HHS Public Access
Author manuscript
Biometrics. Author manuscript; available in PMC 2018 June 27.

Published in final edited form as:
Biometrics. 2018 June ; 74(2): 488–497. doi:10.1111/biom.12774.

A semiparametric likelihood-based method for regression analysis of mixed panel-count data

Liang Zhu[1,*], **Ying Zhang**[2], **Yimei Li**[3], **Jianguo Sun**[4], and **Leslie L. Robison**[5]

[1]Division of Clinical and Translational Sciences, Department of Internal Medicine, University of Texas Health Science Center at Houston, Houston, TX, 77030

[2]Department of Biostatistics, Indiana University, Indianapolis, IN, 46202; School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai, P. R. China, 200240

[3]Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, TN, 38105

[4]Department of Statistics, University of Missouri, Columbia, MO, 65211; School of Mathematics, Jilin University, Changchun, P. R. China, 130012

[5]Department of Epidemiology and Cancer Control, St. Jude Children's Research Hospital, Memphis, TN, 38105

Summary

Panel-count data arise when each study subject is observed only at discrete time points in a recurrent event study, and only the numbers of the event of interest between observation time points are recorded (Sun and Zhao, 2013). However, sometimes the exact number of events between some observation times is unknown and what we know is only whether the event of interest has occurred. In this paper, we will refer this type of data to as mixed panel-count data and propose a likelihood-based semiparametric regression method for their analysis by using the nonhomogeneous Poisson process assumption. However, we establish the asymptotic properties of the resulting estimator by employing the empirical process theory and without using the Poisson assumption. Also we conduct an extensive simulation study, which suggests that the proposed method works well in practice. Finally, the method is applied to a Childhood Cancer Survivor Study that motivated this study.

Keywords

Maximum likelihood method; Panel-binary data; Panel-count data; Semiparametric estimation efficiency; Semiparametric regression analysis

---

* liang.zhu@uth.tmc.edu.

Supplementary Materials
Web Appendices and the code referenced in Sections 2, 3 and 4 are available with this paper at the Biometrics website on Wiley Online Library.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

1. Introduction

Panel-count data arise when a recurrent event is under investigation and each study subject is observed only at discrete time points. In this situation, the observed data include only the numbers of the occurrences of the event of interest between observation time points, and the exact occurrence times of the event are unknown (Sun and Zhao, 2013). The fields that often see such data include medical follow-up studies, reliability studies, and social sciences (Thall and Lachin, 1988; Sun and Kalbfleisch, 1995; Wellner and Zhang, 2000; Sun and Wei, 2000; Sun and Fang, 2003; Wellner and Zhang 2007). In practice, however, sometimes we may face another situation, where instead of knowing the number of the events of interest between some observation times, one knows only whether or not the event of interest has occurred. In other words, we observe the number of occurrences of the event between some observation time points, while knowing only if the event has occurred between others. In the following, we will refer such data to as mixed panel-count data and it seems that there does not exist an established approach for regression analysis of such data, the focus of this paper.

This study was motivated by the Childhood Cancer Survivor Study (CCSS), a long-term follow-up study that followed more than 14,000 childhood cancer survivors who were diagnosed between 1970 and 1986 and had survived more than 5 years since diagnosis. It also follows a random sample of their siblings as a control group. One of the primary objectives of the CCSS is to determine the long-term effects, if any, of childhood cancer and its treatments on the pregnancy process or pregnancy outcomes. Questionnaires about pregnancy were distributed during 1995–1996, 2000–2001, 2002–2003, and 2006–2007. The questions asked included: 1. Have you and a partner ever become pregnant (since last follow-up)? 2. Including live births, stillbirths, miscarriages, and abortions, how many times have you become pregnant or had a woman become pregnant by you? The answer to the first question is yes/no, and the answer to the second question is a count. As expected, some participants answered all questions, but others answered only the first question or none of them. In other words, only mixed panel-count data are available for the pregnancy count or process. It is apparent whether a subject provided a yes/no or count answer here can be regarded to be independent of the questionnaire timings or observation times, which were fixed by the study design, as well as the potential counts.

For the analysis of mixed panel-count data, a naive approach is to ignore the binary part of the data and analyze only the available panel-count data. However, this may not be efficient, as a part of the information is ignored, and more seriously, the expected cumulative number of events at a given observation time would be underestimated. It is easy to see that the analysis based only on the binary part of the data would have the same issues. To address these issues, we propose a semiparametric maximum likelihood estimation (SPMLE) method that uses the proportional mean model (Cheng et al., 2011; Sun and Zhao, 2013; Wellner and Zhang, 2007). The proposed method will make use of all available information in the observed data; thus, it is expected to be efficient. Note that a similar type of data, also arising from panel-count data, is the mixed recurrent-event and panel-count data (Zhu et al., 2015). For this latter situation, there exist some observation time points between which, instead of the number of the events, the exact occurrence times of the event of interest are known.

Author Manuscript

The remainder of the article is organized as follows. Section 2 will introduce some notation and describe the structure of the observed data and the proportional mean model, along with the computing procedure for the proposed SPMLE method. Section 3 will establish the asymptotic properties of the resulting estimator under some mild regularity conditions. Some simulation results are given in Section 4, which indicate that the proposed method seems to work well in practical situations. Section 5 will apply the proposed method to the CCSS described above, and Section 6 contains some discussions and concluding remarks. Technical details are included in the Web Supplementary Materials.

2. Semiparametric Maximum Likelihood Estimation

2.1 Notation

Suppose that $\mathbb{N}(t) = \{N(t) : t \geq 0\}$ is a univariate counting process. In a standard panel-count study, we observe the counting process $\mathbb{N}(t)$ at a random number K of random times denoted by $0 \equiv T_{K,0} < T_{K,1} < \cdots < T_{K,K}$. We will assume that the observation process is conditionally independent of the counting process $\mathbb{N}(t)$ given the time-independent covariate vector Z.

In the following, we will assume that one only observes mixed panel-count data where for some observation periods, the number of the occurrences of the event during the period is recorded, while for other observation periods, we only know if there exist some occurrences of the event over the period. In other words, for the latter situation, the exact count is unknown and only a binary outcome (yes or no occurrence) is available. To describe this, define $r(s, t]$ as the observation type for any two time points $s < t$ such that $r(s, t] = 1$ if the count of the events during the observation period $(s, t]$, $\Delta N(s, t] = N(t) - N(s)$, is recorded and $r(s, t] = 0$ if only a binary outcome is recorded. That is, $r(s, t]$ indicates whether the actual event count over the observation period $(s, t]$ is known. Also define the binary outcome $D(s, t] = (\Delta N(s, t] > 0)$ and write $\underline{T}_K \equiv (T_{K,1}, \cdots, T_{K,K})$, $\underline{r}_K = (r_{K,1}, r_{K,2}, \cdots, r_{K,K})$, $\underline{N}_K = (\Delta N_{K,1}, \Delta N_{K,2}, \cdots, \Delta N_{K,K})$, and $\underline{D}_K = (D_{K,1}, D_{K,2}, \cdots, D_{K,K})$, where $r_{K,j} = r(T_{K,j-1}, T_{K,j}]$, $\Delta N_{K,j} = \Delta N(T_{K,j-1}, T_{K,j}]$, and $D_{K,j} = D(T_{K,j-1}, T_{K,j}]$ for $j = 1, 2, \cdots, K$. As a result, the observation for each subject consists of

$$X = \{Z, K, \underline{T}_K, \underline{r}_K, \underline{r}_K \underline{N}_K + (1 - \underline{r}_K)\underline{D}_K\}.$$

