Supplemental Appendix of
“A Nondegenerate Vuong Test and A Post
Selection Confidence Interval for
Semi/Nonparametric Models”

In this supplemental appendix, we present supporting materials for the main paper:

- Section C verifies the high-level assumptions in the main paper in a nonparametric mean-regression example.
- Section D includes the illustration of the high-level theory and verification of the high-level assumptions in the main paper in a nonparametric quantile-regression example.
- Section E proves the lemmas in Section 6 of the main paper.
- Section F includes some auxiliary results and their proofs.
- Section G conducts simulation studies to compare the finite sample properties of our nonparametric nondegenerate test with the test proposed in Shi (2015b).

C Verification of the High-level Assumptions in the Mean-Regression Example

The assumptions needed for verifying the high-level assumptions in Section 4 of the main paper in the mean-regression example in Section 5 of the main paper are given below. Assumption C.1 imposes conditions on the data structure and the finite-dimensional approximation. These are commonly used conditions in the literature (see, e.g., Andrews 1991, Newey 1997, Chen 2007, and Belloni, Chernozhukov, Chetverikov, and Kato 2015).

Assumption C.1. There exist positive constants $C_1$, $r_1$, $r_2$ and nondecreasing sequences $\{\xi_{k_1}\}_{k_1 \geq 1}$ and $\{\xi_{k_2}\}_{k_2 \geq 1}$, such that, for any $F_0 \in \mathcal{F}$ and for $j = 1, 2$:

(i) $\{Z_i\}_{i \geq 1}$ are i.i.d. draws from $F_0$;
(ii) $\sup_{x_j \in \mathcal{X}_j} \left| \alpha_j^* (x_j) - P_{k_j} (x_j)' \beta_{k_j,F_0}^* \right| \leq C_1 k_j^{-r_j}$ where $r_j > 0$ and $\beta_{k_j,F_0}^*$ is defined in (5.4); (iii) $\sup_{x_j \in \mathcal{X}_j} \| P_{k_j} (x_j) \|^2 \leq \xi_{k_j}$; and
(iv) the eigenvalues of $E_{F_0} \left[ P_{k_j} (X_j) P_{k_j} (X_j)' \right]$ lie in the interval $[C_1^{-1}, C_1]$. 


Assumption C.2 below imposes condition on the error terms. For \( j = 1, 2 \), define \( u_j = Y - \alpha_j^*(X_j) \). It is useful to note here that \( \omega_{F_0,*}^2 = Var F_0 (u_1^2 - u_2^2) / 4 \) and \( \sigma_{F_0,n}^2 = \omega_{F_0,*}^2 + (2n^2)^{-1}(n - 1)tr((H_{F_0,k}^{-1}D_{F_0,k})^2) \) for this example.

**Assumption C.2** There exist constants \( b \geq 4 \) and \( C_2 > 0 \) such that for any \( F_0 \in \mathcal{F} \):

(i) \( E_{F_0} \left[ \omega_{F_0,*}^1(u_1^2 - u_2^2 - E_{F_0}(u_1^2 - u_2^2)) \right]^4 < C_2 \) whenever \( \omega_{F_0,*}^2 > 0 \); and

(ii) \( E_{F_0} \left[ |u_j|^b \right] X_j = x_j \) \( \leq C_2 \) for all \( x_j \in X_j \) for \( j = 1, 2 \).

**Assumption C.3** For any \( F_0 \in \mathcal{F}_0 \), we have \( F_0 \in \mathcal{F} \) and \( E_{F_0}[u_1^2 - u_2^2] = 0 \).

Let \( \xi_k = \xi_{k1} + \xi_{k2} \). Assumption C.4 imposes conditions on the numbers of series terms in the finite-dimensional approximation, and on the divergence of \( n\sigma_{F_0,n}^2 \).

**Assumption C.4** For any sequence of DGP’s \( \{F_n\}_{n \geq 1} \) such that \( F_n \in F \), we have, for the constant \( b \) in Assumption C.2 and the constants \( r_1, r_2 \) in Assumption C.1

(i) \( (n^{1/2}\sigma_{F_n,n})^{-1} \left( 1 + |k| (\xi_k \log(n))^{1/2}n^{-1/2} \right) = o(1) \);

(ii) \( \max_{j=1,2}(n\sigma_{F_n,n}^2)^{-1} \xi_k \log(n))^{1/2} \left( k_j^{1/2}n^{-1/2} + n^{-1/2} \right) = o(1) \);

(iii) \( \xi_k \log(n)n^{-1/2} = o(1) \);

(iv) \( \max_{j=1,2} \left( nk_j^{-2r_j} + \xi_{k_j}k_j \log(n)n^{-1} \right) = O(1) \).

The following theorem summarizes the main results in this section.

**Theorem C.1** Assumptions C.1-C.4 together imply Assumptions 4.1-4.4 in the main paper.

This theorem is an immediate consequence of the following Lemmas C.1-C.3 below. Each lemma is proved immediately after it is stated.

**Lemma C.1** Under Assumptions C.1 and C.2 Assumption 4.1 in the main paper holds with

\[ C = \max\{8C_2, 2C_1C_2\} \] .

**Proof of Lemma C.1** Assumption 4.1(a) is implied by Assumption C.1(i). Assumption 4.1(b) holds because \( E_{F_0}[\ell(Z; \alpha_c^*)] \) is a quadratic function of \( \beta_k \) for any \( F_0 \). Assumption 4.1(c) holds by the first order condition of \( \alpha_c^* \). For Assumption 4.1(d), first note that

\[ E_{F_0} \left[ \ell(Z; \alpha_c^*)^2 \right] = E_{F_0} \left[ (u_2^2 - u_1^2)^2 \right] / 4 \leq E_{F_0} \left[ u_1^4 \right] / 2 + E_{F_0} \left[ u_2^4 \right] / 2 \leq C_2 \] \hspace{1cm} (C.1)

where the second inequality is by Assumption C.2(ii) with \( C_2 \) specified in that assumption. Since \( u_{k_j} = u_j + \alpha_j^*(x_j) - \alpha_{k_j}^*(x_j) \), by Assumptions C.2(ii) and C.1(ii),

\[ E_{F_0} \left[ u_{k_j}^4 \right] X_j = x_j \leq 8E_{F_0} \left[ u_j^4 \right] X_j = x_j + 8 \left| \alpha_j^*(x_j) - \alpha_{k_j}^*(x_j) \right|^4 \]
\[ \leq 8C_2 + 8C_1^4 k_j^{-4r_j} \leq C \quad \text{(C.2)} \]

for \( j = 1, 2 \) and uniformly over \( x_j \in X_j \). Moreover,

\[
E_{F_0} \left[ \| \ell_{\alpha, k}(Z; \alpha_k^*) \|^4 \right] \leq 2 \sum_{j=1,2} E_{F_0} \left[ u_{k_j}^j | P_{k_j} (X_j)' P_{k_j} (X_j) |^2 \right] \leq C \sum_{j=1,2} \xi_{k_j} \text{tr} (-H_{F_0,k_j}) \leq C \sum_{j=1,2} \xi_{k_j} k_j = C \xi_k |k|, \quad \text{(C.3)}
\]

where the second inequality is by (C.2) and Assumption (C.1)(iii), and the third inequality is by Assumption (C.1)(iv). Hence Assumption 4.1(d) is also satisfied. Assumption 4.1(e) holds by Assumption (C.2)(i). For Assumption 4.1(f), the first part of it holds by Assumption (C.1)(iv) and equation (5.8). To show the second part, first note that, under Assumptions (C.1)(iv) and (C.2),

\[
\rho_{\text{max}} \left( E_{F_0} \left[ u_{k_j}^2 P_{k_j} (X_j) P'_{k_j} (X_j) \right] \right) \leq C \rho_{\text{max}} (-H_{F_0,k_j}) \leq CC_1 \text{ for } j = 1, 2, \quad \text{(C.4)}
\]

which together with the form of \( D_{F_0,k} \) in (5.7) and the Aronszajn’s Inequality (see, e.g., Theorem III.2.9 in Bhatia (1997)) implies that \( \rho_{\text{max}}(D_{F_0,k}) \leq 2CC_1 \leq C \). This verifies Assumption 4.1(f).

**Lemma C.2** Under Assumptions (C.1), (C.2), (C.4)(i) and (C.4)(iv), Assumption 4.3 of the main paper holds.

**Lemma C.3** Under Assumptions (C.1), (C.2) and (C.4), Assumption 4.4 of the main paper holds.

The proof of Lemmas C.2 and C.3 makes use of the following three lemmas. The first one of which follows from Lemma 6.2 of Belloni, Chernozhukov, Chetverikov, and Kato (2015), and the other two are proved at the end of this section.

In the rest of this section, we use the new notation \( U_{j,n} = [u_{j,1}, \ldots, u_{j,n}]' \) for \( j = 1, 2 \). Let \( \beta_{k_j,F_n}^* \) be abbreviated as \( \beta_{k_j}^* \). Let \( \alpha_{k_j}^* (\cdot) = P_{k_j} (\cdot)' \beta_{k_j}^* \).

**Lemma C.4** Suppose that Assumptions (C.1)(i), (iii) and (iv) holds. Then under any sequence \( \{F_n\}_{n \geq 1} \) such that \( F_n \in \mathcal{F} \) for all \( n \)

\[
\left\| n^{-1} P'_{k_j,n} P_{k_j,n} - E_{F_n} [P_{k_j} (X_j) P_{k_j} (X_j)'] \right\|^2 \leq O_p(\xi_{k_j} \log(k_j)n^{-1}). \quad \text{(C.5)}
\]

Using Lemma C.4 Assumptions (C.1)(iv) and (C.4)(iv), we have

\[
C_1^{-1}/2 \leq \rho_{\text{min}}(n^{-1} P'_{k_j,n} P_{k_j,n}) \leq \rho_{\text{max}}(n^{-1} P'_{k_j,n} P_{k_j,n}) \leq 2C_1 \quad \text{(C.6)}
\]
for \(j = 1, 2\) with probability approaching 1.

**Lemma C.5** Suppose that Assumptions \[\text{C.1, C.2, C.4(ii), C.4(iv)}\] hold. Then under any sequence \(\{F_n\}_{n \geq 1}\) such that \(F_n \in \mathcal{F}\) for all \(n\) and for \(j = 1, 2\), we have

(a) \(\|\hat{\beta}_{k_j} - \beta_{k_j}^*\|^2 = O_p(k_j n^{-1})\), and

(b) \(\gamma_n(\hat{\beta}_{k_j} - \beta_{k_j}^*) = O_p(n^{-1/2}\|\gamma_n\|)\) for any deterministic \(k_j \times 1\) vector sequence \(\{\gamma_n\}_{n \geq 1}\).

(c) \((\hat{\beta}_k - \beta_k^*)' D_{F_n,k} (\hat{\beta}_k - \beta_k^*) \approx \tilde{\ell}_{\alpha,n} (\alpha_k^*)' (H_{F_n,k}^{-1} D_{F_n,k} H_{F_n,k}^{-1}) \tilde{\ell}_{\alpha,n} (\alpha_k^*) + o_p(\sigma_{F_n,n}^2).\)

**Lemma C.6** Suppose that Assumptions \[\text{C.1, C.2 and C.4(iv)}\] hold. Then under any sequence \(\{F_n\}_{n \geq 1}\) such that \(F_n \in \mathcal{F}\) for all \(n\), we have

(a) for \(\delta_{n,k} = (k_1 \sqrt{\xi_{k_1}} + k_2 \sqrt{\xi_{k_2}}) (\log(n))^{1/2} n^{-3/2} + (k_1^{-r_1} + k_2^{-r_2}) n^{-1/2} + k_1^{-2r_1} + k_2^{-2r_2},\)

\[
\bar{\ell}_n(\hat{\alpha}_k) = (2n)^{-1} \sum_{i=1}^n (u_{1,i}^2 - u_{2,i}^2) - \frac{\bar{\ell}_{\alpha,n} (\alpha_k^*)' H_{F_n,k}^{-1} \bar{\ell}_{\alpha,n} (\alpha_k^*)}{2} + O_p(\delta_{n,k}).
\]

Now we present the proofs of the lemmas above.

**Proof of Lemma C.2.** First we verify Assumption 4.3(a), which is a quadratic expansion of \(\bar{\ell}_n(\hat{\alpha}_k)\). In Lemma C.6(b), we have derived the second order expansion where the remainder term is of the order

\[
\delta_{n,k} = (k_1 \sqrt{\xi_{k_1}} + k_2 \sqrt{\xi_{k_2}}) (\log(n))^{1/2} n^{-3/2} + (k_1^{-r_1} + k_2^{-r_2}) n^{-1/2} + k_1^{-2r_1} + k_2^{-2r_2}. \tag{C.7}
\]

Observe that

\[
\frac{n^{1/2} \delta_{n,k}}{\sigma_{F_n,n}} \leq 2 \max_{j=1,2} \left[ \frac{k_j (\xi_{k_j} \log(n))^{1/2}}{n \sigma_{F_n,n}} + \frac{1}{k_j r_j} \right] \leq 1 + \frac{k_j^{-r_j}}{n^{1/2}} \quad \Rightarrow \quad o(1), \tag{C.8}
\]

where the last equality holds by Assumptions C.4(i) and (iv). This combined with Lemma C.6(b) implies Assumption 4.3(a).

Assumption 4.3(b) is directly implied by Assumptions C.4(i).

**Proof of Lemma C.3.** We start with Assumption 4.4(a). First, Lemma C.4 implies that \(\|\hat{H}_n - H_{F_n,k}\|^2 = O_p(\xi_k \log(n) n^{-1})\), where \(\xi_k = \xi_{k_1} + \xi_{k_2}\). Then for \(\delta_n\) in Assumption 4.4(a), the first part of Assumption 4.4(a) follows the derivation: for \(j = 1, 2\)

\[
\frac{\xi_{k_j} \log(n) n^{-1}}{\delta_n^2} \leq \max \left\{ \frac{\xi_{k_j} \log(n)}{n^2 \sigma_{F_n,n}^2}, \frac{\xi_{k_j} \log(n)}{n} \right\}
\]
\[
\left\{ \frac{1}{n\sigma_{F,n}^2} \frac{\xi_{k_j}|k|^2 \log(n)}{n}, \frac{\xi_{k_j} \log(n)}{n} \right\} = o(1), \tag{C.9}
\]

where the last equality holds by Assumptions C.4(i) and (iii).

For the second and the third parts of Assumption 4.4(a), note that by arguments similar to those in the proof of Theorem 4.6 in Belloni, Chernozhukov, Chetverikov, and Kato (2015), we have

\[
\| \hat{D}_n - \hat{D}_n(\alpha^*_k) \|^2 = O_p \left( \frac{\xi_k \log(n)}{n^{1-2/b}} \right) \quad \text{and} \quad \| \hat{D}_n(\alpha^*_k) - D_{F,n,k} \|^2 = O_p \left( \frac{\xi_k \log(n)}{n^{1-2/b}} \right). \tag{C.10}
\]

The second and the third parts of Assumption 4.4(a) follows from the derivation:

\[
\frac{\xi_k \log(n)}{n^{1-2/b} \delta^2_n} = \max \left\{ \frac{1}{n\sigma_{F,n}^2} \frac{\xi_k |k|^2 \log(n)}{n^{1-2/b}}, \frac{\xi_k \log(n)}{n^{1-2/b}} \right\} = o(1),
\]

where the last equality holds by Assumption C.4(i) and (iii). Thus Assumption 4.4(a) is verified.

