WHAT PRICE SEMIPARAMETRIC COX REGRESSION?

Supplementary Material

This is supplementary material to the paper Jullum and Hjort (2018) which below is referred to as the main paper. The material has three sections. Section 1 contains proofs of Lemma 1 and Theorem 1 in the main paper. Section 2 contains detailed derivations leading to the precise formulae for the variance and covariance quantities $K, \nu(t)$ and G associated with Lemma 1 in the main paper. Section 3 provides explicit variance and covariance formulae related to Theorem 1 when the focus parameter is the (conditional) life time quantile of the survival distribution. Numbering of equations, theorems and lemmas refer to the main paper unless they are prefixed with an S.

1. Proofs of results from the main paper

Proof of Lemma 1. Andersen et al. (1993, Theorem VII.2.2 and Theorem VII.2.3) ensure that the semiparametric part of (26) converges to $Z_{\rm cox}(\cdot)$ while Hjort (1992, Theorem 6.1) and an additional delta method argument ensure that the parametric part of (26) converges to $Z_{\rm pm}(\cdot)$. In order to show that there is joint convergence of these, we shall use bits and pieces of the proofs in these theorems to rewrite the left hand side of (26). By the central limit theorem, a tightness argument, the delta method and Slutsky's theorem, we then show asymptotic equivalence with the right hand side.

From the proof of Andersen et al. (1993, Theorem VII.2.2 and Theorem VII.2.3) it follows that we may write

$$\begin{aligned} \sqrt{n}\{\widehat{A}_{\rm cox}(\cdot) - A_{\rm true}(\cdot)\} &= \overline{W}_n(\cdot) - F(\cdot)J_{\rm cox}^{-1}\sqrt{n}\overline{U}_{n,{\rm cox}} + o_p(1), \\ \sqrt{n}\{\widehat{\beta}_{\rm cox} - \beta_{\rm true}\} &= J_{\rm cox}^{-1}\sqrt{n}\overline{U}_{n,{\rm cox}} + o_p(1), \end{aligned} \tag{S1}$$

where $o_p(1)$ denotes quantities converging (uniformly) in probability to zero, and

$$\begin{aligned} \sqrt{n}\overline{W}_n(t) &= \sqrt{n}\frac{1}{n}\sum_{i=1}^n \int_0^t \frac{{\rm d}M_i(s)}{n^{-1}R_n^{(0)}(s;\beta_{\rm true})}, \\ \sqrt{n}\overline{U}_{n,{\rm cox}} &= \sqrt{n}\frac{1}{n}\sum_{i=1}^n \int_0^\tau \{X_i - E_n(s;\beta_{\rm true})\}\,{\rm d}M_i(s). \end{aligned} \tag{S2}$$

Further, from the proof of Hjort (1992, Theorem 6.1 (with details extended from Theorem 2.1)) we learn that

$$\sqrt{n}\begin{pmatrix} \widehat{\theta} - \theta_0 \\ \widehat{\beta}_{\rm pm} - \beta_0 \end{pmatrix} = J^{-1}\overline{U}_n + o_p(1),$$

with $\sqrt{n}\overline{U}_n = n^{-1/2}\sum_{i=1}^n U_i(\gamma_0)$ as described in (12). A delta method motivated Taylor expansion of $A_{\rm pm}(t;\widehat{\theta})$ around θ_0 allows us to write

$$\sqrt{n}\begin{pmatrix} A_{\rm pm}(t;\widehat{\theta}) - A_{\rm pm}(t;\theta_0) \\ \widehat{\beta}_{\rm pm} - \beta_0 \end{pmatrix} = {\rm block}(A_{\rm pm}^{\rm d}(t;\theta_0)^{\rm t}, \mathcal{I}_q)J^{-1}\sqrt{n}\overline{U}_n + o_p(1). \tag{S3}$$

Thus, it remains to show that

$$\sqrt{n}\{\overline{W}_n(\cdot), \overline{U}_{n,{\rm cox}}^{\rm t}, \overline{U}_n^{\rm t}\}^{\rm t} \overset{d}{\to} \{W(\cdot), U_{\rm cox}^{\rm t}, U^{\rm t}\}^{\rm t}, \tag{S4}$$

and that this limit process gives the desired covariance function.

To show that (S4) holds, we shall first show that the integrands in (S2) may be replaced by their limits in probability, i.e. that $\sqrt{n}\overline{W}_n(t) = n^{-1/2}\sum_{i=1}^n W_i^{(0)}(t) + o_p(1)$ and $\sqrt{n}\overline{U}_{n,{\rm cox}} = n^{-1/2}\sum_{i=1}^n U_{i,{\rm cox}}^{(0)} + o_p(1)$, where

$$W_i^{(0)}(t) = \int_0^t \frac{\mathrm{d}M_i(s)}{r^{(0)}(s;\beta_{\rm true})},$$
$$U_{i,\rm cox}^{(0)} = \int_0^\tau \{X_i - E(s;\beta_{\rm true})\}\,\mathrm{d}M_i(s).$$

