Supplementary Material for “Smoothed and Corrected Score Approach to Censored Quantile Regression With Measurement Errors”

Yuanshan Wu, Yanyuan Ma, and Guosheng Yin

This supplementary material consists of theoretical proofs and additional numerical results.

Section A: Theoretical Proofs

We provide the proofs of Theorems 1 and 2 presented in Section 2.4 of the paper.

**Lemma 1** Under Assumptions 1, 2, and 4, \( \| \mu(b; h_n) - \mu_0(b) \| \leq O(h_n^r) \) and \( \| \tilde{\mu}(b; h_n) - \tilde{\mu}_0(b) \| \leq O(h_n^r) \) uniformly over \( b \in \mathcal{B}(d_0) \) as \( n \to \infty \).

**Proof.** We rewrite

\[
\| \mu(b; h_n) - \mu_0(b) \| \\
= \left\| E \left[ \Delta \left\{ W + \frac{\partial G(O, b; h_n)}{\partial b} \right\} \right] - \mu_0(b) \right\| \\
= \left\| E(\Delta W) + E \left[ \Delta \frac{\partial E\{G(O, b; h_n) \mid U\}}{\partial b} \right] - \mu_0(b) \right\| \\
= \left\| \frac{\partial}{\partial b} E \left\{ \Delta(X - Z^Tb)K \left( \frac{X - Z^Tb}{h_n} \right) \right\} + E\{\Delta Z I(X > Z^Tb)\} \right\},
\]

where the first term can be expanded as

\[
\frac{\partial}{\partial b} E \left\{ \Delta(X - Z^Tb)K \left( \frac{X - Z^Tb}{h_n} \right) \right\} \\
= -E \left\{ \Delta ZK \left( \frac{X - Z^Tb}{h_n} \right) \right\} - E \left\{ \Delta Z \frac{X - Z^Tb}{h_n} K^{(1)} \left( \frac{X - Z^Tb}{h_n} \right) \right\}.
\]
Thus, we have
\[
\|\mu(b; h_n) - \mu_0(b)\| \leq \left\| E \left\{ \Delta Z K \left( \frac{X - Z^T b}{h_n} \right) \right\} - E\{\Delta ZI(X > Z^T b)\} \right\| \\
+ \left\| E \left\{ \Delta Z \frac{X - Z^T b}{h_n} K^{(1)} \left( \frac{X - Z^T b}{h_n} \right) \right\} \right\| \\
\equiv \|I_{n1}(b)\| + \|I_{n2}(b)\|. \quad (S.1)
\]

It is easy to show that
\[
I_{n2}(b) = E\left[ Z E \left\{ \Delta \frac{X - Z^T b}{h_n} K^{(1)} \left( \frac{X - Z^T b}{h_n} \right) \right\} | Z \right]\]
\[
= E \left\{ Z \int_{-\infty}^{\infty} K^{(1)}(t)h_n t f_T(h_n t + Z^T b|Z)S_C(h_n t + Z^T b|Z) dt \right\}.
\]

Using the Taylor series expansion, we have that
\[
f_T(h_n t + Z^T b|Z) = \sum_{i=0}^{r_0 - 2} \frac{1}{i!} f_T^{(i)}(Z^T b|Z)(h_n t)^i + \frac{1}{(r_0 - 1)!} f_T^{(r_0 - 1)}(t^*|Z)(h_n t)^{r_0 - 1},
\]
and for every \(i = 0, \ldots, r_0 - 2\),
\[
S_C(h_n t + Z^T b|Z) = \sum_{j=0}^{r_0 - i - 2} \frac{1}{j!} S_C^{(j)}(Z^T b|Z)(h_n t)^j + \frac{1}{(r_0 - i - 1)!} S_C^{(r_0 - i - 1)}(t_i^*|Z)(h_n t)^{r_0 - i - 1},
\]

where both \(t^*\) and \(t_i^*\) are between \(h_n t\) and \(h_n t + Z^T b\). Further computation leads to
\[
h_n t f_T(h_n t + Z^T b|Z)S_C(h_n t + Z^T b|Z)\]
\[
= \left\{ \frac{1}{(r_0 - 1)!} f_T^{(r_0 - 1)}(t^*|Z)S_C(h_n t + Z^T b|Z) \right. \\
+ \sum_{i=0}^{r_0 - 2} \frac{1}{i!(r_0 - i - 1)!} f_T^{(i)}(Z^T b|Z)S_C^{(r_0 - i - 1)}(t_i^*|Z)(h_n t)^{r_0 - i - 1} \right\} (h_n t)^{r_0} \\
+ \sum_{i=0}^{r_0 - 2} \sum_{j=0}^{r_0 - i - 2} \frac{1}{i!j!} f_T^{(i)}(Z^T b|Z)S_C^{(j)}(Z^T b|Z)(h_n t)^{i+j+1}. \quad (S.2)
\]

For any \(\eta > 0\), we decompose \(I_{n2}(b) = I_{n21}(b) + I_{n22}(b)\), where
\[
I_{n21}(b) = E\left\{ \int_{[h_n t] \leq \eta} K^{(1)}(t)h_n t f_T(h_n t + Z^T b|Z)S_C(h_n t + Z^T b|Z) dt \right\} \quad (S.3)
\]
and
\[ I_{n22}(b) = E \left\{ Z \int_{|h_n t| > \eta} K^{(1)}(t) h_n t f_T(h_n t + Z^T b | Z) S_C(h_n t + Z^T b | Z) dt \right\}. \] (S.4)

Substituting (S.2) into (S.3) and using Assumptions 1-(ii) and 2, and the Lebesgue dominated convergence theorem, we have that
\[ h_n^{-r_0} I_{n21}(b) \rightarrow \int_{-\infty}^{\infty} t^{r_0} K^{(1)}(t) dt \sum_{i=0}^{r_0-1} \frac{1}{i!} \frac{1}{(r_0 - 1 - i)!} E \left\{ Z f_T^{(i)}(Z^T b | Z) S_C^{(r_0-i-1)}(Z^T b | Z) \right\} \]
uniformly over \( b \in B(d_0) \), as \( n \rightarrow \infty \). Furthermore, in view of (S.4) and Assumption 1-(iii), we can find some finite constant \( M > 0 \) such that
\[ \sup_{b \in B(d_0)} h_n^{-r_0} \| I_{n21}(b) \| \leq M \| Z \| \int_{|h_n t| > \eta} |t^{r_0} K^{(1)}(t)| dt \rightarrow 0. \]

Hence, we have shown that
\[ \sup_{b \in B(d_0)} \| I_{n2}(b) \| = O(h_n^{r_0}) \] (S.5)
under Assumptions 2, 4-(i), and 4-(ii).

