Statistica Sinica Preprint No: SS-2016-0348	
Title	Maximum Partial-Rank Correlation Estimation for Left-Truncated and Right-Censored Survival Data
Manuscript ID	SS-2016-0348
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202016.0348
Complete List of Authors	Shao-Hsuan Wang and Chin-Tsang Chiang
Corresponding Author	Chin-Tsang Chiang
E-mail	chiangct@ntu.edu.tw
Notice: Accepted version subject to English editing.	

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)

Maximum Partial-Rank Correlation Estimation for Left-Truncated and Right-Censored Survival Data

Shao-Hsuan Wang and Chin-Tsang Chiang
Institute of Applied Mathematical Sciences, National Taiwan University

February 24, 2018

Abstract

This article presents a general single-index hazards 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. Meanwhile, an efficient computational algorithm is offered to carry out the maximization of the constructed target function. Further, the developed approach can be extended to deal with right-truncation and left-censoring under a reverse time hazards regression model. Following the theoretical development of the maximum rank correlation estimator in the literature, we can also 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 its competitors. As for the application of our methodology, it is illustrated by data from the US Health and Retirement Study.

1 Introduction

In survival analysis, the incident and prevalent cohort sampling schemes have been widely adopted to collect a survival sample. Due to cost and time constraints, prevalent cohort approach is generally more efficient than the incident cohort one to accumulate enough failure cases, especially for the

Keywords: asymptotic normality; consistency; left-censoring; left-truncation; partial-rank correlation estimation; rank correlation estimation; random weighted bootstrap; right-censoring; right-truncation; U-statistic.

study of a rare disease. As the analyzed data in the US Health and Retirement Study (HRS), initially non-institutionalized persons were recruited by a cross-sectional sampling criterion between 1992 and 1993. White non-Hispanic men and women in such survival data have already experienced initiating events (birth) but have not experienced failure events (death) before the recruitment. 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 have experienced failure events before the recruitment are not observed and, thus, their failure times are left-truncated, i.e. $\{T^* \leq A^*\}$. Since some individuals might be lost to follow-up or drop out during the study period, their failure times are right-censored. Under a general single-index hazards regression model, our research aims to 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 with the conditional distribution of (T^*, A^*, Z^*) on $\{T^* \geq A^*\}$, where $Z^* = (Z_1^*, \ldots, Z_p^*)^\top$ are the covariates of interest with the support $\mathcal{Z}$. A rather mild covariate-dependent truncation and censoring mechanism is further assumed in our study. In addition, this article explores the relation between a continuous failure time T^* and covariates Z^* through an appealing single-index hazards 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 non-negative bivariate function and is strictly decreasing in u for each t, $\beta_0^\top z$ is a single-index with β_0 being the index coefficients, and τ is the end of the study period. Due to the identifiability of β_0 up to scale, a parametric system is adopted with $\beta_0 = (1, \theta_0^\top)^\top$ and $\theta_0 = (\theta_{01}, \ldots, \theta_{0p-1})^\top$ being 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 inter-

preted as the relative effect of Z_k^*, compared to Z_1^*, on the hazard function, $k = 2, \ldots, p$. In fact, several specific forms of model (1.1) include the proportional hazards regression $\lambda_0(t)\exp(\beta_0^\top z)$ (Cox (1972)), the additive hazards 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 the Cox's proportional hazards model and the assumption of covariate-dependent truncation and censoring, Wang, Brookmeyer, and Jewell (1993) proposed the maximum partial likelihood estimator of β_0 based on left-truncated and right-censored survival data. In the spirit of the estimation by Lin and Ying (1994) for the additive hazards model with censored survival data, Huang and Qin (2013) further took into account left-truncation and presented the modified conditional estimating equation estimator. Currently, there is still no estimation approach for the transformation model when the distributions of truncation and censoring times are covariate-dependent.