To prove that these hold, let $f(\cdot)$ be some real scalar function from the domain of

$$h_n(s) = \{n^{-1}R_n^{(k)}(s;\beta_{\rm true})\}_{k=0,1}$$

and its limit $h_0(s) = \{r^{(k)}(s;\beta_{\rm true})\}_{k=0,1}$, for $s \in [0,\tau]$, which also is continuous and bounded in a neighbourhood around $h_0(s)$, uniformly in s. Consider $n^{-1/2}\sum_{i=1}^n \int_0^\tau \{f(h_n(s)) - f(h_0(s))\}\,\mathrm{d}M_i(s)$. This is a martingale having variance equal to the expectation of

$$\frac{1}{n}\sum_{i=1}^n \int_0^\tau \{f(h_n(s)) - f(h_0(s))\}^2\,\mathrm{d}N_i(s) \le \sup_s\{f(h_n(s)) - f(h_0(s))\}^2 \frac{1}{n}\sum_{i=1}^n N_i(\tau),$$

which is further bounded by the expectation of $C_n = \sup_s\{f(h_n(s)) - f(h_0(s))\}^2$. Since f is continuous in $h_n(s)$, the continuous mapping theorem and the uniform convergence of h_n in condition (A) ensure that $C_n \to_p 0$. Since $f(h_0(s))$ is bounded uniformly in s, the C_n is also bounded, and $\mathrm{E}\,C_n \to 0$. Thus, $n^{-1/2}\sum_{i=1}^n \int_0^\tau f(h_n(s))\,\mathrm{d}M_i(s)$ has the same limit as $n^{-1/2}\sum_{i=1}^n \int_0^\tau f(h_0(s))\,\mathrm{d}M_i(s)$. Now, since both integrands in (S2) fulfil the conditions on the f-function above, we may rewrite $\sqrt{n}\overline{W}_n(t)$ and $\sqrt{n}\overline{U}_{n,\rm cox}$ as described.

With this formulation, the left hand side of (S4) is just $\sqrt{n}$ times the mean of i.i.d. processes plus terms converging to zero in probability. To prove (S4), we need to demonstrate both finite-dimensional convergence and tightness, cf. Billingsley (1968, Ch. 2). Finite-dimensional convergence follows from the multidimensional central limit theorem, and tightness is secured since we already know by Andersen et al. (1993, Theorem VII.2.3.) that $\sqrt{n}\overline{W}_n(\cdot) \to_d W(\cdot)$. Thus, there is joint convergence in (S4), and the covariance function of this limit process is

$$\Gamma(s,t) = \mathrm{Cov}\left(\begin{pmatrix} U_{\rm cox} \\ W(s) \\ U \end{pmatrix}, \begin{pmatrix} U_{\rm cox} \\ W(t) \\ U \end{pmatrix}^{\rm t}\right) = \mathrm{Cov}\left(\begin{pmatrix} U_{i,\rm cox}^{(0)} \\ W_i^0(s) \\ U_i(\gamma_0) \end{pmatrix}, \begin{pmatrix} U_{i,\rm cox}^{(0)} \\ W_i^0(t) \\ U_i(\gamma_0) \end{pmatrix}^{\rm t}\right)$$
$$= \begin{pmatrix} J_{\rm cox} & 0_{q\times 1} & G \\ 0_{1\times q} & \sigma^2(\min(s,t)) & \nu(s) \\ G^{\rm t} & \nu(t)^{\rm t} & K \end{pmatrix},$$

with quantities given in the lemma and formulae which are proven in the supplementary material. Combine now the non-o_p-terms on the right hand side of (S1) and (S3) to a vector function of

$$\sqrt{n}(\overline{W}_n(\cdot), \overline{U}_{n,\rm cox}^{\rm t}, \overline{U}_n^{\rm t})^{\rm t}.$$

The continuous mapping theorem applied to this function then gives the desired limit distribution, which by some algebraic efforts is seen to take the form of (26). Slutsky's theorem takes care of the o_p-terms and completes the proof. ■

Proof of Theorem 1. By Lemma 1, the limit results in (26) hold. Since T is Hadamard differentiable at $\{A_{\rm true}(\cdot), \beta_{\rm true}\}$ and $\{A(\cdot;\theta_0), \beta_0\}$, the functional which combines the two into a two-dimensional vector with their individual mappings, has the same property. Application of the functional delta method (van der Vaart, 2000, Theorem 20.8) then gives the convergence result. That this limit is Gaussian follows by van der Vaart and Wellner (1996, Lemma 3.9.8) since $Z_{\rm cox}$ and $Z_{\rm pm}$ are both Gaussian. ■

2. Details on explicit variance and covariance formulae

The goal of this section is to derive the precise formulae for $K, \nu(t)$ and G, which are given by

$$K = \mathrm{Var}(U) = \mathrm{Var}(U_i(\gamma_0)) = \mathrm{E}\{U_i(\gamma_0)U_i(\gamma_0)^{\rm t}\},$$
$$\nu(t) = \mathrm{Cov}(W(t), U^{\rm t}) = \mathrm{E}\{W(t)U^{\rm t}\} = \mathrm{E}\{W_i^{(0)}(t)U_i(\gamma_0)^{\rm t}\},$$
$$G = \mathrm{Cov}(U_{\rm cox}, U^{\rm t}) = \mathrm{E}(U_{\rm cox}U^{\rm t}) = \mathrm{E}(U_{i,\rm cox}U_i(\gamma_0)^{\rm t}),$$

where

$$U_i(\gamma_0) = \int_0^\tau \begin{pmatrix}\psi(s;\theta_0)\\ X_i\end{pmatrix} \{Y_i(s)q(s;\gamma_0|X_i)\,\mathrm{d}s + \mathrm{d}M_i(s)\}\,,$$
$$W_i^{(0)}(t) = \int_0^\tau \frac{\mathbf{1}_{\{s<t\}}\mathrm{d}M_i(s)}{r^{(0)}(s;\beta_{\rm true})},$$
$$U_{i,\rm cox}^{(0)} = \int_0^\tau \{X_i - E(s;\beta_{\rm true})\}\,\mathrm{d}M_i(s).$$

