Supplemental Materials

Yuan Jiang and Chunming Zhang

We present the proofs for Theorems 3–4. Throughout the supplemental materials, $\|\cdot\|$ is used only for the L_2 norm, and $\|\cdot\|_1$ and $\|\cdot\|_{\infty}$ denote the L_1 norm and the L_{∞} norm respectively.

Proof of Theorem 3

Let $V_{ij}(\alpha) = q_1(Y_{ni}; X_{ij}\alpha)X_{ij}$. $\hat{\beta}_{n,j}^{\text{CR}}$ are the solutions of the following equations

$$
h_{n,j}(\alpha) = \frac{1}{n} \sum_{i=1}^{n} V_{ij}(\alpha) = 0, \quad j = 1, \ldots, p_n.
$$

With Condition 1(b), $h_{n,j}(\cdot)$ is an increasing function.

Part 1: $j \in \{1, ..., s_n\}.$

To prove $\hat{w}_{\text{max}}^{(I)} = O_P(1/\gamma_n^{(I)})$, it suffices to prove that there exists a small enough $\delta > 0$ such that

$$
P\left(\min_{1\leq j\leq s_n}|\widehat{\beta}_{n,j}^{\text{CR}}| \leq \gamma_n^{(1)}\delta\right) = o(1). \tag{1}
$$

Since $h_{n,j}(\cdot)$ is an increasing function and $h_{n,j}(\widehat{\beta}_{n,j}^{\text{CR}}) = 0$,

$$
P(|\widehat{\beta}_{n,j}^{\text{CR}}| \le \gamma_n^{(1)}\delta) \le P\{h_{n,j}(-\gamma_n^{(1)}\delta) \le 0 \le h_{n,j}(\gamma_n^{(1)}\delta)\}.
$$
 (2)

By Taylor's expansion,

$$
V_{ij}(\pm \gamma_n^{(1)}\delta) = q_1(Y_{ni}; 0)X_{ij} + (\pm \gamma_n^{(1)}\delta)q_2(Y_{ni}; \pm \gamma_n^{(1)}\delta^* X_{ij})X_{ij}^2,
$$

with $\delta^* \in (0, \delta)$. Let $\mu_0 = F^{-1}(0)$ and $C_0 = q''(\mu_0)/F'(\mu_0) \neq 0$, thus

$$
E\{V_{1j}(\pm \gamma_n^{(1)}\delta)\} = C_0 E(Y_n X_j) + (\pm \gamma_n^{(1)}\delta) E\{q_2(Y_n; \pm \gamma_n^{(1)}\delta^* X_j)X_j^2\}.
$$

Because $|E(Y_n X_j)| \geq c\gamma_n^{(1)}$, and $\max_{1 \leq j \leq s_n} |E\{q_2(Y_n; \pm \gamma_n^{(1)} \delta^* X_j) X_j^2\}|$ is bounded, we can choose δ small enough, such that for $1 \leq j \leq s_n$,

$$
|E\{V_{1j}(\pm \gamma_n^{(I)}\delta)\}| \ge \frac{1}{2}|C_0||E(Y_nX_j)| \ge \frac{c}{2}|C_0|\gamma_n^{(I)}
$$

and
$$
\text{sign}[E\{V_{1j}(\gamma_n^{(I)}\delta)\}] = \text{sign}[E\{V_{1j}(-\gamma_n^{(I)}\delta)\}].
$$

Assuming $E\{V_{1j}(\gamma_n^{(1)}\delta)\} < 0$ and $E\{V_{1j}(-\gamma_n^{(1)}\delta)\} < 0$ without loss of generality,

$$
P\{h_{n,j}(-\gamma_n^{(I)}\delta) \le 0 \le h_{n,j}(\gamma_n^{(I)}\delta)\}
$$

\n
$$
\le P\left(\sum_{i=1}^n [V_{ij}(\gamma_n^{(I)}\delta) - E\{V_{ij}(\gamma_n^{(I)}\delta)\}] \ge -nE\{V_{1j}(\gamma_n^{(I)}\delta)\}\right)
$$

\n
$$
\le P\left(\sum_{i=1}^n [V_{ij}(\gamma_n^{(I)}\delta) - E\{V_{ij}(\gamma_n^{(I)}\delta)\}] \ge \frac{c}{2}|C_0|n\gamma_n^{(I)}\right)
$$

