \sum_{t=1}^{T}\left\{\frac{x_{1,t-1}}{(1-\alpha_1)}\left\{\frac{1}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\left[P(x_{1t}-2,x_{2t}-1|x_{1,t-1}-1,x_{2,t-1})\right.\right.\right.\\
&-\ \left.P(x_{1t}-2,x_{2t}|x_{1,t-1}-1,x_{2,t-1})-P(x_{1t}-1,x_{2t}-1|x_{1,t-1}-1,x_{2,t-1})\right]\\
&-\ \frac{P(x_{1t}-1,x_{2t}|x_{1,t-1}-1,x_{2,t-1})}{P^2(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\left[P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1})\right.\\
&-\ \left.\left.\left.P(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})-P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})\right]\right\}\right\}
\end{aligned}$$

$$
\begin{aligned}
\ddot{\ell}_{\alpha_2\phi} &= \sum_{t=1}^{T}\left\{\frac{x_{2,t-1}}{(1-\alpha_2)}\left\{\frac{1}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\left[P(x_{1t}-1,x_{2t}-2|x_{1,t-1},x_{2,t-1}-1)\right.\right.\right.\\
&- \left. P(x_{1t},x_{2t}-2|x_{1,t-1},x_{2,t-1}-1)-P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1}-1)\right]\\
&- \frac{P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1}-1)}{P^2(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\left[P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1})\right.\\
&- \left.\left.\left. P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})-P(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})\right]\right\}\right\}
\end{aligned}
$$

$$
\begin{aligned}
\ddot{\ell}_{\lambda_1\lambda_2} &= \sum_{t=1}^{T}\left\{\frac{1}{P^2(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\left\{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1})\right.\right.\\
&- \left.\left. P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})P(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})\right\}\right\}
\end{aligned}
$$

$$
\begin{aligned}
\ddot{\ell}_{\lambda_1\phi} &= \sum_{t=1}^{T}\left\{\frac{1}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\left\{P(x_{1t}-2,x_{2t}-1|x_{1,t-1},x_{2,t-1})\right.\right.\\
&- \left. P(x_{1t}-2,x_{2t}|x_{1,t-1},x_{2,t-1})-P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1})\right\}\\
&- \frac{P(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})}{P^2(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\left\{P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1})\right.\\
&- \left.\left. P(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})-P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})\right\}\right\}
\end{aligned}
$$

$$
\begin{aligned}
\ddot{\ell}_{\lambda_2\phi} &= \sum_{t=1}^{T}\left\{\frac{1}{P(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\left\{P(x_{1t}-1,x_{2t}-2|x_{1,t-1},x_{2,t-1})\right.\right.\\
&- \left. P(x_{1t},x_{2t}-2|x_{1,t-1},x_{2,t-1})-P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1})\right\}\\
&- \frac{P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})}{P^2(x_{1t},x_{2t}|x_{1,t-1},x_{2,t-1})}\left\{P(x_{1t}-1,x_{2t}-1|x_{1,t-1},x_{2,t-1})\right.\\
&- \left.\left. P(x_{1t}-1,x_{2t}|x_{1,t-1},x_{2,t-1})-P(x_{1t},x_{2t}-1|x_{1,t-1},x_{2,t-1})\right\}\right\}
\end{aligned}
$$

Note that, in contrast to the univariate case, the information (as well as the scores) of the Poisson BINAR(1) model cannot be decomposed into quantities associated with each component of the model seperately. This barrier is just due to the model's structure, i.e. to its bivariate nature.

The Fisher information matrix i can then be calculated as usual:

$$i = -E\left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}|\boldsymbol{\theta}\right] = -E\left[\ddot{\ell}_{\boldsymbol{\theta}}|\boldsymbol{\theta}\right]$$

where $\ell(\boldsymbol{\theta})$ is the log-likelihood of the Poisson BINAR(1) model.

Appendix II. Simulation Results

Table I. Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 50$.

		Bias			**SD ratios**		**% Extra Datasets**
$(\alpha_1, \alpha_2, \lambda_1^\star, \lambda_2^\star, \phi)$		MoM	YW	ML	MoM/ML	YW/ML	
(0.3, 0.3, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.034	-0.036	-0.015	1.212	1.005	19.00
	$\hat{\alpha}_2$	-0.023	-0.029	-0.009	1.327	1.065	
	$\hat{\lambda}_1^\star$	0.066	0.065	0.014	1.324	1.186	
	$\hat{\lambda}_2^\star$	0.026	0.035	-0.011	1.387	1.213	
	$\hat{\phi}$	0.003	0.007	0.022	1.065	1.087	
(0.3, 0.3, 1, 1, 1)	$\hat{\alpha}_1$	-0.043	-0.042	-0.019	1.234	1.038	23.60
	$\hat{\alpha}_2$	-0.031	-0.029	-0.007	1.225	1.039	
	$\hat{\lambda}_1^\star$	0.150	0.146	0.035	1.343	1.198	
	$\hat{\lambda}_2^\star$	0.124	0.116	0.008	1.361	1.210	
	$\hat{\phi}$	-0.055	-0.051	0.009	1.045	1.061	
(0.3, 0.3, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.029	-0.030	-0.011	1.197	1.002	39.40
	$\hat{\alpha}_2$	-0.025	-0.029	-0.012	1.231	0.996	
	$\hat{\lambda}_1^\star$	0.009	0.009	-0.049	1.122	1.039	
	$\hat{\lambda}_2^\star$	0.060	0.079	-0.006	1.254	1.056	
	$\hat{\phi}$	0.052	0.055	0.085	0.944	0.955	
(0.3, 0.3, 1, 3, 1)	$\hat{\alpha}_1$	-0.051	-0.041	-0.014	1.053	0.951	24.60
	$\hat{\alpha}_2$	-0.028	-0.031	-0.008	1.254	1.015	
	$\hat{\lambda}_1^\star$	0.168	0.140	0.031	1.095	1.041	
	$\hat{\lambda}_2^\star$	0.159	0.175	0.028	1.371	1.148	
	$\hat{\phi}$	-0.060	-0.058	-0.007	0.975	0.984	
(0.3, 0.3, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.025	-0.031	-0.009	1.257	0.982	42.40
	$\hat{\alpha}_2$	-0.017	-0.026	-0.009	1.267	0.999	
	$\hat{\lambda}_1^\star$	-0.160	-0.136	-0.192	1.225	1.043	
	$\hat{\lambda}_2^\star$	-0.222	-0.182	-0.210	1.242	1.072	
	$\hat{\phi}$	0.263	0.270	0.249	0.961	0.979	
(0.3, 0.3, 3, 3, 1)	$\hat{\alpha}_1$	-0.034	-0.038	-0.016	1.207	0.962	22.20
	$\hat{\alpha}_2$	-0.014	-0.027	-0.013	1.237	0.996	
	$\hat{\lambda}_1^\star$	0.010	0.019	-0.066	1.282	1.123	
	$\hat{\lambda}_2^\star$	-0.121	-0.056	-0.092	1.228	1.071	
	$\hat{\phi}$	0.155	0.167	0.158	1.010	1.031	

Table I (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 50$.

