Annals of the Institute of Statistical Mathematics manuscript No.
(will be inserted by the editor)

Supplementary Material for "Semiparametric Methods for Left-Truncated and Right-Censored Survival Data with Covariate Measurement Error"

Li-Pang Chen · Grace Y. Yi

Received: date / Revised: date

Abstract This supplement contains the discussion of the length-biased sampling, proofs of all the theoretical results in the main paper, and additional numerical results.

A Length-Biased Sampling Data with Measurement Error

A.1 Length-Biased Sampling

In the foregoing development, we leave the distribution of left-truncation A^* discussed in Section 2 unspecified. If we impose certain assumptions on A^*, the preceding development carry through and the new results can then generalize existing work. For instance, considered by Wang (1991) and De Uña-Alvarez (2004), suppose the incidence of disease onset follows a stationary Poisson distribution, then the truncation time follows a uniform distribution. Under this situation, the survival time in the prevalent cohort has a length-biased sampling distribution, because the probability of a survival time is proportional to the length of survival time (e.g., Huang and Qin 2011; Huang et al. 2012).

Consistent with Huang et al. (2012), assume the following conditions for the calendar time of the initial event:

(A1) The variable (T^*, V^*) is independent of u, where u is the time of the occurance of the disease incidence.

(A2) Disease incidence occurs over calendar time at a constant rate.

Then given $V = v$, the conditional density function of (T, A) is

$$\frac{f\,(t|v)}{\int_0^\infty \alpha f(\alpha|v)d\alpha}$$

Li-Pang Chen
Department of Statistical and Actuarial Sciences, University of Western Ontario,
1151 Richmond St, London, ON N6A 3K7
E-mail: lchen723@uwo.ca

Corresponding Author: Grace Y. Yi
Department of Statistical and Actuarial Sciences, Department of Computer Science, University of Western Ontario,
1151 Richmond St, London, ON N6A 3K7
E-mail: gyi5@uwo.ca

(Lancaster 1990; Huang et al. 2012), and the survival time T has a length-biased conditional density function:

$$\frac{tf(t|v)}{\int_0^\infty \alpha f(\alpha|v)d\alpha}.$$

Let C_i be the censoring time for subject i. Then we have $Y_i = \min\{T_i, A_i + C_i\}$ and $\Delta_i = \min\{T_i, A_i + C_i\}$. Noting that $\int_0^\infty \alpha f(\alpha|v)d\alpha = \int_0^\infty S(\alpha|v)\,d\alpha$, by Assumptions (A1) and (A2) and the independent censoringship, the likelihood function of (Y_i, A_i, Δ_i) given V_i can be constructed as

$$L_{LB} \propto \prod_{i=1}^{n} \frac{f(y_i|v_i)^{\delta_i} S(y_i|v_i)^{1-\delta_i}}{\int_0^\infty S(\alpha|v_i)d\alpha}, \tag{A.1}$$

which can be decomposed as the product of

$$L_{C,LB} = \prod_{i=1}^{n} \frac{f(y_i|v_i)^{\delta_i} S(y_i|v_i)^{1-\delta_i}}{S(a_i|v_i)}$$

and

$$L_{M,LB} = \prod_{i=1}^{n} \frac{S(a_i|v_i)}{\int_0^\infty S(\alpha|v_i)d\alpha}.$$

Compared with the likelihood function (4), the likelihood function ($A.1$) does not involve the estimation procedure of density function $h(a)$, which can be thought of as a degenerate version of (4), agreeing with the standard view that the length-biased sampling is regarded as a special case of the LTRC data (e.g., Asgharian et al. 2002; Qin and Shen 2010). To develop estimating procedures using ($A.1$), we need only to deal with $\Lambda_0(\cdot)$ and β but not $h(\cdot)$ as (4). In the absence of covariate measurement error, many authors developed methods to handle length-biased data. For example, Qin and Shen (2010) proposed the weighted estimating equation approach, and Huang et al. (2012) explored a pseudo-profile likelihood method. Here, we further accommodate the feature of covariate measurement error for length-biased data and develop a valid inference method.

A.2 Estimation of Parameters for Survival Data

From the decomposition of ($A.1$), we can see that the conditional likelihood $L_{C,LB}$ is the same as those of (4). Hence, the estimator of the conditional likelihood, $\widehat{\beta}_{LB}$, can be derived from (14). To emphasize the different setting, let $\widehat{\ell}^*_{C,LB}$ denote the corrected conditional log likelihood under the length-biased sampling, which leads to an estimator of β:

$$\widehat{\beta}_{LB} = \underset{\beta}{\operatorname{argmax}}\, \widehat{\ell}^*_{C,LB}. \tag{A.2}$$

On the other hand, for the marginal likelihood $L_{M,LB}$, there is no density function $h(\cdot)$, so we simply apply the regression calibration (18) to replace the error-prone covariate X_i. Hence, the corrected marginal log likelihood, $\widehat{\ell}^*_{M,LB}$, has a similar form to (21) except for the estimate of $H(\cdot)$. As a result, an estimator of β is given by

$$\widetilde{\beta}_{LB} = \underset{\beta}{\operatorname{argmax}}(\widehat{\ell}^*_{C,LB} + \widehat{\ell}^*_{M,LB}). \tag{A.3}$$

A.3 Asymptotic Results

Let $\mu_{LB}\left(\widetilde{x}_i, z_i\right) = \int_0^\tau \exp\left\{-\Lambda_0(u)\exp(\widetilde{x}_i^\top\beta_{x0} + z_i^\top\beta_{z0})\right\}du$ be the function of $(\widetilde{x}_i, z_i)$. Define

$$
\begin{aligned}
&\Psi_{LB}\left(w_i, \widetilde{x}_i, z_i, y_i, a_i\right) \\
&= \int_0^\tau \left\{\widetilde{v}_i - \frac{\mathcal{S}^{(1)}(u,\beta_0)}{\mathcal{S}^{(0)}(u,\beta_0)} + \begin{pmatrix} \Sigma_\epsilon\beta_{x0} \\ 0_q \end{pmatrix}\right\} dN_i(u) \\
&\quad - \int_0^\tau \frac{\exp\left(\widetilde{v}_i^\top\beta_0\right) I(a_i \le u \le y_i)}{\mathcal{S}^{(0)}(u,\beta_0)} \left(\widetilde{v}_i - \frac{\mathcal{S}^{(1)}(u,\beta_0)}{\mathcal{S}^{(0)}(u,\beta_0)}\right) dE\left\{N_i(u)\right\} \\
&\quad - \left[\int_{-\infty}^\infty \int_0^\tau \frac{\partial}{\partial\beta_0}\left\{\frac{dN_i(u)}{\mathcal{S}^{(0)}(u,\beta_0)} + \frac{d\mathcal{N}(u)\exp\left(\widetilde{v}_i^\top\beta_0\right)I(a_i \le u \le y_i)}{\mathcal{S}^{(0)}(u,\beta_0)^2}\right\} m(\beta_{x0}) \right. \\
&\quad \left. \times \exp\left(\widehat{v}^\top\beta_0\right) I(u \le a \le \tau) dG(a, \widehat{v})\right] \\
&\quad + \left[\int_{-\infty}^\infty \int_0^\tau \left\{\frac{1}{\mu_{LB}\left(\widetilde{x}, z\right)} \frac{\partial}{\partial\beta_0}\psi_{LB,i}(\beta_0|\widetilde{x}, z) \right.\right. \\
&\quad \left.\left. - \frac{\partial\mu_{LB}\left(\widetilde{x}, z\right)}{\partial\beta_0} \frac{1}{\mu_{LB}^2\left(\widetilde{x}, z\right)} \psi_{LB,i}(\beta_0|\widetilde{x}, z)\right\} dG(a, \widehat{v})\right] \\
&\quad - \frac{\partial}{\partial\beta_0}\Lambda_0(a_i)\exp\left(\widehat{v}_i^\top\beta_0\right) - \frac{1}{\mu_{LB}\left(\widetilde{x}_i, z_i\right)} \frac{\partial\mu_{LB}\left(\widetilde{x}_i, z_i\right)}{\partial\beta_0},
\end{aligned}
$$

where $\psi_{LB,i}(\beta_0|\widetilde{x}, z)$ is similarly defined by (25) with the integral relative to $dH(\cdot)$ removed, i.e.,

$$
\begin{aligned}
\psi_{LB,i}(\beta_0|\widetilde{x}, z) = &\int_0^\tau \int_0^\tau S(\xi|\widetilde{x}, z)\left\{\frac{dN_i(u)}{\mathcal{S}^{(0)}(u,\beta_0)} \right. \\
&\left. - \frac{d\mathcal{N}(u)\exp\left(w_i^{*\top}\beta_{x0} + z_i^\top\beta_{z0}\right)I(a_i \le u \le y_i)}{\left\{\mathcal{S}^{(0)}(u,\beta_0)\right\}^2}\right\} m(\beta_{x0}) \\
&\times \exp\left(\widetilde{x}^\top\beta_{x0} + z^\top\beta_{z0}\right) d\xi + o_p(1).
\end{aligned}
$$

We now establish the following results whose proof is deferred to Section **??**.

Theorem A.1 *Under regularity conditions given in Appendix in the main paper, we have that $n \to \infty$,*

(1) $\widetilde{\beta}_{LB} \xrightarrow{p} \beta_0$;
(2) $\sqrt{n}\left(\widetilde{\beta}_{LB} - \beta_0\right) \xrightarrow{d} N(0, \mathcal{A}_{LB}^{-1}\mathcal{B}_{LB}\mathcal{A}_{LB}^{-1})$,

where $\mathcal{B}_{LB} = E(\Psi_{LB,i}^{\otimes 2})$ *with* $\Psi_{LB,i} = \Psi_{LB}\left(W_i, \widetilde{X}_i, Z_i, Y_i, A_i\right)$, *and*

$$
\begin{aligned}
&\mathcal{A}_{LB} \\
&= \int_0^\tau \left[\left\{\frac{\mathcal{S}^{(2)}(u,\beta_0)}{\mathcal{S}^{(0)}(u,\beta_0)} - \left(\frac{\mathcal{S}^{(1)}(u,\beta_0)}{\mathcal{S}^{(0)}(u,\beta_0)}\right)^{\otimes 2}\right\} - \begin{pmatrix} \Sigma_\epsilon & 0_{p\times q} \\ 0_{q\times p} & 0_{q\times q}\end{pmatrix}\right] dE\left\{N_i(u)\right\} \\
&\quad + E\left[\frac{\partial^2}{\partial\beta_0\partial\beta_0^\top}\Lambda_0(A_i)\exp\left(\widehat{V}_i^\top\beta_0\right) \right. \\
&\quad \left. + \left\{\mu_{LB}\left(\widetilde{X}_i, Z_i\right)\right\}^{-2}\left\{\mu_{LB}\left(\widetilde{X}_i, Z_i\right)\frac{\partial\left\{\mu_{LB}\left(\widetilde{X}_i, Z_i\right)\right\}^2}{\partial\beta_0\partial\beta_0^\top} - \left(\frac{\partial\mu_{LB}\left(\widetilde{X}_i, Z_i\right)}{\partial\beta_0}\right)^{\otimes 2}\right\}\right].
\end{aligned}
$$

The proof of Theorem A.1 (1) is similar to the proof of Theorem 3 (1) (presented in the manuscript) except for the inference of function $dH(a)$. The proof of Theorem A.1 (2) can be done by modifying the proof of Theorem 3 (2) (presented in the manuscript) and that of Huang et al. (2012) who developed asymptotic normality under the length-biased sampling setting.

B Preliminary Results

In this Appendix, we present the lemmas that are useful for proving the theorems in the manuscript.

Lemma B.1 *Let*

$$L_P^* = \prod_{i=1}^n \{m(\beta_x)\}^{\delta_i} \left[\frac{\exp\left(\widetilde{v}_i^\top \beta\right)}{\sum\limits_{j=1}^n \exp\left(\widetilde{v}_j^\top \beta\right) I(a_j \le y_i \le y_j)} \right]^{\delta_i}$$

and

$$\begin{aligned} L_R^* &= \prod_{i=1}^n \left[\lambda_0(y_i) \sum_{j=1}^n \exp\left(\widetilde{v}_j^\top \beta\right) \{m(\beta_x)\}^{-1} I(a_j \le y_i \le y_j) \right]^{\delta_i} \\ &\quad \times \exp\left[-\int \lambda_0(u) \sum_{j=1}^n \exp\left(\widetilde{v}_j^\top \beta\right) \{m(\beta_x)\}^{-1} I(a_j \le u \le y_j) du \right]. \end{aligned}$$

Then

(1) $L_C^* = L_P^* \times L_R^*$*;*
(2) L_R^* *is ancillary which does not convey the information of* β*.*

Proof:
Let $L_C^* = \exp\left(\ell_C^*\right)$ where ℓ_C^* is given by (10), and let

$$\mathcal{U}_i = \left[\sum_{j=1}^n \exp\left(\widetilde{v}_j^\top \beta\right) \{m(\beta_x)\}^{-1} I(a_j \le y_i \le y_j) \right]^{\delta_i}.$$

Then L_C^* can be written as

$$\begin{aligned} L_C^* &= \prod_{i=1}^n \frac{\left\{\lambda_0(y_i)\exp\left(\widetilde{v}_i^\top \beta\right)\right\}^{\delta_i} \exp\left[-\Lambda_0(y_i)\exp\left(\widetilde{v}_i^\top \beta\right)\{m(\beta_x)\}^{-1}\right]}{\exp\left[-\Lambda_0(a_i)\exp\left(\widetilde{v}_i^\top \beta\right)\{m(\beta_x)\}^{-1}\right]} \\ &= \prod_{i=1}^n \frac{\left\{\lambda_0(y_i)\exp\left(\widetilde{v}_i^\top \beta\right)\right\}^{\delta_i}}{\mathcal{U}_i} \times \prod_{i=1}^n \frac{\mathcal{U}_i \exp\left[-\Lambda_0(y_i)\exp\left(\widetilde{v}_i^\top \beta\right)\{m(\beta_x)\}^{-1}\right]}{\exp\left[-\Lambda_0(a_i)\exp\left(\widetilde{v}_i^\top \beta\right)\{m(\beta_x)\}^{-1}\right]} \\ &= L_P^* \times L_R^*. \end{aligned}$$

Analogous to the derivations of Wang et al. (1993), we can show that L_R^* is ancillary which does not convey the information of β, and hence, it is sufficient to obtain the estimator of β by maximizing L_P^*, or equivalently, deriving an estimator of β from $\log(L_C^*)$ is equivalent to deriving an estimator from using $\log(L_P^*)$ alone. □

Let

$$
\begin{aligned}
\widehat{\ell}^*_P &= \log\left(L^*_P\right) \\
&= \sum_{i=1}^n \int_0^\tau \left[\widetilde{v}_i^\top \beta + \frac{1}{2}\beta_x^\top \Sigma_\epsilon \beta_x - \log\left\{\sum_{j=1}^n \exp(\widetilde{v}_j^\top \beta) I(a_j \le u \le y_j)\right\}\right] dN_i(u).
\end{aligned}
$$

Lemma B.1 indicates that inference based on ℓ^*_C, defined in (10), is equivalent to that based on $\widehat{\ell}^*_P$ in the sense that ℓ^*_C and $\widehat{\ell}^*_P$ provide the same maximum likelihood estimator of β.

Hence, in the maximization of (23), using $\widehat{\ell}^*_C + \widehat{\ell}^*_M$ to perform inference about β is equivalent to using $\widehat{\ell}^* = \widehat{\ell}^*_P + \widehat{\ell}^*_M$. Corresponding to $\widehat{\ell}^*_P$ and $\widehat{\ell}^*_M$, we let

$$
\widetilde{\ell}^*_P = \sum_{i=1}^n \int_0^\tau \left[\widetilde{v}_i^\top \beta + \frac{1}{2}\beta_x^\top \Sigma_\epsilon \beta_x - \log\left\{E\left(\exp(\widetilde{V}_i^\top \beta) I(A_i \le u \le Y_i)\right)\right\}\right] dN_i(u)
$$

and

$$
\begin{aligned}
\widetilde{\ell}^*_M = \sum_{i=1}^n &\Big[\log\left\{dH(a_i)\right\} - \Lambda_0(a_i)\exp\left(\widehat{v}_i^\top \beta\right) \\
&- \log\left\{\int_0^\tau \exp\left\{-\Lambda_o(u)\exp(\widetilde{x}_i^\top \beta_x + z_i^\top \beta_z)\right\} dH(u)\right\}\Big],
\end{aligned}
$$

where $\widetilde{x}_i$ is defined in (17). Define $\widetilde{\ell}^* = \widetilde{\ell}^*_P + \widetilde{\ell}^*_M$.