It is apparent that the recording of a binary outcome between two observation times can be caused by many reasons and one main reason, as in the CCSS discussed above, is due to the design of a study. In the following, we will assume that the observation-type indicator $r(s, t]$ is independent of both the underlying event process $\mathbb{N}(t)$ and the observation process. Note that in addition to the CCSS, this also holds in many other studies. For example, in the Study of Osteoporotic Fractures, the participants at the baseline were asked whether they ever had a broken or fractured bone and if so, to list every broken bone and age when broken. But in the following visits, they were only asked whether they had a broken or fractured bone without enough details for the occurrences of fractures. In other words, at the beginning, we have the detailed counts of events, but for the subsequent follow-ups, only binary outcomes are available.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2.2 Model, Likelihood, and Estimation

To describe the covariate effect, we will consider the following proportional mean model

$$\Lambda(t \mid Z) = E\{N(t) \mid Z\} = \Lambda_0(t) e^{\beta^T Z} \quad (1)$$

(Lin et al., 2000; Cai and Schaubel, 2004; Wellner and Zhang, 2007; Cheng et al., 2011), where β is the regression parameter and Λ_0 is a monotone nondecreasing baseline mean function. Note that for the likelihood-based analysis of either recurrent-event data or panel-count data, it is often assumed that conditionally on Z, $\mathbb{N}(t)$ is a nonhomogeneous Poisson process with the mean function given by (1). But in general, such analysis will still be valid even if the Poisson assumption is violated. Actually Wellner and Zhang (2000, 2007) applied this working model to the analyses of panel-count data and demonstrated the robustness to the Poisson assumption, as long as the proportional mean assumption (1) holds.

Also note that if we observe $\Delta N(s, t]$ over $(s, t]$, then the Poisson distribution of this increment is

$$P(\Delta N(s,t] = k \mid Z) = \frac{\{\Delta\Lambda_0(s,t] e^{\beta_0^T Z}\}^K}{k!} \exp\{-\Delta\Lambda_0(s,t] e^{\beta_0^T Z}\},$$

where $\Delta\Lambda_0(s, t] = \Lambda_0(t) - \Lambda_0(s)$. On the other hand, if only the binary outcome $D(s, t]$ is observed, then we have

$$P(D(s,t] = k \mid Z) = \{1 - e^{-\Delta\Lambda_0(s,t] e^{\beta_0^T Z}}\}^k \{e^{-\Delta\Lambda_0(s,t] e^{\beta_0^T Z}}\}^{1-k} \qquad \text{for } k = 0, 1.$$

Thus under the assumptions above and based on the conditional independence of the increments of N, the likelihood for a sample of n independent and identically distributed copies of X can be expressed as

$$L_n(\beta, \Lambda) = \prod_{i=1}^{n} \prod_{j=1}^{K_i} \left[\frac{\{\Delta\Lambda_{K_{i,j}} e^{\beta^T Z_i}\}^{\Delta N_{K_{i,j}}}}{\Delta N_{K_{i,j}}!} e^{-\Delta\Lambda_{K_{i,j}} e^{\beta^T Z_i}} \right]^{r_{K_{i,j}}} \times \left[\{1 - e^{-\Delta\Lambda_{K_{i,j}} e^{\beta^T Z_i}}\}^{D_{K_{i,j}}} \{e^{-\Delta\Lambda_{K_{i,j}} e^{\beta^T Z_i}}\}^{1 - D_{K_{i,j}}} \right]^{1 - r_{K_{i,j}}}.$$

It follows that the full log-likelihood function of (β, Λ) is given by

$$l_n(\beta, \Lambda) = \sum_{i=1}^{n} \sum_{j=1}^{K_i} \left(r_{K_{i,j}} \{\Delta N_{K_{i,j}} \log \Delta\Lambda_{K_{i,j}} + \Delta N_{K_{i,j}} \beta^T Z_i - \Delta\Lambda_{K_{i,j}} e^{\beta^T Z_i}\} + (1 - r_{K_{i,j}}) \left[D_{K_{i,j}} \log\{1 - e^{-\Delta\Lambda_{K_{i,j}} e^{\beta^T Z_i}}\} - (1 - D_{K_{i,j}}) \Delta\Lambda_{K_{i,j}} e^{\beta^T Z_i} \right] \right),$$

where

$$\Delta\Lambda_{K_{i,j}} = \Lambda(T_{K_{i,j}}) - \Lambda(T_{K_{i,j-1}})$$

for $i = 1, 2, \cdots, n$; $j = 1, 2, \cdots, K_i$.

For estimation of model (1) or (β, Λ), it is natural to use the SPMLE defined as $(\hat{\beta}_n, \widehat{\Lambda}_n) \equiv \arg\max_{(\beta, \Lambda) \in \mathscr{R} \times \mathscr{F}} l_n(\beta, \Lambda)$. Here $\mathscr{R}$ is a compact subset in the d-dimensional Euclidean space R^d and $\mathscr{F}$ is a class of monotone nondecreasing functions defined in $(0, \infty)$ that are bounded in a finite interval $(0, \tau]$ with the difference of the function values at any two points t_1, t_2 separated by s being bounded away from zero. In the next subsection, we will present an algorithm that can be used to compute the estimator $\hat{\theta}_n = (\hat{\beta}_n, \widehat{\Lambda}_n)$.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

2.3 Computation Algorithm

To compute the SPMLE of (β, Λ) defined above, we suggest to employ the following two-step algorithm in the spirit of profile likelihood. That is, for each fixed value of β, we set $\widehat{\Lambda}_n(\cdot, \beta) \equiv \arg\max_{\Lambda \in \mathscr{F}} l_n(\beta, \Lambda)$ and define the profile log-likelihood $l_n^p(\beta) \equiv l_n(\beta, \widehat{\Lambda}_n(\cdot, \beta))$. Then $\widehat{\beta}_n = \arg\max_{\beta \in \mathscr{R}} l_n^p(\beta)$ and $\widehat{\Lambda}_n = \widehat{\Lambda}_n(\cdot, \widehat{\beta}_n)$. Here the estimator $\widehat{\Lambda}_n$ is defined to have jumps only at the distinct observation time points in the collection of all observation times $\{T_{i,j}: i = 1, 2, \cdots, n; j = 1, 2, \cdots, K_i\}$. In summary, the two-step algorithm can be implemented by the following doubly iterative algorithm.

