SUPPLEMENTARY MATERIAL FOR ESTIMATION FROM CROSS-SECTIONAL DATA UNDER A SEMIPARAMETRIC TRUNCATION

Cédric Heuchenne, Jacobo De uña Alvarez, Géraldine Laurent

REPRINT | 2021 / 33

UCLouvain

LIDAM
ISBA | Louvain Institute of Data Analysis and Modeling in economics and statistics

Biometrika (201.), .., .., *pp*. 1–15
© 201. Biometrika Trust
Printed in Great Britain

Advance Access publication on ..

Supplementary material for Estimation from cross-sectional data under a semiparametric truncation model

BY J. DE UÑA-ÁLVAREZ
Department of Statistics and Operations Research & CINBIO, University of Vigo, Campus Lagoas-Marcosende, 36310 Vigo, Spain
jacobo@uvigo.es

C. HEUCHENNE
HEC-Liege, University of Liege and Institute of Statistics, Biostatistics and Actuarial Sciences, Université catholique de Louvain, Rue Louvrex, 14, 4000 Liege, Belgium
C.Heuchenne@ulg.ac.be

G. LAURENT
HEC-Liege, University of Liege, Rue Louvrex, 14, 4000 Liege, Belgium
G.Laurent@ulg.ac.be

1. ADDITIONAL ANALYSIS OF MISSPECIFIED TRUNCATION DISTRIBUTIONS

In this subsection, we analyze cases where a wrong well-known family of parametric distributions is chosen to fit the truncation distribution. In Table 1, the truncation distribution is either a log-normal one or a gamma one but it is supposed to be a Weibull one where its parameters are unknown and estimated with the two proposed methods. The gamma density is $f(t) = (\lambda^\alpha/\Gamma(\alpha))t^{\alpha-1}\exp(-(\lambda t))$ and the log-normal one is $f(t) = \exp(-(\ln(x) - \mu)/(2\sigma^2))/(x\sigma\sqrt{2\pi})$.

In Table 1, we observe results similar to Section 4. Once again, the full likelihood estimator $F_{T,\tilde{\theta}}$ gets also worse more than $F_{T,\hat{\theta}}$ but leads to better results for $F(\cdot)$. In the proposed simulations, the impact of choosing the wrong truncation distribution seems very weak (see also Section 5 where these distributions are chosen); most of the time, the semiparametric estimators of F_{θ_0} outperform the product limit estimators.

2. ADDITIONAL ANALYSIS OF CONFIDENCE INTERVALS

In Table 2, basic and percentile bootstrap methods have been tested for (symmetric) two-sided confidence intervals with confidence level of 95%.

In Fig. 1 below, quantile-quantile plots are constructed to better assess the bootstrap approximation for the proposed confidence intervals for θ. The x-axis corresponds to the quantiles of the empirical distribution constructed with the estimations (minus the true value of the considered parameter) obtained from the $R = 1000$ samples while the y-axis corresponds to these quantiles but where the empirical distribution is obtained from the estimations in the bootstrap samples taken from one particular sample (minus the estimated parameter in this sample). For the percentile bootstrap case, the opposite of these quantiles are taken due to the symmetry assumption of this method. The graphs hereunder show that both distributions are close to each

Table 1. *Misspecification with a fitted Weibull truncation distribution*

γ	Cens. %	IMSE($F_{T,\hat\theta}$) RIMSE$_{T,\hat\theta,\tilde\theta}$	IMSE($F_{T,\tilde\theta}$)	IMSE($\widehat{F}_{\hat\theta}$) RIMSE$_{Y,\hat\theta,n}$	IMSE($\widehat{F}_{\tilde\theta}$) RIMSE$_{Y,\tilde\theta,n}$	IMSE(F_n)
		$n=100$ - T ~Log-normal$(0.50;0.43)$ - Trunc. %$=71.35$				
1	26.67	14.4	15.2	11.3	10.2	10.6
		0.95		1.07	0.96	
3	49.50	15.7	19.6	12.1	9.9	11.1
		0.80		1.09	0.89	
5	61.03	16.1	21.6	12.6	10.2	11.9
		0.75		1.06	0.86	
7	68.17	18.2	25.9	13.6	10.4	12.1
		0.70		1.12	0.86	
9	73.07	19.6	31.4	14.5	10.6	12.4
		0.62		1.17	0.85	
		$n=100$ - T ~Gamma$(6.47;2.34)$ - Trunc. %$=71.85$				
1	25.62	29.7	29.4	15.8	15.2	15.7
		1.01		1.01	0.97	
3	48.05	35.0	38.5	16.2	15.5	16.2
		0.91		1.00	0.96	
5	59.58	39.4	46.8	17.8	16.2	17.4
		0.84		1.02	0.93	
7	66.81	42.2	53.7	17.6	16.5	17.9
		0.79		0.99	0.92	
9	71.80	45.4	60.5	18.9	17.3	18.7
		0.75		1.01	0.92	
		$n=100$ - T ~Log-normal$(0.09;0.43)$ - Trunc. %$=54.96$				
1	27.60	6.8	6.8	5.5	5.1	6.6
		0.99		0.83	0.78	
3	50.75	7.0	7.9	6.6	5.3	7.1
		0.88		0.94	0.74	
5	62.16	7.1	8.7	7.6	5.5	7.1
		0.82		1.08	0.77	
7	69.17	7.9	9.4	8.2	5.7	7.3
		0.84		1.12	0.79	
9	73.91	8.4	11.6	9.8	6.2	7.9
		0.73		1.24	0.79	
		$n=100$ - T ~Gamma$(6.47;5.55)$ - Trunc. %$=54.45$				
1	27.56	4.4	4.5	5.6	4.9	6.8
		0.97		0.83	0.73	
3	50.64	4.5	4.9	5.2	4.8	7.3
		0.92		0.71	0.65	
5	62.07	4.6	5.6	7.1	5.5	7.5
		0.82		0.96	0.74	
7	69.06	5.0	6.0	7.8	6.0	7.8
		0.83		1.01	0.78	
9	73.79	5.3	6.9	8.1	6.3	8.0
		0.77		1.02	0.79	

Distributions, Y ~Weibull$(0.75;1.25)$ and $C-T \sim 5\times$Beta$(0.75;\gamma)$; Trunc. % and Cens. %, the truncation and censoring percentages; IMSE, the estimated integrated mean squared error ($\times10^{-3}$ for $F_{T,\hat\theta}$ and $F_{T,\tilde\theta}$, and $\times10^{-2}$ for $\widehat{F}_{\hat\theta}$, $\widehat{F}_{\tilde\theta}$ and F_n); $F_{T,\hat\theta}$ and $F_{T,\tilde\theta}$, the truncation distributions based on the conditional and full maximum likelihood estimators; $\widehat{F}_{\hat\theta}$ and $\widehat{F}_{\tilde\theta}$, the misspecified semiparametric estimators of F based on the conditional and full maximum likelihood estimators; F_n, the product-limit estimator of F; RIMSE$_{T,\hat\theta,\tilde\theta}$, the ratio of the estimated integrated mean squared errors of $F_{T,\hat\theta}$ and $F_{T,\tilde\theta}$; RIMSE$_{Y,\widehat{\theta},n}$ (respectively RIMSE$_{Y,\tilde\theta,n}$) the ratio of the estimated integrated mean squared errors of $\widehat{F}_{\hat\theta}$ (respectively $\widehat{F}_{\tilde\theta}$) and F_n. The standard errors for the integrated squared errors are bounded by 1.8×10^{-3} for the truncation distributions and 8.0×10^{-3} for the estimators of F.