		Bias			**SD ratios**		**% Extra Datasets**
$(\alpha_1, \alpha_2, \lambda_1^\star, \lambda_2^\star, \phi)$		MoM	YW	ML	MoM/ML	YW/ML	
(0.3, 0.5, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.034	-0.035	-0.011	1.210	1.007	27.00
	$\hat{\alpha}_2$	-0.083	-0.067	-0.020	1.580	1.123	
	$\hat{\lambda}_1^\star$	0.041	0.041	-0.031	1.229	1.116	
	$\hat{\lambda}_2^\star$	0.179	0.127	-0.018	1.735	1.348	
	$\hat{\phi}$	0.023	0.024	0.060	0.964	0.980	
(0.3, 0.5, 1, 1, 1)	$\hat{\alpha}_1$	-0.040	-0.036	-0.007	1.162	0.974	29.20
	$\hat{\alpha}_2$	-0.093	-0.072	-0.021	1.543	1.153	
	$\hat{\lambda}_1^\star$	0.199	0.188	0.003	1.302	1.154	
	$\hat{\lambda}_2^\star$	0.403	0.318	0.043	1.856	1.476	
	$\hat{\phi}$	-0.096	-0.096	0.015	1.015	1.032	
(0.3, 0.5, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.046	-0.046	-0.028	1.179	0.986	44.00
	$\hat{\alpha}_2$	-0.037	-0.050	-0.015	1.860	1.193	
	$\hat{\lambda}_1^\star$	0.005	0.000	-0.008	1.034	0.982	
	$\hat{\lambda}_2^\star$	0.077	0.164	0.012	1.831	1.247	
	$\hat{\phi}$	0.090	0.095	0.078	0.960	0.971	
(0.3, 0.5, 1, 3, 1)	$\hat{\alpha}_1$	-0.036	-0.033	-0.008	1.138	0.943	36.40
	$\hat{\alpha}_2$	-0.071	-0.066	-0.016	1.738	1.150	
	$\hat{\lambda}_1^\star$	0.156	0.144	0.001	1.139	1.049	
	$\hat{\lambda}_2^\star$	0.517	0.480	0.090	1.742	1.248	
	$\hat{\phi}$	-0.067	-0.063	0.021	0.978	0.992	
(0.3, 0.5, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.028	-0.035	-0.015	1.261	0.994	45.60
	$\hat{\alpha}_2$	-0.055	-0.058	-0.019	1.609	1.086	
	$\hat{\lambda}_1^\star$	-0.240	-0.221	-0.196	1.206	1.048	
	$\hat{\lambda}_2^\star$	-0.071	-0.062	-0.163	1.584	1.188	
	$\hat{\phi}$	0.355	0.365	0.267	0.947	0.958	
(0.3, 0.5, 3, 3, 1)	$\hat{\alpha}_1$	-0.023	-0.032	-0.018	1.223	0.986	31.60
	$\hat{\alpha}_2$	-0.061	-0.064	-0.023	1.647	1.083	
	$\hat{\lambda}_1^\star$	-0.097	-0.061	-0.050	1.202	1.073	
	$\hat{\lambda}_2^\star$	0.187	0.197	0.019	1.559	1.158	
	$\hat{\phi}$	0.193	0.208	0.148	0.945	0.966	

Table I (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 50$.

		Bias			**SD ratios**		**% Extra Datasets**
$(\alpha_1,\ \alpha_2,\ \lambda_1^\star,\ \lambda_2^\star,\ \phi)$		MoM	YW	ML	MoM/ML	YW/ML	
(0.5, 0.3, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.082	-0.067	-0.019	1.633	1.155	26.40
	$\hat{\alpha}_2$	-0.041	-0.042	-0.017	1.210	1.000	
	$\hat{\lambda}_1^\star$	0.188	0.141	-0.005	1.689	1.296	
	$\hat{\lambda}_2^\star$	0.058	0.061	-0.016	1.215	1.100	
	$\hat{\phi}$	0.008	0.009	0.043	1.020	1.033	
(0.5, 0.3, 1, 1, 1)	$\hat{\alpha}_1$	-0.105	-0.081	-0.021	1.602	1.214	25.80
	$\hat{\alpha}_2$	-0.046	-0.043	-0.013	1.175	1.026	
	$\hat{\lambda}_1^\star$	0.447	0.356	0.046	1.813	1.481	
	$\hat{\lambda}_2^\star$	0.167	0.159	-0.017	1.362	1.245	
	$\hat{\phi}$	-0.069	-0.071	0.037	1.058	1.071	
(0.5, 0.3, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.096	-0.078	-0.026	1.572	1.155	50.00
	$\hat{\alpha}_2$	-0.020	-0.027	-0.007	1.224	0.981	
	$\hat{\lambda}_1^\star$	0.127	0.076	-0.066	1.326	1.098	
	$\hat{\lambda}_2^\star$	-0.019	0.017	-0.077	1.220	1.030	
	$\hat{\phi}$	0.088	0.088	0.108	0.911	0.912	
(0.5, 0.3, 1, 3, 1)	$\hat{\alpha}_1$	-0.108	-0.083	-0.027	1.510	1.110	36.60
	$\hat{\alpha}_2$	-0.041	-0.043	-0.019	1.253	0.983	
	$\hat{\lambda}_1^\star$	0.384	0.288	0.019	1.434	1.170	
	$\hat{\lambda}_2^\star$	0.216	0.228	0.039	1.328	1.109	
	$\hat{\phi}$	-0.011	-0.014	0.074	0.940	0.942	
(0.5, 0.3, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.064	-0.068	-0.025	1.731	1.151	44.20
	$\hat{\alpha}_2$	-0.033	-0.037	-0.017	1.173	0.949	
	$\hat{\lambda}_1^\star$	0.016	0.029	-0.102	1.633	1.210	
	$\hat{\lambda}_2^\star$	-0.232	-0.223	-0.196	1.137	1.019	
	$\hat{\phi}$	0.361	0.370	0.278	0.963	0.979	
(0.5, 0.3, 3, 3, 1)	$\hat{\alpha}_1$	-0.069	-0.064	-0.018	1.646	1.116	29.20
	$\hat{\alpha}_2$	-0.024	-0.028	-0.008	1.176	0.970	
	$\hat{\lambda}_1^\star$	0.278	0.229	0.008	1.586	1.213	
	$\hat{\lambda}_2^\star$	-0.120	-0.103	-0.121	1.157	1.045	
	$\hat{\phi}$	0.203	0.212	0.150	0.987	1.004	

Table I (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 50$.

