Joint Analysis of Panel Count Data with an Informative Observation Process and a Dependent Terminal Event

Jie Zhou[1], Haixiang Zhang[2*], Liuquan Sun[3], Jianguo Sun[4]

[1] School of Mathematical Sciences, Capital Normal University, Beijing 100048, China

[2*] Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

[3]Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing 100190, China

[4]Department of Statistics, University of Missouri, Columbia, MO 65211, U.S.A

Abstract Panel count data occur in many clinical and observational studies, and in many situations, the observation process may be informative and also there may exist a terminal event such as death which stops the follow-up. In this article, we propose a new joint model for the analysis of panel count data in the presence of both informative observation process and a dependent terminal event via two latent variables. For the inference on the proposed model, a class of estimating equations is developed and the resulting estimators are shown to be consistent and asymptotically normal. In addition, a lack-of-fit test is provided for assessing the adequacy of the model. Simulation studies suggest that the proposed approach works well for practical situations. Also an illustrative example from a bladder cancer clinical trial is used to illustrate the methods.

Keywords: Estimating equation; Informative observation process; Joint modeling; Panel count data; Terminal event.

[2*]Corresponding author: zhx_math@163.com (H. Zhang)

1 Introduction

Panel count data usually occur in longitudinal follow-up studies that concern occurrence rates of certain recurrent events. This kind of data usually arise from event history studies that concern some recurrent events and in which subjects are monitored or observed only at discrete time points instead of continuously. The fields in which one often sees such data include demographical and epidemiological studies, medical researches, reliability experiments, tumorgenicity experiments and sociological studies (Kalbfleisch and Lawless, 1985; Thall and Lachin, 1988; Sun, 2006).

Many authors have investigated the analysis of panel count data. For example, Sun and Kalbfleisch (1995) considered the estimation of the mean function of the underlying point process that yields panel count data. Balakrishnan and Zhao (2009, 2010, 2011), Park, et al. (2007), Sun and Fang (2003) and Zhao and Sun (2011) presented some nonparametric test procedures for the comparison of the mean functions of counting processes based on panel count data. Hu et al. (2003) and Sun and Wei (2000) developed some estimating equation-based methods for regression analysis of panel count data. Wellner and Zhang(2007) and Zhang (2002) also discussed regression analysis of panel count data and gave some likelihood-based approaches. Furthermore, Huang et al. (2006) and Sun et al. (2007) considered regression analysis of panel count data with dependent observation times. Li et al. (2010) proposed a class of semiparametric transformation models for panel count data with dependent observation process. Zhang, et al. (2013) presented a robust joint model for multivariate panel count data via latent variables. Tong, et al. (2009) and Zhang, et al. (2013) considered the variable selection issues on panel count data. A relatively complete references on panel count data can be found in Sun and Zhao (2013).

Most of the existing methods for panel count data assume that there is no terminal event and the observation process is independent of the underlying recurrent event process unconditionally or conditional on the covariates. In many situations, however, the follow-up of the study subjects could be stopped by a terminal event, such as death, which precludes further recurrent events. For example, in a tumourigenicity study, tumours would not develop after death. Furthermore, it is often the case that the terminal event is strongly correlated with the recurrent events of interest as well as the observation process. In the presence of terminal events, there exists considerable work on the analysis of recurrent event data and longitudinal data analysis, and two approaches are commonly adopted. One is the marginal model approach (Cook and Lawless, 1997; Ghosh and Lin, 2002; Zhao, et al., 2011), and the other is the frailty model approach (Huang and Wang, 2004; Liu et al., 2004; Ye et al. 2007; Zeng and Cai, 2010, Sun, et al., 2012). However, the problem is much harder for panel count data. To deal with these problems, we propose a new joint model for the analysis of panel count data in the presence of both informative observation times and a dependent terminal event via two latent variables. The association among the recurrent event process, the observation times, and the terminal event is modeled nonparametrically. The proposed joint model is flexible and robust in that the distributions of the latent variables and the dependence structures are left unspecified.

The rest of this paper is organized as follows. In Section 2, we introduce some notation and describe the proposed models that will be used throughout the paper. Specifically, we will describe the joint model of the recurrent event process, the observation times, and the terminal event through two latent variables. In Section 3, an estimating equation approach is developed for the estimation of the regression parameters. Also we establish the asymptotic properties of the proposed estimates. In Section 4, we develop a technique for checking the adequacy of the proposed model. Section 5 reports some results from simulation studies conducted for evaluating the proposed methods. In Section 6, we apply our proposed method to a bladder cancer study and some concluding remarks are provided in Section 7. Details of the proof are give in the Appendix.

2 Notation and Models

Consider a recurrent event study and let $N(t)$ denote the number of the occurrences of the recurrent event of interest up to time t, $0 \le t \le \tau$, where τ is a known constant representing the end of study. For each subject, suppose that a $d \times 1$ vector of covariates X is observed and let D be the time of the terminal event, such as death, and C be the censoring time. Define $T = C \wedge D$ and $\delta = I(D \le C)$, where $a \wedge b = \min(a, b)$, $I(\cdot)$ is the indicator function. Let u and v be two latent variables which are independent of X. For any time t, suppose that given (u, v, X) and $D \ge t$, the mean function of $N(t)$ has the form,

$$E\{N(t)|X, D \ge t, u, v\} = \mu_N(t; u) \exp(X'\beta_0), \tag{1}$$

where $\mu_N(t; u)$ is an unknown baseline mean function, and β_0 is a vector of unknown regression parameters.

Let $H(t)$ denote the observation process, and assume that $H(t)$ is independent of $N(t)$ conditional on (u, v, X) and $D \ge t$, and follows the rate model,

$$E\{dH(t)|X, D \ge t, u, v\} = \exp(X'\gamma_0) d\mu_H(t; v), \tag{2}$$

where γ_0 is a vector of unknown regression parameters, and $\mu_H(t; v)$ is an unknown baseline mean function with $\mu_H(0; v) = 0$. The recurrent event and observation processe are related to the terminal event through latent variables u and v, respectively. The condition $D \ge t$ is used because it is of interest in many studies to make inference for subjects who are currently alive (Ye et al., 2007; Zeng and Cai, 2010; Zhao et al., 2011).

For the terminal event, we assume that it follows the Cox model,

$$\log \Lambda_0(D) = -X'\eta_0 + \epsilon, \tag{3}$$

where η_0 is a vector of unknown regression parameters, $\Lambda_0(t)$ is an unspecified baseline cumulative hazard function, and ϵ is a random error with extreme-value distribution. In the following, the joint distribution of u, v and ϵ will be left unspecified. Hence the joint model (1), (2) and (3) are extensive since $\mu_N(t;u)$ and $\mu_H(t;v)$ are both nonparametric and depend on latent random variables which are associated with the terminal event via ϵ in an arbitrary way. In what follows, we assume that given X, the censoring time C is independent of $\{u, v, D, N(\cdot), H(\cdot)\}$. For a sample of n subjects, the observed data consist of $\{N_i(t)dH_i(t), T_i, \delta_i, X_i, H_i(t), 0 \le t \le T_i, i = 1, \cdots, n\}$.

3 Inference Procedure

Now we discuss the estimation of the parameters β_0 and γ_0. For this, note that D can be censored and the latent variables u and v are unobservable and thus it is impossible to make inference for the parameters of interest directly. To overcome this problem, we first consider the observed mean function given the observed endpoint T, in which the resulting nonparametric component depends on latent variables. Then we will derive an expression of the nonparametric component which can be estimated using the observed data for given β and γ. Then the resulting estimator makes the latent variables disappear. The details are as follows: Let $\mathcal{A}_0(t; u, v) = \int_0^t \mu_N(z; u) d\mu_H(z; v)$ and define $dR(t, s) = E\{d\mathcal{A}_0(t; u, v) | \epsilon \ge s\}$. Following the assumption that (u, v, ϵ) is independent of (X, C), we obtain that

$$E\{N(t)dH(t)|X, T \ge t\} = \exp\{X'(\beta_0 + \gamma_0)\} dR(t, \log \Lambda_0(t) + X'\eta_0), \tag{4}$$

and

$$dR(t, s) = \frac{E[N(t)dH(t)I\{\log \Lambda_0(T) + X'\eta_0 \ge s \ge \log \Lambda_0(t) + X'\eta_0\}]}{E[\exp\{X'(\beta_0 + \gamma_0)\}I\{\log \Lambda_0(T) + X'\eta_0 \ge s \ge \log \Lambda_0(t) + X'\eta_0\}]}. \tag{5}$$

So $\log \Lambda_0(T) + X'\eta_0 \ge s \ge \log \Lambda_0(t) + X'\eta_0$ implies $T \ge t$. The derivation of (4) and (5) is given in the Appendix.

