Brazilian Journal of Probability and Statistics
2014, Vol. 28, No. 2, 241–254
DOI: 10.1214/12-BJPS204

Remarks on asymptotic efficient estimation for regression effects in stationary and nonstationary models for panel count data

Brajendra C. Sutradhar[a], Vandna Jowaheer[b] and R. Prabhakar Rao[c]

[a]*Memorial University*
[b]*University of Mauritius*
[c]*Sri Sathya Sai University, India*

Abstract. In a panel count data setup, repeated counts of an individual are assumed to be influenced by the individual's random effect. Consequently, conditional on the random effect, the repeated responses of the individual are assumed to be serially correlated. Under the assumption that the random effects of the individuals follow a normal distribution, Jowaheer and Sutradhar (*Statist. Probab. Letters* **79** (2009) 1928–1934) have demonstrated that the generalized quasi-likelihood (GQL) estimation approach produces more efficient estimates than the so-called generalized method of moments (GMM) approach for both regression effects and the variance component of the normal random effects. For the cases where the distribution of the random effects is unknown, there exist two estimation approaches, namely the conditional maximum likelihood (CML) and instrumental variables based GMM (IVGMM) approaches, for the estimation of the regression effects. The purpose of this paper is to examine the asymptotic efficiency performances of the CML and IVGMM approaches as compared to the GQL approach for the regression estimation. When the covariates are stationary, that is, time independent, it is, however, known that the CML and IVGMM approaches are useless for the regression estimation, whereas the GQL approach does not encounter any such limitations. For the general case, that is, when the covariates are time dependent, the IVGMM approach appears to be computationally expensive and hence it is not included in efficiency comparison. Between the CML and GQL approaches, it is found through exact asymptotic variance calculations that the GQL approach is asymptotically more efficient than the CML approach in estimating the regression effects. This makes the GQL as a unified efficient approach irrespective of the cases whether the panel count data are stationary or nonstationary.

1 Introduction

Let y_{it} denote the count response for an individual i $(i = 1, \dots, I)$ at time t $(t = 1, \dots, T)$. Also let x_{it} be the p-dimensional vector of fixed covariates corre-

Key words and phrases. Conditional autocorrelations, consistency and efficiency, generalized quasi-likelihood estimating equations, nonstationary counts, random effects causing overdispersion, regression parameters.

Received December 2011; accepted July 2012.

sponding to y_{it}, and β denote the p-dimensional vector of regression parameters. It is also likely that a random effect, say γ_i for the ith individual, will influence the expectation of y_{it}, for all $t = 1, \ldots, T$. Furthermore, conditional on this individual random effect γ_i, any two responses of the ith individual, say y_{iu} and y_{it} for $u < t$, will be serially correlated.

Suppose that y_{i1} conditional on γ_i has the Poisson density, that is,

$$f_{i1}(y_{i1}|\gamma_i) = \frac{\exp(-\mu^*_{i1})\mu^{*y_{i1}}_{i1}}{y_{i1}!}, \tag{1.1}$$

where $\mu^*_{i1} = \exp(x'_{i1}\beta + \gamma_i)$. Next suppose that for $t = 2, \ldots, T$, y_{it} and $y_{i,t-1}$ has the dynamic relationship

$$[y_{it}|\gamma_i] = \rho * [y_{i,t-1}|\gamma_i] + [d_{it}|\gamma_i] \tag{1.2}$$

(Sutradhar (2011, Section 8.1), Sutradhar and Bari (2007)), where $\rho * y_{i,t-1} = \sum_{s=1}^{y_{i,t-1}} b_s(\rho)$ with $\Pr[b_s(\rho) = 1] = \rho$ and $\Pr[b_s(\rho) = 0] = 1 - \rho$. In (1.1), $[y_{i1}|\gamma_i] \sim \text{Poi}(\mu^*_{i1})$, and in (1.2), for $t = 2, \ldots, T$, $[d_{it}|\gamma_i] \sim \text{Poi}(\mu^*_{it} - \rho\mu^*_{i,t-1})$, with $\mu^*_{it} = \exp(x'_{it}\beta + \gamma_i)$, for $t = 1, \ldots, T$. It then follows that the conditional mean and variance of y_{it} conditional on the random effect γ_i are given by

$$E[y_{it}|\gamma_i] = \text{var}[Y_{it}|\gamma_i] = \mu^*_{it} = \exp\bigl(x'_{it}\beta + \gamma_i\bigr), \tag{1.3}$$

and for $u < t$, the lag $(t - u)$ correlation conditional on γ_i has the formula

$$\text{corr}[Y_{iu}, Y_{it}|\gamma_i] = \rho^{t-u}\sqrt{\frac{\mu^*_{iu}}{\mu^*_{it}}}, \tag{1.4}$$

which is free from γ_i. Further suppose that $\gamma_i \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2_\gamma)$ (Breslow and Clayton (1993), Jiang (1998), Sutradhar (2004)). Since by (1.2), y_{it} conditional on γ_i follows the Poisson distribution with parameter μ^*_{it} given in (1.3), y_{it} then unconditionally has the mean and the variance given by

$$\mu_{it}(\beta, \sigma^2_\gamma) = E[Y_{it}] = E_{\gamma_i}E[Y_{it}|\gamma_i] = E_{\gamma_i}\bigl[\mu^*_{it}\bigr] = \exp\bigl[x'_{it}\beta + \sigma^2_\gamma/2\bigr], \tag{1.5}$$

