# Longitudinal Mixed Models with *t* Random Effects for Repeated Count and Binary Data

**R. Prabhakar Rao, Brajendra C. Sutradhar, and V.N. Pandit**

**Abstract** Unlike the estimation for the parameters in a linear longitudinal mixed model with independent *t* errors, the estimation of parameters of a generalized linear longitudinal mixed model (GLLMM) for discrete such as count and binary data with independent *t* random effects involved in the linear predictor of the model, may be challenging. The main difficulty arises in the estimation of the degrees of freedom parameter of the *t* distribution of the random effects involved in such models for discrete data. This is because, when the random effects follow a heavy tailed t-distribution, one can no longer compute the basic properties analytically, because of the fact that moment generating function of the *t* random variable is unknown or can not be computed, even though characteristic function exists and can be computed. In this paper, we develop a simulations based numerical approach to resolve this issue. The parameters involved in the numerically computed unconditional mean, variance and correlations are estimated by using the well known generalized quasi-likelihood (GQL) and method of moments approach. It is demonstrated that the marginal GQL estimator for the regression effects asymptotically follow a multivariate Gaussian distribution. The asymptotic properties of the estimators for the rest of the parameters are also indicated.

**Keywords** Asymptotic normal distribution • Consistent estimation • Count and binary panel data • Generalized quasi-likelihood • Regression effects • *t* random effects • Simulating t observations • Stationary and non-stationary covariates • Unconditional mean • Variance and correlations

---

R. Prabhakar Rao (✉) • V.N. Pandit
Department of Economics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam, Andhra Pradesh, India
e-mail: rprabhakarrao@gmail.com; vnpandit@gmail.com

B.C. Sutradhar
Department of Mathematics and Statistics, Memorial University, St. John's, NL, Canada A1C5S7
e-mail: bsutradh@mun.ca



B.C. Sutradhar (ed.), *Advances and Challenges in Parametric and Semi-parametric Analysis for Correlated Data*, Lecture Notes in Statistics 218,
DOI 10.1007/978-3-319-31260-6_2

## 1 Introduction

Let $(y_{i1},\ldots,y_{it},\ldots,y_{iT})$ denote the $T$ repeated count or binary responses for the *ith* subject, $i=1,\ldots,K$. Also, let $x_{it}$ be the $p\times 1$ vector of covariates corresponding to $y_{it}$, and $\beta$ is the $p\times 1$ regression effects of $x_{it}$ on $y_{it}$. Next suppose that in addition to $x_{it}$, the repeated responses of the $i$th individual are also influenced by one random effect $\gamma_i^*$. Conditional on this random effect $\gamma_i^*$, some authors have modeled the longitudinal correlations of the repeated counts and binary data by using lag 1 dynamic relationships. More specifically, Sutradhar and Bari (2007) have used an AR(1) (auto-regressive order 1) type dynamic relationship to model the longitudinal correlations for repeated count data. Similarly, Sutradhar et al. (2008) [see also Amemiya (1985, p. 353), Manski (1987), and Honore and Kyriazidou (2000, p. 84)] have used a lag 1 dynamic binary mixed logit (BDML) model to accommodate the correlations of the repeated binary data. The unconditional correlation structures in both of these papers have been computed under the normality assumption for the random effects, specifically correlations are obtained by assuming that $\gamma_i^* \stackrel{iid}{\sim} N(0,\sigma_\gamma^2)$. For convenience, we provide these correlation structures in brief for count and binary data as follows.

### 1.1 Conditional and Unconditional (Normality Based) Correlation Structures for Repeated Count Data

Suppose that

$$\begin{aligned} y_{i1}|\gamma_i^* &\sim \text{Poi}(\mu_{i1}^*) \;\text{ with }\; \mu_{i1}^* = \exp(x_{i1}'\beta + \gamma_i^*) \\ y_{it}|\gamma_i^* &= \rho \circ [y_{i,t-1}|\gamma_i^*] + [d_{it}|\gamma_i^*],\; \text{for } t=2,\ldots,T, \end{aligned} \tag{1}$$

where $\text{Poi}(\mu_{it}^*)$ refers to the Poisson distribution with mean parameter $\mu_{it}^*$, and $\rho\circ y_{i,t-1} = \sum_{s=1}^{y_{i,t-1}} b_s(\rho)$ with $Pr[b_s(\rho)=1]=\rho$, $Pr[b_s(\rho)=0]=1-\rho$, and $[d_{it}|\gamma_i^*] \sim Poi(\mu_{it}^* - \rho\mu_{i,t-1}^*)$, with $\mu_{it}^* = \exp(x_{it}'\beta + \sigma_\gamma\gamma_i^*)$. This model in (1) is referred to as the Poisson AR(1) model which produces the correlation between $y_{iu}$ and $y_{it}$ as

$$\text{corr}(Y_{iu}, Y_{it}|\gamma_i^*) = \rho^{|t-u|}\left[\frac{\mu_{iu}^*}{\mu_{it}^*}\right]^{\frac{1}{2}}, \tag{2}$$

which is free from $\gamma_i^*$, but depends on the time dependent covariates and on $\rho$, a correlation index parameter.

Note that the likelihood inference for the AR(1) model (1) is extremely complicated. This is because under this model, one writes

$$f((y_{i1},\ldots,y_{it},\ldots,y_{iT})|\gamma_i^*) = f(y_{i1}|\gamma_i^*)\Pi_{t=2}^T[f_{it|t-1}(y_{it}|y_{i,t-1},\gamma_i^*)] \tag{3}$$

where the conditional distribution, namely $f_{it|t-1}(y_{it}|y_{i,t-1},\gamma_i^*)$ has a complicated form given by

$$f_{it|t-1}(y_{it}|y_{i,t-1},\gamma_i^*) = \exp[-(\mu_{it}^* - \rho\mu_{i,t-1}^*)] \times \sum_{s_{it}=0}^{\min(y_{it},y_{i,t-1})} \frac{y_{i,t-1}!\rho^{s_{it}}(1-\rho)^{y_{it}-s_{it}}(\mu_{it}^* - \rho\mu_{i,t-1}^*)^{y_{it}-s_{it}}}{s_{it}!(y_{i,t-1}-s_{it})!(y_{it}-s_{it})!} \quad (4)$$

(Freeland and McCabe 2004). Furthermore, the integration of the conditional likelihood function (3) over the Gaussian distribution of the random effects, i.e., $\gamma_i^* \overset{iid}{\sim} N(0,\sigma_\gamma^2)$, is an additional complex problem. As opposed to the generalized linear longitudinal mixed model (GLLMM) setup, Over the last two decades many researchers, for example, Breslow and Clayton (1993), Lee and Nelder (1996), Jiang (1998), Sutradhar (2004), among others have used the normality assumption for the random effects in a generalized linear mixed model (GLMM) setup, and discussed the estimation of $\beta$ and $\sigma_\gamma^2$. In the present GLLMM setup (1)–(2), there is an additional correlation index parameter $\rho$ to estimate.

When the normality assumption for the random effect $\gamma_i^*$ is used in the count panel data setup, the mean, variance and correlations of the repeated counts contain three unknown parameters, namely $\beta$, $\sigma_\gamma^2$, and $\rho$. To be specific, by using the moment generating function (mgf) of $\gamma_i^* \overset{iid}{\sim} N(0,\sigma_\gamma^2)$, that is, $E_{\gamma_i^*}(\exp(a\gamma_i^*)) = \exp[\frac{1}{2}a^2\sigma_\gamma^2]$, $a$ being an auxiliary parameter, one obtains the three basic properties of the count panel data as follows (see Sutradhar 2011, Sect. 8.1.1):

$$\mu_{it} = E[Y_{it}] = E_{\gamma_i^*}E[Y_{it}|\gamma_i^*] = \exp(x_{it}'\beta)E_{\gamma_i^*}(\exp(\gamma_i^*)) = \exp[x_{it}'\beta + \frac{1}{2}\sigma_\gamma^2] \quad (5)$$

$$\begin{aligned}\sigma_{itt} &= \text{var}[Y_{it}] = E_{\gamma_i^*}\text{var}[Y_{it}|\gamma_i^*] + \text{var}_{\gamma_i^*}E[Y_{it}|\gamma_i^*] = E_{\gamma_i^*}\mu_{it}^* + \text{var}_{\gamma_i^*}(\mu_{it}^*)\\ &= \exp(x_{it}'\beta)E_{\gamma_i^*}\exp(\gamma_i^*) + \exp(2x_{it}'\beta)\text{var}_{\gamma_i^*}(\exp(\gamma_i^*))\\ &= \mu_{it} + \exp(2x_{it}'\beta)[\exp(2\sigma_\gamma^2) - \exp(\sigma_\gamma^2)]\\ &= \mu_{it} + [\exp(\sigma_\gamma^2) - 1]\mu_{it}^2 \end{aligned} \quad (6)$$

and for $u < t$, the unconditional covariance between $y_{iu}$ and $y_{it}$, is given by

$$\begin{aligned}\sigma_{iut} &= \text{cov}[Y_{iu},Y_{it}] = E_{\gamma_i^*}[\text{cov}\{(Y_{iu},Y_{it})|\gamma_i^*\}] + \text{cov}_{\gamma_i^*}[\mu_{iu}^*,\mu_{it}^*]\\ &= \rho^{t-u}\exp(x_{iu}'\beta)E_{\gamma_i^*}[\exp(\gamma_i^*)] + \exp([x_{iu}+x_{it}]'\beta)\text{var}_{\gamma_i^*}\{\exp(\gamma_i^*)\}\\ &= \rho^{t-u}\mu_{iu} + [\exp(\sigma_\gamma^2) - 1]\mu_{iu}\mu_{it}, \end{aligned} \quad (7)$$

yielding the lag $t-u$ correlation

$$\mathrm{corr}(Y_{iu}, Y_{it}) = \frac{\rho^{t-u}\mu_{iu} + [\exp(\sigma_\gamma^2) - 1]\mu_{iu}\mu_{it}}{[\{\mu_{iu} + [\exp(\sigma_\gamma^2) - 1]\mu_{iu}^2\}\{\mu_{it} + [\exp(\sigma_\gamma^2) - 1]\mu_{it}^2\}]^{\frac{1}{2}}}. \tag{8}$$

Notice that the unconditional mean (5) and the unconditional variance (6) are functions in $\beta$ and $\sigma_\gamma^2$, whereas the unconditional covariances (7) and correlations (8) are functions in $\beta,\ \sigma_\gamma^2$, as well as the dynamic dependence or correlation index parameter $\rho$. Remark that Sutradhar and Bari (2007), among others, have exploited the aforementioned moments (5)–(8) to develop a four-moments based generalized quasi-likelihood (GQL) approach for the estimation of these parameters $\beta,\ \sigma_\gamma^2$, and $\rho$.

## *1.2 Conditional and Unconditional (Normality Based) Correlation Structures for Repeated Binary Data*

As indicated earlier, over the last three decades, many econometricians such as Heckman (1981), Amemiya (1985, p. 353), Manski (1987), and Honore and Kyriazidou (2000, p. 844) have made attempts to accommodate the dynamic nature of the repeated binary responses by using a binary dynamic mixed logit (BDML) model given by

$$Pr(y_{it} = 1|\gamma_i, y_{i,t-1}) = \begin{cases} \frac{\exp(x_{i1}^{'}\beta+\sigma_\gamma\gamma_i)}{1+\exp(x_{i1}^{'}\beta+\sigma_\gamma\gamma i)} & \text{for } t = 1 \\ \frac{\exp(x_{it}^{\prime}\beta+\theta y_{i,t-1}+\sigma_\gamma\gamma_i)}{1+\exp(x_{it}^{\prime}\beta+\theta y_{i,t-1}+\sigma_\gamma\gamma_i)} & \text{for } t = 2, \ldots, T, \end{cases} \tag{9}$$

where $\beta$ is the effect of the covariates similar to the Poisson model, $\theta$ is referred to as the dynamic dependence parameter, and $\gamma_i = [\gamma_i^*/\sigma_\gamma] \overset{iid}{\sim} (0, 1)$. Note that the distribution of $\gamma_i$ is unknown. Also note that even if it is assumed that $\gamma_i$ follows the Gaussian distribution, that is, $\gamma_i \overset{iid}{\sim} N(0, 1)$, obtaining the likelihood estimates for $\beta,\ \theta$, and $\sigma_\gamma^2$ is complicated. Honore and Kyriazidou (2000, p. 844) attempted to avoid the estimation difficulty by estimating the $\beta$ and $\theta$ parameters based on the transformed observations, such as the first differences of the responses $y_{i1}-y_{i0},\ y_{i2}-y_{i1}, \ldots$, which are approximately independent of $\gamma_i$. They have used an approximate weighted log likelihood estimation approach, which however puts some impractical restrictions on covariates such as assuming $x_{i3} = x_{i4}$, for the $T = 4$ case.

