Conditional Quantile Correlation Learning for Ultrahigh Dimensional Varying Coefficient Models and Its Application in Survival Analysis

Xiaochao Xia\textsuperscript{1}, Jialiang Li\textsuperscript{2,3,4}, Bo Fu\textsuperscript{5,6}

\textsuperscript{1}Huazhong Agricultural University
\textsuperscript{2}National University of Singapore, \textsuperscript{3}Duke-NUS Graduate Medical School, \textsuperscript{4}Singapore Eye Research Institute
\textsuperscript{5}Fudan University, \textsuperscript{6}University College London

Supplementary Material

In this supplementary file, we will present some additional simulation examples for illustration and give the detailed proofs to Theorems 2.1-2.3. The proofs of Theorems 3.1-3.3 follow similar lines to Theorems 2.1-2.3.

S1 Additional Simulations and Results

In addition to Examples 1 in the main paper, we provide some additional examples below.

\textit{Example 2}. We generate response data as

\[ H(Y) = 2T \cdot X_1 + (2T - 1)^2 \cdot X_2 + \sin(2\pi T) \cdot X_3 + \varepsilon, \]
where we consider three monotone transform functions as: (T1) $H(y) = \log(0.5(e^{2y} - 1))$; (T2) $H(y) = \log(2y)$; and (T3) $H(y) = (2y)^{1/3}$. The rest of the model setup are the same as Example 1. Here we consider two types for model error distributions:

- **Case (2a):** the same as Case (1a), i.e., $\varepsilon \sim N(0, 1)$;
- **Case (2b):** the same as Case (1b), i.e., $\varepsilon \sim C(0, 1)$.

The corresponding results are reported in Table 1. From the tabled results we can draw a similar conclusion to that for Example 1. Our method performs more satisfactorily than the NIS. This example illustrates the invariance of the CQCSIS under monotone transformation on the response variable.

<table>
<thead>
<tr>
<th>$H(y)$</th>
<th>Case</th>
<th>Method($\tau$)</th>
<th>$s_n$</th>
<th>$\theta = 0$</th>
<th>$\theta = 0.4$</th>
<th>$\theta = 0.8$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MMS</td>
<td>RSD</td>
<td>PS</td>
</tr>
<tr>
<td>(T1)</td>
<td>(2a)</td>
<td>CQCSIS(0.50)</td>
<td>3</td>
<td>9</td>
<td>17</td>
<td>0.850</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CQCSIS(0.75)</td>
<td>3</td>
<td>21</td>
<td>42</td>
<td>0.750</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NIS</td>
<td>3</td>
<td>5</td>
<td>12</td>
<td>0.900</td>
</tr>
<tr>
<td>(2b)</td>
<td></td>
<td>CQCSIS(0.50)</td>
<td>3</td>
<td>36</td>
<td>115</td>
<td>0.580</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CQCSIS(0.75)</td>
<td>3</td>
<td>132</td>
<td>272</td>
<td>0.375</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NIS</td>
<td>3</td>
<td>587</td>
<td>374</td>
<td>0.190</td>
</tr>
<tr>
<td>(T2)</td>
<td>(2a)</td>
<td>CQCSIS(0.50)</td>
<td>3</td>
<td>7</td>
<td>20</td>
<td>0.830</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CQCSIS(0.75)</td>
<td>3</td>
<td>15</td>
<td>63</td>
<td>0.710</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NIS</td>
<td>3</td>
<td>270</td>
<td>296</td>
<td>0.170</td>
</tr>
<tr>
<td>(2b)</td>
<td></td>
<td>CQCSIS(0.50)</td>
<td>3</td>
<td>63</td>
<td>132</td>
<td>0.515</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CQCSIS(0.75)</td>
<td>3</td>
<td>158</td>
<td>271</td>
<td>0.290</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NIS</td>
<td>3</td>
<td>643</td>
<td>631</td>
<td>0.270</td>
</tr>
<tr>
<td>(T3)</td>
<td>(2a)</td>
<td>CQCSIS(0.50)</td>
<td>3</td>
<td>8</td>
<td>28</td>
<td>0.845</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CQCSIS(0.75)</td>
<td>3</td>
<td>21</td>
<td>55</td>
<td>0.720</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NIS</td>
<td>3</td>
<td>68</td>
<td>114</td>
<td>0.480</td>
</tr>
<tr>
<td>(2b)</td>
<td></td>
<td>CQCSIS(0.50)</td>
<td>3</td>
<td>46</td>
<td>117</td>
<td>0.605</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CQCSIS(0.75)</td>
<td>3</td>
<td>142</td>
<td>249</td>
<td>0.370</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NIS</td>
<td>3</td>
<td>820</td>
<td>170</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 1: Results of the median of minimum model size (MMS), its robust standard deviation (RSD) and the proportion of truly active covariates selected (PS) with a pre-specified threshold size $d_n = \lfloor n / \log n \rfloor$ for Example 2.

**Example 3.** This example is modified from Example 3 in [Fan et al. (2011)](Fan et al., 2011). Specifi-
CQCSIS and its application in survival analysis

ically, let \( \{W_1, \ldots, W_p\} \) be independent and standard normal random variables, and \( \{U_1, U_2\} \) be independent random variables with unit uniform distribution. We generate response variable from the following model:

\[
Y = 2X_1 + 3TX_2 + (T + 1)^2X_3 + \frac{4\sin(2\pi T)}{2 - \sin(2\pi T)}X_4 + \exp(T)X_5 + 2\sqrt{T}X_6 + \varepsilon,
\]

where \( X_j = (W_j + t_1U_1)/(1 + t_1) \), \( j = 1, \ldots, p \), and \( T = (U_2 + t_2U_1)/(1 + t_2) \), implying that \( \text{Corr}(X_j, X_k) = t_2^2/(12 + t_1^2) \) for \( j \neq k \) and \( \text{Corr}(X_j, T) = t_1t_2/((12 + t_1^2)(1 + t_2^2))^{1/2} \) independent of \( j \). Regarding the distribution of error \( \varepsilon \), we consider four scenarios:

- **Case (3a):** the same as Case (1a), i.e., \( \varepsilon \sim N(0, 1) \);
- **Case (3b):** the same as Case (1b), i.e., \( \varepsilon \sim C(0, 1) \);
- **Case (3c):** the error follows the scaled Cauchy distribution, i.e.,
  \[
  \varepsilon = 0.5(\frac{\exp(T)}{1+\exp(T)})X_7 + (2T - 1)^2X_8 + \cos(2\pi T)X_9 + \sqrt{T + 1}X_{10})C(0, 1);
  \]
- **Case (3d):** the error follows a scaled Chi-squared distribution, that is
  \[
  \varepsilon = 0.5(\frac{\exp(T)}{1+\exp(T)})X_7 + (2T - 1)^2X_8 + \cos(2\pi T)X_9 + \sqrt{T + 1}X_{10}) \cdot (\varepsilon - Q_{\varepsilon, r}) \text{ with } \varepsilon \sim \chi^2(1).
  \]

The results are summarized in Table 2. Eyeballing the table, we observe that CQCSIS and NIS perform similarly as in previous examples.

**Example 4.** We consider survival time data in a setting where the proportional hazard assumption is not satisfied. Specifically, we generate the response variable

\[
Y = 2X_1 + 3TX_2 + (T + 1)^2X_3 + \frac{4\sin(2\pi T)}{2 - \sin(2\pi T)}X_4 + \varepsilon,
\]

where \( X_j \)'s and \( T \) are generated in the
same way as Example 3 with \((t_1, t_2) = (0, 0)\). Let \(\tilde{Y} = \min(Y, Z)\), where censoring time \(Z\) is simulated from a mixture of normal distributions: (S1) \(0.2N(-5, 4) + 0.1N(5, 1) + 0.7N(55, 1)\) and (S2) \(0.4N(-5, 4) + 0.1N(5, 1) + 0.5N(55, 1)\). We consider two scenarios for the distribution of error \(\varepsilon\):

- Case (4a): the same as Case (3a), i.e., \(\varepsilon \sim N(0, 1)\);

- Case (4b): the same as Case (3b), i.e., \(\varepsilon \sim C(0, 1)\).

Note that in scenarios (S1) and (S2), the censoring rates are roughly 20% and 35%, respectively.

**Example 5.** We consider an alternative setting where the censoring depends on covariate. The model setup is the same as that in Example 4 only except that the censoring time \(Z\) is generated from the covariate dependent mixture distribution: (S3) \(0.4N(-5, 4) + 0.1N(5, 1) + 0.5N(55, 1)\) if \(X_1 > 0\) and \(0.5N(-5, 4) + 0.2N(5, 1) +

<table>
<thead>
<tr>
<th>Case</th>
<th>Method((\tau))</th>
<th>(s_n)</th>
<th>((t_1, t_2) = (0, 0))</th>
<th>((t_1, t_2) = (1, 0))</th>
<th>((t_1, t_2) = (1, 1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3a)</td>
<td>CQCSIS(0.50)</td>
<td>6</td>
<td>7 4 0.940</td>
<td>6 6 0.950</td>
<td>26 32 0.795</td>
</tr>
<tr>
<td></td>
<td>CQCSIS(0.75)</td>
<td>6</td>
<td>13 13 0.895</td>
<td>13 19 0.905</td>
<td>38 41 0.700</td>
</tr>
<tr>
<td></td>
<td>NIS</td>
<td>6</td>
<td>6 0 1.000</td>
<td>6 1 1.000</td>
<td>6 0 1.000</td>
</tr>
<tr>
<td>(3b)</td>
<td>CQCSIS(0.50)</td>
<td>6</td>
<td>14 19 0.850</td>
<td>26 40 0.750</td>
<td>64 85 0.505</td>
</tr>
<tr>
<td></td>
<td>CQCSIS(0.75)</td>
<td>6</td>
<td>33 53 0.685</td>
<td>73 111 0.485</td>
<td>152 176 0.255</td>
</tr>
<tr>
<td></td>
<td>NIS</td>
<td>6</td>
<td>655 367 0.050</td>
<td>770 236 0.005</td>
<td>790 228 0.000</td>
</tr>
<tr>
<td>(3c)</td>
<td>CQCSIS(0.50)</td>
<td>10</td>
<td>823 156 0.000</td>
<td>813 215 0.000</td>
<td>843 160 0.000</td>
</tr>
<tr>
<td></td>
<td>CQCSIS(0.75)</td>
<td>10</td>
<td>832 164 0.000</td>
<td>837 148 0.000</td>
<td>844 147 0.000</td>
</tr>
<tr>
<td></td>
<td>NIS</td>
<td>10</td>
<td>828 183 0.000</td>
<td>867 129 0.000</td>
<td>855 152 0.000</td>
</tr>
<tr>
<td>(3d)</td>
<td>CQCSIS(0.50)</td>
<td>6</td>
<td>11 18 0.850</td>
<td>10 11 0.905</td>
<td>26 42 0.720</td>
</tr>
<tr>
<td></td>
<td>CQCSIS(0.75)</td>
<td>10</td>
<td>815 177 0.000</td>
<td>871 137 0.000</td>
<td>883 147 0.000</td>
</tr>
</tbody>
</table>

Table 2: Results of the median of minimum model size (MMS), its robust standard deviation (RSD) and the proportion of truly active covariates selected (PS) with a pre-specified threshold size \(d_n = \lfloor n / \log n \rfloor\) for Example 3.
$0.3N(55, 1)$ if $X_1 < 0$. The censoring rates are about 40%. Error distributions the same as cases (4a) and (4b).

In both examples, we apply the proposed CQCSIS method at the median and the first quartile, respectively. Here, we compare our censored version of the CQCSIS (denoted by the CQCSIScens) in Section 4 with four existing approaches: the CQCSISnaive, under which we treat the censored observations as complete observations and then apply the CQCSIS, the quantile-adaptive screening (QaSIS) by He et al. (2013) which treats varying coefficients as constants, the conditional quantile sure independent screening (CQSIS) by Wu and Yin (2015) as well as the censored rank sure independent screening (CRSIS) by Song et al. (2014a). In earlier publications all these approaches were shown to be quite robust when the proportional hazards assumption is violated. Other screening methods in the survival literature are known to be less effective and thus are not chosen for comparison in this paper.

Table 3 presents the simulation results. Eyeballing the table, we can see that CQCSIScens performs much better than all other procedures. In this case, QaSIS and CQSIS are unsatisfactory because they fail to acknowledge the varying coefficients. In addition, the results of Example 5 displayed in Table 3 indicate that our proposed CQCSIScens also performs fairly well for various model errors even though the censoring mechanism is covariate-dependent.

Finally, in order to evaluate the performance of proposed two-stage variable selec-
Table 3: Results of the median of minimum model size (MMS), its robust standard deviation (RSD) and the proportion of truly active covariates selected (PS) with a pre-specified threshold size $d_{n} = \lfloor n / \log n \rfloor$ for Examples 4 and 5.
tion procedures, we consider one more example.

**Example 6** We generate the response variable from the following model

\[ Y = 2T X_1 + 5(2T - 1)^2 X_2 + 3 \sin(2\pi T) X_3 + \varepsilon, \]

where \( \varepsilon \) is considered from following two cases:

- Case (6a) (Normal): \( \varepsilon \sim 0.8N(0,1); \)

- Case (6b) (Cauchy): \( \varepsilon \sim \frac{2}{15} C(0,1). \)

The remaining of the model setup is the same as Example 1.

For Examples 6, we first conduct a variable screening procedure using the proposed CQCSIS method and then we carry out SCAD-penalized variable selection procedure. For this example, we compare the group SCAD-penalized quantile regression method with \( \tau = 0.5 \) stated in Section 4, abbreviated as RQSCAD(0.5), with the group SCAD-penalized mean regression method of Fan et al. (2011), abbreviated as LSSCAD.

We consider two cases for screening tuning parameter: \( d_n = 10 \) and 20. For each situation, we report: (C) the average of true positives (the average number of true nonzero coefficients that are correctly estimated to be nonzero), (I) the average of false positives (the average number of true zero coefficients that are incorrectly estimated to be nonzero), (CF) the proportion of correctly fitted models, (UF) the proportion of under-fitted models, (OF) the proportion of over-fitted models, and (MADE) the
median of absolute deviation error (ADE) over \( N \) simulations, where ADE is defined as

\[
ADE(\hat{\beta}) = \frac{1}{n_{\text{grid}}} \sum_{j=1}^{p} \sum_{l=1}^{n_{\text{grid}}} |\hat{\beta}_j(t_l) - \beta_j(t_l)|,
\]

where the \( t_i \)'s are the equally spaced grid points on the support of \( T \) with \( n_{\text{grid}} = 100 \).

