# A presmoothing approach for estimation in the semiparametric Cox mixture cure model

## Supplementary Material

Eni Musta[*,†], Valentin Patilea[‡] and Ingrid Van Keilegom[*]

[*] *KU Leuven,* [†] *University of Amsterdam* [‡] *Ensai*

This supplement is organized as follows. Appendix A contains technical lemmas and proofs. Appendix B collects additional simulation results, that were omitted from the main paper due to page limits.

### Appendix A: Technical lemmas and proofs

**Lemma 1.** *Let $B$ be a Bernoulli random variable, $T_0$ a nonnegative random variable and let $T = T_0$ if $B = 1$ and $T = \infty$ if $B = 0$. Let $X$ and $Z$ be two real-valued random vectors. Then*

$$T_0 \perp (C, X) \mid Z \quad \text{and } B \perp (C, T_0, Z) \mid X \quad \Longrightarrow \quad T \perp C \mid (X, Z)$$

*Proof.* This lemma is similar to Lemma 8.1 in [2]. We provide the proof for completeness. By elementary properties of conditional independence we have

$$B \perp (C, T_0, Z) \mid X \quad \Longleftrightarrow \quad B \perp C \mid (X, Z, T_0) \quad \text{and} \quad B \perp T_0 \mid (X, Z) \quad \text{and} \quad B \perp Z \mid X$$

and

$$T_0 \perp (C, X) \mid Z \quad \Longleftrightarrow \quad T_0 \perp C \mid (X, Z) \quad \text{and} \quad T_0 \perp X \mid Z.$$

Then,

$$C \perp B \mid (X, Z, T_0) \quad \text{and} \quad C \perp T_0 \mid (X, Z) \quad \Longleftrightarrow \quad (B, T_0) \perp C \mid (X, Z)$$

The result follows from the fact that $T$ is completely determined by $B$ and $T_0$. □

#### *A.1. Identifiability with restricted survival times*

For any $0 < \tau^* \leq \tau_0$, let

$$T_0^* = \min(T_0, \tau^*), \qquad T^* = BT_0^* + (1 - B)\infty \qquad \text{and} \qquad C^* = \min(C, \tau^*).$$

Moreover, let

$$Y^* = \min(T^*, C^*) \quad \text{ and } \Delta^* = \mathbb{1}_{\{T^* \leq C^*\}}.$$

A first aspect to study is the identifiability of the true values of the parameter when $(Y, \Delta)$ is replaced by $(Y^*, \Delta^*)$. Here, identifiability means that the true values $\beta_0$ and $\Lambda_0$ of the parameters maximize the expectation of the criterion maximized to obtain the estimators. This issue is addressed in Lemma 2. Let us introduce some additional notation: for any $0 < \tau^* \leq \tau_0$ and $\Lambda \in \mathcal{H}$, $\Lambda_{|\tau^*}$ is defined as

$$\Lambda_{|\tau^*}(t) = \Lambda(t), \;\; \forall t \in [0, \tau^*) \quad \text{ and } \quad \Delta\Lambda_{|\tau^*}(\tau^*) = \Lambda_{|\tau^*}(\{\tau^*\}) = 1. \tag{A1}$$

The dominating measure for the model of $T_0$ changes with such a stopped cumulative hazard measure to allow for a positive mass at $\tau^*$. Then, $\ell$ defined in (13) becomes

$$\begin{aligned}\ell(y, d, x, z; \beta, \Lambda_{|\tau^*}, \gamma) = \mathbb{1}_{\{y<\tau^*\}} & \left[ d \log f_u(y|z; \beta, \Lambda) \right. \\ & \left. +(1-d) \log \{1 - \phi(\gamma, x) + \phi(\gamma, x) S_u(y|z; \beta, \Lambda)\} \right] \\ & + \mathbb{1}_{\{y \geq \tau^*\}} \left[ d \log S_u(\tau^*|z; \beta, \Lambda) + (1-d) \log \{1 - \phi(\gamma, x)\} \right]. \end{aligned} \tag{A2}$$

**Lemma 2.** *Let $0 < \tau^* \leq \tau_0$. Assume that for any $\tilde\beta \in B$ and $\tilde\Lambda \in \mathcal{H}$,*

$$S_u(t|z; \tilde\beta, \tilde\Lambda_{|\tau^*}) = S_u(t|z; \beta_0, \Lambda_{0|\tau^*}), \; \forall t \in [0, \tau^*) \quad \Longrightarrow \quad \tilde\beta = \beta_0 \;\; \text{and} \;\; \tilde\Lambda_{|\tau^*} = \Lambda_{0|\tau^*}. \tag{A3}$$

*Then $(\beta_0, \Lambda_{0|\tau^*})$ is the unique solution of*

$$\max_{\beta \in B, \Lambda \in \mathcal{H}} \mathbb{E}\left[ \ell(Y^*, \Delta^*, X, Z; \beta, \Lambda_{|\tau^*}, \gamma_0) \right]. \tag{A4}$$

Condition (A3) is a minimal requirement of identification of the true value of the parameters in the model for the uncured subjects if the variable $T_0 \wedge C$ was observed and only the events in a subset of the support of $T_0$ are considered. In the Cox PH model (A3) is guaranteed by the requirement that $Var(Z)$ has full rank.

*Proof of Lemma 2.* First, let

$$H_k([0,t]|x,z) = \mathbb{P}(Y \leq t, \Delta = k | X = x, Z = z), \quad k \in \{0, 1\}, \quad t \in [0, \infty),$$

and let $H_k(dt|x,z)$ be the associated conditional measures. These conditional measures characterize the distribution of $(Y, \Delta)$ given $X = x$ and $Z = z$. By the model and independence assumptions, for any $t \geq 0$,

$$H_1(dt|x,z) = \phi(\gamma_0, x) F_C([t, \infty)|x,z) f_u(t|z; \beta_0, \Lambda_0) dt, \tag{A5}$$

and

$$H_0(dt|x,z) = \{1 - \phi(\gamma_0, x) + \phi(\gamma_0, x) S_u(t|z; \beta_0, \Lambda_0)\} F_C(dt|x,z). \tag{A6}$$

Following an usual notation abuse, herein we treat $dt$ not just as the length of a small interval but also as the name of the interval itself. Note that up to additive terms which do not depend on the parameters $\beta, \Lambda$,

$$(y, d) \mapsto d \log f_u(y|z; \beta_0, \Lambda_0) + (1-d) \log \{1 - \phi(\gamma_0, x) + \phi(\gamma_0, x) S_u(y|z; \beta_0, \Lambda_0)\},$$

is the conditional log-density of $(Y, \Delta)$ given $X = x$ and $Z = z$. From this and Kullback information inequality one can deduce that the expectation of $\ell$ defined in (13) is maximized by $\beta_0, \Lambda_0$ and $\gamma_0$.

Let $0 < \tau^* \leq \tau_0$. Note that

$$H_1([\tau^*, \tau_0]|x,z) = H_1([\tau^*, \infty)|x,z) = \phi(\gamma_0, x) \int_{[\tau^*, \tau_0]} F_C([t, \infty)|x,z) f_u(t|z; \beta_0, \Lambda_0) dt,$$

and

$$H_0([\tau^*, \infty)|x,z) = \phi(\gamma_0, x) \int_{[\tau^*, \tau_0]} S_u(t|z; \beta_0, \Lambda_0) F_C(dt|x,z) + \{1 - \phi(\gamma_0, x)\} F_C([\tau^*, \infty)|x,z).$$

Moreover,

$$\begin{aligned} d(x,z;\tau^*) := \phi(\gamma_0, x) &\int_{[\tau^*, \tau_0]} F_C([t, \infty)|x,z) f_u(t|z; \beta_0, \Lambda_0) dt \\ &+ \phi(\gamma_0, x) \int_{[\tau^*, \tau_0]} S_u(t|z; \beta_0, \Lambda_0) F_C(dt|x,z) \\ &= \phi(\gamma_0, x) F_C([\tau^*, \infty)|x,z) S_u(\tau^*|z; \beta_0, \Lambda_0) \\ &= \mathbb{P}(T_0 \wedge C \geq \tau^*, B = 1). \end{aligned}$$

In the limit case of no cure, $d(x,z;\tau^*) = H_1([\tau^*, \infty)|x,z) + H_0([\tau^*, \infty)|x,z)$. By construction we have $Y^* = \min(Y, \tau^*)$, and

$$\mathbb{P}(Y^* = \tau^*, \Delta^* = 1 | X = x, Z = z) = d(x,z;\tau^*).$$

Next, let

$$H_k^*([0,t]|x,z) = \mathbb{P}(Y^* \leq t, \Delta^* = k | X = x, Z = z), \quad k \in \{0,1\}, \quad t \in [0, \infty),$$

and let $H_k^*(dt|x,z)$ be the associated conditional measures. This means for any $t \in [0, \tau^*)$,

$$H_1^*(dt|x,z) = H_1(dt|x,z) \quad \text{and} \quad H_0^*(dt|x,z) = H_0(dt|x,z).$$

Moreover,

$$H_1^*(\{\tau^*\}|x,z) = H_1^*([\tau^*, \infty)|x,z) = d(x,z;\tau^*),$$

and

$$H_0^*(\{\tau^*\}|x,z) = H_0^*([\tau^*, \tau_0]|x,z) = \{1 - \phi(\gamma_0, x)\} F_C([\tau^*, \infty)|x,z).$$

Now, according to the inversion formulae of [6], without any reference to a model, one can solve the set of equations

$$\begin{aligned} H_1^*(dt|x,z) &= \phi^*(x,z) F_C^*([t, \infty)|x,z) F_u^*(dt|x,z), \\ H_0^*(dt|x,z) &= \{1 - \phi^*(x,z) + \phi^*(x,z) S_u^*(t|x,z)\} F_C^*(dt|x,z), \end{aligned} \tag{A7}$$

where $F_u^* = 1 - S_u^*$. Solving (A7) for $F_C^*$, $S_u^*$ and $\phi^*$, the functional $S_u^*$ is a proper survival function which puts mass only on sets where $H_1^*$ does. Note that solving the similar system with $H_1, H_0$ instead of $H_1^*, H_0^*$, one gets the true $F_C$, $S_u$ and $\phi$. If $\Lambda_C^*$ denotes the cumulative hazard function associated to the solution $F_C^*$, then

$$\Lambda_C^*(dt|x,z) = \frac{H_0^*(dt|x,z)}{H_1^*((t,\infty)|x,z) + H_0^*([t,\infty)|x,z)}, \quad t \geq 0,$$

and thus, by construction, we have $F_C(dt|x,z) = F_C^*(dt|x,z)$ on $[0,\tau^*)$, for any $x, z$. Then, by (A6) and the second equation in (A7) we deduce

$$\phi^*(x,z)F_u^*(t|x,z) = \phi(\gamma_0,x)F_u(t|z;\beta_0,\Lambda_0), \quad \forall t \in [0,\tau^*), \forall x,z.$$

Next, taking into account that $S_u^*(t|x,z) = 0$, $\forall t \geq \tau^*$, $\forall x,z$, and integrating the second equation (A7) on $[\tau^*,\infty)$, we obtain

$$\{1-\phi^*(x,z)\}F_C^*([\tau^*,\infty)|x,z) = H_0^*(\{\tau^*\}|x,z) = \{1-\phi(x,z)\}F_C([\tau^*,\infty)|x,z).$$

Since $F_C^*([0,\tau^*)|x,z) = F_C([0,\tau^*)|x,z)$, we deduce that $\phi^*(x,z) = \phi(\gamma_0,x)$ and thus

$$F_u^*(t|x,z) = F_u(t|z;\beta_0,\Lambda_0) = F_u(t|z;\beta_0,\Lambda_{0|\tau^*}), \quad \forall t \in [0,\tau^*), \forall x,z. \tag{A8}$$

The second equality in the last display is by the construction of the survival function from the cumulative hazard function: only the values of $\Lambda_0$ on $[0,t]$ contribute to obtain $F_u(t|z;\beta_0,\Lambda_0)$. Since the inversion formula necessarily yields $F_u^*([0,\tau^*]|x,z) \equiv 1$, we deduce

$$F_u^*(\{\tau^*\}|x,z) = S_u(\tau^*|z;\beta_0,\Lambda_0) = S_u(\tau^*|z;\beta_0,\Lambda_{0|\tau^*}). \tag{A9}$$

Finally, we can write

$$\begin{aligned} E\left[\ell(Y^*,\Delta^*,X,Z;\beta,\Lambda_{|\tau^*},\gamma_0)\right] &= \iiint \log \ell(t,1,x,z;\beta,\Lambda_{|\tau^*},\gamma_0)H_1^*(dt|x,z)G(dx,dz) \\ &\quad + \iiint \log \ell(t,0,x,z;\beta,\Lambda_{|\tau^*},\gamma_0)H_0^*(dt|x,z)G(dx,dz). \end{aligned}$$

To obtain the identifiability result it remains to apply Kullback information inequality. More precisely, it suffices to notice that here, up to additive terms which do not depend on the parameters, $\ell$ defined in (A2) considered with $\beta_0, \Lambda_{0|\tau^*}$ corresponds to the log-density of the conditional law of $(Y^*,\Delta^*)$ given $X = x$ and $Z = z$. (Note that the dominated measure changed as we introduce jumps at $\tau^*$.) This follows from (A8) and (A9). Thus $\beta_0, \Lambda_{0|\tau^*}$ is solution of the problem (A4). The unicity of the solution is guaranteed by (A3). □

### *A.2. Consistency*

Proof of Theorem 1.. We follow the idea of [7]. Since we are interested in almost sure convergence, we work with fixed realizations of the data, $\omega$ that will lie in a set of probability one. Let $\Omega$ be the abstract probability space where the random vector $(B, T_0, C, X, Z)$ is defined (for example we can take $\Omega = \{0,1\}\times[0,\tau_0]\times[0,\tau]\times\mathcal{X}\times\mathcal{Z}$ and $(B, T_0, C, X, Z)(\omega) = \omega$. Let $N \subset \Omega$ be a set of probability one $\mathbb{P}(N) = 1$ and fix $\omega \in N$. We will show that each subsequence $\hat{\gamma}_{n_k}$ has a subsequence that converges to $\gamma_0$. As a bounded sequence in $\mathbb{R}^p$, $\hat{\gamma}_{n_k}$ has a convergent subsequence $\hat{\gamma}_{m_k} \to \gamma^*$. It suffices to show that $\gamma^* = \gamma_0$. Since $\hat{\gamma}_{m_k}$ maximizes $\log \hat{L}_{m_k,1}$, we have

$$\begin{aligned} 0 &\leq \frac{1}{m_k}\log \hat{L}_{m_k,1}(\hat{\gamma}_{m_k}) - \frac{1}{m_k}\log \hat{L}_{m_k,1}(\gamma_0)\\ &= \frac{1}{m_k}\sum_{i=1}^{m_k}\left[\{1-\hat{\pi}(X_i)\}\log\frac{\phi(\hat{\gamma}_{m_k},X_i)}{\phi(\gamma_0,X_i)} + \hat{\pi}(X_i)\log\frac{1-\phi(\hat{\gamma}_{m_k},X_i)}{1-\phi(\gamma_0,X_i)}\right]\\ &= \frac{1}{m_k}\sum_{i=1}^{m_k}\left[\{1-\pi_0(X_i)\}\log\frac{\phi(\gamma^*,X_i)}{\phi(\gamma_0,X_i)} + \pi_0(X_i)\log\frac{1-\phi(\gamma^*,X_i)}{1-\phi(\gamma_0,X_i)}\right] + o(1) \end{aligned} \tag{A10}$$

if $N \subset \{\omega : \sup_x |\hat{\pi}(x) - \pi_0(x)| \to 0\}$. Note that the remainder term $o(1)$ in the previous display depends on $\omega$ and converges to zero as $\hat{\pi}$ converges to $\pi_0$. Next we will show that, for an appropriate choice of $N$, the first term converges to

$$\mathbb{E}\left[\{1-\pi_0(X)\}\log\frac{\phi(\gamma^*,X)}{\phi(\gamma_0,X)} + \pi_0(x)\log\frac{1-\phi(\gamma^*,X)}{1-\phi(\gamma_0,X)}\right] \tag{A11}$$

where the expectation is taken with respect to $X$ and $\gamma^* \in \mathbb{R}^p$ (for a fixed $\omega$). Since here we are dealing with a simple parametric model, this convergence follows easily from the uniform law of large numbers. However, we follow a longer argument to explain the idea that will be used also in the proof of Theorem 2 (where the model is semiparametric). It is obvious, by the law of large numbers, that

$$\begin{aligned} &\frac{1}{m_k}\sum_{i=1}^{m_k}\left[\{1-\pi_0(X_i)\}\log\phi(\gamma_0,X_i) + \pi_0(X_i)\log\left(1-\phi(\gamma_0,X_i)\right)\right]\\ &\to \mathbb{E}\left[\{1-\pi_0(X)\}\log\phi(\gamma_0,X) + \pi_0(x)\log\left(1-\phi(\gamma_0,X)\right)\right] \text{ a.s.} \end{aligned}$$

and, at first sight it seems that the same holds when $\gamma_0$ is replaced by $\gamma^*$. However, the proof is more delicate because $\gamma^*$ depends on $\omega$ and thus also the event of probability one where the strong law of large numbers holds for this average. To avoid this we consider a countable dense subset of $G$, $\{\tilde{\gamma}_l\}_{l\geq 1}$ (for example the subset for which all components of $\gamma$ are rational numbers). Now, consider the countable collection of the probability one sets $\{N_l\}_{l\geq 1}$ where

$$\begin{aligned}&\frac{1}{m_k}\sum_{i=1}^{m_k}\left[\{1-\pi_0(X_i)\}\log\phi(\tilde{\gamma}_l,X_i)+\pi_0(X_i)\log\left(1-\phi(\tilde{\gamma}_l,X_i)\right)\right]\\&\to\mathbb{E}\left[\{1-\pi_0(X)\}\log\phi(\tilde{\gamma}_l,X)+\pi_0(x)\log\left(1-\phi(\tilde{\gamma}_l,X)\right)\right]\quad\forall l\geq 1.\end{aligned}$$

If $N\subseteq(\cap_{l\geq 1}N_l)$, we can write

$$\begin{aligned}&\left|\frac{1}{m_k}\sum_{i=1}^{m_k}\left[\{1-\pi_0(X_i)\}\log\phi(\gamma^*,X_i)+\pi_0(X_i)\log\left(1-\phi(\gamma^*,X_i)\right)\right]\right.\\&\qquad\left.-\,\mathbb{E}\left[\{1-\pi_0(X)\}\log\phi(\gamma^*,X)+\pi_0(x)\log\left(1-\phi(\gamma^*,X)\right)\right]\right|\\&\leq\left|\frac{1}{m_k}\sum_{i=1}^{m_k}\left[\{1-\pi_0(X_i)\}\log\frac{\phi(\gamma^*,X_i)}{\phi(\tilde{\gamma}_l,X_i)}+\pi_0(X_i)\log\frac{(1-\phi(\gamma^*,X_i))}{(1-\phi(\tilde{\gamma}_l,X_i))}\right]\right|\\&\quad+\left|\frac{1}{m_k}\sum_{i=1}^{m_k}\left[\{1-\pi_0(X_i)\}\log\phi(\tilde{\gamma}_l,X_i)+\pi_0(X_i)\log\left(1-\phi(\tilde{\gamma}_l,X_i)\right)\right]\right.\\&\qquad\left.-\,\mathbb{E}\left[\{1-\pi_0(X)\}\log\phi(\tilde{\gamma}_l,X)+\pi_0(x)\log\left(1-\phi(\tilde{\gamma}_l,X)\right)\right]\right|\\&\quad+\left|\mathbb{E}\left[\{1-\pi_0(X)\}\log\frac{\phi(\gamma^*,X)}{\phi(\tilde{\gamma}_l,X)}+\pi_0(x)\log\frac{1-\phi(\gamma^*,X)}{1-\phi(\tilde{\gamma}_l,X)}\right]\right|.\end{aligned}$$

Since $\tilde{\gamma}_l$ can be taken arbitrarily close to $\gamma^*$, by properties of $\phi$ in assumptions (AC3)-(AC4), it can be easily derived that, for an appropriate choice of $\tilde{\gamma}_l$, the first and the third term on the right hand side in the previous equation converge to zero. Moreover, the second term also converges to zero in the set of probability one that we are considering. As a result, we can conclude that

$$\begin{aligned}0&\leq\frac{\log\hat{L}_{m_k,1}(\hat{\gamma}_{m_k})-\log\hat{L}_{m_k,1}(\gamma_0)}{m_k}\\&=\mathbb{E}\left[\{1-\pi_0(X)\}\log\frac{\phi(\gamma^*,X)}{\phi(\gamma_0,X)}+\pi_0(x)\log\frac{1-\phi(\gamma^*,X)}{1-\phi(\gamma_0,X)}\right]+o(1)\end{aligned}$$

For each $x\in\mathcal{X}$, consider the function

$$g_x(z)=\phi(\gamma_0,x)\log\frac{z}{\phi(\gamma_0,x)}+\{1-\phi(\gamma_0,x)\}\log\frac{1-z}{1-\phi(\gamma_0,x)},\quad z\in(0,1).$$

It is easy to check that $g_x(z)\leq 0$ and the equality holds only if $z=\phi(\gamma_0,x)$. Hence, the expectation in (A11) is smaller or equal to zero. Due to the inequality in (A10), it must be equal to zero, which means that $\phi(\gamma^*,X)=\phi(\gamma_0,X)$. By the identifiability assumption (5), this is possible only if $\gamma^*=\gamma_0$. □

