Vertical Quantile Comparison Functions Estimation in Location Scale Families

Kristian Nestor
Sundarraman Subramanian
Department of Mathematical Sciences
New Jersey Institute of Technology
Newark
USA

Abstract

Two-sample location-scale refers to a situation in which a pair of random variables are linearly related to a base random variable that has mean 0 and variance 1. Using a formulation that leverages the location-scale structure, a semiparametric estimator of the vertical quantile comparison function is proposed and its large-sample properties are derived. Its efficacy relative to that of a nonparametric estimator as well as its robustness to departures from location-scale models is investigated through numerical studies.

Key words: Compactly differentiable; Empirical process; Functional delta method; Kaplan–Meier integral; Empirical coverage probability; Quantile functions

1 Introduction

Suppose that there are two distribution functions F_1 and F_2 induced by X_1 and X_2 respectively, a pair of independent random variables. Let $p \in (0,1)$. Associated with $F_i, i = 1,2$, are $Q_i(p)$, where $Q_i(p) = \inf\{x : F_i(x) \geq p\}$ are the quantile functions that yield values, called the p-quantiles, on the real line. The vertical quantile comparison (VQC) function, $\Upsilon(p)$, maps the interval $(0,1)$ back to itself through $\Upsilon : p \mapsto (F_2 \circ Q_1)(p) = F_2(Q_1(p))$. As p runs over the interval $(0,1)$, the graph of the VQC function traces a curve that may broadly overlap the 45-degree line, or may traffic in the upper or lower regions away from the 45-degree line, or may intersect it at one or several points. The 45-degree line is the benchmark signaling the equality of F_1 and F_2.

The vertical quantile function has broad applicability in many areas. For instance, in clinical trials, if F_1 and F_2 are the distribution functions of the time to an outcome for control and treatment groups respectively, then $\Upsilon(0.5)$ represents the probability that a treatment group individual has a shorter survival time than the median for the control group. A low value indicates that individuals from the treatment group have a higher likelihood of survival beyond the control median, suggesting that the treatment is effective and may be pursued. A high value would not provide sufficient confidence in the treatment, calling for its termination. In this way, the value of the VQC function at various designated p quantiles (signposts) offers a road map of treatment efficacy. The measure provides a robust alternative to hazard ratios with lesser computational overheads. The same idea pervades in quality control and reliability. The VQC function helps quantify the frequency with which a new product outperforms an existing one at some of its chosen p quantiles. In economics, the VQC function is used to compare incomes between groups, offering interpretable summaries of relative advantage at specific quantiles (Chernozhukov, Fernandez-Val, and Melly, 2013). For censored data, Li, Tiwari, and Wells (1996) proposed a nonparametric estimator for $\Upsilon(p)$, which employed the Kaplan–Meier (KM) estimator of F_2 and the quantile function estimator associated with F_1. The estimator was reported to perform well in numerical studies.

Suppose that $X_i = \mu_i + \sigma_i W_i$, $i = 1,2$, where W_i are independent copies of W having the distribution function F with mean 0 and variance 1. We refer to W as having the base distribution F. For example, W may be standard normal. The parameters μ_i and σ_i are the means and standard deviations of X_i. Note that

$$F_i(x) = F\left(\frac{x - \mu_i}{\sigma_i}\right), \quad i = 1,2. \tag{1.1}$$

It is said that F_1 and F_2 belong to a location-scale (LS) family of distributions with base distribution function F. LS models are widely used in medical research, where outcome distributions are paramount for evaluating treatment efficacy and clinical decision-making. More specifically, differences in disease progression often appears as a shift or scale of a health related outcome distribution such as blood pressure, survival time, biomarker levels, among others. LS models are essential in medical studies since they naturally handle heterogeneity between patient groups. In survival analysis, accelerated failure time models assume that the log-survival times follow an LS family of distributions such as the log-normal or log-logistic, allowing clinicians to interpret treatment effects as multiplicative changes in median survival (Kalbfleischa and Prentice, 2002).

In this paper, we are concerned with estimation of the VQC functions in two-sample LS families. Note that the distribution of $(X_i - \mu_i)/\sigma_i, i = 1,2$, is F, so, a basic rationale is that the LS framework admits the pooling of standardized values from both samples, inflating the effective sample size that one expects would improve the efficiency of the estimate. The proposed estimator, as will be seen below, is semiparametric.

The quantile functions associated with F_i are linearly related to the quantile function associated with F:

$$Q_i(p) \quad = \quad \mu_i + \sigma_i Q(p), \;\; i = 1,2. \tag{1.2}$$

From Eq. (1.1) and Eq. (1.2), the VQC function under the two-sample LS framework is

$$\Upsilon(p) \quad = \quad F\left(\frac{\mu_1 - \mu_2}{\sigma_2} + \frac{\sigma_1}{\sigma_2} Q(p)\right), \tag{1.3}$$

a formulation that leverages the LS structure to express $\Upsilon(p)$ in terms of the base distribution and its quantile function. Whereas the nonparametric estimator of $\Upsilon(p)$ uses the separate samples for the distribution and quantile functions estimation, the proposed approach exploits Eq. (1.3) and uses the combined sample for estimating F and Q embedded in the LS-adjusted formula for $\Upsilon(p)$. The apparent advantage that an inflated sample gives may come with a price as we now need estimates of the finite dimensional parameters μ_i and σ_i. This paper investigates whether the trade off leads to efficiency gains while preserving some of the flexibility of the nonparametric approach.

In proposing to exploit Eq. (1.3), we first consider the case of no censoring. Later we show how to extend the approach to the censored case. Let $X_{i1}, \ldots, X_{i,n_i}$ be a random sample from F_i, $i = 1, 2$. Estimate $\boldsymbol{\theta}_i = (\mu_i, \sigma_i)^\top$ and form the estimated standardized values $\hat{W}_{ij} = (X_{ij} - \hat{\mu}_i)/\hat{\sigma}_i, j = 1, \ldots, n_i, i = 1, 2$, where $\hat{\mu}_i$ and $\hat{\sigma}_i$ are the sample means and standard deviations respectively. Let $\hat{F}_{\hat{\boldsymbol{\theta}}}$ be the empirical distribution function of the ordered $\hat{W}_{ij}$ and let $\hat{Q}_{\hat{\boldsymbol{\theta}}}$ be the empirical quantile function, see Section 2 for the notations. Our estimator of $\Upsilon(p)$ is

$$\hat{\Upsilon}(p) = \hat{F}_{\hat{\boldsymbol{\theta}}}\left(\frac{\hat{\mu}_1 - \hat{\mu}_2}{\hat{\sigma}_2} + \frac{\hat{\sigma}_1}{\hat{\sigma}_2}\hat{Q}_{\hat{\boldsymbol{\theta}}}(p)\right). \tag{1.4}$$

When there is right censoring, Eq. (1.4) still applies with $\hat{F}_{\hat{\boldsymbol{\theta}}}$ taken as the Kaplan–Meier (KM) estimator of the pooled estimated standardized values, and $\hat{\mu}_i$ and $\hat{\sigma}_i$ are the estimates obtained via Stute's (1995) KM integrals. A technical justification that pooling the samples in the presence of censoring does not lead to internal conflicts will be offered when we address the censored case in Section 2.3.

The proposed estimator is indexed by estimated (finite dimensional) parameters. A technical challenge is to replace the estimated parameters by their true ones, for which we exploit an empirical processes approach employed by Stute and Zhu (2005). The decomposition of the (centered) estimator's large-sample representation as a sum of centered processes indexed by "known parameters" – for which large-sample representations can be readily given from available results – and the centered finite-dimensional quantities $\hat{\mu}_i - \mu_i$ and $\hat{\sigma}_i - \sigma_i$ are derived. Due to the presence of estimated parameters (the plug-ins), the proposed estimator has an intractable asymptotic variance covariance function. The bootstrap is deployed to construct pointwise confidence intervals for the VQC function using the proposed estimator.

The paper is organized as follows. Section 2.1 details a generic large-sample analysis. The cases of no-censoring and censoring are treated in Sections 2.2 and 2.3. Section 3 contains numerical results. Some concluding discussion is given in Section 4. Certain technical derivations can be found in the Appendix.

2 The estimator and its large-sample study

We first derive a generic large sample representation for the proposed estimator. In the subsections that follow, we fine-tune the representations for the cases of no-censoring and censoring.

2.1 Generic large-sample representation

For $i = 1, 2$, let $X_{ij}, j = 1, \ldots, n_i$, be independent and identically distributed (iid) random variables with distribution function F_i. The random variable $W_{ij} = (X_{ij} - \mu_i)/\sigma_i$, is notional (theoretical construct; unobservable), hence we must use the estimated W_{ij} obtained by substituting estimates $\hat{\mu}_i$ and $\hat{\sigma}_i$ for the unknown parameters. With this in mind, let $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$, where $\boldsymbol{\theta}_i = (\mu_i, \sigma_i)^\top, i = 1, 2$. Let $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \boldsymbol{\gamma}_2^\top)^\top$ be in a neighborhood of $\boldsymbol{\theta}$, where $\boldsymbol{\gamma}_i = (\mu_i', \sigma_i')^\top, i = 1, 2$. Let $W_{ij}^{\boldsymbol{\gamma}_i} = (X_{ij} - \mu_i')/\sigma_i'$ and let $F_{W_i}^{\boldsymbol{\gamma}_i}$ be the distribution of $W_i^{\boldsymbol{\gamma}_i}$. Note that $W_{ij}^{\boldsymbol{\theta}_i} \equiv W_{ij}, j = 1, \ldots, n_i$, are iid with common distribution F. Let $W_{\boldsymbol{\gamma}}$ be a random variable from the collection $W_{ij}^{\boldsymbol{\gamma}_i}$. To determine $F_{\boldsymbol{\gamma}}$, its distribution function, let ξ, independent of X_1 and X_2, be Bernoulli with success probability ρ, indicating membership in the first population. Then

$$W_{\boldsymbol{\gamma}} \;=\; \left(\frac{X_1 - \mu_1'}{\sigma_1'}\right)^{\xi}\left(\frac{X_2 - \mu_2'}{\sigma_2'}\right)^{1-\xi}. \tag{2.1}$$

Note that when $\mu_i' = \mu_i$ and $\sigma_i' = \sigma_i$, $W_{\boldsymbol{\gamma}} \equiv W_{\boldsymbol{\theta}} := W$ has distribution function F. We adapt the conditions of Stute and Zhu (2005) to impose the following: (**C1**) For some $\zeta > 2$, $E\left(|X_i|^\zeta\right) < \infty$; (**C2**) For all $\boldsymbol{\gamma}$ satisfying

$\|\boldsymbol{\gamma} - \boldsymbol{\theta}\| = O(n^{-1/2})$, the random variables $W_i^{\boldsymbol{\gamma}_i}$ have continuously differentiable densities, and both the densities and their derivatives are bounded and bounded away from 0 over the support, uniformly in $\boldsymbol{\gamma}_i$; (**C3**) The distribution functions of $W_i^{\boldsymbol{\gamma}_i}$ are continuous in $\boldsymbol{\gamma}_i$ at $\boldsymbol{\gamma}_i = \boldsymbol{\theta}_i$, meaning that $\boldsymbol{\gamma}_i \to \boldsymbol{\theta}_i \implies F_{W_i}^{\boldsymbol{\gamma}_i}(\cdot) \to F(\cdot)$.

Lemma 1. *Let* $[a,b] \subset \mathbb{R}^1$. *Then,* $\sup_{\omega \in [a,b]} |F_{\boldsymbol{\gamma}}(w) - F(w)| = O(n^{-1/2})$.