To obtain an estimate of $dR(t, s)$, we have to give the estimates of η_0 and $\Lambda_0(t)$ from model (3). According to Fleming and Harrington (1991), we can get the maximum partial likelihood estimator $\hat{\eta}$ and the Breslow estimator $\hat{\Lambda}_0(t)$. Then given β and γ, we have

$$d\hat{R}(t,s;\beta,\gamma) = \frac{\sum_{i=1}^{n} N_i(t)dH_i(t)I\{\log\hat{\Lambda}_0(T_i) + X_i'\hat{\eta} \geq s \geq \log\hat{\Lambda}_0(t) + X_i'\hat{\eta}\}}{\sum_{j=1}^{n} \exp\{X_j'(\beta+\gamma)\}I\{\log\hat{\Lambda}_0(T_j) + X_j'\hat{\eta} \geq s \geq \log\hat{\Lambda}_0(t) + X_j'\hat{\eta}\}}.$$

For given γ, to estimate β_0, motivated by (4) and the generalized estimating equation approach (Liang and Zeger, 1986), we propose the following estimating function for β_0,

$$\begin{aligned} U(\beta;\gamma) &= \sum_{i=1}^{n}\int_0^{\tau} W(t)\{X_i - \bar{X}_i(t;\beta,\gamma)\}\Delta_i(t)\Big[N_i(t)dH_i(t) - \exp\{X_i'(\beta+\gamma)\} \\ &\quad \times \frac{\sum_{j=1}^{n} N_j(t)dH_j(t)\hat{\Phi}_j(t,X_i)}{\sum_{j=1}^{n}\exp\{X_j'(\beta+\gamma)\}\hat{\Phi}_j(t,X_i)}\Big], \end{aligned} \tag{6}$$

where $\Delta_i(t) = I(T_i \geq t)$, $W(t)$ is a possibly data-dependent weight function,

$$\hat{\Phi}_j(t,X_i) = I\{\log\hat{\Lambda}_0(T_j) + X_j'\hat{\eta} \geq \log\hat{\Lambda}_0(t) + X_i'\hat{\eta} \geq \log\hat{\Lambda}_0(t) + X_j'\hat{\eta}\},$$

$$\bar{X}_i(t;\beta,\gamma) = \frac{\sum_{j=1}^{n} X_j\exp\{X_j'(\beta+\gamma)\}\hat{\Phi}_j(t,X_i)}{\sum_{j=1}^{n}\exp\{X_j'(\beta+\gamma)\}\hat{\Phi}_j(t,X_i)}.$$

Of course in reality, γ_0 is unknown. From Zeng and Cai (2010), we propose the following estimating equation for γ_0,

$$\tilde{U}(\gamma) = \sum_{i=1}^{n}\int_0^{\tau} Q(t)\{X_i - \bar{X}_i^*(t;\gamma)\}\Delta_i(t)\{dH_i(t) - \exp(X_i'\gamma)d\bar{H}_i(t;\gamma)\} = 0, \tag{7}$$

where $Q(t)$ is a possibly data-dependent weight function, and

$$\bar{X}_i^*(t;\gamma) = \frac{\sum_{j=1}^{n} X_j\exp(X_j'\gamma)\hat{\Phi}_j(t,X_i)}{\sum_{j=1}^{n}\exp(X_j'\gamma)\hat{\Phi}_j(t,X_i)},$$

$$d\bar{H}_i(t;\gamma) = \frac{\sum_{j=1}^{n} dH_j(t)\hat{\Phi}_j(t,X_i)}{\sum_{j=1}^{n}\exp(X_j'\gamma)\hat{\Phi}_j(t,X_i)}.$$

Denote $\hat{\gamma}$ as the solution to estimating equation (7), and $\hat{\beta}$ as the solution to $U(\beta; \hat{\gamma}) = 0$. According to the law of large numbers and the consistency of $\hat{\eta}$ and $\hat{\Lambda}_0(t)$, we can obtain that $\hat{\beta}$ and $\hat{\gamma}$ are consistent. The following theorem presents the asymptotic normality of $\hat{\beta}$ and $\hat{\gamma}$, and the proof details are given in the Appendix.

Theorem 1. *Assume that the regularity conditions C.1-C.3 stated in the Appendix hold, then $n^{1/2}(\hat{\beta} - \beta_0)$ and $n^{1/2}(\hat{\gamma} - \gamma_0)$ have an asymptotic multivariate normal distribution with mean zero and covariance matrix $A^{-1}\Sigma(A^{-1})'$, where A and Σ are defined in the Appendix.*

We need to estimate the asymptotic covariance of $\hat{\beta}$ and $\hat{\gamma}$. First, A can be consistently estimated by $\hat{A}$, where

$$\hat{A} = -n^{-1} \begin{pmatrix} \partial U(\hat{\beta}; \hat{\gamma})/\partial \beta & \partial U(\hat{\beta}; \hat{\gamma})/\partial \gamma \\ 0 & \partial \tilde{U}(\hat{\gamma})/\partial \gamma \end{pmatrix}.$$

Our next goal is to estimate Σ, but Σ is complicated and involves the Hadamard derivatives of $d\bar{M}_0(t, X; \eta, \Lambda)$ and $d\bar{H}_0(t, X; \eta, \Lambda)$ with respect to Λ, thus direct estimation of Σ is not feasible. Here $d\bar{M}_0(t, X; \eta, \Lambda)$ and $d\bar{H}_0(t, X; \eta, \Lambda)$ are defined in the Appendix. To deal with this problem, we propose the following Monte Carlo method: from the proof of Theorem 1, we know that the variation of $U(\beta_0; \gamma_0)$ comes from $dM_i(t) - \exp\{X_i'(\beta_0 + \gamma_0)\}d\bar{M}(t, X_i; \hat{\eta}, \hat{\Lambda}_0)$, the empirical summation in the numerator and denominator of $\bar{M}(t, X_i; \hat{\eta}, \hat{\Lambda}_0)$ and the plug-in estimator $(\hat{\eta}, \hat{\Lambda}_0)$. Here $dM_i(t)$ and $d\bar{M}(t, X; \eta, \Lambda)$ are defined in the Appendix. We will use resampling approach to capture all this variation. We generate n independent and identically distributed random variables $\mathcal{Z}_1, \cdots, \mathcal{Z}_n$ from the standard normal distribution. Then the three sources of variation of $U(\beta_0; \gamma_0)$ can be expressed by the following functions of $\mathcal{Z}_1, \cdots, \mathcal{Z}_n$,

$$\begin{aligned} \Omega_1^* &= \sum_{i=1}^n \mathcal{Z}_i \int_0^\tau W(t)\{X_i - \bar{X}_i(t; \hat{\beta}, \hat{\gamma})\}\Delta_i(t)\Big[N_i(t)dH_i(t) - \exp\{X_i'(\hat{\beta} + \hat{\gamma})\} \\ &\quad \times \frac{\sum_{j=1}^n N_j(t)dH_j(t)\hat{\Phi}_j(t, X_i)}{\sum_{j=1}^n \exp\{X_j'(\hat{\beta} + \hat{\gamma})\}\hat{\Phi}_j(t, X_i)}\Big], \end{aligned}$$

$$\Omega_2^* = \sum_{i=1}^n \int_0^\tau W(t)\{X_i - \bar{X}_i(t;\hat{\beta},\hat{\gamma})\}\Delta_i(t)\exp\{X_i'(\hat{\beta}+\hat{\gamma})\}\Big[-\frac{\sum_{j=1}^n \mathcal{Z}_j N_j(t)dH_j(t)\hat{\Phi}_j(t,X_i)}{\sum_{j=1}^n \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j(t,X_i)}$$
$$+\frac{\sum_{j=1}^n N_j(t)dH_j(t)\hat{\Phi}_j(t,X_i)}{[\sum_{j=1}^n \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j(t,X_i)]^2}\sum_{j=1}^n \mathcal{Z}_j \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j(t,X_i)\Big].$$

Define

$$\hat{\eta}^* = \hat{\eta} + \hat{\Omega}^{-1}n^{-1}\sum_{i=1}^n \mathcal{Z}_i \int_0^\tau \{X_i - \bar{X}^D(t,\hat{\eta})\}d\hat{M}_i^D(t),$$

and

$$\hat{\Lambda}_0^*(t) = \hat{\Lambda}_0(t) + n^{-1}\sum_{i=1}^n \mathcal{Z}_i \int_0^t \frac{d\hat{M}_i^D(z)}{S^{(0)}(z,\hat{\eta})} - \int_0^t \bar{X}^D(z,\hat{\eta})'d\hat{\Lambda}_0(z)(\hat{\eta}^* - \hat{\eta}),$$

where $H_i^D(t) = I(T_i \le t, \delta_i = 1)$, $\hat{M}_i^D(t) = H_i^D(t) - \int_0^t \Delta_i(z)\exp(X_i'\hat{\eta})d\hat{\Lambda}_0(z)$, $S^{(0)}(t,\eta) = n^{-1}\sum_{i=1}^n \Delta_i(t)\exp(X_i'\eta)$, $S^{(1)}(t,\eta) = n^{-1}\sum_{i=1}^n \Delta_i(t)X_i\exp(X_i'\eta)$, $\bar{X}^D(t,\hat{\eta}) = S^{(1)}(t,\hat{\eta})/S^{(0)}(t,\hat{\eta})$, $S^{(2)}(t,\eta) = n^{-1}\sum_{i=1}^n \Delta_i(t)X_i^{\otimes 2}\exp(X_i'\eta)$, $\hat{\Omega} = n^{-1}\sum_{i=1}^n \int_0^\tau [\frac{S^{(2)}(t,\hat{\eta})}{S^{(0)}(t,\hat{\eta})} - \{\frac{S^{(1)}(t,\hat{\eta})}{S^{(0)}(t,\hat{\eta})}\}^{\otimes 2}]dH_i^D(t)$, and for a vector a, $a^{\otimes 2} = aa'$. Furthermore, the pure variation due to $(\hat{\eta},\hat{\Lambda}_0)$ is characterized by

$$\Omega_3^* = \sum_{i=1}^n \int_0^\tau W(t)\{X_i - \bar{X}_i(t;\hat{\beta},\hat{\gamma})\}\Delta_i(t)\exp\{X_i'(\hat{\beta}+\hat{\gamma})\}$$
$$\times\left[\frac{\sum_{j=1}^n N_j(t)dH_j(t)\hat{\Phi}_j(t,X_i)}{\sum_{j=1}^n \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j(t,X_i)} - \frac{\sum_{j=1}^n N_j(t)dH_j(t)\hat{\Phi}_j^*(t,X_i)}{\sum_{j=1}^n \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j^*(t,X_i)}\right],$$

