Statistica Sinica **33** (2023), 2117-2136
doi:https://doi.org/10.5705/ss.202020.0359

WEIGHTED RANK ESTIMATION FOR RANDOM-EFFECTS MONOTONIC INDEX MODELS WITH PANEL COUNT DATA

Tianqing Liu[1], Xiaohui Yuan[2] and Jianguo Sun[3]

[1]*Jilin University,* [2]*Changchun University of Technology and* [3]*University of Missouri*

Abstract: Panel count data arise when study subjects who may experience certain recurrent events are only observed intermittently at discrete examination times. In addition to the underlying recurrent event process of interest, there usually exist two other nuisance processes, namely the observation and follow-up processes, which may be correlated with the recurrent event process of interest. We propose a general class of random-effects monotonic index models for regression analysis of such panel count data. In order to estimate the regression parameters, we develop a weighted rank (WR) estimation procedure and and establish the consistency and asymptotic normality of the resulting WR estimator. A numerical study and an application of the proposed methodology show that it works well in practice.

Key words and phrases: Informative follow-up time, informative observation process, panel count data, random-effects monotonic index models, weighted rank estimation.

1. Introduction

Panel count data occur when study subjects who may experience certain recurrent events are examined only at discrete time points, rather than continuously, owing to cost, feasibility, or other practical considerations (Kalbfleisch and Lawless (1985); Thall and Lachin (1988); Sun and Zhao (2013); Chiou et al. (2019)). Here, in addition to the underlying recurrent event process of interest, there usually exist two other nuisance processes, namely the observation and follow-up processes (Wellner and Zhang (2000); Lin et al. (2000); Zhang (2002); Cai and Schaubel (2004); Lu, Zhang and Huang (2007)). Furthermore, the latter processes may be correlated with the recurrent event process of interest, leading to so-called informative examination and follow-up times or processes. See Sun and Zhao (2013) and Chiou et al. (2019) for a comprehensive review of panel count data.

Corresponding author: Tianqing Liu, Center for Applied Statistical Research and School of Mathematics, Jilin University, Changchun, Jilin 130012, China. E-mail: tianqingliu@gmail.com.

Many statistical methods have been developed for the analysis of panel count data. In general, existing approaches can be classified broadly into three types: pseudo-likelihood estimations, the estimating equation approach, and nonparametric estimations. In pseudo-likelihood procedures, Cox-type models are commonly used, and the pseudo-likelihood function is constructed based on nonhomogeneous Poisson process assumptions. The likelihood and pseudo-likelihood methods are robust against departures from the Poisson assumption, as long as the proportional rates model holds (Zhang (2002); Wellner and Zhang (2000); Lu, Zhang and Huang (2007); Zhu et al. (2018)).

The estimating equation approaches are computationally convenient, but may be inefficient. Sun and Wei (2000) and Hu, Sun and Wei (2003) considered such approaches by modeling cumulative event counts at different time points, and Hua and Zhang (2012) improved the estimation efficiency of these approaches using generalized estimating equations.

Kernel smoothing and spline approximation are two popular nonparametric approaches to analyzing of panel count data. For example, Zhao, Tu and Yu (2018) investigated the B-splined pseudo-likelihood method for a time-varying coefficients model of panel count data. Wang and Yu (2021) employed kernel smoothing to study a time-varying coefficients panel count model under the assumption of a nonhomogeneous Poisson process. However, nonparametric estimations may be difficult for problems with high-dimensional covariates, owing to the well-known "curse of dimensionality".

When analyzing panel count data, the observation process and the follow-up process may be informative about the recurrent event process, even after conditioning on available covariates, which presents a challenge. For example, patients with higher cancer recurrence rates may have more frequent clinical examinations and a longer follow-up time, because they may require more medical assessments and attention (Li et al. (2011); Sun and Zhao (2013); Ma and Sundaram (2018)). Informative examination and follow-up times are often encountered in panel count data, and falsely treating them as noninformative could result in biased regression coefficient estimations and misleading conclusions.

One way of accounting for informative examination and follow-up times is to specify joint models for all three processes (Kim (2006); Sun, Tong and He (2007); He, Tong and Sun (2009); Buzkova (2010); Zhao and Tong (2011); Zhou et al. (2017); Ma and Sundaram (2018); Jiang, Su, and Zhao (2020)). For example, Huang, Wang and Zhang (2006) and Wang, Ma and Yan (2013) postulated frailty proportional rates models for recurrent event processes, leaving the distributions of the frailty and the possibly correlated examination times unspecified.

Chiou et al. (2018) considered a semiparametric accelerated mean model for recurrent event processes, and allowed the examination time process and the underlying recurrent event process to be correlated through a shared frailty.

In the following, we propose a general class of random-effects monotonic index models for a recurrent event process of interest in the presence of informative examination and follow-up times. A major advantage of these models is their flexibility and generality, because they include many popular models as special cases, such as the proportional means model (Lin et al. (2000); Sun and Wei (2000); Zhang et al. (2013)) and the accelerated mean model (Xu et al. (2017); Chiou et al. (2018)). To estimate the regression parameters, we develop a weighted-rank (WR) estimation method, that is insensitive to the choice of the observation process and the censoring mechanism.

Many works have developed rank-estimation methods for different types of regression models. For example, Han (1987) proposed distribution-free maximum-rank correlation estimators for generalized regression models. Cavanagh and Sherman (1998) proposed monotone rank estimators for monotonic linear index models. Lin et al. (2017) provided a maximum-rank correlation estimator for random-effects transformation models when the random-effects distribution is symmetric. Liu, Yuan and Sun (2021a) proposed a WR estimator for nonparametric transformation models with nonignorable missing data; see Abrevaya (1999), Khan and Tamer (2007), Wang and Chiang (2019) and Liu, Yuan and Sun (2021b,c) for applications of rank-estimation methods in other contexts. To the best of our knowledge, there are no rank-estimation procedures available for the analysis of a recurrent event process in the presence of informative examination and follow-up times.

The remainder of the paper is organized as follows. In Section 2, we introduce the random-effects monotonic index model and present the proposed WR estimators for the regression parameters. We also establish the consistency and asymptotic normality of the WR estimators and provide a random weighting resampling scheme for approximating the distribution of the WR estimators. The simulation results presented in Section 3 suggest that the proposed method works well in practical situations. In Section 4, the approach is applied to real panel count data, and Section 5 concludes the paper. The proofs of the asymptotic results are provided in Section S1 of the Supplementary Material.

2. A WR Estimation for Random-Effects Monotonic Index Models

In this section, we first describe a class of random-effects monotonic index models, and then present the proposed WR estimation procedure.

2.1. Random-effects monotonic index models

Consider a study involving N subjects, who may experience recurrent events. For subject $i \in \{1, \ldots, N\}$, let $M_i(t)$ represent the cumulative number of events that have occurred before time t, for $0 \le t \le \tau$, where τ is a known constant time point. For subject i, suppose there is a $(p+1)$-dimensional vector W_i of covariates, the effects of which on $M_i(t)$ are of primary interest. For $i = 1, \ldots, N$, let K_i denote the number of observation times for the ith subject. We allow K_i to take the value zero when there are no observations for the ith subject. Even if the follow-up time for each subject is τ, the counting process $M_i(t)$ is observed only at finite time points $T_{i(1)} < \cdots < T_{i(K_i)}$, which are order statistics of $\{T_{i1}, \ldots, T_{iK_i}\}$. Naturally, not every subject can be followed until τ. Let C_i be the follow-up time for the ith subject. Then, $M_i(T_{ik})$ cannot be observed when $C_i < T_{ik} < \tau$. That is, we only observe panel count data given by

$$\{(T_{ik}, M_{ik} = M_i(T_{ik}), W_i^{\mathsf{T}}, C_i)^{\mathsf{T}} : T_{ik} \le C_i, k = 1, \ldots, K_i, i = 1, \ldots, N\}. \quad (2.1)$$

Let Z_i be a q-dimensional random vector of latent variables that is independent of W_i. We assume that given $\{W_i, Z_i\}$, the conditional mean function (CMF) of $M_i(t)$ follows the random-effects monotonic index model:

$$E\{M_i(t)|W_i, Z_i\} = \mu(W_i^{\mathsf{T}}\boldsymbol{\beta}^*, Z_i, t), \quad (2.2)$$

where $\|\boldsymbol{\beta}^*\| = 1$, and for fixed Z_i and t, $\mu(W_i^{\mathsf{T}}\boldsymbol{\beta}^*, Z_i, t)$ is an unspecified strictly increasing function of $W_i^{\mathsf{T}}\boldsymbol{\beta}^*$. If $M_i(t)$ is a Poisson process, equation (2.2) is a conventional model for the mean of a Poisson variable on the interval $(0, \tau)$. We are interested in making inferences about $\boldsymbol{\beta}^*$ in model (2.2), without assuming that $M_i(\cdot)$ is a homogeneous or non-homogeneous Poisson process.