Table 2. *Confidence intervals for the truncation parameters*

n	γ	Boot. Method	Av. Length λ_T Cond.	Av. Length λ_T Full	Cov. λ_T Cond.	Cov. λ_T Full	Av. Length α_T Cond.	Av. Length α_T Full	Cov. α_T Cond.	Cov. α_T Full
50	3	Basic	0.35	0.34	0.96	0.96	2.07	1.86	0.94	0.95
		Perc.	0.35	0.34	0.99	0.98	2.07	1.86	0.97	0.97
	5	Basic	0.32	0.31	0.92	0.92	2.12	1.87	0.90	0.93
		Perc.	0.32	0.31	0.98	0.96	2.12	1.87	0.97	0.95
	7	Basic	0.41	0.41	0.90	0.91	2.28	1.97	0.87	0.90
		Perc.	0.41	0.41	0.99	0.97	2.28	1.97	0.97	0.95
	9	Basic	0.41	0.40	0.88	0.89	2.35	1.99	0.84	0.90
		Perc.	0.41	0.40	1.00	0.96	2.35	1.99	0.96	0.95
100	3	Basic	0.25	0.23	0.95	0.96	1.43	1.30	0.98	0.97
		Perc.	0.25	0.23	0.99	0.99	1.43	1.29	0.97	0.98
	5	Basic	0.25	0.23	0.94	0.93	1.51	1.32	0.98	0.97
		Perc.	0.25	0.23	0.98	0.97	1.51	1.32	0.98	0.97
	7	Basic	0.28	0.25	0.92	0.94	1.62	1.40	0.97	0.97
		Perc.	0.28	0.25	0.99	0.97	1.62	1.40	0.99	0.97
	9	Basic	0.30	0.26	0.91	0.94	1.73	1.46	0.95	0.95
		Perc.	0.30	0.26	1.00	0.95	1.73	1.46	0.97	0.96
200	3	Basic	0.11	0.11	0.94	0.93	0.88	0.83	0.99	0.97
		Perc.	0.11	0.11	0.99	0.99	0.88	0.83	0.98	0.98
	5	Basic	0.18	0.16	0.94	0.90	1.04	0.92	0.99	0.98
		Perc.	0.18	0.16	1.00	0.98	1.04	0.92	0.98	0.98
	7	Basic	0.23	0.20	0.94	0.90	1.15	0.98	0.98	0.99
		Perc.	0.23	0.20	0.99	0.99	1.15	0.98	0.99	0.98
	9	Basic	0.26	0.22	0.93	0.90	1.24	1.03	0.97	0.97
		Perc.	0.26	0.22	1.00	0.98	1.24	1.03	0.98	0.97

Distributions, Y ~Weibull$(0.75; 1.25)$, T ~Weibull$(0.5; 3)$ and $C - T \sim 5\times$Beta$(0.75; \gamma)$; Boot. Method, the bootstrap method; Av. Length, the average length of the 1000 obtained confidence intervals; Cov., the percentage of confidence intervals covering the true values of the parameters (coverage); λ_T (respectively α_T), columns related to confidence intervals for λ_T (respectively α_T); Cond. (respectively Full), columns related to confidence intervals constructed with the conditional (respectively full) maximum likelihood method; Basic, the basic bootstrap method; Perc., the percentile bootstrap method. The standard error is bounded by 1.2×10^{-2} for the coverages and 6×10^{-3} (respectively 1.1×10^{-2}) for the average lengths of the confidence intervals for λ_T (respectively α_T).

other suggesting that both bootstrap procedures perform reasonably well. In addition, we also counted how many times the true value of the parameters was observed above the upper bound of the confidence interval and under its lower bound. If for numerous simulations, we observed approximately equal percentages on both sides, it is not always the case: percentages may be unbalanced according to the parameter, the basic or percentile bootstrap or the likelihood technique.

3. Proofs

To develop the proofs of the results displayed or mentioned in Section 3, we need to introduce some notations. We use capital letters to denote cumulative distribution functions and lower case letters to denote probability density functions. For $F_{T,\theta}(\cdot) = \mathrm{pr}_\theta(T \leq \cdot)$ a family of distributions indexed by $\theta \in \Theta$, a compact subset of $\mathbb{R}^d$, $d \geq 1$, including θ_0 ($F_T(\cdot) = F_{T,\theta_0}(\cdot)$),

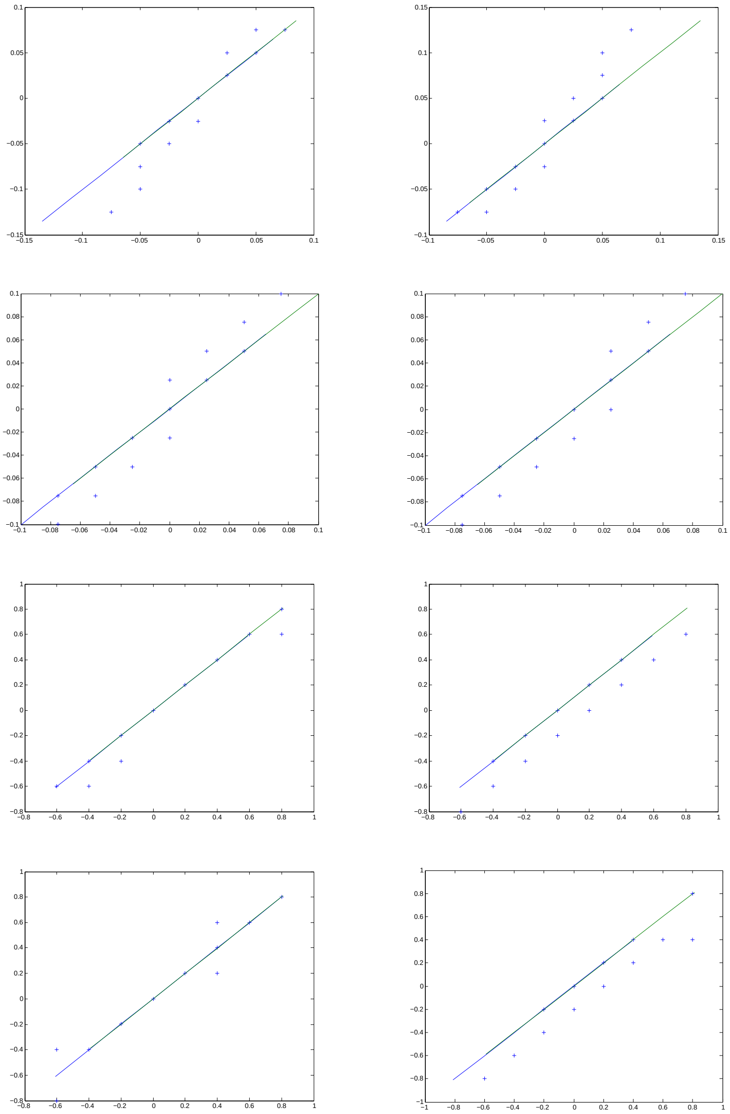

Fig. 1. Quantile-quantile plot for the bootstrap distribution of $\hat{\lambda}_T - \lambda_T$ (the 4 upper graphs) and $\hat{\alpha} - \alpha_T$ (the 4 lower graphs); $Y \sim$Weibull$(0.75; 1.25)$, $T \sim$Weibull$(0.5; 3)$, $C - T \sim 5\times$Beta$(0.75; \gamma)$; x-axis, distribution based on the simulated samples; y-axis, bootstrap distribution for one particular sample; left column, basic bootstrap; right column, percentile bootstrap; first and third rows, conditional likelihood; second and fourth rows, full likelihood method.