Let f be the density of F. For any $w \in \mathbb{R}^1$, let w_i^* be between w and $\sigma_i^{-1}\{(\mu_i' - \mu_i) + \sigma_i' w\}$. Then .

$$\begin{aligned}
F_{\boldsymbol{\gamma}}(w) := \mathbb{P}(W_{\boldsymbol{\gamma}} \leq w) &= \mathbb{P}\left(\frac{X_1 - \mu_1'}{\sigma_1'} \leq w\right)\rho + \mathbb{P}\left(\frac{X_2 - \mu_2'}{\sigma_2'} \leq w\right)(1-\rho) \\
&= \mathbb{P}\left(W_1 \leq \frac{\mu_1' - \mu_1}{\sigma_1} + \frac{\sigma_1'}{\sigma_1}w\right)\rho + \mathbb{P}\left(W_2 \leq \frac{\mu_2' - \mu_2}{\sigma_2} + \frac{\sigma_2'}{\sigma_2}w\right)(1-\rho) \\
&= F\left(\frac{\mu_1' - \mu_1}{\sigma_1} + \frac{\sigma_1'}{\sigma_1}w\right)\rho + F\left(\frac{\mu_2' - \mu_2}{\sigma_2} + \frac{\sigma_2'}{\sigma_2}w\right)(1-\rho) \qquad (2.2) \\
&= F(w) + \frac{\rho}{\sigma_1}\{(\mu_1' - \mu_1) + (\sigma_1' - \sigma_1)w\} f(w_1^*) \\
&\quad + \frac{1-\rho}{\sigma_2}\{(\mu_2' - \mu_2) + (\sigma_2' - \sigma_2)w\} f(w_2^*). \qquad (2.3)
\end{aligned}$$

By condition **C2**, uniformly for $w \in [a,b]$, $|w_i^* - w| = O(n^{-1/2})$. In turn, $|w_1^* - w_2^*| = O(n^{-1/2})$. By the mean value theorem and condition **C2**, $|f(w_1^*) - f(w_2^*)| = O(n^{-1/2})$. Eq. (2.3) completes the proof. □

For $\alpha \in (0,1)$ and $\beta \in (0,1)$, and $p \in [\alpha, \beta]$, $Q_{\boldsymbol{\gamma}}(p) := \inf\{w : F_{\boldsymbol{\gamma}}(w) \geq p\}$. Note that $Q_{\boldsymbol{\theta}}(p) \equiv Q(p)$. Let c_i be between $Q(p)$ and $\sigma_i^{-1}\{(\mu_i' - \mu_i) + \sigma_i' Q_{\boldsymbol{\gamma}}(p)\}$. Let $d_1 = \rho f(c_1)/(\rho f(c_1) + (1-\rho)f(c_2)) = 1 - d_2$.

Lemma 2. *Suppose for all* $p \in [\alpha, \beta]$, $Q_{\boldsymbol{\gamma}}(p) \in [a,b]$ *uniformly over* $\boldsymbol{\gamma}$, *where* $[a,b]$ *is a finite interval. Then*

$$Q_{\boldsymbol{\gamma}}(p) - Q(p) = -\frac{d_1}{\sigma_1}(\mu_1' - \mu_1) - \frac{d_2}{\sigma_2}(\mu_2' - \mu_2) - Q_{\boldsymbol{\gamma}}(p)\frac{d_1}{\sigma_1}(\sigma_1' - \sigma_1) - Q_{\boldsymbol{\gamma}}(p)\frac{d_2}{\sigma_2}(\sigma_2' - \sigma_2). \qquad (2.4)$$

Proof Since $F_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p)) = p$, apply Eq. (2.2) to obtain

$$\rho F\left(\frac{\mu_1' - \mu_1}{\sigma_1} + \frac{\sigma_1'}{\sigma_1}Q_{\boldsymbol{\gamma}}(p)\right) + (1-\rho) F\left(\frac{\mu_2' - \mu_2}{\sigma_2} + \frac{\sigma_2'}{\sigma_2}Q_{\boldsymbol{\gamma}}(p)\right) = p.$$

Applying the mean value theorem, we must have

$$\begin{aligned}
&\rho\left[F(Q(p)) + \left(\frac{\mu_1' - \mu_1}{\sigma_1} + \frac{\sigma_1'}{\sigma_1}Q_{\boldsymbol{\gamma}}(p) - Q(p)\right) f(c_1)\right] \\
&+(1-\rho)\left[F(Q(p)) + \left(\frac{\mu_2' - \mu_2}{\sigma_2} + \frac{\sigma_2'}{\sigma_2}Q_{\boldsymbol{\gamma}}(p) - Q(p)\right) f(c_2)\right] = p.
\end{aligned}$$

Then, after some elementary transpositions, it is easy to show that

$$\begin{aligned}
(\rho f(c_1) + (1-\rho) f(c_2))(Q_{\boldsymbol{\gamma}}(p) - Q(p)) &= -\rho f(c_1)\frac{\mu_1' - \mu_1}{\sigma_1} - (1-\rho) f(c_2)\frac{\mu_2' - \mu_2}{\sigma_2} \\
&\quad - Q_{\boldsymbol{\gamma}}(p)\left(\rho f(c_1)\frac{\sigma_1' - \sigma_1}{\sigma_1} + (1-\rho) f(c_2)\frac{\sigma_2' - \sigma_2}{\sigma_2}\right).
\end{aligned}$$

Eq. (2.4) follows. The proof of the lemma is completed. □

Remark 1 For $p \in [\alpha, \beta]$, let $\mathbb{A}_{\boldsymbol{\gamma}}(p) = \{w : F_{\boldsymbol{\gamma}}(w) \geq p\}$. Note that $\inf \mathbb{A}_{\boldsymbol{\gamma}}(p) = Q_{\boldsymbol{\gamma}}(p)$. By condition C2 and Eq. (2.3), $|F_{\boldsymbol{\gamma}}(w) - F(w)| \leq K\|\boldsymbol{\theta} - \boldsymbol{\gamma}\|$, where K is a generic constant. Since $\|\boldsymbol{\gamma} - \boldsymbol{\theta}\| = O(n^{-1/2})$, the upper bound $K\|\boldsymbol{\theta} - \boldsymbol{\gamma}\|$ can be chosen so that $p \pm K\|\boldsymbol{\gamma} - \boldsymbol{\theta}\| \in [\alpha, \beta]$. If $w \in \mathbb{A}_{\boldsymbol{\gamma}}(p)$, then $F(w) = F_{\boldsymbol{\gamma}}(w) + (F(w) - F_{\boldsymbol{\gamma}}(w)) \geq p - K\|\boldsymbol{\theta} - \boldsymbol{\gamma}\|$. Hence $\mathbb{A}_{\boldsymbol{\gamma}}(p) \subset \mathbb{A}_{\boldsymbol{\theta}}(p - K\|\boldsymbol{\theta} - \boldsymbol{\gamma}\|)$. It follows that $Q(p - K\|\boldsymbol{\theta} - \boldsymbol{\gamma}\|) \leq Q_{\boldsymbol{\gamma}}(p)$. On the other hand, if $w \in \mathbb{A}_{\boldsymbol{\theta}}(p + K\|\boldsymbol{\theta} - \boldsymbol{\gamma}\|)$, then $F(w) \geq p + K\|\boldsymbol{\theta} - \boldsymbol{\gamma}\|$ and hence that $F_{\boldsymbol{\gamma}}(w) \geq F(w) - K\|\boldsymbol{\gamma} - \boldsymbol{\theta}\| \geq p$.

Hence $\mathbb{A}_{\boldsymbol{\theta}}(p+K\|\boldsymbol{\theta}-\boldsymbol{\gamma}\|)\subset\mathbb{A}_{\boldsymbol{\gamma}}(p)$. It follows that $Q(p-K\|\boldsymbol{\theta}-\boldsymbol{\gamma}\|)\le Q_{\boldsymbol{\gamma}}(p)\le Q(p+K\|\boldsymbol{\theta}-\boldsymbol{\gamma}\|)$. By the mean value theorem, uniformly for $p\in[\alpha,\beta]$, $|Q_{\boldsymbol{\gamma}}(p)-Q(p)|\le K\|\boldsymbol{\theta}-\boldsymbol{\gamma}\| = O(n^{-1/2})$.

Let $\hat{F}_{\boldsymbol{\gamma}}$ be the notional EDF of $\{W_{ij}^{\boldsymbol{\gamma}_i}, j=1,\dots,n_i, i=1,2\}$. As already mentioned, when $\boldsymbol{\gamma}=\boldsymbol{\theta}$, the random variables $W_{ij}^{\boldsymbol{\theta}_i}\equiv W_{ij}, j=1,\dots,n_i, i=1,2$, are iid with distribution function F. Combine all W_{ij} and let $\hat{F}_{\boldsymbol{\theta}}\equiv\hat{F}$ be the (notional) EDF of the combined sample. We are concerned with the special case $\boldsymbol{\gamma}=\hat{\boldsymbol{\theta}}$, where $\hat{\boldsymbol{\theta}}=(\hat{\boldsymbol{\theta}}_1^\top,\hat{\boldsymbol{\theta}}_2^\top)^\top$, $\hat{\boldsymbol{\theta}}_1=(\hat{\mu}_1,\hat{\sigma}_1)^\top$ and $\hat{\boldsymbol{\theta}}_2=(\hat{\mu}_2,\hat{\sigma}_2)^\top$, and $\hat{\mu}_i$ and $\hat{\sigma}_i$ are the mean and standard deviation estimators. Combine all $\hat{W}_{ij}=(X_{ij}-\hat{\mu}_i)/\hat{\sigma}_i$ and let $\hat{F}_{\hat{\boldsymbol{\theta}}}$ be the (computable) EDF of the combined sample.

Set $\hat{C}_{\hat{\boldsymbol{\theta}}}(p)=\left(\hat{\mu}_1-\hat{\mu}_2+\hat{\sigma}_1\hat{Q}_{\hat{\boldsymbol{\theta}}}(p)\right)/\hat{\sigma}_2$ and $C(p)=(\mu_1-\mu_2+\sigma_1 Q(p))/\sigma_2$. From Eq. (1.3), the VQC function is $\Upsilon(p)=F\circ C(p)=F(C(p))$. From Eq. (1.4), the proposed estimator is $\hat{F}_{\hat{\boldsymbol{\theta}}}(\hat{C}_{\hat{\boldsymbol{\theta}}}(p))$. The estimator $\hat{\Upsilon}$ is a composition of $\hat{F}_{\hat{\boldsymbol{\theta}}}$, an estimator of F, and $\hat{C}_{\hat{\boldsymbol{\theta}}}(p)$, which shifts and scales an estimator of Q using the mean and standard deviation estimators. Using the $n^{1/2}$ consistency of $\hat{\mu}_i$ and $\hat{\sigma}_i$, it can be shown that

$$\begin{aligned}\hat{B}(p) \;:=\; \hat{C}_{\hat{\boldsymbol{\theta}}}(p)-C(p) &= \frac{\sigma_1}{\sigma_2}(\hat{Q}_{\hat{\boldsymbol{\theta}}}(p)-Q(p))+\frac{1}{\sigma_2}\{(\hat{\mu}_1-\mu_1)+(\hat{\sigma}_1-\sigma_1)Q(p)-(\hat{\mu}_2-\mu_2)\}\\ &\quad -\frac{1}{\sigma_2^2}\{(\mu_1-\mu_2)+\sigma_1 Q(p)\}(\hat{\sigma}_2-\sigma_2)+o_{\mathbb{P}}(n^{-1/2}). \qquad (2.5)\end{aligned}$$

In propositions 1 and 2 we derive large sample representations for $\hat{A}(s):=\hat{F}_{\hat{\boldsymbol{\theta}}}(s)-F(s)$ and $\hat{Q}_{\hat{\boldsymbol{\theta}}}(p)-Q(p)$.

