Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

ARTICLE TYPE

Supplementary material for "Semiparametric analysis of a generalized linear model with multiple covariates subject to detection limits"

Ling-Wan Chen[1] | Jason P. Fine[2] | Eric Bair[3] | Victor S. Ritter[1] | Thomas F. McElrath[4] | David E. Cantonwine[4] | John D. Meeker[5] | Kelly K. Ferguson[6] | Shanshan Zhao*[1]

[1]Biostatistics & Computational Biology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA
[2]Departments of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[3]Sciome LLC, Durham, NC, USA
[4]Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[5]Department of Environmental Health Sciences, University of Michigan School of Public Health, Ann Arbor, MI, USA
[6]Epidemiology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

Correspondence
*Shanshan Zhao, 111 TW Alexander Dr, Research Triangle Park, NC 27709, USA.
Email: shanshan.zhao@nih.gov

Summary

This is the supplementary material for "Semiparametric analysis of a generalized linear model with multiple covariates subject to detection limits", including the proof of Theorem 1, and additional Tables and Figures.

1 | LEMMAS

To derive the asymptotic properties of $\hat{\theta}$ for arbitrary p from $p = 1$, we require five technical lemmas (Lemmas 1-5), followed by two important lemmas about the general Z-estimation theory (Lemmas 6-7). For the simplicity of notation and the ease of proofs for lemmas, we present the case $p = 2$ and sketch the extension to arbitrary p in each lemma.

Let $\boldsymbol{\alpha}_0 = (\alpha_{10}, \alpha_{20})^T$ be the true coefficients for AFT models, $\eta_0(t_1, t_2)$ be the true cumulative joint distribution for the error term in AFT models, and $(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)$ be the true value of $(\theta, \phi, \boldsymbol{\alpha}, \eta)$. Denote $\Psi(\theta, \phi, \boldsymbol{\alpha}, \eta)$ as a deterministic function, which is defined by

$$\Psi(\theta, \phi, \boldsymbol{\alpha}, \eta) = E\{\psi(Y, \mathbf{X}, \mathbf{V}, \boldsymbol{\Delta}; \theta, \phi, \boldsymbol{\alpha}, \eta)\}.$$

Let p^* be the "in outer probability" and P^* be the outer probability where $P^*(B) = \inf\{P(A) : A \supset B, A \in \mathcal{A}\}$ for any subset B of Ω in a probability space $\{\Omega, \mathcal{A}, P\}$. For a vector u, define $u^{\otimes 2} = uu^T$ and $|u|$ as its Euclidean norm. Let $\| \eta - \eta_0 \| = \sup_{\mathbf{t}} |\eta(\mathbf{t}) - \eta_0(\mathbf{t})|$ and $\nu\{(\phi, \boldsymbol{\alpha}, \eta), (\phi_0, \boldsymbol{\alpha}_0, \eta_0)\} = | \phi - \phi_0 | + | \boldsymbol{\alpha} - \boldsymbol{\alpha}_0 | + \| \eta - \eta_0 \|$.

Lemma 1. If $\ddot{b}(\cdot)$ is a bounded Lipschitz function, then

$$\{\Delta_1\Delta_2[Y - \dot{b}\{\theta^T D_{(\mathbf{X},\mathbf{t})}\} D_{(\mathbf{X},\mathbf{t})}] : \theta \in \Theta, \mathbf{t} = (t_1, t_2)^T \in \mathcal{T} \subset \mathbb{R}^2\}$$

is Donsker.

Proof. Under Condition C8 where the $\ddot{b}(\cdot)$ is bounded, the $\dot{b}(\cdot)$ is then a Lipschitiz function. With the indicator functions Δ_1 and Δ_2 and Theorem 2.10.6 in van der Vaart and Wellner,[1] we have that $D_{(\mathbf{X},\mathbf{t})}$ and $\dot{b}\{\theta^T D_{(\mathbf{X},\mathbf{t})}\}$ are Donsker. Therefore, the class $\{\Delta_1\Delta_2[Y - \dot{b}\{\theta^T D_{(\mathbf{X},\mathbf{t})}\}D_{(\mathbf{X},\mathbf{t})}] : \theta \in \Theta, \mathbf{t} = (t_1, t_2) \in \mathcal{T} \subset \mathbb{R}^2\}$ is Donsker. We can extend Lemma 1 to arbitrary p by modifying $\Delta_1\Delta_2$ to $\Delta_1\Delta_2 \cdots \Delta_p$ and $\mathbf{t} = (t_1, t_2)^T$ to $\mathbf{t} = (t_1, ..., t_p)^T \subset \mathbb{R}^p$.

Lemma 2. Let $\mathcal{X}$ and $\mathcal{A}$ be the bounded covariate and parameter spaces, and $\mathcal{H}$ be a collection of distribution functions satisfying Conditions 3 and 4. We have $\mathcal{F} = \{\eta(t_1 - \alpha_1^T x, t_2 - \alpha_2^T x) : (t_1, t_2) \in \mathcal{T}^2 \subset \mathbb{R}^2, x \in \mathcal{X}, (\alpha_1, \alpha_2) \in \mathcal{A}, \eta \in \mathcal{H}\}$ is Donsker.

Proof. Let $\mathcal{F}_1 = \{\eta(\mathbf{t}), \mathbf{t} = (t_2, t_2)\}$. From Corollary 2.7.2 in van der Vaart and Wellner,[1] let the number of brackets $[L_i, U_i]$ such that $L_i(\mathbf{t}) \le \eta(\mathbf{t}) \le U_i(\mathbf{t})$ for any nontrivial ϵ, with $0 < \epsilon < 1$ and $\int\int |U_i(t_1, t_2) - L_i(t_1, t_2)| d\eta(t_1, t_2) < \epsilon^s$, satisfies $\log N_{[]}(\epsilon, \mathcal{F}_1, L_1(P)) \le K_1/\epsilon^s$ with some constants $K_1 < \infty$ and $0 < s < 1$. Let α_1 and α_2 be one-dimensional for the brevity of notation. Since the parameter space $\mathcal{A}$ is bounded, $\mathcal{A}$ can be partitioned by a set of rectangular regions $[l_{k_1}, u_{k_1}) \times [l_{k_2}, u_{k_2})$ such that $|u_{k_1} - l_{k_1}| \le \epsilon^{s/2}$ and $|u_{k_2} - l_{k_2}| \le \epsilon^{s/2}$. Thus, the number of rectangular regions is bounded by K_2/ϵ^s with some constants $K_2 < \infty$ and $0 < s < 1$.

We now consider $\mathcal{F} = \{\eta(t_1 - \alpha_1^T x, t_2 - \alpha_2^T x)\}$ and find the bracketing number for $\mathcal{F}$. Let

$$\begin{aligned} O_{ik}(t_1, t_2, x) = \min\{&L_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x), L_i(t_1 - u_{k_1}x, t_2 - u_{k_2}x), \\ &L_i(t_1 - l_{k_1}x, t_2 - u_{k_2}x), L_i(t_1 - u_{k_1}x, t_2 - l_{k_2}x)\}, \end{aligned}$$

and

$$\begin{aligned} S_{ik}(t_1, t_2, x) = \max\{&U_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x), U_i(t_1 - u_{k_1}x, t_2 - u_{k_2}x), \\ &U_i(t_1 - l_{k_1}x, t_2 - u_{k_2}x), U_i(t_1 - u_{k_1}x, t_2 - l_{k_2}x)\}. \end{aligned}$$

We have

$$\begin{aligned} &O_{ik}(t_1, t_2, x) \\ &\le \min\{\eta_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x), \eta_i(t_1 - u_{k_1}x, t_2 - u_{k_2}x), \\ &\qquad \eta_i(t_1 - l_{k_1}x, t_2 - u_{k_2}x), \eta_i(t_1 - u_{k_1}x, t_2 - l_{k_2}x)\} \\ &\le \eta(t_1 - \alpha_1^T x, t_2 - \alpha_2^T x) \\ &\le \max\{\eta_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x), \eta_i(t_1 - u_{k_1}x, t_2 - u_{k_2}x), \\ &\qquad \eta_i(t_1 - l_{k_1}x, t_2 - u_{k_2}x), \eta_i(t_1 - u_{k_1}x, t_2 - l_{k_2}x)\} \\ &\le S_{ik}(t_1, t_2, x). \end{aligned}$$

Therefore,

$$P|S_{ik}(t_1,t_2,x) - O_{ik}(t_1,t_2,x)|$$

$$\leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |U_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x) - L_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x) dF_X(x) \tag{S.1}$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |U_i(t_1 - u_{k_1}x, t_2 - u_{k_2}x) - L_i(t_1 - u_{k_1}x, t_2 - u_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x) dF_X(x) \tag{S.2}$$

$$\ldots$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |U_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x) - L_i(t_1 - u_{k_1}x, t_2 - l_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x) dF_X(x) \tag{S.3}$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |U_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x) - L_i(t_1 - l_{k_1}x, t_2 - u_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x) dF_X(x) \tag{S.4}$$

$$\ldots$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |U_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x) - L_i(t_1 - u_{k_1}x, t_2 - u_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x) dF_X(x). \tag{S.5}$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |U_i(t_1 - u_{k_1}x, t_2 - u_{k_2}x) - L_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x) dF_X(x). \tag{S.6}$$

The above calculation involves 16 integral equations, which can be grouped to three different types. Let $U_i(s_{11}, s_{12})$ and $L_i(s_{21}, s_{22})$, and define: Type 1. $s_{11} = s_{21}$ and $s_{12} = s_{22}$, i.e. equations (S.1) and (S.2); Type 2. either $s_{11} \neq s_{21}$ or $s_{12} \neq s_{22}$, i.e. equations (S.3) and (S.4); Type 3. $s_{11} \neq s_{21}$ and $s_{12} \neq s_{22}$, i.e. equations (S.5) and (S.6). In this calculation, there are 4, 8 and 4 equations for Types 1-3 integral. We calculated the equations by considering their types. For Type 1. integral, since $[L_i, U_i]$ are brackets for $\mathcal{F}_1$, we have those integrals $\leq \epsilon^s$. Now we calculate Type 2. integral using equation (S.3) as an example. Under Condition C4 for $\dot{\eta}_{0,\boldsymbol{\alpha}_0}$ where $\dot{\eta}_{0,\boldsymbol{\alpha}_0} < c$ and some calculations, we have that

$$
\begin{aligned}
(S.3) = &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Big| U_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x) - \eta(t_1 - l_{k_1}x, t_2 - l_{k_2}x) + \eta(t_1 - l_{k_1}x, t_2 - l_{k_2}x) \\
&\qquad - \eta(t_1 - u_{k_1}x, t_2 - l_{k_2}x) + \eta(t_1 - u_{k_1}x, t_2 - l_{k_2}x) - L_i(t_1 - u_{k_1}x, t_2 - l_{k_2}x)\Big| \\
&\qquad \times d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x)dF_X(x) \\
\le\ & 2\epsilon^s + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\eta(t_1 - l_{k_1}x, t_2 - l_{k_2}x) - \eta(t_1 - l_{k_1}x, t_2 - u_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x)dF_X(x) \\
=\ & 2\epsilon^s + \int_{0}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{\eta(t_1 - l_{k_1}x, t_2 - l_{k_2}x) - \eta(t_1 - l_{k_1}x, t_2 - u_{k_2}x)\} d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x)dF_X(x) \\
&+ \int_{-\infty}^{0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{\eta(t_1 - l_{k_1}x, t_2 - u_{k_2}x) - \eta(t_1 - l_{k_1}x, t_2 - l_{k_2}x)\} d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x)dF_X(x) \\
=\ & 2\epsilon^s + \int_{0}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{\eta_0(t_1 + \alpha_{10}x + l_{k_1}x, t_2 + \alpha_{20}x + u_{k_2}x) \\
&\qquad - \eta_0(t_1 + \alpha_{10}x + l_{k_1}x, t_2 + \alpha_{20}x + l_{k_2}x)\} d\eta(t_1, t_2)dF_X(x) \\
&+ \int_{-\infty}^{0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{\eta_0(t_1 + \alpha_{10}x + l_{k_1}x, t_2 + \alpha_{20}x + l_{k_2}x) \\
&\qquad - \eta_0(t_1 + \alpha_{10}x + l_{k_1}x, t_2 + \alpha_{20}x + u_{k_2}x)\} d\eta(t_1, t_2)dF_X(x) \\
\le\ & 2\epsilon^s + \int_{0}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c_1(u_{k_2} - l_{k_2})x d\eta(t_1, t_2)dF_X(x) - \int_{-\infty}^{0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c_1(u_{k_2} - l_{k_2})x d\eta(t_1, t_2)dF_X(x) \\
\le\ & 2\epsilon^s + cE|X|\epsilon^{s/2} = K_3\epsilon^s,
\end{aligned}
$$

where $K_3 = 2 + c_1E|X|\epsilon^2 < \infty$ and $0 < s < 1$. Therefore, those equations, which belong to Type 2. integral, have an upper bound $K_3\epsilon^s$. Finally, we show that Type 3. integral also has an upper bound using equation (S.5) as an example, where

$$
\begin{aligned}
(S.5) = &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |U_i(t_1 - l_{k_1}x, t_2 - l_{k_2}x) - \eta(t_1 - l_{k_1}x, t_2 - l_{k_2}x) + \eta(t_1 - l_{k_1}x, t_2 - l_{k_2}x) \\
&- \eta(t_1 - u_{k_1}x, t_2 - u_{k_2}x) + \eta(t_1 - u_{k_1}x, t_2 - u_{k_2}x) - L_i(t_1 - u_{k_1}x, t_2 - u_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x) dF_X(x) \\
\le\ & 2\epsilon^s + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\eta(t_1 - l_{k_1}x, t_2 - l_{k_2}x) - \eta(t_1 - u_{k_1}x, t_2 - u_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x) dF_X(x) \\
=\ & 2\epsilon^s + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\eta(t_1 - l_{k_1}x, t_2 - l_{k_2}x) - \eta(t_1 - l_{k_1}x, t_2 - u_{k_2}x) + \eta(t_1 - l_{k_1}x, t_2 - u_{k_2}x) \\
&\qquad - \eta(t_1 - u_{k_1}x, t_2 - u_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x) dF_X(x) \\
\le\ & 2\epsilon^s + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\eta(t_1 - l_{k_1}x, t_2 - l_{k_2}x) - \eta(t_1 - l_{k_1}x, t_2 - u_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x) dF_X(x) \\
&+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\eta(t_1 - l_{k_1}x, t_2 - u_{k_2}x) - \eta(t_1 - u_{k_1}x, t_2 - u_{k_2}x)| d\eta_0(t_1 + \alpha_{10}x, t_2 + \alpha_{20}x) dF_X(x) \\
\le\ & 2\epsilon^s + 2c_1 E|X|\epsilon^{s/2} = K_4\epsilon^s,
\end{aligned}
$$

and $K_4 = 2 + 2c_1 E|X|\epsilon^2 < \infty$. Hence, $N_{[]}\{\epsilon, \mathcal{F}, L_1(P)\} \le \exp\{K_1(4 + 8K_3 + 4K_4)/\epsilon^s\}\{K_2(4 + 8K_3 + 4K_4)/\epsilon^s\} \le \exp\{(K_1 + K_2)(4 + 8K_3 + 4K_4)/\epsilon^s\}$, i.e. $\log N_{[]}\{\epsilon, \mathcal{F}, L_1(P)\} \le K_5/\epsilon^s$, where $K_5 = (K_1 + K_2)(4 + 8K_3 + 4K_4) < \infty$ and $0 < s < 1$. We obtain that $\mathcal{F}$ is P-Donsker.