and

$$\begin{aligned}\sigma_{itt}(\beta, \sigma^2_\gamma) &= \text{var}[Y_{it}] = E\bigl[Y^2_{it}\bigr] - \bigl[E(Y_{it})\bigr]^2 = E_{\gamma_i}E\bigl[Y^2_{it}|\gamma_i\bigr] - \mu^2_{it} \\ &= E_{\gamma_i}\bigl[\mu^*_{it} + \mu^{*2}_{it}\bigr] - \mu^2_{it} = \mu_{it} + \bigl[\exp(\sigma^2_\gamma) - 1\bigr]\mu^2_{it},\end{aligned} \tag{1.6}$$

respectively, and the lag $t - u$ unconditional correlation between y_{iu} and y_{it} has the formula

$$\text{corr}(Y_{iu}, Y_{it}) = \frac{\rho^{t-u}(1/\mu_{it}) + [\exp(\sigma^2_\gamma) - 1]}{[\{[\exp(\sigma^2_\gamma) - 1] + 1/\mu_{iu}\}\{[\exp(\sigma^2_\gamma) - 1] + 1/\mu_{it}\}]^{1/2}}. \tag{1.7}$$

In (1.6) and (1.7), μ_{it} is used for $\mu_{it}(\beta, \sigma^2_\gamma)$, without any loss of generality.

Note that the mean (1.5), variance (1.6), and correlations of the responses (1.7) are functions of both β and σ_γ^2. Consequently, to understand the influence of the covariates x_{it} on the mean and the variance of the data, it is not enough to estimate β, rather, both β and σ_γ^2 have to be estimated. Some researchers have, however, gone ahead with the computation of β only to understand the effect of x_{it} on the count response y_{it}. They have exploited certain estimation techniques for β which are unaffected by γ_i, and these techniques are suitable when one cannot assume any distributions for the random effects γ_i. For example, we refer to the conditional maximum likelihood (CML) approach discussed by Wooldridge (1999, eqn. (2.6), p. 79), and the so-called instrumental variables based GMM (IVGMM) approach considered by Montalvao (1997, eqn. (32), p. 85). Further discussions on these approaches are available in Sutradhar (2011, Sections 8.2.1.3 and 8.2.1.4). For convenience of computing the asymptotic variances of these estimators for β, we reproduce these two approaches in brief in the following two sections.

1.1 CML approach for β estimation

Note that when the correlations between any two responses of an individual are ignored, that is when $\rho = 0$ is used in model (1.2), it follows that y_{it} conditional on γ_i has the marginal Poisson distribution given by

$$f_{it}(y_{it}|\gamma_i) = \frac{\exp(-\mu_{it}^*)\mu_{it}^{*y_{it}}}{y_{it}!}, \tag{1.8}$$

with $\mu_{it}^* = \exp(x_{it}'\beta + \gamma_i)$. In the CML approach, conditional on total count $\sum_{t=1}^T y_{it} = n_i$, one first writes a conditional likelihood for the repeated responses under the ith firm as

$$\begin{aligned} L_i(\beta|n_i) &= f_i(y_{i1}, \dots, y_{iT}|n_i) \\ &= \frac{n_i!}{y_{i1}! \cdots y_{i,T-1}!(n_i - \sum_{t=1}^{T-1} y_{it})!} p_{i1}^{y_{i1}} \cdots p_{i,T-1}^{y_{i,T-1}} p_{iT}^{n_i - \sum_{t=1}^{T-1} y_{it}}, \end{aligned} \tag{1.9}$$

where $p_{it} = \mu_{it}^* / \sum_{t=1}^T \mu_{it}^*$. Note that because

$$p_{it} = \frac{\exp(x_{it}'\beta + \gamma_i)}{\exp(\gamma_i)\sum_{t=1}^T \exp(x_{it}'\beta)} = \frac{\exp(x_{it}'\beta)}{\sum_{t=1}^T \exp(x_{it}'\beta)} \tag{1.10}$$

is free from γ_i, the conditional likelihood in (1.9) is also free from γ_i. Consequently, one may estimate β (Montalvo (1997, Section 1), Wooldridge (1999, eqn. (2.6), p. 79)) by maximizing the log-likelihood

$$L^*(\beta) = \log \prod_{i=1}^I L_i(\beta|n_i) = k_0 + \sum_{i=1}^I \sum_{t=1}^T y_{it} \log(p_{it}), \tag{1.11}$$

where k_0 is a constant free from β, and $y_{iT} = n_i - \sum_{t=1}^{T-1} y_{it}$. Further note that the maximization of the log-likelihood function (1.11) for β is equivalent to solve the likelihood estimating equation given by

$$\frac{\partial L^*(\beta)}{\partial \beta} = \sum_{i=1}^{I}\sum_{t=1}^{T} y_{it}\left[x_{it} - \sum_{t=1}^{T} p_{it}x_{it}\right] = 0. \tag{1.12}$$

The CML estimate of β obtained from (1.12) is expected to be consistent but inefficient. The inefficiency arises mainly because of conditioning on the cluster total as well as for ignoring the serial correlations. More specially, when the data are serially correlated, the independence assumption based conditional likelihood (1.11) is no longer a valid likelihood. Hence, it is bound to produce inefficient estimate.