We denote $\{H_i(t), t \ge 0\}$ as the observation process or examination time process, given by the point process

$$H_i(t) = \sum_{k=1}^{K_i} I(T_{ik} \le t), \quad (2.3)$$

representing the cumulative visit numbers up to time t. Throughout this paper, for any real numbers $\{a_i\}_{i \ge 1}$, we set $\sum_{k=1}^{0} a_i \equiv 0$. Thus, if $K_i = 0$, we have

$H_i(t) \equiv 0$. Let

$$O_i = O_i(\cdot) = (W_i^{\mathsf{T}}(\cdot), Z_i^{\mathsf{T}}(\cdot), C_i(\cdot), M_i(\cdot), H_i(\cdot))^{\mathsf{T}}$$
$$= (W_i^{\mathsf{T}}, Z_i^{\mathsf{T}}, C_i, M_i(\cdot), H_i(\cdot))^{\mathsf{T}}, \quad i = 1, \dots, N,$$

be independent copies of $O = O(\cdot) = (W^{\mathsf{T}}(\cdot), Z^{\mathsf{T}}(\cdot), C(\cdot), M(\cdot), H(\cdot))^{\mathsf{T}} = (W^{\mathsf{T}}, Z^{\mathsf{T}}, C, M(\cdot), H(\cdot))^{\mathsf{T}}$, where $W = (X_1, \dots, X_{p+1})^{\mathsf{T}}$. We assume that C, $H(\cdot)$, and $M(\cdot)$ are conditionally independent given $\{W, Z\}$. For the observation process and follow-up time, we assume that

$$E\{H(t)|W, Z\} = \varsigma(W)\nu(Z, t), \tag{2.4}$$
$$P(C > t|W, Z) = \kappa_1(W, t)\kappa_2(Z, t), \tag{2.5}$$

where $\varsigma(\cdot)$, $\nu(\cdot,\cdot)$, $\kappa_1(\cdot,\cdot)$, and $\kappa_2(\cdot,\cdot)$ are unspecified functions such that $\varsigma(\cdot) \geq 0$, $\nu(\cdot,\cdot) \geq 0$, $0 \leq \kappa_1(\cdot,\cdot)\kappa_2(\cdot,\cdot) \leq 1$, and for all fixed $\{Z, W\}$, $\partial\nu(Z,t)/\partial t \geq 0$ and $\partial\{\kappa_1(W,t)\kappa_2(W,t)\}/\partial t \leq 0$.

Remark 1. Let $H(t) = \sum_{k=1}^{K} I(T_k \leq t)$. Suppose K and $\{T_k : k \geq 1\}$ are conditionally independent given $\{W, Z\}$, $E(K|W, Z) = \varsigma(W)\nu_1(Z)$ and $E\{I(T_k \leq t)|W, Z\} = \nu_2(Z, t)$, for all $k \geq 1$. Then, it is easy to verify that

$$E\{H(t)|W, Z\} = \varsigma(W)\nu_1(Z)\nu_2(Z, t),$$

which shows that (2.4) is satisfied with $\varsigma(W)$ and $\nu(Z, t) = \nu_1(Z)\nu_2(Z, t)$.

Example 1. Assume that, given $\{W, Z\}$, K has a Poisson distribution with mean $\varsigma(W)\nu_1(Z) > 0$, and T_k, for $k \geq 1$, are independent and identically distributed (i.i.d.) with distribution function $\{F(t)\}^{\sigma(Z)}$, where $\sigma(z) > 0$ for $z \in \mathbb{R}^q$ and $F(t)$ is a distribution function such that $f(t) = dF(t)/dt > 0$ for $t > 0$. Under these assumptions, $E\{H(t)|W, Z\} = \varsigma(W)\nu(Z, t)$, where $\nu(Z, t) = \nu_1(Z)\nu_2(Z, t)$ and $\nu_2(Z, t) = \{F(t)\}^{\sigma(Z)}$.

Remark 2. Suppose $G(C) = G_1(\zeta_1, Z) \wedge G_2(\zeta_2, W)$, where ζ_1, ζ_2, Z, and W are independent, $G(\cdot)$ is an unknown increasing function, $G_1 \wedge G_2 = \min(G_1, G_2)$, $G_1(\cdot,\cdot)$, and $G_2(\cdot,\cdot)$ are unknown functions. Then, (2.5) is satisfied.

Example 2. Assume that $C = \Lambda_a \wedge \Lambda_b$, where Λ_a and Λ_b are conditionally independent given $\{W, Z\}$. Here, Λ_a and Λ_b follow exponential distributions with means $1/\lambda_1(W) > 0$ and $1/\lambda_2(Z) > 0$, respectively. Under these assumptions, $P(C > t|W, Z) = \kappa_1(W, t)\kappa_2(Z, t)$, where $\kappa_1(W, t) = \exp\{-\lambda_1(W)t\}$ and $\kappa_2(Z, t) = \exp\{-\lambda_2(Z)t\}$.

Remark 3. The frailty proportional rates models (Huang, Wang and Zhang (2006); Wang, Ma and Yan (2013)) and semiparametric accelerated mean models (Chiou et al. (2018)) are important cases of the proposed random-effects monotonic index models. Let Z denote a latent nonnegative frailty variable, the distribution of which is unspecified and satisfies $E(Z|W) = 1$. The frailty proportional rates model specifies the mean of the event process $M(t)$, given the frailty variable Z and covariate W, as $\mu_1(W^{\top}\boldsymbol{\beta}^*, Z, t) = Z\Lambda(t)\exp(W^{\top}\boldsymbol{\beta}^*)$, where $\Lambda(\cdot)$ is a completely unspecified baseline mean function. The semiparametric accelerated mean model assumes that the recurrent event process $M(t)$, conditioning on the latent nonnegative frailty variable Z and covariate W, has the mean function $\mu_2(W^{\top}\boldsymbol{\beta}^*, Z, t) = Z_i\Lambda_0\{t\exp(W^{\top}\boldsymbol{\beta}^*)\}$, where $\Lambda_0(t) = \int_0^t \lambda_0(u)du$ and $\lambda_0(u)$ is an unspecified, absolutely continuous baseline rate function.

Remark 4. Motivated by the frailty proportional rates models and semiparametric accelerated mean models, it is possible to relax the independence assumption of Z and W. Specifically, assume that given $\{W, Z\}$, C, $H(\cdot)$ and $M(\cdot)$ are conditionally independent, and the conditional mean function of $M(t)$ follows the special random-effects monotonic index model: $E\{M(t)|W, Z\} = g(Z)\pi(W^{\top}\boldsymbol{\beta}^*, t)$, where $\|\boldsymbol{\beta}^*\| = 1$, $g(Z)$ is an unspecified nonnegative function of Z such that $E\{g(Z)|W\} = 1$, and for fixed t, $\pi(W^{\top}\boldsymbol{\beta}^*, t)$ is an unspecified strictly increasing function of $W^{\top}\boldsymbol{\beta}^*$. For the observation process and follow-up time, assume that $E\{H(t)|W, Z\} = \varsigma(W)\sigma(Z)\nu(t)$ and $P(C > t|W, Z) = \kappa(W, t)$, where $\varsigma(\cdot)$, $\sigma(\cdot)$, $\nu(\cdot)$, and $\kappa(\cdot, \cdot)$ are unspecified functions such that $\varsigma(\cdot) \geq 0$, $\sigma(\cdot) \geq 0$, $E\{\sigma(Z)|W\} = 1$, $\nu(\cdot) \geq 0$, $0 \leq \kappa(\cdot, \cdot) \leq 1$, $\partial\nu(t)/\partial t \geq 0$, and for fixed W, $\partial\kappa(W, t)/\partial t \leq 0$. Then, under certain mild conditions, the consistency of the WR estimator $\hat{\boldsymbol{\beta}}$ (See subsection 2.2) to $\boldsymbol{\beta}^*$ still holds.

