REVSTAT – Statistical Journal
Volume 18, Number 1, January 2020, 27–45

A TRANSITION MODEL FOR ANALYSIS OF ZERO-INFLATED LONGITUDINAL COUNT DATA USING GENERALIZED POISSON REGRESSION MODEL

Authors: Taban Baghfalaki
– Department of Statistics, Faculty of Mathematical Sciences, Tarbiat Modares University, Tehran, Iran
t.baghfalaki@modares.ac.ir

Mojtaba Ganjali
– Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran
m-ganjali@sbu.ac.ir

Received: September 2016 Revised: May 2017 Accepted: September 2017

Abstract:

- In most of the longitudinal studies, involving count responses, excess zeros are common in practice. Usually, the current response measurement in a longitudinal sequence is a function of previous outcomes. For example, in a study about acute renal allograft rejection, the number of acute rejection episodes for a patient in current time is a function of this outcome at previous follow-up times. In this paper, we consider a transition model for accounting the dependence of current outcome on the previous outcomes in the presence of excess zeros. We propose the use of the generalized Poisson distribution as a flexible distribution for considering overdispersion (or underdispersion). The maximum likelihood estimates of the parameters are obtained using the EM algorithm. Some simulation studies are performed for illustration of the proposed methods. Also, analysis of a real data set of a kidney allograft rejection study illustrates the application of the proposed model.

Key-Words:

- *count data; EM algorithm; generalized Poisson distribution; longitudinal data; transition models; zero-inflated models.*

AMS Subject Classification:

- 62J99, 62P10.

1. INTRODUCTION

In modeling many count longitudinal clinical studies, the excess of zero is a common problem. For example, in a study about acute renal allograft rejection, many patients may have no acute rejection episodes at some follow-up times or in an asthma-related study, if the response variable is the number of asthma-related hospitalizations at each follow-up time, many patients may report no asthma-related hospitalizations. In these examples, the response variable for the patient can be considered as a count variable which may be recorded with extra zeros. Useful models for describing these kinds of data sets are zero-inflated models. In these models a special probability is allocated to zero observations (see Section 2 for more details).

Several approaches are proposed for analyzing these data sets. For example, hurdle model [25, 15, 16] and zero-inflated Poisson (ZIP) model [19, 12] are two well-known approaches for analysing zero-inflated (ZI) count data. Also, zero-inflated generalized Poisson (ZIGP) and zero-inflated negative binomial (ZINB) models are two other well-known approaches for considering overdispersion of which ZIGP model can also consider underdispersion to analyse inflated count data. [7] proposed a ZIGP model to analyse the data set of outsourcing of patent applications.

The analysis of longitudinal ZI count data are discussed frequently in literature. [4] proposed ZIP and ZINB models for analysing data of a study of growth. They describe their approaches as mixture models with a proportion P of subjects not at risk, and a proportion of 1–P at risk subjects who take on outcome values following a Poisson or negative binomial distribution. [21] used the ZIP and ZINB models to analyze longitudinal studies in epidemiology. [23] proposed a random effect model to analysis the ZI longitudinal count data. [28] discussed application of the ZI and hurdle models for longitudinal studies concerning vaccination safety. [14] used ZIP regression for analysing longitudinal data. [2] proposed a two-part regression model for analysing ZI longitudinal count data. They used their proposed approach for analysing an healthcare utilization data set. [26] discussed a Bayesian paradigm for ZIP and ZINB model for analysing data set of a study of psychiatric outpatient service. [27] provided a review of the literature and tests the Poisson, the ZIP, the negative binomial (NB) and the ZINB models in the context of longitudinal count data. [3] give many examples of the use of ZI distributions to model longitudinal data and consider this approach as a conventional one. [22] described a mixed-effect hurdle model for ZI longitudinal count data, where a baseline variable is included in the model specification. They used their proposed approach to analyse a healthcare utilization data.

A common problem in the practice of studying count data is overdispersion or underdispersion. The use of Poisson distribution to analyze count data has a lack of fit because of ignoring to consider these problems. To deal with overdispersion the use of NB distribution is proposed. But, this distribution has a lack of fit for considering the possible underdispersion. A distribution function which considers both the overdispersion and underdispersion is the generalized Poisson distribution [6, 5]. Note that the zero-inflation generally involve overdispersion or underdispersion. Here, the use of ZIGP distribution is recommended to consider both problems of underdispersion and overdispersion. Underdispersion is rarely occurred in practice. Therefore, the most concern of this paper is on the overdispersion in zero-inflated longitudinal data.

Three main modeling families are introduced to model longitudinal data: marginal models, subject-specified models and conditionally specified models [9, 24]. In a marginal model, marginal distributions are used to describe the longitudinal outcomes vector given a set of predictor variables. The correlation among the components of the longitudinal measurements can be captured by a fully parametric approach or by modeling a limited number of lower-order moments such as generalized estimating equations (GEE). In random effects or subject-specified models the longitudinal outcome vector is modeled by a vector of random effects. Several software and programs, for instance SAS and Mplus, make it possible to fit ZIP and ZINB distributions to longitudinal ZI data using random effects models. Finally in a conditionally specified model any response within the sequence of longitudinal measurements is modeled conditional upon the outcome on the previous time or a subset of previous outcomes. A particular relevant class of conditional models is the so-called autoregressive or transition models. In a transition model a current measurement in a longitudinal study is described as a function of the previous outcomes [9]. In this paper, our focus is on transition models. For some applications of the transition models in repeated measurement outcomes see [1, 18, 11]. Also, for reviews of transition models for analyzing the longitudinal data see [9], [30] and [10].

In this paper, we use the ZIGP transition models to analyze longitudinal count data with extra zeros. We use the usual EM algorithm for parameters estimation. The proposed model is illustrated using some simulation studies, where the performance of the proposed distributional assumption for transition model is compared with ZIP, ZINB, NB and GP distributional assumptions. Also, the proposed method is used for analyzing a real data set of a kidney allograft rejection study in application section where the best fitting model is selected by using Akaike information criterion (AIC), Bayesian information criterion (BIC) and Hannan–Quinn criterion (HQC).

This paper is organized as follows: Section 2 is a review on generalized Poisson and zero-inflated generalized Poisson distributions and the relation of these distributions with Poisson and zero-inflated Poisson distributions. Section 3 includes some notation, definitions of models, likelihood functions, the EM algorithm and our illustration of the proposed transition model for analyzing zero-inflated longitudinal data. In Section 4, some simulation studies are performed. In this section four different structures are considered for generating data and performance of ZIGP, ZINB, ZIP, NB and GP transition models are compared for each structure. The description and the analysis of a real data set using the proposed model are given and comparison of the performance of our approach with some other distributional assumptions is given in Section 5. The last section includes some conclusions and discussions.

2. ZERO-INFLATED GENERALIZED POISSON DISTRIBUTION

The random variable Y is said to have a generalized Poisson distribution, if its probability mass function is given by

$$f(y;\xi,\omega) = \frac{\xi(\xi+\omega y)^{y-1}}{y!}\, e^{-(\xi+\omega y)}, \qquad y = 0,1,2,... \tag{2.1}$$

where $\xi > 0$ and $\max(-1, -\xi/4) < \omega < 1$ [13]. The mean and variance of this distribution are given by

$$E(Y) = \frac{\xi}{1-\omega}, \qquad \text{Var}(Y) = \frac{\xi}{(1-\omega)^3} = \frac{1}{(1-\omega)^2} E(Y),$$

therefore, the term $\frac{1}{(1-\omega)^2}$ plays the role of a dispersion factor. Clearly, when $\omega = 0$, the generalized Poisson distribution reduces to the usual Poisson distribution with parameter ξ. Further, when $\omega > 0$, we have overdispersion in the model; when $\omega < 0$, we have underdispersion.

A parameterization of this distribution is given by setting $\lambda = \frac{\xi}{1-\omega}$ and $\phi = \frac{\omega}{\xi}$, denoted by $Y \sim GP(\lambda, \phi)$, and its probability mass function is given by

$$f_{GP}(y; \lambda, \phi) = \left(\frac{\lambda}{1+\phi\lambda}\right)^y \frac{(1+\phi y)^{y-1}}{y!} \exp\left(\frac{-\lambda(1+\phi y)}{1+\phi\lambda}\right), \quad y = 0, 1, 2, \ldots, \quad \lambda > 0, \tag{2.2}$$

where ϕ is a real value parameter such that for all y, $1 + \phi y > 0$ and $1 + \phi\lambda > 0$. These restrictions are confirmed by the restriction on the distribution (2.1). The generalized Poisson distribution (2.2) is a natural extension of the Poisson distribution. If $\phi = 0$, then the probability function (2.2) reduces to the Poisson distribution, denoted by $Y \sim P(\lambda)$. By the above mentioned parameterization, the mean of Y is given by $E(Y) = \lambda$ and the variance of Y is given by $\text{Var}(Y) = \lambda(1 + \phi\lambda)^2$. In the generalized Poisson distribution, the ϕ parameter is called dispersion parameter. When $\phi > 0$, the overdispersion is presented in the model, whereas when $\phi < 0$, the underdispersion is included in the model. The generalized Poisson distribution is a more flexible distribution than the negative binomial distribution for considering possibility of underdispersion or overdispersion. This property is one of the well-known properties of generalized Poisson distribution. [17] proved that the generalized Poisson distribution, the same as negative binomial distribution, can be considered as a mixture of the Poisson distribution. [17] show that there are some differences between the fits of the generalized Poisson and negative binomial distributions. When the first two moments are fixed, the negative binomial distribution have larger mass at zero than the generalized Poisson distribution. This means their zero-inflated variations tend to have larger discrepancy. However, the fits of their zero-inflated variations may differ when there is a large zero fraction [17]. For more details about generalized Poisson distribution see [6] and [5]. Also, *VGAM* and *HMMpa* packages of R can be applied to use the generalized Poisson distribution.

