REGRESSION ANALYSIS OF PANEL COUNT DATA WITH BOTH TIME-DEPENDENT COVARIATES AND TIME-VARYING EFFECTS

Yuanyuan Guo[1*], Dayu Sun[2*] and Jianguo Sun[1]

[1]*University of Missouri* and [2]*Emory University*

Abstract: Panel count data occur in many fields, including clinical, demographic, and industrial studies, and an extensive body of literature has been established for their regression analysis. However, most existing methods apply only to situations in which both the covariates and their effects are constant or one of them may be time dependent. This study considers the situation in which both the covariates and their effects may be time dependent, and we develop an estimating equation-based approach to estimate these time-varying effects. The proposed method uses the B-splines to approximate the time-dependent coefficients, and we establish the asymptotic properties of the proposed estimators. To assess the finite-sample performance of the proposed estimators, we conduct an extensive simulation study, showing that the proposed method works well in practical situations. Lastly, we demonstrate our method by applying it to data from the China Health and Nutrition Survey.

Key words and phrases: B-spline, panel count data, proportional mean model, time-dependent effect.

1. Introduction

Event-history studies that examine the times to recurrent event occurrences appear in many fields, including clinical, demographic, and industrial studies. Such studies give rise to two types of data namely recurrent event data and panel count data (Cook and Lawless (2007); Sun and Zhao (2013)). The former means that all study subjects can be observed or followed continuously and, thus, one has complete data on the occurrence times of the event of interest. In contrast, the latter means that study subjects can be observed only at discrete time points, yielding only incomplete information on the occurrence times. Despite the missing information, panel count data frequently arise in practice, because subjects usually cannot be followed continuously. Thus, it is necessary to develop

---

Corresponding author: Jianguo Sun, Department of Statistics, University of Missouri, 146 Middlebush Hall, Columbia, MO 65211, USA. E-mail: sunj@missouri.edu.
*The first two authors contributed equally.

statistical methods for such data.

An example of panel count data that motivated this study is given by the China Health and Nutrition Survey (CHNS) (Tian (2018); Oliveira (2016)). The surveys were conducted every two to four years between 1989 and 2015 to obtain the fertility histories of the female participants, as well as their demographic, education, income, and health information. Owing to the intermittent survey times, the exact dates of the pregnancies or childbirths may not be available, leading to typical panel count data on the pregnancy number of the study subjects. Among others, one objective of the CHNS is to determine the relationship between the pregnancy process and various factors or covariates, including the income, location, education level, and health status of the female subject. It is apparent that some of these factors may change over time, as might their effects on the pregnancy process. For instance, the fertility difference between women with a good and a poor education may be more considerable for older women, because the education effect could accumulate over time. Furthermore, the health condition could also have a larger impact on older women, because they tend to be less resilient to diseases. As pointed out by many authors (Tian, Zucker and Wei (2005); Yu and Lin (2010); Perperoglou (2014); Lin, Fei and Li (2016)), an analysis of such data that ignores the time-dependent nature of the covariates and/or their effects would be less efficient. In other words, we need statistical methods that can accommodate both time-dependent covariates and time-varying effects.

A large body of literature has been established for regression analyses of panel count data with time-independent covariates and covariate effects. Other methods have been proposed for cases in which either the covariates are time-dependent or their effects vary over time. For example, Li, Sun and Sun (2010) considered a semi-parametric transformation mean model for panel count data with time-dependent covariates. Zhao, Tu and Yu (2018) and Wang and Yu (2021) considered time-varying coefficients and proposed some pseudo-likelihood methods for estimation. The former used a B-spline function approximation, and the latter employed a local polynomial approximation. However, both methods can apply only to constant covariate situations. To the best of our knowledge, there is no established method for regression analyses of panel count data that allows both time-dependent covariates and time-varying effects. As such, we propose a relevant estimating equation-based method.

The remainder of this paper is organized as follows. After introducing the notation and assumptions that we use throughout the paper, we present an estimating equation procedure in Section 2 for estimating covariate effects. In the proposed method, the conditional mean model is employed for the underly-

ing recurrent event process, and B-spline functions are used to approximate the time-varying covariate effects. The asymptotic properties of the proposed estimators, namely, the consistency, convergence rate, and asymptotic distribution, are established in Section 3. Section 4 presents the results obtained from an extensive simulation study conducted to assess the finite-sample performance of the proposed method, showing that it works well for practical situations. In Section 5, we apply the proposed approach to the CHNS, and Section 6 concludes the paper.

2. Estimation of Time-varying Covariate Effects

Consider a recurrent event study consisting of n independent subjects, and let $N_i(t)$ denote the underlying recurrent event process representing the total number of the occurrences of the recurrent event of interest up to time t for subject i. Assume that $N_i(t)$ is potentially observed only at $0 < t_{i,1} < \cdots < t_{i,m_i} < \tau$ and define $H_i^*(t) = \sum_{j=1}^{m_i} I(t_{i,j} \leq t)$, the underlying observation process. In practice, for each subject, there usually exists a following or stopping time C_i, and it follows that the realized censored observation process on the ith subject is given by $H_i(t) = H_i^*\{\min(C_i, t)\}$. In other words, one observes panel count data such that $N_i(t)$ is observed only at the time points where $H_i(t)$ jumps, for $i = 1, \ldots, n$.

For each subject, suppose that there exist two vectors of covariates, denoted by $\mathbf{W}_i(t) = (W_{i1}(t), \ldots, W_{ip_1}(t))^T$ and $\mathbf{Z}_i(t) = (Z_{i1}(t), \ldots, Z_{ip_2}(t))^T$, that are possibly time dependent. The former $\mathbf{W}_i(t)$ represents the covariates assumed to have constant effects, and $\mathbf{Z}_i(t)$ denotes the covariates assumed to have time-varying effects. To describe the covariate effects on $N_i(t)$, we assume that, given $\mathbf{W}_i(t)$ and $\mathbf{Z}_i(t)$, the recurrent event process $N_i(t)$ follows the conditional multiplicative mean model

$$E\{N_i(t)|\mathbf{W}_i(t), \mathbf{Z}_i(t)\} = \Lambda(t) \exp\{\boldsymbol{\gamma}^T \mathbf{W}_i(t) + \boldsymbol{\beta}^T(t)\mathbf{Z}_i(t)\} \ , \qquad (2.1)$$

where, $\Lambda(t)$ denotes an unknown baseline mean function, and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_{p_1})^T$ and $\boldsymbol{\beta}(t) = (\beta_1(t), \ldots, \beta_{p_2}(t))^T$ represent the constant and time-dependent coefficients, respectively. Futhermore, we assume that, given $\mathbf{W}_i(t)$ and $\mathbf{Z}_i(t)$, $N_i(t)$ and $H_i(t)$ are conditionally independent, and some comments on this are given in Section 6.

Let $\mathcal{M}_j$ denote the parameter space for β_j, for $j = 1, \ldots, p_2$. Because of the infinite dimension of each $\mathcal{M}_j$, the estimation of $\boldsymbol{\beta}(t)$ is usually infeasible. To deal with this, we propose employing the sieve approach to approximate $\boldsymbol{\beta}(t)$ by a linear combination of a finite number of basis functions, such as B-splines

(He et al. (2017)). More specifically, define a sequence of knots $\mathcal{T} = \{t_j\}_{j=1}^{m_n+2l}$, with $0 = t_1 = \cdots = t_l < t_{l+1} < \cdots < t_{m_n+l} < t_{m_n+l+1} = \cdots = t_{m_n+2l} = \tau$, that partition $[0, \tau]$ into $K_n + 1$ subintervals $[t_{l+j}, t_{l+j+1}]$, for $j = 0, \ldots, K_n$, where $K_n = O(n^\nu)$ and $\max_{0<j<m_n} |t_{j+1} - t_j| = O(n^{-\nu})$ for $\nu \in (0, 0.5)$. For simplicity, we assume the same knots for all $\beta_j(t)$ for $j = 1, \ldots, p_2$. Then, we can construct a sieve space to approximate $\mathcal{M}_j$ as

$$\mathcal{M}_{nj} = \left\{ \beta_{nj}(t) = \alpha_{j0} + \sum_{k=1}^{q_n} \alpha_{jk} B_k(t) = \mathbf{B}_n^T(t)\boldsymbol{\alpha}_j, \|\boldsymbol{\alpha}_j\|_\infty < M_n \right\},$$

which is the class of B-splines of order l with the knots sequence $\mathcal{T}$. In the above, M_n is some large number with $M_n \to \infty$ as $n \to \infty$, $q_n = K_n + l$, $\mathbf{B}_n(t) = \{1, B_1(t), \ldots, B_k(t)\}^T$ is a class of B-spline bases, and $\boldsymbol{\alpha}_{nj} = (\alpha_{nj0}, \alpha_{nj1}, \ldots, \alpha_{njq_n})$. Because the dimension of $\mathcal{M}_{nj}$, that is, q_n, is finite, we can now estimate the parameters in a much easier way by adopting existing methods for panel count data with a finite number of parameters.