where $\hat{\Phi}_j^*(t,X_i)$ is defined the same way as $\hat{\Phi}_j(t,X_i)$ except that $(\hat{\eta},\hat{\Lambda}_0)$ is replaced with $(\hat{\eta}^*,\hat{\Lambda}_0^*)$. In a similar way, the variation of $\tilde{U}(\gamma_0)$ comes from $dH_i(t) - \exp(X_i'\gamma_0)d\bar{H}_i(t;\gamma_0)$, the empirical summations in the numerator and denominator of $\bar{H}_i(t;\gamma_0)$ and $(\hat{\eta},\hat{\Lambda}_0)$. Similarly, the variation of $\tilde{U}(\gamma_0)$ can be characterized by the following three terms,

$$\Omega_4^* = \sum_{i=1}^n \mathcal{Z}_i \int_0^\tau Q(t)\{X_i - \bar{X}_i^*(t;\hat{\gamma})\}\Delta_i(t)\{dH_i(t) - \exp(X_i'\hat{\gamma})d\bar{H}_i(t;\hat{\gamma})\},$$

$$\Omega_5^* = \sum_{i=1}^n \int_0^\tau Q(t)\{X_i - \bar{X}_i^*(t;\hat\gamma)\}\Delta_i(t)\exp(X_i'\hat\gamma)\Big[-\frac{\sum_{j=1}^n \mathcal{Z}_j dH_j(t)\hat\Phi_j(t,X_i)}{\sum_{j=1}^n \exp(X_j'\hat\gamma)\hat\Phi_j(t,X_i)}$$
$$+\frac{\sum_{j=1}^n dH_j(t)\hat\Phi_j(t,X_i)}{[\sum_{j=1}^n \exp(X_j'\hat\gamma)\hat\Phi_j(t,X_i)]^2}\sum_{j=1}^n \mathcal{Z}_j\exp(X_j'\hat\gamma)\hat\Phi_j(t,X_i)\Big],$$

and

$$\Omega_6^* = \sum_{i=1}^n \int_0^\tau Q(t)\{X_i - \bar{X}_i^*(t;\hat\gamma)\}\Delta_i(t)\exp(X_i'\hat\gamma)[d\bar{H}_i(t;\hat\gamma) - d\bar{H}_i^*(t;\hat\gamma)],$$

where $\bar{H}_i^*(t;\hat\gamma)$ is defined the same way as $\bar{H}_i(t;\hat\gamma)$ except that $(\hat\eta,\hat\Lambda_0)$ is replaced with $(\hat\eta^*,\hat\Lambda_0^*)$. Define $\hat\Upsilon = (\hat\Upsilon_1',\hat\Upsilon_2')'$, where $\hat\Upsilon_1 = n^{-1/2}(\Omega_1^* + \Omega_2^* + \Omega_3^*)$ and $\hat\Upsilon_2 = n^{-1/2}(\Omega_4^* + \Omega_5^* + \Omega_6^*)$. Given the observed data $\{N_i(t)dH_i(t), T_i, \delta_i, X_i, H_i(t)\}$, we can estimate Σ by the empirical covariance matrix of $\hat\Upsilon$ with the help of repeating generation of the random samples $(\mathcal{Z}_1,\cdots,\mathcal{Z}_n)$. The following theorem justifies the Monte Carlo method. The proof is given in the Appendix.

Theorem 2. *Let $E_{\mathcal{Z}}$ denotes the conditional expectation with respect to $\mathcal{Z}_1,\cdots,\mathcal{Z}_n$ given the observed data. Then $E_{\mathcal{Z}}(\hat\Upsilon^{\otimes 2}) \xrightarrow{P} \Sigma$, where $\xrightarrow{P}$ denotes convergence in probability.*

4 Model Diagnostics

In this section, we will propose some graphical and numerical procedures for checking the adequacy of model (1). Following Lin et al. (2000), we propose the following cumulative sums of residual,

$$\mathcal{F}(t,x) = n^{-1/2}\sum_{i=1}^n \int_0^t I(X_i \le x)d\hat{M}_i^*(z), \tag{8}$$

where $I(X_i \le x)$ means that each of the components of X_i is no larger than the corresponding component of x, and

$$d\hat{M}_i^*(t) = \Delta_i(t)\Big[N_i(t)dH_i(t) - \exp\{X_i'(\hat{\beta}+\hat{\gamma})\} \times \frac{\sum_{j=1}^n N_j(t)dH_j(t)\hat{\Phi}_j(t,X_i)}{\sum_{j=1}^n \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j(t,X_i)}\Big].$$

Here the null hypothesis H_0 is defined as the correct specification of model (1). Similarly to $U(\beta_0;\gamma_0)$, the variation of $\mathcal{F}(t,x)$ can be characterized by the following three terms:

$$\begin{aligned}\Omega_7^*(t,x) &= \sum_{i=1}^n \mathcal{Z}_i \int_0^t I(X_i \le x)\Delta_i(z)\Big[N_i(z)dH_i(z) - \exp\{X_i'(\hat{\beta}+\hat{\gamma})\} \\ &\quad \times \frac{\sum_{j=1}^n N_j(z)dH_j(z)\hat{\Phi}_j(z,X_i)}{\sum_{j=1}^n \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j(z,X_i)}\Big].\end{aligned}$$

$$\begin{aligned}\Omega_8^*(t,x) &= \sum_{i=1}^n \int_0^t I(X_i \le x)\Delta_i(z)\exp\{X_i'(\hat{\beta}+\hat{\gamma})\}\Big[-\frac{\sum_{j=1}^n \mathcal{Z}_j N_j(z)dH_j(z)\hat{\Phi}_j(z,X_i)}{\sum_{j=1}^n \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j(z,X_i)} \\ &\quad + \frac{\sum_{j=1}^n N_j(z)dH_j(z)\hat{\Phi}_j(z,X_i)}{[\sum_{j=1}^n \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j(z,X_i)]^2}\sum_{j=1}^n \mathcal{Z}_j\exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j(z,X_i)\Big].\end{aligned}$$

$$\begin{aligned}\Omega_9^*(t,x) &= \sum_{i=1}^n \int_0^t I(X_i \le x)\Delta_i(z)\exp\{X_i'(\hat{\beta}+\hat{\gamma})\}\Big[\frac{\sum_{j=1}^n N_j(z)dH_j(z)\hat{\Phi}_j(z,X_i)}{\sum_{j=1}^n \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j(z,X_i)} \\ &\quad - \frac{\sum_{j=1}^n N_j(z)dH_j(z)\hat{\Phi}_j^*(z,X_i)}{\sum_{j=1}^n \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j^*(z,X_i)}\Big].\end{aligned}$$

Define

$$\hat{\Gamma}_1(t,x) = n^{-1}\sum_{i=1}^{n}\int_0^t I(X_i \le x) \triangle_i (z)\{X_i - \bar{X}_i(z;\hat{\beta},\hat{\gamma})\} \exp\{X_i'(\hat{\beta}+\hat{\gamma})\} \times \frac{\sum_{j=1}^{n} N_j(z) dH_j(z)\hat{\Phi}_j(z,X_i)}{\sum_{j=1}^{n} \exp\{X_j'(\hat{\beta}+\hat{\gamma})\}\hat{\Phi}_j(z,X_i)},$$

and $\hat{\Gamma}(t,x) = (\hat{\Gamma}_1(t,x)', \hat{\Gamma}_1(t,x)')'$. Then the null distribution of $\mathcal{F}(t,x)$ can be obtained from the following theorem with the proof presented in the Appendix.

Theorem 3. *Suppose that the conditions in Theorem 1 hold, then under H_0, the null distribution of $\mathcal{F}(t,x)$ can be approximated by the following zero-mean Gaussian process,*

$$\hat{\mathcal{F}}(t,x) = n^{-1/2}\{\Omega_7^*(t,x) + \Omega_8^*(t,x) + \Omega_9^*(t,x) - \hat{\Gamma}(t,x)'\hat{A}^{-1}\hat{\Upsilon}\}. \tag{9}$$

Based on Theorem 3, it is easy to see that we can obtain a large number of realization of $\mathcal{F}(t,x)$ by repeatedly generating the standard normal random sample $(\mathcal{Z}_1, \cdots, \mathcal{Z}_n)$ while fixing the observed data. To assess the adequacy of model (1), we can plot these realizations of $\hat{\mathcal{F}}(t,x)$ along with the observed $\mathcal{F}(t,x)$ and examine any unusual pattern of $\hat{\mathcal{F}}(t,x)$ compared with the $\mathcal{F}(t,x)$. More formally, we can apply the supremum test statistic $\sup_{t,x}|\mathcal{F}(t,x)|$ to conduct the lack-of-fit test, where the p-value can be obtained by comparing the observed value of $\sup_{t,x}|\mathcal{F}(t,x)|$ to a large number of realization from $\sup_{t,x}|\hat{\mathcal{F}}(t,x)|$.

5 A Simulation Study

In this section we will present some results from an extensive simulation study. In the s-tudy, the covariate X_i was generated from a Bernoulli distribution with success probability 0.5. The terminal event time was generated through $\log(D_i/4) = -0.5X_i + \epsilon_i$, where ϵ_i was generated from the extreme-value distribution. And the censoring time C_i was taken as $\min(C_i^*, \tau)$, where C_i^* followed a uniform distribution over (2,10) and $\tau = 6$, which yielded

23% censoring for the terminal event. Let $u_i = \exp(\phi_1\epsilon_i/5)$, $v_i = \rho_i \exp(-\phi_2\epsilon_i/5)$ where ϕ_1=−1, 0 or 1, ϕ_2=−1, 0 or 1, and ρ_i followed a uniform distribution over $(0.5, 1.5)$. Given X_i, u_i, v_i and $T_i = \min(C_i, D_i)$, we generated the observation process from a nonhomogeneous Poisson process with intensity function

$$\lambda_i(t) = v_i \exp(\gamma_0 X_i) I(T_i \geq t).$$

The average number of observations per subject is about 3.