A zero-inflated generalized Poisson distribution for a positive value π $(0 \leq \pi \leq 1)$ is defined as follows:

$$f_{ZIGP}(y; \lambda, \phi, \pi) = \begin{cases} \pi + (1-\pi) f_{GP}(0; \lambda, \phi), & y = 0, \\ (1-\pi) f_{GP}(y; \lambda, \phi), & y > 0, \end{cases} \tag{2.3}$$

where $f_{GP}(\cdot; \lambda, \phi)$ is the probability mass function of (2.2). We will use the notation $Y \sim ZIGP(\lambda, \phi, \pi)$ to denote the distribution of (2.3). The mean and variance of this distribution are given by $E(Y) = (1-\pi)\lambda$ and $\text{var}(Y) = E(Y)[(1+\phi\lambda)^2 + \pi\lambda]$, respectively. The variance of this distribution shows that for $\pi > 0$ and $\phi > 0$ the distribution of Y exhibits overdispersion. The distribution (2.3) reduced to the generalized Poisson distribution when $\pi = 0$ and it reduced to zero-inflated Poisson distribution when $\phi = 0$, denoted by $Y \sim ZIP(\lambda, \pi)$. When π is allowed to be negative, the distribution (2.3) presents a zero-deflated generalized Poisson distribution which rarely occurs in practice.

3. ZERO-INFLATED TRANSITION MODELS FOR COUNT RESPONSES

Suppose N individuals are participated in a longitudinal study and for each individual n_i $(i = 1, 2, ..., N)$ repeated measurements are recorded as response variables. Also, let Y_{ij}, $i = 1, 2, ..., N$ and $j = 1, 2, ..., n_i$ be the longitudinal measurements for the i^{th} individual at j^{th} time point and let W_{ij}, $i = 1, 2, ..., N$ and $j = 1, 2, ..., n_i$ be indicator variables as follows:

$$W_{ij} = \begin{cases} 1\,, & Y_{ij} \text{ is from the perfect state}\,, \\ 0\,, & Y_{ij} \text{ is from the Poisson state}\,. \end{cases}$$

where by perfect we means that the sample is from a degenerated distribution at 0. It is clear that W_{ij} is a latent variable. Also, let $\boldsymbol{h}_{ij} = (Y_{i1}, ..., Y_{i,j-1})$ be the previous outcomes up to time j or in other words history of outcomes for the i^{th} individual.

In a transition model, the outcome Y_{ij} is modeled in term of $\boldsymbol{h}_{ij}$ [9]. The order of a transition model is the number of the previous measurements that are considered for modeling the measurement of the current time. We consider a first order zero-inflated transition model as follows:

$$\begin{aligned} &P_{ZI}\Big(Y_{ij} = y_{ij} \,|\, \pi_{ij}, \lambda_{ij}, \phi, \boldsymbol{x}_{ij}, \boldsymbol{z}_{ij}, y_{i,j-1}\Big) = \\ &\qquad = \begin{cases} \pi_{ij} + (1-\pi_{ij})\, P\big(Y_{ij} = y_{ij} \,|\, \lambda_{ij}, \phi, \boldsymbol{x}_{ij}, y_{i,j-1}\big)\,, & y_{ij} = 0\,, \\ (1-\pi_{ij})\, P\big(Y_{ij} = y_{ij} \,|\, \lambda_{ij}, \phi, \boldsymbol{x}_{ij}, y_{i,j-1}\big)\,, & y_{ij} \neq 0\,, \end{cases} \end{aligned} \tag{3.1}$$

where

$$\log(\lambda_{i1}) \;=\; \boldsymbol{x}'_{i1}\boldsymbol{\beta}\,, \tag{3.2}$$

$$\text{logit}(\pi_{i1}) \;=\; \boldsymbol{z}'_{i1}\boldsymbol{\alpha}\,, \tag{3.3}$$

and, for $j = 2, 3, ..., n_i$,

$$\log(\lambda_{ij}) \;=\; \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \gamma_1 I_{\{0\}}(Y_{i,j-1}) + \gamma_2 y_{i,j-1}(1 - I_{\{0\}}(Y_{i,j-1}))\,, \tag{3.4}$$

$$\text{logit}(\pi_{ij}) \;=\; \boldsymbol{z}'_{ij}\boldsymbol{\alpha} + \tau_1 I_{\{0\}}(Y_{i,j-1}) + \tau_2 y_{i,j-1}(1 - I_{\{0\}}(Y_{i,j-1}))\,, \quad j = 2, ..., n_i\,, \tag{3.5}$$

where $\pi_{ij} = P(Y_{ij} = 0 \,|\, \boldsymbol{\alpha}, \boldsymbol{z}_{ij}, \boldsymbol{h}_{ij}) = P(Y_{ij} = 0 \,|\, \boldsymbol{\alpha}, \boldsymbol{z}_{ij}, y_{i,j-1})$ is the rate of zeros given some covariates and the history of outcomes. In this model the effect of the previous zero response on current measurement (γ_1) and the effect of the non-zero previous response on current mean (γ_2) are separately considered. This is due to the fact that one expects to have the current mean to be close to the previous values of responses.

We will use the notation $Y \sim ZIGP(\lambda_{ij}, \pi_{ij}, \phi)$ to denote model (3.1). Note that (3.1) is reduced to zero-inflated Poisson model when $\phi = 0$. We will use the notation $Y \sim ZIP(\lambda_{ij}, \pi_{ij})$ to denote model (3.1) when $\phi = 0$. Let $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\tau}, \phi)$ be the vector of all the unknown parameters in the model where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)'$ and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)'$. The likelihood

function of the model can be written as:

$$\begin{aligned} L(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{x},\boldsymbol{z}) &= \prod_{i=1}^{N}\left\{ f(y_{i1}) \times \prod_{j=2}^{n_i} f(y_{ij}|y_{i,j-1}) \right\} \\ &= \prod_{i=1}^{N}\prod_{j=1}^{n_i} \Big(\pi_{ij} + (1-\pi_{ij})\, P\big(Y_{ij}=0 \,|\, \lambda_{ij},\phi,\boldsymbol{x}_{ij},\boldsymbol{h}_{ij}\big)\Big)^{I(y_{ij}=0)} \\ &\quad \times \Big((1-\pi_{ij})\, P\big(Y_{ij}\neq 0 \,|\, \lambda_{ij},\phi,\boldsymbol{x}_{ij},\boldsymbol{h}_{ij}\big)\Big)^{1-I(y_{ij}=0)}, \end{aligned}$$

where $h_{i1}=0$ and it will not be considered in the model. This likelihood function can be maximized using some numerical methods such as Newton–Raphson [20].

Another approach for obtaining parameter estimates is the use of the Expectation-Maximization (EM) [8] algorithm. To obtain the EM estimates of the parameters, we use the indicator variable, W_{ij}, $i=1,2,...,N$, $j=1,2,...,n_i$. As mentioned earlier W_{ij} is a latent variable for indicating the perfect state versus the Poisson state outcome. Therefore, the log-likelihood function of $(\boldsymbol{Y},\boldsymbol{W})$ as complete data is given by

$$\begin{aligned} \ell_c(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{w},\boldsymbol{x},\boldsymbol{z}) &= \sum_{i=1}^{N}\sum_{j=1}^{n_i} w_{ij}\log(\pi_{ij}) + \sum_{i=1}^{N}\sum_{j=1}^{n_i}(1-w_{ij})\log(1-\pi_{ij}) \\ &\quad + \sum_{i=1}^{N}\sum_{j=1}^{n_i}(1-w_{ij})\left\{ y_{ij}\log(\lambda_{ij}) - y_{ij}\log(1+\phi\lambda_{ij}) \right. \\ &\quad \left. + (y_{ij}-1)\log(1+\phi y_{ij}) - \log(y_{ij}!) - \lambda_{ij}\frac{1+\phi y_{ij}}{1+\phi\lambda_{ij}} \right\}. \end{aligned}$$

The EM algorithm contains two steps: in the first step (E-step), the expectation of the complete likelihood function (here $\ell_c(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{w},\boldsymbol{x},\boldsymbol{z})$) given the observed data (here $\boldsymbol{Y}$) and the current value of the parameters in the r^{th} step (called $\boldsymbol{\theta}^{(r)}$) is calculated, by defining $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)}) = E\big[\ell_c(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{w},\boldsymbol{x},\boldsymbol{z}) \,|\, \boldsymbol{y},\boldsymbol{x},\boldsymbol{z},\boldsymbol{\theta}^{(r)}\big]$. We have