\n
$$
\le 2 \exp\left(\frac{-c^2 C_0^2 n^2 \gamma_n^{(I)2}/4}{C_1 n + C_2 c|C_0|n\gamma_n^{(I)}/2}\right),
$$
\n(3)

where the last inequality can be obtained similar to proving Lemma 1 (with possibly different C_1 and C_2). By (2), (3) and Bonferroni inequality, for a small enough $\delta > 0$,

$$
P\Big(\min_{1 \le j \le s_n} |\widehat{\beta}_{n,j}^{\text{CR}}| \le \gamma_n^{(1)} \delta\Big) \le 2s_n \exp\left(\frac{-c^2 C_0^2 n^2 \gamma_n^{(1)2} / 4}{C_1 n + C_2 c |C_0| n \gamma_n^{(1)} / 2}\right) = o(1). \tag{4}
$$

The equality in (4) follows from $\gamma_n^{(I)} = O(1)$, $\sqrt{n} \gamma_n^{(I)} \to \infty$ and $\log(s_n) = o(n \gamma_n^{(I)2})$.

Part 2: $j \in \{s_n+1, \ldots, p_n\}.$ To prove $(\hat{w}_{\min}^{(II)})^{-1} = o_P(\gamma_n^{(II)})$, it suffices to prove that for any $\epsilon > 0$,

$$
P\left(\max_{s_n+1\leq j\leq p_n}|\widehat{\beta}_{n,j}^{\text{CR}}|\geq \gamma_n^{\text{(II)}}\epsilon\right)=o(1). \tag{5}
$$

As $h_{n,j}(\widehat{\beta}_{n,j}^{\text{CR}}) = 0$, it follows that

$$
P\left(|\widehat{\beta}_{n,j}^{\text{CR}}| \ge \gamma_n^{\text{(II)}} \epsilon\right) \le P\{h_{n,j}(\gamma_n^{\text{(II)}} \epsilon) \le 0\} + P\{h_{n,j}(-\gamma_n^{\text{(II)}} \epsilon) \ge 0\}.
$$
 (6)

Similar to Part 1,

$$
E\{V_{1j}(\gamma_n^{(II)}\epsilon)\} = C_0 E(Y_n X_j) + \epsilon \gamma_n^{(II)} E\{q_2(Y_n; \gamma_n^{(II)}\epsilon^* X_j) X_j^2\},
$$

with $\epsilon^* \in (0, \epsilon)$. Since $E(Y_n X_j) = o(\gamma_n^{(\text{II})})$ and $E\{q_2(Y_n; \gamma_n^{(\text{II})} \epsilon^* X_j) X_j^2\} \ge \eta$,

$$
E\{V_{1j}(\gamma_n^{(\text{II})}\epsilon)\}\geq \epsilon\eta\gamma_n^{(\text{II})}/2, \text{ as } n\to\infty.
$$

Again by an application of Bernstein's inequality as in (3) , for large n,

$$
P\{h_{n,j}(\gamma_n^{(\text{II})}\epsilon) \le 0\} = P\left(\frac{1}{n}\sum_{i=1}^n [V_{ij}(\gamma_n^{(\text{II})}\epsilon) - E\{V_{ij}(\gamma_n^{(\text{II})}\epsilon)\}] \le -E\{V_{1j}(\gamma_n^{(\text{II})}\epsilon)\}\right)
$$

$$
\le P\left(\sum_{i=1}^n [V_{ij}(\gamma_n^{(\text{II})}\epsilon) - E\{V_{ij}(\gamma_n^{(\text{II})}\epsilon)\}] \le -\epsilon \eta n \gamma_n^{(\text{II})}/2\right)
$$

$$
\le 2 \exp\left(\frac{-\epsilon^2 \eta^2 n^2 \gamma_n^{(\text{II})} / 4}{C_1 n + C_2 \epsilon \eta n \gamma_n^{(\text{II})}/2}\right).
$$
 (7)

Similarly,

$$
P\{h_{n,j}(-\gamma_n^{(\text{II})}\epsilon) \ge 0\} \le 2 \exp\left(\frac{-\epsilon^2 \eta^2 n^2 \gamma_n^{(\text{II})2} / 4}{C_1 n + C_2 \epsilon \eta n \gamma_n^{(\text{II})} / 2}\right).
$$
 (8)