Proposition 1. *Under conditions* **C1–C3**, $\hat{A}(s):=\hat{F}_{\hat{\boldsymbol{\theta}}}(s)-F(s)$ *admits the large sample representation*

$$\begin{aligned}\hat{A}(s) &= \hat{F}(s)-F(s)+\frac{\rho}{\sigma_1}f(s)\{(\hat{\mu}_1-\mu_1)+s(\hat{\sigma}_1-\sigma_1)\}\\ &\quad+\frac{1-\rho}{\sigma_2}f(s)\{(\hat{\mu}_2-\mu_2)+s(\hat{\sigma}_2-\sigma_2)\}+o_{\mathbb{P}}(n^{-1/2}). \qquad (2.6)\end{aligned}$$

Proof We write $\hat{F}_{\hat{\boldsymbol{\theta}}}(s)=\hat{F}_{\hat{\boldsymbol{\theta}}}(s)-F_{\hat{\boldsymbol{\theta}}}(s)-\hat{F}(s)+F(s)+F_{\hat{\boldsymbol{\theta}}}(s)+\hat{F}(s)-F(s)$. Apply Eq. (A.1) in Lemma 3 to obtain $\hat{F}_{\hat{\boldsymbol{\theta}}}(s)-F(s)=\hat{F}(s)-F(s)+F_{\hat{\boldsymbol{\theta}}}(s)-F(s)+o_{\mathbb{P}}(n^{-1/2})$. Let s_i^* be between s and $\sigma_i^{-1}\{(\hat{\mu}_i-\mu_i)+\hat{\sigma}_i s\}$. Applying Eq. (2.3) to $F_{\hat{\boldsymbol{\theta}}}(s)-F(s)$, we have, modulus a remainder term $o_{\mathbb{P}}(n^{-1/2})$,

$$\hat{A}(s) = \hat{F}(s)-F(s)+\frac{\rho}{\sigma_1}\{(\hat{\mu}_1-\mu_1)+s(\hat{\sigma}_1-\sigma_1)\}f(s_1^*)+\frac{1-\rho}{\sigma_2}\{(\hat{\mu}_2-\mu_2)+s(\hat{\sigma}_2-\sigma_2)\}f(s_2^*).$$

Since $s_i^*\xrightarrow{\mathbb{P}}s$ and since f is continuous, Eq. (2.6) follows immediately. □

We assume (**C4**): For all $p\in[\alpha,\beta]$, $Q_{\boldsymbol{\gamma}}(p)\in[a,b]$ uniformly over $\boldsymbol{\gamma}$, where $[a,b]$ is a finite interval.

Proposition 2. *Under conditions* **C1–C4**, $\hat{Q}_{\hat{\boldsymbol{\theta}}}(p)-Q(p)$ *admits the large-sample representation*

$$\begin{aligned}\hat{Q}_{\hat{\boldsymbol{\theta}}}(p)-Q(p) &= \hat{Q}(p)-Q(p)-\frac{\rho}{\sigma_1}\{(\hat{\mu}_1-\mu_1)+Q(p)(\hat{\sigma}_1-\sigma_1)\}\\ &\quad-\frac{(1-\rho)}{\sigma_2}\{(\hat{\mu}_2-\mu_2)+Q(p)(\hat{\sigma}_2-\sigma_2)\}+o_{\mathbb{P}}(n^{-1/2}). \qquad (2.7)\end{aligned}$$

Proof Substitute $\boldsymbol{\gamma}=\hat{\boldsymbol{\theta}}$ in Eq. (2.4) to obtain

$$Q_{\hat{\boldsymbol{\theta}}}(p)-Q(p) = -\frac{d_1}{\sigma_1}(\hat{\mu}_1-\mu_1)-\frac{d_2}{\sigma_2}(\hat{\mu}_2-\mu_2)-Q_{\hat{\boldsymbol{\theta}}}(p)\frac{d_1}{\sigma_1}(\hat{\sigma}_1-\sigma_1)-Q_{\hat{\boldsymbol{\theta}}}(p)\frac{d_2}{\sigma_2}(\hat{\sigma}_2-\sigma_2), \qquad (2.8)$$

where $d_1=\rho f(c_1)/(\rho f(c_1)+(1-\rho)f(c_2))=1-d_2$, and c_i are between $Q(p)$ and $\sigma_i^{-1}\{(\hat{\mu}_i-\mu_i)+\hat{\sigma}_i Q_{\hat{\boldsymbol{\theta}}}(p)\}$. The consistency of $\hat{\mu}_i$ and $\hat{\sigma}_i$ along with conditions **C1–C4** implies through Eq. (2.8) that $Q_{\hat{\boldsymbol{\theta}}}(p)\xrightarrow{\mathbb{P}}Q(p)$ and that $c_i\xrightarrow{\mathbb{P}}Q(p), i=1,2$. In turn, it follows that $d_1\xrightarrow{\mathbb{P}}\rho$ and $d_2\xrightarrow{\mathbb{P}}1-\rho$. From Eq. (2.8) it now follows that

$$\begin{aligned}Q_{\hat{\boldsymbol{\theta}}}(p)-Q(p) &= -\frac{\rho}{\sigma_1}\{(\hat{\mu}_1-\mu_1)+Q(p)(\hat{\sigma}_1-\sigma_1)\}\\ &\quad-\frac{(1-\rho)}{\sigma_2}\{(\hat{\mu}_2-\mu_2)+Q(p)(\hat{\sigma}_2-\sigma_2)\}+o_{\mathbb{P}}(n^{-1/2}). \qquad (2.9)\end{aligned}$$

Applying a decomposition analogous to that used in the proof of Proposition 1, and then applying Lemma 4,

$$\begin{aligned}\hat{Q}_{\hat{\boldsymbol{\theta}}}(p) - Q(p) &= \hat{Q}(p) - Q(p) + Q_{\hat{\boldsymbol{\theta}}}(p) - Q(p) + \hat{Q}_{\hat{\boldsymbol{\theta}}}(p) - Q_{\hat{\boldsymbol{\theta}}}(p) - \hat{Q}(p) + Q(p) \\ &= \hat{Q}(p) - Q(p) + Q_{\hat{\boldsymbol{\theta}}}(p) - Q(p) + o_{\mathbb{P}}(n^{-1/2}).\end{aligned}$$

Then apply Eq. (2.9) to obtain Eq. (2.7). $\square$

Combining Eqs. (2.5) and (2.7), we obtain

$$\begin{aligned}\hat{B}(p) &= \frac{1}{\sigma_2}\left\{(\hat{\mu}_1 - \mu_1) + (\hat{\sigma}_1 - \sigma_1)Q(p) - (\hat{\mu}_2 - \mu_2)\right\} \\ &\quad - \frac{1}{\sigma_2^2}\left\{(\mu_1 - \mu_2) + \sigma_1 Q(p)\right\}(\hat{\sigma}_2 - \sigma_2) + \frac{\sigma_1}{\sigma_2}(\hat{Q}_{\hat{\boldsymbol{\theta}}}(p) - Q(p)) + o_{\mathbb{P}}(n^{-1/2}) \\ &= \frac{1}{\sigma_2}\left\{(\hat{\mu}_1 - \mu_1) + (\hat{\sigma}_1 - \sigma_1)Q(p) - (\hat{\mu}_2 - \mu_2)\right\} - \frac{1}{\sigma_2^2}\left\{(\mu_1 - \mu_2) + \sigma_1 Q(p)\right\}(\hat{\sigma}_2 - \sigma_2) \\ &\quad + \frac{\sigma_1}{\sigma_2}\left(\hat{Q}(p) - Q(p)\right) - \frac{\rho}{\sigma_2}\left\{(\hat{\mu}_1 - \mu_1) + Q(p)\,(\hat{\sigma}_1 - \sigma_1)\right\} \\ &\quad - \frac{\sigma_1(1-\rho)}{\sigma_2^2}\left\{(\hat{\mu}_2 - \mu_2) + Q(p)\,(\hat{\sigma}_2 - \sigma_2)\right\} + o_{\mathbb{P}}(n^{-1/2}).\end{aligned} \tag{2.10}$$

Combining the coefficients for the various centered quantities in Eq. (2.10), we obtain

$$\begin{aligned}\hat{B}(p) &= \frac{\sigma_1}{\sigma_2}\left(\hat{Q}(p) - Q(p)\right) + \frac{1-\rho}{\sigma_2}(\hat{\mu}_1 - \mu_1) - \frac{1}{\sigma_2}\left\{1 + \frac{\sigma_1(1-\rho)}{\sigma_2}\right\}(\hat{\mu}_2 - \mu_2) \\ &\quad + \frac{1-\rho}{\sigma_2}Q(p)(\hat{\sigma}_1 - \sigma_1) - \frac{1}{\sigma_2^2}\left[(\mu_1 - \mu_2) + \sigma_1\,(2-\rho)\,Q(p)\right](\hat{\sigma}_2 - \sigma_2) + o_{\mathbb{P}}(n^{-1/2}).\end{aligned} \tag{2.11}$$

Let $D[a,b]$ be the class of càdlàg functions on $[a,b]$ equipped with the supremum norm. Let $\hat{\boldsymbol{T}}(\cdot) = \left(\hat{F}_{\hat{\boldsymbol{\theta}}}(\cdot), \hat{C}(\cdot)\right)^\top$ be the sequence of random elements and $\boldsymbol{T}(\cdot) = (F(\cdot), C(\cdot))^\top$ be the fixed point, both in $D[a,b] \times D[\alpha,\beta]$. Suppose that $n^{1/2}(\hat{\boldsymbol{T}} - \boldsymbol{T}) \xrightarrow{\mathcal{D}} \boldsymbol{W}$ as $n \to \infty$ on $D[a,b] \times D[\alpha,\beta]$, where $\boldsymbol{W} = (\mathbb{W}_1, \mathbb{W}_2)^\top$ is a random element of $D[a,b] \times D[\alpha,\beta]$. Let $C[a,b]$ be the subspace of continuous functions in $D[a,b]$. Define

$$\begin{aligned}\hat{D}(p) &= n^{1/2}\left(\hat{F}(C(p)) - F(C(p))\right) + \frac{\sigma_1}{\sigma_2}f(C(p))\,n^{1/2}\left(\hat{Q}(p) - Q(p)\right) \\ &\quad + \left(\frac{\rho}{\sigma_1} + \frac{1-\rho}{\sigma_2}\right)f(C(p))\left\{n^{1/2}(\hat{\mu}_1 - \mu_1) - \frac{\sigma_1}{\sigma_2}\,n^{1/2}(\hat{\mu}_2 - \mu_2)\right\} \\ &\quad + \frac{1}{\sigma_2}\left(\frac{\mu_1 - \mu_2}{\sigma_1}\rho + Q(p)\right)f(C(p))\left\{n^{1/2}(\hat{\sigma}_1 - \sigma_1) - \frac{\sigma_1}{\sigma_2}\,n^{1/2}(\hat{\sigma}_2 - \sigma_2)\right\}.\end{aligned} \tag{2.12}$$

Large-sample representations for $n^{1/2}\left(\hat{F}(C(\cdot)) - F(C(\cdot))\right)$, $n^{1/2}\left(\hat{Q}(\cdot) - Q(\cdot)\right)$, and for $n^{1/2}(\hat{\mu}_i - \mu_i), i = 1,2,$ and $n^{1/2}(\hat{\sigma}_i - \sigma_i), i = 1,2,$ found in Eq. (2.12) will be given in Appendixes A.2 and A.3.