Lemma 2 can be extended to arbitrary p using $\eta(t_1 - \alpha_1^T x, \ldots, t_p - \alpha_p^T x)$, and can be proved by considering all combinations of pairs for $\boldsymbol{\alpha}$ in L_i and U_i. We demonstrate that there are 16 equations for $p = 2$. For $p > 2$, it is still doable but requires more steps. In addition, Corollary 2.7.2 in van der Vaart and Wellner[1] is for arbitrary p, and the follow-up discussion after Corollary 2.7.2 on Page 157 in van der Vaart and Wellner[1] details the conditions for showing P-Donsker in an arbitrary p.

Lemma 3. Suppose Conditions 1-3 and 6-10 hold, we have that

$$
\left\{ \frac{\int_{C_1-\alpha_1^T x}^{\tau_1} f_y(y|t_1 + \alpha_1^T x, T_2, x)\big[y - \dot{b}\{\theta^T D_{(x,t_1+\alpha_1^T x,T_2)}\}\big] D_{(x,t_1+\alpha_1^T x,T_2)} \eta(dt_1, dt_2 = T_2 - \alpha_2^T x)}{\int_{C_1-\alpha_1^T x}^{\tau_1} f_y(y|t_1 + \alpha_1^T x, T_2, x)\eta(dt_1, dt_2 = T_2 - \alpha_2^T x)} : \right.
$$
$$
\left. \theta \in \Theta, |1/a(\phi)| < l, (\alpha_1, \alpha_2) \in \mathcal{A}, v\{(\hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\alpha}}), (\phi_0, \boldsymbol{\alpha}_0, \eta_0)\} < \delta_1, x \in \mathcal{X}, y \in \mathcal{Y} \right\}, \tag{S.7}
$$

$$
\left\{ \frac{\int_{C_2-\alpha_2^T x}^{\tau_2} f_y(y|t_2 + \alpha_2^T x, T_1, x)\big[y - \dot{b}\{\theta^T D_{(x,T_1,t_2+\alpha_2^T x)}\}\big] D_{(x,T_1,t_2+\alpha_2^T x)} \eta(dt_1 = T_1 - \alpha_1^T x, dt_2)}{\int_{C_2-\alpha_2^T x}^{\tau_2} f_y(y|t_2 + \alpha_2^T x, T_1, x)\eta(dt_1 = T_1 - \alpha_1^T x, dt_2)} : \right.
$$
$$
\left. \theta \in \Theta, |1/a(\phi)| < l, (\alpha_1, \alpha_2) \in \mathcal{A}, v\{(\hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\alpha}}), (\phi_0, \boldsymbol{\alpha}_0, \eta_0)\} < \delta_2, x \in \mathcal{X}, y \in \mathcal{Y} \right\}, \tag{S.8}
$$

and

$$
\left\{ \left(\int_{C_1-\alpha_1^T x}^{\tau_1} \int_{C_2-\alpha_2^T x}^{\tau_2} f_y(y|t_1 + \alpha_1^T x, t_2 + \alpha_2^T x, x)\big[y - \dot{b}\{\theta^T D_{(x,t_1+\alpha_1^T x,t_2+\alpha_2^T x)}\}\big] \right. \right.
$$
$$
\left. D_{(x,t_1+\alpha_1^T x,t_2+\alpha_2^T x)} \eta(dt_1, dt_2) \right) \Big/ \int_{C_1-\alpha_1^T x}^{\tau_1} \int_{C_2-\alpha_2^T x}^{\tau_2} f_y(y|t_1 + \alpha_1^T x, t_2 + \alpha_2^T x, x)\eta(dt_1, dt_2) :
$$
$$
\left. \theta \in \Theta, |1/a(\phi)| < l, (\alpha_1, \alpha_2) \in \mathcal{A}, v\{(\hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\alpha}}), (\phi_0, \boldsymbol{\alpha}_0, \eta_0)\} < \delta_3, x \in \mathcal{X}, y \in \mathcal{Y} \right\} \tag{S.9}
$$

are Donsker.

Proof: This lemma is an extension of Lemma 6 in Kong and Nan[2] for $p = 1$. Since the first two classes (S.7) and (S.8) are for the subset where the subjects have only one covariate censored among two covariates, (S.7) and (S.8) can be shown as Donsker using Lemma 6 in Kong and Nan.[2] We here focus on showing the class (S.9) is Donsker.

Under Condition10 where $\int_{C_1-\alpha_1^T x}^{\tau_1}\int_{C_2-\alpha_2^T x}^{\tau_1} f_y(y|t_1+\alpha_1^T x, t_2+\alpha_2^T x, x)\eta(dt_1, dt_2)$ is bounded away from 0, the proof can be completed by showing the numerator and denominator in (S.9) are both Donsker via Theorem 2.10.6 and Example 2.10.8 in van der Vaart and Wellner.[1] Using the integration by parts, we have

$$
\begin{aligned}
&\int_{C_1-\alpha_1^T x}^{\tau_1}\int_{C_2-\alpha_2^T x}^{\tau_2} f_y(y|t_1+\alpha_1^T x, t_2+\alpha_2^T x, x)\eta(dt_1, dt_2)\\
&\approx \eta(\tau_1,\tau_2) f_y(y|\tau_1+\alpha_1^T x, \tau_2+\alpha_2^T x, x) - \eta(C_1-\alpha_1^T x, \tau_2) f_y(y|C_1, \tau_2+\alpha_2^T x, x)\\
&\quad - \eta(\tau_1, C_2-\alpha_2^T x) f_y(y|\tau_1+\alpha_1^T x, C_2, x) + \eta(C_1-\alpha_1^T x, C_2-\alpha_2^T x) f_y(y|C_1, C_2, x)\\
&\quad - \int_{C_1-c_{31}}^{\tau_1}\int_{C_2-c_{32}}^{\tau_2} I(t_1\geq C_1-\alpha_1^T x, t_2 \geq C_2-\alpha_2^T x)\eta(t_1,t_2) f_y(y|t_1+\alpha_1^T x, t_2+x^T\alpha_2, x)\\
&\qquad \times [y-\dot{b}\{\theta^T D_{(x,t_1+\alpha_1^T x, t_2+\alpha_2^T x)}\}]\gamma^T \dot{h}_1(t_1+\alpha_1^T x)\dot{h}_2(t_2+\alpha_2^T x)/a(\phi) dt_1 dt_2.
\end{aligned}
$$

Under Conditions 3 and 7-9, $f_y(y|t_1+\alpha_1^T x, t_2+x^T\alpha_2, x)$ is Lipschitz function for $(\theta, \phi, \boldsymbol{\alpha})$, and $\dot{h}(\mathbf{t})$ is Lipschitz function for $\mathbf{t}$. By Section 2.6 in van der Vaart and Wellner,[1] the class for the indicator functions $\{I(t_1 \geq C_1-\alpha_1^T x, t_2\geq C_2-\alpha_2^T x)\}$ is a Donsker. From Lemma 2, $\{\eta(C_1-\alpha_1^T x, C_2-\alpha_2^T x)\}$ is Donsker. Hence, by Section 2.10.1 and Theorem 2.10.3 of van der Vaart and Wellner,[1] we have $\{I(t_1\geq C_1-\alpha_1^T x, t_2\geq C_2-\alpha_2^T x)\eta(t_1,t_2) f_y(y|t_1+\alpha_1^T x, t_2+\alpha_2^T x, x)[y-\dot{b}\{\theta^T D_{(x,t_1+\alpha_1^T x,t_2+\alpha_2^T x)}\}]\gamma^T\dot{h}_1(t_1+\alpha_1^T x)\dot{h}_2(t_2+\alpha_2^T x)/a(\phi)\}$ is a Donsker, which implies that the denominator of (S.9) is Donsker. Similarly, we use integration by parts for the numerator in (S.9) and show that the numerator (S.9) is a Donsker class. The steps are similar to the above for the denominator. This completes the proof.

The extension of Lemma 3 for arbitrary p will have $2^p - 1$ classes (i.e. 3 classes for $p = 2$). The proof for arbitrary$p$ is quite tedious because it requires more calculations for integration by parts for each class. Thus, we only show $p = 2$ but the results work for any arbitrary p.

Lemma 4. Suppose Conditions 7-10 hold. When $\theta \to \theta_0$ and $\nu\{(\hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\alpha}}), (\phi_0, \boldsymbol{\alpha}_0, \eta_0)\} \to 0$, we have that $E\{|\psi(\theta,\phi,\alpha,\eta) - \psi(\theta_0,\phi_0,\alpha_0,\eta_0)|^2\} \to 0$

Proof: This lemma is similar to Lemma 7 in Kong and Nan[2] for $p = 1$. The proof is based on the Mean Value Theorem and can be directly shown by simple calculations. We thus omit the details. Lemma 4 for arbitrary p works.

Lemma 5. Suppose Conditions 3, 7-10 hold, we have $E|\psi(\theta_0,\phi_0,\alpha_0,\eta_0)|^2 < \infty$

Proof: Given Condition 3, 7-10, the proof follows a straightforward calculation. Thus, the details are omitted. This lemma also works for arbitrary p and can be proved by direct calculation.

Lemma 6. (Consistency) Suppose θ_0 is the unique solution to $\Psi(\theta,\phi_0,\boldsymbol{\alpha}_0,\eta_0) = 0$ in the parameter space Θ and $(\hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\alpha}})$ are the estimators of $(\phi_0, \boldsymbol{\alpha}_0, \eta_0)$ such that $\nu\{(\hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\alpha}}), (\phi_0,$
$\boldsymbol{\alpha}_0, \eta_0)\} = o_{p^*}(1)$. If

$$
\sup_{\theta\in\Theta, \nu\{(\phi,\alpha,\eta),(\phi_0,\alpha_0,\eta_0)\}\leq\delta_n} \frac{|\Psi_n(\theta,\phi,\boldsymbol{\alpha},\eta) - \Psi(\theta,\phi_0,\boldsymbol{\alpha}_0,\eta_0)|}{1+|\Psi_n(\theta,\phi,\boldsymbol{\alpha},\eta)| + |\Psi(\theta,\phi_0,\boldsymbol{\alpha}_0,\eta_0)|} = o_{p^*}(1)
$$

for every sequence $\{\delta_n \downarrow 0\}$, then $\hat{\theta}$ satisfying $\Psi_n(\hat{\theta}, \hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\alpha}}) = o_{p^*}(1)$ converges in outer probability to θ_0.

This lemma is proposed in Nan and Wellner[3] where they introduced a general asymptotic theory for semiparametric Z-estimation with bundled parameters. We thus reformat their Lemma 2.1 as our Lemma 6. The details of proof and discussion for this lemma can be found in Nan and Wellner.[3] This lemma works for arbitrary p by modifying $\boldsymbol{\alpha}$ and η. Hence, to prove the consistency of $\hat{\theta}$, we only need to verify the condition in this lemma. We detail the proof of consistency in Section 2.

Lemma 7. (Rate of convergence and asymptotic representation) Suppose that $\hat{\theta}$ satisfying $\Psi_n(\hat{\theta}, \hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\alpha}}) = o_{p^*}(n^{-1/2})$ is a consistent estimator of θ_0 that is the unique solution to $\Psi_n(\theta,\phi_0,\boldsymbol{\alpha}_0,\eta_0)$ in Θ, and that $(\hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\alpha}})$ are the estimators of $(\phi, \boldsymbol{\alpha}, \eta)$ such that $\nu\{(\hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\alpha}}), (\phi_0,\boldsymbol{\alpha}_0,\eta_0)\} = O_{p^*}(n^{1/2})$. Suppose that the following four conditions are satisfied.

i. (Stochastic equicontinuity)

$$\frac{|n^{1/2}(\Psi_n - \Psi)(\hat{\theta}, \hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}) - n^{1/2}(\Psi_n - \Psi)(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)|}{1 + n^{1/2}|\Psi_n(\hat{\theta}, \hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\boldsymbol{\alpha}}})| + n^{1/2}|\Psi(\hat{\theta}, \hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\boldsymbol{\alpha}}})|} = o_{p^*}(1)$$

ii. $n^{1/2}\Psi_n(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0) = O_{p^*}(1)$;

iii. (Smoothness) there exists continuous matrices $\dot{\Psi}_1(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)$, $\dot{\Psi}_2(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)$, $\dot{\Psi}_3(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)$ and a continuous linear functional $\dot{\Psi}_4(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)$ such that

$$\begin{aligned}
&[\Psi(\hat{\theta}, \hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}) - \Psi(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0) - \dot{\Psi}_1(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\theta} - \theta_0) - \dot{\Psi}_2(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\phi} - \phi_0) \\
&\qquad - \dot{\Psi}_3(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) - \dot{\Psi}_4(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\eta}_{\hat{\boldsymbol{\alpha}}} - \eta_0)] \\
&= o(|\hat{\theta} - \theta_0|) + o[\nu\{(\hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}), (\phi_0, \boldsymbol{\alpha}_0, \eta_0)\}],
\end{aligned}$$

where we assume that the matrix $\dot{\Psi}_1(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)$ is non-singular; and

iv. $n^{1/2}\dot{\Psi}_2(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\phi} - \phi_0) = O_{p^*}(1)$, $n^{1/2}\dot{\Psi}_3(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) = O_{p^*}(1)$ and $\dot{\Psi}_4(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\eta}_{\hat{\boldsymbol{\alpha}}} - \eta_0) = O_{p^*}(1)$.

Then $\hat{\theta}$ is $n^{1/2}-$consistent and

$$\begin{aligned}
\sqrt{n}(\hat{\theta} - \theta_0) =& \{-\dot{\Psi}_1(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)\}^{-1}\sqrt{n}\{(\Psi_n - \Psi)(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0) + \dot{\Psi}_2(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\phi} - \phi_0) \\
&+ \dot{\Psi}_3(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0) + \dot{\Psi}_4(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\eta}_{\hat{\boldsymbol{\alpha}}} - \eta_0)\} + o_{p^*}(1).
\end{aligned}$$

This lemma is again studied in Nan and Wellner[3] for semiparametric Z-estimation with bundled parameters. Since our nuisance parameters are free of the parameter of interest θ, we rewrite their Corollary 2.1 as our Lemma 7. The detailed proof and discussion for this lemma can also be found in Nan and Wellner.[3] This lemma can be extended to arbitrary p by modifying those parts involving $\boldsymbol{\alpha}$ and η. With this result, we can show the asymptotic normality of $\hat{\theta}$ by verifying four conditions i-iv in this lemma in Section 3.