1.2 Instrumental variables based GMM estimation approach

As opposed to the conditioning on the total count as in the last section, Montalvo (1997, eqn. 32, p. 85) considered the lag 1 based differences, namely

$$\psi_{it}(\beta) = y_{it} - y_{i,t-1}\exp\left[(x_{it} - x_{i,t-1})'\beta\right] \qquad \text{for } t = 2, \ldots, T$$

which is unbiased for zero irrespective of the distribution of γ_i. Next, by exploiting the $(T-1) \times 1$ vector

$$\psi_i(\beta) = \left[\psi_{i2}(\beta), \ldots, \psi_{it}(\beta), \ldots, \psi_{iT}(\beta)\right]',$$

Montalvo (1997, eqn. 36) obtained a GMM estimate for β by minimizing the quadratic distance function

$$D = \left[\sum_{i=1}^{I} Z_i'\psi_i(\beta)\right]' W^{-1}\left[\sum_{i=1}^{I} Z_i'\psi_i(\beta)\right], \tag{1.13}$$

where Z_i is the $(T-1) \times p\{\frac{T(T+1)}{2} - 1\}$ instrumental matrix given by

$$Z_i = \begin{bmatrix} z_{i2} & 0 & 0 & \cdots & 0 \\ 0 & z_{i3} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & z_{iT} \end{bmatrix}, \tag{1.14}$$

with $z_{it} = [x_{it}', x_{i(t-1)}', \ldots, x_{i1}']$, and where

$$W = \frac{1}{I}\sum_{i=1}^{I} Z_i'\psi_i(\beta)\psi_i'(\beta)Z_i.$$

Note that obtaining β by minimizing the distance function D in (1.13) is equivalent to solve the estimating equation

$$\left[\sum_{i=1}^{I} \frac{\partial \psi_i'}{\partial \beta} Z_i\right] W^{-1}\left[\sum_{i=1}^{I} Z_i'\psi_i(\beta)\right] = 0 \tag{1.15}$$

for β, where $\frac{\partial \psi_i'}{\partial \beta}$ is obtained by using the formula for the general element

$$\frac{\partial \psi_{it}}{\partial \beta} = -y_{i,t-1}[x_{it} - x_{i,t-1}] \exp\big[(x_{it} - x_{i,t-1})'\beta\big].$$

But, the use of a sandwich type covariance matrix estimate $\hat{W}(\hat{\beta}_r)$ in the distance function D in (1.13) may cause bias and hence inconsistency, because of the repeated use of iterative estimated values for the parameter of interest.

In the next section, we explain the GQL approach in brief for the estimation of β, for known σ_γ^2 and ρ. This we do for the purpose of comparing the asymptotic variances of the estimators of β only.

1.3 GQL estimation of β

Let $y_i = (y_{i1}, \ldots, y_{it}, \ldots, y_{iT})'$ be the $T \times 1$ response vector, and $\mu_i = (\mu_{i1}, \ldots, \mu_{it}, \ldots, \mu_{iT})'$ is the mean and $\Sigma_i(\rho, \sigma_\gamma^2) = (\sigma_{iut})$ is the covariance matrix of y_i. Recall that the formulas for μ_{it} and σ_{itt} are given in (1.5) and (1.6), respectively. Also it follows by (1.7) that σ_{iut} has the formula

$$\sigma_{iut} = \mathrm{corr}(Y_{iu}, Y_{it})[\sigma_{iuu}\sigma_{itt}]^{1/2}.$$

Note that for known σ_γ^2 and ρ, one may obtain the GQL estimate of β by solving the GQL estimating equation

$$\sum_{i=1}^{I} \frac{\partial \mu_i'}{\partial \beta} \Sigma_i^{-1}(\rho, \sigma^2)(y_i - \mu_i) = 0 \tag{1.16}$$

(Sutradhar (2003)).

Note that the GQL estimation for β by solving (1.16) requires the knowledge of the distribution of the random effects γ_i, whereas the CML estimation by (1.12) and the IVGMM estimation by solving (1.15) do not require this distribution because these later equations are technically constructed free of γ_i's. This however does not imply that the properties such as the variances of the CML and IVGMM estimators can be unaffected by the distribution of the random effects. This is evident from the formulas for the variances of the estimators those we derive in the next section. A further empirical study by considering nonnormal random effects and examining their effects on the GQL estimation and on the variances of all estimators would have been more revealing, but this type of robust analysis is beyond the scope of the present paper.