Theorem 1. *Under conditions* **C1–C4**, $n^{1/2}(\hat{\Upsilon} - \Upsilon)$ *is asymptotically equivalent to* $\hat{D}$ *given by Eq. (2.12).*

Proof Define $\varphi : D[a,b] \times D[\alpha,\beta] \to D[a,b]$ by $\varphi(\eta,\psi) := \eta \circ \psi \equiv \eta(\psi)$ for every $(\eta,\psi) \in D[a,b] \times D[\alpha,\beta]$. Note that $\Upsilon(\cdot) = \varphi(\boldsymbol{T}(\cdot)^\top)$ and $\hat{\Upsilon}(\cdot) = \varphi(\hat{\boldsymbol{T}}(\cdot)^\top)$. Therefore,

$$n^{1/2}\left\{\hat{\Upsilon}(\cdot) - \Upsilon(\cdot)\right\} = n^{1/2}\left\{\varphi\left(\hat{T}(\cdot)^\top\right) - \varphi\left(\boldsymbol{T}(\cdot)^\top\right)\right\}$$

By Proposition II.8.8 of Andersen, Borgan, Gill, and Keiding (1993), φ is compactly differentiable tangentially to $C[a,b] \times D[\alpha,\beta]$ at $\boldsymbol{T}$ with derivative $d\,\varphi(\boldsymbol{T}^\top)\cdot(h,k)$. The operator $d\,\varphi(\boldsymbol{T}^\top)$ acting on (h,k) produces

$$d\,\varphi(\boldsymbol{T}^\top)\cdot(h,k) \;\equiv\; d\,\varphi(F,C)\cdot(h,k) \;=\; h \circ C + F' \circ C \cdot k, \tag{2.13}$$

where F' is the ordinary continuous derivative of F and the last "$\cdot$" is ordinary multiplication. Theorem II.8.1 of Andersen et al. (1993), the functional delta method, can be now applied to deduce that

$$n^{1/2}\left(\hat{\Upsilon}(\cdot)-\Upsilon(\cdot)\right) \equiv n^{1/2}\left(\varphi(\hat{\boldsymbol{T}}^{\top})-\varphi(\boldsymbol{T}^{\top})\right) \qquad \text{and} \qquad d\,\varphi(\boldsymbol{T}^{\top})\cdot n^{1/2}\left(\hat{\boldsymbol{T}}^{\top}-\boldsymbol{T}^{\top}\right)$$

are asymptotically equivalent. Since $n^{1/2}\left(\hat{\boldsymbol{T}}-\boldsymbol{T}\right)$ is asymptotically equivalent to the vector $n^{1/2}\left(\hat{A},\hat{B}\right)^{\top}$, we may apply Eq. (2.13) and evaluate $d\,\varphi(\boldsymbol{T}^{\top})=d\,\varphi(F,C)$ at $h=n^{1/2}\hat{A}(\cdot)$ and $k=n^{1/2}\hat{B}(\cdot)$, to obtain

$$n^{1/2}\left(\hat{\Upsilon}(\cdot)-\Upsilon(\cdot)\right) = n^{1/2}\hat{A}\left(C(\cdot)\right)+f\left(C(\cdot)\right)\times n^{1/2}\hat{B}(\cdot)+o_{\mathbb{P}}(1). \tag{2.14}$$

Plugging in the representations given by Eq. (2.6) and Eq. (2.11) into Eq. (2.14), it can be checked that the first two terms of Eq. (2.12) is immediate. It can be checked that $n^{1/2}(\hat{\mu}_1-\mu_1)$ is scaled by the factor $\{\rho/\sigma_1+(1-\rho)/\sigma_2\}f(C(p))$; that $n^{1/2}(\hat{\mu}_2-\mu_2)$ is scaled by the factor

$$\frac{1}{\sigma_2}\left\{(1-\rho)-(1+(1-\rho)\frac{\sigma_1}{\sigma_2}\right\}f(C(p)) = -\frac{\sigma_1}{\sigma_2}\left(\frac{\rho}{\sigma_1}+\frac{1-\rho}{\sigma_2}\right)f(C(p)).$$

Recall that $C(p)=(\mu_1-\mu_2+\sigma_1 Q(p))/\sigma_2$. The centered $n^{1/2}(\hat{\sigma}_1-\sigma_1)$ is scaled by the factor

$$\begin{aligned}\left\{\frac{\rho}{\sigma_1}C(p)+\frac{1-\rho}{\sigma_2}Q(p)\right\}f(C(p)) &= \left\{\frac{\mu_1-\mu_2}{\sigma_1\sigma_2}\rho+\frac{\rho\,Q(p)}{\sigma_2}+\frac{1-\rho}{\sigma_2}Q(p)\right\}f(C(p))\\ &= \frac{1}{\sigma_2}\left\{\frac{\mu_1-\mu_2}{\sigma_1}\rho+Q(p)\right\}f(C(p)).\end{aligned}$$

Finally, the centered $n^{1/2}(\hat{\sigma}_2-\sigma_2)$ is scaled by the factor

$$\begin{aligned}&\left\{\frac{1-\rho}{\sigma_2}C(p)-\frac{1}{\sigma_2^2}\left[(\mu_1-\mu_2)+\{\sigma_1+(1-\rho)\sigma_1\}Q(p)\right]\right\}f(C(p))\\ &=\left\{\frac{1-\rho}{\sigma_2^2}(\mu_1-\mu_2)+\frac{\sigma_1}{\sigma_2^2}Q(p)(1-\rho)-\frac{1}{\sigma_2^2}\left[(\mu_1-\mu_2)+\{\sigma_1+(1-\rho)\sigma_1\}Q(p)\right]\right\}f(C(p))\\ &=-\frac{\sigma_1}{\sigma_2^2}\left\{\frac{\mu_1-\mu_2}{\sigma_1}\rho+Q(p)\right\}f(C(p)).\end{aligned}$$

It follows that $n^{1/2}\left(\hat{\Upsilon}(p)-\Upsilon(p)\right)$ is asymptotically equivalent to $\hat{D}(p)$ given by Eq. (2.12). $\square$

From the proof of Theorem 1, we can deduce that the weak convergence of $n^{1/2}\left(\hat{\Upsilon}(\cdot)-\Upsilon(\cdot)\right)$ in $D[\alpha,\beta]$ to a zero-mean Gaussian process follows from the weak convergence of $n^{1/2}\left(\hat{\boldsymbol{T}}(\cdot)-\boldsymbol{T}(\cdot)\right)$ and the functional delta method. In the next two subsections we obtain final expressions for $\hat{D}(p)$.

2.2 Large-sample representation for the uncensored case

Let $n_1/n \to \kappa$ as $n_1\to\infty$ and $n_2\to\infty$ and let $\kappa_1=\kappa, \kappa_2=1-\kappa$. In Appendix A.2 we show that

$$\begin{aligned}\hat{D}(p) &= n^{-1/2}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\{1_{\{W_{ij}\le C(p)\}}-F(C(p))\}-\frac{\sigma_1}{\sigma_2}f(C(p))\,n^{-1/2}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\frac{1_{\{W_{ij}\le Q(p)\}}-p}{f(Q(p))}\\ &\quad+\sigma_1\left(\frac{\rho}{\sigma_1}+\frac{1-\rho}{\sigma_2}\right)f(C(p))\left[\kappa^{-1/2}\left(n_1^{-1/2}\sum_{j=1}^{n_1}W_{1j}\right)-(1-\kappa)^{-1/2}\left(n_2^{-1/2}\sum_{j=1}^{n_2}W_{2j}\right)\right]\\ &\quad+\frac{1}{2}\left\{\frac{(\mu_1-\mu_2)\rho+\sigma_1 Q(p)}{\sigma_2}\right\}f(C(p))\\ &\quad\times\left\{\kappa^{-1/2}\left(n_1^{-1/2}\sum_{j=1}^{n_1}(W_{1j}^2-1)\right)-(1-\kappa)^{-1/2}\left(n_2^{-1/2}\sum_{j=1}^{n_2}(W_{2j}^2-1)\right)\right\}+o_{\mathbb{P}}(1).\end{aligned} \tag{2.15}$$

The first two terms of $\hat{D}(p)$ result when μ_i and σ_i are known. The additional terms are the price of estimating the unknown μ_i and σ_i. Due to the lengthy expression for $\hat{D}(p)$ above, the limiting covariance function of the Gaussian process would have a complicated form. In Section 3, we use bootstrap to approximate the variance of $\hat{\Upsilon}(p)$, using which we compute pointwise confidence intervals for the VQC function.

2.3 Large-sample representation for the censored case

For each $i = 1, 2$, let $C_{ij}, j = 1, \ldots, n_i$ be a random sample from G_i, where the G_i are the distribution functions of the censoring variables. The observed data consists of $\{(Z_{ij}, \delta_{ij}), j = 1, \ldots, n_i, i = 1, 2\}$, where $Z_{ij} = \min(X_{ij}, C_{ij})$ and $\delta_{ij} = I(X_{ij} \leq C_{ij})$. So that we handle a single (pooled array), let $\hat{W} = (Z - \hat{\mu})/\hat{\sigma}$, where $\hat{\mu}$ may be $\hat{\mu}_1$ or $\hat{\mu}_2$ and likewise for $\hat{\sigma}$.

Standardizing the observed minimum artificially shifts and scales the censored observations. We show that pooling standardized values continues to provide correct estimates of F. Let $\Lambda_{\hat{\boldsymbol{\theta}}}$ be the associated cumulative hazard function (CHF). Let $\tilde{G}_i$ be the distribution functions of $(C_i - \hat{\mu}_i)/\hat{\sigma}_i$. Define

$$\mathcal{D}_t(\alpha, \beta) = \int_{-\infty}^{t} d\,\alpha(s)/\beta(s).$$

It is standard to show under the LS framework that $H_{\hat{W},1}(t) = P(\hat{W} \leq t, \delta = 1)$ and $H_{\hat{W}}(t) = P(\hat{W} \leq t)$ are

$$\begin{aligned} H_{\hat{W},1}(t) &= \rho\,\mathcal{D}_t\left(F_{\hat{\boldsymbol{\theta}}}, 1/(1 - \tilde{G}_1)\right) + (1-\rho)\,\mathcal{D}_t\left(F_{\hat{\boldsymbol{\theta}}}, 1/(1 - \tilde{G}_2)\right); \\ dH_{\hat{W},1}(t) &= \left(\rho(1 - \tilde{G}_1(t)) + (1-\rho)(1 - \tilde{G}_2(t))\right) dF_{\hat{\boldsymbol{\theta}}}(t); \qquad (2.16) \\ 1 - H_{\hat{W}}(t) &= (1 - F_{\hat{\boldsymbol{\theta}}}(t))\left(\rho(1 - \tilde{G}_1(t)) + (1-\rho)(1 - \tilde{G}_2(t))\right). \qquad (2.17) \end{aligned}$$

