Statistica Sinica **29** (2019), 2141-2161
doi:https://doi.org/10.5705/ss.202016.0348

MAXIMUM PARTIAL-RANK CORRELATION ESTIMATION FOR LEFT-TRUNCATED AND RIGHT-CENSORED SURVIVAL DATA

Shao-Hsuan Wang and Chin-Tsang Chiang

National Taiwan University

Abstract: This article presents a general single-index hazard regression model to assess the effects of covariates on a failure time. Based on left-truncated and right-censored survival data, a new partial-rank correlation function is proposed to estimate the index coefficients in the presence of covariate-dependent truncation and censoring. Furthermore, an efficient computational algorithm is proposed for the computation that maximizes the constructed target function. The developed approach can be extended to include right-truncation and left-censoring under a reverse-time hazard regression model. Based on the maximum rank correlation estimator in the literature, we establish the consistency and asymptotic normality of the maximum partial-rank correlation estimator. A series of simulations shows that the proposed estimator has satisfactory finite-sample performance compared with that of its competitors. Lastly, we demonstrate our methodology by applying it to data from the US Health and Retirement Study.

Key words and phrases: Asymptotic normality, consistency, left-censoring, left-truncation, partial-rank correlation estimation, rank correlation estimation, random weighted bootstrap, right-censoring, right-truncation, U-statistic.

1. Introduction

In survival analyses, samples are often collected using the incident and prevalent cohort sampling schemes. Owing to cost and time constraints, the prevalent cohort approach is generally the more efficient of the two in terms of collecting sufficient failure cases, especially when studying a rare disease. For example, the US Health and Retirement Study (HRS) initially recruited noninstitutionalized persons using a cross-sectional sampling criterion between 1992 and 1993. Prior to the recruitment, the white nonHispanic men and women in these data had experienced initiating events (birth), but not failure events (death). In the context of this research, the survival time T^* and the truncation time A^* are defined as the times from the calendar date of the initiating event to the calendar dates of the failure event and recruitment, respectively. On the other hand, individuals who had experienced failure events before the recruitment are not observed and, thus, their failure times are left-truncated (i.e. $\{T^* \le A^*\}$). If individuals are not available for follow-up or if they drop out during the study period, their failure times are right-censored. Under a general single-index hazard regression model, we develop a new approach to estimate the index coefficients based on left-truncated and right-censored survival data.

To simplify the presentation, the observed survival time, truncation time, and covariates are denoted by the triplets (T, A, Z), and their joint distribution is the same as the conditional distribution of (T^*, A^*, Z^*) on $\{T^* \ge A^*\}$, where $Z^* = (Z_1^*, \ldots, Z_p^*)^\top$ are the covariates of interest with support $\mathcal{Z}$. We also assume a mild covariate-dependent truncation and censoring mechanism in our study. In addition, we explore the relation between the continuous failure time T^* and the covariates Z^* using the following single-index hazard regression model:

$$\lambda(t|z) = \lambda(t, \beta_0^\top z), \quad t \in [0, \tau], \tag{1.1}$$

where $\lambda(t|z)$ is the hazard function of T^*, given $Z^* = z$, $\lambda(t, u)$ is an unknown nonnegative bivariate function that is strictly decreasing in u for each t, $\beta_0^\top z$ is a single index, with β_0 denoting the index coefficients, and τ is the end of the study period. Owing to the identifiability of β_0 up to a scale, a parametric system is adopted, where $\beta_0 = (1, \theta_0^\top)^\top$ and $\theta_0 = (\theta_{01}, \ldots, \theta_{0p-1})^\top$ is an interior point of the compact parameter space $\Theta \subseteq \mathbb{R}^{p-1}$. As a result, the coefficient β_{0k} $(= \theta_{0k-1})$ is interpreted as the relative effect of Z_k^*, compared with Z_1^*, on the hazard function, for $k = 2, \ldots, p$. Specific forms of model (1.1) include the proportional hazard regression $\lambda_0(t)\exp(\beta_0^\top z)$ (Cox (1972)), the additive hazard regression $\lambda_0(t) + \beta_0^\top z$ (Aalen (1980)), and the transformation regression model $H(T^*) = -\beta_0^\top Z^* + \varepsilon$ (Cheng, Wei and Ying (1995)) with a monotonic hazard function of ε, where $\lambda_0(t)$ is an unknown baseline hazard function, $H(\cdot)$ is an unknown increasing function, and ε is a random error. Under Cox's proportional hazard model and the assumption of covariate-dependent truncation and censoring, Wang, Brookmeyer and Jewell (1993) proposed a maximum partial likelihood estimator of β_0 based on left-truncated and right-censored survival data. Lin and Ying (1994) provided an estimation for the additive hazard model with censored survival data, which Huang and Qin (2013) extended to include left-truncation, proposing a modified conditional estimating equation estimator. Currently, there is no estimation approach for the transformation model when the distributions of the truncation and censoring times are covariate-dependent.

In the data analysis of the HRS, the violation of the covriate-dependent truncation is supported by the proposed Hausman-type test. Therefore, the existing approaches in the literature for left-truncated and right-censored survival data with stationary or nonstationary disease incidence are not appropriate for describing the effects of body mass index (BMI), level of education, and smoking status on life expectancy. In addition, our testing procedure confirms the inadequacy of the proportional and additive hazard regression models. Thus, the general formulation in model (1.1) and its statistical inferences become necessary in application. Based on a new partial-rank correlation function, we develop an approach to estimate β_0 in model (1.1). We also propose an efficient algorithm is provided to compute the maximum partial-rank correlation estimator. The partial-rank correlation estimation of Khan and Tamer (2007) for right-censored survival data shows that each pair of units should be comparable in the constructed estimation criterion. In our estimation, the compared units are further required to come from the same truncated population. Moreover, the consistency and asymptotic normality of the proposed estimator can be established using the theoretical frameworks in Han (1987) and Sherman (1993). Interestingly, the developed approach can also be extended to estimate the regression coefficients in a reverse-time hazard regression model

$$\lambda_r(t|z) = \lambda_r(t, \beta_0^\top z) \tag{1.2}$$

with right-truncated and left-censored survival data, where $\lambda_r(t|z) = f(t|z)/(1 - S(t|z))$. Here, $f(t|z)$ and $S(t|z)$ are the conditional density and survival functions, respectively, of $T^* = t$ on $Z^* = z$, and $\lambda_r(t, u)$ is an unknown nonnegative bivariate function and is strictly decreasing in u for each t.

The remainder of the paper is organized as follows. In Section 2, a partial-rank correlation function is proposed as the basis for estimating model (1.1) with left-truncated and right-censored survival data. The index coefficients β_0 are further shown to be the unique maximizer of the constructed partial-rank correlation function. Moreover, an extension to model (1.2) with right-truncated and left-censored survival data is given in this section. Section 3 outlines the maximum partial-rank correlation estimation and the corresponding computational algorithm. Then, we establish the consistency and asymptotic normality of the estimator and the bootstrap approximation of the sampling distribution of the estimator. In Section 4, Monte Carlo simulations are used to investigate the finite-sample performance of the proposed estimator and its competitors. The HRS data are analyzed in Section 5 to show the usefulness of our methodology. Section 6 summarizes our findings and remarks on possible future research. The proofs of the main results are relegated to the appendix.

2. Partial-Rank Correlation Function and its Extension

For survival data with left-truncation and right-censoring, we develop an approach to estimate β_0 based on a new partial-rank correlation function. Under model (1.1) and some suitable conditions, β_0 is further shown to be the unique maximizer of this target function. In fact, the proposed estimation criterion includes several specific cases and can be reasonably extended to right-truncation and left-censoring. To simplify the presentation, let C represent the residual censoring time after the recruitment, $Y = \min\{T, A + C\}$ be the last observed time, and $\delta = \mathrm{I}(T \leq A + C)$ be a noncensoring indicator with $\mathrm{I}(\cdot)$ as the indicator function. The notations $\wedge$ and $\vee$ denote the minimum and maximum, respectively.