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)}) &= \sum_{i=1}^{N}\sum_{j=1}^{n_i} E[W_{ij}|\boldsymbol{y},\boldsymbol{x},\boldsymbol{z},\boldsymbol{\theta}^{(r)}]\log(\pi_{ij}) \\ &\quad + \sum_{i=1}^{N}\sum_{j=1}^{n_i}(1-E[W_{ij}|\boldsymbol{y},\boldsymbol{x},\boldsymbol{z},\boldsymbol{\theta}^{(r)}])\log(1-\pi_{ij}) \\ &\quad + \sum_{i=1}^{N}\sum_{j=1}^{n_i}(1-E[W_{ij}|\boldsymbol{y},\boldsymbol{x},\boldsymbol{z},\boldsymbol{\theta}^{(r)}])\left\{ y_{ij}\log(\lambda_{ij}) - y_{ij}\log(1+\phi\lambda_{ij}) \right. \\ &\quad \left. + (y_{ij}-1)\log(1+\phi y_{ij}) - \log(y_{ij}!) - \lambda_{ij}\frac{1+\phi y_{ij}}{1+\phi\lambda_{ij}} \right\}. \end{aligned}$$

For computing the EM algorithm, the following expectation is needed:

$$\begin{aligned} E[W_{ij}|\boldsymbol{y},\boldsymbol{w},\boldsymbol{x},\boldsymbol{z},\boldsymbol{\theta}^{(r)}] &= P(W_{ij}=1|y_{i,j-1},\boldsymbol{x},\boldsymbol{z},\boldsymbol{\theta}^{(r)}) \\ &= \begin{cases} \frac{\pi_{ij}^{(r)}}{\pi_{ij}^{(r)}+(1-\pi_{ij}^{(r)})P(Y_{ij}=0|\lambda_{ij}^{(r)},\phi^{(r)},\boldsymbol{x}_{ij},y_{i,j-1})}, & y_{ij}=0\,, \\ 0\,, & y_{ij}\neq 0\,, \end{cases} \end{aligned}$$

where $\pi_{ij}^{(r)} = P(Y_{ij}=0|\boldsymbol{z}_{ij},\boldsymbol{\theta}^{(r)},y_{i,j-1})$ and $\lambda_{ij}^{(r)}$ is the current Poisson rate at the r^{th} iteration.

In the second step (M-step), we define

$$\boldsymbol{\theta}^{(r+1)} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\max}\, Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})\,.$$

The algorithm is converged and is stopped when

$$\left\|\boldsymbol{\theta}^{(r)} - \boldsymbol{\theta}^{(r+1)}\right\| < \varepsilon\,,$$

where $\|\cdot\|$ is a pre-specified measure.

4. SIMULATION STUDIES

In this section some simulation studies are performed for investigating the performance of the proposed approach. At first, the data are generated from ZIGP and the performance of ZIGP, GP, ZINB, NB and ZIP are compared on analyzing these data. Two other simulated data are generated under ZINB and ZIP where the performance of analyzing ZIGP, ZINB and ZIP are investigated in each case. Note that ZIP model is a ZIGP model with $\phi = 0$. The last simulation study is used to examine the performance of ZIGP, ZINB and ZIP in the presence of underdispersion.

4.1. Zero-inflated generalized Poisson model

In this simulation study the data set is generated from a transition model under ZIGP. The simulation study contains two sample sizes $N = 100$ and 500 where $M = 1000$ iterations are performed. For generating data, we consider a ZIGP model as follows:

$$Y_{ij}|\lambda_{ij},\pi_{ij} \sim ZIGP(\lambda_{ij},\pi_{ij},\phi)\,, \tag{4.1}$$

where

$$\begin{aligned}
\log(\lambda_{i1}) &= \beta_0 + \beta_1 x_i + \beta_2 t_1\,,\\
\operatorname{logit}(\pi_{i1}) &= \alpha_0 + \alpha_1 x_i + \alpha_2 t_1\,,\\
\log(\lambda_{ij}) &= \beta_0 + \beta_1 x_i + \beta_2 t_j + \gamma_1 I_{\{0\}}(Y_{i,j-1}) + \gamma_2 y_{i,j-1}(1 - I_{\{0\}}(Y_{i,j-1}))\,,\quad j=2,3,4\,,\\
\operatorname{logit}(\pi_{ij}) &= \alpha_0 + \alpha_1 x_i + \alpha_2 t_j + \tau_1 I_{\{0\}}(Y_{i,j-1}) + \tau_2 y_{i,j-1}(1 - I_{\{0\}}(Y_{i,j-1}))\,,\quad j=2,3,4\,.
\end{aligned} \tag{4.2}$$

For this simulation study, two sets of real values are considered as follows:

1) $\alpha_0 = -1,\ \alpha_1 = 1,\ \alpha_2 = 0,\ \beta_0 = -3,\ \beta_1 = \beta_2 = 1,\ \gamma_1 = -1,\ \gamma_2 = 0,\ \tau_1 = 0,\ \tau_2 = 1$ and $\phi = 1$.

2) $\alpha_0 = -1, \alpha_1 = -1, \alpha_2 = 0, \beta_0 = -3, \beta_1 = \beta_2 = 1, \gamma_1 = 1, \gamma_2 = -1, \tau_1 = 1, \tau_2 = -1$ and $\phi = 0.5$.

The results of these simulation studies are summarized in Tables 1 and 2, respectively.

Table 1: Results of simulation study for generated data under ZIGP model, estimate (Est.), standard error (S.E.), relative bias (Bias) and mean square error (MSE) for $M=$ 1000 simulated data with sample sizes 100 and 500 and the first set of real values.

N	Para.	Real	ZIGP			GP			ZINB			NB			ZIP		
			Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE
100	α_0	-1.00	$-1.26\,(0.39)$	0.26	0.46	—	—	—	$-15.35\,(28.71)$	14.35	102.37	—	—	—	$0.53\,(0.71)$	-1.53	2.84
	α_1	1.00	$1.18\,(0.97)$	0.18	0.91	—	—	—	$8.51\,(18.13)$	7.51	3.82	—	—	—	$0.05\,(0.12)$	-0.94	0.90
	α_2	0.00	$-0.04\,(0.39)$	$*$	0.14	—	—	—	$0.48\ \ (2.81)$	$*$	8.11	—	—	—	$0.02\,(0.03)$	$*$	0.00
	τ_1	0.00	$0.14\,(0.35)$	$*$	0.48	—	—	—	$1.19\,(23.82)$	$*$	564.39	—	—	—	$0.16\,(0.13)$	$*$	0.21
	τ_2	-1.00	$-0.93\,(0.69)$	-0.06	0.85	—	—	—	$-3.42\ \ (9.77)$	2.42	100.75	—	—	—	$-0.02\,(0.04)$	-0.97	0.95
	β_0	-3.00	$-2.98\,(0.59)$	-0.00	0.32	$-3.35\,(0.46)$	0.11	0.33	$-3.33\ \ (0.64)$	0.11	0.53	$-3.16\,(0.50)$	0.05	0.26	$-0.05\,(0.29)$	-0.98	8.76
	β_1	1.00	$1.03\,(0.42)$	0.03	0.17	$0.62\,(0.28)$	-0.37	0.21	$1.02\ \ (0.41)$	0.02	0.16	$0.64\,(0.27)$	-0.35	0.19	$0.11\,(0.18)$	-0.88	0.81
	β_2	1.00	$0.99\,(0.21)$	-0.00	0.04	$1.02\,(0.21)$	0.02	0.04	$1.04\ \ (0.18)$	0.04	0.03	$1.04\,(0.14)$	0.04	0.02	$0.36\,(0.15)$	-0.63	0.42
	γ_1	-1.00	$-0.96\,(0.45)$	-0.03	0.19	$-1.17\,(0.55)$	0.17	0.33	$-1.07\ \ (0.48)$	0.07	0.23	$-1.34\,(0.31)$	0.34	0.21	$-0.05\,(0.20)$	-0.94	0.93
	γ_2	0.00	$0.09\,(0.39)$	$*$	0.15	$-0.00\,(0.02)$	$*$	0.00	$-0.02\ \ (0.09)$	$*$	0.10	$0.01\,(0.08)$	$*$	0.00	$0.01\,(0.07)$	$*$	0.00
	ϕ	1.00	$0.85\,(0.16)$	-0.14	0.04	$1.86\,(0.27)$	0.86	0.86	$0.27\ \ (0.10)$	-0.72	0.54	$0.16\,(0.02)$	-0.83	0.69	—	—	—
500	α_0	-1.00	$-0.95\,(0.23)$	-0.05	0.39	—	—	—	$-11.92\ \ (7.01)$	10.92	166.61	—	—	—	$0.09\,(0.21)$	-1.09	1.23
	α_1	1.00	$1.06\,(0.40)$	0.06	0.16	—	—	—	$10.01\ \ (7.39)$	9.01	133.75	—	—	—	$0.08\,(0.14)$	-0.91	0.84
	α_2	0.00	$-0.03\,(0.10)$	$*$	0.01	—	—	—	$0.41\ \ (0.24)$	$*$	0.22	—	—	—	$-0.08\,(0.17)$	$*$	0.03
	τ_1	0.00	$0.09\,(0.40)$	$*$	0.16	—	—	—	$-0.51\ \ (1.07)$	$*$	1.37	—	—	—	$0.24\,(0.31)$	$*$	0.15
	τ_2	-1.00	$-0.97\,(0.45)$	-0.03	0.25	—	—	—	$-1.89\ \ (2.22)$	0.89	5.53	—	—	—	$-0.02\,(0.03)$	-0.97	0.95
	β_0	-3.00	$-2.99\,(0.31)$	-0.00	0.09	$-3.21\,(0.28)$	0.07	0.26	$-3.35\ \ (0.19)$	0.11	0.16	$-3.33\,(0.26)$	0.11	0.18	$-0.40\,(0.65)$	-0.86	7.16
	β_1	1.00	$1.00\,(0.17)$	0.00	0.03	$0.65\,(0.26)$	-0.34	0.18	$0.92\ \ (0.14)$	-0.07	0.02	$0.71\,(0.16)$	-0.28	0.10	$0.27\,(0.33)$	-0.72	0.62
	β_2	1.00	$1.01\,(0.08)$	0.01	0.00	$1.02\,(0.21)$	0.02	0.04	$1.08\ \ (0.06)$	0.08	0.01	$1.07\,(0.06)$	0.07	0.01	$0.49\,(0.21)$	-0.50	0.30
	γ_1	-1.00	$-1.02\,(0.18)$	0.02	0.03	$-1.17\,(0.57)$	0.17	0.34	$-1.23\ \ (0.16)$	0.23	0.08	$-1.28\,(0.17)$	0.28	0.10	$-0.25\,(0.39)$	-0.74	0.70
	γ_2	0.00	$-0.00\,(0.02)$	$*$	0.00	$0.27\,(0.52)$	$*$	0.34	$-0.00\ \ (0.02)$	$*$	0.00	$0.00\,(0.03)$	$*$	0.00	$0.01\,(0.06)$	$*$	0.00
	ϕ	1.00	$0.93\,(0.11)$	-0.06	0.01	$1.76\,(0.28)$	0.76	0.66	$0.21\ \ (0.02)$	-0.78	0.62	$0.45\,(0.01)$	-0.55	0.52	—	—	—