$\theta_{0,c}$ and $\theta_{0,p}$, we write $\dot f_{T,\theta}(t) = \partial f_{T,\theta}(t)/\partial\theta$, $f'_{T,\theta}(t) = \partial f_{T,\theta}(t)/\partial t$ and similarly for higher order derivatives. For the Hessian matrix of any function $m(\theta, x)$ depending on θ and other variables assumed to be two times differentiable with respect to the components of θ, we write $\partial^2 m(\theta,x)/\partial\theta^T\partial\theta = \ddot m(\theta, x)$ and $\ddot m(\theta_0, x)$ denotes its value at the point $\theta = \theta_0$. In addition,

$$\xi^f_{3,\theta}(y,z,\delta,t) = \left(\int_0^{+\infty} w^{-1}_{\theta,\mathcal{G}}(x)d\mathcal{H}^1(x)\right)^{-1} \xi^f_{2,\theta}(y,z,\delta,t) - \frac{\int_0^t f(x)w^{-1}_{\theta,\mathcal{G}}(x)d\mathcal{H}^1(x)}{(\int_0^{+\infty} w^{-1}_{\theta,\mathcal{G}}(x)d\mathcal{H}^1(x))^2}\xi_{2,\theta}(y,z,\delta,+\infty),$$

$$\Omega_{1,\theta} = -E\left\{\frac{\ddot f_{T,\theta}(T)}{f_{T,\theta}(T)} - \frac{\dot f_{T,\theta}(T)}{f_{T,\theta}(T)}\frac{\dot f^T_{T,\theta}(T)}{f_{T,\theta}(T)} - \Delta\left(\frac{\ddot w_{\theta,\mathcal{G}}(X)}{w_{\theta,\mathcal{G}}(X)} - \frac{\dot w_{\theta,\mathcal{G}}(X)}{w_{\theta,\mathcal{G}}(X)}\frac{\dot w^T_{\theta,\mathcal{G}}(X)}{w_{\theta,\mathcal{G}}(X)}\right) -(1-\Delta)\left(\frac{\int_0^{X\wedge\mathcal{T}} \ddot f_{T,\theta}(X-t)d\mathcal{G}(t)}{\int_0^{X\wedge\mathcal{T}} f_{T,\theta}(X-t)d\mathcal{G}(t)} - \frac{\left(\int_0^{X\wedge\mathcal{T}} \dot f_{T,\theta}(X-t)d\mathcal{G}(t)\right)\left(\int_0^{X\wedge\mathcal{T}} \dot f^T_{T,\theta}(X-t)d\mathcal{G}(t)\right)}{\left(\int_0^{X\wedge\mathcal{T}} f_{T,\theta}(X-t)d\mathcal{G}(t)\right)^2}\right) \mid X \geq T\right\}.$$

To avoid too long expressions, we also denote the vector of first (respectively second) derivatives of $w^{-1}_{\theta,\mathcal{G}}(\cdot)$ by $\dot w^{-1}_{\theta,\mathcal{G}}(\cdot)$ (respectively $\ddot w^{-1}_{\theta,\mathcal{G}}(\cdot)$). We then finally define

$$\begin{aligned} &DL_\theta(T) \\ &= \frac{\ddot f_{T,\theta}(T)}{f_{T,\theta}(T)} - \frac{\dot f_{T,\theta}(T)}{f_{T,\theta}(T)}\frac{\dot f^T_{T,\theta}(T)}{f_{T,\theta}(T)} \\ &- \frac{\int_0^\infty \left\{\ddot F_{T,\theta}(t)w^{-1}_{\theta,\mathcal{G}}(t) + \dot F_{T,\theta}(t)\dot w^{-1^T}_{\theta,\mathcal{G}}(t) + \dot w^{-1}_{\theta,\mathcal{G}}(t)\dot F^T_{T,\theta}(t) + \ddot w^{-1}_{\theta,\mathcal{G}}(t)F_{T,\theta}(t)\right\}d\mathcal{H}^1(t)}{\int_0^\infty F_{T,\theta}(t)w^{-1}_{\theta,\mathcal{G}}(t)d\mathcal{H}^1(t)} \\ &+ \frac{\int_0^\infty \left\{\dot F_{T,\theta}(t)w^{-1}_{\theta,\mathcal{G}}(t) + F_{T,\theta}(t)\dot w^{-1}_{\theta,\mathcal{G}}(t)\right\}}{\left(\int_0^\infty F_{T,\theta}(t)w^{-1}_{\theta,\mathcal{G}}(t)d\mathcal{H}^1(t)\right)^2} \\ &\qquad \times \int_0^\infty \left\{\dot F^T_{T,\theta}(t)w^{-1}_{\theta,\mathcal{G}}(t) + F_{T,\theta}(t)\dot w^{-1^T}_{\theta,\mathcal{G}}(t)\right\}d\mathcal{H}^1(t) \end{aligned}$$

and

$$\Omega_{2,\theta} = -E\left[\Delta\left\{DL_\theta(T) - \frac{\ddot w_{\theta,\mathcal{G}}(X)}{w_{\theta,\mathcal{G}}(X)} + \frac{\dot w_{\theta,\mathcal{G}}(X)}{w_{\theta,\mathcal{G}}(X)}\frac{\dot w^T_{\theta,\mathcal{G}}(X)}{w_{\theta,\mathcal{G}}(X)}\right\} +(1-\Delta)I(X<\tilde{\mathcal{T}})\left\{DL_\theta(T) + \frac{\int_X^\infty \ddot w^{-1}_{\theta,\mathcal{G}}(t)d\mathcal{H}^1(t)}{\int_X^\infty w^{-1}_{\theta,\mathcal{G}}(t)d\mathcal{H}^1(t)} - \frac{\int_X^\infty \dot w^{-1}_{\theta,\mathcal{G}}(t)d\mathcal{H}^1(t)\int_X^\infty \dot w^{-1^T}_{\theta,\mathcal{G}}(t)d\mathcal{H}^1(t)}{\left(\int_X^\infty w^{-1}_{\theta,\mathcal{G}}(t)d\mathcal{H}^1(t)\right)^2}\right\} \mid X \geq T\right].$$

LEMMA 1. *Under assumption* $1-4$ *, for a given function* $h(X)$ *such that* $E[h^3(X) \mid X \geq T] < +\infty$,

$$\int_0^t h(z)(w_{\theta_0,\mathcal{G}}(z) - w_{\theta_0,\mathcal{G}_n}(z))d\mathcal{H}_n^1(z)$$
$$= \frac{1}{n}\sum_{i=1}^n \int_0^t h(z)\int_0^z \xi(X_i - T_i, \Delta_i, (z-s)\wedge \mathcal{T})dF_{T,\theta_0}(s)d\mathcal{H}^1(z) + r_{1,n}(t),$$

where $\sup_t |r_{1,n}(t)| = o_P(n^{-1/2})$. *If* $E\left[\frac{\dot{f}^2_{T,\theta_0}(T)}{f^2_{T,\theta_0}(T)}\right] < +\infty$,

$$\int_0^t h(z)(\dot{w}_{\theta_0,\mathcal{G}}(z) - \dot{w}_{\theta_0,\mathcal{G}_n}(z))d\mathcal{H}_n^1(z)$$
$$= \frac{1}{n}\sum_{i=1}^n \int_0^t h(z)\int_0^z \xi(X_i - T_i, \Delta_i, (z-s)\wedge \mathcal{T})\dot{f}_{T,\theta_0}(s)dsd\mathcal{H}^1(z) + r_{2,n}(t),$$

where $\sup_t |r_{2,n}(t)| = o_P(n^{-1/2})$.