2.2. WR estimators

To construct a consistent rank regression estimator of $\boldsymbol{\beta}^*$ analogous to the maximum-rank correlation estimator in Han (1987) or the monotone rank estimator in Cavanagh and Sherman (1998), we construct a function $\Psi(O_i, O_j)$ that satisfies the property

$$E\{\Psi(O_i, O_j)|W_i, W_j\} \geq E\{\Psi(O_j, O_i)|W_i, W_j\} \quad \text{if and only if} \quad W_i^{\top}\boldsymbol{\beta}^* > W_j^{\top}\boldsymbol{\beta}^*. \tag{2.6}$$

For panel count data, we set

$$U_{ij} = \sum_{k=1}^{K_i} I(T_{ik} \le C_i \wedge C_j) M_{ik}, \quad V_{ij} = \sum_{l=1}^{K_j} I(T_{jl} \le C_i \wedge C_j), \tag{2.7}$$

$$\Psi(O_i, O_j) = Y_{ij} = U_{ij} V_{ij}, \quad i, j = 1, \ldots, N. \tag{2.8}$$

Here, $Y_{ij} = 0$ if $K_i K_j = 0$. Let $\Theta = \{\boldsymbol{\beta} \in \mathbb{R}^{p+1} : \|\boldsymbol{\beta}\| = 1\}$. The proposed WR estimator of $\boldsymbol{\beta}^*$ is defined as

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta} \in \Theta}{\operatorname{argmax}}\, Q_N(\boldsymbol{\beta}), \tag{2.9}$$

where

$$Q_N(\boldsymbol{\beta}) = \frac{1}{N^2 - N} \sum_{i \neq j}^{N} Y_{ij} I(W_i^{\mathsf{T}} \boldsymbol{\beta} > W_j^{\mathsf{T}} \boldsymbol{\beta}). \tag{2.10}$$

Note that $\{Q_N(\boldsymbol{\beta}) : \boldsymbol{\beta} \in \Theta\}$ is a U-process of order two. The consistency and asymptotic normality of $\hat{\boldsymbol{\beta}}$ are given by the following theorems.

Theorem 1. *Under the conditions* C0–C5 *in the Appendix, we have* $\hat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}^*$ *as* $N \to \infty$.

Note that the maximization of $Q_N(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ in (2.9) is subject to the constraint $\|\boldsymbol{\beta}\| = 1$. Following Liu, Yuan and Sun (2021a), we derive the asymptotic distribution of $\hat{\boldsymbol{\beta}}$ by first reparameterizing $\boldsymbol{\beta}^* = \boldsymbol{\beta}(\boldsymbol{\vartheta}^*) = \boldsymbol{\beta}(\boldsymbol{\theta}^*, \alpha^*) =: (\boldsymbol{\theta}^{*\mathsf{T}}, \alpha^* \sqrt{1 - \|\boldsymbol{\theta}^*\|^2})^{\mathsf{T}}$, where $\boldsymbol{\vartheta}^* = (\boldsymbol{\theta}^{*\mathsf{T}}, \alpha^*)^{\mathsf{T}}$ and $\boldsymbol{\theta}^* = (\theta_1^*, \ldots, \theta_p^*)^{\mathsf{T}}$. Let $\operatorname{sgn}(\cdot)$ denote the sign function, that is, $\operatorname{sgn}(u) = I(u > 0) - I(u < 0)$, for $u \in \mathbb{R}$. Obviously, we can write $\boldsymbol{\vartheta}^* = \boldsymbol{\vartheta}(\boldsymbol{\beta}^*) =: (\beta_1^*, \ldots, \beta_p^*, \operatorname{sgn}(\beta_{p+1}^*))^{\mathsf{T}}$, where $\alpha^* = \operatorname{sgn}(\beta_{p+1}^*)$ and $\theta_j^* = \beta_j^*$, for $j = 1, \ldots, p$. Define

$$\hat{\boldsymbol{\vartheta}} = (\hat{\boldsymbol{\theta}}^{\mathsf{T}}, \hat{\alpha})^{\mathsf{T}} = \underset{\boldsymbol{\vartheta} \in \Upsilon}{\operatorname{argmax}}\, Q_N\{\boldsymbol{\beta}(\boldsymbol{\vartheta})\}, \tag{2.11}$$

where $\Upsilon = \{\boldsymbol{\vartheta}(\boldsymbol{\beta}) : \boldsymbol{\beta} \in \Theta\}$. Then, it is easy to see that $\hat{\boldsymbol{\theta}} = (\hat{\beta}_1, \ldots, \hat{\beta}_p)^{\mathsf{T}}$ and $\hat{\alpha} = \operatorname{sgn}(\hat{\beta}_{p+1})$. From Theorem 1 and the condition C4 in the Appendix, $\hat{\boldsymbol{\vartheta}}$ is a consistent estimator of $\boldsymbol{\vartheta}^*$.

Theorem 2. *Define*

$$Q(\boldsymbol{\beta}) = E[\Psi(O_1, O_2) I(W_1^{\mathsf{T}} \boldsymbol{\beta} > W_2^{\mathsf{T}} \boldsymbol{\beta})], \quad A(\boldsymbol{\vartheta}) = -\frac{\partial^2 Q\{\boldsymbol{\beta}(\boldsymbol{\vartheta})\}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^{\mathsf{T}}},$$

where $\boldsymbol{\vartheta} = (\boldsymbol{\theta}^\mathsf{T}, \alpha)^\mathsf{T}$. *Let* U *be a normal random vector with mean* $0_{p\times 1}$ *and covariance matrix* $\Sigma = \mathrm{cov}\{h(O_i, \boldsymbol{\vartheta}^*)\}$, *where*

$$\begin{aligned} h(O_i, \boldsymbol{\vartheta}) &= \frac{A^{-1}(\boldsymbol{\vartheta})\partial b_i(\boldsymbol{\vartheta})}{\partial \boldsymbol{\theta}}, \\ b_i(\boldsymbol{\vartheta}) &= E[a_{ij}(\boldsymbol{\vartheta}) + a_{ji}(\boldsymbol{\vartheta}) - 2E\{a_{ij}(\boldsymbol{\vartheta})\}|O_i], \\ a_{ij}(\boldsymbol{\vartheta}) &= \Psi(O_i, O_j)[I\{W_i^\mathsf{T}\boldsymbol{\beta}(\boldsymbol{\vartheta}) > W_j^\mathsf{T}\boldsymbol{\beta}(\boldsymbol{\vartheta})\} \\ &\quad -I\{W_i^\mathsf{T}\boldsymbol{\beta}(\boldsymbol{\theta}^*, \alpha) > W_j^\mathsf{T}\boldsymbol{\beta}(\boldsymbol{\theta}^*, \alpha)\}]. \end{aligned}$$

If the conditions C0–C6 *in the Appendix are satisfied, we have* $N^{1/2}(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}^*) = N^{1/2}((\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)^\mathsf{T}, (\hat{\alpha} - \alpha^*))^\mathsf{T} \xrightarrow{d} (U^\mathsf{T}, 0)^\mathsf{T}$ *as* $N \to \infty$. *Define* $B(\boldsymbol{\vartheta}) = \partial\boldsymbol{\beta}(\boldsymbol{\vartheta})/\partial\boldsymbol{\vartheta}^\mathsf{T}$. *Then, as* $N \to \infty$, $N^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) = N^{1/2}\{\boldsymbol{\beta}(\hat{\boldsymbol{\vartheta}}) - \boldsymbol{\beta}(\boldsymbol{\vartheta}^*)\} \xrightarrow{d} B(\boldsymbol{\vartheta}^*)(U^\mathsf{T}, 0)^\mathsf{T}$, *which is a normal random vector with mean* $0_{(p+1)\times 1}$ *and covariance matrix* $\Omega = \mathrm{cov}\{B(\boldsymbol{\vartheta}^*)(U^\mathsf{T}, 0)^\mathsf{T}\}$.