Under the sieve space, by replacing $\boldsymbol{\beta}(t)$ with $\boldsymbol{\beta}_n(t)$, model (2.1) can be rewritten as

$$E\{N_i(t)|\mathbf{W}_i(t), \mathbf{Z}_i(t)\} = \Lambda_0(t) \exp\left\{ \boldsymbol{\gamma}^T \mathbf{W}_i(t) + \sum_{j=1}^{p_2} \left(\mathbf{B}_n^T(t)\boldsymbol{\alpha}_{nj} \right) Z_{ij}(t) \right\}$$

$$= \Lambda_0(t) \exp\left\{ \boldsymbol{\gamma}^T \mathbf{W}_i(t) + \boldsymbol{\alpha}_n^T \tilde{\mathbf{Z}}_i(t) \right\}$$

$$= \Lambda_0(t) \exp\left\{ \boldsymbol{\theta}_n^T \mathbf{X}_i(t) \right\}. \tag{2.2}$$

Here, $\tilde{\mathbf{Z}}_i(t) = \left(Z_{i1}(t)\mathbf{B}_n^T(t), Z_{i2}(t)\mathbf{B}_n^T(t), \ldots, Z_{ip_2}(t)\mathbf{B}_n^T(t) \right)^T$, $\boldsymbol{\alpha}_n = (\boldsymbol{\alpha}_{n1}^T, \boldsymbol{\alpha}_{n1}^T, \ldots, \boldsymbol{\alpha}_{np_2}^T)^T$, $\mathbf{X}_i(t) = (\mathbf{W}_i^T(t), \tilde{\mathbf{Z}}_i^T(t))^T$, and $\boldsymbol{\theta}_n = (\boldsymbol{\gamma}, \boldsymbol{\alpha}_n)^T$. Note that model (2.2) involves only time-independent covariate effects, and many existing methods can be used to estimate $\boldsymbol{\theta}_n$ without much modification. More specifically, motivated by Hu, Sun and Wei (2003), we propose using the estimating equation

$$\frac{1}{n} \sum_{i=1}^n \int_0^\tau Y_i(t) N_i(t) \left\{ \mathbf{X}_i(t) - \bar{\mathbf{X}}(t; \boldsymbol{\theta}_n) \right\} dH_i(t) = 0. \tag{2.3}$$

In the above, $Y_i(t) = I(C_i \geq t)$ is the at-risk indicator and $\bar{\mathbf{X}}(t; \boldsymbol{\theta}_n) = \mathbf{S}_1(t; \boldsymbol{\theta}_n)/S_0(t; \boldsymbol{\theta}_n)$, where

$$S_u(t; \boldsymbol{\theta}) = \frac{1}{n} \sum_{i=1}^n Y_i(t) \mathbf{X}_i^{\otimes u}(t) \exp\left(\boldsymbol{\theta}^T X_i(t) \right) dH_i(t),$$

$u = 0, 1, 2$ for $0 \leq t \leq \tau$, with $a^{\otimes 0} = 1$, $a^{\otimes 1} = a$, and $a^{\otimes 2} = aa^T$, for some vector a.

Let $\hat{\boldsymbol{\theta}}_n$ denote the estimator of $\boldsymbol{\theta}_n$ given by the solution to equation (2.3). Then, one can estimate $\beta_j(t)$ using $\hat{\beta}_j(t) = \mathbf{B}_n^T(t)\hat{\boldsymbol{\alpha}}_{nj}$. In practice, we may also be interested in estimating the baseline mean function $\Lambda(t)$, for which a natural estimator is given by the Breslow-type estimator

$$\hat{\Lambda}(t, \hat{\boldsymbol{\theta}}_n) = \sum_{i=1}^{n} \frac{Y_i(u)N_i(u)dH_i(u)}{nS_0(u; \hat{\boldsymbol{\theta}}_n)}.$$

3. Asymptotic Properties

Now, we establish the asymptotic properties of the estimators proposed in the previous section, including the consistency, convergence rate, and asymptotic normality. Let $\boldsymbol{\vartheta} = (\boldsymbol{\gamma}, \boldsymbol{\beta}, \Lambda)$, $\hat{\boldsymbol{\vartheta}}_n = (\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\beta}}_n, \hat{\Lambda})$, and $\boldsymbol{\vartheta}_0 = (\boldsymbol{\gamma}_0, \boldsymbol{\beta}_0, \Lambda_0)$ denote the true value of $\boldsymbol{\vartheta}$. In addition for convenience, let $\mathbf{V}(t) = \left(\mathbf{W}^T(t), \mathbf{Z}^T(t)\right)^T$, the population version of the combined covariates, and define the parameter space $\Theta = \mathcal{A} \times \mathcal{M} \times \mathcal{F}$, where $\mathcal{A}$, $\mathcal{M} = \prod_{j=1}^{p_2} \mathcal{M}_j$, and $\mathcal{F}$ denote the parameter spaces of γ, $\boldsymbol{\beta}$, and Λ, respectively. Let $\mathfrak{B}^d$ denote the collection of Borel sets in $\mathbb{R}^d$ and $\mathcal{L}_2[0, \tau]$ the collection of Borel sets in $\mathcal{L}_2$ on $[0, \tau]$, and define $\mathfrak{B}^1[0, \tau] = \left\{B \cap [0, \tau] : B \in \mathfrak{B}^1\right\}$ and $\mathfrak{L}_2^d[0, \tau] = \underbrace{\mathcal{L}_2[0, \tau] \times \cdots \times \mathcal{L}_2[0, \tau]}_{d}$.

In addition, define the measure

$$v_1(B_1 \times B_2 \times B_3) = \int_{B_3 \times B_2} \int_{B_1} dE[Y(t)H(t)] d\mu_{Z \times W}$$

for $B_1 \in \mathfrak{B}^1[0, \tau]$, $B_2 \in \mathfrak{L}_2^{p_2}[0, \tau]$, and $B_3 \in \mathfrak{L}_2^{p_1}[0, \tau]$, where $\mu_{Z \times W}$ denotes the joint probability measure for $\mathbf{W}(t)$ and $\mathbf{Z}(t)$. Alternatively, from the definition of $\mathbf{V}$, we have $\mu_V = \mu_{Z \times W}$ and can rewrite $v_1(B_1 \times B_2 \times B_3)$ as

$$v_1(B_1 \times B_4) = \int_{B_4} \int_{B_1} dE[Y(t)H(t)] d\mu_V,$$

for $B_1 \in \mathfrak{B}^1[0, \tau]$ and $B_4 \in \mathfrak{L}_2^p[0, \tau]$, and $\mu_1(B) = v_1(B \times \mathfrak{L}_2^p[0, \tau])$. Define the L_2 metric $d(\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2)$ on Θ as

$$d(\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2)^2 = \|\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2\|_2^2 + \int \|\boldsymbol{\beta}_1(U) - \boldsymbol{\beta}_2(U)\|_2^2 d\mu_1(U)$$
$$+ \int (\Lambda_1(U) - \Lambda_2(U))^2 d\mu_1(U).$$

To establish the asymptotic results, we need the following regularity conditions:

(C1) The observation process has the rate function $E\left[dH^*(t)|\mathbf{W}(t), \mathbf{Z}(t), C\right] = \omega(t)dt$, where $\omega(t)$ is a bounded, nonnegative, and continuous function on $[0, \tau]$. There exists a positive integer D_0 such that $\Pr\left(H^*(\tau) < D_0\right) = 1$. That is, the total observation number is finite. Moreover, the support of $\omega(t)$ is $[\tau_0, \tau]$, with $\tau_0 > 0$ and $\Lambda_0(\tau_0) > 0$, for some constant τ_0.

(C2) The measure $\mu_1 \times \mu_V$ is absolutely continuous with respect to υ_1 and $\mu_1(\{\tau\}) > 0$.

(C3) The parameter space of Λ, that is, $\mathcal{F}$, consists of bounded nondecreasing functions in $\mathcal{L}_2$ over $[0, \tau]$.