Next, we verify Assumption 4.4(b). By Lemma C.5(a) and Assumption C.1(iii), we have

\[
\sup_{x_j \in X_j} \left| \hat{\alpha}_{k_j}(x_j) - \alpha^*_{k_j}(x_j) \right|^2 = O_p(\xi_{k_j}k_j n^{-1}). \tag{C.11}
\]

Using Assumption C.1(ii) and Lemma C.6(a) and (C.11), we get

\[
n^{-1} \sum_{i=1}^{n} \left| \hat{\alpha}_{k_j}(X_{j,i}) - \alpha^*_{k_j}(X_{j,i}) \right|^4 \leq \frac{\sup_{x_j \in X_j} \left| \hat{\alpha}_{k_j}(x_j) - \alpha^*_{k_j}(x_j) \right|^2 \sum_{i=1}^{n} \left| \hat{\alpha}_{k_j}(X_{j,i}) - \alpha^*_{k_j}(X_{j,i}) \right|^2}{n} = O_p(\xi_{k_j}k_j n^{-1})O_p(k_j n^{-1}) = \sigma_{F,n}^2 O_p \left( \frac{1}{n\sigma_{F,n}^2} \frac{\xi_{k_j}k_j^2}{n} \right) = o_p(\sigma_{F,n}^2), \tag{C.12}
\]

where the last equality holds by Assumption C.4(i). Using Assumptions 4.1(f), C.4(iii) and (C.10), we have

\[
0 \leq \rho_{\max}(\hat{D}_n(\alpha^*_k)) \leq 2C \quad \text{and} \quad 0 \leq \rho_{\max}(\hat{D}_n(\alpha^*_k)) \leq 2C. \tag{C.13}
\]
with probability approaching 1. By Lemma C.5(a) and (C.13), we have

$$n^{-1} \sum_{i=1}^{n} u_{k_j,i}^2 \left| \hat{\alpha}_{kj}(X_{j,i}) - \alpha^*_{kj}(X_{j,i}) \right|^2 = O_p(k_j n^{-1}).$$  \hfill (C.14)

Using the Cauchy-Schwarz inequality, we get

$$\left| n^{-1} \sum_{i=1}^{n} u_{k_j,i} \left[ \hat{\alpha}_{kj}(X_{j,i}) - \alpha^*_{kj}(X_{j,i}) \right] \right|^2 \leq n^{-1} \sum_{i=1}^{n} u_{k_j,i}^2 \left| \hat{\alpha}_{kj}(X_{j,i}) - \alpha^*_{kj}(X_{j,i}) \right|^2 \times n^{-1} \sum_{i=1}^{n} \left| \hat{\alpha}_{kj}(X_{j,i}) - \alpha^*_{kj}(X_{j,i}) \right|^4$$

$$= O_p(k_j n^{-1}) O_p(\xi_k k_j^2 n^{-2}) = O_p(\xi_k k_j^3 n^{-3}) = o_p(\sigma^4_{F_{n,n}})$$ \hfill (C.15)

where the first equality is by the third line of (C.12) and by (C.14), the second equality is by the last line of (C.12), and the last equality holds by Assumptions C.4(ii). Similarly, we can show that

$$\left| n^{-1} \sum_{i=1}^{n} u_{k_j,i} \left[ \hat{\alpha}_{kj}(X_{j,i}) - \alpha^*_{kj}(X_{j,i}) \right] \left| \hat{\alpha}_{k_j-j}(X_{j-i,j}) - \alpha^*_{k_j-j}(X_{j-i,j}) \right| \right|^2 = o_p(\sigma^4_{F_{n,n}})$$ \hfill (C.16)

for \((j, -j) = (1, 2)\) or \((2, 1)\). Also, by definition,

$$\ell(Z, \hat{\alpha}_k) - \ell(Z, \alpha_k^*) = \left[ \hat{\alpha}_{k_1}(X_2) - \alpha^*_{k_1}(X_2) \right]^2 / 2 - \left[ \hat{\alpha}_{k_1}(X_1) - \alpha^*_{k_1}(X_1) \right]^2 / 2$$

$$+ u_{k_1} \left[ \hat{\alpha}_{k_1}(X_1) - \alpha^*_{k_1}(X_1) \right] - u_{k_2} \left[ \hat{\alpha}_{k_2}(X_2) - \alpha^*_{k_2}(X_2) \right].$$ \hfill (C.17)

which together with (C.12), (C.15) and (C.16) implies that

$$n^{-1} \sum_{i=1}^{n} \left| \ell(Z, \hat{\alpha}_k) - \ell(Z, \alpha_k^*) \right|^2$$

$$= n^{-1} \sum_{i=1}^{n} \left[ u_{k_1} \left[ \hat{\alpha}_{k_1}(X_1) - \alpha^*_{k_1}(X_1) \right] - u_{k_2} \left[ \hat{\alpha}_{k_2}(X_2) - \alpha^*_{k_2}(X_2) \right] \right]^2 + o_p(\sigma^2_{F_{n,n}}).$$ \hfill (C.18)

By (C.14),

$$n^{-1} \sum_{i=1}^{n} \left[ u_{k_1} \left[ \hat{\alpha}_{k_1}(X_1) - \alpha^*_{k_1}(X_1) \right] - u_{k_2} \left[ \hat{\alpha}_{k_2}(X_2) - \alpha^*_{k_2}(X_2) \right] \right]^2 = O_p((k_1 + k_2) n^{-1}).$$ \hfill (C.19)
Next note that
\[
\ell(Z, \alpha^*_k) - \ell(Z, \alpha^*_n) = \frac{u_1^2 - u_{k_1}^2}{2} - \frac{u_2^2 - u_{k_2}^2}{2} = \frac{(u_1 - u_{k_1})^2}{2} + (u_1 - u_{k_1})u_{k_1} - \frac{(u_2 - u_{k_2})^2}{2} - (u_2 - u_{k_2})u_{k_2}.
\]

By Assumption C.1(ii),
\[
n^{-1} \sum_{i=1}^{n} (u_{j,i} - u_{k,j,i})^4 = n^{-1} \sum_{i=1}^{n} \left| \alpha^*_{k,j} (X_{j,i}) - \alpha^*_{j,n} (X_{j,i}) \right|^4 = O_p(k_j^{-4r_j}).
\]

By Assumption C.1(ii), (C.2) and the Markov inequality,
\[
n^{-1} \sum_{i=1}^{n} (u_{j,i} - u_{k,j,i})^2 u_{k,j,i}^2 = n^{-1} \sum_{i=1}^{n} u_{k,j,i}^2 \left| \alpha^*_{k,j} (X_{j,i}) - \alpha^*_{j,n} (X_{j,i}) \right|^2 = O_p(k_j^{-2r_j}).
\]

Combining the results in (C.20), (C.21) and (C.22), we get
\[
n^{-1} \sum_{i=1}^{n} \left| \ell(Z, \alpha^*_k) - \ell(Z, \alpha^*_n) \right|^2 = O_p(k_1^{-2r_1} + k_2^{-2r_2}).
\]

By Assumptions C.1(i), C.1(ii) and C.1(iv), (C.2) and (C.20)
\[
E_{F_n} \left[ \left\| n^{-1} \sum_{i=1}^{n} \left[ (\ell(Z, \alpha^*_k) - \ell(Z, \alpha^*_n))P_{k,j} (X_j) - E_{F_n} ((\ell(Z, \alpha^*_k) - \ell(Z, \alpha^*_n))P_{k,j} (X_j)) \right] \right\|^2 \right] \\
\leq n^{-2} \sum_{i=1}^{n} E_{F_n} \left[ (\ell(Z, \alpha^*_k) - \ell(Z, \alpha^*_n))^2 P_{k,j} (X_j)P_{k,j} (X_j) \right] \\
\leq Cn^{-1}(k_1^{-2r_1} + k_2^{-2r_2})E_{F_n} \left[ P_{k,j} (X_j)P_{k,j} (X_j) \right] = O((k_1^{-2r_1} + k_2^{-2r_2})kjn^{-1}).
\]

By Assumptions C.1(ii) and C.1(iv), (C.2) and (C.20)
\[
\left\| E_{F_n} \left[ (\ell(Z, \alpha^*_k) - \ell(Z, \alpha^*_n))P_{k,j} (X_j) \right] \right\|^2 \\
= \rho_{\max}(E_{F_n} \left[ P_{k,j} (X_j)P_{k,j} (X_j) \right])E_{F_n} \left[ (\ell(Z, \alpha^*_k) - \ell(Z, \alpha^*_n))^2 \right] \leq C(k_1^{-2r_1} + k_2^{-2r_2}).
\]

Using Lemma C.5(a)-(b), (C.24), (C.25), the Markov inequality, the Cauchy-Schwarz inequality
and the triangle inequality, we get

\[
\left| n^{-1} \sum_{i=1}^{n} (\ell(Z, \hat{\alpha}_k) - \ell(Z, \alpha_k^*)) (\ell(Z, \alpha_k^*) - \ell(Z, \alpha_n^*)) \right|
\]

\[
= \left| n^{-1} \sum_{i=1}^{n} (\ell(Z, \alpha_k^*) - \ell(Z, \alpha_n^*)) P_{k_j}(X_j) (\hat{\beta}_k - \beta_k^*) \right|
\]

\[
\leq \left| E_{F_n} \left( (\ell(Z, \alpha_k^*) - \ell(Z, \alpha_n^*)) P_{k_j}(X_j) (\hat{\beta}_k - \beta_k^*) \right) \right| + O((k_1^{-r_1} + k_2^{-r_2})|k|n^{-1})
\]

\[
= O_p(\|E_{F_n} \left( (\ell(Z, \alpha_k^*) - \ell(Z, \alpha_n^*)) P_{k_j}(X_j) \right) \| n^{-1/2}) + O((k_1^{-r_1} + k_2^{-r_2})|k|n^{-1})
\]

\[
= O_p((k_1^{-r_1} + k_2^{-r_2})(n^{-1/2} + |k|n^{-1})) = o_p(\sigma_F^2 n, n)
\]

where the last equality is by Assumptions A4(i) and A4(iv). By Assumption A4, (18), (23) and (24),

\[
\sum_{i=1}^{n} |\ell(Z, \hat{\alpha}_k) - \ell(Z, \alpha_n^*)|^2
\]

\[
= \sum_{i=1}^{n} |u_{k_1} [\hat{\alpha}_{k_1}(X_1) - \alpha_{k_1}^*(X_1)] - u_{k_2} [\hat{\alpha}_{k_2}(X_2) - \alpha_{k_2}^*(X_2)]|^2 + o_p(\sigma_F^2 n, n)
\]

\[
= (\hat{\beta}_k - \beta_k^*)' D_n(\alpha_k^*) (\hat{\beta}_k - \beta_k^*) + o_p(\sigma_F^2 n, n).
\]

By the Cauchy-Schwarz inequality,

\[
\left| (\hat{\beta}_k - \beta_k^*)' (\hat{\beta}_k - \beta_k^*) \right| 
\]

\[
\leq \left\| \beta_{k,n} - \beta_{k,n}^* \right\|^2 \left\| D_n(\alpha_k^*) - D_{F,n,k} \right\|
\]

\[
= O_p \left( \left( \frac{\xi_k \log(n)}{n^{1/2}} \right) \frac{1}{n^{1/2}} \right) O_p((k_1 + k_2)n^{-1}) = o_p(\sigma_F^2 n, n)
\]

where the first equality holds by (C.10) and Lemma C.5(a), and the second equality holds by Assumptions C.4(ii). Combining (C.27) with (C.28), we have

\[
\frac{1}{n} \sum_{i=1}^{n} |\ell(Z_i, \hat{\alpha}_n) - \ell(Z_i, \alpha_n^*)|^2 = (\hat{\beta}_k - \beta_k^*)' D_{F,n,k} (\hat{\beta}_k - \beta_k^*) + o_p(\sigma_F^2 n, n),
\]

which together with Lemma C.5(c) implies that

\[
\frac{1}{n} \sum_{i=1}^{n} |\ell(Z_i, \hat{\alpha}_k) - \ell(Z_i, \alpha_n^*)|^2 = \epsilon_{\alpha,n}(\alpha_k^*) H_{F,n,k}^{-1} D_{F,n,k} H_{F,n,k}^{-1} \epsilon_{\alpha,n}(\alpha_k^*) + o_p(\sigma_F^2 n, n).
\]
This verifies Assumption 4.4(b).

We next verify Assumption 4.4(c). First, notice that by the Cauchy-Schwarz inequality,
\[
\frac{\sum_{i=1}^{n}(\ell(Z_i, \alpha_n^*) - \ell_{F_n}(\alpha_n^*)) \langle \tilde{\alpha}_k(X_{j,i}) - \alpha_k^*(X_{j,i}) \rangle^2}{n\sigma_{F_n,n}^2} \leq \frac{\sum_{i=1}^{n}(\ell(Z_i, \alpha_n^*) - \ell_{F_n}(\alpha_n^*))^2 \sum_{i=1}^{n}|\tilde{\alpha}_k(X_{j,i}) - \alpha_k^*(X_{j,i})|^2}{n\sigma_{F_n,n}^2} = o_p(1) \quad (C.31)
\]
where the equality holds by the Markov inequality, Assumption C.1(i), \( E_{F_n}[(\ell(Z_i, \alpha_n^*) - \ell_{F_n}(\alpha_n^*))^2] \leq \sigma_{F_n,n}^2 \), and (C.12). That together with (C.17) implies that
\[
\sum_{i=1}^{n}(\ell(Z_i, \alpha_n^*) - \ell_{F_n}(\alpha_n^*)) \langle \tilde{\alpha}_k(X_{j,i}) - \alpha_k^*(X_{j,i}) \rangle \leq \sum_{i=1}^{n}(\ell(Z_i, \alpha_n^*) - \ell_{F_n}(\alpha_n^*)) \ell_{\alpha,k}(Z_i, \alpha_n^*) \langle \beta_k - \beta_k^* \rangle + o_p(1)
\]
\[
\sum_{i=1}^{n}(\ell(Z_i, \alpha_n^*) - \ell_{F_n}(\alpha_n^*)) \ell_{\alpha,k}(Z_i, \alpha_n^*) \langle \beta_k - \beta_k^* \rangle + \frac{\sum_{i=1}^{n}[\ell(Z_i, \alpha_n^*) - \ell_{F_n}(\alpha_n^*)) \ell_{\alpha,k}(Z_i, \alpha_n^*) - E_{F_n}[\ell(Z_i, \alpha_n^*) \ell_{\alpha,k}(Z_i, \alpha_n^*)] \langle \beta_k - \beta_k^* \rangle}{n\sigma_{F_n,n}^2} + o_p(1).
\]

Let \( \gamma_n = E_{F_n}[\ell(Z; \alpha_n^*) \ell_{\alpha,k}(Z; \alpha_n^*)]/\sigma_{F_n,n} \). Then
\[
\gamma_n' \gamma_n = \frac{\omega_{F_n,n}^2}{\sigma_{F_n,n}^2} E_{F_n} \left[ \frac{\ell(Z; \alpha_n^*)}{\omega_{F_n,n}} \ell_{\alpha,k}(Z; \alpha_n^*) (D_{F_n,k}^{1/2})^+ \right] D_{F_n,k} E_{F_n} \left[ (D_{F_n,k}^{1/2})^+ \ell_{\alpha,k}(Z; \alpha_n^*) \ell(Z; \alpha_n^*) / \omega_{F_n,n} \right] 
\]
\[
\leq C E_{F_n} \left[ \frac{\ell(Z; \alpha_n^*)}{\omega_{F_n,n}} \ell_{\alpha,k}(Z; \alpha_n^*) (D_{F_n,k}^{1/2})^+ \right] E_{F_n} \left[ \frac{\ell(Z; \alpha_n^*)}{\omega_{F_n,n}} \ell_{\alpha,k}(Z; \alpha_n^*) (D_{F_n,k}^{1/2})^+ \right] \leq C, \quad (C.33)
\]
where the first inequality holds by Assumption 4.4(f) which has been verified in Lemma C.1 and the second inequality holds by Lemma F.1. Equation (C.33) together with Lemma C.5(b) implies that
\[
E_{F_n}[\ell(Z; \alpha_n^*) \ell_{\alpha,k}(Z; \alpha_n^*) \langle \beta_k - \beta_k^* \rangle / \sigma_{F_n,n}^2 = O_p(n^{-1/2} \sigma_{F_n,n}^{-1})) = o(1). \quad (C.34)
\]
where the second equality holds by Assumption 4.4(i). Now consider,
\[
\frac{E_{F_n}[[\ell(Z; \alpha_n^*) - \ell_{F_n}(\alpha_n^*)) \ell_{\alpha,k}(Z; \alpha_k^*)]^2]}{\sigma_{F_n,n}^2} \]
Using the Hölder inequality, Assumptions C.1(iii)-(iv) and (C.2), we get

\[
E_{F_n} \left[ \frac{\ell(Z, \alpha_n^*) - \ell_{F_n}(\alpha_n^*)}{\sigma_{F_n,n}^2} \sum_{j=1,2} u_{k_j}^2 P_{k_j}(X_j)'P_{k_j}(X_j) \right].
\] (C.35)

which together with the Markov inequality implies that

\[
\| \sum_{i=1}^n [(\ell(Z, \alpha_n^*) - \ell_{F_n}(\alpha_n^*))\ell_{\alpha,k}(Z_i, \alpha_n^*) - E_{F_n} [(\ell(Z, \alpha_n^*)\ell_{\alpha,k}(Z_i, \alpha_n^*))]] \| = O_p \left( \sum_{j=1,2} \xi_{k_j} \right).
\] (C.37)

Then we have

\[
\left| \sum_{i=1}^n [(\ell(Z; \alpha_n^*) - \ell_{F_n}(\alpha_n^*))\ell_{\alpha,k}(Z_i; \alpha_n^*) - E_{F_n} [(\ell(Z; \alpha_n^*)\ell_{\alpha,k}(Z_i; \alpha_n^*))]](\hat{\beta}_k - \beta_k) \right|
\]

\[
\leq O_p \left( \sum_{j=1,2} \xi_{k_j} \right) \frac{\| \hat{\beta}_k - \beta_k \|}{n^{1/2} \sigma_{F_n,n}}
\]

\[
= O_p \left( \sum_{j=1,2} \xi_{k_j} \right) O_p \left( |k|^{1/2} \sigma_{F_n,n}^{-1} n^{-1} \right)
\]

\[
= O_p \left( \frac{1}{n^{1/2} \sigma_{F_n,n}} \right) O_p \left( \sum_{j=1,2} \xi_{k_j} |k|^{3/2} \right) = o_p(1),
\] (C.38)
By the Cauchy-Schwarz inequality,

$$\left| \sum_{i=1}^{n} (\ell(Z_i, \alpha_n^*) - \ell_{F_n}(\alpha_n^*)) (\ell(Z_i, \alpha_k^*) - \ell(Z_i, \alpha_n^*)) \right|^2 \leq n^2 \sigma_{F_{n,n}}^2$$

\[
\leq n^{-1} \sum_{i=1}^{n} \frac{(\ell(Z_i, \alpha_n^*) - \ell_{F_n}(\alpha_n^*)^2)}{\sigma_{F_{n,n}}^2} \times n^{-1} \sum_{i=1}^{n} \frac{(\ell(Z_i, \alpha_k^*) - \ell(Z_i, \alpha_n^*))^2}{\sigma_{F_{n,n}}^2} = O_p((k_1^{-2r_1} + k_2^{-2r_2}) \sigma_{F_{n,n}}^{-2}) = o_p(1) \tag{C.40}
\]

where the equality is by (C.23), Assumptions C.2(i) and C.4(iv), and the Markov inequality. Assumption 4.4(c) is implied by (C.39) and (C.40).

Assumption 4.4(d) is implied by Assumption C.4(iv).