2.3. Coordinate-wise algorithm

In this section, following Wu and Stefanski (2015) and Liu, Yuan and Sun (2021a), we present a coordinate-wise optimization algorithm to optimize the objective function $Q_N(\boldsymbol{\beta})$. The idea of the algorithm is to maximize one coordinate at a time, with the other fixed. Define

$$\ell_k(\gamma|\boldsymbol{\zeta}, \boldsymbol{\eta}) = \sum_{(i,j)\in\mathscr{S}} p_{ij} I\{(\boldsymbol{\zeta}^\mathsf{T}, \gamma, \boldsymbol{\eta}^\mathsf{T})(W_i - W_j) < 0\}, \quad k = 1, \ldots, p+1,$$

where $\boldsymbol{\zeta}^\mathsf{T} = (\zeta_1, \ldots, \zeta_{k-1})$, $\boldsymbol{\eta}^\mathsf{T} = (\eta_1, \ldots, \zeta_{p+1-k})$, $p_{ij} = Y_{ji}$, and $\mathscr{S} = \{(i,j) : p_{ij} \neq 0\}$. Note that we set $\boldsymbol{\zeta} = \emptyset$ if $k = 1$ and $\boldsymbol{\eta} = \emptyset$ if $k = p+1$. To this end, we can write $\ell_k(\gamma|\boldsymbol{\zeta}, \boldsymbol{\eta})$ as

$$\begin{aligned} \ell_k(\gamma|\boldsymbol{\zeta}, \boldsymbol{\eta}) &= \sum_{(i,j)\in\mathscr{S}} p_{ij} I(a_{ij,k}\gamma < b_{ij,k}) \\ &= \sum_{(i,j)\in\mathscr{S}} p_{ij}\{I(\gamma < c_{ij,k})I(a_{ij,k} > 0) + I(\gamma > c_{ij,k})I(a_{ij,k} < 0)\} \\ &= \sum_{l=1}^{L} p_l\{I(\gamma < c_{l,k})I(a_{l,k} > 0) + I(\gamma > c_{l,k})I(a_{l,k} < 0)\}, \end{aligned} \tag{2.12}$$

where $c_{ij,k} = b_{ij,k}/a_{ij,k}$, $a_{ij,k} = X_{ik} - X_{jk}$, $b_{ij,k} = -\sum_{l=1}^{k-1}(X_{il} - X_{jl})\zeta_l - \sum_{l=k+1}^{p+1}(X_{il} - X_{jl})\eta_{l-k}$ and

$$\{(a_{ij,k}, b_{ij,k}, c_{ij,k}, p_{ij,k}) : (i,j) \in \mathscr{S}\}$$

$$= \{(a_{l,k}, b_{l,k}, c_{l,k}, p_{l,k}) : c_{1,k} < \cdots < c_{L,k},\ L = |\mathscr{S}|\}.$$

Here, for ease of presentation, the dependence of $(b_{ij,k}, c_{ij,k})$ and $(b_{l,k}, c_{l,k})$ on $(\boldsymbol{\zeta}, \boldsymbol{\eta})$ is suppressed. Note that the objective function $\ell_k(\gamma|\boldsymbol{\zeta}, \boldsymbol{\eta})$ is just a piecewise constant function with cut-off values $\{c_l\}_{l=1}^L$. Furthermore, Liu, Yuan and Sun (2021a) showed that

$$\ell_k(c_{s+1,k}|\boldsymbol{\zeta}, \boldsymbol{\eta}) = \ell_k(c_{s,k}|\boldsymbol{\zeta}, \boldsymbol{\eta}) - p_{s+1} I(a_{s+1,k} > 0) + p_s I(a_{s,k} < 0),$$
$$s = 1, \ldots, L-1. \tag{2.13}$$

Using (2.13), the values of $\{\ell_k(c_{l,k}|\boldsymbol{\zeta}, \boldsymbol{\eta})\}_{l=1}^L$ can be easily obtained. Thus, $\ell_k(\gamma|\boldsymbol{\zeta}, \boldsymbol{\eta})$ is very easy to maximize. It follows that

$$\hat{\gamma}_k(\boldsymbol{\zeta}, \boldsymbol{\eta}) = \operatorname*{argmax}_{\gamma \in \mathbb{R}} \ell_k(\gamma|\boldsymbol{\zeta}, \boldsymbol{\eta}) = \operatorname*{argmax}_{\gamma \in \{c_{1,k}, \ldots, c_{L,k}\}} \ell_k(\gamma|\boldsymbol{\zeta}, \boldsymbol{\eta}). \tag{2.14}$$

To this end, the proposed coordinate optimization algorithm is as follows:

1. Set the initial value $\boldsymbol{\beta}^{(0)} = (\beta_1^{(0)}, \ldots, \beta_{p+1}^{(0)})^\top$;

2. Given $\boldsymbol{\beta}^{(m)} = (\beta_1^{(m)}, \ldots, \beta_{p+1}^{(m)})^\top$, for $k = 1, \ldots, p+1$, compute

$$\beta_k^{(m+1)} = \hat{\gamma}_k((\beta_1^{(m+1)}, \ldots, \beta_{k-1}^{(m+1)})^\top, (\beta_{k+1}^{(m)}, \ldots, \beta_{p+1}^{(m)})^\top),$$

   where $\hat{\gamma}_k(\boldsymbol{\zeta}, \boldsymbol{\eta})$ is defined in (2.14). Then, set $\boldsymbol{\beta}^{(m+1)} = (\beta_1^{(m+1)}, \ldots, \beta_{p+1}^{(m+1)})^\top$;

3. Repeat step 2 till $|Q_N(\boldsymbol{\beta}^{(m+1)}) - Q_N(\boldsymbol{\beta}^{(m)})| < 10^{-6}$. Denote the final value of $\boldsymbol{\beta}$ by $\boldsymbol{\beta}^{(\infty)}$;

4. Set $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}^{(\infty)}/\|\boldsymbol{\beta}^{(\infty)}\|$.

Obviously, the proposed coordinate-wise optimization algorithm guarantees a monotone increasing of $Q_N(\boldsymbol{\beta})$ at each iteration step in a very efficient manner. That is, we always have $Q_N(\boldsymbol{\beta}^{(m+1)}) \geq Q_N(\boldsymbol{\beta}^{(m)})$, for $m = 1, 2, \ldots$.

To apply Theorem 2 to make an inference about $\boldsymbol{\beta}^*$, we estimate the limiting covariance matrix Ω of $\hat{\boldsymbol{\beta}}$. Inspired by the method of Jin, Ying and Wei (2001), we develop a resampling scheme to approximate the distribution of $\hat{\boldsymbol{\beta}}$, particularly its covariance matrix. Let $\lambda_1, \ldots, \lambda_N$ be i.i.d. exponential random variables with mean one, that is, Exp(1). The resampling WR estimator of $\boldsymbol{\beta}^*$ is defined as

$$\hat{\boldsymbol{\beta}}^* = \operatorname*{argmax}_{\boldsymbol{\beta} \in \Theta} Q_N^*(\boldsymbol{\beta}), \tag{2.15}$$

where

$$Q_N^*(\boldsymbol{\beta}) = \frac{1}{N^2 - N} \sum_{i \neq j}^{N} \lambda_i \lambda_j Y_{ij} I(W_i^{\mathsf{T}} \boldsymbol{\beta} > W_j^{\mathsf{T}} \boldsymbol{\beta}). \tag{2.16}$$

Then, we have the following proposition.

Proposition 1. *Under the conditions* C0–C6 *in the Appendix, given* $\{O_i, i = 1, \ldots, N\}$, $N^{1/2}(\hat{\boldsymbol{\beta}}^* - \hat{\boldsymbol{\beta}}) \xrightarrow{d} N(0_{(p+1)\times 1}, \Omega)$ *as* $N \to \infty$, *which is the limiting distribution of* $N^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)$.