Lemma B.2 *Under regularity conditions in Appendix in the main paper,*

$$
\sup_{\beta \in \Theta, t \in [0,\tau]} \left|\widehat{\Lambda}_0(t) - \Lambda_0(t)\right| \xrightarrow{a.s.} 0 \text{ as } n \to \infty.
$$

Lemma B.3 *Under regularity conditions in Appendix in the main paper,*

$$
\sup_{\beta \in \Theta, t \in [0,\tau]} \left|\frac{1}{n}\widehat{\ell}^* - \frac{1}{n}\widetilde{\ell}^*\right| \xrightarrow{a.s.} 0 \text{ as } n \to \infty.
$$

Proof:

This can be shown by proving the following claims.

Claim 1: $\sup_{\beta \in \Theta} \left|\frac{1}{n}\widehat{\ell}^*_P - \frac{1}{n}\widetilde{\ell}^*_P\right| \xrightarrow{a.s.}$ as $n \to \infty$.

Claim 2: $\sup_{\beta \in \Theta} \left|\frac{1}{n}\widehat{\ell}^*_M - \frac{1}{n}\widetilde{\ell}^*_M\right| \xrightarrow{a.s.}$ as $n \to \infty$. □

We also present Theorem 5.7 of van der Vaart (1998) here as the following lemma which will be used in subsequent proof.

Lemma B.4 *Let $M_n(\cdot)$ be random functions and let $M(\cdot)$ be a real-valued function of θ. Let θ_0 be the true value of θ. Suppose that for any $\epsilon > 0$,*

$$
\begin{aligned}
&\sup_{\theta \in \Theta} |M_n(\theta) - M(\theta)| \xrightarrow{p} 0; \\
&\sup_{\theta : d(\theta,\theta_0) \ge \epsilon} M(\theta) < M(\theta_0);
\end{aligned}
$$

where $d(\theta, \theta_0) = \|\theta - \theta_0\|$ is the Euclidean distance between θ and θ_0. Then any sequence of estimators $\widehat{\theta}_n$ with $M_n(\widehat{\theta}_n) \ge M_n(\theta_0) - o_P(1)$ converges in probability to θ_0.

C Proofs of the Theorems in Section 3

C.1 Proof of Theorem 1

The uniformly consistency of $\widehat{\Lambda}_0(t)$ comes from Lemma B.2. □

C.2 Proof of Theorem 2

Proof of Theorem 2 (1):
By Conditions (C3) and (C4), $\widetilde{\ell}^*_P$ is the sum of i.i.d. random functions. Then by Lemma 2 in the main paper and the Law of Large Numbers, we have that as $n \to \infty$,

$$\frac{1}{n}\widetilde{\ell}^*_P \xrightarrow{p} \kappa_P$$

for every β. By Claim 1 in Lemma B.3, we have that as $n \to \infty$,

$$\sup_{\beta \in \Theta, t \in [0,\tau]} \left| n^{-1}\widehat{\ell}^*_P - \kappa_P \right| \xrightarrow{a.s.} 0.$$

Therefore, by Lemma B.4, we have that as $n \to \infty$,

$$\widehat{\beta} \xrightarrow{p} \beta_0. \tag{C.1}$$

Proof of Theorem 2 (2):
Since $\widehat{\ell}^*_P = \sum\limits_{i=1}^{n} \int_0^\tau \left[\widetilde{v}_i^\top \beta + \frac{1}{2}\beta_x^\top \Sigma_{\epsilon;0}\beta_x - \log\left\{S^{(0)}(u,\beta)\right\}\right] dN_i(u)$, taking the derivative of $\widehat{\ell}^*_P$ with respect of β gives that

$$U_P(\beta) \triangleq \frac{\partial \widehat{\ell}^*_P}{\partial \beta} = \sum_{i=1}^{n} \int_0^\tau \left\{ \widetilde{v}_i + \begin{pmatrix} \Sigma_{\epsilon;0}\beta_x \\ 0_q \end{pmatrix} - \frac{S^{(1)}(u,\beta)}{S^{(0)}(u,\beta)} \right\} dN_i(u). \tag{C.2}$$

Since $\widehat{\beta}$ is the estimator satisfying $U_P(\widehat{\beta}) = 0$ and $\widehat{\beta}$ is the consistent estimator of β by (C.1), to show the asymptotic distribution of $\widehat{\beta}$, we consider the Taylor series expansion of $U_P(\widehat{\beta})$ around the true parameter β_0:

$$0 = U_P(\widehat{\beta}) = U_P(\beta_0) + \frac{\partial U_P(\beta_0)}{\partial \beta_0}(\widehat{\beta} - \beta_0) + o_p\left(\frac{1}{\sqrt{n}}\right), \tag{C.3}$$

yielding that

$$\sqrt{n}(\widehat{\beta} - \beta_0) = -\left(\frac{1}{n}\frac{\partial U_P(\beta_0)}{\partial \beta_0}\right)^{-1} \times \frac{1}{\sqrt{n}} U_P(\beta_0) + o_p(1). \tag{C.4}$$

To work out the asymptotic distribution of $\sqrt{n}\left(\widehat{\beta} - \beta_0\right)$, it suffices to determine the asymptotic behavior of $\frac{\partial U_P(\beta_0)}{\partial \beta_0}$ and $U_P(\cdot)$. To this end, we proceed with the following two steps.

Step 1: To examine the convergence of $\frac{\partial U_P(\beta_0)}{\partial \beta_0}$, we first note that $\{N_i(t) : t \in [0,\tau]\}$ is a Glivenko-Cantelli class (van der Vaart and Wellner 1996, Example 2.4.2), which gives that as $n \to \infty$,

$$\frac{1}{n}\sum_{i=1}^{n} dN_i(t) \xrightarrow{a.s.} dE\left\{N_i(t)\right\}$$

uniformly (van der Vaart 1998, Theorem 19.1). Then by (C.2) and the Uniform Law of Large Numbers, we have that as $n \to \infty$,

$$\begin{aligned} &-\frac{1}{n}\frac{\partial U_P(\beta_0)}{\partial \beta_0} \\ &= \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau}\left[\left\{\frac{S^{(2)}(u,\beta_0)}{S^{(0)}(u,\beta_0)} - \left(\frac{S^{(1)}(u,\beta_0)}{S^{(0)}(u,\beta_0)}\right)^{\otimes 2}\right\} - \begin{pmatrix} \Sigma_{\epsilon;0} & 0_{p\times q} \\ 0_{q\times p} & 0_{q\times q}\end{pmatrix}\right] dN_i(u) \\ &\xrightarrow{p} \mathcal{I}, \end{aligned} \tag{C.5}$$

where $\mathcal{I}$ is given by (16).

Step 2: To determine the asymptotic distribution of $U_P(\beta_0)$, we sort out the leading term of $U_P(\beta_0)$ which can be expressed as a sum of i.i.d. random variables. By the similar derivations for Theorem 2.1 of Lin and Wei (1989), we express

$$\frac{1}{\sqrt{n}}U_P(\beta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Phi\left(W_i, Z_i, Y_i, A_i\right) + o_p(1), \tag{C.6}$$

where

$$\begin{aligned} \Phi\left(w_i, z_i, y_i, a_i\right) = &\int_0^{\tau}\left\{\widetilde{v}_i - \frac{\mathcal{S}^{(1)}(u,\beta_0)}{\mathcal{S}^{(0)}(u,\beta_0)} + \begin{pmatrix} \Sigma_{\epsilon;0}\beta_{x0} \\ 0_q \end{pmatrix}\right\} dN_i(u) \\ &- \int_0^{\tau}\frac{\exp\left(\widetilde{v}_i^{\top}\beta_0\right) I(a_i \le u \le y_i)}{\mathcal{S}^{(0)}(u,\beta_0)}\left(\widetilde{v}_i - \frac{\mathcal{S}^{(1)}(u,\beta_0)}{\mathcal{S}^{(0)}(u,\beta_0)}\right) dE\left\{N_i(u)\right\}. \end{aligned} \tag{C.7}$$

Since the $\Phi\left(W_i, Z_i, Y_i, A_i\right)$ are i.i.d. random functions, applying the Central Limit Theorem yields that as $n \to \infty$,

$$\frac{1}{\sqrt{n}}U_P(\beta_0) \xrightarrow{d} N\left(0, \mathcal{J}\right), \tag{C.8}$$

where $\mathcal{J} = E\left(\Phi_i^{\otimes 2}\right)$ and $\Phi_i = \Phi\left(W_i, Z_i, Y_i, A_i\right)$.

Finally, combining (C.5) and (C.8) with (C.4) and applying the Slutsky's Theorem, we conclude that as $n \to \infty$,

$$\sqrt{n}(\widehat{\beta} - \beta_0) \xrightarrow{d} N\left(0, \mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}\right).$$

□

D Proofs of the Theorems in Section 4

The derivations in this appendix are in principle analogous to those of Appendix C. However, the technical details are a lot more complex than those of Appendix C, because no infinite dimensional parameters are involved with the key estimating function $U_P(\beta)$ in Appendix C while such parameters are contained in the estimating function considered here.

D.1 Proof of Theorem 3

Proof of Theorem 3 (1):

The proof is the same as that of Theorem 2 (2) except that $\widehat{\ell}_P^*$ and κ_P are replaced by $\widehat{\ell}^*$ and κ, respectively.

Proof of Theorem 3 (2):

To find the asymptotic distribution of $\widetilde{\beta}$, we note that $\widetilde{\beta}$ solves $U(\beta) = 0$, where

$$U(\beta) = \frac{\partial \widehat{\ell}^*}{\partial \beta} = \frac{\partial \widehat{\ell}_P^*}{\partial \beta} + \frac{\partial \widehat{\ell}_M^*}{\partial \beta}. \tag{D.1}$$

Considering the Taylor series expansion of $U(\widetilde{\beta})$ around β_0 gives that

$$0 = U(\widetilde{\beta}) = U(\beta_0) + \frac{\partial U(\beta_0)}{\partial \beta_0}(\widetilde{\beta} - \beta_0) + o_p\left(\frac{1}{\sqrt{n}}\right), \tag{D.2}$$

or equivalently,

$$\sqrt{n}(\widetilde{\beta} - \beta_0) = -\left(\frac{1}{n}\frac{\partial U(\beta_0)}{\partial \beta_0}\right)^{-1}\left(\frac{1}{\sqrt{n}}U(\beta_0)\right) + o_p(1).$$

Analogous to the examination of (C.4) in Appendix C.2, we proceed with the following two steps, separately examining $\frac{\partial U(\beta_0)}{\partial \beta_0}$ and $U(\beta_0)$. By (D.1), we note that the main difficulty here is caused by the involvement of the term $U_M(\beta) = \frac{\partial \widehat{\ell}_M^*}{\partial \beta}$, while $\frac{\partial \widehat{\ell}_P^*}{\partial \beta}$ is examined in Appendix C.2.

Step 1: To show the convergence of $\frac{\partial U(\beta_0)}{\partial \beta_0}$, we first define

$$\widehat{\mu}(\widehat{x}_i, z_i) = \int_0^\tau \exp\left\{-\widehat{\Lambda}_0(u)\exp(\widehat{x}_i^\top \beta_{x0} + z_i^\top \beta_{z0})\right\} d\widehat{H}(u), \tag{D.3}$$

and we let $\widehat{\mu}_i$ denote $\widehat{\mu}(\widehat{x}_i, z_i)$ for ease of notation. Since $\widehat{\mu}_{W^*}$ and $\widehat{\Sigma}_{W^*}$ in (18) are two consistent estimates, i.e., $\widehat{\mu}_{W^*} = \mu_{W^*} + o_p(1)$ and $\widehat{\Sigma}_{W^*} = \Sigma_{W^*} + o_p(1)$, then we have that $\widehat{X}_i = \widetilde{X}_i + o_p(1)$. In addition, $\widehat{\beta}$ and $\widehat{\Lambda}_0(\cdot)$ are consistent estimators due to Theorems 1 and 2 (a). Therefore, we have that $\widehat{H}(\cdot)$ determined by (20) is still the consistent estimator of $H(\cdot)$ (e.g., Wang 1991). Hence, by Theorem 1, we conclude that $\widehat{\mu}_i \xrightarrow{p} \mu_i$ as $n \to \infty$, where

$$\mu_i = \mu(\widetilde{x}_i, z_i) = \int_0^\tau \exp\left\{-\Lambda_0(u)\exp(\widetilde{x}_i^\top \beta_{x0} + z_i^\top \beta_{z0})\right\} dH(u).$$

Noting that

$$\begin{aligned}\frac{\partial U_M(\beta_0)}{\partial \beta_0} &= -\sum_{i=1}^n \frac{\partial^2}{\partial \beta_0 \partial \beta_0^\top}\widehat{\Lambda}_0(a_i)\exp\left(\widehat{x}_i^\top \beta_{x0} + z_i^\top \beta_{z0}\right) \\ &\quad -\sum_{i=1}^n \left\{\frac{1}{\widehat{\mu}_i}\frac{\partial^2 \widehat{\mu}_i}{\partial \beta_0 \partial \beta_0^\top} - \frac{1}{\widehat{\mu}_i^2}\left(\frac{\partial \widehat{\mu}_i}{\partial \beta_0}\right)^{\otimes 2}\right\},\end{aligned}$$

we obtain, by Theorem 2 and the Law of Large Numbers, that as $n \to \infty$,

$$-\frac{1}{n}\frac{\partial U(\beta_0)}{\partial \beta_0} \xrightarrow{p} \mathcal{A}, \tag{D.4}$$

where $\mathcal{A} = \mathcal{I} + \mathcal{A}_M$, and $\mathcal{I}$ and $\mathcal{A}_M$ are given by (16) and (27), respectively.

Step 2: We now derive the asymptotic distribution of $\frac{1}{\sqrt{n}}U(\beta_0)$. Since a sum of i.i.d. random variables of $\frac{\partial \widehat{\ell}_P^*}{\partial \beta}$ is established in (C.6) of Appendix C.2, it remains to examine $U_M(\beta_0) = \frac{\partial \widehat{\ell}_M^*}{\partial \beta}$ by (D.1). To this end, we make an important comment. Different from the partial likelihood score function $U_P(\beta)$ which involves the parameter β only, $U_M(\cdot)$ involves not only the parameter β but also the infinite dimensional parameter $\Lambda_0(\cdot)$. The goal here is to sort out the key term in $U_M(\beta_0)$ which can be expressed as a sum of i.i.d. random functions.
Define

$$\begin{aligned}\widetilde{U}_M(\beta_0) &= -\sum_{i=1}^n \frac{\partial}{\partial \beta_0}\Lambda_0(a_i)\exp\left(\widehat{v}_i^\top \beta_0\right) - \sum_{i=1}^n \frac{1}{\mu_i}\frac{\partial \mu_i}{\partial \beta_0} \\ &\triangleq -\sum_{i=1}^n U_{M,i},\end{aligned}$$

and write the difference between $U_M(\beta_0)$ and $\widetilde{U}_M(\beta_0)$ as

$$\frac{1}{\sqrt{n}}\left\{U_M(\beta_0) - \widetilde{U}_M(\beta_0)\right\} = U_1 + U_2, \tag{D.5}$$

where

$$U_1 = -\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\partial}{\partial \beta_0}\left\{\widehat{\Lambda}_0(a_i) - \Lambda_0(a_i)\right\}\exp\left(\widehat{v}_i^\top \beta_0\right) \tag{D.6}$$

and

$$U_2 = -\frac{1}{\sqrt{n}}\sum_{i=1}^n \left\{\frac{1}{\widehat{\mu}_i}\frac{\partial \widehat{\mu}_i}{\partial \beta} - \frac{1}{\mu_i}\frac{\partial \mu_i}{\partial \beta}\right\}. \tag{D.7}$$