Remark that recently Bartolucci and Nigro (2010, Eq. (5), Sect. 3) have constructed a random effects free conditional likelihood for a binary model which is different from (9). More specifically, they exploited the conditional approach for a quadratic exponential type model (Cox 1972; Zhao and Prentice 1990) given by

$$Pr(y_{i1}\ldots,y_{iT}|\gamma_i^*,x_{i1}\ldots,x_{iT}) = \Delta_i^{-1}\exp[y_i'\xi_i + \theta g_i'(y_i)1_{T(T-1)/2} + c_i(y_i) + \gamma_i y_i' 1_T] \tag{10}$$

where $y_i = [y_{i1},\ldots,y_{iT}]'$, $g_i(y_i) = [y_{i1}y_{i2},\ldots,y_{iT-1}y_{iT}]'$, and $\xi_i = [\xi_{i1},\ldots,\xi_{iT}]'$, with $\xi_{it} = x_{it}'\beta$. In (10), $1_n$, for example, is an n-dimensional unit vector, $\Delta_i$ is a normalizing constant defined as

$$\Delta_i = \sum \exp[y_i'\xi_i + \theta g_i'(y_i)1_{T(T-1)/2} + c_i(y_i) + \gamma_i y_i' 1_T],$$

with summation overall $2^T$ possible values of $y_i$. Also in (10), $c_i(y_i)$ is referred to as a shape function that can be expressed as a linear combination of products of three or more of the elements of $y_i$. By ignoring $c_i(y_i)$, i.e., $c_i(y_i) = 0$, it can be shown that for a given total score $\sum_{t=1}^{T} y_{it} = y_{i+}$, the conditional distribution of $y_{i1},\ldots,y_{iT}$ may be written as

$$Pr(y_{i1},\ldots,y_{iT}|y_{i+},\gamma_i^*,x_{i1},\ldots,x_{iT}) = \Delta_i^{*-1}\exp\left[y_i'\xi_i + \theta\sum_{t=2}^{T} y_{i,t-1}y_{it}\right] \tag{11}$$

where $\Delta_i^* = \Delta_i$ evaluated at $\sum_{t=1}^{T} y_{it} = y_{i+}$, i.e., $y_{iT} = y_{i+} - \sum_{t=1}^{T-1} y_{it}$. Because the conditional distribution in (11) is free from $\gamma_i$, Bartolucci and Nigro (2010) used this conditional distribution to estimate the main parameters $\beta$ and $\theta$.

We now turn back to the desired binary dynamic mixed model (9). It is clear that even if one is interested to estimate $\beta$ and $\theta$, neither the aforementioned weighted likelihood approach of Honore and Kyriazidou (2000), nor the conditional likelihood approach of Bartolucci and Nigro (2010) can be used to remove the random effects from dynamic mixed model (9) for easier estimation of $\beta$ and $\theta$. Moreover, for binary panel data analysis following (9), one, in fact, is interested to understand the mean and variance of the data, which, however, can not be computed by removing the random effects $\gamma_i$ from the model. In stead, the computation of the moments require averaging over certain functions in $\gamma_i$ over its distribution. Thus, rather than making any attempt to remove $\gamma_i$ from (9), many authors such as Breslow and Clayton (1993), Lee and Nelder (1996), Jiang (1998), and Sutradhar (2004) have studied the inferences for the model (9) under the assumption that $\gamma_i^* \overset{iid}{\sim} N(0,\sigma_\gamma^2)$.

Under this normality assumption, one may obtain the conditional and unconditional means, variance and covariances as follows (see Sutradhar 2011, Sect. 9.2.1). First, conditional on $\gamma_i$, the means of the repeated binary responses under model (9) are given by

$$\pi^*_{it}(\gamma_i) = E[Y_{it}|\gamma_i] = \begin{cases} \frac{\exp(x'_{i1}\beta+\sigma_\gamma\gamma_i)}{1+\exp(x'_{i1}\beta+\sigma_\gamma\gamma_i)}, & \text{for } i=1,\ldots,K;\ t=1 \\ p_{it0}+\pi^*_{i,t-1}(p_{it1}-p_{it0}), & \text{for } i=1,\ldots,k;\ t=2,\ldots,T \end{cases} \tag{12}$$

where

$$p_{it1} = \frac{\exp(x'_{it}\beta+\theta+\sigma_\gamma\gamma_i)}{[1+\exp(x'_{it}\beta+\theta+\sigma_\gamma\gamma_i)]} \text{ and } p_{it0} = \frac{\exp(x'_{it}\beta+\sigma_\gamma\gamma_i)}{[1+\exp(x'_{it}\beta+\sigma_\gamma\gamma_i)]}.$$

Subsequently, one obtains the unconditional means as

$$\begin{aligned} \mu_{it} &= E(Y_{it}) = Pr(y_{it}=1) \\ &= M^{-1}\sum_{w=1}^{M}\pi^*_{it}(\gamma_{iw}) \\ &= M^{-1}\sum_{w=1}^{M}[p_{it0}+\pi^*_{i,t-1}(p_{it1}-p_{it0})]_{|\gamma_i=\gamma_{iw}} \end{aligned} \tag{13}$$

(Jiang 1998; Sutradhar 2004) where $\gamma_{iw}$ is the $w$th ($w=1,\ldots,M$) realized value of $\gamma_i$ generated from the standard normal distribution. Here $M$ is a sufficiently large number, such as $M = 5000$. By (12), the $p_{it1,w}$ involved in (13), for example, is written as

$$p_{it1,w} = \frac{\exp(x'_{it}\beta+\theta+\sigma_\gamma\gamma_{iw})}{[1+\exp(x'_{it}\beta+\theta+\sigma_\gamma\gamma_{iw})]}.$$

Next, conditional on $\gamma_i$, for $u < t$, the second-order expectation may be written as

$$E(Y_{iu}Y_{it}|\gamma_i) = \lambda^*_{iut}(\gamma_i) = \text{cov}(Y_{iu},Y_{it}|\gamma_i)+\pi_{iu}\pi_{it} = \sigma^*_{iut}+\pi^*_{iu}\pi^*_{it}, \tag{14}$$

where the conditional covariance between $y_{iu}$ and $y_{it}$, conditional on $\gamma_i$, has the formula

$$\sigma^*_{iut} = \text{cov}(Y_{iu},Y_{it}|\gamma_i) = \pi^*_{iu}(\gamma_i)(1-\pi^*_{iu}(\gamma_i))\Pi^t_{j=u+1}(p_{ij1}-p_{ij0}). \tag{15}$$

It then follows that the unconditional second-order raw moments have the formula

$$\phi_{iut} = E(Y_{iu}Y_{it}) = M^{-1}\sum_{w=1}^{M}\big[\pi^*_{iu}(\gamma_{iw})(1-\pi^*_{iu}(\gamma_{iw})) \times \Pi^t_{j=u+1}(p_{ij1,w}-p_{ij0,w}) + \pi^*_{iu}(\gamma_{iw})\pi^*_{it}(\gamma_{iw})\big], \quad (16)$$

yielding the unconditional covariance as

$$\sigma_{iut} = \phi_{iut} - \mu_{iu}\mu_{it}, \quad (17)$$

with $\mu_{it}$ is the unconditional mean given by (13).

### 1.3 Plan of the Paper Under the Proposed *t* Random Effects with Unknown Degrees of Freedom $\nu$

In this paper, as opposed to the Gaussian distribution, we consider a wider class of $t$ distributions for the random effects $\{\gamma_i^*\}$, with mean 0, a scale parameter $\lambda^2_\gamma$, and shape or degrees of freedom parameter $\nu$, i.e, $\gamma_i^* \overset{iid}{\sim} t_\nu(0,\lambda^2_\gamma,\nu)$, with its probability density given by

$$f(\gamma_i^*) = \frac{\nu^{\frac{1}{2}}\Gamma\frac{\nu+1}{2}}{\Gamma\frac{\nu}{2}}(\lambda^2_\gamma)^{\frac{-1}{2}}\left[\nu + \frac{\gamma_i^{*2}}{\lambda^2_\gamma}\right]^{-\frac{\nu+1}{2}}. \quad (18)$$

This $t$ distribution exhibits heavy symmetric tails when $\nu$ is small, and it reduces to the normal distribution $N(0,\sigma^2_\gamma)$ for $\nu \to \infty$. Note, however, that one can not compute the mgf, that is, $E_{\gamma_i^*}(\exp(a\gamma_i^*))$ under this $t$ distribution (18). As a remedy, the moments of this $t$ distribution (18) are computed either from the characteristic function (cf) (Sutradhar 1986) or by direct integrations over the distribution. For $\nu > 4$, the first four moments, for example, are given by

$$E(\gamma_i^*) = 0,\ \text{var}(\gamma_i^*) = \frac{\nu}{\nu-2}\lambda^2_\gamma = \sigma^2_\gamma$$
$$E(\gamma_i^{*3}) = 0,\ E(\gamma_i^{*4}) = \frac{3\lambda^4{}_\gamma\nu^2}{(\nu-2)(\nu-4)} = 3\sigma^4_\gamma[\frac{\nu-2}{\nu-4}]. \quad (19)$$

But, it follows from (5)–(8) that in the present longitudinal mixed model setup for count data with $\gamma_i^* \overset{iid}{\sim} t_\nu(0,\lambda^2_\gamma,\nu)$, one requires the result for the mgf $E_{\gamma_i^*}(\exp(a\gamma_i^*))$, which however can not be computed analytically under the $t_\nu$ distribution (18). A similar but different problem arises in the longitudinal mixed model setup for binary data, where for (13)–(15), one needs to generate random effect values $\gamma_{iw}$ from standard $t$ distribution $t(0, 1, \nu)$ with $\nu$ degrees of freedom which is however unknown in practice.

As a remedy, in this paper, we offer a simulation-based numerical approach to compute the mgf, and develop a GQL estimation approach for the estimation of all parameters of the models including the degrees of freedom parameter $\nu > 4$. More specifically, in Sects. 2 and 3, we discuss the Poisson mixed model with $t_\nu$ random effects and the desired inferences. The binary model and the inferences with $t_\nu$ random effects are provided in Sects. 4 and 5. Some concluding remarks are given in Sect. 6.

## 2 Poisson Mixed Model with $t_\nu$ Random Effects

### *2.1 Basic Properties of the Poisson Mixed Model: Unconditional Mean and Variance*

In the present setup, $\gamma_i^* \overset{iid}{\sim} t_\nu(0, \lambda_\gamma^2, \nu)$. Now because, similar to (5)–(6), the unconditional mean and variance have the formulas

$$\mu_{it} = E[Y_{it}] = E_{\gamma_i^*} E[Y_{it}|\gamma_i^*] = E_{\gamma_i^*} \mu_{it}^* = \exp(x_{it}'\beta) E_{\gamma_i^*} \{\exp(\gamma_i^*)\} \tag{20}$$

$$\sigma_{itt} = \text{var}[Y_{it}] = E_{\gamma_i^*} \text{var}[Y_{it}|\gamma_i^*] + \text{var}_{\gamma_i^*} E[Y_{it}|\gamma_i^*] \tag{21}$$

$$= \exp(x_{it}'\beta) E_{\gamma_i^*} \{\exp(\gamma_i^*)\} + \exp(2x_{it}'\beta) \left[ E_{\gamma_i^*} \{\exp(2\gamma_i^*)\} - \left[ E_{\gamma_i^*} \{\exp(\gamma_i^*)\} \right]^2 \right],$$

they could be evaluated numerically by simulating $\gamma_{iw},\ w = 1, \ldots, W$, for a large $W$ such as $W = 5000$, from $\gamma_{iw} \overset{iid}{\sim} t_\nu(0, 1, \nu)$, and using

$$E_{\gamma_i^*} \{\exp(a\gamma_i^*)\} = E_{\gamma_i} \{\exp(a\lambda_\gamma \gamma_i)\} \approx \frac{1}{W} \sum_{w=1}^{W} \left[\exp(a\lambda_\gamma \gamma_{iw})\right], \tag{22}$$

in (20)–(21) for $a = 1, 2$, provided $\nu$ were known. Note that for known $\nu$, this simulated approximation in (22) is quite similar to the simulation approximation used by Sutradhar (2008, Sect. 3) [see also Sutradhar et al. (2008, Eq. (2.6))] for the binary case with random effects generated from $N(0, 1)$ distribution. However, because in the present case, $\nu$ is unknown and requires to be estimated, we resolve this simulation issue by generating $\gamma_{iw},\ w = 1, \ldots, W$ first from a reference $t_4(0, 1, 4)$ distribution (equivalent to standard normal reference distribution) and using the transformation from following Lemma 2.3 so that these $\gamma_{iw}, w = 1, \ldots, W$ subsequently follow the $t_\nu(0, 1, \nu)$ distribution as desired. Lemmas 2.1 and 2.2 below are needed to write the Lemma 2.3.