2.1. Partial-rank correlation function

Given any two independent units (T_1^*, Z_1^*) and (T_2^*, Z_2^*), an essential element of our partial-rank correlation function is given by

$$Q(z_1, z_2, a, c) = \mathrm{P}(T_1^* > T_2^* > a, T_2^* < c | Z_1^* = z_1, Z_2^* = z_2), \tag{2.1}$$

which is easily shown to be

$$\int_a^c S(u|z_1)S(u|z_2)\lambda(u|z_2)du \quad \forall\ c > a \geq 0. \tag{2.2}$$

The reason for adopting a truncation value a and a censoring value c in $\mathrm{Q}(z_1, z_2; a, c)$ is to adjust for the sampling bias caused by the left-truncation and to make each pair of units comparable in the presence of right-censoring. By the symmetric feature of $S(t|z_1)S(t|z_2)$ with respect to (z_1, z_2) and assumption **A1** ($\inf_{\{z\in\mathcal{Z}\}} S(\tau|z) > 0$ and $\sup_{\{z\in\mathcal{Z}\}} S_{A^*}(\tau|z) < 1$), where $S_{A^*}(a|z)$ is a survival function of A^*, given $Z^* = z$, model (1.1) further implies that $\lambda(t|z_2) > \lambda(t|z_1)$ whenever $\beta_0^\top z_1 > \beta_0^\top z_2$. Thus, the following lemma is a direct consequence.

Lemma 1. *Suppose that model* (1.1) *is valid and assumption A1 is satisfied. Then, for any* $z_1, z_2 \in \mathcal{Z}$ *and* $\tau > c > a \geq 0$

$$Q(z_1, z_2; a, c) - Q(z_2, z_1; a, c) > 0 \quad \textit{whenever} \quad \beta_0^\top z_1 > \beta_0^\top z_2.$$

From the proof of the maximizer of the rank correlation function in Han (1987), it follows from (2.1), Lemma 1, and the equality

$$\mathrm{E}[Q(Z_1^*, Z_2^*; a, c)\mathrm{I}(\beta_0^\top Z_1^* > \beta_0^\top Z_2^*) - Q(Z_1^*, Z_2^*; a, c)\mathrm{I}(\beta^\top Z_1^* > \beta^\top Z_2^*)]$$

$$= \frac{1}{2}\mathrm{E}[(Q(Z_1^*, Z_2^*; a, c) - Q(Z_2^*, Z_1^*; a, c))(\mathrm{I}(\beta_0^\top Z_1^* > \beta_0^\top Z_2^*) - \mathrm{I}(\beta^\top Z_1^* > \beta^\top Z_2^*))]$$

that β_0 is a maximizer of

$$\mathrm{E}[Q(Z_1^*, Z_2^*; a, c)\mathrm{I}(\beta^\top Z_1^* > \beta^\top Z_2^*)] = \mathrm{P}(T_1^* > T_2^* > a, T_2^* < c, \beta^\top Z_1^* > \beta^\top Z_2^*) \quad \forall\ \tau > c > a \geq 0. \tag{2.3}$$

For general forms of censoring, Khan and Tamer (2007) proposed a partial-rank correlation estimation for β_0. Moreover, the authors showed that β_0 is the unique maximizer of their partial-rank correlation function and that the rank correlation estimation criterion by Han (1987) is infeasible for censored survival data. For right-censored survival data, their partial-rank correlation function is constructed using $\mathrm{P}(Y_1 > Y_2, \delta_2 = 1, \beta^\top Z_1 > \beta^\top Z_2)$. In terms of $Q(z_1, z_2; a, c)$ in (2.1), this can be expressed as

$$\mathrm{E}[Q(Z_1^*, Z_2^*, 0, C_1 \wedge C_2)\mathrm{I}(\beta^\top Z_1^* > \beta^\top Z_2^*)]. \tag{2.4}$$

Instead of imposing the assumption of independent censoring, this approach relies on subjects whose failure times are comparable. More precisely, T_1^* and T_2^* are said to be comparable if the indicator status $\mathrm{I}(T_1^* > T_2^*)$ can be fully determined based on (Y_ℓ, δ_ℓ), for $\ell = 1, 2$. In conjunction with the presence of left-truncation, we provide a more general covariate-dependent truncation and censoring assumption **A2** ($A^* \perp T^* | Z^*$ and $C \perp (T, A) | Z$). By adjusting for the truncation bias, the following partial-rank correlation function is proposed as the basis for the estimation of β_0:

$$\mathrm{C}(\beta) = \mathrm{E}\left[Q(Z_1, Z_2; A_1 \vee A_2, (C_1 + A_1) \wedge (C_2 + A_2))\mathrm{I}(\beta^\top Z_1 > \beta^\top Z_2)\right]. \tag{2.5}$$

Coupled with the expression of $Q(z_1, z_2; a, c)$ in (2.1) and the equality $P(Y_1 > Y_2 > (A_1 \vee A_2), \delta_2 = 1 | Z_1 = z_1, Z_2 = z_2) = P(T_1 > T_2 > (A_1 \vee A_2), (A_1 + C_1) > (A_2 + C_2) | Z_1 = z_1, Z_2 = z_2)$, an alternative probability representation can be derived as

$$\mathrm{C}(\beta) = \mathrm{P}(Y_1 > Y_2 > (A_1 \vee A_2), \delta_2 = 1, \beta^\top Z_1 > \beta^\top Z_2). \tag{2.6}$$

With the observable random quantities (Y, δ, A, Z) in left-truncated and right-censored survival data, our approach requires that each pair of units is comparable and comes from the same truncated population. Figure 1 displays the relative positions of the calendar times of the initiating events, recruitment, failure events, and censoring events of the two independent units. The resulting truncation, failure, and censoring times satisfy the constraints $Y_1 > Y_2 > (A_1 \vee A_2)$ and $\delta_2 = 1$ in (2.6).

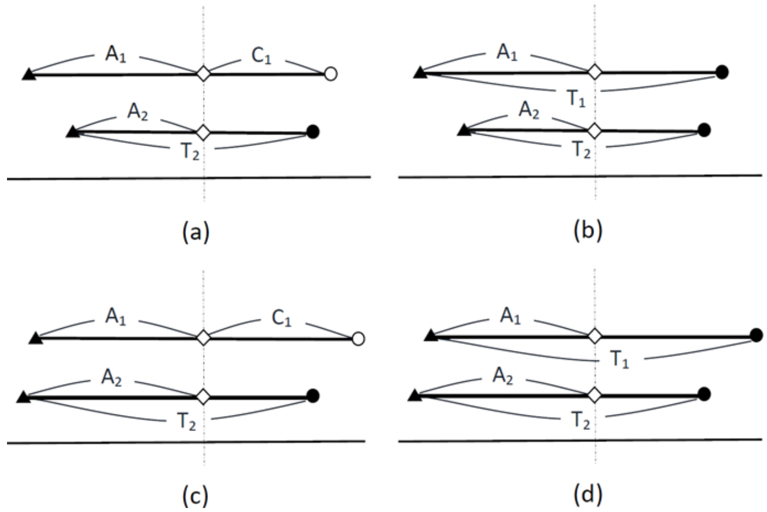

Figure 1. Relative positions of the calendar times of the initiating event (▲), recruitment (◇), the failure event (●), and the censoring event (○).

To derive the main results, we require an additional assumption:

A3. $\mathcal{Z}$ is not contained in any proper linear subset of $\mathbb{R}^p$ and Z^* has a positive Lebesgue density everywhere.

As in the context of rank correlation estimation, assumption A3 ensures the uniqueness of β_0. Under some suitable conditions, β_0 is shown to be the unique maximum of $\mathrm{C}(\beta)$ as follows.

Theorem 1. *Under model* (1.1) *and assumptions* A1–A3*,*

$$\beta_0 = \underset{\{\beta(\theta):\theta\in\Theta\}}{\operatorname{argmax}}\ \mathrm{C}(\beta).$$

Proof. See the Appendix.

Note that the device $Q(z_1, z_2; a, c)$ in (2.1) can also accommodate the following particular cases:

Case 1. (complete failure time data) For such data, the conditions $A_1 = A_2 = 0$ and $C_1 = C_2 = \infty$ are naturally set in (2.5) and assumption A2 is automatically satisfied. The rank correlation function $\mathrm{P}(T_1 > T_2, \beta^\top Z_1 > \beta^\top Z_2)$ of Han (1987) is easily derived as

$$\mathrm{E}[Q(Z_1, Z_2; 0, \infty)\mathrm{I}(\beta^\top Z_1 > \beta^\top Z_2)]. \tag{2.7}$$

Case 2. (right-censored survival data) In the presence of right-censoring, A_1 and A_2 are set to zero in (2.5) and assumption A2 can be simplified to $C \perp$

$(T, A)|Z$. The partial-rank correlation function $\mathrm{P}(Y_1 > Y_2, \delta_2 = 1, \beta^\top Z_1 > \beta^\top Z_2)$ of Khan and Tamer (2007) is the same, with the following form:

$$\mathrm{E}[Q(Z_1, Z_2; 0, C_1 \wedge C_2)\mathrm{I}(\beta^\top Z_1 > \beta^\top Z_2)]. \tag{2.8}$$