1. Choose the initial $\beta^{(0)}$.
2. For given $\beta^{(p)}$ (p=0, 1, 2, ⋯), update the estimate of Λ_0, $\Lambda^{(p)}$, by the modified iterative convex minorant (MICM) algorithm proposed by Jongbloed (1998) on the log-likelihood $l_n(\beta^{(p)}, \Lambda)$.
3. For given $\Lambda^{(p)}$, update the estimate of β, $\beta^{(p+1)}$, by optimizing $l_n(\beta, \Lambda^{(p)})$ using the Newton-Raphson method.
4. Repeat steps 2 and 3 until the following convergence criterion is satisfied:

$$\left| \frac{l_n(\beta^{(p+1)}, \Lambda^{(p+1)}) - l_n(\beta^{(p)}, \Lambda^{(p)})}{l_n(\beta^{(p)}, \Lambda^{(p)})} \right| \le \eta .$$

Remark 2.1—The details about the MICM algorithm for computing the MLE with panel-count data can be found in Wellner and Zhang (2000).

Remark 2.2—One can easily calculate that

$$\nabla_\beta^2 l_n(\beta, \Lambda) = -\sum_{i=1}^{n} \sum_{j=1}^{K_i} \left\{ \left[r_{K_{i,j}} + (1 - r_{K_{i,j}}) \left(\frac{D_{K_{i,j}} ee_{i,j} \{ e_{i,j} - 1 + ee_{i,j} \}}{\left(1 - ee_{i,j}\right)^2} + (1 - D_{K_{i,j}}) \right) \right] e_{i,j} \right\} Z_i Z_i^T,$$

where $e_{i,j} = \Delta\Lambda_{K_{i,j}} \exp(\beta^T Z_i)$ and $ee_{i,j} = \exp(-\Delta\Lambda_{K_{i,j}} \exp(\beta^T Z_i))$. Since

$$f(y) = y - 1 + e^{-y} > 0 \text{ for } y > 0,$$

$\nabla_\beta^2 l_n(\beta, \Lambda)$ is a negative definite matrix for any (β, Λ) if Z is not linearly dependent. Also based on this fact and the following two inequalities

$$l_n(\beta^{(p+1)}, \Lambda^{(p+1)}) - l_n(\beta^{(p+1)}, \Lambda^{(p)}) \ge 0,$$

and

$$l_n(\beta^{(p+1)}, \Lambda^{(p)}) - l_n(\beta^{(p)}, \Lambda^{(p)}) > 0,$$

one can easily prove that

$$l_n(\beta^{(p+1)}, \Lambda^{(p+1)}) - l_n(\beta^{(p)}, \Lambda^{(p)}) = l_n(\beta^{(p+1)}, \Lambda^{(p+1)}) - l_n(\beta^{(p+1)}, \Lambda^{(p)}) + l_n(\beta^{(p+1)}, \Lambda^{(p)}) - l_n(\beta^{(p)}, \Lambda^{(p)}) > 0$$

for $p = 1, 2, \cdots$. This suggests that the log-likelihood increases after each iteration and the algorithm behaves similar to an EM-algorithm, which may converge slowly in some scenarios. For the case here, however, the algorithm appears to work well and has no convergence issue in the numerical experiments below. More comments on this are given below.

In the next section, we will establish the asymptotic properties of the proposed estimator $\hat{\theta}_n = (\hat{\beta}_n, \hat{\Lambda}_n)$ with the proofs provided in the Web Supplementary Materials.

3. Asymptotic Results

In this section, we will establish the asymptotic properties of the estimator $\hat{\theta}_n = (\hat{\beta}_n, \hat{\Lambda}_n)$ with the proof given in the Web Supplementary Materials. For this, we need to adopt and reintroduce some notation and working model assumptions used in Wellner and Zhang (2007) to make the presentation of this article self-contained. Denote $\mathscr{B}_d$ and $\mathscr{B}$ as the collection of Borel sets in R^d and R^+, respectively, and let $B_{[0,\tau]} = \{B \cap [0, \tau]: \ B \in \mathscr{B}\}$ and $B^2_{[0,\tau]} = B_{[0,\tau]} \times B_{[0,\tau]}$. For any $B \in B_{[0,\tau]}$ and $C \in \mathscr{B}_d$, define the measure

$$\nu(B \times C) = \int_C \sum_{k=1}^{\infty} P(K = k | Z = z) \sum_{j=1}^{k} P(T_{K,j} \in B | K = k, Z = z) dH(z),$$

where H is the distribution function of Z and measure μ as $\mu(B) = \nu(B \times R^d)$. Define the L_2-metric $d(\theta_1, \theta_2)$ in the parameter space $\Theta = \mathscr{R} \times \mathscr{F}$ as

$$d(\theta_1, \theta_2) = \{|\beta_1 - \beta_2|^2 + \|\Lambda_1 - \Lambda_2\|^2_{L_2(\mu)}\}^{1/2}.$$

For the asymptotic properties, we also need the following regularity conditions.

C1. The true model parameter $\theta_0 = (\beta_0, \Lambda_0) \in \mathscr{R}^\circ \times \mathscr{F}$, where $\mathscr{R}^\circ$ is the interior of $\mathscr{R}$.

C2. The observation times $T_{K,j}$, for all $j = 1, \ldots, K$, $K = 1, 2, \ldots$, are random variables, taking values in the bounded interval $[\sigma, \tau]$ for some $\sigma, \tau \in (0, \infty)$ with $\Lambda_0(\sigma) > 0$, and the measure $\mu \times H$ on $B_{[0,\tau]} \times \mathscr{B}_d$ is absolutely continuous with respect to ν.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

C3. The true baseline mean function Λ_0 satisfies $\Lambda_0(\tau) \le M$ for some $M \in (0, \infty)$.

C4. The support of H, $\mathscr{Z} \equiv \mathrm{supp}(H)$, is a bounded set in R^d. (Thus, there exists $z_0 > 0$ such that $P(|Z| \le z_0) = 1$.)

C5. For all $a \in R^d$, $a \neq 0$, and $c \in R$, $P(a^T Z \neq c) > 0$.

C6. For some $k_0 < \infty$, we have $P(K \le k_0) = 1$.

C7. For some $v_0 \in (0, \infty)$, the function $Z \mapsto E(e^{v_0 \mathbb{N}(\tau)} | Z)$ is uniformly bounded for $Z \in \mathscr{Z}$.

C8. The observation time points are s_0−separated: that is, there exists a constant $s_0 > 0$ such that $P(T_{K,j} - T_{K,j-1} \ge s_0$ for all $j = 1, \ldots, K) = 1$. Furthermore, the measure μlutely continuous with respect to Lebesgue measure λ, with a derivative $\dot{\mu}$ satisfying $\dot{\mu}_1(t) \ge c_0 > 0$ for some positive constant c_0.