**Lemma 2.1.** *Suppose that* $\psi_i^* \sim N_p(0, \lambda_\gamma^2)$. *Next, suppose that* $\xi_i^{*2}$ *a scalar random variable which follows the well known* $\chi_\nu^2$ *distribution with* $\nu$ *degrees of freedom, that is,* $\xi_i^{*2} \sim \chi_\nu^2$, *and* $\psi_i^*$ *and* $\xi_i^{*2}$ *are independent. Then, for* $\psi_i = \frac{\psi_i^*}{\lambda_\gamma}$, *the ratio variable* $\gamma_i^*$ *defined as*

$$\gamma_i^* = \lambda_\gamma \psi_i / [\sqrt{(\xi_i^{*2}/\nu)}] = \lambda_\gamma \gamma_i \tag{23}$$

*has the* $t_\nu(0, \lambda_\gamma^2, \nu)$ *distribution given by (18).*

However, even though $\psi_i$ in (23) is a parameter free normal variable, an observation $\gamma_{iw}^*$ following the *t*-distribution (18) for $\gamma_i^*$ (20)–(22), can not be drawn yet, because the distribution of $\xi_i^{*2}$ is parameter $\nu$ dependent. Because $\nu > 4$ in (18), to resolve this issue, we suggest to use a *t*-distribution with 4 degrees of freedom as a reference distribution. Suppose that $\xi_i^2$ is generated from this $\chi_4^2$ distribution. One may then generate a $\xi_i^{*2}$ from $\chi_\nu^2$ approximately for any $\nu > 4$, by using the relation between $\xi_i^2$ and $\xi_i^{*2}$ as in Lemma 2.2 below.

**Lemma 2.2.** *If* $\xi_i^2$ *is generated from the* $\chi_4^2$ *distribution, one may then generate* $\xi_i^{*2}$ *by using the relationship*

$$\xi_i^{*2} = \sqrt{2\nu}\left[\frac{\xi_i^2 - 4}{\sqrt{8}}\right] + \nu = \frac{1}{2}\sqrt{\nu}\left[\xi_i^2 - 4\right] + \nu, \tag{24}$$

*which has the same first two moments as that of* $\chi_\nu^2$.

One may then generate an observation from a $t_\nu$ distribution as in Lemma 2.3.

**Lemma 2.3.** *For* $w = 1, \ldots, W$, *with* $W = 5000$ (*say*), *the w-th observation* $\gamma_{iw}^*$ *from the* $t_\nu$ *distribution may be generated by applying Lemma 2.2 to Lemma 2.1. That is,*

$$\gamma_{iw}^* = \lambda_\gamma \{2\nu\}^{\frac{1}{2}} \frac{\psi_{iw}}{\left[\sqrt{\nu}\left(\xi_{iw}^2 - 4\right) + 2\nu\right]^{\frac{1}{2}}} = \lambda_\gamma \gamma_{iw}, \tag{25}$$

*where* $\psi_{iw}$ *and* $\xi_{iw}^2$ *are observations from the standard normal* $N(0,1)$ *and* $\chi_4^2$ *distributions, respectively.*

Consequently, by applying (25) under Lemma 2.3, to (22) and (20), one computes the unconditional mean as

$$\begin{aligned}\mu_{it}(\beta, \lambda_\gamma, \nu) = E[Y_{it}] &= \exp(x_{it}'\beta) E_{\gamma_i^*}\left\{\exp(\gamma_i^*)\right\} \\ &= \exp(x_{it}'\beta) E_{\gamma_i^*}\left\{\exp(\lambda_\gamma \gamma_i)\right\}\end{aligned}$$

$$
\begin{aligned}
&= \exp(x'_{it}\beta)\frac{1}{W}\sum_{w=1}^{W}\{\exp(\lambda_\gamma \gamma_{iw})\} \\
&= \frac{1}{W}\exp(x'_{it}\beta)\sum_{w=1}^{W}\exp\left[\frac{\lambda_\gamma \psi_{iw}}{\left\{\frac{1}{2\sqrt{(\nu)}}(\xi_{iw}^2 - 4) + 1\right\}}\right] \\
&= \frac{1}{W}\exp(x'_{it}\beta)\sum_{w=1}^{W}\exp\left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}, \qquad (26)
\end{aligned}
$$

where for $w = 1, \ldots, W$, $\psi_{iw}$ are generated from standard normal $N(0, 1)$ distribution, and $\xi_{iw}^2$ are generated from $\chi_4^2$ distributions. Furthermore, $\psi_{iw}$ and $\xi_{iw}^2$ are independent.

In order to compute the unconditional variance, use $\mu_{it}^* = E_{\gamma_i^*}\left[\exp(x'_{it}\beta + \gamma_i^*)\right]$ from (20), and first compute $\phi_{i,tt} = E[Y_{it}^2]$ as follows:

$$
\begin{aligned}
\phi_{i,tt}(\beta, \lambda_\gamma, \nu) =\ & E[Y_{it}^2] = E_{\gamma_i^*}\left[\mu_{it}^* + \mu_{it}^{*2}\right] \\
=\ & \frac{1}{W}\sum_{w=1}^{W}\left[\exp[x'_{it}\beta + \left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}]\right. \\
& \left. + \exp[2x'_{it}\beta + \left\{2R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}]\right], \qquad (27)
\end{aligned}
$$

where $R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)$ is defined in (26). Hence, the unconditional variance has the formula

$$
\sigma_{i,tt}(\beta, \lambda_\gamma, \nu) = \phi_{i,tt}(\beta, \lambda_\gamma, \nu) - \mu_{it}^2(\beta, \lambda_\gamma, \nu). \qquad (28)
$$

Note that this variance formula can be obtained from (21) as well. We remark that unlike for the Poisson-normal mixed model, the mean and variance under the Poisson-$t_\nu$ mixed model are functions of the regression effects $\beta$, and variance parameter $\lambda_\gamma$ and shape parameter $\nu$ of the random effect distribution.

## *2.2 Correlation Properties of the Poisson Mixed Model: Unconditional Covariances*

To compute the unconditional covariance between $y_{iu}$ and $y_{it}$ ($u < t$), we first observe from (1)–(2) that their covariance conditional on the random effects $\gamma_i^*$ is not zero. Specifically, by (2), the conditional covariance is given by

$$
\text{cov}[(Y_{iu}, Y_{it})|\gamma_i^*] = \rho^{t-u}\mu_{iu}^*, \qquad (29)
$$

implying that

$$E[Y_{iu}Y_{it}|\gamma_i^*] = \rho^{t-u}\mu_{iu}^* + \mu_{iu}^*\mu_{it}^*. \tag{30}$$

Consequently,

$$\begin{aligned}\delta_{i,ut}(\beta,\lambda_\gamma,\nu,\rho) &= E[Y_{iu}Y_{it}]\\ &= E_{\gamma_i^*}E[Y_{iu}Y_{it}|\gamma_i^*]\\ &= E_{\gamma_i^*}\left[\rho^{t-u}\mu_{iu}^* + \mu_{iu}^*\mu_{it}^*\right]\\ &= \rho^{t-u}\mu_{iu}(\beta,\lambda_\gamma,\nu) + E_{\gamma_i^*}\left[\mu_{iu}^*\mu_{it}^*\right],\end{aligned} \tag{31}$$

where the unconditional mean $\mu_{iu}(\beta,\lambda_\gamma,\nu)$ has the formula similar to that of (26). Next, by similar computation as in (27), one obtains

$$\begin{aligned}\delta_{i,ut}(\beta,\lambda_\gamma,\nu,\rho) &= \rho^{t-u}\mu_{iu}(\beta,\lambda_\gamma,\nu) + E_{\gamma_i^*}\left[\exp\{x_{iu}+x_{it}\}'\beta + 2\gamma_i^*\right]\\ &= \rho^{t-u}\mu_{iu}(\beta,\lambda_\gamma,\nu) + \exp\left[\{x_{iu}+x_{it}\}'\beta\right]\frac{1}{W}\sum_{w=1}^{W}\exp\left[2R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right].\end{aligned} \tag{32}$$

Hence, for $u < t$, the unconditional covariance between $y_{iu}$ and $y_{it}$ is given by

$$\sigma_{i,ut}(\beta,\lambda_\gamma,\nu,\rho) = \delta_{i,ut}(\beta,\lambda_\gamma,\nu,\rho) - \mu_{iu}(\beta,\lambda_\gamma,\nu)\mu_{it}(\beta,\lambda_\gamma,\nu), \tag{33}$$

where $\delta_{i,ut}(\beta,\lambda_\gamma,\nu,\rho)$ has the formula given by (32) and $\mu_{it}(\beta,\lambda_\gamma,\nu)$ is given by (26).

## 3 GQL Estimation for the Parameters of the Poisson Mixed Model

The estimation of the parameters of the model will be done in cycle of iterations. In Sect. 3.1, we discuss a generalized quasi-likelihood (GQL) (Sutradhar 2003, Sect. 3) estimation approach for the estimation of the main regression parameter $\beta$ under the assumption that other parameters $(\rho,\lambda_\gamma,\nu)$ are known or their consistent estimates are available. In subsequent sections, we discuss their consistent estimation.

### 3.1 GQL Estimation for the Regression Effects $\beta$

For $\beta$ estimation, we exploit the first order responses, namely $y_i = [y_{i1}, \ldots, y_{it}, \ldots, y_{iT}]'$. Suppose that $\mu_i = E[Y_i]$. This mean vector is given by $\mu_i(\beta, \lambda_\gamma, \nu) = [\mu_{i1}, \ldots, \mu_{it}, \ldots, \mu_{iT}]'$, where, by (26), $\mu_{it}(\beta, \lambda_\gamma, \nu)$ has the formula

$$\mu_{it}(\beta, \lambda_\gamma, \nu) = \exp(x'_{it}\beta)\frac{1}{W}\sum_{w=1}^{W}\exp\left\{R(\psi_{iw}, \xi^2_{iw}; \lambda_\gamma, \nu)\right\} : T \times 1.$$

Next by using the formulas for the variances $\sigma_{i,tt}(\beta, \lambda_\gamma, \nu)$ from (28), and the covariances $\sigma_{i,ut}(\beta, \lambda_\gamma, \nu, \rho)$ from (33), we construct the $T \times T$ covariance matrix as

$$\Sigma_i(\beta, \lambda_\gamma, \nu, \rho) = (\sigma_{i,ut}(\beta, \lambda_\gamma, \nu, \rho)) : T \times T, \text{ for } u = t; \text{and } u \neq t.$$

Note that under the present model $\sigma_{i,tt}(\cdot)$ does not follow from $\sigma_{i,ut}(\cdot)$ as a special case. More specifically, $\sigma_{i,ut}(\cdot)$'s are constructed for $u < t$. The GQL estimating equation for $\beta$ is then given by

$$\sum_{i=1}^{K}\frac{\partial \mu'_i(\beta, \lambda_\gamma, \nu)}{\partial \beta}\Sigma_i^{-1}(\beta, \lambda_\gamma, \nu, \rho)\left(y_i - \mu_i(\beta, \lambda_\gamma, \nu)\right) = 0, \tag{34}$$

(Sutradhar 2003, 2004) where $\frac{\partial \mu'_i(\beta, \lambda_\gamma, \nu)}{\partial \beta}$ may be computed by using the formula for $\frac{\partial \mu_{it}(\beta, \lambda_\gamma, \nu)}{\partial \beta}$ for all $t = 1, \ldots, T$. This derivative follows from (26), and is given by

$$\frac{\partial \mu_{it}(\beta, \lambda_\gamma, \nu)}{\partial \beta} = \mu_{it}(\beta, \lambda_\gamma, \nu)x_{it}.$$

Consequently, the GQL estimating equation in (34) reduces to

$$\sum_{i=1}^{K}X'_iA_i\Sigma_i^{-1}(\beta, \lambda_\gamma, \nu, \rho)\left(y_i - \mu_i(\beta, \lambda_\gamma, \nu)\right) = 0, \tag{35}$$

where $X'_i = (x_{i1}, \ldots, x_{it}, \ldots, x_{iT})$ is the $p \times T$ covariate matrix for the $i$th individual, and

$$A_i = \text{diag}[\mu_{i1}(\beta, \lambda_\gamma, \nu), \ldots, \mu_{it}(\beta, \lambda_\gamma, \nu), \ldots, \mu_{iT}(\beta, \lambda_\gamma, \nu)] : \; T \times T.$$

### 3.1.1 Asymptotic Properties of the GQL Estimator of $\beta$

For true $\beta$, define

$$\bar{f}_K(\beta) = \frac{1}{K}\sum_{i=1}^{K} f_i(\beta) = \frac{1}{K}\sum_{i=1}^{K} \frac{\partial \mu_i'(\beta, \lambda_\gamma, \nu)}{\partial \beta} \Sigma_i^{-1}(\beta, \lambda_\gamma, \nu, \rho)\left(y_i - \mu_i(\beta, \lambda_\gamma, \nu)\right), \tag{36}$$

where $y_1, \ldots, y_i, \ldots, y_K$ are independent to each other as they are collected from $K$ independent individuals, but they are not identically distributed because

$$Y_i \sim (\mu_i(\beta, \lambda_\gamma, \nu), \Sigma_i(\beta, \lambda_\gamma, \nu, \rho)), \tag{37}$$

where the mean vectors and covariance matrices vary for the individuals $i = 1, \ldots, K$. By (37), it follows from (36) that

$$\begin{aligned}
E[\bar{f}_K(\beta)] &= 0 \\
\text{cov}[\bar{f}_K(\beta)] &= \frac{1}{K^2}\sum_{i=1}^{K} \text{cov}[f_i(\beta)] \\
&= \frac{1}{K^2}\sum_{i=1}^{K} \frac{\partial \mu_i'(\beta, \lambda_\gamma, \nu)}{\partial \beta} \Sigma_i^{-1}(\beta, \lambda_\gamma, \nu, \rho) \frac{\partial \mu_i(\beta, \lambda_\gamma, \nu)}{\partial \beta'} \\
&= \frac{1}{K^2}\sum_{i=1}^{K} V_i(\beta, \lambda_\gamma, \nu, \rho) = \frac{1}{K^2} V_K^*(\beta, \lambda_\gamma, \nu, \rho).
\end{aligned} \tag{38}$$