Performing the operation $\mathcal{D}_t(H_{\hat{W},1}, 1 - H_{\hat{W}})$, it is seen from Eq. (2.16) and Eq. (2.17) that

$$\mathcal{D}_t(H_{\hat{W},1}, 1 - H_{\hat{W}}) = \int_{-\infty}^{t} \frac{1}{1 - H_{\hat{W}}(s)} dH_{\hat{W},1}(s) = \int_{-\infty}^{t} \frac{dF_{\hat{\boldsymbol{\theta}}}(s)}{1 - F_{\hat{\boldsymbol{\theta}}}(s)} \equiv \Lambda_{\hat{\boldsymbol{\theta}}},$$

hence correctly estimating the CHF and, in turn, the distribution function $F_{\hat{\boldsymbol{\theta}}}$. In Appendix A.3 we show that

$$\begin{aligned} \hat{D}(p) = {} & S(C(p))\, n^{-1/2} \sum_{j=1}^{n} I_j^{(1)}(C(p)) - \frac{\sigma_1}{\sigma_2} \frac{f(C(p))}{f(Q(p))} (1-p)\, n^{-1/2} \sum_{j=1}^{n} I_j^{(1)}(Q(p)) \\ & + \left[\frac{\rho}{\sigma_1} + \frac{1-\rho}{\sigma_2} - \frac{\mu_1}{\sigma_1}.\frac{1}{\sigma_2}\left\{\frac{\mu_1 - \mu_2}{\sigma_1}\rho + Q(p)\right\}\right] f(C(p))\, \kappa^{-1/2} \left(n_1^{-1/2} \sum_{j=1}^{n_1} I_{1j}^{(2)}\right) \\ & - \frac{\sigma_1}{\sigma_2}\left[\frac{\rho}{\sigma_1} + \frac{1-\rho}{\sigma_2} - \frac{\mu_2}{\sigma_2}.\frac{1}{\sigma_2}\left\{\frac{\mu_1 - \mu_2}{\sigma_1}\rho + Q(p)\right\}\right] f(C(p))\,(1-\kappa)^{-1/2} \left(n_2^{-1/2} \sum_{j=1}^{n_2} I_{2j}^{(2)}\right) \\ & + \frac{1}{2\sigma_1\sigma_2}\left\{\frac{\mu_1 - \mu_2}{\sigma_1}\rho + Q(p)\right\} f(C(p))\, \kappa^{-1/2} \left(n_1^{-1/2} \sum_{j=1}^{n_1} I_{1j}^{(3)}\right) \\ & - \frac{\sigma_1}{2\sigma_2^3}\left\{\frac{\mu_1 - \mu_2}{\sigma_1}\rho + Q(p)\right\} f(C(p))\,(1-\kappa)^{-1/2} \left(n_2^{-1/2} \sum_{j=1}^{n_2} I_{2j}^{(3)}\right) + o_{\mathbb{P}}(1). \qquad (2.18) \end{aligned}$$

The influence functions $I_l^{(1)}(\cdot), l = 1, \ldots, n$, $I_{ij}^{(2)}, j = 1, \ldots, n_i$, and $I_{ij}^{(3)}, j = 1, \ldots, n_i$ are given by Eq. (A.13), Eq. (A.18) and Eq. (A.20) respectively.

3 Simulation studies

3.1 Numerical results for the uncensored case

We carried out comparison studies between the proposed and Li et al. (1996) estimators. The comparison was based on the mean integrated squared error (MISE), where the integrating variable p was taken over a fine grid of values in the interval $(0.05, 0.95)$. The integrand is the square of the difference of the estimator from the true VQC function. The percent improvement of the proposed estimator relative to the nonparametric estimator was computed.

The failure times for both samples were generated from distributions belonging to some location-scale family. Formally, each random variable is defined as $X_i = \mu_i + \sigma_i W, i = 1, 2$, where W follows some chosen baseline distribution such as standard normal, standardized logistic, or a standardized t distribution. When $W \equiv Z$ is standard normal, the VQC function is $\Upsilon(p) = \Phi\left(\frac{\mu_1-\mu_2}{\sigma_2} + \frac{\sigma_1}{\sigma_2}Q(p)\right)$, where $\Phi(.)$ and $Q(.)$ are the distribution and quantile functions of the standard normal distribution respectively. When W is the standardized logistic, then $W = \sqrt{\frac{3}{\pi}}V$, where V is standard logistic. The VQC function in this case is

$$\Upsilon(p) = F_V\left(\frac{\mu_1-\mu_2}{\sigma_2}\sqrt{\frac{\pi}{3}} + \frac{\sigma_1}{\sigma_2}Q_V(p)\right),$$

where $F_V(.)$ and $Q_V(.)$, are the distribution and quantile functions respectively of the standard logistic distribution. When W is the standardized t, then $W = \sqrt{\frac{\nu-2}{\nu}}t_\nu$, where t_ν is the student's distribution with ν degrees of freedom. The VQC function in this case is

$$\Upsilon(p) = F_\nu\left(\frac{\mu_1-\mu_2}{\sigma_2}\sqrt{\frac{\nu}{\nu-2}} + \frac{\sigma_1}{\sigma_2}Q_\nu(p)\right),$$

where $F_\nu(.)$ and $Q_\nu(.)$ are the distribution and quantile functions respectively of t_ν.

The simulations were carried out for sample sizes $n_1 = n_2 = 25, 50, 75$ and 100. The MISE was based on 5,000 replications. The results are reported in Table 1. In all cases, the proposed estimator offered a relative reduction between 6% and 25% over the nonparametric estimator.

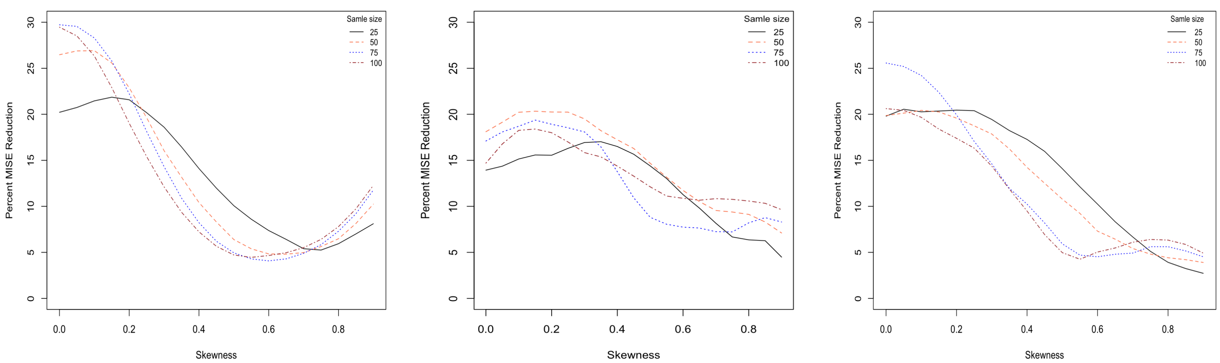
Table 1: Percent reduction in MISE of the proposed estimator relative to the Li et al. (1996) estimator

		$\boldsymbol{\theta} = (\mu_1, \sigma_1, \mu_2, \sigma_2)^\top$		
Baseline distribution	$\boldsymbol{n_1 = n_2}$	$(-0.5, 1, 3, 1.5)^\top$	$(5, 1, 2, 2)^\top$	$(1, 3.5, 0, 3)^\top$
Gaussian	25	13.76	12.25	20.98
	50	15.00	13.13	24.34
	75	15.56	13.41	24.47
	100	16.45	14.47	24.81
Logistic	25	19.84	10.01	6.99
	50	20.23	10.05	9.32
	75	18.91	12.97	6.07
	100	21.52	13.26	7.86
		$\nu = 5$	$\nu = 9$	$\nu = 13$
t_ν	25	15.98	8.49	13.00
	50	12.13	11.86	15.30
	75	11.21	11.93	15.32
	100	13.15	10.99	15.55

A study of the proposed estimator's performance was conducted when the location and scale assumption was violated. Let W indicate a generic base distribution such as standard normal, a standardized logistic, or a standardized t_ν, where ν is the degrees of freedom. Note that W has mean 0 and variance 1. Let W_1, W_2 and W_3 be independent random variables each having the same distribution as W. We considered $X_1 = \mu_1 + \sigma_1 W_1$ and $X_2 = \mu_2 + \sigma_2 Y$, where $Y = \delta|W_2| + (1-\delta^2)^{1/2} W_3$. Here, the second sample is drawn from a skewed distribution that matches the first sample distribution only when δ, the skewness parameter, is 0. Note that $F_2(t) = P(X_2 \le t) = F_Y((t-\mu_2)/\sigma_2)$. Let $Q_1(p)$ be the quantile function of X_1. Then $Q_1(p) = \mu_1 + \sigma_1 Q_{W_1}(p)$. After some calculations, the true VQC function is $\Upsilon(p) = F_2(Q_1(p)) = F_Y((Q_1(p) - \mu_2)/\sigma_2)$, where

$$F_Y(y) = \int_{-\infty}^{\infty} F_W\left(\frac{y - \delta|x|}{\sqrt{1-\delta^2}}\right) f_W(x)dx, \qquad y \in \mathbb{R}^1. \tag{3.1}$$

When δ is 0, X_1 and X_2 belong to an LS family. Substituting $\delta = 0$ on the right hand side of Eq. (3.1), $F_Y(y) = F_W(y)$. In particular, when $\delta = 0$, $\Upsilon(p) = F_W(Q_1(p)) = F_W(\mu_1 + \sigma_1 Q_W(p))$. When δ is away from 0, the two samples are not from an LS family. We, however, continue to pool both samples and compute the proposed estimator, which continues to outperform the nonparametric estimator significantly. The percent reduction in MISE of the proposed estimator relative to the nonparametric estimator is shown in Figure 1.



(a) Normal $\boldsymbol{\theta} = (0, 1, 0, 1)^\top$ (b) t_4 $\boldsymbol{\theta} = (-0.5, 1, 3, 1.5)^\top$ (c) Logistic $\boldsymbol{\theta} = (-0.5, 1, 3, 1.5)^\top$

Figure 1: Robustness study for some distributions

The empirical coverage probability (ECP) of the 95% pointwise confidence intervals for $\Upsilon(p)$ is the proportion of 1,000 intervals that include $\Upsilon(p)$. They were computed from the asymptotic normality of $\hat{\Upsilon}(p)$. The standard error of $\hat{\Upsilon}(p)$ was estimated by $B = 1,000$ bootstrap samples. The ECPs are reported in Table 2.

Table 2: Empirical coverage probability of 95% confidence intervals for $\Upsilon(p)$ for $\boldsymbol{\theta} = (0, 1, 0, 1)^\top$

		p				
Baseline distribution	$\boldsymbol{n_1 = n_2}$	.35	.45	.50	.65	.75
Gaussian	25	.920	.928	.932	.930	.923
	50	.936	.947	.944	.942	.938
	75	.951	.956	.959	.937	.947
	100	.954	.954	.942	.948	.940
Logistic	25	.936	.930	.939	.931	.920
	50	.938	.940	.944	.938	.938
	75	.944	.948	.944	.940	.945
	100	.954	.945	.948	.944	.950

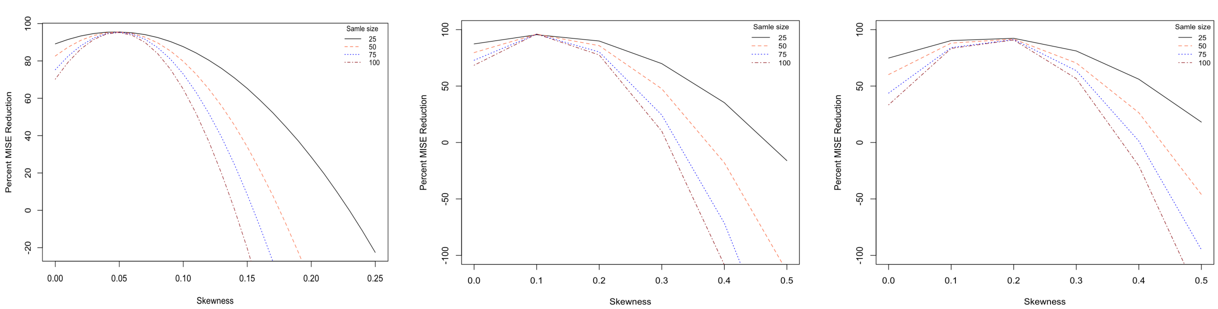
3.2 Numerical results for the censored case

The three scenarios that we considered for the uncensored case were replicated here. Failure times were either normal or logistic or t_ν. The censoring was taken as $C_i = (\mu_i + k_1) + (\sigma_i + k_2)W_i$, where W_i were independent standard normal, standardized logistic, or standardized t_ν. When $W_i \equiv Z_i$ were standard normal, $k_1 = 3$ and $k_2 = 3$. When W_i were standardized logistic, $k_1 = 2$ and $k_2 = 2$. When W_i were standardized t_ν, $k_1 = 3$ and $k_2 = 1$. The percentage reduction in MISE relative to the nonparametric estimator are reported in Table 3. The proposed estimator gave improved estimates for mild to moderate censoring rates.