In the data analysis of the HRS, the violation of 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 non-stationary disease incidence are not appropriate to describe the effects of body mass index (BMI), level of education, and smoking status on the life expectancy. In addition to this, our another testing procedure confirms the inadequacy of proportional and additive hazards regression models. A more general formulation in model (1.1) and its statistical inferences become necessary in application. Based on a new partial-rank correlation function, an approach is developed to estimate β_0 in model (1.1). Meanwhile, an effective computational algorithm is provided to compute the presented maximum partial-rank correlation estimator. As we can observe in the partial-rank correlation estimation of Khan and Tamer (2007) for right-censored survival data, 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 similarly established according to 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 hazards 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 $\lambda_r(t|z) = f(t|z)/(1-S(t|z))$ with $f(t|z)$ and $S(t|z)$ being the conditional density and survival functions of $T^* = t$ on $Z^* = z$, and $\lambda_r(t,u)$ is an unknown non-negative bivariate function and is strictly decreasing in u for each t.

This article is organized as follows. In Section 2, a partial-rank correlation function is proposed as a basis of estimation for 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. Meanwhile, 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 were conducted to investigate the finite-sample performance of the proposed estimator and its competitors. The HRS data were also analyzed in Section 5 to show the usefulness of our methodology. Section 6 summarizes the findings in this study and makes some remarks for future research. As for the proofs of the main results, they are relegated to the appendix.

2 Partial-Rank Correlation Function and Its Extension

For survival data with left-truncation and right-censoring, an approach is developed 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 covers some particular cases and can be reasonably extended to deal with 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 the non-censoring indicator with $\mathrm{I}(\cdot)$ being the indicator function. The notations $\wedge$ and $\vee$ are also used to stand for 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 of adopting a truncation value a and a censoring value c in $\mathrm{Q}(z_1, z_2; a, c)$ is mainly to adjust for the sampling bias caused by left-truncation and 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 $Z^* = z$, it is further implied by model (1.1) 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 $\tau > c > a \geq 0$*

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

□

Following the proof of a maximizer of the rank correlation function in Han (1987), it can also be ensured by (2.1), Lemma 1, and the equality

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

that β_0 is a maximizer of

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

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 indicated 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 was constructed by $\mathrm{P}(Y_1 > Y_2, \delta_2 = 1, \beta^\top Z_1 > \beta^\top Z_2)$. In terms of $Q(z_1, z_2; a, c)$ in (2.1), it 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), $\ell = 1, 2$. In conjunction with the presence of left-truncation, we further address 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:

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

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 calendar times of initiating events, recruitment, failure events, and censoring events of two independent units. Apparently, 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).

In the course of deriving the main results, an additional assumption is further made:

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

As in the context of rank correlation estimation, assumption A3 is drawn for 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 2.1. *Under model* (1.1) *and assumptions A1-A3,*

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

Proof. See Appendix. □

It is noteworthy 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 type of 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 to be

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

Case 2. (right-censored survival data) In the presence of right-censoring, A_1 and A_2 are set to be 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 only subject to left-truncation, it is natural to specify $C_1 = C_2 = \infty$ and modify assumption A2 as $A^* \perp T^*|Z^*$. In this setup, our partial-rank correlation function can be rewritten as

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

2.2 An Extension to Right-Truncated and Left-Censored Data

In insurance applications and AIDS cohort studies (cf. Kaminsky (1987) and Kalbleisch 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 time period $[0, \tau]$, i.e. $X_0^* + T^* \leq \tau$. As shown in the Australian AIDS data (cf. Cui (1999)), $(X_0^* + T^*)$ may not be recorded before the chronological time X_1, which can be a determined or random time, with $X_1 \leq \tau$. 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 $\{D^* \geq T^*\}$. It was indicated by Lagakos, Barraj, and Gruttola (1988) and Cui (1999) that the reverse survival time $S^* = \tau - T^*$, X_0^*, and $\tau - X_1$ can be regarded as the roles of failure time, truncation time, and residual censoring time 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. For this reason, a formulation in hazards regression model (1.1) is adopted in reverse time hazards regression model (1.2).