2 Asymptotic variances of the regression estimators

Recall from Section 1.1 that for known σ_γ^2 and ρ, the CML estimate of β is obtained by solving the likelihood estimating equation (1.12). Let $\hat{\beta}_{\mathrm{CML}}$ denote this estimate, which we obtain by using the iterative equation

$$
\begin{aligned}
\hat{\beta}_{\mathrm{CML}}(r+1) &= \hat{\beta}_{\mathrm{CML}}(r) - \left[\left\{\frac{\partial^2 L^*(\beta)}{\partial\beta\,\partial\beta'}\right\}^{-1} \frac{\partial L^*(\beta)}{\partial\beta}\right]\Bigg|_{\beta=\hat{\beta}_{\mathrm{CML}}(r)} \\
&= \hat{\beta}_{\mathrm{CML}}(r) \\
&\quad + \left[\left\{\sum_{i=1}^{I}\left(\sum_{t=1}^{T} y_{it}\right)\left(\sum_{t=1}^{T} p_{it}x_{it}x_{it}' - \sum_{t=1}^{T} p_{it}x_{it}\sum_{t=1}^{T} p_{it}x_{it}'\right)\right\}^{-1}\right. \\
&\quad \left.\times \sum_{i=1}^{I}\sum_{t=1}^{T} y_{it}\left\{x_{it} - \sum_{t=1}^{T} p_{it}x_{it}\right\}\right]\Bigg|_{\beta=\hat{\beta}_{\mathrm{ML}}(r)},
\end{aligned}
\tag{2.1}
$$

where $\hat{\beta}_{\mathrm{CML}}(r)$ is the value of β at rth iteration. It then follows from Theorem 3.4 of Newey and McFadden (1993), for example, that with probability approaching 1, the solution $\hat{\beta}_{\mathrm{CML}}$ obtained from (2.1) would be unique satisfying

$$
\begin{aligned}
&\sqrt{I}(\hat{\beta}_{\mathrm{CML}} - \beta) \\
&\quad = E\left[I^{-1}\left\{\sum_{i=1}^{I}\left(\sum_{t=1}^{T} y_{it}\right)\left(\sum_{t=1}^{T} p_{it}x_{it}x_{it}' - \sum_{t=1}^{T} p_{it}x_{it}\sum_{t=1}^{T} p_{it}x_{it}'\right)\right\}\right]^{-1} \\
&\qquad \times I^{-1/2}\left[\sum_{i=1}^{I}\sum_{t=1}^{T} y_{it}\left\{x_{it} - \sum_{t=1}^{T} p_{it}x_{it}\right\}\right] + o_p(1),
\end{aligned}
\tag{2.2}
$$

implying the consistency of $\hat{\beta}_{\mathrm{CML}}$ for β. Next, by using the Lindeberg–Feller central limit theorem (see Amemiya (1985, Theorem 3.3.6, p. 92)), for example, it follows from (2.2) that asymptotically (as $I \to \infty$)

$$
\sqrt{I}(\hat{\beta}_{\mathrm{CML}} - \beta) \sim N_p(0, IV_{\mathrm{CML}}), \tag{2.3}
$$

where

$$
V_{\mathrm{CML}} = \left[\sum_{i=1}^{I} X_i^{*\prime} A_i X_i^*\right]^{-1}\left(\sum_{i=1}^{I} X_i^{*\prime}\Sigma_i X_i^*\right)\left[\sum_{i=1}^{I} X_i^{*\prime} A_i X_i^*\right]^{-1}, \tag{2.4}
$$

where

$$
X_i^{*\prime} = [x_{i1}^* \dots x_{it}^* \dots x_{iT}^*], \qquad A_i = \mathrm{diag}[\mu_{i1}, \dots, \mu_{iT}] \quad \text{and} \quad \Sigma_i = (\sigma_{iut}),
$$

with $x_{it}^* = [x_{it} - \sum_{t=1}^{T} p_{it}x_{it}]$, $p_{it} = \frac{\mu_{it}}{\sum_{t=1}^{T}\mu_{it}}$, where $\mu_{it} = \exp(x_{it}'\beta + \frac{1}{2}\sigma_\gamma^2)$, $\sigma_{itt} = \mu_{it} + [\exp(\sigma_\gamma^2) - 1]\mu_{it}^2$ and $\sigma_{iut} = \rho^{t-u}\mu_{iu} + [\exp(\sigma_\gamma^2) - 1]\mu_{iu}\mu_{it}$ for $u < t$.

Similarly, for known σ_γ^2 and ρ, the IVGMM estimate of β, say $\hat{\beta}_{\mathrm{IVGMM}}$, is obtained by solving the estimating equation (1.15). This may be achieved by using the iterative equation

$$
\begin{aligned}
&\hat{\beta}_{\mathrm{IVGMM}}(r+1) \\
&\quad = \hat{\beta}_{\mathrm{IVGMM}}(r) - \left[\left[\left\{\sum_{i=1}^{I} \frac{\partial \psi_i'}{\partial \beta} Z_i\right\} W^{-1} \left\{\sum_{i=1}^{I} Z_i' \frac{\partial \psi_i(\beta)}{\partial \beta'}\right\}\right]^{-1}\right. \\
&\qquad \left.\left. \times \left[\sum_{i=1}^{I} \frac{\partial \psi_i'}{\partial \beta} Z_i\right] W^{-1} \left[\sum_{i=1}^{I} Z_i' \psi_i(\beta)\right]\right]\right|_{\beta=\hat{\beta}_{\mathrm{IVGMM}}(r)},
\end{aligned} \tag{2.5}
$$

where $\hat{\beta}_{\mathrm{IVGMM}}(r)$ is the value of β at rth iteration. By similar calculations as in (2.1)–(2.3), it follows that

$$
\sqrt{I}(\hat{\beta}_{\mathrm{IVGMM}} - \beta) \sim N_p(0, I V_{\mathrm{IVGMM}}), \tag{2.6}
$$