Thus by (6), (7), (8) and Bonferroni inequality,

$$
P\left(\max_{s_n+1\leq j\leq p_n}|\widehat{\beta}_{n,j}^{\text{CR}}|\geq \gamma_n^{\text{(II)}}\epsilon\right)\leq 4(p_n-s_n)\exp\left(\frac{-\epsilon^2\eta^2n^2\gamma_n^{\text{(II)}2}/4}{C_1n+C_2\epsilon\eta n\gamma_n^{\text{(II)}}/2}\right)=o(1). \tag{9}
$$

The equality in (9) follows from the conditions $\sqrt{n}\gamma_n^{(II)} \to \infty$, $\log(p_n - s_n) = o(n\gamma_n^{(II)})$ and $\log(p_n - s_n) = o(n\gamma_n^{\text{(II)}2})$.

Proof of Theorem 4

Minimizing (4.3) is equivalent to minimizing the following criterion functions,

$$
\ell_{n,j}^{\text{PCR}}(\alpha) = \frac{1}{n} \sum_{i=1}^{n} Q(Y_{ni}, F^{-1}(X_{ij}\alpha)) + \kappa_n |\alpha|, \quad j = 1, \ldots, p_n.
$$

Part 1: $j \in \{1, ..., s_n\}$.

Similar to the proof of Theorem 3, we prove that for a small enough $\delta > 0$, there exist local minimizers $\beta_{n,j}^{\text{PCR}}$ of $\ell_{n,j}^{\text{PCR}}(\alpha)$ such that

$$
P\Big(\min_{1\leq j\leq s_n}|\widehat{\beta}_{n,j}^{\text{PCR}}| > \gamma_n^{(1)}\delta\Big) \to 1. \tag{10}
$$

It suffices to prove that for a small enough $\delta > 0$ and some large enough $C_n > 0$, there exist some β_j with $|\beta_j| = 2\delta$ such that

$$
P\Big(\min_{1\leq j\leq s_n}\Big\{\inf_{|\alpha|\leq \delta} \ell_{n,j}^{\text{PCR}}(\gamma_n^{(1)}\alpha) - \ell_{n,j}^{\text{PCR}}(\gamma_n^{(1)}\beta_j)\Big\} > 0\Big) \to 1,\tag{11}
$$

and

$$
P\Big(\min_{1\leq j\leq s_n}\Big\{\inf_{|\alpha|\geq C_n} \ell_{n,j}^{\text{PCR}}(\gamma_n^{(1)}\alpha) - \ell_{n,j}^{\text{PCR}}(\gamma_n^{(1)}\beta_j)\Big\} > 0\Big) \to 1. \tag{12}
$$

(11) and (12) imply that with probability tending to one, there must exist local minimizers $\widehat{\beta}_{n,j}^{\text{PCR}}$ of $\ell_{n,j}^{\text{PCR}}(\alpha)$ such that $\gamma_n^{(1)}\delta < |\widehat{\beta}_{n,j}^{\text{PCR}}| < \gamma_n^{(1)}C_n$ for $1 \le j \le s_n$, and this implies (10). First, we prove (12). We notice that for every $n \geq 1$, when $|\alpha| \to \infty$,

$$
\min_{1 \le j \le s_n} \left\{ \ell_{n,j}^{\text{PCR}}(\gamma_n^{(\text{I})}\alpha) - \ell_{n,j}^{\text{PCR}}(\gamma_n^{(\text{I})}\beta_j) \right\} \ge \kappa_n \gamma_n^{(\text{I})}|\alpha| - \max_{1 \le j \le s_n} \ell_{n,j}^{\text{PCR}}(\gamma_n^{(\text{I})}\beta_j) \to \infty.
$$

Thus (12) holds.