Proof. Define

$$I(T \leq Y)g(Y-T)f(T),$$

where $f(\cdot)$ and $g(\cdot) : \mathbb{R}^+ \to [0,1]$ are respectively a known and a monotone function. By Theorem 2.7.5 in van der Vaart and Wellner (1996), the bracketing number of the class $\mathcal{J}_m$ of $g(\cdot)$ functions is $m^J = N_{[]}(\varepsilon, \mathcal{J}_m, L_6(\mathcal{H}_{Y-T})) = \exp(K\varepsilon^{-1})$, for some constant $K > 0$. Now define the class

$$\mathcal{W} = \Big\{ r \to \int_0^{+\infty} I(0 \leq r-s)g(r-s)f(s)dF_{T,\theta_0}(s), r \in \mathbb{R}^+, t \to g(t) : \mathbb{R}^+ \to [0,1] \text{ is a monotone function and } s \to f(s) \text{ is a known function with } E[f^2(T)] < +\infty \Big\} \quad (1)$$

and for $i = 1, \ldots, m^J$, $[g_i^\ell, g_i^u]$, the m^J brackets for the class $\mathcal{J}_m$. Next, for each i, $i = 1, \ldots, m^J$, define

$$w_i^u(Y) = \int_0^Y \max(g_i^\ell(Y-t)f(t), g_i^u(Y-t)f(t))dF_{T,\theta_0}(t)$$

and similarly for $w_i^\ell(Y)$ with a minimum instead of a maximum. Each function of $\mathcal{W}$ for which $g_i^\ell(Y-T) \leq g(Y-T) \leq g_i^u(Y-T)$ is included in $[w_i^\ell(Y), w_i^u(Y)]$. Moreover, it is easily checked that $\int (w_i^u(z) - w_i^\ell(z))^6 d\mathcal{H}_Y(z) = O(\varepsilon^6)$ such that $m^J = N_{[]}(\varepsilon, \mathcal{W}, L_6(\mathcal{H}_Y))$ brackets suffice to cover $\mathcal{W}$ (where $\mathcal{H}_Y(x) = \text{pr}(Y \leq x \mid X \geq T)$). Next, for a given function $h(X)$ with $E[h^3(X) \mid X \geq T] < +\infty$, we prove that

$$\int_0^t h(z)(w_{\theta_0,\mathcal{G}}(z) - w_{\theta_0,\mathcal{G}_n}(z))d\mathcal{H}_n^1(z)$$
$$= \int_0^t h(z)(w_{\theta_0,\mathcal{G}}(z) - w_{\theta_0,\mathcal{G}_n}(z))d\mathcal{H}^1(z) + r_n(t), \quad (2)$$

where $\sup_t |r_n(t)| = o_P(n^{-1/2})$. First, we decompose the integral on the left hand side of the above expression into

$$\int_0^t h(z)I(h(z) \leq n^\delta)(w_{\theta_0,\mathcal{G}}(z) - w_{\theta_0,\mathcal{G}_n}(z))d\mathcal{H}_n^1(z)$$
$$+\int_0^t h(z)I(h(z) > n^\delta)(w_{\theta_0,\mathcal{G}}(z) - w_{\theta_0,\mathcal{G}_n}(z))d\mathcal{H}_n^1(z),$$

for a small $\delta > 0$. The second term is simply treated by the Markov inequality to show that it is $o_P(n^{-1/2})$. Note that this step involves $\text{pr}(h(X) > n^\delta \mid X \geq T)$, that is also treated with the Markov inequality. Next, with $f(s) = 1$, $w_{\theta_0,\mathcal{G}_n}(X)$ is a sequence of random functions that take their values in the class $\mathcal{W}$ defined in (1). Since $I(X \leq t)$ is a monotone function and $h(X)$ is purely random with $E[h^3(X) \mid X \geq T] < +\infty$, Lemma 19.36 of van der Vaart (1998) can be used. Indeed, the class of functions defined by $x \to I(x \leq t)h(x)w(x)$ is Donsker. We use this lemma with a restricted class of this type of functions where $w(X) = O_P(n^{-1/2}(\log\log n)^{1/2})$ and $h(X) < n^\delta$. We therefore obtain

$$\int_0^t h(z)I(h(z) \leq n^\delta)(w_{\theta_0,\mathcal{G}}(z) - w_{\theta_0,\mathcal{G}_n}(z))d\mathcal{H}_n^1(z)$$
$$= \int_0^t h(z)I(h(z) \leq n^\delta)(w_{\theta_0,\mathcal{G}}(z) - w_{\theta_0,\mathcal{G}_n}(z))d\mathcal{H}^1(z) + \tilde{r}_n(t),$$

where $\sup_t |\tilde{r}_n(t)| = o_P(n^{-1/2})$ and the above result (2) follows. Finally, the asymptotic representation of the Kaplan–Meier estimator given by Lo and Singh (1986) provides the first result in the statement of the lemma.
Now, in the same way,

$$\int_0^t h(z)(\dot{w}_{\theta_0,\mathcal{G}}(z) - \dot{w}_{\theta_0,\mathcal{G}_n}(z))d\mathcal{H}_n^1(z) = \int_0^t h(z)(\dot{w}_{\theta_0,\mathcal{G}}(z) - \dot{w}_{\theta_0,\mathcal{G}_n}(z))d\mathcal{H}^1(z) + r_n^*(t),$$

where $\sup_t |r_n^*(t)| = o_P(n^{-1/2})$, is obtained following the lines of the above results with $f(s) = \dot{f}_{T,\theta_0}(s)/f_{T,\theta_0}(s)$. The asymptotic representation in the second statement of the above lemma is then deduced and this finishes the proof. □

LEMMA 2. *Under assumptions* $1-5$,

$$\sup_t |\widehat{F}_{\theta_0}(t) - F_{\theta_0}(t)| = O(n^{-1/2}(\log\log n)^{1/2}) \ \textit{a.s.}$$

Proof. This result immediately follows from the uniform consistency properties of the Kaplan–Meier estimator and the empirical distribution function (see for example Cai (1998), Theorem 1, Lemma 2 and the expression below (15) in this paper). □

LEMMA 3. *Under assumptions* $1-5$, $\widehat{F}_{\theta_0}(t) - F_{\theta_0}(t) = \frac{1}{n}\sum_{i=1}^n \xi_{3,\theta_0}(X_i, X_i - T_i, \Delta_i, t) + r_{3,n}(t)$, *where* $\sup_t |r_{3,n}(t)| = o_P(n^{-1/2})$ *a.s.*

Proof. First, rewrite the numerator of $\widehat{F}_{\theta_0}(t) - F_{\theta_0}(t)$ as

$$\int_0^t (w^{-1}_{\theta_0,\mathcal{G}_n}(x) - w^{-1}_{\theta_0,\mathcal{G}}(x))d\mathcal{H}_n^1(x) + \int_0^t w^{-1}_{\theta_0,\mathcal{G}}(x)d(\mathcal{H}_n^1(x) - \mathcal{H}^1(x))$$
$$= E_{1,n}(t) + E_{2,n}(t).$$

Using Lemma 1 with $h(X) = 1/w^2_{\theta_0,\mathcal{G}}(X)$, the term $E_{1,n}(t)$ can be rewritten

$$\int_0^t \frac{w_{\theta_0,\mathcal{G}}(x) - w_{\theta_0,\mathcal{G}_n}(x)}{w^2_{\theta_0,\mathcal{G}}(x)} d\mathcal{H}^1_n(x) + o_P(n^{-1/2})$$
$$= \frac{1}{n}\sum_{i=1}^n \int_0^t \frac{\int_0^x \xi(X_i - T_i, \Delta_i, (x-s)\wedge\mathcal{T}) dF_{T,\theta_0}(s)}{w^2_{\theta_0,\mathcal{G}}(x)} d\mathcal{H}^1(x) + o_P(n^{-1/2}), \tag{3}$$

where the last terms (on both sides of the above expression) are uniform in t. Next, using standard arguments about the uniform consistency of Kaplan–Meier estimators and empirical processes,

$$\begin{aligned}\widehat{F}_{\theta_0}(t) - F_{\theta_0}(t) = {} & \frac{E_{1,n}(t) + E_{2,n}(t)}{\int_0^{+\infty} w^{-1}_{\theta_0,\mathcal{G}}(x) d\mathcal{H}^1(x)} \\ & - \frac{\int_0^t w^{-1}_{\theta_0,\mathcal{G}}(x) d\mathcal{H}^1(x)}{\left(\int_0^{+\infty} w^{-1}_{\theta_0,\mathcal{G}}(x) d\mathcal{H}^1(x)\right)^2} \left(E_{1,n}(+\infty) + E_{2,n}(+\infty)\right) + o_P(n^{-1/2}),\end{aligned}$$

where the last term is uniform in t. Applying the development (3) to $E_{1,n}(+\infty)$ finishes the proof. □

LEMMA 4. *Assume that assumptions $1-5$ and 6(ii),(iii) are met, and* $E[|X| \mid X \geq T] < +\infty$. *Under assumptions 6(i), 8(i) and 10,*

$$\sup_{\theta\in\Theta} \left|\mathcal{L}_{c,1}(\theta, \mathcal{G}_n) - E\left[\mathcal{L}_{c,1}(\theta, \mathcal{G})\right]\right| \to 0 \text{ a.s.},$$

which induces $\widehat{\theta} - \theta_{0,c} = o_P(1)$, *and under assumption 8(ii) and* $\sup_{\theta,t} |\partial F_{T,\theta}(t)/\partial\theta_k| < +\infty$ *($k = 1, \ldots, d$),*

$$\sup_{\theta\in\Theta} \left|\mathcal{L}_p(\theta, \widehat{F}_\theta, \mathcal{G}_n, \mathcal{H}^1_n) - E\left[\mathcal{L}_p(\theta, F_\theta, \mathcal{G}, \mathcal{H}^1)\right]\right| \to 0 \text{ a.s.},$$

which induces $\tilde{\theta} - \theta_{0,p} = o_P(1)$.