For working conditions and detailed expressions not repeated below, see Section 2 in the main paper. Recall that $M_i(t) = N_i(t) - \int_0^t Y_i(s)\alpha_{\rm true}(s|X_i)\,\mathrm{d}s$ is a martingale and that

$$q(s;\gamma_0|x) = \alpha_{\rm true}(s)\exp(x^{\rm t}\beta_{\rm true}) - \alpha_{\rm pm}(s;\theta_0)\exp(x^{\rm t}\beta_0).$$

With slight abuse of notation, omitting the dependence on γ_0, write $U_i(\gamma_0) = U_i^{(1)} + U_i^{(2)}$ where

$$U_i^{(1)} = \int_0^\tau \begin{pmatrix}\psi(s;\theta_0)\\ X_i\end{pmatrix} Y_i(s)q(s;\gamma_0|X_i)\,\mathrm{d}s \quad\text{and}\quad U_i^{(2)} = \int_0^\tau \begin{pmatrix}\psi(s;\theta_0)\\ X_i\end{pmatrix}\mathrm{d}M_i(s). \tag{S5}$$

Recall also the form of the individual risk function $R_{(i)}^{(0)}(s;\beta) = Y_i(s)\exp(X_i^{\rm t}\beta)$, and its first and second order β derivatives

$$R_{(i)}^{(1)}(s;\beta) = Y_i(s)\exp(X_i^{\rm t}\beta)X_i \quad\text{and}\quad R_{(i)}^{(2)}(s;\beta) = Y_i(s)\exp(X_i^{\rm t}\beta)X_iX_i^{\rm t},$$

which by our working conditions have expectations $\mathrm{E}\{R_{(i)}^{(k)}(s;\beta)\} = r^{(k)}(s;\beta)$. As in the main paper, we also write

$$g^{(k)}(s;\beta) = \alpha_{\rm true}(s)r^{(k)}(s;\beta_{\rm true}+\beta) - \alpha_{\rm pm}(s;\theta_0)r^{(k)}(s;\beta_0+\beta),\ \text{for } k=0,1,2,$$

for the expectation of $R_{(i)}^{(k)}(s;\beta)q(s;\gamma_0|X_i)$. The following lemma will be helpful when deriving the precise expressions for the variance and covariance terms.

Lemma S1. *Let $h(s,t;X_i)$ be a function of s and t which is stochastically dependent only on X_i. Under our working conditions, we then have that*

$$\mathrm{E}\{Y_i(s)Y_i(t)h(s,t;X_i)\} = \mathrm{E}\{R_{(i)}^{(0)}(\max(s,t);\beta_{\rm true})\exp(-X_i^{\rm t}\beta_{\rm true})h(s,t;X_i)\},$$
$$\mathrm{E}\{\mathrm{d}M_i(s)Y_i(t)h(s,t;X_i)\} = -\mathrm{E}\{R_{(i)}^{(0)}(t;\beta_{\rm true})h(s,t;X_i)\}\alpha_{\rm true}(s)\mathbf{1}_{\{s<t\}}\,\mathrm{d}s.$$

Proof. The former follows from the definition of Y_i by noting that

$$Y_i(s)Y_i(t) = \mathbf{1}_{\{T_i>s\}}\mathbf{1}_{\{T_i>t\}} = \mathbf{1}_{\{T_i>\max(s,t)\}} = Y_i(\max(s,t)).$$

In proving the latter, note first that

$$\mathrm{E}\{\mathrm{d}M_i(s)Y_i(t)h(s,t;X_i)\} = \mathrm{E}[\mathrm{E}\{\mathrm{d}M_i(s)Y_i(t)|X_i\}h(s,t;X_i)].$$

The inner expectation here may be written as the difference between $\mathrm{E}\{\mathrm{d}N_i(s)Y_i(t)|X_i\}$ and $\mathrm{E}\{Y_i(s)Y_i(t)|X_i\}\exp(X_i^{\rm t}\beta_{\rm true})\alpha_{\rm true}(s)\,\mathrm{d}s = \mathrm{E}\{R_{(i)}^{(0)}(\max(s,t);\beta_{\rm true})|X_i\}\alpha_{\rm true}(s)\,\mathrm{d}s$. We also have that

$$\begin{aligned}\mathrm{E}\{\mathrm{d}N_i(s)Y_i(t)|X_i\} &= \mathrm{E}\{\mathrm{d}\mathbf{1}_{\{T_i\le s,D_i=1\}}(s)\mathbf{1}_{\{T_i\ge t\}}|X_i\} = \mathrm{E}\{\mathrm{d}N_i(s)|X_i\}\mathbf{1}_{\{s\ge t\}}\\ &= \mathrm{E}\{R_{(i)}^{(0)}(s;\beta_{\rm true})|X_i\}\alpha_{\rm true}(s)\mathbf{1}_{\{s\ge t\}}\,\mathrm{d}s.\end{aligned}$$

Thus, $\mathrm{E}\{\mathrm{d}M_i(s)Y_i(t)h(s,t;X_i)\} = -\mathrm{E}\{R_{(i)}^{(0)}(t;\beta_{\rm true})h(s,t;X_i)\}\alpha_{\rm true}(s)\mathbf{1}_{\{s<t\}}\,\mathrm{d}s.$ ∎