Case 3. (left-truncated survival data) For data subject to left-truncation only, it is natural to specify $C_1 = C_2 = \infty$ and then to modify assumption A2 as $A^* \perp T^*|Z^*$. In this setup, our partial-rank correlation function can be rewritten as

$$\begin{aligned}&\mathrm{E}\Big[Q(Z_1, Z_2; A_1 \vee A_2, \infty)\mathrm{I}(\beta^\top Z_1 > \beta^\top Z_2)\Big]\\ &= \mathrm{P}(T_1 > T_2 > (A_1 \vee A_2), \beta^\top Z_1 > \beta^\top Z_2).\end{aligned} \tag{2.9}$$

2.2. An extension to right-truncated and left-censored data

In insurance applications and AIDS cohort studies (e.g., Kaminsky (1987) and Kalbfleisch and Lawless (1991)), the chronological times of initiating and consequent events, say X_0^* and $(X_0^* + T^*)$, of individuals are available only if $(X_0^* + T^*)$ falls within some chronological period $[0, \tau]$, that is, $X_0^* + T^* \le \tau$. As shown in the Australian AIDS data (Cui (1999)), $(X_0^* + T^*)$ may not be recorded before the chronological time X_1, which can be a determined or random time, with $X_1 \le \tau$. Thus, the lag T^* between events is right-truncated by $D^* = \tau - X_0^*$ and left-censored by $C_\tau = X_1 - X_0^*$, and the triplets (T^*, D^*, Z^*) are observed only if $\{D^* \ge T^*\}$. Lagakos, Barraj and Gruttola (1998) and Cui (1999) show that the reverse survival time $S^* = \tau - T^*$, X_0^*, and $\tau - X_1$ can be regarded as the failure time, truncation time, and residual censoring time, respectively, in left-truncated and right-censored survival data. As a result, the reverse-time hazard function $\lambda_r(t|z)$ is conveniently approached and explained. Thus, model (1.1) is adopted in the reverse-time hazard regression model (1.2).

Let $(T, D = \tau - X_0, Z)$ represent the observed lag, right-truncated time, and covariates, with the joint distribution of (T, D, Z) being the same as the conditional distribution of (T^*, D^*, Z^*) on $\{D^* \ge T^*\}$. This can be transferred to the setup of the triplets (S, X_0, Z), which have the same joint distribution of (S^*, X_0^*, Z^*), given $\{S^* \ge X_0^*\}$. By substituting $(S_\ell^*, \tau - X_{1\ell}, X_{0\ell})$ for $(T_\ell^*, C_\ell, A_\ell)$ in (2.5) and $(Y_{c\ell}, \delta_{c\ell}, X_{0,\ell})$ for $(Y_\ell, \delta_\ell, A_\ell)$ in (2.6), where $Y_{c\ell} = \min\{S_\ell, \tau - X_{1\ell} + X_{0\ell}\} = \min\{S_\ell, \tau - C_{\tau\ell}\}$ and $\delta_{c\ell} = I(S_\ell \le \tau - X_{1\ell} + X_{0\ell}) = I(S_\ell \le \tau - C_{\tau\ell})$, for $\ell = 1, 2$, we have the following partial-rank correlation function:

$$\mathrm{C}_\tau(\beta) = \mathrm{E}\Big[Q(Z_1, Z_2; X_{01} \vee X_{02}, (\tau - X_{11} + X_{01})$$

$$\left. \wedge (\tau - X_{12} + X_{02}))\mathrm{I}(\beta^\top Z_1 > \beta^\top Z_2) \right]$$

$$= \mathrm{P}(Y_{c1} > Y_{c2} > (X_{01} \vee X_{02}), \delta_{c2} = 1, \beta^\top Z_1 > \beta^\top Z_2). \tag{2.10}$$

In terms of the definition of (S, X_0, X_1), an alternative probability representation of $\mathrm{C}_\tau(\beta)$ can be derived as

$$P(Y_{\tau 1} < Y_{\tau 2} < (D_1 \wedge D_2), \delta_{\tau 2} = 1, \beta^\top Z_1 > \beta^\top Z_2), \tag{2.11}$$

where $Y_{\tau\ell} = \max\{T_\ell, C_{\tau\ell}\}$ and $\delta_{\tau\ell} = I(T_\ell \geq C_{\tau\ell})$, for $\ell = 1, 2$. Under model (1.2) and assumptions A1, **A2*** ($T^* \perp D^* | Z^*$ and $(D - C_\tau) \perp (T, D)|Z$), and A3, β_0 is immediately shown to be the unique maximizer of $\mathrm{C}_\tau(\beta)$. For data subject to right-truncation only, assumption A2* can be modified as $T^* \perp D^* | Z^*$ and the partial-rank correlation function in (2.11) can be rewritten as

$$P(T_1 < T_2 < (D_1 \wedge D_2), \beta^\top Z_1 > \beta^\top Z_2). \tag{2.12}$$

3. Statistical Inferences

The maximum partial-rank correlation estimator of β_0 is proposed as a maximizer of a sample analogue of $\mathrm{C}(\beta)$. An effective computational algorithm is provided to implement such an optimization problem. In addition, we establish the consistency and asymptotic normality of the estimator and a weighted bootstrap approximation of the sampling distribution of the estimator.

3.1. Estimation and computational algorithm

Based on the constructed partial-rank correlation function in (2.6) and left-truncated and right-censored survival data of the form $\{(Y_i, \delta_i, A_i, Z_i)\}_{i=1}^n$, a sample analogue of $\mathrm{C}(\beta)$ is naturally given by a U-statistic of the form:

$$\mathrm{C}_n(\beta) = \frac{1}{n(n-1)} \sum_{i \neq j} \mathrm{I}(Y_i > Y_j > (A_i \vee A_j), \delta_j = 1, \beta^\top Z_i > \beta^\top Z_j). \tag{3.1}$$

Given that β_0 is a maximizer of $\mathrm{C}(\beta)$, we estimate β using the maximizer

$$\hat{\beta} \in \underset{\{\beta(\theta):\theta \in \Theta\}}{\operatorname{argmax}} \ \mathrm{C}_n(\beta). \tag{3.2}$$

In application, an easily implemented numerical algorithm is necessary to compute the maximum partial-rank correlation estimator $\hat{\beta}$. For the constrained optimization of $\mathrm{C}_n(\beta)$, a direct maximization is usually impractical and difficult because the target function is not differentiable with respect to β.

Using complete failure time data, Wang and Chiang (2017) provided an effective procedure to compute the maximization of the rank correlation function with respect to the coefficients of a generalized single-index. Indeed, their algorithm can also be adopted to compute a maximizer $\hat{\beta}$ of $\mathrm{C}_n(\beta)$. Let a smoothed counterpart of $\mathrm{C}_n(\beta)$ be defined as

$$\mathrm{C}_{n\sigma}(\beta) = \frac{1}{n(n-1)} \sum_{i\neq j}^{n} \mathrm{I}(Y_i > Y_j > (A_i \vee A_j), \delta_j = 1) s\left(\frac{\beta^\top Z_i - \beta^\top Z_j}{\sigma}\right), \quad (3.3)$$

where $s(v) = 1/(1+\exp(-v))$ is a sigmoid function and σ is a tuning parameter. In addition, let $g_{n\sigma}(\beta)$ denote the gradient function of $\mathrm{C}_{n\sigma}(\beta)$, $(\epsilon_1, \epsilon_2, r)$ be prechosen positive values with $0 < r < 1$, and $\|\cdot\|$ be the Euclidean norm of a vector. Provided that $\sigma = o(1/\sqrt{n})$, Ma and Huang (2005) showed that the maximizer of $\mathrm{C}_{n\sigma}(\beta)$ and $\hat{\beta}$ have the same asymptotic distribution. The following computational algorithm is shown by Wang and Chiang (2017) to be theoretically valid and practically feasible for computing $\hat{\beta}$:

Step 1. Set the initial values of (β, σ) as $(\hat{\beta}^{(0)}, \sigma^{(0)})$ and the step length as α.

Step 2. Refine $\sigma^{(k)}$ as $\sigma^{(k+1)} = r\sigma^{(k)}$ if $|\rho^{(k)} - 1| > \epsilon_1$, and set $\sigma^{(k+1)}$ as $\sigma^{(k)}$ otherwise, where $\rho^{(k)} = \mathrm{C}_{0n}(\hat{\beta}^{(k)})/\mathrm{C}_{n\sigma^{(k)}}(\hat{\beta}^{(k)})$ for $k \geq 0$.