Proof of Lemma C.5. Proof of the result in part (a) is standard and follows the similar arguments in Newey (1997). We include it for completeness. By the definition of \( \hat{\beta}_{k_j} \)

\[
\hat{\beta}_{k_j} - \beta_{k_j} = (P'_{k_j,n} P_{k_j,n})^{-1} P'_{k_j,n} U_{k_j,n} \tag{C.41}
\]

where \( U_{k_j,n} = [u_{k_{j1}}, \ldots, u_{k_{jn}}]' \). Therefore

\[
\| \hat{\beta}_{k_j} - \beta_{k_j} \| = U_{k_j,n}' P_{k_j,n} (P'_{k_j,n} P_{k_j,n})^{-2} P'_{k_j,n} U_{k_j,n} \leq \rho_{\min} (n^{-1} P'_{k_j,n} P_{k_j,n})^{-2} n^{-2} U_{k_j,n}' P_{k_j,n} P'_{k_j,n} U_{k_j,n} \tag{C.42}
\]

By Assumptions C.1(i) and C.2,

\[
E_{F_n} \left[ \sum_{i=1}^{n} \left( \sum_{i=1}^{n} p_{j,i} (X_{j,i} u_{k_{j,i}}) \right)^2 \right] = \sum_{i=1}^{k_j} \sum_{i=1}^{n} E_{F_n} \left[ (p_{j,i} (X_{j,i} u_{k_{j,i}})^2 \right] \leq n Ctr (-H_{F_{0,k_j}}) \tag{C.43}
\]

which together with the Markov inequality and Assumptions C.1(iv) implies that

\[
n^{-2} U_{k_j,n}' P_{k_j,n} P'_{k_j,n} U_{k_j,n} = O_p(k_j n^{-1}). \tag{C.44}
\]

The result in part (a) follows by (C.6), (C.42) and (C.44).

For part (b), first observe that

\[
\gamma_n' (\hat{\beta}_{k_j} - \beta_{k_j}^*) = \gamma_n' (P'_{k_j,n} P_{k_j,n})^{-1} P'_{k_j,n} U_{k_j,n} = n^{-1} \gamma_n' (-H_{F_{n,k_j}})^{-1} P'_{k_j,n} U_{k_j,n} + n^{-1} \gamma_n' ((n^{-1} P'_{k_j,n} P_{k_j,n})^{-1} - (-H_{F_{n,k_j}})^{-1}) P'_{k_j,n} U_{k_j,n} \tag{C.45}
\]

11
By the first order condition of $\alpha_{k_j}^*$,

$$E_{F_n} \left[ u_{k_j} P_{k_j}(X_j) \right] = 0_{k_j \times 1}. \quad (C.46)$$

By the definition of $H_{F_n,k_j}$, (C.46), Assumptions C.1(i) and C.1(iv), and (C.2)

$$E_{F_n} \left[ \left| \gamma_n'(-H_{F_n,k_j})^{-1} P'_{k_j,n} U_{k_j,n} \right|^2 \right] = E_{F_n} \left[ \left| \sum_{i=1}^{n} \gamma_n'(-H_{F_n,k_j})^{-1} P_{k_j}(X_{j,i}) u_{k_j,i} \right|^2 \right]
$$

$$= \sum_{i=1}^{n} E_{F_n} \left[ \left| \gamma_n'(-H_{F_n,k_j})^{-1} P_{k_j}(X_{j,i}) u_{k_j,i} \right|^2 \right]
$$

$$\leq nCE_{F_n} \left[ \left| \gamma_n'(-H_{F_n,k_j})^{-1} P_{k_j}(X_{j,i}) \right|^2 \right]
$$

$$= nC\gamma_n'(-H_{F_n,k_j})^{-1} \gamma_n = O(n\|\gamma_n\|^2), \quad (C.47)$$

which together with the Markov inequality implies that

$$n^{-1}\gamma_n'(-H_{F_n,k_j})^{-1} P'_{k_j,n} U_{k_j,n} = O_p(n^{-1/2}\|\gamma_n\|). \quad (C.48)$$

By the Cauchy-Schwarz inequality,

$$\left| n^{-1}\gamma_n'((n^{-1}P'_{k_j,n} P_{k_j,n})^{-1} - (-H_{F_n,k_j})^{-1}) P'_{k_j,n} U_{k_j,n} \right|
$$

$$\leq \|\gamma_n\| \left| n^{-1}P'_{k_j,n} P_{k_j,n} + H_{F_n,k_j} \right| \left( \rho_{\min}(n^{-1}P'_{k_j,n} P_{k_j,n}) \right)^{1/2} \rho_{\min}(-H_{F_n,k_j})
$$

$$= O_p((\xi_{k_j} k_j \log(k_j))^{1/2} n^{-1}\|\gamma_n\|) = O_p(n^{-1/2}\|\gamma_n\|), \quad (C.49)$$

where the first equality is by Assumption C.1(iv), (C.5), (C.6) and (C.44), the last equality is by Assumption C.4(iv). Combining the results in (C.45), (C.48) and (C.49), we prove the result in part (b).

For part (c), by the definitions of $\widehat{\beta}_k$ and $\beta_k^*$, we have

$$\widehat{\beta}_k - \beta_k^* = -(\widehat{H}_n)^{-1}\ell_{\alpha,n}(\alpha_k^*). \quad (C.50)$$

Hence, we can write

$$(\widehat{\beta}_k - \beta_k^*)' D_{F_n,k}(\widehat{\beta}_k - \beta_k^*) = \ell_{\alpha,n}(\alpha_k^*)'((\widehat{H}_n)^{-1} D_{F_n,k}(\widehat{H}_n)^{-1})\ell_{\alpha,n}(\alpha_k^*) \quad (C.51)$$
Using (C.44) and the Markov inequality, we get
\[
\ell_{a,n}(\alpha_k^*)\ell_{a,n}(\alpha_k^*) = \sum_{j=1,2} n^{-2} U'_{k_j,n} P_{k_j,n} P'_{k_j,n} U_{k_j,n} = O_p((k_1 + k_2)n^{-1}).
\] (C.52)

Note that
\[
\ell_{a,n}(\alpha_k^*)'((\hat{H}_n)^{-1}D_{F_n,k}(\hat{H}_n)^{-1} - H_{F_n,k}^{-1}D_{F_n,k}H_{F_n,k}^{-1})\ell_{a,n}(\alpha_k^*)
= \ell_{a,n}(\alpha_k^*)'((H_{F_n,k} - \hat{H}_n)H_{F_n,k}^{-1}D_{F_n,k}(\hat{H}_n)^{-1}\ell_{a,n}(\alpha_k^*)
+ \ell_{a,n}(\alpha_k^*)'H_{F_n,k}^{-1}D_{F_n,k}(\hat{H}_n)^{-1}(H_{F_n,k} - \hat{H}_n)H_{F_n,k}^{-1}\ell_{a,n}(\alpha_k^*). \quad (C.53)
\]

By the Cauchy-Schwarz inequality,
\[
\left| \ell_{a,n}(\alpha_k^*)'((H_{F_n,k} - \hat{H}_n)H_{F_n,k}^{-1}D_{F_n,k}(\hat{H}_n)^{-1}\ell_{a,n}(\alpha_k^*) \right|
\leq \left| \ell_{a,n}(\alpha_k^*)'((\hat{H}_n)^{-1}D_{F_n,k}(\hat{H}_n)^{-1}\ell_{a,n}(\alpha_k^*) \right|^{1/2} \times \left| (H_{F_n,k} - \hat{H}_n) \right| \times
\left| \ell_{a,n}(\alpha_k^*)'((H_{F_n,k} - \hat{H}_n)H_{F_n,k}^{-1}D_{F_n,k}(\hat{H}_n)^{-1}\ell_{a,n}(\alpha_k^*) \right|^{1/2}
\leq \rho_{\text{max}}(D_{F_n,k}) \rho_{\text{max}}^2(H_{F_n,k}^{-2}) \rho_{\text{max}}((\hat{H}_n)^{-2}) \left| \ell_{a,n}(\alpha_k^*)'\ell_{a,n}(\alpha_k^*) \right| \times \left| (H_{F_n,k} - \hat{H}_n) \right|
= O_p((k_1 + k_2)n^{-1})O_p((\xi_n \log(n)n^{-1})^{1/2}) = o_p(\sigma_{F_n,n}^2) \quad (C.54)
\]

where the first equality is by Assumption 4.1(f) (which has been verified in Lemma C.1), Lemma C.4, (C.6) and (C.52), the last equality is by Assumptions C.4(i). Similarly, we can show that
\[
\left| \ell_{a,n}(\alpha_k^*)'H_{F_n,k}^{-1}D_{F_n,k}(\hat{H}_n)^{-1}(H_{F_n,k} - \hat{H}_n)H_{F_n,k}^{-1}\ell_{a,n}(\alpha_k^*) \right| = o_p(\sigma_{F_n,n}^2) \quad (C.55)
\]

which together with (C.53) and (C.55) implies that
\[
\ell_{a,n}(\alpha_k^*)'((\hat{H}_n)^{-1}D_{F_n,k}(\hat{H}_n)^{-1} - H_{F_n,k}^{-1}D_{F_n,k}H_{F_n,k}^{-1})\ell_{a,n}(\alpha_k^*) = o_p(\sigma_{F_n,n}^2). \quad (C.56)
\]

Then part(c) is proved by collecting the results in (C.51) and (C.56). ■

**Proof of Lemma C.6.** Proof of the results in part (a) and part (b) is standard and follows the similar arguments in Newey (1997). We include it for completeness. Let \( U_{k_j,n} = [u_{k_j,1}, \ldots, u_{k_j,n}]' \).

By the definition of \( \hat{\alpha}_{k_j} \) and \( \alpha_{k_j} \),
\[
\hat{\alpha}_{k_j}(X_{j,i}) - \alpha_{k_j}(X_{j,i}) = P_{k_j}(X_{j,i})'(\hat{\beta}_{k_j} - \beta_{k_j}). \quad (C.57)
\]
which implies that

\[ n^{-1} \sum_{i=1}^{n} \left| \tilde{\alpha}_{k_j}(X_{j,i}) - \alpha^*_k(X_{j,i}) \right|^2 \]

\[ = (\hat{\beta}_{k_j} - \beta^*_j)^t (n^{-1} \mathbf{P}'_{k_j,n} \mathbf{P}_{k_j,n}) (\hat{\beta}_{k_j} - \beta^*_j) \]

\[ \leq \rho_{\max} (n^{-1} \mathbf{P}'_{k_j,n} \mathbf{P}_{k_j,n}) \| \hat{\beta}_{k_j} - \beta^*_j \|^2 = O_p(k_j n^{-1}) \quad (C.58) \]

where the second equality is by Assumption C.1(iv), (C.6) and Lemma C.5(a). This shows part (a).

Now we show part (b). By (C.41) and (C.57),

\[ n^{-1} \sum_{i=1}^{n} \left| Y_i - \tilde{\alpha}_{k_j}(X_{j,i}) \right|^2 = n^{-1} \sum_{i=1}^{n} \left[ u_{k_j,i} - (\tilde{\alpha}_{k_j}(X_{j,i}) - \alpha^*_k(X_{j,i})) \right]^2 \]

\[ = n^{-1} \sum_{i=1}^{n} u_{k_j,i}^2 + n^{-1} \sum_{i=1}^{n} \left| \tilde{\alpha}_{k_j}(X_{j,i}) - \alpha^*_k(X_{j,i}) \right|^2 \]

\[ -2n^{-1} \sum_{i=1}^{n} (\tilde{\alpha}_{k_j}(X_{j,i}) - \alpha^*_k(X_{j,i})) u_{k_j,i} \]

\[ = n^{-1} \sum_{i=1}^{n} u_{k_j,i}^2 - n^{-1} U'_{k_j,n} \mathbf{P}_{k_j,n} (\mathbf{P}'_{k_j,n} \mathbf{P}_{k_j,n})^{-1} \mathbf{P}'_{k_j,n} U_{k_j,n}. \quad (C.59) \]

Since \( u_{k_j,i} - u_i = \alpha^*_j(X_{j,i}) - \alpha^*_k(X_{j,i}) \),

\[ n^{-1} \sum_{i=1}^{n} u_{k_j,i}^2 - n^{-1} \sum_{i=1}^{n} u_i^2 \]

\[ = n^{-1} \sum_{i=1}^{n} (u_{k_j,i} - u_i)^2 - 2n^{-1} \sum_{i=1}^{n} (u_{k_j,i} - u_i) u_i \]

\[ = n^{-1} \sum_{i=1}^{n} (\alpha^*_j(X_{j,i}) - \alpha^*_j(X_{j,i}))^2 + 2n^{-1} \sum_{i=1}^{n} (\alpha^*_k(X_{j,i}) - \alpha^*_j(X_{j,i})) u_i. \quad (C.60) \]

The pseudo-true value of the parameter can be written as the limit of a sequence of sieve approximation: \( \alpha^*_j(x_j) = \sum_{i=1}^{\infty} p_{j,i}(x_j) \beta^*_j(\epsilon) \), where \( (\beta^*_j(\epsilon))_{\epsilon=1}^{\infty} = \arg\min_{\beta_{j,\ell} \in \mathbb{R}_{\geq 0}} \left[ |Y - \sum_{i=1}^{\infty} p_{j,i}(X_j) \beta_{j,\ell}|^2 \right] \). By the definition of \( u_j \) and the first order optimality condition of \( (\beta^*_j(\epsilon))_{\epsilon=1}^{\infty} \),

\[ E_{F_n} [u_j p_{j,\ell}(X_j)] = 0 \quad (C.61) \]

for \( j = 1, 2 \) and for any \( \ell = 1, 2, \ldots \), which implies that \( E_{F_n} [u_j \alpha^*_j(X_j)] = 0 \) and \( E_{F_n} [u_j \alpha^*_k(X_j)] = 0 \).
for $j = 1, 2$. Therefore

$$
E_{F,n} \left[ n^{-1} \sum_{i=1}^{n} (\alpha_{k_j}(X_{j,i}) - \alpha_{j}^{*}(X_{j,i}))u_i \right]^2
$$

\[= n^{-2} \sum_{i=1}^{n} E_{F,n} \left[ (\alpha_{k_j}(X_{j,i}) - \alpha_{j}^{*}(X_{j,i}))^2 u_i^2 \right] = O(n^{-1}k_{j}^{-2r_j}), \tag{C.62}
\]

where the first equality is by Assumption C.1(i), the second equality is by Assumptions C.1(ii) and C.2(ii). Combining the results in (C.59), (C.60), (C.62) and applying Assumption C.1(ii) and the Markov inequality, we deduce that

$$
\frac{1}{2} \sum_{i=1}^{n} \left[ |Y_i - \tilde{\alpha}_{k_j}(X_{j,i})|^2 - |Y_i - \alpha_{j}^{*}(X_{j,i})|^2 \right] \leq \frac{1}{2} U'_{k_j,n} P_{k_j,n} (P'_{k_j,n} P_{k_j,n})^{-2} P'_{k_j,n} U_{k_j,n} + O_p(k_{j}^{-r_j}n^{1/2}) + O(nk_{j}^{-2r_j}). \tag{C.63}
\]

By (C.6) and (C.44),

$$
U'_{k_j,n} P_{k_j,n} (P'_{k_j,n} P_{k_j,n})^{-2} P'_{k_j,n} U_{k_j,n} \leq \rho_{\max}((n^{-1}P'_{k_j,n} P_{k_j,n})^{-2})n^{-2} \sum_{i=1}^{n} U'_{k_j,n} P_{k_j,n} (P'_{k_j,n} P_{k_j,n})^{-2} P'_{k_j,n} U_{k_j,n} = O_p(k_{j}^{-1}). \tag{C.64}
\]

and similarly

$$
n^{-1}U'_{k_j,n} P_{k_j,n} H_{F_n,k_j}^{-2} P'_{k_j,n} U_{k_j,n} = O_p(k_{j}). \tag{C.65}
\]

Using the Cauchy-Schwarz inequality, Lemma C.4, (C.64) and (C.65), we have

$$
\left| \frac{U'_{k_j,n} P_{k_j,n} (P'_{k_j,n} P_{k_j,n})^{-1} P'_{k_j,n} U_{k_j,n} - U'_{k_j,n} P_{k_j,n} (-H_{F_n,k_j})^{-1} P'_{k_j,n} U_{k_j,n}}{n} \right|^2
\]

$$
= \left| \frac{U'_{k_j,n} P_{k_j,n} (P'_{k_j,n} P_{k_j,n})^{-1} (n^{-1}P_{k_j,n} P_{k_j,n} + H_{F_n,k_j})(-H_{F_n,k_j})^{-1} P'_{k_j,n} U_{k_j,n}}{n^2} \right|^2
\]

$$
\leq \left| n^{-1}P_{k_j,n} P_{k_j,n} + H_{F_n,k_j} \right|^2 \times \left( U'_{k_j,n} P_{k_j,n} (P'_{k_j,n} P_{k_j,n})^{-2} P'_{k_j,n} U_{k_j,n} \right) \times
$$

$$
\frac{U'_{k_j,n} P_{k_j,n} H_{F_n,k_j}^{-2} P'_{k_j,n} U_{k_j,n}}{n^2} = O_p(\log(k_{j})k_{j}^2n^{-3}). \tag{C.66}
\]

That together with (C.63) implies that

$$
\frac{1}{2n} \sum_{i=1}^{n} \left[ |Y_i - \tilde{\alpha}_{k_j}(X_{j,i})|^2 - |Y_i - \alpha_{j}^{*}(X_{j,i})|^2 \right]
$$

15
\[= - \frac{U'_{k_j,n} P_{k_j,n} (-H_{F_n,k_j})^{-1} P'_{k_j,n} U_{k_j,n}}{2n^2} + O_p((\log(k_j)\xi_{k_j})^{1/2}k_jn^{-3/2} + k_j^{-\tau_j} n^{-1/2} + k_j^{-2\tau_j}). \quad (C.67)\]

Now observe that \(\bar{\ell}_n(\hat{\alpha}_{k,n})\) can be decomposed as

\[
\bar{\ell}_n(\hat{\alpha}_{k,n}) = \frac{1}{2n} \sum_{i=1}^{n} \left[ |Y_i - \alpha_1^*(X_{1,i})|^2 - |Y_i - \alpha_2^*(X_{2,i})|^2 \right] \\
+ \frac{1}{2n} \sum_{i=1}^{n} \left[ |Y_i - \hat{\alpha}_{k_1}(X_{1,i})|^2 - |Y_i - \alpha_1^*(X_{1,i})|^2 \right] \\
- \frac{1}{2n} \sum_{i=1}^{n} \left[ |Y_i - \hat{\alpha}_{k_2}(X_{2,i})|^2 - |Y_i - \alpha_2^*(X_{2,i})|^2 \right] \quad (C.68)
\]

which combined with (C.67) proves the lemma. □

D Illustration II: Nonparametric Quantile-Regression

For \(j = 1, 2\), consider the model

\[Y = \alpha_j^*(X_j) + u_j, \quad E_{F_0}[I\{u_j \leq 0\}|X_j] = \tau, \quad (D.1)\]

where \(\tau \in (0, 1)\), \(u_j\) is a unobservable error term, \(\alpha_j^*(x)\) is a unknown function and \(F_0\) denotes the joint distribution of \(Z \equiv (Y, X_1, X_2)\). The regressors \(X_1\) and \(X_2\) in two equations may be nested, over-lapped or strictly non-nested. Note that even when the regressors are strictly non-nested in the intuitive sense (i.e., \(X_1\) and \(X_2\) do not contain any common variable), models represented by the two equations in general are still overlapping according to the definitions in Section 2.2 of the main paper because it may be possible that \(\alpha_1^*(X_1) = \alpha_2^*(X_2) = \text{Constant}\).