**Lemma 3.** *Assume (AC2),(AC5) hold and $\tau^*$ is such that* (20) *is satisfied. Then $\sup_n\hat{\Lambda}_n(\tau^*)<\infty$ almost surely.*

*Proof.* By definition

$$\hat{\Lambda}_n(\tau^*) = \frac{1}{n}\sum_{i=1}^n \frac{\Delta_i \mathbb{1}_{\{Y_i<\tau^*\}}}{\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{Y_i\leq Y_j\leq \tau_0\}} \exp(\hat{\beta}_n' Z_j)\left\{\Delta_j + (1-\Delta_j) g_j(Y_j,\hat{\Lambda}_n,\hat{\beta}_n,\hat{\gamma}_n)\right\}}.$$

From assumptions (AC2) and (AC5) we have

$$\begin{aligned}
&\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\tau^*\leq Y_j\leq \tau_0\}} \exp(\hat{\beta}_n' Z_j)\left\{\Delta_j + (1-\Delta_j) g_j(Y_j,\hat{\Lambda}_n,\hat{\beta}_n,\hat{\gamma}_n)\right\}\\
&\geq \frac{1}{n}\sum_{j=1}^n \Delta_j \mathbb{1}_{\{\tau^*\leq Y_j\leq \tau_0\}} \exp(\hat{\beta}_n' Z_j)\\
&\geq c\frac{1}{n}\sum_{j=1}^n \Delta_j \mathbb{1}_{\{\tau^*\leq Y_j\leq \tau_0\}},
\end{aligned}$$

for some $c>0$. Since $\frac{1}{n}\sum_{j=1}^n \Delta_j \mathbb{1}_{\{\tau^*\leq Y_j\leq \tau_0\}} \xrightarrow{a.s.} \mathbb{P}\left(Y\geq \tau^*, \Delta=1\right)>0$, it follows that $\frac{1}{n}\sum_{j=1}^n \Delta_j \mathbb{1}_{\{\tau^*\leq Y_j\leq \tau_0\}}$ is bounded from below away from zero almost everywhere. As a result

$$\begin{aligned}
\sup_n \hat{\Lambda}_n(\tau^*) &\leq \sup_n \frac{1}{n}\sum_{i=1}^n \frac{\Delta_i \mathbb{1}_{\{Y_i<\tau^*\}}}{\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\tau^*\leq Y_j\leq \tau_0\}} \exp(\hat{\beta}_n' Z_j)\left\{\Delta_j + (1-\Delta_j) g_j(Y_j,\hat{\Lambda}_n,\hat{\beta}_n,\hat{\gamma}_n)\right\}}\\
&\leq \sup_n \frac{1}{n}\sum_{i=1}^n \frac{\Delta_i \mathbb{1}_{\{Y_i<\tau_0\}}}{c\frac{1}{n}\sum_{j=1}^n \Delta_j \mathbb{1}_{\{\tau^*\leq Y_j\leq \tau_0\}}}\\
&\leq \frac{1}{c}\left(\inf_n \frac{1}{n}\sum_{j=1}^n \Delta_j \mathbb{1}_{\{\tau^*\leq Y_j\leq \tau_0\}}\right)^{-1}
\end{aligned}$$

is bounded almost surely. Note that, if (19) is satisfied, then we can take $\tau^* = \tau_0$. □

Proof of Theorem 2.. Let $0<\tau^*\leq\tau_0$ and

$$\hat{l}_n^*(\beta,\Lambda_{|\tau^*},\hat{\gamma}_n) = \frac{1}{n}\sum_{i=1}^n \ell(Y_i^*,\Delta_i^*,X_i,Z_i;\beta,\Lambda_{|\tau^*},\hat{\gamma}_n),$$

with $\ell$ defined in (A2). If we consider the Cox PH model for the conditional law of $T_0$, then

$$\begin{aligned}
\hat{l}_n^*(\beta,\Lambda_{|\tau^*},\hat{\gamma}_n) &= \frac{1}{n}\sum_{i=1}^n \Delta_i\left[\mathbb{1}_{\{Y_i<\tau^*\}}\left\{\log\Delta\Lambda(Y_i) + \beta' Z_i - \Lambda(Y_i)e^{\beta' Z_i}\right\}\right]\\
&+\frac{1}{n}\sum_{i=1}^n (1-\Delta_i)\mathbb{1}_{\{Y_i<\tau^*\}}\log\left\{1-\phi(\hat{\gamma}_n,X_i)+\phi(\hat{\gamma}_n,X_i)\exp\left(-\Lambda(Y_i)e^{\beta' Z_i}\right)\right\}\\
&-\frac{\Lambda(\tau^*-)}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i\geq\tau^*\}}\mathbb{1}_{\{B_i=1\}}e^{\beta' Z_i}+\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i\geq\tau^*\}}\mathbb{1}_{\{B_i=0\}}\log\left\{1-\phi(\hat{\gamma}_n,X_i)\right\},
\end{aligned}$$

and has to be maximized with respect to $\beta$ and $\Lambda$ in the class of step functions $\Lambda$ with jumps of size $\Delta\Lambda$ at the event times in $[0,\tau^*)$. As in [5], it can be shown that the maximizer $(\hat{\Lambda}_n^*, \hat{\beta}_n^*)$ of $\hat{l}_n^*$ exists and it is finite. Moreover, for $t \in [0,\tau_0]$, $\hat{\Lambda}_n^* = \Lambda^*_{n,\hat{\beta}_n^*,\hat{\gamma}_n}$ where

$$\Lambda^*_{n,\beta,\gamma}(t) = \frac{1}{n}\sum_{i=1}^n \frac{\Delta_i \mathbb{1}_{\{Y_i \le t, Y_i < \tau^*\}}}{\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{Y_j \ge Y_i\}} \exp(\beta' Z_j)\left\{\Delta_j^* + (1-\Delta_j)\mathbb{1}_{\{Y_j<\tau^*\}} g_j(Y_j, \Lambda^*_{n,\beta,\gamma}, \beta, \gamma)\right\}},$$

$\Delta^* = \mathbb{1}_{\{T_0^* \le C^*\}} = \Delta \mathbb{1}_{\{Y_j < \tau^*\}} + \mathbb{1}_{\{Y_i \ge \tau^*\}}\mathbb{1}_{\{B_i=1\}}$ and $g_j(t,\Lambda,\beta,\gamma)$ defined in (17).

Let

$$\tilde{\Lambda}_{0,n}(t) = \frac{1}{n}\sum_{i=1}^n \frac{\Delta_i \mathbb{1}_{\{Y_i \le t, Y_i < \tau_0\}}}{\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{Y_j \ge Y_i, Y_j \le \tau_0\}} \exp(\beta_0' Z_j)\left\{\Delta_j + (1-\Delta_j) g_j(Y_j, \Lambda_0, \beta_0, \gamma_0)\right\}}. \tag{A12}$$

We want to prove that $\hat{\beta}_n \xrightarrow{a.s.} \beta_0$, and $\sup_{t\in[0,\bar{\tau}]} |\hat{\Lambda}_n(t) - \Lambda_0(t)| \xrightarrow{a.s.} 0$ for any $\bar{\tau} < \tau_0$. We suppose that the previous statement is false, i.e $\hat{\beta}_n$ does not converge almost surely to $\beta_0$ or there exists $\bar{\tau}$ such that $\sup_{t\in[0,\bar{\tau}]} |\hat{\Lambda}_n(t) - \Lambda_0(t)|$ does not converge to zero almost surely. This means that, there exist $\epsilon > 0$ and $\bar{\tau} < \tau_0$ such that

$$\mathbb{P}[A_1(\bar{\tau},\epsilon)] > 0, \quad \text{with} \quad A_1(\bar{\tau},\epsilon) = \left\{ \limsup_{n\to\infty} \left[ \left\| \hat{\beta}_n - \beta_0 \right\| + \sup_{t\in[0,\bar{\tau}]} \left| \hat{\Lambda}_n(t) - \Lambda_0(t) \right| \right] > \epsilon \right\}.$$

On the other hand, since $(\hat{\Lambda}_n, \hat{\beta}_n)$ maximizes $\hat{\ell}_n(\Lambda, \beta, \hat{\gamma}_n)$, for any realization $\omega$ of the data we have

$$\hat{l}_n(\hat{\beta}_n, \hat{\Lambda}_n, \hat{\gamma}_n) - \hat{l}_n(\beta_0, \tilde{\Lambda}_{0,n}, \hat{\gamma}_n) \ge 0. \tag{A13}$$

Then the idea for creating the contradiction is to show that the previous inequality is not satisfied for any $\omega$ in some event of positive probability. We argue for a fixed realization $\omega$ of the data. As a bounded sequence in $\mathbb{R}^q$, $\hat{\beta}_n$ has a convergent subsequence $\hat{\beta}_{n_k} \to \bar{\beta}$. Let $(\tau_i)_{i\ge 1}$ be an increasing sequence such that $\lim_{i\to\infty} \tau_i = \tau_0$. Since for all $\tau < \tau_0$, $\hat{\Lambda}_n(\tau) < \infty$ almost surely (see Lemma 3), by Helly's selection theorem ([1]), there exists a subsequence $\hat{\Lambda}_{m_k}$ of $\hat{\Lambda}_{n_k}$, converging pointwise to a function $\bar{\Lambda}$ on $[0,\tau_1]$. Repeating the same argument, we can extract a further subsequence converging pointwise to a function $\bar{\Lambda}$ on $[0,\tau_2]$ and so on. Hence, there exist a subsequence $\hat{\Lambda}_{r_k}$ converging pointwise to a function $\bar{\Lambda}$ on all compacts of $[0,\tau_0]$ that do not include $\tau_0$. This defines a monotone function $\bar{\Lambda}$ on $[0,\tau_0)$, which could be extended at $\tau_0$ by taking the limit. As in Lemma 2 of [5], it can be shown that $\bar{\Lambda}$ is absolutely continuous and pointwise convergence of monotone functions to a continuous monotone function implies uniform convergence on compacts. Note that the chosen subsequence and the limits $\bar{\beta}$ and $\bar{\Lambda}$ depend on $\omega$. To keep the notation simple, in what follows we use the index $n$ instead of the chosen subsequence $r_k$. For any $\tau^* < \tau_0$, we can write

$$\begin{aligned} 0 &\le \hat{l}_n(\hat{\beta}_n, \hat{\Lambda}_n, \hat{\gamma}_n) - \hat{l}_n(\beta_0, \tilde{\Lambda}_{0,n}, \hat{\gamma}_n) \\ &= \hat{l}^*_n(\hat{\beta}_n, \hat{\Lambda}_{n|\tau^*}, \hat{\gamma}_n) + D_{1n} - \hat{l}^*_n(\beta_0, \tilde{\Lambda}_{0,n|\tau^*}, \hat{\gamma}_n) - D_{2n} \\ &= \mathbb{E}[\ell(Y^*, \Delta^*, X, Z; \bar{\beta}, \bar{\Lambda}_{|\tau^*}, \gamma_0)] + D_{1n} + R_{1n} \\ &\quad - \mathbb{E}[\ell(Y^*, \Delta^*, X, Z; \beta_0, \Lambda_{0|\tau^*}, \gamma_0)] - D_{2n} - R_{2n}, \end{aligned} \tag{A14}$$

where

$$D_{1n} = \hat{l}_n(\hat{\beta}_n, \hat{\Lambda}_n, \hat{\gamma}_n) - \hat{l}^*_n(\hat{\beta}_n, \hat{\Lambda}_{n|\tau^*}, \hat{\gamma}_n), \tag{A15}$$

$$D_{2n} = \hat{l}_n(\beta_0, \tilde{\Lambda}_{0,n}, \hat{\gamma}_n) - \hat{l}^*_n(\beta_0, \tilde{\Lambda}_{0,n|\tau^*}, \hat{\gamma}_n), \tag{A16}$$

$$R_{1n} = \hat{l}^*_n(\hat{\beta}_n, \hat{\Lambda}_{n|\tau^*}, \hat{\gamma}_n) - \mathbb{E}[\ell(Y^*, \Delta^*, X, Z; \bar{\beta}, \bar{\Lambda}_{|\tau^*}, \gamma_0)], \tag{A17}$$

$$R_{2n} = \hat{l}^*_n(\beta_0, \tilde{\Lambda}_{0,n|\tau^*}, \hat{\gamma}_n) - \mathbb{E}[\ell(Y^*, \Delta^*, X, Z; \beta_0, \Lambda_{0|\tau^*}, \gamma_0)]. \tag{A18}$$

Note that the limit of $(\hat{\beta}_n, \hat{\Lambda}_n)$ depends on $\omega$, but here the expectation is taken with respect to $(Y^*, \Delta^*, X, Z)$ for fixed $(\bar{\beta}, \bar{\Lambda})$. We now define the event $A_3(\tau^*) = \{|R_{1n} - R_{2n}| \to 0\}$. By Lemma 4, for any $\tau^* < \tau_0$, we have $\mathbb{P}[A_1(\bar{\tau}, \epsilon) \cap A_3(\tau^*)] = \mathbb{P}[A_1(\bar{\tau}, \epsilon)]$. Next, for $\bar{\tau} < \tau_0$ and $\epsilon > 0$ such that $\mathbb{P}[A_1(\bar{\tau}, \epsilon)] > 0$, by Lemma 6 there exist $0 < c_1 < 1$ and $\delta > 0$ such that we have

$$c = \inf \Bigg\{ \mathbb{E}[\ell(Y^*, \Delta^*, X, Z; \beta_0, \Lambda_{0|\tau^*}, \gamma_0)] - \mathbb{E}[\ell(Y^*, \Delta^*, X, Z; \beta, \Lambda_{|\tau^*}, \gamma_0)] : \\ \bar{\tau} + \delta \le \tau^* < \tau_0, \quad \|\beta - \beta_0\| \ge c_1\epsilon/2 \quad \text{or} \quad \sup_{t \in [0, \bar{\tau}]} |\Lambda(t) - \Lambda_0(t)| \ge (1 - c_1)\epsilon/2 \Bigg\} > 0.$$

Note that if $\omega \in A_1(\bar{\tau}, \epsilon)$ and $\bar{\beta}$ and $\bar{\Lambda}$ are the limits for $\hat{\beta}_n$ and $\hat{\Lambda}_n$, respectively, then necessarily, either $\|\bar{\beta} - \beta_0\| \ge c_1\epsilon/2$, or $\sup_{t \in [0, \bar{\tau}]} |\bar{\Lambda}(t) - \Lambda_0(t)| \ge (1 - c_1)\epsilon/2$, and consequently

$$\mathbb{E}[\ell(Y^*, \Delta^*, X, Z; \bar{\beta}, \bar{\Lambda}_{|\tau^*}, \gamma_0)] - \mathbb{E}[\ell(Y^*, \Delta^*, X, Z; \beta_0, \Lambda_{0|\tau^*}, \gamma_0)] \ge c, \quad \forall \bar{\tau} + \delta \le \tau^* < \tau_0.$$

Finally, we define

$$A_2(\tau^*) = \left\{ \limsup_{n \to \infty} |D_{1n} - D_{2,n}| \le c/2 \right\},$$

with $D_{1n}$ and $D_{2n}$ defined in (A15) and (A16), and choose $\tau^* \in [\bar{\tau} + \delta, \tau_0)$ such that

$$c_b \left\{ \mathbb{P}(T_0 \ge \tau^*) \log\{1/\mathbb{P}(T_0 \ge \tau^*)\} + \mathbb{P}(C \in [\tau^*, \tau_0]) \right\} < c/2,$$

with $c_b$ the constant from Lemma 5. Then we have $\mathbb{P}[A_2(\tau^*)] = 1$. Gathering facts, we deduce that by a suitable choice of $\tau^* \in [\bar{\tau} + \delta, \tau_0)$, we necessarily have

$\mathbb{P}[A_1(\bar{\tau},\epsilon)\cap A_2(\tau^*)\cap A_3(\tau^*)] > 0$. Moreover, with such a suitable $\tau^*$, for any $\omega \in A_1(\bar{\tau},\epsilon)\cap A_2(\tau^*)\cap A_3(\tau^*)$, we have

$$\limsup_{n\to\infty}\left[\hat{l}_n(\hat{\beta}_n,\hat{\Lambda}_n,\hat{\gamma}_n)-\hat{l}_n(\beta_0,\tilde{\Lambda}_{0,n},\hat{\gamma}_n)\right]\leq -c/2<0.$$

We deduce that (A13) is violated on an event of positive probability, which by definition is impossible. Thus $\hat{\beta}_n\xrightarrow{a.s.}\beta_0$, and $\sup_{t\in[0,\bar{\tau}]}|\hat{\Lambda}_n(t)-\Lambda_0(t)|\xrightarrow{a.s.}0$ for any $\bar{\tau}<\tau_0$.

If condition (19) is satisfied, we want to show in addition that $|\hat{\Lambda}_n(\tau_0)-\Lambda_0(\tau_0)|\xrightarrow{a.s.}0$. In that case, $\hat{\Lambda}_n(\tau_0)<\infty$ almost surely and as a result, for any realization $\omega$, there exists a subsequence $\hat{\Lambda}_{r_k}$ converging to some absolutely continuous function $\bar{\Lambda}$ uniformly on $[0,\tau_0]$. Since we already showed that $|\hat{\Lambda}_n(t)-\Lambda_0(t)|\xrightarrow{a.s.}0$ for any $t<\tau_0$ and $\Lambda_0(\tau_0)=\lim_{t\uparrow\tau_0}\Lambda_0(t)$, we necessarily have $\bar{\Lambda}=\Lambda_0$ on the whole interval $[0,\tau_0]$. This concludes the proof of the Theorem. □

**Lemma 4.** *Consider a realization of the data $\omega$ and assume that $\hat{\beta}_n(\omega)\to\bar{\beta}$ and $\hat{\Lambda}_n(\omega)(t)\to\bar{\Lambda}(t)$ for any $t\in[0,\tau_0)$, for some absolutely continuous function $\bar{\Lambda}$. Let $0<\tau^*<\tau_0$ and let $R_{1n}$, $R_{2n}$ be defined as in* (A17) *and* (A18)*, respectively. There exists an event $A_3(\tau^*)$ of probability one such that, for any $\omega\in A_3(\tau^*)$,*

$$R_{1n}(\omega)-R_{2n}(\omega)\to 0.$$

*Proof.* Let us consider some $0<\tau^*<\tau_0$. From Theorem 1 and Lemma 2 in [5] it follows that the event

$$A_3^1(\tau^*)=\left\{\hat{\gamma}_n\to\gamma_0 \quad\text{and}\quad \sup_{t\in[0,\tau^*]}|\tilde{\Lambda}_{0,n}(t)-\Lambda_0(t)|\to 0\right\}$$

has probability one. Next we argue for the given realization of the data $\omega\in A_3^1(\tau^*)$ and will determine the event $A_3(\tau^*)$ appropriately. By the triangular inequality we can write

$$\begin{aligned}|R_{1n}-R_{2n}|\leq&\left|\left\{\hat{l}_n^*(\hat{\beta}_n,\hat{\Lambda}_{n|\tau*},\hat{\gamma}_n)-\hat{l}_n^*(\beta_0,\tilde{\Lambda}_{0,n|\tau*},\hat{\gamma}_n)\right\}-\left\{\hat{l}_n^*(\bar{\beta},\bar{\Lambda}_{|\tau*},\gamma_0)-\hat{l}_n^*(\beta_0,\Lambda_{0|\tau*},\gamma_0)\right\}\right|\\&+\left|\hat{l}_n^*(\beta_0,\Lambda_{0|\tau*},\gamma_0)-\mathbb{E}\left[l(Y^*,\Delta^*,X,Z;\beta_0,\Lambda_{0|\tau^*},\gamma_0)\right]\right|\\&+\left|\hat{l}_n^*(\bar{\beta},\bar{\Lambda}_{|\tau*},\gamma_0)-\mathbb{E}\left[l(Y^*,\Delta^*,X,Z;\bar{\beta},\bar{\Lambda}_{|\tau^*},\gamma_0)\right]\right|.\end{aligned}\quad\text{(A19)}$$

Since $\bar{\Lambda}$ is absolutely continuous, it is differentiable almost everywhere. Let

$\bar{\lambda}(t) = \mathrm{d}\bar{\Lambda}(t)/\mathrm{d}t$. By definition we have

$$
\begin{aligned}
&\hat{l}_n^*(\hat{\beta}_n, \hat{\Lambda}_{n|\tau*}, \hat{\gamma}_n) - \hat{l}_n^*(\beta_0, \tilde{\Lambda}_{0,n|\tau*}, \hat{\gamma}_n) \\
&= \frac{1}{n}\sum_{i=1}^n \Delta_i \mathbb{1}_{\{Y_i<\tau^*\}} \left\{ \log \frac{\Delta\hat{\Lambda}_n(Y_i)}{\Delta\tilde{\Lambda}_{0,n}(Y_i)} + (\hat{\beta}_n - \beta_0)' Z_i - \hat{\Lambda}_n(Y_i) e^{\hat{\beta}_n' Z_i} + \tilde{\Lambda}_{0,n}(Y_i) e^{\beta_0' Z_i} \right\} \\
&\quad + \frac{1}{n}\sum_{i=1}^n (1-\Delta_i)\mathbb{1}_{\{Y_i<\tau^*\}} \log \frac{1-\phi(\hat{\gamma}_n, X_i) + \phi(\hat{\gamma}_n, X_i)\exp\left(-\hat{\Lambda}_n(Y_i) e^{\hat{\beta}_n' Z_i}\right)}{1-\phi(\hat{\gamma}_n, X_i) + \phi(\hat{\gamma}_n, X_i)\exp\left(-\tilde{\Lambda}_{0,n}(Y_i) e^{\beta_0' Z_i}\right)} \\
&\quad - \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i\geq\tau^*\}} \mathbb{1}_{\{B_i=1\}} \left\{ \hat{\Lambda}_n(\tau^*) e^{\hat{\beta}_n' Z_i} - \tilde{\Lambda}_{0,n}(\tau^*) e^{\beta_0' Z_i} \right\}.
\end{aligned}
$$