The recurrent event process $N_i(t)$ was generated from a Poisson process $N_i^*(t)$ with the following intensity function:

$$\lambda_i^*(t|X_i, u_i, v_i, \omega_i) = \omega_i u_i \exp(\beta_0 X_i),$$

where ω_i was an independent gamma random variable with mean 1 and variance σ^2. Specifically, let $(t_{i,1}, \cdots, t_{i,K_i})$ be the observation times for the ith subject, then $N_i(t_{i,j}), j = 1, \cdots, K_i$ were generated piecewisely by generating $N_i(t_{i,j}) - N_i(t_{i,j-1})$ from a Poisson distribution with the mean functions $\omega_i(t_{i,j} - t_{i,j-1})u_i \exp(\beta_0 X_i)$, where $t_{i,0}$ was set to be 0. It is easy to verify that $N_i(t) = N_i^*(t \wedge D)$ and $N_i(t)$ satisfies (1). Note that ϕ_1 and ϕ_2 reflected the dependence among the recurrent event process, the observation times and the terminal event. For example, $\phi_1 = 0$ and $\phi_2 = 0$ implied that the recurrent event process, the observation times and the terminal event were independent, while $\phi_1 \neq 0$ and $\phi_2 \neq 0$ reflected that the three processes were related with each other. In the simulation study, we set $\beta_0 = 0.5$ and -0.5, $\gamma_0 = 0.5$, $W(t) = Q(t) = 1$. We found that 100 resamplings is enough for the variance estimation. All the results reported below were computed based on 1000 replications with sample sizes $n = 100$ and $n = 200$.

Tables 1 and 2 report the simulation results on the estimates for β_0 and γ_0, where include the bias (BIAS) given by the difference of sample means of estimator and the true value, the sampling standard errors (SSE), the sampling means of the estimated standard errors(SEE), and the 95% empirical coverage probabilities(CP). It can be seen from the table that the estimators seem to be unbiased and the proposed variance estimation seems to work well. Also the coverage probabilities are reasonable and consistent with the normal levels.

For comparison, we also considered the original method of Sun and Wei (2000) (denoted by SW). Here we generated the data in the same way as before except that we consider the parameter $\eta = 0$ and 0.5. Hence, when $\eta = 0$ and $(\phi_1, \phi_2) = (0, 0)$, the model of Sun and Wei (2000) was satisfied. However, it is easy to find that the SW estimator is consistent when $\eta = 0$, whatever values of (ϕ_1, ϕ_2). Table 3 gives the comparison results for estimation of β_0, As expected, the SW's method yielded consistent estimators when $\eta = 0$, with smaller bias and larger variance. However, when $\eta = 0.5$, for which the model of SW was violated, the SW's method results in biased estimates. Thus, the proposed method seems to be much more efficient and reliable than the SW method.

6 An Application

In this section, we applied the proposed methods to the bladder cancer data that have been discussed by Sun and Wei (2000) among others. This study was conducted by the Veterans Administration Cooperative Urological Research Group. The patients were randomly assigned to placebo and thiotepa treatment groups at the beginning of the study. The tumors were removed at the patients' clinic visit, but may recurrent again. For each patient, the number of initial tumors before entering the study and the size of the largest initial tumor were measured as baseline covariates. In the following, the observed information includes the clinical visit times (in month) and the number of bladder tumors that occurred between clinical visits, where 85 bladder cancer patients were included in the study, among those, 47 in the placebo group and 38 in the thiotepa treatment group. The follow up of 22 patients were terminated by death, 12 in the treatment group and 10 in the placebo group. Here we focus on the effects of thiotepa treatment and the number of initial tumors on the panel count of cumulative number of tumors in the presence of both informative observation times and a dependent terminal event.

For the analysis, we define $N_i(t)$ as the cumulative number of observed tumors at time t, $i = 1, ..., 85$. Let $X_{i1} = 1$ if the patient was in the thiotepa group and 0 if the patient was in the placebo group, and X_{i2} to be the logarithm of the number of the initial tumors plus 1. Let τ be the longest observation time being 53 months. The application of the proposed method in Section 3 with $W(t) = Q(t) = 1$ yielded $\hat{\beta}_1 = -1.5594$ and $\hat{\beta}_2 = 1.2991$ with the estimated standard errors of 0.3817 and 0.3456, respectively. These results imply that both the thiotepa treatment and initial number of tumors have significant effects on the tumor occurrence process. Specifically, the thiotepa treatment significantly reduced the bladder tumor occurrence rate, and the patients with higher number of initial tumors tend to have a higher tumor occurrence rate. These results are consistent with Sun and Wei (2000).

To assess the adequacy of the proposed model for the bladder cancer data, we applied the model checking techniques presented in Section 4. We calculated the statistic $\mathcal{F}(t, x)$, and obtained $\sup_{x,t} |\mathcal{F}(x, t)| = 19.3408$ with p-value of 0.842 based on 1000 realizations of the corresponding statistic $\sup_{x,t} |\hat{\mathcal{F}}(x, t)|$, which indicates that the proposed model fits the bladder cancer data well.

7 Concluding Remarks

This paper discussed the analysis of panel count data in the presence of both informative observation times and a terminal event. A joint model was proposed to describe the recurrent event process, the observation times and the terminal event together via two latent variables. An estimating equation-based inference procedure was proposed for the estimation of parameters. Also a goodness-of-fit procedure was presented for assessing the appropriateness of the proposed models and a simulation study was conducted and indicated that the estimation procedure works well in practical situations. In addition, an illustrative example was given.

There remain several topics to study in the future. First note that we only considered the time-independent covariates. In practice, they may be varying with time and thus it is desirable to extend the proposed procedure to the situation with time-dependent covariates. However, this is clearly not straightforward. For any regression problem, variable selection is always an important issue and it is apparent that this is true too for the problem discussed above. In other words, it would be useful to develop some procedures in this aspect for the proposed joint model. For this, one possible way is to use non-concave penalized estimating function approach (Tong, et al., 2009) based on (7) and it is apparent that a lot of research efforts are needed for it.

Acknowledgments

This research was partly supported by the National Natural Science Foundation of China Grants (Nos. 11301355, 11231010, 11171330 and 11301212), China Postdoctoral Science Foundation (No. 2014M550861), and Key Laboratory of RCSDS, CAS (No. 2008DP173182).

Appendix: Derivation of (4),(5) and proofs of Theorems 1, 2 and 3

To obtain the asymptotic distribution of $\hat{\beta}$ and $\hat{\gamma}$, we need the following regularity conditions:

C.1. $\{N_i(\cdot), H_i(\cdot), T_i, \delta_i, X_i\}$, $i = 1, \cdots, n$ are independent and identically distributed.

C.2. $H(\tau)$ and X are bounded almost surely, $N(t)$ is of bounded variation and $P(T \geq \tau) > 0$.

C.3. A is nonsingular, where

$$\begin{aligned}
A &= \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \\
A_{11} &= E\left[\int_0^\tau W(t)\{X_i - \bar{x}(t, X_i)\}^{\otimes 2}\Delta_i(t)\exp\{X_i'(\beta_0+\gamma_0)\}dR(t, \log\Lambda_0(t) + X_i'\eta_0)\right], \\
A_{22} &= E\left[\int_0^\tau Q(t)\{X_i - \bar{x}_i^*(t, X_i)\}^{\otimes 2}\Delta_i(t)\exp(X_i'\gamma_0)d\bar{H}_i(t;\gamma)\right], \\
A_{12} &= A_{11},
\end{aligned}$$

where $\bar{x}(t, X_i)$ and $\bar{x}^*(t, X_i)$ are the limits of $\bar{X}_i(t; \beta_0, \gamma_0)$ and $\bar{X}_i^*(t; \gamma_0)$ conditional on X_i, respectively.