Suppose that the unknown function \(\alpha_j^*(\cdot)\) of model \(j\) belongs to the set \(A_j\). We use the finite-dimensional approximations:

\[A_{j,k_j} = \{\alpha_{k_j}(\cdot)|\alpha_{k_j}(\cdot) = \alpha_j(\beta_{k_j}) = P_{k_j}(\cdot)'\beta_{k_j}|\beta_{k_j} \in \mathbb{R}^{k_j}\} \quad (D.2)\]

where \(P_{k_j}(\cdot) = [p_{j,1}(\cdot), \ldots, p_{j,k_j}(\cdot)]'\) is a \((k_j\text{-dimensional})\) vector of basis functions. Let \(\rho_{\tau}(u) = (I\{u \leq 0\} - \tau) u\). Define

\[\beta_{k_j,F_0}^* = \arg \min_{\beta_{k_j} \in \mathbb{R}^{k_j}} E_{F_0} [\rho_{\tau}(Y - P_{k_j}(X_j)'\beta_{k_j})] \quad \text{for } j = 1, 2. \quad (D.3)\]

We use \(\beta_{k_j}^*\) to denote the counterpart of \(\beta_{k_j,F_0}^*\), when the expectation in (D.3) is taken with respect
to any DGP $F \in \mathcal{F}$. We suppress the dependence of $\beta_{kj}^*$ on $F$ for notational convenience. The approximate M-estimator in this example is defined as

$$\hat{\beta}_{kj} = \arg\max_{\beta_{kj} \in \mathbb{R}^j} n^{-1} \sum_{i=1}^n \rho_\tau(Y_i - P_{kj}(X_{j,i})') \beta_{kj} \text{ for } j = 1, 2. \quad (D.4)$$

In this example, the pseudo-density ratio function is

$$\ell(Z; \alpha) = \rho_\tau(Y - \alpha_1(X_1)) - \rho_\tau(Y - \alpha_2(X_2)). \quad (D.5)$$

The score function $\ell_{\alpha,k}(Z; \alpha)$ in this example is

$$\ell_{\alpha,k}(Z; \alpha) = \left( \begin{array}{c} (\tau - I\{Y \leq \alpha_1(X_1)\}) P_{k_1}(X_1) \\ (I\{Y \leq \alpha_2(X_2)\} - \tau) P_{k_2}(X_2) \end{array} \right), \quad (D.6)$$

which combined with the formula in (3.10) defines

$$D_{F_0,k} = E_{F_0} \begin{bmatrix} u_{k_1}^2 P_{k_1}(X_1) P_{k_1}(X_1)' & -u_{k_1} u_{k_2} P_{k_1}(X_1) P_{k_2}(X_2)' \\ -u_{k_1} u_{k_2} P_{k_2}(X_2) P_{k_1}(X_1)' & u_{k_2}^2 P_{k_2}(X_2) P_{k_2}(X_2)' \end{bmatrix}, \quad (D.7)$$

where $u_{kj} = I\{Y \leq \alpha_{kj}^*(X_j)\} - \tau$. The Hessian matrix $H_{F_0,kj}$ is

$$H_{F_0,kj} = -E_{F_0} \left[ f_{uj}(\Delta(\alpha_{kj}^*)|X_j) P_{kj}(X_j) P_{kj}(X_j)' \right] \text{ for } j = 1, 2 \quad (D.8)$$

where $\Delta(\alpha_j) = \alpha_j(X_j) - \alpha_{kj}^*(X_j)$ and $f_{uj}(\cdot|X_j)$ denotes the conditional density of $u_j$ given $X_j$. In this example, $H_{F_0,k}$ can be constructed using (3.14) and (D.8).

We next provide the sufficient conditions for Theorem 4.1 in this example.

**Assumption D.1** For any $F_0 \in \mathcal{F}$ and for $j = 1, 2$: (i) $f_{uj}(u_j|x_j) < C$ uniformly in $u_j$ and $x_j$; (ii) $\partial f_{uj}(u_j|x_j)/\partial u_j < C$ uniformly in $u_j$ and $x_j$; (iii) $f_{uj}(0|x_j) > C^{-1}$ uniformly in $x_j$; (iv) $\xi_{kj} k_j(\log(n))^{2-1} + k_j^{-r_j} \log(n) = o(1)$; (v) $E_{F_0}[\rho_\tau(u_1) - \rho_\tau(u_2)] = 0$ when $F_0 \in \mathcal{F}$; (vi) $E_{F_0}[u_j^2] < C$ for $j = 1, 2$; (vii) there exists a $\delta > 0$ such that $E \left[ \omega_{F_0,*}^{-1}(\rho_\tau(u_1) - \rho_\tau(u_2)) \right]^{2+\delta} < C$ whenever $\omega_{F_0,*}^2 > 0$.

Let $\delta_{0,j,n} = (k_j \log(n) n^{-1})^{1/2}$, $\delta_{1,j,n} = \delta_{0,j,n} + k_j^{-r_j}$ and $\delta_{2,j,n} = \delta_{0,j,n} \xi_{kj}^{1/2} + k_j^{-r_j}$.

**Theorem D.1** Suppose that Assumptions $C.1$ and $D.1$ hold, where $\beta_{kj,F_0}^*$ in Assumption $C.1(ii)$
is defined in \([D.3]\). If for any sequence of DGP's \(\{F_n\}_{n\geq 1}\) such that \(F_n \in \mathcal{F}_0\), we also have

\[
\frac{1}{n\sigma^2_{F_n,n}} + \max_{j=1,2} \left\{ \frac{\delta_{1,j,n}(k_j \delta_{2,j,n} \log(n))^{1/2}}{\sigma_{F_n,n}} + \frac{1 + k_j^{-r_j} n^{1/2}}{k_j^{r_j} \sigma_{F_n,n}} \right\} = o(1), \tag{D.9}
\]

then under \(F_n\) for all \(n\),

\[
\frac{n \tilde{\ell}_n(\hat{\alpha}_k) + (1/2) tr(\hat{D}_n(\alpha^*_k) H^{-1}_{F_n,k})}{\sqrt{n \omega_{F_0,*}^2 + \frac{n-1}{2n} \omega_{F_n,k}^2}} \rightarrow_d N(0, 1), \tag{D.10}
\]

where \(\hat{D}_n(\alpha^*_k) = n^{-1} \sum_{i=1}^n P_{u,k}(X_i) P_{u,k}(X_i)'\) and \(P_{u,k}(X_i)' = [u_{k1} P_{k1} (X_{1,i})', u_{k2} P_{k2} (X_{2,i})']\).

To prove Theorem \(D.1\) we need to derive some useful lemmas. For this purpose, we introduce some notations and preliminary results in the literature.

Under Assumptions \(C.1\) and \(D.1\) we can invoke Theorem 1 in Belloni, Chernozhukov, Chetverikov, and Fernández-Val (2011) to get

\[
\|\hat{\beta}_{kj} - \beta^*_{kj}\| = O_p \left( k_{j}^{1/2} n^{-1/2} \right) \text{ for } j = 1, 2. \tag{D.11}
\]

Let \(\mathcal{N}_{j,n} = \{\beta_{kj} \in \mathbb{R}^{k_j} : \|\beta_{kj} - \beta^*_{kj}\| \leq C \delta_{0,j,n}\}\) where \(C\) is some fixed and large constant. Then we have \(\hat{\beta}_{kj} \in \mathcal{N}_{j,n}\) with probability approaching 1. Using (D.11), Assumptions \(C.1\)(ii) and (iv), we have

\[
\sup_{\beta_{kj} \in \mathcal{N}_{j,n}} E_{\mathcal{F}_0} \left[ \left| P'_{k_j}(X_j) \beta_{kj} - \alpha^*_j(X_j) \right|^2 \right] \leq C \sup_{\beta_{kj} \in \mathcal{N}_{j,n}} \left\| \beta_{kj} - \beta^*_{kj} \right\|^2 + k_{j}^{-2r_j} = O(\delta_{1,j,n}^2) \tag{D.12}
\]

for any \(\mathcal{F}_0 \in \mathcal{F}\). Moreover, using (D.11), Assumptions \(C.1\)(ii), (iii) and (iv), we get

\[
\sup_{\beta_{kj} \in \mathcal{N}_{j,n}} \left| P'_{k_j}(x_j) \beta_{kj} - \alpha^*_j(x_j) \right| \leq \sup_{\beta_{kj} \in \mathcal{N}_{j,n}} \left\| \beta_{kj} - \beta^*_{kj} \right\| \xi_{k_j}^{1/2} + C k_{j}^{-r_j} = O(\delta_{2,j,n}) \tag{D.13}
\]

uniformly over \(x_j\).

For ease of notations, we define \(\Delta_i(\alpha_j) = \alpha_j(X_{j,i}) - \alpha^*_j(X_{j,i})\) and \(\Delta(\alpha_j) = \alpha_j(X_j) - \alpha^*_j(X_j)\) for \(j = 1, 2\). Then by definition,

\[
\rho_r(Y - \alpha_j(X_j)) = (I\{u_j \leq \Delta(\alpha_j)\} - \tau) (u_j - \Delta(\alpha_j)) = (I\{u_j \leq 0\} - \tau) u_j - (I\{u_j \leq 0\} - \tau) \Delta(\alpha_j) - (I\{u_j \leq 0\} - I\{u_j \leq \Delta(\alpha_j)\}) (u_j - \Delta(\alpha_j)), \tag{D.14}
\]
which together with \( \rho_r(Y - \alpha^*_j(X_j)) = (I\{u_j \leq 0\} - \tau) \) implies that

\[
\rho_r(Y - \alpha^*_j(X_j)) - \rho_r(Y - \alpha_j(X_j)) = (I\{u_j \leq 0\} - \tau) \Delta(\alpha_j)
\]

\[
= (I\{u_j \leq 0\} - I\{u_j \leq \Delta(\alpha_j)\}) (u_j - \Delta(\alpha_j)). \tag{D.15}
\]

The above expression is useful to derive the expansion of the QLR statistic.

Let \( \mu_n [g(Z)] = n^{-1} \sum_{i=1}^n [g(Z_i) - E_{F_0}[g(Z_i)]] \) denote the empirical process indexed by function \( g \).

\textbf{Lemma D.1} \textit{Under Assumptions C.1 and D.1 we have}

\[
\mu_n [(I\{u_j \leq \Delta(\bar{\alpha}_k)\} - I\{u_j \leq 0\}) u_j] = O_p \left( \delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n}^{1/2} \right). \tag{D.16}
\]

\textbf{Proof of Lemma D.1} Define \( F_{1,j,n} = \{ (I\{u_j \leq \Delta(P_k(X_j)'\beta_{kj})\} - I\{u_j \leq 0\}) u_j : \beta_{kj} \in \mathcal{N}_{j,n} \} \) for \( j = 1, 2 \). It is clear that the VC-dimension of \( F_{1,j,n} \) is \( O(k_j) \). By \( \text{D.13} \), we have

\[
\sup_{\beta_{kj} \in \mathcal{N}_{j,n}} |(I\{u_j \leq \Delta(P_k(X_j)'\beta_{kj})\} - I\{u_j \leq 0\}) u_j| \leq \sup_{\beta_{kj} \in \mathcal{N}_{j,n}} |\Delta(P_k(X_j)'\beta_{kj})| \leq C \delta_{2,j,n}. \tag{D.17}
\]

By definition,

\[
E_{F_n} \left[ \left| (I\{u_j \leq \Delta(P_k(X_j)'\beta_{kj})\} - I\{u_j \leq 0\}) u_j \right|^2 \right]
\]

\[
= E_{F_n} \left[ I\{0 \leq u_j \leq \Delta(P_k(X_j)'\beta_{kj})\} u_j^2 \right] + E_{F_n} \left[ I\{\Delta(P_k(X_j)'\beta_{kj}) \leq u_j \leq 0\} u_j^2 \right]. \tag{D.18}
\]

Using Assumption \( \text{D.1(i)} \), we get

\[
E_{F_n} \left[ I\{0 \leq u_j \leq \Delta(\alpha_j)\} u_j^2 \right] \leq \int_0^{\Delta(\alpha_j)} u^2 f_{u_j}(u|X_j) du \leq C |\Delta(\alpha_j)|^3. \tag{D.19}
\]

which together with \( \text{D.12} \) and \( \text{D.13} \) implies that for any \( \beta_{kj} \in \mathcal{N}_{j,n} \),

\[
E_{F_n} \left[ I\{0 \leq u_j \leq \Delta(P_k(X_j)'\beta_{kj})\} u_j^2 \right] \leq CE_{F_n} \left[ |\Delta(P_k(X_j)'\beta_{kj})|^3 \right] \leq CE_{F_n} \left[ \Delta(P_k(X_j)'\beta_{kj})^2 \right] \delta_{2,j,n} \leq C \delta_{1,j,n}^{1/2} \delta_{2,j,n}. \tag{D.20}
\]

Similarly, we have

\[
E_{F_n} \left[ I\{\Delta(P_k(X_j)'\beta_{kj}) \leq u_j \leq 0\} u_j^2 \right] \leq C \delta_{1,j,n}^{1/2} \delta_{2,j,n} \tag{D.21}
\]
Combining the results in (D.25) and (D.27) and the VC-dimension of $F$ which together with Assumption C.1(ii) and (D.12) implies that for any $\beta$

$$E_{F_n} \left[ \left( I\{u_j \leq 0\} - I\{u_j \leq \Delta(P_{k_j}(X_j)\beta_{k_j})\} \right) u_j \right] \leq C\delta_{1,j,n}^2 \delta_{2,j,n}. \quad (D.22)$$

Combining the results in (D.17), (D.22) and the VC-dimension of $F_{1,j,n}$ is $O(k_j)$, we can invoke Lemma 22.3 and Lemma 23 in Belloni, Chernozhukov, Chetverikov, and Fernández-Val (2011) to show that

$$\mu_n \left[ \left( I\{u_j \leq \Delta(\hat{\alpha}_{k_j})\} - I\{u_j \leq 0\} \right) \Delta(\alpha_{k_j}^*) \right] = O_p(\delta_{0,j,n} \delta_{1,j,n}^{\frac{1}{2}} \delta_{2,j,n}^2 + \delta_{0,j,n}^{\frac{1}{2}} \delta_{2,j,n}). \quad (D.23)$$

That together with $\delta_{0,j,n} \delta_{2,j,n} \leq \delta_{0,j,n} \delta_{1,j,n}^{\frac{1}{2}} \delta_{2,j,n} = o(1)$ (which is implied by Assumption D.1(iv)) implies (D.16).

**Lemma D.2** Under Assumptions C.1 and D.1 we have

$$\mu_n \left[ \left( I\{u_j \leq \Delta(\hat{\alpha}_{k_j})\} - I\{u_j \leq 0\} \right) \Delta(\alpha_{k_j}^*) \right] = O_p(\delta_{0,j,n} \delta_{1,j,n}^{\frac{1}{2}} \delta_{2,j,n}^2). \quad (D.24)$$

**Proof of Lemma D.2** For $j = 1, 2$, we define

$$F_{2,j,n} = \left\{ \left( I\{u_j \leq \Delta(P_{k_j}(X_j)\beta_{k_j})\} - I\{u_j \leq 0\} \right) \Delta(\alpha_{k_j}^*) : \beta_{k_j} \in \mathcal{N}_{j,n} \right\}.$$

It is clear that the VC-dimension of $F_{2,j,n}$ is $O(k_j)$. By Assumption C.1(ii), we have

$$\sup_{\beta_{k_j} \in \mathcal{N}_{j,n}} \left| \left( I\{u_j \leq \Delta(P_{k_j}(X_j)\beta_{k_j})\} - I\{u_j \leq 0\} \right) \Delta(\alpha_{k_j}^*) \right| \leq Ck_j^{-r_j}. \quad (D.25)$$

Under Assumption D.1(i),

$$E_{F_n} \left[ \left| I\{u_j \leq \Delta(\alpha_{j})\} - I\{u_j \leq 0\} \right|^2 \right| X_j \right] = E_{F_n} \left[ I\{0 \leq u_j \leq \Delta(\alpha_{j})\} \right] + E_{F_n} \left[ I\{\Delta(\alpha_{j}) \leq u_j \leq 0\} \right] \right| X_j \right] \leq 2 \left| \int_0^{\Delta(\alpha_{j})} f_{u_j}(u|X_j)du \right| \leq C|\Delta(\alpha_{j})|, \quad (D.26)$$

which together with Assumption C.1(ii) and (D.12) implies that for any $\beta_{k_j} \in \mathcal{N}_{j,n}$

$$E_{F_n} \left[ \left| I\{u_j \leq \Delta(P_{k_j}(X_j)\beta_{k_j})\} - I\{u_j \leq 0\} \right|^2 \Delta^2(\alpha_{k_j}^*) \right] \leq Ck_j^{-2r_j} \delta_{1,j,n}. \quad (D.27)$$

Combining the results in (D.25), (D.27) and the VC-dimension of $F_{2,j,n}$ is $O(k_j)$, we can invoke Lemma 22.3 and Lemma 23 in Belloni, Chernozhukov, Chetverikov, and Fernández-Val (2011) to
get
\[
\mu_n \left[ (I\{u_j \leq \Delta(\hat{\alpha}_{kj})\} - I\{u_j \leq 0\}) \Delta(\alpha^*_{kj}) \right] = O_p \left( \delta_{0,j,n}^{1/2} \frac{\delta_{1,j,n}^{1/2} k^{-r_j}}{k_j} + \delta_{0,j,n}^{2} \frac{\delta_{1,j,n}^{2} k^{-r_j}}{k_j} \right).
\]  
(D.28)

That together with \( \delta_{0,j,n} = o(1) \) (which is implied by Assumption D.1(iv)) implies (D.24).

\[ \blacksquare \]

**Lemma D.3** Under Assumptions C.1 and D.1 we have
\[
\sup_{\lambda_j \in S_{kj}} \mu_n \left[ (I\{u_j \leq \Delta(\hat{\alpha}_{kj})\} - I\{u_j \leq 0\}) \lambda_j^2 P_{kj}(X_j) \right] = O_p(\delta_{0,j,n}^{1/2}),
\]
where \( S_{kj} = \{ \lambda_j \in R^{k_j} : \|\lambda_j\| = 1 \} \).