where

$$
\begin{aligned}
V_{\mathrm{IVGMM}} = &\left[\left\{\sum_{i=1}^{I} \frac{\partial \psi_i'}{\partial \beta} Z_i\right\} W^{-1} \left\{\sum_{i=1}^{I} Z_i' \frac{\partial \psi_i(\beta)}{\partial \beta'}\right\}\right]^{-1} \\
&\times \left[\sum_{i=1}^{I} \frac{\partial \psi_i'}{\partial \beta} Z_i\right] W^{-1} \left[\sum_{i=1}^{I} Z_i' \operatorname{cov}(\psi_i(\beta)) Z_i\right] \\
&\times W^{-1} \left[\sum_{i=1}^{I} Z_i' \frac{\partial \psi_i(\beta)}{\partial \beta'}\right] \\
&\times \left[\left\{\sum_{i=1}^{I} \frac{\partial \psi_i'}{\partial \beta} Z_i\right\} W^{-1} \left\{\sum_{i=1}^{I} Z_i' \frac{\partial \psi_i(\beta)}{\partial \beta'}\right\}\right]^{-1}.
\end{aligned} \tag{2.7}
$$

Next, for known values of σ_γ^2 and ρ, we solve the GQL estimating equation (1.16) for the regression effects β. Let $\hat{\beta}_{\mathrm{GQL}}$ be the solution of (1.16), which may be obtained by using the iterative equation

$$
\begin{aligned}
\hat{\beta}_{\mathrm{GQL}}(r+1) = \hat{\beta}_{\mathrm{GQL}}(r) + &\left[\left[\left\{\sum_{i=1}^{I} \frac{\partial \mu_i'}{\partial \beta} \Sigma_i^{-1}(\rho, \sigma_\gamma^2) \frac{\partial \mu_i}{\partial \beta'}\right\}^{-1}\right.\right. \\
&\left.\left.\left. \times \left\{\sum_{i=1}^{I} \frac{\partial \mu_i'}{\partial \beta} \Sigma_i^{-1}(\rho, \sigma_\gamma^2)(y_i - \mu_i)\right\}\right]\right]\right|_{\beta=\hat{\beta}_{\mathrm{GQL}}(r)},
\end{aligned} \tag{2.8}
$$

where $\hat{\beta}_{\mathrm{GQL}}(r)$ is the value of β at rth iteration. Once again following (2.1)–(2.3), one obtains by (2.8) that

$$
\sqrt{I}(\hat{\beta}_{\mathrm{GQL}} - \beta) \sim N_p(0, I V_{\mathrm{GQL}}), \tag{2.9}
$$

where

$$V_{\mathrm{GQL}} = \left\{ \sum_{i=1}^{I} \frac{\partial \mu_i'}{\partial \beta} \Sigma_i^{-1}(\rho, \sigma_\gamma^2) \frac{\partial \mu_i}{\partial \beta'} \right\}^{-1}. \tag{2.10}$$

Note that the computation of $\hat{\beta}_{\mathrm{IVGMM}}$ by (2.5) (see also (1.15)) is much more involved than the computation of $\hat{\beta}_{\mathrm{CML}}$ by (2.1) (see also (1.12)). Furthermore, the IVGMM estimate obtained from (1.15) will in fact produce less efficient estimate than the CML approach. This is because, the IVGMM estimating equation (1.15) uses only lag 1 pair-wise responses, whereas the CML approach uses all possible responses in the cluster to form the likelihood function (1.9). Consequently, we will not include the IVGMM approach for empirical asymptotic variance comparison in the next section. More specifically, in the next section, we compare the asymptotic variances for the CML and GQL estimators of β empirically. This we do under two scenarios, first, for the stationary case when covariates are time independent, and then for the general case under the nonstationary model.

3 Exact asymptotic variance computation

It is of interest here to compare the asymptotic variances of the estimators for the regression effects β under different methods. This we do for known values of σ_γ^2 and ρ. Note that even though the estimating equations for β under any methods involve the estimates of other parameters, namely $\hat{\sigma}_\gamma^2$ and $\hat{\rho}$, the asymptotic variances of the estimators such as the formulas in (2.4), (2.6) and (2.10), for CML, IVGMM, and GQL estimators, respectively, are obtained for known σ_γ^2 and ρ. This is quite reasonable because the nuisance parameters estimates are supposed to be consistent and the variances of the main regression estimators are computed by using the converged (in probability) values of the estimators $\hat{\sigma}_\gamma^2$ and $\hat{\rho}$. We also remark that unlike in some linear model studies, the computation for the exact variance formulas for regression estimators using the nuisance parameter estimates in their equations, would be extremely complicated, which is avoided by using the true values of the nuisance parameters in the equations while computing the variances. Further remark that the variances may however be consistently computed by using back the consistent estimates of the nuisance parameters.