Step 3. Set $\hat{\beta}^{(k+1)}$ as $\hat{\beta}^{(k)}$ if $g_{n\sigma^{(k+1)}}(\hat{\beta}^{(k)}) = 0$ or $g_{n\sigma^{(k+1)}}(\hat{\beta}^{(k)}) \propto \hat{\beta}^{(k)}$, and set $\hat{\beta}^{(k+1)}$ as $\hat{\beta}^{(k)} + \alpha p^{(k)}/\|p^{(k)}\|$ otherwise, where $p^{(k)} = \big(I_p - \hat{\beta}^{(k)}\hat{\beta}^{(k)\top}/\|\hat{\beta}^{(k)}\|^2\big)$ $g_{n\sigma^{(k+1)}}(\hat{\beta}^{(k)})$.

Step 4. Repeat Steps 2–3 until $|\rho^{(K)} - 1| < \epsilon_1$ and $\|\hat{\beta}^{(K+1)} - \hat{\beta}^{(K)}\|/\|\hat{\beta}^{(K)}\| < \epsilon_2$ for some integer K, and compute $\hat{\beta}$ as $\hat{\beta}^{(K)}/\hat{\beta}_1^{(K)}$, where $\hat{\beta}_1^{(K)}$ is a coefficient estimator of Z_1.

With an appropriate choice of ϵ_1, the rate of $\sigma^{(k)}$ can be adjusted to $o(1/\sqrt{n})$ after some iterations. The R code of the above algorithm can also be found in Chen and Chiang (2018) at the Biometrics website on Wiley Online Library.

Remark 1. In the spirit of our estimation, an estimator can also be proposed for the index coefficients in model (1.2). Based on right-truncated and left-censored survival data of the form $\{(Y_{\tau i}, \delta_{\tau i}, D_i, Z_i)\}_{i=1}^n$, β_0 is estimated by a maximizer of the following sample analogue of $\mathrm{C}_\tau(\beta)$ in (2.11):

$$\mathrm{C}_{\tau n}(\beta) = \frac{1}{n(n-1)} \sum_{i\neq j} \mathrm{I}(Y_{\tau i} < Y_{\tau j} < (D_i \wedge D_j), \delta_{\tau j} = 1, \beta^\top Z_i > \beta^\top Z_j). \quad (3.4)$$

3.2. Consistency, asymptotic normality, and bootstrap approximation

Let $\mathcal{N}_\theta$ be a neighborhood of θ in Θ, X denote the vector $(T, C, A, Z^\top)^\top$, $\mathcal{X}$ be the support of X, $V_0 = \mathrm{E}[\partial_\theta^2 \tau(X, \theta_0)]/2$, $\Delta_0 = \mathrm{E}[\partial_\theta \tau(X, \theta_0)\partial_\theta^\top \tau(X, \theta_0)]$, and

$\Sigma_0 = V_0^{-1}\triangle_0 V_0^{-1}$, with

$$\tau(x,\theta) = \mathrm{P}(T > t > (A \vee a), (C+A) \wedge (c+a) > t, Z^\top(1,\theta) > z^\top(1,\theta)) \\ +\mathrm{P}(t > T > (A \vee a), (C+A) \wedge (c+a) > T, z^\top(1,\theta) > Z^\top(1,\theta)),$$

for all $(x,\theta) \in \mathcal{X} \times \Theta$. We assume the following conditions.

A4. $\mathrm{E}\|\partial_\theta \tau(X,\theta_0)\| < \infty$.

A5. $\|\partial_\theta^2 \tau(x,\theta) - \partial_\theta^2 \tau(x,\theta_0)\| \le M\|\theta - \theta_0\|$ for some positive constant M independent of $(x,\theta) \in \mathcal{X} \times \mathcal{N}_{\theta_0}$.

A6. $\sup_{\{(x,\theta)\in\mathcal{X}\times\mathcal{N}_{\theta_0}\}} \|\partial_x^2 \tau(x,\theta)\| < \infty$ and $\sum_{i_1,i_2} \mathrm{E}\big[|(\partial_\theta^2 \tau(X,\theta_0))_{i_1,i_2}|\big] < \infty$.

A7. V_0 is positive-definite.

Following the proofs in Han (1987) and Sherman (1993), we establish the consistency and asymptotic normality of $\hat\theta$ as follows:

Theorem 2. *Under model* (1.1) *and assumptions* A1-A7, $\hat\theta \xrightarrow{p} \theta_0$ *and* $\sqrt{n}(\hat\theta - \theta_0) \xrightarrow{d} N(0,\Sigma_0)$.

Using a consistent estimator of Σ_0, the asymptotic normality of $\hat\theta$ can be applied to develop the related inference procedures. Instead of directly estimating Σ_0 using a smoothing estimation technique (cf., Sherman (1993)), a weighted bootstrap approximation of the sampling distribution of $\hat\theta$ is preferred in practical implementations.

Let $D_n = \{(Y_i,\delta_i,A_i,Z_i)\}_{i=1}^n$ be the collected left-truncated and right-censored survival data. Independent of D_n, the random quantities $\xi_1,\ldots,\xi_n$ are independently generated from a common population with $\mathrm{P}(\xi = 0) < 1$. A weighted bootstrap analogue of $\mathrm{C}_n(\beta)$ is given by

$$\mathrm{C}_n^\omega(\beta) = \sum_{i=1}^n\sum_{j=1}^n w_i w_j \mathrm{I}(Y_i > Y_j > (A_i \vee A_j), \delta_j = 1, \beta^\top Z_i > \beta^\top Z_j), \quad (3.5)$$

where $w_i = \xi_i / \sum_{j=1}^n \xi_j$, for $i = 1,\ldots,n$, and the counterpart, say $\hat\beta^\omega$, of $\hat\beta$ is defined as a maximizer of $\mathrm{C}_n^\omega(\beta)$. In the following theorem, we establish the asymptotic equivalence of $\rho(\hat\theta^\omega - \hat\theta)$ and $(\hat\theta - \theta_0)$, where $\rho = \mathrm{E}[\xi]/\sqrt{\mathrm{var}(\xi)}$ is a scale-factor modification for the variability in the weights.

Theorem 3. *Under model* (1.1) *and assumptions* A1-A7,

$$\sup_{u\in\mathrm{R}} |\mathrm{P}(\sqrt{n}\rho(\hat\theta^\omega - \hat\theta) \le u|D_n) - \mathrm{P}(\sqrt{n}(\hat\theta - \theta_0) \le u)| \xrightarrow{p} 0.$$

Proof. See the Appendix.

Let $\sigma^{\omega}(\hat{\theta}_k)$ and $q^{\omega}_{\varsigma}(\hat{\theta}_k)$ be the standard deviation and 100ςth quantile ($0 < \varsigma < 1$) of $\rho(\hat{\theta}^{\omega} - \hat{\theta})$, $k = 1, \ldots, (p-1)$, respectively, and z_{ς} be the 100ςth quantile of the standard normal distribution. It follows from Theorems 2–3 that approximate $100(1-\alpha)\%$ quantile-type and normal-type bootstrap confidence intervals of θ_{0k}, for $0 < \alpha < 1$, can be constructed by

$$(\hat{\theta}_k - q^{\omega}_{1-\alpha/2}(\hat{\theta}_k), \hat{\theta}_k - q^{\omega}_{\alpha/2}(\hat{\theta}_k)) \text{ and } (\hat{\theta}_k - z_{1-\alpha/2}\sigma^{\omega}(\hat{\theta}_k), \hat{\theta}_k + z_{1-\alpha/2}\sigma^{\omega}(\hat{\theta}_k)), \tag{3.6}$$

respectively. According to our empirical results, the quantile-type bootstrap interval estimator outperforms the normal-type interval estimator, in general, in terms of the length and coverage probability.

4. Simulations

In this section, we describe the simulation experiments that we used to investigate the finite-sample performance of the proposed estimator and its competitors. To ensure numerical stability, the simulation results are based on 1,000 replications with sample sizes (n) of 200 and 400. The bootstrap inferences are drawn from 500 bootstrap samples with $\xi_1, \ldots, \xi_n \overset{i.i.d.}{\sim} Gamma(4, 2)$. Three hazard models with $Z^* = (Z^*_1, Z^*_2, Z^*_3)^{\top}$ and the same index coefficients $\beta_0 = (1, 1, 1)^{\top}$ are studied under different setups of left-truncation and right-censoring. Moreover, conditional on $Z^* = z$, the residual censoring time $C = r_0(U(0, z_2) + 0.1)$ is generated independently, where r_0 is specified to produce censoring rates ($c.r.$) of 20% and 40%.