Proof. To prove this lemma, we will use Theorem 5.7 in van der Vaart (1998, p. 45). First, using assumptions 4, 5, 6(i),(ii) and 10

$$\frac{1}{n}\sum_{i=1}^n \left(\log f_{T,\theta}(T_i) - \Delta_i \log w_{\theta,\mathcal{G}_n}(X_i) - (1-\Delta_i)\log \int_0^{X_i\wedge\mathcal{T}} f_{T,\theta}(X_i - t) d\mathcal{G}_n(t)\right)$$
$$= \frac{1}{n}\sum_{i=1}^n \left(\log f_{T,\theta}(T_i) - \Delta_i \log w_{\theta,\mathcal{G}}(X_i) - (1-\Delta_i)\log \int_0^{X_i\wedge\mathcal{T}} f_{T,\theta}(X_i - t) d\mathcal{G}(t)\right)$$
$$+ o(1) \text{ a.s.}, \tag{4}$$

where the last term is uniform in θ. Next, under assumption 6(iii) and since in addition $\sup_{\theta,t} |\partial f_{T,\theta}(t)/\partial\theta_k| < +\infty$ for all k $(k = 1, \ldots, d)$, $E[|X| \mid X \geq T] < +\infty$ and the above function between brackets is parametric, the bracketing number of the class $\mathcal{F}_1$ of these functions indexed by $\theta \in \Theta \subset \mathbb{R}^d$ is

$$N_{[]}(\varepsilon, \mathcal{F}_1, L_1(P)) = O(\varepsilon^{-d}),$$

where P stands here for the distribution of X, T, Δ given $X \geq T$. As a consequence, the class $\mathcal{F}_1$ is Glivenko–Cantelli, which ensures the uniform consistency in θ.

We first consider $\mathcal{L}_{c,1}(\theta; \mathcal{G})$.

For $\theta_1 \neq \theta_2$ and $\delta = 1$, we assume nonidentifiability, i.e., $\frac{f_{T,\theta_1}(t)I(t\leq y)}{w_{\theta_1,\mathcal{G}}(y)}$ and $\frac{f_{T,\theta_2}(t)I(t\leq y)}{w_{\theta_2,\mathcal{G}}(y)}$ are equal on all nonzero measure sets. Since $f_{T,\theta}(t)$ is an identifiable probability density function, continuous with respect to t, there are at least two small intervals denoted $\delta^1_{t,\theta_{1,2}}$ and $\delta^2_{t,\theta_{1,2}}$, of nonzero measure with upper bound smaller than $\tau_{F_T(\cdot)}$ and for which $\int_{\delta^1_{t,\theta_{1,2}}}(f_{T,\theta_1}(t) - f_{T,\theta_2}(t))dt > 0$ and $\int_{\delta^2_{t,\theta_{1,2}}}(f_{T,\theta_1}(t) - f_{T,\theta_2}(t))dt < 0$. Taking y^*, a point larger than the upper bounds of both $\delta^1_{t,\theta_{1,2}}$ and $\delta^2_{t,\theta_{1,2}}$, and smaller than $\tau_{F_T(\cdot)}$, we have $\int_{\delta^i_{t,\theta_{1,2}}} f_{T,\theta_1}(t)dt \int_{y^*}^{\tau_{F_T(\cdot)}} w^{-1}_{\theta_1,\mathcal{G}}(y)dy = \int_{\delta^i_{t,\theta_{1,2}}} f_{T,\theta_2}(t)dt \int_{y^*}^{\tau_{F_T(\cdot)}} w^{-1}_{\theta_2,\mathcal{G}}(y)dy$, $i = 1, 2$. This leads to a ratio $\int_{y^*}^{\tau_{F_T(\cdot)}} w^{-1}_{\theta_2,\mathcal{G}}(y)dy / \int_{y^*}^{\tau_{F_T(\cdot)}} w^{-1}_{\theta_1,\mathcal{G}}(y)dy$ equal to two different constants (one smaller and one larger than one), which is impossible.

By assumption 8(i) and Theorem 5.7 in van der Vaart (1998, p. 45), $\widehat{\theta} - \theta_{0,c} = o_P(1)$. The log-likelihood of the second proposed procedure is

$$
\begin{aligned}
&\mathcal{L}_p(\theta, \widehat{F}_\theta, \mathcal{G}_n, \mathcal{H}^1_n) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\Delta_i\left\{\log f_{T,\theta}(T_i) - \log\int_0^{+\infty} F_{T,\theta}(t)dF_\theta(t) + \log\frac{w^{-1}_{\theta,\mathcal{G}}(X_i)}{\int w^{-1}_{\theta,\mathcal{G}}(t)d\mathcal{H}^1(t)}\right\}\right. \\
&\left.+I(C_i \leq \tilde{\mathcal{T}})(1-\Delta_i)\left\{\log f_{T,\theta}(T_i) - \log\int_0^{+\infty} F_{T,\theta}(t)dF_\theta(t) + \log(1 - F_\theta(X_i))\right\}\right] \\
&+o(1) \text{ a.s.,}
\end{aligned} \tag{5}
$$

where $\tilde{\mathcal{T}} < \tau_{F(\cdot)}$ and the last term is uniform in θ. Indeed, by assumptions 4, 5, 6(ii), $E[|X| \mid X \geq T] < +\infty$ and $\sup_{\theta,t}|\partial f_{T,\theta}(t)/\partial\theta_k| < +\infty$, for all k,

$$\sup_{\theta,t}|\widehat{F}_\theta(t) - F_\theta(t)| \to 0, \ a.s. \tag{6}$$

This is obtained by the uniform consistency of the Kaplan–Meier estimator combined with the fact that the class

$$\mathcal{F}_2 = \left\{(\delta, x) \to \delta I(x \leq y)w^{-1}_\theta(x); \theta \in \Theta, y \in \mathbf{R}\right\}$$

is Glivenko–Cantelli. Next, using similar arguments and $\sup_{\theta,t}|\partial F_{T,\theta}(t)/\partial\theta_k| < +\infty$, for all k, enables to obtain the second and fifth terms on the right hand side of (5) while the sixth term is treated by (6) since it is easily checked that $\sup_\theta F_\theta(\tilde{\mathcal{T}}) < 1$. Finally, the full sum in (5) is treated similarly to the sum on the right hand side of (4) such that the uniform consistency in θ to

$$
\begin{aligned}
&E\left[\Delta\left\{\log f_{T,\theta}(T) - \log\int_0^{+\infty} F_{T,\theta}(t)dF_\theta(t) + \log\frac{w^{-1}_{\theta,\mathcal{G}}(X)}{\int w^{-1}_{\theta,\mathcal{G}}(t)d\mathcal{H}^1(t)}\right\}\right. \\
&\left.+I(C \leq \tilde{\mathcal{T}})(1-\Delta)\left\{\log f_{T,\theta}(T) - \log\int_0^{+\infty} F_{T,\theta}(t)dF_\theta(t) + \log(1 - F_\theta(X))\right\} \mid X \geq T\right]
\end{aligned}
$$

is obtained. By assumption 8(ii) and Theorem 5.7 in van der Vaart (1998, p. 45), $\tilde{\theta} - \theta_{0,p} = o_P(1)$ and this finishes the proof. □

To simpify notations, assume in the sequel $\theta_{0,c} = \theta_{0,p} = \theta_0$.