3.1 Stationary case

In some situations in practice it may happen that the covariates of an individual are time independent, that is, $x_{it} = \tilde{x}_i$ for all $t = 1, \ldots, T$. There appears to be a serious problem with the estimation of the regression effects β by the CML and IVGMM approaches, in such a stationary situation. In fact, in this special case, β is not estimable by these two approaches. This is because, if $x_{it} = \tilde{x}_i$, then

$$p_{it} = \frac{\exp(x_{it}'\beta)}{\sum_{t=1}^{T} \exp(x_{it}'\beta)} = \frac{\exp(\tilde{x}_i'\beta)}{\sum_{t=1}^{T} \exp(\tilde{x}_i'\beta)} = \frac{1}{T}$$

becomes free from β, and consequently, the CML estimating function

$$\sum_{i=1}^{I}\sum_{t=1}^{T} y_{it}\left[x_{it} - \sum_{t=1}^{T} p_{it}x_{it}\right]$$

in CML estimating equation (1.12) becomes free from β. Thus, the estimation breaks down and no questions arises to compute the asymptotic variance of the estimator by using (2.1). The IVGMM approach also encounters the same problem for β estimation. This is because, in this special stationary case,

$$\psi_{it}(\beta) = y_{it} - y_{i,t-1}\exp\big[(x_{it} - x_{i,t-1})'\beta\big] = y_{it} - y_{i,t-1}$$

becomes free from β. Thus, the IVGMM estimating equation (2.5) is no longer a function of β, and hence β is not estimable.

As opposed to the CML and IVGMM approaches, the GQL approach does not encounter any problems in β estimation. Note that for

$$\tilde{\mu}_i = \exp\left(\tilde{x}_i'\beta + \frac{1}{2}\sigma_\gamma^2\right) \quad \text{and} \quad a_i = \big(\exp(\sigma_\gamma^2) - 1\big)\tilde{\mu}_i^2,$$

the covariance matrix Σ_i (see (1.6) and (1.7)) under the stationarity assumption may be written as

$$\Sigma_i = \tilde{\mu}_i\left[C + \frac{a_i}{\tilde{\mu}_i}1_T 1_T'\right], \tag{3.1}$$

where

$$C = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-2} & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-3} & \rho^{T-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & \rho & 1 \end{bmatrix}. \tag{3.2}$$

One may then simplify the GQL estimating equation for β (1.16) as

$$\sum_{i=1}^{I} \tilde{\mu}_i^{-1} b_i d_i(y)\tilde{x}_i = 0, \tag{3.3}$$

where

$$\begin{aligned} b_i &= \frac{1}{1 + 1_T' C_i^{-1} 1_T (a_i/\tilde{\mu}_i)}, \qquad \text{with} \\ d_i(y) &= 1_T' C^{-1}(y_i - \tilde{\mu}_i 1_T). \end{aligned} \tag{3.4}$$

The asymptotic variance of $\hat{\beta}_{\mathrm{GQL}}$ obtained from (2.10) has the formula

$$\operatorname{cov}[\hat{\beta}_{\mathrm{GQL}}] = \left[\sum_{i=1}^{I} s_i \tilde{\mu}_i^2 \tilde{x}_i \tilde{x}_i'\right]^{-1}, \tag{3.5}$$

where

$$s_i = 1'_T \Sigma_i^{-1} 1_T = \left(\frac{b_i}{\tilde{\mu}_i}\right) 1'_T C^{-1} 1_T,$$

with

$$C^{-1} = \frac{1}{(1-\rho^2)} \begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}. \tag{3.6}$$

In a further special case with $p = 1$, for example, the variance of $\hat{\beta}_{\text{GQL}}$ has a simple formula

$$\text{cov}[\hat{\beta}_{\text{GQL}}] = \left[\sum_{i=1}^{I} s_i \tilde{\mu}_i^2 \tilde{x}_i^2\right]^{-1}. \tag{3.7}$$

In order to have a feel how this variance of the GQL estimator of β may change when longitudinal correlation (ρ) as well as variance of the random effects σ_γ^2 vary, we now compute the exact asymptotic variance by using the formula (3.7) for $I = 200$. We use $T = 3$, and $\beta = 1.0$. Next, we consider $\tilde{x}_i = 1$ for $i = 1, \ldots, I/2$, and $\tilde{x}_i = -1$ for $i = I/2+1, \ldots, 200$. The asymptotic variances were found to be

ρ	σ_γ^2 values	Corresponding asymptotic variances
0.0	0.0, 0.5, 1.0, 2.0	4.59×10^{-4}, 2.79×10^{-4}, 1.69×10^{-4}, 6.21×10^{-5}
0.3	0.0, 0.5, 1.0, 2.0	6.64×10^{-4}, 4.02×10^{-4}, 2.44×10^{-4}, 8.99×10^{-5}
0.5	0.0, 0.5, 1.0, 2.0	8.27×10^{-4}, 5.02×10^{-4}, 3.04×10^{-4}, 1.12×10^{-4}
0.8	0.0, 0.5, 1.0, 2.0	1.13×10^{-3}, 6.84×10^{-4}, 4.15×10^{-4}, 1.53×10^{-4}

The asymptotic variances of the GQL estimator of β appear to decrease as σ_γ^2 increases but they increase as the longitudinal correlation ρ gets larger. Thus, under stationary correlation model, that is when the covariates of an individual are time independent, the CML and the IVGMM approaches break down completely for the estimation of the regression effects, whereas the GQL approach produces consistent estimates with very small asymptotic variances.