Let $(T, D = \tau - X_0, Z)$ represent the observed lag, right-truncated time, and covariates, and the joint distribution of (T, D, Z) is the same with the conditional distribution of (T^*, D^*, Z^*) on $\{D^* \geq T^*\}$. It 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 $\{S^* \geq X_0^*\}$. 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 $\delta_{c\ell} = I(S_\ell \leq \tau - X_{1\ell} + X_{0\ell}) = I(S_\ell \leq \tau - C_{\tau\ell})$, $\ell = 1, 2$, one has the following partial-rank correlation function:

$$\begin{aligned} \mathrm{C}_\tau(\beta) &= \mathrm{E}\Big[Q(Z_1, Z_2; X_{01} \vee X_{02}, (\tau - X_{11} + X_{01}) \wedge (\tau - X_{12} + X_{02}))\mathrm{I}(\beta^\top Z_1 > \beta^\top Z_2)\Big] \\ &= \mathrm{P}(Y_{c1} > Y_{c2} > (X_{01} \vee X_{02}), \delta_{c2} = 1, \beta^\top Z_1 > \beta^\top Z_2). \end{aligned} \tag{2.10}$$

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 $\delta_{\tau\ell} = I(T_\ell \geq C_{\tau\ell})$, $\ell = 1, 2$. 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 only subject to right-truncation, 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 further 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 analog of $\mathrm{C}(\beta)$ is naturally given by a U-statistic of the form:

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

In light of the fact that β_0 is a maximizer of $\mathrm{C}(\beta)$, we estimate β with a 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 becomes necessary to compute the maximum partial-rank correlation estimator $\hat{\beta}$. For the constrained optimization of $\mathrm{C}_n(\beta)$, a direct maximization is generally impractical and difficult because such a target function is not differentiable with respect to β.

In the setup of complete failure time data, Wang and Chiang (2017) provided an effective procedure to carry out the maximization of 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

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

where $s(\upsilon) = 1/(1+\exp(-\upsilon))$ is a sigmoid function and σ is a tuning parameter, $g_{n\sigma}(\beta)$ denote the gradient function of $\mathrm{C}_{n\sigma}(\beta)$, $(\epsilon_1, \epsilon_2, r)$ be pre-chosen positive values with $0 < r < 1$, and $\|\cdot\|$ stand for the Euclidean norm of a vector. Provided that $\sigma = o(1/\sqrt{n})$, Ma and Huang (2005) showed that a maximizer of $\mathrm{C}_{n\sigma}(\beta)$ and $\hat{\beta}$ have the same asymptotic distribution. The following computational algorithm has been justified by Wang and Chiang (2017) to be theoretically valid and practically feasible in 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 update $\hat{\beta}^{(k+1)}$ 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. Repeatedly implement 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 be $o(1/\sqrt{n})$ 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):

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

3.2 Consistency, Asymptotic Normality, and Bootstrap Approximation

Let $\mathcal{N}_\theta$ be a neighborhood of θ in Θ, X stand for 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

$$\begin{aligned} \tau(x, \theta) \quad = \quad & \mathrm{P}(T > t > (A \vee a), (C + A) \wedge (c + a) > t, Z^\top(1, \theta) > z^\top(1, \theta)) \\ & + \mathrm{P}(t > T > (A \vee a), (C + A) \wedge (c + a) > T, z^\top(1, \theta) > Z^\top(1, \theta)) \quad \forall\ (x, \theta) \in \mathcal{X} \times \Theta. \end{aligned}$$

Some regularity conditions are further assumed:

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

A5. $\|\partial_\theta^2 \tau(x, \theta) - \partial_\theta^2 \tau(x, \theta_0)\| \leq M \|\theta - \theta_0\|$ 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 can also establish the consistency and asymptotic normality of $\hat{\theta}$ as follows:

Theorem 3.1. *Under model* (1.1) *and assumptions A1-A7,* $\hat{\theta} \xrightarrow{p} \theta_0$ *and* $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, \Sigma_0)$.

□

Accompanied with 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 through a smoothing estimation technique (cf. Sherman (1993)), a weighted bootstrap approximation of the sampling distribution of $\hat{\theta}$ is generally preferred in practical implementation.