Next if the multivariate version of Lindeberg's condition holds, that is,

$$\lim_{K\to\infty} {V^*}_K^{-1} \sum_{i=1}^{K} \sum_{(f_i' {V^*}_K^{-1} f_i) > \epsilon} f_i f_i' g(f_i) = 0 \tag{39}$$

for all $\epsilon > 0$, $g(\cdot)$ being the probability distribution of $f_i$, then Lindeberg-Feller central limit theorem (Amemiya 1985, Theorem 3.3.6; McDonald 2005, Theorem 2.2) imply that

$$Z_K = K[V_K^*]^{-\frac{1}{2}} \bar{f}_K(\beta) \to N_p(0, I_p). \tag{40}$$

Next because $\hat{\beta}_{GQL}$ is a solution of (34), one writes by (36) that

$$\sum_{i=1}^{K} f_i(\hat{\beta}_{GQL}) = 0, \tag{41}$$

which by first order Taylor's series expansion produces

$$\sum_{i=1}^{K} f_i(\beta) + (\hat{\beta}_{GQL} - \beta) \sum_{i=1}^{K} f_i'(\beta) = 0. \tag{42}$$

That is,

$$\begin{aligned}
\hat{\beta}_{GQL} - \beta &= -\left[\sum_{i=1}^{K} f_i'(\beta)\right]^{-1} \sum_{i=1}^{K} f_i(\beta) \\
&= -\left[-\sum_{i=1}^{K} \frac{\partial \mu_i'(\beta, \lambda_\gamma, \nu)}{\partial \beta} \Sigma_i^{-1}(\beta, \lambda_\gamma, \nu, \rho) \frac{\partial \mu_i(\beta, \lambda_\gamma, \nu)}{\partial \beta'}\right]^{-1} \sum_{i=1}^{K} f_i(\beta) \\
&= \left[V_K^*(\beta, \lambda_\gamma, \nu, \rho)\right]^{-1} K\bar{f}(\beta) \\
&= \left[V_K^*(\beta, \lambda_\gamma, \nu, \rho)\right]^{-\frac{1}{2}} Z_K \to N(0, {V^*}_K^{-1}(\beta, \lambda_\gamma, \nu, \rho)),
\end{aligned} \tag{43}$$

by (40). It then follows that

$$\lim_{K \to \infty} \hat{\beta}_{GQL} \to N(\beta, {V^*}_K^{-1}(\beta, \lambda_\gamma, \nu, \rho)). \tag{44}$$

Also it follows that

$$||[V_K^*(\beta, \lambda_\gamma, \nu, \rho)]^{\frac{1}{2}} [\hat{\beta}_{GQL} - \beta]|| = O_p(\sqrt{p}). \tag{45}$$

## 3.2 GQL Estimation for the Scale and Shape Parameters

Notice from Sect. 2.1 that all three basic moment properties, namely the mean function (26), variances in (28), and the covariances given by (33) contain the scale parameter $\lambda_\gamma$ and the shape parameter $\nu$. Thus, it is sensible to exploit all first and second order responses to estimate these parameters. Note that the second order responses consist of both squared ($ss$) and pair-wise products ($pp$) of all repeated observations. Consequently, we consider a vector $g_i$ consisting of all first order and second order responses. In notation, $g_i$ has the form

$$g_i = [y_i', \; y_{iss}', \; y_{ipp}']' : \frac{T(T+3)}{2} \times 1, \tag{46}$$

where $y_i = [y_{i1}, \ldots, y_{it}, \ldots, y_{iT}]' : T \times 1$, as in (34), and

$$\begin{aligned}
y_{iss} &= [y_{i1}^2, \ldots, y_{it}^2, \ldots, y_{iT}^2]' : T \times 1, \text{ and} \\
y_{ipp} &= [y_{i1}y_{i2}, \ldots, y_{iu}y_{it}, \ldots, y_{i,T-1}y_{iT}]' : \frac{T(T-1)}{2} \times 1.
\end{aligned}$$

Let

$$E[g_i] = [\mu_i'(\beta, \lambda_\gamma, \nu), \phi_i'(\beta, \lambda_\gamma, \nu), \delta_i'(\beta, \lambda_\gamma, \nu, \rho)]'$$
$$= \eta_i(\beta, \lambda_\gamma, \nu, \rho) \text{ (say)}, \tag{47}$$

where

$$\mu_i(\beta, \lambda_\gamma, \nu) = [\mu_{i1}(\beta, \lambda_\gamma, \nu), \ldots, \mu_{it}(\beta, \lambda_\gamma, \nu), \ldots, \mu_{iT}(\beta, \lambda_\gamma, \nu)]'$$
$$\phi_i(\beta, \lambda_\gamma, \nu) = [\phi_{i,11}(\beta, \lambda_\gamma, \nu), \ldots, \phi_{i,tt}(\beta, \lambda_\gamma, \nu), \ldots, \phi_{i,TT}(\beta, \lambda_\gamma, \nu)]'$$
$$\delta_i(\beta, \lambda_\gamma, \nu, \rho) = [\delta_{i,12}(\beta, \lambda_\gamma, \nu, \rho), \ldots, \delta_{i,ut}(\beta, \lambda_\gamma, \nu, \rho), \ldots, \delta_{i,T-1,T}(\beta, \lambda_\gamma, \nu, \rho)]',$$

with $\mu_{it}(\beta, \lambda_\gamma, \nu)$ and $\phi_{i,tt}(\beta, \lambda_\gamma, \nu)$ for all $t = 1, \ldots, T$, are given by (26) and (27), respectively, and $\delta_{i,ut}(\beta, \lambda_\gamma, \nu, \rho)$ for $u < t$, are defined as in (32). Further, let

$$\Omega_i(\beta, \lambda_\gamma, \nu, \rho) = \text{cov}[g_i]$$
$$= \begin{bmatrix} \Sigma_i & \Omega_{i,ss} & \Omega_{i,pp} \\ \Omega_{i,ss}' & \Sigma_{i,ss} & \Omega_{i,sp} \\ \Omega_{i,pp}' & \Omega_{i,sp}' & \Sigma_{i,pp} \end{bmatrix}, \tag{48}$$

where

$$\Sigma_i = \text{cov}[Y_i],\ \ \Sigma_{i,ss} = \text{cov}[Y_{iss}],\ \ \Sigma_{i,pp} = \text{cov}[Y_{ipp}]$$
$$\Omega_{i,ss} = \text{cov}[Y_i, Y_{iss}'],\ \ \Omega_{i,pp} = \text{cov}[Y_i, Y_{ipp}'],\ \ \Omega_{i,sp} = \text{cov}[Y_{iss}, Y_{ipp}'].$$

Further, let $\pi = [\lambda_\gamma, \nu]' : 2 \times 1$, be a vector of the scale and shape parameters of the random effects distribution. Similar to (34), for known $\beta$ and $\rho$, one may then estimate the $\pi$ vector by solving the GQL estimation equation given by

$$\sum_{i=1}^{K} \frac{\partial \eta_i'(\beta, \lambda_\gamma, \nu, \rho)}{\partial \pi} \Omega_i^{-1}(\beta, \lambda_\gamma, \nu, \rho) \left(g_i - \eta_i(\beta, \lambda_\gamma, \nu, \rho)\right) = 0, \tag{49}$$

where

$$\frac{\partial \eta_i'(\beta, \lambda_\gamma, \nu, \rho)}{\partial \pi} = \begin{pmatrix} \frac{\partial \eta_i'(\beta, \lambda_\gamma, \nu, \rho)}{\partial \lambda_\gamma} \\ \frac{\partial \eta_i'(\beta, \lambda_\gamma, \nu, \rho)}{\partial \nu} \end{pmatrix}. \tag{50}$$

Note that the derivatives in (50) may be computed by using the following general derivatives with respect to $\lambda_\gamma$ and $\nu$:

$$\frac{\partial \mu_{it}}{\partial \lambda_\gamma} = \exp(x'_{it}\beta)\frac{1}{W}\sum_{w=1}^{W}\frac{\partial R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)}{\partial \lambda_\gamma}\exp\left\{R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\},$$

$$\frac{\partial \mu_{it}}{\partial \nu} = \exp(x'_{it}\beta)\frac{1}{W}\sum_{w=1}^{W}\frac{\partial R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)}{\partial \nu}\exp\left\{R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\},$$

$$\frac{\partial \phi_{i,tt}}{\partial \lambda_\gamma} = \frac{1}{W}\sum_{w=1}^{W}\frac{\partial R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)}{\partial \lambda_\gamma}\left[\exp[x'_{it}\beta + \left\{R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\}]\right.$$
$$\left. + 2\frac{\partial R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)}{\partial \lambda_\gamma}\exp[2x'_{it}\beta + \left\{2R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\}]\right],$$

$$\frac{\partial \phi_{i,tt}}{\partial \nu} = \frac{1}{W}\sum_{w=1}^{W}\frac{\partial R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)}{\partial \nu}\left[\exp[x'_{it}\beta + \left\{R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\}]\right.$$
$$\left. + 2\frac{\partial R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)}{\partial \nu}\exp[2x'_{it}\beta + \left\{2R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\}]\right],$$

$$\frac{\partial \delta_{i,ut}}{\partial \lambda_\gamma} = \rho^{t-u}\frac{\partial \mu_{iu}(\beta,\lambda_\gamma,\nu)}{\partial \lambda_\gamma} + \exp\left[\{x_{iu}+x_{it}\}'\beta\right]\frac{1}{W}\sum_{w=1}^{W}2\frac{\partial R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)}{\partial \lambda_\gamma}$$
$$\times\exp\left[2R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right] - \left\{\frac{\partial \mu_{iu}(\beta,\lambda_\gamma,\nu)}{\partial \lambda_\gamma}\mu_{it} + \mu_{iu}\frac{\partial \mu_{it}(\beta,\lambda_\gamma,\nu)}{\partial \lambda_\gamma}\right\}$$

$$\frac{\partial \delta_{i,ut}}{\partial \nu} = \rho^{t-u}\frac{\partial \mu_{iu}(\beta,\lambda_\gamma,\nu)}{\partial \nu} + \exp\left[\{x_{iu}+x_{it}\}'\beta\right]\frac{1}{W}\sum_{w=1}^{W}2\frac{\partial R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)}{\partial \nu}$$
$$\times\exp\left[2R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right] - \left\{\frac{\partial \mu_{iu}(\beta,\lambda_\gamma,\nu)}{\partial \nu}\mu_{it} + \mu_{iu}\frac{\partial \mu_{it}(\beta,\lambda_\gamma,\nu)}{\partial \nu}\right\},$$

where

$$\frac{\partial R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)}{\partial \lambda_\gamma} = \left[\frac{\psi_{iw}}{\left\{\frac{1}{2\sqrt{(\nu)}}(\xi_{iw}^2-4)+1\right\}}\right]$$
$$\frac{\partial R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)}{\partial \nu} = \left[\frac{\lambda_\gamma\psi_{iw}\left\{\frac{1}{4\{\sqrt{(\nu)}\}^3}(\xi_{iw}^2-4)\right\}}{\left\{\frac{1}{2\sqrt{(\nu)}}(\xi_{iw}^2-4)+1\right\}^2}\right]. \tag{51}$$

The construction of the GQL estimating equation (49) still requires the computational formula for the weight matrix $\Omega_i(\beta,\lambda_\gamma,\nu,\rho)$. Now because this weight matrix requires the computation of second, third and fourth order moments for the repeated count data, unlike for the Gaussian data the computation of these moments are complicated. Some of the fourth order moments may not be computable without further joint distributional assumption for these repeated counts. Thus, for simplicity and because the consistent estimation of the parameters in $\pi$ does not require the use of exact weight matrix, in the next section, we provide an approximation for the computation of the elements of the weight matrix $\Omega_i(\beta, \lambda_\gamma, \nu, \rho)$ by pretending that the correlation index is zero, that is, $\rho = 0$ in (2). This assumption is equivalent to say that the repeated counts are assumed to conditionally (conditional on the random effects) independent (CI).