We first examine U_1. Recall that $\mathcal{N}(t) = P\left(\Delta_i = 1, Y_i \le t\right)$ and let $d\bar{N}(t) = \frac{1}{n}\sum\limits_{i=1}^n dN_i(t)$. Then we can show that

$$\begin{aligned}U_1 = &-\frac{1}{\sqrt{n}}\sum_{j=1}^n \int_0^\tau \frac{\partial}{\partial \beta_0}\left[\frac{d\bar{N}(u)}{\mathcal{S}^{(0)}(u,\beta_0)} - \frac{d\mathcal{N}(u) S^{(0)}(u,\beta_0)}{\left\{\mathcal{S}^{(0)}(u,\beta_0)\right\}^2}\right] m(\beta_{x0}) \\ &\times \exp\left(\widehat{v}_j^\top \beta_0\right) I(u \le a_j \le \tau) + o_p(1).\end{aligned} \tag{D.8}$$

In addition, since $\frac{1}{n}\sum\limits_{j=1}^n \exp\left(\widehat{V}_j^\top \beta_0\right) I(u \le A_j \le \tau)$ is an average of i.i.d. random variables due to Conditions (C3), (C4), and (C5), we have that by the Law of Large Numbers, as $n \to \infty$, we write (e.g., Jiang 2010, p.61)

$$\begin{aligned}\frac{1}{n}\sum_{j=1}^n \exp\left(\widehat{V}_j^\top \beta_0\right) I(u \le A_j \le \tau) = &\int_{-\infty}^\infty \int_0^\tau \left\{\exp\left(\widehat{v}^\top \beta_0\right) I(u \le a \le \tau)\right\} dG(a, \widehat{v}) \\ &+ O_p\left(\frac{1}{\sqrt{n}}\right).\end{aligned} \tag{D.9}$$

Therefore, combining (D.8) and (D.9) gives

$$\begin{aligned}U_1 = &-\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\partial}{\partial \beta_0}\left[\int_{-\infty}^\infty \int_0^\tau \left\{\frac{dN_i(u)}{\mathcal{S}^{(0)}(u,\beta_0)} - \frac{d\mathcal{N}(u)\exp\left(\widetilde{v}_i^\top \beta_0\right) I(a_i \le u \le y_i)}{\left\{\mathcal{S}^{(0)}(u,\beta_0)\right\}^2}\right\}\right. \\ &\left. \times\, m(\beta_{x0})\exp\left(\widehat{v}^\top \beta_0\right) I(u \le a \le \tau)\right] dG(a, \widehat{v}) + o_p(1) \\ \triangleq &-\frac{1}{\sqrt{n}}\sum_{i=1}^n \Psi_{M1}\left(w_i, \widetilde{x}_i, z_i, y_i, a_i\right) + o_p(1).\end{aligned} \tag{D.10}$$

Next, we examine U_2. By analogy with the derivations of (D.9), (D.7) can be re-written as

$$U_2 = \sqrt{n}\int_{-\infty}^{\infty}\int_0^{\tau}\left(\frac{1}{\widehat{\mu}}\frac{\partial\widehat{\mu}}{\partial\beta} - \frac{1}{\mu}\frac{\partial\mu}{\partial\beta}\right) dG(a,\widehat{v}) + o_p(1), \tag{D.11}$$

where $\widehat{\mu} = \widehat{\mu}(\widehat{x}, z)$ and $\mu = \mu(\widetilde{x}, z)$.
We now express $\sqrt{n}(\widehat{\mu}-\mu)$ as a sum of i.i.d. random functions. Since

$$\begin{aligned}\sqrt{n}(\widehat{\mu}-\mu) = \sqrt{n}\int_0^{\tau} \Big[&\exp\left\{-\widehat{\Lambda}_0(u)\exp(\widetilde{x}^\top\beta_{x0} + z^\top\beta_{z0})\right\} \\ &- \exp\left\{-\Lambda_0(u)\exp(\widetilde{x}^\top\beta_{x0} + z^\top\beta_{z0})\right\}\Big] dH(u) + o_p(1), \end{aligned}\tag{D.12}$$

where the equality is due to $\widehat{X}_j = \widetilde{X}_j + o_p(1)$ and $\widehat{H}(u) = H(u) + o_p(1)$ (e.g., Wang 1991).
Next, we examine the integrand of (D.12) with $\widetilde{x}_j$ and z_j replaced by the corresponding random variables. Applying the Taylor series expansion to $\exp\left\{-\widehat{\Lambda}_0(u)\exp(\widetilde{X}_j^\top\beta_{x0} + Z_j^\top\beta_{z0})\right\}$ with respect to $\Lambda_0(\cdot)$, we obtain that

$$\begin{aligned}&\exp\left\{-\widehat{\Lambda}_0(u)\exp(\widetilde{X}_j^\top\beta_{x0} + Z_j^\top\beta_{z0})\right\} - \exp\left\{-\Lambda_0(u)\exp(\widetilde{X}_j^\top\beta_{x0} + Z_j^\top\beta_{z0})\right\} \\ =& -\exp\left\{-\Lambda_0(u)\exp(\widetilde{X}_j^\top\beta_{x0} + Z_j^\top\beta_{z0})\right\}\left\{\widehat{\Lambda}_0(u) - \Lambda_0(u)\right\} \\ &\times\exp\left(\widetilde{X}_j^\top\beta_{x0} + Z_j^\top\beta_{z0}\right) + o_p\left(\frac{1}{\sqrt{n}}\right). \end{aligned}\tag{D.13}$$

By the similar derivation in (D.8), we have

$$\begin{aligned}&\left\{\widehat{\Lambda}_0(\tau) - \Lambda_0(\tau)\right\} \\ =& \frac{1}{n}\sum_{i=1}^n\int_0^{\tau}\left\{\frac{dN_i(u)}{\mathcal{S}^{(0)}(u,\beta_0)} - \frac{d\mathcal{N}(u)\exp\left(w_i^{*\top}\beta_{x0} + z_i^\top\beta_{z0}\right)I(A_i \le u \le Y_i)}{\left\{\mathcal{S}^{(0)}(u,\beta_0)\right\}^2}\right\} m(\beta_{x0}) \\ &+ o_p(1). \end{aligned}\tag{D.14}$$

Combining (D.13) and (D.14) with (D.12) gives

$$\sqrt{n}\,(\widehat{\mu}-\mu) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_i(\beta_0|\widetilde{x}, z) + o_p\,(1)\,, \tag{D.15}$$

where $\psi_i(\beta_0|\widetilde{x}, z)$ is given by (25), and $S(\xi|\widetilde{x}, z) = \exp\left\{-\Lambda_0(\xi)\exp(\widetilde{x}^\top\beta_{x0} + z^\top\beta_{z0})\right\}$.
Therefore, combining (D.15) and (D.11) yields

$$\begin{aligned}U_2 =& \frac{1}{\sqrt{n}}\sum_{i=1}^n\left[\int_{-\infty}^{\infty}\int_0^{\tau}\left\{\frac{1}{\mu}\frac{\partial}{\partial\beta_0}\psi_i(\beta_0|\widetilde{x}, z) - \frac{\partial\mu}{\partial\beta_0}\frac{1}{\mu^2}\psi_i(\beta_0|\widetilde{x}, z)\right\}\right] dG(a,\widehat{v}) + o_p(1) \\ \triangleq& \frac{1}{\sqrt{n}}\sum_{i=1}^n \Psi_2\left(w_i, \widetilde{x}_i, z_i, y_i, a_i\right) + o_p(1). \end{aligned}\tag{D.16}$$

Finally, combining (C.6), (D.5), (D.10) and (D.16) gives that

$$\frac{1}{\sqrt{n}}U(\beta_0) = \frac{1}{\sqrt{n}}\left\{U_P(\beta_0) + U_M(\beta_0)\right\} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \Psi\left(W_i, \widetilde{X}_i, Z_i, Y_i, A_i\right) + o_p(1),$$

where

$$\begin{aligned}\Psi\left(W_i, \widetilde{X}_i, Z_i, Y_i, A_i\right) =& \Phi\left(W_i, Z_i, Y_i, A_i\right) - \Psi_{M1}\left(W_i, \widetilde{X}_i, Z_i, Y_i, A_i\right) \\ &+ \Psi_2\left(W_i, \widetilde{X}_i, Z_i, Y_i, A_i\right) - U_{M,i}, \end{aligned}$$

shown as in (26).

By the Central Limit Theorem, we have that as $n \to \infty$,

$$\frac{1}{\sqrt{n}} U(\beta_0) \xrightarrow{d} N(0, \mathcal{B}), \tag{D.17}$$

where $\mathcal{B} = E\left\{\Psi^{\otimes 2}\left(W_i, \widetilde{X}_i, Z_i, A_i, Y_i\right)\right\}$. Therefore, using (D.4) and (D.17) and applying the Slutsky's Theorem yields that as $n \to \infty$,

$$\sqrt{n}(\widetilde{\beta} - \beta) \xrightarrow{d} N\left(0, \mathcal{A}^{-1}\mathcal{B}\mathcal{A}^{-1}\right).$$

□

D.2 Proof of Theorem 4

By Theorems 2 and 3 in the main paper, to prove that $\widetilde{\beta}$ is more efficient than $\widehat{\beta}$, it suffices to show that for any non-zero column vector t,

$$t^\top\left(\mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1} - \mathcal{A}^{-1}\mathcal{B}\mathcal{A}^{-1}\right)t > 0.$$

Condition (C7) gives that $\mathcal{A}_M = \mathcal{A} - \mathcal{I}$ is positive definite, which yields that $\mathcal{I}^{-1} - \mathcal{A}^{-1}$ is positive definite (Chandra and Chatterjee 2000, p.15).

Let

$$\begin{aligned} a = & \int_0^\tau \left\{\widetilde{V}_i - \frac{\mathcal{S}^{(1)}(u,\beta_0)}{\mathcal{S}^{(0)}(u,\beta_0)} + \begin{pmatrix} \Sigma_{\epsilon;0}\beta_{x0} \\ 0_q \end{pmatrix}\right\} dN_i(u) \\ & - \int_0^\tau \frac{\exp\left(\widetilde{V}_i^\top\beta_0\right) I(A_i \le u \le Y_i)}{\mathcal{S}^{(0)}(u,\beta_0)} \left\{\widetilde{V}_i - \frac{\mathcal{S}^{(1)}(u,\beta_0)}{\mathcal{S}^{(0)}(u,\beta_0)}\right\} dE\left\{N_i(u)\right\} \end{aligned}$$

and

$$\begin{aligned} b = & \left\{\int_{-\infty}^\infty \int_0^\tau \frac{\partial}{\partial\beta_0}\left[\frac{dN_i(u)}{\mathcal{S}^{(0)}(u,\beta_0)} - \frac{d\mathcal{N}(u)\exp\left(\widetilde{V}_i^\top\beta_0\right) I(A_i \le u \le Y_i)}{\left\{\mathcal{S}^{(0)}(u,\beta_0)\right\}^2}\right] m(\beta_{x0}) \right. \\ & \left. \times \exp\left(\widehat{v}^\top\beta_0\right) I(u \le a \le \tau) dG(a,\widehat{v})\right\} \\ & - \left[\int_{-\infty}^\infty \int_0^\tau \left\{\frac{1}{\mu}\frac{\partial}{\partial\beta_0}\psi_i(\beta_0|\widetilde{x},z) - \frac{\partial\mu}{\partial\beta_0}\frac{1}{\mu^2}\psi_i(\beta_0|\widetilde{x},z)\right\} dG(a,\widehat{v})\right] \\ & + \frac{\partial}{\partial\beta_0}\Lambda_0(A_i)\exp\left(\widehat{V}_i^\top\beta_0\right) + \frac{1}{\mu_i}\frac{\partial}{\partial\beta_0}\mu_i. \end{aligned}$$

Define $\mathcal{J} = E\left(aa^\top\right)$, $\mathcal{B} = E\left\{(a-b)(a-b)^\top\right\}$ and $\mathcal{B}_M = E\left(ab^\top + ba^\top - bb^\top\right)$. Then it is immediate that $\mathcal{B} = \mathcal{J} - \mathcal{B}_M$. Representing the asymptotic covariance matrix related to $U_P(\cdot)$ in (C.2), $\mathcal{J}$ is a positive definite matrix. Since $\mathcal{B}$ is the asymptotic covariance matrix related to the function $U(\cdot)$ in (D.1), $\mathcal{B}$ is a positive definite matrix. Hence, for any vector t, $t^\top\mathcal{B}t = t^\top(\mathcal{J} - \mathcal{B}_M)t > 0$, or equivalently, $t^\top\mathcal{J}t - t^\top\mathcal{B}_M t > 0$.

Finally, for any $t \neq 0$,

$$\begin{aligned} & t^\top\left\{\mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1} - \mathcal{A}^{-1}\mathcal{B}\mathcal{A}^{-1}\right\}t \\ = & t^\top\left\{\mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1} - \mathcal{A}^{-1}\mathcal{J}\mathcal{A}^{-1} + \mathcal{A}^{-1}\mathcal{B}_M\mathcal{A}^{-1}\right\}t \\ \geqslant & t^\top\left\{\left(\mathcal{I}^{-1} - \mathcal{A}^{-1}\right)^\top\mathcal{J}\left(\mathcal{I}^{-1} - \mathcal{A}^{-1}\right) + \mathcal{A}^{-1}\mathcal{B}_M\mathcal{A}^{-1}\right\}t \\ > & 0, \end{aligned}$$

where the last inequality comes from the fact that $\left(\mathcal{I}^{-1} - \mathcal{A}^{-1}\right)^\top\mathcal{J}\left(\mathcal{I}^{-1} - \mathcal{A}^{-1}\right)$ is a positive definite matrix. Hence, the conclusion follows. □

E Proofs of the Theorems in Section 5

The proofs in this appendix are more complicated than those derivations of Appendices C and D, because the parameters in the measurement error model have to be estimated from validation data and the induced variability must be incorporated when establishing asymptotic results.

E.1 Proof of Theorem 5

Proof of Theorem 5 (1):

Since $\widehat{\gamma}$ is a consistent estimator of γ_0, consistency of $\widehat{\beta}_{val}$ can be established following the proof of Theorem 2 (1) in Appendix C.

Proof of Theorem 5 (2):

To derive the asymptotic distribution of $\widehat{\beta}_{val}$, we begin with examining the estimator $\widehat{\gamma}$. By the Taylor series expansion of (31) with respect to γ_0, we obtain that

$$\sqrt{n}\,(\widehat{\gamma}-\gamma_0) = \frac{\sqrt{n}}{m}\sum_{i\in\mathcal{V}}(W_i - X_i) + o_p(1). \tag{E.1}$$

Next, we define

$$\widehat{\ell}^*_{P,val} = \sum_{i=1}^{n}\int_0^\tau \left[\widetilde{v}_i^{*\top}\beta + \frac{1}{2}\beta_x^\top\widehat{\Sigma}_\epsilon\beta_x - \log\left\{\sum_{j=1}^{n}\exp(\widetilde{v}_j^{*\top}\beta)I(a_j\le u\le y_j)\right\}\right]dN_i(u) \tag{E.2}$$

and $\widetilde{v}_i^* = \left((w_i-\widehat{\gamma})^\top, z_i^\top\right)^\top$. Similar to Lemma B.1, we can show that $\widehat{\ell}^*_{C,val} = \widehat{\ell}^*_{P,val} + \widehat{\ell}^*_{R,val}$ and that $\widehat{\ell}^*_{R,val}$ is ancillary, and thus inference about β based on $\widehat{\ell}^*_{C,val}$ is equivalent to that based on $\widehat{\ell}^*_{P,val}$.

Let $U_{P,val}(\beta) = \frac{\partial\widehat{\ell}^*_{P,val}}{\partial\beta}$. Since $\widehat{\beta}_{val}$ solves $U_{P,val}(\beta)=0$, then by the Taylor series expansion of $U_{P,val}(\beta)$ around β_0, we have that

$$\sqrt{n}\left(\widehat{\beta}_{val}-\beta_0\right) = -\left\{\frac{1}{n}\frac{\partial}{\partial\beta_0}U_{P,val}(\beta_0)\right\}^{-1}\times\frac{1}{\sqrt{n}}U_{P,val}(\beta_0)+o_p(1). \tag{E.3}$$

Analogous to the derivation of (C.2) in Appendix C.2, we proceed with the following two steps by examining $\frac{\partial}{\partial\beta_0}U_{P,val}(\beta_0)$ and $U_{P,val}(\beta_0)$, respectively. The main difference here is the involvement of estimators in measurement error model.

Step 1: By the consistency of $\widehat{\gamma}$, we have $\widehat{\gamma}=\gamma_0+o_p(1)$. By the similar derivations of (C.5), we have that as $n\to\infty$,

$$-\frac{1}{n}\frac{\partial}{\partial\beta_0}U_{P,val}(\beta_0)\xrightarrow{p}\mathcal{I}, \tag{E.4}$$

where $\mathcal{I}$ is determined by (16).