Table 2: Results of simulation study for generated data under ZIGP model, estimate (Est.), standard error (S.E.), relative bias (Bias) and mean square error (MSE) for $M =$ 1000 simulated data with sample sizes 100 and 500 and the second set of real values.

N	Para.	Real	ZIGP			GP			ZINB			NB			ZIP		
			Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE
100	α_0	−1.00	−1.19 (0.72)	0.09	0.81	—	—	—	65.76 (198.29)	−66.76	37221.00	—	—	—	0.11 (0.33)	−1.11	1.31
	α_1	−1.00	−0.98 (0.53)	−0.02	0.22	—	—	—	−82.51 (200.41)	81.51	40113.72	—	—	—	−0.58 (0.71)	−0.42	0.56
	α_2	0.00	0.03 (0.17)	*	0.02	—	—	—	0.93 (0.51)	*	1.08	—	—	—	−0.06 (0.12)	*	0.01
	τ_1	1.00	1.14 (0.40)	0.14	0.55	—	—	—	−70.94 (199.79)	−71.94	38439.76	—	—	—	0.49 (0.59)	−0.51	0.52
	τ_2	−1.00	−1.01 (0.21)	−0.01	0.16	—	—	—	−82.87 (193.31)	81.87	37844.46	—	—	—	0.31 (0.34)	−1.31	1.81
	β_0	−3.00	−3.11 (0.40)	0.04	0.14	−3.40 (0.35)	0.13	0.27	−3.62 (0.68)	0.21	0.76	−3.57 (0.65)	0.19	0.53	−0.79 (1.15)	−0.74	5.87
	β_1	1.00	1.08 (0.23)	0.08	0.05	1.22 (0.18)	0.22	0.08	1.22 (0.48)	0.22	0.24	1.40 (0.56)	0.40	0.32	0.20 (0.18)	−0.80	0.66
	β_2	1.00	0.98 (0.15)	−0.02	0.02	1.03 (0.12)	0.03	0.01	1.15 (0.08)	0.15	0.03	0.95 (0.11)	−0.05	0.01	0.55 (0.17)	−0.45	0.23
	γ_1	1.00	1.14 (0.15)	0.14	0.04	0.62 (0.39)	−0.38	0.28	0.75 (0.31)	−0.25	0.14	1.06 (0.22)	0.06	0.03	0.72 (0.76)	−0.28	0.51
	γ_2	−1.00	−1.05 (0.26)	0.05	0.06	−0.98 (0.28)	−0.02	0.07	−1.22 (0.33)	0.22	0.14	−0.84 (0.25)	−0.16	0.06	−0.14 (0.19)	−0.86	0.77
	ϕ	0.50	0.47 (0.03)	−0.06	0.00	0.91 (0.11)	0.82	0.18	0.34 (0.06)	−0.33	0.03	0.28 (0.01)	−0.44	0.05	—	—	—
500	α_0	−1.00	−1.08 (0.32)	0.08	0.37	—	—	—	−20.95 (5.89)	19.95	426.92	—	—	—	1.45 (0.29)	−2.45	6.06
	α_1	−1.00	−1.12 (0.03)	0.12	0.01	—	—	—	−0.84 (0.28)	−0.16	0.09	—	—	—	−1.02 (0.09)	0.02	0.00
	α_2	0.00	−0.10 (0.06)	*	0.01	—	—	—	0.44 (0.12)	*	0.21	—	—	—	−0.25 (0.00)	*	0.06
	τ_1	1.00	1.01 (0.10)	0.01	0.11	—	—	—	18.64 (5.94)	17.64	340.38	—	—	—	0.17 (0.28)	−0.83	0.73
	τ_2	−1.00	−1.04 (0.13)	−0.04	0.11	—	—	—	−0.42 (3.05)	−0.58	8.07	—	—	—	0.23 (0.17)	−1.23	1.53
	β_0	−3.00	−3.00 (0.31)	0.00	0.05	−3.46 (0.04)	0.15	0.21	−3.25 (0.13)	0.08	0.08	−3.39 (0.19)	0.13	0.18	−1.17 (0.03)	−0.61	3.34
	β_1	1.00	1.01 (0.01)	0.01	0.00	1.29 (0.04)	0.29	0.09	1.14 (0.11)	0.14	0.03	1.32 (0.16)	0.32	0.12	0.32 (0.30)	−0.68	0.50
	β_2	1.00	0.98 (0.01)	−0.02	0.00	1.05 (0.01)	0.05	0.00	1.02 (0.07)	0.02	0.00	0.97 (0.04)	−0.03	0.00	0.75 (0.07)	−0.25	0.07
	γ_1	1.00	1.11 (0.25)	0.11	0.04	0.61 (0.06)	−0.39	0.15	0.82 (0.15)	−0.18	0.05	0.79 (0.08)	−0.21	0.05	0.68 (0.05)	−0.32	0.10
	γ_2	−1.00	−0.98 (0.10)	0.02	0.02	−0.98 (0.06)	−0.02	0.00	−0.98 (0.11)	−0.02	0.01	−0.89 (0.08)	−0.11	0.02	−0.72 (0.01)	−0.28	0.08
	ϕ	0.50	0.47 (0.02)	−0.05	0.00	0.92 (0.03)	0.85	0.18	0.38 (0.02)	−0.25	0.02	0.26 (0.01)	−0.49	0.06	—	—	—

The simulated data set are analyzed using NB, GP, ZIP, ZINB and ZIGP models, such that

$$\begin{aligned} Y_{ij}|\lambda_{ij},\phi &\sim NB\Big(\phi,\frac{\phi}{\phi+\lambda_{ij}}\Big),\\ Y_{ij}|\lambda_{ij},\phi &\sim GP\big(\lambda_{ij},\phi\big),\\ Y_{ij}|\lambda_{ij},\pi_{ij} &\sim ZIP\big(\lambda_{ij},\pi_{ij}\big),\\ Y_{ij}|\lambda_{ij},\pi_{ij},\phi &\sim ZINB\Big(\phi,\frac{\phi}{\phi+\lambda_{ij}},\pi_{ij}\Big),\\ Y_{ij}|\lambda_{ij},\pi_{ij},\phi &\sim ZIGP\big(\lambda_{ij},\phi,\pi_{ij}\big). \end{aligned} \tag{4.3}$$

Note that $Y\sim NB(\phi,\kappa)$ if the probability mass function is given by $f_{NB}(y;\phi,\kappa)=\frac{\Gamma(y+\phi)}{\Gamma(\phi)y!}\kappa^{\phi}(1-\kappa)^{y}$, $y=0,1,...,r$ and $r>0$. Also, $Y\sim ZINB(\phi,\kappa,\pi)$ is a zero-inflated negative binomial distribution which can be obtained by (2.3) by replacing $f_{GP}(\cdot;\lambda,\phi)$ by $f_{NB}(\cdot;\phi,\kappa)$. In order to compare the results, the mean of the estimated values, the standard errors, relative biases and mean of square errors (MSEs) are used. The latter two criteria are defined as follows:

$$Bias(\theta)=\frac{1}{M}\sum_{k=1}^{M}\Big(\frac{\hat{\theta}_k}{\theta}-1\Big),$$

$$MSE(\theta)=\frac{1}{M}\sum_{k=1}^{M}\Big(\hat{\theta}_k-\theta\Big)^2,$$

where $\hat{\theta}_k$ is the estimate of θ for the k^{th} sample, $k=1,2,...,M$.