Second, we prove (11). By Taylor's expansion, for $1 \leq j \leq s_n$,

$$
\ell_{n,j}^{\text{PCR}}(\gamma_n^{(I)}\alpha) = \frac{1}{n} \sum_{i=1}^n Q(Y_{ni}, F^{-1}(0)) + \frac{\gamma_n^{(I)}}{n} \alpha \sum_{i=1}^n q_1(Y_{ni}; 0) X_{ij} + \frac{1}{2} \frac{\gamma_n^{(I)2}}{n} \alpha^2 \sum_{i=1}^n q_2(Y_{ni}; \gamma_n^{(I)} \alpha_j^* X_{ij}) X_{ij}^2 + \gamma_n^{(I)} \kappa_n |\alpha|,
$$

with α_j^* between 0 and α . Thus, we have that

$$
\min_{1 \leq j \leq s_n} \left\{ \inf_{|\alpha| \leq \delta} \ell_{n,j}^{\text{PCR}}(\gamma_n^{(1)} \alpha) - \ell_{n,j}^{\text{PCR}}(\gamma_n^{(1)} \beta_j) \right\}
$$
\n
$$
\geq \min_{1 \leq j \leq s_n} \inf_{|\alpha| \leq \delta} \left\{ \frac{\gamma_n^{(1)}}{n} (\alpha - \beta_j) \sum_{i=1}^n q_1(Y_{ni}; 0) X_{ij} \right\}
$$
\n
$$
+ \min_{1 \leq j \leq s_n} \left[\frac{1}{2} \frac{\gamma_n^{(1)2}}{n} \inf_{|\alpha| \leq \delta} \left\{ \alpha^2 \sum_{i=1}^n q_2(Y_{ni}; \gamma_n^{(1)} \alpha_j^* X_{ij}) X_{ij}^2 - \beta_j^2 \sum_{i=1}^n q_2(Y_{ni}; \gamma_n^{(1)} \beta_j^* X_{ij}) X_{ij}^2 \right\} \right]
$$
\n
$$
+ \min_{1 \leq j \leq s_n} \inf_{|\alpha| \leq \delta} \left\{ \gamma_n^{(1)} \kappa_n(|\alpha| - |\beta_j|) \right\}
$$
\n
$$
\equiv I_1 + I_2 + I_3,
$$

with α_j^* between 0 and α , and β_j^* between 0 and $\beta_{n,j}$. Let $\mu_0 = F^{-1}(0)$ and $C_0 = q''(\mu_0)/F'(\mu_0) \neq 0$, for I_1 ,

$$
I_{1} \geq \gamma_{n}^{(1)} \min_{1 \leq j \leq s_{n}} \inf_{|\alpha| \leq \delta} \{ C_{0}(\alpha - \beta_{j}) E(Y_{n} X_{j}) \}
$$

+ $\gamma_{n}^{(1)} \min_{1 \leq j \leq s_{n}} \inf_{|\alpha| \leq \delta} \left[C_{0}(\alpha - \beta_{j}) \frac{1}{n} \sum_{i=1}^{n} \{ Y_{ni} X_{ij} - E(Y_{n} X_{j}) \} \right]$
- $\gamma_{n}^{(1)} \max_{1 \leq j \leq s_{n}} \sup_{|\alpha| \leq \delta} \left\{ C_{0} \mu_{0}(\alpha - \beta_{j}) \frac{1}{n} \sum_{i=1}^{n} X_{ij} \right\}$
\n $\equiv I_{1,1} + I_{1,2} + I_{1,3}.$

We can see that

$$
|I_{1,3}| \leq \gamma_n^{(1)} |C_0 \mu_0| \max_{1 \leq j \leq s_n} \left\{ \sup_{|\alpha| \leq \delta} \left(3\delta \left| \frac{1}{n} \sum_{i=1}^n X_{ij} \right| \right) \right\}
$$

= $O_P(\gamma_n^{(1)} {\log(s_n)}/{n}^{1/2})\delta$,

since $\max_{1 \leq j \leq s_n} |n^{-1} \sum_{i=1}^n X_{ij}| = O_P(\{\log(s_n)/n\}^{1/2})$ from Bernstein's inequality (Lemma 2.2.9 in van der Vaart and Wellner, 1996). Again $|I_{1,2}| = O_P(\gamma_n^{(1)} \{ \log(s_n)/n \}^{1/2})$ δ by a similar argument as in the proof of Theorem 3. Now we choose $\beta_j = -2\delta \text{sign}\{C_0E(Y_nX_j)\}$ satisfying $|\beta_j| = 2\delta$. Then

$$
I_{1,1} = \gamma_n^{(1)} \min_{1 \le j \le s_n} \inf_{|\alpha| \le \delta} \left([\alpha + 2\delta \text{sign}\{C_0 E(Y_n X_j)\}] C_0 E(Y_n X_j) \right)
$$

$$
\ge \gamma_n^{(1)} \min_{1 \le j \le s_n} \{ \delta |C_0 E(Y_n X_j)| \} \ge |C_0| c \gamma_n^{(1)} \mathcal{A}_n \delta.
$$