### 3.2.1 Computation of $\boldsymbol{\Omega_i(CI) \equiv \Omega_i^*(\beta, \lambda_\gamma, \nu)}$

Note that as outlined above, the $\Omega_i(\beta, \lambda_\gamma, \nu, \rho) = \text{cov}(g_i)$ matrix in (49) will be replaced by

$$\text{cov}(g_i|\rho = 0) = \Omega_i^*(\beta, \lambda_\gamma, \nu), \tag{52}$$

which contains moments up to order four under conditionally independence (CI) assumption. More specifically, we compute the $\Omega_i(\cdot)$ matrix in (48), but, under the assumption that $\rho = 0$, that is,

$$\begin{aligned} \Omega_i^*(\beta, \lambda_\gamma, \nu) &= \text{cov}[g_i|\rho = 0] \\ &= \begin{bmatrix} \Sigma_i^* & \Omega_{i,ss}^* & \Omega_{i,pp}^* \\ {\Omega^*}'_{i,ss} & \Sigma_{i,ss}^* & \Omega_{i,sp}^* \\ {\Omega^*}'_{i,pp} & {\Omega^*}'_{i,sp} & \Sigma_{i,pp}^* \end{bmatrix}, \end{aligned} \tag{53}$$

where

$$\begin{aligned} \Sigma_i^* &= \text{cov}[Y_i|\rho = 0],\ \Sigma_{i,ss}^* = \text{cov}[Y_{iss}|\rho = 0],\ \Sigma_{i,pp}^* = \text{cov}[Y_{ipp}|\rho = 0] \\ \Omega_{i,ss}^* &= \text{cov}[(Y_i, Y_{iss}')|\rho = 0],\ \Omega_{i,pp}^* = \text{cov}[(Y_i, Y_{ipp}')|\rho = 0], \\ \Omega_{i,sp}^* &= \text{cov}[(Y_{iss}, Y_{ipp}')|\rho = 0]. \end{aligned}$$

#### (a) Computation of the Second Order Moments Matrix $\Sigma_i^*$ :

Because the variances are not affected by the correlation index parameter, their formulas remain the same as in (28). However, the covariances under $\rho = 0$ will be different than (33). More specifically the formulas for the variances and covariances under the assumption $\rho = 0$ are given by

$$\begin{aligned} \text{var}[Y_{it}|\rho] = &\ \sigma_{i,tt}^*(\beta, \sigma_\gamma^2, \nu) \\ = &\ \sigma_{i,tt}(\beta, \lambda_\gamma, \nu) = \phi_{i,tt}(\beta, \lambda_\gamma, \nu) - \mu_{it}^2(\beta, \lambda_\gamma, \nu),\ \text{by (28)}; \end{aligned} \tag{54}$$

$$\begin{aligned}\text{cov}[(Y_{iu}, Y_{it})|\rho = 0] &= \sigma^*_{i,ut}(\beta, \sigma^2_\gamma, \nu) \\ &= \exp\left[\{x_{iu} + x_{it}\}'\beta\right] \frac{1}{W}\sum_{w=1}^{W} \exp\left[2R(\psi_{iw}, \xi^2_{iw}; \lambda_\gamma, \nu)\right] \\ &\quad -\mu_{iu}(\beta, \lambda_\gamma, \nu)\mu_{it}(\beta, \lambda_\gamma, \nu), \text{ by (32)–(33)}.\end{aligned} \tag{55}$$

**(b) Computation of the Third Order Moments Matrix $\Omega^*_{i,ss}$ :**

To compute this matrix $\Omega^*_{i,ss} = \text{cov}[\{Y_i, Y'_{iss}\}|\rho = 0]$, it is sufficient to compute the elements (i) $\text{cov}[\{Y_{it}, Y^2_{it}\}|\rho = 0]$, and (ii) $\text{cov}[\{Y_{iu}, Y^2_{it}\}|\rho = 0]$, for $u < t$.

(i) Formula for $\text{cov}[\{Y_{it}, Y^2_{it}\}|\rho = 0]$ :

Notice that this formula does not depend on $\rho$, and the conditioning on $\rho = 0$ is not needed. Thus

$$\begin{aligned}\text{cov}[\{Y_{it}, Y^2_{it}\}|\rho = 0] &= E[Y^3_{it}] - E[Y_{it}]E[Y^2_{it}] \\ &= E[Y^3_{it}] - \mu_{it}(\beta, \lambda_\gamma, \nu)\phi_{i,tt}(\beta, \lambda_\gamma, \nu), \text{ by (54)},\end{aligned} \tag{56}$$

where

$$\begin{aligned}E[Y^3_{it}] &= E_{\gamma^*_i}\left[\mu^*_{it} + 3\mu^{*2}_{it} + \mu^{*3}_{it}\right] \\ &= \frac{1}{W}\sum_{w=1}^{W}\left[\exp[x'_{it}\beta + \{R(\psi_{iw}, \xi^2_{iw}; \lambda_\gamma, \nu)\}]\right. \\ &\quad + 3\exp[2x'_{it}\beta + \{2R(\psi_{iw}, \xi^2_{iw}; \lambda_\gamma, \nu)\}] \\ &\quad \left.+ \exp[3x'_{it}\beta + \{3R(\psi_{iw}, \xi^2_{iw}; \lambda_\gamma, \nu)\}]\right],\end{aligned} \tag{57}$$

with $R(\psi_{iw}, \xi^2_{iw}; \lambda_\gamma, \nu)$ as defined in (26).

(ii) Formula for $\text{cov}[\{Y_{iu}, Y^2_{it}\}|\rho = 0]$, for $u < t$ :

Because $u$ and $t$ denote two different times points, the covariance between $y_{iu}$ and $y^2_{it}$ is a function of the correlation index parameter $\rho$. However, we now simplify this covariance formula as follows under the assumption that $\rho = 0$.

$$\begin{aligned}\text{cov}[\{Y_{iu}, Y^2_{it}\}|\rho = 0] &= E[Y_{iu}Y^2_{it}|\rho = 0] - E[Y_{iu}]E[Y^2_{it}] \\ &= E[Y_{iu}Y^2_{it}|\rho = 0] - \mu_{iu}(\beta, \lambda_\gamma, \nu)\phi_{i,tt}(\beta, \lambda_\gamma, \nu), \text{ by (54)},\end{aligned} \tag{58}$$

where, by (30) and (26)–(27), one writes

$$\begin{aligned}E[Y_{iu}Y^2_{it}|\rho = 0] &= E_{\gamma^*_i}[E\{Y_{iu}|\gamma^*_i\}E\{Y^2_{it}|\gamma^*_i\}] \\ &= E_{\gamma^*_i}[\mu^*_{iu}\{\mu^*_{it} + \mu^{*2}_{it}\}]\end{aligned} \tag{59}$$

$$= \frac{1}{W}\sum_{w=1}^{W}\Big[\exp[\{x_{iu}+x_{it}\}'\beta + 2\left\{R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\}] + \exp[\{x_{iu}+2x_{it}\}'\beta + 3\left\{R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\}]\Big]. \quad (60)$$

**(c) Computation of the Third Order Moments Matrix $\Omega^*_{i,pp}$ :**

To compute this matrix $\Omega^*_{i,pp} = \text{cov}[\{Y_i, Y'_{ipp}\}|\rho = 0]$, it is sufficient to compute the elements (i) $\text{cov}[\{Y_{iu}, Y_{iu}Y_{it}\}|\rho = 0]$ for $u < t$ or $u > t$, and (ii) $\text{cov}[\{Y_{iu}, Y_{it}Y_{im}\}|\rho = 0]$, for $u \neq t, u \neq m, t < m$.

(i) Formula for $\text{cov}[\{Y_{iu}, Y_{iu}Y_{it}\}|\rho = 0]$ :
By similar calculations as in (59)–(60), we write

$$\begin{aligned}\text{cov}[\{Y_{iu}, Y_{iu}Y_{it}\}|\rho = 0] &= E_{\gamma_i^*}[E\{Y_{iu}^2Y_{it}\}|\gamma_i^*] - E[Y_{iu}]E[\{Y_{iu}Y_{it}\}|\rho = 0] \\ &= E_{\gamma_i^*}[\{\mu^*_{iu} + \mu^{*\,2}_{iu}\}\mu^*_{it}] - \mu_{iu}(\beta,\lambda_\gamma,\nu)E_{\gamma_i^*}[\mu^*_{iu}\mu^*_{it}] \\ &= \frac{1}{W}\sum_{w=1}^{W}\Big[\exp[\{x_{iu}+x_{it}\}'\beta + 2\left\{R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\}] \\ &+ \exp[\{2x_{iu}+x_{it}\}'\beta + 3\left\{R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\}] \\ &- \mu_{iu}(\beta,\lambda_\gamma,\nu)\exp[\{x_{iu}+x_{it}\}'\beta + 2\left\{R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\}]\Big]. \quad (61)\end{aligned}$$

(ii) Formula for $\text{cov}[\{Y_{iu}, Y_{it}Y_{im}\}|\rho = 0]$ :
By similar calculations as in (i), we write the formula for this covariance as

$$\begin{aligned}&\text{cov}[\{Y_{iu}, Y_{it}Y_{im}\}|\rho = 0] \\ &= \frac{1}{W}\sum_{w=1}^{W}\Big[\exp[\{x_{iu}+x_{it}+x_{im}\}'\beta + 3\left\{R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\}] \\ &- \mu_{iu}(\beta,\lambda_\gamma,\nu)\exp[\{x_{it}+x_{im}\}'\beta + 2\left\{R(\psi_{iw},\xi_{iw}^2;\lambda_\gamma,\nu)\right\}]\Big]. \quad (62)\end{aligned}$$

**(d) Computation of the Fourth Order Moments Matrix $\Sigma^*_{i,ss}$ :**

To compute this fourth order matrix, one needs the formulas for two general elements, namely (i) $\text{var}[Y_{it}^2]$, and (ii) $\text{cov}[\{Y_{iu}^2, Y_{it}^2\}|\rho = 0]$. These formulas are developed as follows:

(i) Recall from (27) that $E[Y_{it}^2] = \phi_{i,tt}(\beta,\lambda_\gamma,\nu)$. Next because

$$E[Y_{it}^4|\gamma_i^*] = \left[\mu^*_{it} + 7\mu^{*2}_{it} + 6\mu^{*3}_{it} + \mu^{*4}_{it}\right],$$

it then follows that

$$\begin{aligned}
\mathrm{var}(Y_{it})^2 &= E[Y_{it}^4] - [E\{Y_{it}^2\}]^2 \\
&= E_{\gamma_i^*} E[Y_{it}^4|\gamma_i^*] - [\phi_{i,tt}(\beta, \lambda_\gamma, \nu)]^2 \\
&= \frac{1}{W}\sum_{w=1}^{W}\left[\exp\left[x_{it}'\beta + \left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}\right] + 7\exp\left[2x_{it}'\beta + \left\{2R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}\right]\right. \\
&+ 6\exp[3x_{it}'\beta + \left\{3R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}] + \exp[4x_{it}'\beta + \left\{4R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}]\Big], \qquad (63)
\end{aligned}$$

with $R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)$ as defined in (26).

(ii) Formula for $\mathrm{cov}[\{Y_{iu}^2, Y_{it}^2\}|\rho = 0]$, for $u < t$ :

By similar calculations, this covariance has the computing formula given by

$$\mathrm{cov}[\{Y_{iu}^2, Y_{it}^2\}|\rho = 0] = E[\{Y_{iu}^2 Y_{it}^2\}|\rho = 0] - \phi_{i,uu}(\beta, \lambda_\gamma, \nu)\phi_{i,tt}(\beta, \lambda_\gamma, \nu), \qquad (64)$$

where

$$\begin{aligned}
E[\{Y_{iu}^2 Y_{it}^2\}|\rho = 0] &= E_{\gamma_i^*}\left[\{\mu_{iu}^* + \mu_{iu}^{*\,2}\}\{\mu_{it}^* + \mu_{it}^{*2}\}\right] \\
&= \frac{1}{W}\sum_{w=1}^{W}\Big[\exp[\{x_{iu} + x_{it}\}'\beta + 2\left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}] \\
&+ \exp\left[\{x_{iu} + 2x_{it}\}'\beta + \left\{3R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}\right] \\
&+ \exp\left[\{2x_{iu} + x_{it}\}'\beta + \left\{3R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}\right] \\
&+ \exp[2\{x_{iu} + x_{it}\}'\beta + \left\{4R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}\Big]. \qquad (65)
\end{aligned}$$

**(e) Computation of the Fourth Order Moments Matrix $\Omega_{i,sp}^*$:**

To compute this fourth order matrix, one needs the formulas for two general covariance elements, namely (i) $\mathrm{cov}[\{Y_{iu}^2, Y_{iu}Y_{it}\}|\rho = 0]$, and (ii) $\mathrm{cov}[\{Y_{iu}^2, Y_{it}Y_{im}\}|\rho = 0]$. These formulas are developed as follows:

(i) Formula for $\mathrm{cov}[\{Y_{iu}^2, Y_{iu}Y_{it}\}|\rho = 0]$ :

$$\begin{aligned}
\mathrm{cov}[\{Y_{iu}^2, Y_{iu}Y_{it}\}|\rho = 0] &= E[\{Y_{iu}^3 Y_{it}\}|\rho = 0] \\
&- \phi_{i,uu}(\beta, \lambda_\gamma, \nu)\frac{1}{W}\sum_{w=1}^{W}\Big[\exp[\{x_{iu} + x_{it}\}'\beta \\
&+ 2\left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}]\Big], \text{ by (27) and (32)}, \qquad (66)
\end{aligned}$$

where

$$\begin{aligned}
E[\{Y_{iu}^3 Y_{it}\}|\rho = \ & 0] = E_{\gamma_i^*}\left[E[Y_{iu}^3|\gamma_i^*]E[Y_{it}|\gamma_i^*]\right] \\
= \ & E_{\gamma_i^*}\left[\{\mu_{iu}^* + 3\mu_{iu}^{*\,2} + \mu_{iu}^{*\,3}\}\{\mu_{it}^*\}\right] \\
= \ & \frac{1}{W}\sum_{w=1}^{W}\left[\exp[\{x_{iu} + x_{it}\}'\beta + 2\left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}]\right. \\
& + 3\exp[\{2x_{iu} + x_{it}\}'\beta + 3\left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}] \\
& \left. + \ \exp[\{3x_{iu} + x_{it}\}'\beta + 4\left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}]\right]. \qquad (67)
\end{aligned}$$