The results of Tables 1 and 2 show that the performance of the ZIGP in parameter estimation is better than those of the other models. The performance of ZINB in estimating parameters of the logistic model is not well while in estimating the other parameters is almost good. The GP and NB models do not have good performances in this simulation study and the ZIP has a good performance in estimating some parameters. The results of the simulation study for ZIGP show that the increase in the sample size is an effective way of decreasing biases and standard deviations of parameters estimates. As shown in these tables, relative biases and MSEs are reduced by increasing the sample size. This suggests that the method in finding estimates is consistent.

4.2. Zero-inflated Poisson model

In this simulation study, we simulate data from the following model:

$$Y_{ij}|\lambda_{ij},\pi_{ij}\sim ZIP(\lambda_{ij},\pi_{ij})\,, \tag{4.4}$$

such that the parameterizations and real values of parameters in λ_{ij} and π_{ij} are the same as the first set of real values and those described in equation (4.2). The results of this simulation study are summarized in Table 3. The results show the well performance of the ZIP model. Also, the results show that the performance of ZIGP is as good as ZIP model. The ZINB model dose not have a good performance when the sample size is 100 while for N=500 has a performance which is as good as the other two models. The overdispersion parameter ϕ

is estimated zero in ZIGP model but it has a large value in ZINB (note that in negative binomial distribution the dispersion index is proportion to ϕ^{-1} and overdispersion presents in the data when the value of ϕ is very large). As a conclusion, this simulation study shows that the use of ZIGP model is preferred to the use of ZINB model. The ZIGP has a similar performance to ZIP and, for moderate sample size, a much better performance than ZINB model.

Table 3: Results of simulation study for generated data under ZIP model, estimate (Est.), standard error (S.E.), relative bias (Bias) and mean square error (MSE) for $M=$ 1000 simulated data with sample sizes 100 and 500.

N	Para.	Real	ZIP			ZIGP			ZINB		
			Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE
100	α_0	-1.00	$-1.06\,(0.86)$	0.06	0.74	$-1.20\,(0.98)$	0.20	0.95	$<-10^3\,(>10^3)$	$>10^3$	$>10^3$
	α_1	1.00	$1.03\,(0.50)$	0.03	0.24	$1.23\,(0.57)$	0.23	0.55	$>10^3\,(>10^3)$	$>10^3$	$>10^3$
	α_2	0.00	$0.02\,(0.20)$	$*$	0.04	$-0.00\,(0.19)$	$*$	0.00	$0.02\,(0.20)$	$*$	0.04
	τ_1	0.00	$-0.07\,(0.56)$	$*$	0.32	$-0.04\,(0.61)$	$*$	0.03	$-0.09\,(0.64)$	$*$	0.42
	τ_2	1.00	$-1.26\,(0.91)$	0.26	0.96	$-1.05\,(0.36)$	0.05	0.01	$-1.07\,(0.37)$	0.07	0.14
	β_0	-3.00	$-3.00\,(0.22)$	0.00	0.05	$-2.98\,(0.33)$	-0.00	0.00	$-3.04\,(0.25)$	0.01	0.06
	β_1	1.00	$1.00\,(0.08)$	0.00	0.00	$0.99\,(0.14)$	-0.01	0.00	$1.01\,(0.08)$	0.01	0.01
	β_2	-1.00	$0.99\,(0.05)$	-0.00	0.00	$0.99\,(0.06)$	-0.00	0.00	$1.00\,(0.05)$	0.00	0.00
	γ_1	-1.00	$-0.99\,(0.08)$	-0.00	0.00	$-0.99\,(0.09)$	-0.00	0.00	$-1.00\,(0.09)$	0.00	0.00
	γ_2	0.00	$-0.00\,(0.01)$	$*$	0.00	$0.00\,(0.01)$	$*$	0.00	$-0.00\,(0.01)$	$*$	0.00
	ϕ	0.00	—	—	—	$-0.00\,(0.00)$	$*$	0.00	$>10^3\,(>10^3)$	$>10^3$	$>10^3$
500	α_0	-1.00	$-1.02\,(0.36)$	0.02	0.13	$-1.02\,(0.36)$	0.02	0.13	$-1.03\,(0.36)$	0.03	0.13
	α_1	1.00	$1.02\,(0.21)$	0.02	0.04	$1.02\,(0.21)$	0.02	0.04	$1.02\,(0.21)$	0.02	0.04
	α_2	0.00	$0.00\,(0.07)$	$*$	0.00	$0.00\,(0.07)$	$*$	0.00	$0.00\,(0.07)$	$*$	0.00
	τ_1	0.00	$-0.02\,(0.24)$	$*$	0.06	$-0.02\,(0.24)$	$*$	0.06	$-0.02\,(0.24)$	$*$	0.06
	τ_2	1.00	$-1.01\,(0.17)$	0.01	0.02	$-1.01\,(0.17)$	0.01	0.02	$-1.02\,(0.17)$	0.02	0.06
	β_0	-3.00	$-3.00\,(0.11)$	0.00	0.01	$-3.00\,(0.11)$	0.00	0.01	$-3.00\,(0.11)$	0.00	0.01
	β_1	1.00	$1.00\,(0.04)$	0.00	0.00	$1.00\,(0.04)$	0.00	0.00	$1.00\,(0.04)$	0.00	0.00
	β_2	-1.00	$0.99\,(0.02)$	-0.00	0.00	$0.99\,(0.02)$	0.00	0.00	$1.00\,(0.02)$	0.00	0.00
	γ_1	-1.00	$-1.00\,(0.03)$	0.00	0.00	$-0.99\,(0.03$	0.00	0.00	$-1.00\,(0.03)$	0.00	0.00
	γ_2	0.00	$-0.00\,(0.00)$	$*$	0.00	$-0.00\,(0.00)$	$*$	0.00	$-0.00\,(0.00)$	$*$	0.00
	ϕ	0.00	—	—	—	$-0.00\,(0.00)$	$*$	0.00	$>10^3\,(>10^3)$	$>10^3$	$>10^3$

4.3. Zero-inflated negative binomial model

In this simulation study, we simulate data from the following model:

$$Y_{ij}|\lambda_{ij},\pi_{ij} \sim ZINB\left(\phi, \frac{\phi}{\phi+\lambda_{ij}}, \pi_{ij}\right), \tag{4.5}$$

such that the parameterizations and real values of parameters in λ_{ij} and π_{ij} are the same as the first set of real values and those described in equation (4.2), also, we consider $\phi = 1$. The results of this simulation study are summarized in Table 4. The results show the well performance of the ZINB model in large sample size. Also, the results show that the performance of ZIGP is as good as ZINB model expect for estimating intercept and the overdispersion parameters. Also, in moderate sample size the performance of ZIGP model is better

than those in ZINB model. The results show that the ZIP model dose not have a good performance.

Table 4: Results of simulation study for generated data under ZINB model, estimate (Est.), standard error (S.E.), relative bias (Bias) and mean square error (MSE) for $M=$ 1000 simulated data with sample sizes 100 and 500.

N	Para.	Real	ZINB			ZIGP			ZIP		
			Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE
100	α_0	-1.00	$-1.47\,(1.29)$	0.47	1.87	$-0.20\,(0.89)$	-0.79	1.41	$0.69\,(0.77)$	-1.69	3.47
	α_1	1.00	$1.24\,(0.91)$	0.24	0.88	$0.67\,(0.51)$	-0.32	0.36	$0.42\,(0.37)$	-0.57	0.47
	α_2	0.00	$0.06\,(0.27)$	$*$	0.08	$-0.09\,(0.17)$	$*$	0.04	$-0.32\,(0.19)$	$*$	0.14
	τ_1	0.00	$-0.03\,(1.10)$	$*$	1.20	$0.10\,(0.64)$	$*$	0.42	$0.59\,(0.51)$	$*$	0.61
	τ_2	-1.00	$-1.61\,(3.28)$	0.61	11.05	$-0.82\,(0.48)$	-0.17	0.26	$-0.37\,(0.26)$	-0.62	0.46
	β_0	-3.00	$-3.13\,(0.45)$	0.04	0.21	$-2.73\,(0.50)$	-0.08	0.32	$-2.13\,(0.63)$	-0.28	1.14
	β_1	1.00	$1.03\,(0.26)$	0.03	0.06	$0.98\,(0.22)$	-0.01	0.05	$0.87\,(0.28)$	-0.12	0.09
	β_2	1.00	$1.02\,(0.09)$	0.02	0.00	$0.95\,(0.10)$	-0.04	0.01	$0.84\,(0.14)$	-0.15	0.04
	γ_1	-1.00	$-0.97\,(0.21)$	-0.02	0.04	$-0.95\,(0.20)$	-0.04	0.04	$-0.85\,(0.28)$	-0.14	0.09
	γ_2	0.00	$-0.00\,(0.03)$	$*$	0.00	$-0.00\,(0.03)$	$*$	0.00	$0.00\,(0.04)$	$*$	0.00
	ϕ	1.00	$1.11\,(0.25)$	0.11	0.07	$0.22\,(0.03)$	-0.77	0.60	—	—	—
500	α_0	-1.00	$-1.00\,(0.50)$	0.00	0.24	$-0.07\,(0.37)$	-0.92	1.00	$0.72\,(0.34)$	-1.72	3.10
	α_1	1.00	$1.03\,(0.33)$	0.00	0.11	$0.66\,(0.22)$	-0.33	0.16	$0.34\,(0.18)$	-0.65	0.46
	α_2	0.00	$-0.01\,(0.10)$	$*$	0.01	$-0.15\,(0.08)$	$*$	0.03	$-0.33\,(0.07)$	$*$	0.11
	τ_1	0.00	$0.02\,(0.34)$	$*$	0.11	$0.25\,(0.24)$	$*$	0.12	$0.69\,(0.20)$	$*$	0.52
	τ_2	-1.00	$-0.99\,(0.29)$	0.00	0.08	$-0.73\,(0.20)$	-0.26	0.11	$-0.29\,(0.10)$	-0.70	0.50
	β_0	-3.00	$-3.00\,(0.20)$	0.00	0.04	$-2.65\,(0.20)$	-0.11	0.16	$-2.20\,(0.33)$	-0.26	0.75
	β_1	1.00	$1.00\,(0.07)$	0.00	0.00	$0.95\,(0.08)$	-0.04	0.00	$0.88\,(0.13)$	-0.11	0.03
	β_2	1.00	$0.99\,(0.04)$	0.00	0.00	$0.93\,(0.04)$	-0.06	0.00	$0.86\,(0.08)$	-0.13	0.02
	γ_1	-1.00	$-0.99\,(0.09)$	$-.00$	0.00	$-0.91\,(0.09)$	-0.08	0.01	$-0.86\,(0.13)$	-0.13	0.03
	γ_2	0.00	$-0.00\,(0.01)$	$*$	0.00	$-0.00\,(0.01)$	$*$	0.00	$0.00\,(0.03)$	$*$	0.00
	ϕ	1.00	$1.01\,(0.10)$	0.00	0.01	$0.23\,(0.01)$	-0.76	0.57	—	—	—