**Proof of Lemma D.3.** Using the same arguments in the proof of Lemma 33 in Belloni, Chernozhukov, Chetverikov, and Fernández-Val (2011), we have
\[
\sup_{\lambda_j \in S_{kj}} \mu_n \left[ (I\{u_j \leq \Delta(\hat{\alpha}_{kj})\} - I\{u_j \leq 0\}) \lambda_j^2 P_{kj}(X_j) \right] = O_p(\delta_{0,j,n}^{3/2} \xi_{kj}^{1/4} + \delta_{0,j,n}^{3/2} \xi_{kj}^{1/2}).
\]  
(D.30)

As \( \delta_{0,j,n}^{2} \xi_{kj} = o(1) \) by Assumption D.1(iv), we get
\[
\frac{\delta_{0,j,n}^{3/2} \xi_{kj}^{1/4}}{\delta_{0,j,n}^{3/2} \xi_{kj}^{1/4}} = \frac{\delta_{0,j,n}^{3/2} \xi_{kj}^{1/4}}{\delta_{0,j,n}^{3/2} \xi_{kj}^{1/4}} = o(1)
\]
which together with (D.30) implies that
\[
\sup_{\lambda_j \in S_{kj}} \mu_n \left[ (I\{u_j \leq \Delta(\hat{\alpha}_{kj})\} - I\{u_j \leq 0\}) \lambda_j^2 P_{kj}(X_j) \right] = O_p(\delta_{0,j,n}^{3/2} \xi_{kj}^{1/4}).
\]  
(D.32)

For \( j = 1, 2 \), we define
\[ \mathcal{F}_{3,j,n} = \left\{ (I\{u_j \leq \Delta(\alpha^*_{kj})\} - I\{u_j \leq 0\}) \lambda_j^2 P_{kj}(X_j) : \lambda_j \in S_{kj} \right\}. \]

It is clear that the VC-dimension of \( \mathcal{F}_{3,j,n} \) is \( O(k_j) \). By Assumption C.1(iii), we have
\[
\sup_{\lambda_j \in S_{kj}} \left| (I\{u_j \leq \Delta(\alpha^*_{kj})\} - I\{u_j \leq 0\}) \lambda_j^2 P_{kj}(X_j) \right| \leq \xi_{kj}^{1/2}.
\]
(D.33)

Under Assumptions C.1(ii), C.1(iv) and D.1(i),
\[
E_{F_n} \left[ \left| (I\{u_j \leq \Delta(\alpha^*_{kj})\} - I\{u_j \leq 0\}) \lambda_j^2 P_{kj}(X_j) \right|^2 \right]
\]
Combining the results in (D.33), (D.34) and the VC-dimension of $J_{3,j,n}$ is $O(k_j)$, we can invoke Lemma 22.3 and Lemma 23 in [Belloni, Chernozhukov, Chetverikov, and Fernández-Val (2011)] to get

$$
\sup_{\lambda_j \in S_{kj}} \mu_n \left[ (I\{u_j \leq \Delta(\alpha_{kj}^*)\} - I\{u_j \leq 0\}) \lambda_j' P_k (X_j) \right] = O_p(k_j^{-r_j/2} + \delta_{0,j,n}^2 k_j^{-1/2}).
$$

(D.35)

which together with (D.32) and $\delta_{0,j,n}^2 k_j^{-1/2} \lambda_{0,j,n}^{1/4} (\delta_{0,j,n}^{1/4})^{-1} = (\delta_{0,j,n}^{1/2})^{1/2} = O(1)$ (which is implied by Assumption [D.1]iv) implies that

$$
\sup_{\lambda_j \in S_{kj}} \mu_n \left[ (I\{u_j \leq \Delta(\hat{\alpha}_k)\} - I\{u_j \leq 0\}) \lambda_j' P_k (X_j) \right] = O_p(\delta_{0,j,n}^{3/2} k_j^{-r_j/2} + \delta_{0,j,n}^{1/2} k_j^{-r_j/2}).
$$

(D.36)

Since $\delta_{0,j,n}^{3/2} k_j^{-r_j/2} = O(\delta_{0,j,n}^{1/2} k_j^{-r_j/2})$, (D.29) follows by (D.36).


\textbf{Lemma D.4} Under Assumptions [C.1] and [D.1] we have

$$
E_{Z,F_n} \left[ (I\{u_j \leq \Delta(\hat{\alpha}_k)\} - I\{u_j \leq 0\}) [u_j - (\hat{\alpha}_k - \alpha_{kj}^*)] \right]
$$

$$
= \frac{(\hat{\beta}_k - \beta_{kj}^*)' H_{F_n,k_j} (\hat{\beta}_k - \beta_{kj}^*)}{2} + O_p(\delta_{0,j,n}^2 + \delta_{0,j,n}^{-1/2} k_j^{-r_j} + k_j^{-r_j} + n^{-1/2} k_j^{-r_j}),
$$

(D.37)

where $E_{Z,F_n} [\cdot]$ denotes expectation taken on $Z$ under the joint distribution $F_n$.

\textbf{Proof of Lemma D.4.} Let $H_{F_n,k_j}^* = -E_{F_n} \left[ f_{u_j}(0|X_j) P_k (X_j) P_k (X_j)' \right]$. By Assumptions [C.1]iv, [D.1]i and [D.1]iii,

$$
-C \leq \rho_{\min} (H_{F_n,k_j}^*) \leq \rho_{\max} (H_{F_n,k_j}^*) \leq -C^{-1}
$$

(D.38)

uniformly in $n$ for $j = 1, 2$. By Assumptions [C.1]ii, [C.1]iv and [D.1]ii,

$$
\| E_{F_n} \left[ (f_{u_j}(0|X_j) - f_{u_j}(\Delta(\alpha_{kj}^*)|X_j)) \lambda' P_k (X_j) P_k (X_j)' \right] \|^2
$$

$$
\leq C E_{F_n} \left[ (f_{u_j}(0|X_j) - f_{u_j}(\Delta(\alpha_{kj}^*)|X_j))^2 \lambda' P_k (X_j) \right]^2
$$

$$
\leq C E_{F_n} \left[ \Delta^2(\alpha_{kj}^*) \lambda' P_k (X_j) \right]^2 \leq C \| \lambda \|^2 k_j^{-2r_j}
$$

(D.39)
which together with Assumption D.1(iv) implies that

$$\|H_{F_n,k_j} - H_{F_n,k_j}^*\| = O(k_j^{-r_j}) = o(1).$$  \hspace{2cm} (D.40)

Combining the results in (D.38) and (D.40), we get

$$-C \leq \rho_{\min}(H_{F_n,k_j}) \leq \rho_{\max}(H_{F_n,k_j}) \leq -C^{-1}. \hspace{2cm} (D.41)$$

Let $U_{k_j,n} = (u_{k_j,1}, \ldots, u_{k_j,n})'$ where $u_{k_j,i} = I\{Y \leq \alpha_{k_j}^*(X_j)\} - \tau$. Under Assumptions C.1 and D.1, we can invoke Lemma 4 in Belloni, Chernozhukov, Chetverikov, and Fernández-Val (2011) to deduce that

$$\lambda'_{k_j}(\hat{\beta}_{k_j} - \beta_{k_j}^*) = \lambda'_{k_j} H_{F_n,k_j}^{-1} P'_{k_j,n} U_{k_j,n} + O_p(\epsilon_{1,j,n} + \epsilon_{2,j,n} + \epsilon_{3,j,n}) \hspace{2cm} (D.42)$$

uniformly over $\lambda_{k_j} \in S_{k_j} = \{\lambda_j \in R^{k_j} : \|\lambda_j\| = 1\}$, where

$$\epsilon_{1,j,n} = \delta_{0,j,n}^{3/2} \xi_{k_j}^{1/4} + \delta_{0,j,n}^2 \xi_{k_j}^{1/2}, \epsilon_{2,j,n} = \delta_{0,j,n}^{3/2} \xi_{k_j}^{1/2} + \delta_{0,j,n} \xi_{k_j}^{-r_j} \text{ and } \epsilon_{3,j,n} = k_j \xi_{k_j}^{1/2} n^{-1}. \hspace{2cm} (D.43)$$

As $k_j n^{-1} = O(\delta_{0,j,n}^2)$ and $\delta_{0,j,n}^{3/2} \xi_{k_j}^{1/4} = o(1)$ by Assumption D.1(iv), we have

$$(k_j \xi_{k_j}^{1/2} n^{-1})(\delta_{0,j,n}^{3/2} \xi_{k_j}^{1/4})^{-1} = \delta_{0,j,n} \xi_{k_j}^{1/2} O(1) = o(1), \hspace{2cm} (D.44)$$

which together with (D.31), (D.42) and (D.43) implies that

$$\lambda'_{k_j}(\hat{\beta}_{k_j} - \beta_{k_j}^*) = n^{-1} \lambda'_{k_j} H_{F_n,k_j}^{-1} P'_{k_j,n} U_{k_j,n} + O_p \left(\delta_{0,j,n}^{3/2} \xi_{k_j}^{1/4} + \delta_{0,j,n} \xi_{k_j}^{-r_j}\right). \hspace{2cm} (D.45)$$

By the first order optimality condition of $\beta_{k_j}^*$,

$$E_{F_n} \left[ P_{k_j}(X_j) u_{k_j} \right] = 0_{k_j \times 1}. \hspace{2cm} (D.46)$$

By Assumptions C.1(i) and C.1(iv), (D.38), (D.40) and (D.41),

$$E_{F_n} \left[ \lambda'_{k_j} (H_{F_n,k_j}^* - H_{F_n,k_j}^{-1}) P'_{k_j,n} U_{k_j,n} \right]^2 = E_{F_n} \left[ \sum_{i=1}^{n} \lambda'_{k_j} (H_{F_n,k_j}^* - H_{F_n,k_j}^{-1}) P_{k_j}(X_{j,i}) u_{k_j,i} \right]^2$$

$$= \sum_{i=1}^{n} E_{F_n} \left[ \lambda'_{k_j} H_{F_n,k_j}^* (H_{F_n,k_j}^* - H_{F_n,k_j}^{-1}) H_{F_n,k_j}^{-1} P_{k_j}(X_{j,i}) u_{k_j,i} \right]^2$$

$$= \sum_{i=1}^{n} E_{F_n} \left[ \lambda'_{k_j} H_{F_n,k_j}^* (H_{F_n,k_j}^* - H_{F_n,k_j}^{-1}) H_{F_n,k_j}^{-1} u_{k_j,i}^2 \right]$$

23
\[ \leq Cn \left\| Y_{k_j}^* H_{F_n,k_j}^{-1} (H_{F_n,k_j}^* - H_{F_n,k_j}) H_{F_n,k_j}^{-1} \right\|^2 = O(nk_j^{-2r_j}) \]  \hspace{1cm} (D.47)

which together with the Markov inequality implies that

\[ n^{-1} \lambda_{k_j}^*(H_{F_n,k_j}^* - H_{F_n,k_j}) P_{k_j,n} U_{k_j,n} = O(n^{-1/2}k_j^{-r_j}). \]  \hspace{1cm} (D.48)

Combining the results in \(D.45\) and \(D.48\), and then applying Assumption \(D.1(iv)\), we deduce that

\[ \lambda_{k_j}^*(\hat{\beta}_{k_j} - \beta_{k_j}^*) = n^{-1} \lambda_{k_j}^* H_{F_n,k_j}^{-1} P_{k_j,n} U_{k_j,n} + O_p \left( \delta_{0,j,n}^{3/2} \xi_k^{1/4} + \delta_{0,j,n} k_j^{-r_j} \right). \]  \hspace{1cm} (D.49)

By Assumptions \(C.1(ii)\) and \(D.1(i)\),

\[ f_{u_j} (\Delta(\gamma_{k_j}^*)|X_j) (\gamma_{k_j} - \gamma_{k_j}^*)^2 - f_{u_j} (\Delta(\gamma_{k_j}^*)|X_j) (\gamma_{k_j} - \gamma_{k_j}^*)^2 \]
\[ = 2 f_{u_j} (\Delta(\gamma_{k_j}^*)|X_j) (\gamma_{k_j} - \gamma_{k_j}^*) \Delta(\gamma_{k_j}^*) + f_{u_j} (\Delta(\gamma_{k_j}^*)|X_j) \Delta(\gamma_{k_j}^*)^2 \]
\[ = 2 (\beta_{k_j} - \beta_{k_j}^* ) \left[ f_{u_j} (\Delta(\gamma_{k_j}^*)|X_j) P_{k_j}(X_j) \Delta(\gamma_{k_j}^*) \right] + O(k_j^{-2r_j}). \]  \hspace{1cm} (D.50)

Let \( \lambda_{*,j,n} = E_{F_n} \left[ f_{u_j} (\Delta(\gamma_{k_j}^*)|X_j) P_{k_j}(X_j) \Delta(\gamma_{k_j}^*) \right] \). Under Assumptions \(C.1(iv)\)

\[ \lambda_{*,j,n} \lambda_{*,j,n} \leq C \lambda_{*,j,n} \left( E_{F_n} \left[ P_{k_j}(X_j) P_{k_j}(X_j) \right] \right)^{-1} \lambda_{*,j,n}, \]  \hspace{1cm} (D.51)

which together with the same arguments in the proof of Lemma \(F.1\) Assumptions \(C.1(ii)\) and \(D.1(i)\) implies that

\[ \lambda_{*,j,n} \lambda_{*,j,n} \leq E_{F_n} \left[ f_{u_j}^2 (\Delta(\gamma_{k_j}^*)|X_j) \Delta(\gamma_{k_j}^*)^2 \right] = O(k_j^{-2r_j}). \]  \hspace{1cm} (D.52)

Using \(D.49\) and \(D.52\), we get

\[ (\hat{\beta}_{k_j} - \beta_{k_j}^*) \left[ E_{F_n} \left[ f_{u_j} (\Delta(\gamma_{k_j}^*)|X_j) P_{k_j}(X_j) \Delta(\gamma_{k_j}^*) \right] \right] \]
\[ = \frac{\lambda_{*,j,n} H_{F_n,k_j}^{-1} P_{k_j,n} U_{k_j,n} U_{k_j,n}}{n} + O_p \left( \delta_{0,j,n}^{3/2} \xi_k^{1/4} k_j^{-r_j} + \delta_{0,j,n} k_j^{-2r_j} \right). \]  \hspace{1cm} (D.53)