(ii) $\text{cov}[\{Y_{iu}^2, Y_{it}Y_{im}\}|\rho = 0]$ :

$$\begin{aligned}
\text{cov}[\{Y_{iu}^2, Y_{it}Y_{im}\}|\rho = 0] = E_{\gamma_i^*}\left[\{\mu_{iu}^* + \mu_{iu}^{*\,2}\}\{\mu_{it}^*\mu_{im}^*\}\right] \\
-\phi_{i,uu}(\beta, \lambda_\gamma, \nu)\frac{1}{W}\sum_{w=1}^{W}\left[\exp[\{x_{it} + x_{im}\}'\beta + 2\left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}]\right], \qquad (68)
\end{aligned}$$

where the first term in the right hand side of (68) has the formula

$$\begin{aligned}
& E_{\gamma_i^*}\left[\{\mu_{iu}^* + \mu_{iu}^{*\,2}\}\{\mu_{it}^*\mu_{im}^*\}\right] \\
& = \frac{1}{W}\sum_{w=1}^{W}\left[\exp[\{x_{iu} + x_{it} + x_{im}\}'\beta + 3\left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}]\right. \\
& \quad \left. + \exp[\{2x_{iu} + x_{it} + x_{im}\}'\beta + 4\left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}]\right]. \qquad (69)
\end{aligned}$$

**(f) Computation of the Fourth Order Moments Matrix $\Omega_{i,pp}^*$ :**

The computation for this matrix requires the formulas for (i) $\text{cov}[\{Y_{iu}Y_{it}, Y_{iu}Y_{it}\}|\rho = 0]$, (ii) $\text{cov}[\{Y_{iu}Y_{it}, Y_{iu}Y_{im}\}|\rho = 0]$, (iii) $\text{cov}[\{Y_{iu}Y_{it}, Y_{iv}Y_{it}\}|\rho = 0]$, and (iv) $\text{cov}[\{Y_{iu}Y_{it}, Y_{iv}Y_{im}\}|\rho = 0]$. The computations for all these four covariances are similar. We, for example, give the formulas for covariances in (i) and (iv).

(i) Formula for $\text{cov}[\{Y_{iu}Y_{it}, Y_{iu}Y_{it}\}|\rho = 0]$ :

$$\begin{aligned}
\text{cov}[\{Y_{iu}Y_{it}, Y_{iu}Y_{it}\}|\rho = 0] = E[\{Y_{iu}^2 Y_{it}^2\}|\rho = 0] \\
-\left[\frac{1}{W}\sum_{w=1}^{W}\left(\exp[\{x_{it} + x_{im}\}'\beta + 2\left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}]\right)\right]^2, \qquad (70)
\end{aligned}$$

where $E[\{Y_{iu}^2 Y_{it}^2\}|\rho = 0]$ is computed in (65).

(iv) Formula for $\text{cov}[\{Y_{iu}Y_{it}, Y_{iv}Y_{im}\}|\rho = 0]$ :

$$\begin{aligned}
&\text{cov}[\{Y_{iu}Y_{it}, Y_{iv}Y_{im}\}|\rho = 0] \\
&= \frac{1}{W}\sum_{w=1}^{W}\left[\exp[\{x_{iu} + x_{it} + x_{iv} + x_{im}\}'\beta + 4\left\{R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right\}]\right] \\
&\quad -\mu_{iu}(\beta, \lambda_\gamma, \nu)\mu_{it}(\beta, \lambda_\gamma, \nu)\mu_{iv}(\beta, \lambda_\gamma, \nu)\mu_{im}(\beta, \lambda_\gamma, \nu). \qquad (71)
\end{aligned}$$

### 3.2.2 Asymptotic Properties of the GQL Estimator $\hat{\pi}_{GQL} = [\hat{\lambda}_{\gamma,GQL}, \hat{\nu}_{GQL}]' : 2 \times 1$

Notice that when $\Omega_i^*(\beta, \lambda_\gamma, \nu)$ from Sect. 3.2.1 is used in (49) for $\Omega_i(\beta, \lambda_\gamma, \nu, \rho)$, one solves the approximate GQL estimating equation

$$\sum_{i=1}^{K}\frac{\partial\eta_i'(\beta, \lambda_\gamma, \nu, \rho)}{\partial\pi}\Omega_i^{*-1}(\beta, \lambda_\gamma, \nu)\left(g_i - \eta_i(\beta, \lambda_\gamma, \nu, \rho)\right) = 0, \qquad (72)$$

for $\pi = (\lambda_\gamma, \nu)'$.

Let $\hat{\pi}_{GQL} = [\hat{\lambda}_{\gamma,GQL}, \hat{\nu}_{GQL}]' : 2\times 1$ be the solution of (72). By similar calculations as in Sect. 3.1.1 (see (44)), it can be shown that

$$\lim_{K\to\infty}\hat{\pi}_{GQL} \to N(\pi, {Q^*}_K^{-1}(\beta, \lambda_\gamma, \nu, \rho)), \qquad (73)$$

where

$$\begin{aligned}
Q_K^* = &\left[\sum_{i=1}^{K}\frac{\partial\eta_i'(\beta, \lambda_\gamma, \nu, \rho)}{\partial\pi}\Omega_i^{*-1}(\beta, \lambda_\gamma, \nu)\frac{\partial\eta_i(\beta, \lambda_\gamma, \nu, \rho)}{\partial\pi'}\right]^{-1} \\
&\times\sum_{i=1}^{K}\frac{\partial\eta_i'(\beta, \lambda_\gamma, \nu, \rho)}{\partial\pi}\Omega_i^{*-1}(\beta, \lambda_\gamma, \nu)\Omega_i(\beta, \lambda_\gamma, \nu, \rho)\Omega_i^{*-1}(\beta, \lambda_\gamma, \nu)\frac{\partial\eta_i(\beta, \lambda_\gamma, \nu, \rho)}{\partial\pi'} \\
&\times\left[\sum_{i=1}^{K}\frac{\partial\eta_i'(\beta, \lambda_\gamma, \nu, \rho)}{\partial\pi}\Omega_i^{*-1}(\beta, \lambda_\gamma, \nu)\frac{\partial\eta_i(\beta, \lambda_\gamma, \nu, \rho)}{\partial\pi'}\right]^{-1}. \qquad (74)
\end{aligned}$$

## 3.3 *Moment Estimation of Correlation Index Parameter $\rho$*

Recall from (33) that

$$\begin{aligned}
&E[(Y_{iu} - \mu_{iu}(\cdot))(Y_{it} - \mu_{it}(\cdot))] = \rho^{t-u}\mu_{iu}(\beta, \lambda_\gamma, \nu) \\
&\quad + \exp\left[\{x_{iu} + x_{it}\}'\beta\right]\frac{1}{W}\sum_{w=1}^{W}\exp\left[2R(\psi_{iw}, \xi_{iw}^2; \lambda_\gamma, \nu)\right] - \mu_{iu}(\beta, \lambda_\gamma, \nu)\mu_{it}(\beta, \lambda_\gamma, \nu).
\end{aligned}$$

Consequently, by using lag 1 based pair-wise product responses, one obtains

$$E\sum_{i=1}^{K}\sum_{t=1}^{T-1}\frac{\{(Y_{it}-\mu_{it}(\cdot))(Y_{i,t+1}-\mu_{i,t+1}(\cdot))\}}{K(T-1)} = \rho\sum_{i=1}^{K}\sum_{t=1}^{T-1}\frac{\mu_{it}(\cdot)}{K(T-1)}$$
$$+\frac{1}{KW}\sum_{i=1}^{K}\sum_{w=1}^{W}\{\exp\left[2R(\psi_{iw},\xi_{iw}^{2};\lambda_{\gamma},\nu)\right]\}\frac{\{\sum_{t=1}^{T-1}\exp\left[\{x_{iu}+x_{it}\}'\beta\right]\}}{T-1}$$
$$-\sum_{i=1}^{K}\sum_{t=1}^{T-1}\frac{\mu_{it}(\cdot)\mu_{i,t+1}(\cdot)}{K(T-1)}. \quad (75)$$

Further, one writes

$$E\sum_{i=1}^{K}\sum_{t=1}^{T}\frac{\{(Y_{it}-\mu_{it}(\cdot))^{2}\}}{KT} = \sum_{i=1}^{K}\sum_{t=1}^{T}\frac{\sigma_{i,tt}(\cdot)}{KT}, \quad (76)$$

where the variance $\sigma_{i,tt}(\cdot)$ has the formula given by (28).

Now by dividing (75) by (76), and using first order approximation, one obtains the unbiased moment estimator $\hat{\rho}_M$ for $\rho$ as

$$\hat{\rho}_M \simeq \left[\frac{\sum_{i=1}^{K}\sum_{t=1}^{T-1}\{(Y_{it}-\mu_{it}(\cdot))(Y_{i,t+1}-\mu_{i,t+1}(\cdot))\}/\{K(T-1)\}}{\sum_{i=1}^{K}\sum_{t=1}^{T}\{(Y_{it}-\mu_{it}(\cdot))^{2}\}/\{KT\}}\right]$$
$$\div\left[\frac{\sum_{i=1}^{K}\sum_{t=1}^{T-1}\mu_{it}(\cdot)/\{K(T-1)\}}{\sum_{i=1}^{K}\sum_{t=1}^{T}\sigma_{i,tt}(\cdot)/\{KT\}}\right]$$
$$-\left[\frac{1}{KW}\sum_{i=1}^{K}\sum_{w=1}^{W}\{\exp\left[2R(\psi_{iw},\xi_{iw}^{2};\lambda_{\gamma},\nu)\right]\}\frac{\{\sum_{t=1}^{T-1}\exp\left[\{x_{iu}+x_{it}\}'\beta\right]\}}{T-1}\right]$$
$$\div\left[\sum_{i=1}^{K}\sum_{t=1}^{T-1}\frac{\mu_{it}(\cdot)}{K(T-1)}\right]+\left[\sum_{i=1}^{K}\sum_{t=1}^{T-1}\frac{\mu_{it}(\cdot)\mu_{i,t+1}(\cdot)}{K(T-1)}\right]\div\left[\sum_{i=1}^{K}\sum_{t=1}^{T-1}\frac{\mu_{it}(\cdot)}{K(T-1)}\right]. \quad (77)$$

Under some regularity conditions on the covariates so that var$[\hat{\rho}_M]$ is bounded by a finite quantity, it follows that the moment estimator $\hat{\rho}_M$ is consistent for $\rho$. This is mainly because $\hat{\rho}_M$ given by (77) is approximately unbiased for $\rho$.

# 4 Binary Dynamic Mixed Logit Model with $t_\nu$ Random Effects

Recall the binary dynamic mixed logit (BDML) model given in (9), that is,

$$Pr(y_{it} = 1|\gamma_i, y_{i,t-1}) = \begin{cases} \frac{\exp(x_{i1}^{'}\beta+\sigma_\gamma\gamma_i)}{1+\exp(x_{i1}^{'}\beta+\sigma_\gamma\gamma i)} & \text{for } t = 1 \\ \frac{\exp(x_{it}^{'}\beta+\theta y_{i,t-1}+\sigma_\gamma\gamma_i)}{1+\exp(x_{it}^{'}\beta+\theta y_{i,t-1}+\sigma_\gamma\gamma_i)} & \text{for } t = 2, \ldots, T, \end{cases}$$

Under the normality assumption for the random effects, i.e., when $\gamma_i \overset{iid}{\sim} N(0, 1)$, the basic properties such as unconditional mean, variance and correlations under such BDML model is given by (13)–(17). In the following subsection, we provide these properties for the BDML model under the assumption that the random effects now follow a $t$-distribution with $\nu$ degrees of freedom.

## *4.1 Basic Properties of the Binary Mixed Model: Unconditional Mean and Variance*

By similar calculations as in the normal case (13), one obtains an approximate unconditional mean based on $t_\nu$ random effects, as

$$\begin{aligned} E[Y_{it}] &= \mu_{it}(\beta,\ \theta,\ \lambda_\gamma,\ \nu) = W^{-1}\sum_{w}^{W}\pi_{it}^{*}(\psi_{iw}, \xi_{iw}^2) \\ &= W^{-1}\sum_{w=1}^{W}\left[p_{it0}(\psi_{iw},\ \xi_{iw}^2) + \pi_{i,t-1}^{*}(\psi_{iw}, \xi_{iw}^2)\left\{p_{it1}(\psi_{iw}, \xi_{iw}^2) - p_{it0}(\psi_{iw}, \xi_{iw}^2)\right\}\right], \quad (78) \end{aligned}$$

where

$$\pi_{i1}^{*}(\psi_{iw}, \xi_{iw}^2) = p_{i10}(\psi_{iw}, \xi_{iw}^2) = \frac{\exp(x_{i1}^{'}\beta + \gamma_{iw}^{*})}{1 + \exp(x_{i1}^{'}\beta + \gamma_{iw}^{*})}, \text{ and}$$

$$p_{ity_{i,t-1}}(\psi_{iw}, \xi_{iw}^2) = \frac{\exp(x_{it}^{'}\beta + \theta y_{i,t-1} + \gamma_{iw}^{*})}{1 + \exp(x_{it}^{'}\beta + \theta y_{i,t-1} + \gamma_{iw}^{*})},$$

with $\gamma_{iw}^{*}$ as the $t_\nu(0, \lambda_\gamma^2, \nu)$ random effect given by (25). Next because $y_{it}$ is a binary observation, it follows that

$$\text{var}[Y_{it}] = \sigma_{i,tt}(\beta,\ \theta,\ \lambda_\gamma,\ \nu) = \mu_{it}(\beta,\ \theta,\ \lambda_\gamma,\ \nu)[1 - \mu_{it}(\beta,\ \theta,\ \lambda_\gamma,\ \nu)], \quad (79)$$

where the unconditional mean $\mu_{it}(\beta,\ \theta,\ \lambda_\gamma,\ \nu)$ has the recursive type formula as in (78).