Tables 4 and 5 summarize the simulated results. Our method performs competitively with the LSSCAD method, and behaves much better for (6b) Cauchy error. This is anticipated since the LSSCAD is based on mean regression model and sensitive to heavy-tailed errors. We note that a larger screening threshold parameter gives higher proportion of correctly fitted models and larger number of correctly identified nonzero coefficients with a slight loss in estimation accuracy for non-vanishing varying coefficients. Obviously, the increase of sample size significantly improves the consistency of the variable selection and the accuracy of estimated functional coefficients.

Example 7 The data generating process of this example is the same as that in Example 1. The only difference is that we here consider four distinct combinations of dimensionality \( p \) and sample size \( n \) to see the performance affected by choosing different \((n, p)\). The screening results are summarized in Table 6. From Table 6, we can see that in normal error setting, the NIS performs better as expected, however, in Cauchy error setting the proposed CQCSIS is more desirable than NIS. This finding is the same as before. In addition, an increase of \( p \) generally makes all the screening methods harder.
<table>
<thead>
<tr>
<th>Case</th>
<th>Method (τ)</th>
<th>No. of estimated nonzeros</th>
<th>Proportion of fitted models</th>
<th>MADE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>I</td>
<td>CF</td>
</tr>
<tr>
<td>(6a)</td>
<td>RQSCAD(0.5)</td>
<td>2.86(0.00)</td>
<td>0.03(0.00)</td>
<td>0.91(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>2.86(0.00)</td>
<td>0.10(0.00)</td>
<td>0.89(0.00)</td>
</tr>
<tr>
<td>(6b)</td>
<td>RQSCAD(0.5)</td>
<td>2.85(0.00)</td>
<td>0.04(0.00)</td>
<td>0.89(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>2.57(0.00)</td>
<td>0.60(0.75)</td>
<td>0.51(0.75)</td>
</tr>
<tr>
<td>(6a)</td>
<td>RQSCAD(0.5)</td>
<td>2.93(0.00)</td>
<td>0.00(0.00)</td>
<td>0.96(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>2.95(0.00)</td>
<td>0.08(0.00)</td>
<td>0.93(0.00)</td>
</tr>
<tr>
<td>(6b)</td>
<td>RQSCAD(0.5)</td>
<td>2.88(0.00)</td>
<td>0.01(0.00)</td>
<td>0.92(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>2.07(1.49)</td>
<td>0.52(0.75)</td>
<td>0.30(0.75)</td>
</tr>
<tr>
<td>(6a)</td>
<td>RQSCAD(0.5)</td>
<td>2.98(0.00)</td>
<td>0.01(0.00)</td>
<td>0.98(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>2.99(0.00)</td>
<td>0.07(0.00)</td>
<td>0.96(0.00)</td>
</tr>
<tr>
<td>(6b)</td>
<td>RQSCAD(0.5)</td>
<td>2.88(0.00)</td>
<td>0.01(0.00)</td>
<td>0.93(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>2.58(0.00)</td>
<td>1.44(1.49)</td>
<td>0.39(0.75)</td>
</tr>
<tr>
<td>(6a)</td>
<td>RQSCAD(0.5)</td>
<td>2.82(0.00)</td>
<td>0.10(0.00)</td>
<td>0.80(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>3.00(0.00)</td>
<td>0.13(0.00)</td>
<td>0.90(0.00)</td>
</tr>
<tr>
<td>(6b)</td>
<td>RQSCAD(0.5)</td>
<td>2.76(0.00)</td>
<td>0.01(0.00)</td>
<td>0.84(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>1.98(1.49)</td>
<td>1.46(1.49)</td>
<td>0.17(0.00)</td>
</tr>
</tbody>
</table>

**Table 4:** Simulated results for two-stage approaches for Example 6, where \( n = 200 \) and robust standard deviation is given in parenthesis.

<table>
<thead>
<tr>
<th>Case</th>
<th>Method (τ)</th>
<th>No. of estimated nonzeros</th>
<th>Proportion of fitted models</th>
<th>MADE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>I</td>
<td>CF</td>
</tr>
<tr>
<td>(6a)</td>
<td>RQSCAD(0.5)</td>
<td>3.00(0.00)</td>
<td>0.01(0.00)</td>
<td>1.00(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>3.00(0.00)</td>
<td>0.02(0.00)</td>
<td>0.99(0.00)</td>
</tr>
<tr>
<td>(6b)</td>
<td>RQSCAD(0.5)</td>
<td>2.96(0.00)</td>
<td>0.06(0.00)</td>
<td>0.97(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>2.58(0.00)</td>
<td>0.51(0.75)</td>
<td>0.56(0.75)</td>
</tr>
<tr>
<td>(6a)</td>
<td>RQSCAD(0.5)</td>
<td>3.00(0.00)</td>
<td>0.01(0.00)</td>
<td>1.00(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>3.00(0.00)</td>
<td>0.01(0.00)</td>
<td>0.99(0.00)</td>
</tr>
<tr>
<td>(6b)</td>
<td>RQSCAD(0.5)</td>
<td>2.96(0.00)</td>
<td>0.03(0.00)</td>
<td>0.98(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>2.05(1.49)</td>
<td>0.37(0.75)</td>
<td>0.35(0.75)</td>
</tr>
<tr>
<td>(6a)</td>
<td>RQSCAD(0.5)</td>
<td>3.00(0.00)</td>
<td>0.00(0.00)</td>
<td>1.00(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>3.00(0.00)</td>
<td>0.06(0.00)</td>
<td>0.96(0.00)</td>
</tr>
<tr>
<td>(6b)</td>
<td>RQSCAD(0.5)</td>
<td>2.98(0.00)</td>
<td>0.01(0.00)</td>
<td>0.99(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>2.63(0.00)</td>
<td>1.12(1.49)</td>
<td>0.48(0.75)</td>
</tr>
<tr>
<td>(6a)</td>
<td>RQSCAD(0.5)</td>
<td>3.00(0.00)</td>
<td>0.00(0.00)</td>
<td>1.00(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>3.00(0.00)</td>
<td>0.06(0.00)</td>
<td>0.96(0.00)</td>
</tr>
<tr>
<td>(6b)</td>
<td>RQSCAD(0.5)</td>
<td>2.94(0.00)</td>
<td>0.01(0.00)</td>
<td>0.96(0.00)</td>
</tr>
<tr>
<td></td>
<td>LSSCAD</td>
<td>2.08(0.93)</td>
<td>1.22(1.49)</td>
<td>0.21(0.00)</td>
</tr>
</tbody>
</table>

**Table 5:** Simulated results for two-stage approaches for Example 6, where \( n = 400 \) and robust standard deviation is given in parenthesis.
<table>
<thead>
<tr>
<th>Case</th>
<th>MMS (n, p = (100, 1000))</th>
<th>RSD (n, p = (100, 1000))</th>
<th>PS (n, p = (100, 1000))</th>
<th>MMS (n, p = (200, 1000))</th>
<th>RSD (n, p = (200, 1000))</th>
<th>PS (n, p = (200, 1000))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1a)</td>
<td>CQCIS(0.5)</td>
<td>11.5 8.2 0.808</td>
<td>19 11.2 0.595</td>
<td>4 1.5 0.990</td>
<td>8 4.5 0.965</td>
<td>8 4.5 0.965</td>
</tr>
<tr>
<td></td>
<td>CQCIS(0.75)</td>
<td>13 9.9 0.735</td>
<td>20.5 12.9 0.520</td>
<td>4 1.5 0.990</td>
<td>8 4.5 0.980</td>
<td>8 4.5 0.980</td>
</tr>
<tr>
<td></td>
<td>NIS</td>
<td>3 0.0 1.000</td>
<td>3 0.0 1.000</td>
<td>3 0.0 1.000</td>
<td>3 0.0 1.000</td>
<td></td>
</tr>
<tr>
<td>(1b)</td>
<td>CQCIS(0.5)</td>
<td>11 7.5 0.815</td>
<td>21 11.2 0.530</td>
<td>4 1.5 0.995</td>
<td>8 4.7 0.985</td>
<td>8 4.7 0.985</td>
</tr>
<tr>
<td></td>
<td>CQCIS(0.75)</td>
<td>14 9.7 0.735</td>
<td>26.5 24.3 0.395</td>
<td>4 1.5 0.990</td>
<td>8.5 5.2 0.950</td>
<td>8.5 5.2 0.950</td>
</tr>
<tr>
<td></td>
<td>NIS</td>
<td>3 31.5 0.730</td>
<td>3 82.6 0.665</td>
<td>3 31.5 0.730</td>
<td>3 82.6 0.665</td>
<td></td>
</tr>
<tr>
<td>(1c)</td>
<td>CQCIS(0.5)</td>
<td>66.5 121.3 0.145</td>
<td>169 383.6 0.020</td>
<td>13 7.8 0.790</td>
<td>20 11.4 0.570</td>
<td>20 11.4 0.570</td>
</tr>
<tr>
<td></td>
<td>CQCIS(0.75)</td>
<td>14 10.1 0.735</td>
<td>27 23.5 0.385</td>
<td>68.5 136.4 0.180</td>
<td>154.5 251.3 0.060</td>
<td>154.5 251.3 0.060</td>
</tr>
<tr>
<td></td>
<td>NIS</td>
<td>3 0.0 1.000</td>
<td>3 0.0 1.000</td>
<td>3 82.6 0.665</td>
<td>3 82.6 0.665</td>
<td></td>
</tr>
<tr>
<td>(1d)</td>
<td>CQCIS(0.5)</td>
<td>13 7.8 0.790</td>
<td>20 14.4 0.570</td>
<td>5 0.0 0.980</td>
<td>5 0.7 0.925</td>
<td>5 0.7 0.925</td>
</tr>
<tr>
<td></td>
<td>CQCIS(0.75)</td>
<td>68.5 136.4 0.180</td>
<td>154.5 251.3 0.060</td>
<td>4 1.5 0.995</td>
<td>33 48.5 0.570</td>
<td>33 48.5 0.570</td>
</tr>
<tr>
<td></td>
<td>NIS</td>
<td>12 84.5 0.575</td>
<td>15.5 339.0 0.535</td>
<td>11 9.7 0.875</td>
<td>6 134.7 0.675</td>
<td>6 134.7 0.675</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5 45.1 0.710</td>
<td>5 134.7 0.675</td>
<td>5 134.7 0.675</td>
</tr>
</tbody>
</table>

Table 6: Simulation results for the Example 1 with $\rho = 0.8$ and different combinations of $(n, p)$.
Example 8 As pointed out by one referee, how CQCSIS can be affected if \( \tau \) varies. To answer this question, in this example we revisit Example 3 with a range of \( \tau \) and \((t_1, t_2) = (0, 0)\). Specifically, we take nine different values of \( \tau \) (= 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9) to see the performance of CQCSIS at various quantile levels. In addition, we also consider a composite CQCSIS, denoted as CCQCSIS below, that uses the average of CQC at a grid of these given quantile levels as the screening utility as in Ma and Zhang (2016), which is defined by

\[
u_j = E\{E[p_{\tau,j}^2(T)]\} = \int_0^1 E[p_{\tau,j}^2(T)]d\tau.
\]

An empirical utility is defined by \(\hat{u}_j = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{n} \sum_{i=1}^{n} \hat{p}_{\tau_k,j}^2(T_i)\) with \(\tau_k = \frac{k-0.5}{N}\) for a large \(N\), for example, we take \(N = 100\) in the simulation. Table 7 reports the results, which indicate that CQCSIS performs satisfactorily across quantile levels in the middle range of the overall interval, while a bit poor at quite low and high quantile levels. Also we can see that the performance of CCQCSIS is quite appealing. Theoretically speaking, the screening property of CCQCSIS can be similarly proved without much difficulty.

Example 9 This example is designed as same as Example 4 with censoring rate of 20%. We here consider four cases of \((n, p) = (100, 1000), (100, 2000), (200, 2000), (200, 5000)\). The results are summarized in Tables 8 and 9. We can make a similar observation as in Example 7.
### Table 7: Simulated results for Example 8.

<table>
<thead>
<tr>
<th></th>
<th>Normal Error</th>
<th>Cauchy Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MMS</td>
<td>RSD</td>
</tr>
<tr>
<td>CQCSIS(0.1)</td>
<td>37.0</td>
<td>119.2</td>
</tr>
<tr>
<td>CQCSIS(0.2)</td>
<td>7.0</td>
<td>7.6</td>
</tr>
<tr>
<td>CQCSIS(0.3)</td>
<td>5.0</td>
<td>1.5</td>
</tr>
<tr>
<td>CQCSIS(0.4)</td>
<td>5.0</td>
<td>0.7</td>
</tr>
<tr>
<td>CQCSIS(0.5)</td>
<td>5.0</td>
<td>1.5</td>
</tr>
<tr>
<td>CQCSIS(0.6)</td>
<td>5.0</td>
<td>1.5</td>
</tr>
<tr>
<td>CQCSIS(0.7)</td>
<td>5.0</td>
<td>1.5</td>
</tr>
<tr>
<td>CQCSIS(0.8)</td>
<td>6.0</td>
<td>2.4</td>
</tr>
<tr>
<td>CQCSIS(0.9)</td>
<td>12.0</td>
<td>20.7</td>
</tr>
<tr>
<td>CCQCSIS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NIS</td>
<td>5.0</td>
<td>0.7</td>
</tr>
</tbody>
</table>