Derivation of (4) and (5): Given (u, v, X) and $D \geq t$, we suppose that the mean function of the recurrent process and the rate of the observation process are independent of D. Since $\mathcal{A}_0(t; u, v) = \int_0^t \mu_N(z; u) d\mu_H(z; v)$. Then from the independent censoring assumption and models (1) and (2), we obtain that

$$\begin{aligned} E\{N(t)dH(t)|X, T \geq t\} &= E\{N(t)dH(t)|X, D \geq t\} \\ &= E[E\{N(t)dH(t)|X, D \geq t, u, v\}|X, D \geq t] \\ &= \exp\{X'(\beta_0 + \gamma_0)\}E\{d\mathcal{A}_0(t; u, v)|X, D \geq t\} \\ &= \exp\{X'(\beta_0 + \gamma_0)\}E\{d\mathcal{A}_0(t; u, v)|X, \epsilon \geq \log \Lambda_0(t) + X'\eta_0\}, \end{aligned}$$

where the last equality is from (3). Define $dR(t, s) = E\{d\mathcal{A}_0(t; u, v)|\epsilon \geq s\}$, then

$$E\{N(t)dH(t)|X, T \geq t\} = \exp\{X'(\beta_0 + \gamma_0)\}dR(t, \log \Lambda_0(t) + X'\eta_0). \tag{10}$$

For any integrable function $g(X, t, s)$, by the assumption that (u, v, ϵ) is independent of (X, C), we obtain that

$$\begin{aligned} dR(t, s) &= E\{d\mathcal{A}_0(t; u, v)|\epsilon \geq s\} \\ &= \frac{E\{d\mathcal{A}_0(t; u, v)I(\epsilon \geq s)\}}{E\{I(\epsilon \geq s)\}} \\ &= \frac{E[d\mathcal{A}_0(t; u, v)I(\epsilon \geq s)I\{\log \Lambda_0(C) + X'\eta_0 \geq s\}g(X, t, s)]}{E[I(\epsilon \geq s)I\{\log \Lambda_0(C) + X'\eta_0 \geq s\}g(X, t, s)]} \\ &= \frac{E[d\mathcal{A}_0(t; u, v)I\{\log \Lambda_0(T) + X'\eta_0 \geq s\}g(X, t, s)]}{E[I\{\log \Lambda_0(T) + X'\eta_0 \geq s\}g(X, t, s)]}. \end{aligned} \tag{11}$$

In particular, we choose $g(X, t, s) = \exp\{X'(\beta_0 + \gamma_0)\}I\{\log \Lambda_0(t) + X'\eta_0 \leq s\}$, then the denominator of (11) becomes $E[\exp\{X'(\beta_0 + \gamma_0)\}\Phi(T, X, t, s)]$, where $\Phi(T, X, t, s) = I\{\log \Lambda_0(T) + X'\eta_0 \geq s \geq \log \Lambda_0(t) + X'\eta_0\}$. Note that $\Phi(T, X, t, s) = 1$ implies $T \geq t$.

Using a similar method as (10), we get

$$
\begin{aligned}
&E\{N(t)dH(t)\Phi(T,X,t,s)\} = E[E\{N(t)dH(t)\Phi(T,X,t,s)|X,\Phi(T,X,t,s),u,v\}] \\
&= E(E[\exp\{X'(\beta_0+\gamma_0)\}d\mathcal{A}_0(t;u,v)\Phi(T,X,t,s)|X,\Phi(T,X,t,s),u,v]) \\
&= E[\exp\{X'(\beta_0+\gamma_0)\}d\mathcal{A}_0(t;u,v)\Phi(T,X,t,s)],
\end{aligned}
$$

which is the numerator of (11). Thus

$$
dR(t,s) = \frac{E[N(t)dH(t)I\{\log\Lambda_0(T)+X'\eta_0 \geq s \geq \log\Lambda_0(t)+X'\eta_0\}]}{E[\exp\{X'(\beta_0+\gamma_0)\}I\{\log\Lambda_0(T)+X'\eta_0 \geq s \geq \log\Lambda_0(t)+X'\eta_0\}]}.
$$

This completes the derivation. □

Proof of Theorem 1. Define

$$
\begin{aligned}
\Phi_i(t,X;\eta,\Lambda) &= I\{\log\Lambda(T_i)+X_i'\eta \geq \log\Lambda(t)+X'\eta \geq \log\Lambda(t)+X_i'\eta\}, \\
dM_i(t) &= N_i(t)dH_i(t), \\
d\bar{M}(t,X;\eta,\Lambda) &= \frac{\sum_{j=1}^{n}\Phi_j(t,X;\eta,\Lambda)dM_j(t)}{\sum_{j=1}^{n}\exp\{X_j'(\beta_0+\gamma_0)\}\Phi_j(t,X;\eta,\Lambda)},
\end{aligned}
$$

$$
d\bar{M}_0(t,X;\eta,\Lambda) = \frac{E\{\Phi_j(t,X;\eta,\Lambda)dM_j(t)|X\}}{E[\exp\{X_j'(\beta_0+\gamma_0)\}\Phi_j(t,X;\eta,\Lambda)|X]}
$$

and

$$
\bar{x}(t,X;\eta,\Lambda) = \frac{E[X_j\exp\{X_j'(\beta_0+\gamma_0)\}\Phi_j(t,X;\eta,\Lambda)|X]}{E[\exp\{X_j'(\beta_0+\gamma_0)\}\Phi_j(t,X;\eta,\Lambda)|X]}.
$$

Denote $\bar{x}(t,X)=\bar{x}(t,X;\eta_0,\Lambda_0)$ and $\Phi_i(t,X)=\Phi_i(t,X;\eta_0,\Lambda_0)$. Then we have

$$
\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau} W(t)\{X_i-\bar{x}(t,X_i)\}\Delta_i(t)\exp\{X_i'(\beta_0+\gamma_0)\}\{d\bar{M}(t,X_i;\hat{\eta},\hat{\Lambda}_0)-d\bar{M}_0(t,X_i;\hat{\eta},\hat{\Lambda}_0)\} \\
&= n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau} W(t)\left[\int\{x-\bar{x}(t,x)\}I(c\geq t)\frac{\exp\{x'(\beta_0+\gamma_0)\}\Phi_i(t,x)}{E[\exp\{X_i'(\beta_0+\gamma_0)\}\Phi_i(t,x)]}dF(x,c)\right]dM_i(t) \\
& -n^{-1/2}\sum_{i=1}^{n}\int\Big[\int_0^{\tau} W(t)\{x-\bar{x}(t,x)\}I(c\geq t)\exp\{x'(\beta_0+\gamma_0)\}\Phi_i(t,x) \\
& \times\frac{E\{\Phi_i(t,x)dM_i(t)\}}{(E[\exp\{X_i'(\beta_0+\gamma_0)\}\Phi_i(t,x)])^2}\Big]\exp\{X_i'(\beta_0+\gamma_0)\}dF(x,c)+o_p(1), \qquad (12)
\end{aligned}
$$

where $F(x,c)$ is the joint probability measure of (X_i,T_i).

Furthermore, according to Fleming and Harrington (1991, Page 299), we obtain that

$$\hat{\eta}-\eta_0=\Omega^{-1}n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\{X_i-\bar{x}^D(t)\}dM_i^D(t)+o_p(n^{-1/2}),$$

$$\hat{\Lambda}_0(t)-\Lambda_0(t)=n^{-1}\sum_{i=1}^{n}\int_0^{t}\frac{dM_i^D(z)}{s^{(0)}(z;\eta_0)}-\int_0^{t}\bar{x}^D(z)'d\Lambda_0(z)(\hat{\eta}-\eta_0)+o_p(n^{-1/2}),$$

where $M_i^D(t)=H_i^D(t)-\int_0^t\Delta_i(z)\exp(X_i'\eta_0)d\Lambda_0(z)$, and Ω, $s^{(0)}(t;\eta_0)$ and $\bar{x}^D(t)$ are the limits of $\hat{\Omega}$, $S^{(0)}(t;\eta_0)$ and $\bar{X}^D(t;\eta_0)$, respectively. Denote $dR_\eta(t,X)$ and $dR_\Lambda(t,X)$ as the derivative and the Hadamard derivative of $d\bar{M}_0(t,X;\eta_0,\Lambda_0)$ with respect to η and Λ, respectively. Then by the functional delta method, we get

$$
\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau} W(t)\{X_i-\bar{x}(t,X_i)\}\Delta_i(t)\exp\{X_i'(\beta_0+\gamma_0)\}\{d\bar{M}_0(t,X_i;\hat{\eta},\hat{\Lambda}_0)-d\bar{M}_0(t,X_i;\eta_0,\Lambda_0)\} \\
&= n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}\left[P_1\Omega^{-1}\{X_i-\bar{x}^D(t)\}+\frac{B_1(t)}{s^{(0)}(t;\eta_0)}\right]dM_i^D(t)+o_p(1), \qquad (13)
\end{aligned}
$$

where

$$
\begin{aligned}
P_1 \;=\; & E\Big[\int_0^{\tau} W(t)\{X_i-\bar{x}(t,X_i)\}\Delta_i(t)\exp\{X_i'(\beta_0+\gamma_0)\}\Big\{dR_\eta(t,X_i) \\
& -\Big(\int_0^{t}\bar{x}^D(z)'d\Lambda_0(z)\Big)dR_\Lambda(t,X_i)\Big\}\Big],
\end{aligned}
$$

$$B_1(t) = E\left[\int_t^\tau W(z)\{X_i - \bar{x}(z, X_i)\}\Delta_i(z)\exp\{X_i'(\beta_0+\gamma_0)\}dR_\Lambda(z, X_i)\right].$$

Then it follows from (12) and (13) that

$$\begin{aligned} n^{-1/2}U(\beta_0;\gamma_0) &= n^{-1/2}\sum_{i=1}^n \int_0^\tau W(t)\{X_i - \bar{X}_i(t;\beta_0,\gamma_0)\}\Delta_i(t)\Big[dM_i(t) - \exp\{X_i'(\beta_0+\gamma_0)\} \\ &\quad \times d\bar{M}_0(t, X_i;\eta_0,\Lambda_0) - \exp\{X_i'(\beta_0+\gamma_0)\}\{d\bar{M}(t, X_i;\hat{\eta},\hat{\Lambda}_0) - d\bar{M}_0(t, X_i;\hat{\eta},\hat{\Lambda}_0)\} \\ &\quad - \exp\{X_i'(\beta_0+\gamma_0)\}\{d\bar{M}_0(t, X_i;\hat{\eta},\hat{\Lambda}_0) - d\bar{M}_0(t, X_i;\eta_0,\Lambda_0)\}\Big] \\ &= n^{-1/2}\sum_{i=1}^n \xi_i + o_p(1), \end{aligned} \tag{14}$$

where

$$\begin{aligned} \xi_i &= \int_0^\tau W(t)\{X_i - \bar{x}(t, X_i)\}\Delta_i(t)[dM_i(t) - \exp\{X_i'(\beta_0+\gamma_0)\}d\bar{M}_0(t, X_i;\eta_0,\Lambda_0)] \\ &\quad - \int_0^\tau W(t)\left[\int \{x - \bar{x}(t,x)\}I(c\geq t)\frac{\exp\{x'(\beta_0+\gamma_0)\}\Phi_i(t,x)}{E[\exp\{X_i'(\beta_0+\gamma_0)\}\Phi_i(t,x)]}dF(x,c)\right]dM_i(t) \\ &\quad + \int \Big[\int_0^\tau W(t)\{x - \bar{x}(t,x)\}I(c\geq t)\exp\{x'(\beta_0+\gamma_0)\}\Phi_i(t,x) \\ &\quad \times \frac{E\{\Phi_i(t,x)dM_i(t)\}}{(E[\exp\{X_i'(\beta_0+\gamma_0)\}\Phi_i(t,x)])^2}\Big]\exp\{X_i'(\beta_0+\gamma_0)\}dF(x,c) \\ &\quad - \int_0^\tau \left[P_1\Omega^{-1}\{X_i - \bar{x}^D(t)\} + \frac{B_1(t)}{s^{(0)}(t;\eta_0)}\right]dM_i^D(t). \end{aligned}$$