Let us start out with K. Since $\mathrm{E}\{U_i^{(1)}\} = \mathrm{E}\{U_i^{(2)}\} = 0$, we may write

$$K = \mathrm{E}\{U_i^{(1)}(U_i^{(1)})^{\rm t}\} + \mathrm{E}\{U_i^{(2)}(U_i^{(2)})^{\rm t}\} + \mathrm{E}\{U_i^{(1)}(U_i^{(2)})^{\rm t}\} + \mathrm{E}\{U_i^{(2)}(U_i^{(1)})^{\rm t}\}. \tag{S6}$$

The second expectation in (S6) is the easiest. By exploiting martingale properties, we get

$$\begin{aligned}
\mathrm{E}\{U_i^{(2)}(U_i^{(2)})^{\mathrm{t}}\} &= \mathrm{E}\bigg\{\int_0^\tau \begin{pmatrix}\psi(s;\theta_0)\\ X_i\end{pmatrix}\begin{pmatrix}\psi(s;\theta_0)\\ X_i\end{pmatrix}^{\mathrm{t}} Y_i(s)\exp(X_i^{\mathrm{t}}\beta_{\mathrm{true}})\alpha_{\mathrm{true}}(s)\,\mathrm{d}s\bigg\}\\
&= \mathrm{E}\bigg\{\int_0^\tau \begin{pmatrix}\psi(s;\theta_0)\psi(s;\theta_0)^{\mathrm{t}}R_{(i)}^{(0)}(s;\beta_{\mathrm{true}}) & \psi(s;\theta_0)R_{(i)}^{(1)}(s;\beta_{\mathrm{true}})^{\mathrm{t}}\\ R_{(i)}^{(1)}(s;\beta_{\mathrm{true}})\psi(s;\theta_0)^{\mathrm{t}} & R_{(i)}^{(2)}(s;\beta_{\mathrm{true}})\end{pmatrix}\times\alpha_{\mathrm{true}}(s)\,\mathrm{d}s\bigg\}\\
&= \int_0^\tau \begin{pmatrix}\psi(s;\theta_0)\psi(s;\theta_0)^{\mathrm{t}}r^{(0)}(s;\beta_{\mathrm{true}}) & \psi(s;\theta_0)r^{(1)}(s;\beta_{\mathrm{true}})^{\mathrm{t}}\\ r^{(1)}(s;\beta_{\mathrm{true}})\psi(s;\theta_0)^{\mathrm{t}} & r^{(2)}(s;\beta_{\mathrm{true}})\end{pmatrix}\alpha_{\mathrm{true}}(s)\,\mathrm{d}s.
\end{aligned}$$

When dealing with the other terms it becomes notationally convenient to introduce

$$Q(s,t;X_i) = \begin{pmatrix}\psi(s;\theta_0)\\ X_i\end{pmatrix}\begin{pmatrix}\psi(t;\theta_0)\\ X_i\end{pmatrix}^{\mathrm{t}}.$$

By the above lemma, the third expectation in (S6) may be written as

$$\begin{aligned}
\mathrm{E}\{U_i^{(1)}(U_i^{(2)})^{\mathrm{t}}\} &= -\int_0^\tau\int_0^\tau \mathrm{E}\{Q(t,s;X_i)q(t;\gamma_0|X_i)R_{(i)}^{(0)}(t;\beta_{\mathrm{true}})\alpha_{\mathrm{true}}(s)\mathbf{1}_{\{s<t\}}\}\,\mathrm{d}s\,\mathrm{d}t\\
&= -\int_0^\tau\int_0^t \mathrm{E}\{Q(t,s;X_i)q(t;\gamma_0|X_i)R_{(i)}^{(0)}(t;\beta_{\mathrm{true}})\alpha_{\mathrm{true}}(s)\}\,\mathrm{d}s\,\mathrm{d}t.
\end{aligned}$$

Since the fourth expectation in (S6) is just the transpose of the third and $Q(t,s;X_i)^{\mathrm{t}} = Q(s,t;X_i)$, we also have

$$\mathrm{E}\{U_i^{(2)}(U_i^{(1)})^{\mathrm{t}}\} = -\int_0^\tau\int_0^t \mathrm{E}\{Q(s,t;X_i)q(t;\gamma_0|X_i)R_{(i)}^{(0)}(t;\beta_{\mathrm{true}})\alpha_{\mathrm{true}}(s)\}\,\mathrm{d}s\,\mathrm{d}t.$$

Now, the first expectation in (S6) may be written as

$$\begin{aligned}
\mathrm{E}\{U_i^{(1)}(U_i^{(1)})^{\mathrm{t}}\} &= \int_0^\tau\int_0^\tau \mathrm{E}\{Q(s,t;X_i)q(s;\gamma_0|X_i)q(t;\gamma_0|X_i)R_{(i)}^{(0)}(\max(s,t);\beta_{\mathrm{true}})\\
&\qquad\times\exp(-X_i^{\mathrm{t}}\beta_{\mathrm{true}})\}\,\mathrm{d}s\,\mathrm{d}t\\
&= \int_0^\tau\int_0^t \mathrm{E}\{Q(s,t;X_i)q(s;\gamma_0|X_i)q(t;\gamma_0|X_i)R_{(i)}^{(0)}(t;\beta_{\mathrm{true}})\\
&\qquad\times\exp(-X_i^{\mathrm{t}}\beta_{\mathrm{true}})\}\,\mathrm{d}s\,\mathrm{d}t\\
&\quad+\int_0^\tau\int_0^s \mathrm{E}\{Q(s,t;X_i)q(s;\gamma_0|X_i)q(t;\gamma_0|X_i)R_{(i)}^{(0)}(s;\beta_{\mathrm{true}})\\
&\qquad\times\exp(-X_i^{\mathrm{t}}\beta_{\mathrm{true}})\}\,\mathrm{d}t\,\mathrm{d}s\\
&= \int_0^\tau\int_0^t \mathrm{E}\bigg[\{Q(s,t;X_i)+Q(t,s;X_i)\}q(t;\gamma_0|X_i)R_{(i)}^{(0)}(t;\beta_{\mathrm{true}})\alpha_{\mathrm{true}}(s)\\
&\qquad\times\bigg(1-\frac{\alpha_{\mathrm{pm}}(s;\theta)\exp(X_i^{\mathrm{t}}\beta_0)}{\alpha_{\mathrm{true}}(s)\exp(X_i^{\mathrm{t}}\beta_{\mathrm{true}})}\bigg)\bigg]\,\mathrm{d}s\,\mathrm{d}t.
\end{aligned}$$