		Bias			**SD ratios**		**% Extra Datasets**
(α_1, α_2, $\lambda_1^\star$, $\lambda_2^\star$, ϕ)		MoM	YW	ML	MoM/ML	YW/ML	
(0.5, 0.5, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.090	-0.073	-0.024	1.678	1.195	32.80
	$\hat{\alpha}_2$	-0.097	-0.076	-0.029	1.475	1.098	
	$\hat{\lambda}_1^\star$	0.171	0.122	-0.020	1.562	1.250	
	$\hat{\lambda}_2^\star$	0.204	0.143	0.002	1.555	1.248	
	$\hat{\phi}$	0.036	0.032	0.061	0.970	0.972	
(0.5, 0.5, 1, 1, 1)	$\hat{\alpha}_1$	-0.100	-0.078	-0.018	1.577	1.157	37.80
	$\hat{\alpha}_2$	-0.098	-0.072	-0.017	1.603	1.172	
	$\hat{\lambda}_1^\star$	0.427	0.351	0.032	1.671	1.339	
	$\hat{\lambda}_2^\star$	0.400	0.313	0.009	1.588	1.308	
	$\hat{\phi}$	-0.064	-0.077	0.049	1.010	1.011	
(0.5, 0.5, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.101	-0.079	-0.025	1.462	1.077	55.20
	$\hat{\alpha}_2$	-0.055	-0.053	-0.009	1.736	1.129	
	$\hat{\lambda}_1^\star$	0.111	0.046	-0.044	1.311	1.113	
	$\hat{\lambda}_2^\star$	0.148	0.144	-0.054	1.702	1.197	
	$\hat{\phi}$	0.160	0.157	0.120	0.905	0.903	
(0.5, 0.5, 1, 3, 1)	$\hat{\alpha}_1$	-0.111	-0.086	-0.029	1.476	1.084	41.20
	$\hat{\alpha}_2$	-0.078	-0.075	-0.029	1.662	1.103	
	$\hat{\lambda}_1^\star$	0.357	0.268	0.036	1.347	1.130	
	$\hat{\lambda}_2^\star$	0.518	0.499	0.168	1.658	1.194	
	$\hat{\phi}$	0.020	0.012	0.054	0.931	0.923	
(0.5, 0.5, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.060	-0.061	-0.018	1.694	1.136	52.80
	$\hat{\alpha}_2$	-0.058	-0.061	-0.022	1.667	1.085	
	$\hat{\lambda}_1^\star$	-0.113	-0.117	-0.138	1.616	1.221	
	$\hat{\lambda}_2^\star$	-0.115	-0.109	-0.106	1.511	1.138	
	$\hat{\phi}$	0.441	0.458	0.259	0.967	0.995	
(0.5, 0.5, 3, 3, 1)	$\hat{\alpha}_1$	-0.073	-0.070	-0.022	1.714	1.144	40.00
	$\hat{\alpha}_2$	-0.052	-0.060	-0.022	1.773	1.175	
	$\hat{\lambda}_1^\star$	0.137	0.090	-0.051	1.508	1.182	
	$\hat{\lambda}_2^\star$	-0.040	0.002	-0.057	1.647	1.277	
	$\hat{\phi}$	0.367	0.389	0.230	1.016	1.021	

Table II. Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 200$.

		Bias			**SD ratios**		**% Extra Datasets**
(α_1, α_2, $\lambda_1^\star$, $\lambda_2^\star$, ϕ)		MoM	YW	ML	MoM/ML	YW/ML	
(0.3, 0.3, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.014	-0.015	-0.009	1.397	1.134	0.00
	$\hat{\alpha}_2$	-0.008	-0.011	-0.007	1.298	1.081	
	$\hat{\lambda}_1^\star$	0.019	0.021	0.010	1.437	1.259	
	$\hat{\lambda}_2^\star$	0.008	0.012	0.007	1.440	1.261	
	$\hat{\phi}$	0.007	0.008	0.009	1.178	1.210	
(0.3, 0.3, 1, 1, 1)	$\hat{\alpha}_1$	-0.006	-0.009	-0.004	1.499	1.172	0.80
	$\hat{\alpha}_2$	-0.003	-0.006	-0.001	1.493	1.159	
	$\hat{\lambda}_1^\star$	0.011	0.017	0.005	1.796	1.541	
	$\hat{\lambda}_2^\star$	0.007	0.009	0.000	1.771	1.522	
	$\hat{\phi}$	0.004	0.009	0.008	1.333	1.396	
(0.3, 0.3, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.014	-0.015	-0.007	1.372	1.094	2.20
	$\hat{\alpha}_2$	-0.010	-0.012	-0.006	1.391	1.108	
	$\hat{\lambda}_1^\star$	0.040	0.041	0.020	1.359	1.227	
	$\hat{\lambda}_2^\star$	0.056	0.062	0.035	1.480	1.226	
	$\hat{\phi}$	-0.017	-0.015	-0.007	1.125	1.145	
(0.3, 0.3, 1, 3, 1)	$\hat{\alpha}_1$	-0.006	-0.007	-0.002	1.340	1.078	1.00
	$\hat{\alpha}_2$	-0.008	-0.009	-0.005	1.363	1.106	
	$\hat{\lambda}_1^\star$	0.029	0.029	0.011	1.521	1.353	
	$\hat{\lambda}_2^\star$	0.048	0.050	0.026	1.505	1.262	
	$\hat{\phi}$	-0.012	-0.009	-0.001	1.162	1.191	
(0.3, 0.3, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.007	-0.010	-0.005	1.459	1.129	9.80
	$\hat{\alpha}_2$	-0.005	-0.009	-0.005	1.392	1.073	
	$\hat{\lambda}_1^\star$	-0.046	-0.035	-0.046	1.417	1.204	
	$\hat{\lambda}_2^\star$	-0.054	-0.039	-0.042	1.347	1.150	
	$\hat{\phi}$	0.065	0.068	0.059	1.098	1.110	
(0.3, 0.3, 3, 3, 1)	$\hat{\alpha}_1$	-0.011	-0.011	-0.004	1.349	1.075	1.00
	$\hat{\alpha}_2$	-0.007	-0.009	-0.003	1.413	1.106	
	$\hat{\lambda}_1^\star$	0.049	0.049	0.007	1.368	1.186	
	$\hat{\lambda}_2^\star$	0.027	0.033	-0.002	1.432	1.236	
	$\hat{\phi}$	0.004	0.007	0.018	1.130	1.147	

Table II (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 200$.