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

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

where $w_i = \xi_i / \sum_{j=1}^n \xi_j$, $i = 1, \ldots, n$, 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 3.2. *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 Appendix. □

Let $\sigma^{\omega}(\hat{\theta}_k)$ and $q_{\varsigma}^{\omega}(\hat{\theta}_k)$ be the standard deviation and 100ςth, $0 < \varsigma < 1$, quantile of $\rho(\hat{\theta}^{\omega} - \hat{\theta})$, $k = 1, \ldots, (p-1)$, and z_{ς} be the 100ςth quantile of the standard normal distribution. It follows from Theorems 3.1-3.2 that approximate $100(1-\alpha)\%$, $0 < \alpha < 1$, quantile-type and normal-type bootstrap confidence intervals of θ_{0k} can be constructed by

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

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

4 Simulations

In this section, we conducted simulation experiments to investigate the finite-sample performance of the proposed estimator and its competitors. To assure numerical stability, the simulation results

were based on 1000 replications with the sample sizes (n) of 200 and 400, and the bootstrap inferences were drawn from 500 bootstrap samples with $\xi_1, \ldots, \xi_n \overset{i.i.d.}{\sim} Gamma(4, 2)$. Three hazards models with $Z^* = (Z_1^*, Z_2^*, Z_3^*)^\top$ and the same index coefficients $\beta_0 = (1, 1, 1)^\top$ were further studied under variant setups of left-truncation and right-censoring. Moreover, conditioning on $Z^* = z$, the residual censoring time $C = r_0(U(0, z_2) + 0.1)$ was independently generated with r_0 being specified to produce the censoring rates (*c.r.*) of 20% and 40%.

Example 4.1. A mixture of discrete and continuous covariate vector Z^* was specified with $Z_1^* \sim N(0, 1)$, $Z_2^* \sim U(0, 1)$, and $Z_3^* \sim U(\{1, 2, \ldots, 10\})$. The following proportional hazards regression model was designed to generate T^*:

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

Conditioning on $Z^* = z$, $A^* = r_1(U(0, 0.5) + |z_1|\mathrm{I}(|z_1| < 0.5))$ was independently generated with the proportions of untruncated data (*p.u.*), i.e. $\mathrm{P}(A^* > T^*)$, 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 (1993) for the proportional hazards regression model and the conditional estimating equation estimator $\bar{\beta}$ of Huang and Qin (2013) for the additive hazards regression model. Table 1 displays the means and standard deviations of 1000 estimates for variant combinations of sample sizes, censoring rates, and proportions of untruncated data. Due to model misspecification, $\bar{\beta}$ has relatively large bias and standard deviation in the model formulation of M1. It is further observed in this table that 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 one shall see in the next two examples, the maximum partial likelihood estimator has very poor performance under 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 1000 bootstrap standard errors and 95% quantile-type bootstrap confidence intervals are found to toward the standard deviations and 95% quantile intervals of 1000 estimates as n increases or $c.r.$ decreases. The empirical coverage probabilities of β_0, which are also exhibited in Table 4, are slightly higher than the nominal level of 0.95 for $(n, c.r.) = (200, 40\%)$ and stay around this nominal level for the rest cases.

Example 4.2. In this simulation scenario, a random vector $(Z_{01}^*, 1/Z_{02}^*, 1/Z_{03}^*)^\top$ was specified to follow a multivariate normal distribution with mean of zero, standard deviation of one, and pairwise correlation of 0.5. Further, the joint distribution of Z^* was designed to be the same with that of $(10Z_{01}^*, 10Z_{02}^*, 10Z_{03}^*)^\top$ on $\{Z_{01}^* + Z_{02}^* + Z_{03}^* > 0\}$. Moreover, T^* was generated from the following additive hazards regression model:

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

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

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

Example 4.3. With the triplets $(Z_{01}^*, Z_{02}^*, Z_{03}^*)^\top$ in Example 4.2, the joint distribution of Z^* was

specified to be the same with that of $(Z_{01}^*, Z_{02}^*, Z_{03}^*)^\top$ on $\{Z_{01}^* + Z_{02}^* + Z_{03}^* > 0\}$. The hazards regression model of T^* on $Z^* = z$ was further designed to be

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

Conditioning on $Z^* = z$, a covariate-dependent truncation time $A^* = U(0, 10) + |z_3|\mathrm{I}(|z_3| < 10)$ was also set with $p.u. = 0.9$. It is noted that the above model is neither the proportional hazards regression model nor the additive hazards 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 1000 maximum partial-rank correlation estimates are very close to β_0. For the standard deviation of $\hat{\beta}$, it decreases as n increases and $c.r.$ decreases. As for the performance of the weighted bootstrap standard error and confidence interval, the conclusions can be drawn as those in Example 4.1.