The proof of Proposition 1 is similar to that of Jin, Ying and Wei (2001), and is thus omitted. Theoretically, the distribution of $N^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)$ can be approximated by the resampling distribution of $N^{1/2}(\hat{\boldsymbol{\beta}}^* - \hat{\boldsymbol{\beta}})$, given the data $\{O_i, i = 1, \ldots, N\}$. In practice, for $b = 1, \ldots, B$, we produce $\hat{\boldsymbol{\beta}}_b^*$ by solving (2.15) with random weights $(\lambda_{b1}, \ldots, \lambda_{bN})$ in (2.16), while holding the data $\{O_i, i = 1, \ldots, N\}$ at their observed values. The distribution of $\hat{\boldsymbol{\beta}}$ can then be approximated by the empirical distribution of $\{\hat{\boldsymbol{\beta}}_b^*\}_{b=1}^B$. The asymptotic covariance matrix Ω can be estimated by $\hat{\Omega} = NB^{-1} \sum_{b=1}^B (\hat{\boldsymbol{\beta}}_b^* - \bar{\boldsymbol{\beta}}^*)(\hat{\boldsymbol{\beta}}_b^* - \bar{\boldsymbol{\beta}}^*)^{\mathsf{T}}$, with $\bar{\boldsymbol{\beta}}^* = B^{-1} \sum_{b=1}^B \hat{\boldsymbol{\beta}}_b^*$. For $j = 1, \ldots, p+1$, the $100 \times (1-\alpha)\%$ confidence interval of β_j^* is given by $[\hat{\beta}_j - N^{-1/2} \hat{\omega}_{jj}^{1/2} C_{1-\alpha/2}, \hat{\beta}_j + N^{-1/2} \hat{\omega}_{jj}^{1/2} C_{1-\alpha/2}]$, where $\hat{\omega}_{jj}$ is the (j,j)-element of $\hat{\Omega}$, and $C_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution.

3. Simulation Study

In this Section, we present the results obtained from a simulation study conducted to investigate the performance of the proposed WR estimator $\hat{\boldsymbol{\beta}}$, and compare it with that of several other estimators given in the literature for the same problem. Additional results for the complicated simulation settings are given in Section S2 of the Supplementary Material.

In the study, we assume that the latent variable Z_i follows a gamma distribution with parameters $(2, 2)$, and K_i follows a uniform distribution over $\{3, 4, 5, 6\}$. We set $\tau = 8$ and $C_i = \min(C_i^*, \tau)$, where C_i^* follows a uniform distribution over $[1, 10]$. Given K_i, the observation times $T_{i1}, \ldots, T_{iK_i}$ are i.i.d. with a uniform distribution over $(0, C_i)$. Let $W_i = (X_{i1}, X_{i2})^{\mathsf{T}}$, where X_{1i} is a standard normal random variable, and X_{i2} is an exponential random variable with mean one. Given W_i, Z_i, K_i, and $(T_{i1}, \ldots, T_{iK_i})$, we generate $M_{ik} = M(T_{ik})$ using the formula

$$M_i(T_{ik}) = M_i(T_{i1}) + \{M_i(T_{i2}) - M_i(T_{i1})\} + \cdots + \{M_i(T_{ik}) - M_i(T_{i,k-1})\}$$

and assume that $M_i(t) - M_i(s)$ follows a Poisson distribution with mean

$$\mu(W_i^\top \boldsymbol{\beta}^*, Z_i, t) - \mu(W_i^\top \boldsymbol{\beta}^*, Z_i, s).$$

For the random-effects monotonic index model, we consider the following four CMFs:

(a) $\mu(W_i^\top \boldsymbol{\beta}^*, Z_i, t) = tg(Z_i)\exp(W_i^\top \boldsymbol{\beta}^*)$;

(b) $\mu(W_i^\top \boldsymbol{\beta}^*, Z_i, t) = t\{g(Z_i) + \exp(W_i^\top \boldsymbol{\beta}^*) + 1\}$;

(c) $\mu(W_i^\top \boldsymbol{\beta}^*, Z_i, t) = t\{g(Z_i) + \exp(W_i^\top \boldsymbol{\beta}^*)\}$;

(d) $\mu(W_i^\top \boldsymbol{\beta}^*, Z_i, t) = tg(Z_i) + \exp(W_i^\top \boldsymbol{\beta}^*)$,

where $\boldsymbol{\beta}^* = (\beta_1^*, \beta_2^*)^\top$ and $g(Z) = \sqrt{Z}$ or $g(Z) = \exp(-Z/5)$.

As mentioned above, several methods have been proposed in the literature for the problem discussed here under the simpler model (a). In the study, we compared the proposed WR estimator with four such estimators, namely the estimating equation-based estimator given in Sun and Wei (2000), the maximum pseudo-likelihood estimator given in Zhang (2002), the maximum likelihood estimator given in Lu, Zhang and Huang (2007), and the augmented estimating equation-based estimator given in Wang, Ma and Yan (2013). In the following, these are referred to as the EE, MPL, ML, and AEE estimators, respectively, and are determined using the R function *panelReg* in the SPEF package (Chiou et al. (2019)). The results given below are based on $N = 100$ or 200 and $B = 200$, with 1,000 replications.

Tables 1 and 2 compare the proposed WR estimator with the four existing estimators, with $\boldsymbol{\beta}^* = (-1, 1)^\top/\sqrt{2}$ and $N = 100$ and 200, and we calculate both the empirical bias (Bias) and the mean squared error (MSE) for each of the five estimators. The WR estimator is always unbiased and is also stable across all the cases. By comparison, the four other estimators seem to be unbiased under model (a), but are biased under models (b), (c), and (d). In addition, the proposed WR estimator appears to have a smaller MSE than the EE, MPL, ML, and AEE estimators under models (b), (c) and (d). Moreover, the MPL, ML, and WR estimators perform similarly in terms of the bias and MSE under model (a). In other words, as expected, the WR estimator applies to more general situations than the existing estimators do.

Tables 3 and 4 present the results obtained for the proposed WR estimator under the simulation settings described above, with $\boldsymbol{\beta}^* = (-1, 1)^\top/\sqrt{2}$ and

Table 1. Simulation results under four different conditional mean functions with $g(Z) = \sqrt{Z}$. Bias: empirical bias; MSE: mean squared error; CMF: conditional mean function.

			β_1^*		β_2^*	
N	CMF	Estimator	Bias	MSE	Bias	MSE
100	a	WR	0.0093	0.0082	0.0004	0.0044
		AEE	0.0010	0.0061	-0.0017	0.0039
		ML	0.0014	0.0063	-0.0026	0.0039
		MPL	0.0017	0.0056	-0.0029	0.0037
		EE	-0.0032	0.0075	-0.0077	0.0073
	b	WR	0.0191	0.0118	0.0056	0.0073
		AEE	0.3368	0.1165	-0.2431	0.0638
		ML	0.3363	0.1162	-0.2428	0.0637
		MPL	0.3449	0.1219	-0.2524	0.0680
		EE	0.4133	0.1763	-0.3122	0.1050
	c	WR	0.0151	0.0085	0.0051	0.0057
		AEE	0.2683	0.0747	-0.1886	0.0388
		ML	0.2679	0.0745	-0.1886	0.0388
		MPL	0.2775	0.0796	-0.1975	0.0420
		EE	0.3524	0.1305	-0.2538	0.0715
	d	WR	0.0066	0.0080	-0.0024	0.0047
		AEE	0.1105	0.0172	-0.0745	0.0098
		ML	0.1104	0.0174	-0.0745	0.0099
		MPL	0.1164	0.0181	-0.0785	0.0101
		EE	0.1797	0.0385	-0.1211	0.0209
200	a	WR	0.0051	0.0050	0.0001	0.0021
		AEE	-0.0009	0.0035	0.0056	0.0033
		ML	0.0006	0.0032	-0.0026	0.0023
		MPL	0.0011	0.0028	-0.0022	0.0021
		EE	-0.0026	0.0036	0.0010	0.0031
	b	WR	0.0042	0.0044	-0.0010	0.0031
		AEE	0.3253	0.1080	-0.2248	0.0538
		ML	0.3250	0.1079	-0.2247	0.0538
		MPL	0.3333	0.1131	-0.2334	0.0575
		EE	0.4130	0.1733	-0.3072	0.0981
	c	WR	0.0075	0.0036	0.0031	0.0026
		AEE	0.2553	0.0670	-0.1711	0.0315
		ML	0.2551	0.0669	-0.1711	0.0315
		MPL	0.2641	0.0714	-0.1793	0.0342
		EE	0.3495	0.1249	-0.2495	0.0651
	d	WR	0.0053	0.0053	-0.0002	0.0024
		AEE	0.0976	0.0126	-0.0629	0.0063
		ML	0.0975	0.0126	-0.0630	0.0064
		MPL	0.1028	0.0133	-0.0672	0.0067
		EE	0.1773	0.0344	-0.1153	0.0163

Table 2. Simulation results under four different conditional mean functions with $g(Z) = \exp(-Z/5)$. Bias: empirical bias; MSE: mean squared error; CMF: conditional mean function.