### Table 8: Simulation results for the Example 9 with different combinations of \((n, p)\).

<table>
<thead>
<tr>
<th></th>
<th>MMS</th>
<th>RSD</th>
<th>PS</th>
<th>MMS</th>
<th>RSD</th>
<th>PS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n, p) = (100, 1000)</td>
<td>(n, p) = (100, 2000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CQCISCens(0.5)</td>
<td>142</td>
<td>157</td>
<td>0.015</td>
<td>310</td>
<td>373</td>
<td>0.000</td>
</tr>
<tr>
<td>CQCISCens(0.25)</td>
<td>238</td>
<td>246</td>
<td>0.000</td>
<td>427</td>
<td>420</td>
<td>0.000</td>
</tr>
<tr>
<td>CQCISNaive(0.5)</td>
<td>277</td>
<td>200</td>
<td>0.005</td>
<td>629</td>
<td>533</td>
<td>0.000</td>
</tr>
<tr>
<td>CQCISNaive(0.25)</td>
<td>523</td>
<td>298</td>
<td>0.000</td>
<td>984</td>
<td>546</td>
<td>0.000</td>
</tr>
<tr>
<td>QaSISCens(0.5)</td>
<td>277</td>
<td>252</td>
<td>0.005</td>
<td>546</td>
<td>566</td>
<td>0.000</td>
</tr>
<tr>
<td>QaSISCens(0.25)</td>
<td>319</td>
<td>296</td>
<td>0.000</td>
<td>635</td>
<td>641</td>
<td>0.000</td>
</tr>
<tr>
<td>CQCSISCens(0.5)</td>
<td>450</td>
<td>375</td>
<td>0.020</td>
<td>731</td>
<td>735</td>
<td>0.005</td>
</tr>
<tr>
<td>CQCSISCens(0.25)</td>
<td>415</td>
<td>347</td>
<td>0.010</td>
<td>837</td>
<td>621</td>
<td>0.000</td>
</tr>
<tr>
<td>CRSIS</td>
<td>711</td>
<td>240</td>
<td>0.000</td>
<td>1407</td>
<td>503</td>
<td>0.000</td>
</tr>
</tbody>
</table>

|                | (n, p) = (200, 2000) | (n, p) = (200, 5000) |
| CQCISCens(0.5) | 24  | 32  | 0.645 | 56  | 73  | 0.360 |
| CQCISNaive(0.25)| 40  | 51  | 0.480 | 82  | 157 | 0.255 |
| CQCISNaive(0.5)| 103 | 151 | 0.200 | 223 | 347 | 0.070 |
| CQCSISCens(0.25)| 459 | 500 | 0.025 | 836 | 1163| 0.010 |
| QaSISCens(0.5) | 288 | 508 | 0.135 | 638 | 1084| 0.015 |
| QaSISCens(0.25)| 375 | 552 | 0.065 | 699 | 1003| 0.010 |
| CQCSISCens(0.5)| 254 | 423 | 0.190 | 566 | 1188| 0.060 |
| CQCSISCens(0.25)| 318 | 520 | 0.135 | 607 | 1165| 0.060 |
| CRSIS          | 1246| 457 | 0.000 | 3389| 1116| 0.000 |
Appendix: proofs

Before proving the main results, we introduce some notation first. For \( i = 1, \ldots, n \) and \( j = 1, \ldots, p \), we denote

\[
\begin{align*}
f_{1,ij} &= I(Y_i - Q_{\tau,Y} > 0)X_{ij},
f_{2,ij} &= I(Y_i - Q_{\tau,Y} > 0),
f_{3,ij} &= X_{ij}^2,
f_{4,ij} &= X_{ij},
f_{w,ij} &= w_i(F)I(\tilde{Y}_i < Q_{\tau,Y})X_{ij},
f_{2w,ij} &= w_i(F)I(\tilde{Y}_i < Q_{\tau,Y}) \quad \text{and} \quad f_{3w,ij} = w_i^2(F)I(\tilde{Y}_i < Q_{\tau,Y}).
\end{align*}
\]

### Appendix A: Some Properties of B-spline

According to the properties of normalized B-splines, we have that for each \( j = 1, \cdots, p \) and \( k = 1, \ldots, L_n \), (i) \( B_k(t) \geq 0 \) and \( \sum_{k=1}^{L_n} B_k(t) = 1 \) for \( t \in \mathcal{T} \); (ii) there exist positive

<table>
<thead>
<tr>
<th>S1 censoring with Cauchy Error case</th>
<th>MMS</th>
<th>RSD</th>
<th>PS</th>
<th>MMS</th>
<th>RSD</th>
<th>PS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n, p) = (100, 1000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CQCISCens(0.5)</td>
<td>207</td>
<td>205</td>
<td>0.005</td>
<td>517</td>
<td>471</td>
<td>0.000</td>
</tr>
<tr>
<td>CQCISCens(0.25)</td>
<td>379</td>
<td>272</td>
<td>0.000</td>
<td>804</td>
<td>594</td>
<td>0.000</td>
</tr>
<tr>
<td>CQCISNaive(0.5)</td>
<td>431</td>
<td>289</td>
<td>0.000</td>
<td>778</td>
<td>611</td>
<td>0.000</td>
</tr>
<tr>
<td>CQCISNaive(0.25)</td>
<td>583</td>
<td>271</td>
<td>0.000</td>
<td>1137</td>
<td>611</td>
<td>0.000</td>
</tr>
<tr>
<td>QaSISCens(0.5)</td>
<td>421</td>
<td>262</td>
<td>0.000</td>
<td>764</td>
<td>605</td>
<td>0.000</td>
</tr>
<tr>
<td>QaSISCens(0.25)</td>
<td>412</td>
<td>340</td>
<td>0.000</td>
<td>940</td>
<td>597</td>
<td>0.000</td>
</tr>
<tr>
<td>CQSISCens(0.5)</td>
<td>498</td>
<td>384</td>
<td>0.000</td>
<td>1039</td>
<td>703</td>
<td>0.000</td>
</tr>
<tr>
<td>CQSISCens(0.25)</td>
<td>555</td>
<td>350</td>
<td>0.000</td>
<td>1134</td>
<td>641</td>
<td>0.000</td>
</tr>
<tr>
<td>CRSIS</td>
<td>780</td>
<td>190</td>
<td>0.000</td>
<td>1581</td>
<td>385</td>
<td>0.000</td>
</tr>
</tbody>
</table>

| (n, p) = (200, 2000)              |      |      |      |      |      |      |
| CQCISCens(0.5)                    | 64   | 118  | 0.370| 147  | 265.1| 0.145|
| CQCISCens(0.25)                   | 161  | 174  | 0.105| 352  | 474.3| 0.035|
| CQCISNaive(0.5)                   | 259  | 299  | 0.065| 587  | 813.8| 0.010|
| CQCISNaive(0.25)                  | 643  | 397  | 0.010| 1579 | 1456.3| 0.000|
| QaSISCens(0.5)                    | 433  | 467  | 0.010| 940  | 1145.9| 0.000|
| QaSISCens(0.25)                   | 525  | 574  | 0.005| 1335 | 1549.8| 0.000|
| CQSISCens(0.5)                    | 500  | 569  | 0.045| 1297 | 1776.9| 0.020|
| CQSISCens(0.25)                   | 646  | 646  | 0.010| 1738 | 1743.5| 0.000|
| CRSIS                             | 1537 | 373  | 0.000| 3900 | 1047.4| 0.000|

Table 9: Simulation results for the Example 9 with different combinations of \((n, p)\).
constants $C_1, C_2$ such that for any $a \in \mathbb{R}^L$,

$$\frac{C_1}{L_n} a'a \leq \int a'B(t)(B(t))'adt \leq \frac{C_2}{L_n} a'a.$$ 

Under property (ii), we have that: (iii) there exist positive constants $C_3$ and $C_4$ such that for $k = 1, \ldots, L_n$,

$$C_3 L_n^{-1} \leq \mathbb{E} \{ (B_k(T))^2 \} \leq C_4 L_n^{-1},$$

where $C_3 = C_1 M_1$ and $C_4 = C_2 M_2$; and (iv)

$$C_3 L_n^{-1} \leq \lambda_{\min}(\mathbb{E} \{ B(T)(B(T))' \}) \leq \lambda_{\max}(\mathbb{E} \{ B(T)(B(T))' \}) \leq C_4 L_n^{-1},$$

Appendix B: Some Technical Lemmas

**Lemma 1.** Suppose that random variable $X$ has a conditional exponential tail $P(|X| > x|T) \leq K_1 \exp(-K_2^{-1}x)$ for positive constants $K_1, K_2$, uniformly on the support of $T$. Then, for all $r \geq 2$, $\mathbb{E}(|X|^r|T) \leq K_1 K_2^r r!$.

**Proof of Lemma 1** By the condition on $X$ and a change of variable, we have

$$\mathbb{E} \{ |X|^r | T \} = \int_0^\infty P(|X| > x^{1/r})dx \leq K_1 \int_0^\infty \exp(-K_2^{-1}x^{1/r})dx$$

$$= r K_1 K_2^r \int_0^\infty \exp(-t)r^{r-1}dt = r K_1 K_2^r \Gamma(r) = K_1 K_2^r r!.$$
Hence, the lemma follows. \( \square \)

According to this lemma, it is easily seen that \( \mathbb{E}\{|X|^2|T\} \) is finite.

**Lemma 2.** (*Bernstein’s inequality, Lemma 2.2.11, van der Vaart and Wellner, 1996*)

For independent random variables \( Y_1, \ldots, Y_n \) with mean zero and \( \mathbb{E}\{|Y_i|^r\} \leq r!K^{r-2}v_i/2 \) for every \( r \geq 2, i = 1, \ldots, n \) and some constants \( K, v_i \). Then, for \( x > 0 \), we have

\[
P(\left|Y_1 + \cdots + Y_n\right| > x) \leq 2 \exp\left(-\frac{x^2}{2(v + Kx)}\right),
\]

for \( v \geq \sum_{i=1}^{n} v_i \).

**Lemma 3.** (*Bernstein’s inequality, Lemma 2.2.9, van der Vaart and Wellner, 1996*)

For independent random variables \( Y_1, \ldots, Y_n \) with mean zero and bounded range \([-M, M]\), then

\[
P(\left|Y_1 + \cdots + Y_n\right| > x) \leq 2 \exp\left(-\frac{x^2}{2(v + Mx/3)}\right),
\]

for \( v \geq \text{Var}(Y_1 + \cdots + Y_n) \).

**Lemma 4.** (*Hoeffding’s inequality, Lemma 14.11, Bühlmann and van de Geer, 2011*)

For independent random variables \( Y_1, \ldots, Y_n \) with mean zero such that \( P(Y_i \in [a_i, b_i]) = \)
1 for some $a_i$ and $b_i$ for all $i = 1, \ldots, n$. Then, for $x > 0$, we have

$$P(|Y_1 + \cdots + Y_n| > x) \leq 2 \exp \left(-\frac{x^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right).$$

**Lemma 5.** Suppose that conditions (C1)-(C3) are satisfied. For any $\delta > 0$, there exist some positive constants $c_1, \ldots, c_8$ and $\tilde{c}_1, \ldots, \tilde{c}_6$ such that for $j = 1, \ldots, p$, and $k = 1, \ldots, L_n$,

$$P\left(\left|\frac{1}{n} \sum_{i=1}^{n} B_k(T_i) f_{1,ij} - \mathbb{E}\{B_k(T)m_{1j}(T)\}\right| \geq \frac{\delta}{n}\right) \leq 2 \exp \left(-\frac{\delta^2}{c_1 L_n^{-1}n + c_2\delta}\right),$$
$$P\left(\left|\frac{1}{n} \sum_{i=1}^{n} B_k(T_i) f_{2,ij} - \mathbb{E}\{B_k(T)m_{2j}(T)\}\right| \geq \frac{\delta}{n}\right) \leq 2 \exp \left(-\frac{\delta^2}{c_3 L_n^{-1}n + c_4\delta}\right),$$
$$P\left(\left|\frac{1}{n} \sum_{i=1}^{n} B_k(T_i) f_{3,ij} - \mathbb{E}\{B_k(T)m_{3j}(T)\}\right| \geq \frac{\delta}{n}\right) \leq 2 \exp \left(-\frac{\delta^2}{c_5 L_n^{-1}n + c_6\delta}\right),$$
$$P\left(\left|\frac{1}{n} \sum_{i=1}^{n} B_k(T_i) f_{4,ij} - \mathbb{E}\{B_k(T)m_{4j}(T)\}\right| \geq \frac{\delta}{n}\right) \leq 2 \exp \left(-\frac{\delta^2}{c_7 L_n^{-1}n + c_8\delta}\right),$$
$$P\left(\left|\frac{1}{n} \sum_{i=1}^{n} B_k(T_i) f_{1w,ij} - \mathbb{E}\{B_k(T)m_{1j,w}(T)\}\right| \geq \frac{\delta}{n}\right) \leq 2 \exp \left(-\frac{\delta^2}{c_1 L_n^{-1}n + \tilde{c}_2\delta}\right),$$
$$P\left(\left|\frac{1}{n} \sum_{i=1}^{n} B_k(T_i) f_{2w,ij} - \mathbb{E}\{B_k(T)m_{2j,w}(T)\}\right| \geq \frac{\delta}{n}\right) \leq 2 \exp \left(-\frac{\delta^2}{c_3 L_n^{-1}n + \tilde{c}_4\delta}\right),$$
$$P\left(\left|\frac{1}{n} \sum_{i=1}^{n} B_k(T_i) f_{3w,ij} - \mathbb{E}\{B_k(T)m_{3j,w}(T)\}\right| \geq \frac{\delta}{n}\right) \leq 2 \exp \left(-\frac{\delta^2}{c_5 L_n^{-1}n + \tilde{c}_6\delta}\right).$$

**Proof of Lemma 5.** We first prove the inequality on $f_{1,ij}$. Denote $W_{ijk} = B_k(T_i) f_{1,ij} -$
\[ \mathbb{E}\{B_k(T)m_{ij}(T)\}. \] Then for every \( r > 2, \)

\[
\mathbb{E}\{|W_{ijk}|^r\} \leq 2^r \mathbb{E}\{|B_k(T_i)I(Y_i - Q_{t,Y} > 0)X_{ij}|^r\} \\
\leq 2^r \mathbb{E}\{B_k(T_i)\mathbb{E}[|X_{ij}|^r|T]\} \leq 2^r \mathbb{E}\{B_k^2(T_i)K_1K_r^2 r!\} \\
\leq 8C_4L_n^{-1}K_1K_2^2(2K_2)^{r-2}r!/2,
\]

where the first inequality is due to the \( C_r \) inequality, the third inequality follows by Lemma 1, and the last line uses the properties of B-spline basis. Thus, an application of Lemma 2 yields that

\[
P\left(\frac{1}{n}\sum_{i=1}^{n} W_{ijk} \geq \frac{\delta}{n}\right) \leq 2 \exp\left(-\frac{\delta^2}{2(8C_4K_1K_2^2L_n^{-1}n + 2K_2\delta)}\right).
\]

Hence, the first result is proved by letting \( c_1 = 16C_4K_1K_2^2 \) and \( c_2 = 4K_2. \)

For the second inequality, write \( U_{ijk} = B_k(T_i)f_{2,ij} - \mathbb{E}\{B_k(T)m_{2j}(T)\}. \) Clearly, it follows from the boundedness of the indicator function that \( \max_{i,j,k}|U_{ijk}| \leq 2 \) and \( \text{Var}(U_{ijk}) \leq \mathbb{E}(|B_k(T_i)f_{2,ij}|^2) \leq C_4L_n^{-1} \) uniformly in \( i, j, k. \) So applying Lemma 3 yields

\[
P\left(\frac{1}{n}\sum_{i=1}^{n} U_{ijk} \geq \frac{\delta}{n}\right) \leq 2 \exp\left(-\frac{\delta^2}{c_3L_n^{-1}n + c_4\delta}\right),
\]

where \( c_3 = 2C_4 \) and \( c_4 = 4/3. \)

Similarly, noting that \( 0 \leq w(F) \leq 1, \) we can prove the remaining inequalities of the lemma. Thus, the lemma follows. \( \square \)
Lemma 6. Suppose that conditions (C1)-(C3) are satisfied. For any $\delta > 0$, there exist some positive constants $c_9, c_{10}$ such that

$$P\left(\left\| \frac{1}{n} \sum_{i=1}^{n} B(T_i)(B(T_i))' - \mathbb{E}\{B(T)(B(T))'\} \right\| \geq \frac{L_0 \delta}{n} \right) \leq 6L_0^2 \exp\left( - \frac{\delta^2}{c_9 L_n^{-1} n + c_{10} \delta} \right),$$

and for any positive constant $h_1$, there exists some positive constant $h_2$ such that

$$P\left(\left\| \left( \frac{1}{n} \sum_{i=1}^{n} B(T_i)(B(T_i))' \right)^{-1} \right\| \geq (1 + h_1)\left\| (\mathbb{E}\{B(T)(B(T))'\})^{-1} \right\| \right) \leq 6L_0^2 \exp\left( - h_2 L_n^{-3} n \right),$$

where $\| \cdot \|$ is the operator norm, that is, $\|A\| = \sqrt{\lambda_{\text{max}}(A'A)}$ for any matrix $A$.