Step 2: By (E.2) and that $U_{P,val}(\beta)=\frac{\partial\widehat{\ell}^*_{P,val}}{\partial\beta}$, we have that

$$\frac{1}{\sqrt{n}}U_{P,val}(\beta_0) = \frac{1}{\sqrt{n}}\sum_{i\in\mathcal{M}}\int_0^\tau\left\{\begin{pmatrix} w_i-\widehat{\gamma}\\ z_i\end{pmatrix}+\widehat{\Sigma}_\epsilon\beta_{x0}-\frac{\widehat{S}^{(1)}(u;\beta_0)}{\widehat{S}^{(0)}(u;\beta_0)}\right\}dN_i(u), \tag{E.5}$$

where

$$\widehat{S}^{(k)}(u;\beta_0)=\frac{1}{n}\sum_{i\in\mathcal{M}}\begin{pmatrix}w_i-\widehat{\gamma}\\ z_i\end{pmatrix}^{\otimes k}\exp\left\{\begin{pmatrix}w_i-\widehat{\gamma}\\ z_i\end{pmatrix}^{\top}\begin{pmatrix}\beta_{x0}\\ \beta_{z0}\end{pmatrix}I\left(a_i\le u\le y_i\right)\right\}$$

for $k=0,1$.

Since (E.5) involves the estimators $\widehat{\gamma}$ and $\widehat{\Sigma}_\epsilon$, so by adding and subtracting γ_0 and $\Sigma_{\epsilon;0}$, (E.5) can be re-written as

$$\begin{aligned}&\frac{1}{\sqrt{n}}U_{P,val}\left(\beta_0\right)\\ &=\frac{1}{\sqrt{n}}\sum_{i\in\mathcal{M}}\int_0^\tau\left[-\begin{pmatrix}\widehat{\gamma}-\gamma_0\\ 0\end{pmatrix}+\left(\widehat{\Sigma}_\epsilon-\Sigma_{\epsilon;0}\right)\beta_{x0}-\left\{\frac{\widehat{S}^{(1)}\left(u;\beta_0\right)}{\widehat{S}^{(0)}\left(u;\beta_0\right)}-\frac{S^{(1)}\left(u;\beta_0\right)}{S^{(0)}\left(u;\beta_0\right)}\right\}\right]dN_i(u)\\ &+\frac{\sqrt{1+\rho}}{\sqrt{m+n}}\sum_{i\in\mathcal{M}\cup\mathcal{V}}\zeta_i\Phi\left(w_i,z_i,y_i,a_i\right)+o_p(1),\end{aligned}\tag{E.6}$$

where $\Phi\left(w_i,z_i,y_i,a_i\right)$ is given by (C.7), and ζ_i is a indicator that $\zeta_i=1$ if $i\in\mathcal{M}$ and $\zeta_i=0$ if $i\in\mathcal{V}$.

We now examine the integral in (E.6), which is done by evaluating each term in the integrand separately.

First,

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{i\in\mathcal{M}}\int_0^\tau-\begin{pmatrix}\widehat{\gamma}-\gamma_0\\ 0\end{pmatrix}dN_i(u)\\ &=-\frac{\sqrt{1+\rho}}{\rho}\frac{1}{\sqrt{n+m}}\sum_{i\in\mathcal{M}\cup\mathcal{V}}(1-\zeta_i)E\left\{N_i(\tau)\right\}\begin{pmatrix}W_i-X_i\\ 0\end{pmatrix}+o_p(1).\end{aligned}\tag{E.7}$$

Using $\sqrt{n}\left(\widehat{\Sigma}_\epsilon-\Sigma_{\epsilon;0}\right)\beta_{x0}\left[\frac{1}{n}\sum\limits_{i\in\mathcal{M}}N_i(\tau)-E\left\{N_i(\tau)\right\}\right]=o_p(1)$ and the consistency of $\widehat{\gamma}$, we can derive

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{i\in\mathcal{M}}\int_0^\tau\left(\widehat{\Sigma}_\epsilon-\Sigma_{\epsilon;0}\right)\beta_{x0}dN_i(u)\\ &=\frac{mE\left\{N_i(\tau)\right\}}{m-1}\frac{\sqrt{1+\rho}}{\rho}\frac{1}{\sqrt{n+m}}\sum_{i\in\mathcal{M}\cup\mathcal{V}}(1-\zeta_i)\left\{\epsilon_i\epsilon_i^\top-(m-1)\Sigma_{\epsilon;0}\right\}\beta_{x0}\\ &+o_p(1).\end{aligned}\tag{E.8}$$

Finally, using $\sqrt{n}\left\{\frac{\widehat{S}^{(1)}(u;\beta_0)}{\widehat{S}^{(0)}(u;\beta_0)}-\frac{S^{(1)}(u;\beta_0)}{S^{(0)}(u;\beta_0)}\right\}\left[\frac{1}{n}\sum\limits_{i\in\mathcal{M}}N_i(\tau)-E\left\{N_i(\tau)\right\}\right]=o_p(1)$, the consistency of $\widehat{\gamma}$, and the Mean Value Theorem to $S^{(1)}(u;\beta_0)$ with respect to γ_0, we can show that

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{i\in\mathcal{M}}\int_0^\tau\left\{\frac{\widehat{S}^{(1)}\left(u;\beta_0\right)}{\widehat{S}^{(0)}\left(u;\beta_0\right)}-\frac{S^{(1)}\left(u;\beta_0\right)}{S^{(0)}\left(u;\beta_0\right)}\right\}dN_i(u)\\ &=\frac{E\left\{N_i(\tau)\right\}}{\mathcal{S}^{(0)}\left(u;\beta_0\right)}\frac{\partial\mathcal{S}^{(1)}\left(u;\beta_0\right)}{\partial\gamma_0}\frac{\sqrt{1+\rho}}{\rho}\frac{1}{\sqrt{n+m}}\sum_{i\in\mathcal{M}\cup\mathcal{V}}\left(1-\zeta_i\right)\left(W_i-X_i\right)+o_p(1).\end{aligned}\tag{E.9}$$

As a consequence, we combine (E.6), (E.7), (E.8), and (E.9) and obtain that

$$\frac{1}{\sqrt{n}}U_{P,val}(\beta_0)=\frac{1}{\sqrt{m+n}}\sum_{i\in\mathcal{M}\cup\mathcal{V}}\mathcal{B}_{val1,i}+o_p(1),\tag{E.10}$$

where $\mathcal{B}_{val1,i}$ is given by (39). By the Central Limit Theorem, we conclude that as $n \to \infty$,

$$\frac{1}{\sqrt{n}} U_{P,val}(\beta_0) \xrightarrow{d} N\left(0, \mathcal{J}_{val}\right), \tag{E.11}$$

where $\mathcal{J}_{val} = E\left\{\left(\mathcal{B}_{val1,i}\right)^{\otimes 2}\right\}$.

Finally, combining (E.3), (E.4) and (E.11) and applying the Slutsky's Theorem, we have that as $n \to \infty$,

$$\sqrt{n}\left(\widehat{\beta}_{val} - \beta_0\right) \xrightarrow{d} N\left(0, \mathcal{I}^{-1}\mathcal{J}_{val}\mathcal{I}^{-1}\right).$$

□

E.2 Proof of Theorem 6

Proof of Theorem 6 (1):
This can be done by following the proof of Theorem 3 in Appendix D.
Proof of Theorem 6 (2) :

Let $U_{P,val}(\beta) = \frac{\partial \widehat{\ell}^*_{P,val}}{\partial \beta}$, $U_{M,val}(\beta) = \frac{\partial \widehat{\ell}^*_{M,val}}{\partial \beta}$, and $U_{val}(\beta) = U_{P,val}(\beta) + U_{M,val}(\beta)$. Since $\widetilde{\beta}_{val}$ solves $U_{val}(\beta) = 0$, then by the Taylor series expansion of $U_{val}(\beta)$ around β_0, we have

$$\sqrt{n}\left(\widetilde{\beta}_{val} - \beta_0\right) = -\left\{\frac{1}{n}\frac{\partial}{\partial \beta_0} U_{val}(\beta_0)\right\}^{-1} \times \frac{1}{\sqrt{n}} U_{val}(\beta_0) + o_p(1). \tag{E.12}$$

Analogous to the proof of Theorem 5 (2) in Appendix E.1, we now examine the asymptotic behaviours of $\frac{\partial}{\partial \beta_0} U_{val}(\beta_0)$ and $U_{val}(\beta_0)$.

First, by the derivations of (D.4) and (E.4), we have that as $n \to \infty$,

$$\frac{-1}{n}\frac{\partial}{\partial \beta_0} U_{val}(\beta_0) \xrightarrow{p} \mathcal{A}, \tag{E.13}$$

where $\mathcal{A}$ is given by Theorem 3. So the remaining part is to examine $U_{val}(\beta_0)$. The additional difficulty here is to deal with $U_{M,val}(\beta_0)$ that is involved in $U_{val}(\beta_0)$ since the derivation of $U_{P,val}(\beta)$ is done in Appendix E.1.

To study $U_{M,val}(\beta)$, similar to the idea of (D.5), we define

$$\begin{aligned}
\widetilde{U}_{M,val}(\beta_0) &= -\sum_{i \in \mathcal{M}} \frac{\partial}{\partial \beta_0} \Lambda_0(A_i) \exp\left(\widehat{V}_i^\top \beta_{x0}\right) - \sum_{i \in \mathcal{M}} \frac{1}{\mu_{val,i}} \frac{\partial \mu_{val,i}}{\partial \beta_0} \\
&= -\sum_{i \in \mathcal{M}} \widetilde{U}_{M,val,i}
\end{aligned} \tag{E.14}$$

and write

$$\frac{1}{\sqrt{n}} \sum_{i \in \mathcal{M}} \left\{U_{M,val}(\beta_0) - \widetilde{U}_{M,val}(\beta_0)\right\} = U_{val,1}(\beta_0) + U_{val,2}(\beta_0), \tag{E.15}$$

where

$$\begin{aligned}
U_{val,1}(\beta_0) = &\frac{-1}{\sqrt{n}} \sum_{i \in \mathcal{M}} \frac{\partial}{\partial \beta_0} \widehat{\Lambda}_0(A_i) \exp\left\{\widehat{X}_i^\top \beta_{x0} + Z_i^\top \beta_{z0}\right\} \\
&+ \frac{1}{\sqrt{n}} \sum_{i \in \mathcal{M}} \frac{\partial}{\partial \beta_0} \Lambda_0(A_i) \exp\left\{\widetilde{X}_i^\top \beta_{x0} + Z_i^\top \beta_{z0}\right\},
\end{aligned}$$

$$U_{val,2}(\beta_0) = -\frac{1}{\sqrt{n}} \sum_{i\in\mathcal{M}} \left\{ \frac{1}{\widehat{\mu}_{val,i}} \frac{\partial \widehat{\mu}_{val,i}}{\partial \beta_0} - \frac{1}{\mu_{val,i}} \frac{\partial \mu_{val,i}}{\partial \beta_0} \right\},$$

$\widehat{\mu}_{val,i} = \widehat{\mu}(\widehat{x}_{val,i}, z_i) = \int_0^\tau \exp\left\{-\widehat{\Lambda}_0(u)\exp\left(\widehat{x}_{val,i}^\top \beta_{x0} + z_i^\top \beta_{z0}\right)\right\} d\widehat{H}_{val}(u)$, and $\mu_{val,i} = \mu(\widetilde{x}_i, z_i)$. Now we carry out the following steps to examine each term of (E.15).

Step 1: We first analyze $U_{val,1}(\beta_0)$.

By the similar derivations of (D.5) and (D.6), we express

$$\begin{aligned}
&U_{val,1}(\beta_0)\\
&= -\left[\frac{1}{\sqrt{n}} \sum_{i\in\mathcal{M}} \frac{\partial}{\partial\beta_0} \int_0^\tau \frac{d\bar{N}(u)}{\widehat{S}^{(0)}(u;\beta_0)} \exp\left\{\widehat{X}_i^\top \beta_{x0} + Z_i^\top \beta_{z0}\right\} I(u \le a_i \le \tau)\right.\\
&\quad \left. - \frac{1}{\sqrt{n}} \sum_{i\in\mathcal{M}} \frac{\partial}{\partial\beta_0} \int_0^\tau \frac{d\bar{N}(u)}{S^{(0)}(u;\beta_0)} \exp\left\{\widetilde{X}_i^\top \beta_{x0} + Z_i^\top \beta_{z0}\right\} I(u \le a_i \le \tau)\right] m(\beta_{x0})\\
&\quad -\left[\frac{1}{\sqrt{n}} \sum_{i\in\mathcal{M}} \frac{\partial}{\partial\beta_0} \int_0^\tau \frac{d\bar{N}(u)}{S^{(0)}(u;\beta_0)} \exp\left\{\widetilde{X}_i^\top \beta_{x0} + Z_i^\top \beta_{z0}\right\} I(u \le a_i \le \tau)\right.\\
&\quad \left. - \frac{1}{\sqrt{n}} \sum_{i\in\mathcal{M}} \frac{\partial}{\partial\beta_0} \int_0^\tau \frac{d\mathcal{N}(u)}{\mathcal{S}^{(0)}(u;\beta_0)} \exp\left\{\widetilde{X}_i^\top \beta_{x0} + Z_i^\top \beta_{z0}\right\} I(u \le a_i \le \tau)\right] m(\beta_{x0})\\
&\triangleq T_1 + T_2.
\end{aligned} \tag{E.16}$$

For T_1 in (E.16), using the Mean Value Theorem on $S^{(0)}(u;\beta)$ with respect to γ, (E.1), and the Law of Large Numbers, we can express

$$\begin{aligned}
T_1 &= E\left[\frac{\partial}{\partial\beta_0} \int_0^\tau d\mathcal{N}(u) \frac{1}{\left\{\mathcal{S}^{(0)}(u;\beta_0)\right\}^2} \frac{\partial \mathcal{S}^{(0)}(u;\beta_0)}{\partial\gamma_0}\right.\\
&\quad \left. \times m(\beta_{x0}) \exp\left\{\widetilde{X}_i^\top \beta_{x0} + Z_i^\top \beta_{z0}\right\} I(u \le A_i \le \tau)\right]\\
&\quad \times \frac{\sqrt{n}}{m} \sum_{i\in\mathcal{V}} (W_i - X_i) + o_p(1)\\
&\triangleq \mathcal{E}_{val,1} \frac{1}{\sqrt{m+n}} \frac{\sqrt{1+\rho}}{\rho} \sum_{i\in\mathcal{M}\cup\mathcal{V}} (1-\zeta_i)(W_i - X_i) + o_p(1).
\end{aligned} \tag{E.17}$$

T_2 in (E.16) is exactly the form in (D.8). Therefore, we directly have

$$\begin{aligned}
&T_2\\
&= -\frac{1}{\sqrt{n}} \sum_{i\in\mathcal{M}} \frac{\partial}{\partial\beta_0} \left[\int_{-\infty}^{\infty} \int_0^\tau \left\{\frac{dN_i(u)}{\mathcal{S}^{(0)}(u,\beta_0)} - \frac{d\mathcal{N}(u)\exp\left(\widetilde{V}_i^\top \beta_0\right) I(A_i \le u \le Y_i)}{\left\{\mathcal{S}^{(0)}(u,\beta_0)\right\}^2}\right\} m(\beta_{x0})\right.\\
&\quad \left. \times \exp\left(\widetilde{v}^\top \beta_0\right) I(u \le a \le \tau)\right] dG(a,\widetilde{v}) + o_p(1)\\
&\triangleq -\frac{1}{\sqrt{n}} \sum_{i\in\mathcal{M}} \Psi_{M1}\left(W_i, \widetilde{X}_i, Z_i, Y_i, A_i\right) + o_p(1)\\
&= \frac{-1}{\sqrt{n+m}} \sqrt{1+\rho} \sum_{i\in\mathcal{M}\cup\mathcal{V}} \zeta_i \Psi_{M1}\left(W_i, \widetilde{X}_i, Z_i, Y_i, A_i\right) + o_p(1).
\end{aligned} \tag{E.18}$$