5 An Analysis of the HRS Data

Our partial-rank correlation estimation was applied 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, who were born between 1931 and 1934, was recruited by a cross-sectional sampling scheme between 1992 and 1993 and was followed up until 2012. By excluding those with missing covariates of interest, a total of 4323 white non-Hispanic men and 4724 white non-Hispanic women were collected in the first interview. For each individual, the birth date, gender, self-reported body mass index (*BMI*), level of education, and smoking status were investigated in this data analysis. On the first reported information, the smoking status was defined as "never smoked" (*nsmok*), "stopped smoking" (*ssmok*), and "currently smoking" (*csmok*), and the 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 the last observed date of studied individuals were further determined by the National Death Index (NDI) and exit interview. Since some individuals have 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* being the reference category and 1 representing *csmok* and *ssmok*, respectively. For the level of education, *hedu* was treated as the reference category and 1 represents *ledu* and *medu* in the dummy variables Z_3^* and Z_4^*. Since the overweight and underweight, which are evaluated in terms of *BMI*, might decrease life expectancy, the designed variables $Z_5^* = BMI_a$ and $Z_6^* = BMI_a^2$ were used in 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 was further considered as a stratification variable. Based on such 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 through a more general hazards regression model (1.1). By means of the partial-rank correlation estimation, it is shown in Table 5 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. As one can see, 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 a high-education person, a low-education one has a significantly higher mortality risk whereas a median-education one is not significantly different in life expectancy. Further, a higher or lower *BMI* tends to increase the hazard rate of transition to death. In addition to this finding, the mortality risk of a overweight individual is generally higher than that of a underweight one.

When the truncation time is covariate-independent, i.e. $f_{A^*}(a|z) = f_{A^*}(a)$, 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 a maximizer of the following sample analogue of $\mathrm{C}_A(\beta) = P(A_1 > A_2, \beta^\top Z_1 > \beta^\top Z_2)$:

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

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. It is noted 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 the corresponding bootstrap standard errors being $(0.000, 0.223, 0.217, 0.119, 0.503, 0.320)$ and $(0.000, 0.310, 0.773, 0.389, 0.648, 0.760)$. Both $\hat{\beta}$ in Table 5 and $\breve{\beta}$ are found to have different explanations on the life expectancy. In addition the values of $\mathcal{T}_0$ are obtained to be 4.45 for men and 3.46 for women. From their bootstrap p-values 0.000 and 0.000, one can 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 hazards regression models was further investigated in this data analysis. 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. 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 $\lambda(t, \beta_0^\top z) = \lambda_0(t) + \beta_0^\top z$. For this research objective, 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) for men and (6.50,2.38) and (0.000,0.003) for women. These numerical results show that the proportional and additive hazards regression models are not appropriate to characterize the effects of covariates on the hazard rates of transition to death. Based on this conclusion, a challenging task remains in 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 hazards regression model with left-truncated and right-censored survival data. Meanwhile, an effective computational algorithm is employed to carry out such a constraint non-differentiable optimization problem. The developed approach can be further extended to a reversed time hazards 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 a general weighted bootstrap approximations of the sampling quantities of interest related to the proposed estimator. The numerical studies also show that our estimator has very satisfactory performance.