(C4) The parameter space of $\boldsymbol{\beta}_j$, that is, $\mathcal{M}_j$, is bounded and convex in $\mathcal{L}_2([0, \tau])$ for each $j = 1, \ldots, p_2$. Each component of the true value of $\boldsymbol{\beta}(t)$, denoted by $\beta_{0j}(t)$, for $j = 1, \ldots, p_2$, is continuously rth differentiable in $[0, \tau]$, for $r + 1 \le l$.

(C5) The parameter space of $\boldsymbol{\gamma}$, that is, $\mathcal{A}$, is bounded and convex in $\mathbb{R}^{p_1}$.

(C6) The covariate vector $\mathbf{V}(t) = \left(\mathbf{W}^T(t), \mathbf{Z}^T(t)\right)^T$ is uniformly bounded over $[0, \tau]$ with the distribution μ_V.

(C7) Given $\mathbf{V}(t)$, for $t \in [0, \tau]$, C and N are independent. Futhermore, with probability one,

$$\inf_{\mathbf{V}(t), t \in [0, \tau]} \Pr\left(C \ge \tau | \mathbf{V}(t) = \mathbf{v}(t), t \in [0, \tau]\right)$$
$$= \inf_{\mathbf{V}(t), t \in [0, \tau]} \Pr\left(C = \tau | \mathbf{V}(t) = \mathbf{v}(t), t \in [0, \tau]\right) > 0.$$

(C8) If $\boldsymbol{\gamma}^T \mathbf{W}(t) + \boldsymbol{\beta}^T(t)\mathbf{Z}(t) \equiv 0$ for $t \in [0, \tau]$ with probability one for some $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$, then $\boldsymbol{\gamma} = 0$ and $\boldsymbol{\beta}(t) = 0$ for $t \in [0, \tau]$ with probability one.

(C9) The random function $M_0(\mathbf{V}) = \int N(t) \log(N(t)) \, dH(t)$ satisfies $E[M_0(\mathbf{V})] < \infty$.

(C10) $E\left[\exp(cN(t))\right]$ is bounded in $[0, \tau]$ for some constant $c > 0$.

(C11) The true baseline mean function Λ_0 is differentiable in $[\tau_0, \tau]$. Moreover, the lower and upper bounds of its first-order derivative are positive and finite on $[\tau_0, \tau]$.

(C12) There exist $\eta_1 \in (0,1)$ such that $a^T \text{Var}\left(\mathbf{V}(U)|U\right)a \geq \eta_1 a^T E(\mathbf{V}(U)\mathbf{V}^T(U)|$
$U)a$, a.s. for all $a \in \mathbb{R}^{p_1+p_2}$, where $(U, \mathbf{V})$ has distribution $\nu_1/\nu_1\left(\mathbb{R}^+ \times \mathcal{V}\right)$,
where $\mathcal{V}$ is the support of $\mathbf{V}$.

Note that conditions (C1) and (C7) are common in observation schemes and similar to the combination of C8, C10, and C11 in Lu, Zhang and Huang (2009). Condition (C2) comes from the conditions in Theorem 1 of Wellner and Zhang (2007) and Theorem 1 of Lu, Zhang and Huang (2009), ensuring that $\hat{\Lambda}$ is bounded, and conditions (C6)–(C11) are common assumptions in semiparametric estimation. Conditions (C2) and (C8) ensure the identifiability of the semiparametric model, and conditions (C9), (C10), and (C11) are technical assumptions, similar to conditions C4, C10, and C12 in Wellner and Zhang (2007). Condition (C12) is needed to prove the convergence rate and can be justified by arguments similar to those in Wellner and Zhang (2007). Now, we are ready to establish the asymptotic properties of the proposed estimator.

Theorem 1 (Consistency)**.** *Assume that the regularity conditions* (C1)–(C9) *given above hold. Then, we have that* $d(\hat{\boldsymbol{\vartheta}}_n, \boldsymbol{\vartheta}_0) \to 0$ *in probability as* $n \to \infty$.

Note that the proof of the consistency does not rely on the continuity of the baseline function Λ, but does require the differentiability of the time-varying coefficients. This is because the B-spline approximation usually works well when the true time-varying coefficient functions are smooth, to some degree. To derive the convergence rate in the next theorem, we need condition (C11), to control the smoothness of Λ, and condition (C12).

Theorem 2 (Rate of Convergence)**.** *Assume that the regularity conditions* (C1)– (C12) *given above hold. Then, we have that*

$$n^{\min\left\{n^{(1-\nu)/3}, n^{r\nu}\right\}} d\left(\hat{\boldsymbol{\vartheta}}_n, \boldsymbol{\vartheta}_0\right) = O_p\left(1\right),$$

with the optimal rate $O_p(n^{-r/(3r+1)})$ *achieved at* $\nu = 1/\left(1 + 3r\right)$.

Note that the order of the optimal rate $n^{-r/(3r+1)}$ is slower than $n^{-r/(2r+1)}$ in Lu, Zhang and Huang (2009) because the nonparametric parameter Λ is essentially estimated using a step function, whereas $\boldsymbol{\beta}(t)$ is estimated using B-splines. Nevertheless, we can still derive the asymptotic distribution of $\hat{\boldsymbol{\gamma}}$ with the rate of convergence $n^{-1/2}$. The next theorem establishes the asymptotic normality of $\hat{\boldsymbol{\theta}}_n$ in a form similar to that of He et al. (2017).

Theorem 3 (Asymptotic Normality)**.** *Assume that the regularity conditions* (C1)–(C12) *given above hold, and that* $(4r)^{-1} < \nu < 2^{-1}$, *with* $r > 1$. *Define*

$\mathcal{H}_1 = \{\boldsymbol{h}_1 : \boldsymbol{h}_1 \in \mathcal{A}, \|\boldsymbol{h}_1\| \leq 1\}$, $\mathcal{H}_2 = \{\boldsymbol{h}_2 : \boldsymbol{h}_2 \in \mathcal{M}, \text{each component of } \boldsymbol{h}_2 \text{ is of bounded total variation.}\}$, and $\mathcal{H}_3 = \{h_3 : h_3 \text{ is a function with bounded total variation in } [0, \tau] \text{ and } h_3(0) = 0\}$. *Then, for some* $(\boldsymbol{h}_1, \boldsymbol{h}_2, h_3) \in \mathcal{H}_1 \times \mathcal{H}_2 \times \mathcal{H}_3$, *we have that*

$$\sqrt{n}\,(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)^T \boldsymbol{h}_1 + \sqrt{n} \int_0^\tau \left(\hat{\boldsymbol{\beta}}(t) - \boldsymbol{\beta}_0(t)\right)^T d\boldsymbol{h}_2(t) + \sqrt{n} \int_0^\tau \left(\hat{\Lambda}(t) - \Lambda_0(t)\right) dh_3(t)$$
$$\to_d N\left(0, \sigma^2\right),$$

where σ^2 *is given in the Supplementary Material.*

Proofs of the above results are provided in the Supplementary Material. These results suggest that one can asymptotically approximate the distribution of $\hat{\boldsymbol{\gamma}}$ by the normal distribution. However, note that, similarly to He et al. (2017), we cannot determine the explicit form of the asymptotic distribution because the explicit forms of $(\boldsymbol{h}_1, \boldsymbol{h}_2, h_3)$ cannot be determined even though they exist. To estimate the asymptotic covariance matrix of $\hat{\boldsymbol{\gamma}}$ and the pointwise asymptotic variance of $\hat{\boldsymbol{\beta}}(t)$ in $t \in [0, \tau]$ or the covariance matrix of $\hat{\boldsymbol{\theta}}_n$, based on the estimating equation (2.3) and by following Amorim et al. (2008) and Hu, Sun and Wei (2003), we propose employing the robust sandwich-type estimator $\hat{\Sigma} = \hat{\mathbf{A}}^{-1} \hat{\mathbf{D}} \hat{\mathbf{A}}^{-1}$. Here,

$$\hat{\mathbf{A}} = \left.\frac{\partial \mathbf{U}(\boldsymbol{\theta}_n)}{\partial \boldsymbol{\theta}_n}\right|_{\boldsymbol{\theta}_n = \hat{\boldsymbol{\theta}}_n} = -\sum_{i=1}^n \int_0^\tau Y_i(t) N_i(t) \left\{ \frac{\mathbf{S}_2(t; \hat{\boldsymbol{\theta}}_n)}{S_0(t; \hat{\boldsymbol{\theta}}_n)} - \left(\frac{\mathbf{S}_1(t; \hat{\boldsymbol{\theta}}_n)}{S_0(t; \hat{\boldsymbol{\theta}}_n)}\right)^{\otimes 2} \right\} dH_i(t),$$

and

$$\hat{\mathbf{D}} = \sum_{i=1}^n \left[\int_0^\tau Y_i(t) \left(N_i(t) - \hat{\Lambda}(t, \hat{\boldsymbol{\theta}}_n) \exp(\hat{\boldsymbol{\theta}}_n^T \mathbf{X}_i(t)) \right) \left\{ \mathbf{X}(t) - \bar{\mathbf{X}}(t; \hat{\boldsymbol{\theta}}_n) \right\} dH_i(t) \right]^{\otimes 2}.$$