			β_1^*		β_2^*	
N	CMF	Estimator	Bias	MSE	Bias	MSE
200	a	WR	0.0073	0.0068	0.0001	0.0034
		AEE	0.0002	0.0058	-0.0048	0.0037
		ML	0.0007	0.0059	-0.0055	0.0037
		MPL	0.0019	0.0056	-0.0068	0.0035
		EE	-0.0045	0.0064	-0.0018	0.0043
	b	WR	0.0098	0.0057	0.0042	0.0023
		AEE	0.2242	0.0517	-0.1459	0.0231
		ML	0.2241	0.0516	-0.1458	0.0231
		MPL	0.2337	0.0559	-0.1548	0.0257
		EE	0.3198	0.1051	-0.2250	0.0538
	c	WR	0.0063	0.0043	0.0020	0.0018
		AEE	0.0958	0.0098	-0.0604	0.0041
		ML	0.0959	0.0099	-0.0607	0.0042
		MPL	0.1027	0.0113	-0.0663	0.0049
		EE	0.1714	0.0326	-0.1093	0.0145
	d	WR	0.0080	0.0061	0.0006	0.0044
		AEE	0.2625	0.0731	-0.1784	0.0369
		ML	0.2622	0.0730	-0.1784	0.0369
		MPL	0.2685	0.0761	-0.1847	0.0389
		EE	0.3544	0.1290	-0.2522	0.0677

$N = 100$ and 200. The tables show the sample standard deviation of the estimates (SD), the average of the estimated standard deviations (ESD), and the 95% empirical coverage probability (CP). As shown, the ESD appears to be close to the SD, and is thus appropriate. In addition, the results for the CP suggest that the normal approximation to the distribution of the WR estimator seems to be reasonable and, as expected, the results become better as the sample size increases. Additional simulation results under other setups presented in the Supplementary Material yield similar results.

4. Data Analysis

In this section, we apply the proposed WR estimation procedure to a well-known bladder cancer data set (Andrews and Herzberg (1985, pp.250-260), Sun and Wei (2000)). A total of 116 patients with bladder tumors are randomized into three treatment groups: placebo, pyridoxine, and thiotepa. The main objective is to assess the effectiveness of the treatment in reducing the occurrence rate of

Table 3. Simulation results under four different conditional mean functions with $g(Z) = \sqrt{Z}$. SD: sample standard deviation of the parameter estimator; ESD: the average of the estimated standard deviations; CP: the empirical coverage probabilities of the 95% confidence interval; CMF: conditional mean function.

		β_1^*			β_2^*		
N	CMF	SD	ESD	CP	SD	ESD	CP
100	a	0.0899	0.0820	0.9550	0.0662	0.0693	0.9500
	b	0.1068	0.1063	0.9710	0.0852	0.0896	0.9490
	c	0.0911	0.0977	0.9610	0.0755	0.0812	0.9510
	d	0.0893	0.0846	0.9700	0.0685	0.0723	0.9640
200	a	0.0704	0.0604	0.9740	0.0456	0.0486	0.9740
	b	0.0664	0.0742	0.9800	0.0553	0.0626	0.9710
	c	0.0597	0.0671	0.9700	0.0512	0.0562	0.9640
	d	0.0729	0.0597	0.9630	0.0494	0.0502	0.9530

Table 4. Simulation results under four conditional mean functions with $g(Z) = \exp(-Z/5)$. SD: sample standard deviation of the parameter estimator; ESD: the average of the estimated standard deviations; CP: the empirical coverage probabilities of the 95% confidence interval; CMF: conditional mean function.

		β_1^*			β_2^*		
n	CMF	SD	ESD	CP	SD	ESD	CP
100	a	0.1031	0.0968	0.9570	0.0805	0.0830	0.9480
	b	0.0938	0.0887	0.9680	0.0712	0.0739	0.9590
	c	0.0778	0.0779	0.9610	0.0590	0.0637	0.9630
	d	0.1187	0.1165	0.9610	0.1018	0.0992	0.9310
200	a	0.0821	0.0720	0.9590	0.0587	0.0600	0.9470
	b	0.0747	0.0626	0.9730	0.0479	0.0514	0.9700
	c	0.0651	0.0558	0.9690	0.0419	0.0445	0.9630
	d	0.0776	0.0833	0.9760	0.0663	0.0708	0.9640

bladder tumors. Because the patients are observed periodically, only the numbers of occurrences of bladder tumors between observation times are available. That is, only panel count data are observed on the underlying process. In addition to the treatment indicators, there exists another covariate, namely the initial number of bladder tumors.

To apply the proposed estimation procedure, let $M(t)$ denote the total number of recurrences of bladder tumors up to time t, and assume that $M(t)$ is described by the index model (2). Define X_1 as the initial number of tumors, $X_2 = 1$ if the patient was given the thiotepa treatment, and zero otherwise, and $X_3 = 1$ if the patient was given the pyridoxine treatment, and zero otherwise.

Table 5 presents the estimated covariate effects with $B = 200$, including

Table 5. Regression analysis results for the Bladder Tumor data: estimates of the regression coefficients, estimated standard deviation (ESD), and P-values for testing $\beta_j^* = 0$, $j = 1, 2, 3$.

	Estimate			ESD			P-value		
	β_1^*	β_2^*	β_3^*	β_1^*	β_2^*	β_3^*	β_1^*	β_2^*	β_3^*
WR	0.6178	-0.6060	0.5010	0.2067	0.2117	0.5056	0.0028	0.0042	0.3217
AEE	0.6354	-0.7656	0.0997	0.2027	0.3486	0.4380	0.0017	0.0280	0.8200
ML	0.6380	-0.7644	0.0973	0.2184	0.3242	0.3980	0.0035	0.0180	0.8100
MPL	0.4211	-0.6284	0.2146	0.1577	0.2811	0.2913	0.0076	0.0250	0.4600
EE	1.0306	-1.2630	0.3681	0.3408	0.4795	0.5730	0.0025	0.0084	0.5200

the WR estimates, estimated standard deviation (ESD) and p-values for testing $\beta_j^* = 0$, for $j = 1, 2, 3$.

For comparison, we also include the results based on the four estimation procedures discussed in the previous section. The results show that the five methods give similar conclusions, and all suggest that the thiotepa treatment reduces the bladder tumor recurrence rate. By comparison, the pyridoxine treatment does not seem to have any effect in terms of reducing the bladder tumor rate. Although the results from the different approaches are similar, the WR method indicates a stronger thiotepa treatment effect than the others do.

5. Conclusion

We have proposed a class of of random-effects monotonic index models for regression analyses of panel count data, which are common in many fields. A challenge for the problem is that, in addition to the underlying recurrent event process of interest, one has to deal with two nuisance processes, namely the observation process and the follow-up process, which may be correlated with the recurrent event process. The proposed random-effects monotonic index models include many popular models as special cases, such as the proportional means model and the accelerated mean model. We have developed a WR estimation procedure for estimation, and established the asymptotic properties of the resulting estimators of the regression parameters. In addition, our numerical results indicate that the proposed approach applies to more general situations than existing methods do.