The above lemma is adapted from Lemma 7 of Fan, Ma and Dai (2014).

Lemma 7. Suppose that $g_1(t), g_2(t)$ are two functions of $t$ such that for $N_1 > 0, N_2 > 0$ such that $\sup_{t \in T} |g_1(t)| \leq N_1, \sup_{t \in T} |g_2(t)| \leq N_2$. For a given $t \in T$, $\hat{G}_1(t)$ and $\hat{G}_2(t)$ are estimators of $g_1(t)$ and $g_2(t)$ based on a sample with size $n$. Let $p_{n1}(x)$ and $p_{n2}(x)$ be two functions of $x$, both of which consist of some exponential forms that might depend on some constants and both are less than 1. If the following two conditions hold for any $\epsilon \in (0, 1)$:

$$\sup_{t \in T} P(|\hat{G}_1(t) - g_1(t)| \geq \epsilon) \leq p_{n1}(\epsilon),$$

$$\sup_{t \in T} P(|\hat{G}_2(t) - g_2(t)| \geq \epsilon) \leq p_{n2}(\epsilon).$$
Then we have
\[
\sup_{t \in T} P \left( |\hat{G}_1(t)\hat{G}_2(t) - g_1(t)g_2(t)| \geq \epsilon \right) \leq p_{n1}(\epsilon) + p_{n1}\left(\frac{\epsilon}{2N_2}\right) + p_{n2}\left(\frac{\epsilon}{2(N_1 + 1)}\right),
\]
\[
\sup_{t \in T} P \left( (|\hat{G}_1(t)|^2 - (g_1(t))^2 | \geq \epsilon \right) \leq p_{n1}(\epsilon) + p_{n1}\left(\frac{\epsilon}{2(N_1 + 1)}\right) + p_{n1}\left(\frac{\epsilon}{2N_1}\right),
\]
\[
\sup_{t \in T} P \left( \{|\hat{G}_1(t) - \hat{G}_2(t)| - |g_1(t) - g_2(t)|| \geq \epsilon \right) \leq p_{n1}(\epsilon/2) + p_{n2}(\epsilon/2).
\]

Furthermore, suppose that \(g_2(u)\) is uniformly bounded away from zero, in other words, there exists \(N_3 > 0\) such that \(\inf_{t \in T} |g_2(t)| > N_3\), then
\[
\sup_{t \in T} P \left( \left|\frac{\hat{G}_1(t)}{\hat{G}_2(t)} - \frac{g_1(t)}{g_2(t)}\right| \geq \epsilon \right) \leq p_{n1}\left(\frac{N_4}{2}\epsilon\right) + 2p_{n2}(\epsilon) + p_{n2}\left(\frac{N_3N_4}{2N_1}\epsilon\right),
\]
where \(N_4 = N_3 - \min(N_3/2, 1)\). In addition, if we further assume that \(g_2(t) > 0\), then
\[
\sup_{t \in T} P \left( \left|\sqrt{\hat{G}_2(t)} - \sqrt{g_2(t)}\right| \geq \epsilon \right) \leq p_{n2}(\epsilon) + p_{n2}\left((\sqrt{N_4} + \sqrt{N_3})\epsilon\right).
\]

**Proof of Lemma 7** We use the technique in the proof of Lemma 4 in Liu, Li and Wu (2014) to prove the above results. We only outline the details. Note that, using the inequality that \(|x - y| \geq |x| - |y|\), we have
\[
\sup_{t \in T} P \left( |\hat{G}_1(t)| \geq N_1 + \epsilon \right) \leq \sup_{t \in T} P \left( |\hat{G}_1(t) - g_1(t)| \geq \epsilon \right) \leq p_{n1}(\epsilon).
\]
It follows that

\[
\sup_{t \in T} P\left( |\hat{G}_1(t)\hat{G}_2(t) - g_1(t)g_2(t)| \geq \epsilon \right)
\]

\[
\leq \sup_{t \in T} P\left( |\hat{G}_1(t)||\hat{G}_2(t) - g_2(t)| \geq \epsilon/2 \right) + \sup_{t \in T} P\left( |\hat{G}_1(t) - g_1(t)||g_2(t)| \geq \epsilon/2 \right)
\]

\[
\leq \sup_{t \in T} P\left( |\hat{G}_1(t)| \geq N_1 + \epsilon \right) + \sup_{t \in T} P\left( |\hat{G}_2(t) - g_2(t)| \geq \epsilon/(2(N_1 + 1)) \right)
\]

\[
+ \sup_{t \in T} P\left( |\hat{G}_1(t) - g_1(t)| \geq \epsilon/(2N_2) \right)
\]

\[
\leq P_{n1}(\epsilon) + P_{n2}\left( \frac{\epsilon}{2(N_1 + 1)} \right) + P_{n1}\left( \frac{\epsilon}{2N_2} \right).
\]

Similarly, we can obtain the result on the square term. Because \( P(|\xi - \zeta| > x) \leq P(|\xi| > x/2) + P(|\zeta| > x/2) \) for any random variables \( \xi \) and \( \zeta \), so the result for the subtraction holds.

Next, we consider the term on the division. Since

\[
\left| \frac{\hat{G}_1(t) - g_1(t)}{\hat{G}_2(t)} \right| = \left| \frac{\{\hat{G}_1(t) - g_1(t)\}g_2(t) - g_1(t)\{\hat{G}_2(t) - g_2(t)\}}{\hat{G}_2(t)g_2(t)} \right|
\]

\[
\leq \left| \frac{\{\hat{G}_1(t) - g_1(t)\}g_2(t)}{\hat{G}_2(t)g_2(t)} \right| + \left| \frac{g_1(t)\{\hat{G}_2(t) - g_2(t)\}}{\hat{G}_2(t)g_2(t)} \right|
\]

we have

\[
\sup_{t \in T} P\left( \left| \frac{\hat{G}_1(t) - g_1(t)}{\hat{G}_2(t)} \right| \geq \epsilon \right)
\]

\[
\leq \sup_{t \in T} P\left( \left| \frac{\{\hat{G}_1(t) - g_1(t)\}}{\hat{G}_2(t)} \right| \geq \epsilon/2 \right) + \sup_{t \in T} P\left( \left| \frac{g_1(t)\{\hat{G}_2(t) - g_2(t)\}}{\hat{G}_2(t)g_2(t)} \right| \geq \epsilon/2 \right) S0.1\)
Notice that for \( t \in T \),

\[
\{ |\hat{G}_2(t) - g_2(t)| \geq \epsilon \} \supset \{ |\hat{G}_2(t)| \leq |g_2(t)| - \epsilon \} \supset \{ |\hat{G}_2(t)| \leq N_3 - \epsilon \},
\]

which implies that

\[
\sup_{t \in T} P(\{ |\hat{G}_2(t)| \leq N_3 - \epsilon \}) \leq \sup_{t \in T} P(\{ |\hat{G}_2(t) - g_2(t)| \geq \epsilon \}) \leq \rho_n(\epsilon).
\]

Hence, letting \( N_4 = N_3 - \min(N_2/2, 1) \), in other words, taking \( \epsilon = \min(N_2/2, 1) \) on the left hand side of the last inequality, we have \( \sup_{t \in T} P(\{ |\hat{G}_2(t)| \leq N_4 \}) \leq \rho_n(\epsilon) \). Thus, it follows that the first term on the right hand side of inequality (50.1) is bounded by

\[
\begin{align*}
&\sup_{t \in T} P(\{ |\hat{G}_1(t) - g_1(t)| \geq \epsilon/2 |\hat{G}_2(t)|, |\hat{G}_2(t)| > N_4 \} + \sup_{t \in T} P(\{ |\hat{G}_2(t)| \leq N_4 \}) \\
&\leq \sup_{t \in T} P(\{ |\hat{G}_1(t) - g_1(t)| \geq N_4 \epsilon/2 \} + \sup_{t \in T} P(\{ |\hat{G}_2(t)| \leq N_4 \}) \leq \rho_n \left( \frac{N_4 \epsilon}{2} \right) + \rho_n(\epsilon),
\end{align*}
\]

and the first term on the right hand side of inequality (50.1)

\[
\begin{align*}
&\leq \sup_{t \in T} P(\{ |\hat{G}_2(t) - g_2(t)| \geq N_3 \epsilon/(2N_1) |\hat{G}_2(t)|, |\hat{G}_2(t)| > N_4 \} + \sup_{t \in T} P(\{ |\hat{G}_2(t)| \leq N_4 \}) \\
&\leq \sup_{t \in T} P(\{ |\hat{G}_2(t) - g_2(t)| \geq N_3 N_4 \epsilon/(2N_1) \} + \sup_{t \in T} P(\{ |\hat{G}_2(t)| \leq N_4 \}) \\
&\leq \rho_n \left( \frac{N_3 N_4 \epsilon}{2N_1} \right) + \rho_n(\epsilon).
\end{align*}
\]

Therefore, this, together with inequality (50.1), concludes the result.
Last, for the root term, it can be similarly derived by observing the fact that
\[ \sqrt{A} - \sqrt{a} = (A - a) / (\sqrt{A} + \sqrt{a}) \]. Thus, the lemma follows. \qed

In what follows, we are going to derive some exponential bounds for \( b_{m_1,j}(t) \), \( k = 1, 2, 3, 4 \).

**Lemma 8.** Under conditions of Theorem 2.1, for any \( \epsilon > 0 \) such that \( L_n^{-1/2 + d} \epsilon \to \infty \) and \( \epsilon L_n^{-3/2} n' \to \infty \) with any \( 0 < \iota < 1/2 \) as \( n \to \infty \), then we have

\[
\sup_{t \in T} P\left( \left| \hat{m}_{1,j}(t) - m_{1,j}(t) \right| > \epsilon \right) \leq 18L_n^2 \exp(-h_2 L_n^{-3} n) + (4L_n + 6L_n^2) \exp \left( - \frac{\delta^2}{c_{15} L_n^{-1} n + c_{16} \delta} \right)
\]

\[
\sup_{t \in T} P\left( \left| \hat{m}_{2,j}(t) - m_{2,j}(t) \right| > \epsilon \right) \leq 18L_n^2 \exp(-h_2 L_n^{-3} n) + (4L_n + 6L_n^2) \exp \left( - \frac{\delta^2}{c_{17} L_n^{-1} n + c_{18} \delta} \right)
\]

\[
\sup_{t \in T} P\left( \left| \hat{m}_{3,j}(t) - m_{3,j}(t) \right| > \epsilon \right) \leq 12L_n^2 \exp(-h_2 L_n^{-3} n) + (2L_n + 6L_n^2) \exp \left( - \frac{\delta^2}{c_{19} L_n^{-1} n + c_{20} \delta} \right)
\]

\[
\sup_{t \in T} P\left( \left| \hat{m}_{4,j}(t) - m_{4,j}(t) \right| > \epsilon \right) \leq 12L_n^2 \exp(-h_2 L_n^{-3} n) + (2L_n + 6L_n^2) \exp \left( - \frac{\delta^2}{c_{21} L_n^{-1} n + c_{22} \delta} \right)
\]

where \( \delta \approx n \epsilon L_n^{-5/2} \), and \( c_{15}, \ldots, c_{22} \) are some positive constants.

**Proof of Lemma** We only give the proof of the first result because the rest can be proved in a similar manner. Note that \( \hat{m}_{1,j}(t) = (B(t))'(B'B)^{-1}B'f_{1,j} \). By the approximation theory of B-spline, we know that there exists a \( \gamma_{i,j}^* \in R^{L_n} \) such that \( m_{1,j}(t) = (B(t))'\gamma_{i,j}^* + \eta_{i,j}(t) \), where \( \eta_{i,j}(t) \) is the approximation error satisfying

\[ \sup_{1 \leq j \leq p} \sup_{t \in T} |\eta_{i,j}(t)| \leq C_5 L_n^{-d} \]
CQCSIS AND ITS APPLICATION IN SURVIVAL ANALYSIS

for some positive constant $C_5$.

