Regression analysis of multivariate panel count data

XIN HE

Department of Statistics, University of Missouri, 146 Middlebush Hall, Columbia, MO 65211-6100, USA

XINGWEI TONG*

Department of Statistics, Beijing Normal University, Beijing 100875, People’s Republic of China
xweitong@bnu.edu.cn

JIANGUO SUN

Department of Statistics, University of Missouri, 146 Middlebush Hall, Columbia, MO 65211-6100, USA

RICHARD J. COOK

Department of Statistics and Actuarial Sciences, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

SUMMARY

We consider panel count data which are frequently obtained in prospective studies involving recurrent events that are only detected and recorded at periodic assessment times. The data take the form of counts of the cumulative number of events detected at each inspection time, along with explanatory covariates. Examples arise in diverse areas such as epidemiological studies, medical follow-up studies, reliability studies, and tumorigenicity experiments. This article is concerned with regression analysis of multivariate panel count data which arise if more than one type of recurrent event is of interest and individuals are only observed intermittently. We present a class of marginal mean models which leave the dependence structures for related types of recurrent events completely unspecified. Estimating equations are developed for regression parameters, and the resulting estimates are shown to be consistent and asymptotically normal. Simulation studies show that the proposed estimation procedures work well for practical situations. The methodology is applied to a motivating study of patients with psoriatic arthritis in which the events of interest are the onset of joint damage according to 2 different criteria.

Keywords: Counting processes; Estimating equations; Marginal mean model; Multivariate recurrent events; Observation processes.

*To whom correspondence should be addressed.
1. INTRODUCTION

Panel count data arise in studies of recurrent events when each subject is observed only at finite discrete time points instead of continuously (Sun and Wei, 2000; Zhang, 2002). In such settings, observations are taken at several distinct time points and only the number of events that occurred between observation times is known; no information is available on subjects between the observation time points. This frequently happens in prospective cohort studies, population-based epidemiological studies, reliability studies, and tumorigenicity experiments since in these situations, it is either impossible or not practical to maintain continuous observation of subjects. This paper discusses regression analysis of multivariate panel count data.

Multivariate panel count data arise in studies involving several types of recurrent events in which patients are examined only at periodic follow-up assessments. Chen and others (2005) described a study of patients with advanced breast cancer where the events are the development of different types of metastatic bone lesions that are only detectable by bone scans of the entire skeleton carried out when patients visited participating clinical centers. The number of examinations varied from patient to patient, and at each examination, the number of new lesions developed since the previous examination was recorded. Three types of bone lesions arise in these patients, and the interest was in making statements about their respective rates of occurrence of the different events and related covariate effects. Another common example arises in tumorigenicity experiments when several types of tumors can occur together and are of interest. We consider a third example arising from a cohort study of patients with psoriatic arthritis conducted at the University of Toronto Psoriatic Arthritis Clinic where the event of interest is the development of joint damage. Clinicians are interested in damage as measured by radiographic changes as well as loss in function as detected by functional examination, and these constitute the 2 types of events.

Several authors have considered the analysis of univariate panel count data. Thall and Lachin (1988) and Sun and Kalbfleisch (1993) discussed the treatment comparison problem when only panel count data are available. Sun and Kalbfleisch (1995) and Wellner and Zhang (2000) investigated nonparametric estimation of the cumulative mean function of the underlying point process that generates panel count data. Sun and Wei (2000) and Zhang (2002) gave some approaches for regression analysis of panel count data. For multivariate panel count data, Chen and others (2005) proposed 2 approaches based on a mixed Poisson model with piecewise constant baseline intensities. One approach assumes that the different types of recurrent event are related through multivariate log-normal random effects and bases inference on the resulting full likelihood, while the other makes use of the marginal model approach. In the following, a marginal model approach is presented that avoids the Poisson and piecewise constant baseline intensity assumptions.

With multivariate panel count data, one may separately apply methods for univariate panel count data to each type of event. As with multivariate failure time data (Cai and Prentice, 1995; Wei and others, 1989), it is apparent that this would be less efficient than conducting a joint or multivariate analysis if the different types of recurrent events are related and associated covariate effects are the same. Multivariate analyses can, however, also be more efficient even if some covariate effects are different. Finally, separate univariate analyses, unlike multivariate analyses, cannot estimate the correlations between different covariate effects. More discussion on this is given below.

The remainder of the paper is organized as follows. First, we will describe the panel count data arising from the University of Toronto Psoriatic Arthritis Clinic in more details in Section 2. Notations and models that are used throughout the paper are introduced in Section 3. In particular, marginal mean models are employed for the underlying counting processes that characterize panel count data and observation times. One major advantage of these models is that they leave the dependence structures for related types of recurrent events completely arbitrary. Estimating equations are proposed in Section 4 that give consistent and asymptotically normal estimates of regression parameters. In Section 5, some
results from simulation studies for the evaluation of the proposed estimators are presented; they suggest that the presented estimator works well for practical situations. The method is applied to the psoriatic arthritis data discussed above in Section 6 and discussion and concluding remarks are given in Section 7.

2. The University of Toronto Psoriatic Arthritis data

The University of Toronto Psoriatic Arthritis Clinic was established in 1978 in order to collect data on the functional and radiological courses of disease for patients with psoriatic arthritis (Gladman and others, 1995). Functional assessments have been scheduled annually, during which patients are to undergo a detailed physical examination including a careful assessment of each of 64 joints. A joint is classified as damaged by a “functional assessment” if there is evidence of deformity or ankylosis, if it flails, or if it becomes damaged to the point that surgery is required. Radiological assessments are scheduled to be performed on patients at 2-year intervals. From the resulting films, a joint is classified as damaged according to “radiological assessment” if there is evidence of surface erosions of the bone in the joint, joint space narrowing, “disorganization” of the joint, or surgery is required. While these 2 types of assessments were scheduled at regular (but different) times, the actual times and frequency of functional and radiological assessments varied considerably from patient to patient, yielding panel count data.