C9. The true baseline mean function Λ_0 is differentiable, and the derivative has positive and finite lower and upper bounds in the observation interval, i.e., there exists a constant $0 < f_0 < \infty$, such that $1/f_0 \le \Lambda_0'(t) \le f_0 < \infty$ for $t \in O[T]$.

C10. For some $\eta \in (0, 1)$, $a^T Var(Z|U, V)a \ge \eta a^T E(ZZ^T|U, V)a$ a.s. for all $a \in R^d$.

C11. For any $u, v \in O[T]$ satisfying $v - u \ge s_0$, the observation type indicator $r(u, v]$ is independent of or non-informative with respect to the underlying counting process. Moreover there exists $0 < p_{1,0}, p_{2,0} \le 1$ such that $p_{1,0} \le p(u, v) \le p_{2,0}$ with $p(u, v) = Er(u, v]$.

Remark 3.1

Note that the conditions C1–C10 are the same as those given in Wellner and Zhang (2007) for the asymptotic properties of the SPMLE based on panel-count data and they also provided some justification in view of the applications. The condition C11 is specifically made for the mixed panel-count data and implies that the chance of observing actual event counts in any observation interval has a non-zero lower bound. This condition is necessary for proving the consistency of the baseline mean function estimator.

Theorem 1 (Consistency and Rate of Convergence)—Suppose that the conditions C1–C6 and C8–C11 hold and the conditional mean function of counting process $\mathbb{N}$(t) is given by (1). Then, the *SPMLE* $\hat{\theta}_n = (\hat{\beta}_n, \hat{\Lambda}_n)$ converges to the true model parameter in probability in metric d, that is $d(\hat{\theta}_n, \theta_0) \to_p 0$. In addition, if the condition C7 holds, we have

$$n^{1/3} d(\hat{\theta}_n, \theta_0) = O_p(1).$$

Theorem 2 (Asymptotic Normality)—Suppose that the regularity conditions C1–C11 and the conditional mean function (1) hold. Then the SPMLE of $\hat{\beta}_n$ is asymptotically normal such as

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\sqrt{n}(\hat{\beta}_n - \beta_0) \rightarrow_d N\left(0, A^{-1}B\left(A^{-1}\right)^T\right),$$

where

$$A = E_{(K, \underline{T}_K, Z)}\left\{\sum_{j=1}^{K} (\Delta\Lambda_{0K,j})^2 e^{\beta_0^T Z} U_j(Z)[Z - R(K, T_{K,j-1}, T_{K,j})]^{\otimes 2}\right\},$$

and

$$B = E_{(K, \underline{T}_K, Z)}\left\{\sum_{j,j'=1}^{K} C_{j,j'}(Z)\left[Z - R(K, T_{K,j-1}, T_{K,j})\right]^{\otimes 2}\right\}.$$

In the above

$$C_{j,j'} = E(A_j A_{j'} \mid K, T_{K,j-1}, T_{K,j'-1}, T_{K,j}, T_{K,j'}, Z),$$

$$U_j(Z) = \frac{p(T_{K,j-1}, T_{K,j})}{\Delta\Lambda_{0K,j}} + \frac{[1 - p(T_{K,j-1}, T_{K,j})]e^{\beta_0^T Z} e^{-\Delta\Lambda_{0K,j} e^{\beta_0^T z}}}{1 - e^{-\Delta\Lambda_{0K,j} e^{\beta_0^T Z}}},$$

$$R(K, T_{K,j-1}, T_{K,j}) = \frac{E(U_j(Z)e^{\beta_0^T Z} Z \mid K, T_{K,j-1}, T_{K,j})}{E(U_j(Z)e^{\beta_0^T Z} \mid K, T_{K,j-1}, T_{K,j})},$$

and $A_j = r_{K,j}\left(\Delta N_{K,j} - \Delta\Lambda_{0K,j} e^{\beta_0^T Z}\right) + (1 - r_{K,j})\left(\frac{D_{K,j}}{1 - e^{-\Delta\Lambda_{0K,j} e^{\beta_0^T Z}}} - 1\right)\Delta\Lambda_{0K,j} e^{\beta_0^T Z}$, for j, j′ = 1, 2, …, K, and $Z^{\otimes 2}$ stands for ZZ^T.

Remark 3.2

It is worth emphasizing again that the nonhomogeneous Poisson process is just the working model to derive the "likelihood" of the observed mixed panel-count data. The validity of the asymptotic results stated above does not require the underlying counting process to be the Poisson process. It is also straightforward to verify that when N(t) is indeed a Poisson process, $C_{j,j'}(Z) = 0$ for $j \neq j'$ and $C_{j,j}(Z) = (\Delta\Lambda_{0K,j})^2 e^{\beta_0^T Z} U_j(Z)$, which result in B = A in

Theorem 2. This means that the SPMLE of $\hat{\beta}_n$ is semiparametrically efficient under the Poisson process. In addition, for a special case $p(T_{K,j-1}, T_{K,j}) \equiv 1$ for all $j = 1, 2, \cdots, k$, i.e., only panel-count data are observed, Theorem 2 is exactly the first part of Theorem 3.3 of Wellner and Zhang (2007). In other words, the estimator defined above can be viewed as the natural generalization of that given in Wellner and Zhang (2007) for panel-count data.

Remark 3.3

Note that although the asymptotic normality can be established as above, it is difficult to estimate the corresponding standard error empirically based on the sample. For this, we suggest to adopt the nonparametric bootstrap method with 100 resamplings, whose validity is warranted by the established asymptotic normality.

4. Simulation Results

We conducted extensive simulation studies to evaluate the finite sample performance of the proposed estimation procedure under different situations. In the study, for the ith subject, $i = 1, 2, \cdots, n$, we assumed that the covariate Z_i is three-dimensional and consists of a uniform variable over (0, 1), a standard normal variable, and a Bernoulli variable with 0.5 success probability. For the underlying counting process, we considered two cases. One is that $\mathbb{N}_i(t)$ was assumed to be a Poisson process with the conditional mean function given by $\Lambda(t|Z_i) = 2te^{\beta^T Z_i}$. The other is that the counting process was a mixed Poisson process with intensity $(2t + \alpha_i)e^{\beta^T Z_i}$, where α is a random effect on the intensity. Here we set $\alpha_i \in \{-0.4, 0, 0.4\}$ with probabilities 0.25, 0.5, and 0.25, respectively, so that the mean function of the counting process given the covariates still satisfies model (1).