Therefore, combining (E.16), (E.17), and (E.18) gives

$$U_{val,1}(\beta_0) = \frac{1}{\sqrt{n+m}} \sum_{i\in\mathcal{M}\cup\mathcal{V}} \mathcal{B}_{val2,i} + o_p(1), \tag{E.19}$$

where $\mathcal{B}_{val2,i}$ is given by (40).
Step 2: We next examine $U_{val,2}(\beta_0)$.
Similar to the derivations in (D.11), we have

$$\begin{aligned} U_{val,2}(\beta_0) &= \frac{1}{\sqrt{n}} \sum_{i\in\mathcal{M}} \left(\frac{1}{\widehat{\mu}_{val,i}} \frac{\partial \widehat{\mu}_{val,i}}{\partial \beta_0} - \frac{1}{\mu_{val,i}} \frac{\partial \mu_{val,i}}{\partial \beta_0} \right) \\ &= \sqrt{n} \int_{-\infty}^{\infty} \int_0^{\tau} \left(\frac{1}{\widehat{\mu}_{val}} \frac{\partial \widehat{\mu}_{val}}{\partial \beta_0} - \frac{1}{\mu_{val}} \frac{\partial \mu_{val}}{\partial \beta_0} \right) dG(a, \widehat{v}) + o_p(1). \end{aligned} \tag{E.20}$$

Similar to the derivations of Theorem 3 (2) in Appendix D, we first derive $\sqrt{n}\,(\widehat{\mu}_{val} - \mu_{val})$, where $\widehat{\mu}_{val} = \widehat{\mu}(\widehat{x}_{val}, z)$ and $\mu_{val} = \mu(\widetilde{x}, z)$.
Note that

$$\begin{aligned} \sqrt{n}(\widehat{\mu}_{val} - \mu_{val}) = \sqrt{n} \int_0^{\tau} \Big[&\exp\Big\{ -\widehat{\Lambda}_0(u) \exp(\widetilde{x}^{\top}\beta_{x0} + z^{\top}\beta_{z0}) \Big\} \\ &- \exp\Big\{ -\Lambda_0(u) \exp(\widetilde{x}^{\top}\beta_{x0} + z^{\top}\beta_{z0}) \Big\} \Big] dH(u) + o_p(1). \end{aligned} \tag{E.21}$$

On the other hand, the difference $\widehat{\Lambda}_0(\tau) - \Lambda_0(\tau)$ can be expressed as

$$\begin{aligned} &\widehat{\Lambda}_0(\tau) - \Lambda_0(\tau) \\ &= \int_0^{\tau} \left\{ \frac{d\bar{N}(t)}{\widehat{S}^{(0)}(t;\beta_0)} - \frac{d\bar{N}(t)}{S^{(0)}(t;\beta_0)} \right\} m(\beta_{x0}) + \int_0^{\tau} \left\{ \frac{d\bar{N}(t)}{S^{(0)}(t;\beta_0)} - \frac{d\mathcal{N}(t)}{\mathcal{S}^{(0)}(t;\beta_0)} \right\} m(\beta_{x0}) \\ &\triangleq A + B. \end{aligned} \tag{E.22}$$

By the Mean Value Theorem on $S^{(0)}(t;\beta)$ with respect to γ in A, we have

$$\begin{aligned} A &= \int_0^{\tau} \frac{-d\mathcal{N}(t)}{\left\{\mathcal{S}^{(0)}(t;\beta_0)\right\}^2} \left\{ \frac{\partial \mathcal{S}^{(0)}(t;\beta_0)}{\partial \gamma_0} \right\} m(\beta_{x0})\,(\widehat{\gamma} - \gamma_0) + o_p(1) \\ &\triangleq \mathbf{A}\,(\widehat{\gamma} - \gamma_0) + o_p(1). \end{aligned} \tag{E.23}$$

Moreover, B in (E.22) is equal to (D.14), so we obtain that

$$\begin{aligned} B &= \frac{1}{n} \sum_{i\in\mathcal{M}} \int_0^{\tau} \left[\frac{dN_i(t)}{\mathcal{S}^{(0)}(t;\beta_0)} \right. \\ &\quad \left. - \frac{d\mathcal{N}(t) \exp\left\{ (W_i - \gamma_0)^{\top}\beta_{x0} + Z_i^{\top}\beta_{z0} \right\} I(A_i \le u \le Y_i)}{\left\{\mathcal{S}^{(0)}(t;\beta_0)\right\}^2} \right] m(\beta_{x0}) \\ &\triangleq \frac{1}{n} \sum_{i\in\mathcal{M}} \mathbf{B}_i + o_p(1). \end{aligned} \tag{E.24}$$

Therefore, combining (E.23) and (E.24) with (E.22) yields

$$\widehat{\Lambda}_0(\tau) - \Lambda_0(\tau) = \left\{ \mathbf{A}\,(\widehat{\gamma} - \gamma_0) + \frac{1}{n} \sum_{i\in\mathcal{M}} \mathbf{B}_i \right\} + o_p(1). \tag{E.25}$$

Applying the Taylor series expansion to $\exp\left\{ -\widehat{\Lambda}_0(u) \exp\left(\widetilde{X}^{\top}\beta_{x0} + Z^{\top}\beta_{z0} \right) \right\}$ with respect to $\Lambda_0(\cdot)$ gives

$$\begin{aligned} &\exp\left\{ -\widehat{\Lambda}_0(u) \exp\left(\widetilde{X}^{\top}\beta_{x0} + Z^{\top}\beta_{z0} \right) \right\} - \exp\left\{ -\Lambda_0(u) \exp\left(\widetilde{X}^{\top}\beta_{x0} + Z^{\top}\beta_{z0} \right) \right\} \\ &= -\exp\left\{ -\Lambda_0(u) \exp\left(\widetilde{X}^{\top}\beta_{x0} + Z^{\top}\beta_{z0} \right) \right\} \left\{ \widehat{\Lambda}_0(u) - \Lambda_0(u) \right\} \\ &\quad \times \exp\left(\widetilde{X}^{\top}\beta_{x0} + Z^{\top}\beta_{z0} \right) + o_p\left(\frac{1}{\sqrt{n}} \right). \end{aligned} \tag{E.26}$$

Therefore, combining (E.25) and (E.26) with (E.21) yields

$$\begin{aligned}\sqrt{n}\,(\widehat{\mu}_{val}-\mu_{val}) = &\frac{1}{\sqrt{n+m}}\sum_{i\in\mathcal{M}\cup\mathcal{V}}\left[\frac{\sqrt{1+\rho}}{\rho}(W_i-X_i)\int_0^{\tau}\{S(\nu|\widetilde{x},z)\mathbf{A}\right.\\ &\left.\times\exp\left(\widetilde{x}^{\top}\beta_{x0}+z^{\top}\beta_{z0}\right)dH(\nu)\right\}+\sqrt{1+\rho}\psi_i(\beta_0|\widetilde{x},z)\Big]+o_p(1)\\ \triangleq &\frac{1}{\sqrt{n+m}}\sum_{i\in\mathcal{M}\cup\mathcal{V}}\varphi_{val,i}+o_p(1). \end{aligned}\tag{E.27}$$

Similar to the derivations for (D.16), combining (E.20) and (E.27) gives

$$\begin{aligned}&U_{val,2}(\beta_0)\\ =&\frac{1}{\sqrt{n+m}}\sum_{i\in\mathcal{M}\cup\mathcal{V}}\left[\int_{-\infty}^{\infty}\int_0^{\tau}\left\{\frac{1}{\mu_{val}}\frac{\partial}{\partial\beta_0}\varphi_{val,i}-\frac{\partial\mu_{val}}{\partial\beta_0}\frac{1}{\mu_{val}^2}\varphi_{val,i}\right\}\right]dG(a,\widehat{v})+o_p(1)\\ \triangleq&\frac{1}{\sqrt{n+m}}\sum_{i\in\mathcal{M}\cup\mathcal{V}}\mathcal{B}_{val3,i}+o_p(1).\end{aligned}\tag{E.28}$$

To summarize, combining (E.10), (E.14), (E.15), (E.19) and (E.28) yields

$$\frac{1}{\sqrt{n}}U_{val}(\beta_0)=\frac{1}{\sqrt{n+m}}\sum_{i\in\mathcal{M}\cup\mathcal{V}}\left(\mathcal{B}_{val1,i}+\mathcal{B}_{val2,i}+\mathcal{B}_{val3,i}+\sqrt{1+\rho}\zeta_i\widetilde{U}_{M,val,i}\right)+o_p(1).\tag{E.29}$$

Finally, by the Central Limit Theorem, we have that as $n\to\infty$,

$$\frac{1}{\sqrt{n}}U_{val}(\beta_0)\xrightarrow{d}N\left(0,\mathcal{B}_{val}\right),\tag{E.30}$$

where $\mathcal{B}_{val}$ is given by (43). Therefore, combining (E.13) and (E.30) with (E.12) and applying the Slutsky's Theorem give that as $n\to\infty$,

$$\sqrt{n}\left(\widetilde{\beta}_{val}-\beta_0\right)\xrightarrow{d}N\left(0,\mathcal{A}^{-1}\mathcal{B}_{val}\mathcal{A}^{-1}\right).$$

□

E.3 Proof of Theorem 7

This is done by the similar derivations in Appendix D.2. □

F Additional Numerical Results

F.1 Performance of Proposed Estimators: α and Σ_ϵ are Known

In this section, we report the supplementary results for the simulation studies conducted in the manuscript which considers the following settings:

Setting 1: $n=200$, Scenario 1 with $\Sigma=\begin{pmatrix}4 & 0.5\\ 0.5 & 36\end{pmatrix}$, and $(\beta_{x0},\beta_{z0})^{\top}=(0.3,1)^{\top}$.
Setting 2: This is the same as Setting 1 except that n is increased to $n=1000$.
Setting 3: This is the same as Setting 1 except that β_{x0} is increased to be $\beta_{x0}=\log 3$.
Setting 4: This is the same as Setting 1 except that Σ is changed to be $\Sigma=\begin{pmatrix}1 & 0.6\\ 0.6 & 1\end{pmatrix}$.

Setting 5: Simultaneously increase the values in Setting 1 to be

$$n = 1000, \Sigma = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix}, \text{ and } \beta_{x0} = \log 3.$$

Setting 6: This is the same as Setting 1 except that the covariates are generated as in Scenario 2.

The numerical results for Settings 1, 5 and 6 are reported in the manuscript, and the results for Settings 2-4 are summarized here in Tables S.1-S.3, respectively. All those tables show similar patterns.

Furthermore, we conduct additional simulations under the measurement error model (6) with $\alpha = 10, 50$ and 100 based on Setting 1. The results are summarized in Tables S.4-S.6, which reveal the same patterns observed by the results in the manuscript.

F.2 Sensitivity Analyses on Simulated Data

In addition to the sensitivity analyses reported in Section 6.4 of the manuscript, here we report other sensitivity studies by setting $\Sigma_X = 10\%\widehat{\Sigma}_W$ and $\alpha = 0, 10, 50$, or 100 and letting R take a value in $[0.6, 0.9]$. The numerical results are summarized in Table S.7. We observe that the results are fairly similar to those in Table 1 of the manuscript

F.3 Assessment of Misspecification of Measurement Error Model

We now study the performance of our estimators when the measurement error model is misspecified. Specifically, we consider two scenarios. In Scenario A, the measurement error model (5) is used to generate data, but we use model (6) to fit the data; in Scenario B, we use (6) with $\alpha = 100$ as the measurement error model to generate data, but we use model (5) to fit the data. We report the average of biases, average of S.E. and mean squared errors (MSE) for estimators $\widehat{\beta}$ and $\widetilde{\beta}$, respectively, obtained from (14) and (23). The results are displayed in Table S.8 for Scenario A and Table S.9 for Scenario B.

As shown in Table S.8, under Scenario A finite sample biases are comparable to those reported in Table 1 of the main paper. In addition, S.E.s and MSEs are very close to those obtained from the situation where the fitting model is correctly used. These results are not surprising since the model we used to generate the data is nested in the model we used to fit the data. On the other hand, in Scenario B where the model used to fit data differs from the model for generating data, i.e., model misspecification is present, biased results are produced, which is evident from Table S.9. This simulation study also shows that with model misspecification considered here, the proposed method performs better than the corrected conditional likelihood approach.

F.4 Performance with Validation Data

In this subsection, we evaluate the performance with validation data in Section 6.3. We summarize the settings $\alpha = 0$, 10, and 50 in Tables S.10-S.12. The results uncover similar findings to those revealed in Section F.1 and demonstrate satisfactory finite sample performance of the proposed estimators $\widehat{\beta}_{val}$ and $\widetilde{\beta}_{val}$. The results also confirm that $\widetilde{\beta}_{val}$ is more efficient than $\widehat{\beta}_{val}$.

F.5 Simulation Study For the Length-Biased Sampling

To show the numerical performance of estimator $\widetilde{\beta}_{LB}$ in contrast to $\widetilde{\beta}$ which is obtained by (23), we conduct a simulation study using the setting in Section 6.1 with the distribution of truncation times taken as the uniform distribution UNIF$[0,1]$ and α is set as 0. In addition to $\widetilde{\beta}_{LB}$ from ($A.3$), we also report the performance of the naive estimator and $\widehat{\beta}_{LB}$ determined by ($A.2$). The results are reported in Table S.13.

Simulation results show that for different Σ_ϵ and censoring rates, our proposed methods yield satisfactory results, and $\widetilde{\beta}_{LB}$ is more efficient than $\widehat{\beta}_{LB}$. The naive estimator incurs considerable biases. The results in Table S.13 are comparable with those reported in Section 6.

Table S.1 Simulation results under Setting 2 and measurement error model (6) with $\alpha = 0$

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	SEE	SEM	MSE	CP (%)	Bias	SEE	SEM	MSE	CP (%)
0.01	0%	Naive	0.108	0.011	0.010	0.012	17.8	0.123	0.025	0.023	0.016	16.5
		Conditional $\left(\widehat{\beta}\right)$	0.012	0.018	0.017	0.001	95.1	0.009	0.033	0.032	0.001	95.4
		Full $\left(\widetilde{\beta}\right)$	0.004	0.016	0.015	0.001	95.1	0.003	0.028	0.026	0.002	95.3
	25 %	Naive	0.126	0.013	0.011	0.016	15.5	0.144	0.025	0.024	0.021	14.3
		Conditional $\left(\widehat{\beta}\right)$	0.012	0.027	0.026	0.001	95.2	0.015	0.044	0.042	0.004	95.0
		Full $\left(\widetilde{\beta}\right)$	0.010	0.025	0.024	0.001	95.0	0.009	0.036	0.035	0.002	95.4
	50%	Naive	0.134	0.019	0.019	0.018	15.1	0.147	0.027	0.025	0.019	14.8
		Conditional $\left(\widehat{\beta}\right)$	0.014	0.028	0.027	0.001	95.0	0.017	0.048	0.047	0.004	95.0
		Full $\left(\widetilde{\beta}\right)$	0.013	0.025	0.025	0.001	95.3	0.011	0.039	0.038	0.002	95.1
0.5	0%	Naive	0.111	0.013	0.013	0.016	9.8	0.123	0.025	0.024	0.016	10.7
		Conditional $\left(\widehat{\beta}\right)$	0.018	0.019	0.018	0.001	94.9	0.012	0.038	0.037	0.001	94.9
		Full $\left(\widetilde{\beta}\right)$	0.005	0.017	0.015	0.001	95.2	0.010	0.029	0.028	0.001	95.3
	25 %	Naive	0.139	0.021	0.020	0.020	9.5	0.147	0.038	0.035	0.018	12.7
		Conditional $\left(\widehat{\beta}\right)$	0.019	0.034	0.032	0.002	94.7	0.017	0.064	0.062	0.004	95.0
		Full $\left(\widetilde{\beta}\right)$	0.012	0.026	0.024	0.001	95.0	0.011	0.042	0.040	0.002	95.2
	50%	Naive	0.143	0.023	0.021	0.021	9.1	0.153	0.034	0.032	0.020	10.3
		Conditional $\left(\widehat{\beta}\right)$	0.023	0.037	0.036	0.002	94.7	0.021	0.069	0.068	0.005	94.8
		Full $\left(\widetilde{\beta}\right)$	0.014	0.026	0.025	0.001	94.9	0.015	0.043	0.043	0.002	95.1
0.75	0%	Naive	0.124	0.013	0.011	0.012	9.6	0.127	0.029	0.027	0.017	12.1
		Conditional $\left(\widehat{\beta}\right)$	0.028	0.020	0.019	0.001	94.5	0.021	0.040	0.039	0.001	94.7
		Full $\left(\widetilde{\beta}\right)$	0.013	0.020	0.019	0.001	95.1	0.017	0.030	0.030	0.001	95.3
	25 %	Naive	0.140	0.024	0.023	0.020	9.3	0.152	0.040	0.038	0.022	11.5
		Conditional $\left(\widehat{\beta}\right)$	0.031	0.035	0.034	0.002	94.1	0.023	0.070	0.070	0.006	94.8
		Full $\left(\widetilde{\beta}\right)$	0.015	0.027	0.027	0.001	94.8	0.018	0.042	0.041	0.002	95.1
	50%	Naive	0.153	0.025	0.025	0.024	8.7	0.157	0.035	0.035	0.020	10.8
		Conditional $\left(\widehat{\beta}\right)$	0.033	0.039	0.039	0.002	94.0	0.027	0.072	0.070	0.006	94.4
		Full $\left(\widetilde{\beta}\right)$	0.018	0.030	0.029	0.001	94.9	0.020	0.048	0.046	0.002	94.8