Figure 1 contains timeline diagrams for a sample of 10 patients. The length of the horizontal lines indicates the duration of time from clinic entry to last contact for functional assessments (solid lines) and radiological assessments (dotted lines).
radiological assessments (dashed lines). The vertical dashes indicate the respective assessment times, and the corresponding numbers indicate the number of newly damaged joints detected since the last assessment by the corresponding method of assessment. The figure reveals that the 2 methods of assessment are sometimes coincident, but often not, and there can be large gaps during which no assessments are available (see individuals A, B, and E). To give an idea about the relationship between the 2 types of events, Figure 2 contains a scatter plot of the crude event rates defined as the number of damaged joints detected from clinic entry to the last assessment divided by the time since clinic entry to the last assessment for all patients (the dashed line has slope 1). There is a slight tendency for the crude rates of radiological damage to be higher than those for functional damage, but it should be noted that there is considerable sampling variability in these rates.

The data on damaged joint counts as defined by functional and radiological criteria form bivariate panel count data; our interest lies in estimating covariate effects on the respective rate functions and estimating the cumulative mean number of damaged joints according to each criterion. Of course, one way for this is to apply the methods developed for univariate panel count data to each type of damaged joint counts separately. On the other hand, Figure 2 indicates that the 2 types of damaged joints are correlated and thus a joint analysis would be preferred as discussed earlier.

3. MODELS AND NOTATIONS

Consider a recurrent event study that involves \( n \) independent subjects and suppose that each subject may experience \( K \) different types of events. For subject \( i \), let \( N_{ik}(t) \) denote the total number of type \( k \) events that have occurred up to time \( t \), \( 0 \leq t \leq \tau \), where \( \tau \) is a known constant representing study length, \( i = 1, \ldots, n \), \( k = 1, \ldots, K \). For the example discussed in Section 2, we have \( K = 2 \) and \( N_{i1}(t) \) and \( N_{i2}(t) \) denote the cumulative counts of damaged joints by radiological and functional assessments, respectively, up to time \( t \) for patient \( i \). Also for each \( i \), suppose that there exists a positive random variable \( C_i \) representing the censoring or follow-up time on subject \( i \) and a \( p \times 1 \) vector of covariates denoted by \( x_i = (x_{i1}, \ldots, x_{ip})' \) that may affect the rate of occurrence of type \( k \) events. Here, for the simplicity of presentation, we assume that the follow-up time or observation period and the covariates that may affect \( N_{ik}(t) \) are the same for different types of recurrent events. The inference approach presented below can
be easily generalized to situations where $C_i$ and $x_i$ may differ for different types of recurrent events. Define $Y_i(t) = I(t \leq C_i)$, indicating if subject $i$ is at risk of experiencing the recurrent events at time $t$, $i = 1, \ldots, n, k = 1, \ldots, K$.

For the effects of covariates on $N_{ik}(t)$, we assume that given $x_i$, the marginal mean function of $N_{ik}(t)$ has the form

$$E[N_{ik}(t)|x_i] = \mu_k(t)g_N(x_i'\beta_0).$$

(3.1)

In the model above, $\mu_k(t)$ is an unknown continuous baseline mean function, $\beta_0$ is a $p \times 1$ vector of regression parameters representing the effect of $x_i$ on $N_{ik}(t)$, and $g_N(\cdot)$ is a known, positive function that is assumed to be strictly increasing and twice differentiable. For the psoriatic arthritis data described above, model (3.1) implies that covariates affect the cumulative damaged joint counts multiplicatively in the scale of $g_N$. One common choice for $g_N(\cdot)$ is $g_N(x) = \exp(x)$, the exponential function. Other functions that are often used include $g_N(x) = 1 + x$ and $g_N(x) = \log(1 + e^x)$. Models similar to (3.1) have been used by Cai and Schaubel (2004b) and Sun and Wei (2000) among others. Model (3.1) assumes that the baseline mean functions can be different for different types of recurrent events, but the effects of covariates on different types of recurrent events are common. Some comments are given below for the situation where these effects may be different. The goal here is to estimate regression parameters $\beta_0$.

For estimation of $\beta_0$, we assume that only panel count data are available for the $N_{ik}(t)$. Specifically, suppose that $N_{ik}(\cdot)$ is observed only at finite time points $T_{ik,1} < \cdots < T_{ik,m_{ik}}$, where $m_{ik}$ denotes the potential or scheduled number of observations on the $k$th type of recurrent event for subject $i$, $i = 1, \ldots, n, k = 1, \ldots, K$. That is, the observed data have the form

$$\{T_{ik,\ell}, N_{ik}(T_{ik,\ell}), C_i, x_i, m_{ik}; \ell = 1, \ldots, m_{ik}, k = 1, \ldots, K, i = 1, \ldots, n\}.$$

For each $i$ and $k$, define $\tilde{N}_{ik}(t) = H_{ik}\{\min(t, C_i)\}$, where $H_{ik}(t) = \sum_{\ell=1}^{m_{ik}} I(T_{ik,\ell} \leq t)$. Then, $\tilde{N}_{ik}(t)$ is a point process characterizing the observation process on subject $i$ with respect to the $k$th-type recurrent event and jumps by one only at the observation times on $N_{ik}$. In the following, we assume that $H_{ik}$ is a counting process with the marginal mean function

$$E\{H_{ik}(t)|x_i\} = v_k(t)g_H(x_i'\gamma_0)$$

(3.2)

given $x_i$. Here, as with model (3.1), $v_k(t)$ is a completely unknown continuous baseline mean function, $\gamma_0$ denotes the effect of covariates on $H_{ik}$, and $g_H$ is a known, positive function that is assumed to be strictly increasing and twice differentiable.

In Section 4, some estimating equations are developed for estimation of $\beta_0$ along with $\gamma_0$. Note that for both $N_{ik}$, the process of interest, and $H_{ik}$, the observation process, only marginal mean functions are specified and no assumption is made about the relationship among $N_{i1}, \ldots, N_{iK}$ or $H_{i1}, \ldots, H_{iK}$. In the following, we assume that the $C_i$ follow the same distribution function.