2 | PROOF OF CONSISTENCY

To show the consistency of $\hat{\theta}$, we verify the condition for Lemma 6. Firstly, $\hat{\theta}$ is the unique solution of $\Psi_n(\theta, \phi, \boldsymbol{\alpha}, \eta) = 0$, where Ψ_n is the derivative of log-likelihood with respect to θ. Under regularity conditions, since $\hat{\psi}$, $\hat{\boldsymbol{\alpha}}$ and $\hat{\eta}_{\hat{\boldsymbol{\alpha}}}$ are $n^{1/2}$-consistent estimates, we have $\nu\{(\hat{\phi}, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}), (\phi_0, \boldsymbol{\alpha}_0, \eta_0)\} = o_{p^*}(1)$. Hence, via Lemma 6, the proof can be completed by showing that

$$\sup_{\theta \in \Theta, \nu\{(\phi, \boldsymbol{\alpha}, \eta), (\phi_0, \boldsymbol{\alpha}_0, \eta_0)\} \le \delta_n} |\Psi_n(\theta, \phi, \boldsymbol{\alpha}, \eta) - \Psi(\theta, \phi_0, \boldsymbol{\alpha}_0, \eta_0)| = o_{p^*}(1), \tag{S.10}$$

for every sequence $\delta_n \to 0$.

We first define the form of Ψ_n for an arbitrary p. Specifically,

$$
\begin{aligned}
&\Psi_n(\theta, \phi, \boldsymbol{\alpha}, \eta) = \frac{1}{n} \sum_{i=1}^{n} \psi(Y_i, \mathbf{X}_i, \mathbf{V}_i, \boldsymbol{\Delta}_i; \theta, \phi, \boldsymbol{\alpha}, \eta) \\
&= \frac{1}{n} \sum_{i=1}^{n} \Bigg(\prod_{j=1}^{p} \Delta_{ji} \big[Y_i - \dot{b}\{\theta^T D_{(\mathbf{X}_i,\mathbf{T}_i)}\}\big] D_{(\mathbf{X},\mathbf{T}_i)} \\
&+ \sum_{j=1}^{p} (1-\Delta_{ji})(\prod_{l\neq j} \Delta_{li}) \Bigg\{ \int_{C_j - \alpha_j^T \mathbf{X}_i}^{\tau_j} f_Y(Y_i | s_j + \alpha_j^T \mathbf{X}_i, \mathbf{T}_{-ji}, \mathbf{X}_i) \eta(ds_j, ds_l = T_{li} - \alpha_l^T \mathbf{X}_i, \forall l \neq j) \Bigg\}^{-1} \\
&\quad \times \int_{C_j - \alpha_j^T \mathbf{X}_i}^{\tau_j} \Big(f_Y(Y_i | s_j + \alpha_j^T \mathbf{X}_i, \mathbf{T}_{-ji}, \mathbf{X}_i) \big[Y_i - \dot{b}\{\theta^T D_{(\mathbf{X}_i, s_j + \alpha_j^T \mathbf{X}_i, \mathbf{T}_{-ji})}\}\big] D_{(\mathbf{X}_i, s_j + \alpha_j^T \mathbf{X}_i, \mathbf{T}_{-ji})} \\
&\qquad\qquad \times \eta(ds_j, ds_l = T_{li} - \alpha_l^T \mathbf{X}_i, \forall l \neq j) \Big) \\
&+ \sum_{j=1}^{p} \sum_{k>j}^{p} (1-\Delta_{ji})(1-\Delta_{ki})(\prod_{l\neq j\neq k} \Delta_{li}) \Bigg\{ \int_{C_j - \alpha_j^T \mathbf{X}_i}^{\tau_j} \int_{C_k - \alpha_k^T \mathbf{X}_i}^{\tau_k} f_Y(Y_i | s_j + \alpha_j^T \mathbf{X}_i, s_k + \alpha_k^T \mathbf{X}_i, \mathbf{T}_{-(j,k)i}, \mathbf{X}_i) \\
&\qquad\qquad \times \eta(ds_j, ds_k, ds_l = T_{li} - \alpha_l^T \mathbf{X}_i, \forall l \neq \{j,k\}) \Bigg\}^{-1} \\
&\quad \times \int_{C_j - \alpha_j^T \mathbf{X}_i}^{\tau_j} \int_{C_k - \alpha_k^T \mathbf{X}_i}^{\tau_k} \Big\{ f_Y(Y_i | s_j + \alpha_j^T \mathbf{X}_i, s_k + \alpha_k^T \mathbf{X}_i, \mathbf{T}_{-(j,k)i}, \mathbf{X}_i) \\
&\qquad\qquad \times \big[Y_i - \dot{b}\{\theta^T D_{(\mathbf{X}_i, s_j + \alpha_j^T \mathbf{X}_i, s_k + \alpha_k^T \mathbf{X}_i, \mathbf{T}_{-(j,k)i}\}})\big] \\
&\qquad\qquad \times D_{(\mathbf{X}_i, s_j + \alpha_j^T \mathbf{X}_i, s_k + \alpha_k^T \mathbf{X}_i, \mathbf{T}_{-(j,k)i})} \eta(ds_j, ds_k, ds_l = T_{li} - \alpha_l^T \mathbf{X}_i, \forall l \neq \{j,k\}) \Big\} \\
&+ \cdots \\
&+ \prod_{j=1}^{p} (1-\Delta_{ji}) \Bigg\{ \int_{C_1 - \alpha_1^T \mathbf{X}_i}^{\tau_1} \cdots \int_{C_p - \alpha_p^T \mathbf{X}_i}^{\tau_p} f_{\theta,\phi}(Y_i | s_1 + \alpha_1^T \mathbf{X}_i, \cdots, s_p + \alpha_p^T \mathbf{X}_i, \mathbf{X}_i) \eta(ds_1, \ldots, ds_p) \Bigg\}^{-1} \\
&\times \int_{C_1 - \alpha_1^T \mathbf{X}_i}^{\tau_1} \cdots \int_{C_p - \alpha_p^T \mathbf{X}_i}^{\tau_p} \Big(f_{\theta,\phi}(Y_i | s_1 + \alpha_1^T \mathbf{X}_i, \cdots, s_p + \alpha_p^T \mathbf{X}_i, \mathbf{X}_i) \big[Y_i - \dot{b}\{\theta^T D_{(\mathbf{X}_i, s_1 + \alpha_1^T \mathbf{X}_i, \cdots, s_p + \alpha_p^T \mathbf{X}_i)}\}\big] \\
&\qquad\qquad \times D_{(\mathbf{X}_i, s_1 + \alpha_1^T \mathbf{X}_i, \cdots, s_p + \alpha_p^T \mathbf{X}_i)} \eta(ds_1, \ldots, ds_p) \Big) \Bigg).
\end{aligned}
$$

Since showing the equation (S.10) for any arbitrary p dimension involves heavy notation which makes the details unreadable, we demonstrate the steps using $p = 2$. Let $E_j(\boldsymbol{\alpha}) = C_j - \alpha_j^T \mathbf{X}$ and $S_j(\boldsymbol{\alpha}) = T_j - \alpha_j^T \mathbf{X}$, for $j = 1, 2$. Define

$$
\begin{aligned}
A_j(s_j, \theta, \phi, \boldsymbol{\alpha}) &= f_Y(Y | s_j + \alpha_j^T \mathbf{X}, T_{3-j}, \mathbf{X}) \big[Y - \dot{b}\{\theta^T D_{(\mathbf{X}, s_j + \alpha_j^T \mathbf{X}, T_{3-j})}\}\big] D_{(\mathbf{X}, s_j + \alpha_j^T \mathbf{X}, T_{3-j})}, \\
B_j(s_j, \theta, \phi, \boldsymbol{\alpha}) &= f_Y(Y | s_j + \alpha_j^T \mathbf{X}, T_{3-j}, \mathbf{X}), \text{ for } j = 1, 2,
\end{aligned}
$$

and

$$
\begin{aligned}
A_3(s_1, s_2, \theta, \phi, \boldsymbol{\alpha}) &= f_Y(Y | s_j + \alpha_j^T \mathbf{X}, j = 1, 2, \mathbf{X}) \big[Y - \dot{b}\{\theta^T D_{(\mathbf{X}, s_j + \alpha_j^T \mathbf{X}, j=1,2)}\}\big] D_{(\mathbf{X}, s_j + \alpha_j^T \mathbf{X}, j=1,2)}, \\
B_3(s_1, s_2, \theta, \phi, \boldsymbol{\alpha}) &= f_Y(Y | s_j + \alpha_j^T \mathbf{X}, j = 1, 2, \mathbf{X}).
\end{aligned}
$$

Next, we show that

$$
\begin{aligned}
&\sup_{\theta\in\Theta, \nu\{(\phi,\boldsymbol{\alpha},\eta),(\phi_0,\boldsymbol{\alpha}_0,\eta_0)\}\le\delta_n} |\Psi_n(\theta,\phi,\boldsymbol{\alpha},\eta) - \Psi(\theta,\phi_0,\boldsymbol{\alpha}_0,\eta_0)| \\
&\le \sup_{\theta\in\Theta} |(\mathbb{P}_n - P)[\Delta_1\Delta_2\{Y - \dot{b}(\theta^T D_{(\mathbf{X},\mathbf{T})})\} D_{(\mathbf{X},\mathbf{T})}]| \qquad \text{(S.11)} \\
&+ \sum_{j=1}^{2}\Bigg[\sup_{\theta\in\Theta, \nu\{(\phi,\boldsymbol{\alpha},\eta),(\phi_0,\boldsymbol{\alpha}_0,\eta_0)\}\le\delta_n}\left|(\mathbb{P}_n - P)(1-\Delta_j)\Delta_{3-j}\frac{\int_{E_j(\boldsymbol{\alpha})}^{\tau_j} A_j(s_j,\theta,\phi,\boldsymbol{\alpha})\eta\{ds_j, ds_{3-j} = S_{3-j}(\boldsymbol{\alpha})\}}{\int_{E_j(\boldsymbol{\alpha})}^{\tau_j} B_j(s_j,\theta,\phi,\boldsymbol{\alpha})\eta\{ds_j, ds_{3-j} = S_{3-j}(\boldsymbol{\alpha})\}}\right| \qquad \text{(S.12)} \\
&\quad + \sup_{\theta\in\Theta, \nu\{(\phi,\boldsymbol{\alpha},\eta),(\phi_0,\boldsymbol{\alpha}_0,\eta_0)\}\le\delta_n} P\left|\frac{\int_{E_j(\boldsymbol{\alpha})}^{\tau_j} A_j(s_j,\theta,\phi,\boldsymbol{\alpha})\eta\{ds_j, ds_{3-j} = S_{3-j}(\boldsymbol{\alpha})\}}{\int_{E_j(\boldsymbol{\alpha})}^{\tau_j} B_j(s_j,\theta,\phi,\boldsymbol{\alpha})\eta\{ds_j, ds_{3-j} = S_{3-j}(\boldsymbol{\alpha})\}}\right. \qquad \text{(S.13)} \\
&\qquad \left. - \frac{\int_{E_j(\boldsymbol{\alpha}_0)}^{\tau_j} A_j(s_j,\theta,\phi,\boldsymbol{\alpha}_0)\eta_0\{ds_j, ds_{3-j} = S_{3-j}(\boldsymbol{\alpha}_0)\}}{\int_{E_j(\boldsymbol{\alpha}_0)}^{\tau_j} B_j(s_j,\theta,\phi,\boldsymbol{\alpha}_0)\eta_0\{ds_1, ds_{3-j} = S_{3-j}(\boldsymbol{\alpha}_0)\}}\right|\Bigg] \\
&+ \sup_{\theta\in\Theta, \nu\{(\phi,\boldsymbol{\alpha},\eta),(\phi_0,\boldsymbol{\alpha}_0,\eta_0)\}\le\delta_n}\left|(\mathbb{P}_n - P)(1-\Delta_1)(1-\Delta_2)\frac{\int_{E_1(\boldsymbol{\alpha})}^{\tau_1}\int_{E_2(\boldsymbol{\alpha})}^{\tau_1} A_3(s_1,s_2,\theta,\phi,\boldsymbol{\alpha})\eta(ds_1,ds_2)}{\int_{E_1(\boldsymbol{\alpha})}^{\tau_1}\int_{E_2(\boldsymbol{\alpha})}^{\tau_2} B_3(s_1,s_2,\theta,\phi,\boldsymbol{\alpha})\eta(ds_1,ds_2)}\right| \qquad \text{(S.14)} \\
&+ \sup_{\theta\in\Theta, \nu\{(\phi,\boldsymbol{\alpha},\eta),(\phi_0,\boldsymbol{\alpha}_0,\eta_0)\}\le\delta_n} P\left|\frac{\int_{E_1(\boldsymbol{\alpha})}^{\tau_1}\int_{E_2(\boldsymbol{\alpha})}^{\tau_2} A_3(s_1,s_2,\theta,\phi,\boldsymbol{\alpha})\eta(ds_1,ds_2)}{\int_{E_1(\boldsymbol{\alpha})}^{\tau_1}\int_{E_2(\boldsymbol{\alpha})}^{\tau_2} B_3(s_1,s_2,\theta,\phi,\boldsymbol{\alpha})\eta(ds_1,ds_2)}\right. \qquad \text{(S.15)} \\
&\qquad \left. - \frac{\int_{E_1(\boldsymbol{\alpha}_0)}^{\tau_1}\int_{E_2(\boldsymbol{\alpha}_0)}^{\tau_2} A_3(s_1,s_2,\theta,\phi,\boldsymbol{\alpha}_0)\eta_0(ds_1,ds_2)}{\int_{E_1(\boldsymbol{\alpha}_0)}^{\tau_1}\int_{E_2(\boldsymbol{\alpha}_0)}^{\tau_2} B_3(s_1,s_2,\theta,\phi,\boldsymbol{\alpha}_0)\eta_0(ds_1,ds_2)}\right|
\end{aligned}
$$

The equation (S.11) is equal to $o_{p^*}(1)$ by Lemma 1, and equations (S.12) and (S.14) are equal to $o_{p^*}(1)$ by Lemma 3. Using the Mean Value Theorem, we obtain that the equations (S.13) and (S.15) are equal to $o_{p^*}(1)$. Combining these results gives the equation (S.10) which completes the proof of consistency. The extension of the proof to arbitrary p has similar steps but requires more calculations to obtain equation (S.10).