Next we show that \( \sup_{b \in B(d_0)} \| I_{n1}(b) \| = O(h_n^{r_0}) \). Using the Taylor series expansion and integration by parts, we have
\[ I_{n1}(b) = E \left[ Z E \left( \Delta \left\{ K \left( \frac{X - Z^T b}{h_n} \right) - I(X > Z^T b) \right| Z \right) \right] \]
\[ = E \left[ Z \int_{-\infty}^{\infty} \left\{ K \left( \frac{t - Z^T b}{h_n} \right) - I(t > Z^T b) \right\} dF_{X,\Delta=1}(t|Z) \right] \]
\[ = E \left[ Z \int_{-\infty}^{\infty} \left\{ K(t) - I(t > 0) \right\} dF_{X,\Delta=1}(h_n t + Z^T b | Z) \right] \]
\[ = E \left[ Z \left\{ F_{X,\Delta=1}(Z^T b | Z) - \int_{-\infty}^{\infty} F_{X,\Delta=1}(h_n t + Z^T b | Z) K^{(1)}(t) dt \right\} \right] \]
\[ = -E \left\{ Z \int_{-\infty}^{\infty} (1/r_0! F_{X,\Delta=1}^{(r_0)}(t | Z)(h_n t)^{r_0} K^{(1)}(t) dt \right\}, \]
where \( t^\dagger \) is between \( h_n t + Z^T b \) and \( Z^T b \). Based on a similar decomposition as in \( I_{n2}(b) \), we also have
\[ \sup_{b \in B(d_0)} \| I_{n1}(b) \| = O(h_n^{r_0}). \] (S.6)
Hence, combining (S.1), (S.5), and (S.6), we conclude that

$$\sup_{b \in B(d_0)} \| \mu(b; h_n) - \mu_0(b) \| \leq O(h_n^m).$$

Similarly, we have

$$\| \tilde{\mu}(b; h_n) - \tilde{\mu}_0(b) \| = \left\| \frac{\partial}{\partial b} E \left\{ (X - Z^T b) K \left( \frac{X - Z^T b}{h_n} \right) \right\} + E \{ Z I(X \geq Z^T b) \} \right\|.$$

Thus, employing the arguments similar to the proof of the first part of the lemma, we have

$$\sup_{b \in B(d_0)} \| \tilde{\mu}(b; h_n) - \tilde{\mu}_0(b) \| \leq O(h_n^m),$$

which completes the proof.

Proof of Theorem 1.

Without loss of generality, assume that 0 = \( \tau_0 < \tau_1 < \cdots < \tau_{q_n} \equiv \tau_U \) < 1 are equally spaced between 0 and \( \tau_U \). Thus, \( a_n = \tau_1 - \tau_0 = \tau_1 \) and \( q_n = \tau_U/a_n \). Let \( b_n = a_n/(1 - \tau_U) \), then \( 0 < H(\tau_j) - H(\tau_{j-1}) \leq b_n \) for \( j = 1, \ldots, q_n \). Since \( a_n \to 0 \) as \( n \to \infty \), we have \( q_n \to \infty \) and \( b_n \to 0 \).

Let \( \alpha_0(\tau) = \mu_0\{ \beta_0(\tau) \}, \hat{\alpha}(\tau) = \mu_0\{ \hat{\beta}(\tau) \} \), and \( \mathcal{A}(d) = \{ \mu_0(b) : b \in \mathcal{B}(d) \} \). We start by establishing that under Assumptions 4-(i), and 4-(ii), \( \mu_0 \) is a one-to-one map from \( \mathcal{B}(d_0) \) to \( \mathcal{A}(d_0) \). To see this, suppose that there are \( b_1 \) and \( b_2 \in \mathcal{B}(d_0) \) such that \( \mu_0(b_1) = \mu_0(b_2) \). Then

$$0 = (b_1 - b_2)^T \{ \mu_0(b_1) - \mu_0(b_2) \}$$

$$= E \left[ (Z^T b_1 - Z^T b_2) \left\{ F_{X, \Delta=1}(Z^T b_1 | Z) - F_{X, \Delta=1}(Z^T b_2 | Z) \right\} \right]$$

$$= E \left\{ (Z^T b_1 - Z^T b_2)^2 f_{X, \Delta=1}(Z^T b^* | Z) \right\},$$

where \( Z^T b^* \) is between \( Z^T b_1 \) and \( Z^T b_2 \). According to Assumption 4-(ii), the above equality holds if and only if \( Z^T b_1 = Z^T b_2 \) with probability one. It then follows from the positive
definiteness of \( E(\mathbf{ZZ}^T) \) in Assumption 4-(i) that \( b_1 = b_2 \). Therefore, there exists an inverse function of \( \mu_0 \), denoted by \( \kappa \), from \( A(d_0) \) to \( B(d_0) \), such that \( \kappa(\mu_0(b)) = b \) for any \( b \in B(d_0) \).

The estimation procedure from (2.6) in the paper yields

\[
n^{-1} \sum_{i=1}^{n} \sum_{i=1}^{n} \Delta_i \tilde{g} \{ \mathcal{O}_i; \hat{\beta}(\tau_j); h_n \} = n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau_j} g \{ \mathcal{O}_i; \hat{\beta}(u); h_n \} dH(u) \tag{S.7}
\]

with \( \hat{\beta}(\tau) = \hat{\beta}(\tau_{j-1}) \) for \( \tau_{j-1} = \tau < \tau_j, j = 1, \ldots, q_n \). The martingale property ensures that

\[
E \left[ ZN \{ Z^T \beta_0(\tau_j) \} \right] = E \left[ Z \int_{0}^{\tau_j} I \{ X \geq Z^T \beta_0(u) \} dH(u) \right]. \tag{S.8}
\]

The difference of the left-hand sides of (S.7) and (S.8) is given by

\[
n^{-1} \sum_{i=1}^{n} \Delta_i \tilde{g} \{ \mathcal{O}_i; \hat{\beta}(\tau_j); h_n \} - E \left[ ZN \{ Z^T \beta_0(\tau_j) \} \right] = v_n \{ \hat{\beta}(\tau_j); h_n \} + \mu \{ \hat{\beta}(\tau_j); h_n \} - \mu_0 \{ \beta_0(\tau_j) \}, \tag{S.9}
\]

where \( v_n(b; h_n) = n^{-1} \sum_{i=1}^{n} \Delta_i \tilde{g} \{ \mathcal{O}_i; b; h_n \} - \mu(b; h_n) \). Similarly, the difference of the right-hand sides of (S.7) and (S.8) is given by

\[
n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau_j} g \{ \mathcal{O}_i; \hat{\beta}(u); h_n \} dH(u) - E \left[ Z \int_{0}^{\tau_j} I \{ X \geq Z^T \beta_0(u) \} dH(u) \right] = \int_{0}^{\tau_j} \tilde{v}_n \{ \hat{\beta}(u); h_n \} dH(u) + \int_{0}^{\tau_j} \left[ \tilde{\mu} \{ \hat{\beta}(u); h_n \} - \tilde{\mu}_0 \{ \beta_0(u) \} \right] dH(u), \tag{S.10}
\]

where \( \tilde{v}_n(b; h_n) = n^{-1} \sum_{i=1}^{n} g \{ \mathcal{O}_i; b; h_n \} - \tilde{\mu}(b; h_n) \).