For the sake of simplicity, let $\tilde{f}_{1i}(Y, X_j)$ be the $i$th component of $f_{ij}$, i.e. $\tilde{f}_{1i}(Y, X_j) = I(Y_i - \tilde{Q}_{r,Y} > 0)X_{ij}$ and denote $A_n = \frac{1}{n} \sum_{i=1}^{n} B(T_i)(B(T_i))'$, $E_n = \frac{1}{n} \sum_{i=1}^{n} B(T_i)|\tilde{f}_{1i}(Y, X_j) - f_{1i,j}|$, $D_n = \frac{1}{n} \sum_{i=1}^{n} B(T_i)f_{1i,j}$, $A = \mathbb{E}\{B(T)(B(T))'\}$ and $D = \mathbb{E}\{B(T)m_{1j}(T)\}$. Define

$$\gamma_j^M = \arg\min_{\gamma_j} \mathbb{E}\{[m_{1j}(T) - (B(T))'\gamma_j]^2\},$$

$$\hat{\gamma}_j = \arg\min_{\gamma_j} \frac{1}{n} \sum_{i=1}^{n} [\tilde{f}_{1i}(Y, X_j) - (B(T_i))'\gamma_j]^2.$$

Then, $\gamma_j^M = A^{-1}D$ and $\hat{\gamma}_j = A_n^{-1}(E_n + D_n)$. Thus, we have

$$\sup_{t \in \mathcal{T}} |\hat{m}_{1j}(t) - m_{1j}(t)| \leq \sup_{t \in \mathcal{T}} |(B(t))'(|\hat{\gamma}_j - \gamma_j^*|) + \sup_{t \in \mathcal{T}} |\eta_j(t)|$$

$$\leq \|\hat{\gamma}_j - \gamma_j^*\| + \sup_{t \in \mathcal{T}} |\eta_j(t)|$$

$$\leq \|A_n^{-1}E_n\| + \|A_n^{-1}D_n - A^{-1}D\|$$

$$+ \|\gamma_j^M - \gamma_j^*\| + \sup_{t \in \mathcal{T}} |\eta_j(t)|$$

$$= I_{n1} + I_{n2} + I_{n3} + I_{n4} \quad \text{(say)},$$

where the third line uses the following fact that

$$\sup_{t \in \mathcal{T}} \|B(t)\| = \sup_{t \in \mathcal{T}} \left( \sum_{k=1}^{L_n} (B_k(t))^2 \right)^{1/2} \leq \sup_{t \in \mathcal{T}} \left( \sum_{k=1}^{L_n} \|B_k(t)\| \|B_k\|_\infty \right) \leq \sup_{t \in \mathcal{T}} \left( \sum_{k=1}^{L_n} B_k(t) \right) \leq 1.$$
On the one hand, we can derive that

\[ I_{n3} = \| \gamma_j^M - \gamma_j^* \| = \| (E\{B(T)(B(T))'\})^{-1}E\{B(T)\eta_j(T)\} \| \]
\[ \leq C_3^{-1}L_n\|E\{B(T)\}\|\max_j\sup_{t \in T}\|\eta_j(T)\| \]
\[ \leq C_3^{-1}L_nC_5L_n^{-d}\|E\{B(T)\}\| \leq C_6L_n^{1/2-d} , \]

where the last inequality is because \( E\{B_k(T)\} \leq C_0L_n^{-1} \) for some constant \( C_0 \), and \( C_6 = C_3^{-1}C_5C_0 \). On the other hand, it is easily shown that \( I_{n4} \leq C_5L_n^{-d} \). Therefore, by setting \( \epsilon/2 > C_6L_n^{1/2-d} + C_5L_n^{-d} \), it follows that

\[ \sup_{t \in T} P\left( |\hat{m}_{1j}(t) - m_{1j}(t)| > \epsilon \right) \leq P\left( \sup_{t \in T} |\hat{m}_{1j}(t) - m_{1j}(t)| > \epsilon \right) \]
\[ \leq P\left( |I_{n1} + I_{n2} + I_{n3} + I_{n4}| > \epsilon \right) \]
\[ \leq P\left( |I_{n1}| + |I_{n2}| > \epsilon/2 \right) . \quad (S0.2) \]

Next, we first consider the term for \( I_{n2} \) in inequality (S0.2). Note that

\[ I_{n2} = \|A_n^{-1}(D_n - D) + A_n^{-1}(A - A_n)A^{-1}D\| \]
\[ \leq \|A_n^{-1}(D_n - D)\| + \|A_n^{-1}(A - A_n)A^{-1}D\| \]
\[ = I_{n2}^{(1)} + I_{n2}^{(2)} \quad \text{(say)}. \quad (S0.3) \]
By Lemma 1, we have

\[ |\mathbb{E}\{B_k(T) m_{1j}(T)\}| \leq |\mathbb{E}\{B_k(T) [\mathbb{E}(|f_{1j}(Y, X_j)|^2|T)]^{1/2}\}| \]
\[ \leq |\mathbb{E}\{B_k(T) [\mathbb{E}(X_j^2|T)]^{1/2}\}| \]
\[ \leq (2K_1K_2^2)^{1/2} \mathbb{E}\{B_k(T)\} \leq C_7 L_n^{-1}, \]

where \(C_7 = (2K_1K_2^2)^{1/2}C_0\). Thus, it follows that

\[ \|D\| = \left( \sum_{k=1}^{L_n} \left[ \mathbb{E}\{B_k(T) m_{1j}(T)\} \right]^2 \right)^{1/2} \leq C_7 L_n^{1/2}. \] \((S0.4)\)

By Lemma 5, we have, for \(\delta > 0\),

\[ P\left( \|D_n - D\| \geq \frac{L_n^{1/2} \delta}{n} \right) \leq \sum_{k=1}^{L_n} P\left( \left| \frac{1}{n} \sum_{i=1}^{n} B_k(T_i) f_{1,i} - \mathbb{E}\{B_k(T) m_{1j}(T)\} \right| \geq \frac{\delta}{n} \right) \]
\[ \leq 2L_n \exp\left( - \frac{\delta^2}{c_1 L_n^{-1} n + c_2 \delta} \right). \] \((S0.5)\)

Furthermore, by Lemma 6, we have

\[ P\left( \|A_n^{-1}\| \geq (1 + h_1)C_3^{-1} L_n \right) \leq 6L_n^2 \exp(-h_2 L_n^{-3} n). \] \((S0.6)\)

Combining \((S0.3)-(S0.6)\), we have

\[ P\left( I_{n2}^{(1)} \geq \frac{C_8 L_n^{3/2} \delta}{n} \right) \leq 2L_n \exp\left( - \frac{\delta^2}{c_1 L_n^{-1} n + c_2 \delta} \right) + 6L_n^2 \exp(-h_2 L_n^{-3} n). \] \((S0.7)\)
where \( C_8 = (1 + h_1)C_3^{-1} \). As a result, by Lemma 6, we conclude that

\[
P\left( \left| I_{n^2}^{(2)} \right| \geq \frac{C_9 L_5^{5/2} \delta}{n} \right) \leq P\left( \left\| A_n^{-1}(A_n - A)D \right\| \geq \frac{C_9 L_5^{5/2} \delta}{n} \right)
\leq P\left( \left\| A_n^{-1}(A_n - A) \right\| \geq \frac{(1 + h_1)C_3^{-1} L_5^2 \delta}{n} \right)
\leq P\left( \left\| A_n^{-1} \right\| \geq (1 + h_1)C_3^{-1} L_n^{-1} \right) + P\left( \left\| A_n - A \right\| \geq \frac{L_n \delta}{n} \right)
\leq 6L_n^2 \exp(-h_2 L_n^{-3}n) + 6L_n^2 \exp\left(-\frac{\delta^2}{c_9 L_n^{-1}n + c_{10} \delta}\right), (S0.8)
\]

where \( C_9 = C_3^{-2}C_7(1 + h_1) \). According to (S0.7) and (S0.8), we obtain, for any \( \delta > 0 \),

\[
P\left( \left| I_{n^2} \right| \geq \frac{C_8 L_5^{3/2} + C_9 L_5^{5/2} \delta}{n} \right)
\leq P\left( \left| I_{n^2}^{(1)} \right| \geq \frac{C_8 L_5^{3/2} \delta}{n} \right) + P\left( \left| I_{n^2}^{(2)} \right| \geq \frac{C_9 L_5^{5/2} \delta}{n} \right)
\leq 12L_n^2 \exp(-h_2 L_n^{-3}n) + (2L_n + 6L_n^2) \exp\left(-\frac{\delta^2}{c_{11} L_n^{-1}n + c_{12} \delta}\right), (S0.9)
\]

where \( c_{11} = \max(c_1, c_9) \) and \( c_{12} = \max(c_2, c_{10}) \).

Last, we consider the term for \( I_{n1} \) in inequality (S0.2). Write \( a_{njk} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} B_k(T_i)\left[ \hat{f}_{1i}(Y, X_j) - \hat{f}_{1,ij} \right] \). Note that \( |Q_{\tau,Y} - \hat{Q}_{\tau,Y}| = O(n^{-1/2}(\log n)^{1/2}) \) almost surely (Serfling, 1980, Section 2.5.1; He et al., 2013, Lemma 8.4) and the fact that for any \( y \) and any \( \epsilon > 0 \), \( \sup_{|y_1 - y| < \epsilon} |I(Y < y_1) - I(Y < y)| \leq I(y - \epsilon < Y < y + \epsilon) \). On the event \( \{|\hat{Q}_{\tau,Y} - Q_{\tau,Y}| \leq M_3 \eta n^{-\epsilon}\} \) with any \( \eta > 0 \), \( 0 < \epsilon < 1/2 \) and some constant \( M_3 > 0 \), we
have

\[
\left| a_{njk} \right| \leq \frac{1}{n} \sum_{i=1}^{n} B_k(T_i) \left| \hat{f}_{1i}(Y, X_j) - f_{1,ij} \right|
\]

\[
\leq \frac{1}{n} \sum_{i=1}^{n} B_k(T_i) |X_{ij}| I(Q_{\tau,Y} - M_3\eta n^{-t} < Y_i < Q_{\tau,Y} + M_3\eta n^{-t})
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} \xi_{ijk} \text{ (say}).
\]

Write \( q_\tau(Y_i, Q_{\tau,Y}) = I(Q_{\tau,Y} - M_3\eta n^{-t} < Y_i < Q_{\tau,Y} + M_3\eta n^{-t}) \). Then, by Taylor’s expansion and condition (C4), we have

\[
\mathbb{E}\{q_\tau(Y_i, Q_{\tau,Y})|X_{ij}, T_i\} = \int_{Q_{\tau,Y} - M_3\eta n^{-t}}^{Q_{\tau,Y} + M_3\eta n^{-t}} f_{Y_i|(X_{ij}, T_i)}(s)ds = 2M_3\eta \int_{Y_i|(X_{ij}, T_i)}(Q_{\tau,Y})n^{-t}\{1 + o(1)\},
\]

which further implies that for sufficiently large \( n \),

\[
\mu_1 \equiv \mathbb{E}\{B_k(T_i)|X_{ij}|\mathbb{E}[q_\tau(Y_i, Q_{\tau,Y})|X_{ij}, T_i]\}
\]

\[
\leq 2M_3\eta n^{-t}\mathbb{E}\{B_k(T_i)|X_{ij}|f_{Y_i|(X_{ij}, T_i)}(Q_{\tau,Y})\}\{1 + o(1)\}
\]

\[
\leq 4M_3\eta M_{f1} n^{-t}\mathbb{E}\{B_k(T_i)|X_{ij}|\}
\]

\[
\leq 4M_3\eta M_{f1} n^{-t} K_1 K_2 C_0 L_n^{-1} = M_4 n^{-t} L_n^{-1},
\]

where condition (C4) implies that there exists a constant \( M_{f1} \) such that \( \sup_{|y - Q_{\tau,y}| < \varepsilon} f_{Y|(X_{j}, T)}(y) \leq M_{f1} < \infty \) for any \( \varepsilon > 0 \), and \( M_4 = 4M_3 M_{f1} K_1 K_2 C_0 \eta \). For every \( r > 2 \), by the \( C_r \)
inequality and the boundedness of an indicator function, we have

\[ E\{|\xi_{ijk} - \mu_1|\} \leq 2^r E\{|B^r_k(T_i)|X_{ij}|\} \leq 2^r E\{|B^r_k(T_i)K_2r!|\} \leq 2^r K_1K_2^r r! C_4 L_n^{-1} = 8K_1K_2^2 C_4 L_n^{-1} (2K_2)^{r-2} r!/2. \]

By choosing suitable constants such that \( \delta > 2M_4 n^{1-t} L_n^{-1} \), we have \( \delta/n - \mu_1 > \delta/(2n) \) for some constant \( M_{f2} \). It follows from Lemma 4 that

\[
\max_{1 \leq j \leq p} \max_{1 \leq k \leq L_n} P\left(\left|a_{njk}\right| > \frac{\delta}{n}\right) \\
\leq \max_{1 \leq j \leq p} \max_{1 \leq k \leq L_n} P\left(\frac{1}{n} \sum_{i=1}^{n} \xi_{ijk} - \mu_1 > \frac{\delta}{2n}\right) \\
\leq \max_{1 \leq j \leq p} \max_{1 \leq k \leq L_n} P\left(\frac{1}{n} \sum_{i=1}^{n} \xi_{ijk} - \mu_1 > \frac{\delta}{2n}\right) \\
\leq 2 \exp\left(- \frac{\delta^2}{c_{13} L_n^{-1} n + c_{14} \delta}\right),
\]

where \( c_{13} = 16K_1K_2^2 C_4 \) and \( c_{14} = 8K_2 \). With the above results, we have

\[
P\left(\left|I_{n1}\right| > \frac{(1 + h_1)C_3^{-1} L_n^{3/2} \delta}{n}\right) \\
\leq P\left(\|A_n^{-1}\| > (1 + h_1)C_3^{-1} L_n\right) + \sum_{k=1}^{L_n} P\left(\left|a_{njk}\right| > \delta/n\right) \\
\leq 6L_n^2 \exp\left(-h_2 L_n^{-3} n\right) + 2L_n \exp\left(- \frac{\delta^2}{c_{13} L_n^{-1} n + c_{14} \delta}\right).
\]  

(S0.10)
CQCSIS AND ITS APPLICATION IN SURVIVAL ANALYSIS

Hence, combining (S0.2), (S0.9), (S0.10) and the condition of lemma and taking \( \frac{(2C_8L_n^{3/2} + C_9L_n^{5/2})\delta}{n} = \epsilon/2 \), we can obtain

\[
\sup_{t \in T} P\left( |\hat{m}_{1j}(t) - m_{1j}(t)| > \epsilon \right) \\
\leq 18L_n^2 \exp(-h_2L_n^{-3}n) + (4L_n + 6L_n^2) \exp \left( -\frac{\delta^2}{c_{15}L_n^{-1}n + c_{16}L_n^{-1}} \right),
\]

where \( c_{15} = \max(c_{11}, c_{13}) \), \( c_{16} = \max(c_{12}, c_{14}) \), and \( \delta = n\epsilon/(4C_8L_n^{3/2} + 2C_9L_n^{5/2}) \propto n\epsilon L_n^{-5/2} \).