Example 1. A mixture of discrete and continuous covariates are specified, with $Z^*_1 \sim N(0, 1)$, $Z^*_2 \sim U(0, 1)$, and $Z^*_3 \sim U(\{1, 2, \ldots, 10\})$. We designed the following proportional hazard regression model to generate T^*:

$$\text{M1. } \lambda(t, z) = \lambda_0(t)\exp(\beta_0^{\top} z) \text{ with } \lambda_0(t) = 4t.$$

Conditional on $Z^* = z$, $A^* = r_1(U(0, 0.5) + |z_1|\mathrm{I}(|z_1| < 0.5))$ is generated independently with the proportions of untruncated data ($p.u.$), that is $\mathrm{P}(A^* > T^*)$, being 0.1, 0.5, and 0.9 for different values of r_1.

Under the setup of covariate-dependent truncation time, we compared the proposed maximum partial-rank correlation estimator $\hat{\beta}$ with the maximum partial likelihood estimator $\tilde{\beta}$ of Wang, Brookmeyer and Jewell (1993) for the proportional hazard regression model and with the estimator $\bar{\beta}$ of Huang and Qin (2013) for the additive hazard regression model. Table 1 displays the means

Table 1. The means (standard deviations) of 1,000 estimates under model M1 for sample sizes (n) of 200 and 400, proportions of untruncated data ($p.u.$) of 0.1, 0.5, and 0.9, and censoring rates ($c.r.$) of 20% and 40%.

		$n = 200$						$n = 400$					
$p.u.$		0.1		0.5		0.9		0.1		0.5		0.9	
$c.r.$		θ_{01}	θ_{02}	θ_{01}	θ_{02}	θ_{01}	θ_{02}	θ_{01}	θ_{02}	θ_{01}	θ_{02}	θ_{01}	θ_{02}
20%	$\hat{\theta}$	1.02	1.00	1.03	1.01	1.03	0.98	1.01	1.00	1.00	1.00	1.00	1.00
		(0.144)	(0.141)	(0.125)	(0.122)	(0.143)	(0.144)	(0.086)	(0.080)	(0.091)	(0.085)	(0.105)	(0.102)
	$\tilde{\theta}$	1.00	1.01	1.01	1.01	1.01	1.01	1.00	1.01	1.00	1.00	1.00	1.01
		(0.087)	(0.087)	(0.083)	(0.083)	(0.091)	(0.090)	(0.054)	(0.055)	(0.058)	(0.055)	(0.063)	(0.064)
	$\bar{\theta}$	1.09	1.08	1.10	1.11	1.14	1.13	1.06	1.06	1.09	1.09	1.13	1.14
		(0.186)	(0.187)	(0.177)	(0.181)	(0.164)	(0.161)	(0.117)	(0.115)	(0.123)	(0.124)	(0.121)	(0.117)
40%	$\hat{\theta}$	1.02	1.00	1.04	1.01	1.03	1.01	1.02	1.00	1.00	1.01	1.00	1.01
		(0.150)	(0.151)	(0.136)	(0.137)	(0.143)	(0.149)	(0.098)	(0.093)	(0.096)	(0.091)	(0.107)	(0.106)
	$\tilde{\theta}$	1.00	1.00	1.02	1.01	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.01
		(0.100)	(0.101)	(0.091)	(0.090)	(0.092)	(0.091)	(0.065)	(0.065)	(0.060)	(0.061)	(0.067)	(0.070)
	$\bar{\theta}$	1.08	1.09	1.12	1.11	1.14	1.15	1.07	1.06	1.10	1.10	1.13	1.14
		(0.257)	(0.268)	(0.208)	(0.214)	(0.185)	(0.194)	(0.152)	(0.147)	(0.136)	(0.137)	(0.128)	(0.127)

Table 2. The means (standard deviations) of 1,000 estimates under model M2 for sample sizes (n) of 200 and 400 and censoring rates ($c.r.$) of 20% and 40%.

	$n = 200$				$n = 400$			
$c.r.$	20%		40%		20%		40%	
	θ_{01}	θ_{02}	θ_{01}	θ_{02}	θ_{01}	θ_{02}	θ_{01}	θ_{02}
$\hat{\theta}$	1.01	1.00	1.02	1.03	1.00	1.00	1.01	1.01
	(0.130)	(0.111)	(0.191)	(0.212)	(0.091)	(0.093)	(0.102)	(0.096)
$\tilde{\theta}$	0.78	0.81	0.83	0.84	0.75	0.74	0.74	0.73
	(0.240)	(0.261)	(0.282)	(0.251)	(0.170)	(0.174)	(0.273)	(0.268)
$\bar{\theta}$	1.01	1.00	1.02	1.02	1.00	1.00	1.00	1.01
	(0.182)	(0.175)	(0.195)	(0.204)	(0.121)	(0.117)	(0.130)	(0.136)

and standard deviations of 1,000 estimates for different combinations of sample sizes, censoring rates, and proportions of untruncated data. As a result of a model misspecification, $\bar{\beta}$ has a relatively large bias and standard deviation in the model formulation of M1. In addition, the bias magnitudes of both $\hat{\beta}$ and $\tilde{\beta}$ are generally small. However, the standard deviation of $\tilde{\beta}$ is slightly smaller than that of $\hat{\beta}$. In addition, the variations of both $\hat{\beta}$ and $\tilde{\beta}$ decrease as n increases, $c.r.$ decreases, and $p.u.$ falls around 0.5. As shown in the next two examples, the maximum partial likelihood estimator has very poor performance under a model misspecification. To simplify the presentation, a weighted bootstrap estimator of the standard deviation and a weighted bootstrap confidence interval of β_0 are assessed in the setting with $p.u. = 0.5$. In Table 4, the averages of 1,000 bootstrap standard errors and 95% quantile-type bootstrap confidence intervals

Table 3. The means (standard deviations) of 1,000 estimates under model M3 for sample sizes (n) of 200 and 400 and censoring rates ($c.r.$) of 20% and 40%.

	$n=200$				$n=400$			
$c.r.$	20%		40%		20%		40%	
	θ_{01}	θ_{02}	θ_{01}	θ_{02}	θ_{01}	θ_{02}	θ_{01}	θ_{02}
$\hat{\theta}$	1.00	1.01	1.02	1.05	1.00	1.00	1.02	1.01
	(0.148)	(0.146)	(0.345)	(0.340)	(0.099)	(0.094)	(0.157)	(0.154)
$\tilde{\theta}$	0.22	0.08	0.21	0.07	0.18	0.06	0.17	0.16
	(0.282)	(0.116)	(0.400)	(0.166)	(0.251)	(0.105)	(0.267)	(0.101)
$\bar{\theta}$	0.61	0.22	0.38	0.10	0.41	0.06	0.26	0.02
	(0.508)	(0.426)	(0.530)	(0.426)	(0.820)	(0.761)	(0.462)	(0.382)

are found to tend toward the standard deviations and 95% quantile intervals of 1,000 estimates as n increases or $c.r.$ decreases. The empirical coverage probabilities of β_0, shown in Table 4, are slightly higher than the nominal level of 0.95 for $(n, c.r.) = (200, 40\%)$, and stay close to this nominal level for the remaining cases.

Example 2. In this simulation scenario, a random vector $(Z_{01}^*, 1/Z_{02}^*, 1/Z_{03}^*)^\top$ is specified to follow a multivariate normal distribution with mean zero, standard deviation one, and pairwise correlation 0.5. Furthermore, the joint distribution of Z^* was designed to be the same as that of $(10Z_{01}^*, 10Z_{02}^*, 10Z_{03}^*)^\top$ on $\{Z_{01}^* + Z_{02}^* + Z_{03}^* > 0\}$. Moreover, T^* is generated from the following additive hazard regression model:

$$\text{M2. } \lambda(t, z) = \lambda_0(t) + \beta_0^\top z \text{ with } \lambda_0(t) = 1.$$

For the truncation time, conditional on $Z^* = z$, $A^* = 0.4(U(0,1) + |z_1|\mathrm{I}(|z_1| < 0.5))$ is set with $p.u. = 0.9$.

Compared with $\hat{\beta}$ and $\bar{\beta}$, $\tilde{\beta}$ has a substantially bias and standard deviation in Table 2. Furthermore, the biases of $\hat{\beta}$ and $\bar{\beta}$ are comparable. Even in the case of the conditional estimating equation approach used for the additive hazard regression model, the standard deviation of $\hat{\beta}$ is, surprisingly, found to be smaller than that of $\bar{\beta}$. Once again, the bootstrap standard error and confidence interval slightly overestimate the asymptotic standard deviation and the quantile interval, respectively, but their accuracies improve significantly as n increases or $c.r.$ decreases. Moreover, the constructed weighted bootstrap confidence intervals have fairly accurate coverage probabilities.