3 | PROOF OF ASYMPTOTIC NORMALITY

We assume $p = 2$ for the brevity of notation. The proof of asymptotic normality is based on the result in Lemma 7. Thus, Conditions i-iv for Lemma 7 are needed to be verified. Under the regularity conditions, using Lemmas 1, 3 and 4, Condition i. in Lemma 7 holds. By the central limit theorem for i.i.d. data with $E|\psi(\theta_0,\phi_0,\boldsymbol{\alpha}_0,\eta_0)|^2 < \infty$ in Lemma 5, we then have Condition ii. Furthermore, we use Taylor expansion to show Condition iii. Let $\dot{\Psi}_1(\theta,\phi,\boldsymbol{\alpha},\eta) = E\{\partial\psi(Y,\mathbf{X},\mathbf{V},\boldsymbol{\Delta};\theta,\phi,\boldsymbol{\alpha},\eta)/\partial\theta\}$, $\dot{\Psi}_2(\theta,\phi,\boldsymbol{\alpha},\eta) = E\{\partial\psi(Y,\mathbf{X},\mathbf{V},\boldsymbol{\Delta};\theta,\phi,\boldsymbol{\alpha},\eta)/\partial\phi\}$, and $\dot{\Psi}_3$ $(\theta,\phi,\boldsymbol{\alpha},\eta) = E\{\partial\psi(Y,\mathbf{X},\mathbf{V},\boldsymbol{\Delta};\theta,\phi,\boldsymbol{\alpha},\eta)/\partial\boldsymbol{\alpha}\}$. Given that $\nu\{(\hat{\phi},\hat{\boldsymbol{\alpha}},\hat{\eta}_{\hat{\boldsymbol{\alpha}}}),(\phi_0,\boldsymbol{\alpha}_0,\eta_0)\} = O_{p^*}$ $(n^{-1/2})$, we take Taylor expansion of Ψ_n for θ, ϕ and $\boldsymbol{\alpha}$ to show that

$$
\begin{aligned}
\Psi_n(\hat{\theta},\hat{\phi},\hat{\boldsymbol{\alpha}},\hat{\eta}_{\hat{\boldsymbol{\alpha}}}) - \Psi(\theta_0,\phi_0,\boldsymbol{\alpha}_0,\eta_0) &= \dot{\Psi}_1(\tilde{\theta},\hat{\phi},\hat{\boldsymbol{\alpha}},\hat{\eta}_{\hat{\boldsymbol{\alpha}}})(\hat{\theta}-\theta_0) - \dot{\Psi}_2(\hat{\theta},\tilde{\phi},\hat{\boldsymbol{\alpha}},\hat{\eta}_{\hat{\boldsymbol{\alpha}}})(\hat{\phi}-\phi_0) \\
&\quad - \dot{\Psi}_3(\hat{\theta},\hat{\phi},\tilde{\boldsymbol{\alpha}},\hat{\eta}_{\hat{\boldsymbol{\alpha}}})(\hat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}_0) - R(\theta_0,\phi_0,\hat{\boldsymbol{\alpha}},\hat{\eta}_{\hat{\boldsymbol{\alpha}}},\eta_0),
\end{aligned}
$$

where $\tilde{\theta}$, $\tilde{\phi}$ and $\tilde{\boldsymbol{\alpha}}$ are the line segments of $(\theta_0,\hat{\theta})$, $(\phi_0,\hat{\phi})$ and $(\boldsymbol{\alpha}_0,\hat{\boldsymbol{\alpha}})$, and $R(\theta_0,\phi_0,\hat{\boldsymbol{\alpha}},\hat{\eta}_{\hat{\boldsymbol{\alpha}}},\eta_0)$ is the remainder term. We then can show that $|\dot{\Psi}_1(\tilde{\theta},\hat{\phi},\hat{\boldsymbol{\alpha}},\hat{\eta}_{\hat{\boldsymbol{\alpha}}}) - \dot{\Psi}_1(\theta_0,\phi_0,\boldsymbol{\alpha}_0,\eta_0)| = o_{p^*}(1)$, $|\dot{\Psi}_2(\hat{\theta},\tilde{\phi},\hat{\boldsymbol{\alpha}},\hat{\eta}_{\hat{\boldsymbol{\alpha}}}) - \dot{\Psi}_2(\theta_0,\phi_0,\boldsymbol{\alpha}_0,\eta_0)| = o_{p^*}(1)$ and $|\dot{\Psi}_3(\hat{\theta},\hat{\phi},\tilde{\boldsymbol{\alpha}},\hat{\eta}_{\hat{\boldsymbol{\alpha}}}) -$

$\dot{\Psi}_3(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)| = o_{p^*}(1)$ by direct calculations. Now, we focus on the remainder term which is written as

$$
\begin{aligned}
&R(\theta_0, \phi_0, \boldsymbol{\alpha}, \eta, \eta_0)\\
&= \sum_{j=1}^{2} P\Bigg[(1-\Delta_j)\Delta_{3-j}\Bigg\{ \frac{\int_{E_j(\boldsymbol{\alpha})}^{\tau_j} A_j(s_j, \theta_0, \phi_0, \boldsymbol{\alpha})\eta\{ds_j, ds_{3-j} = S_{3-j}(\boldsymbol{\alpha})\}}{\int_{E_j(\boldsymbol{\alpha})}^{\tau_j} B_j(s_j, \theta_0, \phi_0, \boldsymbol{\alpha})\eta\{ds_j, ds_{3-j} = S_{3-j}(\boldsymbol{\alpha})\}} \\
&\qquad - \frac{\int_{E_j(\boldsymbol{\alpha})}^{\tau_j} A_j(s_j, \theta_0, \phi_0, \boldsymbol{\alpha})\eta_0\{ds_j, ds_{3-j} = S_{3-j}(\boldsymbol{\alpha})\}}{\int_{E_j(\boldsymbol{\alpha})}^{\tau_j} B_j(s_j, \theta_0, \phi_0, \boldsymbol{\alpha})\eta_0\{ds_j, ds_{3-j} = S_{3-j}(\boldsymbol{\alpha})\}}\Bigg\}\Bigg] \\
&+ P\Bigg[(1-\Delta_1)(1-\Delta_2)\Bigg\{ \frac{\int_{E_1(\boldsymbol{\alpha})}^{\tau_1}\int_{E_2(\boldsymbol{\alpha})}^{\tau_2} A_3(s_1, s_2, \theta_0, \phi_0, \boldsymbol{\alpha})\eta(ds_1, ds_2)}{\int_{E_1(\boldsymbol{\alpha})}^{\tau_1}\int_{E_2(\boldsymbol{\alpha})}^{\tau_2} B_3(s_1, s_2, \theta_0, \phi_0, \boldsymbol{\alpha})\eta(ds_1, ds_2)} \\
&\qquad - \frac{\int_{E_1(\boldsymbol{\alpha})}^{\tau_1}\int_{E_2(\boldsymbol{\alpha})}^{\tau_2} A_3(s_1, s_2, \theta_0, \phi_0, \boldsymbol{\alpha})\eta_0(ds_1, ds_2)}{\int_{E_1(\boldsymbol{\alpha})}^{\tau_1}\int_{E_2(\boldsymbol{\alpha})}^{\tau_2} B_3(s_1, s_2, \theta_0, \phi_0, \boldsymbol{\alpha})\eta_0(ds_1, ds_2)}\Bigg\}\Bigg],
\end{aligned}
$$

where $E_j(\boldsymbol{\alpha}) = C_j - \alpha_j^T\mathbf{X}$ and $S_j(\boldsymbol{\alpha}) = T_j - \alpha_j^T\mathbf{X}$, for $j = 1, 2$.

We further define

$$
\begin{aligned}
&\dot{\Psi}_4(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\eta}_{\hat{\boldsymbol{\alpha}}} - \eta_0) \qquad \text{(S.16)}\\
&= \sum_{j=1}^{2} P\Bigg[(1-\Delta_j)\Delta_{3-j}\Bigg\{ \frac{\int_{E_j(\boldsymbol{\alpha}_0)}^{\tau_j} A_j(s_j, \theta_0, \phi_0, \boldsymbol{\alpha}_0)[\hat{\eta}_{\hat{\boldsymbol{\alpha}}}\{ds_j, ds_{3-j} = S_{3-j}(\hat{\boldsymbol{\alpha}})\} - \eta_0(ds_j, ds_{3-j})]}{\int_{E_j(\boldsymbol{\alpha}_0)}^{\tau_j} B_j(s_j, \theta_0, \phi_0, \boldsymbol{\alpha}_0)\eta_0(ds_j, ds_{3-j})} \\
&- \frac{\int_{E_j(\boldsymbol{\alpha}_0)}^{\tau_j} A_j(s_j, \theta_0, \phi_0, \boldsymbol{\alpha}_0)\eta_0(ds_j, s_{3-j})\int_{E_j(\boldsymbol{\alpha}_0)}^{\tau_j} B_j(s_j, \theta_0, \phi_0, \boldsymbol{\alpha}_0)[\hat{\eta}_{\hat{\boldsymbol{\alpha}}}\{ds_j, ds_{3-j} = S_{3-j}(\hat{\boldsymbol{\alpha}})\} - \eta_0(ds_j, ds_{3-j})]}{\left\{\int_{E_j(\boldsymbol{\alpha}_0)}^{\tau_j} B_j(s_j, \theta_0, \phi_0, \boldsymbol{\alpha}_0)\eta_0(ds_j, ds_{3-j})\right\}^2}\Bigg\}\Bigg] \\
&+ P\Bigg[(1-\Delta_1)(1-\Delta_2)\Bigg\{ \frac{\int_{E_1(\boldsymbol{\alpha}_0)}^{\tau_1}\int_{E_2(\boldsymbol{\alpha}_0)}^{\tau_2} A_3(s_1, s_2, \theta_0, \phi_0, \boldsymbol{\alpha}_0)\{\hat{\eta}_{\hat{\boldsymbol{\alpha}}}(ds_1, ds_2) - \eta_0(ds_1, ds_2)\}}{\int_{E_1(\boldsymbol{\alpha}_0)}^{\tau_1}\int_{E_2(\boldsymbol{\alpha}_0)}^{\tau_2} B_3(s_1, s_2, \theta_0, \phi_0, \boldsymbol{\alpha}_0)\eta_0(ds_1, ds_2)} \\
&- \frac{\int_{E_1(\boldsymbol{\alpha}_0)}^{\tau_1}\int_{E_2(\boldsymbol{\alpha}_0)}^{\tau_2} A_3(s_1, s_2, \theta_0, \phi_0, \boldsymbol{\alpha}_0)\eta_0(ds_1, ds_2)\int_{E_1(\boldsymbol{\alpha}_0)}^{\tau_1}\int_{E_2(\boldsymbol{\alpha}_0)}^{\tau_2} B_3(s_1, s_2, \theta_0, \phi_0, \boldsymbol{\alpha}_0)\{\hat{\eta}_{\hat{\boldsymbol{\alpha}}}(ds_1, ds_2) - \eta_0(ds_1, ds_2)\}}{\left\{\int_{E_1(\boldsymbol{\alpha}_0)}^{\tau_1}\int_{E_2(\boldsymbol{\alpha}_0)}^{\tau_2} B_3(s_1, s_2, \theta_0, \phi_0, \boldsymbol{\alpha}_0)\eta_0(ds_1, ds_2)\right\}^2}\Bigg\}\Bigg],
\end{aligned}
$$

and have that

$$
\begin{aligned}
&|R(\theta_0, \phi_0, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}, \eta_0) - \dot{\Psi}_4(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\eta}_{\hat{\boldsymbol{\alpha}}} - \eta_0)| \\
&= |R(\theta_0, \phi_0, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}, \eta_0) - R(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}, \eta_0) + R(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}, \eta_0) - \dot{\Psi}_4(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\eta}_{\hat{\boldsymbol{\alpha}}} - \eta_0)| \\
&\le |R(\theta_0, \phi_0, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}, \eta_0) - R(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}, \eta_0)| + |R(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}, \eta_0) - \dot{\Psi}_4(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\eta}_{\hat{\boldsymbol{\alpha}}} - \eta_0)|
\end{aligned}
$$

After a straightforward and tedious calculation, it can be shown that $|R(\theta_0, \phi_0, \hat{\boldsymbol{\alpha}}, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}, \eta_0) - R(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}, \eta_0)|$ and $|R(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \hat{\eta}_{\hat{\boldsymbol{\alpha}}}, \eta_0) - \dot{\Psi}_4(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\eta}_{\hat{\boldsymbol{\alpha}}} - \eta_0)|$ both are $o(|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0| + ||\hat{\eta}_{\hat{\boldsymbol{\alpha}}} - \eta_0||)$ (see an example for the unvariate censoring in Kong and Nan[2]). These results give Condition iii. in Lemma 7. Finally, under Conditions i.-iii. and the result where $\hat{\psi}$, $\hat{\boldsymbol{\alpha}}$ and $\hat{\eta}_{\hat{\boldsymbol{\alpha}}}$ are $n^{1/2}$-consistent estimates, Condition iv. holds.

Having verifying Conditions i.-iv. in Lemma 7, we obtain that $\hat{\theta}$ is $n^{1/2}$−consistent estimate and

$$
\begin{aligned}
n^{1/2}(\hat{\theta} - \theta_0) =& \{-\dot{\Psi}_1(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)\}^{-1} n^{1/2}\{(\Psi_n - \Psi)(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0) + \dot{\Psi}_2(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\phi} - \phi_0) \\
&+ \dot{\Psi}_3(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0) + \dot{\Psi}_4(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\eta}_{\hat{\boldsymbol{\alpha}}} - \eta_0)\} + o_{p^*}(1),
\end{aligned}
$$

where $\dot{\Psi}_1(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0) = E\{\partial\psi(Y, \mathbf{X}, \mathbf{V}, \boldsymbol{\Delta}; \theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)/\partial\theta_0\}$, $\dot{\Psi}_2(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0) = E\{\partial\psi(Y, \mathbf{X}, \mathbf{V}, \boldsymbol{\Delta}; \theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)/\partial\phi_0\}$, $\dot{\Psi}_3(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0) = E\{\partial\psi(Y, \mathbf{X}, \mathbf{V}, \boldsymbol{\Delta}; \theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)/\partial\boldsymbol{\alpha}_0\}$ and $\dot{\Psi}_4(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\eta}_{\hat{\boldsymbol{\alpha}}} - \eta_0)$ is given in (S.16). Since the asymptotic distribution for $n^{1/2}(\hat{\phi} - \phi_0)$ and $n^{1/2}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0)$ are known in the standard generalized linear model and semiparametric AFT models, we can define that there exists functions m_2 and m_3 such that $n^{1/2}(\hat{\phi} - \phi_0) = \mathbb{G}_n m_2(\theta_0, \phi_0, Y, \mathbf{X}, \mathbf{V}, \boldsymbol{\Delta}) + o_p(1)$ and $n^{1/2}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0) = \mathbb{G}_n m_3(\boldsymbol{\alpha}_0, \mathbf{X}, \mathbf{V}, \boldsymbol{\Delta}) + o_p(1)$. Thus, under regularity conditions, we have

$$
n^{1/2}\{(\Psi_n - \Psi)(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0) = \mathbb{G}_n\{\psi(Y, \mathbf{X}, \mathbf{V}, \boldsymbol{\Delta}; \theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)\},
$$

$$
n^{1/2}\{\dot{\Psi}_2(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)(\hat{\phi} - \phi_0)\} = \mathbb{G}_n\{\dot{\Psi}_2(\theta_0, \phi_0, \boldsymbol{\alpha}_0, \eta_0)m_2(\theta_0, \phi_0, Y, \mathbf{X}, \mathbf{V}, \boldsymbol{\Delta})\},
$$

$$n^{1/2}\{\dot{\Psi}_3(\theta_0,\phi_0,\boldsymbol{\alpha}_0,\eta_0)(\hat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}_0)\} = \mathbb{G}_n\{\dot{\Psi}_3(\theta_0,\phi_0,\boldsymbol{\alpha}_0,\eta_0)m_3(\boldsymbol{\alpha}_0,\mathbf{X},\mathbf{V},\boldsymbol{\Delta})\},$$

and

$$n^{1/2}\{\dot{\Psi}_4(\theta_0,\phi_0,\boldsymbol{\alpha}_0,\eta_0)(\hat{\eta}_{\hat{\alpha}}-\eta_0)\} = \mathbb{G}_n\{\dot{\Psi}_4(\boldsymbol{\alpha}_0,\eta_0)m_1(\boldsymbol{\alpha}_0,\mathbf{X},\mathbf{V},\boldsymbol{\Delta})\}.$$