Proof of Theorem $1.a$. Straightforward calculations using a first order Taylor expansion around θ_0 of the derivative of $\mathcal{L}_{c,1}(\theta,\mathcal{G}_n)$ at $\theta=\widehat{\theta}$ and treating double sums as in Lemma 3 lead to

$$\begin{aligned}
\widehat{\theta}-\theta_0 =& -\Big(\ddot{\mathcal{L}}_{c,1}(\theta^*,\mathcal{G}_n)\Big)^{-1}\frac{1}{n}\sum_{i=1}^{n}\Bigg\{\frac{\dot{f}_{T,\theta_0}(T_i)}{f_{T,\theta_0}(T_i)}-\Delta_i\frac{\dot{w}_{\theta_0,\mathcal{G}_n}(X_i)}{w_{\theta_0,\mathcal{G}_n}(X_i)} \\
&\qquad -(1-\Delta_i)\frac{\int_0^{X_i\wedge\mathcal{T}}\dot{f}_{T,\theta_0}(X_i-t)d\mathcal{G}_n(t)}{\int_0^{X_i\wedge\mathcal{T}}f_{T,\theta_0}(X_i-t)d\mathcal{G}_n(t)}\Bigg\}+o_P(n^{-1/2}) \\
=& \Big(\ddot{\mathcal{L}}_{c,1}(\theta^*,\mathcal{G}_n)\Big)^{-1}\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\Bigg[\\
&\Delta_i\Bigg\{\frac{\dot{w}_{\theta_0,\mathcal{G}}(X_i)\int_0^{X_i}\xi(X_j-T_j,\Delta_j,(X_i-s)\wedge\mathcal{T})f_{T,\theta_0}(s)ds}{w^2_{\theta_0,\mathcal{G}}(X_i)} \\
&\qquad -\frac{\int_0^{X_i}\xi(X_j-T_j,\Delta_j,(X_i-s)\wedge\mathcal{T})\dot{f}_{T,\theta_0}(s)ds}{w_{\theta_0,\mathcal{G}}(X_i)}\Bigg\} \\
&+(1-\Delta_i)\Bigg\{\frac{\int_0^{X_i\wedge\mathcal{T}}\dot{f}_{T,\theta_0}(X_i-t)d\xi(X_j-T_j,\Delta_j,t)}{\int_0^{X_i\wedge\mathcal{T}}f_{T,\theta_0}(X_i-t)d\mathcal{G}(t)} \\
&\qquad -\frac{\int_0^{X_i\wedge\mathcal{T}}\dot{f}_{T,\theta_0}(X_i-t)dG(t)\int_0^{X_i\wedge\mathcal{T}}f_{T,\theta_0}(X_i-t)d\xi(X_j-T_j,\Delta_j,t)}{\Big(\int_0^{X_i\wedge\mathcal{T}}f_{T,\theta_0}(X_i-t)d\mathcal{G}(t)\Big)^2}\Bigg\}\Bigg] \\
&-\Big(\ddot{\mathcal{L}}_{c,1}(\theta^*,\mathcal{G}_n)\Big)^{-1}\frac{1}{n}\sum_{i=1}^{n}\Bigg\{\frac{\dot{f}_{T,\theta_0}(T_i)}{f_{T,\theta_0}(T_i)}-\Delta_i\frac{\dot{w}_{\theta_0,\mathcal{G}}(X_i)}{w_{\theta_0,\mathcal{G}}(X_i)} \\
&\qquad -(1-\Delta_i)\frac{\int_0^{X_i\wedge\mathcal{T}}\dot{f}_{T,\theta_0}(X_i-t)d\mathcal{G}(t)}{\int_0^{X_i\wedge\mathcal{T}}f_{T,\theta_0}(X_i-t)d\mathcal{G}(t)}\Bigg\}+o_P(n^{-1/2}),
\end{aligned}$$

where each element θ^*_{ij} of θ^* is between θ_{0j} and $\widehat{\theta}_j$ (θ_{0j} and $\widehat{\theta}_j$ denote the jth element of θ_0 and $\widehat{\theta}$ respectively; $i,j=1,\ldots,d$). Weak consistency of $-\ddot{\mathcal{L}}_{c,1}(\theta^*,\mathcal{G}_n)$ as an estimator of Ω_{1,θ_0} is obtained using uniform consistency of the Kaplan–Meier estimator, Lemma (4), assumption 6 (iv), (vi), (vii), assumption 7 and $\inf_{\rho_{\mathcal{G}(\cdot)}\le x\le\tau_{F(\cdot)}}\int_0^{x\wedge\mathcal{T}}f_{T,\theta_0}(x-t)d\mathcal{G}(t)>0$.

Defining for $i \neq j$, $V_i = (X_i, \Delta_i)$, $W_j = (X_j, T_j, \Delta_j)$,

$$
\begin{aligned}
B(V_i, W_j) = \Delta_i & \left\{ \frac{\dot{w}_{\theta_0,\mathcal{G}}(X_i) \int_0^{X_i} \xi(X_j - T_j, \Delta_j, (X_i - s) \wedge \mathcal{T}) f_{T,\theta_0}(s) ds}{w^2_{\theta_0,\mathcal{G}}(X_i)} \right. \\
& \qquad \left. - \frac{\int_0^{X_i} \xi(X_j - T_j, \Delta_j, (X_i - s) \wedge \mathcal{T}) \dot{f}_{T,\theta_0}(s) ds}{w_{\theta_0,\mathcal{G}}(X_i)} \right\} \\
+ (1 - \Delta_i) & \left\{ \frac{\int_0^{X_i \wedge \mathcal{T}} \dot{f}_{T,\theta_0}(X_i - t) d\xi(X_j - T_j, \Delta_j, t)}{\int_0^{X_i \wedge \mathcal{T}} f_{T,\theta_0}(X_i - t) d\mathcal{G}(t)} \right. \\
& \qquad \left. - \frac{\int_0^{X_i \wedge \mathcal{T}} \dot{f}_{T,\theta_0}(X_i - t) dG(t) \int_0^{X_i \wedge \mathcal{T}} f_{T,\theta_0}(X_i - t) d\xi(X_j - T_j, \Delta_j, t)}{\left(\int_0^{X_i \wedge \mathcal{T}} f_{T,\theta_0}(X_i - t) d\mathcal{G}(t) \right)^2} \right\}
\end{aligned}
$$

and $B^*(V_i, W_j) = B(V_i, W_j) - E(B(V_i, W_j) \mid W_j)$, we have

$$
E(B^*(V_i, W_j)) = E(B^*(V_i, W_j) \mid V_i) = E(B^*(V_i, W_j) \mid W_j) = 0
$$

and

$$
\begin{aligned}
\text{pr}\Big(\frac{1}{n(n-1)} \sum_{i \neq j} & B^*(V_i, W_j) > C n^{-(1/2+\delta)}\Big) \\
& \leq \frac{E^2(\sum_{i \neq j} B^*(V_i, W_j))}{(C^2 n^{-(1+2\delta)}) n^2 (n-1)^2} = O(n^{-1+2\delta}),
\end{aligned}
$$

for any given $C > 0$ and $0 < \delta < 1/2$. This latter expression ensures that

$$
\frac{1}{n(n-1)} \sum_{i \neq j} B^*(V_i, W_j) = o_P(n^{-1/2})
$$

and finally, using assumption 9(i), $\widehat{\theta} - \theta_0 = \Omega^{-1}_{1,\theta_0} \frac{1}{n} \sum_{i=1}^n \eta_{1,\theta_0}(T_i, X_i, \Delta_i) + o_P(n^{-1/2})$. □