		Bias			**SD ratios**		**% Extra Datasets**
$(\alpha_1, \alpha_2, \lambda_1^\star, \lambda_2^\star, \phi)$		MoM	YW	ML	MoM/ML	YW/ML	
(0.3, 0.5, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.004	-0.007	-0.004	1.395	1.125	1.00
	$\hat{\alpha}_2$	-0.015	-0.014	-0.003	2.122	1.266	
	$\hat{\lambda}_1^\star$	0.009	0.012	0.003	1.496	1.338	
	$\hat{\lambda}_2^\star$	0.035	0.028	-0.001	2.161	1.513	
	$\hat{\phi}$	-0.001	0.003	0.008	1.210	1.259	
(0.3, 0.5, 1, 1, 1)	$\hat{\alpha}_1$	-0.012	-0.013	-0.008	1.488	1.175	4.80
	$\hat{\alpha}_2$	-0.022	-0.017	-0.005	1.990	1.241	
	$\hat{\lambda}_1^\star$	0.034	0.034	0.009	1.648	1.470	
	$\hat{\lambda}_2^\star$	0.085	0.059	0.007	2.215	1.587	
	$\hat{\phi}$	-0.006	-0.003	0.013	1.209	1.270	
(0.3, 0.5, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.013	-0.014	-0.008	1.361	1.082	3.40
	$\hat{\alpha}_2$	-0.003	-0.012	-0.004	2.368	1.291	
	$\hat{\lambda}_1^\star$	0.004	0.000	0.005	1.320	1.248	
	$\hat{\lambda}_2^\star$	-0.021	0.034	0.010	2.282	1.389	
	$\hat{\phi}$	0.021	0.026	0.011	1.144	1.177	
(0.3, 0.5, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.015	-0.014	-0.007	1.292	1.075	1.20
	$\hat{\alpha}_2$	-0.013	-0.017	-0.007	2.345	1.323	
	$\hat{\lambda}_1^\star$	0.041	0.033	0.019	1.313	1.227	
	$\hat{\lambda}_2^\star$	0.065	0.089	0.031	2.337	1.464	
	$\hat{\phi}$	0.001	0.004	0.005	1.161	1.193	
(0.3, 0.5, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.006	-0.008	-0.001	1.379	1.074	9.20
	$\hat{\alpha}_2$	-0.008	-0.014	-0.005	2.132	1.215	
	$\hat{\lambda}_1^\star$	-0.068	-0.064	-0.064	1.338	1.170	
	$\hat{\lambda}_2^\star$	-0.060	-0.027	-0.041	1.871	1.261	
	$\hat{\phi}$	0.084	0.088	0.059	1.142	1.158	
(0.3, 0.5, 3, 3, 1)	$\hat{\alpha}_1$	-0.012	-0.014	-0.009	1.440	1.104	1.60
	$\hat{\alpha}_2$	-0.017	-0.020	-0.008	2.290	1.291	
	$\hat{\lambda}_1^\star$	0.046	0.051	0.054	1.412	1.243	
	$\hat{\lambda}_2^\star$	0.085	0.101	0.054	2.100	1.401	
	$\hat{\phi}$	0.028	0.034	0.010	1.133	1.163	

Table II (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 200$.

		Bias			**SD ratios**		**% Extra Datasets**
(α_1, α_2, $\lambda_1^\star$, $\lambda_2^\star$, ϕ)		MoM	YW	ML	MoM/ML	YW/ML	
(0.5, 0.3, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.012	-0.014	-0.005	2.253	1.275	1.20
	$\hat{\alpha}_2$	0.000	-0.004	-0.002	1.391	1.094	
	$\hat{\lambda}_1^\star$	0.028	0.031	0.021	2.102	1.448	
	$\hat{\lambda}_2^\star$	0.002	0.006	0.015	1.440	1.271	
	$\hat{\phi}$	-0.003	0.001	-0.007	1.173	1.211	
(0.5, 0.3, 1, 1, 1)	$\hat{\alpha}_1$	-0.012	-0.011	-0.002	2.128	1.251	3.60
	$\hat{\alpha}_2$	-0.011	-0.012	-0.007	1.352	1.112	
	$\hat{\lambda}_1^\star$	0.063	0.058	0.016	2.421	1.687	
	$\hat{\lambda}_2^\star$	0.054	0.052	0.029	1.576	1.426	
	$\hat{\phi}$	-0.032	-0.028	-0.014	1.286	1.350	
(0.5, 0.3, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.015	-0.014	-0.003	2.174	1.283	4.00
	$\hat{\alpha}_2$	-0.016	-0.017	-0.011	1.266	1.052	
	$\hat{\lambda}_1^\star$	0.025	0.019	-0.004	1.820	1.371	
	$\hat{\lambda}_2^\star$	0.058	0.059	0.038	1.291	1.130	
	$\hat{\phi}$	0.018	0.020	0.019	1.173	1.190	
(0.5, 0.3, 1, 3, 1)	$\hat{\alpha}_1$	-0.020	-0.019	-0.009	1.959	1.174	6.20
	$\hat{\alpha}_2$	-0.013	-0.016	-0.009	1.399	1.115	
	$\hat{\lambda}_1^\star$	0.041	0.032	0.004	1.866	1.419	
	$\hat{\lambda}_2^\star$	0.026	0.037	0.011	1.544	1.313	
	$\hat{\phi}$	0.031	0.036	0.034	1.228	1.264	
(0.5, 0.3, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.011	-0.012	0.000	2.212	1.232	13.40
	$\hat{\alpha}_2$	-0.008	-0.011	-0.006	1.414	1.115	
	$\hat{\lambda}_1^\star$	-0.028	-0.028	-0.053	1.956	1.299	
	$\hat{\lambda}_2^\star$	-0.066	-0.054	-0.032	1.360	1.187	
	$\hat{\phi}$	0.095	0.100	0.059	1.128	1.153	
(0.5, 0.3, 3, 3, 1)	$\hat{\alpha}_1$	-0.001	-0.010	-0.002	2.494	1.320	0.60
	$\hat{\alpha}_2$	-0.009	-0.011	-0.005	1.287	1.049	
	$\hat{\lambda}_1^\star$	-0.027	0.033	0.023	2.174	1.434	
	$\hat{\lambda}_2^\star$	0.036	0.035	0.039	1.428	1.272	
	$\hat{\phi}$	0.005	0.015	-0.014	1.211	1.252	

Table II (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 200$.

		Bias			**SD ratios**		**% Extra Datasets**
(α_1, α_2, $\lambda_1^\star$, $\lambda_2^\star$, ϕ)		MoM	YW	ML	MoM/ML	YW/ML	
(0.5, 0.5, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.013	-0.012	-0.003	2.093	1.247	3.00
	$\hat{\alpha}_2$	-0.013	-0.014	-0.006	2.235	1.287	
	$\hat{\lambda}_1^\star$	0.022	0.014	-0.002	2.037	1.512	
	$\hat{\lambda}_2^\star$	0.021	0.020	0.007	1.997	1.470	
	$\hat{\phi}$	0.012	0.019	0.015	1.264	1.319	
(0.5, 0.5, 1, 1, 1)	$\hat{\alpha}_1$	-0.016	-0.014	-0.002	2.138	1.243	8.80
	$\hat{\alpha}_2$	-0.023	-0.019	-0.008	2.092	1.272	
	$\hat{\lambda}_1^\star$	0.069	0.053	0.004	2.270	1.662	
	$\hat{\lambda}_2^\star$	0.091	0.070	0.023	2.223	1.678	
	$\hat{\phi}$	-0.012	-0.005	0.006	1.192	1.289	
(0.5, 0.5, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.020	-0.018	-0.006	2.114	1.231	4.60
	$\hat{\alpha}_2$	-0.012	-0.014	-0.004	2.279	1.300	
	$\hat{\lambda}_1^\star$	0.000	-0.010	-0.008	1.616	1.268	
	$\hat{\lambda}_2^\star$	0.018	0.031	0.008	2.219	1.428	
	$\hat{\phi}$	0.044	0.049	0.021	1.182	1.211	
(0.5, 0.5, 1, 3, 1)	$\hat{\alpha}_1$	-0.022	-0.020	-0.007	2.032	1.238	8.20
	$\hat{\alpha}_2$	-0.007	-0.011	-0.001	2.394	1.308	
	$\hat{\lambda}_1^\star$	0.069	0.055	0.004	1.739	1.408	
	$\hat{\lambda}_2^\star$	0.028	0.053	-0.015	2.256	1.448	
	$\hat{\phi}$	-0.005	0.003	0.012	1.197	1.229	
(0.5, 0.5, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.005	-0.010	-0.002	2.016	1.188	15.40
	$\hat{\alpha}_2$	-0.002	-0.009	-0.001	2.495	1.341	
	$\hat{\lambda}_1^\star$	-0.130	-0.103	-0.052	1.824	1.278	
	$\hat{\lambda}_2^\star$	-0.147	-0.105	-0.059	1.914	1.275	
	$\hat{\phi}$	0.141	0.150	0.061	1.146	1.178	
(0.5, 0.5, 3, 3, 1)	$\hat{\alpha}_1$	-0.014	-0.017	-0.005	2.270	1.263	3.40
	$\hat{\alpha}_2$	-0.012	-0.015	-0.003	2.179	1.263	
	$\hat{\lambda}_1^\star$	0.025	0.037	0.020	1.995	1.396	
	$\hat{\lambda}_2^\star$	0.017	0.027	0.014	1.942	1.391	
	$\hat{\phi}$	0.058	0.069	0.012	1.198	1.240	