Under the i.i.d. assumption,

\[ E_{F_n} \left[ P_{k_j,n} U_{k_j,n} U_{k_j,n} P_{k_j,n} \right] = E_{F_n} \left[ \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} P_{k_j}(X_{j,i_1}) P_{k_j}(X_{j,i_2})' u_{k_j,i_1} u_{k_j,i_2} \right] \]
\[ = E_{F_n} \left[ \sum_{i=1}^{n} P_{k_j}(X_{j,i}) P_{k_j}(X_{j,i})' u_{k_j,i}^2 \right] \]  \hspace{1cm} (D.54)
which together with (D.41), Assumption C.1(iv) and (D.52) implies that
\[
E_{F_n} \left[ X_{s,j,n} H_{F_n,k_j}^{-1} P_{k_j,n} U_{k_j,n} F_{k_j,n} H_{F_n,k_j}^{-1} \lambda_{s,j,n} \right] \leq C \frac{\lambda_{s,j,n}^2}{n^2} = O(k_j^{-2r_j} n^{-1}), \tag{D.55}
\]
which together with the Markov inequality and (D.53) implies that
\[
(\hat{\beta}_{k_j,n} - \beta_{k_j,n}^*) \lambda_{s,j,n} = O_p \left( \delta_{0,j,n}^{3/2} k_j^{-r_j} + \delta_{0,j,n} k_j^{-2r_j} + n^{-1/2} k_j^{-r_j} \right). \tag{D.56}
\]
Combining the results in (D.50) and (D.56), we get
\[
\left| f_{u_j} (\Delta(\alpha_{k_j}) | X_j) \Delta(\hat{\alpha}_{k_j} - \alpha_{k_j})^2 - f_{u_j} (\Delta(\alpha_{k_j}^*) | X_j) (\hat{\alpha}_{k_j} - \alpha_{k_j}^*)^2 \right|
\leq O_p \left( \delta_{0,j,n}^{3/2} k_j^{-r_j} + k_j^{-2r_j} + n^{-1/2} k_j^{-r_j} \right), \tag{D.57}
\]
where we note that \( \delta_{0,j,n} k_j^{-2r_j} = o(k_j^{-2r_j}) \) as \( \delta_{0,j,n} = o(1) \). Let \( E_{X_j,F_n} [\cdot] \) denote the expectation taken with respect to \( X_j \). Note that using (D.56), we also have
\[
E_{X_j,F_n} \left[ f_{u_j} (\Delta(\alpha_{k_j}) | X_j) \Delta(\hat{\alpha}_{k_j} - \alpha_{k_j})^2 - f_{u_j} (\Delta(\alpha_{k_j}^*) | X_j) (\hat{\alpha}_{k_j} - \alpha_{k_j}^*)^2 \right]
\leq E_{X_j,F_n} \left[ f_{u_j} (\Delta(\alpha_{k_j}) | X_j) \Delta(\hat{\alpha}_{k_j} - \alpha_{k_j}^*) \right]
\leq (\hat{\beta}_{k_j} - \beta_{k_j}^*) \lambda_{s,j,n} = O_p \left( \delta_{0,j,n}^{3/2} k_j^{-r_j} + \delta_{0,j,n} k_j^{-2r_j} + n^{-1/2} k_j^{-r_j} \right). \tag{D.58}
\]
Using integration by parts, we can write
\[
\int_0^{\Delta(\alpha_j)} u f_{u_j} (u | x_j) du = \frac{1}{2} \int_0^{\Delta(\alpha_j)} f_{u_j} (u | x_j) du^2 - \frac{f_{u_j} (\Delta(\alpha_j) | x_j) \Delta^2(\alpha_j)}{2} - \int_0^{\Delta(\alpha_j)} u^2 \frac{\partial f_{u_j} (u | x_j)}{\partial u} du, \tag{D.59}
\]
which together with the triangle inequality and Assumption D.1(ii) implies that
\[
\left| E_{F_n} \left[ (I\{u_j \leq \Delta(\alpha_j)\} - I\{u_j \leq 0\} ) u_j | X_j \right] - \frac{f_{u_j} (\Delta(\alpha_{k_j}) | X_j) \Delta^2(\alpha_j)}{2} \right|
\leq \frac{f_{u_j} (\Delta(\alpha_j) | x_j) \Delta^2(\alpha_j)}{2} + \int_0^{\Delta(\alpha_j)} u^2 \frac{\partial f_{u_j} (u | x_j)}{\partial u} du_j \nabla \leq C \left( |\Delta(\alpha_j)|^2 |\alpha_j - \alpha_{k_j}^*| + |\Delta(\alpha_j)|^3 \right) \tag{D.60}
\]
uniformly over \( \beta_{k_j} \in \mathcal{N}_{j,n} \). Using (D.12), (D.13), (D.57), (D.60) and the triangle inequality, we
Using (D.12), (D.13), (D.58), (D.63) and the triangle inequality, we immediately get

$$\left| E_{Z,F_n} \left[ (I\{u_j \leq \Delta(\hat{\alpha}_{k_j})\} - I\{u_j \leq 0\}) u_j - \frac{f_{u_j}(\Delta(\alpha^*_k)|X_j)(\hat{\alpha}_{k_j} - \alpha^*_k)^2}{2} \right] \right|$$

$$\leq CE_{X_j,F_n} \left[ |\Delta(\hat{\alpha}_{k_j})|^2 |\hat{\alpha}_{k_j} - \alpha^*_k| + |\Delta(\hat{\alpha}_{k_j})|^3 \right] + O_p \left( \delta^{3/2}_{0,j,n} k_j^{-r_j} + k_j^{-2r_j} + n^{-1/2}k_j^{-r_j} \right)$$

$$= O_p(\delta^2_{1,j,n} \delta_{2,j,n} + \delta^{3/2}_{0,j,n} k_j^{-r_j} + k_j^{-2r_j} + n^{-1/2}k_j^{-r_j}) \quad (D.61)$$

Similarly, we have

$$\int_0^{\Delta(\alpha_j)} f_{u_j}(u|X_j)du = \Delta(\alpha_j)f_{u_j}(\Delta(\alpha_j)|X_j) - \int_0^{\Delta(\alpha_j)} \frac{\partial f_{u_j}(u|x_j)}{\partial u} du \quad (D.62)$$

which together with the triangle inequality and Assumption D.1(ii) implies that for any \(\alpha_j\),

$$\left| E_{F_n} \left[ (I\{u_j \leq \Delta(\alpha_j)\} - I\{u_j \leq 0\}) |X_j| - f_{u_j}(\Delta(\alpha^*_k)|X_j)\Delta(\alpha_j) \right] \right|$$

$$\leq \left| f_{u_j}(\Delta(\alpha^*_k)|X_j) - f_{u_j}(0|X_j) \right| |\Delta(\alpha_j)| + C |\Delta(\alpha_j)|^2$$

$$\leq C \left( |\Delta(\alpha_j)|^2 + |\Delta(\alpha_j)||\Delta(\alpha^*_k)| \right). \quad (D.63)$$

Using (D.12), (D.13), (D.58), (D.63) and the triangle inequality, we immediately get

$$\left| E_{Z,F_n} \left[ (I\{u_j \leq \Delta(\hat{\alpha}_{k_j})\} - I\{u_j \leq 0\}) (\hat{\alpha}_{k_j} - \alpha^*_k) - f_{u_j}(\Delta(\alpha^*_k)|X_j)(\hat{\alpha}_{k_j} - \alpha^*_k)^2 \right] \right|$$

$$\leq \left| E_{Z,F_n} \left[ (I\{u_j \leq \Delta(\hat{\alpha}_{k_j})\} - I\{u_j \leq 0\}) - f_{u_j}(\Delta(\alpha^*_k)|X_j)\Delta(\hat{\alpha}_{k_j}) \right] \right| (\hat{\alpha}_{k_j} - \alpha^*_k)$$

$$+ \left| E_{X_j,F_n} \left[ f_{u_j}(\Delta(\alpha^*_k)|X_j)\Delta(\hat{\alpha}_{k_j}) - f_{u_j}(\Delta(\alpha^*_k)|X_j)(\hat{\alpha}_{k_j} - \alpha^*_k) - f_{u_j}(\Delta(\alpha^*_k)|X_j)(\hat{\alpha}_{k_j} - \alpha^*_k)^2 \right] \right|$$

$$\leq CE_{X_j,F_n} \left[ |\Delta(\hat{\alpha}_{k_j})|^2 + |\Delta(\hat{\alpha}_{k_j})||\Delta(\alpha^*_k)| \right] + O_p \left( \delta^{3/2}_{0,j,n} k_j^{-r_j} + \delta_{0,j,n} k_j^{-2r_j} + n^{-1/2}k_j^{-r_j} \right)$$

$$= O_p(\delta^2_{1,j,n} \delta_{2,j,n} + \delta^{3/2}_{0,j,n} k_j^{-r_j} + \delta_{0,j,n} k_j^{-2r_j} + n^{-1/2}k_j^{-r_j}). \quad (D.64)$$

Combining the results in (D.61) and (D.64) and applying the triangle inequality, we immediately get the claimed result. ■

**Lemma D.5** Under Assumptions C.1 and D.1 we have

$$\hat{\beta}_{k_j} - \beta^*_{k_j} = \frac{U'_{k_j,n} P'_{k_j,n} H_{F_{n,k_j}^{-1}} P_{k_j,n} U_{k_j,n}}{n^2} + O_p(\delta^{5/2}\xi^{1/4} + \delta_{0,j,n} k_j^{-r_j}). \quad (D.65)$$
Proof of Lemma D.5 From (D.11) and (D.49), we get

\[
(\hat{\beta}_{k_j} - \beta_{k_j}^*)' H_{F_n,k_j} (\hat{\beta}_{k_j} - \beta_{k_j}^*) = \frac{(\beta_{k_j} - \beta_{k_j}^*)' P_{k_j,n} U_{k_j,n}}{n} + O_p \left( \delta_{0,j,n}^{5/2} \xi_{k_j}^{1/4} + \delta_{0,j,n}^2 k_j^{-r_j} \right). \tag{D.66}
\]

Under the i.i.d. assumption, (D.46) and Assumptions C.1(iv), we have

\[
E_{F_n} \left[ \frac{\mathbf{U}'_{k_j,n} P_{k_j,n} P_{k_j,n} U_{k_j,n}}{n^2} \right] = \left[ n^{-2} \sum_{i=1}^n P_{k_j}(X_j) u_{k_j} \right] = \left[ n^{-2} \sum_{i=1}^{k_j} \left[ \sum_{i=1}^n p_{l,j}(X_j) u_{k_j} \right] \right] \leq n^{-2} \sum_{i=1}^{k_j} \sum_{i=1}^n E_{F_n} \left[ p_{l,j}(X_j)^2 \right] = O(k_j n^{-1}) \tag{D.67}
\]

which together with (D.45) and (D.66) immediately implies (D.65). \(\blacksquare\)

Lemma D.6 Under Assumptions C.1 and D.1, we have

\[
n^{-1} \sum_{i=1}^n u_{k_j,i} \Delta_i(\hat{\alpha}_{k_j}) = \frac{\mathbf{U}'_{k_j,n} P_{k_j,n} H_{F_n,k_j}^{-1} P_{k_j,n} U_{k_j,n}}{n^2} + O_p \left( \delta_{0,j,n}^{5/2} \xi_{k_j}^{1/4} + \delta_{0,j,n}^2 k_j^{-r_j} + k_j^{-1} + k_j^{-2r_j} \right). \tag{D.68}
\]

Proof of Lemma D.6 First, we note that

\[
n^{-1} \sum_{i=1}^n u_{k_j,i} \Delta_i(\hat{\alpha}_{k_j}) = n^{-1} \sum_{i=1}^n u_{k_j,i} (\hat{\alpha}_{k_j} - \alpha_{k_j}^*) + n^{-1} \sum_{i=1}^n u_{j,i}^* \Delta_i(\alpha_{k_j}^*) + n^{-1} \sum_{i=1}^n (u_{k_j,i} - u_{j,i}^*) \Delta_i(\alpha_{k_j}^*) \tag{D.69}
\]

where \(u_{j,i}^* = I\{u_{j,i} \leq 0\} - \tau\). We can write

\[
n^{-1} \sum_{i=1}^n u_{k_j,i} (\hat{\alpha}_{k_j} - \alpha_{k_j}^*) = \frac{\mathbf{U}'_{k_j,n} P_{k_j,n} (\hat{\beta}_{k_j} - \beta_{k_j}^*)}{n} \tag{D.70}
\]

which together with (D.49) and (D.67) implies that

\[
n^{-1} \sum_{i=1}^n u_{k_j,i} (\hat{\alpha}_{k_j} - \alpha_{k_j}^*) = \frac{\mathbf{U}'_{k_j,n} P_{k_j,n} H_{F_n,k_j}^{-1} P_{k_j,n} U_{k_j,n}}{n^2} + O_p \left( \delta_{0,j,n}^{5/2} \xi_{k_j}^{1/4} + \delta_{0,j,n}^2 k_j^{-r_j} \right). \tag{D.71}
\]
By the definition of $u_{j,i}$

$$E_{F_n} \left[ u_j^* | X_j \right] = 0.$$  \hfill (D.72)

By the i.i.d. assumption, \[ \text{(D.72)} \] and Assumption \[ \text{C.1(ii)} \], we get

$$E_{F_n} \left[ \left| \sum_{i=1}^{n} u_{j,i}^* (\alpha_{k_j}^* - \alpha_j^*) \right|^2 \right] = n^{-1} E_{F_n} \left[ (u_{j,i}^*)^2 (\alpha_{k_j}^* - \alpha_j^*)^2 \right] \leq C k_{j}^{-2r_j} n^{-1} \hfill (D.73)$$

which together with the Markov inequality implies

$$n^{-1} \sum_{i=1}^{n} u_{j,i}^* (\alpha_{k_j}^* - \alpha_j^*) = O_p \left( k_{j}^{-2r_j} n^{-\frac{1}{2}} \right). \hfill (D.74)$$

By the i.i.d. assumption, the triangle inequality and Assumption \[ \text{C.1(ii)} \],

$$E_{F_n} \left[ \left| \sum_{i=1}^{n} (u_{k_j,i} - u_{j,i}^*) \Delta_i (\alpha_{k_j}^*) \right| \right] \leq E_{F_n} \left[ \left| I\{u_j \leq 0\} - I\{u_j \leq \Delta(\alpha_{k_j}^*)\}\right| \Delta(\alpha_{k_j}^*) \right] \leq \Delta(\alpha_{k_j}^*) \int_0^{\Delta(\alpha_{k_j}^*)} f_{u_j}(u | X_j) du \leq C E_{F_n} \left[ \Delta(\alpha_{k_j}^*)^2 \right] = C k_{j}^{-2r_j} \hfill (D.75)$$

which together with the Markov inequality implies

$$n^{-1} \sum_{i=1}^{n} (u_{k_j,i} - u_{j,i}^*) \Delta_i (\alpha_{k_j}^*) = O_p \left( k_{j}^{-2r_j} \right). \hfill (D.76)$$

Collecting the results in \[ \text{(D.69)}, \text{(D.71)}, \text{(D.74)}, \text{and (D.76)} \], we immediately get \[ \text{(D.68)} \]. \hfill $\blacksquare$

**Lemma D.7** Under Assumptions \[ \text{C.1} \] and \[ \text{D.1} \] we have

$$E_{u_j,F_n} \left[ \left| I\{u_j \leq \Delta(\alpha_{k_j})\} - I\{u_j \leq 0\}\right| \Delta(\alpha_{k_j}) \right] = O_p \left( \delta_{0,j,n} k_{j}^{-r_j} + \delta_{0,j,n} k_{j}^{-2r_j} + n^{-1/2} k_{j}^{-r_j} + \delta_{1,j,n} k_{j}^{-r_j} \right). \hfill (D.77)$$

**Proof of Lemma D.7** From \[ \text{(D.63)} \], we see that

$$E_{F_n} \left[ I\{u_j \leq \Delta(\alpha_j)\} - I\{u_j \leq 0\} | X_j \right] - f_{u_j}(\Delta(\alpha_{k_j}^*) | X_j) \Delta(\alpha_j) \leq C \left( |\Delta^2(\alpha_j)| + |\Delta(\alpha_j)| |\Delta(\alpha_{k_j})| \right) \hfill (D.78)$$

28
for any $\beta_{kj} \in \mathcal{N}_{j,n}$, which combined with the triangle inequality implies that

$$
\sup_{\beta_{kj} \in \mathcal{N}_{j,n}} \left| E_{F_n} \left[ I\{ u_j \leq \Delta(\alpha_j) \} - I\{ u_j \leq 0 \} - f_{u_j}(\Delta(\alpha^*_k)|X_j) \Delta(\alpha_j) \right] \Delta(\alpha^*_k) \right|
\leq \sup_{\beta_{kj} \in \mathcal{N}_{j,n}} \left| E_{F_n} \left[ |\Delta^2(\alpha_j)| + |\Delta(\alpha_j)| |\Delta(\alpha^*_k)||\Delta(\alpha_j)\right] \right| = O(\delta_{1,j,n}^2 k_j^{-r_j}). \tag{D.79}
$$

Invoking (D.56), we have

$$
E_{X_j,F_n} \left[ f_{u_j}(\Delta(\alpha^*_k)|X_j) \Delta(\tilde{\alpha}_k) \Delta(\alpha^*_k) \right] = (\tilde{\beta}_{kj} - \beta^*_k)^T E \left[ f_{u_j}(\Delta(\alpha^*_k)|X_j) P_{kj}(X_j)(\alpha^*_k - \alpha^*_j) \right] = O_p \left( \tilde{\beta}_{kj} - \beta^*_k \right)^T E \left[ f_{u_j}(\Delta(\alpha^*_k)|X_j) P_{kj}(X_j)(\alpha^*_k - \alpha^*_j) \right] = O_p \left( \tilde{\beta}_{kj} - \beta^*_k \right)^T E \left[ f_{u_j}(\Delta(\alpha^*_k)|X_j) P_{kj}(X_j)(\alpha^*_k - \alpha^*_j) \right]. \tag{D.80}
$$

Collecting the results in (D.79) and (D.80), we immediately get (D.77). ■

**Lemma D.8** Under Assumptions C.1 and D.1 we have

$$
n^{-1} \sum_{i=1}^{n} \left[ \rho_r(Y_i - \tilde{\alpha}_{k_1}(X_{1,i})) - \rho_r(Y_i - \tilde{\alpha}_{k_2}(X_{2,i})) \right] = \frac{1}{n} \sum_{i=1}^{n} \left[ \rho_r(u_{1,i}) - \rho_r(u_{2,i}) \right] = \frac{\ell_{\alpha,n}(\alpha^*_k) H_{F_n,k}^{-1} \ell_{\alpha,n}(\alpha^*_k)}{2} + O_p \left( \delta_{n,k} \right) \tag{D.81}
$$

where $\delta_{n,k} = \max_{j=1,2} \left\{ \delta_{0,j,n}^2, \delta_{2,j,n}^2 \right\}$.