Obviously, the right-hand sides of (S.9) and (S.10) are equal. Hence, we have

\[
\mu_0 \{ \hat{\beta}(\tau_j) \} - \mu_0 \{ \beta_0(\tau_j) \} = -v_n \{ \hat{\beta}(\tau_j); h_n \} + \int_{0}^{\tau_j} \tilde{v}_n \{ \hat{\beta}(u); h_n \} dH(u) + \mu_0 \{ \hat{\beta}(\tau_j); h_n \} - \mu_0 \{ \hat{\beta}(\tau_j); h_n \} + \int_{0}^{\tau_j} \left[ \tilde{\mu} \{ \hat{\beta}(u); h_n \} - \tilde{\mu}_0 \{ \beta_0(u) \} \right] dH(u) + \int_{k=1}^{j} \int_{\tau_{k-1}}^{\tau_k} \left[ \tilde{\mu}_0 \{ \hat{\beta}(u) \} - \tilde{\mu}_0 \{ \beta_0(u) \} \right] dH(u). \tag{S.11}
\]

Consider \( \mathcal{G}_1 = \{ \Delta \tilde{g}(\mathcal{O}, b; h_n) \colon b \in \mathbb{R}^p \} \) and \( \mathcal{G}_2 = \{ g(\mathcal{O}, b; h_n) \colon b \in \mathbb{R}^p \} \). Both \( \mathcal{G}_1 \) and \( \mathcal{G}_2 \) are Glivenko–Cantelli (van der Vaart and Wellner, 1996) for any \( h_n \) under Assumptions 1-(i),...
for any \( x \) and 4-(i). It follows from the Glivenko–Cantelli theorem that \( \sup_{b \in \mathbb{R}^p} \| v_n(b; h_n) \| \to 0 \) almost surely and \( \sup_{b \in \mathbb{R}^p} \| \tilde{v}_n(b; h_n) \| \to 0 \) almost surely. Then, for any given \( c_1 > 0 \),

\[
\sup_j \left\| v_n(\hat{\beta}(\tau_j); h_n) - \int_0^{\tau_j} \tilde{v}_n(\hat{\beta}(u); h_n) \, dH(u) \right\| < c_1
\]

with probability one when \( n \) is sufficiently large. Because of Lemma 1, for any \( c_2 > 0 \), we have

\[
\sup_j \left\| \mu_0(\hat{\beta}(\tau_j)) - \mu(\hat{\beta}(\tau_j); h_n) + \int_0^{\tau_j} [\tilde{\mu}(\hat{\beta}(u); h_n) - \tilde{\mu}(\beta(u))] \, dH(u) \right\| \leq c_2 \{1 - \log(1 - \tau_U)\}.
\]

Assumption 3 implies that for some constant \( c_3 > 0 \),

\[
\| \mu_0(\beta_0(\tau)) - \mu_0(\beta_0(\tau')) \| \leq c_3 |\tau - \tau'|
\]

for any \( \tau, \tau' \in (0, \tau_U) \). In addition, under Assumption 4-(iv), there exists \( c_4 > 0 \) such that

\[
\| (B_0(b))^{-1} J_0(b)x \| \leq c_4 \| x \|
\]

for any \( x \in \mathbb{R}^p \) and \( b \in B(d_0) \).

Next we define a sequence \( \epsilon_0, \ldots, \epsilon_{q_n - 1} \) to bound \( \sup_{\tau \leq \tau_U} \| \mu_0(\hat{\beta}(\tau)) - \mu_0(\beta_0(\tau)) \|, l = 0, \ldots, q_n - 1 \). Let \( \epsilon_0 = c_3 a_n \) and \( \epsilon_l = c_1 + c_2 \{1 - \log(1 - \tau_U)\} + \sum_{k=0}^{l-1} \epsilon_k c_4 b_n + c_3 a_n \) for \( l = 1, \ldots, q_n - 1 \). It is easy to see from the definition of \( \epsilon_l \) that \( \epsilon_l - \epsilon_{l-1} = \epsilon_l c_4 b_n \), which implies that \( \epsilon_l = (1 + c_4 b_n)^{l-1} [c_1 + c_2 \{1 - \log(1 - \tau_U)\} + \epsilon_0 c_4 b_n + c_3 a_n] \). Given \( \lim_{n \to \infty} a_n = 0 \), \( q_n = \tau_U/a_n \), and \( b_n = a_n/(1 - \tau_U) \), we have \( \lim_{n \to \infty} (1 + c_4 b_n)^{q_n - 1} = \exp\{c_4 \tau_U/(1 - \tau_U)\} \).

Because \( \epsilon_l \) is increasing with \( l \), it is easy to see that for \( n \) greater than some \( N_0 \), \( c_1 \) can be chosen sufficiently small so that for \( l = 0, \ldots, q_n - 1 \), \( \epsilon_l \leq 2 \exp\{\tau_U/(1 - \tau_U)\} c_1 < d_0 \). To show that \( \sup_{\tau \leq \tau_U} \| \mu_0(\hat{\beta}(\tau)) - \mu_0(\beta_0(\tau)) \| \leq \epsilon_l \) for \( l = 0, \ldots, q_n - 1 \), we consider only \( n \geq N_0 \).