### 4.2 Computation of Unconditional Covariances for BDML Model with $t_\nu$ Random Effects

To compute the covariance between $y_{iu}$ and $y_{it}$ $(u < t)$, we note that under the present dynamic model (9), conditional on the random effects $\gamma_i^*$ defined by (25), $y_{iu}$ and $y_{it}$ are not independent. This is because conditional on $\gamma_i^*$, $y_{it}$ and $y_{i,t-1}$, for example, satisfy the dynamic dependence relationship (9). Next because

$$\begin{aligned} E[\{Y_{iu}Y_{it}\}|\gamma_i^*] &= \text{cov}[\{Y_{iu}, Y_{it}\}|\gamma_i^*] + E[Y_{iu}|\gamma_i^*]E[Y_{iT}|\gamma_i^*] \\ &= \sigma_{i,ut}^*(\psi_i, \xi_i^2) + \pi_{iu}^*(\psi_i, \xi_i^2)\pi_{it}^*(\psi_i, \xi_i^2), \end{aligned}$$

with

$$\sigma_{i,ut}^*(\psi_i, \xi_i^2) = \pi_{iu}^*(\psi_i, \xi_i^2)[1 - \pi_{iu}^*(\psi_i, \xi_i^2)]\Pi_{j=u+1}^{t}\left[p_{ij1}(\psi_i, \xi_i^2) - p_{ij0}(\psi_i, \xi_i^2)\right], \tag{80}$$

(Sutradhar and Farrell 2007), one may compute the covariance between $y_{iu}$ and $y_{it}$, first by computing $E[Y_{iu}Y_{it}]$ using

$$E[Y_{iu}Y_{it}] = W^{-1}\sum_{w=1}^{W}\left[\sigma_{i,ut}^*(\psi_{iw}, \xi_{iw}^2) + \pi_{iu}^*(\psi_{iw}, \xi_{iw}^2)\pi_{it}^*(\psi_{iw}, \xi_{iw}^2)\right] = \tau_{i,ut}, \text{ (say)}, \tag{81}$$

where $\psi_{iw}$ and $\xi_{iw}^2$ are generated from $N(0, 1)$ and $\chi_4^2$, respectively, in order to compute $\gamma_{iw}^*$ by (25), and $\pi_{it}^*(\psi_{iw}, \xi_{iw}^2)$ is computed by (78).

## 5 GQL Estimation for the Parameters of the BDML Model with $t_\nu$ Random Effects

It is clear from Sects. 4.1 and 4.2 that the basic properties of the BDML (binary dynamic mixed logit) model (9), that is, the first and second order moments of the repeated binary responses contain all four parameters, namely $\beta$, $\theta$, $\lambda_\gamma$, and $\nu$, of the model. Consequently, we exploit all first and second order observations, and minimize their generalized distance from their corresponding means to construct a GQL estimating equations [Sutradhar (2010, Sect. 5.4); see also Sutradhar (2011, Sect. 9.2)] for these desired parameters. Note that for the binary data, $y_{it}^2 \equiv y_{it}$. One may, thus, consider a vector of first and second order responses given by

$$v_i = (y_{i1}, \ldots, y_{it}, \ldots, y_{iT}, y_{i1}y_{i2}, \ldots, y_{iu}y_{it}, \ldots, y_{i(T-1)}y_{iT})',$$

for the purpose of constructing the desired estimating equation. Now, denote the $E[V_i]$ by

$$\zeta_i = E[V_i] = \eta_i(\beta, \theta, \lambda_\gamma, \nu) = [\mu_{i1}, \ldots, \mu_{it}, \ldots, \mu_{iT}, \tau_{i,12}, \ldots, \tau_{i,ut}, \ldots, \tau_{i,(T-1)T}]' \\ = [\mu_i', \tau_i']', \tag{82}$$

where the formula for the unconditional mean $\mu_{it}$ for all $t = 1, \ldots, T$, is given by (78), and $\tau_{i,ut} = E[Y_{iu}Y_{it}]$ for all $u < t$, may be computed by (81). Further let $\alpha = (\beta, \theta, \lambda_\gamma, \nu)'$, and $\Omega_i$ denote the $T(T+2)/2 \times T(T+1)/2$ covariance of $v_i$. Following Sutradhar (2010) [see also Sutradhar (2004)], one may then write the GQL estimating equation for $\alpha$ as

$$\sum_{i=1}^{K} \frac{\partial \zeta_i'}{\partial \alpha} \Omega_i^{-1}(v_i - \zeta_i) = 0, \tag{83}$$

which may be solved by using the iterative equation

$$\hat{\alpha}_{GQL}(r+1) = \hat{\alpha}_{GQL}(r) + \left[ \left\{ \sum_{i=1}^{K} \frac{\partial \zeta_i'}{\partial \alpha} \Omega_i^{-1} \frac{\partial \zeta_i}{\partial \alpha'} \right\}^{-1} \sum_{i=1}^{K} \frac{\partial \zeta_i'}{\partial \alpha} \Omega_i^{-1}(v_i - \zeta_i) \right]_{|\alpha = \hat{\alpha}_{GQL}(r)} . \tag{84}$$

Note that to compute the $\Omega_i$ matrix for (83) and (84) , one needs to compute the following elements: **(a)** var$[Y_{it}]$; **(b)** cov$[Y_{iu}, Y_{it}]$; **(c)** var$[Y_{iu}Y_{it}]$; **(d)** cov$[Y_{iu}Y_{it}, Y_{i\ell}Y_{im}]$; and **(e)** cov$[Y_{iu}, Y_{im}Y_{it}]$. However, all these elements through **(a)**–**(e)**, may be computed by using the moments up to order four given in Sects. 4.1, 4.2, and 5.1 below. For example,

$$\text{cov}[Y_{iu}Y_{it}, Y_{i\ell}Y_{im}] = E[Y_{iu}Y_{it}Y_{i\ell}Y_{im}] - E[Y_{iu}Y_{it}]E[Y_{i\ell}Y_{im}] \\ = \tilde{\phi}_{i,ut\ell m} - \tau_{i,ut}\tau_{i,\ell m}, \tag{85}$$

where the formula for $\tilde{\phi}_{i,ut\ell m}$ is given in (96), and the formula for $\tau_{i,ut}$, for example, is given in (81).

**Computation of $\frac{\partial \zeta_i'}{\partial \alpha}$** for (84):

Because $\zeta_i = [\mu_i', \tau_i']'$, the gradients for searching for the estimate of $\alpha = (\beta', \theta, \lambda_\gamma, \nu)'$ can be computed by by using the formulas for $\frac{\partial \mu_{it}}{\partial \alpha}$ and $\frac{\partial \tau_{i,ut}}{\partial \alpha}$, where $\mu_{it}$ and $\tau_{i,ut}$ are given by (78) and (81), respectively. For the purpose, we derive these formulas by using

$$\frac{\partial \mu_{it}}{\partial \alpha} = W^{-1} \sum_{w=1}^{W} \left[ \frac{\partial p_{it0}(\psi_{iw}, \xi_{iw}^2)}{\partial \alpha} + \{ \frac{\partial \pi_{i,t-1}^*(\psi_{iw}, \xi_{iw}^2)}{\partial \alpha} \} \{ p_{it1}(\psi_{iw}, \xi_{iw}^2) - p_{it0}(\psi_{iw}, \xi_{iw}^2) \} \right. \\ \left. + \; \pi_{i,t-1}^*(\psi_{iw}, \xi_{iw}^2) \left\{ \frac{\partial p_{it1}(\psi_{iw}, \xi_{iw}^2)}{\partial \alpha} - \frac{\partial p_{it0}(\psi_{iw}, \xi_{iw}^2)}{\partial \alpha} \right\} \right], \tag{86}$$

$$\frac{\partial \tau_{i,ut}}{\partial \alpha} = W^{-1} \sum_{w=1}^{W} \frac{\partial \left[ \sigma_{i,ut}^*(\psi_{iw}, \xi_{iw}^2) + \pi_{iu}^*(\psi_{iw}, \xi_{iw}^2)\pi_{it}^*(\psi_{iw}, \xi_{iw}^2) \right]}{\partial \alpha}, \tag{87}$$

where

$$\sigma^{*}_{i,ut}(\psi_{iw},\xi^2_{iw}) = \pi^{*}_{iu}(\psi_{iw},\xi^2_{iw})[1-\pi^{*}_{iu}(\psi_{iw},\xi^2_{iw})]\Pi^{t}_{j=u+1}\left[p_{ij1}(\psi_{iw},\xi^2_{iw}) - p_{ij0}(\psi_{iw},\xi^2_{iw})\right].$$

Note that to compute the derivative of the product factor involved in $\sigma^{*}_{i,ut}(\cdot)$, one can use the formula

$$\begin{aligned} &\frac{\partial \Pi^{t}_{j=u+1}\left[p_{ij1}(\psi_{iw},\xi^2_{iw}) - p_{ij0}(\psi_{iw},\xi^2_{iw})\right]}{\partial \alpha} \\ &= \Pi^{t}_{j=u+1}\left[p_{ij1}(\psi_{iw},\xi^2_{iw}) - p_{ij0}(\psi_{iw},\xi^2_{iw})\right] \\ &\quad \times \sum_{j=u+1}^{t} \frac{\partial log\ \left[p_{ij1}(\psi_{iw},\xi^2_{iw}) - p_{ij0}(\psi_{iw},\xi^2_{iw})\right]\}}{\partial \alpha}. \end{aligned} \tag{88}$$

To complete the formulation of the above derivatives, we now give the derivatives for one term, namely $p_{ij1}(\psi_{iw},\xi^2_{iw})$, with respect to each element of $\alpha = (\beta', \theta, \lambda_\gamma, \nu)'$. To be specific,

$$\frac{\partial p_{ij1}(\psi_{iw},\xi^2_{iw})}{\partial \beta} = p_{ij1}(\psi_{iw},\xi^2_{iw})(1-p_{ij1}(\psi_{iw},\xi^2_{iw}))x_{ij}; \tag{89}$$

$$\frac{\partial p_{ij1}(\psi_{iw},\xi^2_{iw})}{\partial \theta} = p_{ij1}(\psi_{iw},\xi^2_{iw})(1-p_{ij1}(\psi_{iw},\xi^2_{iw})); \tag{90}$$

$$\begin{aligned} \frac{\partial p_{ij1}(\psi_{iw},\xi^2_{iw})}{\partial \lambda_\gamma} &= p_{ij1}(\psi_{iw},\xi^2_{iw})(1-p_{ij1}(\psi_{iw},\xi^2_{iw})) \\ &\quad \times \left[\psi_{iw}\{2\nu\}^{\frac{1}{2}}\left[\sqrt{\nu}\left(\xi^2_{iw}-4\right)+2\nu\right]^{-\frac{1}{2}}\right];\ \text{and} \end{aligned} \tag{91}$$

$$\begin{aligned} \frac{\partial p_{ij1}(\psi_{iw},\xi^2_{iw})}{\partial \nu} &= p_{ij1}(\psi_{iw},\xi^2_{iw})(1-p_{ij1}(\psi_{iw},\xi^2_{iw}))\sqrt{2}\lambda_\psi\psi_{iw} \\ &\quad \times \left[\frac{1}{2}\{\nu\}^{-\frac{1}{2}}\left[\sqrt{\nu}\left(\xi^2_{iw}-4\right)+2\nu\right]^{-\frac{1}{2}}\right. \\ &\quad \left. - \frac{1}{2}\{\nu\}^{\frac{1}{2}}\left[\sqrt{\nu}\left(\xi^2_{iw}-4\right)+2\nu\right]^{-\frac{3}{2}}[\frac{1}{2}\{\nu\}^{-\frac{1}{2}}(\xi^2_{iw}-4)+2]\right]. \end{aligned} \tag{92}$$