Table S.2 Simulation results under Setting 3 and measurement error model (6) with $\alpha = 0$

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	SEE	SEM	MSE	CP (%)	Bias	SEE	SEM	MSE	CP (%)
0.01	0%	Naive	0.109	0.062	0.060	0.018	21.3	0.122	0.066	0.064	0.019	24.5
		Conditional $\left(\widehat{\beta}\right)$	0.012	0.092	0.091	0.009	95.1	0.011	0.076	0.074	0.006	95.2
		Full $\left(\widetilde{\beta}\right)$	0.009	0.067	0.066	0.005	95.3	0.005	0.070	0.070	0.005	95.3
	25 %	Naive	0.115	0.076	0.075	0.031	19.8	0.123	0.082	0.080	0.022	22.6
		Conditional $\left(\widehat{\beta}\right)$	0.016	0.110	0.110	0.020	95.2	0.012	0.113	0.112	0.013	95.2
		Full $\left(\widetilde{\beta}\right)$	0.012	0.087	0.086	0.008	95.3	0.009	0.089	0.087	0.009	94.3
	50%	Naive	0.124	0.084	0.082	0.028	19.5	0.145	0.086	0.085	0.028	18.7
		Conditional $\left(\widehat{\beta}\right)$	0.017	0.111	0.109	0.013	95.0	0.013	0.121	0.120	0.015	95.0
		Full $\left(\widetilde{\beta}\right)$	0.015	0.096	0.094	0.010	95.1	0.009	0.095	0.094	0.009	95.2
0.5	0%	Naive	0.122	0.063	0.062	0.019	20.8	0.124	0.068	0.066	0.020	22.6
		Conditional $\left(\widehat{\beta}\right)$	0.020	0.093	0.093	0.009	95.0	0.025	0.091	0.090	0.009	94.9
		Full $\left(\widetilde{\beta}\right)$	0.015	0.076	0.074	0.006	95.2	0.013	0.074	0.073	0.004	95.1
	25 %	Naive	0.127	0.079	0.079	0.022	18.4	0.128	0.082	0.082	0.023	21.7
		Conditional $\left(\widehat{\beta}\right)$	0.022	0.112	0.112	0.016	94.9	0.027	0.119	0.118	0.016	94.8
		Full $\left(\widetilde{\beta}\right)$	0.017	0.104	0.101	0.011	95.1	0.016	0.094	0.093	0.009	95.0
	50%	Naive	0.137	0.084	0.084	0.026	20.4	0.148	0.096	0.094	0.030	21.7
		Conditional $\left(\widehat{\beta}\right)$	0.028	0.128	0.126	0.017	94.9	0.032	0.122	0.122	0.016	94.8
		Full $\left(\widetilde{\beta}\right)$	0.018	0.098	0.097	0.009	95.1	0.017	0.098	0.098	0.007	95.0
0.75	0%	Naive	0.127	0.063	0.063	0.020	18.5	0.130	0.069	0.069	0.020	17.6
		Conditional $\left(\widehat{\beta}\right)$	0.033	0.101	0.100	0.012	94.7	0.029	0.119	0.117	0.015	94.8
		Full $\left(\widetilde{\beta}\right)$	0.021	0.076	0.076	0.006	95.1	0.021	0.074	0.074	0.005	95.1
	25 %	Naive	0.159	0.086	0.084	0.033	16.3	0.134	0.084	0.082	0.025	14.5
		Conditional $\left(\widehat{\beta}\right)$	0.042	0.113	0.111	0.015	94.6	0.031	0.120	0.119	0.015	94.8
		Full $\left(\widetilde{\beta}\right)$	0.021	0.104	0.103	0.011	94.9	0.026	0.095	0.095	0.009	95.0
	50%	Naive	0.166	0.092	0.090	0.036	11.7	0.157	0.099	0.096	0.034	13.3
		Conditional $\left(\widehat{\beta}\right)$	0.048	0.129	0.128	0.016	94.6	0.035	0.125	0.124	0.016	94.5
		Full $\left(\widetilde{\beta}\right)$	0.026	0.118	0.115	0.015	94.9	0.030	0.109	0.108	0.012	95.0

Table S.3 Simulation results under Setting 4 and measurement error model (6) with $\alpha = 0$

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	SEE	SEM	MSE	CP (%)	Bias	SEE	SEM	MSE	CP (%)
0.01	0%	Naive	0.105	0.042	0.040	0.014	53.7	0.118	0.052	0.050	0.017	43.7
		Conditional $\left(\widehat{\beta}\right)$	0.015	0.080	0.080	0.010	95.2	0.015	0.094	0.093	0.010	95.1
		Full $\left(\widetilde{\beta}\right)$	0.011	0.060	0.058	0.005	95.3	0.006	0.065	0.065	0.005	95.4
	25 %	Naive	0.121	0.052	0.051	0.019	47.8	0.126	0.067	0.063	0.020	34.4
		Conditional $\left(\widehat{\beta}\right)$	0.016	0.083	0.082	0.007	95.0	0.016	0.106	0.105	0.011	95.4
		Full $\left(\widetilde{\beta}\right)$	0.012	0.065	0.064	0.005	95.2	0.009	0.078	0.077	0.006	95.3
	50%	Naive	0.124	0.059	0.058	0.020	39.4	0.128	0.076	0.074	0.022	27.7
		Conditional $\left(\widehat{\beta}\right)$	0.020	0.090	0.090	0.009	94.9	0.018	0.116	0.115	0.014	94.8
		Full $\left(\widetilde{\beta}\right)$	0.013	0.077	0.075	0.006	95.1	0.009	0.083	0.082	0.007	95.2
0.5	0%	Naive	0.114	0.048	0.046	0.015	42.5	0.126	0.065	0.063	0.020	38.4
		Conditional $\left(\widehat{\beta}\right)$	0.022	0.096	0.095	0.010	95.1	0.019	0.105	0.104	0.011	95.1
		Full $\left(\widetilde{\beta}\right)$	0.016	0.062	0.059	0.004	95.0	0.013	0.074	0.070	0.006	95.2
	25 %	Naive	0.136	0.066	0.065	0.023	35.1	0.142	0.067	0.064	0.025	29.8
		Conditional $\left(\widehat{\beta}\right)$	0.025	0.099	0.098	0.010	94.9	0.023	0.116	0.114	0.015	95.0
		Full $\left(\widetilde{\beta}\right)$	0.018	0.076	0.074	0.006	95.1	0.017	0.086	0.084	0.008	95.1
	50%	Naive	0.152	0.065	0.065	0.027	28.4	0.161	0.078	0.075	0.032	24.8
		Conditional $\left(\widehat{\beta}\right)$	0.028	0.119	0.118	0.015	94.8	0.024	0.118	0.118	0.018	94.9
		Full $\left(\widetilde{\beta}\right)$	0.018	0.074	0.073	0.005	95.0	0.020	0.087	0.086	0.008	95.1
0.75	0%	Naive	0.152	0.068	0.064	0.031	23.5	0.130	0.074	0.071	0.022	25.9
		Conditional $\left(\widehat{\beta}\right)$	0.029	0.104	0.103	0.012	94.8	0.026	0.117	0.115	0.016	94.8
		Full $\left(\widetilde{\beta}\right)$	0.018	0.070	0.069	0.005	95.0	0.017	0.082	0.080	0.010	95.1
	25 %	Naive	0.153	0.066	0.065	0.028	20.8	0.159	0.078	0.076	0.031	19.6
		Conditional $\left(\widehat{\beta}\right)$	0.030	0.127	0.125	0.017	94.7	0.025	0.117	0.117	0.016	94.8
		Full $\left(\widetilde{\beta}\right)$	0.022	0.084	0.083	0.008	95.0	0.018	0.088	0.087	0.009	95.0
	50%	Naive	0.162	0.069	0.068	0.028	18.5	0.165	0.081	0.080	0.034	12.4
		Conditional $\left(\widehat{\beta}\right)$	0.030	0.127	0.126	0.017	94.5	0.026	0.118	0.117	0.014	94.7
		Full $\left(\widetilde{\beta}\right)$	0.023	0.086	0.085	0.008	95.1	0.022	0.096	0.093	0.010	95.1

Table S.4 Simulation results under Setting 1 and measurement error model (6) with $\alpha = 10$

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	SEE	SEM	MSE	CP (%)	Bias	SEE	SEM	MSE	CP (%)
0.01	0%	Naive	0.110	0.030	0.029	0.013	54.5	0.105	0.046	0.045	0.013	46.5
		Conditional $\left(\widehat{\beta}\right)$	0.012	0.050	0.050	0.003	95.2	0.013	0.080	0.080	0.007	95.1
		Full $\left(\widetilde{\beta}\right)$	0.011	0.041	0.040	0.002	95.3	0.004	0.072	0.071	0.005	95.2
	25%	Naive	0.116	0.037	0.037	0.015	50.3	0.111	0.051	0.050	0.015	41.4
		Conditional $\left(\widehat{\beta}\right)$	0.016	0.061	0.060	0.004	95.1	0.013	0.111	0.110	0.012	95.1
		Full $\left(\widetilde{\beta}\right)$	0.012	0.043	0.042	0.002	95.2	0.006	0.077	0.076	0.006	95.2
	50%	Naive	0.120	0.045	0.044	0.016	46.7	0.123	0.058	0.058	0.018	40.6
		Conditional $\left(\widehat{\beta}\right)$	0.018	0.070	0.069	0.005	95.1	0.014	0.119	0.117	0.014	94.7
		Full $\left(\widetilde{\beta}\right)$	0.012	0.060	0.058	0.004	95.0	0.007	0.083	0.082	0.007	95.1
0.5	0%	Naive	0.114	0.041	0.040	0.015	38.6	0.117	0.057	0.057	0.017	41.7
		Conditional $\left(\widehat{\beta}\right)$	0.017	0.075	0.074	0.006	95.1	0.025	0.087	0.086	0.008	94.9
		Full $\left(\widetilde{\beta}\right)$	0.012	0.056	0.055	0.003	95.1	0.013	0.078	0.078	0.007	95.1
	25%	Naive	0.124	0.047	0.046	0.018	31.5	0.119	0.071	0.070	0.019	37.5
		Conditional $\left(\widehat{\beta}\right)$	0.018	0.089	0.087	0.008	94.7	0.037	0.115	0.115	0.015	94.7
		Full $\left(\widetilde{\beta}\right)$	0.013	0.064	0.064	0.004	95.0	0.024	0.093	0.093	0.009	95.1
	50%	Naive	0.134	0.053	0.053	0.021	21.8	0.130	0.082	0.080	0.024	19.4
		Conditional $\left(\widehat{\beta}\right)$	0.024	0.097	0.097	0.010	94.5	0.040	0.122	0.121	0.016	94.6
		Full $\left(\widetilde{\beta}\right)$	0.014	0.068	0.067	0.005	94.8	0.026	0.098	0.097	0.011	94.9
0.75	0%	Naive	0.189	0.050	0.049	0.038	11.6	0.134	0.066	0.065	0.022	10.1
		Conditional $\left(\widehat{\beta}\right)$	0.024	0.085	0.085	0.008	94.8	0.031	0.110	0.110	0.013	94.7
		Full $\left(\widetilde{\beta}\right)$	0.013	0.066	0.065	0.004	95.0	0.020	0.104	0.104	0.011	94.9
	25%	Naive	0.198	0.068	0.066	0.044	7.9	0.138	0.084	0.084	0.026	7.1
		Conditional $\left(\widehat{\beta}\right)$	0.025	0.100	0.098	0.011	94.7	0.046	0.135	0.133	0.020	94.5
		Full $\left(\widetilde{\beta}\right)$	0.016	0.071	0.070	0.005	94.8	0.025	0.109	0.109	0.013	94.8
	50%	Naive	0.233	0.077	0.075	0.060	5.0	0.174	0.090	0.089	0.038	4.3
		Conditional $\left(\widehat{\beta}\right)$	0.025	0.116	0.115	0.014	94.4	0.050	0.142	0.142	0.023	94.5
		Full $\left(\widetilde{\beta}\right)$	0.018	0.085	0.084	0.007	94.7	0.028	0.117	0.115	0.014	94.8

Table S.5 Simulation results under Setting 1 and measurement error model (6) with $\alpha = 50$

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	SEE	SEM	MSE	CP (%)	Bias	SEE	SEM	MSE	CP (%)
0.01	0%	Naive	0.106	0.031	0.030	0.012	51.4	0.108	0.050	0.050	0.014	50.1
		Conditional $\left(\widehat{\beta}\right)$	0.010	0.054	0.053	0.003	95.3	0.012	0.079	0.078	0.006	95.3
		Full $\left(\widetilde{\beta}\right)$	0.005	0.046	0.045	0.002	95.2	0.011	0.073	0.072	0.005	95.1
	25%	Naive	0.111	0.039	0.039	0.014	47.9	0.110	0.061	0.060	0.014	45.6
		Conditional $\left(\widehat{\beta}\right)$	0.005	0.069	0.068	0.005	95.1	0.023	0.098	0.097	0.010	95.0
		Full $\left(\widetilde{\beta}\right)$	0.017	0.048	0.047	0.003	95.2	0.016	0.077	0.077	0.006	95.0
	50%	Naive	0.118	0.046	0.044	0.016	43.0	0.114	0.069	0.078	0.018	40.4
		Conditional $\left(\widehat{\beta}\right)$	0.017	0.077	0.076	0.006	95.0	0.028	0.109	0.109	0.013	94.8
		Full $\left(\widetilde{\beta}\right)$	0.008	0.052	0.052	0.003	95.1	0.017	0.078	0.077	0.006	95.0
0.5	0%	Naive	0.122	0.043	0.042	0.017	40.5	0.124	0.067	0.065	0.020	40.3
		Conditional $\left(\widehat{\beta}\right)$	0.016	0.073	0.072	0.006	94.8	0.024	0.102	0.102	0.011	94.8
		Full $\left(\widetilde{\beta}\right)$	0.014	0.065	0.065	0.004	94.9	0.015	0.093	0.092	0.009	94.9
	25%	Naive	0.123	0.050	0.049	0.018	38.7	0.125	0.075	0.075	0.021	37.7
		Conditional $\left(\widehat{\beta}\right)$	0.020	0.081	0.078	0.007	94.7	0.026	0.106	0.106	0.012	94.8
		Full $\left(\widetilde{\beta}\right)$	0.018	0.069	0.067	0.005	94.9	0.020	0.099	0.099	0.010	94.8
	50%	Naive	0.125	0.063	0.061	0.020	33.4	0.129	0.082	0.082	0.023	31.1
		Conditional $\left(\widehat{\beta}\right)$	0.025	0.093	0.093	0.009	94.4	0.030	0.115	0.113	0.014	94.6
		Full $\left(\widetilde{\beta}\right)$	0.019	0.076	0.075	0.006	94.8	0.023	0.103	0.101	0.011	94.8
0.75	0%	Naive	0.214	0.056	0.054	0.049	16.0	0.133	0.077	0.076	0.024	32.8
		Conditional $\left(\widehat{\beta}\right)$	0.024	0.092	0.090	0.009	94.7	0.027	0.118	0.116	0.015	94.7
		Full $\left(\widetilde{\beta}\right)$	0.014	0.075	0.074	0.006	94.8	0.017	0.098	0.098	0.010	95.0
	25%	Naive	0.230	0.068	0.067	0.057	10.3	0.172	0.090	0.090	0.038	8.0
		Conditional $\left(\widehat{\beta}\right)$	0.022	0.104	0.104	0.011	94.7	0.033	0.121	0.120	0.016	94.5
		Full $\left(\widetilde{\beta}\right)$	0.019	0.088	0.086	0.008	94.8	0.020	0.109	0.107	0.012	94.7
	50%	Naive	0.254	0.077	0.076	0.070	4.9	0.184	0.095	0.093	0.043	4.6
		Conditional $\left(\widehat{\beta}\right)$	0.030	0.116	0.114	0.014	94.2	0.041	0.126	0.126	0.018	94.5
		Full $\left(\widetilde{\beta}\right)$	0.021	0.096	0.095	0.010	94.6	0.026	0.110	0.109	0.013	94.7