For I_2 and I_3 ,

$$
|I_2| \leq \frac{1}{2} \frac{\gamma_n^{(1)2}}{n} \max_{1 \leq j \leq s_n} \sup_{|\alpha| \leq \delta} \left\{ \alpha^2 \middle| \sum_{i=1}^n q_2(Y_{ni}; X_{ij} \gamma_n^{(1)} \alpha_j^*) X_{ij}^2 \right|
$$

+ $\beta_j^2 \left| \sum_{i=1}^n q_2(Y_{ni}; X_{ij} \gamma_n^{(1)} \beta_j^*) X_{ij}^2 \right|$
= $O_P(\gamma_n^{(1)2}) \delta^2$,

and $|I_3| = O(\gamma_n^{(1)} \kappa_n) \delta$.

Under the conditions $\mathcal{A}_n = \gamma_n^{(1)}$, $\kappa_n = o(\gamma_n^{(1)})$ and $\log(s_n) = o(n\gamma_n^{(1)2})$, we can choose a small enough $\delta > 0$ such that with probability tending to one, $I_{1,1}$ dominates $I_{1,2}$, $I_{1,3}$, I_2 and I_3 . Thus (11) is proved.

Part 2: $j \in \{s_n + 1, \ldots, p_n\}.$

We prove that for any $\epsilon > 0$, there exist local minimizers $\hat{\beta}_{n,j}^{\text{PCR}}$ of $\ell_{n,j}^{\text{PCR}}(\alpha)$ such that

$$
P\Big(\max_{s_n+1\leq j\leq p_n}|\widehat{\beta}_{n,j}^{\text{PCR}}| \leq \gamma_n^{\text{(II)}}\epsilon\Big) \to 1. \tag{13}
$$

It suffices to prove that for any $\epsilon > 0$,

$$
P\Big(\min_{s_n+1\leq j\leq p_n}\Big\{\inf_{|\alpha|=\epsilon} \ell_{n,j}^{\text{PCR}}(\gamma_n^{(\text{II})}\alpha) - \ell_{n,j}^{\text{PCR}}(0)\Big\} > 0\Big) \to 1. \tag{14}
$$

By Taylor's expansion,

$$
\min_{s_n+1 \leq j \leq p_n} \left\{ \inf_{|\alpha|=\epsilon} \ell_{n,j}^{\text{PCR}}(\gamma_n^{(\text{II})}\alpha) - \ell_{n,j}^{\text{PCR}}(0) \right\}
$$
\n
$$
\geq \min_{s_n+1 \leq j \leq p_n} \inf_{|\alpha|=\epsilon} \left\{ \frac{\gamma_n^{(\text{II})}}{n} \alpha \sum_{i=1}^n \mathbf{q}_1(Y_{ni}; 0) X_{ij} \right\}
$$
\n
$$
+ \min_{s_n+1 \leq j \leq p_n} \inf_{|\alpha|=\epsilon} \left\{ \frac{\gamma_n^{(\text{II})2}}{2n} \alpha^2 \sum_{i=1}^n \mathbf{q}_2(Y_{ni}; \gamma_n^{(\text{II})}\alpha_j^* X_{ij}) X_{ij}^2 \right\} + \inf_{|\alpha|=\epsilon} (\gamma_n^{(\text{II})}\kappa_n|\alpha|)
$$
\n
$$
\equiv I_1 + I_2 + I_3,
$$

with $\alpha_j^* \in (0, \alpha)$. For I_1 , it is seen that

$$
|I_{1}| \leq \max_{s_{n}+1 \leq j \leq p_{n}} \sup_{|\alpha|=\epsilon} \left\{ \frac{\gamma_{n}^{(\text{II})}}{n} \bigg| \alpha \sum_{i=1}^{n} C_{0} (Y_{ni} - \mu_{0}) X_{ij} \bigg| \right\}
$$

$$
\leq \frac{\gamma_{n}^{(\text{II})}}{n} |C_{0}| \epsilon \max_{s_{n}+1 \leq j \leq p_{n}} \left| \sum_{i=1}^{n} Y_{ni} X_{ij} \right| + \frac{\gamma_{n}^{(\text{II})}}{n} |C_{0} \mu_{0}| \epsilon \max_{s_{n}+1 \leq j \leq p_{n}} \left| \sum_{i=1}^{n} X_{ij} \right|.
$$