It is easy to see from the definition of \( \hat{\beta}(\tau) \) that \( \sup_{\tau \leq \tau_U} \| \mu_0(\hat{\beta}(\tau)) - \mu_0(\beta_0(\tau)) \| = \)}
\[ \sup_{\tau_0 \leq \tau < \tau_1} \| \mu_0 \{ \beta_0 (\tau) \} \| \leq c_3 a_n = \epsilon_0. \]

It follows from the definition of \( \kappa (\cdot) \) that

\[
\| \tilde{\mu}_0 \{ \tilde{\beta} (\tau) \} - \mu_0 \{ \beta_0 (\tau) \} \| \\
= \| \tilde{\mu}_0 [\kappa (\tilde{\alpha} (\tau))] - \mu_0 [\kappa (\alpha_0 (\tau))] \| \\
= \| J_0 [\kappa (\alpha^* (\tau))] [B_0 [\kappa (\alpha^* (\tau))]^{-1} [\mu_0 \{ \tilde{\beta} (\tau) \} - \mu_0 \{ \beta_0 (\tau) \}] \| \\
\leq c_4 \| \mu_0 \{ \tilde{\beta} (\tau) \} - \mu_0 \{ \beta_0 (\tau) \} \|. \tag{S.12}
\]

where \( \alpha^* (\tau) \) is between \( \tilde{\alpha} (\tau) \) and \( \alpha_0 (\tau) \). With \( j = 1 \), for \( \tau \in [\tau_0, \tau_1) \), (S.12) yields

\[ \| \tilde{\mu}_0 \{ \tilde{\beta} (\tau) \} - \mu_0 \{ \beta_0 (\tau) \} \| \leq c_4 \epsilon_0. \]

This, coupled with other simple algebraic manipulations, can show that the norm of the right side of (S.11) is not greater than \( c_1 + c_2 \{ 1 - \log (1 - \tau_U) \} + \epsilon_0 c_4 b_n \) for \( j = 1 \). Therefore,

\[
\sup_{\tau_1 \leq \tau < \tau_2} \| \mu_0 \{ \tilde{\beta} (\tau) \} - \mu_0 \{ \beta_0 (\tau) \} \| \\
\leq \| \mu_0 \{ \tilde{\beta} (\tau_1) \} - \mu_0 \{ \beta_0 (\tau_1) \} \| + \sup_{\tau_1 \leq \tau < \tau_2} \| \mu_0 \{ \beta_0 (\tau_1) \} - \mu_0 \{ \beta_0 (\tau) \} \| \\
\leq c_1 + c_2 \{ 1 - \log (1 - \tau_U) \} + \epsilon_0 c_4 b_n + c_3 a_n \\
= \epsilon_1.
\]

Inductively, for \( l = 2, \ldots, q_n - 1 \), using (S.12), similar arguments yield

\[
\sup_{\tau_1 \leq \tau < \tau_{l+1}} \| \mu_0 \{ \tilde{\beta} (\tau) \} - \mu_0 \{ \beta_0 (\tau) \} \| \\
\leq \| \mu_0 \{ \tilde{\beta} (\tau_1) \} - \mu_0 \{ \beta_0 (\tau_1) \} \| + \sup_{\tau_1 \leq \tau < \tau_{l+1}} \| \mu_0 \{ \beta_0 (\tau_1) \} - \mu_0 \{ \beta_0 (\tau) \} \| \\
\leq c_1 + c_2 \{ 1 - \log (1 - \tau_U) \} + \sum_{k=1}^{l} \int_{\tau_{k-1}}^{\tau_k} \| \tilde{\mu}_0 \{ \tilde{\beta} (u) \} - \mu_0 \{ \beta_0 (u) \} \| dH (u) + c_3 a_n \\
\leq c_1 + c_2 \{ 1 - \log (1 - \tau_U) \} + \sum_{k=1}^{l} c_4 \int_{\tau_{k-1}}^{\tau_k} \| \mu_0 \{ \tilde{\beta} (u) \} - \mu_0 \{ \beta_0 (u) \} \| dH (u) + c_3 a_n \\
\leq c_1 + c_2 \{ 1 - \log (1 - \tau_U) \} + \sum_{k=1}^{l} c_4 \epsilon_{k-1} b_n + c_3 a_n \\
= \epsilon_1.
\]
Because \(a_n, b_n, c_1, \) and \(c_2\) can be arbitrarily small as \(n\) increases, we have

\[
\sup_{0 \leq \tau \leq \tau_U} \| \mu_0 \{ \hat{\beta}(\tau) \} - \mu_0 \{ \beta_0(\tau) \} \| \to 0 \tag{S.13}
\]

in probability. Using the Taylor expansion of \(\kappa(\tilde{\alpha}(\tau))\) around \(\alpha_0(\tau)\) for \(\tau \in [\nu, \tau_U]\) for any \(0 < \nu \leq \tau_U\), we obtain from Assumption 4-(iii) that

\[
\| \hat{\beta}(\tau) - \beta_0(\tau) \| \leq \| B_0 \{ \beta_0(\tau) \}^{-1} \{ \tilde{\alpha}(\tau) - \alpha_0(\tau) \} \| + \| e_n^*(\tau) \|
\leq c_5 \| \mu_0 \{ \hat{\beta}(\tau) \} - \mu_0 \{ \beta_0(\tau) \} \| + \| e_n^*(\tau) \|
\]

where \(\sup_{\tau \in [\nu, \tau_U]} \| e_n^*(\tau) \| \to 0\) in probability and \(c_5 > 0\) does not depend on \(\tau\). This, coupled with (S.13), completes the proof.

**Proof of Theorem 2.**

Write the estimating equation (2.6) as \(n^{1/2} S_n(\beta; \tau; h_n) = 0\), where

\[
S_n(\beta; \tau; h_n) = n^{-1} \sum_{i=1}^{n} \Delta_i \tilde{g} \{ O_i, \beta(\tau); h_n \} - \int_{0}^{\tau} \tilde{g} \{ O_i, \beta(u); h_n \} dH(u).
\]

The estimating equation yields \(n^{1/2} S_n(\hat{\beta}(\tau_j; h_n) = 0\) for \(j = 1, \ldots, q_n\). Recall that \(\hat{\beta}(\tau) = \hat{\beta}(\tau_j)\) for \(\tau_j \leq \tau < \tau_{j+1}\). Thus, it follows from the integral mean value theorem that

\[
\sup_{\tau \in [\tau_j, \tau_{j+1})} n^{1/2} \left\| S_n(\hat{\beta}; \tau; h_n) - S_n(\hat{\beta}; \tau_j; h_n) \right\|
\]

\[
= \sup_{\tau \in [\tau_j, \tau_{j+1})} \left\| \int_{\tau_j}^{\tau} n^{-1/2} \sum_{i=1}^{n} \tilde{g} \{ O_i, \hat{\beta}(u); h_n \} dH(u) \right\|
\]

\[
\leq \sup_{\tau \in [\tau_j, \tau_{j+1})} \left\| \int_{\tau_j}^{\tau} n^{-1} \sum_{i=1}^{n} \tilde{g} \{ O_i, \hat{\beta}(\tau_r); h_n \} \right\| n^{1/2} \frac{a_n}{1 - \tau_U}, \tag{S.14}
\]

where \(\tau_r\) is on the line segment of \(\tau_j\) and \(\tau\). Using the uniform law of large numbers, we also have that