Table S.6 Simulation results under Setting 1 and measurement error model (6) with $\alpha = 100$

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	SEE	SEM	MSE	CP (%)	Bias	SEE	SEM	MSE	CP (%)
0.01	0%	Naive	0.109	0.040	0.040	0.013	54.4	0.112	0.047	0.047	0.015	57.2
		Conditional $\left(\widehat{\beta}\right)$	0.002	0.053	0.051	0.003	95.2	0.007	0.078	0.076	0.006	95.1
		Full $\left(\widetilde{\beta}\right)$	0.001	0.046	0.045	0.002	95.3	0.007	0.071	0.070	0.005	95.3
	25%	Naive	0.110	0.048	0.047	0.014	45.0	0.123	0.056	0.055	0.018	50.6
		Conditional $\left(\widehat{\beta}\right)$	0.003	0.060	0.057	0.004	95.0	0.010	0.085	0.084	0.007	95.1
		Full $\left(\widetilde{\beta}\right)$	0.002	0.053	0.053	0.003	95.3	0.010	0.078	0.076	0.006	95.2
	50%	Naive	0.112	0.050	0.050	0.015	30.4	0.125	0.057	0.055	0.019	37.0
		Conditional $\left(\widehat{\beta}\right)$	0.005	0.067	0.065	0.005	94.9	0.013	0.099	0.098	0.010	95.0
		Full $\left(\widetilde{\beta}\right)$	0.004	0.060	0.060	0.004	95.1	0.011	0.085	0.085	0.007	95.2
0.5	0%	Naive	0.115	0.043	0.043	0.015	45.8	0.121	0.054	0.053	0.018	52.0
		Conditional $\left(\widehat{\beta}\right)$	0.028	0.058	0.057	0.004	94.9	0.028	0.084	0.083	0.008	94.8
		Full $\left(\widetilde{\beta}\right)$	0.002	0.050	0.050	0.002	95.1	0.012	0.075	0.075	0.006	95.2
	25%	Naive	0.115	0.050	0.050	0.016	31.0	0.127	0.059	0.059	0.020	41.6
		Conditional $\left(\widehat{\beta}\right)$	0.029	0.064	0.063	0.004	94.9	0.027	0.090	0.089	0.009	94.8
		Full $\left(\widetilde{\beta}\right)$	0.005	0.056	0.055	0.003	95.1	0.017	0.080	0.078	0.007	95.0
	50%	Naive	0.129	0.054	0.054	0.020	20.2	0.141	0.064	0.063	0.024	27.4
		Conditional $\left(\widehat{\beta}\right)$	0.034	0.070	0.070	0.006	94.7	0.036	0.115	0.114	0.015	94.8
		Full $\left(\widetilde{\beta}\right)$	0.010	0.063	0.062	0.004	95.0	0.020	0.087	0.086	0.008	95.0
0.75	0%	Naive	0.116	0.047	0.047	0.016	37.3	0.128	0.060	0.059	0.020	29.2
		Conditional $\left(\widehat{\beta}\right)$	0.043	0.065	0.065	0.006	94.8	0.042	0.087	0.087	0.009	94.6
		Full $\left(\widetilde{\beta}\right)$	0.018	0.054	0.053	0.003	95.0	0.019	0.078	0.078	0.006	95.1
	25%	Naive	0.121	0.053	0.053	0.017	26.2	0.151	0.065	0.065	0.027	11.0
		Conditional $\left(\widehat{\beta}\right)$	0.045	0.069	0.068	0.007	94.8	0.044	0.093	0.093	0.011	94.6
		Full $\left(\widetilde{\beta}\right)$	0.021	0.062	0.062	0.004	95.1	0.023	0.083	0.083	0.007	95.0
	50%	Naive	0.126	0.056	0.055	0.019	15.0	0.165	0.067	0.066	0.032	9.4
		Conditional $\left(\widehat{\beta}\right)$	0.048	0.085	0.084	0.010	94.7	0.050	0.122	0.121	0.017	94.5
		Full $\left(\widetilde{\beta}\right)$	0.022	0.066	0.065	0.005	95.0	0.024	0.089	0.089	0.008	95.0

Table S.7 Sensitivity analyses for simulated data under Setting 1 and measurement error model (6) with censoring rate 50% and $\Sigma_\epsilon = 0.75$

α	R	Method	Estimator of β_x					Estimator of β_z				
			Bias	SEE	SEM	MSE	CP (%)	Bias	SEE	SEM	MSE	CP (%)
0	0.65	Conditional $\left(\widehat{\beta}\right)$	0.040	0.098	0.097	0.011	94.6	0.048	0.110	0.110	0.014	94.7
		Full $\left(\widetilde{\beta}\right)$	0.014	0.066	0.065	0.005	95.2	0.017	0.073	0.072	0.006	95.1
	0.75	Conditional $\left(\widehat{\beta}\right)$	0.044	0.103	0.102	0.012	94.5	0.055	0.116	0.114	0.018	94.7
		Full $\left(\widetilde{\beta}\right)$	0.014	0.069	0.068	0.005	95.2	0.021	0.097	0.096	0.009	94.9
	0.85	Conditional $\left(\widehat{\beta}\right)$	0.046	0.107	0.107	0.013	94.5	0.064	0.119	0.118	0.018	94.3
		Full $\left(\widetilde{\beta}\right)$	0.015	0.070	0.070	0.005	95.0	0.028	0.108	0.106	0.012	94.7
10	0.65	Conditional $\left(\widehat{\beta}\right)$	0.021	0.058	0.057	0.004	94.8	0.005	0.099	0.099	0.010	95.1
		Full $\left(\widetilde{\beta}\right)$	0.017	0.051	0.050	0.003	95.3	0.016	0.079	0.078	0.006	95.2
	0.75	Conditional $\left(\widehat{\beta}\right)$	0.028	0.072	0.071	0.006	94.8	0.013	0.111	0.110	0.012	94.9
		Full $\left(\widetilde{\beta}\right)$	0.016	0.065	0.065	0.005	94.9	0.026	0.099	0.098	0.010	95.1
	0.85	Conditional $\left(\widehat{\beta}\right)$	0.034	0.085	0.083	0.008	94.6	0.017	0.119	0.116	0.014	94.9
		Full $\left(\widetilde{\beta}\right)$	0.024	0.072	0.071	0.006	94.8	0.032	0.099	0.099	0.011	95.0
50	0.65	Conditional $\left(\widehat{\beta}\right)$	0.027	0.067	0.065	0.005	95.0	0.008	0.109	0.109	0.012	95.2
		Full $\left(\widetilde{\beta}\right)$	0.014	0.058	0.057	0.004	95.1	0.012	0.096	0.096	0.009	95.3
	0.75	Conditional $\left(\widehat{\beta}\right)$	0.036	0.076	0.075	0.007	94.9	0.014	0.112	0.110	0.013	95.0
		Full $\left(\widetilde{\beta}\right)$	0.021	0.066	0.066	0.005	95.1	0.024	0.097	0.097	0.010	95.0
	0.85	Conditional $\left(\widehat{\beta}\right)$	0.039	0.082	0.080	0.008	94.8	0.018	0.119	0.117	0.014	94.9
		Full $\left(\widetilde{\beta}\right)$	0.028	0.068	0.066	0.005	94.9	0.034	0.099	0.099	0.011	94.8
100	0.65	Conditional $\left(\widehat{\beta}\right)$	0.032	0.092	0.090	0.009	95.0	0.045	0.119	0.118	0.016	95.1
		Full $\left(\widetilde{\beta}\right)$	0.010	0.056	0.056	0.003	95.3	0.014	0.078	0.077	0.006	95.2
	0.75	Conditional $\left(\widehat{\beta}\right)$	0.034	0.098	0.097	0.011	94.7	0.046	0.122	0.122	0.017	94.7
		Full $\left(\widetilde{\beta}\right)$	0.025	0.060	0.060	0.004	95.2	0.027	0.101	0.100	0.011	95.0
	0.85	Conditional $\left(\widehat{\beta}\right)$	0.037	0.105	0.104	0.012	94.5	0.052	0.126	0.125	0.019	94.6
		Full $\left(\widetilde{\beta}\right)$	0.027	0.063	0.062	0.005	95.2	0.026	0.113	0.113	0.013	94.9

Table S.8 Simulation with misspecified measurement error model under Scenario A

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	S.E.	MVE	MSE	CP (%)	Bias	S.E.	MVE	MSE	CP (%)
0.01	0%	Conditional $\left(\widehat{\beta}\right)$	0.004	0.043	0.041	0.002	95.2	0.014	0.071	0.070	0.005	95.2
		Full $\left(\widetilde{\beta}\right)$	0.003	0.042	0.040	0.002	95.3	0.010	0.069	0.068	0.005	95.4
	25%	Conditional $\left(\widehat{\beta}\right)$	0.006	0.050	0.047	0.003	95.0	0.013	0.080	0.079	0.007	95.1
		Full $\left(\widetilde{\beta}\right)$	0.003	0.049	0.046	0.002	95.2	0.014	0.073	0.073	0.006	95.4
	50%	Conditional $\left(\widehat{\beta}\right)$	0.010	0.062	0.057	0.004	95.0	0.017	0.095	0.093	0.009	94.9
		Full $\left(\widetilde{\beta}\right)$	0.008	0.057	0.054	0.003	95.1	0.016	0.089	0.088	0.008	95.2
0.5	0%	Conditional $\left(\widehat{\beta}\right)$	0.024	0.050	0.049	0.003	94.9	0.026	0.073	0.071	0.006	94.9
		Full $\left(\widetilde{\beta}\right)$	0.010	0.047	0.046	0.002	95.1	0.012	0.069	0.067	0.005	95.0
	25%	Conditional $\left(\widehat{\beta}\right)$	0.026	0.058	0.056	0.004	94.8	0.028	0.080	0.079	0.007	94.8
		Full $\left(\widetilde{\beta}\right)$	0.012	0.054	0.054	0.002	95.1	0.015	0.076	0.075	0.006	95.0
	50%	Conditional $\left(\widehat{\beta}\right)$	0.027	0.068	0.066	0.005	94.8	0.034	0.104	0.096	0.012	94.7
		Full $\left(\widetilde{\beta}\right)$	0.013	0.057	0.055	0.003	95.1	0.027	0.098	0.096	0.010	95.1
0.75	0%	Conditional $\left(\widehat{\beta}\right)$	0.040	0.055	0.055	0.005	94.8	0.039	0.078	0.076	0.008	94.8
		Full $\left(\widetilde{\beta}\right)$	0.017	0.052	0.049	0.002	95.0	0.021	0.070	0.069	0.005	95.1
	25%	Conditional $\left(\widehat{\beta}\right)$	0.044	0.066	0.062	0.006	94.7	0.044	0.090	0.088	0.010	94.7
		Full $\left(\widetilde{\beta}\right)$	0.019	0.052	0.050	0.003	95.1	0.024	0.079	0.078	0.007	95.0
	50%	Conditional $\left(\widehat{\beta}\right)$	0.044	0.083	0.080	0.009	94.7	0.045	0.111	0.110	0.014	94.7
		Full $\left(\widetilde{\beta}\right)$	0.022	0.058	0.056	0.003	94.9	0.025	0.090	0.088	0.009	95.0

Table S.9 Simulation with misspecified measurement error model under Scenario B

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	S.E.	MVE	MSE	CP (%)	Bias	S.E.	MVE	MSE	CP (%)
0.01	0%	Conditional $\left(\widehat{\beta}\right)$	0.041	0.046	0.044	0.005	83.4	0.057	0.076	0.074	0.009	83.2
		Full $\left(\widetilde{\beta}\right)$	0.028	0.044	0.043	0.003	86.6	0.053	0.061	0.060	0.006	85.0
	25%	Conditional $\left(\widehat{\beta}\right)$	0.050	0.049	0.047	0.005	80.0	0.069	0.076	0.075	0.010	82.8
		Full $\left(\widetilde{\beta}\right)$	0.038	0.047	0.046	0.004	82.6	0.067	0.075	0.074	0.010	84.6
	50%	Conditional $\left(\widehat{\beta}\right)$	0.060	0.062	0.058	0.007	80.0	0.103	0.096	0.091	0.019	82.0
		Full $\left(\widetilde{\beta}\right)$	0.041	0.056	0.056	0.005	82.0	0.090	0.092	0.090	0.019	84.6
0.5	0%	Conditional $\left(\widehat{\beta}\right)$	0.075	0.052	0.047	0.008	78.8	0.077	0.076	0.075	0.011	83.6
		Full $\left(\widetilde{\beta}\right)$	0.033	0.048	0.043	0.003	85.4	0.059	0.066	0.065	0.008	84.8
	25%	Conditional $\left(\widehat{\beta}\right)$	0.079	0.054	0.053	0.009	70.4	0.082	0.082	0.079	0.013	81.8
		Full $\left(\widetilde{\beta}\right)$	0.040	0.048	0.047	0.004	81.4	0.069	0.080	0.075	0.012	84.6
	50%	Conditional $\left(\widehat{\beta}\right)$	0.084	0.067	0.065	0.011	70.1	0.106	0.111	0.096	0.024	81.4
		Full $\left(\widetilde{\beta}\right)$	0.049	0.056	0.056	0.006	81.8	0.100	0.098	0.095	0.020	84.3
0.75	0%	Conditional $\left(\widehat{\beta}\right)$	0.089	0.061	0.060	0.012	77.2	0.089	0.076	0.076	0.014	82.0
		Full $\left(\widetilde{\beta}\right)$	-0.049	0.050	0.049	0.004	83.4	0.060	0.067	0.066	0.008	85.0
	25%	Conditional $\left(\widehat{\beta}\right)$	0.093	0.066	0.062	0.013	70.0	0.093	0.088	0.085	0.016	81.8
		Full $\left(\widetilde{\beta}\right)$	-0.050	0.050	0.050	0.005	82.4	0.070	0.084	0.082	0.012	84.8
	50%	Conditional $\left(\widehat{\beta}\right)$	0.098	0.080	0.076	0.016	70.2	0.110	0.113	0.103	0.021	79.0
		Full $\left(\widetilde{\beta}\right)$	0.053	0.065	0.065	0.006	82.2	0.101	0.099	0.097	0.020	84.2

Table S.10 Simulation results under Setting 1 and measurement error model (6) with $\alpha = 0$ in the presence of validation data