Table 3: Percentage reduction in MISE of the proposed estimator relative to the Li et al. (1996) estimator

		$\boldsymbol{\theta} = (\mu_1, \sigma_1, \mu_2, \sigma_2)^\top$		
Baseline distribution	$\boldsymbol{n_1 = n_2}$	$(-1, 1, 2, 2)^\top$	$(0, 1, 0, 1)^\top$	$(-2, 9, 5, 3)^\top$
	25	53.71	88.87	36.50
	50	48.83	82.96	38.29
Gaussian	75	41.51	76.44	35.96
	100	31.69	70.29	33.22
Censoring Rate		(10%, 21%)	(10%, 10%)	(41%, 28%)
	25	51.01	74.75	50.90
	50	46.02	58.46	48.94
Logistic	75	40.93	42.93	46.92
	100	35.86	31.66	48.13
Censoring Rate		(18%, 30%)	(18%, 18%)	(44%, 35%)
		$\nu = 9$	$\nu = 13$	$\nu = 17$
	25	49.41	87.23	32.92
	50	48.18	80.13	23.68
t_ν	75	46.30	73.34	15.97
	100	41.56	64.72	18.49
Censoring Rate		(20%, 28%)	(19%, 19%)	(44%, 35%)

For the robustness study, similar to the uncensored case, we considered $X_1 = \mu_1 + \sigma_1 W_1$ and $X_2 = \mu_2 + \sigma_2 Y$, where $Y = \delta|W_2| + (1-\delta^2)^{1/2} W_3$. In each scenario, the censoring distributions were like the ones specified in the first study of this subsection, see above. The results are reported in Figure 2. As in the uncensored case, here too the proposed estimator is robust to departures from the LS assumption.



(a) Normal $\boldsymbol{\theta} = (0, 1, 0, 1)^\top$ (b) t_{13} $\boldsymbol{\theta} = (0, 1, 0, 1)^\top$ (c) Logistic $\boldsymbol{\theta} = (0, 1, 0, 1)^\top$

Figure 2: Censored robustness study of proposed estimator for different models

The ECPs of the 95% confidence intervals for $\Upsilon(p)$ are reported in Table 4. The failure and censoring time distributions were as in the preceding simulation studies in this subsection.

Table 4: Empirical coverage probability of 95% confidence intervals for $\Upsilon(p)$ for $\boldsymbol{\theta} = (-0.5, 1, 3, 1.5)^\top$

Baseline distribution	$n_1 = n_2$	p .35	.45	.50	.65	.75
Gaussian	25	.937	.942	.961	.954	.953
	50	.948	.941	.955	.964	.973
	75	.934	.941	.957	.976	.956
	100	.937	.942	.958	.964	.964
t_{15}	25	.921	.926	.936	.935	.935
	50	.923	.933	.928	.947	.938
	75	.930	.941	.937	.932	.938
	100	.943	.946	.955	.949	.942

4 Conclusion

The intuitive idea that pooling the two standardized samples would increase precision turbocharged our desire to develop the proposed semiparametric estimator. The numerical studies indicate that it should provide a robust alternative to the touchstone nonparametric estimator even in cases when the data do not subscribe to an LS framework. Although the idea is easily entertained, setting up the rigorous mathematical framework, as attempted in this paper, required considerable effort. In particular, we were able to obtain desirable convergence rates for various oscillation moduli by framing the standardization as a projection, which then allowed us to exploit a result derived by Stute and Zhu (2005). The convergence rates were instrumental in killing remainder terms. Using the functional delta method we derived an asymptotic representation for the centered VQC function process, which we then tailored for the cases of no-censoring and censoring. In both cases, however, the representations are unwieldy and the weak limits, although Gaussian, are not distribution free. The pointwise confidence intervals for the VQC function returned ECPs that were close to the nominal 95%. The proposed estimator showed an attractive disposition of performing well under misspecifications.

The paper provides a basic building block that we believe would undergird the setting up of a host of allied procedures such as simultaneous confidence bands (SCBs) for the VQC function under an LS framework, and the provision of a powerful model diagnostic test for checking the adequacy of the LS assumption. Judged on the basis of the numerical studies in this paper, these procedures are expected to offer significant improvements over the existing ones. Work on these is in progress and will be reported as soon as they are completed.

A Appendix

A.1 Local oscillations of distribution and quantile functions estimators

Recall that $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$, where $\boldsymbol{\theta}_i = (\mu_i, \sigma_i)^\top, i = 1, 2$; and that $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \boldsymbol{\gamma}_2^\top)^\top$ is in a neighborhood of $\boldsymbol{\theta}$, where $\boldsymbol{\gamma}_i = (\mu_i', \sigma_i')^\top$. Recall that $F_{\boldsymbol{\gamma}}$ is the df of $W_{\boldsymbol{\gamma}}$, see Eq. (2.1). Note that $\hat{F}_{\boldsymbol{\gamma}}$ is the (notional) EDF of $W_{ij}^{\gamma_i} = (X_{ij} - \mu_i)/\sigma_i, j = 1, \ldots, n_i, i = 1, 2$. We write $\hat{F}_{\boldsymbol{\theta}} = \hat{F}$ and $F_{\boldsymbol{\theta}} = F$.

For the local oscillations of the empirical process $(s, \boldsymbol{\gamma}) \rightarrow \hat{F}_{\boldsymbol{\gamma}}(s) - F_{\boldsymbol{\gamma}}(s)$ in a neighborhood of $\boldsymbol{\theta}$, let $\hat{\mathbb{H}}_{\boldsymbol{\gamma}}(s,t) = \hat{F}_{\boldsymbol{\gamma}}(s) - F_{\boldsymbol{\gamma}}(s) - \hat{F}(t) + F(t)$ for $\boldsymbol{\gamma}$ in a neighborhood of $\boldsymbol{\theta}$, and $s \in \mathbb{R}, t \in \mathbb{R}$ satisfying (i) $\|\boldsymbol{\gamma} - \boldsymbol{\theta}\| = O(n^{-1/2})$ and (ii) $|s - t| = O(n^{-1/2+\alpha})$, where $\alpha < 1/2$. In Lemma 3, we recast Lemma 4.2 of Stute and Zhu (2005), which they noted is a modification of Theorem 37 of Pollard (1984).

Lemma 3. *Under conditions* **C1** $-$ **C3** *and conditions (i) and (ii) above, we have for uncensored data that*

$$\sup_{s,t,\boldsymbol{\gamma}} \left| \hat{\mathbb{H}}_{\boldsymbol{\gamma}}(s,t) \right| = o_{\mathbb{P}}\left(n^{-1/2}\right). \tag{A.1}$$

Proof To prove Eq. (A.1), let $\boldsymbol{\eta} = (1/\sigma_1', -\mu_1'/\sigma_1', 1/\sigma_2', -\mu_2'/\sigma_2')^\top$ and let $\boldsymbol{\eta}_0 = (1/\sigma_1, -\mu_1/\sigma_1, 1/\sigma_2, -\mu_2/\sigma_2)^\top$. The random variable $W_{\boldsymbol{\gamma}}$ defined by Eq. (2.1), can be expressed as the projection $\boldsymbol{\eta}^\top \boldsymbol{U}$, where $\boldsymbol{U}$ is the vector

$(X_1\xi, \xi, X_2(1-\xi), (1-\xi))^\top$. Note that $\mathbb{P}(\boldsymbol{\eta}_0^\top \boldsymbol{U} \le t) = F(t)$. Let s_i^* be between s and $\sigma_i^{-1}\{(\mu_i' - \mu_i) + \sigma_i' s\}$. Since $F_{\boldsymbol{\gamma}}$ is the distribution function of $W_{\boldsymbol{\gamma}}$, we have that $F_{\boldsymbol{\gamma}}(s) = \mathbb{P}(\boldsymbol{\eta}^\top \boldsymbol{U} \le s)$. As in the proof of Lemma 1,

$$\begin{aligned}
F_{\boldsymbol{\gamma}}(s) &= \mathbb{P}\left(\frac{X_1 - \mu_1'}{\sigma_1'}\xi + \frac{X_2 - \mu_2'}{\sigma_2'}(1-\xi) \le s\right) \\
&= F\left(\frac{\mu_1' - \mu_1}{\sigma_1} + \frac{\sigma_1'}{\sigma_1}s\right)\rho + F\left(\frac{\mu_2' - \mu_2}{\sigma_2} + \frac{\sigma_2'}{\sigma_2}s\right)(1-\rho) \\
&= F(s) + \frac{\rho}{\sigma_1}\{(\mu_1' - \mu_1) + (\sigma_1' - \sigma_1)s\} f(s_1^*) + \frac{1-\rho}{\sigma_2}\{(\mu_2' - \mu_2) + (\sigma_2' - \sigma_2)s\} f(s_2^*) \\
&= F(s) + O(n^{-1/2}).
\end{aligned}$$

Note that $F(t) - F(s) = (t-s)f(\tilde{t})$ where $\tilde{t}$ lies between t and s. Therefore $F(t) - F(s) = O(n^{-1/2+\alpha})$. Following the proof of Lemma 4.2 of Stute and Zhu (2005), the half spaces $(-\infty, x]$ form a class with a polynomial covering number. From the above calculations, the maximal measure of the sets included in the class is

$$\mathbb{P}(\boldsymbol{\eta}_0^\top \boldsymbol{U} \le t) - \mathbb{P}(\boldsymbol{\eta}^\top \boldsymbol{U} \le s) \;=\; F(t) - F_{\boldsymbol{\gamma}}(s) \;=\; F(t) - F(s) + O(n^{-1/2}) \;=\; O(n^{-1/2+\alpha}),$$

where $\alpha = 1/\zeta$, and reference to ζ can be found in condition **C1**. The rest of the proof follows as in Stute and Zhu (2005), giving the in-probability bound $O(\delta_n^2 \alpha_n)$, where $\alpha_n^2 = \log n/(n\delta_n^2)$ and $\delta_n^2 = O(n^{-1/2+\alpha})$. It follows that $\sup_{s,t,\boldsymbol{\gamma}} \left|\hat{\mathbb{H}}_{\boldsymbol{\gamma}}(s,t)\right| = O_{\mathbb{P}}\left(\sqrt{n^{-3/2+\alpha}\log n}\right)$, from which Eq. (A.1) follows. □

Remark For censored data, after some elementary calculations, $\hat{H}_{\boldsymbol{\gamma}}(s,t)$ can be expressed as a functional of $\hat{\mathbb{G}}_{\boldsymbol{\gamma}}(s,t) = \hat{H}_{\boldsymbol{\gamma}}(s) - H_{\boldsymbol{\gamma}}(s) - \hat{H}(t) + H(t)$ and $\hat{\mathbb{G}}_{\boldsymbol{\gamma},1}(s,t) = \hat{H}_{\boldsymbol{\gamma},1}(s) - H_{\boldsymbol{\gamma},1}(s) - \hat{H}_1(t) + H_1(t)$, where $H_{\boldsymbol{\gamma}}$ is the distribution function of $(Z_i - \mu_i')/\sigma_i'$ and $H_{\boldsymbol{\gamma},1}$ is the subdistribution function of the $(Z_i - \mu_i')/\sigma_i'$ that are uncensored. By the methods of Lemma 3, it follows that $\sup_{s,t,\boldsymbol{\gamma}} \left|\hat{\mathbb{G}}_{\boldsymbol{\gamma}}(s,t)\right| = o_{\mathbb{P}}\left(n^{-1/2}\right)$ and $\sup_{s,t,\boldsymbol{\gamma}} \left|\hat{\mathbb{G}}_{\boldsymbol{\gamma},1}(s,t)\right| = o_{\mathbb{P}}\left(n^{-1/2}\right)$. Then, it can be shown that Lemma 3 applies for the censored case as well. The details are cumbersome, so we omit them.