Hence, by the central limit theorem, we obtain that the asymptotic distribution of $n^{1/2}(\hat{\theta}-\theta_0)$ is normally distributed with mean-zero and the covariance matrix $\Omega = \mathbf{J}^{-1}\Sigma\mathbf{J}^{-1}$, where $\mathbf{J} = -\dot{\Psi}_1(\theta_0,\phi_0,\boldsymbol{\alpha}_0,\eta_0)$ and

$$\begin{aligned}\Sigma =& \big\{\psi(Y,\mathbf{X},\mathbf{V},\boldsymbol{\Delta};\theta_0,\phi_0,\boldsymbol{\alpha}_0,\eta_0) + \dot{\Psi}_2(\theta_0,\phi_0,\boldsymbol{\alpha}_0,\eta_0)m_2(\theta_0,\phi_0,Y,\mathbf{X},\mathbf{V},\boldsymbol{\Delta})\\ &+ \dot{\Psi}_3(\theta_0,\phi_0,\boldsymbol{\alpha}_0,\eta_0)m_3(\boldsymbol{\alpha}_0,\mathbf{X},\mathbf{V},\boldsymbol{\Delta}) + \dot{\Psi}_4(\boldsymbol{\alpha}_0,\eta_0)m_1(\boldsymbol{\alpha}_0,\mathbf{X},\mathbf{V},\boldsymbol{\Delta})\big\}^{\otimes}.\end{aligned}$$

The proof for any arbitrary p can be shown by modifying the Taylor expansion where the remainder term R will involve 2^p-1 equations. The rest steps are similar to $p=2$.

4 | ADDITIONAL SIMULATION RESULTS AND TABLES FOR DATA ANALYSIS

4.1 | Simulation results for p=2

Let $X_1 \sim Ber(0.5)$ and $X_2 \sim N(1,1)$ be two fully observed covariates, and $\mathbf{X} = (1, X_1, X_2)^T$. We generated two covariates subject to LOD, $\mathbf{Z} = (Z_1, Z_2)$ with $Z_1 = h(T_1)$, $Z_2 = h(T_2)$, where

$$T_j = h^{-1}(Z_j) = \alpha_j^T\mathbf{X} + \xi_j, \text{ for } j = 1, 2$$

and $\alpha_1 = (-0.25, -0.5, -0.25)^T$, $\alpha_2 = (-0.25, -0.25, -0.5)^T$ and the joint residuals $(\xi_1, \xi_2)^T$ follows a bivariate distribution η. The outcome of interest Y was generated by $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \gamma_1 Z_1 + \gamma_2 Z_2 + \epsilon$, where $\beta_0 = \beta_1 = \beta_2 = \gamma_1 = \gamma_2 = 1$ and $\epsilon \sim N(0,1)$. We considered two transformation functions $h(\cdot)$ and two joint distributions η: $z = h(t) = -t$ (i.e. $T_j = -Z_j$) or $z = h(t) = \exp(-t)$ (i.e. $T_j = -\log(Z_j)$), and $\eta = \text{MVN}\{(0,0)^T, \Sigma_1\}$ (multivariate normal) or $\eta = 0.5\text{MVN}\{(0,0)^T, \Sigma_1\} + 0.5\text{MVN}\{(0,0)^T, \Sigma_2\}$ (a mixture of multivariate normals), where

$$\Sigma_1 = 1/4^2\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix} \quad \text{and} \quad \Sigma_2 = 1/8^2\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}$$

and $\rho = 0.5$. We generated samples with size 200 or 400, and repeated 1000 times at marginal censoring rate 25% or 50% for $j = 1, 2$. The results for $T = -\log(Z)$ and the multivariate normal as the error term distribution are given Table 1 (main paper). The results for $T = -Z$ and the multivariate normal or a mixture of multivariate normals as the error term distribution are given in Tables S1 and S2, and the results for $T = -\log(Z)$ and a mixture of multivariate normals as the error term distribution are given in Table S3. We further explored the property of the proposed method with small sample sizes, 50 or 100. The results are given in Table S4.

4.2 | Simulation results for p=10 with $h(T_j) = \exp(-T_j)$

Considering two fully observed covariates $\mathbf{X} = (1, X_1, X_2)^T$ where $X_1 \sim Ber(0.5)$ and $X_2 \sim N(1,1)$, we generated ten left-censored variables $\mathbf{Z} = (Z_1, ..., Z_{10})^T$ with $Z_j = h(T_j) = \exp(-T_j)$, $j = 1, ..., 10$, by

$$\mathbf{T} = \boldsymbol{\alpha}^T\mathbf{X} + \xi,$$

where $\mathbf{T} = (T_1, ..., T_{10})^T$,

$$\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_{10}] = \begin{bmatrix} 0.25 & 0.5 & 0.5 & 0.5 & 0.5 & 0.25 & 0.35 & 0.5 & 0.35 & 0.25\\ -0.5 & -0.25 & -0.5 & -0.25 & -0.5 & -0.25 & -0.25 & -0.25 & -0.25 & -0.25\\ -0.25 & -0.5 & -0.25 & -0.25 & -0.25 & -0.5 & -0.25 & -0.5 & -0.5 & -0.25 \end{bmatrix},$$

and $(\xi_1, \dots, \xi_{10})^T \sim \eta$ =MVN$\{(0, ..., 0)^T, \Sigma\}$ with

$$\Sigma = 1/2^2\begin{pmatrix}\mathbf{R}_1 & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{R}_3\end{pmatrix}.$$

Here $\mathbf{R}_1$ and $\mathbf{R}_3$ are 3×3 correlation matrices with all off-diagonal entries as 0.25 and 0.75 respectively, and $\mathbf{R}_2$ is a 4×4 correlation matrix with all off-diagonal entries as 0.5. We let $Y = \beta^T \mathbf{X} + \gamma^T \mathbf{Z} + \epsilon$, where all elements in β and γ were set as 1 and $\epsilon \sim N(0, 1)$. The marginal censoring rate was set as 20% for all Z_j with the overall censoring rate around 70%. We generated data with sample size 400 and repeated the simulation 1000 times.

For each simulated dataset, we implemented six methods: analysis with the full data, complete-case approach, substitution methods with three replacement values (LOD, LOD/$\sqrt{2}$, LOD/2), and our proposed two-stage approach with marginal approximation. For our proposed two-stage approach with marginal approximation, the MC integration was applied to the estimation, and the standard deviations were estimated using 200 bootstrap replicates each with sample size 100, and then adjusted by a factor of 2. The results are shown in Table S5.

4.3 | Simulation results to mimic real data example

To further explore the performance in the data example, we considered 17 metals and 7 fully observed covariates, including baseline maternal age, race, education, insurance, pre-pregnancy BMI, and gestational age and specific gravity at the third trimester visit. We bootstrapped the demographic covariates from the LIFECODES cohort with replacement (n=252), and generated 17 metals based on the AFT models. The true regression coefficients in both the outcome regression and the AFT models were set to be the estiamtes from the real data analysis. and the residual terms of the AFT models were from a multivariate normal distribution whose mean and covariance matched the distribution of the AFT model residuals using complete-cases of the LIFECODES data.

The censoring rates were set to be similar to the real data: three metals were fully observed, three metals had LOD rate <5%, seven metals had LOD rate within 5-70%, and four metals had LOD rate >70%, where the LOD value for each metal was set to have the similar percentage of values below LOD for the metal in the real data, by generating the corresponding metals with a massive sample size and calculating the percentiles. We repeated the simulation 1000 times and the results were given in Table S6. Five methods were implemented including complete-case approach, substitution method with three substitution values: LOD, LOD/$\sqrt{2}$, LOD/2, and the proposed two-stage approach with marginal approximation for $p = 7$ (2-stage(marg)-p7) where we pre-processed the metals with less than 5% of values below LOD substituted with LOD/$\sqrt{2}$ and the metals with more than 70% below LOD as dichotomized variable. The estimates were consistent with the true parameters, except for the metals with heavy censoring.This was expected because using binary indicator of above LOD or not led to a different interpretation of the coefficients. However the directions of these effects were still the same. The two-stage marginal approach and the substitution methods with LOD and LOD/$\sqrt{2}$ as substitution values had coverage rates close to the nominal level which confirmed our data analysis result. In addition, the substitution method with LOD/2 as substitution value had slightly lower coverage rate in some metals which implied that its performance was sensitive to the choice of substitution values. We further increased the sample size to n=400 and the conclusions remained the same. Thus, using the proposed method only for variables with LOD rate within 5-70% showed satisfying performance while greatly reducing computational time.

4.4 | Simulation results for binary outcomes

We present one example for binary outcomes with $p = 2$. Let $X_1 \sim Ber(0.5)$ and $X_2 \sim N(1, 1)$ be two fully observed covariates and $\mathbf{X} = (1, X_1, X_2)^T$. We generated two covariates subject to LOD, $\mathbf{Z} = (Z_1, Z_2)^T$, by

$$-\log(Z_j) = T_j = \alpha_j^T \mathbf{X} + \xi_j, \text{ for } j = 1, 2,$$

where $\alpha_1 = (0.15, -0.25, -0.45)^T$, $\alpha_2 = (-0.25, 0.55, -0.55)^T$ and $(\xi_1, \xi_2)^T$ follows a bivariate normal distribution with zero means and covariance Σ, where

$$\Sigma = 1/4^2 \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

The outcome of interest Y was generated from a Bernoulli distribution with $\mu = E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \gamma_1 Z_1 + \gamma_2 Z_2$, where $\beta_0 = \beta_2 = \gamma_1 = 1$ and $\beta_1 = \gamma_2 = -1$. The LOD_j was chosen to have 25% or 50% marginal censoring rate for $j = 1, 2$ with the overall censoring rate around 36% or 65%, respectively. We generated samples with size 400 and repeated 1000 times. For each simulated dataset, we implemented eight methods: analysis with the full data, complete-case approach, substitution methods with three replacement values (LOD, LOD/$\sqrt{2}$, LOD/2) and three versions of our proposed two-stage approach. The results are shown in Table S7.

4.5 | Additional results for data analysis

A summary of each metal concentration with LOD value was given in Table S8, and a summary of each demographic variable was given in Table S9. A heat map for the 13 log-transformed metal concentrations with 0-70% below LOD was provided in Figure S1. In addition, a heat map for the estimated pairwise correlations between the residuals in the 7 AFT models was shown in Figure S2.

4.6 | Computing time for two AFT methods

A summary of computing time using two AFT methods, rank-based approach and least-square approach, was given in Table S10. In this comparison, q=2, p=2, 3, 5 and 10, and n=100, 200, 400, 600 and 1000.

References

1. van der Vaart A, Wellner JA. *Weak convergence and empirical processes: with applications to statistics*. Springer . 1996.

2. Kong S, Nan B. Semiparametric approach to regression with a covariate subject to a detection limit. *Biometrika* 2016; 103(1): 161–174.

3. Nan B, Wellner JA. A general semiparametric Z-estimation approach for case-cohort studies. *Statistica Sinica* 2013; 23(3): 1155.

TABLE S1 Simulation results for linear outcome model where $p = q = 2$, $T = -Z$ and the multivariate normal as the error term distribution. Eight methods were implemented, including analysis with full data, complete-case approach, substitution methods with three replacement values: LOD, LOD/$\sqrt{2}$ and LOD/2, two-stage rank-based approach with multivariate KM estimator for η (2-stage(rank)), two-stage least-square approach with multivariate KM estimator for η (2-stage(LS)), and two-stage approach with marginal approximation (2-stage(marg)). Estimated S.D.* was from bootstrap for three version of the proposed method and from regression for full data analysis, complete-case approach and substitution methods.