Moreover, we define the following notations:

$$d\bar{H}(t, X;\eta,\Lambda) = \frac{\sum_{j=1}^n dH_j(t)\Phi_j(t, X;\eta,\Lambda)}{\sum_{j=1}^n \exp(X_j'\gamma_0)\Phi_j(t, X;\eta,\Lambda)},$$

$$d\bar{H}_0(t, X;\eta,\Lambda) = \frac{E[dH_j(t)\Phi_j(t, X;\eta,\Lambda)|X]}{E[\exp(X_j'\gamma_0)\Phi_j(t, X;\eta,\Lambda)|X]},$$

$$\bar{x}^*(t, X;\eta,\Lambda) = \frac{E[X_j\exp(X_j'\gamma_0)\Phi_j(t, X;\eta,\Lambda)|X]}{E[\exp(X_j'\gamma_0)\Phi_j(t, X;\eta,\Lambda)|X]}.$$

Denote $\bar{x}^*(t,X) = \bar{x}^*(t,X;\eta_0,\Lambda_0)$. In a similar manner, we can get

$$
\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\int_0^\tau Q(t)\{X_i - \bar{x}^*(t,X_i)\}\Delta_i(t)\exp(X_i'\gamma_0)\{d\bar{H}(t,X;\hat{\eta},\hat{\Lambda}_0) - d\bar{H}_0(t,X;\hat{\eta},\hat{\Lambda}_0)\} \\
& = n^{-1/2}\sum_{i=1}^{n}\int_0^\tau Q(t)\left[\int \{x - \bar{x}^*(t,x)\}I(c\geq t)\frac{\exp(x'\gamma_0)\Phi_i(t,x)}{E[\exp(X_i'\gamma_0)\Phi_i(t,x)]}dF(x,c)\right]dH_i(t) \\
& -n^{-1/2}\sum_{i=1}^{n}\int\Big[\int_0^\tau Q(t)\{x - \bar{x}^*(t,x)\}I(c\geq t)\exp(x'\gamma_0)\Phi_i(t,x) \\
& \times\frac{E\{\Phi_i(t,x)dH_i(t)\}}{(E[\exp(X_i'\gamma_0)\Phi_i(t,x)])^2}\Big]\exp(X_i'\gamma_0)dF(x,c) + o_p(1), \qquad (15)
\end{aligned}
$$

and

$$
\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\int_0^\tau Q(t)\{X_i - \bar{x}^*(t,X_i)\}\Delta_i(t)\exp(X_i'\gamma_0)\{d\bar{H}_0(t,X_i;\hat{\eta},\hat{\Lambda}_0) - d\bar{H}_0(t,X_i;\eta_0,\Lambda_0)\} \\
& = n^{-1/2}\sum_{i=1}^{n}\int_0^\tau \left[P_2\Omega^{-1}\{X_i - \bar{x}^D(t)\} + \frac{B_2(t)}{s^{(0)}(t;\eta_0)}\right]dM_i^D(t) + o_p(1), \qquad (16)
\end{aligned}
$$

where

$$
\begin{aligned}
P_2 &= E\Big[\int_0^\tau Q(t)\{X_i - \bar{x}^*(t,X_i)\}\Delta_i(t)\exp(X_i'\gamma_0)\Big\{dR_\eta^*(t,X_i) \\
&\quad -\left(\int_0^t \bar{x}^D(z)'d\Lambda_0(z)\right)dR_\Lambda^*(t,X_i)\Big\}\Big],
\end{aligned}
$$

$$
B_2(t) = E\left[\int_t^\tau Q(z)\{X_i - \bar{x}^*(z,X_i)\}\Delta_i(z)\exp(X_i'\gamma_0)dR_\Lambda^*(z,X_i)\right],
$$

and $dR_\eta^*(t,X)$ and $dR_\Lambda^*(t,X)$ as the derivative and the Hadamard derivative of $d\bar{H}_0(t,X;\eta_0,\Lambda_0)$ with respectively to η and Λ, respectively.

Then it follows from (15) and (16), we obtain that

$$
n^{-1/2}\tilde{U}(\gamma_0) = n^{-1/2}\sum_{i=1}^{n}\zeta_i + o_p(1), \qquad (17)
$$

where

$$
\begin{aligned}
\zeta_i &= \int_0^\tau Q(t)\{X_i - \bar{x}^*(t,X_i)\}\Delta_i(t)\exp(X_i'\gamma_0)[dH_i(t) - \exp(X_i'\gamma_0)d\bar{H}_0(t,X_i;\eta_0,\Lambda_0)] \\
&\quad - \int_0^\tau Q(t)\left[\int \{x - \bar{x}^*(t,x)\}I(c \ge t)\frac{\exp(x'\gamma_0)\Phi_i(t,x)}{E[\exp(X_i'\gamma_0)\Phi_i(t,x)]}dF(x,c)\right]dH_i(t) \\
&\quad + \int \Big[\int_0^\tau Q(t)\{x - \bar{x}^*(t,x)\}I(c \ge t)\exp(x'\gamma_0)\Phi_i(t,x) \\
&\quad \times \frac{E\{\Phi_i(t,x)dH_i(t)\}}{(E[\exp(X_i'\gamma_0)\Phi_i(t,x)])^2}\Big]\exp(X_i'\gamma_0)dF(x,c) \\
&\quad - \int_0^\tau \left[P_2\Omega^{-1}\{X_i - \bar{x}^D(t)\} + \frac{B_2(t)}{s^{(0)}(t;\eta_0)}\right]dM_i^D(t).
\end{aligned}
$$

Furthermore, we notice that $-n^{-1}\partial U(\beta_0;\gamma_0)/\partial\beta$, $-n^{-1}\partial U(\beta_0;\gamma_0)/\partial\gamma$ and $-n^{-1}\partial\tilde{U}(\gamma_0)/\partial\gamma$ convergence in probability to A_{11}, A_{12} and A_{22}, respectively. Then since (14) and (17), together with the Taylor expansion and Slutsky theorem that

$$
\begin{aligned}
n^{1/2}\begin{pmatrix}\hat{\beta}-\beta_0 \\ \hat{\gamma}-\gamma_0\end{pmatrix} &= A^{-1}n^{-1/2}\begin{pmatrix}U(\beta_0;\gamma_0) \\ \tilde{U}(\gamma_0)\end{pmatrix} + o_p(1) \\
&= A^{-1}n^{-1/2}\sum_{i=1}^n \begin{pmatrix}\xi_i \\ \zeta_i\end{pmatrix} + o_p(1),
\end{aligned}
$$

thus

$$
n^{1/2}\begin{pmatrix}\hat{\beta}-\beta_0 \\ \hat{\gamma}-\gamma_0\end{pmatrix} \xrightarrow{\mathcal{L}} N(0, A^{-1}\Sigma(A^{-1})'),
$$

where $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution, $\Sigma = E\{(\xi_i', \zeta_i')'^{\otimes 2}\}$. This completes the proof. □

Proof of Theorem 2. Since $\hat{\beta}$, $\hat{\gamma}$, $\hat{\eta}$ and $\hat{\Lambda}_0(t)$ are consistent, we can obtain that conditional on the observed data,

$$
\begin{aligned}
n^{-1/2}\Omega_1^* = \sum_{i=1}^n \mathcal{Z}_i \int_0^\tau & W(t)\{X_i - \bar{x}(t,X_i)\}\Delta_i(t)[dM_i(t) \\
& - \exp\{X_i'(\beta_0+\gamma_0)\}d\bar{M}_0(t,X_i;\eta_0,\Lambda_0)] + o_p(1). \qquad (18)
\end{aligned}
$$

Similarly,

$$n^{-1/2}\Omega_2^* \;=\; -n^{-1/2}\sum_{i=1}^n \mathcal{Z}_i\int_0^\tau W(t)\left[\int\{x-\bar{x}(t,x)\}I(c\geq t)\frac{\exp\{x'(\beta_0+\gamma_0)\}\Phi_i(t,x)}{E[\exp\{X_i'(\beta_0+\gamma_0)\}\Phi_i(t,x)]}dF(x,c)\right]$$

$$\times dM_i(t)+n^{-1/2}\sum_{i=1}^n \mathcal{Z}_i\int\Big[\int_0^\tau W(t)\{x-\bar{x}(t,x)\}I(c\geq t)\exp\{x'(\beta_0+\gamma_0)\}\Phi_i(t,x)$$

$$\times\frac{E\{\Phi_i(t,x)dM_i(t)\}}{(E[\exp\{X_i'(\beta_0+\gamma_0)\}\Phi_i(t,x)])^2}\Big]\exp\{X_i'(\beta_0+\gamma_0)\}dF(x,c)+o_p(1). \tag{19}$$