It follows from the above that the asymptotic covariance matrix of $\hat{\boldsymbol{\gamma}}$ can be estimated by the top-left $p_1 \times p_1$ sub-matrix of $\hat{\Sigma}$, denoted by $\hat{\Sigma}_{\hat{\boldsymbol{\gamma}}}$. Furthermore, because $\hat{\beta}_j = \mathbf{B}_n^T(t) \hat{\boldsymbol{\alpha}}_j$, the pointwise asymptotic variance of $\hat{\beta}_j(t)$ for a given t can be estimated by $\hat{\Omega}_j(t) = \mathbf{B}_n^T(t) \hat{\Sigma}_{\hat{\boldsymbol{\alpha}}_j} \mathbf{B}_n(t)$, where $\hat{\Sigma}_{\hat{\boldsymbol{\alpha}}_j}$ denotes the diagonal block sub-matrix of $\hat{\Sigma}$ formed from the $(p_1 + (j-1)p_2 + 1)$th to the $(p_1 + jp_2)$th rows and columns, corresponding to the estimated covariance matrix of $\hat{\boldsymbol{\alpha}}_j$. The numerical study in the next section suggests that the estimators above work well. As an alternative, of course, one may estimate the asymptotic covariance matrix of $\hat{\boldsymbol{\gamma}}$ using the bootstrap procedure, as shown in the next section, although it is relatively more time consuming.

4. A Simulation Study

In this section, we present the results obtained from an extensive simulation study conducted to evaluate the finite-sample performance of the proposed estimation procedure. In the study, we considered four covariates, two with constant effects and two with time-varying effects. That is, $p_1 = p_2 = 2$. More specifically, we assumed that $W_1(t)$ and $Z_1(t)$ are time-dependent covariates given by $W_1(t) = B_{11}I(t \leq V_1) + B_{12}I(t > V_1)$ and $Z_1(t) = B_{21}I(t \leq V_2) + B_{22}I(t > V_2)$, where B_{j1}, B_{j2}, and V_j are generated independently from uniform distributions over $(0, 0.5)$, $(0.5, 1)$, and $(0, \tau)$, with $\tau = 1$, respectively, for $j = 1, 2$. Furthermore, $W_2(t)$ and $Z_2(t)$ are assumed to be time independent and are generated independently from the uniform distribution over $(0, 1)$.

Note that the covariates generated above are mutually independent. Corresponding to this, we also considered the situation where they are dependent. Here, we first generated $(L_1, \ldots, L_6)$ from the multivariate normal distribution with $E[L_j] = 1$, for $j = 1, \ldots, 6$, and $\text{cov}(L_j, L_k) = 0.25$ if $j \neq k$ and 1 if $j = k$, for $j, k = 1, \ldots, 6$. Then, we generated $W_1(t)$ and $Z_1(t)$ as above, but setting $B_{11} = \Phi(L_1)/2$, $B_{12} = (\Phi(L_1) + \Phi(L_2))/2$, $B_{21} = \Phi(L_3)/2$, and $B_{22} = (\Phi(L_3) + \Phi(L_4))/2$, with V_1 and V_2 generated independently from the uniform distribution over $(0, \tau)$. In addition, we set $W_2 = \Phi(L_5)/2$ and $Z_2 = \Phi(L_6)/2$. For the regression parameters, we set $\gamma_1 = 0.5$ and $\gamma_2 = 0.5$ and considered two setups for $\beta(t)$:

Scenario 1 $\beta_1(t) = t$ and $\beta_2(t) = t^2$;

Scenario 2 $\beta_1(t) = (\sin(4\pi t) + 4\pi t)/12$ and $\beta_2(t) = (\cos(4\pi t) + 4\pi t)/12$.

To generate the panel count data, we first generated the observation times $t_{i,j}$ from the non-homogeneous Poisson process with mean function $3t+4$, and the follow-up times C_i from the uniform distribution over $(0.9\tau, \tau)$. Then, given the covariates and the real observation times, the panel count data were generated by $N_i(t_{i,j}) = \sum_{k=1}^{j} N_i^*(t_{i,k}) - N_i^*(t_{i,k-1})$, with $t_{i,0} = 0$, where $N_i^*(t_{i,k}) - N_i^*(t_{i,k-1})$ follows the Poisson distribution with the mean

$$v_i \Lambda_0(t_{i,k}) e^{(\gamma^T \mathbf{W}_i(t_{i,k}) + \beta^T(t_{i,j})\mathbf{Z}_i(t_{i,j}))} - v_i \Lambda_0(t_{i,j-1}) e^{(\gamma^T \mathbf{W}_i(t_{i,j-1}) + \beta^T(t_{i,j-1})\mathbf{Z}_i(t_{i,j-1}))},$$

given v_i. Here, we set $\Lambda_0(t) = 2t+3$ or $\Lambda_0(t) = (\sin(4\pi t) + 4\pi t)/2$, and considered two settings for v_i. One is that $v_i = 1$ for all i, meaning that N_i are Poisson processes with the mean function $\Lambda_0(t) \exp\{\gamma^T \mathbf{W}_i(t) + \beta^T(t)\mathbf{Z}_i(t)\}$. The other is that v_i follows a gamma distribution with mean one and variance $\sigma^2 = 1$, meaning that $N_i(t)$ are mixed Poisson processes with the same mean function as

Table 1. Simulation results of estimating γ_1 and γ_2, with $\Lambda_0(t) = 2t + 3$.

Poisson Model	Para	Bias	SE	Robust		Bootstrap	
				ESE	CP	ESE	CP
Scenario 1 for $\beta(t)$							
Independent Covariates							
Yes	γ_1	-0.0002	0.0518	0.0479	0.942	0.0494	0.949
	γ_2	-0.0016	0.0529	0.0477	0.928	0.0494	0.935
No	γ_1	0.0079	0.2321	0.2078	0.919	0.2144	0.929
	γ_2	0.0004	0.2327	0.2086	0.917	0.2157	0.921
Dependent Covariates							
Yes	γ_1	-0.0002	0.0635	0.0479	0.928	0.0654	0.936
	γ_2	-0.0023	0.1567	0.1426	0.922	0.1478	0.929
No	γ_1	0.0030	0.2680	0.2466	0.931	0.2537	0.941
	γ_2	0.0003	0.5968	0.5514	0.934	0.5724	0.945
Scenario 2 for $\beta(t)$							
Independent Covariates							
Yes	γ_1	-0.0002	0.0486	0.0449	0.920	0.0463	0.926
	γ_2	-0.0000	0.0495	0.0446	0.918	0.0460	0.924
No	γ_1	-0.0144	0.2238	0.2079	0.928	0.2145	0.944
	γ_2	0.0082	0.2240	0.2080	0.920	0.2146	0.928
Dependent Covariates							
Yes	γ_1	-0.0023	0.0654	0.0607	0.927	0.0624	0.938
	γ_2	0.0011	0.1475	0.1265	0.920	0.1416	0.932
No	γ_1	-0.0020	0.2602	0.2460	0.930	0.2539	0.931
	γ_2	0.0064	0.6156	0.5444	0.921	0.5650	0.929

above. The results given below are based on $n = 300$ with 1,000 replications.

Table 1 presents the results of estimating two time-independent regression coefficients γ_1 and γ_2, with $\Lambda_0(t) = 2t + 3$. Here, we used cubic B-splines with three internal knots equally spaced on $(0, \tau)$. The results include the empirical bias (BIAS), given by the average of the estimates minus the true value, the sampling standard deviation (SE), the average of the estimated standard errors (ESE), and the 95% empirical coverage probability (CP). Note that for the variance estimation, in addition to the robust estimation discussed above, we also applied a simple bootstrap procedure with 200 bootstrapped samples for comparison. The results suggest that the proposed estimators seem to be unbiased and the variance estimations appear to be appropriate. In addition, the results on CP indicate that the normal approximation to the distribution of the proposed estimator $\hat{\gamma}$ seems to be reasonable.