			Bias					Empirical S.E.					Mean estimated S.D.*					Coverage rate of 95% CIs				
%<LOD	n	Method	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2
25%	200	full data	-0.005	-0.005	-0.016	0.005	0.021	0.150	0.207	0.164	0.349	0.342	0.149	0.202	0.161	0.331	0.331	0.943	0.942	0.945	0.933	0.947
		complete cases	-0.005	-0.006	-0.016	0.014	0.015	0.310	0.245	0.201	0.460	0.437	0.309	0.241	0.200	0.445	0.432	0.941	0.950	0.945	0.940	0.953
		LOD	-0.350	0.171	0.252	0.116	-0.239	0.216	0.206	0.145	0.414	0.365	0.215	0.199	0.138	0.396	0.359	0.635	0.841	0.528	0.927	0.893
		LOD/$\sqrt{2}$	-0.198	0.109	0.184	0.074	-0.192	0.186	0.209	0.151	0.381	0.349	0.184	0.202	0.144	0.363	0.341	0.821	0.903	0.734	0.930	0.907
		LOD/2	-0.070	0.088	0.157	0.011	-0.190	0.168	0.209	0.153	0.353	0.331	0.165	0.202	0.146	0.335	0.322	0.923	0.918	0.799	0.933	0.904
		2-stage(rank)	0.001	-0.012	-0.021	0.024	0.012	0.161	0.217	0.176	0.375	0.364	0.162	0.213	0.174	0.360	0.361	0.945	0.947	0.944	0.936	0.951
		2-stage(LS)	0.001	-0.011	-0.021	0.023	0.014	0.160	0.217	0.176	0.376	0.364	0.162	0.213	0.175	0.360	0.362	0.943	0.945	0.943	0.934	0.950
		2-stage(marg)	-0.019	-0.027	-0.029	0.051	0.025	0.163	0.218	0.176	0.375	0.362	0.163	0.213	0.174	0.355	0.357	0.946	0.938	0.943	0.925	0.943
	400	full data	0.003	-0.006	-0.004	0.008	-0.001	0.104	0.142	0.114	0.228	0.236	0.105	0.142	0.113	0.233	0.233	0.951	0.941	0.939	0.949	0.944
		complete cases	-0.011	-0.007	-0.007	0.016	0.006	0.212	0.169	0.142	0.304	0.304	0.216	0.170	0.140	0.312	0.302	0.957	0.942	0.943	0.959	0.952
		LOD	-0.341	0.170	0.257	0.112	-0.250	0.147	0.139	0.101	0.272	0.255	0.151	0.140	0.097	0.278	0.252	0.397	0.764	0.232	0.936	0.840
		LOD/$\sqrt{2}$	-0.189	0.108	0.192	0.072	-0.207	0.128	0.141	0.105	0.248	0.241	0.129	0.142	0.101	0.255	0.240	0.683	0.886	0.506	0.942	0.866
		LOD/2	-0.061	0.088	0.166	0.010	-0.207	0.117	0.142	0.106	0.229	0.228	0.116	0.143	0.102	0.236	0.227	0.907	0.907	0.630	0.951	0.855
		2-stage(rank)	0.008	-0.013	-0.012	0.022	-0.001	0.114	0.146	0.122	0.240	0.253	0.113	0.149	0.121	0.252	0.253	0.947	0.950	0.938	0.958	0.942
		2-stage(LS)	0.008	-0.013	-0.011	0.022	-0.003	0.114	0.146	0.122	0.241	0.253	0.113	0.149	0.121	0.252	0.253	0.947	0.947	0.937	0.954	0.942
		2-stage(marg)	-0.011	-0.029	-0.019	0.048	0.010	0.114	0.147	0.122	0.238	0.251	0.114	0.149	0.121	0.247	0.249	0.950	0.949	0.936	0.957	0.939
50%	200	full data	-0.005	-0.005	-0.016	0.005	0.021	0.150	0.207	0.164	0.349	0.342	0.149	0.202	0.161	0.331	0.331	0.943	0.942	0.945	0.933	0.947
		complete cases	-0.024	-0.021	-0.025	0.035	0.029	0.597	0.347	0.264	0.638	0.583	0.586	0.326	0.266	0.626	0.597	0.943	0.928	0.951	0.943	0.958
		LOD	-0.721	0.394	0.464	0.156	-0.442	0.352	0.188	0.120	0.501	0.412	0.353	0.184	0.117	0.493	0.413	0.474	0.447	0.030	0.929	0.805
		LOD/$\sqrt{2}$	-0.274	0.293	0.377	-0.013	-0.415	0.225	0.195	0.132	0.387	0.339	0.225	0.192	0.126	0.379	0.335	0.774	0.661	0.162	0.942	0.753
		LOD/2	0.040	0.271	0.356	-0.166	-0.460	0.169	0.196	0.134	0.318	0.287	0.168	0.193	0.127	0.311	0.281	0.936	0.705	0.220	0.912	0.620
		2-stage(rank)	-0.032	-0.026	-0.032	0.059	0.032	0.197	0.240	0.200	0.433	0.416	0.207	0.244	0.209	0.443	0.446	0.951	0.949	0.952	0.951	0.960
		2-stage(LS)	-0.033	-0.028	-0.034	0.060	0.035	0.194	0.239	0.200	0.428	0.413	0.204	0.243	0.207	0.441	0.445	0.948	0.950	0.953	0.954	0.963
		2-stage(marg)	-0.062	-0.064	-0.061	0.117	0.064	0.200	0.240	0.202	0.426	0.409	0.207	0.243	0.210	0.431	0.438	0.953	0.947	0.949	0.940	0.956
	400	full data	0.003	-0.006	-0.004	0.008	-0.001	0.104	0.142	0.114	0.228	0.236	0.105	0.142	0.113	0.233	0.233	0.951	0.941	0.939	0.949	0.944
		complete cases	-0.028	-0.009	-0.014	0.026	0.018	0.409	0.224	0.186	0.423	0.426	0.409	0.228	0.184	0.437	0.415	0.955	0.945	0.948	0.953	0.939
		LOD	-0.711	0.391	0.466	0.152	-0.447	0.237	0.127	0.083	0.335	0.289	0.247	0.129	0.082	0.346	0.290	0.178	0.148	0.000	0.932	0.661
		LOD/$\sqrt{2}$	-0.268	0.291	0.379	-0.019	-0.415	0.153	0.134	0.091	0.267	0.236	0.158	0.135	0.088	0.266	0.235	0.606	0.413	0.016	0.947	0.574
		LOD/2	0.044	0.269	0.358	-0.171	-0.459	0.117	0.136	0.092	0.222	0.199	0.118	0.136	0.089	0.218	0.197	0.937	0.479	0.032	0.874	0.356
		2-stage(rank)	-0.016	-0.026	-0.022	0.049	0.015	0.135	0.163	0.137	0.291	0.297	0.139	0.168	0.142	0.305	0.307	0.960	0.954	0.945	0.952	0.950
		2-stage(LS)	-0.015	-0.027	-0.022	0.053	0.012	0.135	0.163	0.137	0.291	0.296	0.137	0.168	0.141	0.305	0.306	0.959	0.955	0.941	0.954	0.954
		2-stage(marg)	-0.046	-0.065	-0.052	0.107	0.049	0.138	0.165	0.140	0.282	0.288	0.140	0.167	0.143	0.291	0.294	0.954	0.932	0.933	0.943	0.946

TABLE S2 Simulation results for linear outcome model where $p = q = 2$, $T = -Z$ and a mixture of multivariate normals as the error term distribution. Eight methods were implemented, including analysis with full data, complete-case approach, substitution methods with three replacement values: LOD, LOD/$\sqrt{2}$ and LOD/2, two-stage rank-based approach with multivariate KM estimator for η (2-stage(rank)), two-stage least-square approach with multivariate KM estimator for η (2-stage(LS)), and two-stage approach with marginal approximation (2-stage(marg)). Estimated S.D.* was from bootstrap for three version of the proposed method and from regression for full data analysis, complete-case approach and substitution methods.

			Bias					Empirical S.E.					Mean estimated S.D.*					Coverage rate of 95% CIs				
%<LOD	n	Method	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2
25%	200	full data	0.002	-0.005	0.008	0.011	-0.019	0.165	0.225	0.193	0.425	0.430	0.162	0.232	0.195	0.421	0.420	0.946	0.951	0.944	0.948	0.941
		complete cases	-0.002	-0.004	0.008	0.010	-0.015	0.319	0.272	0.243	0.554	0.558	0.321	0.280	0.245	0.560	0.543	0.944	0.966	0.950	0.949	0.942
		LOD	-0.225	0.218	0.340	0.067	-0.438	0.242	0.221	0.159	0.491	0.436	0.237	0.222	0.154	0.489	0.423	0.842	0.837	0.402	0.942	0.810
		LOD/$\sqrt{2}$	-0.104	0.138	0.259	0.049	-0.361	0.205	0.228	0.171	0.454	0.423	0.199	0.229	0.164	0.448	0.406	0.915	0.899	0.631	0.940	0.843
		LOD/2	0.014	0.116	0.227	-0.020	-0.340	0.181	0.229	0.175	0.418	0.401	0.175	0.229	0.167	0.410	0.383	0.941	0.925	0.709	0.941	0.836
		2-stage(rank)	0.006	-0.012	-0.001	0.028	-0.021	0.178	0.238	0.211	0.460	0.471	0.176	0.249	0.216	0.463	0.463	0.951	0.951	0.947	0.940	0.929
		2-stage(LS)	0.007	-0.013	0.000	0.030	-0.022	0.178	0.238	0.211	0.461	0.472	0.176	0.249	0.216	0.463	0.463	0.950	0.950	0.947	0.936	0.932
		2-stage(marg)	-0.012	-0.029	-0.008	0.057	-0.011	0.181	0.239	0.212	0.459	0.471	0.178	0.249	0.216	0.458	0.459	0.950	0.949	0.945	0.935	0.929
	400	full data	-0.003	-0.002	-0.008	0.000	0.013	0.112	0.157	0.136	0.297	0.294	0.114	0.162	0.137	0.294	0.294	0.956	0.964	0.948	0.939	0.952
		complete cases	0.004	0.000	-0.006	-0.006	0.008	0.219	0.184	0.170	0.387	0.380	0.225	0.195	0.170	0.391	0.379	0.955	0.957	0.954	0.944	0.943
		LOD	-0.227	0.223	0.332	0.053	-0.419	0.163	0.156	0.118	0.341	0.308	0.167	0.156	0.108	0.344	0.298	0.710	0.702	0.171	0.946	0.697
		LOD/$\sqrt{2}$	-0.107	0.143	0.251	0.037	-0.341	0.136	0.159	0.126	0.310	0.297	0.140	0.160	0.115	0.315	0.286	0.888	0.856	0.433	0.948	0.755
		LOD/2	0.011	0.121	0.218	-0.031	-0.320	0.120	0.159	0.128	0.283	0.279	0.123	0.161	0.118	0.288	0.269	0.955	0.877	0.555	0.946	0.764
		2-stage(rank)	0.000	-0.006	-0.010	0.015	0.004	0.119	0.169	0.150	0.323	0.327	0.123	0.173	0.150	0.322	0.323	0.953	0.954	0.947	0.941	0.945
		2-stage(LS)	0.000	-0.006	-0.011	0.015	0.005	0.119	0.168	0.150	0.320	0.325	0.123	0.173	0.150	0.322	0.323	0.953	0.956	0.951	0.942	0.946
		2-stage(marg)	-0.019	-0.024	-0.021	0.042	0.020	0.121	0.168	0.151	0.316	0.324	0.124	0.173	0.150	0.317	0.319	0.950	0.952	0.948	0.942	0.948
50%	200	full data	0.002	-0.005	0.008	0.011	-0.019	0.165	0.225	0.193	0.425	0.430	0.162	0.232	0.195	0.421	0.420	0.946	0.951	0.944	0.948	0.941
		complete cases	0.007	-0.011	0.013	0.018	-0.031	0.636	0.373	0.325	0.781	0.765	0.614	0.375	0.324	0.772	0.745	0.942	0.950	0.950	0.954	0.937
		LOD	-0.439	0.457	0.540	-0.006	-0.649	0.375	0.190	0.119	0.581	0.459	0.378	0.192	0.121	0.577	0.459	0.782	0.331	0.008	0.951	0.701
		LOD/$\sqrt{2}$	-0.105	0.351	0.459	-0.127	-0.594	0.239	0.207	0.135	0.440	0.378	0.237	0.207	0.134	0.440	0.372	0.919	0.612	0.090	0.940	0.641
		LOD/2	0.150	0.327	0.438	-0.267	-0.612	0.174	0.211	0.139	0.356	0.316	0.170	0.210	0.136	0.356	0.308	0.852	0.651	0.124	0.886	0.483
		2-stage(rank)	-0.015	-0.027	-0.011	0.055	-0.004	0.214	0.270	0.245	0.537	0.553	0.220	0.293	0.265	0.575	0.583	0.955	0.959	0.955	0.949	0.942
		2-stage(LS)	-0.015	-0.026	-0.011	0.053	-0.004	0.214	0.270	0.245	0.538	0.557	0.220	0.293	0.265	0.574	0.583	0.952	0.959	0.954	0.947	0.945
		2-stage(marg)	-0.043	-0.063	-0.041	0.104	0.034	0.219	0.274	0.250	0.530	0.547	0.224	0.295	0.270	0.570	0.583	0.951	0.949	0.954	0.948	0.946
	400	full data	-0.003	-0.002	-0.008	0.000	0.013	0.112	0.157	0.136	0.297	0.294	0.114	0.162	0.137	0.294	0.294	0.956	0.964	0.948	0.939	0.952
		complete cases	0.020	0.006	-0.002	-0.017	-0.001	0.424	0.260	0.224	0.548	0.523	0.424	0.259	0.223	0.534	0.514	0.951	0.947	0.955	0.946	0.945
		LOD	-0.432	0.462	0.537	-0.015	-0.647	0.264	0.136	0.090	0.410	0.323	0.266	0.135	0.086	0.406	0.323	0.626	0.073	0.000	0.943	0.487
		LOD/$\sqrt{2}$	-0.104	0.355	0.454	-0.136	-0.585	0.162	0.147	0.101	0.309	0.264	0.167	0.146	0.094	0.310	0.262	0.915	0.331	0.003	0.924	0.383
		LOD/2	0.151	0.331	0.432	-0.275	-0.603	0.116	0.149	0.103	0.249	0.220	0.120	0.147	0.096	0.251	0.217	0.772	0.399	0.007	0.803	0.219
		2-stage(rank)	-0.019	-0.019	-0.021	0.040	0.020	0.139	0.190	0.168	0.378	0.379	0.148	0.199	0.179	0.395	0.398	0.961	0.955	0.961	0.955	0.951
		2-stage(LS)	-0.019	-0.019	-0.021	0.040	0.020	0.140	0.191	0.168	0.380	0.380	0.148	0.199	0.178	0.394	0.397	0.958	0.950	0.964	0.949	0.951
		2-stage(marg)	-0.048	-0.058	-0.051	0.095	0.056	0.143	0.194	0.172	0.372	0.374	0.151	0.199	0.181	0.382	0.388	0.959	0.944	0.951	0.942	0.951

TABLE S3 Simulation results for linear outcome model where $p = q = 2$, $T = -\log(Z)$ and a mixture of multivariate normals as the error term distribution. Eight methods were implemented, including analysis with full data, complete-case approach, substitution methods with three replacement values: LOD, LOD/$\sqrt{2}$ and LOD/2, two-stage rank-based approach with multivariate KM estimator for η (2-stage(rank)), two-stage least-square approach with multivariate KM estimator for η (2-stage(LS)), and two-stage approach with marginal approximation (2-stage(marg)). Estimated S.D.* was from bootstrap for three version of the proposed method and from regression for full data analysis, complete-case approach and substitution methods.