Example 3. With the triplets $(Z_{01}^*, Z_{02}^*, Z_{03}^*)^\top$ in Example 2, the joint distribution of Z^* is specified to be the same as that of $(Z_{01}^*, Z_{02}^*, Z_{03}^*)^\top$ on $\{Z_{01}^* + Z_{02}^* +$

Table 4. The standard deviations (*s.d.*), bootstrap standard errors (*b.s.e*), 95% quantile interval (*q.i.*), quantile-type bootstrap confidence intervals (*q.b.c.i.*), and empirical coverage probabilities (*c.p.*) of 1,000 estimates.

	c.r.		$n = 200$					$n = 400$				
			s.d.	*b.s.e.*	*q.i.*	*q.b.c.i.*	*c.p.*	*s.d.*	*q.i.*	*q.b.c.i.*	*b.s.e.*	*c.p.*
M1	20%	$\hat{\theta}_{01}$	0.125	0.162	(0.802, 1.307)	(0.776, 1.347)	0.956	0.091	0.093	(0.849, 1.212)	(0.845, 1.216)	0.950
		$\hat{\theta}_{02}$	0.122	0.149	(0.791, 1.303)	(0.761, 1.343)	0.956	0.085	0.090	(0.843, 1.197)	(0.840, 1.217)	0.955
	40%	$\hat{\theta}_{01}$	0.136	0.203	(0.822, 1.383)	(0.769, 1.426)	0.961	0.096	0.098	(0.846, 1.248)	(0.843, 1.250)	0.952
		$\hat{\theta}_{02}$	0.137	0.216	(0.785, 1.303)	(0.739, 1.403)	0.960	0.091	0.094	(0.827, 1.210)	(0.818, 1.221)	0.953
M2	20%	$\hat{\theta}_{01}$	0.130	0.153	(0.816, 1.367)	(0.792, 1.387)	0.954	0.091	0.094	(0.849, 1.205)	(0.843, 1.209)	0.951
		θ_{02}	0.111	0.141	(0.799, 1.253)	(0.761, 1.303)	0.953	0.093	0.094	(0.883, 1.257)	(0.880, 1.260)	0.952
	40%	$\hat{\theta}_{01}$	0.191	0.261	(0.702, 1.486)	(0.669, 1.503)	0.951	0.102	0.106	(0.835, 1.248)	(0.850, 1.258)	0.952
		$\hat{\theta}_{02}$	0.212	0.267	(0.639, 1.453)	(0.605, 1.430)	0.951	0.096	0.101	(0.807, 1.210)	(0.803, 1.231)	0.952
M3	20%	$\hat{\theta}_{01}$	0.148	0.201	(0.696, 1.277)	(0.622, 1.287)	0.958	0.099	0.101	(0.823, 1.225)	(0.820, 1.243)	0.953
		θ_{02}	0.146	0.198	(0.709, 1.273)	(0.631, 1.283)	0.958	0.094	0.102	(0.863, 1.277)	(0.860, 1.280)	0.951
	40%	$\hat{\theta}_{01}$	0.345	0.407	(0.372, 1.733)	(0.309, 1.766)	0.962	0.157	0.159	(0.746, 1.348)	(0.740, 1.378)	0.952
		$\hat{\theta}_{02}$	0.340	0.414	(0.345, 1.693)	(0.289, 1.763)	0.961	0.154	0.157	(0.710, 1.310)	(0.701, 1.331)	0.952

$Z^*_{03} > 0\}$. The hazard regression model of T^* on $Z^* = z$ is further designed to be

$$\text{M3. } \lambda(t, z) = \frac{\beta_0^\top z}{(\beta_0^\top z + t)}.$$

Conditional on $Z^* = z$, a covariate-dependent truncation time $A^* = U(0, 10) + |z_3|\mathrm{I}(|z_3| < 10)$ is set, with $p.u. = 0.9$. Note that the above model is neither the proportional hazard regression model nor the additive hazard regression model.

Our simulation results show that the invalid partial likelihood and conditional estimating equation approaches lead to serious biases and unacceptable variations in $\tilde{\beta}$ and $\bar{\beta}$. In contrast, the means of the 1,000 maximum partial-rank correlation estimates are very close to β_0. The standard deviation of $\hat{\beta}$ decreases as n increases and *c.r.* decreases. The weighted bootstrap standard error and confidence interval perform similarly to those in Example 1.

5. An Analysis of the HRS Data

Here, we apply our partial-rank correlation estimation to the RAND version N of the US HRS data, which are available at the website: `http://hrsonline.isr.umich.edu`. A sample of individuals, born between 1931 and 1934, was recruited by a cross-sectional sampling scheme between 1992 and 1993 and followed until 2012. After excluding those with missing covariates of interest, a total of 4,323 white nonHispanic men and 4,724 white nonHispanic women were interviewed. For each individual, the birth date, gender, self-reported body mass index (*BMI*), level of education, and smoking status were investigated in the data analysis. In the first interview, smoking status was defined as "never smoked" (*nsmok*), "stopped smoking" (*ssmok*), and "currently smoking" (*csmok*), and educational attaintment was classified as "less-than-high-school or general educational development" (*ledu*), "high school graduate and some college" (*medu*), and "college graduate and above" (*hedu*). The vital status and last observed date for the individuals were determined by the National Death Index (NDI) and an exit interview. Because some individuals died before recruitment and were lost to follow-up during the study period, their survival times were subject to left-truncation and right-censoring.

Let Z_1^* and Z_2^* be the dummy variables, with *nsmok* as the reference category and the value one assigned to *csmok* and *ssmok*. For level of education, *hedu* is treated as the reference category, and the value one is assigned to *ledu* and *medu* in the dummy variables Z_3^* and Z_4^*, respectively. Because being overweight or underweight, evaluated in terms of *BMI*, might decrease life expectancy, the designed variables $Z_5^* = BMI_a$ and $Z_6^* = BMI_a^2$ are used in the model fitting, where $BMI_a = \log(BMI/\overline{BMI})$, with $\overline{BMI}$ being the sample mean. In this data analysis, the gender (*gender*) of each person is further considered as a stratification variable. Based on the left-truncated and right-censored survival data, our research aims to identify the effects of these attributes on the death time of males and females using a more general hazard regression model (1.1). Using the partial-rank correlation estimation, Table 5 shows that the estimated effects of smoking status, level of education, and *BMI* on the hazard function of transition to death are very similar for men and women. The mortality risks of a current smoker and a stopped smoker are significantly higher than those of a stopped smoker and a nonsmoker, respectively. Compared with those with a higher level of education, people with a low level of education have a significantly higher mortality risk, whereas a median-education is not significantly different in terms of life expectancy. Furthermore, a higher or lower *BMI* tends to increase the hazard rate of transition to death. In addition, the mortality risk of an overweight individual is, in general, higher than that of an underweight person.

When the truncation time is covariate-independent, that is, $f_{A^*}(a|z) = f_{A^*}(a)$, where $f_{A^*}(a|z)$ and $f_{A^*}(a)$ are the conditional density function of A^* given $Z^* = z$ and the marginal density function of A^*, respectively, Chen and Chiang (2018) developed another approach to estimate β_0 based on a prevalent cohort sample without survival times. The proposed estimator, say $\breve{\beta} = (1, \breve{\theta}^\top)^\top$, is defined as the maximizer of the following sample analogue of $\mathrm{C}_A(\beta) = P(A_1 >$

Table 5. The estimates (standard errors) of index coefficients for HRS data.

gender		variables					
		csmok	*ssmok*	*ledu*	*medu*	BMI_a	BMI_a^2
male	$\hat{\beta}$	**1.00**	**0.39**	**0.35**	0.09	**0.57**	**0.66**
		(0.000)	(0.059)	(0.062)	(0.083)	(0.163)	(0.147)
	$\tilde{\beta}$	**1.00**	**0.39**	**0.29**	0.06	**0.31**	**3.42**
		(0.000)	(0.050)	(0.064)	(0.051)	(0.123)	(0.595)
	$\bar{\beta}$	**1.00**	**0.27**	**0.29**	0.02	0.15	**6.11**
		(0.000)	(0.068)	(0.116)	(0.069)	(0.297)	(1.550)
female	$\hat{\beta}$	**1.00**	**0.25**	**0.62**	0.07	**0.62**	**1.01**
		(0.000)	(0.080)	(0.107)	(0.132)	(0.153)	(0.281)
	$\tilde{\beta}$	**1.00**	**0.29**	**0.58**	**0.22**	**0.52**	**2.73**
		(0.000)	(0.073)	(0.110)	(0.096)	(0.160)	(0.474)
	$\bar{\beta}$	**1.00**	**0.18**	**0.71**	0.14	**0.88**	**4.82**
		(0.000)	(0.083)	(0.191)	(0.099)	(0.331)	(1.335)