4.4. Zero-inflated underdispersion generalized Poisson model

For investigating the performance of the proposed transition model, the data set of this subsection are generated from a zero-inflated underdispersed generalized Poisson model and the performance of the ZIGP, ZINB and ZIP models are compared. The data set are generated from a $ZIGP(\lambda_{ij},\pi_{ij},\phi)$ such that $\log(\lambda_{i1})=\beta_0$, $\text{logit}(p_{i1})=\alpha_0$, $\log(\lambda_{ij})=\beta_0+\gamma_1 I_{\{0\}}(Y_{i,j-1})+\gamma_2 y_{i,j-1}(1-I_{\{0\}}(Y_{i,j-1}))$, $j=2,3,4$, $\text{logit}(p_{ij})=\alpha_0+\tau_1 I_{\{0\}}(Y_{i,j-1})+\tau_2 y_{i,j-1}(1-I_{\{0\}}(Y_{i,j-1}))$, $j=2,3,4$, where $\alpha_0=-1$, $\tau_1=-1$, $\tau_2=1$, $\beta_0=1$, $\gamma_1=0$, $\gamma_2=-1$ and $\phi=-0.3$. Also, two sample sizes N=500 and 1000 are selected where $M=1000$ iterations are performed. The results of this simulation study are summarized in Table 5. These results show the well performance of the ZIGP model as the best fitting model while the performance of ZINB model is poor. Also, the results show that the performance of ZIP is better than those of ZINB model. Note that the underdispersion rarely occur in practice. The well performance of ZIGP model are only satisfied in large sample size as described in this simulation study.

Table 5: Results of simulation study for generated data under ZIGP model in the presence of underdispersion, estimate (Est.), standard error (S.E.), relative bias (Bias) and mean square error (MSE) for $M = 1000$ simulated data with sample sizes 500 and 1000.

N	Para.	Real	ZIGP			ZINB			ZIP		
			Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE	Est. (S.E.)	Bias	MSE
500	α_0	-1.00	$-0.97\,(0.11)$	-0.03	0.01	$<-10^3$ $(>10^3)$	3618.73	$>10^3$	$-4.47\,(4.31)$	3.47	29.19
	τ_1	-1.00	$-1.03\,(0.15)$	0.03	0.02	$<-10^3$ $(>10^3)$	12854.74	$>10^3$	$-18.20\,(9.03)$	17.20	371.03
	τ_2	1.00	$1.15\,(0.17)$	0.15	0.15	$<-10^3$ $(>10^3)$	-63370.51	$>10^3$	$-11.15\,(6.46)$	-12.15	186.23
	β_0	1.00	$0.93\,(0.01)$	-0.07	0.14	$0.84\,(0.05)$	-0.18	0.04	$0.84\,(0.10)$	-0.18	0.15
	γ_1	0.00	$0.00\,(0.01)$	$*$	0.00	$0.16\,(0.04)$	$*$	0.03	$0.13\,(0.05)$	$*$	0.02
	γ_2	-1.00	$-1.20\,(0.75)$	0.20	0.65	$-1.55\,(0.33)$	0.55	0.50	$-1.39\,(0.45)$	0.39	0.46
	ϕ	-0.30	$-0.37\,(0.00)$	0.24	0.03	$>10^3$ $(>10^3)$	$<-10^3$	$>10^3$	—	—	—
1000	α_0	-1.00	$-0.97\,(0.05)$	-0.03	0.00	$<-10^3$ $(>10^3)$	3537.27	$>10^3$	$-3.78\,(3.78)$	2.78	21.80
	τ_1	-1.00	$-1.03\,(0.10)$	0.03	0.01	$<-10^3$ $(>10^3)$	12912.70	$>10^3$	$-18.35\,(4.68)$	17.35	322.60
	τ_2	1.00	$1.06\,(0.09)$	0.06	0.08	$<-10^3$ $(>10^3)$	-83788.10	$>10^3$	$-13.00\,(6.35)$	-14.00	235.61
	β_0	1.00	$0.98\,(0.01)$	-0.02	0.04	$0.92\,(0.06)$	-0.08	0.06	$0.95\,(0.05)$	-0.05	0.06
	γ_1	0.00	$0.00\,(0.01)$	$*$	0.00	$0.17\,(0.03)$	$*$	0.03	$0.14\,(0.02)$	$*$	0.02
	γ_2	-1.00	$-1.12\,(0.32)$	0.12	0.29	$-1.47\,(0.40)$	0.47	0.31	$-1.52\,(0.35)$	0.52	0.44
	ϕ	-0.30	$-0.32\,(0.00)$	0.07	0.03	$>10^3$ $(>10^3)$	$<-10^3$	$>10^3$	—	—	—

5. APPLICATION

The data set of this paper is extracted from a longitudinal study on kidney transplant patients in Imam Khomeini hospital of Urmia in Iran. The data set contains some information about $N = 129$ patients who have kidney transplant in this hospital. The response variable in this study is the number of acute rejections which is count response with extra zeros. The data are recorded in one year period which contain the number of acute rejection each four months. The barchat of the response variable for each time point (month 4, 8 and 12) is showed in Figure 1. In this figure, $Y_k,\ k = 1, 2, 3$, is used for indicating the response variable at the k^{th} time point. The number of extra zeros is clear in these charts.

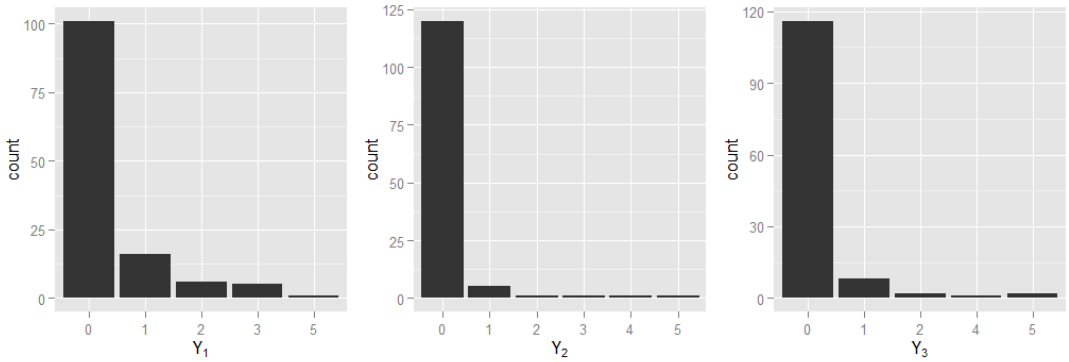

Figure 1: Barcharts of the number of acute rejections for time point at month 4 (first panel), month 8 (middle panel) and month 12 (third panel).

The collected explanatory variables which are considered in our analysis are creatinine index as a continuous covariate and having hyperacute rejection of kidney (rejection in the first 24 hours after surgery) as a categorical covariate. Figure 2 presents the boxplots of the creatinine index versus the number of acute rejections for each time. Also, Table 6 summarizes frequency of the number of acute rejections for each category of this variable for each time point.