Thus, a part of the above expression cancels out all of the third and fourth expectation, leaving the sum of the first, third and fourth expectation to

$$K' = -\int_0^\tau\int_0^t \mathrm{E}\Big[\{Q(s,t;X_i)+Q(t,s;X_i)\}q(t;\gamma_0|X_i)R_{(i)}^{(0)}(t;\beta_0)\alpha_{\mathrm{pm}}(s;\theta_0)\Big]\,\mathrm{d}s\,\mathrm{d}t.$$

Using the $g^{(k)}(s;\beta)$ notation, K' is seen to have the following blocks:

$$\begin{aligned}
K'_{11} &= -\int_0^\tau \int_0^t \{\psi(s;\theta_0)\psi(t;\theta_0)^{\rm t} + \psi(t;\theta_0)\psi(s;\theta_0)^{\rm t}\} g^{(0)}(t;\beta_0)\alpha_{\rm pm}(s;\theta)\,{\rm d}s\,{\rm d}t \\
&= -\int_0^\tau \{A^{\rm d}_{\rm pm}(t;\theta_0)\psi(t;\theta_0)^{\rm t} + \psi(t;\theta_0)A^{\rm d}_{\rm pm}(t;\theta_0)^{\rm t}\} g^{(0)}(t;\beta_0)\,{\rm d}t, \\
K'_{12} = (K'_{21})^{\rm t} &= -\int_0^\tau \int_0^t \{\psi(s;\theta_0) + \psi(t;\theta_0)\} g^{(1)}(t;\beta_0)^{\rm t}\alpha_{\rm pm}(s;\theta_0)\,{\rm d}s\,{\rm d}t \\
&= -\int_0^\tau \{A^{\rm d}_{\rm pm}(t;\theta_0) + \psi(t;\theta_0)A_{\rm pm}(t;\theta_0)\} g^{(1)}(t;\beta_0)^{\rm t}\,{\rm d}t, \\
K'_{22} &= -2\int_0^\tau \int_0^t g^{(2)}(t;\beta_0)\alpha_{\rm pm}(s;\theta_0)\,{\rm d}s\,{\rm d}t = -2\int_0^\tau g^{(2)}(t;\beta_0)A_{\rm pm}(t;\theta_0)\,{\rm d}t.
\end{aligned}$$

Finally, the full expression for $K = {\rm E}\{U_i^{(2)}(U_i^{(2)})^{\rm t}\} + K'$ consists of the following blocks

$$\begin{aligned}
K_{11} = \int_0^\tau &[\psi(s;\theta_0)\psi(s;\theta_0)^{\rm t} r^{(0)}(s;\beta_{\rm true})\alpha_{\rm true}(s) \\
&- \{A^{\rm d}_{\rm pm}(s;\theta_0)\psi(s;\theta_0)^{\rm t} + \psi(s;\theta_0)A^{\rm d}_{\rm pm}(s;\theta_0)^{\rm t}\} g^{(0)}(s;\beta_0)]\,{\rm d}s, \\
K_{12} = K_{21}^{\rm t} = \int_0^\tau &[\psi(s;\theta_0) r^{(1)}(s;\beta_{\rm true})^{\rm t}\alpha_{\rm true}(s) \\
&- \{A^{\rm d}_{\rm pm}(s;\theta_0) + \psi(s;\theta_0)A_{\rm pm}(s;\theta_0)\} g^{(1)}(s;\beta_0)^{\rm t}]\,{\rm d}s, \\
K_{22} = \int_0^\tau &\{r^{(2)}(s;\beta_{\rm true})\alpha_{\rm true}(s) - 2g^{(2)}(s;\beta_0)A_{\rm pm}(s;\theta_0)\}\,{\rm d}s.
\end{aligned}$$

Let us then turn to $\nu(t)$. From the representation in (S5), we may write

$$\nu(t) = {\rm E}[W_i^{(0)}(t)\{U_i^{(1)}\}^{\rm t}] + {\rm E}[W_i^{(0)}(t)\{U_i^{(2)}\}^{\rm t}]. \tag{S7}$$

The second expectation here is once again easy. Exploiting martingale properties gives

$$\begin{aligned}
{\rm E}[W_i^{(0)}(t)\{U_i^{(2)}\}^{\rm t}] &= {\rm E}\left\{\int_0^\tau \frac{\mathbf{1}_{\{s<t\}}}{r^{(0)}(s;\beta_{\rm true})} \begin{pmatrix}\psi(s;\theta_0)\\ X_i\end{pmatrix}^{\rm t} Y_i(s)\exp(X_i^{\rm t}\beta_{\rm true})\alpha_{\rm true}(s)\,{\rm d}s\right\} \\
&= {\rm E}\left\{\int_0^t \begin{pmatrix}R^{(0)}_{(i)}(s;\beta_{\rm true})\psi(s;\theta_0)\\ R^{(1)}_{(i)}(s;\beta_{\rm true})\end{pmatrix}^{\rm t} \alpha_{\rm true}(s)/r^{(0)}(s;\beta_{\rm true}),\,{\rm d}s\right\} \\
&= \int_0^t \begin{pmatrix}\psi(s;\theta_0)\\ E(s;\beta_{\rm true})\end{pmatrix}^{\rm t} \alpha_{\rm true}(s)\,{\rm d}s = \begin{pmatrix}\int_0^t \psi(s;\theta_0)\alpha_{\rm true}(s)\,{\rm d}s\\ F(t)\end{pmatrix}^{\rm t}.
\end{aligned}$$