3.2 Nonstationary case

Since the IVGMM approach appears to be much more complicated computationally than the CML approach, in this section, we make a comparison between the exact asymptotic variances of the GQL and CML estimators of β for a specialized time dependent covariates design with $T = 3$. In the next section, we discuss a simulation study based finite sample efficiency comparison among the three competitive approaches including the IVGMM approach. For the exact asymptotic variance computation, we consider $I = 300$ individuals with same time dependent covariates, for example. Suppose that

$$x_{i1} = -1.0, \qquad x_{i2} = 0 \quad \text{and} \quad x_{i3} = 1.0$$

for all $i = 1, \dots, 300$. Further, let the true values of the parameters are:

$$\beta = 1.0, \qquad \sigma_\gamma^2 = 2.0 \quad \text{and} \quad \rho = 0.0, 0.50 \text{ and } 0.8.$$

These design covariates and parameter values yield

$$\mu_{i1} = 1.0, \qquad \mu_{i2} = e^1 \quad \text{and} \quad \mu_{i3} = e^2;$$

and

$$\Sigma_i = \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & e^1 & \rho e^1 \\ \rho^2 & \rho e^1 & e^2 \end{bmatrix} + \left(e^2 - 1\right) \begin{bmatrix} 1 & e^1 & e^2 \\ e^1 & e^2 & e^3 \\ e^2 & e^3 & e^4 \end{bmatrix} \tag{3.8}$$

for all $i = 1, \dots, 300$. We will use $\Sigma_i = \Sigma$ for all $i = 1, \dots, 300$.

Next, since

$$\begin{aligned} \frac{\partial \mu_i'}{\partial \beta} &= [\mu_{i1} x_{i1}, \mu_{i2} x_{i2}, \mu_{i3} x_{i3}] \\ &= \left[-1.0, 0, e^2\right] \end{aligned}$$

for all $i = 1, \dots, 300$, by (2.10), we compute the asymptotic variance of $\hat{\beta}_{\text{GQL}}$ by using

$$\text{var}(\hat{\beta}_{\text{GQL}}) = \frac{1}{300} \left[(-1.0, 0, e^2) \Sigma^{-1} \begin{pmatrix} -1.0 \\ 0 \\ e^2 \end{pmatrix} \right]^{-1}, \tag{3.9}$$

with Σ matrix as in (3.8).

To compute the asymptotic variance of $\hat{\beta}_{\text{CML}}$ by (2.4), we note that

$$p_{i1} = \frac{e^{-1}}{1 + e^1 + e^{-1}}, \qquad p_{i2} = \frac{1}{1 + e^1 + e^{-1}}, \qquad p_{i3} = \frac{e^1}{1 + e^1 + e^{-1}},$$

yielding

$$\sum_{t=1}^{T} p_{it} x_{it} = \frac{e^1 - e^{-1}}{1 + e^1 + e^{-1}}$$

for all $i = 1, \ldots, 300$. Further, they yield

$$x_{it}^* = E[Z_i Z_j] = x_{it} - \sum_{t=1}^{T} p_{it} x_{it} = \begin{cases} -\dfrac{1+2e^1}{1+e^1+e^{-1}} & \text{for } t=1, \\ \dfrac{e^{-1}-e^1}{1+e^1+e^{-1}} & \text{for } t=2, \\ \dfrac{1+2e^{-1}}{1+e^1+e^{-1}} & \text{for } t=3, \end{cases} \tag{3.10}$$

and

$$\sum_{t=1}^{3} \mu_{it} x_{it}^{*2} = \frac{6e^2+6e^1+e^{-1}+5}{[1+e^1+e^{-1}]^2} \tag{3.11}$$

for all $i = 1, \ldots, 300$. We now follow (2.4) and compute the asymptotic variance of $\hat{\beta}_{\mathrm{CML}}$ by using

$$\begin{aligned} \operatorname{var}(\hat{\beta}_{\mathrm{CML}}) &= \left[\sum_{i=1}^{300}\sum_{t=1}^{3} \mu_{it} x_{it}^{*2}\right]^{-2} \sum_{i=1}^{300} \left[(x_{i1}^*, x_{i2}^*, x_{i3}^*) \Sigma \begin{pmatrix} x_{i1}^* \\ x_{i2}^* \\ x_{i3}^* \end{pmatrix} \right] \\ &= \frac{1}{300} \frac{(1+e^1+e^{-1})^4}{[6e^2+6e^1+e^{-1}+5]^2} \left[(x_{\cdot 1}^*, x_{\cdot 2}^*, x_{\cdot 3}^*) \Sigma \begin{pmatrix} x_{\cdot 1}^* \\ x_{\cdot 2}^* \\ x_{\cdot 3}^* \end{pmatrix} \right], \end{aligned} \tag{3.12}$$

where $x_{\cdot t}^* \equiv x_{it}^*$ for $t = 1, \ldots, 3$, and all $i = 1, \ldots, 300$.

The formulas for the asymptotic variances in (3.9) and (3.12) are computationally straightforward. Under the special case with $I = 300$, $p = 1$, $T = 3$, $\beta = 1.0$ and $\sigma_\gamma^2 = 2.0$, we obtain the exact asymptotic variances as

ρ	$\operatorname{var}(\hat{\beta}_{\mathrm{CML}})$	$\operatorname{var}(\hat{\beta}_{\mathrm{GQL}})$
0.0	1.01×10^{-3}	6.99×10^{-4}
0.5	9.85×10^{-4}	6.78×10^{-4}
0.8	9.05×10^{-4}	6.15×10^{-4}

These results show that the asymptotic variance of the GQL estimator of β is always smaller than that of the CML estimator, irrespective of the value for the longitudinal correlation parameter ρ.