**Proof of Lemma D.8** Using (D.15), we can write

$$
n^{-1} \sum_{i=1}^{n} \left[ \rho_r(u_{j,i}) - \rho_r(Y - \tilde{\alpha}_{k_j}(X_{j,i})) \right] = n^{-1} \sum_{i=1}^{n} \left( I\{ u_{j,i} \leq 0 \} - \tau \right) \Delta_i(\tilde{\alpha}_{k_j}) + n^{-1} \sum_{i=1}^{n} \left( I\{ u_{j,i} \leq 0 \} - I\{ u_{j,i} \leq \Delta_i(\tilde{\alpha}_{k_j}) \} \right) (u_{j,i} - \Delta_i(\tilde{\alpha}_{k_j})). \tag{D.82}
$$

Using (D.16) in Lemma D.1 we have

$$
n^{-1} \sum_{i=1}^{n} \left( I\{ u_{j,i} \leq 0 \} - I\{ u_{j,i} \leq \Delta_i(\tilde{\alpha}_{k_j}) \} \right) u_{j,i} = E_{Z,F_n} \left[ (I\{ u_{j} \leq 0 \} - I\{ u_{j} \leq \Delta(\tilde{\alpha}_{k_j}) \}) u_{j} \right] + O_p(\delta_{0,j,n}^2 \delta_{1,j,n}^2 \delta_{2,j,n}^2). \tag{D.83}
$$
Using (D.24) in Lemma D.2 and (D.77) in Lemma D.7 we have

\[
\begin{align*}
& n^{-1} \sum_{i=1}^{n} \left( I\{u_{j,i} \leq 0\} - I\{u_{j,i} \leq \Delta_i(\hat{\alpha}_{k_j})\} \right) \Delta_i(\hat{\alpha}_{k_j}) \\
& = n^{-1} \sum_{i=1}^{n} \left( I\{u_{j,i} \leq 0\} - I\{u_{j,i} \leq \Delta_i(\hat{\alpha}_{k_j})\} \right) (\hat{\alpha}_{k_j}(X_{j,i}) - \alpha^*_k(X_{j,i})) \\
& + n^{-1} \sum_{i=1}^{n} \left( I\{u_{j,i} \leq 0\} - I\{u_{j,i} \leq \Delta_i(\hat{\alpha}_{k_j})\} \right) \Delta_i(\alpha^*_k) \\
& = E_{Z,F_n} \left[ (I\{u_j \leq 0\} - I\{u_j \leq \Delta(\hat{\alpha}_{k_j})\}) (\hat{\alpha}_{k_j}(X_j) - \alpha^*_k(X_j)) \right] \\
& + (\hat{\beta}_{k_j} - \beta^*_k)^\prime \mu_n \left[ (I\{u_j \leq 0\} - I\{u_j \leq \Delta(\hat{\alpha}_{k_j})\}) P_{k_j}(X_j) \right] \\
& + O_p(\delta_{0,j,n}\delta_{1,j,n}^{1/2}k_j^{-r_j} + \delta_{0,j,n}k_j^{-2r_j} + n^{-1/2}k_j^{-r_j}), \\
& \quad \text{(D.84)}
\end{align*}
\]

where in the last equality, we use

\[
\frac{\delta_{0,j,n}^{2}k_j^{-r_j}}{\delta_{0,j,n}^{1/2}k_j^{-r_j}} = o(1) \quad \text{and} \quad \frac{\delta_{0,j,n}^{3/2}k_j^{-r_j}}{\delta_{0,j,n}^{1/2}k_j^{-r_j}} = O(1). \quad \text{(D.85)}
\]

Using (D.11), and (D.29) in Lemma D.3 we get

\[
(\hat{\beta}_{k_j} - \beta^*_k)^\prime \mu_n \left[ (I\{u_j \leq 0\} - I\{u_j \leq \Delta(\hat{\alpha}_{k_j})\}) P_{k_j}(X_j) \right] = O_p(\delta_{0,j,n}^{2}\delta_{2,j,n}^{1/2}), \quad \text{(D.86)}
\]

which together with (D.84) implies that

\[
\begin{align*}
& n^{-1} \sum_{i=1}^{n} \left( I\{u_{j,i} \leq 0\} - I\{u_{j,i} \leq \Delta_i(\hat{\alpha}_{k_j})\} \right) \Delta_i(\hat{\alpha}_{k_j}) \\
& = E_{Z,F_n} \left[ (I\{u_j \leq 0\} - I\{u_j \leq \Delta(\hat{\alpha}_{k_j,n})\})(\hat{\alpha}_{k_j,n} - \alpha^*_k) \right] \\
& + O_p(\delta_{0,j,n}^{2}\delta_{2,j,n}^{1/2} + \delta_{0,j,n}^{1/2}k_j^{-r_j} + \delta_{0,j,n}k_j^{-2r_j} + n^{-1/2}k_j^{-r_j}). \\
& \quad \text{(D.87)}
\end{align*}
\]

Combining the results in (D.83) and (D.87), we get

\[
\begin{align*}
& n^{-1} \sum_{i=1}^{n} \left( I\{u_{j,i} \leq 0\} - I\{u_{j,i} \leq \Delta_i(\hat{\alpha}_{k_j})\} \right) (u_{j,i} - \Delta_i(\hat{\alpha}_{k_j})) \\
& = E_{Z,F_n} \left[ (I\{u_j \leq 0\} - I\{u_j \leq \Delta(\hat{\alpha}_{k_j})\})(u_j - (\hat{\alpha}_{k_j}(X_j) - \alpha^*_k(X_j))) \right] \\
& + O_p(\delta_{0,j,n}^{2}\delta_{2,j,n}^{1/2} + \delta_{0,j,n}k_j^{-2r_j} + n^{-1/2}k_j^{-r_j}), \\
& \quad \text{(D.88)}
\end{align*}
\]
where we use
\[ \delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n}^{3/2} = \delta_{0,j,n}^{2} \delta_{2,j,n}^{1/2} + \delta_{0,j,n} \delta_{2,j,n}^{1/2} \delta_{j,n}^{r_j} \] and \[ \delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n}^{r_j} = \delta_{1,j,n}^{3/2} < 1. \] (D.89)

Using (D.37) in Lemma D.4 and (D.88), we get
\[ n^{-1} \sum_{i=1}^{n} [I\{u_{j,i} \leq 0\} - I\{u_{j,i} \leq \Delta_i(\hat{\alpha}_{k_j})\}] \left[ u_{j,i} - \Delta_i(\hat{\alpha}_{k_j}) \right] \]
\[ = - \frac{(\beta_{k_j} - \beta_{k_j}^\ast) H_{F_{n,k_j}}(\beta_{k_j} - \beta_{k_j}^\ast)}{2} + O_p(\delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n}^{1/2} + k_j^{-2r_j} + n^{-1/2} k_j^{-r_j}), \] (D.90)
where we use \( \delta_{0,j,n} \delta_{2,j,n}^2 \leq 2 \delta_{0,j,n} \delta_{2,j,n}^2 + 2 \delta_{0,j,n} \delta_{2,j,n}^{1/2} k_j^{-r_j} = O(\delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n}^{1/2}) + O(k_j^{-2r_j}) \) and
\[ \frac{\delta_{0,j,n} \delta_{2,j,n}^{1/2} k_j^{-r_j}}{\delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n}^{1/2}} = \frac{\delta_{0,j,n} \delta_{2,j,n}^{1/2} k_j^{-r_j}}{\delta_{0,j,n} \delta_{2,j,n}^{1/2}} < 1. \] (D.91)

Using (D.65) in Lemma D.5 and (D.90), we get
\[ n^{-1} \sum_{i=1}^{n} [I\{u_{j,i} \leq 0\} - I\{u_{j,i} \leq \Delta_i(\hat{\alpha}_{k_j})\}] \left[ u_{j,i} - \Delta_i(\hat{\alpha}_{k_j}) \right] \]
\[ = - \frac{U'_{k_j,n} P'_{k_j,n} H_{F_{n,k_j}}^{-1} P_{k_j,n} U_{k_j,n}}{2n^2} + O_p(\delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n}^{1/2} + k_j^{-2r_j} + n^{-1/2} k_j^{-r_j}), \] (D.92)
where we use
\[ \delta_{0,j,n} \delta_{2,j,n}^{1/4} < \delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n}^{1/2} \] and \( \delta_{0,j,n} k_j^{-r_j} < \delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n}^{1/2} \). (D.93)

Combining the results in (D.68) of Lemma D.6, (D.82) and (D.92), we get
\[ n^{-1} \sum_{i=1}^{n} [\rho_\tau(u_{j,i}) - \rho_\tau(Y - \hat{\alpha}_{k_j}(X_{j,i}))] \]
\[ = \frac{U'_{k_j,n} P'_{k_j,n} H_{F_{n,k_j}}^{-1} P_{k_j,n} U_{k_j,n}}{2n^2} + O_p(\delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n}^{1/2} + k_j^{-2r_j} + n^{-1/2} k_j^{-r_j}). \] (D.94)

Using (D.94), and the definitions of \( U_{k_j,n} \), \( H_{F_{n,k_j}} \) and \( \ell_{\alpha,k}(Z; \alpha) \), we immediately get the result in (D.81).  

**Proof of Theorem D.1.** We can use Theorem 4.1 to prove the claim. For this purpose, we need to verify Assumptions 4.1 and 4.3. Assumptions 4.1(a) is implied by Assumption C.1(i).
Using (D.14), we can write

\[ m(Z; \alpha_j) = (I\{u_j \leq \Delta(\alpha_j)\} - \tau) [u_j - \Delta(\alpha_j)] \] (D.95)

which implies that

\[
E_{F_0}[m(Z; \alpha_j)|X_j] = E_{F_0}[I\{u_j \leq \Delta(\alpha_j)\}|X_j] - \tau E_{F_0}[u_j - \Delta(\alpha_j)|X_j]
= \int_{-\infty}^{\Delta(\alpha_j)} u f_{u_j}(u|X_j) du - \Delta(\alpha_j) [F_{u_j}(\Delta(\alpha_j)|X_j) - \tau] - \tau E_{F_0}[u_j|X_j]
\] (D.96)

where \( F_{u_j}(\cdot | X_j) \) denotes the conditional CDF of \( u_j \) given \( X_j \). Define

\[
g_j(X_j, \alpha_j) = \int_{-\infty}^{\Delta(\alpha_j)} u f_{u_j}(u|X_j) du - \Delta(\alpha_j) [F_{u_j}(\Delta(\alpha_j)|X_j) - \tau] \]
(D.97)

Since \( \alpha_j(\cdot) = P_{k_j}(\cdot)' \beta_{k_j} \), it is clear that \( g_j(x_j, \alpha_j) \) is continuously differentiable at \( \beta_{k_j} \) for any \( x_j \) with

\[
\frac{\partial g_j(x_j, P_{k_j}(x_j)' \beta_{k_j})}{\partial \beta_{k_j}} = \left[ \tau - F_{u_j}(\Delta(\alpha_j)|x_j) \right] P_{k_j}(x_j)
\] (D.98)

and uniformly over \( x_j \) and \( \beta_{k_j} \),

\[
\left\| \frac{\partial g_j(x_j, P_{k_j}(x_j)' \beta_{k_j})}{\partial \beta_{k_j}} \right\| \leq \zeta_{k_j}^{1/2} < \infty.
\] (D.99)

We can use the dominated convergence theorem to show that the

\[
\frac{\partial E_{F_0}[m(Z; \alpha_j)]}{\partial \beta_{k_j}} = E_{F_0}\left[ \frac{\partial g_j(X_j, \alpha_j)}{\partial \beta_{k_j}} \right].
\] (D.100)

Moreover, \( \partial g_j(x_j, \alpha_j)/\partial \beta_{k_j} \) is continuously differentiable at \( \beta_{k_j} \) for any \( x_j \) with

\[
\frac{\partial^2 g_j(x_j, \alpha_j)}{\partial \beta_{k_j} \partial \beta_{k_j}'} = -f_{u_j}(\Delta(\alpha_j)|X_j) P_{k_j}(X_j) P_{k_j}(X_j)'.
\]

By Assumptions [C.1(iv) and D.1(i)], we have uniformly over \( x_j \) and \( \beta_{k_j} \),

\[
\left\| \frac{\partial^2 g_j(x_j, P_{k_j}(x_j)' \beta_{k_j})}{\partial \beta_{k_j} \partial \beta_{k_j}'} \right\| \leq C \zeta_{k_j}^2 < \infty.
\] (D.101)
Hence we can use the dominated convergence theorem again to show that the
\[
\frac{\partial^2 E_{F_0}[m(Z; \alpha_j)]}{\partial \beta_{kj} \partial \beta_{kj}'} = E_{F_0} \left[ \frac{\partial g_j(X_j, \alpha_j)}{\partial \beta_{kj} \partial \beta_{kj}'} \right]
\]
(D.102)
is well defined. This verifies Assumption 4.1(b). By (D.98) and (D.100), and the first order condition of $\beta_{kj}$,
\[
E_{F_0} \left[ \frac{\partial g_j(X_j, \alpha^*_j)}{\partial \beta_{kj}} \right] = E_{F_0} \left[ (\tau - I\{u_j \leq \Delta(\alpha^*_j)\}) P_{kj}(x_j) \right] = 0_{kj \times 1}.
\]
(D.103)

By the definitions of $\ell_{\alpha, k}(Z; \alpha^*_k)$ and (D.103),
\[
E_{F_0} [\ell_{\alpha, k}(Z; \alpha^*_k)] = E_{F_0} \left[ \left( \left( I\{u_1 \leq \Delta(\alpha^*_k)\} - \tau \right) P_{k1}(X_1) \right) \left( \left( I\{u_2 \leq \Delta(\alpha^*_k)\} \right) P_{k2}(X_2) \right) \right] = 0_{|k|}
\]
(D.104)
which verifies Assumption 4.1(c). By Assumption D.1(vi),
\[
E_{F_0} [\ell(Z; \alpha^*_0)^2] = E_{F_0} \left[ (\rho_\tau(u_1) - \rho_\tau(u_2))^2 \right] \leq 2E_{F_0} \left[ u_1^2 \right] + 2E_{F_0} \left[ u_2^2 \right] \leq C.
\]
(D.105)

Moreover, by Assumptions C.1(ii) and C.1(iii),
\[
E_{F_0} [\|\ell_{\alpha, k}(Z; \alpha^*_k)\|^4] \leq 2E_{F_0} \left[ |P'_{k1}(X_1) P_{k1}(X_1)|^2 + |P'_{k2}(X_2) P_{k2}(X_2)|^2 \right]
\leq C \left[ \xi_{k1} tr \left( E_{F_0} \left[ P_{k1}(X_1)' P_{k1}(X_1) \right] \right) + \xi_{k2} tr \left( E_{F_0} \left[ P_{k2}(X_2)' P_{k2}(X_2) \right] \right) \right]
\leq C(\xi_{k1} + \xi_{k2})(k_1 + k_2) = C\xi_k |k|,
\]
(D.106)
where $\xi_k = \xi_{k1} + \xi_{k2}$. Hence Assumption 4.1(d) is also satisfied. Assumptions 4.1(e) is Assumption D.1(v). Also the bounds on the eigenvalues of $H_{F_0,k}$ in Assumption 4.1(f) can be easily verified using Assumptions C.1(iv), D.1(i) and D.1(iii). By definition,
\[
D_{F_0,k} = E_{F_0} \left[ \begin{array}{cc}
-u^2_{k1} P_{k1}(X_1) P_{k1}(X_1)' & -u_{k1} u_{k2} P_{k1}(X_1) P_{k2}(X_2)'
-u_{k1} u_{k2} P_{k2}(X_2) P_{k1}(X_1)' & u^2_{k2} P_{k2}(X_2) P_{k2}(X_2)'
\end{array} \right].
\]
(D.107)
Under Assumption C.1(iv), we see that
\[
\rho_{\max} \left( E_{F_0} \left[ u_{kj}^2 P_{kj}(X_j) P_{kj}(X_j)' \right] \right) \leq C\rho_{\max} \left( E_{F_0} \left[ P_{kj}(X_j) P_{kj}(X_j)' \right] \right) \leq C \text{ for } j = 1, 2
\]
(D.108)
which together with the form of $D_{F_0, k}$ in (D.107) and the Aronszajn’s Inequality implies that
\[ \rho_{\max}(D_{F_0, k}) \leq 2C. \]
This verifies Assumption 4.1(f).

In Lemma [D.8], we have derived the second order expansion of $\bar{\ell}_n(\hat{\alpha}_k)$, where the remainder term is of the order
\[ \delta_{n,k} = \max_{j=1,2} \left\{ (D) F_0, k \right\} + k^{-2r_j} + n^{-1/2}k^{-r_j} \].
\[ (D.109) \]
By (D.9), we have
\[ n^{1/2} \delta_{n,k} = \frac{n^{1/2} \left( \delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n} + k_j^{-2r_j} + n^{-1/2}k_j^{-r_j} \right)}{\sigma_{F_n,n}} = \frac{n^{1/2} \delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n} + k_j^{-r_j} n^{-1/2} + \frac{1}{k_j} n^{1/2} \sigma_{F_n,n}} = o(1) \]
\[ (D.110) \]
for $j = 1, 2$. As $\xi_k k_j \log(n)n^{-1} = o(1)$ by assumption Assumption D.1(iv), we have
\[ \frac{k_j^{1/2} \xi_k^{1/2}}{n^{1/2} \delta_{0,j,n} \delta_{1,j,n} \delta_{2,j,n} \sigma_{F_n,n}} \leq \frac{\xi_k^{1/4}}{n^{1/4}k_j^{3/4} (\log(n))^{5/4}} = o(1) \]
\[ (D.111) \]
which together with (D.9) implies that $k_j \xi_k (n^2 \sigma_{F_n,n}^2)^{-1} = o(1)$. Assumption 4.3(b) is implied by (D.9). So Assumption 4.3 is also satisfied. \[ \]

E Proofs of the Lemmas in Section 6 of the main paper

Proof of Lemma 6.1. Let $\kappa = (\kappa_1, \kappa_2, \kappa_3)' \in R^3$ be any vector with $\kappa' \kappa = 1$. For ease of notations, let $W_{i,n} = (W_1, W_2, W_3)'$ where
\[ W_1 = V_1 - U_1, \quad W_2 = \psi_{\alpha,k_j}(\alpha_{k_j}^*)'H_{F_n,k_j}^{-1} \ell_{\alpha,k_j}(Z_i; \alpha_{k_j}^*) \] for $j = 1, 2$,
\[ W_{i,n} = \sum_{i=1}^n \kappa' W_{i,n} + o_p(1). \]
\[ (E.2) \]
Let $Z_0 = 0$ and $\mathcal{F}_{i,n}$ be the natural filtration generated by $\{Z_0, Z_1, \ldots, Z_i\}$ under $F_n$ for $i = 0, \ldots, n$. Under Assumptions 4.1(b) and 4.1(d),

$$E_{F_n}[\kappa'W_{i,n} | \mathcal{F}_{i-1,n}] = 0,$$

for all $i = 1, \ldots, n$, which implies that $\{\kappa'W_{i,n}\}_{i \leq n}$ is a martingale difference array. We shall use the martingale central limit theorem (ref. Corollary 3.1 in Hall and Heyde (1980)) to show the desired convergence. It suffices to verify the following two sufficient conditions:

\begin{align*}
\sum_{i=1}^{n} E_{F_n} \left[ |\kappa'W_{i,n}|^2 \mid \mathcal{F}_{i-1,n} \right] &\rightarrow_p \kappa'\Sigma_G \kappa, \text{ and} \\
\sum_{i=1}^{n} E_{F_n} \left[ |\kappa'W_{i,n}|^2 1 \{ |\kappa'W_{i,n}| > \varepsilon \} \right] &\rightarrow_p 0, \ \forall \varepsilon > 0. \tag{E.3}
\end{align*}