If $\omega \in A_3^1(\tau^*)$, we obtain

$$
\begin{aligned}
&\hat{l}_n^*(\hat{\beta}_n, \hat{\Lambda}_{n|\tau*}, \hat{\gamma}_n) - \hat{l}_n^*(\beta_0, \tilde{\Lambda}_{0,n|\tau*}, \hat{\gamma}_n) \\
&= \frac{1}{n}\sum_{i=1}^n \Delta_i \mathbb{1}_{\{Y_i<\tau^*\}} \left\{ \log \frac{\bar{\lambda}(Y_i)}{\lambda_0(Y_i)} + (\bar{\beta} - \beta_0)' Z_i - \bar{\Lambda}(Y_i) e^{\bar{\beta}' Z_i} + \Lambda_0(Y_i) e^{\beta_0' Z_i} \right\} \\
&\quad + \frac{1}{n}\sum_{i=1}^n (1-\Delta_i)\mathbb{1}_{\{Y_i<\tau^*\}} \log \frac{1-\phi(\gamma_0, X_i) + \phi(\gamma_0, X_i)\exp\left(-\bar{\Lambda}(Y_i) e^{\bar{\beta}' Z_i}\right)}{1-\phi(\gamma_0, X_i) + \phi(\gamma_0, X_i)\exp\left(-\Lambda_0(Y_i) e^{\beta_0' Z_i}\right)} \\
&\quad - \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i\geq\tau^*\}} \mathbb{1}_{\{B_i=1\}} \left\{ \bar{\Lambda}(\tau^*) e^{\bar{\beta}' Z_i} - \Lambda_0(\tau^*) e^{\beta_0' Z_i} \right\} + o(1) \\
&= \hat{l}_n^*(\bar{\beta}, \bar{\Lambda}_{|\tau*}, \gamma_0) - \hat{l}_n^*(\beta_0, \Lambda_{0|\tau*}, \gamma_0) + o(1),
\end{aligned}
$$

where the remainder term depends on $\omega$ and converges to zero. Hence, the first term on the right hand side of (A19) converges to zero. Let $A_n^2(\tau^*)$ be the event where

$$
\hat{l}_n^*(\beta_0, \Lambda_{0|\tau*}, \gamma_0) \to \mathbb{E}\left[ l(Y^*, \Delta^*, X, Z; \beta_0, \Lambda_{0|\tau^*}, \gamma_0) \right] \qquad \text{as } n \to \infty.
$$

By the law of large numbers $\mathbb{P}[A_n^2(\tau^*)] = 1$, implying that also the second term on the right hand side of (A19) converges to zero if $\omega \in A_n^2(\tau^*)$. It remains to deal with the third term. Note that here $(\bar{\beta}, \bar{\Lambda})$ depend on $\omega$ and the expectation is taken with respect to $(Y, \Delta, X, Z)$ for fixed $(\bar{\beta}, \bar{\Lambda})$. We have the same issue as in the proof of Theorem 1 when dealing with the terms involving $\bar{\beta}$ and $\bar{\Lambda}$, so we need to consider approximations by elements of a countable dense subset of $\mathcal{B}$ and of the space of bounded, absolutely continuous, increasing functions in $[0, \tau^*]$ (is separable, so such subset exists). The same reasoning is used also in [4, 5, 7]. Hence, there exists a countable collection of probability one sets $\{N_l\}_{l\geq 1}$ where

$$
\hat{l}_n^*(\beta_l, \Lambda_l, \gamma_0) \to \mathbb{E}\left[ l(Y^*, \Delta^*, X, Z; \beta_l, \Lambda_l, \gamma_0) \right] \qquad \text{as } n \to \infty
$$

and $(\beta_l, \Lambda_l)$ can be taken arbitrarily close to $(\bar{\beta}, \bar{\Lambda})$. As a result, if $\omega \in A_n^3(\tau^*) = \bigcap_{l\geq 1} N_l$, then

$$\left| \hat{l}_n^*(\bar{\beta}, \bar{\Lambda}_{|\tau*}, \gamma_0) - \mathbb{E}\left[ l(Y^*, \Delta^*, X, Z; \bar{\beta}, \bar{\Lambda}_{|\tau^*}, \gamma_0) \right] \right| \to 0.$$

To conclude, we define $A_3(\tau^*) = A_3^1(\tau^*) \cap A_3^2(\tau^*) \cap A_3^3(\tau^*)$ and we have $\mathbb{P}[A_3(\tau^*)] = 1$. □

**Lemma 5.** *Let $D_{1n}$ and $D_{2n}$ be defined as in* (A15) *and* (A16)*, respectively, for some $\tau^* < \tau_0$. Then there exists a constant $c_b$ independent of $\tau^*$ such that*

$$\mathbb{P}\left[ \limsup_{n\to\infty} |D_{1n} - D_{2,n}| > c_b \left\{ \mathbb{P}(T_0 \geq \tau^*) \log\{1/\mathbb{P}(T_0 \geq \tau^*)\} + \mathbb{P}(C \in [\tau^*, \tau_0]) \right\} \right] = 0.$$

*Proof.* By definition, for any $\gamma$, $\beta$ and cumulative hazard function $\Lambda$ piecewise constant with jumps at the observed events

$$\begin{aligned}
&l_n(\beta, \Lambda, \gamma) - \hat{l}_n^*(\beta, \Lambda_{|\tau^*}, \gamma) \\
&= \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{\tau^* \leq Y_i < \tau_0\}} \Delta_i \log \Lambda(\{Y_i\}) \\
&\quad + \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{\tau^* \leq Y_i < \tau_0\}} \Delta_i \beta' Z_i \\
&\quad - \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i \geq \tau^*\}} e^{\beta' Z_i} \{\Delta_i \Lambda(Y_i) - \mathbb{1}_{\{B_i = 1\}} \Lambda(\tau^* -)\} \\
&\quad + \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{\tau_0 \geq Y_i \geq \tau^*\}} \Big[ (1-\Delta_i) \log\Big\{ 1 - \phi_i(\gamma) + \phi_i(\gamma) \exp\left(-\Lambda(Y_i) e^{\beta' Z_i}\right) \Big\} \\
&\quad\quad - \mathbb{1}_{\{B_i = 0\}} \log\{1 - \phi_i(\gamma)\} \Big] \\
&=: r_{1n}(\Lambda; \tau^*) + r_{2n}(\beta; \tau^*) - r_{3n}(\Lambda, \beta; \tau^*) + r_{4n}(\Lambda, \beta, \gamma; \tau^*),
\end{aligned}$$

where $\phi_i(\gamma)$ is a short notation for $\phi(\gamma, X_i)$. For proving the Lemma, we have to suitably bound $r_{1n}, \ldots, r_{4n}$. For this purpose, let us notice that, by definition, all the cumulative hazard functions we have to consider ($\hat{\Lambda}_n$, $\tilde{\Lambda}_{0,n}$,...) have bounded jumps at the event times. More precisely, because the parameter space $\mathcal{B}$ and $Z$ are supposed bounded, there exist constants $0 < c_l < c_u$ such that

$$c_l \leq \exp(\beta' Z) \leq c_u.$$

Then the largest jump of any of the cumulative hazard functions we need to consider is bounded by $1/c_l$ (which is located at the last uncensored observation), the second largest one (and is located at the before last uncensored observation) is bounded by $1/2c_l$,...

To control $r_{1n}(\Lambda; \tau^*)$, one would look for a suitable lower bound forthe jumps of $\Lambda$. However, no meaningful lower bound could be derived for these jumps. More precisely, such a bound is necessarily of order $1/n$, so that the sequence of the logarithm of the jumps is unbounded. Fortunately, for our purposes it suffices to find a bound for

$$\left|r_{1n}(\hat{\Lambda}_n;\tau^*) - r_{1n}(\tilde{\Lambda}_{0,n};\tau^*)\right| = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{\tau^*\leq Y_i<\tau_0\}}\Delta_i \left|\log\frac{\hat{\Lambda}_n(\{Y_i\})}{\tilde{\Lambda}_{0,n}(\{Y_i\})}\right|,$$

where

$$\frac{\hat{\Lambda}_n(\{Y_i\})}{\tilde{\Lambda}_{0,n}(\{Y_i\})} = \frac{\sum_{j=1}^n \mathbb{1}_{\{\tau_0\geq Y_j\geq Y_i\}}\exp(\beta_0' Z_j)\left\{\Delta_j + (1-\Delta_j)g_j(Y_j,\Lambda_0,\beta_0,\hat{\gamma}_n)\right)\}}{\sum_{j=1}^n \mathbb{1}_{\{\tau_0\geq Y_j\geq Y_i\}}\exp(\beta' Z_j)\left\{\Delta_j + (1-\Delta_j)g_j(Y_j,\Lambda_{n,\hat{\beta},\hat{\gamma}_n},\beta,\hat{\gamma}_n)\right\}}.$$

Since all $g_j$'s are between 0 and 1, it is easy to see that for any uncensored $Y_i\geq\tau^*$,

$$\frac{1}{\rho_n(Y_i)}\frac{c_l}{c_u} \leq \frac{\hat{\Lambda}_n(\{Y_i\})}{\tilde{\Lambda}_{0,n}(\{Y_i\})} \leq \frac{c_u}{c_l}\rho_n(Y_i),$$

where

$$\rho_n(t) = \frac{\sum_{j=1}^n \mathbb{1}_{\{\tau_0\geq Y_j\geq t\geq\tau^*\}}}{\sum_{j=1}^n \Delta_j\mathbb{1}_{\{\tau_0\geq Y_j\geq t\geq\tau^*\}}}, \qquad t\in[\tau^*,\tau_0].$$

Thus, since all $\rho_n(Y_i)$'s are larger than 1, it suffices to suitably bound

$$0\leq A_n = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i\geq\tau^*\}}\Delta_i\log(\rho_n(Y_i)),$$

which we decompose as

$$A_n = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i\in[\tau^*,\tau_0-a_n]\}}\Delta_i\log(\rho_n(Y_i)) + \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i\in[\tau_0-a_n,\tau_0]\}}\Delta_i\log(\rho_n(Y_i)) =: A_{1n}+A_{2n},$$

for some sequence of real numbers $a_n$, $n\geq 1$, decreasing to zero. The rate of $a_n$ should be taken such that, on one hand, for any constant $C>0$,

$$\mathbb{P}(\limsup_{n\to\infty} A_{2n} > C) = 0, \tag{A20}$$

and, on the other hand, the lim sup of $A_{1n}$ could be controlled by a function of $\tau^*$ almost surely. More precisely, since

$$A_{2n} \leq \frac{\log n}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i\in[\tau_0-a_n,\tau_0]\}}\Delta_i,$$

we take $a_n$ such that $p_n\log n\to 0$ and $p_n\log^2 n\to\infty$, where

$$p_n = \mathbb{P}(Y\in[\tau_0-a_n,\tau_0],\Delta=1) = \mathbb{E}\left[\phi(\gamma_0,X)\int_{[\tau_0-a_n,\tau_0]} F_C([t,\infty)|X,Z)F_u(dt|Z)\right].$$

Then, by Theorem 1(i) from [9], we have

$$\lim_{n\to\infty} \frac{1}{p_n} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{Y_i \in [\tau_0 - a_n, \tau_0]\}} \Delta_i = 1, \quad a.s.,$$

which implies (A20). On the other hand, we have

$$A_{1n} \leq \log \left( \sup_{t \in [\tau^*, \tau_0 - a_n]} \rho_n(t) \right) \times \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{Y_i \in [\tau^*, \tau_0]\}} \Delta_i.$$

By the same Theorem 1(i) from [9],

$$\lim_{n\to\infty} \sup_{t \in [\tau^*, \tau_0 - a_n]} \left[ \rho_n(t) \frac{\mathbb{P}(Y \in [t, \tau_0 - a_n], \Delta = 1)}{\mathbb{P}(Y \in [t, \tau_0 - a_n])} \right] = 1, \quad a.s.$$

By our assumptions, there exists a constant $C_r$, independent of $\tau^*$, $\beta$, $\gamma$ and $\Lambda$, such that

$$1 < \inf_{t \in [\tau^*, \tau_0 - a_n]} \frac{\mathbb{P}(Y \in [t, \tau_0 - a_n])}{\mathbb{P}(Y \in [t, \tau_0 - a_n], \Delta = 1)} < \sup_{t \in [\tau^*, \tau_0 - a_n]} \frac{\mathbb{P}(Y \in [t, \tau_0 - a_n])}{\mathbb{P}(Y \in [t, \tau_0 - a_n], \Delta = 1)} \leq C_r.$$

Gathering facts, deduce with probability 1, for sufficiently large $n$,

$$\left| r_{1n}(\hat{\Lambda}_n; \tau^*) - r_{1n}(\tilde{\Lambda}_{0,n}; \tau^*) \right| \leq c \frac{N^*}{n},$$

where $N^*$ be the number of uncensored observations in $[\tau^*, \tau_0]$ and $c$ is some constant (independent of $\tau^*$, $\beta$, $\gamma$ and $\Lambda$). Here, $N^*$ is a binomial random variable with $n$ trials and success probability

$$p^* = \mathbb{P}(Y \geq \tau^*, \Delta = 1) = \mathbb{E}\left[ \phi(\gamma_0, X) \int_{[\tau^*, \tau_0]} F_C([t, \infty)|X, Z) F_u(dt|Z) \right] \\ \leq \sup_x \phi(\gamma_0, x) \mathbb{P}(Y \geq \tau^*).$$

To bound $r_{3n} = r_{3n}(\Lambda, \beta; \tau^*)$, we note that $\mathbb{1}_{\{B_i=1\}} = \Delta_i + (1 - \Delta_i)\mathbb{1}_{\{B_i=1\}}$ and rewrite

$$r_{3n} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{Y_i \geq \tau^*\}} e^{\beta' Z_i} \Delta_i \{\Lambda(Y_i) - \Lambda(\tau^* -)\} \\ - \frac{\Lambda(\tau^* -)}{n} \sum_{i=1}^{n} \mathbb{1}_{\{Y_i \geq \tau^*\}} e^{\beta' Z_i} \mathbb{1}_{\{B_i=1\}} (1 - \Delta_i) = r_{3an} - r_{3bn}.$$

On one hand,

$$r_{3an} \leq \frac{c_u}{c_l} \times \frac{N^*}{n}.$$

The last inequality is obtained by bounding the jumps of $\Lambda$ and using the following identity: for any integer $M \geq 1$,

$$\sum_{k=1}^{M}\sum_{j=k}^{M}\frac{1}{j} = \sum_{j,k=1}^{M}\frac{\mathbb{1}_{\{k\leq j\}}}{j} = M.$$

To bound $r_{3bn}$, let us note that

$$\Lambda(\tau^*-) \leq \frac{1}{c_l}\sum_{j=N^*+1}^{N}\frac{1}{j} \leq c_1 \log\frac{N}{N^*},$$

with $c_1$ some constant depending only on $c_l$ and the maximal value of the convergent sequence

$$\sum_{j=1}^{m}\frac{1}{j} - \log m, \quad m \geq 1.$$

Here, $N = \sum_{i=1}^{n}\Delta_i$ is a binomial random variable with $n$ trials and success probability

$$p = \mathbb{P}(\Delta = 1) = \mathbb{E}\left[\phi(\gamma_0, X)\int_{[0,\tau_0]} F_C([t,\infty)|X,Z)F_u(dt|Z)\right].$$

Thus

$$r_{3bn} \leq c\log\frac{N}{N^*}\times\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{\{Y_i\geq\tau^*\}}\mathbb{1}_{\{B_i=1\}}(1-\Delta_i) = c\log\frac{N/n}{N^*/n}\times\frac{Q^*}{n},$$

where $Q^*$ is a binomial variable with $n$ trials and success probability

$$\begin{aligned}
q^* = \mathbb{E}&\left[\phi(\gamma_0, X)\int_{[\tau^*,\tau_0]} F_u([t,\tau_0]|X,Z)F_C(dt|X,Z)\right] \\
&\leq \mathbb{E}\left[\phi(\gamma_0,X)F_C([\tau^*,\tau_0]|X,Z)F_u([\tau^*,\tau_0]|Z)\right] \\
\leq &\left[\sup_x \phi(\gamma_0,x)\right]\left[\sup_{x,z}\frac{F_C([\tau^*,\tau_0]|X=x,Z=z)}{\tau_0-\tau^*}\right]\times(\tau^*-\tau_0)\times\mathbb{P}(T_0\geq\tau^*) \\
&\leq c(\tau^*-\tau_0)\times\mathbb{P}(T_0\geq\tau^*),
\end{aligned}$$

and $c$ is some constant. By the strong Law of Large Numbers,

$$\lim_{n\to\infty}\log\frac{N}{N^*} = \log\frac{p}{p^*},\ a.s.$$

Next, to bound $r_{2n} = r_{2n}(\beta;\tau^*)$, we write

$$r_{2n} = \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{\{\tau^*\leq T_0<\tau_0\}}e^{\beta' Z_i}\Delta_i \leq c_u\frac{N^*}{n}.$$

Finally, to control $r_{4n} = r_{4n}(\Lambda, \beta, \gamma; \tau^*)$, since $\mathbb{1}_{\{B_i=0\}} = (1-\Delta_i)\mathbb{1}_{\{B_i=0\}}$ and $\log(1+u) \le u$, $\forall u \ge 0$, we have

$$r_{4n} = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{\tau_0 \ge Y_i \ge \tau^*\}}\mathbb{1}_{\{B_i=0\}} \log\left\{1 + \frac{\phi_i(\gamma)\exp\left(-\Lambda(Y_i)e^{\beta' Z_i}\right)}{1-\phi_i(\gamma)}\right\}$$

$$+\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i \ge \tau^*\}}(1-\Delta_i)\mathbb{1}_{\{B_i=1\}} \log\left\{1 - \phi_i(\gamma) + \phi_i(\gamma)\exp\left(-\Lambda(Y_i)e^{\beta' Z_i}\right)\right\}.$$

Thus

$$\begin{aligned}|r_{4n}| \le \sup_{\gamma,x}\left|\frac{\phi(\gamma,x)}{1-\phi(\gamma,x)}\right| \exp\left(-c_l\Lambda(\tau^*-)\right)\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i\ge\tau^*\}}\mathbb{1}_{\{B_i=0\}} \\ + \sup_{\gamma,x}\left|\log\{1-\phi(\gamma,x)\}\right| \times \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Y_i\ge\tau^*\}}(1-\Delta_i)\mathbb{1}_{\{B_i=1\}} \\ = c_1\exp\left(-c_l\Lambda(\tau^*-)\right)\frac{R^*}{n} + c_2\frac{Q^*}{n},\end{aligned}$$

where $R^*$ is a binomial variable with $n$ trials and success probability

$$r^* = \mathbb{E}\left[\{1-\phi(\gamma_0,X)\}F_C([\tau^*,\tau_0]|X,Z)\right],$$

and $c_1$ and $c_2$ are some constants. Deduce that there exists a constant $c_3$ such that

$$|r_{4n}| \le c_3\left(\frac{R^*}{n} + \frac{Q^*}{n}\right).$$

Gathering facts, there exists a constants $C^*$ and $c^*$, independent of $\tau^*$, $\beta$, $\gamma$ and $\Lambda$, such that

$$\left|\hat{l}_n(\beta,\Lambda,\gamma) - \hat{l}_n^*(\beta,\Lambda_{|\tau^*},\gamma)\right| \le C^*\left\{\frac{N^*}{n} + \frac{Q^*}{n}\left[1+\log\frac{n}{N^*}\right] + \frac{R^*}{n}\right\} + o_{a.s.}(1),$$

where $N^*$, $Q^*$ and $R^*$ are binomial with $n$ trials and success probabilities $p^*$, $q^*$ and $r^*$, respectively, and

$$p^* + q^* \le c^*\mathbb{P}(T_0 \ge \tau^*) \quad \text{and} \quad r^* \le c^*\mathbb{P}(C\in[\tau^*,\tau_0]).$$

□

**Lemma 6.** *Assume that for any $x$ and $z$, the conditional distribution of the censoring times given $X = x$ and $Z = z$ is such that there exists a constant $C > 0$ such that*

$$\inf_{[t_1,t_2]\subset[0,\tau_0]}\inf_{x,z}\{F_C(t_2|x,z) - F_C(t_1|x,z)\} > C(t_2-t_1), \qquad \forall\delta > 0.$$

*Let $0 < \bar\tau < \tau_0$ and $\epsilon > 0$. There exist $c_1, c_2 > 0$, $\delta > 0$ such that $c_1 + c_2 = 1$ and*

$$\inf\Bigg\{\mathbb{E}[\ell(Y^*,\Delta^*,X,Z;\beta_0,\Lambda_{0|\tau^*},\gamma_0)] - \mathbb{E}[\ell(Y^*,\Delta^*,X,Z;\beta,\Lambda_{|\tau^*},\gamma_0)] :$$
$$\bar\tau+\delta \le \tau^* < \tau_0, \quad \|\beta-\beta_0\| \ge c_1\epsilon \quad \text{or} \quad \sup_{t\in[0,\bar\tau]}|\Lambda(t)-\Lambda_0(t)| \ge c_2\epsilon\Bigg\} > 0$$

*Proof.* Note that, for any $\tau^* \in (\bar\tau, \tau_0)$,

$$\mathbb{E}[\ell(Y^*,\Delta^*,X,Z;\beta_0,\Lambda_{0|\tau^*},\gamma_0)] - \mathbb{E}[\ell(Y^*,\Delta^*,X,Z;\beta,\Lambda_{|\tau^*},\gamma_0)]$$

is the Kullback-Leibler divergence $KL(\mathbb{P}|Q)$, where $\mathbb{P}$ and $Q$ are the probability measures of $(Y^*,\Delta^*,X,Z)$ when the true parameters are $(\Lambda_0,\beta_0,\gamma_0)$ and $(\Lambda,\beta,\gamma_0)$ respectively. By Pinsker's inequality, we have

$$KL(\mathbb{P}|Q) \ge 2\delta(\mathbb{P},Q)^2,$$

where $\delta(\mathbb{P},Q)$ is the total variation distance between the two probability measures, defined as

$$\delta(\mathbb{P},Q) = \sup_A |\mathbb{P}(A) - Q(A)|,$$

where the supremum is taken over all measurable sets $A$. We want to find a positive lower bound for $\delta(\mathbb{P},Q)$ independent of $\tau^*$ and $Q$, for all $Q$ such that $\|\beta-\beta_0\| \ge c_1\epsilon$ or $\sup_{t\in[0,\bar\tau]}|\Lambda(t)-\Lambda_0(t)| \ge c_2\epsilon$. Hence, it is sufficient to find $k > 0$ and for each such $Q$ an event $A$, which could depend on $Q$, for which $|\mathbb{P}(A) - Q(A)| > k$. Without loss of generality we can assume that the covariate vector $Z$ has mean zero.