Define

$$d\hat{M}(t,X;\eta,\Lambda)=\frac{\sum_{j=1}^n\Phi_j(t,X;\eta,\Lambda)dM_j(t)}{\sum_{j=1}^n\exp\{X_j'(\hat\beta+\hat\gamma)\}\Phi_j(t,X;\eta,\Lambda)}.$$

Thus,

$$\Omega_3^*=\sum_{i=1}^n\int_0^\tau W(t)\{X_i-\bar{X}_i(t;\hat\beta,\hat\gamma)\}\Delta_i(t)\exp\{X_i'(\hat\beta+\hat\gamma)\}\{d\hat{M}(t,X;\hat\eta,\hat\Lambda_0)-d\hat{M}(t,X;\hat\eta^*,\hat\Lambda_0^*)\}.$$

Moreover, we notice that

$$d\hat{M}(t,X;\hat\eta,\hat\Lambda_0)-d\hat{M}(t,X;\hat\eta^*,\hat\Lambda_0^*)=\{d\hat{M}(t,X;\hat\eta,\hat\Lambda_0)-d\bar{M}_0(t,X;\hat\eta,\hat\Lambda_0)\}$$
$$-\{d\hat{M}(t,X;\hat\eta^*,\hat\Lambda_0^*)-d\bar{M}_0(t,X;\hat\eta^*,\hat\Lambda_0^*)\}+\{d\bar{M}_0(t,X;\hat\eta,\hat\Lambda_0)-d\bar{M}_0(t,X;\hat\eta^*,\hat\Lambda_0^*)\}.$$

Using a similar method to the proof of (12), we get

$$n^{-1/2}\sum_{i=1}^n\int_0^\tau W(t)\{X_i-\bar{X}_i(t;\hat\beta,\hat\gamma)\}\Delta_i(t)\exp\{X_i'(\hat\beta+\hat\gamma)\}\Big[\{d\hat{M}(t,X;\hat\eta,\hat\Lambda_0)$$
$$-d\bar{M}_0(t,X;\hat\eta,\hat\Lambda_0)\}-\{d\hat{M}(t,X;\hat\eta^*,\hat\Lambda_0^*)-d\bar{M}_0(t,X;\hat\eta^*,\hat\Lambda_0^*)\}\Big]=o_p(1)$$

Similar to (13), we obtain that conditional on the observed data,

$$n^{-1/2}\sum_{i=1}^n\int_0^\tau W(t)\{X_i-\bar{X}_i(t;\hat\beta,\hat\gamma)\}\Delta_i(t)\exp\{X_i'(\hat\beta+\hat\gamma)\}\{d\bar{M}_0(t,X;\hat\eta^*,\hat\Lambda_0^*)-d\bar{M}_0(t,X;\hat\eta,\hat\Lambda_0)\}$$

$$=n^{-1/2}\sum_{i=1}^n\mathcal{Z}_i\int_0^\tau\left[P_1\Omega^{-1}\{X_i-\bar{x}^D(t)\}+\frac{B_1(t)}{s^{(0)}(t;\eta_0)}\right]dM_i^D(t)+o_p(1). \tag{20}$$

From (18), (19) and (20), we have that

$$\hat{\Upsilon}_1 = n^{-1/2}(\Omega_1^* + \Omega_2^* + \Omega_3^*) = n^{-1/2}\sum_{i=1}^{n} \mathcal{Z}_i \xi_i + o_p(1).$$

In a similar way, we obtain that

$$\hat{\Upsilon}_2 = n^{-1/2}(\Omega_4^* + \Omega_5^* + \Omega_6^*) = n^{-1/2}\sum_{i=1}^{n} \mathcal{Z}_i \zeta_i + o_p(1).$$

Thus, by the Theorem 3.6.13 of van der Vaart and Wellner (1996), $E_{\mathcal{Z}}(\hat{\Upsilon}^{\otimes 2}) \xrightarrow{P} \Sigma$. This completes the proof of this theorem. □

Proof of Theorem 3. We notice that

$$\begin{aligned}\mathcal{F}(t,x) &= n^{-1/2}\sum_{i=1}^{n}\int_0^t I(X_i \le x)\Delta_i(z)\Big[N_i(z)dH_i(z) - \exp\{X_i'(\beta_0+\hat{\gamma})\} \\ &\quad \times \frac{\sum_{j=1}^{n} N_j(z)dH_j(z)\hat{\Phi}_j(z,X_i)}{\sum_{j=1}^{n}\exp\{X_j'(\beta_0+\hat{\gamma})\}\hat{\Phi}_j(z,X_i)}\Big] - n^{1/2}\hat{\Gamma}_1(t,x)'(\hat{\beta}-\beta_0). \qquad (21)\end{aligned}$$

Using a similar method as the proof of Theorem 1, we can have that the first term on the right-hand side of (21) is

$$n^{-1/2}\sum_{i=1}^{n}\Psi_i(t,x) - n^{1/2}\Gamma_1(t,x)'(\hat{\gamma}-\gamma_0) + o_p(1), \qquad (22)$$

where

$$\begin{aligned}\Psi_i(t,x) &= \int_0^t I(X_i \le x)\Delta_i(z)[dM_i(z) - \exp\{X_i'(\beta_0+\gamma_0)\}d\bar{M}_0(t,X_i;\eta_0,\Lambda_0)] \\ &\quad - \int_0^t \left[\int I(s \le x)I(c \ge z)\frac{\exp\{s'(\beta_0+\gamma_0)\}\Phi_i(z,s)}{E[\exp\{X_i'(\beta_0+\gamma_0)\}\Phi_i(z,s)]}dF(s,c)\right]dM_i(z) \\ &\quad + \int\Big[\int_0^t I(s \le x)I(c \ge z)\exp\{s'(\beta_0+\gamma_0)\}\Phi_i(z,s) \\ &\quad \times \frac{E\{\Phi_i(z,s)dM_i(z)\}}{(E[\exp\{X_i'(\beta_0+\gamma_0)\}\Phi_i(z,s)])^2}\Big]\exp\{X_i'(\beta_0+\gamma_0)\}dF(s,c) \\ &\quad - P_1^*(t,x)\Omega^{-1}\int_0^\tau \left[\{X_i - \bar{x}^D(z)\} + \frac{B_1^*(t,z,x)}{s^{(0)}(z;\eta_0)}\right]dM_i^D(z),\end{aligned}$$

and

$$
\begin{aligned}
P_1^*(t,x) &= E\Big[\int_0^t I(X_i \le x)\Delta_i(s)\exp\{X_i'(\beta_0+\gamma_0)\}\Big\{dR_\eta(s,X_i) \\
&\quad - \Big(\int_0^s \bar{x}^D(z)' d\Lambda_0(z)\Big) dR_\Lambda(s,X_i)\Big\}\Big],
\end{aligned}
$$

$$
B_1^*(t,z,x) = E\left[\int_z^t I(X_i \le x)\Delta_i(s)\exp\{X_i'(\beta_0+\gamma_0)\} dR_\Lambda(s,X_i)\right],
$$

where $\Gamma_1(t,x)$ is the limit of $\hat{\Gamma}_1(t,x)$. Furthermore, we know that the second term on the right-hand side of (21) is equivalent to

$$
-n^{1/2}\Gamma_1(t,x)(\hat{\beta}-\beta_0) + o_p(1). \tag{23}
$$

Then, from (21),(22), (23) and Theorem 1, we obtain that

$$
\mathcal{F}(t,x) = n^{-1/2}\sum_{i=1}^n \Big\{\Psi_i(t,x) - \Gamma(t,x)'A^{-1}(\xi_i', \zeta_i')'\Big\} + o_p(1), \tag{24}
$$

where $\Gamma(t,x) = (\Gamma_1(t,x)', \Gamma_1(t,x)')'$. By the multivariate central limit theorem, we have that $\mathcal{F}(t,x)$ converges in distribution for finite dimensions. It is easy to see that $\mathcal{F}(t,x)$ is tight, then using a similar method as the proof of Theorem 2, we obtain that $\mathcal{F}(t,x)$ converges weakly to a zero-mean Gaussian process and the asymptotic distribution can be approximated by (9). This completes the proof. □

References

Balakrishnan, N. and Zhao, X. (2009). New multi-sample nonparametric tests for panel count data. *Annals of Statistics*, **37**, 1112-1149.

Balakrishnan, N. and Zhao, X. (2010). A nonparametric test for the equality of counting processes with panel count data. *Computational Statistics and Data Analysis*, **54**, 135-142.

Cook, R. J. and Lawless, J. F. (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in Medicine*, **16**, 911-924.

Fleming, T. and Harrington, D. (1991). *Counting processes and Survival Analysis.* NewYork: Wiley.

Ghosh, D. and Lin, D. Y. (2002). Marginal regression models for recurrent and terminal events. *Statistica Sinica*, **12**, 663-688.

Huang, C. Y. and Wang, M. C. (2004). Joint modeling and estimation of recurrent event processes and failure time. *Journal of the American Statistical Association*, **99**, 1153-1165.

Huang, C., Wang, M. and Zhang, Y. (2006). Analysing panel count data with informative observation times. *Biometrika*, **93**, 763-775.

Hu, X. , Sun, J. and Wei, L. J. (2003). Regression parameter estimation from panel counts. *Scandinavian Journal of Statistics*, **30**, 25-43.

Kalbfleisch, J. D. and Lawless, J. F. (1985). The analysis of panel data under a Markov assumption. *Journal of the American Statistical Association*, **80**, 863-871.