Lemma 4. *Let $\Delta_{\boldsymbol{\gamma}}(p) = \hat{Q}_{\boldsymbol{\gamma}}(p) - Q_{\boldsymbol{\gamma}}(p) - \hat{Q}(p) + Q(p)$. Under conditions* **C1–C4***, and (i) and (ii) above,*

$$\sup_{p\in[\alpha,\beta],\boldsymbol{\gamma}} |\Delta_{\boldsymbol{\gamma}}(p)| = o_{\mathbb{P}}(n^{-1/2}). \tag{A.2}$$

Proof For each $p \in [\alpha, \beta]$ and each $\boldsymbol{\gamma} \in \Gamma$, consider the expression $D_{\boldsymbol{\gamma}}(p)$ [cf. Eq. (10), Lo and Singh, 1986]

$$D_{\boldsymbol{\gamma}}(p) = \hat{Q}_{\boldsymbol{\gamma}}(p) - Q_{\boldsymbol{\gamma}}(p) + Q_{\boldsymbol{\gamma}}(\hat{F}_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p))) - Q_{\boldsymbol{\gamma}}(F_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p))).$$

Let $R_{\boldsymbol{\gamma}}(p) = Q(\hat{F}(Q(p))) - Q(F(Q(p))) - Q_{\boldsymbol{\gamma}}(\hat{F}_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p))) + Q_{\boldsymbol{\gamma}}(F_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p)))$. Then $\Delta_{\boldsymbol{\gamma}}(p) = D_{\boldsymbol{\gamma}}(p) - D_{\boldsymbol{\theta}}(p) + R_{\boldsymbol{\gamma}}(p)$. We will show that each is bounded above in absolute value by $\sup_{s,t,\boldsymbol{\gamma}} |\hat{\mathbb{H}}_{\boldsymbol{\gamma}}(s,t)|$. Following Lo and Singh (1986), we can write $D_{\boldsymbol{\gamma}}(p) = D_{\boldsymbol{\gamma}}^{(1)}(p) + D_{\boldsymbol{\gamma}}^{(2)}(p)$, where for each $\boldsymbol{\gamma}$,

$$\begin{aligned}
D_{\boldsymbol{\gamma}}^{(1)}(p) &= Q_{\boldsymbol{\gamma}}(F_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p))) - Q_{\boldsymbol{\gamma}}(\hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p))) + Q_{\boldsymbol{\gamma}}(\hat{F}_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p))) - Q_{\boldsymbol{\gamma}}(F_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p))), \\
D_{\boldsymbol{\gamma}}^{(2)}(p) &= Q_{\boldsymbol{\gamma}}(\hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p))) - Q_{\boldsymbol{\gamma}}(p).
\end{aligned}$$

All terms of $D_{\boldsymbol{\gamma}}^{(1)}(p)$ are operated by the composition mappings $Q_{\boldsymbol{\gamma}} \circ \hat{F}_{\boldsymbol{\gamma}}$ or $Q_{\boldsymbol{\gamma}} \circ F_{\boldsymbol{\gamma}}$, which act on $\hat{Q}_{\boldsymbol{\gamma}}(p)$ or $Q_{\boldsymbol{\gamma}}(p)$. By Lemma 3 of Lo and Singh (1986) $\sup_{p\in[\alpha,\beta]} |\hat{Q}_{\boldsymbol{\gamma}}(p) - Q_{\boldsymbol{\gamma}}(p)| = O(n^{-1/2}(\log n)^{1/2})$ a.s. $= o(n^{-1/2+\alpha})$ a.s., where $0 < \alpha < 1/2$. Apply the mean value theorem to the two segments of $D_{\boldsymbol{\gamma}}^{(1)}(p)$ to obtain

$$D_{\boldsymbol{\gamma}}^{(1)}(p) = \frac{F_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p)) - \hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p))}{f_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p^*))} + \frac{\hat{F}_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p)) - F_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p))}{f_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p^{**})}, \tag{A.3}$$

where p^* lies between $\hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p))$ and $F_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p))$, and p^{**} lies between $\hat{F}_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p))$ and $F_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p))$. We show that we can replace the denominators of the two ratios in Eq. (A.3) by $f_{\boldsymbol{\gamma}}(Q(p^*))$ and $f_{\boldsymbol{\gamma}}(Q(p^{**}))$ respectively. This is because the conditions on $f_{\boldsymbol{\gamma}}$ and its derivative, along with **Remark 1** following Lemma 2 imply that, for instance, the difference $d = f_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p^*)) - f_{\boldsymbol{\gamma}}(Q(p^*))$ is $O(n^{-1/2})$ uniformly over $\boldsymbol{\gamma}$ in the $n^{-1/2}$ neighborhood of $\boldsymbol{\theta}$. Then we add and subtract $\hat{F}(\cdot)$ and $F(\cdot)$ to the numerator of the ratio that occurs in the remainder term. With K denoting a generic constant, it follows that $\mathbb{R}_n$, the remainder term, is

$$\mathbb{R}_n \leq K \left\{ \sup_{s,t,\boldsymbol{\gamma}} |\hat{\mathbb{H}}_{\boldsymbol{\gamma}}(s,t)| + O\left(n^{-1/2}(\log n)^{1/2}\right) \right\} \times O(n^{-1/2}), \tag{A.4}$$

the upper bound free of $\boldsymbol{\gamma}$. Likewise for the second ratio. Let $K = \inf_{p\in[\alpha,\beta],\boldsymbol{\gamma}} f_{\boldsymbol{\gamma}}(Q(p))$. It follows that

$$\begin{aligned} \sup_{p\in[\alpha,\beta]} |D^{(1)}_{\boldsymbol{\gamma}}(p)| &\leq \frac{1}{K} \sup_{x\in[Q_{\boldsymbol{\gamma}}(\alpha),Q_{\boldsymbol{\gamma}}(\beta)],|x-y|\leq cn^{-1/2+\alpha}} |\hat{F}_{\boldsymbol{\gamma}}(x) - F_{\boldsymbol{\gamma}}(x) - \hat{F}_{\boldsymbol{\gamma}}(y) + F_{\boldsymbol{\gamma}}(y)| + \mathbb{R}_n \\ &= O_{\mathbb{P}}\left(\sup_{s,t,\boldsymbol{\gamma}} \left|\hat{\mathbb{H}}_{\boldsymbol{\gamma}}(s,t)\right|\right), \end{aligned}$$

a uniform bound for all $\boldsymbol{\gamma}$. By similar techniques, we can show that $\sup_{p\in[\alpha,\beta]} |R_{\boldsymbol{\gamma}}(p)| = O_{\mathbb{P}}\left(\sup_{s,t,\boldsymbol{\gamma}} \left|\hat{\mathbb{H}}_{\boldsymbol{\gamma}}(s,t)\right|\right)$. To handle $D^{(2)}_{\boldsymbol{\gamma}}(p)$, we apply the mean value theorem and obtain

$$D^{(2)}_{\boldsymbol{\gamma}}(p) = \frac{\hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p) - p}{f_{\boldsymbol{\gamma}}(Q_{\boldsymbol{\gamma}}(p^*))}, \tag{A.5}$$

where p^* is bracketed between $\hat{Q}_{\boldsymbol{\gamma}}(p)$ and p. For the numerator of the ratio in Eq. (A.5), we use an inequality that appears in the proof of Theorem 2 of Lo and Singh (1986), which is

$$\begin{aligned} |\hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p))) - p| &\leq |\hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p)) - \hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p-))| \\ &= |\hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p)) - F_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p)) + F_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p-)) - \hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p-))|. \end{aligned} \tag{A.6}$$

We replace the denominator of the ratio in Eq. (A.5) by $f_{\boldsymbol{\gamma}}(Q(p^*))$. From Eq. (A.5) and Eq. (A.6), therefore

$$|D^{(2)}_{\boldsymbol{\gamma}}(p)| \leq \frac{1}{f_{\boldsymbol{\gamma}}(Q(p^*))} |\hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p)) - F_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p)) + F_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p-)) - \hat{F}_{\boldsymbol{\gamma}}(\hat{Q}_{\boldsymbol{\gamma}}(p-))| + \mathbb{R}_n,$$

where $\mathbb{R}_n$ admits a rate like Eq. (A.4). It follows as before that $\sup_{p\in[\alpha,\beta]} |D^{(2)}_{\boldsymbol{\gamma}}(p)| = O_{\mathbb{P}}\left(\sup_{s,t,\boldsymbol{\gamma}} \left|\hat{\mathbb{H}}_{\boldsymbol{\gamma}}(s,t)\right|\right)$. Therefore, $\sup_{p\in[\alpha,\beta],\boldsymbol{\gamma}} |\Delta_{\boldsymbol{\gamma}}(p)| = O_{\mathbb{P}}\left(\sqrt{n^{-3/2+\alpha}\log n}\right)$, from which Eq. (A.2) follows. □

A.2 Uncensored case details

Recall that $n_1/n \to \kappa$ as $n_1 \to \infty$ and $n_2 \to \infty, \kappa_1 = \kappa = 1 - \kappa_2$. Since $W_{ij} = (X_{ij} - \mu_i)/\sigma_i$, we have

$$\begin{aligned} n^{1/2}(\hat{\mu}_i - \mu_i) &= \left(\frac{n}{n_i}\right)^{1/2} \sigma_i \times \left(n_i^{-1/2} \sum_{j=1}^{n_i} W_{ij}\right) \\ &= \kappa_i^{-1/2} \sigma_i \times \left(n_i^{-1/2} \sum_{j=1}^{n_i} W_{ij}\right) + o_{\mathbb{P}}(1). \end{aligned} \tag{A.7}$$

Because $\hat{\mu}_i - \mu_i = O_{\mathbb{P}}(n_i^{-1/2})$, and $\sum_{i=1}^{n_j} W_{ij} = o_{\mathbb{P}}(n_i)$, we have

$$\begin{aligned}\hat{\sigma}_i^2 &= \frac{1}{n_i}\sum_{j=1}^{n_i}(X_{ij}-\hat{\mu}_i)^2 = \frac{1}{n_i}\sum_{j=1}^{n_i}(X_{ij}-\mu_i+\mu_i-\hat{\mu}_i)^2 \\ &= \sigma_i^2\frac{1}{n_i}\sum_{j=1}^{n_i}W_{ij}^2 - 2\sigma_i(\hat{\mu}_i-\mu_i)\frac{1}{n_i}\sum_{j=1}^{n_i}W_{ij} + (\mu_i-\hat{\mu}_i)^2 \\ &= \sigma_i^2\frac{1}{n_i}\sum_{j=1}^{n_i}W_{ij}^2 + o_{\mathbb{P}}(n_i^{-1/2}). \end{aligned} \tag{A.8}$$

Applying the delta method, it is easy to see from Eq. (A.8) that

$$\hat{\sigma}_i - \sigma_i = \frac{1}{2\sigma_i}(\hat{\sigma}_i^2-\sigma_i^2) + o_{\mathbb{P}}(n_i^{-1/2}) = \frac{\sigma_i}{2}\times\frac{1}{n_i}\sum_{j=1}^{n_i}(W_{ij}^2-1) + o_{\mathbb{P}}(n_i^{-1/2}).$$

Then, it follows that

$$n^{1/2}(\hat{\sigma}_i - \sigma_i) = \kappa_i^{-1/2}\,\frac{\sigma_i}{2}\left(n_i^{-1/2}\sum_{j=1}^{n_i}(W_{ij}^2-1)\right) + o_{\mathbb{P}}(1). \tag{A.9}$$

By Corollary 21.5 of van der Vaart (1998), if F is differentiable at $Q(p)$ with positive derivative $f(Q(p))$, then

$$n^{1/2}\left(\hat{Q}(p) - Q(p)\right) = -n^{-1/2}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\frac{1_{\{W_{ij}\le Q(p)\}} - p}{f(Q(p))} + o_{\mathbb{P}}(1). \tag{A.10}$$

Finally, it is easy to see that

$$n^{1/2}\left(\hat{F}(s) - F(s)\right) = n^{-1/2}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\{1_{\{W_{ij}\le s\}} - F(s)\} + o_{\mathbb{P}}(1). \tag{A.11}$$

We apply the representations given by Eq. (A.7) and Eqs. (A.9)–(A.11) to the RHS of Eq. (2.12) to obtain Eq. (2.15).