Proof of Theorem $1.b$. In the same way as in Theorem 1.a.,

$$
\begin{aligned}
\tilde{\theta} - \theta_0 = & -\left(\ddot{\mathcal{L}}_p(\tilde{\theta}^*, \widehat{F}_{\tilde{\theta}^*}, \mathcal{G}_n, \mathcal{H}_n^1)\right)^{-1} \frac{1}{n} \sum_{i=1}^{n} \Bigg[\Delta_i \Bigg\{ \frac{\dot{f}_{T,\theta_0}(T_i)}{f_{T,\theta_0}(T_i)} \\
& + \frac{\int_0^{+\infty} \left(\dot{f}_{T,\theta_0}(t) \widehat{F}_{\theta_0}(t) + f_{T,\theta_0}(t) \dot{\widehat{F}}_{\theta_0}(t) \right) dt}{\int_0^{+\infty} F_{T,\theta_0}(t) d\widehat{F}_{\theta_0}(t)} \\
& - \frac{\dot{w}_{\theta_0,\mathcal{G}_n}(X_i) \int w_{\theta_0,\mathcal{G}_n}^{-1}(t) d\mathcal{H}_n^1(t) + w_{\theta_0,\mathcal{G}_n}(X_i) \int \dot{w}_{\theta_0,\mathcal{G}_n}^{-1}(t) d\mathcal{H}_n^1(t)}{w_{\theta_0,\mathcal{G}_n}(X_i) \int w_{\theta_0,\mathcal{G}_n}^{-1}(t) d\mathcal{H}_n^1(t)} \Bigg\} \\
& + (1-\Delta_i) I(X_i \leq \tilde{\mathcal{T}}) \Bigg\{ \frac{\dot{f}_{T,\theta_0}(T_i)}{f_{T,\theta_0}(T_i)} + \frac{\int_0^{+\infty} \left(\dot{f}_{T,\theta_0}(t) \widehat{F}_{\theta_0}(t) + f_{T,\theta_0}(t) \dot{\widehat{F}}_{\theta_0}(t) \right) dt}{\int_0^{+\infty} F_{T,\theta_0}(t) d\widehat{F}_{\theta_0}(t)} \\
& - \frac{\dot{\widehat{F}}_{\theta_0}(X_i)}{(1 - \widehat{F}_{\theta_0}(X_i))} \Bigg\} \Bigg] \\
= & \ \Omega_{2,\theta_0}^{-1} \frac{1}{n} \sum_{i=1}^{n} \Bigg[\Bigg\{ \Delta_i \left(\frac{\dot{f}_{T,\theta_0}(T_i)}{f_{T,\theta_0}(T_i)} - \frac{\dot{w}_{\theta_0,\mathcal{G}}(X_i)}{w_{\theta_0,\mathcal{G}}(X_i)} \right) \\
& + (1-\Delta_i) I(X_i \leq \tilde{\mathcal{T}}) \left(\frac{\dot{f}_{T,\theta_0}(T_i)}{f_{T,\theta_0}(T_i)} - \frac{\dot{F}_{\theta_0}(X_i)}{(1 - F_{\theta_0}(X_i))} \right) \Bigg\} \\
& + \int \Bigg(\frac{\int_0^z \xi(X_i - T_i, \Delta_i, (z-s) \wedge \mathcal{T}) \dot{f}_{T,\theta_0}(s) ds}{w_{\theta_0,\mathcal{G}}(z)} \\
& - \frac{\dot{w}_{\theta_0,\mathcal{G}}(z) \int_0^z \xi(X_i - T_i, \Delta_i, (z-s) \wedge \mathcal{T}) f_{T,\theta_0}(s) ds}{w_{\theta_0,\mathcal{G}}^2(z)} \Bigg) d\mathcal{H}^1(z)
\end{aligned}
$$

$$
\begin{aligned}
&-\int_0^{\tilde{\mathcal{T}}}\left(\frac{\dot{\xi}_{3,\theta_0}(X_i,X_i-T_i,\Delta_i,z)}{1-F_{\theta_0}(z)}+\frac{\dot{F}_{\theta_0}(z)\xi_{3,\theta_0}(X_i,X_i-T_i,\Delta_i,z)}{(1-F_{\theta_0}(z))^2}\right)d\mathcal{H}^0(z)\\
&+\left(\mathrm{pr}(\Delta=1)+\int_0^{\tilde{\mathcal{T}}}d\mathcal{H}^0(z)\right)\\
&\times\left\{\frac{\int_0^{+\infty}f_{T,\theta_0}(z)\dot{\xi}_{3,\theta_0}(X_i,X_i-T_i,\Delta_i,z)dz-\xi_{3,\theta_0}^{\dot{F}_{T,\theta_0}}(X_i,X_i-T_i,\Delta_i,+\infty)}{\int_0^{+\infty}F_{T,\theta_0}(z)dF_{\theta_0}(z)}\right.\\
&\qquad\left.-\frac{\xi_{3,\theta_0}^{F_{T,\theta_0}}(X_i,X_i-T_i,\Delta_i,+\infty)\int_0^{+\infty}\left(\dot{f}_{T,\theta_0}(z)F_{\theta_0}(z)+f_{T,\theta_0}(z)\dot{F}_{\theta_0}(z)\right)dz}{\left(\int_0^{+\infty}F_{T,\theta_0}(z)dF_{\theta_0}(z)\right)^2}\right\}\\
&-\mathrm{pr}(\Delta=1)\left(\frac{\dot{\xi}_{2,\theta_0}(X_i,X_i-T_i,\Delta_i,\infty)}{\int_0^{+\infty}w_{\theta_0,\mathcal{G}}^{-1}(z)d\mathcal{H}^1(z)}\right.\\
&\qquad\left.\left.-\frac{\int_0^{\infty}\dot{w}_{\theta_0,\mathcal{G}}^{-1}(z)d\mathcal{H}^1(z)}{(\int_0^{+\infty}w_{\theta_0,\mathcal{G}}^{-1}(z)d\mathcal{H}^1(z))^2}\xi_{2,\theta_0}(X_i,X_i-T_i,\Delta_i,+\infty)\right)\right]\\
&+\Omega_{2,\theta_0}^{-1}\left\{\frac{\int_0^{+\infty}\left(\dot{f}_{T,\theta_0}(z)F_{\theta_0}(z)+f_{T,\theta_0}(z)\dot{F}_{\theta_0}(z)\right)dz}{\int_0^{+\infty}F_{T,\theta_0}(z)dF_{\theta_0}(z)}\frac{1}{n}\sum_{i=1}^n(\Delta_i+(1-\Delta_i)I(X_i\le\tilde{\mathcal{T}}))\right.\\
&\left.-\frac{\int\dot{w}_{\theta_0,\mathcal{G}}^{-1}(z)d\mathcal{H}^1(z)}{\int w_{\theta_0,\mathcal{G}}^{-1}(z)d\mathcal{H}^1(z)}\frac{1}{n}\sum_{i=1}^n\Delta_i\right\}\\
&+o_P(n^{-1/2}),
\end{aligned}
$$

where each element $\tilde{\theta}_{ij}^*$ of $\tilde{\theta}^*$ is between θ_{0j} and $\tilde{\theta}_j$ (θ_{0j} and $\tilde{\theta}_j$ denote the jth element of θ_0 and $\tilde{\theta}$ respectively; $i,j=1,\dots,d$; $\dot{w}_{\theta_0,\mathcal{G}_n}^{-1}(z)=-\dot{w}_{\theta_0,\mathcal{G}_n}(z)/w_{\theta_0,\mathcal{G}_n}^2(z)$ and similarly for $\mathcal{G}$ instead of $\mathcal{G}_n$). Weak consistency of $-\partial^2\mathcal{L}_p(\tilde{\theta}^*,\widehat{F}_{\tilde{\theta}^*},\mathcal{G}_n,\mathcal{H}_n^1)/\partial\theta^T\partial\theta$ as an estimator of Ω_{2,θ_0} is obtained using uniform consistency of the Kaplan–Meier estimator, Lemma 4, assumption 6 (iv), (v), (vi) and assumption 7. Since $\sup_{\theta,t}|\dot{f}_{T,\theta}(t)|<+\infty$, assumption 7 together with Lemma 1 enable to adapt Lemma 3 to obtain

$$\dot{\widehat{F}}_{\theta_0}(t)-\dot{F}_{\theta_0}(t)=\frac{1}{n}\sum_{i=1}^n\dot{\xi}_{3,\theta_0}(X_i,X_i-T_i,\Delta_i,t)+o_P(n^{-1/2}),$$

while the asymptotic representation

$$\int_0^{+\infty}f(t)d(\widehat{F}_{\theta_0}(t)-F_{\theta_0}(t))=\frac{1}{n}\sum_{i=1}^n\xi_{3,\theta_0}^f(X_i,X_i-T_i,\Delta_i,+\infty)+o_P(n^{-1/2})$$