Figure 1 gives the averages of the estimated $\beta_1(t)$ and $\beta_2(t)$ where $N_i(t)$ is a Poisson process with independent covariates over 1,000 equal-spaced grid points on the time axis. For comparison, the true function is also included in

Table 2. Simulation results of estimating γ_1 and γ_2, with $\Lambda_0(t) = (\sin(4\pi t) + 4\pi t)/2$.

Poisson Model	Para	Bias	SE	Robust		Bootstrap	
				ESE	CP	ESE	CP
			Scenario 1 for $\beta(t)$				
			Independent Covariates				
Yes	γ_1	-0.0009	0.0521	0.0485	0.933	0.0501	0.941
	γ_2	0.0012	0.0531	0.0485	0.925	0.0501	0.929
No	γ_1	-0.0023	0.2484	0.2217	0.919	0.2289	0.928
	γ_2	-0.0028	0.2508	0.2212	0.926	0.2291	0.932
			Dependent Covariates				
Yes	γ_1	0.0016	0.0659	0.0631	0.931	0.0651	0.939
	γ_2	0.0013	0.1490	0.1372	0.923	0.1420	0.936
No	γ_1	0.0035	0.2766	0.2573	0.928	0.2659	0.938
	γ_2	0.0027	0.6337	0.5610	0.916	0.5852	0.927
			Scenario 2 for $\beta(t)$				
			Independent Covariates				
Yes	γ_1	-0.0002	0.0476	0.0449	0.931	0.0501	0.950
	γ_2	-0.0001	0.0489	0.0449	0.922	0.0501	0.943
No	γ_1	0.0104	0.2505	0.2199	0.918	0.2270	0.920
	γ_2	-0.0157	0.2442	0.2200	0.921	0.2273	0.923
			Dependent Covariates				
Yes	γ_1	0.0009	0.0662	0.0608	0.929	0.0628	0.932
	γ_2	-0.0005	0.1436	0.1344	0.942	0.1390	0.941
No	γ_1	-0.0009	0.2750	0.2559	0.930	0.2641	0.940
	γ_2	-0.0166	0.6369	0.5594	0.923	0.5825	0.930

the figure; once again, the proposed procedure seems to yield unbiased estimates. Furthermore, Figure 1 shows the respective averages of the estimated pointwise standard errors by the robust and bootstrap methods, along with the pointwise empirical standard errors of $\boldsymbol{\beta}(t)$. Again they indicate that the proposed method appears to give reasonable variance estimates. The same can be seen from Figure 2, which provides the results of estimating $\beta_1(t)$ and $\beta_2(t)$, as in Figure 1, except for a mixed Poisson process with independent covariates.

Table 2 provides the results of estimating two time-independent regression coefficients γ_1 and γ_2, obtained under the same setup as Table 1, except that $\Lambda_0(t) = (\sin(4\pi t) + 4\pi t)/2$. The corresponding results of estimating $\beta_1(t)$ and $\beta_2(t)$ are given in Figures 3–4. It is apparent that the conclusions are similar. Futhermore, the same is true for the dependent covariates, for which the results are provided in the Supplementary Material, owing to the space limitations. We also considered other setups, including different degrees of B-spline functions, different numbers of interior knots, and other types of covariates, and obtained similar results.
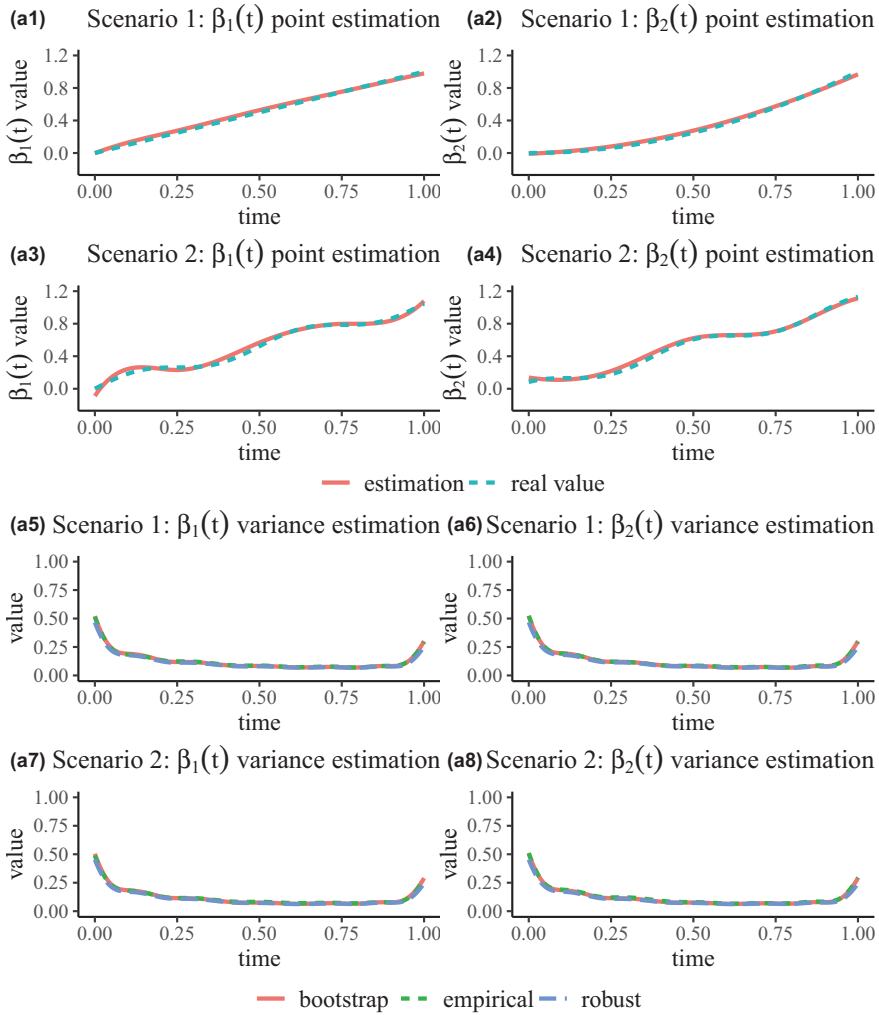
Figure 1. Simulation results of estimating $\beta_1(t)$ and $\beta_2(t)$ under the Poisson process with $\Lambda_0(t) = 2t + 3$ and independent covariates. (a1) to (a4): results of estimating $\beta_1(t)$ and $\beta_2(t)$; (a5) to (a8): results on variance estimation of $\hat{\beta}_1(t)$ and $\hat{\beta}_2(t)$.