### 4. Estimation Procedures

For estimation of $\beta_0$, we first consider situations in which covariates have no effect on the observation process, that is, $\gamma_0 = 0$ or $g_H$ is constant. The estimation procedure is then generalized to situations where the observation process may depend on covariates. For both situations, the focus will be on developing estimating equations for the regression parameters that do not involve the baseline mean functions $\mu_k$ and $v_k$. 
4.1 Estimation with independent observation processes

For simplicity, here we assume that $g_H(x) = 1$ and, hence, that the observation process is independent of the covariates. For each $i$ and $k$, we define

$$
\tilde{N}_{ik} = \sum_{\ell=1}^{m_{ik}} N_{ik}(T_{ik,\ell}) I(T_{ik,\ell} \leq C_i) = \int_0^r N_{ik}(t) d\tilde{N}_{ik}(t),
$$

$i = 1, \ldots, n, k = 1, \ldots, K$. Then, conditional on $x_i$ and under models (3.1) and (3.2), we have

$$
E\{\tilde{N}_{ik}|x_i\} = \alpha_k g_N(x_i'\beta_0),
$$

where $\alpha_k = \int_0^r \mu_k(t) P(C_i \geq t) dv_k(t)$. Without loss of generality, suppose that the covariates $x_i$ are centered. Then, by following Sun and Wei (2000), a natural, unbiased estimating function for $\beta_0$ is given by

$$
U_n(\beta) = \frac{1}{\sqrt{n}} \sum_{k=1}^K \sum_{i=1}^n x_i x_i' \tilde{N}_{ik} g_N^{(1)}(x_i'\beta)\{g_N(x_i'\beta)\}^{-2},
$$

which is the same as that given in Sun and Wei (2000) if $K = 1$ and $g_N(t) = \exp(t)$.

Define the estimate $\hat{\beta}_1$ of $\beta_0$ as the solution to $U_n(\beta) = 0$. Let $g_N^{(0)} = g_N$ and $g_N^{(r)}$ denote the $r$th derivative of $g_N$, $r = 1, 2$, so that

$$
\frac{\partial U_n(\beta)}{\partial \beta} = -\frac{1}{\sqrt{n}} \sum_{k=1}^K \sum_{i=1}^n x_i x_i' \tilde{N}_{ik} g_N^{(1)}(x_i'\beta)\{g_N(x_i'\beta)\}^{-2},
$$

which is strictly negative definite since $g_N$ is strictly increasing. Thus, the equation $U_n(\beta) = 0$ has a unique solution and $\hat{\beta}_1$ is consistent.

To derive the asymptotic distribution of $\hat{\beta}_1$, we define $U^*_i(\beta) = \sum_{k=1}^K x_i \tilde{N}_{ik} g_N^{(1)}(x_i'\beta)\{g_N(x_i'\beta)\}^{-2}$. Then, $U_n(\beta) = n^{-1/2} \sum_{i=1}^n U^*_i(\beta)$, which is the summation of i.i.d random variables. Thus, it can be easily shown that $\sqrt{n}(\hat{\beta}_1 - \beta_0)$ converges in distribution to a multivariate normal vector with mean zero and covariance matrix $\Sigma_1 = \phi^{-1}\Sigma_u\phi^{-1}$, where $\phi = -E\{\partial U_n^*(\beta_0)/\partial \beta\}$ and $\Sigma_u = \text{cov}(U^*_i(\beta_0))$. A consistent estimate of $\Sigma_1$ is given by $\hat{\Sigma}_1 = \hat{\phi}^{-1}\hat{\Sigma}_u\hat{\phi}^{-1}$ (Wei and others, 1989), where

$$
\hat{\phi} = \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^K x_i x_i' \tilde{N}_{ik} g_N^{(1)}(x_i'\hat{\beta}_1)\{g_N(x_i'\hat{\beta}_1)\}^{-2}
$$

and

$$
\hat{\Sigma}_u = \frac{1}{n} \sum_{i=1}^n \left[ \sum_{k=1}^K x_i \tilde{N}_{ik} g_N(x_i'\hat{\beta}_1)\{g_N(x_i'\hat{\beta}_1)\}^{-1} \right] \otimes a
$$

with $a \otimes a = aa'$ for a vector $a$.

4.2 Estimation with dependent observation processes

Consider now the more general situation in which the observation process $H_{ik}(t)$ is affected by the covariates as $N_{ik}(t)$. If $\bar{N}_{ik}$ is defined as before, under models (3.1) and (3.2), we have

$$
E\{\bar{N}_{ik}|x_i\} = \alpha_k g_N(x_i'\beta_0)g_H(x_i'\gamma_0)
$$
given \( x_i \), where \( \alpha_k \) is as defined earlier. If \( \gamma_0 \) is known, this along with \( U_n(\beta) \) suggests to estimate \( \beta_0 \) based on the estimating equation \( U_n(\beta, \gamma_0) = 0 \), where

\[
U_n(\beta, \gamma) = \frac{1}{\sqrt{n}} \sum_{k=1}^{K} \sum_{i=1}^{n} x_i \tilde{N}_{ik} (g_N(x'_i \beta))^{-1} (g_H(x'_i \gamma))^{-1}.
\]

(4.2)

In practice, of course, the parameter \( \gamma_0 \) is unknown. Unlike for \( \tilde{N}_{ik}(t) \), however, here we have complete recurrent event data (Cai and Schaubel, 2004a) for the observation process \( \{\tilde{N}_{ik}(s), 0 < s < C_i\} \), making estimation of \( \gamma_0 \) relatively easy. To see this, define

\[
S^{(d)}_k(t; \gamma) = \frac{1}{n} \sum_{i=1}^{n} Y_i(t) x_i^\otimes dx^{(d)}_H(x'_i \gamma),
\]

for \( d = 0, 1, 2 \), and

\[
S^{(3)}_k(t; \gamma) = \frac{1}{n} \sum_{i=1}^{n} Y_i(t) x_i^\otimes dx^{(1)}_H(x'_i \gamma)^2 (g_H(x'_i \gamma))^{-1},
\]

\( k = 1, \ldots, K \). Also for \( k = 1, \ldots, K \), define

\[
E_k(t; \gamma) = \frac{S^{(1)}_k(t; \gamma)}{S^{(0)}_k(t; \gamma)}
\]

and

\[
V_k(t; \gamma) = \frac{S^{(3)}_k(t; \gamma)}{S^{(0)}_k(t; \gamma)} - E_k(t; \gamma)^\otimes 2,
\]

and suppose that the limits of \( S^{(d)}_k(t; \gamma) \), \( E_k(t; \gamma) \), and \( V_k(s; \gamma) \) exist.