			Bias					Empirical S.E.					Mean estimated S.D.*					Coverage rate of 95% CIs				
%<LOD	n	Method	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2
25%	200	full data	0.004	-0.006	0.001	-0.002	0.001	0.205	0.200	0.137	0.163	0.102	0.205	0.213	0.143	0.163	0.103	0.947	0.971	0.964	0.954	0.959
		complete cases	-0.004	-0.007	0.007	0.001	-0.001	0.298	0.249	0.213	0.179	0.120	0.300	0.262	0.217	0.183	0.121	0.949	0.956	0.952	0.945	0.951
		LOD	-0.349	0.282	0.392	0.002	-0.128	0.235	0.196	0.125	0.172	0.101	0.236	0.205	0.128	0.170	0.101	0.679	0.717	0.138	0.951	0.758
		LOD/$\sqrt{2}$	0.221	0.066	0.126	-0.071	-0.059	0.195	0.208	0.138	0.157	0.101	0.195	0.217	0.140	0.157	0.100	0.804	0.945	0.859	0.925	0.909
		LOD/2	0.703	0.038	0.014	-0.196	-0.032	0.179	0.218	0.151	0.146	0.103	0.173	0.225	0.151	0.146	0.100	0.016	0.951	0.955	0.732	0.928
		2-stage(rank)	0.009	-0.007	-0.006	0.000	0.002	0.211	0.206	0.142	0.166	0.104	0.214	0.220	0.149	0.168	0.107	0.939	0.966	0.958	0.946	0.947
		2-stage(LS)	0.008	-0.007	-0.005	0.000	0.002	0.212	0.206	0.141	0.166	0.104	0.214	0.220	0.149	0.168	0.107	0.942	0.966	0.961	0.942	0.945
		2-stage(marg)	-0.013	-0.018	-0.004	0.011	0.001	0.212	0.205	0.142	0.166	0.103	0.214	0.220	0.149	0.168	0.106	0.943	0.967	0.959	0.943	0.946
	400	full data	0.000	-0.003	0.002	0.001	-0.001	0.147	0.147	0.100	0.114	0.069	0.143	0.149	0.099	0.114	0.071	0.945	0.952	0.945	0.945	0.963
		complete cases	-0.003	-0.009	-0.001	0.004	0.000	0.206	0.182	0.149	0.128	0.081	0.209	0.182	0.150	0.127	0.083	0.955	0.953	0.949	0.949	0.960
		LOD	-0.356	0.286	0.392	0.004	-0.128	0.168	0.143	0.093	0.120	0.069	0.164	0.144	0.089	0.119	0.070	0.415	0.477	0.014	0.950	0.560
		LOD/$\sqrt{2}$	0.212	0.068	0.125	-0.067	-0.059	0.140	0.150	0.102	0.110	0.069	0.136	0.152	0.097	0.110	0.069	0.646	0.935	0.743	0.907	0.865
		LOD/2	0.694	0.036	0.010	-0.191	-0.031	0.127	0.157	0.110	0.105	0.070	0.121	0.158	0.105	0.102	0.069	0.000	0.951	0.929	0.514	0.934
		2-stage(rank)	0.004	-0.004	-0.006	0.004	0.000	0.151	0.150	0.103	0.116	0.070	0.148	0.153	0.103	0.117	0.073	0.940	0.954	0.942	0.945	0.962
		2-stage(LS)	0.004	-0.004	-0.006	0.004	0.000	0.151	0.149	0.103	0.115	0.070	0.148	0.153	0.103	0.117	0.073	0.940	0.954	0.941	0.947	0.963
		2-stage(marg)	-0.017	-0.015	-0.004	0.014	-0.001	0.152	0.149	0.103	0.116	0.070	0.148	0.153	0.103	0.117	0.072	0.936	0.952	0.940	0.939	0.963
50%	200	full data	0.004	-0.006	0.001	-0.002	0.001	0.205	0.200	0.137	0.163	0.102	0.205	0.213	0.143	0.163	0.103	0.947	0.971	0.964	0.954	0.959
		complete cases	0.007	-0.003	0.005	-0.006	0.002	0.520	0.359	0.299	0.225	0.143	0.520	0.361	0.304	0.221	0.144	0.952	0.949	0.955	0.945	0.953
		LOD	-1.202	0.704	0.804	0.011	-0.203	0.317	0.183	0.109	0.193	0.105	0.317	0.191	0.115	0.188	0.105	0.041	0.035	0.000	0.946	0.508
		LOD/$\sqrt{2}$	0.158	0.396	0.524	-0.145	-0.196	0.213	0.197	0.126	0.155	0.099	0.207	0.207	0.127	0.152	0.095	0.885	0.507	0.012	0.834	0.479
		LOD/2	1.070	0.328	0.437	-0.295	-0.234	0.174	0.210	0.140	0.134	0.100	0.162	0.217	0.136	0.130	0.088	0.000	0.663	0.117	0.368	0.268
		2-stage(rank)	-0.022	-0.023	-0.009	0.020	-0.001	0.245	0.229	0.149	0.182	0.107	0.250	0.243	0.160	0.186	0.113	0.952	0.953	0.960	0.944	0.954
		2-stage(LS)	-0.023	-0.023	-0.009	0.020	0.000	0.246	0.229	0.150	0.182	0.108	0.250	0.243	0.160	0.186	0.113	0.947	0.959	0.964	0.940	0.955
		2-stage(marg)	-0.069	-0.059	-0.027	0.047	0.004	0.248	0.229	0.150	0.182	0.107	0.251	0.242	0.160	0.184	0.112	0.947	0.956	0.957	0.935	0.952
	400	full data	0.000	-0.003	0.002	0.001	-0.001	0.147	0.147	0.100	0.114	0.069	0.143	0.149	0.099	0.114	0.071	0.945	0.952	0.945	0.945	0.963
		complete cases	-0.005	-0.013	0.006	0.006	-0.003	0.356	0.250	0.205	0.149	0.093	0.359	0.251	0.208	0.152	0.097	0.951	0.953	0.952	0.948	0.959
		LOD	-1.206	0.709	0.802	0.009	-0.201	0.225	0.135	0.081	0.134	0.071	0.220	0.134	0.080	0.131	0.072	0.001	0.001	0.000	0.943	0.202
		LOD/$\sqrt{2}$	0.148	0.400	0.521	-0.146	-0.191	0.148	0.147	0.092	0.109	0.068	0.144	0.145	0.088	0.106	0.066	0.816	0.212	0.000	0.712	0.175
		LOD/2	1.059	0.331	0.432	-0.296	-0.227	0.121	0.156	0.104	0.096	0.070	0.114	0.153	0.095	0.091	0.061	0.000	0.412	0.010	0.111	0.050
		2-stage(rank)	-0.014	-0.016	-0.006	0.019	-0.003	0.168	0.161	0.108	0.123	0.073	0.169	0.167	0.110	0.127	0.076	0.938	0.951	0.950	0.946	0.957
		2-stage(LS)	-0.014	-0.017	-0.007	0.019	-0.003	0.169	0.162	0.108	0.124	0.073	0.169	0.167	0.110	0.127	0.076	0.942	0.949	0.947	0.946	0.959
		2-stage(marg)	-0.060	-0.051	-0.024	0.045	0.001	0.169	0.161	0.108	0.122	0.072	0.170	0.165	0.110	0.125	0.075	0.930	0.945	0.946	0.932	0.960

TABLE S4 Simulation results for linear outcome model where $p = q = 2$, $T = -\log(Z)$ and the multivariate normal as the error term distribution. Eight methods were implemented, including analysis with full data, complete-case approach, substitution methods with three replacement values: LOD, LOD/$\sqrt{2}$ and LOD/2, two-stage rank-based approach with multivariate KM estimator for η (2-stage(rank)), two-stage least-square approach with multivariate KM estimator for η (2-stage(LS)), and two-stage approach with marginal approximation (2-stage(marg)). Estimated S.D.* was from bootstrap for three version of the proposed method and from regression for full data analysis, complete-case approach and substitution methods.

		Bias					Empirical S.E.					Mean estimated S.D.					Coverage rate of 95% CIs				
n	Method	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2
50	full data	-0.006	-0.005	-0.008	0.011	-0.003	0.370	0.401	0.268	0.268	0.190	0.382	0.397	0.271	0.273	0.190	0.954	0.944	0.946	0.951	0.948
	complete cases	0.026	0.004	-0.061	0.000	0.018	1.213	0.792	0.650	0.447	0.314	1.134	0.742	0.592	0.416	0.284	0.928	0.924	0.922	0.927	0.908
	LOD	-1.478	0.624	0.706	0.104	-0.152	0.644	0.389	0.240	0.347	0.210	0.626	0.388	0.239	0.345	0.211	0.343	0.651	0.178	0.937	0.873
	LOD/$\sqrt{2}$	0.012	0.340	0.443	-0.075	-0.166	0.412	0.403	0.263	0.266	0.185	0.409	0.401	0.252	0.270	0.182	0.949	0.856	0.539	0.941	0.840
	LOD/2	0.965	0.269	0.358	-0.224	-0.210	0.349	0.421	0.283	0.233	0.182	0.331	0.417	0.266	0.232	0.168	0.183	0.895	0.688	0.828	0.764
	2-stage(rank)	-0.084	-0.059	-0.048	0.059	0.008	0.507	0.472	0.315	0.324	0.210	0.602	0.527	0.372	0.382	0.260	0.969	0.958	0.956	0.961	0.966
	2-stage(LS)	-0.085	-0.056	-0.046	0.058	0.007	0.500	0.471	0.314	0.324	0.210	0.596	0.524	0.370	0.382	0.260	0.976	0.958	0.956	0.963	0.965
	2-stage(marg)	-0.126	-0.089	-0.061	0.081	0.011	0.507	0.472	0.315	0.326	0.210	0.611	0.530	0.374	0.391	0.263	0.974	0.957	0.953	0.958	0.970
100	full data	-0.001	-0.003	-0.003	-0.001	0.003	0.260	0.269	0.185	0.186	0.130	0.265	0.275	0.185	0.187	0.126	0.961	0.952	0.950	0.950	0.947
	complete cases	0.023	0.002	-0.001	-0.006	0.001	0.726	0.480	0.372	0.264	0.183	0.727	0.470	0.378	0.265	0.175	0.942	0.933	0.944	0.934	0.948
	LOD	-1.455	0.614	0.709	0.081	-0.141	0.417	0.267	0.167	0.226	0.139	0.428	0.269	0.163	0.233	0.138	0.068	0.375	0.012	0.948	0.808
	LOD/$\sqrt{2}$	0.004	0.333	0.442	-0.089	-0.151	0.281	0.272	0.182	0.178	0.126	0.283	0.277	0.172	0.185	0.122	0.957	0.759	0.287	0.932	0.768
	LOD/2	0.954	0.267	0.355	-0.237	-0.195	0.243	0.283	0.194	0.157	0.122	0.229	0.288	0.183	0.159	0.113	0.010	0.840	0.495	0.686	0.597
	2-stage(rank)	-0.053	-0.028	-0.026	0.029	0.007	0.328	0.304	0.211	0.206	0.140	0.354	0.321	0.218	0.222	0.146	0.952	0.962	0.944	0.953	0.948
	2-stage(LS)	-0.056	-0.029	-0.026	0.031	0.007	0.326	0.304	0.211	0.207	0.141	0.352	0.321	0.218	0.222	0.146	0.953	0.960	0.946	0.953	0.946
	2-stage(marg)	-0.102	-0.062	-0.043	0.054	0.012	0.331	0.303	0.212	0.207	0.140	0.358	0.320	0.219	0.222	0.145	0.952	0.954	0.942	0.946	0.948

TABLE S5 Simulation results for linear outcome model where $p = 10$, $q = 2$, $T = -\log(Z)$, the multivariate normal as the error term distribution and n=400. Six methods were implemented, including analysis with full data, complete-case approach, substitution methods with three replacement values: LOD, LOD/$\sqrt{2}$ and LOD/2 and two-stage approach with marginal approximation (2-stage(marg)). Estimated S.D.* was from bootstrap with sample size 100 for the two-stage marginal approach and from regression for the full data analysis, complete-case approach and substitution methods.

	Method	β_0	β_1	β_2	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}
Bias	full data	0.006	0.002	0.005	-0.002	-0.002	0.000	-0.002	0.000	0.001	-0.004	0.000	0.001	-0.002
	complete cases	0.025	0.003	-0.003	-0.004	-0.001	-0.016	0.004	-0.003	0.002	0.001	0.003	0.002	-0.010
	LOD	-0.921	0.183	0.347	0.009	-0.025	0.058	0.056	0.031	-0.040	0.052	-0.055	-0.048	0.145
	LOD/$\sqrt{2}$	0.026	0.039	0.102	-0.012	-0.018	0.001	-0.004	-0.003	-0.017	-0.006	-0.021	-0.019	0.020
	LOD/2	0.713	-0.031	-0.047	-0.035	-0.017	-0.041	-0.053	-0.033	-0.004	-0.055	0.001	0.000	-0.081
	2-stage(marg)	-0.087	-0.017	-0.006	0.004	0.003	0.014	0.012	0.009	0.001	0.010	-0.002	0.000	0.017
Empirical S.E.	full data	0.168	0.131	0.093	0.061	0.075	0.127	0.111	0.090	0.055	0.094	0.092	0.080	0.106
	complete cases	0.439	0.260	0.215	0.098	0.129	0.278	0.183	0.138	0.085	0.155	0.150	0.125	0.177
	LOD	0.195	0.137	0.097	0.066	0.078	0.144	0.120	0.098	0.057	0.101	0.097	0.084	0.112
	LOD/$\sqrt{2}$	0.173	0.135	0.095	0.062	0.076	0.132	0.111	0.091	0.055	0.094	0.093	0.081	0.105
	LOD/2	0.166	0.138	0.098	0.061	0.076	0.129	0.108	0.090	0.055	0.092	0.093	0.081	0.104
	2-stage(marg)	0.179	0.139	0.097	0.064	0.077	0.133	0.115	0.094	0.056	0.097	0.094	0.082	0.110
Mean estimated S.D.*	full data	0.162	0.131	0.092	0.060	0.075	0.127	0.109	0.093	0.053	0.094	0.092	0.079	0.105
	complete cases	0.429	0.259	0.208	0.095	0.127	0.264	0.177	0.144	0.083	0.152	0.144	0.124	0.177
	LOD	0.184	0.137	0.093	0.066	0.080	0.138	0.120	0.101	0.057	0.103	0.099	0.085	0.114
	LOD/$\sqrt{2}$	0.164	0.134	0.092	0.061	0.076	0.128	0.111	0.094	0.054	0.096	0.093	0.080	0.106
	LOD/2	0.157	0.136	0.095	0.061	0.076	0.127	0.108	0.093	0.054	0.094	0.093	0.080	0.103
	2-stage(marg)	0.200	0.156	0.113	0.076	0.096	0.153	0.135	0.115	0.069	0.117	0.118	0.102	0.130
Coverage rate of 95% CIs	full data	0.941	0.947	0.951	0.943	0.956	0.949	0.950	0.960	0.946	0.954	0.942	0.945	0.947
	complete cases	0.941	0.954	0.942	0.944	0.950	0.943	0.944	0.952	0.947	0.951	0.943	0.953	0.947
	LOD	0.006	0.728	0.058	0.946	0.948	0.917	0.933	0.952	0.890	0.923	0.912	0.915	0.749
	LOD/$\sqrt{2}$	0.943	0.931	0.779	0.943	0.949	0.941	0.956	0.960	0.939	0.949	0.939	0.942	0.946
	LOD/2	0.006	0.939	0.919	0.910	0.950	0.935	0.932	0.943	0.944	0.914	0.947	0.957	0.871
	2-stage(marg)	0.959	0.967	0.979	0.980	0.985	0.965	0.979	0.982	0.980	0.976	0.976	0.981	0.975

TABLE S6 Simulation results for 17 metals under the real data setting where $p = 7$, $q = 24$, $n = 252$. and the multivariate normal as the error term distribution. Five methods were implemented, including analysis with complete-case approach (CC), substitution methods with three replacement values: LOD, LOD/$\sqrt{2}$ and LOD/2, and two-stage approach with marginal approximation (MG). Estimated S.D.* was from bootstrap for the proposed method and from regression for complete-case approach and substitution methods.