For the second result of the lemma, we can complete the proof in a similar manner, where the proof of (S0.10) is distinct and has to modify as follows. Write \( b_{njk} = \frac{1}{n} \sum_{i=1}^{n} B_k(T_i)[\hat{f}_2(Y, X_j) - f_{2,ij}] \). On the event \( \{|\hat{Q}_{\tau,Y} - Q_{\tau,Y}| \leq M_3\eta n^{-1}\} \), we have

\[
|b_{njk}| \leq \frac{1}{n} \sum_{i=1}^{n} B_k(T_i)I(Q_{\tau,Y} - M_3\eta n^{-1} < Y_i < Q_{\tau,Y} + M_3\eta n^{-1})
\]

\[= \frac{1}{n} \sum_{i=1}^{n} \zeta_{ijk} \text{ (say)}.\]

Also, condition (C4) implies that \( \sup_{|y-Q_{\tau,Y}|<\epsilon} f_{Y|T}(y) \leq M_2 < \infty \). Similarly, we derive that \( \mu_2 \hat{=} \mathbb{E}\{\zeta_{ijk}\} = \mathbb{E}\{B_k(T_i)\mathbb{E}[q_{r}(Y_i, Q_{\tau,Y})|T_i]\} \leq 4\eta M_3M_2C_0n^{-1}L_n^{-1} \) and

\[
\text{Var}(\zeta_{ijk} - \mu_2) \leq \mathbb{E}(|\zeta_{ijk}|^2) = \mathbb{E}\{B_k^2(T_i)q_{r}(Y_i, Q_{\tau,Y})\}
\]

\[\leq 4M_3\eta n^{-1}M_2\mathbb{E}\{B_k^2(T_i)\} \leq M_5 n^{-1}L_n^{-1},\]
where $M_5 = 4\eta M_3 M_{f_2} C_4$. In addition, observe that $\max_{i,j,k} |\zeta_{ijk} - \mu_2| \leq 2$. With an application of Lemma 8 and choosing $\delta > 8\eta M_3 M_{f_2} n^{1-\epsilon} L_n^{-1}$, we get

$$
\max_{1 \leq j \leq p} \max_{1 \leq k \leq L_n} P\left(\left| b_{njk} \right| > \frac{\delta}{n}\right)
\leq \max_{1 \leq j \leq p} \max_{1 \leq k \leq L_n} P\left(\left| \frac{1}{n} \sum_{i=1}^{n} \zeta_{ijk} - \mu_2 \right| > \frac{\delta}{2n}\right)
\leq 2 \exp\left( - \frac{\delta^2}{2(65n^{-1} L_n^{-1} + 2\delta/3)} \right)
\leq 2 \exp\left( - \frac{\delta^2}{8M_5 L_n^{-1} n + 16\delta/3} \right).
$$

Thus, we can obtain the second result. As in derivation of the first result, the rest of results of the lemma can be proved.

The lemmas given below are useful for establishing screening properties for censoring data.

**Lemma 9.** Let $\mathcal{F}$ be a class of distribution functions whose support is the same as that of $F$, and let $\mathcal{Y}$ be the support of $Y$. For any $\varepsilon > 0$, define $W(\varepsilon) = \{ F^* \in \mathcal{F} : \| F^* - F \|_\infty \equiv \sup_{y \in \mathcal{Y}} |F^*(y) - F(y)| \leq \varepsilon, |Q_{\tau,Y^*} - Q_{\tau,Y}| \leq \varepsilon \}$, where $Y^*$ is generated from the distribution $F^*$. Then, we have (i)

$$
\sup_{W(\varepsilon)} \left| w(F^*) I\{ \tilde{Y} \leq Q_{\tau,Y^*} \} - w(F) I\{ \tilde{Y} \leq Q_{\tau,Y} \} \right|
\leq \frac{\varepsilon}{1 - \tau} + I\{ Q_{\tau,Y} - \varepsilon < Y \leq Q_{\tau,Y} + \varepsilon \}
+ 3 I\{ Q_{\tau,Y} - \varepsilon < Z \leq Q_{\tau,Y} + \varepsilon \} + I\{ F^{-1}(\tau - \varepsilon) \leq Z < F^{-1}(\tau + \varepsilon) \},
$$
and (ii)

\[
\sup_{w(\varepsilon)} \left| w^2(F^*)I\{\tilde{Y} \leq Q_{\tau,Y^*}\} - w^2(F)I\{\tilde{Y} \leq Q_{\tau,Y}\} \right|
\]

\[
\leq \frac{4\varepsilon}{1 - \tau} + I\{Q_{\tau,Y} - \varepsilon < Y \leq Q_{\tau,Y} + \varepsilon\}
\]

\[
+ 3I\{Q_{\tau,Y} - \varepsilon < Z \leq Q_{\tau,Y} + \varepsilon\} + 2I\{F^{-1}(\tau - \varepsilon) \leq Z < F^{-1}(\tau + \varepsilon)\}.
\]

**Proof of Lemma** Since the part (i) is adapted immediately from Wu and Yin (2015), it remains to show part (ii). By using similar arguments in (A.2) of Wang and Wang (2009), we can derive that

\[
w^2(F)I\{\tilde{Y} \leq Q_{\tau,Y}\}
\]

\[
= I\{Y \leq Z, Z \leq Q_{\tau,Y}\} + I\{Y \leq Q_{\tau,Y}, Q_{\tau,Y} < Z\} + I\{Y > Z, Z < Q_{\tau,Y}\}
\]

\[
\times \left[ 1 + \left( \frac{1 - \tau}{1 - F(Z)} \right)^2 I\{F(Z) < \tau\} - \frac{2(1 - \tau)}{1 - F(Z)} I\{F(Z) < \tau\} \right].
\]

Thus, it follows that

\[
w^2(F^*)I\{\tilde{Y} \leq Q_{\tau,Y^*}\} - w^2(F)I\{\tilde{Y} \leq Q_{\tau,Y}\} = I_1 + I_2 + I_3,
\]
where

\[ I_1 = I\{Y \leq Z, Z \leq Q_{\tau,Y}\} - I\{Y \leq Z, Z \leq Q_{\tau,Y}\}, \]

\[ I_2 = I\{Y \leq Q_{\tau,Y^*}, Q_{\tau,Y^*} < Z\} - I\{Y \leq Q_{\tau,Y}, Q_{\tau,Y} < Z\}, \]

\[ I_3 = I\{Y > Z, Z < Q_{\tau,Y}\} \left[ 1 + \left( \frac{1 - \tau}{1 - F^*(Z)} \right)^2 I\{F^*(Z) < \tau\} - \frac{2(1 - \tau)}{1 - F^*(Z)} I\{F^*(Z) < \tau\} \right] - I\{Y > Z, Z < Q_{\tau,Y}\} \left[ 1 + \left( \frac{1 - \tau}{1 - F(Z)} \right)^2 I\{F(Z) < \tau\} - \frac{2(1 - \tau)}{1 - F(Z)} I\{F(Z) < \tau\} \right]. \]

Then, we have

\[ \sup_{W(\epsilon)} |I_1| \leq \sup_{W(\epsilon)} |I\{Z \leq Q_{\tau,Y}\} - I\{Z \leq Q_{\tau,Y}\}| \leq \max\{Q_{\tau,Y} - \epsilon < Z \leq Q_{\tau,Y} + \epsilon\}, \]

\[ \sup_{W(\epsilon)} |I_2| \leq \sup_{W(\epsilon)} |I\{Y \leq Q_{\tau,Y}\} - I\{Y \leq Q_{\tau,Y}\}| + |I\{Q_{\tau,Y^*} < Z\} - I\{Q_{\tau,Y} < Z\}| \]

\[ \leq \max\{Q_{\tau,Y} - \epsilon < Y \leq Q_{\tau,Y} + \epsilon\} + \max\{Q_{\tau,Y} - \epsilon < Z \leq Q_{\tau,Y} + \epsilon\}, \]

\[ \sup_{W(\epsilon)} |I_3| \leq \sup_{W(\epsilon)} |I\{Y > Z, Z < Q_{\tau,Y}\} - I\{Y > Z, Z < Q_{\tau,Y}\}| \]

\[ + \sup_{W(\epsilon)} \left| \left( \frac{1 - \tau}{1 - F^*(Z)} \right)^2 I\{F^*(Z) < \tau\} - \left( \frac{1 - \tau}{1 - F(Z)} \right)^2 I\{F(Z) < \tau\} \right| \]

\[ + \sup_{W(\epsilon)} \left| \frac{2(1 - \tau)}{1 - F^*(Z)} I\{F^*(Z) < \tau\} - \frac{2(1 - \tau)}{1 - F(Z)} I\{F(Z) < \tau\} \right| \]

\[ = I_{31} + I_{32} + I_{33} \text{ (say)}, \]

where we further have

\[ I_{31} \leq I\{Q_{\tau,Y} - \epsilon < Z \leq Q_{\tau,Y} + \epsilon\}, \]
and

\[
I_{32} \leq \sup_{W(\epsilon)} I\{F^*(Z) < \tau, F(Z) < \tau\} \left( \frac{1 - \tau}{1 - F^*(Z)} \right)^2 - \left( \frac{1 - \tau}{1 - F(Z)} \right)^2 \\
+ \sup_{W(\epsilon)} I\{F^*(Z) < \tau < F(Z)\} \left( \frac{1 - \tau}{1 - F^*(Z)} \right)^2 \\
+ \sup_{W(\epsilon)} I\{F(Z) < \tau < F^*(Z)\} \left( \frac{1 - \tau}{1 - F(Z)} \right)^2 \\
\leq \frac{2}{1 - \tau} \sup_{W(\epsilon)} |F^*(Z) - F(Z)| + \sup_{W(\epsilon)} I\{F^*(Z) < \tau < F(Z)\} \\
+ \sup_{W(\epsilon)} I\{F(Z) < \tau < F^*(Z)\} \\
\leq \frac{2\epsilon}{1 - \tau} + I\{F^{-1}(\tau - \epsilon) < Z \leq F^{-1}(\tau + \epsilon)\},
\]

and, similarly,

\[
I_{33} \leq \frac{2\epsilon}{1 - \tau} + I\{F^{-1}(\tau - \epsilon) < Z \leq F^{-1}(\tau + \epsilon)\}.
\]

Therefore, gathering the above terms yields the desired result. \(\square\)

**Lemma 10.** Under conditions of Theorem 3.1, for any \(\epsilon > 0\) such that \(L_n^{-1/2 + \delta} \epsilon \to \infty\) and \(\epsilon L_n^{-3/2} n^\iota \to \infty\) with any \(0 < \iota < 1/2\) as \(n \to \infty\), then we have

\[
\sup_{t \in \mathcal{T}} P(|\hat{m}_{1j,w}(t) - m_{1j,w}(t)| > \epsilon) \leq 18L_n^2 \exp(-h_2 L_n^{-3} n) + (10L_n + 6L_n^2) \exp \left( -\frac{\delta^2}{\hat{c}_7 L_n^{-1} n + \hat{c}_8 \delta} \right)
\]

\[
\sup_{t \in \mathcal{T}} P(|\hat{m}_{2j,w}(t) - m_{2j,w}(t)| > \epsilon) \leq 18L_n^2 \exp(-h_2 L_n^{-3} n) + (10L_n + 6L_n^2) \exp \left( -\frac{\delta^2}{\hat{c}_9 L_n^{-1} n + \hat{c}_{10} \delta} \right)
\]

\[
\sup_{t \in \mathcal{T}} P(|\hat{m}_{3j,w}(t) - m_{3j,w}(t)| > \epsilon) \leq 18L_n^2 \exp(-h_2 L_n^{-3} n) + (10L_n + 6L_n^2) \exp \left( -\frac{\delta^2}{\hat{c}_{11} L_n^{-1} n + \hat{c}_{12} \delta} \right)
\]
where \( \delta \asymp n \epsilon L_n^{-5/2} \), and \( \tilde{c}_7, \ldots, \tilde{c}_{12} \) are some positive constants.

**Proof of Lemma 10** By the theory of spline approximation, there exists a vector \( \gamma_{j,w}^* \) such that \( m_{1,j,w}(t) = (B(t))' \gamma_{j,w}^* + \eta_{j,w}(t) \), where \( \eta_{j,w}(t) \) is the approximation error. Write \( A_n = \frac{1}{n} \sum_{i=1}^{n} B(T_i)(B(T_i))' \), \( E_{n,w} = \frac{1}{n} \sum_{i=1}^{n} B(T_i)X_{ij}w_i(F)I(\bar{Y}_i < \bar{Q}_{\tau,Y}) \), \( D_{n,w} = \frac{1}{n} \sum_{i=1}^{n} B(T_i)X_{ij}w_i(F)I(\bar{Y}_i < \bar{Q}_{\tau,Y}) \), \( A = \mathbb{E}\{B(T)(B(T))'\} \), \( D_w = \mathbb{E}\{B(T)m_{1,j,w}(T)\} \). Then, following the proof of Lemma 8, we have

\[
\sup_{t \in \mathcal{T}} |\hat{m}_{1,j,w}(t) - m_{1,j,w}(t)| \\
\leq \| A_n^{-1} E_{n,w} \| + \| A_n^{-1} D_{n,w} - A^{-1} D_w \| + \| \hat{\gamma}_{j,w} - \gamma_{j,w}^* \| + \sup_{t \in \mathcal{T}} |\eta_{j,w}(t)| \\
= II_{n1} + II_{n2} + II_{n3} + II_{n4} \quad \text{(say),} \tag{S0.11}
\]

where \( \gamma_{j,w}^M = A^{-1}D_w \). Following the derivation of upper bounds of \( I_{n3} \) and \( I_{n4} \) in Lemma 8, we get that \( II_{n3} \leq C_6 L_n^{-1/2-d} \) and \( II_{n4} \leq C_5 L_n^{-d} \).

Using similar arguments to those in the derivation of exponential tail probability of \( I_{n2} \) in Lemma 8 and invoking Lemma 9, we obtain that

\[
P\left( |II_{n2}| \geq \frac{(C_8 L_n^{3/2} + C_9 L_n^{5/2}) \delta}{n} \right) \\
\leq 12L_n^2 \exp(-h_2 L_n^{-3} n) + (2L_n + 6L_n^2) \exp\left( -\frac{\delta^2}{c_{17} L_n^{-1} n + c_{18} \delta} \right) \tag{S0.12}
\]

for some positive constants \( c_{17} \) and \( c_{18} \). It remains to derive an upper bound of the exponential tail probability of \( II_{n1} \).
Under conditions (D2) and (D3), \( \| \hat{F} - F \|_\infty = O(n^{-1/2}(\log n)^{1/2}) \) and \( |\hat{Q}_{\tau,Y} - Q_{\tau,Y}| = O(n^{-1/2}(\log n)^{1/2}) \) almost surely (see He, Wang and Hong, 2013, Lemma 8.4).