Since $|\sum_{i=1}^{n} Y_{ni} X_{ij}| \leq |\sum_{i=1}^{n} \{Y_{ni} X_{ij} - E(Y_{n} X_{j})\}| + |\sum_{i=1}^{n} E(Y_{n} X_{j})|$ and similar to Part 1,

$$
\max_{s_n+1 \le j \le p_n} \left| \frac{1}{n} \sum_{i=1}^n X_{ij} \right| = O_P(\{\log(p_n - s_n)/n\}^{1/2}),
$$

$$
\max_{s_n+1 \le j \le p_n} \left| \frac{1}{n} \sum_{i=1}^n \{Y_{ni} X_{ij} - E(Y_n X_j)\} \right| = O_P(\{\log(p_n - s_n)/n\}^{1/2}),
$$

we have that $|I_1| \leq O_P(\gamma_n^{(\text{II})} \{ \log(p_n - s_n)/n \}^{1/2}) \epsilon + o_P(\gamma_n^{(\text{II})} \mathcal{B}_n) \epsilon.$

The proof may be separated for case (1) and case (2) in Theorem 4 from here. Case (1): For I_2 , we have that

$$
|I_2| \leq \max_{s_n+1 \leq j \leq p_n} \sup_{|\alpha|=\epsilon} \left\{ \frac{\gamma_n^{(\text{II})2}}{2n} \alpha^2 \bigg| \sum_{i=1}^n q_2(Y_{ni}; \gamma_n^{(\text{II})} \alpha_j^* X_{ij}) X_{ij}^2 \bigg| \right\} = O_P(\gamma_n^{(\text{II})2}) \epsilon^2.
$$

Thus $I_3 = \gamma_n^{(II)} \kappa_n \epsilon$ dominates I_1 and I_2 with probability tending to one, since $\log(p_n - \epsilon)$ $(s_n) = o(n\kappa_n^2),$ $\mathcal{B}_n = O(\kappa_n)$ and $\gamma_n^{(II)} = o(\kappa_n)$. So (14) is proved.

Case (2) : I_2 is always positive with Condition 1(b). Moreover,

$$
I_2 \geq \frac{1}{2} \gamma_n^{(\text{II})2} \min_{s_n+1 \leq j \leq p_n} \inf_{|\alpha|=\epsilon} \left[\alpha^2 E\{ \mathbf{q}_2(Y_n; \gamma_n^{(\text{II})} \alpha_j^* X_j) X_j^2 \} \right] - \frac{1}{2} \gamma_n^{(\text{II})2} \times \max_{s_n+1 \leq j \leq p_n} \sup_{|\alpha|=\epsilon} \left(\alpha^2 \left| \frac{1}{n} \sum_{i=1}^n \left[\mathbf{q}_2(Y_{ni}; \gamma_n^{(\text{II})} \alpha_j^* X_{ij}) X_{ij}^2 - E\{ \mathbf{q}_2(Y_n; \gamma_n^{(\text{II})} \alpha_j^* X_j) X_j^2 \} \right] \right| \right)
$$

$$
\geq \frac{1}{2} \gamma_n^{(\text{II})2} \epsilon^2 \eta - \frac{1}{2} \gamma_n^{(\text{II})2} \epsilon^2 O_P(\{\log(p_n - s_n)/n\}^{1/2}).
$$

Thus, $I_3 = \gamma_n^{(II)} \kappa_n \epsilon$ and the term $\frac{1}{2} \gamma_n^{(II)2} \epsilon^2 \eta$ in I_2 dominate all other terms with probability tending to one, since $\log(p_n - s_n) = o\{\max(n\kappa_n^2, n\gamma_n^{\text{(II)2}})\}, \mathcal{B}_n = O\{\max(\kappa_n, \gamma_n^{\text{(II)}})\}\$ and $log(p_n - s_n) = o(n)$. So (14) is also proved.