5. Analysis of the CHNS

In this section, we apply the proposed methodology to data from the CHNS, an international collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute for Nutrition and Health at the Chinese Center for Disease Control and Prevention. The survey took place over a seven-day period using a multistage, random cluster process to draw a sample of over 11,000 households, representing over 42,000 individuals participating in 15 provinces and municipal cities that vary substan-
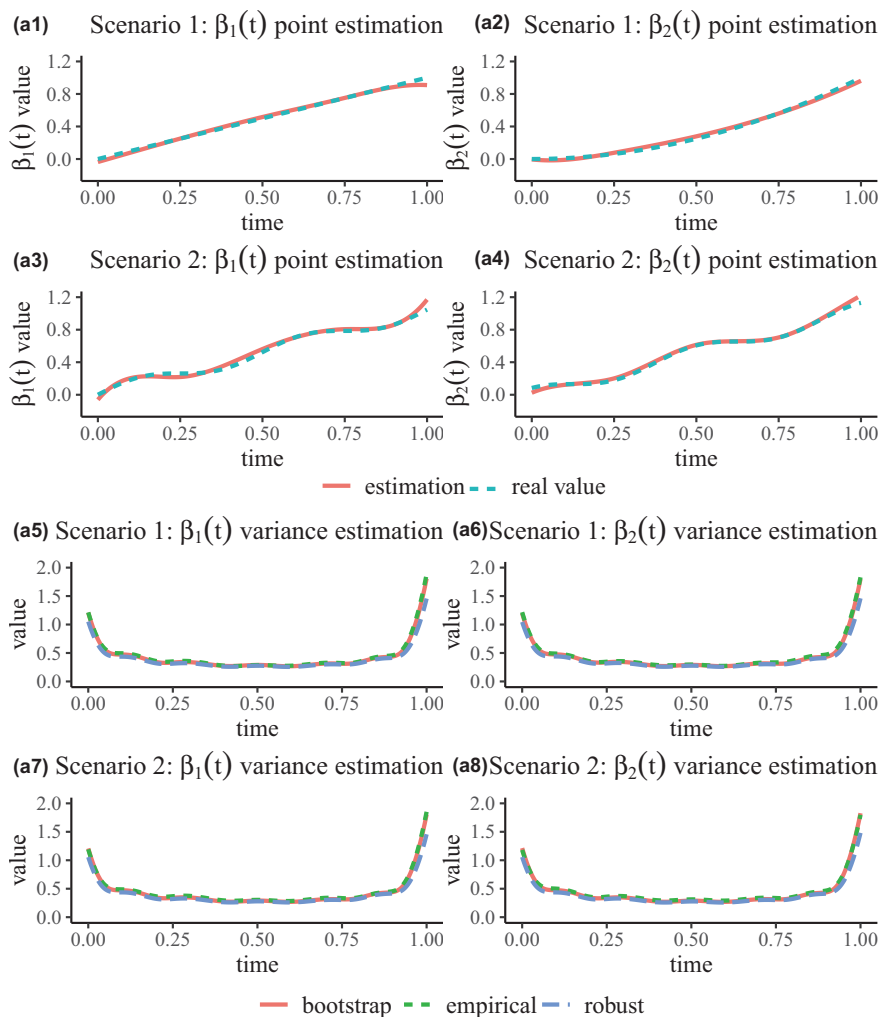
(a1) Scenario 1: $\beta_1(t)$ point estimation   **(a2)** Scenario 1: $\beta_2(t)$ point estimation

(a3) Scenario 2: $\beta_1(t)$ point estimation   **(a4)** Scenario 2: $\beta_2(t)$ point estimation

— estimation - - real value

(a5) Scenario 1: $\beta_1(t)$ variance estimation   **(a6)** Scenario 1: $\beta_2(t)$ variance estimation

(a7) Scenario 2: $\beta_1(t)$ variance estimation   **(a8)** Scenario 2: $\beta_2(t)$ variance estimation

— bootstrap - - empirical -· robust

Figure 2. Simulation results of estimating $\beta_1(t)$ and $\beta_2(t)$ under the nonPoisson process with $\Lambda_0(t) = 2t + 3$ and independent covariates. (a1) to (a4): results of estimating $\beta_1(t)$ and $\beta_2(t)$; (a5) to (a8): results on variance estimation of $\hat{\beta}_1(t)$ and $\hat{\beta}_2(t)$.

tially in terms of geography, economic development, public resources, and health indicators. Villages and townships within counties and urban/suburban neighborhoods within cities were selected randomly.

Among others, one objective of the CHNS is to assess the relationship between the pregnancy process of the female participants and various factors. Furthermore, owing to the nature of the study, only panel count data are available for the pregnancy occurrence process. For the analysis below, we focus on the 2,537 female participants with complete information on the following four covariates:
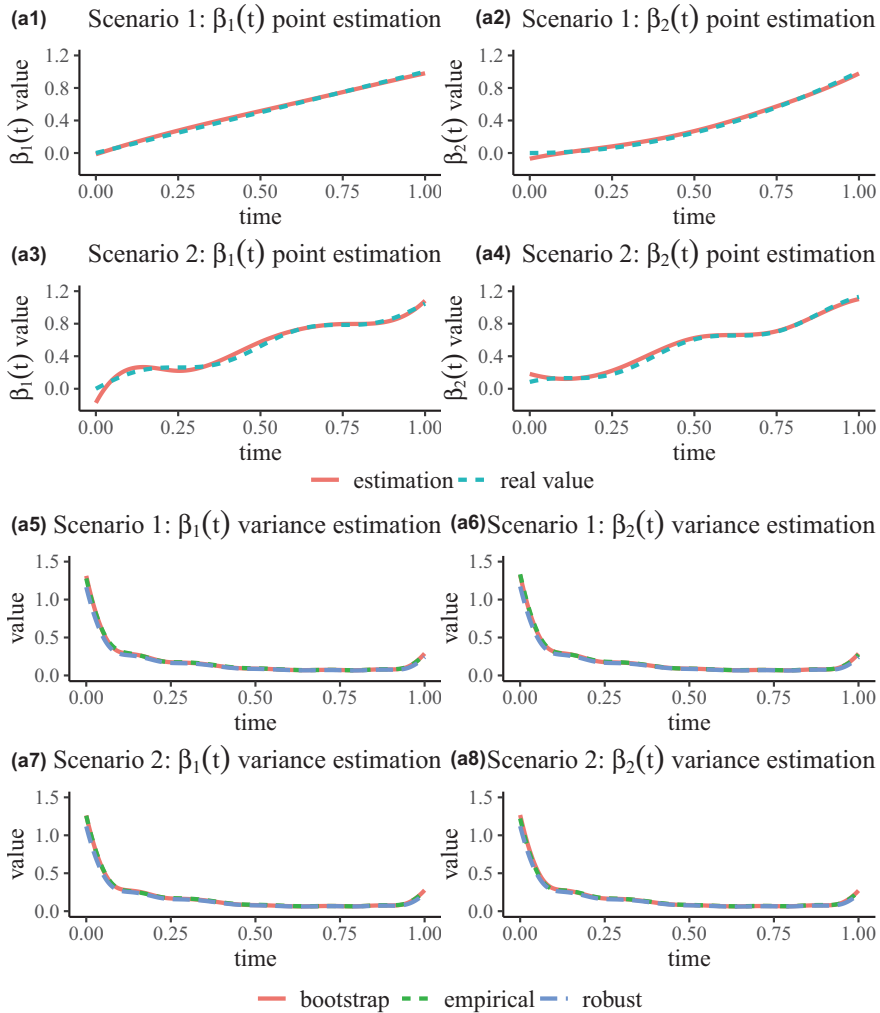
(a1)   Scenario 1: $\beta_1(t)$ point estimation   **(a2)** Scenario 1: $\beta_2(t)$ point estimation

(a3)   Scenario 2: $\beta_1(t)$ point estimation   **(a4)** Scenario 2: $\beta_2(t)$ point estimation

— estimation - - real value

(a5) Scenario 1: $\beta_1(t)$ variance estimation **(a6)**Scenario 1: $\beta_2(t)$ variance estimation

(a7) Scenario 2: $\beta_1(t)$ variance estimation **(a8)**Scenario 2: $\beta_2(t)$ variance estimation

— bootstrap - - empirical — robust

Figure 3. Simulation results of estimating $\beta_1(t)$ and $\beta_2(t)$ under the Poisson process with $\Lambda_0(t) = (\sin(4\pi t) + 4\pi t)/2$ and independent covariates. (a1) to (a4): results of estimating $\beta_1(t)$ and $\beta_2(t)$; (a5) to (a8): results on variance estimation of $\hat{\beta}_1(t)$ and $\hat{\beta}_2(t)$.

whether the mother came from a urban or a rural area (0: urban, 1: rural), the average monthly wage in the previous year, the completed formal education level in regular school (0: No school; 1: Primary school; 2: Middle school; 3: Technical school; 4: College)), and the current health status (1: Excellent, 2: Good, 3: Fair, 4: Poor). Table 3 gives a descriptive summary of the four covariates at the baseline, and the range of the pregnancies is from one to eight.