Table III. Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 500$.

		Bias			**SD ratios**		**% Extra Datasets**
$(\alpha_1, \alpha_2, \lambda_1^\star, \lambda_2^\star, \phi)$		MoM	YW	ML	MoM/ML	YW/ML	
(0.3, 0.3, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.001	-0.002	0.000	1.468	1.169	0.00
	$\hat{\alpha}_2$	-0.004	-0.005	-0.004	1.413	1.132	
	$\hat{\lambda}_1^\star$	0.005	0.006	0.001	1.577	1.351	
	$\hat{\lambda}_2^\star$	0.001	0.003	0.001	1.438	1.258	
	$\hat{\phi}$	-0.001	0.000	0.001	1.210	1.237	
(0.3, 0.3, 1, 1, 1)	$\hat{\alpha}_1$	-0.001	-0.002	0.000	1.509	1.202	0.00
	$\hat{\alpha}_2$	-0.003	-0.004	-0.002	1.467	1.189	
	$\hat{\lambda}_1^\star$	-0.001	0.001	-0.003	1.823	1.574	
	$\hat{\lambda}_2^\star$	0.007	0.007	0.001	1.788	1.546	
	$\hat{\phi}$	0.005	0.007	0.008	1.308	1.369	
(0.3, 0.3, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.008	-0.008	-0.005	1.387	1.132	0.20
	$\hat{\alpha}_2$	-0.002	-0.003	-0.001	1.375	1.100	
	$\hat{\lambda}_1^\star$	0.025	0.024	0.016	1.402	1.282	
	$\hat{\lambda}_2^\star$	0.023	0.026	0.015	1.480	1.229	
	$\hat{\phi}$	-0.014	-0.014	-0.009	1.176	1.188	
(0.3, 0.3, 1, 3, 1)	$\hat{\alpha}_1$	-0.007	-0.008	-0.006	1.342	1.100	0.00
	$\hat{\alpha}_2$	-0.002	-0.003	-0.002	1.414	1.122	
	$\hat{\lambda}_1^\star$	0.015	0.016	0.016	1.493	1.362	
	$\hat{\lambda}_2^\star$	0.005	0.010	0.011	1.513	1.266	
	$\hat{\phi}$	0.000	0.001	-0.002	1.179	1.206	
(0.3, 0.3, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.004	-0.005	-0.003	1.348	1.084	1.80
	$\hat{\alpha}_2$	0.002	0.001	0.003	1.424	1.107	
	$\hat{\lambda}_1^\star$	0.010	0.015	0.005	1.267	1.120	
	$\hat{\lambda}_2^\star$	-0.022	-0.017	-0.031	1.330	1.150	
	$\hat{\phi}$	0.015	0.016	0.019	1.081	1.087	
(0.3, 0.3, 3, 3, 1)	$\hat{\alpha}_1$	0.000	-0.001	0.001	1.359	1.100	0.00
	$\hat{\alpha}_2$	-0.007	-0.007	-0.004	1.448	1.124	
	$\hat{\lambda}_1^\star$	0.002	0.008	0.003	1.392	1.206	
	$\hat{\lambda}_2^\star$	0.044	0.044	0.034	1.421	1.226	
	$\hat{\phi}$	-0.005	-0.003	-0.007	1.098	1.111	

Table III (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 500$.

		Bias			**SD ratios**		**% Extra Datasets**
(α_1, α_2, $\lambda_1^\star$, $\lambda_2^\star$, ϕ)		MoM	YW	ML	MoM/ML	YW/ML	
(0.3, 0.5, 1, 1, 0.5)	$\hat{\alpha}_1$	0.003	0.001	0.001	1.447	1.153	0.00
	$\hat{\alpha}_2$	-0.003	-0.004	-0.001	2.398	1.343	
	$\hat{\lambda}_1^\star$	-0.014	-0.012	-0.008	1.529	1.380	
	$\hat{\lambda}_2^\star$	-0.003	0.001	-0.003	2.307	1.545	
	$\hat{\phi}$	0.004	0.007	0.004	1.230	1.278	
(0.3, 0.5, 1, 1, 1)	$\hat{\alpha}_1$	-0.004	-0.004	-0.001	1.373	1.122	0.00
	$\hat{\alpha}_2$	-0.013	-0.009	-0.003	2.335	1.331	
	$\hat{\lambda}_1^\star$	0.013	0.011	-0.004	1.654	1.493	
	$\hat{\lambda}_2^\star$	0.048	0.033	0.002	2.551	1.769	
	$\hat{\phi}$	0.001	0.001	0.009	1.291	1.371	
(0.3, 0.5, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.005	-0.005	-0.003	1.397	1.121	0.40
	$\hat{\alpha}_2$	-0.008	-0.007	-0.001	2.266	1.289	
	$\hat{\lambda}_1^\star$	0.008	0.008	0.004	1.335	1.241	
	$\hat{\lambda}_2^\star$	0.043	0.032	-0.002	2.169	1.387	
	$\hat{\phi}$	0.001	0.001	0.002	1.199	1.214	
(0.3, 0.5, 1, 3, 1)	$\hat{\alpha}_1$	-0.002	-0.004	-0.002	1.491	1.162	0.00
	$\hat{\alpha}_2$	-0.004	-0.006	-0.002	2.504	1.362	
	$\hat{\lambda}_1^\star$	0.002	0.002	-0.004	1.631	1.505	
	$\hat{\lambda}_2^\star$	0.021	0.037	0.014	2.475	1.556	
	$\hat{\phi}$	0.000	0.005	0.005	1.264	1.323	
(0.3, 0.5, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.001	-0.002	0.000	1.407	1.097	2.20
	$\hat{\alpha}_2$	-0.006	-0.006	-0.002	2.260	1.270	
	$\hat{\lambda}_1^\star$	-0.020	-0.016	-0.010	1.367	1.206	
	$\hat{\lambda}_2^\star$	0.022	0.021	0.009	2.040	1.356	
	$\hat{\phi}$	0.018	0.020	0.006	1.184	1.203	
(0.3, 0.5, 3, 3, 1)	$\hat{\alpha}_1$	-0.008	-0.008	-0.005	1.426	1.126	0.20
	$\hat{\alpha}_2$	-0.002	-0.005	-0.002	2.393	1.318	
	$\hat{\lambda}_1^\star$	0.062	0.059	0.034	1.388	1.236	
	$\hat{\lambda}_2^\star$	0.032	0.057	0.032	2.088	1.375	
	$\hat{\phi}$	-0.013	-0.010	-0.003	1.177	1.196	