*Case 1.* If $\sup_{t\in[0,\bar\tau]}|\Lambda(t)-\Lambda_0(t)| \ge c_2\epsilon$, there exists $\bar t \in [0,\bar\tau]$ such that either

$$\Lambda(\bar t) \ge \Lambda_0(\bar t) + c_2\epsilon \tag{A21}$$

or

$$\Lambda(\bar t) \le \Lambda_0(\bar t) - c_2\epsilon. \tag{A22}$$

We first consider (A21) and define

$$\delta = \min\left\{\frac{\tau_0-\bar\tau}{2},\ \frac{c_2\epsilon}{2\sup_{t\in[0,(\bar\tau+\tau_0)/2]}\lambda_0(t)}\right\}.$$

It follows that for all $t \in [\bar t, \bar t+\delta] \subset [0,(\bar\tau+\tau_0)/2] \subset [0,\tau_0)$ we have $\Lambda(t) \ge \Lambda_0(t) + \frac{1}{2}c_2\epsilon$. Indeed, we can write

$$\Lambda(t) \ge \Lambda(\bar t) \ge \Lambda_0(\bar t) + c_2\epsilon \ge \Lambda_0(t) - \delta \sup_{u\in[0,(\bar\tau+\tau_0)/2]}\lambda_0(u) + c_2\epsilon \ge \Lambda_0(t) + \frac{1}{2}c_2\epsilon, \quad \forall t \in [\bar t, \bar t+\delta].$$

Since $Z$ has mean zero, $(\beta-\beta_0)'Z$ also has zero mean. Moreover, since $B$ is compact and $Z$ is bounded non degenerated variance, we have

$$\inf_{\beta\in B}\mathbb{P}((\beta-\beta_0)'Z\geq 0)>0 \quad \text{and} \quad \inf_{\beta\in B}\mathbb{P}((\beta-\beta_0)'Z\leq 0)>0 \tag{A23}$$

(see proof below). Let $A_\beta$ be the event $\{\Delta^*=0, Y^*\in[\bar{t},\bar{t}+\delta], (\beta-\beta_0)'Z\geq 0\}$, which depends on $\beta$ and thus on $Q$. However, by (A23) and the construction of the model, the event $A_\beta$ has positive probability which stays bounded away from zero. Moreover, we have

$$\begin{aligned}\mathbb{P}(A_\beta)-Q(A_\beta)=\iint_{(\beta-\beta_0)'z\geq 0}&\int_{\bar{t}}^{\bar{t}+\delta}\phi(\gamma_0,x)\\ &\times\left\{\exp\left(-\Lambda_0(t)e^{\beta_0'z}\right)-\exp\left(-\Lambda(t)e^{\beta'z}\right)\right\}F_C(\mathrm{d}t|x,z)G(\mathrm{d}x,\mathrm{d}z).\end{aligned}$$

Whenever $(\beta-\beta_0)'z\geq 0$, by the mean value theorem, we obtain

$$\begin{aligned}\exp\left(-\Lambda_0(t)e^{\beta_0'z}\right)-\exp\left(-\Lambda(t)e^{\beta'z}\right)&=\left\{\Lambda(t)e^{\beta'z}-\Lambda_0(t)e^{\beta_0'z}\right\}e^{-\xi}\\ &=\left[\{\Lambda(t)-\Lambda_0(t)\}e^{\beta_0'z}+\Lambda(t)\{e^{\beta'z}-e^{\beta_0'z}\}\right]e^{-\xi}\\ &\geq\{\Lambda(t)-\Lambda_0(t)\}e^{\beta_0'z}e^{-\xi},\end{aligned}$$

for some $\xi>0$ such that $|\xi-\Lambda_0(t)e^{\beta_0'z}|\leq|\Lambda(t)e^{\beta'z}-\Lambda_0(t)e^{\beta_0'z}|$, $t\in[\bar{t},\bar{t}+\delta]$. Now, let

$$M(t)=\Lambda_0(t)\frac{\sup_{\beta,z}e^{\beta'z}}{\inf_{\beta,z}e^{\beta'z}}+\frac{\log 2}{\inf_{\beta,z}e^{\beta'z}},\quad t\in[\bar{t},\bar{t}+\delta].$$

Then, for $(\beta-\beta_0)'z\geq 0$ and $t\in[\bar{t},\bar{t}+\delta]$, such that $\Lambda(t)\leq M(t)$ we simply use (A21) and write

$$\exp\left(-\Lambda_0(t)e^{\beta_0'z}\right)-\exp\left(-\Lambda(t)e^{\beta'z}\right)\geq\frac{1}{2}c_2\epsilon e^{\beta_0'z}e^{-\xi}\geq k_1\epsilon,$$

for some constant $k_1>0$ independent of $\Lambda$, $\beta$ and the event $A_\beta$, because $M(t)$ is uniformly bounded on $[0,(\bar{\tau}+\tau_0)/2]$ and thus $e^{-\xi}$ is bounded away from zero. On the other hand, for $t\in[\bar{t},\bar{t}+\delta]$ such that $\Lambda(t)>M(t)$, we have

$$\exp\left(-\Lambda(t)e^{\beta'z}\right)\leq\exp\left(-M(t)\inf_{\beta,z}e^{\beta'z}\right)\leq\frac{1}{2}\exp\left(-\Lambda_0(t)e^{\beta_0'z}\right).$$

Consequently,

$$\begin{aligned}\exp\left(-\Lambda_0(t)e^{\beta_0'z}\right)-\exp\left(-\Lambda(t)e^{\beta'z}\right)&\geq\frac{1}{2}\exp\left(-\Lambda_0(t)e^{\beta_0'z}\right)\\ &\geq\frac{1}{2}\exp\left(-\Lambda_0((\bar{\tau}+\tau_0)/2)\sup_{\beta,z}e^{\beta_0'z}\right)=k_2(\bar{\tau})>0.\end{aligned}$$

We conclude that, for any $t \in [\bar{t}, \bar{t}+\delta]$,

$$\exp\left(-\Lambda_0(t)e^{\beta_0' z}\right) - \exp\left(-\Lambda(t)e^{\beta' z}\right) \geq \min\{k_1\epsilon, k_2(\bar{\tau})\} > 0.$$

It follows that

$$\begin{aligned}|\mathbb{P}(A_\beta) - Q(A_\beta)| &\geq \min\{k_1\epsilon, k_2(\bar{\tau})\} \inf_x \phi(\gamma_0, x) \\ &\qquad \times \iint_{(\beta-\beta_0)'z \geq 0} \{F_C(\bar{t}+\delta|x,z) - F_C(\bar{t}|x,z)\}\, G(\mathrm{d}x, \mathrm{d}z).\end{aligned}$$

By assumption we have

$$\inf_{x,z} F_C(\bar{t}+\delta|x,z) - F_C(\bar{t}|x,z) \geq C\delta,$$

yielding that there exist another constant $k_3 > 0$ independent of $\Lambda$, $\beta$ and the event $A_\beta$ (but depending on $\epsilon$ and $\bar{\tau}$) such that

$$\forall \beta \in B, \quad |\mathbb{P}(A_\beta) - Q(A_\beta)| \geq k_3 \inf_{\beta \in B} \mathbb{P}((\beta-\beta_0)'Z \geq 0) > 0.$$

Note that the uniform lower bound holds for any choice of the constants $c_1$ and $c_2$ in the statement of the Lemma.

We next consider (A22). Let

$$\bar{\delta} = \min\left\{\frac{\bar{\tau}}{2}, \frac{c_2\epsilon}{2\sup_{t\in[0,\bar{\tau}]}\lambda_0(t)}\right\}.$$

It follows that for all $t \in [\,\bar{t}-\bar{\delta}, \bar{t}\,]$ we have $\Lambda(t) \leq \Lambda_0(t) - \frac{1}{2}c_2\epsilon$. Indeed, we can write

$$\begin{aligned}\Lambda(t) &\leq \Lambda(\bar{t}) \leq \{\Lambda_0(\bar{t}) - \Lambda_0(t)\} + \Lambda_0(t) - c_2\epsilon \\ &\leq \bar{\delta} \sup_{u\in[0,\bar{\tau}]} \lambda_0(u) + \Lambda_0(t) - c_2\epsilon \leq \Lambda_0(t) - \frac{1}{2}c_2\epsilon, \quad \forall t \in [\,\bar{t}-\bar{\delta}, \bar{t}\,].\end{aligned}$$

Next we redefine $A_\beta$ as the event $\{\Delta^* = 0, Y^* \in [\bar{t}-\bar{\delta}, \bar{t}], (\beta-\beta_0)'Z \leq 0\}$, which depends on $\beta$ and thus on $Q$. However, by (A23) and the construction of the model, the event $A_\beta$ has positive probability which stays bounded away from zero. Moreover, we have

$$\begin{aligned}\mathbb{P}(A_\beta) - Q(A_\beta) = &\iint_{(\beta-\beta_0)'z \leq 0} \int_{\bar{t}-\bar{\delta}}^{\bar{t}} \phi(\gamma_0, x) \\ &\times \left\{\exp\left(-\Lambda_0(t)e^{\beta_0' z}\right) - \exp\left(-\Lambda(t)e^{\beta' z}\right)\right\} F_C(\mathrm{d}t|x,z) G(\mathrm{d}x, \mathrm{d}z).\end{aligned}$$

Whenever $(\beta-\beta_0)'z \leq 0$, by the mean value theorem, we obtain

$$\exp\left(-\Lambda_0(t)e^{\beta_0' z}\right) - \exp\left(-\Lambda(t)e^{\beta' z}\right) \leq \{\Lambda(t) - \Lambda_0(t)\}e^{\beta_0' z}e^{-\xi} \leq -\frac{1}{2}c_2\epsilon e^{\beta_0' z}e^{-\xi},$$

for some $\xi > 0$ such that $|\xi - \Lambda_0(t)e^{\beta_0' z}| \leq |\Lambda(t)e^{\beta' z} - \Lambda_0(t)e^{\beta_0' z}| \leq 2\Lambda_0(t)e^{\beta_0' z}$, $t \in [\bar{t} - \bar{\delta}, \bar{t}]$. Thus necessarily $0 < \xi \leq 2\Lambda_0(\bar{\tau})e^{\beta_0' z}$, and thus $e^{-\xi}$ stays away from zero. Using arguments as we used for the case (A22), we deduce that $\mathbb{P}(A_\beta) - Q(A_\beta)$ is negative and away from zero. Thus we obtain the result with $\bar{\tau} \leq \tau^* < \tau_0$ instead of $\bar{\tau} + \delta \leq \tau^* < \tau_0$. Finally, it remains to recall that inf is a decreasing function of nested sets. Now the arguments for *Case 1* are complete for any choice of the constants $c_1$ and $c_2$ in the statement of the Lemma.

*Case 2.* If $\sup_{t\in[0,\bar{\tau}]} |\Lambda(t) - \Lambda_0(t)| \leq c_2\epsilon$, then necessarily $\|\beta - \beta_0\| \geq c_1\epsilon$. In particular we also have that $\Lambda(\bar{\tau}) \leq \Lambda_0(\bar{\tau}) + c_2\epsilon$, so all such functions $\Lambda$ are uniformly bounded on $[0, \bar{\tau}]$. Without loss of generality we can also assume that $\Lambda_0(\bar{\tau}/2) \geq 1$ (otherwise we can take a larger $\bar{\tau}$). Note that

$$Var((\beta - \beta_0)'Z) = (\beta - \beta_0)'Var(Z)(\beta - \beta_0) \geq (c_1\epsilon)^2\lambda_{min},$$

with $\lambda_{min}$ the smallest eigenvalue of $Var(Z)$. From this lower bound for the variance of $(\beta - \beta_0)'Z$, and since $Z$ is centered and has a bounded support, we have

$$\inf_{|\beta-\beta_0|\geq c_1\epsilon} [\mathbb{P}((\beta - \beta_0)'Z \geq z_0) + \mathbb{P}((\beta - \beta_0)'Z \leq -z_0)] > \frac{(c_1\epsilon)^2\lambda_{min}}{2\sup\|Z\|^2}, \quad \text{(A24)}$$

for $z_0 = c_1\epsilon\lambda_{min}^{1/2}/2$ (see proof below). If

$$\inf_{|\beta-\beta_0|\geq c_1\epsilon} \mathbb{P}\left((\beta - \beta_0)'Z \geq z_0\right) > \frac{(c_1\epsilon)^2\lambda_{min}}{2\sup\|Z\|^2},$$

let $A_\beta$ be the event $\{\Delta^* = 0, Y^* \in [\bar{\tau}/2, \bar{\tau}], (\beta - \beta_0)'Z \geq z_0\}$. By (A24) and the construction of the model, the event $A_\beta$ has positive probability which stays bounded away from zero. Next, as in *Case 1*, we write

$$\begin{aligned}\mathbb{P}(A_\beta) - Q(A_\beta) = &\iint_{(\beta-\beta_0)'z\geq z_0} \int_{\frac{1}{2}\bar{\tau}}^{\bar{\tau}} \phi(\gamma_0, x) \\ &\times \left\{\exp\left(-\Lambda_0(t)e^{\beta_0' z}\right) - \exp\left(-\Lambda(t)e^{\beta' z}\right)\right\} F_C(\mathrm{d}t|x, z)\, G(\mathrm{d}x, \mathrm{d}z),\end{aligned}$$

and

$$\exp\left(-\Lambda_0(t)e^{\beta_0' z}\right) - \exp\left(-\Lambda(t)e^{\beta' z}\right) = \left[\{\Lambda(t) - \Lambda_0(t)\}e^{\beta_0' z} + \Lambda(t)\{e^{\beta' z} - e^{\beta_0' z}\}\right] e^{-\xi},$$

for some $\xi > 0$ such that $|\xi - \Lambda_0(t)e^{\beta_0' z}| \leq |\Lambda(t)e^{\beta' z} - \Lambda_0(t)e^{\beta_0' z}|$. From the boundedness of $\beta$, $z$, $\Lambda$ and $\Lambda_0$ on $[0, \bar{\tau}]$, it follows that $e^{-\xi} \geq k_4 > 0$ for some $k_4$ independent of $\Lambda$, $\beta$ and the event $A_\beta$ (but depending on $\bar{\tau}$). Moreover, since for $t \in [\bar{\tau}/2, \bar{\tau}]$, $(\beta - \beta_0)'z \geq z_0$,

$$|\Lambda(t) - \Lambda_0(t)|\, e^{\beta_0' z} \leq c_2\epsilon e^{\beta_0' z},$$

and

$$\Lambda(t)\{e^{\beta' z} - e^{\beta_0' z}\} \geq \Lambda_0(\bar{\tau}/2) e^{\beta_0' z}\{e^{(\beta-\beta_0)' z} - 1\} \geq e^{\beta_0' z} z_0 = e^{\beta_0' z}\lambda_{min}^{1/2} c_1 \epsilon/2,$$

we obtain

$$\exp\left(-\Lambda_0(t)e^{\beta_0' z}\right) - \exp\left(-\Lambda(t)e^{\beta' z}\right) \geq \epsilon\left[\lambda_{min}^{1/2} c_1/2 - c_2\right] e^{\beta_0' z} e^{-\xi}.$$

Define

$$c_1 = \frac{4}{\lambda_{min}^{1/2} + 4} \qquad \text{and} \qquad c_2 = \frac{\lambda_{min}^{1/2}}{\lambda_{min}^{1/2} + 4},$$

such that $0 < c_1, c_2 < 1$ and $c_1 + c_2 = 1$, and deduce that

$$\exp\left(-\Lambda_0(t)e^{\beta_0' z}\right) - \exp\left(-\Lambda(t)e^{\beta' z}\right) \geq \epsilon\lambda_{min}^{1/2} e^{\beta_0' z} k_4 \left\{\lambda_{min}^{1/2} + 4\right\}^{-1}.$$

Deduce that, for any $\|\beta - \beta_0\| \geq c_1\epsilon$,

$$\begin{aligned}\mathbb{P}(A_\beta) - Q(A_\beta) \geq \epsilon\left[\lambda_{min}^{1/2} e^{\beta_0' z} k_4 \left\{\lambda_{min}^{1/2} + 4\right\}^{-1}\right] \inf_z e^{\beta_0' z} \inf_x \phi(\gamma_0, x) \\ \times \inf_{|\beta-\beta_0|\geq c_1\epsilon} \mathbb{P}(C \in [\bar{\tau}/2, \bar{\tau}], (\beta - \beta_0)' Z \geq z_0) > 0.\end{aligned}$$

Note that this bound does not depend on $\delta$ in the statement of the Lemma. Finally it is easy to see that similar arguments apply when $\inf_{|\beta-\beta_0|\geq c_1\epsilon} \mathbb{P}((\beta - \beta_0)' Z \leq -z_0) > (c_1\epsilon)^2\lambda_{min}/\{4\sup\|Z\|^2\}$, for the same expression of $z_0$. In this case, we define $A_\beta = \{\Delta^* = 0, Y^* \in [\bar{\tau}/2, \bar{\tau}], (\beta - \beta_0)' Z \leq -z_0\}$ and follow the same steps as above to show that

$$\exp\left(-\Lambda_0(t)e^{\beta_0' z}\right) - \exp\left(-\Lambda(t)e^{\beta' z}\right) < 0,$$

and the difference of exponentials stays away from zero. Now, the proof of Lemma 6 is complete. ☐

Proof of Equation (A23).. For $\beta = \beta_0$ we have $\mathbb{P}((\beta - \beta_0)' Z \geq 0) = 1$ and thus we only have to study $\beta \neq \beta_0$. Since $Var(Z)$ is non degenerated, $\mathbb{P}((\beta - \beta_0)' Z = 0) < 1$ for any $\beta \neq \beta_0$. Next, we could write

$$\begin{aligned}0 = \mathbb{E}\left(\frac{(\beta - \beta_0)' Z}{\|\beta - \beta_0\|}\right) &= \mathbb{E}\left(\frac{(\beta - \beta_0)' Z}{\|\beta - \beta_0\|}\mathbb{1}_{\{(\beta-\beta_0)' Z \geq 0\}}\right) + \mathbb{E}\left(\frac{(\beta - \beta_0)' Z}{\|\beta - \beta_0\|}\mathbb{1}_{\{(\beta-\beta_0)' Z < 0\}}\right) \\ &\leq \|Z\|\mathbb{P}((\beta - \beta_0)' Z \geq 0) + \mathbb{E}\left(\frac{(\beta - \beta_0)' Z}{\|\beta - \beta_0\|}\mathbb{1}_{\{(\beta-\beta_0)' Z < 0\}}\right).\end{aligned}$$

It remains to notice that

$$\sup_{\|\tilde{\beta}\|=1} \mathbb{E}\left(\tilde{\beta}' Z \mathbb{1}_{\{\tilde{\beta}' Z < 0\}}\right) < 0.$$

Indeed, if $\mathbb{E}\left(\tilde\beta' Z \mathbb{1}_{\{\tilde\beta' Z<0\}}\right)$, which is negative, could be arbitrarily close to zero, and since $\mathbb{E}\left(\tilde\beta' Z \mathbb{1}_{\{\tilde\beta' Z<0\}}\right) = -\mathbb{E}\left(\tilde\beta' Z \mathbb{1}_{\{\tilde\beta' Z\geq 0\}}\right)$, we deduce that $\mathbb{E}(|\tilde\beta' Z|)$ could be arbitrarily close to zero, for suitable $\tilde\beta$ with unit norm. Since the support of $Z$ is bounded and

$$\lambda_{min}(Var(Z)) \leq \mathbb{E}(|\tilde\beta' Z|^2) \leq \mathbb{E}(|\tilde\beta' Z|)\sup\|Z\|, \quad \forall \|\tilde\beta\| = 1,$$

we thus get a contradiction with the assumption that $\lambda_{min}$, the smallest eigenvalue of $Var(Z)$, is positive. Deduce that (A23) holds true. □

Proof of Equation (A24).. It suffices to prove the following property. Let $U$ be a centered variable such that $|U| \leq M$ for some constant $M$ and $Var(U)$ is bounded from below by some constant $0 < C < M^2$. Then there exists $z_0 > 0$ such that $\mathbb{P}(|U| \geq z_0) > C/2M^2$ with $z_0$ depending on $M$ and $C$ but independent of the law of $U$.