		Estimate					Empirical S.E.					Mean estimated S.D.*					Coverage rate of 95% CIs				
Metal	θ	CC	LOD	$\frac{LOD}{\sqrt{2}}$	$\frac{LOD}{2}$	MG	CC	LOD	$\frac{LOD}{\sqrt{2}}$	$\frac{LOD}{2}$	MG	CC	LOD	$\frac{LOD}{\sqrt{2}}$	$\frac{LOD}{2}$	MG	CC	LOD	$\frac{LOD}{\sqrt{2}}$	$\frac{LOD}{2}$	MG
As	0.002	-0.001	0.007	0.004	0.009	0.004	0.162	0.082	0.082	0.082	0.091	0.160	0.080	0.080	0.080	0.095	0.941	0.942	0.938	0.939	0.955
Mo	0.043	0.035	0.049	0.046	0.050	0.042	0.346	0.168	0.170	0.170	0.188	0.330	0.163	0.165	0.164	0.196	0.940	0.940	0.938	0.937	0.954
Zn	0.033	0.022	0.037	0.031	0.037	0.031	0.304	0.143	0.146	0.146	0.166	0.293	0.143	0.145	0.145	0.180	0.943	0.957	0.949	0.950	0.962
Se	0.173	0.179	0.195	0.190	0.205	0.192	0.529	0.250	0.249	0.247	0.282	0.512	0.253	0.255	0.253	0.304	0.946	0.950	0.949	0.949	0.959
Ba	0.035	0.026	0.043	0.035	0.044	0.031	0.262	0.125	0.131	0.131	0.149	0.260	0.122	0.126	0.125	0.157	0.941	0.935	0.942	0.945	0.956
Mn	0.118	0.121	0.123	0.121	0.123	0.124	0.159	0.078	0.078	0.078	0.088	0.161	0.081	0.081	0.081	0.100	0.950	0.962	0.961	0.959	0.967
Sn	0.007	0.003	0.005	0.005	0.006	0.024	0.160	0.080	0.078	0.075	0.085	0.161	0.083	0.080	0.077	0.092	0.947	0.960	0.964	0.967	0.954
Cu	0.628	0.653	0.611	0.568	0.459	0.658	0.572	0.293	0.268	0.232	0.303	0.561	0.288	0.266	0.230	0.328	0.941	0.948	0.941	0.881	0.955
Hg	0.106	0.094	0.111	0.105	0.107	0.129	0.202	0.104	0.100	0.094	0.110	0.197	0.102	0.098	0.093	0.116	0.930	0.952	0.954	0.955	0.960
Ni	0.056	0.071	0.052	0.057	0.038	0.120	0.466	0.241	0.228	0.199	0.257	0.450	0.231	0.217	0.190	0.266	0.933	0.934	0.940	0.933	0.941
Tl	-0.144	-0.143	-0.152	-0.137	-0.115	-0.088	0.301	0.154	0.139	0.122	0.152	0.285	0.154	0.139	0.123	0.166	0.929	0.946	0.943	0.946	0.943
Pb	0.040	0.048	0.035	0.036	0.032	0.096	0.286	0.168	0.146	0.125	0.157	0.295	0.166	0.144	0.122	0.174	0.949	0.952	0.942	0.932	0.955
Cd	-0.034	-0.039	-0.049	-0.023	0.001	-0.035	0.280	0.172	0.145	0.120	0.141	0.271	0.165	0.139	0.115	0.221	0.936	0.927	0.933	0.929	0.985
W	-0.088	-0.057	-0.045	-0.044	-0.044	-0.048	0.316	0.174	0.174	0.174	0.183	0.313	0.170	0.170	0.170	0.197	0.954	0.948	0.946	0.948	0.958
Cr	-0.140	-0.065	-0.066	-0.066	-0.066	-0.069	0.404	0.212	0.212	0.213	0.230	0.403	0.211	0.211	0.211	0.247	0.948	0.936	0.937	0.934	0.949
U	0.153	0.137	0.121	0.121	0.122	0.122	0.429	0.228	0.228	0.229	0.240	0.417	0.223	0.223	0.223	0.258	0.949	0.951	0.952	0.953	0.947
Be	-0.214	-0.087	-0.100	-0.100	-0.099	-0.104	0.426	0.217	0.217	0.218	0.237	0.428	0.228	0.228	0.228	0.266	0.935	0.931	0.936	0.934	0.949

TABLE S7 Simulation results for logistic outcome model where $p = q = 2$, $n = 400$, $T = -\log(Z)$ and the multivariate normal as the error term distribution. Eight methods were implemented, including analysis with full data, complete-case approach, substitution methods with three replacement values: LOD, LOD/$\sqrt{2}$ and LOD/2, two-stage rank-based approach with multivariate KM estimator for η (2-stage(rank)), two-stage least-square approach with multivariate KM estimator for η (2-stage(LS)), and two-stage approach with marginal approximation (2-stage(marg)). Estimated S.D.* was from bootstrap for three version of the proposed method and from regression for full data analysis, complete-case approach and substitution methods.

		Bias					Empirical S.E.					Mean estimated S.D.*					Coverage rate of 95% CIs				
%<LOD	Method	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2	β_0	β_1	β_2	γ_1	γ_2
25%	full data	0.010	-0.026	0.007	0.033	-0.020	0.349	0.367	0.247	0.332	0.194	0.342	0.364	0.248	0.333	0.195	0.951	0.951	0.958	0.957	0.954
	complete cases	0.013	-0.052	0.020	0.060	-0.039	0.492	0.534	0.393	0.400	0.242	0.484	0.536	0.404	0.395	0.243	0.952	0.964	0.960	0.943	0.962
	LOD	-0.102	0.275	-0.026	-0.065	0.063	0.375	0.329	0.218	0.323	0.178	0.367	0.324	0.218	0.324	0.180	0.933	0.852	0.955	0.941	0.917
	LOD/$\sqrt{2}$	0.023	0.124	-0.007	-0.102	0.056	0.332	0.349	0.237	0.292	0.179	0.328	0.345	0.236	0.301	0.181	0.949	0.933	0.953	0.937	0.926
	LOD/2	0.098	0.082	0.018	-0.195	0.086	0.308	0.358	0.247	0.261	0.177	0.306	0.353	0.246	0.274	0.178	0.946	0.944	0.953	0.887	0.900
	2-stage(rank)	0.019	-0.038	0.008	0.026	-0.018	0.352	0.372	0.251	0.335	0.195	0.357	0.380	0.261	0.350	0.206	0.952	0.957	0.960	0.956	0.962
	2-stage(LS)	0.019	-0.038	0.008	0.027	-0.018	0.353	0.372	0.251	0.336	0.195	0.357	0.380	0.261	0.350	0.206	0.951	0.958	0.961	0.955	0.962
	2-stage(marg)	0.014	-0.005	0.010	0.006	-0.007	0.353	0.368	0.250	0.331	0.192	0.359	0.375	0.261	0.344	0.203	0.949	0.957	0.960	0.951	0.961
50%	full data	0.010	-0.026	0.007	0.033	-0.020	0.349	0.367	0.247	0.332	0.194	0.342	0.364	0.248	0.333	0.195	0.951	0.951	0.958	0.957	0.954
	complete cases	0.009	-0.032	0.070	0.070	-0.063	0.803	0.804	0.589	0.487	0.303	0.834	0.816	0.610	0.507	0.308	0.969	0.961	0.968	0.963	0.957
	LOD	-0.220	0.606	-0.083	-0.091	0.113	0.491	0.292	0.185	0.361	0.175	0.475	0.288	0.188	0.355	0.178	0.927	0.425	0.930	0.929	0.882
	LOD/$\sqrt{2}$	-0.024	0.448	-0.045	-0.251	0.158	0.338	0.310	0.205	0.274	0.162	0.339	0.306	0.206	0.280	0.165	0.946	0.680	0.951	0.836	0.814
	LOD/2	0.020	0.413	-0.013	-0.401	0.229	0.282	0.317	0.215	0.225	0.152	0.284	0.312	0.215	0.232	0.154	0.949	0.724	0.957	0.578	0.658
	2-stage(rank)	0.019	-0.026	0.005	0.009	-0.007	0.355	0.379	0.253	0.354	0.199	0.369	0.392	0.269	0.375	0.212	0.960	0.961	0.966	0.954	0.964
	2-stage(LS)	0.019	-0.025	0.005	0.007	-0.006	0.355	0.379	0.254	0.354	0.198	0.368	0.390	0.268	0.372	0.210	0.958	0.958	0.966	0.951	0.964
	2-stage(marg)	0.024	0.029	0.016	-0.034	0.011	0.359	0.372	0.257	0.347	0.195	0.372	0.380	0.270	0.358	0.206	0.958	0.955	0.965	0.949	0.957

TABLE S8 Summary of 17 urinary trace metals among 252 women who delivered full term in the LIFECODES cohort

Metal	LOD(ppb)	n<LOD	%<LOD
As	0.3	0	0
Mo	0.3	0	0
Zn	2	0	0
Se	5	1	0.4
Ba	0.1	4	1.6
Mn	0.08	5	2.0
Sn	0.1	14	5.6
Cu	2.5	17	6.7
Hg	0.05	21	8.3
Ni	0.8	35	13.9
Tl	0.02	39	15.5
Pb	0.1	65	25.8
Cd	0.06	147	58.3
W	0.2	200	79.4
Cr	0.4	222	88.1
U	0.01	225	89.3
Be	0.04	227	90.1

TABLE S9 Summary of demographc variables among 252 women who delivered full term in the LIFECODES cohort

Demographic variables		n/Mean (S.D.)
Age	< 25	30
	[25, 30)	51
	[30, 35)	103
	≥ 35	68
Race	White	154
	Black	36
	Other	62
Education	High school degree or less	35
	Technical college or some college	37
	College graduate	78
	Graduate school	102
Insurance	Private	202
	Public	50
BMI	< 25	143
	[25, 30)	66
	≥ 30	43
Gestational age at the third trimester visit		26.14 (1.24)
Specific gravity at the third trimester visit		1.02 (0.01)

TABLE S10 Computing time (second) of the rank-based and the least-square approaches under different p and n

	rank-based approach					least-square approach				
	n					n				
p	100	200	400	600	1000	100	200	400	600	1000
2	<1	<1	<1	2.58	9.19	<1	<1	3.32	8.08	21.66
3	<1	<1	1.26	2.93	9.78	1.48	2.50	11.4	19.95	63.84
5	<1	<1	2.25	4.82	15.92	5.62	13.69	58.22	130.78	400.24
10	<1	<1	4.02	9.96	31.39	45.02	180.54	661.67	1491.66	5068.79

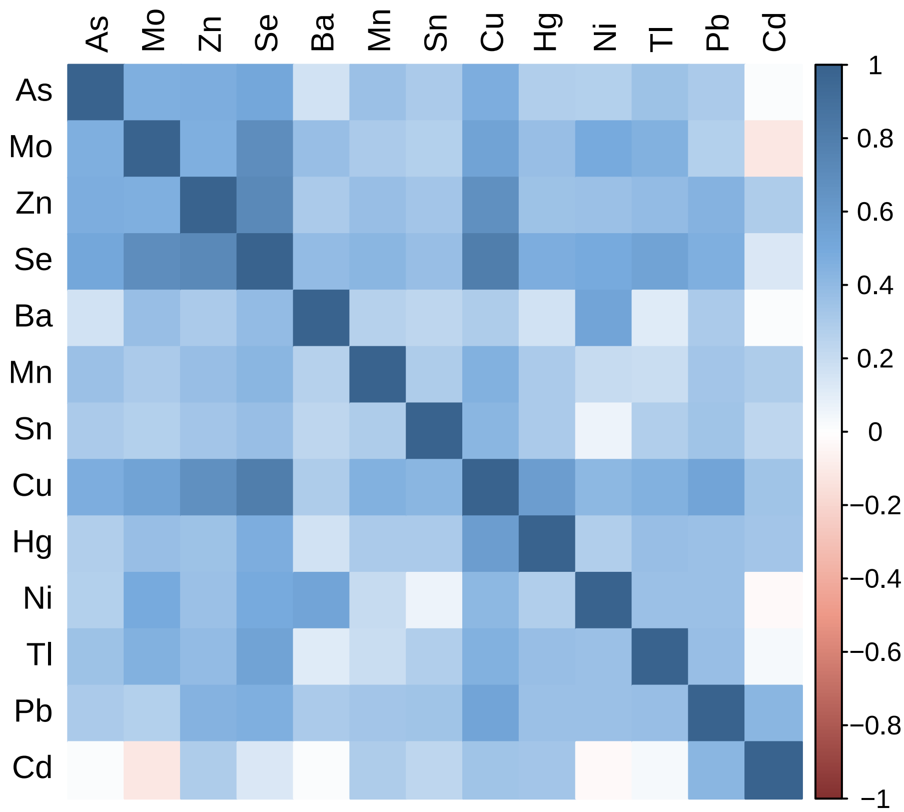

FIGURE S1 The heat mat for pairwise correlation matrix of 13 log-transformed metal concentrations.

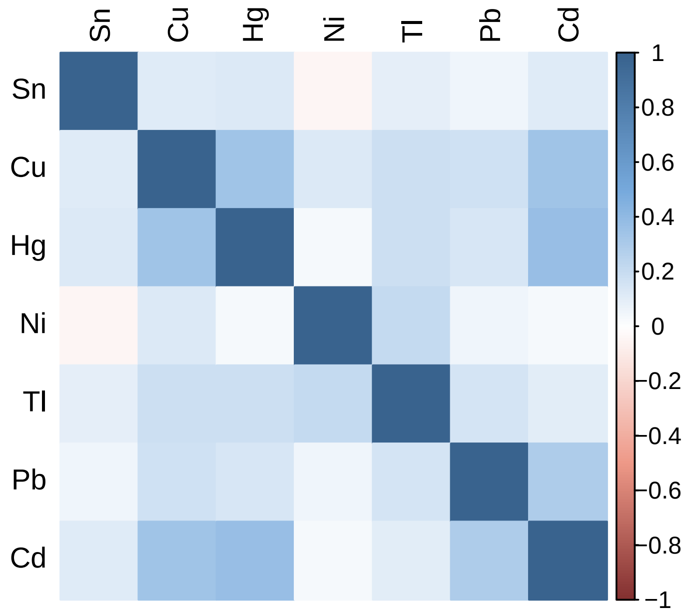

FIGURE S2 The heat mat for the estimated pairwise correlation in the 7 AFT models.

How to cite this article: Chen LW, Fine JP, Bair E, Ritter VS, McElrath TF, Cantonwine DE, Meeker JD, Ferguson KK, and Zhao S (202x), Semiparametric analysis of a generalized linear model with multiple covariates subject to detection limits, *Statistics in Medicine*, *xxxx*.