Table III (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 500$.

		Bias			**SD ratios**		**% Extra Datasets**
$(\alpha_1, \alpha_2, \lambda_1^\star, \lambda_2^\star, \phi)$		MoM	YW	ML	MoM/ML	YW/ML	
(0.5, 0.3, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.004	-0.006	-0.003	2.260	1.302	0.00
	$\hat{\alpha}_2$	-0.006	-0.005	-0.002	1.368	1.109	
	$\hat{\lambda}_1^\star$	0.002	0.008	0.000	2.196	1.513	
	$\hat{\lambda}_2^\star$	0.008	0.004	-0.003	1.520	1.377	
	$\hat{\phi}$	0.002	0.003	0.005	1.287	1.323	
(0.5, 0.3, 1, 1, 1)	$\hat{\alpha}_1$	-0.006	-0.005	0.000	2.314	1.278	0.40
	$\hat{\alpha}_2$	-0.004	-0.004	-0.001	1.468	1.162	
	$\hat{\lambda}_1^\star$	0.029	0.024	0.003	2.531	1.702	
	$\hat{\lambda}_2^\star$	0.012	0.011	0.000	1.667	1.475	
	$\hat{\phi}$	-0.005	-0.004	0.001	1.268	1.336	
(0.5, 0.3, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.003	-0.005	-0.002	2.303	1.265	0.60
	$\hat{\alpha}_2$	-0.003	-0.004	-0.002	1.423	1.108	
	$\hat{\lambda}_1^\star$	0.002	0.007	0.009	1.873	1.396	
	$\hat{\lambda}_2^\star$	0.012	0.017	0.016	1.447	1.205	
	$\hat{\phi}$	0.002	0.004	-0.004	1.235	1.257	
(0.5, 0.3, 1, 3, 1)	$\hat{\alpha}_1$	-0.011	-0.008	-0.002	2.318	1.307	1.20
	$\hat{\alpha}_2$	0.000	-0.002	0.000	1.428	1.116	
	$\hat{\lambda}_1^\star$	0.035	0.020	0.000	2.111	1.548	
	$\hat{\lambda}_2^\star$	-0.003	0.005	0.002	1.485	1.280	
	$\hat{\phi}$	-0.001	0.001	0.001	1.266	1.313	
(0.5, 0.3, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.003	-0.006	-0.003	2.103	1.210	2.60
	$\hat{\alpha}_2$	-0.005	-0.006	-0.004	1.389	1.091	
	$\hat{\lambda}_1^\star$	-0.018	-0.003	-0.005	1.821	1.262	
	$\hat{\lambda}_2^\star$	-0.004	-0.002	0.000	1.321	1.166	
	$\hat{\phi}$	0.030	0.032	0.020	1.184	1.195	
(0.5, 0.3, 3, 3, 1)	$\hat{\alpha}_1$	-0.009	-0.008	-0.002	2.268	1.294	0.00
	$\hat{\alpha}_2$	-0.006	-0.006	-0.004	1.367	1.110	
	$\hat{\lambda}_1^\star$	0.070	0.055	0.007	2.030	1.386	
	$\hat{\lambda}_2^\star$	0.023	0.025	0.008	1.374	1.249	
	$\hat{\phi}$	0.001	0.002	0.010	1.193	1.212	

Table III (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 500$.

		Bias			**SD ratios**		**% Extra Datasets**
(α_1, α_2, $\lambda_1^\star$, $\lambda_2^\star$, ϕ)		MoM	YW	ML	MoM/ML	YW/ML	
(0.5, 0.5, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.003	-0.004	-0.001	2.324	1.291	0.00
	$\hat{\alpha}_2$	-0.007	-0.006	-0.001	2.314	1.324	
	$\hat{\lambda}_1^\star$	0.015	0.014	0.006	2.164	1.539	
	$\hat{\lambda}_2^\star$	0.024	0.020	0.008	2.145	1.542	
	$\hat{\phi}$	-0.014	-0.011	-0.010	1.280	1.353	
(0.5, 0.5, 1, 1, 1)	$\hat{\alpha}_1$	-0.009	-0.006	0.000	2.389	1.341	0.20
	$\hat{\alpha}_2$	-0.008	-0.006	0.000	2.268	1.315	
	$\hat{\lambda}_1^\star$	0.028	0.016	-0.005	2.435	1.738	
	$\hat{\lambda}_2^\star$	0.030	0.019	0.001	2.502	1.824	
	$\hat{\phi}$	-0.002	-0.001	0.000	1.263	1.380	
(0.5, 0.5, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.007	-0.007	-0.003	2.297	1.315	0.20
	$\hat{\alpha}_2$	-0.004	-0.007	-0.004	2.450	1.347	
	$\hat{\lambda}_1^\star$	0.027	0.023	0.011	1.807	1.439	
	$\hat{\lambda}_2^\star$	0.016	0.036	0.020	2.299	1.469	
	$\hat{\phi}$	0.002	0.005	0.006	1.263	1.296	
(0.5, 0.5, 1, 3, 1)	$\hat{\alpha}_1$	-0.004	-0.004	-0.001	2.326	1.290	1.20
	$\hat{\alpha}_2$	-0.005	-0.006	-0.001	2.428	1.344	
	$\hat{\lambda}_1^\star$	0.003	-0.001	0.006	2.032	1.609	
	$\hat{\lambda}_2^\star$	0.019	0.023	0.006	2.383	1.595	
	$\hat{\phi}$	0.009	0.014	-0.003	1.296	1.351	
(0.5, 0.5, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.005	-0.006	-0.002	2.418	1.321	3.20
	$\hat{\alpha}_2$	0.001	-0.003	0.000	2.265	1.246	
	$\hat{\lambda}_1^\star$	-0.007	-0.004	0.006	2.051	1.400	
	$\hat{\lambda}_2^\star$	-0.051	-0.027	-0.005	1.904	1.314	
	$\hat{\phi}$	0.036	0.041	0.007	1.238	1.271	
(0.5, 0.5, 3, 3, 1)	$\hat{\alpha}_1$	-0.005	-0.007	-0.003	2.237	1.251	0.20
	$\hat{\alpha}_2$	-0.004	-0.004	0.000	2.309	1.260	
	$\hat{\lambda}_1^\star$	0.018	0.028	0.036	2.027	1.474	
	$\hat{\lambda}_2^\star$	0.005	0.002	0.006	2.032	1.457	
	$\hat{\phi}$	0.016	0.022	-0.008	1.341	1.373	

Table IV. Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 1000$.