(also obtained in a similar way as Lemma 3) is applied to $f(t)=F_{T,\theta_0}(t)$ and $f(t)=\dot{F}_{T,\theta_0}(t)$ (since $\sup_t|\dot{F}_{T,\theta_0}(t)|<+\infty$). Alternatively, under the same assumptions, we can also obtain the asymptotic representation for $\tilde{\theta}-\theta_{0,p}$ defined by $\Omega_{2,\theta_{0,p}}^{-1}n^{-1}\sum_{i=1}^n\eta_{2,\theta_{0,p}}(T_i,X_i,\Delta_i)$, where the remainder term is of order $o_P(n^{-1/2})$ (or we can derive it from the above calculations using basic arguments). $\square$

Proof of Theorem 2. Since $\sup_{t,\theta} |\partial f_{T,\theta}(t)/\partial\theta_i| < +\infty$, $\sup_{t,\theta} |\partial^2 f_{T,\theta}(t)/\partial\theta_i\partial\theta_j| < +\infty$ $(i,j = 1,\ldots,d)$ and $E[|X| \mid X \geq T] < +\infty$,

$$\begin{aligned}\widehat{F}_{\theta_n}(t) - F_{\theta_0}(t) &= \widehat{F}_{\theta_n}(t) - \widehat{F}_{\theta_0}(t) + \widehat{F}_{\theta_0}(t) - F_{\theta_0}(t)\\ &= \frac{\partial F_{\theta_0}(t)}{\partial\theta^T}(\theta_n - \theta_0) + \frac{1}{n}\sum_{i=1}^{n} \xi_{3,\theta_0}(X_i, X_i - T_i, \Delta_i, t) + o_P(n^{-1/2}).\end{aligned}$$

We treat $\xi_{3,\theta_0}(X, X-T, \Delta, t)$ using Theorem 2.5.6 in van der Vaart and Wellner (1996), i.e., we show that

$$\int_0^{+\infty} \left(\log N_{[]}(\varepsilon, \mathcal{F}, L_2(P))\right)^{1/2} d\varepsilon < +\infty, \tag{7}$$

where $N_{[]}$ is the bracketing number, P the probability measure corresponding to the joint distribution of $(X, X-T, \Delta)$, given $X \geq T$, $L_2(P)$ is the L_2- norm, and

$$\mathcal{F} = \left\{\xi_{3,\theta_0}(X, X-T, \Delta, t), -\infty < t < +\infty\right\}.$$

The function $\xi_{3,\theta_0}(X, X-T, \Delta, t)$ can be decomposed into

$$\begin{aligned}\xi_{3,\theta_0}(X, X-T, \Delta, t) &= \left(\int_0^{+\infty} w_{\theta_0,\mathcal{G}}^{-1}(x)d\mathcal{H}^1(x)\right)^{-1} \xi_{2,\theta_0}(X, X-T, \Delta, t)\\ &\quad - \frac{\int_0^t w_{\theta_0,\mathcal{G}}^{-1}(x)d\mathcal{H}^1(x)}{(\int_0^{+\infty} w_{\theta_0,\mathcal{G}}^{-1}(x)d\mathcal{H}^1(x))^2}\xi_{2,\theta_0}(X, X-T, \Delta, +\infty)\\ &= B_1 * \xi_{2,\theta_0}(X, X-T, \Delta, t) + B_2(t) * \xi_{2,\theta_0}(X, X-T, \Delta, +\infty),\end{aligned}$$

where $B_1 < +\infty$, $B_2(t)$ is a uniformly bounded function of t and $\xi_{2,\theta_0}(X, X-T, \Delta, +\infty)$ is a uniformly bounded purely random function. To establish weak convergence of the process $\sqrt{n}(\widehat{F}_{\theta_n}(t) - F_{\theta_0}(t))$, we only need to study $\xi_{2,\theta_0}(X, X-T, \Delta, t)$. Indeed, $(\partial F_{\theta_0}(t)/\partial\theta^T)(\theta_n - \theta_0)$ is also a product of a vector of uniformly bounded functions (only depending on t) times a purely random vector for which each component has a bounded second moment. Let then rewrite

$$\begin{aligned}&\xi_{2,\theta_0}(X, X-T, \Delta, t)\\ =&\int_0^t \int_0^x \frac{\xi_{1a}(X-T, (x-s)\wedge\mathcal{T}) + I(X-T \leq (x-s)\wedge\mathcal{T})\xi_{1b}(X-T, \Delta)}{w_{\theta_0,\mathcal{G}}^2(x)} dF_T(s)d\mathcal{H}^1(x)\\ &+I(X \leq t)\delta w_{\theta_0,\mathcal{G}}^{-1}(X) - \int_0^t w_{\theta_0,\mathcal{G}}^{-1}(x)d\mathcal{H}^1(x),\end{aligned}$$

where $\xi_{1a}(z,u) = -(1-\mathcal{G}(u))\int_0^{z\wedge u}(1-\mathcal{H}(s))^{-2}d\mathcal{H}^0(s)$, $\xi_{1b}(z,\delta) = I(z \leq \mathcal{T}, \delta = 0)/(1 - \mathcal{H}(z))$. Since $z \to \xi_{1a}(z,u)$ is a decreasing function for all u, the class of bounded functions $z \to \int_0^t\int_0^x \xi_{1a}(z, (x-s)\wedge\mathcal{T})/w_{\theta_0,\mathcal{G}}^2(x)dF_T(s)d\mathcal{H}^1(x)$ is also decreasing in z. Therefore, its bracketing number is $m = O(\exp(K\varepsilon^{-1}))$ by Theorem 2.7.5 in van der Vaart and Wellner (1996). The same theorem is applied to the class of bounded functions $z \to \int_0^t\int_0^x I(z \leq (x-s)\wedge\mathcal{T})/w_{\theta_0,\mathcal{G}}^2(x)dF_T(s)d\mathcal{H}^1(x)$, whereas $\xi_{1b}(X-T,\Delta)$ is a purely random bounded factor. Finally, the two last terms of $\xi_{2,\theta_0}(X, X-T, \Delta, t)$ are treated similarly (a decreasing function times a bounded purely random factor and a bounded function only depending on t). This finishes the proof since each term (except a purely random one) is bounded. That means that the domain of integration of (7) can be restricted to a finite upper bound. □

REFERENCES

CAI, Z. (1998). Asymptotic Properties of Kaplan-Meier estimator for Censored Dependent Data. *Stat. Probabil. Lett.* **37**, 381–389.

LO, S. H. & SINGH, K. (1986). The product-limit estimator and the bootstrap: some asymptotic representations. *Probab. Theory Rel.* **71**, 455–465.

VAN DER VAART, A. W. (1998). *Asymptotic statistics*. Cambridge: Cambridge University Press.

VAN DER VAART, A. W. & WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes*. New York: Springer-Verlag.

[*Received. Revised*]