The number of observations K_i was sampled randomly from the set {1, 2, 3, 4, 5, 6} with equal probabilities. Given K_i, the observation time points $T_{K_i, 1}, \cdots, T_{K_i, K_i}$ were taken to be ordered values generated from the uniform distribution over (1, 10) and rounded to the second decimal point to allow ties in observation times. Within each observation period $(T_{K_{i, j-1}}, T_{K_{i, j}}]$ for $j = 1, \cdots, K_i$, we generated the observation type indicator $r_{K_i, j}$ from the Bernoulli distribution with the probability of success p, where p was chosen to give the appropriate proportion of records in counts. If $r_{K_i, j} = 1$, then the event count over the interval $(T_{K_{i, j-1}}, T_{K_{i, j}}]$, $\Delta N_{K_{i, j}}$, was assumed to be known and recorded. Otherwise, it was assumed that only the binary outcome, $D_{K_{i, j}} = (\Delta N_{K_{i, j}} > 0)$, was recorded. The percentage of count-type observations, p, was set to be 0.1, 0.2, 0.5, 0.8, 0.9, or 0.95, and the true value of $\beta_0 = (\beta_1, \beta_2, \beta_3)^T$ was taken to be (0, 0, 0) or (−1, 0.5, 1.5). The results given below are based on the sample sizes of 100 and 200 with 1000 replications.

Tables 1 and 2 present the results on estimation of regression parameters obtained with the underlying recurrent event process being the Poisson process and $n = 100$ and $n = 200$, respectively. The results corresponding to the mixed Poisson process are given in Table 3

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

with $n = 100$ and the results with $n = 200$ are similar and thus omitted to save the space. In the tables, we include the average of the estimated β (EST), the average of the estimated asymptotic standard errors (ASE), the sampled standard error (SSE), and the 95% coverage probability (CP). Note that for the variance estimation, as mentioned above, we adopt the bootstrap procedure. The results suggest that the proposed estimator seems to be unbiased and the ASEs are close to the SSEs. Also the 95% confidence interval based on the asymptotic normal distribution seems to have the right CP and the ratios of SSEs between $n = 100$ and $n = 200$ are all close to $\sqrt{2}$, indicating that the normal approximation to the distribution of the proposed estimate seems to work properly. In addition, as expected, when the percentage of count-type records increased, the standard error of the estimates decreased because the counts provided more information than did the binary outcomes regarding the counting process. Moreover, the results in Table 2 suggest that the proposed SPMLE method is robust against the Poisson process assumption, though the estimated standard errors in the mixed Poisson scenario were larger than those in the Poisson scenario as expected. Note that in the study here and also the application below, we used $\beta = 0$ as the initial value in the algorithm. We did try different initial values and the algorithm seems always to converge and work well.

As mentioned above, a naive approach for analyzing mixed panel-count data is to ignore the binary part of the data and analyze only the available panel-count data. To compare the proposed method to this, in the simulation studies, we also obtained the results by using the naive approach based on the method given in Wellner and Zhang (2007) for panel-count data and include those results in Tables 1 – 3. One can clearly see that the naive estimation of the regression parameters still seems to be unbiased as expected. This is reasonable because β represents the time-invariant effect of the covariates on the process; thus it does not require the data from the whole process to provide the essential information for the inference. On the other hand, it is apparent that the estimate is much less efficient than that given by the proposed method as discussed above. Also as expected, the efficiency loss compared to the proposed method becomes more manifested when p decreases from 0.95 to 0.1.

To further compare the naive approach and the proposed method, we obtained and plotted the estimates of the baseline cumulative mean functions given by both methods for the Poisson scenario in Figure 1. One can see that the estimates given by the proposed method seem to be unbiased under all different scenarios of missing count probability, but the naive estimates seem to underestimate the cumulative mean functions. The lower the percentage of count-type records, the bigger the deviation. The downward bias is anticipated because the total number of counts for a given length of observation times in each subject used for panel-count data model is potentially smaller than the actual number. Also the discrepancy became larger when the proportion of binary outcome increased. This is true for the mixed Poisson scenario too, which is not presented here. In other words, these results suggest that one should apply the proposed method instead of the naive approach to analyze mixed panel-count data.

Note that in the proposed method, it has been assumed that the observation type indicators $r_{K_i, j}$'s follow the same distribution and are independent of other processes. To follow a

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

suggestion of a reviewer, we performed additional simulation studies to investigate the proposed estimation procedure when the $r_{K_i, j}$'s may depend on covariates. Table 4 presents some results on estimation of regression parameters obtained when there existed one covariate Z_i following the Bernoulli distribution with the success probability of 0.5 and the other set-ups being the same as in Tables 1 – 3. Here we set p =0.1, 0.4, or 0.7 for the subjects with $Z_i = 0$ and 0.3, 0.6, or 0.9 otherwise, and on the average, p is still equal to 0.2, 0.5, or 0.8. One can see from the table that they seem to give the similar conclusions to Tables 1 – 3 and suggest that the proposed method is still valid for the situation where the observation type is related but still conditionally independent of the underlying counting process given the covariates.

5. An Application

In this section, we apply the proposed method to analyze the CCSS data described in Section 1. As mentioned above, the CCSS follows over 14,000 childhood cancer survivors and a random sample of about 4000 healthy siblings as a control group. In the following, we will focus on a subgroup of 1801 women who were at least 25 years old in 1996; 611 participants are leukemia survivors, and 1190 participants are healthy controls. Since they are subgroups of CCSS, these leukemia survivors and siblings are no longer paired. In fact, among 1801 participants, only 66 pairs of survivors and siblings are from the same family, so we consider them as two independent groups here. For the analysis, define $X_i = 1$ if the ith subject is a leukemia survivor and $X_i = 0$ otherwise. Both leukemia survivor and healthy control groups have an average of around 3 observations per subject. In total, there are 5494 records. For the question,"Have you and a partner ever become pregnant since last follow up?" 1930 records indicated "yes"(35%), with 549 records from leukemia survivors and 1381 records from healthy controls. Among the 1930 records with answer "yes" to the first question, 85 did not provide an answer to the second question, "Including live births, stillbirths, miscarriages, and abortions, how many times have you become pregnant?" As a result, the count of events in that record is missing, and we have only mixed panel-count data.

For the analysis, we first ignore the binary part of the data (the 85 records). In this case, the averaged pregnancy counts are 1.60 and 2.31 for the leukemia survivor and healthy control groups, respectively. By applying the method given in Wellner and Zhang (2007), as in the simulation study, the estimate of the regression parameter is −0.365 with the estimated standard error being 0.058. However, we know that true averages should be higher than 1.60 and 2.31 in the two groups because there should be at least 1 pregnancy per record in the 85 records. By applying the proposed estimation procedure, we obtained the estimate of −0.356 with the estimated standard error being 0.042. The p-values from both methods are less than 0.0001 for testing no pregnancy process difference between the leukemia survivor and healthy control groups, suggesting that the leukemia survivors had significantly lower pregnancy rates than the healthy controls. However, the standard error from using panel-count data alone is larger than that from using mixed panel-count data, which indicates an efficiency loss. Figure 2 provides the estimates of the expected cumulative number of pregnancies for leukemia survivors and healthy controls, respectively. As seen from the simulation studies, the estimated expected cumulative number of pregnancies when the

Author Manuscript

binary part of the data is ignored, is slightly lower than that when all available data are used. However, the underestimation is minor in this case because the binary data comprise only 4.4% of the total recorded data.