		Bias			**SD ratios**		**% Extra Datasets**
$(\alpha_1,\ \alpha_2,\ \lambda_1^\star,\ \lambda_2^\star,\ \phi)$		MoM	YW	ML	MoM/ML	YW/ML	
(0.3, 0.3, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.003	-0.004	-0.003	1.393	1.135	0.00
	$\hat{\alpha}_2$	-0.002	-0.001	0.001	1.452	1.137	
	$\hat{\lambda}_1^\star$	0.007	0.008	0.004	1.510	1.305	
	$\hat{\lambda}_2^\star$	0.002	0.000	-0.006	1.541	1.315	
	$\hat{\phi}$	-0.001	-0.001	0.002	1.187	1.212	
(0.3, 0.3, 1, 1, 1)	$\hat{\alpha}_1$	-0.003	-0.003	-0.001	1.538	1.228	0.00
	$\hat{\alpha}_2$	-0.003	-0.002	-0.001	1.521	1.190	
	$\hat{\lambda}_1^\star$	0.014	0.015	0.005	1.826	1.578	
	$\hat{\lambda}_2^\star$	0.012	0.011	0.003	1.799	1.537	
	$\hat{\phi}$	-0.006	-0.005	0.000	1.321	1.378	
(0.3, 0.3, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.001	-0.001	0.000	1.351	1.097	0.00
	$\hat{\alpha}_2$	-0.003	-0.003	-0.001	1.406	1.102	
	$\hat{\lambda}_1^\star$	-0.001	0.000	-0.005	1.335	1.215	
	$\hat{\lambda}_2^\star$	0.012	0.010	0.000	1.397	1.165	
	$\hat{\phi}$	0.001	0.002	0.005	1.147	1.158	
(0.3, 0.3, 1, 3, 1)	$\hat{\alpha}_1$	-0.001	-0.002	-0.001	1.432	1.137	0.00
	$\hat{\alpha}_2$	-0.002	-0.004	-0.003	1.429	1.136	
	$\hat{\lambda}_1^\star$	-0.001	-0.001	0.004	1.478	1.322	
	$\hat{\lambda}_2^\star$	-0.004	0.005	0.008	1.582	1.299	
	$\hat{\phi}$	0.007	0.008	0.003	1.176	1.206	
(0.3, 0.3, 3, 3, 0.5)	$\hat{\alpha}_1$	0.001	0.000	0.000	1.440	1.122	0.40
	$\hat{\alpha}_2$	-0.004	-0.004	-0.002	1.393	1.088	
	$\hat{\lambda}_1^\star$	-0.006	-0.002	-0.005	1.358	1.168	
	$\hat{\lambda}_2^\star$	0.015	0.014	0.005	1.309	1.135	
	$\hat{\phi}$	0.002	0.003	0.003	1.118	1.122	
(0.3, 0.3, 3, 3, 1)	$\hat{\alpha}_1$	-0.003	-0.003	-0.001	1.406	1.102	0.00
	$\hat{\alpha}_2$	-0.001	-0.001	0.000	1.378	1.096	
	$\hat{\lambda}_1^\star$	0.014	0.012	0.003	1.370	1.184	
	$\hat{\lambda}_2^\star$	0.002	0.003	-0.005	1.375	1.182	
	$\hat{\phi}$	-0.004	-0.003	-0.001	1.131	1.146	

Table IV (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 1000$.

		Bias			**SD ratios**		**% Extra Datasets**
(α_1, α_2, $\lambda_1^\star$, $\lambda_2^\star$, ϕ)		MoM	YW	ML	MoM/ML	YW/ML	
(0.3, 0.5, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.001	-0.001	-0.001	1.432	1.144	0.00
	$\hat{\alpha}_2$	0.000	-0.002	-0.001	2.286	1.256	
	$\hat{\lambda}_1^\star$	-0.002	-0.003	0.000	1.499	1.336	
	$\hat{\lambda}_2^\star$	-0.002	0.002	0.004	2.233	1.499	
	$\hat{\phi}$	0.003	0.004	0.000	1.236	1.271	
(0.3, 0.5, 1, 1, 1)	$\hat{\alpha}_1$	-0.003	-0.003	-0.002	1.526	1.214	0.00
	$\hat{\alpha}_2$	0.000	-0.002	0.000	2.451	1.354	
	$\hat{\lambda}_1^\star$	0.012	0.011	0.008	1.630	1.489	
	$\hat{\lambda}_2^\star$	0.007	0.011	0.005	2.544	1.719	
	$\hat{\phi}$	-0.005	-0.003	-0.001	1.235	1.309	
(0.3, 0.5, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.001	-0.002	-0.001	1.346	1.103	0.00
	$\hat{\alpha}_2$	-0.001	-0.004	-0.003	2.393	1.283	
	$\hat{\lambda}_1^\star$	0.002	0.003	0.004	1.310	1.241	
	$\hat{\lambda}_2^\star$	0.006	0.022	0.021	2.244	1.374	
	$\hat{\phi}$	-0.005	-0.004	-0.006	1.213	1.222	
(0.3, 0.5, 1, 3, 1)	$\hat{\alpha}_1$	0.003	0.001	0.000	1.557	1.207	0.00
	$\hat{\alpha}_2$	0.001	-0.001	0.000	2.483	1.375	
	$\hat{\lambda}_1^\star$	-0.014	-0.012	-0.002	1.577	1.447	
	$\hat{\lambda}_2^\star$	-0.024	-0.006	-0.005	2.500	1.547	
	$\hat{\phi}$	0.007	0.010	0.004	1.212	1.258	
(0.3, 0.5, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.001	-0.002	-0.001	1.394	1.111	0.40
	$\hat{\alpha}_2$	-0.005	-0.004	-0.001	2.144	1.206	
	$\hat{\lambda}_1^\star$	-0.008	-0.004	0.007	1.355	1.232	
	$\hat{\lambda}_2^\star$	0.014	0.008	0.004	2.004	1.382	
	$\hat{\phi}$	0.015	0.015	0.003	1.243	1.251	
(0.3, 0.5, 3, 3, 1)	$\hat{\alpha}_1$	0.00	-0.001	-0.001	1.380	1.113	0.00
	$\hat{\alpha}_2$	-0.002	-0.003	-0.001	2.430	1.334	
	$\hat{\lambda}_1^\star$	0.004	0.008	0.001	1.421	1.291	
	$\hat{\lambda}_2^\star$	0.017	0.026	0.008	2.222	1.479	
	$\hat{\phi}$	-0.007	-0.005	0.000	1.221	1.249	

Table IV (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 1000$.