Li, N., Sun, L. and Sun, J. (2010). Semiparametric transformation models for panel count data with dependent observation process. *Statistics in Biosciences*, **2**, 191-210.

Liang, K. and Zeger, S. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, **73**, 13-22.

Lin, D. Y., Wei, L. J., Yang, I. and Ying, Z. (2000). Semiparametric regression for the mean and rate function of recurrent events. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **69**, 711-730.

Liu, L., Wolfe, R. A., and Huang, X. (2004). Shared frailty models for recurrent events and a terminal event. *Biometrics*, **60**, 747-756.

Park, D., Sun, J, Zhao, X. (2007). A class of two-sample nonparametric tests for panel count data. *Communications in Statistics - Theory and Methods*, **36**, 1611-1625.

Sun, J. (2006). *The statistical analysis of interval-censored failure time data*. Springer, New York.

Sun, J. and Fang, H. (2003). A nonparametric test for panel count data. *Biometrika*, **90**, 199-208.

Sun, J. and Kalbfleisch, J. D. (1995). Estimation of the mean function of point processes based on panel count data. *Statistica Sinica*, **5**, 279-290.

Sun, J., Tong, X. and He, X. (2007). Regression analysis of panel count data with dependent observation times. *Biometrics*, **63**, 1053-1059.

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

Sun, L., Song, X., Zhou, J. and Liu, L. (2012). Joint analysis of longitudinal data with informative observation times and a dependent terminal event. *Journal of the American Statistical Association*, **107**, 688-700.

Sun, J. and Zhao, X. (2013). *The Statistical Analysis of Panel Count Data*. New York: Springer.

Thall, P. F. and Lachin, J. M. (1988). Analysis of recurrent events: nonparametric methods for random-interval count data. *Journal of the American Statistical Association*, **83**, 339-347.

Tong, X., He, X., Sun, L. and Sun, J. (2009). Variable selection for panel count data via non-concave penalized estimating function. *Scandinavian Journal of Statistics*, **36**, 620-635.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. New York: Springer.

Wellner, J. A. and Zhang, Y. (2007). Two likelihood-based semiparametric estimation methods for panel count data with covariates. *Annals of Statistics*, **35**, 2106-2142.

Ye, Y., Kalbfleisch, J. and Schaubel, D. (2007). Semiparametric analysis of correlated recurrent and terminal events. *Biometrics*, **63**, 78-87.

Zeng, D. and Cai, J. (2010). A semiparametric additive rate model for recurrent events with an informative terminal event. *Biometrika*, **97**, 699-712.

Zhang, H., Sun, J. and Wang, D. (2013). Variable selection and estimation for multivariate panel count data via the seamless-L0 penalty. *The Canadian Journal of Statistics*, **41**, 368-385.

Zhang, H., Zhao, H., Sun, J., Wang, D. and Kim, K. (2013). Regression analysis of multivariate panel count data with an informative observation process. *Journal of Multivariate Analysis*, **119**, 71-80.

Zhang, Y. (2002). A semiparametric pseudolikelihood estimation method for panel count data. *Biometrika*, **89**, 39-48.

Zhao, X. and Sun, J. (2011). Nonparametric comparison for panel count data with unequal observation processes. *Biometrics*, **67**, 770-779.

Zhao, X., Zhou, J. and Sun, L. (2011). Semiparametric transformation models with time-varying coefficients for recurrent and terminal events. *Biometrics*, **67**, 404-414.

Table 1. Simulation results for estimation of β_0 and γ_0 when $\beta_0 = -0.5$

		β_0				γ_0			
n	(ϕ_1, ϕ_2)	BIAS	SSE	ESE	CP	BIAS	SSE	ESE	CP
100	(-1, -1)	-0.0102	0.3433	0.3265	0.928	0.0480	0.1708	0.1714	0.943
	(-1, 0)	-0.0145	0.3419	0.3235	0.929	0.0537	0.1668	0.1718	0.944
	(-1, 1)	-0.0182	0.3202	0.3250	0.947	0.0478	0.1723	0.1755	0.950
	(0, -1)	-0.0080	0.3544	0.3305	0.921	0.0475	0.1741	0.1729	0.947
	(0, 0)	0.0122	0.3437	0.3263	0.937	0.0383	0.1724	0.1702	0.947
	(0, 1)	0.0146	0.3426	0.3227	0.942	0.0380	0.1727	0.1741	0.947
	(1, -1)	0.0085	0.3589	0.3314	0.928	0.0554	0.1697	0.1714	0.945
	(1, 0)	0.0103	0.3448	0.3293	0.938	0.0430	0.1662	0.1694	0.952
	(1, 1)	0.0112	0.3470	0.3297	0.928	0.0320	0.1694	0.1732	0.955
200	(-1, -1)	-0.0099	0.2311	0.2310	0.946	0.0290	0.1146	0.1166	0.950
	(-1, 0)	0.0136	0.2378	0.2299	0.950	0.0212	0.1147	0.1134	0.938
	(-1, 1)	-0.0083	0.2346	0.2295	0.936	0.0185	0.1175	0.1181	0.946
	(0, -1)	-0.0126	0.2414	0.2317	0.936	0.0234	0.1177	0.1161	0.948
	(0, 0)	0.0120	0.2389	0.2307	0.941	0.0193	0.1046	0.1142	0.964
	(0, 1)	-0.0025	0.2402	0.2294	0.937	0.0206	0.1199	0.1177	0.946
	(1, -1)	0.0271	0.2466	0.2357	0.936	0.0320	0.1139	0.1163	0.939
	(1, 0)	0.0079	0.2411	0.2348	0.938	0.0259	0.1095	0.1142	0.959
	(1, 1)	0.0198	0.2457	0.2349	0.938	0.0265	0.1202	0.1181	0.948

Table 2. Simulation results for estimation of β_0 and γ_0 when $\beta_0 = 0.5$

		β_0				γ_0			
n	(ϕ_1, ϕ_2)	BIAS	SSE	ESE	CP	BIAS	SSE	ESE	CP
100	(-1, -1)	0.0720	0.3372	0.3381	0.947	0.0419	0.1682	0.1690	0.944
	(-1, 0)	0.0755	0.3484	0.3366	0.954	0.0336	0.1602	0.1684	0.951
	(-1, 1)	0.0803	0.3346	0.3398	0.955	0.0435	0.1736	0.1736	0.940
	(0, -1)	0.0729	0.3426	0.3335	0.944	0.0497	0.1640	0.1699	0.949
	(0, 0)	0.0825	0.3481	0.3357	0.945	0.0447	0.1581	0.1687	0.947
	(0, 1)	0.0714	0.3423	0.3400	0.962	0.0504	0.1737	0.1756	0.945
	(1, -1)	0.0871	0.3557	0.3502	0.952	0.0480	0.1721	0.1700	0.937
	(1, 0)	0.1074	0.3465	0.3473	0.947	0.0469	0.1637	0.1685	0.948
	(1, 1)	0.0803	0.3530	0.3406	0.942	0.0446	0.1708	0.1735	0.952
200	(-1, -1)	0.0487	0.2248	0.2263	0.943	0.0227	0.1176	0.1158	0.937
	(-1, 0)	0.0426	0.2164	0.2230	0.953	0.0206	0.1151	0.1141	0.945
	(-1, 1)	0.0373	0.2178	0.2232	0.951	0.0213	0.1163	0.1177	0.950
	(0, -1)	0.0342	0.2220	0.2243	0.947	0.0312	0.1141	0.1154	0.943
	(0, 0)	0.0289	0.2296	0.2250	0.928	0.0188	0.1097	0.1133	0.958
	(0, 1)	0.0311	0.2311	0.2217	0.937	0.0118	0.1138	0.1183	0.965
	(1, -1)	0.0390	0.2376	0.2325	0.936	0.0232	0.1121	0.1151	0.947
	(1, 0)	0.0427	0.2354	0.2283	0.944	0.0114	0.1090	0.1131	0.959
	(1, 1)	0.0424	0.2266	0.2285	0.951	0.0193	0.1189	0.1172	0.935

Table 3. Comparison results for estimation of $\beta_0 = 0.5$ when $n = 200$

		Ours		SW	
η_0	(ϕ_1, ϕ_2)	BIAS	SSE	BIAS	SSE
0	(-1 , -1)	0.0352	0.1998	0.0121	0.2795
	(-1 , 0)	0.0301	0.2126	0.0076	0.2928
	(-1 , 1)	0.0293	0.2027	0.0019	0.2705
	(0 , -1)	0.0310	0.2116	0.0051	0.2876
	(0 , 0)	0.0172	0.2132	-0.0031	0.2940
	(0 , 1)	0.0341	0.1982	0.0167	0.2818
	(1 , -1)	0.0311	0.2146	-0.0068	0.2985
	(1 , 0)	0.0307	0.2096	0.0275	0.2984
	(1 , 1)	0.0277	0.2042	0.0074	0.2922
0.5	(-1 , -1)	0.0436	0.2333	-0.5175	0.3028
	(-1 , 0)	0.0320	0.2293	-0.5313	0.2987
	(-1 , 1)	0.0407	0.2269	-0.5207	0.2912
	(0 , -1)	0.0598	0.2318	-0.4706	0.3199
	(0 , 0)	0.0358	0.2232	-0.5058	0.3103
	(0 , 1)	0.0412	0.2278	-0.4917	0.3054
	(1 , -1)	0.0449	0.2352	-0.4614	0.3249
	(1 , 0)	0.0381	0.2277	-0.4573	0.3265
	(1 , 1)	0.0411	0.2287	-0.4406	0.3230