Under condition (D1), there exists a constant \( c_f \) such that \( 1/f(y) \leq c_f \) uniformly in a neighbourhood of \( Q_{\tau,Y} \). For any \( \eta > 0, 0 < \iota < 1/2 \) and some constants \( \ddot{a}_2, \dot{a}_3, \dddot{a}_4 \), we write \( \phi_{ij}^{(1)} = \eta n^{-\iota} \ddot{a}_2^{-1} B_k(T_i)|X_{ij}|/(1 - \tau), \phi_{ij}^{(2)} = B_k(T_i)|X_{ij}|I\{Q_{\tau,Y} - \eta n^{-\iota} \dot{a}_3^{-1} < Y_i \leq Q_{\tau,Y} + \eta n^{-\iota} \dot{a}_3^{-1}\}, \phi_{ij}^{(3)} = 3B_k(T_i)|X_{ij}|I\{Q_{\tau,Y} - \eta n^{-\iota} \ddot{a}_2^{-1} < Z_i \leq Q_{\tau,Y} + \eta n^{-\iota} \dddot{a}_4^{-1}\}, \phi_{ij}^{(4)} = B_k(T_i)|X_{ij}|I\{F^{-1}(\tau - \eta n^{-\iota} c_f^{-1} \dddot{a}_4^{-1}) \leq Z < F^{-1}(\tau + \eta n^{-\iota} c_f^{-1} \dddot{a}_4^{-1})\} \), and denote \( \mu^{(l)}_{jk} = \mathbb{E}\{\phi_{ij}^{(l)}\}, l = 1, 2, 3 \). Note that there exist positive constants \( M_{f_1}, M_{h_1} \) such that \( f_{Y|(X_j,Y)}(y) \leq M_{f_1} < \infty \) and \( h_{Z|(X_j,Y)}(y) \leq M_{h_1} < \infty \) uniformly in a neighbourhood of \( Q_{\tau,Y} \), due to conditions (D1) and (D2). Combining these with Lemma II, we obtain that \( \max_{j,k} \mu^{(1)}_{jk} \leq K_1 K_2 C_0 \eta \ddot{a}_2^{-1} n^{-\iota} L_n^{-1} / (1 - \tau) \), \( \max_{j,k} \mu^{(2)}_{jk} \leq 4K_1 K_2 C_0 \eta \dddot{a}_4^{-1} M_{f_1} n^{-\iota} L_n^{-1} \), \( \max_{j,k} \mu^{(3)}_{jk} \leq 12K_1 K_2 C_0 \eta \dddot{a}_3^{-1} M_{h_1} n^{-\iota} L_n^{-1} \) and for every \( r > 2 \),

\[
\mathbb{E}\{|\phi_{ij}^{(1)} - \mu_{jk}^{(1)}|^r\} \leq 2^r \mathbb{E}\{|B_{k}^r(T_i)|X_{ij}|^r\} \leq (2K_2)^{r-2} r!(8K_1 K_2^2 C_4 L_n^{-1})/2,
\]

\[
\mathbb{E}\{|\phi_{ij}^{(2)} - \mu_{jk}^{(2)}|^r\} \leq 2^r \mathbb{E}\{|B_{k}^r(T_i)|X_{ij}|^r\} \leq (2K_2)^{r-2} r!(8K_1 K_2^2 C_4 L_n^{-1})/2,
\]

\[
\mathbb{E}\{|\phi_{ij}^{(3)} - \mu_{jk}^{(3)}|^r\} \leq 2^r \mathbb{E}\{|3^r B_{k}^r(T_i)|X_{ij}|^r\} \leq (6K_2)^{r-2} r!(72K_1 K_2^2 C_4 L_n^{-1})/2.
\]

Therefore, taking \( \delta > M_0 n^{-\iota} L_n^{-1} \) with \( M_0 = 2K_1 K_2 C_0 \eta \max(\ddot{a}_2^{-1} / (1 - \tau), 4\dddot{a}_2^{-1} M_{f_1}, 12\dddot{a}_3^{-1} M_{h_1}, 4\dddot{a}_4^{-1} M_{h_1}), \)
it follows from Lemma 2 that

\[ \max_{1 \leq j \leq p} \max_{1 \leq k \leq L} P\left( \frac{1}{n} \sum_{i=1}^{n} \phi_{ijk}^{(1)} \geq \frac{\delta}{n} \right) \leq \max_{1 \leq j \leq p} \max_{1 \leq k \leq L} P\left( \frac{1}{n} \sum_{i=1}^{n} \phi_{ijk}^{(1)} - \mu_{jk}^{(1)} \geq \frac{\delta}{2n} \right) \leq 2 \exp \left( - \frac{\delta^2}{64 K_1 K_2^2 C_4 L_n^{-1} n + 8 K_2 \delta} \right), \]  

\[(S0.13)\]

and, similarly,

\[ \max_{1 \leq j \leq p} \max_{1 \leq k \leq L} P\left( \frac{1}{n} \sum_{i=1}^{n} \phi_{ijk}^{(2)} \geq \frac{\delta}{n} \right) \leq 2 \exp \left( - \frac{\delta^2}{64 K_1 K_2^2 C_4 L_n^{-1} n + 8 K_2 \delta} \right), \]  

\[(S0.14)\]

and

\[ \max_{1 \leq j \leq p} \max_{1 \leq k \leq L} P\left( \frac{1}{n} \sum_{i=1}^{n} \phi_{ijk}^{(3)} \geq \frac{\delta}{n} \right) \leq 2 \exp \left( - \frac{\delta^2}{576 K_1 K_2^2 C_4 L_n^{-1} n + 24 K_2 \delta} \right), \]  

\[(S0.15)\]

An application of Taylor’s expansion of\( F^{-1}(\tau - \eta \mu^{-4} c_f^{-1} \bar{a}^{-1}_4) \) and\( F^{-1}(\tau + \eta \mu^{-4} c_f^{-1} \bar{a}^{-1}_4) \) around\( \tau \) gives \( I\{ F^{-1}(\tau - \eta \mu^{-4} c_f^{-1} \bar{a}^{-1}_4) \leq Z < F^{-1}(\tau + \eta \mu^{-4} c_f^{-1} \bar{a}^{-1}_4) \} \leq I\{ Q_{r,Y} - \eta \mu^{-4} \bar{a}^{-1}_4 \leq Z < Q_{r,Y} + \eta \mu^{-4} \bar{a}^{-1}_4 \} \}. Using this and following the derivation of\( (S0.13) \), we get

\[ \max_{1 \leq j \leq p} \max_{1 \leq k \leq L} P\left( \frac{1}{n} \sum_{i=1}^{n} \phi_{ijk}^{(4)} \geq \frac{\delta}{n} \right) \leq 2 \exp \left( - \frac{\delta^2}{64 K_1 K_2^2 C_4 L_n^{-1} n + 8 K_2 \delta} \right). \]  

\[(S0.16)\]

Let \( d_{ijk} = \frac{1}{n} \sum_{i=1}^{n} B_k(T_i) X_{ij} [w_i(\bar{F}) I(\bar{Y}_i \leq \bar{Q}_{r,Y}) - w_i(F) I(\bar{Y}_i < Q_{r,Y})] \). Taking \( \bar{a}_5^{-1} = \)
min(\tilde{a}_1^{-1}, \tilde{a}_2^{-1}, \tilde{a}_3^{-1}, c_f^{-1} \tilde{a}_4^{-1}, \tilde{a}_4^{-1})$, applying part (i) of Lemma 4 with $\varepsilon = m^{-1} \tilde{a}_5^{-1}$ and combining (S0.13)-(S0.14), we obtain

$$
\max_{1 \leq j \leq p} \max_{1 \leq k \leq L_n} P\left(\left| \frac{1}{n} \sum_{i=1}^{n} d_{ijk} \right| \geq \frac{4\delta}{n} \right)
\leq \max_{1 \leq j \leq p} \max_{1 \leq k \leq L_n} P\left(\frac{1}{n} \sum_{i=1}^{n} \phi_{ijk}^{(1)} \geq \frac{\delta}{n} \right) + \max_{1 \leq j \leq p} \max_{1 \leq k \leq L_n} P\left(\frac{1}{n} \sum_{i=1}^{n} \phi_{ijk}^{(2)} \geq \frac{\delta}{n} \right)
\leq 8 \exp \left( - \frac{\delta^2}{576K_1K_2^2C_4L_n^{-1}n + 24K_2\delta} \right).
$$

As in (S0.11), we get

$$
P\left(\|II_{n1}\| > \frac{4(1 + h_1)C_3^{-1}L_n^{3/2}\delta}{n} \right)
\leq P\left(\|A_n^{-1}\| > (1 + h_1)C_3^{-1}L_n^{1} \right) + \sum_{k=1}^{L_n} P\left(|d_{njk}| > 4\delta/n \right)
\leq 6L_n^2 \exp(-h_2L_n^{-3}n) + 8L_n \exp\left( - \frac{\delta^2}{c_{19}L_n^{-1}n + c_{20}\delta} \right). \tag{S0.17}
$$

where $c_{19} = 576K_1K_2^2C_4$ and $c_{20} = 24K_2$.

Under the condition of the lemma that $\varepsilon^{-1}L_n^{1/2-d} \to 0$, we can choose large $n$ such that $C_6L_n^{1/2-d} + C_5L_n^{-2} \leq \epsilon/2$. Combining (S0.11), (S0.12) and (S0.17) and taking
(5C_8L_n^{2/2} + C_9L_n^{5/2})\delta/n = \epsilon/2$, we get

$$
\sup_{t \in \mathcal{T}} P\left( |\hat{m}_{1j,w}(t) - m_{1j,w}(t)| > \epsilon \right) \leq P\left( |II_{n1}| + |II_{n2}| > \epsilon/2 \right) \\
\leq P\left( |II_{n2}| \geq \frac{(C_8L_n^{3/2} + C_9L_n^{5/2})\delta}{n} \right) + P\left( |II_{n1}| > \frac{4(1 + h_1)C_3^{-1}L_n^{3/2}\delta}{n} \right) \\
\leq 18L_n^2 \exp(-h_2L_n^{-3}n) + (10L_n + 6L_n^2) \exp\left( -\frac{\delta^2}{c_7L_n^{-1}n + c_8}\right),
$$

where $c_7 = \max(c_{17}, c_{19})$ and $c_8 = \max(c_{18}, c_{20})$.

Because the second result of the lemma can be proved as same as the proof of the third result. So, it suffices to show the third one. To this end, we may follow the above step with applying Lemmas 3, 5, 6 and part (ii) of Lemma 4. The main modification of the proof is about the derivation of inequality \((S0.17)\), whereas the term $E_{n,w}$ in \((S0.11)\) is replaced with $E_{n,w} = \frac{1}{n} \sum_{i=1}^{n} B_i(T_i)[w_i^2(\hat{F})I(\hat{Y}_i < \hat{Q}_{i}, Y_i < \hat{Q}_{i,Y})]$. Let $\varphi_{ik} = \frac{1}{n} \sum_{i=1}^{n} B_k(T_i)[w_i^2(\hat{F})I(\hat{Y}_i < \hat{Q}_{i,Y}) - w_i^2(F)I(\hat{Y}_i < \hat{Q}_{i,Y})]$, $\varphi_{ik}^{(1)} = 4\eta n^{-i}\tilde{b}_1^{-1}B_k(T_i)/(1 - \tau)$, $\varphi_{ik}^{(2)} = B_k(T_i)I\{Q_{i,Y} - \eta n^{-i}\tilde{b}_2^{-1} < Y_i \leq \hat{Q}_{i,Y} - \eta n^{-i}\tilde{b}_2^{-1}\}$, $\varphi_{ik}^{(3)} = 3B_k(T_i)I\{Q_{i,Y} - \eta n^{-i}\tilde{b}_3^{-1} < Z_i \leq Q_{i,Y} - \eta n^{-i}\tilde{b}_3^{-1}\}$, $\varphi_{ik}^{(4)} = 2B_k(T_i)I\{F^{-1}(\tau - \eta n^{-i}\tilde{c}_f^{-1}\tilde{b}_4^{-1}) \leq Z < F^{-1}(\tau + \eta n^{-i}\tilde{c}_f^{-1}\tilde{b}_4^{-1})\}$ and denote $\mu_k^{(l)} = \mathbb{E}\{\varphi_{ik}^{(l)}\}, l = 1, 2, 3$, where $\tilde{b}_1, \tilde{b}_2, \tilde{b}_3, \tilde{b}_4$ are some positive constants. Note that under conditions (D1) and (D2), there exist positive constants $M_{f2}, M_{h2}$ such that $f_{Y|T}(y) \leq M_{f1} < \infty$ and $h_{Z|T}(y) \leq M_{h2} < \infty$ uniformly in a neighbourhood of $Q_{i,Y}$, and also we can choose $\eta$ or large $n$ such that $4\eta n^{-i}\tilde{b}_1^{-1}/(1 - \tau) \leq 1$. Thus, we have $\max_k \mu_k^{(l)} \leq C_0\eta \tilde{b}_1^{-1}n^{-i}L_n^{-1}/(1 - \tau)$,
CQCSIS AND ITS APPLICATION IN SURVIVAL ANALYSIS

\[ \max_k \mu_k^{(2)} \leq 4C_0 \eta \tilde{b}_2^{-1} M_{f2} n^{-\tau} L_n^{-1}, \quad \max_k \mu_k^{(3)} \leq 12C_0 \eta \tilde{b}_3^{-1} M_{h2} n^{-\tau} L_n^{-1} \] and \[ |\varphi_{ik}^{(1)} - \mu_k^{(1)}| \leq 2, \quad |\varphi_{ik}^{(2)} - \mu_k^{(2)}| \leq 2, \quad |\varphi_{ik}^{(3)} - \mu_k^{(3)}| \leq 6 \] as well as \[ \text{Var}(\varphi_{ik}^{(1)}) \leq C_4 L_n^{-1}, \quad \text{Var}(\varphi_{ik}^{(2)}) \leq C_4 L_n^{-1}, \quad \text{Var}(\varphi_{ik}^{(3)}) \leq 9C_4 L_n^{-1}. \]