$A_2, \beta^\top Z_1 > \beta^\top Z_2$):

$$\mathrm{C}_{nA}(\beta) = \frac{1}{n(n-1)} \sum_{i \neq j} I(A_i > A_j, \beta^\top Z_i > \beta^\top Z_j). \tag{5.1}$$

The maximizer β_A of $\mathrm{C}_A(\beta)$ is further shown to be β_0 whenever model (1.1) is correct. For the hypotheses

$$\begin{cases} H_{00} : \{f_{A^*}(a|z) = f_{A^*}(a)\} \text{ or } \{f_{A^*}(a|z) \neq f_{A^*}(a), \beta_A = \beta_0\}, \\ H_{0A} : f_{A^*}(a|z) \neq f_{A^*}(a), \end{cases} \tag{5.2}$$

a Hausman-type statistic $\mathcal{T}_0 = (\hat{\theta} - \breve{\theta})^\top (\rho^2 Var(\hat{\theta}^\omega - \breve{\theta}^\omega | D_n))^{-1} (\hat{\theta} - \breve{\theta})$ is introduced to test whether the truncation time is covariate-dependent. Note that β_A can be β_0, even if the truncation time is covariate-dependent. Based on data of the form $\{(A_i, Z_i)\}_{i=1}^n$, $\breve{\beta}$ is computed to be $(1.00, -0.46, -1.25, -0.50, 1.90, 0.26)$ for men and $(1.00, -0.24, -2.21, -1.42, -0.07, 1.98)$ for women, with corresponding bootstrap standard errors of $(0.000, 0.223, 0.217, 0.119, 0.503, 0.320)$ and $(0.000, 0.310, 0.773, 0.389, 0.648, 0.760)$. Both $\hat{\beta}$ in Table 5 and $\breve{\beta}$ have different explanations for life expectancy. In addition, the values of $\mathcal{T}_0$ 4.45 for men and 3.46 for women. From their bootstrap p-values, 0.000 and 0.000, we conclude that the truncation time should be covariate-dependent.

Although the explanations of $\tilde{\beta}$ and $\bar{\beta}$ in Table 5 are similar to that of $\hat{\beta}$, the magnitudes of their coefficient estimates of BMI_a^2 are very different. The appropriateness of multiplicative and additive hazard regression models is investigated further. Under model (1.1), let $\lambda_0^*(t) \exp(\beta_0^{*\top} z)$ and $\lambda_0^{**}(t) + \beta_0^{**\top} z$ be the corresponding maximizer and solution of the asymptotic equivalent functions of the partial likelihood function and the conditional estimating equation, respectively. It follows that $(\lambda_0^*(t), \beta_0^*) = (\lambda_0(t), \beta_0)$ when $\lambda(t, \beta_0^\top z) = \lambda_0(t)\exp(\beta_0^\top z)$, and $(\lambda_0^{**}(t), \beta_0^{**}) = (\lambda_0(t), \beta_0)$ when $\lambda(t, \beta_0^\top z) = \lambda_0(t) + \beta_0^\top z$. We consider the hypotheses

$$\begin{cases} H_{10} : \{\lambda(t, \beta_0^\top z) = \lambda_0(t)\exp(\beta_0^\top z)\} \text{ or } \{\lambda(t, \beta_0^\top z) \neq \lambda_0^*(t)\exp(\beta_0^{*\top} z), \beta_0^* = \beta_0\}, \\ H_{1A} : \{\lambda(t, \beta_0^\top z) \neq \lambda_0^*(t)\exp(\beta_0^{*\top} z), \beta_0^* \neq \beta_0\}, \end{cases} \tag{5.3}$$

and

$$\begin{cases} H_{20} : \{\lambda(t, \beta_0^\top z) = \lambda_0(t) + \beta_0^\top z\} \text{ or } \{\lambda(t, \beta_0^\top z) \neq \lambda_0^{**}(t) + \beta_0^{**\top} z, \beta_0^{**} = \beta_0\}, \\ H_{2A} : \{\lambda(t, \beta_0^\top z) \neq \lambda_0^{**}(t) + \beta_0^{**\top} z, \beta_0^{**} \neq \beta_0\}. \end{cases} \tag{5.4}$$

It follows that $\hat{\beta}$ is a consistent estimator of β_0 under the hypotheses in (5.3)–(5.4), whereas $\tilde{\beta}$ and $\bar{\beta}$ are not consistent estimators of β_0 under H_{1A} and H_{2A}, respectively. Hausman-type test statistics $\mathcal{T}_1 = (\hat{\theta}-\tilde{\theta})^\top(\rho^2 Var(\hat{\theta}^\omega - \tilde{\theta}^\omega | D_n))^{-1}(\hat{\theta}-\tilde{\theta})$ and $\mathcal{T}_2 = (\hat{\theta}-\bar{\theta})^\top(\rho^2 Var(\hat{\theta}^\omega - \bar{\theta}^\omega | D_n))^{-1}(\hat{\theta}-\bar{\theta})$ are naturally proposed, and the hypotheses H_{01} and H_{02} are rejected if the corresponding values of $\mathcal{T}_1$ and $\mathcal{T}_2$ are greater than the critical values at the specified significance levels. The values of $(\mathcal{T}_1, \mathcal{T}_2)$ and their bootstrap p-values are computed to be (6.76, 6.05) and (0.000, 0.005), respectively, for men and (6.50, 2.38) and (0.000, 0.003), respectively, for women. These numerical results show that the proportional and additive hazard regression models are not appropriate for characterizing the effects of covariates on the hazard rates of transition to death. Based on this conclusion, a challenging task remains of examining the correctness of model (1.1) or exploring a potential formulation of $\lambda(t, u)$ in model (1.1).

6. Conclusion and Discussion

A partial-rank correlation estimator is proposed to estimate the index coefficients of a general single-index hazard regression model with left-truncated and right-censored survival data. An efficient computational algorithm is employed to perform this constraint nondifferentiable optimization problem. The developed approach can be extended to a reversed-time hazard regression model (1.2) with right-truncation and left-censoring. Moreover, we establish the consistency and asymptotic normality of the proposed maximum partial-rank correlation estimator and introduce weighted bootstrap approximations of the sampling quantities of interest related to the proposed estimator. The numerical studies also show that our estimator performs satisfactorily.

In terms of the constructed partial-rank correlation function, the single-index $\beta_0^\top Z^*$ is shown to exhibit existence, optimality, and uniqueness up to a scale and location. Unfortunately, the proposed estimation criterion cannot be directly applied to a more general single-index survival model of the form:

$$S(t|z) = S(t, \beta_0^\top z), \tag{6.1}$$

where $S(t,u)$ is an unknown nonnegative bivariate function, strictly increasing in u for each t. This is because the monotonicity of $S(t,u)$ in u for each t does not imply the monotonicity of $\lambda(t,u)$ in u for each t. Currently, there are no estimation and inference procedures for such a model formulation with left-truncated and right-censored survival data. The methodological challenge of estimating the index coefficients remains for future research. In our data analysis, testing procedures based on Hausman-type test statistics are built to examine the distribution of the truncation time and the related model structures. When there is no strong evidence to reject the null hypotheses (H_{00}, H_{10}, and H_{20}) in (5.2)–(5.4), we cannot conclude the adequacy of the covariate-independent truncation, proportional hazard model, and additive hazard model. A thorough study would be worthwhile for the null hypotheses $H_{00}^* : f_{A^*}(a|z) = f_{A^*}(a)$, $H_{10}^* : \lambda(t, \beta_0^\top z) = \lambda_0(t)\exp(\beta_0^\top z)$, and $H_{20}^* : \lambda(t, \beta_0^\top z) = \lambda_0(t) + \beta_0^\top z$.

Acknowledgment

The corresponding author's research was partially supported by the Ministry of Science and Technology grant 106-2118-M-002-007 (Taiwan). We would like to thank the associate editor and two reviewers for their constructive comments.