To apply the proposed estimation procedure, we first assume that all four covariates have time-varying effects; Figure 5 presents the estimated covariate
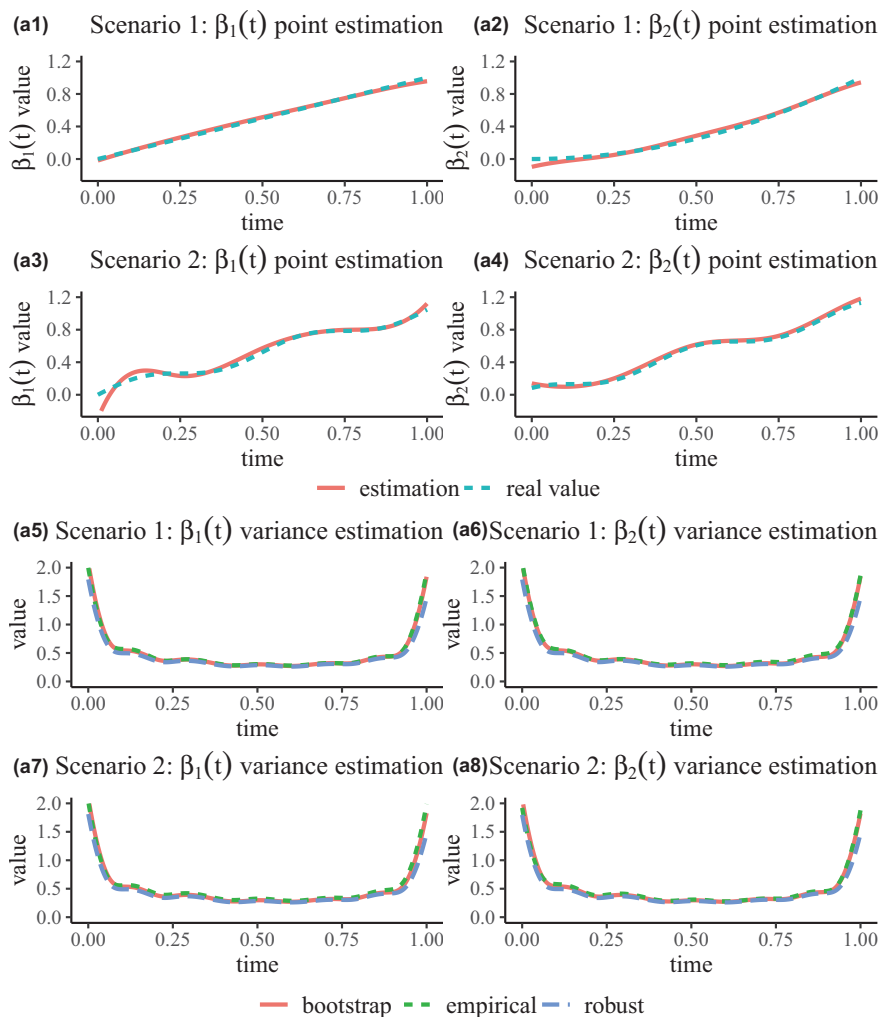
Figure 4. Simulation results of estimating $\beta_1(t)$ and $\beta_2(t)$ under the nonPoisson process with $\Lambda_0(t) = (\sin(4\pi t) + 4\pi t)/2$ and independent covariates. (a1) to (a4): results of estimating $\beta_1(t)$ and $\beta_2(t)$; (a5) to (a8): results on variance estimation of $\hat{\beta}_1(t)$ and $\hat{\beta}_2(t)$.

effects with three interior knots for the B-splines. The results suggest that the mother's location seems to have a significantly positive relationship with the fertility occurrence rate and the effect magnitude appears to increase with a mothers' age. In other words, mothers from rural areas are more likely to have more children than those from urban areas, and the differences between the number of pregnancies increases as the women get older. In contrast, the mother's education level seems to have a significantly negative effect on the number of pregnancies,

indicating that well-educated mothers tend to have fewer children, and the effect increases as the women become older. Furthermore, it seems that the average monthly wage and mother's health status have no significant time-varying effects, or might only have constant effects on the pregnancy process.

To further investigate the above estimated effects, we repeat the analysis by adding the interaction effect between location and education level; the results are presented in Figure 6. The results are similar for the four individual factors, suggesting that there are significantly negative interaction effects for mid-aged mothers. In other words, the education level effect magnitude increases faster for female subjects in rural areas. Note that one may not want to pay much attention to the estimated effects after age 60, owing to the sparsity of the observed data.

Based on the results above, we again apply the proposed estimation procedure, but assume that only the location, the education level, and their interaction may have time-varying effects. In other words, it is supposed that the average monthly wage and health status have only constant effects on the pregnancy process. Table 4 gives the estimated covariate effects for the average monthly wage and current health status with three, five, or seven interior knots for the B-splines. The estimated time-dependent effects based on three interior knots are presented in Figure 7; the results with five or seven interior knots are similar, and are provided in the Supplementary Material. One can see from Table 4 that a mother's average monthly wage is significantly negatively correlated with her fertility rate, and mothers with lower income tend to have more children. In contrast, the health status level has positive effects on the fertility rate, and all results are consistent with respect to the number of interior knots. Finally, the results given in Figure 7 are similar to those given in Figure 6, and again indicate the existence of the time-varying effect and the necessity of using the proposed method.

6. Conclusion

In this paper, we have discussed the regression analysis of panel count data in the presence of both time-dependent covariates and time-varying covariate effects. We developed a spline-based estimating equation procedure and established the asymptotic properties of the proposed estimators. An extensive numerical study indicated that the proposed method works well in practical situations. In addition, the usefulness and necessity of the proposed estimation procedure was illustrated by applying it to data from the CHNS, identifying some time-varying covariate effects.

Table 3. Summary of the four covariates at the baseline for the CHNS.

Covariate	Mean ± SD	Median	Range
Average Monthly Wage Last Year (Yuan)	$1,635.13 \pm 3,741.61$	800	$8 - 99,999$

	Category	Count	Percentage
Location			
	Rural	1,366	53.84%
	Urban	1,171	46.16%
Education Level			
	0: No school	89	3.51%
	1: Primary school	359	14.15%
	2: Middle school	1,373	54.12%
	3: Technical school	330	13.01%
	4: College	386	15.21%
Current Health Status			
	1:Excellent	410	16.16%
	2:Good	1,351	54.25%
	3:Fair	713	28.10%
	4:Poor	63	2.48%

Table 4. The estimated time-constant effects of the average monthly wage (wage) and the current health status (health) for the CHNS.

# of interior knots	Covariate	Estimated effect	SD	95% CI
3	wage	-0.0301	0.0152	(-0.0598, -0.0003)
	health	0.0287	0.0111	(0.0069, 0.0505)
5	wage	-0.0300	0.0150	(-0.0595, -0.0004)
	health	0.0292	0.0111	(0.0074, 0.0510)
7	wage	-0.0300	0.0150	(-0.0590, -0.0004)
	health	0.0290	0.0111	(0.0072, 0.0508)

There exist several directions for future research. First, in the proposed method, we have assumed that the observation process does not depend on covariates and is also independent of the underlying recurrent event process of interest. In practice, this may not be true (Sun and Zhao (2013)). To develop a valid estimation procedure, one may need to model the three processes together. A second assumption is that the underlying recurrent event process follows the proportional mean model, meaning that the mean functions associated with any two sets of covariate values are proportional over time. Sometimes, this may be too restrictive (Lin, Wei and Ying (2001)) and as such, one may want to consider other models, such as the class of semiparametric transformation mean models, when developing estimation procedures (Li, Sun and Sun (2010)).