6. Discussion and Concluding Remarks

This article considered regression analysis of mixed panel-count data. Such data arise when the numbers of events are known during some follow-up periods while during other follow-up periods, only binary outcomes are known (e.g., whether the events of interest have occurred since the last follow-up). As discussed above, mixed panel-count data can occur in many fields. Although much literature exists for panel-count data, little work has been done for the type of data considered here. For this problem, the SPMLE method based on the working model of the Poisson process was developed, and the asymptotic properties of the proposed estimator were established. Also, the numerical results suggest that the proposed method works well in finite samples, even if the Poisson assumption does not hold. It is worth to stress that the proposed method has advantages in making semiparametric inference on the underlying counting process because the existing SPMLE method for panel-count data can be less efficient in making inferences for regression parameters for mixed panel-count data. More seriously, it could cause misleading results in the estimation of the expected cumulative number of events.

There are several directions for future research. One is that as discussed above, in the preceding sections, we have focused on the situation where the observation type indicators $r_{K,j}$'s follow the same distribution and are independent of the underlying counting process and the observation process. As shown in Section 4, one can relax the assumption to allow the $r_{K,j}$'s to depend on covariates or the two processes through covariates. However, in practice, there may exist situations where they are still correlated even given covariates and it would be useful to generalize the proposed method to the situation. Sun et al. (2007) discussed a similar situation and showed that ignoring the dependence could lead to biased results. On the other hand, it does not seem to be straightforward to incorporate the potential correlation in the methodology development and it can be very technically challenging.

In addition to the mixed panel-count data discussed above, other types of mixed panel-count data may exist and one is that panel-count data may be presented as panel-ordinal data. For example, a question on hospitalization could still be, "How many different times were you in a hospital at least overnight?", but multiple-choice items such as "A. 0; B. 1–2 times; C. 3–5 times; D. >5 times" would be provided to participants to avoid recall bias. It is easy to see that this type of data has more information than just panel-binary data but less information than panel-count data and could be mixed with them as well. The proposed method could be conceptually applied to this type of mixed panel-ordinal data. In addition, panel-count data can also be mixed with recurrent-event data, as mentioned above. For the CCSS study, the hospitalization data contain panel-binary data, panel-count data, and recurrent-event data. The pregnancy data in the CCSS have some recurrent-event data as well. Among the four types of data, recurrent-event data contain the most information, followed by panel-count data, panel-ordinal data, and panel-binary data. It would be very useful and instructive to develop a general method that allows the analysis of mixed data of any of the four types.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors wish to thank the Editor, an Associate Editor and a reviewer for their many insightful and useful comments and suggestions that greatly improved the paper. This work was supported in part by the National Institutes of Health Grant (1R03CA219450-01).

References

Cai J, Schaubel DE. Marginal means/rates models for multiple type recurrent event data. Lifetime Data Analysis. 2004; 10:121–138. [PubMed: 15293628]

Cheng G, Zhang Y, Lu L. Efficient algorithms for computing the non and semi-parametric maximum likelihood estimates with panel count data. Journal of Nonparametric Statistics. 2011; 23:567–579.

Jongbloed G. The iterative convex minorant algorithm for nonparametric estimation. Journal of Computational and Graphical Statistics. 1998; 7:310–321.

Lin DY, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2000; 62:711–730.

Sun J, Fang H. A nonparametric test for panel count data. Biometrika. 2003; 90:199–208.

Sun J, Kalbfleisch JD. Estimation of the mean function of point processes based on panel count data. Statistica Sinica. 1995; 5:279–290.

Sun J, Wei LJ. Regression analysis of panel count data with covariate-dependent observation and censoring times. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2000; 62:293–302.

Sun J, Tong X, He X. Regression Analysis of Panel Count Data with Dependent Observation Times. Biometrics. 2007; 63:1053–1059. [PubMed: 18078478]

Sun, J., Zhao, X. Statistical Analysis of Panel Count Data. Springer; 2013.

Thall PF, Lachin JM. Analysis of recurrent events: Nonparametric methods for random-interval count data. Journal of the American Statistical Association. 1988; 83:339– 347.

Wellner JA, Zhang Y. Two Estimators of the Mean of a Counting Process with Panel Count Data. The Annals of Statistics. 2000; 28:779–814.

Wellner JA, Zhang Y. Two likelihood-based semiparametric estimation methods for panel count data with covariates. The Annals of Statistics. 2007; 35:2106–2142.

Zhu L, Zhao H, Sun J, Leisenring W, Robison LL. Regression Analysis of Mixed Recurrent-Event and Panel-Count Data with Additive Rate Models. Biometrics. 2015:71–79. [PubMed: 25345405]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

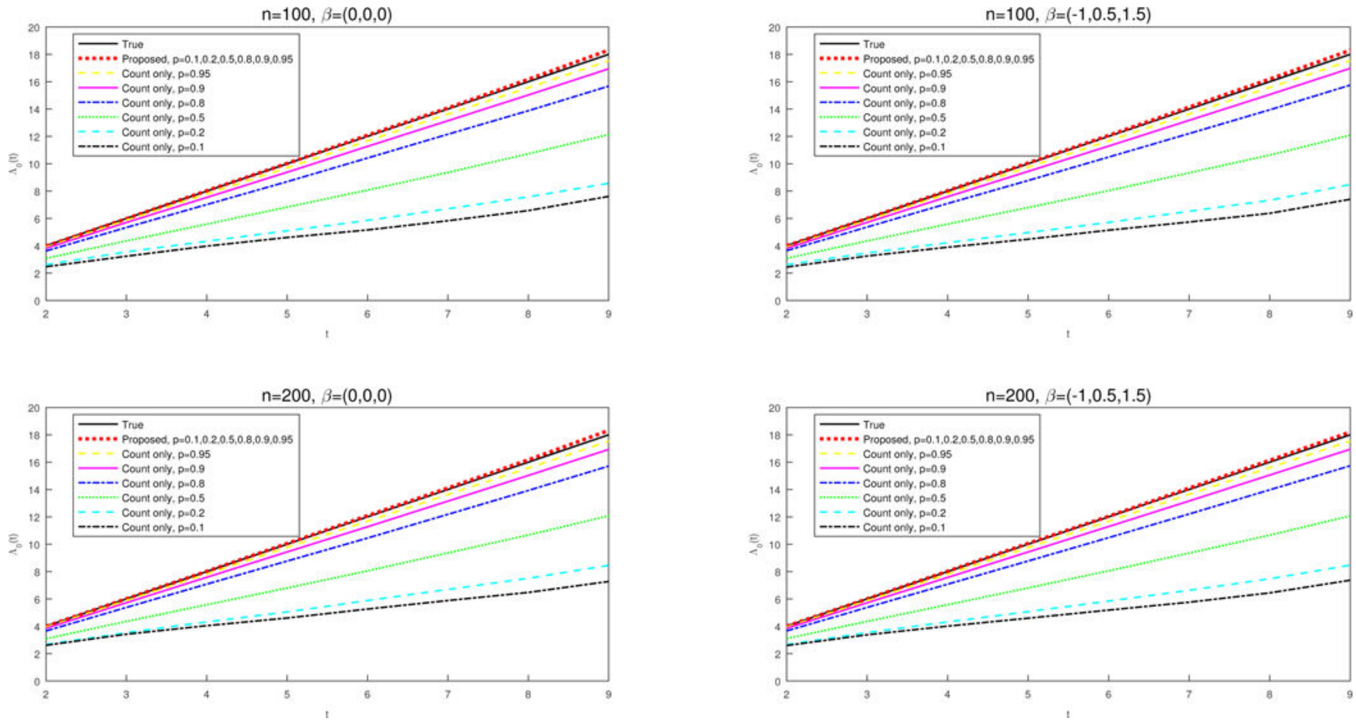

Fig. 1.