### 5.1 Computation Higher Order Moments to Construct $\Omega_i$ in (84)

Note that when the first and second order responses are used to construct distance functions for the estimation of the parameters $\beta,\ \theta,\ \lambda_\gamma,$ and $\nu$, one requires the third and fourth order moments which are used in the weight matrix to develop the estimating equations. The first (mean) and the second order moments are computed in (78), (79) and (81) using suitable close form expressions. For higher order such as the third and fourth order moments, it is convenient to compute them numerically (Sutradhar et al. 2008). To be specific, for the computation of the third order moments, let $\sum_{(y_{iu},y_{i\ell},y_{it})\ni s}$ indicates the summation over all binary variables in the sample space s that contain $T-3$ elements out of T elements except $y_{iu}, y_{il}, y_{it}$. one may then compute the third order moments as

$$E[Y_{iu}Y_{i\ell}Y_{it}] = P[y_{iu}=1, y_{i\ell}=1, y_{it}=1] \equiv \tilde{\delta}_{i,u\ell t} \quad (93)$$
$$= W^{-1}\sum_{w=1}^{W}\sum_{y_{iu},y_{il},y_{it}\ni s}\left[f(y_{i1}|\gamma^*_{iw})\Pi_{j=2}^{T} f(y_{ij}|y_{i,j-1},\gamma^*_{iw})\right]_{y_{iu}=1,y_{i\ell}=1,y_{it}=1}$$

where by (78)

$$f(y_{i1}|\gamma^*_{iw}) = [p_{i10}(\gamma^*_{iw})]^{y_{i1}}[1-p_{i10}(\gamma^*_{iw})]^{1-y_{i1}}$$
$$f(y_{ij}|y_{i,j-1},\gamma^*_{iw}) = [p_{ijy_{i,j-1}}(\gamma^*_{iw})]^{y_{ij}}[1-p_{ijy_{i,j-1}}(\gamma^*_{iw})]^{1-y_{ij}}. \quad (94)$$

After an algebra, one may simplify the third order moments in (93) as

$$\tilde{\delta}_{i,u\ell t} = W^{-1}\sum_{w=1}^{W}\sum_{y_{iu},y_{i\ell},y_{it}\ni s}\left[\tilde{p}_{i10}(y_{i1},\gamma^*_{iw})\Pi_{j=2}^{T}\tilde{p}_{ij1}(y_{ij},y_{i,j-1},\gamma^*_{iw})\right]_{y_{iu}=1,y_{i\ell}=1,y_{it}=1}, \quad (95)$$

with $\tilde{p}_{i10}(y_{i1},\gamma^*_{iw}) = \frac{\exp\{y_{i1}(x'_{i1}\beta+\gamma^*_{iw})\}}{1+\exp(x'_{i1}\beta+\gamma^*_{iw})}$, and $\tilde{p}_{ij1}(y_{ij},y_{i,j-1},\gamma^*_{iw}) = \frac{\exp\{y_{ij}(x'_{ij}\beta+\theta y_{i,j-1}+\gamma^*_{iw})\}}{1+\exp(x'_{ij}\beta+\theta y_{i,j-1}+\gamma^*_{iw})}$,

where $\gamma^*_{iw} \equiv \gamma^*_{iw}(\lambda_\gamma, \nu; \psi_{iw}, \xi^2_{iw})$ as defined by (25).

The computation for the fourth order moments is similar to that of the third order moments. Let $\sum_{(y_{iu},y_{i\ell},y_{im},y_{it})\ni s^*}$ indicates the summation over all binary variables in the sample space $s^*$ that contain $T-4$ elements out of T elements except $y_{iu}, y_{il}, y_{im}, y_{it}$. Now by implementing this summation, following (95), one writes the formula for the fourth order moments as

$$E[Y_{iu}Y_{i\ell}Y_{im}Y_{it}] = W^{-1}\sum_{w=1}^{W}\sum_{y_{iu},y_{i\ell},y_{im},y_{it}\ni s^*}\left[\tilde{p}_{i10}(y_{i1},\gamma^*_{iw})\right.$$
$$\left.\times\ \Pi_{j=2}^{T}\tilde{p}_{ij1}(y_{ij},y_{i,j-1},\gamma^*_{iw})\right]_{y_{iu}=1,y_{i\ell}=1,y_{im}=1,y_{it}=1}$$
$$= \tilde{\phi}_{i,u\ell mt},\ \text{(say)}. \quad (96)$$

This completes the computation of all moments up to order four. These moments were exploited to construct the GQL estimating Eq. (84) for all parameters involved in the model, namely $\beta,\ \theta,\ \lambda_\gamma,$ and $\nu$.

### 5.2 Asymptotic Normality and Consistency of $\hat{\alpha}_{GQL}$

Following (83), for true $\alpha$, define

$$\bar{g}_K(\alpha) = \frac{1}{K}\sum_{i=1}^{K} g_i(\alpha) = \frac{1}{K}\sum_{i=1}^{K}\frac{\partial \zeta_i'}{\partial \alpha}\Omega_i^{-1}(v_i - \zeta_i), \tag{97}$$

where $v_1, \ldots, v_i, \ldots, v_K$ are independent to each other as they are collected from $K$ independent individuals, but they are not identically distributed because

$$v_i \sim (\zeta_i(\beta, \theta, \lambda_\gamma, \nu), \Omega_i(\beta, \theta, \lambda_\gamma, \nu)), \tag{98}$$

where the mean vectors in (82) and also the covariance matrices in (83) are different for different individuals.

Now one may derive the asymptotic distribution of $\hat{\alpha}_{GQL}$ by using the same technique as for the derivation of the asymptotic distribution of $\hat{\beta}_{GQL}$ given in Sect. 3.1.1 for the Poisson mixed model. Thus, it can be shown that

$$\lim_{K\to\infty} \hat{\alpha}_{GQL} \to N(\alpha, \tilde{V}_K^{-1}(\beta, \theta, \lambda_\gamma, \nu)), \tag{99}$$

or equivalently

$$||[\tilde{V}_K(\beta, \theta, \lambda_\gamma, \nu)]^{\frac{1}{2}}[\hat{\alpha}_{GQL} - \alpha]|| = O_p(\sqrt{p+3}), \tag{100}$$

where

$$\tilde{V}_K(\beta, \theta, \lambda_\gamma, \nu) = \sum_{i=1}^{K}\frac{\partial(\zeta_i)'}{\partial \alpha}[\Omega_i(\beta, \theta, \lambda_\gamma, \nu)]^{-1}\frac{\partial(\zeta_i)}{\partial \alpha'}.$$

This establishes the consistency of $\hat{\alpha}_{GQL}$ for $\alpha$.

## 6 Discussion

It has been assumed in various econometric studies for count and binary panel data that the distribution of the random effects involved in the model is unknown. This makes the estimation of the regression effects $\beta$ and dynamic dependence parameter $\rho$ under the Poisson dynamic mixed model, and the estimation of $\beta$ and the dynamic dependence parameter $\theta$ under the binary dynamic mixed model, very difficult. As a remedy, some authors such as Wooldridge (1999) and Montalvo (1997) developed certain estimation techniques those automatically remove the random effects from the model and estimate rest of the parameters, $\beta$ and $\rho$ in the Poisson case. However as demonstrated by Sutradhar et al. (2014), these estimation approaches have two drawbacks. First, the conditional maximum likelihood (CML) method used by Wooldridge (1999) and the instrumental variables based GMM (IVGMM) method used by Montalvo (1997) become useless for the estimation of the regression effects $\beta$ when covariates are stationary (time independent), even though they are able to remove the random effects. Second, when the random effects $\gamma_i$ or $\gamma_i^*$ are removed technically, the estimates of $\beta$ and the dynamic dependence parameter ($\rho$) alone are not sufficient to compute the mean, variance and correlations of the data, which is a major drawback from the view point of data understanding/analysis. One encounters similar problems with the weighted kernel likelihood approach of Honore and Kyriazidou (2000, p. 84) for the inferences in binary dynamic mixed logit models.

The aforementioned inference issues do not arise when one can assume a suitable distribution for the random effects. Because the random effects appear in the linear predictive function of the generalized linear model, many studies mainly in statistics literature have considered normality as a reasonable assumption for the distribution of the random effects. Thus, under the assumption that the random effects involved in the longitudinal mixed models follow $N(0, \sigma_\gamma^2)$, the GQL estimation of the regression effects $\beta$ and $\sigma_\gamma^2$ and moment estimation of the longitudinal correlation index parameter $\rho$ were developed by Sutradhar and Bari (2007), for example, for longitudinal count data, and by Sutradhar (2008) for longitudinal binary data. See also Breslow and Clayton (1993), Breslow and Lin (1995), Lin and Breslow (1996), Jiang (1998), and Sutradhar and Qu (1998).

However, in this paper we have provided an extension of the normal latent effects based longitudinal mixed models for count and binary data to the $t_\nu$ latent effects based models. The inference for these extended models have been complex not only because of an additional degrees of freedom parameter but also for the difficulty that unlike simulation of $N(0, 1)$ random effects in the Gaussian case, the simulation of $t_\nu(0, 1)$ is not possible when $\nu$ is unknown. In this paper we have resolved this issue through a new transformation which helps to generate data from a $t_4(0, 1)$ distribution for the purpose and then proceed for estimation of the $\nu$ parameter. In summary, we have developed a GQL estimation technique for the estimation of all parameters involved in the models including the degrees of freedom parameter.

**Acknowledgements** The authors thank a referee for comments and suggestions. The second author presented a part of this paper in the symposium as a part of his Key Note address part I. Special thanks go to the audience of the symposium for their feedback.

## References

Amemiya, T.: Advanced Econometrics. Harvard University Press, Cambridge, MA (1985)

Bartolucci, F., Nigro, V.: A dynamic model for binary panel data with unobserved heterogenity admitting a $\sqrt{n}$-consistent conditional estimator. Econometrica **78**, 719–733 (2010)

Breslow, N.E., Clayton, D.G.: Approximate inference in generalized linear mixed models. J. Am. Stat. Assoc. **88**, 9–25 (1993)

Breslow, N.E., Lin, X.: Bias correction in generalized linear models with a single component of dispersion. J. Am. Stat. Assoc. **82**, 81–92 (1995)

Cox, D.R.: The analysis of mulitivariate binary data. Appl. Stat. **21**, 113–120 (1972)

Freeland, R.K., McCabe, B.P.M.: Forecasting discrete valued low count time series. Int. J. Forecast. **20**, 427–434 (2004)

Heckman, J.J.: Statistical models for discrete panel data. In: McFadden, D., Manski, C.F. (eds.) Structural Analysis of Discrete Data with Econometric Applications. MIT Press, Cambridge (1981)

Honore, B.E., Kyriazidou, E.: Panel data discrete choice models with lagged dependent variables. Econometrica **68**, 839–874 (2000)

Jiang, J.: Consistent estimators in generalized linear mixed models. J. Am. Stat. Assoc. **93**, 720–729 (1998)

Lee, Y., Nelder, J.A.: Hierarchical generalized linear models. J. R. Stat. Soc. B **58**, 619–678 (1996)

Lin, X., Breslow, N.E.: Bias correction in generalized linear mixed models with multiple components of dispersion. J. Am. Stat. Assoc. **91**, 1007–1016 (1996)

Manski, C.F.: Semi-parametric analysis of random effects linear models from binary panel data. Econometrica **55**, 357–362 (1987)

McDonald, D.R.: The local limit theorem: a historical perspective. J. Iran. Stat. Soc. **4**, 73–86 (2005)

Montalvo, J.G.: GMM estimation of count-panel data models with fixed effects and predetermined instruments. J. Bus. Econ. Stat. **15**, 82–89 (1997)

Sutradhar, B.C.: On the Characteristics function of multivariate student t-distribution. Can. J. Stat. **14**, 329–337 (1986)

Sutradhar, B.C.: An overview on regression models for discrete longitudinal responses. Stat. Sci. **18**, 377–393 (2003)

Sutradhar, B.C.: On exact quasilikelihood inference in generalized linear mixed models. Sankhya: Indian J. Stat. **66**, 261–289 (2004)

Sutradhar, B.C.: On auto-regression type dynamic mixed models for binary panel data. Metron **56**, 205–217 (2008)

Sutradhar, B.C.: Inferences in generalized linear longitudinal mixed models. Can. J. Stat. **38**, 174–196 (2010) [Special issue Ed. B.C. Sutradhar]

Sutradhar, B.C.: Dynamic Mixed Models for Familial Longitudinal Data. Springer, New York (2011)

Sutradhar, B.C., Bari, W.: On generalized quasilikelihood inference in longitudinal mixed models for count data. Sankhya: Indian J. Stat. **69**, 671–699 (2007)

Sutradhar, B.C., Farrell, P.J.: On optimal lag 1 dependence estimation for dynamic binary models with application to asthma data. Sankhya B: Indian J. Stat. **69**, 448–467 (2007)

Sutradhar, B.C., Qu, Z.: On approximate likelihood inference in Poisson mixed model. Can. J. Stat. **26**, 169–186 (1998)

Sutradhar, B.C., Prabhakar Rao, R., Pandit, V.N.: Generalized method of moments versus generalized quasilikelihood inferences in binary panel data. Sankhya B: Indian J. Stat. **70**, 34–62 (2008)

Sutradhar, B.C., Jowaheer, V., Rao, R.P.: Remarks on asymptotic efficient estimation for regression effects in stationary and non-stationary models for panel count data. Br. J. Probab. Stat. **28**, 241–254 (2014)

Wooldridge, J.: Distribution-free estimation of some non-linear panel data models. J. Econ. **90**, 77–97 (1999)

Zhao, L.P., Prentice, R.L.: Correlated binary regression using a quadratic exponential model. Biometrika **77**, 642–648 (1990)