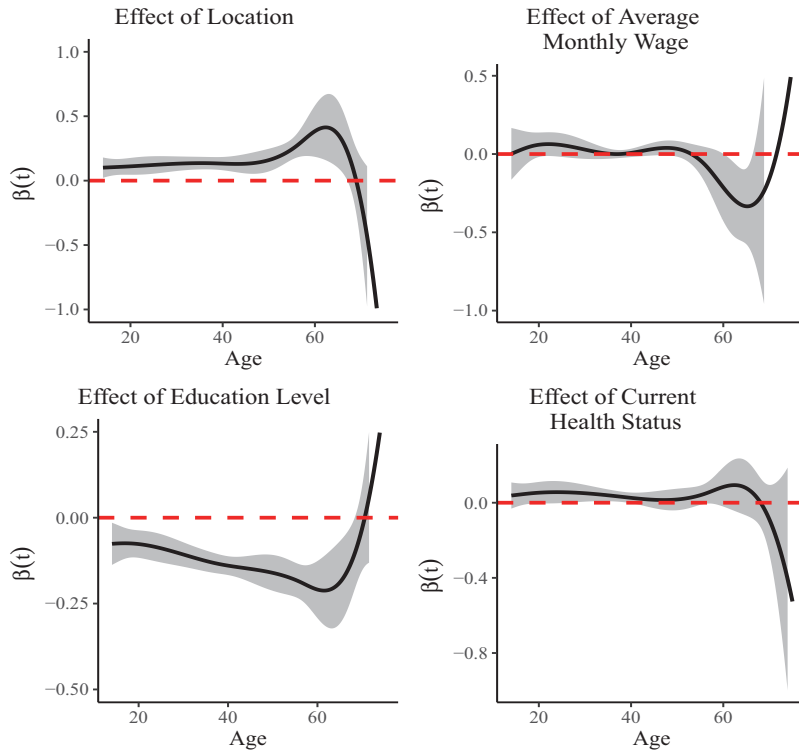In the proposed estimation procedure, we also assumed that one knows which
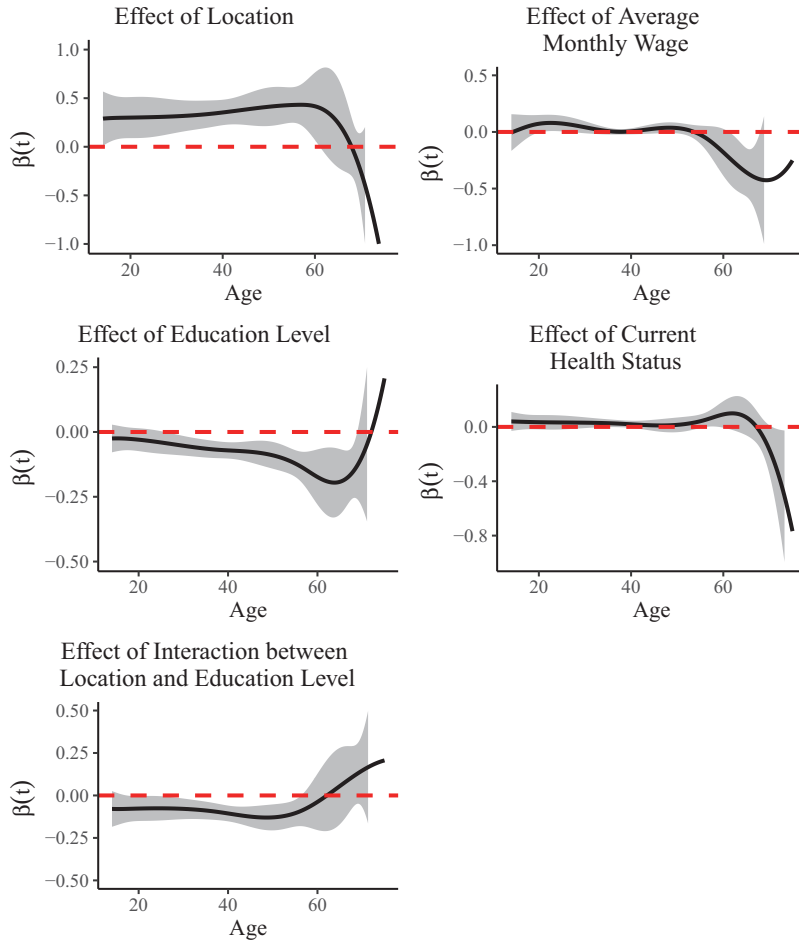
Figure 5. Estimated time-varying effects of all four covariates (solid curves) and corresponding pointwise 95% confidence intervals (ribbons).

covariates have time-varying effects or time-constant effects. In practice, however, this is usually unknown. A simple approach is to try different combinations, as shown in the application above. On the other hand, it would be useful to develop a data-driven procedure to separate the two types of covariates. In practice, a common problem of interest is hypothesis testing about $\beta_0(\cdot)$. Theoretically, one could use the results given in Theorem 3 to develop a test procedure. However, this may not be straightforward because we cannot derive the explicit relationship between $\boldsymbol{h}_2$ and $\boldsymbol{h}_1$, as well as that between $\boldsymbol{h}_2$ and h_3. In other words, we cannot construct a variance estimation for $\beta(t)$ based on Theorem 3 directly. An alternative to this is to apply the bootstrap procedure.

Supplementary Material

The online Supplementary Material includes extra simulation results, additional real-data analysis results, and proofs of the asymptotic theorems.

Figure 6. Estimated time-varying effects of all four covariates and the interaction (solid curves) between them, and corresponding pointwise 95% confidence intervals (ribbons).
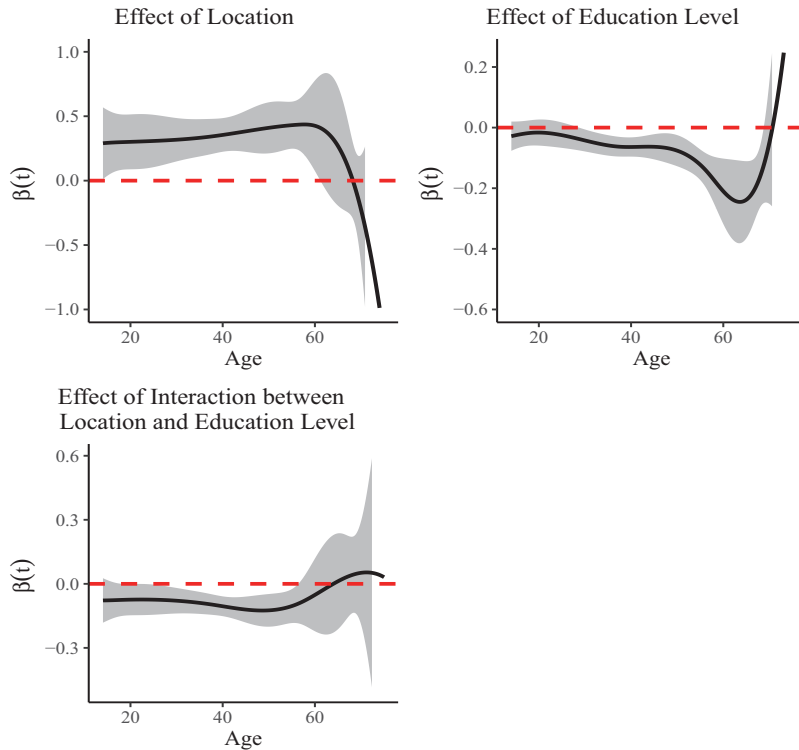
Acknowledgments

Figure 7. Estimated time-varying effects of the location and education level and the interaction between them (solid curves), and corresponding pointwise 95% confidence intervals (ribbons) with three interior knots.

References

Amorim, L. D., Cai, J., Zeng, D. and Barreto, M. L. (2008). Regression splines in the time-dependent coefficient rates model for recurrent event data. *Statistics in Medicine* **27**, 5890–5906.

Cook, R. and Lawless, J. (2007). *The Statistical Analysis of Recurrent Events*. Springer-VerIagn, New York.

He, X., Feng, X., Tong, X. and Zhao, X. (2017). Semiparametric partially linear varying coefficient models with panel count data. *Lifetime Data Analysis* **23**, 439–466.

Hu, X. J., Sun, J. and Wei, L.-J. (2003). Regression parameter estimation from panel counts.

Scandinavian Journal of Statistics **30**, 25–43.

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.

Lin, D. Y., Wei, L. J. and Ying, Z. (2001). Semiparametric transformation models for point processes. *Journal of the American Statistical Association* **96**, 620–628.

Lin, H., Fei, Z. and Li, Y. (2016). A semiparametrically efficient estimator of the time-varying effects for survival data with time-dependent treatment. *Scandinavian Journal of Statistics* **43**, 649–663.

Lu, M., Zhang, Y. and Huang, J. (2009). Semiparametric estimation methods for panel count data using monotone B-splines. *Journal of the American Statistical Association* **104**, 1060–1070.

Oliveira, J. (2016). The value of children: Inter-generational support, fertility, and human capital. *Journal of Development Economics* **120**, 1–16.

Perperoglou, A. (2014). Cox models with dynamic ridge penalties on time-varying effects of the covariates. *Statistics in Medicine* **33**, 170–180.

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

Tian, L. (2018). Less is more: The effects of children on parents' income, assets and intergenerational transfers. *Innovation in Aging* **2**, 468.

Tian, L., Zucker, D. and Wei, L. J. (2005). On the Cox model with time-varying regression coefficients. *Journal of the American Statistical Association* **100**, 172–183.

Wang, Y. and Yu, Z. (2021). A kernel regression model for panel count data with time-varying coefficients. *Statistica Sinica* **31**, 1–19.

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

Yu, Z. and Lin, X. (2010). Semiparametric regression with time-dependent coefficients for failure time data analysis. *Statistica Sinica* **20**, 853–869.

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

Yuanyuan Guo

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

E-mail: yg882@mail.missouri.edu

Dayu Sun

Department of Biostatistics and Bioinformatics, Emory University, 1518 Clifton Rd, Atlanta, GA 30322, USA.

E-mail: dayu.sun@emory.edu

Jianguo Sun

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

E-mail: sunj@missouri.edu