For a finite sample based relative performances of the GQL, CML, as well as IVGMM estimators of the regression effect β, we, for example, refer to the simulation study reported by Sutradhar (2011, Table 8.1, p. 331). This finite sample study also shows the efficiency gain of the GQL approach over the CML approach. More specifically, as opposed to $I = 300$ considered above under an asymptotic study, we now consider $I = 100$ individuals each with $p = 2$ time dependent covariates over a period of $T = 4$ time points. These covariates are given by

$$x_{it1} = \begin{cases} 0.0 & \text{for } i = 1, \ldots, I/2; t = 1, 2, \\ 1.0 & \text{for } i = 1, \ldots, I/2; t = 3, T, \\ 1.0 & \text{for } i = I/2 + 1, \ldots, I; t = 1, 2, \\ 1.5 & \text{for } i = I/2 + 1, \ldots, I; t = 3, T \end{cases}$$

and

$$x_{it2} = \begin{cases} 0.05 + 0.10(t-1) & \text{for } i = 1, \ldots, I/4; t = 1, \ldots, \mathrm{T}, \\ \frac{t}{4} & \text{for } i = I/4 + 1, \ldots, I/2; t = 1, \ldots, T, \\ 0.0 & \text{for } i = I/2 + 1, \ldots, 3I/4; t = 1, 2, \\ 1.0 & \text{for } i = I/2 + 1, \ldots, 3I/4; t = 3, T, \\ -1.0 & \text{for } i = 3I/4 + 1, \ldots, I; t = 1, 2, \\ 1.0 & \text{for } i = 3I/4 + 1, \ldots, I; t = 3, T. \end{cases}$$

Further consider $\gamma_i \overset{\text{i.i.d.}}{\sim} N(0, \sigma_\gamma^2 = 1.0)$, two regression parameters as $\beta_1 = \beta_2 = 0.0$, and the longitudinal correlation index parameter $\rho = 0.5$. By using the above design covariates and design parameters, the longitudinal count data $\{y_{it}, i = 1, \ldots, I; t = 1, \ldots, T\}$ are generated following the dynamic model (1.1)–(1.2). Next by applying the data $\{y_{it}\}$ and the covariates x_{it1}, x_{it2}, the CML, IVGMM, and GQL estimates were obtained by using the iterative equations (2.1), (2.5) and (2.8), respectively. The data generation and hence estimates were obtained 500 times. For known $\sigma_\gamma^2 = 1.0$ and $\rho = 0.5$, it was reported in Sutradhar (2011) that the simulated standard errors (SSEs) of these estimates for β_1 were found to be 0.109, 1.163 and 0.066, respectively. Similarly, the simulated standard errors (SSEs) of the CML, IVGMM and GQL estimates for β_2 were found to be 0.138, 1.434 and 0.127, respectively. When the SSEs of the CML and GQL estimates are compared it is clear that the GQL approach produces regression estimates with smaller standard errors as compared to the CML approach. Thus, this finite sample based results are in agreement with the relative asymptotic efficiencies discussed above.

Acknowledgments

The authors are grateful to Bhagavan Sri Sathya Sai Baba for providing opportunities to carry out a part of this research at the Sri Sathya Sai University, India. The authors would like to thank the editor and a referee for valuable comments and suggestions leading to improvement of the paper.

References

Amemiya, T. (1985). *Advanced Econometrics*. Cambridge: Harvard Univ. Press.

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of American Statistical Association* **88**, 9–25.

Jiang, J. (1998). Consistent estimators in generalized linear mixed models. *Journal of the American Statistical Association* **93**, 720–729. MR1631373

Jowaheer, V. and Sutradhar, B. C. (2009). GMM versus GQL inferences for panel count data. *Statistics & Probability Letters* **79**, 1928–1934. MR2750820

Montalvo, J. G. (1997). GMM estimation of count-panel-data models with fixed effects and predetermined instruments. *Journal of Business & Economic Statistics* **15**, 82–89.

Newey, W. K. and McFadden, D. (1993). Estimation in large samples. In *Handbook of Economics, Vol. 4* (D. McFadden and R. Engler, eds.). Amsterdam: North-Holland.

Sutradhar, B. C. (2003). An overview of regression models for discrete longitudinal responses. *Statistical Science* **18**, 377–393. MR2056579

Sutradhar, B. C. (2004). On exact quasilikelihood inference in generalized linear mixed models. *Sankhya B* **66**, 261–289. MR2090978

Sutradhar, B. C. (2011). *Dynamic Mixed Models for Familial Longitudinal Data*. New York: Springer. MR2777359

Sutradhar, B. C. and Bari, W. (2007). On generalized quasilikelihood inference in longitudinal mixed model for count data. *Sankhya B* **69**, 671–699. MR2521228

Wooldridge, J. (1999). Distribution-free estimation of some non-linear panel data models. *Journal of Econometrics* **90**, 77–97. MR1682401

B. C. Sutradhar
Department of Mathematics and Statistics
Memorial University
St. John's A1C 5S7
Canada
E-mail: bsutradh@mun.ca

V. Jowaheer
Department of Mathematics
University of Mauritius
Reduit
Mauritius

R. P. Rao
Department of Economics
Sri Sathya Sai University
Prasanthi Nilayam, Andhra Pradesh 515134
India