By the above lemma, the first expectation in (S7) may be written as

$$\begin{aligned}
{\rm E}[W_i^{(0)}(t)\{U_i^{(1)}\}^{\rm t}] &= {\rm E}\left\{\int_0^\tau \int_0^\tau \frac{\mathbf{1}_{\{s<t\}}{\rm d}M_i(s)}{r^{(0)}(s;\beta_{\rm true})} \begin{pmatrix}\psi(u;\theta_0)\\ X_i\end{pmatrix}^{\rm t} Y_i(u)q(u;\gamma_0\,|\,X_i)\,{\rm d}s\,{\rm d}u\right\} \\
&= -{\rm E}\left\{\int_0^\tau \int_0^{\min(t,u)} \begin{pmatrix}\psi(u;\theta_0)\\ X_i\end{pmatrix}^{\rm t} R^{(0)}_{(i)}(u;\beta_{\rm true})q(u;\gamma_0\,|\,X_i) \times \frac{\alpha_{\rm true}(s)}{r^{(0)}(s;\beta_{\rm true})}\,{\rm d}s\,{\rm d}u\right\} \\
&= -\int_0^\tau \int_0^{\min(t,u)} \begin{pmatrix}g^{(0)}(u;\beta_{\rm true})\psi(u;\theta_0)\\ g^{(1)}(u;\beta_{\rm true})\end{pmatrix}^{\rm t} \frac{\alpha_{\rm true}(s)}{r^{(0)}(s;\beta_{\rm true})}\,{\rm d}s\,{\rm d}u \\
&= -\int_0^\tau \begin{pmatrix}g^{(0)}(u;\beta_{\rm true})\psi(u;\theta_0)\\ g^{(1)}(u;\beta_{\rm true})\end{pmatrix}^{\rm t} \sigma^2(\min(t,u))\,{\rm d}u.
\end{aligned}$$

Thus, the final expression for $\nu(t)$ becomes

$$\nu(t) = \begin{pmatrix}\int_0^t \psi(s;\theta_0)\alpha_{\rm true}(s)\,{\rm d}s\\ F(t)\end{pmatrix}^{\rm t} - \int_0^\tau \begin{pmatrix}g^{(0)}(u;\beta_{\rm true})\psi(u;\theta_0)\\ g^{(1)}(u;\beta_{\rm true})\end{pmatrix}^{\rm t} \sigma^2(\min(t,u))\,{\rm d}u.$$

Let us finally turn to G. From the representation in (S5), we may write

$$G = {\rm E}\{U_{i,{\rm cox}}^{(0)}(U_i^{(1)})^{\rm t}\} + {\rm E}\{U_{i,{\rm cox}}^{(0)}(U_i^{(2)})^{\rm t}\}. \tag{S8}$$

The second expectation here is once again easy as martingale properties give

$$\begin{aligned}
\mathrm{E}\{U_{i,\mathrm{cox}}^{(0)}(U_i^{(2)})^{\mathrm{t}}\} &= \mathrm{E}\left\{\int_0^\tau \{X_i - E(s;\beta_{\mathrm{true}})\} \begin{pmatrix}\psi(s;\theta_0)\\ X_i\end{pmatrix}^{\mathrm{t}} Y_i(s)\exp(X_i^{\mathrm{t}}\beta_{\mathrm{true}})\alpha_{\mathrm{true}}(s)\,\mathrm{d}s\right\}\\
&= \mathrm{E}\left\{\int_0^\tau \begin{pmatrix}\psi(s;\theta_0)\{R_{(i)}^{(1)}(s;\beta_{\mathrm{true}})^{\mathrm{t}} - R_{(i)}^{(0)}(s;\beta_{\mathrm{true}})E(s;\beta_{\mathrm{true}})^{\mathrm{t}}\}\\ R_{(i)}^{(2)}(s;\beta_{\mathrm{true}}) - R_{(i)}^{(1)}(s;\beta_{\mathrm{true}})E(s;\beta_{\mathrm{true}})^{\mathrm{t}}\end{pmatrix}^{\mathrm{t}} \times \alpha_{\mathrm{true}}(s), \mathrm{d}s\right\}\\
&= \int_0^\tau \begin{pmatrix}0_{p\times q}\\ r^{(2)}(s;\beta_{\mathrm{true}})/r^{(0)}(s;\beta_{\mathrm{true}}) - E(s;\beta_{\mathrm{true}})E(s;\beta_{\mathrm{true}})^{\mathrm{t}}\end{pmatrix}^{\mathrm{t}}\\
&\qquad \times r^{(0)}(s;\beta_{\mathrm{true}})\alpha_{\mathrm{true}}(s)\,\mathrm{d}s\\
&= \begin{pmatrix}0_{p\times q}\\ J_{\mathrm{cox}}\end{pmatrix}^{\mathrm{t}}.
\end{aligned}$$