\[
\sup_{b \in B(d_0)} \left\| n^{-1} \sum_{i=1}^{n} g \{ O_i, b; h_n \} - E[Z \{ 1 - F_X(Z^T b | Z) \}] + O(h_n^r) \right\| \to 0 \tag{S.15}
\]
almost surely. Let \( o_f(a) \) denote a term that converges uniformly to zero in probability in \( \tau \in I \) after being divided by \( a \). Observing that \( n^{1/2}a_n \to 0, \ h_n \to 0 \), and the boundness of \( Z \) under Assumption 4-(i), and combing (S.14) with (S.15), we have

\[
\sup_{\tau \in [\tau_j, \tau_{j+1}]} n^{1/2} \left\| S_n(\hat{\beta}; \tau; h_n) - S_n(\hat{\beta}; \tau_j; h_n) \right\| \leq \{E\|Z\| + O(h_n^a)\} n^{1/2}a_n + o_{[\tau_j, \tau_{j+1}]}(1) \to 0
\]

in probability. We thus have obtained \( n^{1/2}S_n(\hat{\beta}; \tau; h_n) = o_{[\nu, \tau_U]}(1) \). Then we can rewrite

\[
-n^{1/2}S_n(\beta_0, \tau; h_n) = n^{-1/2} \sum_{i=1}^n \left[ \Delta_i \hat{g}(\mathcal{O}_i, \hat{\beta}(\tau); h_n) - \Delta_i \hat{g}(\mathcal{O}_i, \beta_0(\tau); h_n) - \mu_0(\hat{\beta}(\tau)) + \mu_0(\beta_0(\tau)) \right]
\]

\[-n^{-1/2} \sum_{i=1}^n \int_0^\tau \left[ g(\mathcal{O}_i, \hat{\beta}(u); h_n) - g(\mathcal{O}_i, \beta_0(u); h_n) - \tilde{\mu}_0(\hat{\beta}(u)) + \tilde{\mu}_0(\beta_0(u)) \right] dH(u)
\]

\[-\int_0^\tau \left| J_0(\beta_0(u)) \right| B_0(\beta_0(u))^{-1} n^{1/2}[\mu_0(\hat{\beta}(u)) - \mu_0(\beta_0(u))] dH(u)
\]

\[+n^{1/2}[\mu_0(\hat{\beta}(\tau)) - \mu_0(\beta_0(\tau))] + o_{[\nu, \tau_U]}(1).
\]

(S.16)

Under Assumptions 1-(i), 3, and 4-(i), \( \mathcal{G}_3 = \{\Delta \hat{g}(\mathcal{O}, \beta_0(\tau); h_n): \tau \in (0, \tau_U)\} \) and \( \mathcal{G}_4 = \{g(\mathcal{O}, \beta_0(\tau); h_n): \tau \in (0, \tau_U)\} \) are both the Donsker classes for any \( h_n \). Thus, \( -n^{1/2}S_n(\beta_0, \tau; h_n) \) converges weakly to a zero-mean Gaussian process, \( \xi(\tau) \) say, over \( \tau \in [\nu, \tau_U] \).

Denote \( \sigma^2(\mathbf{b}) = \text{trace}(\text{Var}[\Delta \hat{g}(\mathcal{O}, \mathbf{b}; h_n) - \Delta \hat{g}(\mathcal{O}, \beta_0(\tau); h_n) - \mu_0(\mathbf{b}) + \mu_0(\beta_0(\tau))]) \). Next we will show that \( \sup_{\tau \in [\nu, \tau_U]} \sigma^2(\hat{\beta}(\tau)) \xrightarrow{p} 0 \) for any \( \nu \in (0, \tau_U) \).

Under Assumption 5-(i), we can find some \( C_\nu > 0 \) and \( N_\nu \) such that for \( n > N_\nu \),

\[
\sup_{\tau \in [\nu, \tau_U]} \|E[\hat{g}(\mathcal{O}, \beta_0(\tau); h_n)]\|^2 \leq C_\nu^2/2.
\]

Furthermore, it follows from Theorem 1 that \( \sup_{\tau \in [\nu, \tau_U]} \|\hat{\beta}(\tau) - \beta_0(\tau)\| \xrightarrow{p} 0 \). Hence, for any \( \epsilon \) and \( \eta > 0 \), there exists \( N_{\epsilon, \eta} \) such that for \( n > N_{\epsilon, \eta} \),

\[
P\left( \sup_{\tau \in [\nu, \tau_U]} \|\hat{\beta}(\tau) - \beta_0(\tau)\| > \sqrt{\eta}/C_\nu \right) < \epsilon.
\]
Consider the case sup_{τ ∈ [ν, τ_U]} ∥\hat{β}(τ) - β_0(τ)∥ ≤ \sqrt{n}/C_ν. Using the Taylor expansion and the equivalence of the finite-dimensional norms of the matrix, we have

\begin{align*}
\sigma^2\{\hat{β}(τ)\} &\leq \{\hat{β}(τ) - β_0(τ)\}^T E [g\{O, \beta^*_n(\tau); h_n\}] \{\hat{β}(τ) - β_0(τ)\} \\
&\leq \|E [g\{O, \beta^*_n(\tau); h_n\}] \| \|\hat{β}(τ) - β_0(τ)\|^2,
\end{align*}

where β^*_n(τ) is between \hat{β}(τ) and β_0(τ). Under Assumption 5-(ii), we have

\begin{align*}
sup_{τ ∈ [ν, τ_U]} E [g\{O, \beta^*_n(\tau); h_n\}] - E [g\{O, β_0(\tau); h_n\}] \leq C_ν^2/2.
\end{align*}

Hence, we have

\begin{align*}
sup_{τ ∈ [ν, τ_U]} \sigma^2\{\hat{β}(τ)\} \leq η.
\end{align*}

Therefore, we have shown that for any \epsilon and η > 0, there exists N^*_ε,η = max(N_ν, N_ε,η) such that for n > N^*_ε,η,

\begin{align*}
P\left( sup_{τ ∈ [ν, τ_U]} \left| \sigma^2\{\hat{β}(τ)\} - \eta \right| > \epsilon \right) < \epsilon.
\end{align*}

Hence, we can conclude that sup_{τ ∈ [ν, τ_U]} \sigma^2\{\hat{β}(τ)\} \xrightarrow{P} 0 for any ν ∈ (0, τ_U]. Following the arguments of Alexander (1984) and Lai and Ying (1988), we have

\begin{align*}
sup_{τ ∈ [ν, τ_U]} \left\| n^{-1/2} \sum_{i=1}^n \left[ \Delta_i g\{O_i, \hat{β}(τ); h_n\} - \Delta_i g\{O_i, β_0(τ); h_n\} - \mu_0(\hat{β}(τ)) + \mu_0(β_0(τ)) \right] \right\| \xrightarrow{P} 0.
\end{align*}