		Bias			**SD ratios**		**% Extra Datasets**
$(\alpha_1,\ \alpha_2,\ \lambda_1^\star,\ \lambda_2^\star,\ \phi)$		MoM	YW	ML	MoM/ML	YW/ML	
(0.5, 0.3, 1, 1, 0.5)	$\hat{\alpha}_1$	0.003	0.000	0.001	2.322	1.307	0.00
	$\hat{\alpha}_2$	-0.003	-0.004	-0.002	1.430	1.141	
	$\hat{\lambda}_1^\star$	-0.005	0.001	-0.002	2.258	1.499	
	$\hat{\lambda}_2^\star$	0.007	0.007	0.002	1.425	1.277	
	$\hat{\phi}$	-0.003	-0.002	0.000	1.183	1.217	
(0.5, 0.3, 1, 1, 1)	$\hat{\alpha}_1$	-0.002	-0.003	-0.001	2.518	1.386	0.00
	$\hat{\alpha}_2$	-0.002	-0.002	0.000	1.527	1.178	
	$\hat{\lambda}_1^\star$	0.012	0.015	0.008	2.750	1.845	
	$\hat{\lambda}_2^\star$	0.016	0.013	0.008	1.781	1.579	
	$\hat{\phi}$	-0.006	-0.005	-0.003	1.276	1.367	
(0.5, 0.3, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.004	-0.003	-0.001	2.333	1.341	
	$\hat{\alpha}_2$	0.002	0.000	0.001	1.477	1.160	
	$\hat{\lambda}_1^\star$	0.009	0.007	0.001	1.782	1.369	
	$\hat{\lambda}_2^\star$	-0.014	-0.007	-0.009	1.462	1.243	
	$\hat{\phi}$	0.001	0.002	0.003	1.192	1.203	
(0.5, 0.3, 1, 3, 1)	$\hat{\alpha}_1$	0.001	-0.002	-0.001	2.387	1.333	0.00
	$\hat{\alpha}_2$	0.000	0.000	0.001	1.403	1.125	
	$\hat{\lambda}_1^\star$	-0.002	0.004	0.003	2.151	1.574	
	$\hat{\lambda}_2^\star$	-0.001	0.000	-0.002	1.576	1.350	
	$\hat{\phi}$	-0.003	-0.001	-0.001	1.286	1.342	
(0.5, 0.3, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.002	-0.002	0.000	2.366	1.297	0.00
	$\hat{\alpha}_2$	-0.003	-0.003	-0.002	1.458	1.117	
	$\hat{\lambda}_1^\star$	0.000	0.000	-0.015	2.037	1.376	
	$\hat{\lambda}_2^\star$	0.004	0.003	-0.001	1.415	1.259	
	$\hat{\phi}$	0.008	0.008	0.007	1.235	1.244	
(0.5, 0.3, 3, 3, 1)	$\hat{\alpha}_1$	-0.002	-0.003	-0.001	2.364	1.284	0.00
	$\hat{\alpha}_2$	0.000	-0.001	-0.001	1.394	1.112	
	$\hat{\lambda}_1^\star$	-0.005	-0.002	0.000	2.108	1.435	
	$\hat{\lambda}_2^\star$	-0.014	-0.008	0.007	1.422	1.276	
	$\hat{\phi}$	0.010	0.013	-0.001	1.221	1.241	

Table IV (cont.). Bias for the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\lambda}_1^\star$, $\hat{\lambda}_2^\star$ and $\hat{\phi}$ and SD ratios of the MoM and YW estimators to the respective ML estimators of the Poisson BINAR(1) model at $n = 1000$.

		Bias			**SD ratios**		**% Extra Datasets**
(α_1, α_2, $\lambda_1^\star$, $\lambda_2^\star$, ϕ)		MoM	YW	ML	MoM/ML	YW/ML	
(0.5, 0.5, 1, 1, 0.5)	$\hat{\alpha}_1$	-0.005	-0.005	-0.002	2.184	1.248	0.00
	$\hat{\alpha}_2$	-0.001	-0.002	0.000	2.393	1.354	
	$\hat{\lambda}_1^\star$	0.010	0.007	0.000	2.073	1.506	
	$\hat{\lambda}_2^\star$	0.000	0.001	-0.003	2.115	1.528	
	$\hat{\phi}$	0.003	0.005	0.005	1.298	1.364	
(0.5, 0.5, 1, 1, 1)	$\hat{\alpha}_1$	-0.001	-0.002	0.000	2.545	1.397	0.00
	$\hat{\alpha}_2$	-0.001	-0.002	-0.001	2.499	1.382	
	$\hat{\lambda}_1^\star$	-0.012	-0.011	-0.009	2.404	1.718	
	$\hat{\lambda}_2^\star$	-0.009	-0.007	-0.001	2.548	1.826	
	$\hat{\phi}$	0.009	0.013	0.005	1.243	1.352	
(0.5, 0.5, 1, 3, 0.5)	$\hat{\alpha}_1$	-0.006	-0.005	-0.003	2.424	1.343	0.00
	$\hat{\alpha}_2$	0.000	-0.002	-0.001	2.491	1.375	
	$\hat{\lambda}_1^\star$	0.004	0.000	0.000	1.709	1.372	
	$\hat{\lambda}_2^\star$	-0.017	0.001	0.004	2.447	1.539	
	$\hat{\phi}$	0.010	0.012	0.005	1.304	1.330	
(0.5, 0.5, 1, 3, 1)	$\hat{\alpha}_1$	-0.002	-0.003	-0.001	2.365	1.350	0.00
	$\hat{\alpha}_2$	0.002	-0.002	-0.001	2.525	1.349	
	$\hat{\lambda}_1^\star$	-0.006	-0.009	-0.005	1.944	1.571	
	$\hat{\lambda}_2^\star$	-0.029	-0.007	0.001	2.578	1.619	
	$\hat{\phi}$	0.014	0.019	0.011	1.307	1.390	
(0.5, 0.5, 3, 3, 0.5)	$\hat{\alpha}_1$	-0.003	-0.004	-0.002	2.329	1.289	0.80
	$\hat{\alpha}_2$	0.002	-0.002	-0.001	2.401	1.319	
	$\hat{\lambda}_1^\star$	-0.014	-0.011	-0.006	1.920	1.380	
	$\hat{\lambda}_2^\star$	-0.041	-0.019	-0.007	1.999	1.415	
	$\hat{\phi}$	0.031	0.034	0.017	1.295	1.315	
(0.5, 0.5, 3, 3, 1)	$\hat{\alpha}_1$	0.002	-0.001	0.000	2.431	1.323	0.00
	$\hat{\alpha}_2$	0.003	0.000	0.000	2.391	1.291	
	$\hat{\lambda}_1^\star$	-0.036	-0.023	-0.010	2.142	1.534	
	$\hat{\lambda}_2^\star$	-0.042	-0.023	-0.004	2.056	1.467	
	$\hat{\phi}$	0.016	0.022	0.005	1.311	1.354	