Following Cai and Schaubel (2004b), we propose to estimate \( \gamma_0 \) by \( \hat{\gamma} \), the solution to the estimating equation \( H_n(\gamma) = 0 \), where

\[
H_n(\gamma) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{k=1}^{K} \int_{0}^{\tau} \left\{ x_i g^{(1)}_H(x'_i \gamma) - g^{(1)}_H(x'_i \gamma) \right\} d\tilde{N}_{ik}(s).
\]

(4.3)

Note that estimation of \( \gamma_0 \) depends only on the observed information on the observation process \( \tilde{N}_{ik} \). It is then natural to estimate \( \beta_0 \) by \( \hat{\beta}_2 \) defined as the solution to \( U_n(\beta, \hat{\gamma}) = 0 \).

For the asymptotic properties of \( \hat{\beta}_2 \) and \( \hat{\gamma} \), Cai and Schaubel (2004b) proved that under mild regularity conditions, \( \hat{\gamma} \) is unique and consistent and asymptotically follows a normal distribution. For \( \hat{\beta}_2 \), it can be easily shown that it is unique and consistent as \( \hat{\beta}_1 \). For the asymptotic distribution of \( \hat{\beta}_2 \), we show in the Appendix that under some regularity conditions described there, \( \sqrt{n}(\hat{\beta}_2 - \beta_0) \) asymptotically follows a multivariate normal distribution with mean zero and variance matrix that can be consistently estimated by

\[
\hat{\Sigma}_2 = \hat{F}^{-1} \hat{G}_1 \hat{G}_1' \hat{F}^{-1}.
\]

(4.4)
In this formula, \( \hat{G}_1 = (I_p, -\hat{D} \hat{A}^{-1}(\hat{\gamma})) \),

\[
\hat{F} = -\frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{K} (g_N(x'_i \hat{\beta}_2))^{-2} g_N^{(1)}(x'_i \hat{\beta}_2)(g_H(x'_i \hat{\gamma}))^{-1} \tilde{N}_{ik} x_i x'_i,
\]

\[
\hat{D} = -\frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{K} (g_N(x'_i \hat{\beta}_2))^{-1} g_N^{(1)}(x'_i \hat{\gamma})(g_H(x'_i \hat{\gamma}))^{-2} \tilde{N}_{ik} x_i x'_i,
\]

\[
\hat{A}(\gamma) = -\frac{1}{n} \sum_{k=1}^{K} \int_0^t V_k(t; \gamma) d \tilde{N}_k(t),
\]

and \( \hat{\Gamma} \) is given in the Appendix, where \( I_p \) denotes the \( p \times p \) identity matrix and \( \tilde{N}_k(t) = \sum_{i=1}^{n} \tilde{N}_{ik}(t) \).

One may sometimes be interested in the joint distribution of \( \hat{\beta}_2 \) and \( \hat{\gamma} \) and estimating the baseline mean functions \( \mu_k(t) \) and \( \nu_k(t) \), \( k = 1, \ldots, K \). In the Appendix, we also show that one can asymptotically approximate the joint distribution of \( \sqrt{n}(\hat{\beta}_2 - \beta_0) \) and \( \sqrt{n}(\hat{\gamma} - \gamma_0) \) by a multivariate normal distribution with mean zero and covariance matrix

\[
-\hat{F}^{-1} \hat{G}_0 \hat{G}_0^T,
\]

where \( \hat{G}_0 = (0_p, -\hat{A}^{-1}(\hat{\gamma})) \) with \( 0_p \) denoting the \( p \times p \) zero matrix.

Given \( \hat{\beta}_2 \) and \( \hat{\gamma} \), one may want to estimate the baseline mean functions \( \mu_k(t) \) and \( \nu_k(t) \), \( k = 1, \ldots, K \). The estimation of \( \nu_k(t) \) is relatively easy since one has recurrent event data and, in particular, a consistent estimate of it is given by

\[
\hat{\nu}_k(t; \hat{\gamma}) = \int_0^t \frac{d \tilde{N}_k(s)}{n S_k^{(0)}(s; \hat{\gamma})}
\]

(Andersen and others, 1993), \( k = 1, \ldots, K \). For \( \mu_k(t) \), first consider the estimation with respect to each subject. Given \( x_i \) and under model (3.1), a natural estimate of the rate function \( d\mu_k(t) \) is the empirical estimate

\[
d\hat{\mu}_{ik}(t; \hat{\beta}_2) = \sum_{\ell=1}^{m_{ik}} \frac{N_{ik}(t_{ik,\ell}) - N_{ik}(t_{ik,\ell-1})}{t_{ik,\ell} - t_{ik,\ell-1}} (g_N(x'_i \hat{\beta}_2))^{-1} I(t_{ik,\ell-1} < t \leq t_{ik,\ell}),
\]

where \( t_{ik,0} = 0, i = 1, \ldots, n, k = 1, \ldots, K \). Thall and Lachin (1988) considered the same estimate. This leads to an estimate of \( \mu_k(t) \) given by

\[
\hat{\mu}_k(t; \hat{\beta}_2) = \int_0^t d\hat{\mu}_k(s) = \int_0^t \frac{\sum_{i=1}^{n} d\hat{\mu}_{ik}(s)}{\sum_{i=1}^{n} I(s \leq t_{ik,m_{ik}})}
\]