Likewise,

\begin{align*}
sup_{τ ∈ [ν, τ_U]} \left\| n^{-1/2} \sum_{i=1}^n \left[ g\{O_i, \hat{β}(τ); h_n\} - g\{O_i, β_0(τ); h_n\} - \tilde{μ}_0(\hat{β}(τ)) + \tilde{μ}_0(β_0(τ)) \right] \right\| \xrightarrow{P} 0.
\end{align*}

Therefore, (S.16) reduces to

\begin{align*}
-n^{1/2} S_n(β_0, τ; h_n) = - \int_0^τ [J_0(β_0(u))B_0(β_0(u)^{-1})] n^{1/2} [\mu_0(\hat{β}(u)) - μ_0(β_0(u))] dH(u) \\
+ n^{1/2} [\mu_0(\hat{β}(τ)) - μ_0(β_0(τ))] + o_{[ν, τ_U]}(1).
\end{align*}
Using the production integration theory (Gill and Johansen, 1990) and (S.17), we have

\[ n^{1/2}[\mu_0\{\hat{\beta}(\tau)\} - \mu_0\{\beta_0(\tau)\}] = \phi\{-n^{1/2}S_n(\beta_0, \tau; h_n)\} + o_{[\nu, \tau_U]}(1), \]  

(S.18)

where \( \phi \) is a linear map from \( \mathcal{F} \) to \( \mathcal{F} \) such that for any \( m \in \mathcal{F} \),

\[ \phi(m)(\tau) = \int_0^\tau I(s, \tau)dm(s) \]

with \( I(s, t) = \prod_{u \in (s, t]} I_p + J_0(\beta_0(u))B_0(\beta_0(u))^{-1}dH(u) \) and \( \mathcal{F} = \{m: [0, \tau_U] \rightarrow \mathbb{R}^p, \ m \) is left-continuous with right limit, \( m(0) = 0 \}, \) and \( I_p \) being the \( p \times p \) identity matrix. Assumption 4-(iii) implies that \( B_0(\beta_0(\tau))^{-1} \) is bounded away from zero and infinity for all \( \tau \in [\nu, \tau_U] \). It follows from (S.18), the Taylor expansion, and the continuous mapping theorem that \( n^{1/2}\{\hat{\beta}(\tau) - \beta_0(\tau)\} \) converges weakly to \( B_0(\beta_0(\tau))^{-1}\phi\{\xi(\tau)\} \) over \( \tau \in [\nu, \tau_U] \), which is also a mean zero Gaussian process using the linearity of the operator \( \phi \).

Section B: Additional Numerical Results

When the sample size was \( n = 200 \), we observed some non-convergent cases among the 500 simulations, which although were rare. Furthermore, non-convergence mainly occurred with bootstrap samples and, as a result, we present the robust summary statistics. In addition, when the sample size was increased to \( n = 600 \), there was no non-convergent case. The results with \( n = 200 \) are in the main text, while we present in Table S1 the simulation results for the log-transformed censored quantile regression model with \( n = 600 \) and normally heteroscedastic model errors. It can be seen that the smoothed and corrected quantile regression method works well with different distributions of covariate measurement errors. Alternatively, one can compromise by replacing the root with the solution that minimizes the \( L_2 \)-norm of the estimating function, hence brings the estimating equation value as close to zero as possible. Fixing \( n = 200 \) and employing the Nelder–Mead algorithm, we obtain the parameter estimate that minimizes the \( L_2 \)-norm of the estimating function, for which we did not encounter any non-convergent case. As shown in Figures S1 and S2, the estimates...
by minimizing the $L_2$-norm generally agree well with the proposed estimates, though, due to censoring, it shows some volatilities at the quantile levels higher than 0.7.

Since the extreme value error distribution is left-skewed and thus most of the data are concentrated on the right side, the estimates at lower quantile levels are less accurate, which would in turn affect the accuracy of the quantile estimates at the subsequent levels. However, when the sample size is increased to $n = 400$, the performance of the proposed method improved substantially. We summarize the results in Table S2, where we also present the estimates at the extreme quantile $\tau = 0.1$ as well as the results of $n = 200$ for convenience of comparison.

He, Yi, and Xiong (2007) proposed a simulation and extrapolation (SIMEX) method for the accelerated failure time (AFT) model with covariates subject to measurement errors. Their AFT model is postulated as

$$\log eT = TZ + \epsilon/\alpha,$$

where $\beta$ is the unknown regression parameter and $\epsilon$ is the model error independent of $Z$. Their method requires the distribution of $\epsilon$ completely known, while the scaling parameter $\alpha$ is unknown. Furthermore, the covariate measurement error $U$ is restricted to be normal with mean zero and a known diagonal covariance-variance matrix, which essentially leads to a homogeneously parametric regression model.

We generated the survival time $\tilde{T}$ from the model

$$\log \tilde{T} = -0.5 + Z + \sigma(Z)\epsilon,$$

where $\epsilon$ was generated from the standard normal or the extreme value distribution. We also considered the homogeneous error with $\sigma(Z) \equiv 1$ and the heteroscedastic error with $\sigma(Z) = 1 + 0.2Z$. Thus, $\beta_0 = -0.5$, $\beta_1 = 1$, and $\alpha = 1$ if $\sigma(Z) \equiv 1$. The remaining setups were kept the same as before. The simulation results are summarized in Table S3, from which we can see the SIMEX method produces reasonable estimates when the model
error is homogeneous. For the heteroscedastic model error, the coverage probability for the regression slope parameter is lower than the 95% nominal level and the scale parameter $\alpha$ is overestimated. This is mainly due to naively treating the heteroscedastic errors as homogeneous ones.

Transforming model (S.19) into quantile regression, we also applied our proposed smoothed and corrected quantile regression method, and the simulation results are presented in Table S4. It can be seen that our proposed method performs well throughout, while the naive method is still seriously biased even for the homogeneous model error.

When the measurement error is normally distributed, the proposed corrected function is a summation of an infinite series. In practice, as recommended by Stefanski (1989), the first two summands are used as an approximation, which is found to be adequate in our simulation studies. Alternatively, Wang, Stefanski, and Zhu (2012) proposed another corrected function for the normal measurement error. Their correction function involves an integral that has no closed form and thus requires numerical integration. Under the normal model error, we plot in Figure S3 the mean squared errors (MSEs) of the resulting estimators respectively using the naive method, the proposed correction method, and the integral correction method of Wang, Stefanski, and Zhu (2012). It can be seen that the proposed and integral methods are quite comparable, while the naive method produces seriously biased results.