The averages of the maximum likelihood estimates of the baseline mean function under different scenarios from the proposed method and from the method that uses only the panel-count data.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

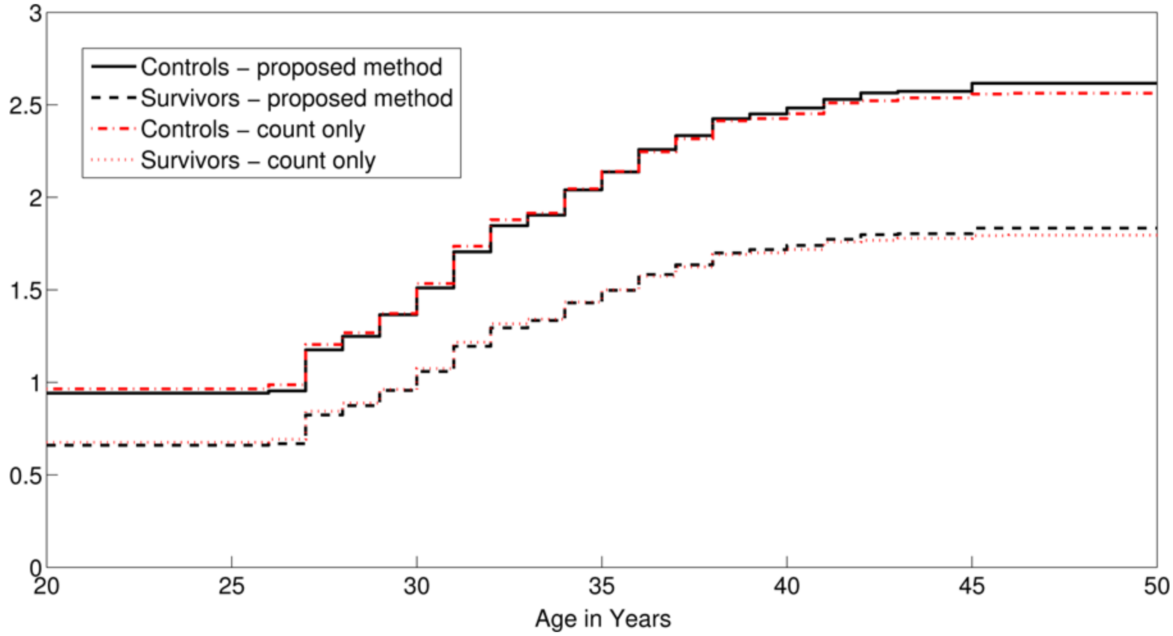

Fig. 2.
Estimated cumulative average numbers of pregnancies

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Table 1

Results on estimation of regression parameters based on the Poisson process with n = 100

	Proposed method				Naive method	
p	**EST**	**ASE**	**SSE**	**CP**	**EST**	**SSE**
			β=(0, 0, 0)			
0.1	0.072	0.163	0.164	0.911	−0.017	0.684
	−0.005	0.074	0.075	0.931	−0.006	0.203
	0.024	0.129	0.128	0.942	0.008	0.395
0.2	0.018	0.123	0.124	0.935	−0.010	0.433
	0.001	0.055	0.055	0.931	0.001	0.127
	0.010	0.098	0.098	0.948	−0.007	0.246
0.5	−0.018	0.092	0.098	0.937	−0.010	0.231
	−0.002	0.036	0.037	0.939	0.000	0.067
	−0.010	0.066	0.067	0.932	−0.008	0.129
0.8	−0.019	0.077	0.080	0.939	−0.006	0.131
	−0.001	0.029	0.030	0.932	−0.002	0.041
	−0.010	0.054	0.053	0.949	−0.004	0.078
0.9	−0.011	0.072	0.072	0.957	−0.008	0.108
	−0.001	0.027	0.028	0.931	−0.001	0.034
	−0.008	0.050	0.050	0.945	−0.006	0.065
0.95	−0.007	0.069	0.069	0.948	−0.008	0.100
	−0.001	0.027	0.028	0.942	−0.001	0.031
	−0.006	0.049	0.048	0.947	−0.006	0.058
			β = (−1, 0.5, 1.5)			
0.1	−0.940	0.158	0.153	0.928	−0.966	0.755
	0.495	0.067	0.066	0.948	0.501	0.244
	1.499	0.110	0.107	0.954	1.505	0.422
0.2	−0.975	0.126	0.123	0.940	−0.961	0.519
	0.497	0.048	0.046	0.949	0.498	0.154
	1.491	0.078	0.079	0.942	1.492	0.274

	Proposed method				Naive method	
p	**EST**	**ASE**	**SSE**	**CP**	**EST**	**SSE**
0.5	−1.002	0.082	0.082	0.951	−1.007	0.253
	0.501	0.029	0.029	0.933	0.497	0.084
	1.495	0.049	0.047	0.944	1.487	0.132
0.8	−1.000	0.064	0.064	0.946	−1.011	0.141
	0.500	0.022	0.021	0.950	0.499	0.046
	1.499	0.037	0.037	0.947	1.491	0.076
0.9	−0.999	0.060	0.061	0.943	−1.011	0.101
	0.500	0.020	0.020	0.947	0.499	0.033
	1.499	0.034	0.034	0.944	1.493	0.058
0.95	−0.998	0.059	0.059	0.938	−1.005	0.083
	0.500	0.020	0.020	0.948	0.499	0.027
	1.499	0.033	0.033	0.946	1.493	0.050

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2

Results on estimation of regression parameters based on the Poisson process with n = 200