for \( 0 \leq t \leq \max \{ t_{ik,m_{ik}} \}, k = 1, \ldots, K \). Note that a similar, natural estimate of \( \mu_k(t) \) is given by

\[
\sum_{i=1}^{n} \sum_{\ell=1}^{m_{ik}} \frac{N_{ik}(t_{ik,\ell}) (g_N(x'_i \hat{\beta}_2))^{-1} I(t_{ik,\ell-1} < t \leq t_{ik,\ell})}{\sum_{i=1}^{n} I(t \leq t_{ik,m_{ik}})}
\]

or

\[
\sum_{i=1}^{n} \sum_{\ell=1}^{m_{ik}} \frac{N_{ik}(t_{ik,\ell-1}) (g_N(x'_i \hat{\beta}_2))^{-1} I(t_{ik,\ell-1} < t \leq t_{ik,\ell})}{\sum_{i=1}^{n} I(t \leq t_{ik,m_{ik}})},
\]

the weighted sample mean estimates. However, unless the \( m_{ik} \) are large, the former can seriously overestimate \( \mu_k(t) \) while the latter can underestimate \( \mu_k(t) \).
5. Simulation Studies

This section reports some results obtained from simulation studies conducted to assess the performance of the estimation procedures proposed in Section 4 under various situations. In the simulation studies, we focused on the 2-sample comparison problem with \( x_i = -1 \) for half subjects and 1 for the others and first generated the follow-up times \( C_i \) from the uniform distribution over \((\tau/3, \tau)\), where \( \tau \) is a positive constant. For the observation times \( T_{ik,\ell} \), it was assumed that \( H_i(t) \) follows the mixed-effects marginal mean model

\[
E\{H_i(t)|x_i, P_i\} = P_i g_H(x_i; \gamma_0) v_k(t)
\]
given \( x_i \) and random effect \( P_i, i = 1, \ldots, n, k = 1, 2 \). In the model above, the \( P_i \) were generated from the gamma distribution with mean 1 and variance \( \sigma^2_P \), \( v_1(t) = 0.5t \), \( v_2(t) = t \), and \( g_H(x) \) was set to be the identity function for the case in which the observation time is independent of covariates and \( g_H(x) = e^x \) if the observation time is dependent of covariates. Then, by following Cai and Schaubel (2004b), the observation times \((T_{ik,1}, \ldots, T_{ik,m_{ik}})\) were defined as

\[
T_{ik,\ell+1} = T_{ik,\ell} - \log(U_{ik,\ell+1})[P_i g_H(x_i'; \gamma_0)dv_k]^{-1},
\]
where \( T_{ik,0} = 0, U_{ik,\ell} \sim U(0, 1) \), and \( m_{ik} \) is the number of the \( T_{ik,\ell} \) that fall within \((0, C_i)\).

Given \( m_{ik} \) and \((T_{ik,1}, \ldots, T_{ik,m_{ik}})\), we generated the panel count data \( N_{ik}(T_{ik,\ell}) \) from the mixed Poisson processes as

\[
N_{ik}(T_{ik,\ell}) = N^*_{ik}[\mu_k(T_{ik,1})] + N^*_{ik}[\mu_k(T_{ik,2}) - \mu_k(T_{ik,1})] + \cdots + N^*_{ik}[\mu_k(T_{ik,\ell}) - \mu_k(T_{ik,\ell-1})],
\]
\( \ell = 1, \ldots, m_{ik}, k = 1, 2, i = 1, \ldots, n \). In the above, \( N^*_{ik}[\mu_k(t)] \) denotes the random number generated from the Poisson distribution with mean \( Q_i \mu_k(t) \exp(x_i; \beta_0) \), where \( Q_i \) is the random number arising from the gamma distribution with mean 1 and variance \( \sigma^2_Q \), \( \mu_1(t) = t \), and \( \mu_2(t) = t^2 \). The results given below are based on \( n = 100 \), or \( 200 \), \( \tau = 12 \), \( \sigma_P = 1 \), \( \gamma_0 = 1 \), and \( 1000 \) replications.

Table 1 presents the simulation results obtained for independent observation process situations with \( \beta_0 = -1, 0, \) or 1 and \( \sigma^2_Q = 1 \) or 2. The table includes the estimated bias (BIAS) given by the mean of the point estimates \( \hat{\beta}_1 \) minus the true value of \( \beta_0 \), the averages of the standard deviation estimates (SEE), the sample standard deviations of \( \hat{\beta}_1 \) (SSE), and the empirical 95% coverage probability (CP) for \( \beta_0 \).

<table>
<thead>
<tr>
<th>Sample size</th>
<th>( \sigma^2_Q )</th>
<th>( \beta_0 )</th>
<th>BIAS</th>
<th>SEE</th>
<th>SSE</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n = 100 )</td>
<td>1</td>
<td>-1</td>
<td>-0.0025</td>
<td>0.2276</td>
<td>0.2330</td>
<td>0.945</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>0.0045</td>
<td>0.2273</td>
<td>0.2331</td>
<td>0.946</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>0.0020</td>
<td>0.2271</td>
<td>0.2378</td>
<td>0.935</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-1</td>
<td>0.0058</td>
<td>0.2600</td>
<td>0.2796</td>
<td>0.930</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>0.0133</td>
<td>0.2618</td>
<td>0.2849</td>
<td>0.923</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>-0.0042</td>
<td>0.2619</td>
<td>0.2853</td>
<td>0.927</td>
</tr>
<tr>
<td>( n = 200 )</td>
<td>1</td>
<td>-1</td>
<td>-0.0095</td>
<td>0.1683</td>
<td>0.1616</td>
<td>0.957</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>0.0009</td>
<td>0.1698</td>
<td>0.1621</td>
<td>0.952</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-1</td>
<td>-0.0054</td>
<td>0.1993</td>
<td>0.2012</td>
<td>0.939</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>0.0045</td>
<td>0.1963</td>
<td>0.1996</td>
<td>0.941</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>-0.0032</td>
<td>0.1995</td>
<td>0.2084</td>
<td>0.947</td>
</tr>
</tbody>
</table>
Table 2. Simulation results for dependent observation processes