In terms of the constructed partial-rank correlation function, the single-index $\beta_0^\top Z^*$ is shown to enjoy the existence, optimality, and uniqueness up to 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 non-negative bivariate function and is strictly increasing in u for each t. This is because the monotonicity of $S(t, u)$ in u for each t cannot imply the monotonicity of $\lambda(t, u)$ in u for each t. Currently, there is still no estimation and inference procedures for such a model formulation with left-truncated and right-censored survival data. A methodological challenge for estimating the index coefficients remains for future research. In our data analysis, some testing procedures, which are based on Hausman-type test statistics, are built to examine the distribution feature 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), one cannot conclude the adequacy of covariate-independent truncation, proportional hazards model, and additive hazards model. A more 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 $H_{20}^* : \lambda(t, \beta_0^\top z) = \lambda_0(t) + \beta_0^\top z$.

Acknowledgements. 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.

REFERENCES

Aalen, O. (1980). *A model for nonparametric regression analysis of counting processes.* Lecture Notes in Statistics 2, 1-25, Springer, New York.

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

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* DOI: 10.1111/biom.12839.

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. Roy. Statist. Soc.* B **34** 187-220.

Cox D. R. (1975). Partial likelihood. *Biometrika* **62** 269-276.

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. Stat. Plan. 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. *Stat. 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.

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.

APPENDIX

Proof of Theorem 2.1. 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 $(z_{01}, z_{02}) = (z_2, z_1)$, it is ensured by assumption A2 that

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

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

An application of Lemma 1 further leads to

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

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

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

By the law of iterated expectation, one also has $\mathrm{C}(\beta) = \mathrm{E}[\Gamma(Z_1, Z_2)\ \mathrm{I}(\beta^\top Z_1 > \beta^\top Z_2)]$. 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 $\mathrm{C}(\beta)$. The proof is completed. □

Proof of Theorem 3.2. By the equation (7) in Sherman (1993) and assumptions A1-A7, it follows that

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

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 of zero and variance-covariance matrix of Δ_0. An application of Theorem 2 in Sherman (1993) further leads to

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

As 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 $W_n = \sum_{j=1}^n V_0^{-1} \xi_j u_j/(\sqrt{n}\mathrm{E}[\xi])$. Since $\hat{\beta}^\omega = (1, \hat{\theta}^\omega)^\top$ is a maximizer of $\mathrm{C}_n^\omega(\beta)$, it yields that

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

Coupled with (A.7), one further has

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

which implies that

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

By (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 of zero and variance-covariance matrix of Σ_0. By Theorem 3.1, (A.12), and the probability inequality, Theorem 3.2 is, thus, established. □

TABLE 1

The means (standard deviations) of 1000 estimates under model M1 for the sample sizes (n) of 200 and 400, the proportions of untruncated data (p.u.) of 0.1, 0.5, and 0.9, and the 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)	(0101)	(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 1000 estimates under model M2 for the sample sizes (n) of 200 and 400 and the 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)

TABLE 3

The means (standard deviations) of 1000 estimates under model M3 for the sample sizes (n) of 200 and 400 and the 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)

TABLE 4

The standard deviations (s.d.), the bootstrap standard errors (b.s.e), 95% quantile interval (q.i.), quantile-type bootstrap confidence intervals (q.b.c.i.), and the empirical coverage probabilities (c.p.) of 1000 *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

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)

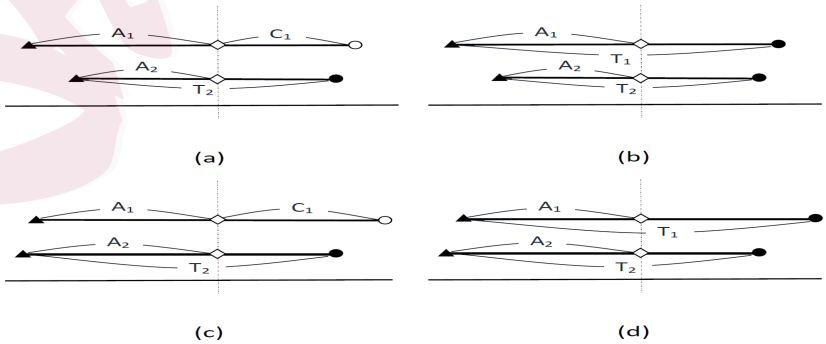

Figure 1: Relative positions of the calendar times of initiating event (▲), recruitment (◇), failure event (●), and censoring event (○).