By the above lemma, the first expectation in (S8) may be written as

$$\begin{aligned}
\mathrm{E}\{U_{i,\mathrm{cox}}^{(0)}(U_i^{(1)})^{\mathrm{t}}\} &= \mathrm{E}\left\{\int_0^\tau\int_0^\tau \{X_i - E(t;\beta_{\mathrm{true}})\}\,\mathrm{d}M_i(t) \begin{pmatrix}\psi(s;\theta_0)\\ X_i\end{pmatrix}^{\mathrm{t}} Y_i(s)q(s;\gamma_0\,|\,X_i)\,\mathrm{d}s\right\}\\
&= -\mathrm{E}\left\{\int_0^\tau\int_0^s \{X_i - E(t;\beta_{\mathrm{true}})\} \begin{pmatrix}\psi(s;\theta_0)\\ X_i\end{pmatrix}^{\mathrm{t}} R_{(i)}^{(0)}(s;\beta_{\mathrm{true}})q(s;\gamma_0\,|\,X_i)\times\alpha_{\mathrm{true}}(t)\,\mathrm{d}t\,\mathrm{d}s\right\}\\
&= -\int_0^\tau\int_0^s \begin{pmatrix}\psi(s;\theta_0)\{g^{(1)}(s;\beta_{\mathrm{true}})^{\mathrm{t}} - g^{(0)}(s;\beta_{\mathrm{true}})E(t;\beta_{\mathrm{true}})^{\mathrm{t}}\}\\ g^{(2)}(s;\beta_{\mathrm{true}}) - g^{(1)}(s;\beta_{\mathrm{true}})E(t;\beta_{\mathrm{true}})^{\mathrm{t}}\end{pmatrix}^{\mathrm{t}} \times\alpha_{\mathrm{true}}(t)\,\mathrm{d}t\,\mathrm{d}s\\
&= -\int_0^\tau \begin{pmatrix}\psi(s;\theta_0)\{A_{\mathrm{true}}(s)g^{(1)}(s;\beta_{\mathrm{true}})^{\mathrm{t}} - g^{(0)}(s;\beta_{\mathrm{true}})F(s)^{\mathrm{t}}\}\\ A_{\mathrm{true}}(s)g^{(2)}(s;\beta_{\mathrm{true}}) - g^{(1)}(s;\beta_{\mathrm{true}})F(s)^{\mathrm{t}}\end{pmatrix}^{\mathrm{t}} \mathrm{d}s.
\end{aligned}$$

Thus, the final expression for G becomes

$$G = \begin{pmatrix}0_{p\times q}\\ J_{\mathrm{cox}}\end{pmatrix}^{\mathrm{t}} - \int_0^\tau \begin{pmatrix}\psi(s;\theta_0)\{A_{\mathrm{true}}(s)g^{(1)}(s;\beta_{\mathrm{true}})^{\mathrm{t}} - g^{(0)}(s;\beta_{\mathrm{true}})F(s)^{\mathrm{t}}\}\\ A_{\mathrm{true}}(s)g^{(2)}(s;\beta_{\mathrm{true}}) - g^{(1)}(s;\beta_{\mathrm{true}})F(s)^{\mathrm{t}}\end{pmatrix}^{\mathrm{t}} \mathrm{d}s.$$

Hence, we have derived the expressions for all the desired quantities.

3. The life time quantile as a focus parameter

The u-quantile of the life time for an individual with covariate values corresponding to x is given by

$$\mu = \phi_{u,x} = T(A(\cdot),\beta;t,x) = A^{-1}\left(\frac{-\log(1-u)}{\exp(x^{\mathrm{t}}\beta)}\right) = A^{-1}(-\log(1-u)\,|\,x), \tag{S9}$$

for some $u \in (0,1)$ where $A(\phi_{u,x})$ is continuous. This focus parameter has semiparametric and fully parametric estimators given by respectively $\widehat{\mu}_{\mathrm{cox}} = \widehat{A}_{\mathrm{cox}}^{-1}(-\log(1-u)\,|\,x)$ and $\widehat{\mu}_{\mathrm{pm}} = \widehat{A}_{\mathrm{pm}}^{-1}(-\log(1-u)\,|\,x)$, i.e. the inverse of the estimators defined in (21), evaluated at $-\log(1-u)$. These estimators are consistent for respectively $\mu_{\mathrm{true}} = A_{\mathrm{true}}^{-1}(-\log(1-u)\,|\,x)$ and $\mu_0 = A_0^{-1}(-\log(1-u)\,|\,x)$, being inverses of the quantities in (25). van der Vaart (2000, Lemma 21.3) states that if $\sqrt{n}\{H_n(\cdot) - H(\cdot)\} \to_d Z(\cdot)$, for some scalar function $H_n(\cdot)$ with non-decreasing limit function $H(\cdot)$, one also has $\sqrt{n}\{H_n^{-1}(u) - H^{-1}(u)\} \to_d -Z(H^{-1}(u))/h(H^{-1}(u))$, for any u in the range of $H(\cdot)$ – provided the derivative h of H exists and is positive at $H^{-1}(u)$. Thus, finding the equivalent of (30) for the quantile focus parameter in (S9) amounts 'simply' to inverting the equation in (33). Omitting the notational dependence on u and x, let $\phi_{\mathrm{cox}} = A_{\mathrm{true}}^{-1}(-\log(1-u)\,|\,x)$ and $\phi_{\mathrm{pm}} = A_0^{-1}(-\log(1-u)\,|\,x)$. Recall from the main paper that $\zeta_{\mathrm{cox}}(\cdot) = (\exp(x^{\mathrm{t}}\beta_{\mathrm{true}}), A_{\mathrm{true}}(\cdot)\exp(x^{\mathrm{t}}\beta_{\mathrm{true}})x^{\mathrm{t}})^{\mathrm{t}}$ and $\zeta_{\mathrm{pm}}(\cdot) = (\exp(x^{\mathrm{t}}\beta_0), A_{\mathrm{pm}}(\cdot;\theta_0)\exp(x^{\mathrm{t}}\beta_0)x^{\mathrm{t}})^{\mathrm{t}}$. Now, since $A_{\mathrm{true}}(s\,|\,x)$ and $A_0(s\,|\,x)$ have derivatives $h_{\mathrm{cox}}(s) = \alpha_{\mathrm{true}}(s)\exp(x^{\mathrm{t}}\beta_{\mathrm{true}})$ and $h_{\mathrm{pm}}(s) = \alpha_{\mathrm{pm}}(s;\theta_0)$ $\exp(x^{\mathrm{t}}\beta_0)$, we have