<table>
<thead>
<tr>
<th>Sample size</th>
<th>$\sigma_Q^2$</th>
<th>$\beta_0$</th>
<th>BIAS</th>
<th>SEE</th>
<th>SSE</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 100$</td>
<td>1</td>
<td>-1</td>
<td>-0.0056</td>
<td>0.2042</td>
<td>0.2059</td>
<td>0.949</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.0072</td>
<td>0.2045</td>
<td>0.2015</td>
<td>0.952</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.0111</td>
<td>0.2045</td>
<td>0.2111</td>
<td>0.940</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-1</td>
<td>0.0097</td>
<td>0.2409</td>
<td>0.2628</td>
<td>0.933</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.0103</td>
<td>0.2408</td>
<td>0.2630</td>
<td>0.922</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-0.0033</td>
<td>0.2413</td>
<td>0.2669</td>
<td>0.921</td>
<td></td>
</tr>
<tr>
<td>$n = 200$</td>
<td>1</td>
<td>-1</td>
<td>-0.0004</td>
<td>0.1518</td>
<td>0.1467</td>
<td>0.959</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.0029</td>
<td>0.1520</td>
<td>0.1441</td>
<td>0.955</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.0116</td>
<td>0.1538</td>
<td>0.1449</td>
<td>0.956</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-1</td>
<td>0.0149</td>
<td>0.1825</td>
<td>0.1889</td>
<td>0.940</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.0093</td>
<td>0.1827</td>
<td>0.1878</td>
<td>0.928</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.0120</td>
<td>0.1824</td>
<td>0.1880</td>
<td>0.951</td>
<td></td>
</tr>
</tbody>
</table>

can be seen that the estimated covariate effect seems unbiased, and the 2 standard deviation estimates seem quite close to each other, suggesting that the proposed variance estimate is reasonable. The table also shows that as expected, the results become better when the sample size increases. Note that $\sigma_Q^2$ measures the correlation between the 2 recurrent processes $N_{i1}$ and $N_{i2}$. The simulation results indicate that as expected, the covariate effect can be estimated more efficiently for smaller $\sigma_Q^2$, meaning that the 2 recurrent processes are less related.

The simulation results for dependent observation processes or $\hat{\beta}_0$ are given in Table 2, which presents the same quantities as in Table 1. Note that here we only give the results about $\beta_0$ because as shown in Cai and Schaubel (2004b), the estimation procedure for $\gamma_0$ seems to work as well as the estimation procedure for $\beta_0$. Table 2 basically gives the same results as in Table 1 and suggests that as that given in Section 4.1, the estimation procedure proposed in Section 4.2 also seems to work reasonably well for the situations considered here. To evaluate the normal approximation to the finite-sample distribution of $\hat{\beta}_1$ or $\hat{\beta}_2$, we studied the quantile plots of the standardized estimates against the standard normal distribution. These plots, which are not given here, indicate that the normal approximation seems good.

6. ANALYSIS OF THE UNIVERSITY OF TORONTO PSORIATIC ARTHRITIS DATA

We now consider the analysis of the data from the University of Toronto Psoriatic Arthritis Clinic described in Section 2. In the study, 3 covariates of primary interest are the presence of a family history of psoriasis (yes/no), arthritis duration (years), and the number of active (defined as tender or swollen) joints at clinic entry. We restrict our attention to 177 female patients having a baseline, at least one follow-up assessment, and complete covariate data.

For the analysis, define $N_{i1}(t)$ and $N_{i2}(t)$ as the cumulative numbers of radiologically and functionally damaged joints up to time $t$ for patient $i$, respectively, $i = 1, \ldots, 177$. Also for the $i$th patient, define $x_{i1} = 1$ if he or she had a family history of psoriasis and 0 otherwise, and $x_{i2}$ and $x_{i3}$ to be equal to the arthritis duration and the number of active joints at clinic entry, respectively. Table 3 gives the joint analysis results obtained by the application of the method proposed in Section 4.2 and includes the point estimates $\hat{\gamma}$ and $\hat{\beta}_2$, their estimated standard errors, and the $p$-values for testing these parameters equal to zero. The first 3 columns of numbers pertain to the models for the observation process and the last 3 to the joint damage processes $N_{i1}(t)$ and $N_{i2}(t)$. For comparison, we also performed univariate analyses...
Table 3. Results of joint and univariate analyses of radiological and functional joint damage data from the University of Toronto Psoriatic Arthritis Clinic

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Observation process</th>
<th>Joint damage process</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\dot{\gamma}$</td>
<td>SE($\dot{\gamma}$)</td>
</tr>
<tr>
<td>Multivariate analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family history of psoriasis</td>
<td>0.1689</td>
<td>0.1165</td>
</tr>
<tr>
<td>Duration of PsA in years</td>
<td>−0.0015</td>
<td>0.0057</td>
</tr>
<tr>
<td>Number of active joints</td>
<td>0.0030</td>
<td>0.0060</td>
</tr>
<tr>
<td>Univariate analysis—radiological damage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family history of psoriasis</td>
<td>0.1375</td>
<td>0.1403</td>
</tr>
<tr>
<td>Duration of PsA in years</td>
<td>−0.0043</td>
<td>0.0063</td>
</tr>
<tr>
<td>Number of active joints</td>
<td>−0.0079</td>
<td>0.0075</td>
</tr>
<tr>
<td>Univariate analysis—functional damage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family history of psoriasis</td>
<td>0.1609</td>
<td>0.1200</td>
</tr>
<tr>
<td>Duration of PsA in years</td>
<td>−0.0007</td>
<td>0.0058</td>
</tr>
<tr>
<td>Number of active joints</td>
<td>0.0048</td>
<td>0.0059</td>
</tr>
</tbody>
</table>

PsA, psoriatic arthritis; SE, standard error.

and included the results in Table 3. The univariate results involve separate modeling of covariate effects on the rate functions for radiologically damaged joint counts and functionally damaged joint counts.

