Appendix

Proof of Theorem 1

We first show that for any given $\eta > 0$, there exists a large constant C such that

$$P\left(\inf_{||u||=C} L(\beta_0 + n^{-1/2}u) > L(\beta_0)\right) > 1 - \eta. \tag{A.1}$$

By Taylor's expansion and condition (C1), we have

$$\begin{aligned}
&L(\beta_0 + n^{-1/2}u) - L(\beta_0) = \sum_{i=1}^{n} \rho_{\tau,\lambda}\left(\frac{Y_i - X_i^\top(\beta_0 + n^{-1/2}u)}{S_n}\right) - \sum_{i=1}^{n} \rho_{\tau,\lambda}\left(\frac{Y_i - X_i^\top\beta_0}{S_n}\right) \\
&= -\frac{1}{\sqrt{n}}\frac{1}{S_n}\sum_{i=1}^{n}\psi_{\tau,\lambda}\left(\frac{Y_i - X_i^\top\beta_0}{S_n}\right)X_i^T u + \frac{1}{2S_n^2}u^T\left[\frac{1}{n}\sum_{i=1}^{n}\psi'_{\tau,\lambda}\left(\frac{Y_i - X_i^\top\beta_0}{S_n}\right)X_i^T X_i + o_p(1)\right]u \\
&\triangleq I_1 + I_2.
\end{aligned}$$

By $S_n \xrightarrow{P} \sigma > 0$, the law of large numbers, and classical central limit theorem, we obtain

$$\frac{1}{n}\sum_{i=1}^{n}\psi'_{\tau,\lambda}\left(\frac{Y_i - X_i^\top\beta_0}{S_n}\right)X_i^T X_i = E\left[\psi'_{\tau,\lambda}(\epsilon/\sigma)\right]E(X^T X) + o_p(1),$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_{\tau,\lambda}\left(\frac{Y_i - X_i^\top\beta_0}{S_n}\right)X_i = O_p(1).$$

Therefore, by choosing a sufficiently large C, I_2 dominates I_1 in $||u|| = C$. Since $E\left[\psi'_{\tau,\lambda}(\epsilon/\sigma)\right] > 0$, this completes the proof of Equation (A.1). Equation (A.1) implies with probability at least $1-\eta$ that exists a local minimum of $L(\beta)$ in the ball $\{\beta_0 + n^{-1/2}u : ||u|| \le C\}$. The proof of Theorem 1 is completed.

Proof of Theorem 2

Let

$$\phi_n(\theta) = \frac{1}{nS_n}\sum_{i=1}^{n}\psi_{\tau,\lambda}\left(\frac{Y_i - X_i^T\theta}{S_n}\right)X_i.$$

By Taylor's expansion, there exists a vector $\acute{\beta}_n$ on the line segment between β_0 and $\hat{\beta}_n$ such that

$$\phi_n(\hat{\beta}_n) = \phi_n(\beta_0) + \dot{\phi}_n(\beta_0)(\hat{\beta}_n - \beta_0) + \frac{1}{2}(\hat{\beta}_n - \beta_0)^\top\ddot{\phi}_n(\acute{\beta}_n)(\hat{\beta}_n - \beta_0),$$

where $\dot{\phi}_n(\cdot)$ and $\ddot{\phi}_n(\cdot)$ are the first-order derivative and the second-order derivatives of $\phi_n(\cdot)$. From (2.2), we have $\phi_n(\hat{\beta}_n) = 0$. By condition (C1) and Theorem 1, we have

$$\frac{S_n}{\sqrt{n}}\sum_{i=1}^{n}\psi_{\tau,\lambda}\left(\frac{Y_i - X_i^T\beta_0}{S_n}\right)X_i = \sqrt{n}(\hat{\beta}_n - \beta_0)\frac{1}{n}\sum_{i=1}^{n}\psi'_{\tau,\lambda}\left(\frac{Y_i - X_i^T\beta_0}{S_n}\right)X_i^T X_i + o_p(1).$$

Since $S_n \xrightarrow{P} \sigma$ as $n \to \infty$, thus,

$$\sqrt{n}(\hat{\beta}_n - \beta_0) \xrightarrow{D} N\left(\mathbf{0}, \frac{E\psi^2_{\tau,\lambda}(\epsilon/\sigma)}{(E\psi'_{\tau,\lambda}(\epsilon/\sigma))^2}\sigma^2(EX^TX)^{-1}\right).$$

Proof of Theorem 3

According to the definition of c_{n-m}, for all $||\beta|| = 1$, we can obtain

$$\sharp\{i : m+1 \le i \le n, \text{ and } |X_i^\top \theta| > 0\}/(n-m) \ge 1 - c_{n-m}.$$