For any $0 < z_0 < M$ we can write

$$C \leq \mathbb{E}(U^2) = \mathbb{E}\left(U^2 \mathbb{1}_{\{|U|\geq z_0\}}\right) + \mathbb{E}\left(U^2 \mathbb{1}_{\{|U|< z_0\}}\right) \leq M^2 \mathbb{P}(|U| \geq z_0) + z_0^2\{1-\mathbb{P}(|U| \geq z_0)\}.$$

Deduce that

$$\mathbb{P}(|U| \geq z_0) \geq \frac{C - z_0^2}{M^2 - z_0^2}.$$

Finally, it suffices to take $z_0^2 \leq CM^2/(2M^2 - C)$ in order to obtain

$$\mathbb{P}(|U| \geq z_0) \geq \frac{C}{2M^2}.$$

Since $C$ could be arbitrarily small, we could take $z_0^2 = C/4$. □

### *A.3. Asymptotic normality*

Proof of Theorem 3.. Let $m$ be as in (21) and define $M_n(\gamma, \pi) = \frac{1}{n}\sum_{i=1}^n m(X_i; \gamma, \pi)$ and $M(\gamma, \pi) = \mathbb{E}[m(X; \gamma, \pi)]$. Note that $m(x; \gamma_0, \pi_0) = 0$ and $M(\gamma_0, \pi_0) = 0$. If we take partial derivatives with respect to $\gamma$ of the vector $M(\gamma, \pi)$ evaluated at $(\gamma_0, \pi_0)$ we obtain a matrix $\nabla_\gamma M(\gamma_0, \pi_0)$ with elements

$$\nabla_\gamma M(\gamma_0, \pi_0)_{kl} = \mathbb{E}\left[\frac{\partial}{\partial\gamma_l}\left\{\frac{1-\pi(X)}{\phi(\gamma, X)} - \frac{\pi(X)}{1-\phi(\gamma, X)}\right\}\frac{\partial}{\partial\gamma_k}\phi(\gamma, X)\right]\Bigg|_{\gamma_0,\pi_0},$$

for $k, l \in \{1, \ldots, p\}$, because

$$\frac{1-\pi_0(x)}{\phi(\gamma_0, x)} - \frac{\pi_0(x)}{1-\phi(\gamma_0, x)} = 0 \quad \forall x \in \mathcal{X}.$$

Hence,

$$\Gamma_1 := \nabla_\gamma M(\gamma_0, \pi_0) = -\mathbb{E}\left[W(X)\nabla_\gamma \phi(\gamma_0, X)\nabla_\gamma \phi(\gamma_0, X)'\right] \tag{A25}$$

where

$$W(X) = \frac{1-\pi_0(X)}{\phi(\gamma_0,X)^2} + \frac{\pi_0(X)}{(1-\phi(\gamma_0,X))^2} = \frac{1}{\phi(\gamma_0,X)} + \frac{1}{1-\phi(\gamma_0,X)} > 0.$$

We will also use the Gateaux derivative of $M(\gamma, \pi_0)$ in a direction $[\pi - \pi_0]$ given by

$$\begin{aligned}
\Gamma_2(\gamma,\pi_0)[\pi-\pi_0] &:= \nabla_\pi M(\gamma,\pi_0)[\pi-\pi_0] \\
&= \lim_{h\to 0}\frac{1}{h}\left[M(\gamma,\pi_0+h(\pi-\pi_0)) - M(\gamma,\pi_0)\right] \\
&= -\lim_{h\to 0}\frac{1}{h}\mathbb{E}\left[\left\{\frac{h\left\{\pi(X)-\pi_0(X)\right\}}{\phi(\gamma,X)} + \frac{h\left\{\pi(X)-\pi_0(X)\right\}}{1-\phi(\gamma,X)}\right\}\nabla_\gamma\phi(\gamma,X)\right] \\
&= -\mathbb{E}\left[\left\{\pi(X)-\pi_0(X)\right\}\left\{\frac{1}{\phi(\gamma,X)} + \frac{1}{1-\phi(\gamma,X)}\right\}\nabla_\gamma\phi(\gamma,X)\right].
\end{aligned} \tag{A26}$$

We apply Theorem 2 in [3] so we need to verify its conditions. Consistency of $\hat{\gamma}_n$ is shown in Theorem 1, while condition (2.1) in [3] is satisfied by construction since

$$\|M_n(\hat{\gamma}_n, \hat{\pi}_n)\| = 0 = \inf_{\gamma\in G}\|M_n(\gamma, \hat{\pi}_n)\|.$$

Note that assumption (AC1) was needed in Theorem 1 in order to obtain almost sure convergence. However, here we only need convergence in probability for which (AN4)-(ii) suffices. For condition (2.2) in [3], the derivative of $M$ with respect to $\gamma$ is computed in (A25) and the matrix is negative definite (as a result also full rank) because of our assumption (AN3). Moreover the directional derivative was computed in (A26) and for $(\gamma,\pi) \in G_{\delta_n} \times \Pi_{\delta_n}$ with $G_{\delta_n} = \{\gamma \in G : \|\gamma - \gamma_0\| \le \delta_n\}$, $\Pi_{\delta_n} = \{\pi \in \Pi : \|\pi - \pi_0\|_\infty \le \delta_n\}$, $\delta_n = o(1)$, we have

$$\begin{aligned}
&\|M(\gamma,\pi) - M(\gamma,\pi_0) - \Gamma_2(\gamma,\pi_0)[\pi-\pi_0]\| \\
&= \left\|\mathbb{E}\left[\left\{\frac{\pi_0(X)-\pi(X)}{\phi(\gamma,X)} + \frac{\pi_0(X)-\pi(X)}{1-\phi(\gamma,X)}\right\}\nabla_\gamma\phi(\gamma,X)\right]\right. \\
&\quad \left. + \mathbb{E}\left[\left\{\pi(X)-\pi_0(X)\right\}\left\{\frac{1}{\phi(\gamma,X)} + \frac{1}{1-\phi(\gamma,X)}\right\}\nabla_\gamma\phi(\gamma,X)\right]\right\| = 0,
\end{aligned}$$

which means that condition (2.3i) is satisfied. For condition (2.3ii), we have

$$\begin{aligned}
&\Gamma_2(\gamma,\pi_0)[\pi-\pi_0] - \Gamma_2(\gamma_0,\pi_0)[\pi-\pi_0] \\
&= -\mathbb{E}\left[\left\{\pi(X)-\pi_0(X)\right\}\left\{\left(\frac{1}{\phi(\gamma,X)} + \frac{1}{1-\phi(\gamma,X)}\right)\nabla_\gamma\phi(\gamma,X)\right.\right. \\
&\qquad \left.\left. - \left(\frac{1}{\phi(\gamma_0,X)} + \frac{1}{1-\phi(\gamma_0,X)}\right)\nabla_\gamma\phi(\gamma_0,X)\right\}\right].
\end{aligned}$$

Then, from $\sup_x |\pi(x) - \pi_0(x)| \leq \delta_n$, $|\gamma - \gamma_0| \leq \delta_n \to 0$ and (AN1), it follows that

$$\|\Gamma_2(\gamma, \pi_0)[\pi - \pi_0] - \Gamma_2(\gamma_0, \pi_0)[\pi - \pi_0]\| \leq o(1)\delta_n.$$

Conditions (2.4) and (2.6) in [3] are satisfied thanks to our assumption (AN4) because

$$\begin{aligned}
&M_n(\gamma_0, \pi_0) + \Gamma_2(\gamma_0, \pi_0)[\hat{\pi} - \pi_0] \\
&= \frac{1}{n}\sum_{i=1}^n \left\{ \frac{1 - \pi_0(X_i)}{\phi(\gamma_0, X_i)} - \frac{\pi_0(X_i)}{1 - \phi(\gamma_0, X_i)} \right\} \nabla_\gamma \phi(\gamma_0, X_i) \\
&\quad + \mathbb{E}^* \left[ \{\hat{\pi}(X) - \pi_0(X)\} \left( \frac{1}{\phi(\gamma_0, X)} + \frac{1}{1 - \phi(\gamma_0, X)} \right) \nabla_\gamma \phi(\gamma_0, X) \right] \\
&= \mathbb{E}^* \left[ \{\hat{\pi}(X) - \pi_0(X)\} \left( \frac{1}{\phi(\gamma_0, X)} + \frac{1}{1 - \phi(\gamma_0, X)} \right) \nabla_\gamma \phi(\gamma_0, X) \right].
\end{aligned} \tag{A27}$$

Then we conclude by central limit theorem that

$$\sqrt{n}\left(M_n(\gamma_0, \pi_0) + \Gamma_2(\gamma_0, \pi_0)[\hat{\pi} - \pi_0]\right) \xrightarrow{d} N(0, V)$$

where $V = Var(\Psi(Y, \Delta, X, Z))$. It remains to deal with condition (2.5), which is a consequence of Theorem 3 in [3] and assumption (AN2) because from (AN1) we have

$$\begin{aligned}
\|m(x; \gamma_1, \pi_1) - m(x; \gamma_2, \pi_2)\| \leq &\left\| \left( \frac{1 - \pi_1(x)}{\phi(\gamma_1, x)} + \frac{\pi_1(x)}{1 - \phi(\gamma_1, X)} \right) \nabla_\gamma \phi(\gamma_1, X) \right. \\
&\left. - \left( \frac{1 - \pi_2(x)}{\phi(\gamma_2, x)} + \frac{\pi_2(x)}{1 - \phi(\gamma_2, X)} \right) \nabla_\gamma \phi(\gamma_2, X) \right\| \\
&\leq C_1\|\gamma_1 - \gamma_2\| + C_2\|\pi_1 - \pi_2\|_\infty.
\end{aligned}$$

Finally, the asymptotic normality follows from Theorem 2 in [3] and the asymptotic covariance matrix is given by

$$\Sigma_\gamma = (\Gamma_1'\Gamma_1)^{-1}\Gamma_1' V \Gamma_1 (\Gamma_1'\Gamma_1)^{-1} = \Gamma_1^{-1} V \Gamma_1^{-1} \tag{A28}$$

□

Proof of Theorem 4.. We show that conditions 1 and 4 of Theorem 4 in [5] are satisfied. Define $S_n$ as the version of $\hat{S}_n$ where $\hat{\gamma}_n$ is replaced by $\gamma_0$

$$\begin{aligned}
S_n(\hat{\Lambda}_n, \hat{\beta}_n)(h_1, h_2) = &\frac{1}{n}\sum_{i=1}^n \Delta_i \mathbb{1}_{\{Y_i < \tau_0\}} [h_1(Y_i) + h_2' Z_i] \\
&- \frac{1}{n}\sum_{i=1}^n \left\{ \Delta_i + (1 - \Delta_i)\mathbb{1}_{\{Y_i \leq \tau_0\}} g_i(Y_i, \hat{\Lambda}_n, \hat{\beta}_n, \gamma_0) \right\} \\
&\quad \times \left\{ e^{\hat{\beta}_n' Z_i} \int_0^{Y_i} h_1(s) \mathrm{d}\hat{\Lambda}_n(s) + e^{\hat{\beta}_n' Z_i} \hat{\Lambda}_n(Y_i) h_2' Z_i \right\}
\end{aligned}$$

*Condition 1.* We start by writing

$$\hat{S}_n(\Lambda_0,\beta_0)(h_1,h_2)-S(\Lambda_0,\beta_0)(h_1,h_2)=\left[\hat{S}_n(\Lambda_0,\beta_0)(h_1,h_2)-S_n(\Lambda_0,\beta_0)(h_1,h_2)\right] + \left[S_n(\Lambda_0,\beta_0)(h_1,h_2)-S(\Lambda_0,\beta_0)(h_1,h_2)\right]. \tag{A29}$$

For the second term on the right hand side we have

$$S_n(\Lambda_0,\beta_0)(h_1,h_2)-S(\Lambda_0,\beta_0)(h_1,h_2)=\int f_h(y,\delta,x,z)\,\mathrm{d}(\mathbb{P}_n-\mathbb{P})(y,\delta,x,z) \tag{A30}$$

where

$$\begin{aligned} f_h(y,\delta,x,z) &= h_2'z\left\{\delta\mathbb{1}_{\{y<\tau_0\}}-\left[\delta-(1-\delta)\mathbb{1}_{\{y\le\tau_0\}}g(y,\Lambda_0,\beta_0,\gamma_0)\right]e^{\beta_0'z}\Lambda_0(y)\right\}\\ &\quad+\delta\mathbb{1}_{\{y<\tau_0\}}h_1(y)-\left[\delta-(1-\delta)\mathbb{1}_{\{y\le\tau_0\}}g(y,\Lambda_0,\beta_0,\gamma_0)\right]e^{\beta_0'z}\int_0^y h_1(s)\mathrm{d}\Lambda_0(s). \end{aligned}$$

The classes $\{h_2\in\mathbb{R}^q,\|h_2\|\le\mathfrak{m}\}$, $\{h_1\in BV[0,\tau_0],\,\|h_1\|_v\le\mathfrak{m}\}$ and

$$\left\{\int_0^y h_1(t)\,\mathrm{d}\Lambda_0(t),\ h_1\in BV[0,\tau_0],\|h_1\|_v\le\mathfrak{m}\right\}$$

are Donsker classes (the last one because it consists of monotone bounded functions). As in [5], because of the boundedness of the covariates and $\Lambda_0$, it follows that $\{f_h(y,\delta,x,z),\,h\in\mathcal{H}_\mathfrak{m}\}$ is also a Donsker class since it is sum of products of Donsker classes with fixed uniformly bounded functions.

On the other hand, for the first term on the right hand side of (A29), we have

$$\begin{aligned} &\left[\hat{S}_n(\Lambda_0,\beta_0)(h_1,h_2)-S_n(\Lambda_0,\beta_0)(h_1,h_2)\right]\\ &=-\frac{1}{n}\sum_{i=1}^n(1-\Delta_i)\mathbb{1}_{\{Y_i\le\tau_0\}}\left\{e^{\beta_0'Z_i}\int_0^{Y_i}h_1(s)\mathrm{d}\Lambda_0(s)+e^{\beta_0'Z_i}\Lambda_0(Y_i)h_2'Z_i\right\}\\ &\qquad\times\left\{g_i(Y_i,\Lambda_0,\beta_0,\hat{\gamma}_n)-g_i(Y_i,\Lambda_0,\beta_0,\gamma_0)\right\}\\ &=-\frac{1}{n}\sum_{i=1}^n(1-\Delta_i)\mathbb{1}_{\{Y_i\le\tau_0\}}e^{\beta_0'Z_i}\left\{\int_0^{Y_i}h_1(s)\mathrm{d}\Lambda_0(s)+\Lambda_0(Y_i)h_2'Z_i\right\}\\ &\qquad\times\frac{\partial g_i}{\partial\phi}(Y_i,\Lambda_0,\beta_0,\gamma_0)\left\{\phi(\hat{\gamma}_n,X_i)-\phi(\gamma_0,X_i)\right\}+o_P(n^{-1/2}), \end{aligned} \tag{A31}$$

where

$$\begin{aligned} \frac{\partial g_i}{\partial\phi}(Y_i,\Lambda_0,\beta_0,\gamma_0) &= \frac{\exp\left(-\Lambda(Y_i)e^{\beta'Z_j}\right)}{1-\phi(\gamma,X_j)+\phi(\gamma,X_j)\exp\left(-\Lambda(Y_i)e^{\beta'Z_j}\right)}\\ &\quad+\frac{\phi(\gamma,X_j)\exp\left(-\Lambda(Y_i)e^{\beta'Z_j}\right)\left[\exp\left(-\Lambda(Y_i)e^{\beta'Z_j}\right)-1\right]}{\left[1-\phi(\gamma,X_j)+\phi(\gamma,X_j)\exp\left(-\Lambda(Y_i)e^{\beta'Z_j}\right)\right]^2}. \end{aligned}$$

In order to conclude that the remainder term is of order $o_P(n^{-1/2})$ we use

$$\sup_x |\phi(\hat{\gamma}_n, x) - \phi(\gamma_0, x)| \leq \sup_{\gamma \in G, x \in \mathcal{X}} \|\nabla_\gamma \phi(\gamma, x)\| \, |\hat{\gamma}_n - \gamma_0| = O_P(n^{-1/2})$$

and the fact that $\frac{\partial^2 g_i}{\partial \phi^2}(Y_i, \Lambda_0, \beta_0, \gamma)$ and

$$(1-\Delta_i)\mathbb{1}_{\{Y_i \leq \tau_0\}} e^{\beta_0' Z_i} \left\{ \int_0^{Y_i} h_1(s) \mathrm{d}\Lambda_0(s) + \Lambda_0(Y_i) h_2' Z_i \right\}$$

are uniformly bounded functions thanks to our assumptions on $Z$, $\Lambda$, $\Phi$ and $h$. From the same assumptions we also obtain

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n (1-\Delta_i)\mathbb{1}_{\{Y_i \leq \tau_0\}} e^{\beta_0' Z_i} \left\{ \int_0^{Y_i} h_1(s) \mathrm{d}\Lambda_0(s) + \Lambda_0(Y_i) h_2' Z_i \right\} \\
&\qquad\qquad \times \frac{\partial g_i}{\partial \phi}(Y_i, \Lambda_0, \beta_0, \gamma_0) \left\{ \phi(\hat{\gamma}_n, X_i) - \phi(\gamma_0, X_i) \right\} \\
&= \frac{1}{n}\sum_{i=1}^n (1-\Delta_i) e^{\beta_0' Z_i} \left\{ \int_0^{Y_i} h_1(s) \mathrm{d}\Lambda_0(s) + \Lambda_0(Y_i) h_2' Z_i \right\} \\
&\qquad\qquad \times \frac{\partial g_i}{\partial \phi}(Y_i, \Lambda_0, \beta_0, \gamma_0) \nabla_\gamma \phi(\gamma_0, X_i)' \left\{ \hat{\gamma}_n - \gamma_0 \right\} + o_P(n^{-1/2}) \\
&= \mathbb{E}\left[ (1-\Delta)\mathbb{1}_{\{Y \leq \tau_0\}} e^{\beta_0' Z} \left\{ \int_0^{Y} h_1(s) \mathrm{d}\Lambda_0(s) + \Lambda_0(Y) h_2' Z \right\} \right. \\
&\qquad\qquad \left. \times \frac{\partial g}{\partial \phi}(Y, \Lambda_0, \beta_0, \gamma_0) \nabla_\gamma \phi(\gamma_0, X)' \right] (\hat{\gamma}_n - \gamma_0) + o_P(n^{-1/2}).
\end{aligned} \tag{A32}$$

and the expectation term is uniformly bounded. To prove the asymptotic normality of $\hat{\gamma}_n - \gamma_0$ in Theorem 3 we used Theorem 2 in [3]. Going through the proof of Theorem 2 in [3], we actually have

$$(\hat{\gamma}_n - \gamma_0) = -(\Gamma_1'\Gamma_1)^{-1}\Gamma_1' \left\{ M_n(\gamma_0, \pi_0) + \Gamma_2(\gamma_0, \pi_0)[\hat{\pi} - \pi_0] \right\} + o_P(n^{-1/2})$$

where $\Gamma_1$ is defined in (A25) and $\Gamma_2$ in (A26). From Assumption (AN4-iii) and (A27), it follows that

$$(\hat{\gamma}_n - \gamma_0) = -(\Gamma_1'\Gamma_1)^{-1}\Gamma_1' \int \Psi(y, \delta, x) \, \mathrm{d}(\mathbb{P}_n - \mathbb{P})(y, \delta, x, z) + o_P(n^{-1/2}). \tag{A33}$$

Putting together (A29)-(A33), we have

$$\begin{aligned}
&\hat{S}_n(\Lambda_0, \beta_0)(h_1, h_2) - S(\Lambda_0, \beta_0)(h_1, h_2) \\
&= \int \left\{ f_h(y, \delta, x, z) - Q_h \Gamma_1^{-1} \Psi(y, \delta, x) \right\} \mathrm{d}(\mathbb{P}_n - \mathbb{P})(y, \delta, x, z) + o_P(n^{-1/2})
\end{aligned}$$

where

$$Q_h = \mathbb{E}\left[(1-\Delta)\mathbb{1}_{\{Y\le\tau_0\}}e^{\beta_0'Z}\left\{\int_0^Y h_1(s)\mathrm{d}\Lambda_0(s)+\Lambda_0(Y)h_2'Z\right\}\frac{\partial g}{\partial\phi}(Y,\Lambda_0,\beta_0,\gamma_0)\nabla_\gamma\phi(\gamma_0,X)'\right].$$

In order to conclude the convergence of $\sqrt{n}(\hat{S}_n(\Upsilon_0)-S(\Upsilon_0))$ to a Gaussian process $G^*$, we need to have that $\{Q_h\Gamma_1^{-1}\Psi(y,\delta,x), h\in\mathcal{H}_{\mathfrak{m}}\}$ is a bounded Dosker class of functions (since sum of bounded Donsker classes is also Donsker). We can write

$$\begin{aligned}
&Q_h\Gamma_1^{-1}\Psi(y,\delta,x)\\
&= h_2'\mathbb{E}\left[(1-\Delta)\mathbb{1}_{\{Y\le\tau_0\}}Ze^{\beta_0'Z}\Lambda_0(Y)\frac{\partial g}{\partial\phi}(Y,\Lambda_0,\beta_0,\gamma_0)\nabla_\gamma\phi(\gamma_0,X)'\right]\Gamma_1^{-1}\Psi(y,\delta,x)\\
&\quad+\int_0^{\tau_0}\mathbb{E}\left[(1-\Delta)\mathbb{1}_{\{Y\le\tau_0\}}e^{\beta_0'Z}\frac{\partial g}{\partial\phi}(Y,\Lambda_0,\beta_0,\gamma_0)\nabla_\gamma\phi(\gamma_0,X)'\right]h_1(s)\mathrm{d}\Lambda_0(s)\Gamma_1^{-1}\Psi(y,\delta,x)
\end{aligned}$$

By assumption (AN1) and $\inf_x H((\tau_0,\infty)|x)>0$, $\Lambda_0(\tau_0)<\infty$ and the boundedness of the covariates, we have that

$$\mathbb{E}\left[(1-\Delta)\mathbb{1}_{\{Y\le\tau_0\}}Ze^{\beta_0'Z}\Lambda_0(Y)\frac{\partial g}{\partial\phi}(Y,\Lambda_0,\beta_0,\gamma_0)\nabla_\gamma\phi(\gamma_0,X)'\right]\Gamma_1^{-1}\Psi(y,\delta,x)$$

is uniformly bounded. Hence

$$\left\{h_2'\mathbb{E}\left[(1-\Delta)\mathbb{1}_{\{Y\le\tau_0\}}Ze^{\beta_0'Z}\Lambda_0(Y)\frac{\partial g}{\partial\phi}(Y,\Lambda_0,\beta_0,\gamma_0)\nabla_\gamma\phi(\gamma_0,X)'\right]\Gamma_1^{-1}\Psi(y,\delta,x):\right.$$
$$\left. h_2\in\mathbb{R}^q, \|h_2\|_{L_1}\le\mathfrak{m}\right\}$$

is a Donsker class (see Example 2.10.10 in [8]). It can also be shown that, since $h_1$ belongs to the class of bounded functions with bounded variation and all the other terms are uniformly bounded, that

$$\left\{\int_0^{\tau_0}\mathbb{E}\left[(1-\Delta)\mathbb{1}_{\{Y\le\tau_0\}}e^{\beta_0'Z}\frac{\partial g}{\partial\phi}(Y,\Lambda_0,\beta_0,\gamma_0)\nabla_\gamma\phi(\gamma_0,X)'\right]h_1(s)\mathrm{d}\Lambda_0(s)\Gamma_1^{-1}\Psi(y,\delta,x),\right.$$
$$\left. h_1\in BV[0,\tau_0],\ \|h_1\|_v\le\mathfrak{m}\right\}$$

is also a bounded Donsker class (covering numbers of order $\epsilon$ of $\{h_1\in BV[0,\tau_0], \|h_1\|_v\le\mathfrak{m}\}$ correspond to covering numbers of order $c\epsilon$ for some constant $c>0$).