The results based on the joint analysis suggest that all 3 covariates had significant effects on the rates of both radiological and functional damage in joints. Specifically, the patients with a family history of psoriasis seem to have lower rates of damage, suggesting that the skin component of the disease is more active in these patients. Both a longer history of psoriatic arthritis and the higher number of active joints at clinic entry imply an increased damage rate. The results also indicate that all 3 covariates seem to have no effects on the observation process.

The univariate analyses gave similar conclusions for most of the covariates effects except that they significantly underestimated the effect of arthritis duration on the rate of radiologically damaged joints. It can be seen that the estimated effects of all covariates based on the joint analysis are intermediate between those based on the 2 univariate analyses. Also the estimated covariate effects from the 2 univariate analyses are reasonably close, which suggests that the assumption of the same covariate effects on the 2 types of joint damage seems reasonable.

Figure 3 presents the estimates given in (4.6) of the baseline cumulative mean functions for the numbers of joints damaged according to the radiological and functional criteria. The figure includes the estimates obtained based on both joint and univariate analyses. The close agreement between the mean functions from the univariate and bivariate models again reflects the plausibility of the assumption of common regression coefficients. It is clear that one can expect a greater number of joints to be classified as damaged by the radiological criteria than by the functional criteria. This is consistent with the notion that damage tends to be detected earlier by radiological assessment compared to functional assessment (Sianni and others, 2006). Also the joint analysis suggested a larger difference between the 2 criteria than the univariate analysis.

7. CONCLUDING REMARKS AND DISCUSSION

Multivariate panel count data often arise in periodic follow-up studies that concern recurrent events. In the preceding sections, some marginal mean models were presented for regression analysis and estimating
equation approaches were proposed for inference about regression parameters. One main advantage of the proposed methodology is that it leaves the relationship of different types of recurrent events completely unspecified. Also compared to the parametric approach proposed in Chen and others (2005), it does not rely on the Poisson process and piecewise constant assumptions and can be relatively easily implemented. The application of the proposed methodology to the psoriatic arthritis data suggests that all 3 covariates, a family history of psoriasis, arthritis duration, and the number of active joints at clinic entry, had significant effects on the rate of the damaged joints.

When facing the analysis of multi-type data, it is often natural to assess whether joint or multivariate analyses are warranted over separate univariate analyses. This is particularly true for models which do not provide parametric estimates of association parameters. Joint methods are helpful when interest lies in the global assessment of covariate effects across 2 or more processes. More discussion on this aspect can be found in Wei and others (1989) for multivariate failure time data and Chen and others (2005) for the current setting. Multivariate methods are also helpful when interest lies in obtaining common estimates of covariate effects. In general, separate univariate analyses will give different estimates, but when it appears reasonable and the estimates are comparable, constraining the regression coefficients to be common can simplify discussion. Also under the assumption of common effects, as discussed before, the use of joint analysis will often produce more efficient estimates of the effects. That is, the common effects can be estimated more precisely based on the joint analysis than based on either of the separate analysis. In contrast, the separate analysis cannot give a direct estimation of the common effects.

The method presented in the previous sections can be generalized in several directions. One is that models (3.1) and (3.2) assume that the covariates that affect different types of recurrent events are same. However, in practice, there may exist type-specific covariates for each particular type of recurrent events and for this, methods similar to those given above can be developed. Note that models (3.1) and (3.2) also assume that the covariate effects on different types of recurrent events are identical. Sometimes this may not be true. In this case, one can redefine a larger, type-specific covariate vector that corresponds to a new and larger vector of covariate effects that includes all different covariate effects but is the same for different types of recurrent events. Another generalization would be to incorporate weight functions or
processes into the estimating functions defined in (4.1), (4.2), and (4.3); this could result in more efficient estimates of regression parameters. One may also generalize the proposed method to situations where the distribution of the follow-up times may depend on covariates. In this case, one may need to specify a model for the covariate effects on the \( C_i \) and develop some joint estimation procedures for all regression parameters together.

**Funding**

National Science Foundation (10671106) to X.T.

**Acknowledgments**

We wish to thank the editors, an associate editor, and a referee for their helpful comments and suggestions that greatly improved the paper. **Conflict of Interest:** None declared.

**Appendix**

*Proof of the asymptotic normality of \( \hat{\beta}_2 \) and \( \hat{\gamma} \)

Let \( N_{ik}(t), \tilde{N}_{ik}(t), Y_i(t), S_k^{(d)}(t; \gamma), E_k(t; \gamma), \) and \( V_k(s; \gamma) \) be defined as in the previous sections and \( s_k^{(d)}(t; \gamma), e_k(t; \gamma), \) and \( v_k(s; \gamma) \) denote the limit processes of \( S_k^{(d)}(t; \gamma), E_k(t; \gamma), \) and \( V_k(s; \gamma), \) respectively. Define

\[
A(\gamma) = \sum_{k=1}^{K} \int_0^\tau v_k(t; \gamma)s_k^{(0)}(t; \gamma)dv_k(t)
\]

and

\[
F(\beta, \gamma) = E \left[ \sum_{k=1}^{K} \left\{g_N(x_i^\prime \beta)\right\}^{-2} g_N^{(1)}(x_i^\prime \beta) g_H(x_i^\prime \gamma) \right]^{-1} \tilde{N}_{ik} x_i x_i^\prime
\]

For the asymptotic normality of \( \hat{\beta}_2 \) and \( \hat{\gamma}, \) we assume that the regularity conditions given in Cai and Schaubel (2004b) hold. Also, we need the following conditions:

a) \( \{N_{ik}(\cdot), Y_i(\cdot), x_i, H_{ik}(\cdot)\}_{k=1}^{K} \) are i.i.d. for \( i = 1, \ldots, n. \)

b) \( P(Y_i(\tau) = 1) > 0. \)

c) \( |x_{il}| < c_1 \) almost surely for all \( i \) and \( l, \) where \( c_1 \) is a positive constant and \( x_{il} \) denotes the \( l \)th component of \( x_i. \)

d) Both matrices \( A(\gamma) \) and \( F(\beta, \gamma) \) are positive definite.