In this study, we have focused on the estimation of the regression parameters. However, one may sometimes also be interested in estimating the mean function μ in (2.2), as well as the relationship between the recurrent event process of interest and the observation process. It is apparent that a new method has to be developed to solve these issues. Another possible direction for future research is the case in which there exist several correlated recurrent event processes and one observes multivariate panel count data (Sun and Zhao (2013)). In this case, a key issue is how to model the correlation between these processes. Of course, one could choose to leave the correlation structure arbitrary if only the marginal effects are of interest.

Supplementary Material

The online Supplementary Material contains the proofs of Theorems 1 and 2 and additional simulation results.

Acknowledgments

There are no conflicts of interest.

Appendix

In this appendix, we will give the conditions for the proof of the asymptotic results in Theorems 1 and 2. For this, let $O = (W^{\mathsf{T}}, Z^{\mathsf{T}}, C, M(\cdot), H(\cdot))^{\mathsf{T}}$ denote an observation from the distribution P on the set $\mathcal{O} \subseteq \mathbb{R}^{p+1} \times \mathbb{R}^q \times \mathbb{R}_+ \times D(0,\infty) \times D(0,\infty)$, where $D(0,\infty)$ is the space of real-valued cadlag functions on $[0,\infty)$. For each $o = o(\cdot) = (w, z, c, m(\cdot), h(\cdot))^{\mathsf{T}}$ in $\mathcal{O}$ and each $\boldsymbol{\beta}$ in Θ, define

$$\varrho(o,\boldsymbol{\beta}) = E\{\Psi(O,o)I(W^{\mathsf{T}}\boldsymbol{\beta} > w^{\mathsf{T}}\boldsymbol{\beta})\} + E\{\Psi(o,O)I(w^{\mathsf{T}}\boldsymbol{\beta} > W^{\mathsf{T}}\boldsymbol{\beta})\},$$

where $w = (x_1, \ldots, x_{p+1})^{\mathsf{T}}$ and $\Psi(O_1, O_2)$ is defined in (2.8). Write ∇_m for the mth partial derivative operator of the function $\varrho(z,\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{p+1})^{\mathsf{T}} \in \mathbb{R}^{p+1}$, and let

$$|\nabla_k|\varrho(o,\boldsymbol{\beta}) = \sum_{i_1,\ldots,i_k \in \{1,\ldots,p+1\}} \left| \frac{\partial^k}{\partial\beta_{i_1} \cdots \partial\beta_{i_k}} \varrho(o,\boldsymbol{\beta}) \right|.$$

Let $\mathcal{W}$, $\mathcal{Z}$ and $\mathcal{C}$ denote the supports of W, Z and C, respectively. To establish the asymptotic properties, we require the following conditions.

C0 (a) $O_i = (W_i^{\mathsf{T}}, Z_i^{\mathsf{T}}, C_i, M_i(\cdot), H_i(\cdot))^{\mathsf{T}}$, $i = 1, \ldots, N$ are independent copies of

$$O = (W^{\mathsf{T}}, Z^{\mathsf{T}}, C, M(\cdot), H(\cdot))^{\mathsf{T}};$$

(b) C, $M(\cdot)$ and $H(\cdot)$ are conditionally independent given $\{W, Z\}$; (c) W and Z are independent; (d) $E\{M(t)|W, Z\} = \mu(W^\mathsf{T}\boldsymbol{\beta}^*, Z, t)$ and for fixed $\{Z, t\}$, $\mu(W^\mathsf{T}\boldsymbol{\beta}^*, Z, t)$ is an unspecified strictly increasing function of $W^\mathsf{T}\boldsymbol{\beta}^*$; (e) $E\{H(t)|W, Z\} = \varsigma(W)\nu(Z, t)$, $P(C > t|W, Z) = \kappa_1(W, t)\kappa_2(Z, t)$, where $\varsigma(\cdot)$, $\nu(\cdot)$, $\kappa_1(\cdot, \cdot)$ and $\kappa_2(\cdot, \cdot)$ are unspecified functions such that $\varsigma(\cdot) \geq 0$, $\nu(\cdot, \cdot) \geq 0$, $0 \leq \kappa_1(\cdot, \cdot)\kappa_2(\cdot, \cdot) \leq 1$, and for all fixed $Z \in \mathcal{Z}$, $W \in \mathcal{W}$ and $t > 0$, $\partial\nu(Z, t)/\partial t \geq 0$ and $\partial\{\kappa_1(W, t)\kappa_2(W, t)\}/\partial t \leq 0$;

C1 Assume that for all $(w_1, w_2) \in \mathcal{W} \times \mathcal{W}$, $\varsigma(w_1)\varsigma(w_2) > 0$. For all $(w_1, w_2) \in \mathcal{W} \times \mathcal{W}$, there exist compact subsets $\mathcal{Z}_{w_1,w_2} \subseteq \mathcal{Z}$ and $\mathcal{C}_{w_1,w_2} \subseteq \mathcal{C}$ such that (a) $P(Z \in \mathcal{Z}_{w_1,w_2})\mathscr{L}(\mathcal{C}_{w_1,w_2}) > 0$, where $\mathscr{L}(\cdot)$ is the Lebesgue measure on $\mathbb{R}$; (b) for all $r \in \mathcal{Z}_{w_1,w_2}$, $(s, r) \in \mathcal{Z}_{w_1,w_2} \times \mathcal{Z}_{w_1,w_2}$ and $a \in \mathcal{C}_{w_1,w_2}$,

$$\frac{\partial\{1 - \kappa_1(w_1, a)\kappa_2(s, a)\kappa_1(w_2, a)\kappa_2(r, a)\}}{\partial a} > 0, \tag{A.1}$$

$$\int_0^a \frac{\partial\nu(r, t)}{\partial t} dt > 0; \tag{A.2}$$

C2 The set $\mathcal{W}$ is not contained in any proper linear subspace of $\mathbb{R}^{p+1}$;

C3 The vector of covariates, $W = (X_1, \ldots, X_{p+1})^\mathsf{T}$, is of full rank, and there exists $j \in \{1, \ldots, p+1\}$ such that X_j has an everywhere-positive Lebesgue density conditional on $X_{-j} = (X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_{p+1})^\mathsf{T}$;

C4 The unknown parameter $\boldsymbol{\beta}^* = (\beta_1^*, \ldots, \beta_{p+1}^*)^\mathsf{T}$, where $\boldsymbol{\beta}^* \in \Theta$, $\beta_j^* \neq 0$, and j is defined in condition C3; without loss of generality, assume that $j = p+1$ in conditions C3-C4;

C5 $E(Y_{12}^2) = E\{\Psi^2(O_1, O_2)\} < +\infty$.

C6 (a) Let $\mathcal{B}$ denote a neighborhood of $\boldsymbol{\beta}^*$. For each o, all mixed second partial derivatives of $\varrho(o, \boldsymbol{\beta})$ exist on $\mathcal{B}$. There is a function $\omega(o)$ such that $E\{\omega(O)\} < +\infty$ and for all $o \in \mathcal{O}$ and $\boldsymbol{\beta}$ in $\mathcal{B}$,

$$\|\nabla_2\varrho(o, \boldsymbol{\beta}) - \nabla_2\varrho(o, \boldsymbol{\beta}^*)\| \leq \omega(o)\|\boldsymbol{\beta} - \boldsymbol{\beta}^*\|;$$

(b) $E\{\|\nabla_1\varrho(O, \boldsymbol{\beta}^*)\|^2\} < +\infty$;

(c) $E\{|\nabla_2|\varrho(O, \boldsymbol{\beta}^*)\} < +\infty$;

(d) $E\{\partial^2\varrho\{O, \boldsymbol{\beta}(\boldsymbol{\theta}, \alpha^*)\}/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\mathsf{T}\}|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}$ is negative definite.

Most of the above conditions are assumed for rank estimation of a standard random-effects monotonic index model. Additional conditions are on the observation process, censoring mechanism and the distribution assumptions on $(C, W^{\mathsf{T}}, Z^{\mathsf{T}})^{\mathsf{T}}$. C0 defines the structure which generates the observations. In particular, C0(d) contains the key monotonicity assumption. Conditions C0-C5 guarantee the identifiability of $\boldsymbol{\beta}^*$. Specifically, conditions C0-C5 ensure that $Q(\boldsymbol{\beta}) = E\{Q_N(\boldsymbol{\beta})\}$ is uniquely maximized over Θ at $\boldsymbol{\beta}^*$, which can be proved by verifying that the inequality in (2.6) holds if and only if $\boldsymbol{\beta} = \boldsymbol{\beta}^*$.