	Proposed method				Naive method	
p	**EST**	**ASE**	**SSE**	**CP**	**EST**	**SSE**
			β=(0, 0, 0)			
0.1	0.034	0.106	0.107	0.928	−0.005	0.442
	0.000	0.048	0.050	0.936	0.001	0.125
	0.013	0.085	0.085	0.936	−0.008	0.250
0.2	0.004	0.085	0.085	0.939	−0.004	0.278
	0.000	0.037	0.037	0.946	−0.002	0.082
	0.004	0.067	0.065	0.947	0.004	0.170
0.5	−0.022	0.067	0.069	0.938	−0.008	0.150
	0.000	0.025	0.025	0.944	−0.003	0.046
	−0.006	0.047	0.047	0.950	−0.001	0.089
0.8	−0.015	0.058	0.053	0.966	−0.011	0.089
	0.000	0.020	0.021	0.929	0.001	0.029
	−0.004	0.038	0.038	0.945	−0.004	0.056
0.9	−0.009	0.050	0.046	0.971	−0.012	0.077
	0.000	0.019	0.020	0.931	0.000	0.024
	−0.001	0.035	0.035	0.948	−0.003	0.045
0.95	−0.006	0.047	0.044	0.963	−0.011	0.070
	0.000	0.019	0.019	0.933	0.000	0.021
	−0.001	0.034	0.034	0.948	−0.002	0.040
			β=(−1, 0.5, 1.5)			
0.1	−0.979	0.104	0.103	0.943	−1.021	0.510
	0.498	0.041	0.040	0.950	0.488	0.154
	1.507	0.068	0.069	0.931	1.510	0.259
0.2	−0.992	0.082	0.080	0.947	−1.016	0.352
	0.499	0.030	0.029	0.946	0.493	0.110
	1.499	0.050	0.049	0.951	1.507	0.176

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

	Proposed method				Naive method	
p	EST	ASE	SSE	CP	EST	SSE
0.5	−1.001	0.056	0.056	0.942	−1.035	0.175
	0.500	0.019	0.020	0.931	0.499	0.053
	1.497	0.033	0.033	0.941	1.491	0.096
0.8	−0.999	0.044	0.045	0.940	−0.992	0.090
	0.500	0.015	0.015	0.938	0.494	0.034
	1.498	0.025	0.026	0.938	1.491	0.054
0.9	−0.999	0.041	0.042	0.937	−1.003	0.073
	0.500	0.014	0.014	0.934	0.499	0.023
	1.499	0.024	0.024	0.942	1.491	0.041
0.95	−0.999	0.040	0.042	0.938	−1.005	0.060
	0.500	0.013	0.013	0.946	0.498	0.019
	1.499	0.023	0.024	0.939	1.492	0.035

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3

Results on estimation of regression parameters based on the mixed Poisson process with n = 100

	Proposed method				Naive method	
p	**EST**	**ASE**	**SSE**	**CP**	**EST**	**SSE**
			β=(0, 0, 0)			
0.1	0.055	0.170	0.175	0.919	0.010	0.689
	−0.001	0.077	0.080	0.930	0.008	0.203
	0.021	0.134	0.137	0.934	0.032	0.398
0.2	0.014	0.136	0.139	0.927	0.003	0.424
	0.001	0.059	0.059	0.937	0.010	0.123
	0.008	0.104	0.108	0.929	0.020	0.260
0.5	−0.024	0.102	0.109	0.932	−0.001	0.236
	0.003	0.040	0.040	0.936	0.003	0.067
	−0.006	0.075	0.076	0.936	0.007	0.141
0.8	−0.023	0.088	0.091	0.939	−0.006	0.146
	0.001	0.033	0.033	0.946	0.001	0.044
	−0.007	0.061	0.060	0.953	−0.002	0.082
0.9	−0.016	0.083	0.079	0.954	−0.007	0.127
	0.001	0.031	0.031	0.943	0.002	0.037
	−0.005	0.058	0.057	0.947	−0.002	0.071
0.95	−0.012	0.079	0.075	0.946	−0.007	0.112
	0.002	0.030	0.030	0.948	0.001	0.033
	−0.003	0.056	0.056	0.947	−0.001	0.064
			β = (−1, 0.5, 1.5)			
0.1	−0.967	0.169	0.171	0.943	−1.017	0.802
	0.494	0.072	0.074	0.939	0.496	0.239
	1.505	0.122	0.117	0.947	1.522	0.422
0.2	−0.991	0.142	0.137	0.935	−1.018	0.526
	0.496	0.055	0.054	0.935	0.493	0.157
	1.497	0.090	0.089	0.939	1.491	0.266

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

	Proposed method				Naive method	
p	**EST**	**ASE**	**SSE**	**CP**	**EST**	**SSE**
0.5	−1.000	0.097	0.099	0.933	−0.993	0.273
	0.498	0.036	0.038	0.927	0.493	0.086
	1.493	0.060	0.060	0.943	1.482	0.140
0.8	−0.999	0.082	0.085	0.928	−1.006	0.158
	0.500	0.029	0.030	0.925	0.499	0.050
	1.497	0.050	0.051	0.925	1.487	0.086
0.9	−0.999	0.078	0.081	0.931	−1.005	0.127
	0.500	0.028	0.029	0.926	0.499	0.040
	1.498	0.047	0.048	0.931	1.486	0.068
0.95	−0.999	0.076	0.079	0.929	−1.007	0.107
	0.500	0.027	0.029	0.930	0.499	0.035
	1.498	0.046	0.046	0.932	1.489	0.059

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Results on estimation of regression parameters with the observation type dependent on covariates

		***n* = 100**				***n* = 200**			
Poisson Process									
β	p	EST	ASE	SSE	CP	EST	ASE	SSE	CP
−1	0.2	−1.019	0.113	0.105	0.953	−1.031	0.086	0.081	0.946
	0.5	−1.039	0.098	0.092	0.933	−1.048	0.078	0.073	0.909
	0.8	−1.033	0.085	0.078	0.942	−1.027	0.066	0.056	0.938
0	0.2	0.001	0.074	0.073	0.947	−0.006	0.051	0.050	0.950
	0.5	−0.017	0.057	0.058	0.941	−0.023	0.043	0.044	0.923
	0.8	−0.015	0.048	0.049	0.945	−0.011	0.034	0.033	0.961
1	0.2	1.001	0.058	0.051	0.955	0.997	0.045	0.031	0.961
	0.5	0.993	0.046	0.033	0.955	0.996	0.037	0.022	0.969
	0.8	0.997	0.040	0.025	0.970	0.998	0.032	0.017	0.968
mixed-Poisson Process									
−1	0.2	−1.021	0.116	0.106	0.953	−1.034	0.090	0.075	0.962
	0.5	−1.041	0.100	0.098	0.938	−1.047	0.081	0.070	0.917
	0.8	−1.032	0.088	0.082	0.943	−1.025	0.070	0.056	0.956
0	0.2	0.001	0.078	0.080	0.940	−0.011	0.056	0.056	0.945
	0.5	−0.021	0.063	0.065	0.938	−0.022	0.047	0.047	0.944
	0.8	−0.015	0.054	0.055	0.956	−0.011	0.039	0.037	0.956
1	0.2	1.000	0.067	0.056	0.950	0.996	0.050	0.040	0.929
	0.5	0.994	0.053	0.042	0.945	0.993	0.042	0.030	0.955
	0.8	0.997	0.047	0.034	0.953	0.998	0.036	0.024	0.946