e) \( N_{ik}(\tau) < c_2 \) and \( H_{ik}(\tau) < c_2 \) almost surely for all \( i = 1, \ldots, n \) and \( k = 1, \ldots, K, \) where \( c_2 \) is a positive constant.

f) There exist some neighborhoods of \( \beta_0 \) and \( \gamma_0, \) respectively, within which \( g_N(x_i^\prime \beta) \geq c_3 \) and \( g_H(x_i^\prime \gamma) \geq c_3, \) and \( s_k^{(d)}(t; \gamma) \) is uniformly continuous with respect to \( \gamma \) and \( t \in [0, \tau], \) where again \( c_3 \) is a positive constant, \( d = 0, 1, 2. \)

For each \( i \) and \( k \) and given \( \gamma, \) define

\[
d\tilde{M}_{ik}(t; \gamma) = d\tilde{N}_{ik}(t) - Y_{ik}(t)g_H(x_i^\prime \gamma)d\hat{\nu}_k(t; \gamma),
\]
which is a zero-mean process under model \((3.2)\). Also define

\[
\hat{\Sigma}_U = \frac{1}{n} \sum_{i=1}^{n} \left[ \sum_{k=1}^{K} x_i \tilde{N}_i k \{g_N(x'_i \hat{\beta}_2)\}^{-1} \{g_H(x'_i \hat{\gamma})\}^{-1} \right] ^{\otimes 2},
\]  

(A.1)

\[
\hat{\Sigma}_H = \frac{1}{n} \sum_{i=1}^{n} \left[ \sum_{k=1}^{K} \int_{0}^{\tau} \left\{ x_i \frac{g_H^{(1)}(x'_i \hat{\gamma})}{g_H(x'_i \hat{\gamma})} - E_k(t; \hat{\gamma}) \right\} \, d\hat{M}_{ik}(t; \hat{\gamma}) \right] ^{\otimes 2},
\]  

(A.2)

and

\[
\hat{\Sigma}_\gamma = A^{-1}(\hat{\gamma}) \hat{\Sigma}_H A^{-1}(\hat{\gamma}).
\]

The asymptotic normality of \(\hat{\gamma}\) is given in Theorem 1 in Cai and Schaubel (2004b). In particular, they showed that \(\sqrt{n}(\hat{\gamma} - \gamma_0)\) converges in distribution to a multivariate normal vector with mean zero and covariance matrix that can be consistently estimated by \(\hat{\Sigma}_\gamma\). For the asymptotic normality of \(\hat{\beta}_2\), using the Taylor series expansions of \(U_n(\hat{\beta}_2, \hat{\gamma})\) and \(H_n(\hat{\gamma})\) around \(\beta_0\) and \(\gamma_0\) and based on the regularity conditions given above, one can easily show that \(\sqrt{n}(\hat{\beta}_2 - \beta_0)\) has the same asymptotic distribution as

\[
-F^{-1}\{U_n(\beta_0, \gamma_0) - DA^{-1}(\gamma_0)H_n(\gamma_0)\},
\]

(A.3)

where

\[
F = F(\beta_0, \gamma_0) = E \left\{ \frac{\partial U_i^*(\beta, \gamma)}{\partial \beta} \right\}_{\beta = \beta_0},
\]

\[
D = D(\beta_0, \gamma_0) = E \left\{ \frac{\partial U_i^*(\beta_0, \gamma)}{\partial \gamma} \right\}_{\gamma = \gamma_0},
\]

and

\[
U_i^*(\beta, \gamma) = \sum_{k=1}^{K} x_i \tilde{N}_i k \{g_N(x'_i \beta)\}^{-1} \{g_H(x'_i \gamma)\}^{-1}.
\]

For \(U_n(\beta_0, \gamma_0)\) and \(H_n(\gamma_0)\), as in Sun and Wei (2000), it can be easily shown that they asymptotically have a joint multivariate normal distribution with mean zero and covariance matrix that can be approximated by

\[
\hat{\Gamma} = \left( \begin{array}{cc} \hat{\Sigma}_U & \hat{\Sigma}_{UH} \\ \hat{\Sigma}_{UH}' & \hat{\Sigma}_H \end{array} \right),
\]

where \(\hat{\Sigma}_U\) and \(\hat{\Sigma}_H\) are defined in (A.1) and (A.2), respectively, and

\[
\hat{\Sigma}_{UH} = \frac{1}{n} \sum_{i=1}^{n} \left[ \sum_{k=1}^{K} x_i \tilde{N}_i k \{g_N(x'_i \hat{\beta}_2)\}^{-1} \{g_H(x'_i \hat{\gamma})\}^{-1} \right] 
\]

\[
\times \left[ \sum_{k=1}^{K} \int_{0}^{\tau} \left\{ x_i \frac{g_H^{(1)}(x'_i \hat{\gamma})}{g_H(x'_i \hat{\gamma})} - E_k(t; \hat{\gamma}) \right\} \, d\hat{M}_{ik}(t; \hat{\gamma}) \right].
\]
Thus, it follows from (A.3) that $\sqrt{n}(\hat{\beta}_2 - \beta_0)$ has an asymptotic normal distribution with mean zero and covariance matrix that can be consistently estimated by the quantity given in (4.4).

The joint asymptotic normality of $\hat{\beta}_2$ and $\hat{\gamma}$ directly follows from (A.3) and the fact that $\sqrt{n}(\hat{\gamma} - \gamma_0)$ has the same asymptotic distribution as

$$G_0(U_n(\beta_0, \gamma_0)'', H_n(\gamma_0)'')'.$$

These also prove the consistency of the estimate given in (4.5) for their covariance matrix.

REFERENCES


[Received May 24, 2006; revised February 2, 2007; accepted for publication May 3, 2007]