The limit process $G^*$ has mean zero because

$$\mathbb{E}[f_h(y,\delta,x,z)] = S(\Upsilon_0)(h)=0 \qquad \text{and} \qquad \mathbb{E}[\Psi(Y,\Delta,X)]=0.$$

The covariance process of $G^*$ is

$$
\begin{aligned}
&Cov\left(G^*(h), G^*(\tilde{h})\right)\\
&= \mathbb{E}\left[\left\{f_h(Y,\Delta,X,Z) - Q_h\Gamma_1^{-1}\Psi(Y,\Delta,X)\right\}\left\{f_{\tilde{h}}(Y,\Delta,X,Z) - Q_{\tilde{h}}\Gamma_1^{-1}\Psi(Y,\Delta,X)\right\}\right]\\
&= \mathbb{E}[f_h(Y,\Delta,X,Z)f_{\tilde{h}}(Y,\Delta,X,Z)] - Q_h\Gamma_1^{-1}\mathbb{E}[f_{\tilde{h}}(Y,\Delta,X,Z)\Psi(Y,\Delta,X)]\\
&\quad - Q_{\tilde{h}}\Gamma_1^{-1}\mathbb{E}[f_h(Y,\Delta,X,Z)\Psi(Y,\Delta,X)] + Q_{\tilde{h}}\Gamma_1^{-1}\mathbb{E}[\Psi(Y,\Delta,X)\Psi(Y,\Delta,X)']\Gamma_1^{-1}Q_h'.
\end{aligned}
\tag{A34}
$$

*Condition 4 of Theorem 4 in [5].* As for condition 1, we write

$$
\begin{aligned}
\sqrt{n}\left\{(\hat{S}_n - S)(\Upsilon_n) - (\hat{S}_n - S)(\Upsilon_0)\right\} &= \sqrt{n}\left\{(S_n - S)(\Upsilon_n) - (S_n - S)(\Upsilon_0)\right\}\\
&\quad + \sqrt{n}\left\{(\hat{S}_n - S_n)(\Upsilon_n) - (\hat{S}_n - S_n)(\Upsilon_0)\right\}
\end{aligned}
\tag{A35}
$$

For the second term in the right hand side of (A35), similarly to (A31)-(A32), we have

$$
\begin{aligned}
&(\hat{S}_n - S_n)(\Upsilon_n) - (\hat{S}_n - S_n)(\Upsilon_0)\\
&= \frac{1}{n}\sum_{i=1}^n (1-\Delta_i)\mathbb{1}_{\{Y_i\leq\tau_0\}} e^{\hat{\beta}_n' Z_i}\left\{\int_0^{Y_i} h_1(s)\mathrm{d}\hat{\Lambda}_n(s) + \hat{\Lambda}_n(Y_i)h_2' Z_i\right\}\\
&\qquad\qquad \frac{\partial g_i}{\partial\phi}(Y_i,\hat{\Lambda}_n,\hat{\beta}_n,\gamma_0)\nabla_\gamma\phi(\gamma_0,X_i)'\left\{\hat{\gamma}_n - \gamma_0\right\}\\
&\quad - \frac{1}{n}\sum_{i=1}^n (1-\Delta_i)\mathbb{1}_{\{Y_i\leq\tau_0\}} e^{\beta_0' Z_i}\left\{\int_0^{Y_i} h_1(s)\mathrm{d}\Lambda_0(s) + \Lambda_0(Y_i)h_2' Z_i\right\}\\
&\qquad\qquad \frac{\partial g_i}{\partial\phi}(Y_i,\Lambda_0,\beta_0,\gamma_0)\nabla_\gamma\phi(\gamma_0,X_i)'\left\{\hat{\gamma}_n - \gamma_0\right\} + o_P(n^{-1/2})
\end{aligned}
$$

Using the boundedness in probability of $\hat{\beta}_n$ and $\hat{\Lambda}_n(\tau_0)$, the boundedness of the covariates, $\beta_0$, $\Lambda_0(\tau)$, $\nabla_\gamma\phi(\gamma,x)$ and the consistency results in Theorem 2, it follows that

$$
\begin{aligned}
&\left|\frac{1}{n}\sum_{i=1}^n (1-\Delta_i)\mathbb{1}_{\{Y_i\leq\tau_0\}}\left\{e^{\hat{\beta}_n' Z_i}\left(\int_0^{Y_i} h_1(s)\mathrm{d}\hat{\Lambda}_n(s) + \hat{\Lambda}_n(Y_i)h_2' Z_i\right)\frac{\partial g_i}{\partial\phi}(Y_i,\hat{\Lambda}_n,\hat{\beta}_n,\gamma_0)\right.\right.\\
&\quad\left.\left. - e^{\beta_0' Z_i}\left(\int_0^{Y_i} h_1(s)\mathrm{d}\Lambda_0(s) + \Lambda_0(Y_i)h_2' Z_i\right)\frac{\partial g_i}{\partial\phi}(Y_i,\Lambda_0,\beta_0,\gamma_0)\right\}\nabla_\gamma\phi(\gamma_0,X_i)\right| = o_P(1)
\end{aligned}
$$

As a consequence, since $\hat{\gamma}_n - \gamma_0 = O_P(n^{-1/2})$, we obtain

$$
\sqrt{n}\left\{(\hat{S}_n - S_n)(\Upsilon_n) - (\hat{S}_n - S_n)(\Upsilon_0)\right\} = o_P(1).
$$

It remains to deal with the first term on the right hand side of (A35). It suffices to show that, for any sequence $\epsilon_n \to 0$,

$$
\sup_{|\Lambda-\Lambda_0|_\infty\leq\epsilon_n,\,\|\beta-\beta_0\|\leq\epsilon_n} \frac{|(S_n - S)(\Upsilon) - (S_n - S)(\Upsilon_0)|}{n^{-1/2}\vee\|\beta-\beta_0\|\vee|\Lambda-\Lambda_0|_\infty} = o_P(1).
$$

Let

$$a_1(y,\delta,z,x) = \delta e^{\beta' z}\int_0^y h_1(s)\mathrm{d}\Lambda(s) - \delta e^{\beta_0' z}\int_0^y h_1(s)\mathrm{d}\Lambda_0(s),$$

$$a_2(y,\delta,z,x) = \delta e^{\beta' z}\Lambda(y)h_2' z - \delta e^{\beta_0' z}\Lambda_0(y)h_2' z,$$

and

$$\begin{aligned}a_3(y,\delta,z,x) &= (1-\delta)\mathbb{1}_{\{y\leq\tau_0\}}g(y,\Lambda,\beta,\gamma_0)\left\{e^{\beta' z}\int_0^y h_1(s)\mathrm{d}\Lambda(s) + e^{\beta' z}\Lambda(y)h_2' z\right\}\\ &\quad - (1-\delta)\mathbb{1}_{\{y\leq\tau_0\}}g(y,\Lambda_0,\beta_0,\gamma_0)\left\{e^{\beta_0' z}\int_0^y h_1(s)\mathrm{d}\Lambda_0(s) + e^{\beta_0' z}\Lambda_0(y)h_2' z\right\}\end{aligned}$$

Then, we have

$$\begin{aligned}(S_n - S)(\Upsilon) - (S_n - S)(\Upsilon_0) = &-\frac{1}{n}\sum_{i=1}^n \{a_1(Y_i,\Delta_i,Z_i,X_i) - \mathbb{E}\left[a_1(Y,\Delta,Z,X)\right]\}\\ &-\frac{1}{n}\sum_{i=1}^n \{a_2(Y_i,\Delta_i,Z_i,X_i) - \mathbb{E}\left[a_2(Y,\Delta,Z,X)\right]\}\\ &-\frac{1}{n}\sum_{i=1}^n \{a_3(Y_i,\Delta_i,Z_i,X_i) - \mathbb{E}\left[a_3(Y,\Delta,Z,X)\right]\}\end{aligned}$$

Next we consider the first term. The other two can be handled similarly. From a Taylor expansion we have

$$\begin{aligned}&\delta e^{\beta' z}\int_0^y h_1(s)\mathrm{d}\Lambda(s) - \delta e^{\beta_0' z}\int_0^y h_1(s)\mathrm{d}\Lambda_0(s)\\ &= (\beta-\beta_0)' z\delta e^{\beta_0' z}\int_0^y h_1(s)\mathrm{d}\Lambda(s) + \delta e^{\beta_0' z}\int_0^y h_1(s)\mathrm{d}(\Lambda-\Lambda_0)(s) + o(\|\beta-\beta_0\|).\end{aligned}$$

Hence

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^n \{a_1(Y_i,\Delta_i,Z_i,X_i) - \mathbb{E}\left[a_1(Y,\Delta,Z,X)\right]\}\\ &\leq (\beta-\beta_0)\left\{\frac{1}{n}\sum_{i=1}^n Z_i\Delta_i e^{\beta_0' Z_i}\int_0^{Y_i} h_1(s)\mathrm{d}\Lambda(s) - \mathbb{E}\left[Z\Delta e^{\beta_0' Z}\int_0^Y h_1(s)\mathrm{d}\Lambda(s)\right] + o(1)\right\}\\ &\quad + \frac{1}{n}\sum_{i=1}^n\left\{\Delta_i e^{\beta_0' Z_i}\int_0^{Y_i} h_1(s)\mathrm{d}(\Lambda-\Lambda_0)(s) - \mathbb{E}\left[\Delta e^{\beta_0' Z}\int_0^Y h_1(s)\mathrm{d}(\Lambda-\Lambda_0)(s)\right]\right\}.\end{aligned} \tag{A36}$$

By the law of large numbers

$$\frac{1}{n}\sum_{i=1}^n\left\{Z_i\Delta_i e^{\beta_0' Z_i}\int_0^{Y_i} h_1(s)\mathrm{d}\Lambda(s) - \mathbb{E}\left[Z\Delta e^{\beta_0' Z}\int_0^Y h_1(s)\mathrm{d}\Lambda(s)\right]\right\} = o_P(1)$$

and as a result, the first term in the right hand side of (A36) is $o_P(\|\beta-\beta_0\|)$. The second term can be rewritten as

$$\int_0^{\tau_0} D_n(s)h_1(s)\mathrm{d}(\Lambda-\Lambda_0)(s)$$

where

$$D_n(s)=\frac{1}{n}\sum_{i=1}^n\left\{\Delta_i\mathbb{1}_{\{Y_i\geq s\}}e^{\beta_0'Z_i}-\mathbb{E}\left[\Delta\mathbb{1}_{\{Y\geq s\}}e^{\beta_0'Z}\right]\right\}.$$

By integration by parts and the chain rule we have

$$\begin{aligned}
&\int_0^{\tau_0} D_n(s)h_1(s)\mathrm{d}(\Lambda-\Lambda_0)(s)\\
&=D_n(\tau_0)h_1(\tau_0)(\Lambda-\Lambda_0)(\tau_0)-\int_0^{\tau_0}(\Lambda-\Lambda_0)(s)\mathrm{d}\left[D_n(s)h_1(s)\right]\\
&=D_n(\tau_0)h_1(\tau_0)(\Lambda-\Lambda_0)(\tau_0)-\int_0^{\tau_0}(\Lambda-\Lambda_0)(s)D_n(s)\mathrm{d}h_1(s)\\
&\quad+\frac{1}{n}\sum_{i=1}^n\left\{\Delta_i(\Lambda-\Lambda_0)(Y_i)h_1(Y_i)e^{\beta_0'Z_i}-\mathbb{E}\left[(\Lambda-\Lambda_0)(Y)h_1(Y)\Delta e^{\beta_0'Z}\right]\right\}
\end{aligned}$$

Note that

$$\begin{aligned}
\mathbb{E}\left[(\Lambda-\Lambda_0)(Y)h_1(Y)\Delta e^{\beta_0'Z}\right]&=\mathbb{E}\left[e^{\beta_0'Z}\int_0^{\tau_0}(\Lambda-\Lambda_0)(s)h_1(s)\mathrm{d}H_1(s|X,Z)\middle|X,Z\right]\\
&=-\int_0^{\tau_0}(\Lambda-\Lambda_0)(s)h_1(s)\mathrm{d}\mathbb{E}\left[\Delta\mathbb{1}_{\{Y\geq s\}}e^{\beta_0'Z}\right].
\end{aligned}$$

It can be shown that $\sqrt{n}D_n$ converges weakly to a tight, mean zero Gaussian process $D$ in $l^\infty([0,\tau_0])$. Since $h_1$ is bounded, it follows

$$\frac{|D_n(\tau_0)h_1(\tau_0)(\Lambda-\Lambda_0)(\tau_0)|}{\|\Lambda-\Lambda_0\|_\infty}=o_P(1)$$

Moreover, since $D_n\to 0$ and $h_1$ is of bounded variation

$$\frac{\left|\int_0^{\tau_0}(\Lambda-\Lambda_0)(s)D_n(s)\mathrm{d}h_1(s)\right|}{\|\Lambda-\Lambda_0\|_\infty}\leq\sup_{t\in[0,\tau_0]}|D_n(s)|\int_0^{\tau_0}|\mathrm{d}h(s)|\to 0.$$

Finally, since $\left\{g_\Lambda(y,\delta,z)=\delta(\Lambda-\Lambda_0)(y)h_1(y)e^{\beta_0'Z}:\,\|\Lambda-\Lambda_0\|_\infty\leq\epsilon_n\right\}$ is a Donsker class (product of bounded variation functions, uniformly bounded) and

$$\mathbb{E}\left[(\Lambda-\Lambda_0)(Y)^2h_1(Y)^2\Delta e^{2\beta_0'Z}\right]=O(\epsilon_n^2)=o(1),$$

we have that

$$\sqrt{n}\frac{1}{n}\sum_{i=1}^n\left\{\Delta_i(\Lambda-\Lambda_0)(Y_i)h_1(Y_i)e^{\beta_0'Z_i}-\mathbb{E}\left[(\Lambda-\Lambda_0)(Y)h_1(Y)\Delta e^{\beta_0'Z}\right]\right\}$$

converges to zero in probability. So we obtain

$$\frac{\int_0^{\tau_0} D_n(s)h_1(s)\mathrm{d}(\Lambda-\Lambda_0)(s)}{n^{-1/2}\vee\|\Lambda-\Lambda_0\|_\infty}=o_P(1)$$

The other two terms related to $a_2$ and $a_3$ can be treated similarly.

This concludes the verification of conditions of Theorem 4 in [5] (or Theorem 3.3.1 in [8]). Hence, the weak convergence of $\sqrt{n}(\Upsilon_n-\Upsilon_0)$ to a tight, mean zero Gaussian process $G$ follows. Next we compute the covariance process of $G$. From Theorem 3.3.1 in [8] we have

$$-\sqrt{n}\dot{S}(\Upsilon_0)(\Upsilon_n-\Upsilon_0)(h)=\sqrt{n}(\hat{S}_n(\Upsilon_0)-S(\Upsilon_0))(h)+o_P(1). \tag{A37}$$

Moreover, in [5] it is computed that

$$\dot{S}(\Upsilon_0)(\Upsilon_n-\Upsilon_0)(h)=\int_0^{\tau_0}\sigma_1(h)(t)\mathrm{d}(\hat{\Lambda}_n(t)-\Lambda_0(t))+(\hat{\beta}_n-\beta_0)'\sigma_2(h) \tag{A38}$$

where $\sigma=(\sigma_1,\sigma_2)$ is a continuous linear operator from $\mathcal{H}_{\mathfrak{m}}$ to $\mathcal{H}_{\mathfrak{m}}$ of the form

$$\begin{aligned}\sigma_1(h)(t)&=\mathbb{E}\left[\mathbb{1}_{\{Y\geq t\}}V(t,\Upsilon_0)(h)g(t,\Upsilon_0)e^{\beta_0'Z}\right]\\&\quad-\mathbb{E}\left[\int_t^{\tau_0}\mathbb{1}_{\{Y\geq s\}}V(t,\Upsilon_0)(h)g(s,\Upsilon_0)\{1-g(s,\Upsilon_0)\}e^{2\beta_0'Z}\mathrm{d}\Lambda_0(s)\right]\end{aligned}$$

and

$$\sigma_2(h)(t)=\mathbb{E}\left[\int_0^{\tau_0}\mathbb{1}_{\{Y\geq t\}}W(t,\Upsilon_0)V(t,\Upsilon_0)(h)g(t,\Upsilon_0)e^{\beta_0'Z}\mathrm{d}\Lambda_0(t)\right]$$

where

$$V(t,\Upsilon_0)(h)=h_1(t)-\{1-g(t,\Upsilon_0)\}\,e^{\beta_0'Z}\int_0^t h_1(s)\mathrm{d}\Lambda_0(s)+h_2'W(t,\Upsilon_0)$$

and

$$W(t,\Upsilon_0)=\left[1-\{1-g(t,\Upsilon_0)\}\,e^{\beta_0'Z}\Lambda_0(t)\right]Z$$

In [5], it is also shown that $\sigma$ is invertible with inverse $\sigma^{-1}=(\sigma_1^{-1},\sigma_2^{-1})$. Hence, for all $g\in\mathcal{H}_{\mathfrak{m}}$, let $h=\sigma^{-1}(g)$. If in (A38) we replace $h$ by $\sigma^{-1}(g)$ and use (A37), we obtain

$$\begin{aligned}&\int_0^{\tau_0}g_1(t)\mathrm{d}\sqrt{n}(\Lambda_n(t)-\Lambda_0(t))+\sqrt{n}(\hat{\beta}_n-\beta_0)'g_2\\&=-\sqrt{n}(\hat{S}_n(\Upsilon_0)-S(\Upsilon_0))(\sigma^{-1}(g))+o_P(1)\xrightarrow{d}-G^*(\sigma^{-1}(g)).\end{aligned}$$

Since the previous results holds for all $g\in\mathcal{H}_{\mathfrak{m}}$, it follows that $(\sqrt{n}(\hat{\Lambda}_n-\Lambda_0),\sqrt{n}(\hat{\beta}_n-\beta_0))$ converges to a tight mean zero Gaussian process $G$ with covariance

$$Cov(G(g),G(\tilde{g}))=Cov\left(G^*(\sigma^{-1}(g)),G^*(\sigma^{-1}(\tilde{g}))\right) \tag{A39}$$

and the covariance of $G^*$ is given in (A34). □

## Appendix B: Additional simulation results

In this section we report the simulation results for scenario 2 of the models 1-4, $n = 200, 400$, that were omitted from the main paper and the results for sample size $n = 1000$ (all models and scenarios). In addition, Table A2 complements Tables 4 and 5 containing results for $\hat{\beta}$.