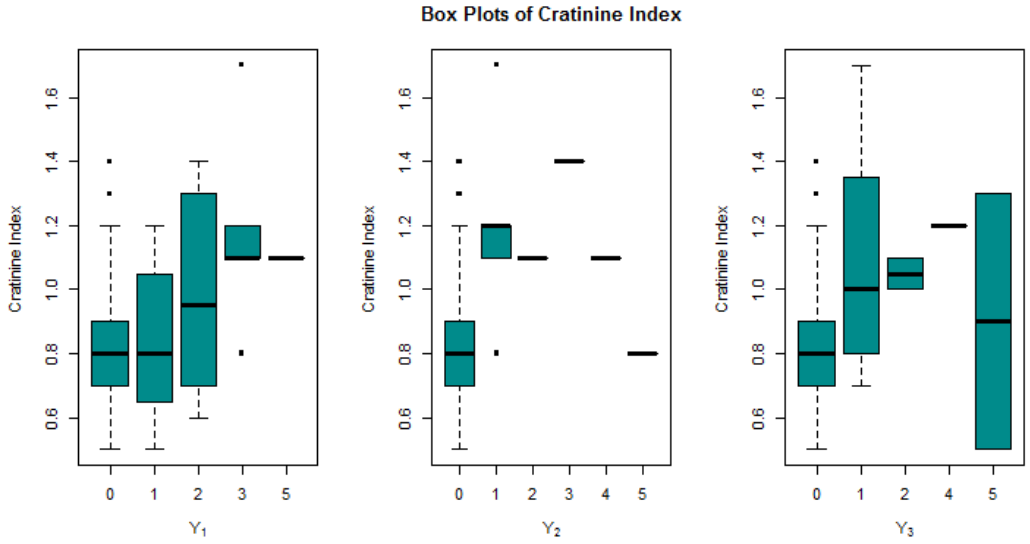

Figure 2: Boxplots of creatinine index versus the number of acute rejections for all time points.

For analyzing this data set, we use the proposed zero-inflated generalized Poisson transition model, also, Poisson (PM), negative binomial (NBM), generalized Poisson (GPM), zero-inflated Poisson (ZIPM), zero-inflated generalized Poisson (ZIGPM) and zero-inflated negative binomial (ZINBM) models under the transition structure are used for analyzing the data set. The explanatory variables which are considered for analysing the data are creatinine index (CRAT), having early acute rejection (EAR) and time ($t = 4, 8, 12$).

Table 6: Frequency of early acute rejection of kidney on the total number of acute rejection at each time point. "Yes" is used for having early acute rejection and "No" is used for not having early acute rejection.

Number	Early acute rejection					
	1st time point		2nd time point		3rd time point	
	Yes	No	Yes	No	Yes	No
0	19	82	31	89	28	88
1	7	9	1	4	5	3
2	5	1	1	0	1	1
3	3	2	1	0	0	0
4	0	0	1	0	0	1
5	1	0	0	1	1	1

We consider models (4.3) for analysing this data, where

$$\begin{aligned}\log(\lambda_{ij}) \;=\; & \beta_0 + \beta_1 CRAT_i + \beta_2 Time_j + \beta_3 EAR_i \\ & + \; \gamma_1 I_{\{0\}}(Y_{i,j-1}) + \gamma_2\big(1 - I_{\{0\}}(Y_{i,j-1})\big) y_{i,j-1}\,,\end{aligned} \tag{5.1}$$

and

$$\begin{aligned}\operatorname{logit}(\pi_{ij}) \;=\; & \alpha_0 + \alpha_1 CRAT_i + \alpha_2 Time_j + \alpha_3 EAR_i \\ & + \; \tau_1 I_{\{0\}}(Y_{i,j-1}) + \tau_2\big(1 - I_{\{0\}}(Y_{i,j-1})\big) y_{i,j-1}\,.\end{aligned} \tag{5.2}$$

For model comparison, we evaluate different model fits by considering some information criteria. These criteria are AIC, BIC and HQC, which are defined as follows:

Let $\boldsymbol{\theta}$ be the vector of unknown parameters, then

$$\begin{aligned} AIC \;&=\; -2\ell(\hat{\boldsymbol{\theta}}|\boldsymbol{Y}) + 2|\boldsymbol{\theta}|\,,\\ BIC \;&=\; -2\ell(\hat{\boldsymbol{\theta}}|\boldsymbol{Y}) + |\boldsymbol{\theta}|\ln(N)\,,\\ HQC \;&=\; -2\ell(\hat{\boldsymbol{\theta}}|\boldsymbol{Y}) + 2\ln(\ln(N))\,,\end{aligned}$$

where $|\boldsymbol{\theta}|$ is the number of unknown parameters in vector $\boldsymbol{\theta}$, N is the number of subjects and $\hat{\boldsymbol{\theta}}$ is the vector of parameters estimates. The smaller values of AIC, BIC and HQC indicate a better fitting model.

We use the EM algorithm, as described in Section 3, for parameters estimation of zero-inflated models, also, the usual maximum likelihood approach is used for parameter estimation of other models. In the EM algorithm, the initial values for unknown parameters were set equal to the estimates obtained by analysing separate models. The results of the above mentioned models are summarized in Table 7. This table contains parameter estimates and their standard errors for the first order transition model where standard deviations for zero-inflated models are estimated using a Bootstrap approach with 10000 iterations and for the others we use inverse of the Hessian matrix. The results show, based on the values of different criteria, that for this data set, the performance of ZIGP and ZINB models are similar and the difference between them is negligible. After them ZIP has the best fitting model and the worst fitting model based on these criteria is the PM. The results show some evidence for existence of mild overdispersion.

The results show that for zero-inflated models creatinine index (CRAT), having early acute rejection (EAR) and time are significant variables such that the more the creatinine index is, the larger is the estimated probability of nonzeros. Also, two covariates time and early acute rejection are positively significant, i. e. by increasing them the probability of zero increases. The results of zero-inflated models also show that only transition parameter τ_1 is significant. The results show that significant covariates in non-inflated models are similar to those in modeling zero probability in zero-inflated models, that is, the significant parameters in modeling zero probability of zero-inflated models have similar interpretation to those in modeling the rate of distributions in non-inflated models. Also, ϕ and τ_1 are the other significant parameters in these models.

Note that in a first order transition model the first response of each individual should be modeled given its previous response which is not recorded. How to face this issue, called

the initial condition problem [11, 10]. This problem does not exit in this study, because the number of acute rejections before the time of study is zero. In other words, the patients have been entered in the study from the time of kidney transplant and they have been followed for one year. Also, in this paper, we consider the first order transition model for modeling the data set, because the number of replications in our real data is three and a first order transition model for considering between-group dependence in data is adequate.

Table 7: Results of fitting (parameter estimations and standard errors in parenthesis) the Poisson model (PM), negative binomial model (NBM), generalized Poisson model (GPM), zero-inflated Poisson model (ZIPM), zero-inflated generalized Poisson model (ZIGPM) and zero-inflated negative binomial model (ZINBM) to kidney transplant study (significant parameters are highlighted in bold).

Parameter	ZIGPM	ZINBM	ZIPM	GPM	NBM	PM
	Est. (S.E.)	Est. (S.E.)	Est. (S.E.)	Est. (S.E.)	Est. (S.E.)	Est. (S.E.)
α_0	**2.66 (1.13)**	**2.62 (1.16)**	**3.10 (0.91)**	—	—	—
α_1 (CRAT)	**−3.43 (1.09)**	**−3.42 (1.09)**	**−3.35 (0.92)**	—	—	—
α_2 (Time)	**0.11 (0.04)**	**0.12 (0.05)**	**0.11 (0.05)**	—	—	—
α_3 (EAR)	**1.10 (0.58)**	**1.12 (0.59)**	**1.01 (0.44)**	—	—	—
β_0	−0.71 (0.82)	−0.72 (0.83)	−0.18 (0.54)	**−3.57 (0.94)**	**−3.19 (0.77)**	**−2.49 (0.46)**
β_1 (CRAT)	0.55 (0.67)	0.58 (0.69)	0.32 (0.52)	**2.19 (0.73)**	**2.08 (0.66)**	**1.99 (0.37)**
β_2 (Time)	−0.25 (0.40)	−0.24 (0.39)	−0.21 (0.29)	**−0.93 (0.34)**	**−0.87 (0.32)**	**−0.79 (0.23)**
β_3 (EAR)	0.79 (1.35)	0.70 (1.28)	0.60 (0.98)	**3.03 (1.93)**	**2.10 (1.27)**	**0.41 (0.65)**
τ_1	**1.83 (0.62)**	**1.85 (0.63)**	**1.69 (0.49)**	—	—	—
τ_2	0.04 (0.31)	0.04 (0.32)	0.02 (0.25)	—	—	—
γ_1	−0.35 (1.01)	−0.29 (0.96)	−0.24 (0.72)	**−2.82 (1.18)**	**−2.27 (0.78)**	**−1.34 (0.41)**
γ_2	0.00 (0.19)	0.00 (0.19)	−0.01 (0.14)	−0.17 (0.28)	−0.13 (0.23)	−0.02 (0.14)
ϕ	**0.21 (0.07)**	**2.11 (0.99)**	—	**1.22 (0.36)**	**0.34 (0.10)**	—
AIC	384.04	384.98	386.19	392.96	392.51	441.87
BIC	419.16	419.20	417.64	412.98	412.53	459.03
HQC	364.00	364.05	367.35	382.13	381.67	433.03

6. CONCLUSION AND DISCUSSION

In this paper, we have discussed a new transition model for analysing longitudinal outcomes with extra zeros. We compare the performance of different distributional assumptions: zero-inflated generalized Poisson, zero-inflated negative binomial and zero-inflated Poisson and we conclude that zero-inflated generalized Poisson is a flexible distributional assumption.