From C0, we can write

$$P(C_1 \wedge C_2 > a | W_1, W_2, Z_1, Z_2) = \kappa_1(W_1, a)\kappa_2(Z_1, a)\kappa_1(W_2, a)\kappa_2(Z_2, a).$$

Thus, the conditional density of $C_1 \wedge C_2$ given $\{W_1, W_2, Z_1, Z_2\}$ can be written as

$$f_{C_1 \wedge C_2}(a | W_1, W_2, Z_1, Z_2) = \frac{\partial\{1 - \kappa_1(W_1, a)\kappa_2(Z_1, a)\kappa_1(W_2, a)\kappa_2(Z_2, a)\}}{\partial a}. \tag{A.3}$$

Observing (A.3), (A.1) in C1 is easily satisfied. Note that (A.2) in C1 is easily satisfied if we assume $\partial\nu(r, t)/\partial t > 0$ for all $r \in \mathcal{Z}$ and $t > 0$. For the consistency proof, condition C5 can be weakened to a finite first moment for Y_{12}, but the normality result requires a finite second moment. Conditions C6 contains standard regularity conditions sufficient to support an argument based on a Taylor expansion of $\varrho\{o; \boldsymbol{\beta}(\boldsymbol{\theta}, \alpha^*)\}$ about $\boldsymbol{\theta}$.

References

Abrevaya, J. (1999). Rank estimation of a transformation model with observed truncation. *Econom. J.* **2**, 292–305.

Andrews, D. F. and Herzberg, A. M. (1985). *Data: A Collection of Problems from Many Fields for the Student and Research Worker*. Springer-Verlag, New York.

Buzkova, P. (2010). Panel count data regression with informative observation times. *Int. J. Biostat.* **6**, Article 30.

Cai, J. and Schaubel, D. E. (2004). Analysis of recurrent event data. *Handbook of Statistics* **23**, 603–623.

Cavanagh, C. and Sherman, R. P. (1998). Rank estimators for monotonic index models. *J. Econometrics* **84**, 351–381.

Chiou, S. H., Xu, G., Yan, J. and Huang C.-Y. (2018). Semiparametric estimation of the accelerated mean model with panel count data under informative examination times. *Biometrics* **74**, 944–953.

Chiou, S. H., Huang, H.-Y., Xu, G. and Yan, J. (2019). Semiparametric regression analysis of panel count data: A practical review. *Int. Stat. Rev.* **87**, 24–43.

Han, A. K. (1987). Non-parametric analysis of a generalized regression model: The maximum rank correlation estimator. *J. Econometrics* **35**, 303–316.

He, X., Tong, X. and Sun, J. (2009). Semiparametric analysis of panel count data with correlated observation and follow-up times. *Lifetime Data Anal.* **15**, 177–196.

Hu, X. J., Sun, J. and Wei, L.-J. (2003). Regression parameter estimation from panel counts. *Scand. J. Stat.* **30**, 25–43.

Hua, L. and Zhang, Y. (2012). Spline-based semiparametric projected generalized estimating equation method for panel count data. *Biostatistics* **13**, 440–454.

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

Jiang, H., Su, W. and Zhao, X. (2020). Robust estimation for panel count data with informative observation times and censoring times. *Lifetime Data Anal.* **26**, 65–84.

Jin, Z., Ying, Z. and Wei, L. J. (2001). A simple resampling method by perturbing the minimand. *Biometrika* **88**, 381–390.

Kalbfleisch, J. D. and Lawless, J. F. (1985). The analysis of panel data under a Markov assumption. *J. Am. Statist. Assoc.* **80**, 863–871.

Khan, S. and Tamer, E. (2007). Partial rank estimation of duration models with general forms of censoring. *J. Econometrics* **136**, 251–280.

Kim, Y.-J. (2006). Analysis of panel count data with dependent observation times. *Comm. Statist. Simulation Comput.* **35**, 983–990.

Li, N., Park, D.-H., Sun, J. and Kim, K. (2011). Semiparametric transformation models for multivariate panel count data with dependent observation process. *Canad. J. Statist.* **39**, 458–474.

Lin, D. Y., Wei, L. J., Yang, I. and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *J. R. Statist. Soc. B* **62**, 711–730.

Lin, Y., Luo, Y., Xie, S. and Chen, K. (2017). Robust rank estimation for transformation models with random effects. *Biometrika* **104**, 971–986.

Liu, T., Yuan, X. and Sun, J. (2021a). Weighted rank estimation for nonparametric transformation models with nonignorable missing data. *Comput. Statist. Data Anal.* **153**, 107061.

Liu, T., Yuan, X. and Sun, J. (2021b). Weighted rank estimation for nonparametric transformation models with doubly truncated data. *J. Korean Statist. Soc.* **50**, 1–24.

Liu, T., Yuan, X. and Sun, J. (2021c). Weighted rank estimation of nonparametric transformation models with case-1 and case-2 interval-censored failure time data. *J. Nonparametr. Stat.* **33**, 225-248.

Lu, M., Zhang, Y. and Huang, J. (2007). Estimation of the mean function with panel count data using monotone polynomial splines. *Biometrika* **94**, 705–718.

Ma, L. and Sundaram, R. (2018). Analysis of gap times based on panel count data with informative observation times and unknown start time. *J. Am. Statist. Assoc.* **113**, 294–305.

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. *J. R. Statist. Soc. B* **62**, 293–302.

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

Xu, G., Chiou, S. H., Huang, C.-Y., Wang, M.-C. and Yan, J. (2017). Joint scale-change models for recurrent events and failure time. *J. Am. Statist. Assoc.* **112**, 794–805.

Thall, P. F. and Lachin, J. M. (1988). Analysis of recurrent events: Nonparametric methods for random-interval count data. *J. Am. Statist. Assoc.* **83**, 339–347.

Wang, S.-H. and Chiang, C.-T. (2019). Maximum partial-rank correlation estimation for left-truncated and right-censored survival data. *Statist. Sinica* **29**, 2141-2161.

Wang, X., Ma, S. and Yan, J. (2013). Augmented estimating equations for semiparametric panel count regression with informative observation times and censoring time. *Statist. Sinica* **23**, 359–381.

Wu, Y. and Stefanski, L. A. (2015). Automatic structure recovery for additive models. *Bometrika* **102**, 381–395.

Wang, Y. and Yu, Z. (2021). A kernel regression model for panel count data with time-varying coefficients. *Statistica Sinica* **31**, 1707–1725.

Wellner, J. A. and Zhang Y. (2000). Two estimators of the mean of a counting process with panel count data. *Ann. Statist.* **28**, 779–814.

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

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

Zhao, X. and Tong, X. (2011). Semiparametric regression analysis of panel count data with informative observation times. *Comput. Statist. Data Anal.* **55**, 291–300.

Zhao, H., Tu, W. and Yu, Z. (2018). A nonparametric time-varying coefficient model for panel count data. *J. Nonparametr. Stat.* **30**, 640–661.

Zhou, J., Zhang, H., Sun, L. and Sun, J. (2017). Joint analysis of panel count data with an informative observation process and a dependent terminal event. *Lifetime Data Anal.* **23**, 560–584.

Zhu, L., Zhang, Y., Li, Y., Sun, J. and Robison, L. L. (2018). A semiparametric likelihood-based method for regression analysis of mixed panel-count data. *Biometrics* **74**, 488–497.

Tianqing Liu

Center for Applied Statistical Research and School of Mathematics, Jilin University, Changchun, Jilin 130012, China.

E-mail: tianqingliu@gmail.com

Xiaohui Yuan

School of Mathematics and Statistics, Changchun University of Technology, Changchun, Jilin 130012, China.

E-mail: yuanxh@ccut.edu.cn

Jianguo Sun

Department of Statistics, University of Missouri, Columbia, MO 65211, USA.

E-mail: sunj@missouri.edu

(Received September 2020; accepted December 2021)