We also considered the case with multiple covariates, some of which are measured with errors and some are not. The survival times $\tilde{T}$ were generated from the log-transformed linear model with heteroscedastic errors,

$$
\log \tilde{T} = -\frac{1}{2} + Z_2 + \frac{1}{3} Z_3 + \left(1 + \frac{1}{2} Z_1 + \frac{1}{5} Z_2 + \frac{1}{4} Z_3 \right) \epsilon,
$$

where the model error $\epsilon$ was generated from the standard normal distribution. The corresponding censored quantile regression model is given by

$$
Q_T(\tau|Z) = \beta_0(\tau) + \beta_1(\tau) Z_1 + \beta_2(\tau) Z_2 + \beta_3(\tau) Z_3,
$$
where $\mathbf{Z} = (1, Z_1, Z_2, Z_3)^T$, $\beta_0(\tau) = -0.5 + Q_\epsilon(\tau)$, $\beta_1(\tau) = Q_\epsilon(\tau)/2$, $\beta_2(\tau) = 1 + Q_\epsilon(\tau)/5$, and $\beta_3(\tau) = 1/3 + Q_\epsilon(\tau)/4$. The censoring times were generated in the same manner as before to yield a censoring rate of 20%. Let $\mathbf{W} = \mathbf{Z} + \mathbf{U}$, where

$$
\mathbf{U} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.25 \\ 0 & 0 & 0.25 & 0.25 \end{pmatrix}\right),
$$

which implies $Z_1$ is precisely observed while both $Z_2$ and $Z_3$ are measured with correlated errors. The comparisons in terms of biases and MSEs are presented in Figure S4. It can be seen that the proposed method remains satisfactory in this situation, while the naive method yields biased estimates for $\beta_2(\cdot)$ and $\beta_3(\cdot)$. It is also found that the naive method produces an unbiased estimator for $\beta_1(\cdot)$, which corresponds to the effect of the error-free covariate $Z_1$.

References

Alexander, K. S. (1984), “Probability Inequalities for Empirical Processes and a Law of the Iterated Logarithm,” *The Annals of Probability*, 12, 1041–1067.

Gill, R. D., and Johansen, S. (1990), “A Survey of Product-Integration With a View Toward Application in Survival Analysis,” *The Annals of Statistics*, 18, 1501–1555.

He, W., Yi, G. Y., and Xiong, J. (2007), “Accelerated Failure Time Models With Covariates Subject to Measurement Error,” *Statistics in Medicine*, 26, 4817–4832.

Lai, T. L., and Ying, Z. (1988), “Stochastic Integrals of Empirical-Type Process With Applications to Censored Data,” *Journal of Multivariate Analysis*, 27, 334–358.

Stefanski, L. A. (1989), “Unbiased Estimation of a Nonlinear Function of a Normal-Mean
With Application to Measurement Error Models,” *Communications in Statistics-Theory and Methods*, 18, 4335–4358.

van der Vaart, A. W., and Wellner, J. A. (1996), *Weak Convergence and Empirical Processes*, New York: Springer-Verlag.

Wang, H., Stefanski, L. A., and Zhu, Z. (2012), “Corrected-Loss Estimation for Quantile Regression With Covariate Measurement Errors,” *Biometrika*, 99, 405–421.
Table S1: Simulation results for the log-transformed censored quantile regression model with sample size \( n = 600 \), three different distributions of covariate measurement errors, and heteroscedastic model errors for \( \epsilon \sim N(0, 1) \)

| \( \tau \) | \( \beta_0(\cdot) \) | \( \beta_1(\cdot) \) |
|---|---|---|
| | True | Est | SE/ESE | CP(%) | True | Est | SE/ESE | CP(%) |
| \( U \sim L_1(0, 0.5^2) \) |
| 0.2 | -1.342 | -1.269 | 0.986 | 92.4 | 0.832 | 0.829 | 1.000 | 94.8 |
| 0.3 | -1.024 | -0.987 | 0.977 | 93.0 | 0.895 | 0.889 | 1.025 | 93.2 |
| 0.4 | -0.753 | -0.731 | 1.038 | 93.6 | 0.949 | 0.949 | 1.012 | 92.0 |
| 0.5 | -0.500 | -0.491 | 1.007 | 93.8 | 1.000 | 1.013 | 1.024 | 92.8 |
| 0.6 | -0.247 | -0.248 | 0.964 | 92.8 | 1.051 | 1.070 | 0.977 | 92.8 |
| 0.7 | 0.024 | 0.035 | 1.014 | 94.6 | 1.105 | 1.119 | 0.901 | 93.0 |
| \( U \sim N(0, 0.5^2) \) |
| 0.2 | -1.342 | -1.285 | 0.998 | 92.4 | 0.832 | 0.832 | 0.999 | 94.2 |
| 0.3 | -1.024 | -0.973 | 1.023 | 92.4 | 0.895 | 0.883 | 1.013 | 93.6 |
| 0.4 | -0.753 | -0.724 | 1.047 | 93.0 | 0.949 | 0.941 | 1.065 | 93.2 |
| 0.5 | -0.500 | -0.483 | 1.039 | 93.2 | 1.000 | 0.999 | 1.100 | 94.4 |
| 0.6 | -0.247 | -0.242 | 1.076 | 93.8 | 1.051 | 1.059 | 1.071 | 93.2 |
| 0.7 | 0.024 | 0.025 | 1.021 | 95.0 | 1.105 | 1.115 | 0.977 | 94.0 |
| \( U \sim \text{Unif}(-\sqrt{3}/2, \sqrt{3}/2) \) |
| 0.2 | -1.342 | -1.262 | 1.066 | 92.4 | 0.832 | 0.829 | 0.987 | 95.4 |
| 0.3 | -1.024 | -0.976 | 1.054 | 92.6 | 0.895 | 0.887 | 0.987 | 94.8 |
| 0.4 | -0.753 | -0.719 | 1.023 | 92.8 | 0.949 | 0.941 | 1.053 | 93.0 |
| 0.5 | -0.500 | -0.479 | 1.016 | 93.4 | 1.000 | 0.996 | 1.013 | 93.6 |
| 0.6 | -0.247 | -0.224 | 1.015 | 93.2 | 1.051 | 1.053 | 0.975 | 92.6 |
| 0.7 | 0.024 | 0.049 | 1.044 | 92.6 | 1.105 | 1.105 | 1.023 | 94.0 |