We have used the EM algorithm for parameter estimation. For illustration of the proposed models some simulation studies have been conducted. Also, a real data set of a kidney allograft rejection study has been analyzed as an illustrative example. Based on the results the creatinine index, having early acute rejection and time are significant covariates such that the more the creatinine index is, the larger is the estimated probability of nonzeros acute rejection. Also, two covariates time and early acute rejection are positively significant, i. e. by increasing them the probability of zero acute rejection increases. The results show

that the significant parameters in modeling zero probability of zero-inflated models have similar effect to parameters in the modeling rate of distributions in non-inflated models. We have considered a first order transition model for considering within-group dependence in longitudinal measurements, because the number of repeated longitudinal measurements has been small in our real data set. As a future work, illustration of the proposed approach for higher order of transition model for analyzing data set with larger number of repeated measures and comparison of the performance of it with that of the first order transition model may be performed. For this purpose (3.2) and (3.4) can be improve to be $\log(\lambda_{i1}) = \boldsymbol{x}'_{i1}\boldsymbol{\beta}$, $\text{logit}(\pi_{i1}) = \boldsymbol{z}'_{i1}\boldsymbol{\alpha}$, $\log(\lambda_{ij}) = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{\gamma}'\boldsymbol{h}_{i,j}$ and $\text{logit}(\pi_{ij}) = \boldsymbol{z}'_{ij}\boldsymbol{\alpha} + \boldsymbol{\tau}'\boldsymbol{h}_{i,j}, j = 2, ..., n_i$. Another parameterizations for λ_{ij} and π_{ij} of (3.2) and (3.4) may be the use of the first order transition model along with some random effects, that is, $\log(\lambda_{i1}) = \boldsymbol{x}'_{i1}\boldsymbol{\beta} + b_{i1}$, $\text{logit}(\pi_{i1}) = \boldsymbol{z}'_{i1}\boldsymbol{\alpha} + b_{i2}$, $\log(\lambda_{ij}) = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \gamma_1 I_{\{0\}}(Y_{i,j-1}) + \gamma_2 y_{i,j-1}(1 - I_{\{0\}}(Y_{i,j-1})) + b_{i1}$ and $\text{logit}(\pi_{ij}) = \boldsymbol{z}'_{ij}\boldsymbol{\alpha} + \tau_1 I_{\{0\}}(Y_{i,j-1}) + \tau_2 y_{i,j-1}(1 - I_{\{0\}}(Y_{i,j-1})) + b_{i2}, j = 2, ..., n_i$. where $\boldsymbol{b}_i = (b_{i1}, b_{i2})'$ is a bivariate random effects. As a parameterization for the random effects, one can write $b_{1i} \sim N(0, \sigma_1^2)$, $b_{2i}|b_{1i} \sim N(\psi b_{1i}, \sigma_2^2)$. We have used the EM algorithm for parameter estimation, one can use a Bayesian paradigm using MCMC for parameter estimation [29]. The priors elicitation are an important issue for performing this paradigm. The data set which analyzed in this paper has not had any missing values. The proposed method can be extended for modeling data sets in the presence of missing values as a future work. For this purpose, an ignorable or non-ignorable missing mechanism should be selected. The modeling of missing data mechanism for modeling non-ignorable missing data mechanism is necessary and these a sensitivity analysis is commonly suggested.

ACKNOWLEDGMENTS

This work has been supported by the grant number 96000139 from Iranian National Science Foundation (INSF). The authors would like to thank the INSF. The authors also acknowledge the valuable suggestions from the referees.

REFERENCES

[1] Agresti, A. (1999). Modeling ordered categorical data: recent advances and future challenges, *Statistics in medicine*, **18**, 2191–2207.

[2] Alfo, M. and Maruotti, A. (2014). Two-part regression models for longitudinal zero-inflated count data, *The Canadian Journal of Statistics*, **38**, 197–216.

[3] Buu, A.; Li, R.; Tan, X. and Zucker, R.A. (2012). Statistical models for longitudinal zero-inflated count data with applications to the substance abuse field, *Statistics in medicine*, **31**(29), 4074–4086.

[4] Cheung, Y.B. (2002). Zero-inflated models for regression analysis of count data: a study of growth and development, *Statistics in Medicine*, **21**, 1461–1469.

[5] Consul, P.C. (1989). *Generalized Poisson distribution: Properties and Applications*, Marcel Dekker, New York.

[6] Consul, P.C. and Famoye, F. (1992). Generalized Poisson regression model, *Communications in Statistics – Theory and Methods*, **21**, 81–109.

[7] Czado, C.; Erhardt, V.; Min, A. and Wagner, S. (2007). Zero-inflated generalized Poisson models with regression effects on the mean dispersion and zero-inflation level applied to patent outsourcing rates, *Statistical Modelling*, **7**(2), 125–153.

[8] Dempster, A.P.; Laird, N.M. and Rubin, D.B. (1977). Maximum likelihood for incomplete data via the EM algorithm, *Journal of the Royal Statistical Society B*, **39**, 1–38.

[9] Diggle, P.J.; Heagerty, P.; Liang, K.Y. and Zeger, S.L. (2002). *Analysis of Longitudinal Data*, Oxford: Oxford University Press.

[10] Ganjali, M.; Baghfalaki, T. and Ghahrodi, Z.R. (2017). Transitional Ordinal Modeling, *Wiley StatsRef: Statistics Reference Online*, Free Trial, 1–13.

[11] Ganjali, M. and Rezaee, Z. (2007). A Transition Model for Analysis of Repeated Measure Ordinal Response Data to Identify the Effects of Different Treatments, *Drug Information Journal*, **41**, 527–534.

[12] Hall, D.B. (2000). Zero-Inflated Poisson and Binomial Regression with Random Effects: A Case Study, *Biometrics*, **56**, 1030–1039.

[13] Harris, T.; Yang, Z. and Hardin, J.W. (2012). Modeling underdispersed count data with generalized Poisson regression, *Stata Journal*, **12**(1), 736–747.

[14] Hasan, M.T. and Sneddon, G. (2009). Zero-Inflated Poisson Regression for Longitudinal Data, *Communications in StatisticsÛSimulation and Computation*, **38**(3), 638–653.

[15] Heilbron, D.C. (1994). Zero-altered and other regression models for count data with added zeros, *Biometrical Journal*, **36**, 531–347.

[16] Hu, M.C.; Pavlicova, M. and Nunes, E.V. (2011). Zero-inflated and hurdle models of count data with extra zeros: examples from an HIV-risk reduction intervention trial, *The American Journal of Drug and Alcohol Abuse*, **37**, 367–375.

[17] Joe, H. and Zhu, R. (2005). Generalized Poisson distribution: the property of mixture of Poisson and comparison with negative binomial distribution, *Biometrical Journal*, **47**(2), 219–229.

[18] Lall, R.; Campbell, M.J.; Walters, S.J. and Morgan, K. (2002). A review of ordinal regression models applied on health-related quality of life assessments, *Statistical Methods in Medical Research*, **11**, 49–67.

[19] Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing, *Technometrics*, **34**, 1–14.

[20] Lange, K. (2004). *Optimization*, Springer-Verlag, New York.

[21] Lewsey, J.D. and Thomson, W.M. (2004). The utility of the zero-inflated Poisson and zero-inflated negative binomial models: a case study of cross-sectional and longitudinal DMF data examining the effect of socio-economic status, *Community of Dentistry and Oral Epidemiology*, **32**, 183–189.

[22] Maruottiab, A. and Raponic, V. (2014). On Baseline Conditions for Zero-Inflated Longitudinal Count Data, *Communications in Statistics – Simulation and Computation*, **43**, 743–760.

[23] Min, Y. and Agresti, A. (2005). Random effect models for repeated measures of zero-inflated count data, *Statistical Modeling*, **5**, 1–19.

[24] Molenberghs, G. and Verbeke, G. (2005). *Models for Discrete Longitudinal Data*, Springer-Verlag.

[25] Mullahy, J. (1986). Specification and testing of some modified count data models, *Journal of Econometrics*, **33**, 341–365.

[26] Neelon, B.H.; OMalley, A.J. and Normand, S.L. (2010). A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use, *Statistical Modelling*, **10**, 421–439.

[27] Reinecke, J. and Seddig, D. (2011). Growth mixture models in longitudinal research, *AStA Advances in Statistical Analysis*, **95**(4), 415–434.

[28] Rose, C.E.; Martin, S.W.; Wannemuehler, K.A. and Plikaytis, B.D. (2006). On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data, *Journal of Biopharmaceutical Statistics*, **16**, 463–481.

[29] Silva, G.L.; Juarez-Colunga, E. and Dean, C. (2015). A joint analysis of counts and severity with zero-inflated longitudinal data, *CEB-EIB 2015*, Bilbao, 23–25 September.

[30] Song, P.X.K. (2007). *Correlated Data Analysis*, Springer-Verlag, New York.