Then, taking \( \delta > M_7 n^{1-\tau} L_n^{-1} \) with \( M_7 = 2C_0 \eta \max(\tilde{b}_1^{-1}/(1-\tau), 4\tilde{b}_2^{-1} M_{f2}, 12\tilde{b}_3^{-1} M_{h2}, 8\tilde{b}_4^{-1} M_{h2}) \), it follows from Lemma \( \text{Lemma 3} \) that

\[ \max_{1 \leq k \leq L_n} P \left( \frac{1}{n} \sum_{i=1}^{n} \varphi_{ik}^{(1)} \geq \frac{\delta}{n} \right) \leq \max_{1 \leq k \leq L_n} P \left( \frac{1}{n} \sum_{i=1}^{n} \varphi_{ik}^{(1)} - \mu_k^{(1)} \geq \frac{\delta}{2n} \right) \leq 2 \exp \left( - \frac{\delta^2}{8C_4 L_n^{-1} n + 8\delta / 3} \right), \] \( \text{(S0.18)} \)

and

\[ \max_{1 \leq k \leq L_n} P \left( \frac{1}{n} \sum_{i=1}^{n} \varphi_{ik}^{(2)} \geq \frac{\delta}{n} \right) \leq 2 \exp \left( - \frac{\delta^2}{8C_4 L_n^{-1} n + 8\delta / 3} \right), \] \( \text{(S0.19)} \)

\[ \max_{1 \leq k \leq L_n} P \left( \frac{1}{n} \sum_{i=1}^{n} \varphi_{ik}^{(3)} \geq \frac{\delta}{n} \right) \leq 2 \exp \left( - \frac{\delta^2}{72C_4 L_n^{-1} n + 8\delta} \right). \] \( \text{(S0.20)} \)

Also, similar to \( \text{(S0.16)} \), we get

\[ \max_{1 \leq k \leq L_n} P \left( \frac{1}{n} \sum_{i=1}^{n} \varphi_{ik}^{(4)} \geq \frac{\delta}{n} \right) \leq 2 \exp \left( - \frac{\delta^2}{32C_4 L_n^{-1} n + 16\delta / 3} \right). \] \( \text{(S0.21)} \)

Hence, taking \( \tilde{b}_5^{-1} = \min(\tilde{b}_1^{-1}, \tilde{b}_2^{-1}, \tilde{b}_3^{-1}, c_f^{-1} \tilde{b}_4^{-1}, \tilde{b}_4^{-1}) \), applying part (ii) of Lemma \( \text{Lemma 4} \) with
\( \varepsilon = \eta n^{-i \delta^{-1}} \) and combining \((S0.18)-(S0.21)\), we obtain

\[
\max_{1 \leq k \leq L_n} P\left( \left| \frac{1}{n} \sum_{i=1}^{n} e_{ik} \right| \geq \frac{4\delta}{n} \right) \\
\leq \max_{1 \leq k \leq L_n} P\left( \frac{1}{n} \sum_{i=1}^{n} \varphi_{ik}^{(1)} \geq \frac{\delta}{n} \right) + \max_{1 \leq k \leq L_n} P\left( \frac{1}{n} \sum_{i=1}^{n} \varphi_{ik}^{(2)} \geq \frac{\delta}{n} \right) \\
+ \max_{1 \leq k \leq L_n} P\left( \frac{1}{n} \sum_{i=1}^{n} \varphi_{ik}^{(3)} \geq \frac{\delta}{n} \right) + \max_{1 \leq k \leq L_n} P\left( \frac{1}{n} \sum_{i=1}^{n} \varphi_{ik}^{(4)} \geq \frac{\delta}{n} \right) \\
\leq 8 \exp\left( -\frac{\delta^2}{72C_4 L_n^{-1} n + 8\delta} \right).
\]

Thus, the second result of the lemma is verified. Likewise, we can prove the second result.

\( \square \)

**Appendix C: Proofs of Theorems 2.1-2.3 and 3.1-3.3**

**Proof of Theorem 2.1:** Since the proof of this theorem may complete by following the proof of Theorem 1 in Liu, Li and Wu (2014), which consists of three steps. We here just give an outline of the proof.

*Step 1.* For any \( \epsilon > 0 \), derive an exponential upper bound of \( \sup_{t \in T} P(\left| \hat{\rho}_j(t) - \rho_j(t) \right| > \epsilon) \). To this end, observe that

\[
\hat{\rho}_j(t) = \frac{\hat{m}_{1j}(t) - \hat{m}_{2j}(t)\hat{m}_{4j}(t)}{\sqrt{\{\hat{m}_{2j}(t) - [\hat{m}_{2j}(t)]^2\} \{\hat{m}_{3j}(t) - [\hat{m}_{4j}(t)]^2\}}}
\]

and recall the definition of \( \rho(t) \). Iteratively applying Lemma 8, some tedious calcula-
tions yield that

\[
\sup_{t \in T} P\left( |\hat{\rho}_j(t) - \rho_j(t)| > \epsilon \right) \\
\leq 1218 L_n^2 \exp(-h_2 L_n^{-3} n) + (254 L_n + 456 L_n^2) \exp \left( -\frac{\delta^2}{c_{19} L_n^{-1} n + c_{20} \delta} \right) \\
= p_n(\epsilon) \quad \text{(say)} \quad \quad \quad \quad (S0.22)
\]

for some positive constants \( c_{19}, c_{20} \), where \( \delta \approx \epsilon n L_n^{-5/2} \).

\textbf{Step 2.} For any \( \epsilon > 0 \), derive the upper bound of \( P\left( \max_{1 \leq j \leq p} |\bar{u}_j - u_j| > \epsilon \right) \). For this, define \( u_{nj} = n^{-1} \sum_{i=1}^n \tilde{\rho}_j^2(T_i) \). Since \( |\bar{u}_j - u_j| \leq |\bar{u}_j - u_{nj}| + |u_{nj} - u_j| \), we have

\[
P\left( |\bar{u}_j - u_j| > \epsilon \right) \leq P\left( |\bar{u}_j - u_{nj}| > \epsilon/2 \right) + P\left( |u_{nj} - u_j| > \epsilon/2 \right), \quad \quad (S0.23)
\]

where the second term can be bounded by \( 2 \exp(-n\epsilon^2/64) \) according to Lemma \textbf{H} and the first term

\[
P\left( |\bar{u}_j - u_{nj}| > \epsilon/2 \right) \leq P\left( \frac{1}{n} \sum_{i=1}^n \left| \tilde{\rho}_j^2(T_i) - \rho_j^2(T_i) \right| > \epsilon/2 \right) \\
\leq \sum_{i=1}^n P\left( \left| \tilde{\rho}_j^2(T_i) - \rho_j^2(T_i) \right| > \epsilon/2 \right) \\
\leq n \sup_{t \in T} P\left( |\tilde{\rho}_j^2(t) - \rho_j^2(t)| > \epsilon/2 \right).
\]
Thus, by Lemmas 4, 8 and (S0.22), we obtain

\[
\max_{1 \leq j \leq p} P(|\tilde{u}_j - u_{nj}| > \epsilon/2) \\
\leq n[p_n(\epsilon/2) + p_n(\epsilon/8) + p_n(\epsilon/4)] \\
\leq 3654nL_n^2 \exp(-h_2L_n^{-3}n) + (762L_n + 1368L_n^2)n \exp\left(-\frac{\delta^2}{c_21L_n^{-1}n + c_22}\right)
\]

where \(c_{21}, c_{22}\) are positive constants, and \(\delta \approx \epsilon nL_n^{-5/2}\). Hence, under condition (C7), invoking (S0.23) and setting \(\epsilon = CL_n^{-2\kappa}\), we have

\[
P\left(\max_{1 \leq j \leq p} |\tilde{u}_j - u_j| > CL_n^{-2\kappa}\right) \\
\leq \sum_{j=1}^{p} P(|\tilde{u}_j - u_j| > CL_n^{-2\kappa}) \\
\leq O\left(pn\{L_n^2 \exp(-h_2L_n^{-3}n) + L_n \exp(-C_{10}L_n^{-2}n^{1-4\kappa})\}\right), \quad (S0.24)
\]

where \(C_{10}\) is some positive constant. This proves result (i).

**Step 3.** Prove result (ii) under condition (C6). Recall the definition of \(\hat{\mathcal{M}}\) and set

\(\nu_n = CL_n^{-2\kappa}\). Then, it follows from condition (C6) that

\[
P(\mathcal{M}_* \subset \hat{\mathcal{M}}) \\
\geq P\left(\min_{j \in \mathcal{M}_*} \tilde{u}_j > \nu_n\right) = P\left(\min_{j \in \mathcal{M}_*} u_j - \min_{j \in \mathcal{M}_*} \tilde{u}_j < \min_{j \in \mathcal{M}_*} u_j - \nu_n\right) \\
\geq P\left(\min_{j \in \mathcal{M}_*} (\tilde{u}_j - u_j) > \nu_n - \min_{j \in \mathcal{M}_*} u_j\right) \\
\geq P\left(\min_{j \in \mathcal{M}_*} u_j - \max_{j \in \mathcal{M}_*} |\tilde{u}_j - u_j| > \nu_n\right) \\
= 1 - P\left(\max_{j \in \mathcal{M}_*} |\tilde{u}_j - u_j| > \min_{j \in \mathcal{M}_*} u_j - \nu_n\right) \\
\geq 1 - O\left(s_nn\{L_n^2 \exp(-h_2L_n^{-3}n) + L_n \exp(-C_{10}L_n^{-2}n^{1-4\kappa})\}\right),
\]
where the last inequality is due to result (i). Thus, the result (ii) follows.

\[ \square \]

**Proof of Theorem 2.2:** According to the condition of Theorem 2.2, there exists some \( \delta_0 > 0 \) such that \( \min_{j \in \mathcal{M}} u_j - \max_{j \notin \mathcal{M}} u_j = \delta_0 \). Note that by Fatou’s Lemma, it follows that

\[
P \left( \liminf_{n \to \infty} \left\{ \min_{j \in \mathcal{M}_*} \hat{u}_j - \max_{j \notin \mathcal{M}_*} \hat{u}_j \right\} \leq 0 \right) \leq \lim_{n \to \infty} P \left( \min_{j \in \mathcal{M}_*} \hat{u}_j \leq \max_{j \notin \mathcal{M}_*} \hat{u}_j \right). \tag{S0.25} \]

For the term on the right hand side of last inequality, we have

\[
P \left( \min_{j \in \mathcal{M}_*} \hat{u}_j \leq \max_{j \notin \mathcal{M}_*} \hat{u}_j \right) = P \left( \left\{ \min_{j \in \mathcal{M}_*} \hat{u}_j - \min_{j \in \mathcal{M}_*} u_j \right\} \leq \left\{ \max_{j \notin \mathcal{M}_*} \hat{u}_j - \max_{j \notin \mathcal{M}_*} u_j \right\} - \delta_0 \right) \leq P \left( \min_{j \in \mathcal{M}_*} \left| \hat{u}_j - u_j \right| + \max_{j \notin \mathcal{M}_*} \left| \hat{u}_j - u_j \right| \geq \delta_0 \right) \leq P \left( \max_{1 \leq j \leq p} \left| \hat{u}_j - u_j \right| \geq \delta_0/2 \right),
\]

where, similar to (S0.24), we can derive

\[
P \left( \max_{1 \leq j \leq p} \left| \hat{u}_j - u_j \right| \geq \delta_0/2 \right) \leq \sum_{j=1}^{p} P \left( \left| \hat{u}_j - u_j \right| \geq \delta_0/2 \right) \leq O \left( pn \{ L_n^2 \exp(-h_2 L_n^{-3} n) + L_n \exp(-C_{11} \delta_0^2 L_n^{-4} n) \} \right) + 2p \exp(-\delta_0^2 n/64) \leq O \left( pL_n^2 n \exp(-C_{11} \delta_0^2 L_n^{-4} n) \right)
\]

for some positive constant \( C_{11} \). The last inequality goes to 0 as long as \( n \to \infty \) and
log \( p < C_1 L^{-4} n - 2 \log L - \log n \). This together with (S0.25) implies

\[
P \left( \liminf_{n \to \infty} \left\{ \min_{j \in \mathcal{M}_*} \hat{u}_j - \max_{j \not\in \mathcal{M}_*} \hat{u}_j \right\} > 0 \right) = 1.
\]

Hence, the ranking consistency is proved. \( \square \)

**Proof of Theorem 2.3:** By definition of \( \mu_j \) and condition (C5), we have

\[
\sum_{j=1}^{p} u_j = \sum_{j=1}^{p} \mathbb{E} \{ \rho_j^2(T) \} \leq (K_3 K_4)^{-1} \sum_{j=1}^{p} \mathbb{E} \{ \text{Cov} \{ I(Y_i > Q_{\tau, \nu}), X_j|T \} \}^2 \leq (K_3 K_4)^{-1} \mathbb{E} \{ \| b \|^2 \},
\]

which implies that the size of \( \{ j : u_j > \delta L_n n^{-2\kappa} \} \) cannot exceed \( O \left( L_n^{-1} n^{-2\kappa} \mathbb{E} \{ \| b \|^2 \} \right) \) for any \( \delta > 0 \). Thus, it follows that on the set \( \mathcal{A}_n = \left\{ \max_{1 \leq j \leq p} |\hat{u}_j - u_j| \leq \delta L_n n^{-2\kappa} \right\} \), the size of \( \{ j : \hat{u}_j > 2\delta L_n n^{-2\kappa} \} \) cannot exceed the size of \( \{ j : u_j > \delta L_n n^{-2\kappa} \} \), which is bounded by \( O \left( L_n^{-1} n^{-2\kappa} \mathbb{E} \{ \| b \|^2 \} \right) \). Then, taking \( \delta = C/2 \), we have

\[
P \left( |\hat{\mathcal{M}}| \leq O \left( n^{2\kappa} L_n^{-1} \mathbb{E} \{ \| b \|^2 \} \right) \right) \geq P(\mathcal{A}_n) \geq 1 - P \left( \max_{1 \leq j \leq p} |\hat{u}_j - u_j| > CL_n n^{-2\kappa} / 2 \right).
\]

The desired conclusion follows from part (i) of Theorem 2.1. \( \square \)

**Proofs of Theorems 3.1-3.3:** Because the proofs can be complete directly through following the steps of the proofs of Theorems 2.1-2.3. Hence, we omit details. \( \square \)
Bibliography


Xiaochao Xia (Corresponding author), College of Science, Huazhong Agricultural University, Wuhan, China.

E-mail: (xia_xiao_chao@126.com)

Jialiang Li, Department of Statistics and Applied Probability, National University of Singapore; Duke-NUS Graduate Medical School; Singapore Eye Research Institute.

E-mail: (stalj@nus.edu.sg)

Bo Fu, School of Data Science, Fudan University, Shanghai, China; Administrative Data Research Centre for England, University College London, UK.

E-mail: (fu@fudan.edu.cn)