Note: SE/ESE is the ratio of the sampling standard error and the estimated (bootstrap) standard error, and CP is the coverage probability.
Table S2: Simulation results for the log-transformed censored quantile regression model with the heteroscedastic model error for $\epsilon$ from an extreme value distribution and measurement error $U \sim N(0, 0.5^2)$

| $\tau$ | $\beta_0(\cdot)$ True | Est | SE/ESE | CP(%) | $\beta_1(\cdot)$ True | Est | SE/ESE | CP(%) |
|--------|------------------------|-----|---------|-------|------------------------|-----|---------|-------|
| $n = 200$ |
| 0.1 | -2.750 | -2.332 | 1.073 | 79.4 | 0.550 | 0.534 | 1.036 | 89.4 |
| 0.2 | -2.000 | -1.823 | 1.008 | 85.6 | 0.700 | 0.680 | 1.042 | 94.0 |
| 0.3 | -1.531 | -1.453 | 1.026 | 91.0 | 0.794 | 0.778 | 1.050 | 94.4 |
| 0.4 | -1.172 | -1.154 | 0.957 | 93.8 | 0.866 | 0.863 | 1.018 | 94.6 |
| 0.5 | -0.867 | -0.895 | 0.950 | 94.2 | 0.927 | 0.946 | 0.957 | 94.6 |
| 0.6 | -0.587 | -0.635 | 0.923 | 93.2 | 0.983 | 1.018 | 1.013 | 93.0 |
| 0.7 | -0.314 | -0.349 | 0.876 | 95.2 | 1.037 | 1.066 | 0.974 | 94.8 |
| $n = 400$ |
| 0.1 | -2.750 | -2.546 | 0.988 | 86.0 | 0.550 | 0.541 | 0.975 | 92.0 |
| 0.2 | -2.000 | -1.898 | 1.082 | 89.2 | 0.700 | 0.699 | 1.101 | 94.2 |
| 0.3 | -1.531 | -1.474 | 1.004 | 93.2 | 0.794 | 0.777 | 1.014 | 94.6 |
| 0.4 | -1.172 | -1.121 | 0.949 | 94.0 | 0.866 | 0.849 | 1.032 | 94.4 |
| 0.5 | -0.867 | -0.843 | 0.935 | 94.0 | 0.927 | 0.921 | 1.050 | 94.8 |
| 0.6 | -0.587 | -0.593 | 0.947 | 94.4 | 0.983 | 1.000 | 0.984 | 94.2 |
| 0.7 | -0.314 | -0.329 | 1.074 | 94.4 | 1.037 | 1.073 | 1.023 | 94.6 |
Table S3: Simulation results of the simulation-extrapolation method in He, Yi, and Xiong (2007) for the accelerated failure time model with the covariate measurement error $U \sim N(0, 0.5^2)$

| $\epsilon$   | Error   | $\beta_0 = -0.5$ | $\beta_1 = 1$ | $\alpha = 1$ |
|--------------|---------|------------------|----------------|--------------|
|              | Est     | SE/ESE           | CP(%)          | Est          | SE/ESE | CP(%)          | Est  | SE   |
| $N(0, 1)$    |         |                  |                |              |        |                |      |      |
| Homo.        | -0.449  | 1.025            | 93.0           | 0.973        | 1.061   | 92.2           | 1.015 | 0.071 |
| Hetero.      | -0.402  | 0.938            | 93.8           | 0.923        | 1.032   | 87.4           | 1.313 | 0.088 |
| Extreme      |         |                  |                |              |        |                |      |      |
| Homo.        | -0.436  | 1.065            | 90.8           | 0.968        | 1.079   | 92.0           | 1.013 | 0.076 |
| Hetero.      | -0.464  | 0.964            | 95.2           | 0.952        | 1.021   | 91.8           | 1.301 | 0.095 |
Table S4: Simulation results of the proposed method under model (S.19) with the homogeneous model error (i.e., $\sigma(Z) \equiv 1$) and covariate measurement error $U \sim N(0, 0.5^2)$

| $\tau$ | $\beta_0(\cdot)$ | $\beta_1(\cdot)$ |
|--------|-------------------|-------------------|
|        | True | Est | SE/ESE | CP(%) | True | Est | SE/ESE | CP(%) |
|        | 0.2  | 1.342 | 1.281 | 0.953 | 93.2 | 1.000 | 0.965 | 0.947 | 93.2 |
|        | 0.3  | 1.024 | 1.001 | 0.975 | 94.0 | 1.000 | 0.984 | 0.901 | 94.4 |
|        | 0.4  | 0.753 | 0.747 | 0.965 | 94.0 | 1.000 | 1.004 | 0.946 | 94.4 |
|        | 0.5  | 0.500 | 0.511 | 0.990 | 94.0 | 1.000 | 1.021 | 0.974 | 93.8 |
|        | 0.6  | 0.247 | 0.257 | 1.020 | 93.2 | 1.000 | 1.032 | 0.940 | 93.2 |
|        | 0.7  | 0.024 | 0.029 | 0.990 | 92.6 | 1.000 | 1.035 | 0.983 | 93.0 |

$\epsilon \sim N(0, 1)$

| $\tau$ | $\beta_0(\cdot)$ | $\beta_1(\cdot)$ |
|--------|-------------------|-------------------|
|        | True | Est | SE/ESE | CP(%) | True | Est | SE/ESE | CP(%) |
|        | 0.2  | 2.000 | 1.828 | 1.039 | 86.2 | 1.000 | 0.949 | 1.025 | 92.4 |
|        | 0.3  | 1.531 | 1.478 | 1.034 | 92.2 | 1.000 | 0.974 | 0.958 | 93.0 |
|        | 0.4  | 1.172 | 1.189 | 0.967 | 94.0 | 1.000 | 1.012 | 0.882 | 94.8 |
|        | 0.5  | 0.867 | 0.929 | 0.943 | 94.0 | 1.000 | 1.041 | 0.864 | 94.8 |
|        | 0.6  | 0.587 | 0.660 | 0.959 | 93.4 | 1.000 | 1.057 | 0.847 | 93.2 |
|        | 0.7  | 0.314 | 0.372 | 0.941 | 94.4 | 1.000 | 1.065 | 0.885 | 93.8 |

$\epsilon \sim$ Extreme Value
Figure S1: Bias of the estimated quantile regression intercept using the proposed method (solid lines), that by minimizing the $L_2$-norm of the estimating function (dot-dashed lines), and the naive method (dashed lines) under three different model error distributions: normal, extreme value and $t_2$, and three different measurement error distributions: Laplace, normal and uniform, respectively.
Figure S2: Bias of the estimated quantile regression slope using the proposed method (solid lines), the estimate that minimizes the $L_2$-norm of the estimating function (dot-dashed lines), and the naive method (dashed lines) under three different model error distributions: normal, extreme value and $t_2$, and three different measurement error distributions: Laplace, normal and uniform, respectively.
Figure S3: Comparisons of mean squared errors (MSEs) of the estimated quantile regression intercept (left panel) and slope (right panel) for the proposed correction method (solid lines), the integral correction method (dot-dashed lines), and the naive method (dashed lines) for the normal covariate measurement error, when the model error is normally distributed.
Figure S4: Comparisons of biases and mean squared errors (MSEs) of the estimated quantile regression slope for the proposed method (solid lines) and the naive method (dashed lines), when both the covariate measurement error and model error are normally distributed.