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	SEE	SEM	MSE	CP (%)	Bias	SEE	SEM	MSE	CP (%)
0.01	0%	Naive	0.069	0.030	0.030	0.006	16.5	0.106	0.040	0.038	0.013	14.6
		Conditional $\left(\widehat{\beta}_{val}\right)$	0.006	0.043	0.041	0.002	95.1	0.012	0.063	0.061	0.004	95.1
		Full $\left(\widetilde{\beta}_{val}\right)$	0.004	0.040	0.039	0.002	95.3	0.010	0.045	0.043	0.002	95.2
	25%	Naive	0.091	0.034	0.033	0.009	16.3	0.109	0.048	0.046	0.014	12.7
		Conditional $\left(\widehat{\beta}_{val}\right)$	0.010	0.045	0.044	0.002	95.1	0.013	0.070	0.070	0.005	95.0
		Full $\left(\widetilde{\beta}_{val}\right)$	0.008	0.042	0.040	0.002	95.0	0.012	0.060	0.058	0.004	95.1
	50%	Naive	0.127	0.040	0.040	0.018	14.0	0.106	0.052	0.050	0.014	11.2
		Conditional $\left(\widehat{\beta}_{val}\right)$	0.011	0.054	0.052	0.003	95.0	0.013	0.087	0.086	0.008	94.9
		Full $\left(\widetilde{\beta}_{val}\right)$	0.010	0.047	0.047	0.002	95.0	0.013	0.063	0.060	0.004	95.0
0.5	0%	Naive	0.115	0.035	0.034	0.014	14.5	0.117	0.046	0.045	0.016	13.7
		Conditional $\left(\widehat{\beta}_{val}\right)$	0.018	0.047	0.045	0.003	94.9	0.017	0.067	0.066	0.005	94.8
		Full $\left(\widetilde{\beta}_{val}\right)$	0.010	0.042	0.041	0.002	95.1	0.012	0.058	0.056	0.004	95.1
	25%	Naive	0.127	0.038	0.038	0.018	13.8	0.123	0.051	0.050	0.018	10.0
		Conditional $\left(\widehat{\beta}_{val}\right)$	0.023	0.050	0.049	0.003	94.8	0.027	0.076	0.075	0.007	94.6
		Full $\left(\widetilde{\beta}_{val}\right)$	0.015	0.046	0.045	0.002	95.0	0.017	0.068	0.066	0.005	95.0
	50%	Naive	0.134	0.046	0.044	0.020	11.6	0.160	0.053	0.052	0.028	9.6
		Conditional $\left(\widehat{\beta}_{val}\right)$	0.028	0.056	0.055	0.004	94.8	0.030	0.081	0.080	0.007	94.7
		Full $\left(\widetilde{\beta}_{val}\right)$	0.018	0.050	0.050	0.003	95.1	0.022	0.074	0.074	0.006	94.9
0.75	0%	Naive	0.133	0.038	0.037	0.019	12.2	0.125	0.050	0.050	0.018	13.0
		Conditional $\left(\widehat{\beta}_{val}\right)$	0.032	0.051	0.050	0.004	94.8	0.031	0.073	0.073	0.006	94.9
		Full $\left(\widetilde{\beta}_{val}\right)$	0.016	0.046	0.045	0.002	95.1	0.017	0.064	0.064	0.004	95.2
	25%	Naive	0.137	0.040	0.039	0.020	9.3	0.138	0.058	0.057	0.022	10.6
		Conditional $\left(\widehat{\beta}_{val}\right)$	0.035	0.060	0.060	0.005	94.7	0.033	0.085	0.084	0.008	94.6
		Full $\left(\widetilde{\beta}_{val}\right)$	0.021	0.051	0.050	0.003	95.0	0.023	0.076	0.076	0.006	95.0
	50%	Naive	0.144	0.048	0.047	0.023	6.6	0.149	0.065	0.065	0.026	8.8
		Conditional $\left(\widehat{\beta}_{val}\right)$	0.038	0.083	0.082	0.008	94.6	0.040	0.093	0.093	0.010	94.6
		Full $\left(\widetilde{\beta}_{val}\right)$	0.022	0.054	0.053	0.003	95.0	0.026	0.082	0.082	0.007	94.9

Table S.11 Simulation results under Setting 1 and measurement error model (6) with $\alpha = 10$ in the presence of validation data

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	SEE	SEM	MSE	CP (%)	Bias	SEE	SEM	MSE	CP (%)
0.01	0%	Naive	0.107	0.028	0.027	0.012	56.3	0.109	0.041	0.040	0.014	44.5
		Conditional $\left(\widehat{\beta}\right)$	0.012	0.045	0.045	0.003	95.2	0.009	0.075	0.075	0.006	95.3
		Full $\left(\widetilde{\beta}\right)$	0.010	0.039	0.039	0.002	95.4	0.009	0.069	0.068	0.005	95.2
	25%	Naive	0.113	0.030	0.030	0.014	51.9	0.117	0.048	0.047	0.016	40.7
		Conditional $\left(\widehat{\beta}\right)$	0.015	0.055	0.055	0.003	95.2	0.013	0.105	0.104	0.012	95.1
		Full $\left(\widetilde{\beta}\right)$	0.010	0.040	0.040	0.002	95.4	0.012	0.075	0.074	0.006	95.2
	50%	Naive	0.115	0.040	0.040	0.015	45.3	0.125	0.052	0.050	0.018	40.1
		Conditional $\left(\widehat{\beta}\right)$	0.017	0.067	0.065	0.005	95.3	0.014	0.112	0.111	0.013	94.9
		Full $\left(\widetilde{\beta}\right)$	0.011	0.056	0.055	0.003	95.1	0.013	0.078	0.077	0.006	95.1
0.5	0%	Naive	0.115	0.035	0.034	0.014	43.1	0.116	0.049	0.048	0.016	40.3
		Conditional $\left(\widehat{\beta}\right)$	0.015	0.064	0.063	0.004	95.0	0.026	0.077	0.076	0.007	94.9
		Full $\left(\widetilde{\beta}\right)$	0.012	0.052	0.052	0.003	95.2	0.012	0.074	0.074	0.006	95.1
	25%	Naive	0.126	0.041	0.040	0.018	30.5	0.120	0.066	0.064	0.019	38.1
		Conditional $\left(\widehat{\beta}\right)$	0.015	0.068	0.066	0.005	94.8	0.034	0.109	0.109	0.013	94.8
		Full $\left(\widetilde{\beta}\right)$	0.012	0.054	0.054	0.003	95.2	0.017	0.086	0.085	0.008	95.1
	50%	Naive	0.137	0.046	0.044	0.021	22.4	0.128	0.071	0.070	0.021	21.4
		Conditional $\left(\widehat{\beta}\right)$	0.022	0.075	0.073	0.006	94.8	0.042	0.114	0.113	0.015	94.8
		Full $\left(\widetilde{\beta}\right)$	0.017	0.058	0.058	0.004	94.9	0.018	0.090	0.090	0.008	94.9
0.75	0%	Naive	0.196	0.039	0.039	0.040	10.9	0.130	0.054	0.053	0.020	11.5
		Conditional $\left(\widehat{\beta}\right)$	0.025	0.077	0.075	0.007	94.7	0.027	0.104	0.104	0.012	94.9
		Full $\left(\widetilde{\beta}\right)$	0.017	0.058	0.058	0.004	95.2	0.018	0.098	0.097	0.010	95.2
	25%	Naive	0.195	0.048	0.046	0.040	9.6	0.132	0.078	0.077	0.024	8.7
		Conditional $\left(\widehat{\beta}\right)$	0.028	0.088	0.086	0.008	94.7	0.049	0.129	0.128	0.019	94.6
		Full $\left(\widetilde{\beta}\right)$	0.019	0.060	0.060	0.004	95.1	0.020	0.099	0.099	0.010	94.8
	50%	Naive	0.230	0.056	0.055	0.056	5.4	0.163	0.084	0.082	0.034	4.6
		Conditional $\left(\widehat{\beta}\right)$	0.028	0.103	0.100	0.011	94.5	0.047	0.133	0.132	0.020	94.6
		Full $\left(\widetilde{\beta}\right)$	0.020	0.061	0.061	0.004	95.0	0.021	0.106	0.105	0.012	94.8

Table S.12 Simulation results under Setting 1 and measurement error model (6) with $\alpha = 50$ in the presence of validation data

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	SEE	SEM	MSE	CP (%)	Bias	SEE	SEM	MSE	CP (%)
0.01	0%	Naive	0.107	0.027	0.026	0.012	53.6	0.106	0.046	0.045	0.013	52.7
		Conditional $\left(\widehat{\beta}\right)$	0.003	0.043	0.042	0.002	95.4	0.012	0.069	0.068	0.005	95.3
		Full $\left(\widetilde{\beta}\right)$	0.001	0.041	0.040	0.002	95.3	0.009	0.065	0.064	0.004	95.4
	25%	Naive	0.109	0.029	0.028	0.013	50.7	0.108	0.050	0.048	0.012	48.6
		Conditional $\left(\widehat{\beta}\right)$	0.004	0.059	0.058	0.004	95.5	0.018	0.094	0.093	0.009	95.3
		Full $\left(\widetilde{\beta}\right)$	0.002	0.042	0.041	0.002	95.3	0.013	0.067	0.067	0.005	95.2
	50%	Naive	0.120	0.030	0.030	0.016	44.1	0.108	0.058	0.056	0.015	46.3
		Conditional $\left(\widehat{\beta}\right)$	0.013	0.068	0.067	0.005	95.2	0.019	0.101	0.100	0.010	95.3
		Full $\left(\widetilde{\beta}\right)$	0.006	0.048	0.046	0.003	95.1	0.015	0.068	0.067	0.005	95.2
0.5	0%	Naive	0.117	0.036	0.036	0.015	50.2	0.113	0.056	0.055	0.016	50.1
		Conditional $\left(\widehat{\beta}\right)$	0.015	0.063	0.062	0.004	95.0	0.022	0.093	0.093	0.009	95.1
		Full $\left(\widetilde{\beta}\right)$	0.008	0.053	0.053	0.003	95.2	0.012	0.088	0.087	0.008	95.0
	25%	Naive	0.118	0.040	0.039	0.016	46.3	0.122	0.064	0.062	0.019	47.8
		Conditional $\left(\widehat{\beta}\right)$	0.018	0.070	0.068	0.005	95.1	0.043	0.096	0.096	0.011	95.0
		Full $\left(\widetilde{\beta}\right)$	0.012	0.058	0.057	0.004	95.0	0.014	0.089	0.089	0.008	95.0
	50%	Naive	0.124	0.055	0.055	0.016	43.1	0.126	0.072	0.072	0.021	40.5
		Conditional $\left(\widehat{\beta}\right)$	0.031	0.075	0.074	0.007	94.7	0.055	0.109	0.107	0.015	94.8
		Full $\left(\widetilde{\beta}\right)$	0.014	0.063	0.063	0.004	94.8	0.015	0.091	0.090	0.009	95.1
0.75	0%	Naive	0.228	0.043	0.042	0.054	13.6	0.145	0.062	0.060	0.025	38.6
		Conditional $\left(\widehat{\beta}\right)$	0.016	0.075	0.073	0.006	94.9	0.037	0.107	0.106	0.013	94.8
		Full $\left(\widetilde{\beta}\right)$	0.010	0.056	0.054	0.003	95.0	0.015	0.088	0.088	0.008	95.1
	25%	Naive	0.241	0.049	0.047	0.060	9.6	0.159	0.081	0.080	0.032	10.7
		Conditional $\left(\widehat{\beta}\right)$	0.020	0.086	0.086	0.008	94.9	0.050	0.116	0.114	0.016	94.8
		Full $\left(\widetilde{\beta}\right)$	0.013	0.058	0.055	0.004	95.0	0.016	0.097	0.095	0.010	94.9
	50%	Naive	0.252	0.057	0.056	0.067	5.3	0.168	0.085	0.083	0.035	8.4
		Conditional $\left(\widehat{\beta}\right)$	0.033	0.096	0.094	0.010	94.7	0.058	0.119	0.119	0.018	94.6
		Full $\left(\widetilde{\beta}\right)$	0.017	0.069	0.067	0.005	94.8	0.021	0.098	0.097	0.010	94.9

Table S.13 Simulation results with length-biased sampling

Σ_ϵ	cr	Method	Estimator of β_x					Estimator of β_z				
			Bias	S.E.	MVE	MSE	CP (%)	Bias	S.E.	MVE	MSE	CP (%)
0.01	0%	Naive	0.085	0.027	0.027	0.008	19.4	0.084	0.045	0.045	0.008	66.1
		Conditional $\left(\widehat{\beta}_{LB}\right)$	0.000	0.044	0.041	0.002	95.4	0.007	0.071	0.071	0.005	95.2
		Full $\left(\widetilde{\beta}_{LB}\right)$	-0.000	0.035	0.034	0.002	95.3	0.007	0.071	0.069	0.005	95.4
	25%	Naive	0.094	0.042	0.041	0.012	14.1	0.094	0.056	0.052	0.011	57.2
		Conditional $\left(\widehat{\beta}_{LB}\right)$	0.003	0.057	0.054	0.003	95.0	0.009	0.093	0.090	0.009	95.1
		Full $\left(\widetilde{\beta}_{LB}\right)$	0.002	0.051	0.050	0.003	95.2	0.010	0.092	0.088	0.009	95.3
	50%	Naive	0.112	0.051	0.050	0.018	12.7	0.116	0.072	0.072	0.016	38.7
		Conditional $\left(\widehat{\beta}_{LB}\right)$	0.006	0.081	0.077	0.007	95.0	0.025	0.136	0.126	0.019	95.1
		Full $\left(\widetilde{\beta}_{LB}\right)$	0.003	0.067	0.066	0.005	95.2	0.020	0.116	0.114	0.014	95.2
0.5	0%	Naive	0.090	0.027	0.027	0.011	13.6	0.098	0.054	0.051	0.011	66.1
		Conditional $\left(\widehat{\beta}_{LB}\right)$	0.019	0.048	0.047	0.003	95.0	0.030	0.075	0.073	0.007	94.9
		Full $\left(\widetilde{\beta}_{LB}\right)$	-0.000	0.036	0.036	0.001	95.3	0.004	0.071	0.069	0.005	95.5
	25%	Naive	0.100	0.042	0.042	0.013	11.8	0.108	0.056	0.053	0.013	54.1
		Conditional $\left(\widehat{\beta}_{LB}\right)$	0.024	0.065	0.062	0.005	94.9	0.036	0.098	0.094	0.011	94.8
		Full $\left(\widetilde{\beta}_{LB}\right)$	-0.002	0.053	0.052	0.002	95.2	0.008	0.093	0.093	0.008	95.3
	50%	Naive	0.118	0.051	0.051	0.017	11.7	0.116	0.079	0.077	0.016	20.8
		Conditional $\left(\widehat{\beta}_{LB}\right)$	0.036	0.096	0.090	0.010	94.8	0.051	0.140	0.136	0.022	94.8
		Full $\left(\widetilde{\beta}_{LB}\right)$	0.006	0.068	0.066	0.005	95.2	0.015	0.118	0.113	0.014	95.0
0.75	0%	Naive	0.108	0.028	0.028	0.012	9.1	0.113	0.058	0.057	0.014	53.5
		Conditional $\left(\widehat{\beta}_{LB}\right)$	0.024	0.056	0.056	0.005	94.9	0.044	0.087	0.079	0.010	94.8
		Full $\left(\widetilde{\beta}_{LB}\right)$	-0.019	0.037	0.037	0.002	95.2	-0.011	0.072	0.070	0.005	95.1
	25%	Naive	0.116	0.044	0.043	0.016	6.6	0.126	0.059	0.059	0.018	43.4
		Conditional $\left(\widehat{\beta}_{LB}\right)$	0.032	0.073	0.073	0.007	94.9	0.045	0.105	0.102	0.013	94.8
		Full $\left(\widetilde{\beta}_{LB}\right)$	-0.017	0.056	0.055	0.002	95.1	-0.011	0.095	0.094	0.007	95.1
	50%	Naive	0.125	0.054	0.053	0.017	8.4	0.123	0.083	0.082	0.017	14.5
		Conditional $\left(\widehat{\beta}_{LB}\right)$	0.035	0.128	0.121	0.021	94.8	0.081	0.176	0.168	0.038	94.7
		Full $\left(\widetilde{\beta}_{LB}\right)$	0.018	0.070	0.069	0.005	95.0	0.004	0.119	0.113	0.014	95.1

References

Asgharian, M., M'Lan, C. E. & Wolfson, D. B. (2002). Length-Biased Sampling with Right Censoring: An Unconditional Approach. *Journal of the American Statistical Association* **97**, 201-209.

Chandra, T. K., and Chatterjee, D. (2000). *A First Course in Probability.* Narosa.

De Una-Alvarez, J. (2004). Nonparametric estimation under length-biased sampling and Type I censoring: a moment based approach. *Ann. Inst. Statist. Math* **56**, 667-681.

Huang, C. Y. and Qin J. (2011). Nonparametric estimation for length-biased and right-censored data. *Biometrika* **98**, 177-186.

Huang, C. Y., Qin J., and Follmann, D. A. (2012). A maximum pseudo-profile likelihood estimator for the Cox model under length-biased sampling. *Biometrika* **99**, 199-210.

Jiang, J. (2010). *Large Sample Techniques for Statistics.* Springer, New York.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data.* Cambridge University Press, New York.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data.* Wiley, New York.

Lin, D. Y. and Wei, L. J. (1989). The robust inference for the cox proportional hazards model. *Journal of American Statistical Association* **84**, 1074-1078.

Qin, J. and Shen, Y. (2010). Statistical methods for analyzing right-censored length-biased data under Cox model. *Biometrics* **66**, 382-392.

van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge University Press, New York.

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

Wang, M. C. (1991). Nonparametric estimation from cross-sectional survival data. *Journal of the American Statistical Association* **86**, 130-143.

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