Appendix

Proof of Theorem 1. Let $f_C(c|z)$ denote the density of C, given $Z = z$, and $\Gamma(z_{01}, z_{02}) = \mathrm{E}[Q(Z_1, Z_2; A_1 \vee A_2, (C_1 + A_1) \wedge (C_2 + A_2))/(S(A_1|Z_1)S(A_2|Z_2))|Z_1 = z_{01}, Z_2 = z_{02}]$. By specifying $(z_{01}, z_{02}) = (z_1, z_2)$ and $(z_{01}, z_{02}) = (z_2, z_1)$, assumption A2 ensures that

$$\Gamma(z_1, z_2) = \int \cdots \int Q(z_1, z_2; a_1 \vee a_2, (c_1 + a_1) \wedge (c_2 + a_2)) \prod_{\ell=1}^{2} \frac{f_{A^*}(a_\ell|z_\ell) f_C(c_\ell|z_\ell)}{P(T^* > a_\ell|z_\ell)} da_\ell dc_\ell \text{ and}$$

$$\Gamma(z_2, z_1) = \int \cdots \int Q(z_2, z_1; a_1 \vee a_2, (c_1 + a_1) \wedge (c_2 + a_2)) \prod_{\ell=1}^{2} \frac{f_{A^*}(a_\ell|z_\ell) f_C(c_\ell|z_\ell)}{P(T^* > a_\ell|z_\ell)} da_\ell dc_\ell. \tag{A.1}$$

An application of Lemma 1 further leads to

$$\Gamma(z_1, z_2) > \Gamma(z_2, z_1) \text{ whenever } \beta_0^\top z_1 > \beta_0^\top z_2. \tag{A.2}$$

Moreover, the following property is an implication of assumption A3:

$$\mathrm{E}[\mathrm{I}(\beta_0^\top Z_1 > \beta_0^\top Z_2, \beta^\top Z_1 < \beta^\top Z_2)] > 0 \ \forall \beta \neq \beta_0. \tag{A.3}$$

By the law of iterated expectation, we have $\mathrm{C}(\beta) = \mathrm{E}[\Gamma(Z_1, Z_2)\ \mathrm{I}(\beta^\top Z_1 > \beta^\top Z_2)]$. Coupled with (A.2)–(A.3) and the equality

$$\mathrm{C}(\beta_0) - \mathrm{C}(\beta) = \mathrm{E}[(\Gamma(Z_1, Z_2) - \Gamma(Z_2, Z_1))\mathrm{I}(\beta_0^\top Z_1 > \beta_0^\top Z_2, \beta^\top Z_1 < \beta^\top Z_2)], \tag{A.4}$$

β_0 can be shown to be the unique maximizer of $\mathrm{C}(\beta)$.

Proof of Theorem 3. From equation (7) in Sherman (1993) and assumptions A1–A7, it follows that

$$\mathrm{C}_n(\beta) - \mathrm{C}_n(\beta_0) = (\theta - \theta_0)^\top \left(\Psi_n - \frac{V_0}{2}(\theta - \theta_0) \right)(1 + o_p(1)) + o_p\left(\frac{1}{n}\right) \tag{A.5}$$

uniformly over $o_p(1)$ neighborhoods of θ_0, where $\Psi_n = \sum_{j=1}^n u_j/n$ with $u_1, \ldots, u_n$ being independent and identically distributed random variables from a population with mean zero and variance-covariance matrix Δ_0. An application of Theorem 2 in Sherman (1993) further leads to

$$\sqrt{n}(\hat{\theta} - \theta_0) = \frac{1}{\sqrt{n}} \sum_{j=1}^n V_0^{-1} u_j + o_p(1). \tag{A.6}$$

For the weighted bootstrap analogue $\mathrm{C}_n^\omega(\beta)$ of $\mathrm{C}_n(\beta)$, the argument of Sherman (1993) enables us to derive that

$$\mathrm{C}_n^\omega(\beta) - \mathrm{C}_n^\omega(\beta_0) = (\theta - \theta_0)^\top \left(\frac{\mathrm{E}[\xi]}{n} \sum_{j=1}^n \xi_j u_j - \frac{(\mathrm{E}[\xi])^2 V_0}{2}(\theta - \theta_0) \right)(1 + o_{\tilde{p}}(1)) + o_{\tilde{p}}\left(\frac{1}{n}\right) \tag{A.7}$$

uniformly over $o_{\tilde{p}}(1)$ neighborhoods of θ_0, where $\tilde{P}$ is the probability measure generated by $D_n \times \{\xi_1, \ldots, \xi_n\}$. Let $W_n = \sum_{j=1}^n V_0^{-1} \xi_j u_j / (\sqrt{n}\mathrm{E}[\xi])$. Because

$\hat{\beta}^{\omega} = (1, \hat{\theta}^{\omega})^{\top}$ is a maximizer of $\mathrm{C}_n^{\omega}(\beta)$, we have

$$\mathrm{C}_n^{\omega}(\hat{\beta}^{\omega}) - \mathrm{C}_n^{\omega}\left(\left(1, \theta_0 + \frac{W_n}{\sqrt{n}}\right)^{\top}\right) \geq 0. \tag{A.8}$$

Coupled with (A.7), we further have

$$-\big(\sqrt{n}(\hat{\theta}^{\omega} - \theta_0) - W_n\big)^{\top} V_0 \big(\sqrt{n}(\hat{\theta}^{\omega} - \theta_0) - W_n\big)(1 + o_{\tilde{p}}(1)) \geq 0, \tag{A.9}$$

which implies that

$$\sqrt{n}(\hat{\theta}^{\omega} - \theta_0) = W_n + o_{\tilde{p}}(1). \tag{A.10}$$

From (A.6) and (A.10), the following property can be obtained:

$$\sqrt{n}(\hat{\theta}^{\omega} - \hat{\theta}) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} V_0^{-1}\left(1 - \frac{\xi_j}{\mathrm{E}[\xi]}\right) u_j + o_{\tilde{p}}(1). \tag{A.11}$$

In the spirit of the proof in Janssen (1994), the Lindeberg-Feller central limit theorem can be applied to show that

$$\sup_{u \in \mathrm{R}} |\mathrm{P}(\sqrt{n}\rho(\hat{\theta}^{\omega} - \hat{\theta}) \leq u | D_n) - \Phi_{\Sigma_0}(u)| \xrightarrow{p} 0, \tag{A.12}$$

where $\Phi_{\Sigma_0}(u)$ is a multivariate normal distribution with mean vector zero and variance-covariance matrix Σ_0. By Theorem 2, (A.12), and the probability inequality, Theorem 3 is established.

References

Aalen, O. (1980). *A Model for Nonparametric Regression Analysis of Counting Processes.* Lecture Notes in Statistics 2, 1-25, Springer, New York.

Chen, S. W. and Chiang, C. T. (2018). General single-index survival regression models for incident-prevalent covariate data and prevalent data without follow-up. *Biometrics* **74**, 881–890.

Cheng, S. C., Wei, L. J. and Ying, Z. (1995). Analysis of transformation models with censored data. *Biometrika* **82**, 835–845.

Cox, D. R. (1972). Regression models and life-tables. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* B **34**, 187–220.

Cui, J. (1999). Nonparametric estimation of a delay distribution based on left-censored and right-truncated data. *Biometrics* **55**, 345–349.

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

Huang, C. Y. and Qin, J. (2013). Semiparametric estimation for the additive hazards model with left-truncated and right-censored data. *Biometrika* **100**, 1–12.

Janssen, P. (1994). Weighted bootstrapping of U-statistics. *J. Statist. Plann. Inference* **38**, 31–42.

Khan, S. and Tamer, E. (2007). Partial rank estimation of duration models with general forms of censoring. *J. Econometrics* **1**, 251–280.

Kalbfleisch, J. D. and Lawless, J. F. (1991). Regression models for right truncated data with applications to AIDS incubation times and reporting lags. *Statist. Sinica* **1**, 19–32.

Kaminsky, K. (1987). Prediction of IBNR claim counts by modeling the distribution of report lags. *Insurance Math. Econom.* **6**, 151–159.

Lagakos, S. W., Barraj, L. M. and Gruttola, V. DE (1998). Nonparametric analysis of truncated survival data, with application to AIDS. *Biometrika* **75**, 515–523.

Lin, D. Y. and Ying, Z. (1994). Semiparametric analysis of the additive risk model. *Biometrika* **81**, 61–71.

Ma and Huang, J. (2005). Regularized ROC estimation: with applications to disease classication using microarray data. Technical Report.

Sherman, R. P. (1993). The limiting distribution of the maximum rank correlation estimator. *Econometrica* **61**, 123–138.

Wang, S. H. and Chiang, C. T. (2017). Concordance-gradient-based estimation for the optimal sufficient dimension reduction score. *Technical Report*.

Wang, M. C., Brookmeyer, R. and Jewell, N. P. (1993). Statistical models for prevalent cohort data. *Biometrics* **49**, 1–11.

Institute of Applied Mathematical Sciences, National Taiwan University, Taipei 10617, Taiwan.

E-mail: warmsay@yahoo.com.tw

Institute of Applied Mathematical Sciences, National Taiwan University, Taipei 10617, Taiwan.

E-mail: chiangct@ntu.edu.tw

(Received July 2016; accepted February 2018)