## References


[1] Ash, R. B. (1972) Real Analysis and Probability. *Academic Press*, New York.

[2] Burke, K. and Patilea, V. (2020) Supplementary Material for "A likelihood approach for cure regression models". *TEST*.

[3] Chen, X., Linton, O. and Van Keilegom, I. (2003) Estimation of semiparametric models when the criterion function is not smooth. *Econometrica*, 71 (5), 1591–1608.

[4] Murphy, S. A et. al. (1994) Consistency in a proportional hazards model incorporating a random effect. *Ann. Stat.*, 22 (2), 712–731.

[5] Lu, W. (2008) Maximum likelihood estimation in the proportional hazards cure model. *Ann. I. Stat. Math.*, 60 (3), 545–574.

[6] Patilea, V. and Van Keilegom, I. (2020) A general approach for cure models in survival analysis. *To apperar in: Ann. Stat.*

[7] Scharfstein, D. O., Tsiatis, A. A. and Gilbert, P. B. (1998) Semiparametric efficient estimation in the generalized odds-rate class of regression models for right-censored time-to-event data. *Lifetime Data Anal.*, **4** (4), 355–391.

[8] van der Vaart, A. W. and Wellner, J. A. (1996) Weak convergence and empirical processes. *Springer Series in Statistics*, Springer-Verlag, New York,With applications to statistics.

[9] Wellner, J. A. (1978) Limit theorems for the ratio of the empirical distribution function to the true distribution function. *Z. Wahrsch. Verw. Gebiete*, **45** (1), 73–88.

TABLE A1
*Bias, variance and MSE of $\hat{\gamma}$ and $\hat{\beta}$ for* `smcure` *(second rows) and our approach (first rows) in Model 1 and 2 (scenario 2).*

| | | | | Cens. level 1 | | | Cens. level 2 | | | Cens. level 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mod. | n | scen. | Par. | Bias | Var. | MSE | Bias | Var. | MSE | Bias | Var. | MSE |
| 1 | 200 | 2 | $\gamma_1$ | 0.005 | 0.032 | 0.032 | −0.005 | 0.037 | 0.037 | −0.005 | 0.045 | 0.045 |
| | | | | 0.013 | 0.032 | 0.032 | 0.009 | 0.038 | 0.038 | 0.015 | 0.046 | 0.046 |
| | | | $\gamma_2$ | −0.024 | 0.093 | 0.094 | −0.035 | 0.116 | 0.118 | −0.018 | 0.137 | 0.137 |
| | | | | 0.017 | 0.096 | 0.096 | 0.016 | 0.123 | 0.123 | 0.046 | 0.144 | 0.146 |
| | | | $\beta$ | 0.019 | 0.033 | 0.033 | 0.025 | 0.035 | 0.036 | 0.011 | 0.041 | 0.041 |
| | | | | 0.018 | 0.033 | 0.033 | 0.023 | 0.036 | 0.036 | 0.007 | 0.042 | 0.042 |
| | 400 | 2 | $\gamma_1$ | 0.002 | 0.016 | 0.016 | 0.003 | 0.018 | 0.018 | 0.009 | 0.021 | 0.021 |
| | | | | 0.008 | 0.016 | 0.016 | 0.011 | 0.019 | 0.019 | 0.020 | 0.021 | 0.022 |
| | | | $\gamma_2$ | −0.017 | 0.046 | 0.047 | −0.016 | 0.054 | 0.054 | −0.013 | 0.070 | 0.070 |
| | | | | 0.012 | 0.046 | 0.047 | 0.018 | 0.055 | 0.055 | 0.025 | 0.068 | 0.068 |
| | | | $\beta$ | 0.001 | 0.014 | 0.014 | 0.019 | 0.019 | 0.019 | 0.001 | 0.020 | 0.020 |
| | | | | 0.000 | 0.014 | 0.014 | 0.017 | 0.019 | 0.019 | −0.001 | 0.020 | 0.020 |
| 2 | 200 | 2 | $\gamma_1$ | −0.004 | 0.033 | 0.033 | −0.016 | 0.039 | 0.039 | −0.032 | 0.050 | 0.051 |
| | | | | 0.020 | 0.034 | 0.034 | 0.037 | 0.042 | 0.044 | 0.079 | 0.064 | 0.070 |
| | | | $\gamma_2$ | −0.016 | 0.041 | 0.041 | −0.037 | 0.044 | 0.045 | −0.054 | 0.057 | 0.060 |
| | | | | 0.029 | 0.044 | 0.045 | 0.031 | 0.052 | 0.053 | 0.064 | 0.075 | 0.079 |
| | | | $\beta$ | 0.004 | 0.016 | 0.016 | 0.008 | 0.017 | 0.017 | 0.008 | 0.018 | 0.018 |
| | | | | 0.002 | 0.016 | 0.016 | 0.003 | 0.018 | 0.018 | −0.005 | 0.019 | 0.019 |
| | 400 | 2 | $\gamma_1$ | 0.000 | 0.018 | 0.018 | −0.002 | 0.022 | 0.022 | −0.015 | 0.028 | 0.028 |
| | | | | 0.012 | 0.018 | 0.018 | 0.027 | 0.022 | 0.023 | 0.043 | 0.030 | 0.032 |
| | | | $\gamma_2$ | −0.019 | 0.021 | 0.021 | −0.022 | 0.025 | 0.026 | −0.045 | 0.031 | 0.033 |
| | | | | 0.007 | 0.022 | 0.022 | 0.018 | 0.027 | 0.027 | 0.023 | 0.033 | 0.033 |
| | | | $\beta$ | 0.005 | 0.007 | 0.007 | 0.010 | 0.008 | 0.008 | 0.009 | 0.010 | 0.010 |
| | | | | 0.004 | 0.007 | 0.007 | 0.007 | 0.008 | 0.008 | 0.001 | 0.011 | 0.011 |

Table A2

*Bias, variance and MSE of $\hat{\beta}$ for* `smcure` *(second rows) and our approach (first rows) in Model 3 and 4.*

| | | | | Cens. level 1 | | | Cens. level 2 | | | Cens. level 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mod. | n | scen. | Par. | Bias | Var. | MSE | Bias | Var. | MSE | Bias | Var. | MSE |
| 3 | 200 | 1 | $\beta_1$ | −0.013 | 0.007 | 0.007 | −0.007 | 0.006 | 0.006 | −0.009 | 0.008 | 0.008 |
| | | | | −0.014 | 0.007 | 0.007 | −0.009 | 0.006 | 0.007 | −0.014 | 0.008 | 0.008 |
| | | | $\beta_2$ | 0.007 | 0.003 | 0.003 | 0.006 | 0.003 | 0.003 | 0.005 | 0.004 | 0.004 |
| | | | | 0.007 | 0.003 | 0.003 | 0.008 | 0.003 | 0.003 | 0.007 | 0.004 | 0.004 |
| | | | $\beta_3$ | 0.016 | 0.043 | 0.043 | 0.004 | 0.043 | 0.043 | 0.007 | 0.053 | 0.053 |
| | | | | 0.017 | 0.043 | 0.043 | 0.006 | 0.043 | 0.043 | 0.015 | 0.053 | 0.054 |
| | | 3 | $\beta_1$ | 0.024 | 0.014 | 0.014 | 0.024 | 0.014 | 0.015 | 0.017 | 0.017 | 0.017 |
| | | | | 0.023 | 0.014 | 0.014 | 0.025 | 0.014 | 0.015 | 0.024 | 0.017 | 0.018 |
| | | | $\beta_2$ | −0.005 | 0.004 | 0.004 | −0.002 | 0.005 | 0.005 | −0.003 | 0.006 | 0.006 |
| | | | | −0.005 | 0.004 | 0.004 | −0.003 | 0.005 | 0.005 | −0.004 | 0.006 | 0.006 |
| | | | $\beta_3$ | 0.018 | 0.066 | 0.066 | −0.002 | 0.081 | 0.081 | 0.021 | 0.099 | 0.100 |
| | | | | 0.018 | 0.066 | 0.066 | 0.001 | 0.083 | 0.083 | 0.033 | 0.104 | 0.105 |
| | 400 | 1 | $\beta_1$ | −0.007 | 0.003 | 0.003 | −0.004 | 0.003 | 0.003 | −0.002 | 0.003 | 0.003 |
| | | | | −0.007 | 0.003 | 0.003 | −0.005 | 0.003 | 0.003 | −0.006 | 0.003 | 0.003 |
| | | | $\beta_2$ | 0.004 | 0.002 | 0.002 | 0.003 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 |
| | | | | 0.004 | 0.002 | 0.002 | 0.003 | 0.002 | 0.002 | 0.003 | 0.002 | 0.002 |
| | | | $\beta_3$ | 0.006 | 0.020 | 0.020 | 0.007 | 0.022 | 0.022 | 0.002 | 0.024 | 0.024 |
| | | | | 0.006 | 0.020 | 0.020 | 0.008 | 0.022 | 0.022 | 0.006 | 0.024 | 0.024 |
| | | 3 | $\beta_1$ | 0.013 | 0.006 | 0.006 | 0.009 | 0.007 | 0.007 | 0.003 | 0.008 | 0.008 |
| | | | | 0.012 | 0.006 | 0.006 | 0.009 | 0.007 | 0.007 | 0.007 | 0.008 | 0.008 |
| | | | $\beta_2$ | −0.002 | 0.002 | 0.002 | −0.001 | 0.003 | 0.003 | 0.000 | 0.003 | 0.003 |
| | | | | −0.002 | 0.002 | 0.002 | −0.001 | 0.003 | 0.003 | −0.001 | 0.003 | 0.003 |
| | | | $\beta_3$ | 0.008 | 0.030 | 0.030 | 0.000 | 0.035 | 0.035 | −0.007 | 0.045 | 0.045 |
| | | | | 0.007 | 0.030 | 0.030 | 0.001 | 0.035 | 0.035 | 0.001 | 0.047 | 0.047 |
| 4 | 200 | 1 | $\beta_1$ | −0.012 | 0.005 | 0.005 | −0.010 | 0.005 | 0.005 | −0.006 | 0.005 | 0.005 |
| | | | | −0.013 | 0.005 | 0.005 | −0.013 | 0.005 | 0.005 | −0.013 | 0.005 | 0.006 |
| | | | $\beta_2$ | 0.003 | 0.003 | 0.003 | 0.002 | 0.003 | 0.003 | 0.001 | 0.003 | 0.003 |
| | | | | 0.003 | 0.003 | 0.003 | 0.002 | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 |
| | | | $\beta_3$ | 0.002 | 0.032 | 0.032 | 0.000 | 0.036 | 0.036 | 0.000 | 0.040 | 0.040 |
| | | | | 0.002 | 0.032 | 0.032 | 0.002 | 0.037 | 0.037 | 0.005 | 0.041 | 0.041 |
| | | 3 | $\beta_1$ | 0.013 | 0.005 | 0.005 | 0.010 | 0.005 | 0.005 | 0.002 | 0.007 | 0.007 |
| | | | | 0.014 | 0.005 | 0.005 | 0.013 | 0.005 | 0.006 | 0.010 | 0.007 | 0.007 |
| | | | $\beta_2$ | −0.003 | 0.004 | 0.004 | −0.002 | 0.005 | 0.005 | −0.002 | 0.006 | 0.006 |
| | | | | −0.004 | 0.004 | 0.004 | −0.003 | 0.005 | 0.005 | −0.004 | 0.006 | 0.006 |
| | | | $\beta_3$ | −0.001 | 0.057 | 0.057 | −0.008 | 0.067 | 0.067 | −0.023 | 0.085 | 0.085 |
| | | | | 0.001 | 0.057 | 0.057 | −0.001 | 0.068 | 0.068 | −0.005 | 0.089 | 0.089 |
| | 400 | 1 | $\beta_1$ | −0.005 | 0.002 | 0.002 | −0.003 | 0.002 | 0.002 | 0.001 | 0.003 | 0.003 |
| | | | | −0.005 | 0.002 | 0.002 | −0.005 | 0.002 | 0.002 | −0.003 | 0.003 | 0.003 |
| | | | $\beta_2$ | 0.000 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 | −0.001 | 0.001 | 0.001 |
| | | | | 0.000 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 | −0.001 | 0.001 | 0.001 |
| | | | $\beta_3$ | 0.002 | 0.015 | 0.015 | 0.000 | 0.017 | 0.017 | −0.002 | 0.018 | 0.018 |
| | | | | 0.002 | 0.015 | 0.015 | 0.000 | 0.017 | 0.017 | −0.001 | 0.018 | 0.018 |
| | | 3 | $\beta_1$ | 0.006 | 0.002 | 0.002 | 0.004 | 0.003 | 0.003 | 0.000 | 0.003 | 0.003 |
| | | | | 0.006 | 0.002 | 0.002 | 0.006 | 0.003 | 0.003 | 0.006 | 0.004 | 0.004 |
| | | | $\beta_2$ | −0.004 | 0.002 | 0.002 | −0.004 | 0.003 | 0.003 | −0.004 | 0.003 | 0.003 |
| | | | | −0.005 | 0.002 | 0.002 | −0.005 | 0.003 | 0.003 | −0.005 | 0.003 | 0.003 |
| | | | $\beta_3$ | 0.001 | 0.028 | 0.028 | −0.004 | 0.033 | 0.033 | −0.013 | 0.046 | 0.046 |
| | | | | 0.001 | 0.028 | 0.028 | −0.001 | 0.034 | 0.034 | 0.000 | 0.048 | 0.048 |

TABLE A3
*Bias, variance and MSE of $\hat{\gamma}$ and $\hat{\beta}$ for* `smcure` *(second rows) and our approach (first rows) in Model 3 (Scenario 2).*

| | | | | Cens. level 1 | | | Cens. level 2 | | | Cens. level 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mod. | n | scen. | Par. | Bias | Var. | MSE | Bias | Var. | MSE | Bias | Var. | MSE |
| 3 | 200 | 2 | $\gamma_1$ | −0.056 | 0.261 | 0.264 | −0.215 | 0.344 | 0.390 | −0.379 | 0.422 | 0.566 |
| | | | | 0.091 | 0.299 | 0.308 | 0.090 | 0.454 | 0.462 | 0.127 | 0.755 | 0.771 |
| | | | $\gamma_2$ | −0.055 | 0.124 | 0.127 | −0.123 | 0.137 | 0.152 | −0.198 | 0.165 | 0.204 |
| | | | | 0.183 | 0.178 | 0.212 | 0.275 | 0.268 | 0.344 | 0.390 | 0.485 | 0.638 |
| | | | $\gamma_3$ | 0.023 | 0.401 | 0.402 | 0.087 | 0.512 | 0.519 | 0.162 | 0.612 | 0.639 |
| | | | | 0.161 | 0.488 | 0.514 | 0.278 | 0.729 | 0.806 | 0.427 | 1.159 | 1.341 |
| | | | $\gamma_4$ | −0.049 | 0.321 | 0.324 | −0.093 | 0.421 | 0.429 | −0.164 | 0.543 | 0.570 |
| | | | | 0.036 | 0.367 | 0.369 | 0.086 | 0.547 | 0.554 | 0.062 | 0.795 | 0.799 |
| | | | $\beta_1$ | 0.017 | 0.008 | 0.009 | 0.017 | 0.009 | 0.009 | 0.013 | 0.010 | 0.010 |
| | | | | 0.016 | 0.008 | 0.009 | 0.018 | 0.009 | 0.009 | 0.015 | 0.010 | 0.011 |
| | | | $\beta_2$ | 0.004 | 0.004 | 0.004 | 0.007 | 0.004 | 0.004 | 0.005 | 0.005 | 0.005 |
| | | | | 0.004 | 0.004 | 0.004 | 0.009 | 0.004 | 0.004 | 0.008 | 0.005 | 0.005 |
| | | | $\beta_3$ | 0.022 | 0.059 | 0.059 | 0.008 | 0.059 | 0.060 | −0.003 | 0.074 | 0.074 |
| | | | | 0.024 | 0.059 | 0.060 | 0.015 | 0.060 | 0.060 | 0.013 | 0.074 | 0.075 |
| | 400 | 2 | $\gamma_1$ | −0.037 | 0.120 | 0.121 | −0.145 | 0.174 | 0.195 | −0.255 | 0.240 | 0.305 |
| | | | | 0.047 | 0.122 | 0.124 | 0.029 | 0.179 | 0.179 | 0.042 | 0.265 | 0.267 |
| | | | $\gamma_2$ | −0.059 | 0.056 | 0.060 | −0.123 | 0.075 | 0.090 | −0.177 | 0.077 | 0.0108 |
| | | | | 0.090 | 0.068 | 0.076 | 0.110 | 0.098 | 0.110 | 0.173 | 0.125 | 0.155 |
| | | | $\gamma_3$ | −0.020 | 0.181 | 0.182 | 0.026 | 0.245 | 0.246 | 0.080 | 0.337 | 0.344 |
| | | | | 0.066 | 0.194 | 0.198 | 0.124 | 0.261 | 0.276 | 0.212 | 0.380 | 0.425 |
| | | | $\gamma_4$ | −0.021 | 0.151 | 0.152 | −0.069 | 0.220 | 0.225 | −0.122 | 0.315 | 0.329 |
| | | | | 0.023 | 0.156 | 0.157 | 0.022 | 0.218 | 0.218 | 0.033 | 0.309 | 0.310 |
| | | | $\beta_1$ | 0.011 | 0.004 | 0.004 | 0.008 | 0.004 | 0.004 | 0.013 | 0.005 | 0.005 |
| | | | | 0.011 | 0.004 | 0.004 | 0.007 | 0.004 | 0.004 | 0.013 | 0.005 | 0.005 |
| | | | $\beta_2$ | 0.003 | 0.002 | 0.002 | 0.003 | 0.002 | 0.002 | 0.001 | 0.002 | 0.002 |
| | | | | 0.003 | 0.002 | 0.002 | 0.004 | 0.002 | 0.002 | 0.003 | 0.002 | 0.002 |
| | | | $\beta_3$ | 0.019 | 0.028 | 0.029 | 0.002 | 0.029 | 0.029 | −0.001 | 0.033 | 0.033 |
| | | | | 0.020 | 0.028 | 0.029 | 0.006 | 0.029 | 0.029 | 0.009 | 0.034 | 0.034 |

TABLE A4
*Bias, variance and MSE of $\hat{\gamma}$ and $\hat{\beta}$ for `smcure` (second rows) and our approach (first rows) in Model 4 (Scenario 2).*

| | | | | Cens. level 1 | | | Cens. level 2 | | | Cens. level 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mod. | n | scen. | Par. | Bias | Var. | MSE | Bias | Var. | MSE | Bias | Var. | MSE |
| 4 | 200 | 2 | $\gamma_1$ | −0.121 | 0.117 | 0.131 | −0.240 | 0.140 | 0.198 | −0.375 | 0.179 | 0.320 |
| | | | | 0.006 | 0.134 | 0.134 | −0.011 | 0.170 | 0.170 | −0.035 | 0.231 | 0.232 |
| | | | $\gamma_2$ | 0.053 | 0.014 | 0.017 | 0.091 | 0.017 | 0.026 | 0.122 | 0.018 | 0.033 |
| | | | | 0.041 | 0.016 | 0.018 | 0.056 | 0.020 | 0.024 | 0.077 | 0.027 | 0.033 |
| | | | $\gamma_3$ | −0.195 | 0.137 | 0.175 | −0.348 | 0.156 | 0.277 | −0.511 | 0.165 | 0.426 |
| | | | | 0.119 | 0.169 | 0.183 | 0.122 | 0.216 | 0.231 | 0.139 | 0.273 | 0.293 |
| | | | $\gamma_4$ | 0.096 | 0.156 | 0.165 | 0.157 | 0.189 | 0.213 | 0.240 | 0.229 | 0.287 |
| | | | | 0.072 | 0.182 | 0.187 | 0.096 | 0.216 | 0.226 | 0.137 | 0.288 | 0.307 |
| | | | $\gamma_5$ | −0.038 | 0.161 | 0.163 | −0.078 | 0.185 | 0.191 | −0.149 | 0.218 | 0.240 |
| | | | | 0.046 | 0.175 | 0.177 | 0.058 | 0.220 | 0.223 | 0.047 | 0.254 | 0.257 |
| | | | $\beta_1$ | 0.014 | 0.007 | 0.007 | 0.009 | 0.008 | 0.008 | 0.001 | 0.008 | 0.008 |
| | | | | 0.017 | 0.007 | 0.007 | 0.016 | 0.008 | 0.008 | 0.014 | 0.008 | 0.008 |
| | | | $\beta_2$ | 0.005 | 0.004 | 0.004 | 0.003 | 0.004 | 0.004 | 0.002 | 0.004 | 0.005 |
| | | | | 0.005 | 0.004 | 0.004 | 0.005 | 0.004 | 0.004 | 0.005 | 0.004 | 0.005 |
| | | | $\beta_3$ | 0.010 | 0.056 | 0.056 | −0.003 | 0.063 | 0.063 | −0.022 | 0.071 | 0.072 |
| | | | | 0.015 | 0.057 | 0.057 | 0.011 | 0.064 | 0.064 | 0.006 | 0.072 | 0.072 |
| | 400 | 2 | $\gamma_1$ | −0.068 | 0.057 | 0.062 | −0.169 | 0.073 | 0.102 | −0.279 | 0.089 | 0.167 |
| | | | | 0.015 | 0.061 | 0.061 | 0.001 | 0.078 | 0.078 | −0.022 | 0.097 | 0.098 |
| | | | $\gamma_2$ | 0.033 | 0.007 | 0.008 | 0.068 | 0.008 | 0.013 | 0.099 | 0.010 | 0.019 |
| | | | | 0.019 | 0.008 | 0.008 | 0.030 | 0.010 | 0.011 | 0.044 | 0.012 | 0.014 |
| | | | $\gamma_3$ | −0.180 | 0.073 | 0.105 | −0.311 | 0.084 | 0.181 | −0.450 | 0.093 | 0.296 |
| | | | | 0.043 | 0.079 | 0.081 | 0.038 | 0.096 | 0.097 | 0.030 | 0.119 | 0.120 |
| | | | $\gamma_4$ | 0.050 | 0.078 | 0.080 | 0.119 | 0.096 | 0.110 | 0.177 | 0.118 | 0.149 |
| | | | | 0.028 | 0.082 | 0.083 | 0.049 | 0.101 | 0.104 | 0.066 | 0.124 | 0.129 |
| | | | $\gamma_5$ | −0.060 | 0.077 | 0.081 | −0.106 | 0.096 | 0.107 | −0.150 | 0.101 | 0.124 |
| | | | | 0.003 | 0.082 | 0.082 | 0.001 | 0.099 | 0.099 | 0.003 | 0.116 | 0.116 |
| | | | $\beta_1$ | 0.006 | 0.003 | 0.003 | 0.002 | 0.004 | 0.004 | −0.004 | 0.004 | 0.004 |
| | | | | 0.008 | 0.003 | 0.003 | 0.008 | 0.004 | 0.004 | 0.007 | 0.004 | 0.004 |
| | | | $\beta_2$ | 0.002 | 0.002 | 0.002 | 0.000 | 0.002 | 0.002 | 0.000 | 0.002 | 0.002 |
| | | | | 0.002 | 0.002 | 0.002 | 0.001 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 |
| | | | $\beta_3$ | 0.003 | 0.026 | 0.026 | −0.010 | 0.030 | 0.030 | −0.023 | 0.034 | 0.034 |
| | | | | 0.007 | 0.026 | 0.026 | 0.002 | 0.031 | 0.031 | −0.001 | 0.034 | 0.034 |