By the assumption of Theorem 3, there exists $c_{n1} > c_{n-m}$ and $c_{n2} > c_{n-m}$ such that

$$\epsilon < (1-2c_{n1})/(2-2c_{n1}), a(\tau,\lambda) < (1-\epsilon)(2-2c_{n2}).$$

We take $c_n^* = \min\{c_{n1}, c_{n2}\}$, then $c_n^* > c_{n-m}$, and have

$$\epsilon < (1-2c_n^*)/(2-2c_n^*), a(\tau,\lambda) < (1-\epsilon)(2-2c_n^*). \tag{A.2}$$

By using a compacity argument (Yohai, 1987), we can find $\delta > 0$ such that

$$\inf_{||\beta||=1} \sharp\{i : m+1 \le i \le n, \text{ and } |X_i^\top \beta| > \delta\}/(n-m) \ge 1 - c_n^*.$$

According to Equation (A.2), we can obtain $1-\epsilon > 1/(2-2c_n^*)$. Therefore, we can find ζ such that $(1-\epsilon)(1-c_n^*) > 1-\zeta > 1/2$. Take $\Delta > 0$ which satisfies

$$1 < 1 + \Delta < \min\left\{\frac{(1-\epsilon)(1-c_n^*)}{1-\zeta}, \frac{(1-\epsilon)(2-2c_n^*)}{a(\tau,\lambda)}\right\}.$$

Denote

$$a_0 = \frac{(1-\zeta)a(\tau,\lambda)(1+\Delta)}{(1-\epsilon)(2-2c_n^*)}.$$

Then, we have $a_0 < \min\{1-\zeta, a(\tau,\lambda)/2\}$. Therefore, $m/n \le \epsilon$ implies

$$a_0(n-m)/n \ge (1-\epsilon)a_0 > (1-\zeta)a(\tau,\lambda)/(2-2c_n^*).$$

Since $\rho_{\tau,\lambda}(t)$ is a bounded, continuous, and even function, there exists $k_2 \ge 0$ such that $\rho_{\tau,\lambda}(k_2) = a_0/(1-\zeta)$. Let $C = (k_2 S_n + \max_{m+1\le i\le n}|Y_i|)/\delta$. Hence, $m/n \le \epsilon$ implies

$$\begin{aligned}\inf_{||\beta||\ge C}\sum_{i=1}^n \rho_{\tau,\lambda}(r_i(\beta)) &\ge \inf_{||\theta||=1}\sum_{i\in A}\rho_{\tau,\lambda}\left(\frac{|Y_i| - C|X_i^\top\theta|}{S_n}\right) \ge (n-m)(1-c_n^*)\rho_{\tau,\lambda}(k_2)\\ &= (n-m)(1-c_n^*)a_0/(1-\zeta) > na(\tau,\lambda)/2 \ge \sum_{i=1}^n \rho_{\tau,\lambda}(r_i(\bar{\beta}_n)),\end{aligned}$$

where $A = \{i : m+1 \le i \le n \textit{ and } |X_i^\top\theta| > \delta\}$.

For a contaminated sample $\mathbf{D}_n$ with $m/n \le \epsilon$, if $||\hat{\beta}_n|| \ge C$, we have

$$\sum_{i=1}^n \rho_{\tau,\lambda}(r_i(\hat{\beta}_n)) > \sum_{i=1}^n \rho_{\tau,\lambda}(r_i(\bar{\beta}_n)).$$

This is a contradiction with the fact that $\hat{\beta}_n$ minimizes $\sum_{i=1}^n \rho_{\tau,\lambda}(r_i(\beta))$ for $\beta \in R^p$. Note that $m/n \le \epsilon$, $\epsilon < (1-2c_{n-m})/(2-2c_{n-m})$, and $a(\tau,\lambda) < (1-\epsilon)(2-2c_{n-m})$. Therefore, we obtain

$$BP(\hat{\beta}_n; \mathbf{D}_{n-m}, \tau, \lambda) \ge \min\left\{BP(\bar{\beta}_n; \mathbf{D}_{n-m}), \frac{1-2c_{n-m}}{2-2c_{n-m}}, 1 - \frac{a(\tau,\lambda)}{2-2c_{n-m}}\right\}.$$

Since the sample $\mathbf{D}_n$ is in general position, we have $c_{n-m} = (p-1)/(n-m)$. Because $\bar{\beta}_n$ is a robust estimator with asymptotic breakdown point 1/2, therefore,

$$BP(\hat{\beta}_n; \mathbf{D}_{n-m}, \tau, \lambda) \ge \min\left\{1 - \frac{\kappa}{2}, 1 - \frac{\kappa}{2} a(\tau,\lambda)\right\}.$$