A.3 Censored case details

We first obtain representations for each of the centered quantities on the RHS of Eq. (2.12). From Major and Rejtö (1988) and the functional delta method (Theorem II.8.1. of Andersen et al. (1993), it follows that

$$n^{1/2}(\hat{F}(x) - F(x)) = S(x)\,n^{-1/2}\sum_{l=1}^{n} I_l^{(1)}(x) + o_{\mathbb{P}}(1), \tag{A.12}$$

where

$$I_l^{(1)}(x) = \frac{I(W_l\le x, \delta_l = 1)}{\mathbb{P}(W > W_l)} - \int_0^x \frac{I(W_l > y)d\Lambda(y)}{\mathbb{P}(W>y)}. \tag{A.13}$$

By Proposition II.8.4 of Andersen et al. (1993), $\phi(F) = Q(p)$ is (tangentially) compactly differentiable at F with derivative $d\phi(F)\cdot h = -h(Q(p))/f(Q(p))$. Equivalently, $n^{1/2}(\hat{Q}(p) - Q(p))$ is asymptotically equivalent to $(1-p)\,d\phi(F)\cdot n^{1/2}(\hat{F}-F)$. Applying the operator $d\phi(F)$ on $h = n^{1/2}(\hat{F}-F)$, Eq. (A.12) implies that

$$n^{1/2}(\hat{Q}(p) - Q(p)) = -\frac{(1-p)}{f(Q(p))}\,n^{-1/2}\sum_{l=1}^{n} I_l^{(1)}(Q(p)) + o_{\mathbb{P}}(1). \tag{A.14}$$

For $i=1,2$, let $\hat{S}_i(\cdot)$ be the KM estimator of $S_i(x)=P(X_i>x)$ and let $\hat{F}_i=1-\hat{S}_i$. Let $\nu_i=\inf\{u: H_i(u)=1\}\le\infty$, where H_i is the distribution function of $Z_{il}, l=1,\ldots,n_i$. Let A_{H_i} be the set of all atoms of H_i, which may possibly be empty. For any measurable function φ with $\int\varphi\,dF_i<\infty$, with probability 1 and in the mean, $\int\varphi\,d\hat{F}_i$ converges to $\int\varphi\,d\tilde{F}_i$, where $\tilde{F}_i(x)=F_i(x)I(x<\nu_i)+\{F_i(\nu_i-)+I(\nu_i\in A_{H_i})F(\nu_i)\}I(x\ge\nu_i)$ (Stute and Wang, 1993). It will be assumed that $\nu_i<\infty$ and that ν_i is not an atom of F_i (Stute and Wang, 1993; Stute, 1995). Then, $\tilde{F}_i=F_i$. When $\varphi(x)$ is x or x^2, one obtains the estimators of the first two moments.

Let $T_{i1}<\ldots<T_{im_i}$ be the distinct uncensored Z_{ij}'s. Let d_{il} and r_{il} be the number of observed failures at T_{il} and the number of Z_{ij} that are greater than or equal to T_{il}. Note that $\Delta\hat{S}_i(T_{il})\equiv\hat{S}_i(T_{il})-\hat{S}_i(T_{i(l-1)})=-\hat{S}_i(T_{i(l-1)})d_{il}/r_{il}$, and $\hat{S}_i(T_{i0})\equiv 1$. The KM integral estimates of μ_i and ψ_i, the second moment of F_i, are

$$\hat{\mu}_i \;=\; -\sum_{l=1}^{m_i}T_{il}\Delta\hat{S}_i(T_{il});\qquad \hat{\psi}_i \;=\; -\sum_{l=1}^{m_i}T_{il}^2\Delta\hat{S}_i(T_{il}). \tag{A.15}$$

Let $C_{G_i}(t)=\mathcal{D}_t(\Lambda_{G_i},1-H_i)$, where Λ_{G_i} is the CHF associated with G_i. Stute (1995) proved that

$$\int\varphi(x)d\hat{F}_i(x) \;=\; \frac{1}{n_i}\sum_{j=1}^{n_i}\{\varphi(Z_{ij})\beta_{i0}(Z_{ij})\delta_{ij}+\beta_{i1}(Z_{ij})(1-\delta_{ij})-\beta_{i2}(Z_{ij})\}+o_{\mathbb{P}}(n_i^{-1/2}), \tag{A.16}$$

where $\beta_{im}(x), m=0,1,2$, are

$$\begin{aligned}\beta_{i0}(x) &= \frac{1}{1-G_i(x)};\\ \beta_{i1}(x) &= \frac{1}{1-H_i(x)}\int_x^\infty\varphi(w)dF_i(w);\\ \beta_{i2}(x) &= \int_{-\infty}^\infty\int_{-\infty}^{x\wedge w}\frac{d\Lambda_{G_i}(v)}{1-H_i(v)}\varphi(w)dF_i(w) \;=\; \int_{-\infty}^\infty C_{G_i}(x\wedge w)\varphi(w)dF_i(w).\end{aligned}$$

The expectations of the second and third summands in Eq. (A.16) cancel out (Stute, 1995). The first summand in Eq. (A.16) has expectation equal to $\int\varphi dF_i$. When $\varphi(x)=x$, $\int\varphi dF_i=\mu_i$. We obtain from Eq. (A.16) that

$$n^{1/2}(\hat{\mu}_i-\mu_i) \;=\; \kappa_i^{-1/2}\left(n_i^{-1/2}\sum_{j=1}^{n_i}I_{ij}^{(2)}\right)+o_{\mathbb{P}}(1), \tag{A.17}$$

where, replacing $\varphi(w)$ with w in the definition of $\beta_{i1}(x)$ and $\beta_{i2}(x)$ above, we obtain the influence function

$$I_{ij}^{(2)} \;=\; \{Z_{ij}\beta_{i0}(Z_{ij})\delta_{ij}-\mu_i\}+\frac{1-\delta_{ij}}{1-H_i(Z_{ij})}\int_{Z_{ij}}^\infty w\,dF_i(w)-\int_{-\infty}^\infty C_{G_i}(Z_{ij}\wedge w)w\,dF_i(w). \tag{A.18}$$

Note that $\mathbb{E}(I_{ij}^{(2)})=0$ for $j=1,\ldots,n_i$ and $i=1,2$. The estimate of σ_i will be through the first two KM integral estimates of μ_i and ψ_i. Accordingly, if we let $\varphi(w)=w^2$ then $\int\varphi(w)dF_i=\psi_i$, and from Eq. (A.16),

$$n^{1/2}(\hat{\psi}_i-\psi_i) \;=\; \kappa_i^{-1/2}\left(n_i^{-1/2}\sum_{j=1}^{n_i}I_{ij}^{(3)}\right)+o_{\mathbb{P}}(1), \tag{A.19}$$

where, replacing $\varphi(w)$ with w^2 in the definition of $\beta_{i1}(x)$ and $\beta_{i2}(x)$ above, we obtain the influence function

$$I_{ij}^{(3)} \;=\; \left\{Z_{ij}^2\beta_{i0}(Z_{ij})\delta_{ij}-\psi_i\right\}+\frac{1-\delta_{ij}}{1-H_i(Z_{ij})}\int_{Z_{ij}}^\infty w^2\,dF_i(w)-\int_{-\infty}^\infty C_{G_i}(Z_{ij}\wedge w)w^2\,dF_i(w). \tag{A.20}$$

Note that $\mathbb{E}(I_{ij}^{(3)})=0$ for $j=1,\ldots,n_i$ and $i=1,2$.

Applying the delta method yields

$$\begin{aligned} n_i^{1/2}\left(\hat{\sigma}_i - \sigma_i\right) &= \frac{1}{2\sigma_i}\, n_i^{1/2}\left(\hat{\sigma}_i^2 - \sigma_i^2\right) + o_{\mathbb{P}}(1) \\ &= \frac{1}{2\sigma_i}\, n_i^{1/2}\left[(\hat{\psi}_i - \psi_i) - 2\mu_i(\hat{\mu}_i - \mu_i)\right] + o_{\mathbb{P}}(1). \end{aligned} \tag{A.21}$$

Eq. (A.21) indicates that, for $n^{1/2}(\hat{\mu}_1 - \mu_1)$, the scaling factor computed from Eq. (2.12) will be modified to

$$\left[\frac{\rho}{\sigma_1} + \frac{1-\rho}{\sigma_2} - \frac{\mu_1}{\sigma_1}.\frac{1}{\sigma_2}\left\{\frac{\mu_1 - \mu_2}{\sigma_1}\rho + Q(p)\right\}\right] f(C(p)).$$

For $n^{1/2}(\hat{\mu}_2 - \mu_2)$, the modified scaling factor is

$$-\frac{\sigma_1}{\sigma_2}\left[\frac{\rho}{\sigma_1} + \frac{1-\rho}{\sigma_2} - \frac{\mu_2}{\sigma_2}.\frac{1}{\sigma_2}\left\{\frac{\mu_1 - \mu_2}{\sigma_1}\rho + Q(p)\right\}\right] f(C(p)).$$

For $n^{1/2}(\hat{\psi}_1 - \psi_1)$, the modified scaling factor is

$$\frac{1}{2\sigma_1\sigma_2}\left\{\frac{\mu_1 - \mu_2}{\sigma_1}\rho + Q(p)\right\} f(C(p)).$$

For $n^{1/2}(\hat{\psi}_2 - \psi_2)$, the modified scaling factor is

$$-\frac{\sigma_1}{2\sigma_2^3}\left\{\frac{\mu_1 - \mu_2}{\sigma_1}\rho + Q(p)\right\} f(C(p)).$$

Apply the representations in Eqs. (A.12), (A.14), (A.17), and (A.19) to the RHS of Eq. (2.12) to get Eq. (2.18).

References

Andersen, P. K., Borgan, O., Gill, R. D., Keiding, N. (1993). *Statistical models based on counting processes.* Springer, New York.

Chernozhukov, V., Fernandez-Val, I., Melly, B. (2013). Inference on counterfactual distributions. *Econometrica.* **81** 2205—2268.

Kalbfleisch, J. D., and Prentice, R. L. (2002). *The statistical analysis of failure time data.* 2nd Edition, John Wiley and Sons, New York.

Li, G., Tiwari, R. C., and Wells, M. D. (1996). Quantile comparison functions in two-sample problems, with application to comparisons of diagnostic markers. *J. Amer. Statist. Assoc.* **91** 689–698.

Lo, SH. and Singh, K. (1986). The product-limit estimator and the bootstrap: some asymptotic representations. *Probab. Th. Rel. Fields* **71** 455–465. https://doi.org/10.1007/BF01000216

Major, P. and Rejtö, L. (1988). Strong embedding of the estimator of the distribution function under random censorship. *Ann. Statist,* **16** 1113–1132.

Pollard, D. (1984). *Convergence of stochastic processes.* Springer-Verlag, New York.

Stute, W. (1995). The central limit theorem under random censorship, *Ann. Statist,* **23** 422–439.

Stute, W. and Wang, J.-L. (1993). A strong law under random censorship. *Ann. Statist.* **21** 1591–1607.

Stute, W. and Zhu, L.-X. (2005). Nonparametric checks for single-index models. *Ann. Statist.* **33** 1048–1083.

van der Vaart, A.W. (1998). *Asymptotic statistics.* Cambridge University Press, New York. https://doi.org/10.1017/CBO9780511802256