Table A5
*Bias, variance and MSE of $\hat{\gamma}$ and $\hat{\beta}$ for* `smcure` *(second rows) and our approach (first rows) in Model 1 and 2 ($n = 1000$).*

| | | | | Cens. level 1 | | | Cens. level 2 | | | Cens. level 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mod. | n | scen. | Par. | Bias | Var. | MSE | Bias | Var. | MSE | Bias | Var. | MSE |
| 1 | 1000 | 1 | $\gamma_1$ | 0.007 | 0.011 | 0.011 | 0.008 | 0.013 | 0.013 | 0.010 | 0.015 | 0.015 |
| | | | | 0.016 | 0.011 | 0.012 | 0.019 | 0.013 | 0.013 | 0.022 | 0.015 | 0.016 |
| | | | $\gamma_2$ | −0.006 | 0.032 | 0.032 | −0.006 | 0.037 | 0.037 | −0.006 | 0.041 | 0.041 |
| | | | | 0.023 | 0.033 | 0.034 | 0.027 | 0.038 | 0.039 | 0.028 | 0.043 | 0.044 |
| | | | $\beta$ | 0.001 | 0.005 | 0.005 | 0.002 | 0.006 | 0.006 | 0.002 | 0.006 | 0.006 |
| | | | | 0.000 | 0.005 | 0.005 | 0.001 | 0.006 | 0.006 | 0.000 | 0.006 | 0.006 |
| | | 2 | $\gamma_1$ | 0.005 | 0.007 | 0.007 | 0.006 | 0.008 | 0.008 | 0.007 | 0.010 | 0.010 |
| | | | | 0.008 | 0.007 | 0.007 | 0.011 | 0.008 | 0.008 | 0.014 | 0.010 | 0.010 |
| | | | $\gamma_2$ | −0.008 | 0.019 | 0.019 | −0.006 | 0.022 | 0.022 | −0.007 | 0.026 | 0.026 |
| | | | | 0.012 | 0.019 | 0.020 | 0.016 | 0.023 | 0.023 | 0.018 | 0.027 | 0.027 |
| | | | $\beta$ | 0.000 | 0.006 | 0.006 | −0.001 | 0.007 | 0.007 | 0.000 | 0.008 | 0.008 |
| | | | | −0.001 | 0.006 | 0.006 | −0.002 | 0.007 | 0.007 | −0.002 | 0.008 | 0.008 |
| | | 3 | $\gamma_1$ | 0.001 | 0.011 | 0.011 | −0.008 | 0.012 | 0.012 | −0.003 | 0.018 | 0.018 |
| | | | | 0.004 | 0.011 | 0.011 | −0.002 | 0.013 | 0.013 | 0.009 | 0.019 | 0.019 |
| | | | $\gamma_2$ | −0.064 | 0.093 | 0.097 | −0.076 | 0.123 | 0.129 | −0.107 | 0.163 | 0.175 |
| | | | | 0.046 | 0.099 | 0.102 | 0.058 | 0.130 | 0.134 | 0.072 | 0.167 | 0.173 |
| | | | $\beta$ | 0.011 | 0.018 | 0.018 | 0.005 | 0.022 | 0.022 | 0.004 | 0.026 | 0.026 |
| | | | | 0.009 | 0.018 | 0.018 | 0.001 | 0.022 | 0.022 | −0.005 | 0.026 | 0.026 |
| 2 | 1000 | 1 | $\gamma_1$ | 0.006 | 0.008 | 0.008 | 0.000 | 0.009 | 0.009 | 0.007 | 0.012 | 0.012 |
| | | | | 0.009 | 0.008 | 0.008 | 0.007 | 0.009 | 0.009 | 0.022 | 0.013 | 0.013 |
| | | | $\gamma_2$ | −0.010 | 0.007 | 0.007 | −0.011 | 0.009 | 0.009 | −0.018 | 0.011 | 0.011 |
| | | | | −0.002 | 0.007 | 0.007 | −0.002 | 0.009 | 0.009 | −0.005 | 0.012 | 0.012 |
| | | | $\beta$ | −0.001 | 0.002 | 0.002 | 0.000 | 0.003 | 0.003 | 0.001 | 0.003 | 0.003 |
| | | | | −0.001 | 0.002 | 0.002 | −0.001 | 0.003 | 0.003 | −0.001 | 0.003 | 0.003 |
| | | 2 | $\gamma_1$ | −0.004 | 0.006 | 0.006 | −0.010 | 0.007 | 0.007 | −0.004 | 0.010 | 0.010 |
| | | | | 0.002 | 0.006 | 0.006 | 0.003 | 0.008 | 0.008 | 0.022 | 0.010 | 0.011 |
| | | | $\gamma_2$ | −0.012 | 0.007 | 0.008 | −0.023 | 0.009 | 0.010 | −0.031 | 0.012 | 0.012 |
| | | | | 0.004 | 0.008 | 0.008 | 0.000 | 0.009 | 0.009 | 0.002 | 0.012 | 0.012 |
| | | | $\beta$ | 0.005 | 0.003 | 0.003 | 0.006 | 0.003 | 0.003 | 0.005 | 0.004 | 0.004 |
| | | | | 0.004 | 0.003 | 0.003 | 0.005 | 0.003 | 0.003 | 0.001 | 0.004 | 0.004 |
| | | 3 | $\gamma_1$ | −0.011 | 0.014 | 0.014 | −0.029 | 0.014 | 0.015 | −0.054 | 0.019 | 0.022 |
| | | | | 0.002 | 0.016 | 0.016 | 0.015 | 0.018 | 0.018 | 0.028 | 0.029 | 0.029 |
| | | | $\gamma_2$ | −0.314 | 0.144 | 0.242 | −0.583 | 0.189 | 0.529 | −0.841 | 0.237 | 0.945 |
| | | | | 0.045 | 0.148 | 0.151 | 0.068 | 0.175 | 0.180 | 0.137 | 0.223 | 0.242 |
| | | | $\beta$ | 0.004 | 0.006 | 0.006 | 0.004 | 0.008 | 0.008 | 0.006 | 0.010 | 0.010 |
| | | | | 0.002 | 0.006 | 0.006 | 0.001 | 0.008 | 0.008 | 0.003 | 0.010 | 0.010 |

Table A6
*Bias, variance and MSE of $\hat{\gamma}$ and $\hat{\beta}$ for `smcure` and our approach in Model 3 ($n = 1000$).*

| | | | Cens. level 1 | | | Cens. level 2 | | | Cens. level 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| n | scen. | Par. | Bias | Var. | MSE | Bias | Var. | MSE | Bias | Var. | MSE |
| 1000 | 1 | $\gamma_1$ | 0.000 | 0.026 | 0.026 | 0.010 | 0.035 | 0.035 | 0.013 | 0.050 | 0.050 |
| | | | 0.001 | 0.026 | 0.026 | 0.003 | 0.031 | 0.031 | 0.006 | 0.040 | 0.040 |
| | | $\gamma_2$ | −0.009 | 0.008 | 0.008 | −0.021 | 0.001 | 0.011 | −0.051 | 0.014 | 0.017 |
| | | | −0.018 | 0.008 | 0.008 | −0.022 | 0.010 | 0.010 | −0.036 | 0.013 | 0.014 |
| | | $\gamma_3$ | 0.019 | 0.071 | 0.072 | 0.022 | 0.094 | 0.094 | 0.022 | 0.128 | 0.128 |
| | | | 0.041 | 0.069 | 0.070 | 0.060 | 0.078 | 0.082 | 0.062 | 0.105 | 0.109 |
| | | $\gamma_4$ | 0.004 | 0.061 | 0.061 | −0.030 | 0.087 | 0.088 | −0.095 | 0.124 | 0.133 |
| | | | 0.026 | 0.058 | 0.058 | −0.002 | 0.077 | 0.078 | 0.033 | 0.098 | 0.099 |
| | | $\beta_1$ | −0.002 | 0.001 | 0.001 | −0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| | | | −0.002 | 0.001 | 0.001 | −0.002 | 0.001 | 0.001 | −0.002 | 0.001 | 0.001 |
| | | $\beta_2$ | 0.002 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| | | | 0.002 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| | | $\beta_3$ | 0.001 | 0.007 | 0.007 | −0.006 | 0.008 | 0.008 | −0.003 | 0.009 | 0.009 |
| | | | 0.001 | 0.007 | 0.007 | −0.006 | 0.0018 | 0.008 | −0.001 | 0.009 | 0.009 |
| | 2 | $\gamma_1$ | −0.025 | 0.046 | 0.047 | −0.071 | 0.077 | 0.082 | −0.142 | 0.117 | 0.137 |
| | | | 0.020 | 0.046 | 0.047 | 0.007 | 0.066 | 0.066 | 0.005 | 0.090 | 0.090 |
| | | $\gamma_2$ | −0.057 | 0.022 | 0.025 | −0.097 | 0.030 | 0.040 | −0.108 | 0.038 | 0.050 |
| | | | 0.032 | 0.023 | 0.024 | 0.037 | 0.031 | 0.032 | 0.056 | 0.038 | 0.041 |
| | | $\gamma_3$ | −0.041 | 0.068 | 0.070 | −0.024 | 0.125 | 0.126 | 0.015 | 0.169 | 0.170 |
| | | | 0.009 | 0.066 | 0.066 | 0.044 | 0.104 | 0.106 | 0.059 | 0.132 | 0.136 |
| | | $\gamma_4$ | −0.019 | 0.059 | 0.060 | −0.025 | 0.093 | 0.093 | −0.037 | 0.127 | 0.128 |
| | | | 0.005 | 0.058 | 0.058 | 0.020 | 0.081 | 0.081 | 0.039 | 0.101 | 0.102 |
| | | $\beta_1$ | 0.003 | 0.001 | 0.001 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 |
| | | | 0.003 | 0.001 | 0.001 | 0.001 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 |
| | | $\beta_2$ | 0.001 | 0.001 | 0.001 | 0.002 | 0.001 | 0.001 | −0.002 | 0.001 | 0.001 |
| | | | 0.001 | 0.001 | 0.001 | 0.002 | 0.001 | 0.001 | −0.002 | 0.001 | 0.001 |
| | | $\beta_3$ | 0.003 | 0.010 | 0.010 | 0.004 | 0.011 | 0.011 | −0.010 | 0.013 | 0.013 |
| | | | 0.003 | 0.010 | 0.010 | 0.005 | 0.011 | 0.011 | −0.004 | 0.013 | 0.013 |
| | 3 | $\gamma_1$ | 0.004 | 0.030 | 0.031 | −0.008 | 0.050 | 0.050 | −0.058 | 0.093 | 0.096 |
| | | | −0.001 | 0.029 | 0.029 | −0.008 | 0.043 | 0.043 | −0.022 | 0.069 | 0.070 |
| | | $\gamma_2$ | −0.014 | 0.009 | 0.009 | −0.011 | 0.014 | 0.014 | 0.031 | 0.022 | 0.023 |
| | | | 0.015 | 0.009 | 0.009 | 0.018 | 0.012 | 0.012 | 0.045 | 0.018 | 0.020 |
| | | $\gamma_3$ | 0.000 | 0.044 | 0.044 | −0.004 | 0.072 | 0.072 | 0.049 | 0.123 | 0.125 |
| | | | 0.017 | 0.043 | 0.043 | 0.016 | 0.058 | 0.058 | 0.043 | 0.089 | 0.091 |
| | | $\gamma_4$ | −0.002 | 0.037 | 0.037 | 0.002 | 0.060 | 0.060 | −0.020 | 0.099 | 0.099 |
| | | | −0.007 | 0.036 | 0.036 | 0.003 | 0.051 | 0.051 | −0.012 | 0.071 | 0.072 |
| | | $\beta_1$ | 0.007 | 0.002 | 0.002 | 0.008 | 0.003 | 0.003 | 0.005 | 0.003 | 0.003 |
| | | | 0.006 | 0.002 | 0.002 | 0.007 | 0.003 | 0.003 | 0.006 | 0.003 | 0.003 |
| | | $\beta_2$ | 0.000 | 0.001 | 0.001 | −0.002 | 0.001 | 0.001 | −0.001 | 0.001 | 0.001 |
| | | | 0.000 | 0.001 | 0.001 | −0.002 | 0.001 | 0.001 | −0.001 | 0.001 | 0.001 |
| | | $\beta_3$ | 0.000 | 0.001 | 0.001 | 0.014 | 0.013 | 0.013 | −0.004 | 0.017 | 0.017 |
| | | | 0.000 | 0.001 | 0.001 | 0.014 | 0.012 | 0.013 | −0.001 | 0.016 | 0.016 |

Table A7
*Bias, variance and MSE of $\hat{\gamma}$ and $\hat{\beta}$ for* `smcure` *and our approach in Model 4 ($n = 1000$).*

| | | | Cens. level 1 | | | Cens. level 2 | | | Cens. level 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| n | scen. | Par. | Bias | Var. | MSE | Bias | Var. | MSE | Bias | Var. | MSE |
| 1000 | 1 | $\gamma_1$ | −0.002 | 0.027 | 0.027 | 0.007 | 0.033 | 0.033 | 0.010 | 0.038 | 0.038 |
| | | | 0.004 | 0.027 | 0.027 | 0.008 | 0.032 | 0.032 | 0.012 | 0.038 | 0.038 |
| | | $\gamma_2$ | 0.013 | 0.007 | 0.007 | −0.001 | 0.008 | 0.008 | −0.005 | 0.009 | 0.009 |
| | | | −0.015 | 0.007 | 0.007 | −0.024 | 0.009 | 0.009 | −0.029 | 0.011 | 0.012 |
| | | $\gamma_3$ | −0.192 | 0.041 | 0.078 | −0.149 | 0.037 | 0.059 | −0.196 | 0.046 | 0.084 |
| | | | 0.020 | 0.037 | 0.037 | 0.029 | 0.047 | 0.048 | 0.035 | 0.058 | 0.059 |
| | | $\gamma_4$ | −0.054 | 0.063 | 0.066 | −0.086 | 0.076 | 0.084 | −0.155 | 0.090 | 0.114 |
| | | | 0.030 | 0.067 | 0.068 | 0.047 | 0.083 | 0.085 | 0.053 | 0.104 | 0.107 |
| | | $\gamma_5$ | −0.043 | 0.060 | 0.062 | −0.108 | 0.075 | 0.087 | −0.188 | 0.083 | 0.118 |
| | | | 0.025 | 0.061 | 0.062 | 0.035 | 0.075 | 0.076 | 0.039 | 0.092 | 0.093 |
| | | $\beta_1$ | −0.001 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 | 0.002 | 0.001 | 0.001 |
| | | | −0.002 | 0.001 | 0.001 | −0.002 | 0.001 | 0.001 | −0.001 | 0.001 | 0.001 |
| | | $\beta_2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 |
| | | | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 |
| | | $\beta_3$ | 0.001 | 0.006 | 0.006 | 0.002 | 0.006 | 0.006 | 0.002 | 0.007 | 0.007 |
| | | | 0.001 | 0.006 | 0.006 | 0.002 | 0.007 | 0.007 | 0.003 | 0.007 | 0.007 |
| | 2 | $\gamma_1$ | −0.049 | 0.024 | 0.027 | −0.116 | 0.031 | 0.044 | −0.207 | 0.040 | 0.083 |
| | | | 0.006 | 0.024 | 0.024 | −0.004 | 0.031 | 0.031 | −0.020 | 0.039 | 0.039 |
| | | $\gamma_2$ | 0.019 | 0.003 | 0.003 | 0.049 | 0.003 | 0.006 | 0.079 | 0.004 | 0.011 |
| | | | 0.006 | 0.003 | 0.003 | 0.011 | 0.004 | 0.004 | 0.018 | 0.005 | 0.005 |
| | | $\gamma_3$ | −0.142 | 0.031 | 0.051 | −0.229 | 0.039 | 0.091 | −0.330 | 0.045 | 0.154 |
| | | | 0.017 | 0.032 | 0.033 | 0.021 | 0.040 | 0.040 | 0.017 | 0.046 | 0.046 |
| | | $\gamma_4$ | 0.029 | 0.033 | 0.034 | 0.079 | 0.041 | 0.047 | 0.144 | 0.051 | 0.072 |
| | | | 0.012 | 0.033 | 0.033 | 0.026 | 0.040 | 0.040 | 0.039 | 0.047 | 0.049 |
| | | $\gamma_5$ | −0.034 | 0.033 | 0.034 | −0.068 | 0.039 | 0.044 | −0.103 | 0.048 | 0.058 |
| | | | 0.005 | 0.032 | 0.032 | 0.006 | 0.036 | 0.036 | 0.007 | 0.041 | 0.042 |
| | | $\beta_1$ | 0.001 | 0.001 | 0.001 | −0.003 | 0.001 | 0.001 | −0.008 | 0.001 | 0.001 |
| | | | 0.002 | 0.001 | 0.001 | 0.002 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| | | $\beta_2$ | 0.000 | 0.001 | 0.001 | −0.001 | 0.001 | 0.001 | −0.001 | 0.001 | 0.001 |
| | | | 0.001 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| | | $\beta_3$ | 0.000 | 0.010 | 0.010 | −0.009 | 0.011 | 0.011 | −0.019 | 0.012 | 0.012 |
| | | | 0.002 | 0.010 | 0.010 | 0.000 | 0.011 | 0.011 | −0.001 | 0.012 | 0.012 |
| | 3 | $\gamma_1$ | −0.002 | 0.015 | 0.015 | −0.012 | 0.019 | 0.019 | −0.029 | 0.026 | 0.027 |
| | | | 0.004 | 0.015 | 0.015 | 0.001 | 0.018 | 0.018 | 0.001 | 0.023 | 0.023 |
| | | $\gamma_2$ | −0.005 | 0.001 | 0.001 | −0.001 | 0.002 | 0.002 | 0.006 | 0.002 | 0.002 |
| | | | 0.007 | 0.002 | 0.002 | 0.004 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 |
| | | $\gamma_3$ | 0.051 | 0.016 | 0.019 | 0.034 | 0.015 | 0.017 | 0.041 | 0.021 | 0.022 |
| | | | −0.009 | 0.014 | 0.014 | −0.006 | 0.016 | 0.016 | −0.006 | 0.021 | 0.021 |
| | | $\gamma_4$ | 0.000 | 0.020 | 0.020 | 0.009 | 0.024 | 0.024 | 0.018 | 0.031 | 0.031 |
| | | | −0.001 | 0.019 | 0.019 | 0.004 | 0.023 | 0.023 | 0.004 | 0.028 | 0.028 |
| | | $\gamma_5$ | −0.006 | 0.019 | 0.019 | −0.011 | 0.024 | 0.024 | −0.015 | 0.033 | 0.033 |
| | | | −0.005 | 0.019 | 0.019 | −0.007 | 0.022 | 0.022 | −0.011 | 0.028 | 0.028 |
| | | $\beta_1$ | 0.003 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | −0.001 | 0.001 | 0.001 |
| | | | 0.002 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| | | $\beta_2$ | 0.000 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 |
| | | | −0.001 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 |
| | | $\beta_3$ | 0.003 | 0.010 | 0.010 | 0.001 | 0.012 | 0.012 | −0.005 | 0.016 | 0.016 |
| | | | 0.003 | 0.010 | 0.010 | 0.002 | 0.012 | 0.012 | −0.001 | 0.016 | 0.016 |