$$\sqrt{n}\begin{pmatrix}\widehat{\mu}_{\mathrm{cox}} - \mu_{\mathrm{true}}\\ \widehat{\mu}_{\mathrm{pm}} - \mu_0\end{pmatrix} \overset{d}{\to} -\begin{pmatrix}Z_{A,\mathrm{cox}}(\phi_{\mathrm{cox}})/h_{\mathrm{cox}}(\phi_{\mathrm{cox}})\\ Z_{A,\mathrm{pm}}(\phi_{\mathrm{pm}})/h_{\mathrm{pm}}(\phi_{\mathrm{pm}})\end{pmatrix} = -\begin{pmatrix}\zeta_{\mathrm{cox}}(\phi_{\mathrm{cox}})^{\mathrm{t}}Z_{\mathrm{cox}}(\phi_{\mathrm{cox}})/h_{\mathrm{cox}}(\phi_{\mathrm{cox}})\\ \zeta_{\mathrm{pm}}(\phi_{\mathrm{pm}})^{\mathrm{t}}Z_{\mathrm{pm}}(\phi_{\mathrm{pm}})/h_{\mathrm{pm}}(\phi_{\mathrm{pm}})\end{pmatrix},$$

as long as $\alpha_{\mathrm{true}}(\phi_{\mathrm{cox}})$ and $\alpha_{\mathrm{pm}}(\phi_{\mathrm{pm}};\theta)$ are positive, the latter for all $\theta \in N(\theta_0)$. Consequently, for the focus parameter in (S9), Σ_μ of (31) has elements v_{cox}, v_c and v_{pm} given by

$$\begin{aligned}
v_{\mathrm{cox}} &= \{\zeta_{\mathrm{cox}}(\phi_{\mathrm{cox}})^{\mathrm{t}}\Sigma_{11}(\phi_{\mathrm{cox}},\phi_{\mathrm{cox}})\zeta_{\mathrm{cox}}(\phi_{\mathrm{cox}})\}/\{h_{\mathrm{cox}}(\phi_{\mathrm{cox}})^2\},\\
v_c &= \{\zeta_{\mathrm{cox}}(\phi_{\mathrm{cox}})^{\mathrm{t}}\Sigma_{12}(\phi_{\mathrm{cox}},\phi_{\mathrm{pm}})\zeta_{\mathrm{pm}}(\phi_{\mathrm{pm}})\}/\{h_{\mathrm{cox}}(\phi_{\mathrm{cox}})h_{\mathrm{pm}}(\phi_{\mathrm{pm}})\}, \qquad \text{(S10)}\\
v_{\mathrm{pm}} &= \{\zeta_{\mathrm{pm}}(\phi_{\mathrm{pm}})^{\mathrm{t}}\Sigma_{22}(\phi_{\mathrm{pm}},\phi_{\mathrm{pm}})\zeta_{\mathrm{pm}}(\phi_{\mathrm{pm}})\}/\{h_{\mathrm{pm}}(\phi_{\mathrm{pm}})^2\}.
\end{aligned}$$

Remark S1. Under model conditions, where $\phi_{\rm cox} = \phi_{\rm pm} = \phi_{u,x}$, then the ARE for the u-quantile (conditioned on covariates given by x) takes the form

$$\frac{v_{\rm pm}}{v_{\rm cox}} = \frac{\zeta_{\rm cox}(\phi_{u,x})^{\rm t}\Sigma_{22}(\phi_{u,x},\phi_{u,x})\zeta_{\rm cox}(\phi_{u,x})}{\zeta_{\rm cox}(\phi_{u,x})^{\rm t}\Sigma_{11}(\phi_{u,x},\phi_{u,x})\zeta_{\rm cox}(\phi_{u,x})},$$

which is identical to the ARE for $A(\phi_{u,x})$, i.e. for $A(t\,|\,x)$ such that $u = 1 - S(t\,|\,x)$. An analogue holds without covariates.

References

Andersen PK, Borgan Ø, Gill RD, Keiding N (1993) Statistical Models Based on Counting Processes. Springer-Verlag, Berlin

Billingsley P (1968) Convergence of Probability Measures. Wiley, New York

Hjort NL (1992) On inference in parametric survival data models. International Statistical Review 60:355–387

Jullum M, Hjort NL (2018) What price semiparametric Cox regression? Lifetime Data Analysis

van der Vaart A, Wellner J (1996) Weak Convergence and Empirical Processes. Springer, Berlin

van der Vaart A (2000) Asymptotic Statistics. Cambridge University Press, Cambridge