In order to show (E.3), it suffices to show that

\begin{align*}
\sum_{i=1}^{n} E_{F_n} \left[ W_{i,n} W'_{i,n} \mid \mathcal{F}_{i-1,n} \right] - \Sigma_G = o_p(1). \tag{E.5}
\end{align*}

Note that in part (a) of equation (A.6) in the proof of Theorem 4.1 we have shown that

\begin{align*}
\sum_{i=1}^{n} E_{F_n} \left[ W_{i,n}^2 \mid \mathcal{F}_{i-1,n} \right] \rightarrow_p 1. \tag{E.6}
\end{align*}

Also, we have, for $j = 1, 2$,

\begin{align*}
\sum_{i=1}^{n} E_{F_n} \left[ W_{j,i,n}^2 \mid \mathcal{F}_{i-1,n} \right] \\
= n^{-1} \sum_{i=1}^{n} \psi_{\alpha,k_j}(\alpha'_{k_j})' H_{F_n,k_j}^{-1} E_{F_n} \left[ \ell_{\alpha,k_j}(Z_i; \alpha'_{k_j}) \ell_{\alpha,k_j}(Z_i; \alpha'_{k_j})' \mid \mathcal{F}_{i-1,n} \right] H_{F_n,k_j}^{-1} \psi_{\alpha,k_j}(\alpha'_{k_j}) \\
= n^{-1} \sum_{i=1}^{n} \psi_{\alpha,k_j}(\alpha'_{k_j})' H_{F_n,k_j}^{-1} D_{F_n,k_j} H_{F_n,k_j}^{-1} \psi_{\alpha,k_j}(\alpha'_{k_j}) \psi_{\alpha,k_j}(\alpha'_{k_j}) \\
= 1, \tag{E.7}
\end{align*}

where the second equality is by Assumption 4.1(b) and the definition of $D_{F_n,k_j}$, the last equality is by the definition of $\psi_{\alpha,k_j}$. Next, note that, for $j = 1, 2$,

\begin{align*}
\sum_{i=1}^{n} E_{F_n} \left[ W_{i,n} W_{j,i,n} \mid \mathcal{F}_{i-1,n} \right] = \sum_{i=1}^{n} E_{F_n} \left[ V_{i,n} W_{j,i,n} \mid \mathcal{F}_{i-1,n} \right] - \sum_{i=2}^{n} E_{F_n} \left[ U_{i,n} W_{j,i,n} \mid \mathcal{F}_{i-1,n} \right]
\end{align*}
\[
\begin{align*}
&= n^{-1} \sum_{i=1}^{n} E_{F_n} \left[ \frac{\ell(Z_i; \alpha_n^*) \ell_{\alpha, k_j}(Z_i; \alpha_{k_j}^*) H_{F_n, k_j}^{-1} \psi_{\alpha, k_j}(\alpha_{k_j}^*)}{\sigma_{F_n, n} v_{\psi, k_j}} \right] F_{i-1, n} \\
&= n^{-1} \sum_{i=1}^{n} \rho_{0j,n} = \rho_{0j,n},
\end{align*}
\]

where the second equality is by the definitions of \(V_{i,n}, U_{i,n}\) and \(W_{j,i,n}\) and \(E_{F_n} [U_{i,n} W_{j,i,n} | F_{i-1, n}] = 0\) which follows by the i.i.d. assumption and Assumption 4.1(c), the third equality is by the definition of \(\rho_{0j,n}\) and the i.i.d. assumption. Similarly

\[
\begin{align*}
&= \frac{1}{n} \sum_{i=1}^{n} E_{F_n} \left[ W_{i,n} W_{i,n}' \right] F_{i-1, n} \\
&= \frac{1}{n} \sum_{i=1}^{n} E_{F_n} \left[ \frac{\psi_{\alpha, k_1}(\alpha_{k_1}^*) H_{F_n, k_1}^{-1} \ell_{\alpha, k_1}(Z_i; \alpha_{k_1}^*) H_{F_n, k_2}^{-1} \psi_{\alpha, k_2}(\alpha_{k_2}^*)}{n^2 v_{\psi, k_1} n^2 v_{\psi, k_2}} \right] F_{i-1, n} \\
&= \frac{1}{n} \sum_{i=1}^{n} \rho_{12,n} = \rho_{12,n}.
\end{align*}
\]

Collecting the results in (E.6), (E.7), (E.8) and (E.9), we get

\[
\sum_{i=1}^{n} E_{F_n} \left[ W_{i,n} W_{i,n}' | F_{i-1, n} \right] - \Sigma_G = \left( \sum_{i=1}^{n} E_{F_n} \left[ W_{i,n}^2 | F_{i-1, n} \right] - 1 \right) \begin{pmatrix} 0_{1 \times 2} \\ 0_{2 \times 1} \end{pmatrix} = o_p(1)
\]

which proves (E.5) and hence (E.3).

We next verify (E.4). First, by the monotonicity of expectation and the \(C_r\) inequality,

\[
\begin{align*}
&= \sum_{i=1}^{n} E_{F_n} \left[ |\kappa' W_{i,n}|^2 1 \{ |\kappa' W_{i,n}| > \varepsilon \} \right] \le \sum_{i=1}^{n} E_{F_n} \left[ |\kappa' W_{i,n}|^4 \varepsilon^{-2} \right] \\
&\le \frac{C}{\varepsilon^2} \sum_{i=1}^{n} E_{F_n} \left[ |\kappa_1 W_{i,n}|^4 + |\kappa_2 W_{i,n}|^4 + |\kappa_3 W_{i,n}|^4 \right].
\end{align*}
\]

In equation (A.24) in the proof of Theorem 4.1 we have shown that

\[
\frac{1}{\varepsilon^2} \sum_{i=1}^{n} E_{F_n} \left[ |W_{i,n}|^4 \right] = o(1).
\]

36
Under Assumptions 4.1(a)-(b) and 6.1, we have

$$\frac{1}{\varepsilon^2} \sum_{i=1}^{n} E_{F_n} \left[ |\kappa_1 W_{1,i,n}|^4 + |\kappa_2 W_{2,i,n}|^4 \right]$$

$$\leq \frac{1}{\varepsilon^2 n^2} \sum_{i=1}^{n} \sum_{j=1}^{2} E_{F_n} \left[ \frac{\psi_{\alpha,k_j}(\alpha_{k_j})}{v_{\alpha,k_j}} \right] \left[ \left| H_{F_n,k_j}^{-1} \ell_{\alpha,k_j}(Z_i; \alpha_{k_j}) \right|^4 \right]$$

$$= \frac{1}{\varepsilon^2 n^2} \sum_{j=1}^{2} E_{F_n} \left[ \frac{\psi_{\alpha,k_j}(\alpha_{k_j})}{v_{\alpha,k_j}} \right] \left[ \left| H_{F_n,k_j}^{-1} \ell_{\alpha,k_j}(Z_i; \alpha_{k_j}) \right|^4 \right] = o(1) \quad (E.12)$$

which combined with (E.10) and (E.11) proves (E.4). □

Proof of Lemma 6.2. (a) Note that under Assumption 6.2(f), we have

$$\| \hat{H}_{k_j,n} - H_{F_n,k_j} \| = o_p(|k|^{-1/2}) \quad \text{and} \quad \| \hat{D}_{k_j,n} - D_{F_n,k_j} \| = o_p(|k|^{-1/2}), \quad (E.13)$$

which together with Assumption 4.1(f) imply that

$$0 \leq \rho_{\max}(\hat{D}_{k_j,n}) \leq 2C \quad \text{and} \quad (2C)^{-1} \leq |\rho_{\min}(\hat{H}_{k_j,n})| \leq |\rho_{\max}(\hat{H}_{k_j,n})| \leq 2C \quad (E.14)$$

with probability approaching 1. By definition, we can write

$$\frac{\hat{v}_{\psi,k_j}^2 - v_{\psi,k_j}^2}{v_{\psi,k_j}^2} = \left[ \psi_{\alpha,k_j}(\hat{\alpha}_{k_j}) - \psi_{\alpha,k_j}(\alpha_{k_j}) \right] \left[ \frac{\hat{H}_{k_j,n}}{v_{\psi,k_j}^2} \right] \left[ \psi_{\alpha,k_j}(\hat{\alpha}_{k_j}) - \psi_{\alpha,k_j}(\alpha_{k_j}) \right]$$

$$+ 2 \left[ \psi_{\alpha,k_j}(\hat{\alpha}_{k_j}) - \psi_{\alpha,k_j}(\alpha_{k_j}) \right] \left[ \frac{\hat{H}_{k_j,n} - H_{F_n,k_j}}{v_{\psi,k_j}^2} \right] \left[ \psi_{\alpha,k_j}(\hat{\alpha}_{k_j}) - \psi_{\alpha,k_j}(\alpha_{k_j}) \right]$$

$$+ \psi_{\alpha,k_j}(\alpha_{k_j}) H_{F_n,k_j}^{-1} \left[ \hat{H}_{k_j,n} - H_{F_n,k_j} \right] \left[ \frac{\hat{D}_{k_j,n} - D_{F_n,k_j}}{v_{\psi,k_j}^2} \right] \psi_{\alpha,k_j}(\alpha_{k_j})$$

$$+ \psi_{\alpha,k_j}(\alpha_{k_j}) H_{F_n,k_j}^{-1} \left[ \hat{D}_{k_j,n} - D_{F_n,k_j} \right] H_{F_n,k_j}^{-1} \left[ \frac{\hat{H}_{k_j,n} - H_{F_n,k_j}}{v_{\psi,k_j}^2} \right] \psi_{\alpha,k_j}(\alpha_{k_j}) \quad (E.15)$$

By (E.14) and Assumption 6.2(a),

$$\left[ \psi_{\alpha,k_j}(\hat{\alpha}_{k_j}) - \psi_{\alpha,k_j}(\alpha_{k_j}) \right] \left[ \frac{\hat{H}_{k_j,n}}{v_{\psi,k_j}^2} \right] \left[ \psi_{\alpha,k_j}(\hat{\alpha}_{k_j}) - \psi_{\alpha,k_j}(\alpha_{k_j}) \right]$$

37
By the Cauchy-Schwarz inequality, (6.2(a)), Assumptions 6.2(a) and 6.2(e),

\[
\left| \psi_{\alpha,k_j}(\alpha_{k_j}^*) \right|^2 \leq \left( \frac{H_{k_j,n}^{-1}}{v_{\psi,k_j}^2} \right)^2 \left| \psi_{\alpha,k_j}(\alpha_{k_j}^*) \right|^2 = o_p(1). (E.16)
\]

By the Cauchy-Schwarz inequality, (E.13), (E.14), Assumptions 6.2(a) and 6.2(e),

\[
\left| \psi_{\alpha,k_j}(\alpha_{k_j}^*) \right|^2 \leq \left( \frac{H_{k_j,n}^{-1}}{v_{\psi,k_j}^2} \right)^2 \left| \psi_{\alpha,k_j}(\alpha_{k_j}^*) \right|^2 = o_p(1). (E.17)
\]

By the Cauchy-Schwarz inequality, (E.13), (E.14), Assumptions 6.2(a) and 6.2(e),

\[
\left| \psi_{\alpha,k_j}(\alpha_{k_j}^*) \right|^2 \leq \left( \frac{H_{k_j,n}^{-1}}{v_{\psi,k_j}^2} \right)^2 \left| \psi_{\alpha,k_j}(\alpha_{k_j}^*) \right|^2 = o_p(1). (E.18)
\]

Using the same arguments for showing (E.18), we can prove that

\[
\left| \psi_{\alpha,k_j}(\alpha_{k_j}^*) \right|^2 \leq \left( \frac{H_{k_j,n}^{-1}}{v_{\psi,k_j}^2} \right)^2 \left| \psi_{\alpha,k_j}(\alpha_{k_j}^*) \right|^2 = o_p(1) (E.19)
\]

Collecting the results in (E.15)–(E.19), we immediately get the claimed result.

(b) By the consistency of $\hat{\rho}_{12,n}^2$ and the fact that $|\rho_{12,n}| \leq 1$, it is sufficient to show that

\[
\hat{\rho}_{12,n}^2 = \frac{\psi_{\alpha,k_1}(\hat{\alpha}_k_1) \hat{H}_{k_1,n}^{-1}}{v_{\psi,k_1}^2} \frac{\hat{H}_{k_2,n}^{-1} \psi_{\alpha,k_2}(\hat{\alpha}_k_2)}{v_{\psi,k_2}^2} = o_p(1). (E.20)
\]
By Assumption 6.2(f), we have

\[
\| \hat{D}_{k_1,k_2,n} - D_{F_{k_1,k_2}} \| = o_p(|k|^{-1/2}). \tag{E.21}
\]

Note that

\[
\hat{\rho}_{12,n} - \rho_{12,n} = \left[ \frac{\psi_{\alpha,k_1}(\hat{\alpha}_{k_1}) - \psi_{\alpha,k_1}(\alpha_{k_1}^*)}{v_{\psi,k_1}} \right] H_{k_1,n}^{-1} \hat{D}_{k_1,k_2,n} H_{k_2,n}^{-1} \left[ \frac{\psi_{\alpha,k_2}(\hat{\alpha}_{k_2}) - \psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}} \right]
+ 2 \left[ \frac{\psi_{\alpha,k_1}(\hat{\alpha}_{k_1}) - \psi_{\alpha,k_1}(\alpha_{k_1}^*)}{v_{\psi,k_1}} \right] H_{k_1,n}^{-1} \hat{D}_{k_1,k_2,n} H_{k_2,n}^{-1} \left[ \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}} \right]
+ \left[ \frac{\psi_{\alpha,k_1}(\alpha_{k_1}^*)}{v_{\psi,k_1}} \right] H_{k_1,n}^{-1} \hat{D}_{k_1,k_2,n} H_{k_2,n}^{-1} \left[ \frac{\psi_{\alpha,k_2}(\hat{\alpha}_{k_2}) - \psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}} \right]
+ \left[ \frac{\psi_{\alpha,k_1}(\alpha_{k_1}^*)}{v_{\psi,k_1}} \right] H_{k_1,n}^{-1} \hat{D}_{k_1,k_2,n} H_{k_2,n}^{-1} \left[ \frac{\psi_{\alpha,k_2}(\hat{\alpha}_{k_2}) - \psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}} \right]. \tag{E.22}
\]

Lemma F.2 together with Assumption 4.1(f) implies that

\[
\rho_{\max}(D_{F_{k_1,k_2}} D_{F_{k_1,k_2}}) \leq C. \tag{E.23}
\]

Similarly, Lemma F.2 together with Assumptions 4.1(f) and (E.21) implies that

\[
\rho_{\max}(\hat{D}_{k_1,k_2,n} \hat{D}_{k_1,k_2,n}) \leq C \tag{E.24}
\]

with probability approaching 1. Using the Cauchy-Schwarz inequality, we get

\[
\frac{\left[ \frac{\psi_{\alpha,k_1}(\hat{\alpha}_{k_1}) - \psi_{\alpha,k_1}(\alpha_{k_1}^*)}{v_{\psi,k_1}} \right] H_{k_1,n}^{-1} \hat{D}_{k_1,k_2,n} H_{k_2,n}^{-1} \left[ \frac{\psi_{\alpha,k_2}(\hat{\alpha}_{k_2}) - \psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}} \right]^2}{\rho_{\min}(\hat{H}_{k_1,n})^2 \rho_{\min}(\hat{H}_{k_2,n})^2} \leq o_p(1), \tag{E.25}
\]

39
where the last equality is by Assumption 6.2(a), (E.14) and (E.24). Similarly,

\[
\left| \frac{\left[ \psi_{\alpha,k_1}(\hat{\alpha}_{k_1}) - \psi_{\alpha,k_1}(\alpha_{k_1}^*) \right]'}{v_{\psi,k_1}^*} \tilde{H}_{k_1,n}^{-1} \tilde{D}_{k_1,k_2,n} \tilde{H}_{k_2,n}^{-1} \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}^*} \right|^2 \\
\leq \frac{\left[ \psi_{\alpha,k_1}(\hat{\alpha}_{k_1}) - \psi_{\alpha,k_1}(\alpha_{k_1}^*) \right]'}{v_{\psi,k_1}^*} \tilde{H}_{k_1,n}^{-2} \left[ \psi_{\alpha,k_1}(\hat{\alpha}_{k_1}) - \psi_{\alpha,k_1}(\alpha_{k_1}^*) \right] \\
\times \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)'}{v_{\psi,k_2}^*} \tilde{H}_{k_2,n}^{-1} \tilde{D}_{k_1,k_2,n} \tilde{D}_{k_1,k_2,n} \tilde{H}_{k_2,n}^{-1} \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}^*} \\
\leq \frac{\rho_{\max}(\tilde{D}_{k_1,k_2,n} \tilde{D}_{k_1,k_2,n})}{\rho_{\min}((\tilde{H}_{k_1,n})^2)^2} \left\| \psi_{\alpha,k_1}(\hat{\alpha}_{k_1}) - \psi_{\alpha,k_1}(\alpha_{k_1}^*) \right\| ^2 \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_1}^*} \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}^*} = o_p(1) \quad (E.26)
\]

where the last equality is by Assumptions 6.2(a), 6.2(e), (E.14) and (E.24). Moreover,

\[
\left| \frac{\psi_{\alpha,k_1}(\alpha_{k_1}^*)'}{v_{\psi,k_1}^*} H_{F_{n,k_1}}^{-1} (\tilde{H}_{k_1,n} - H_{F_{n,k_1}}) \tilde{H}_{k_1,n}^{-1} \tilde{D}_{k_1,k_2,n} \tilde{H}_{k_2,n}^{-1} \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}^*} \right|^2 \\
\leq \left\| \tilde{H}_{k_1,n} - H_{F_{n,k_1}} \right\| ^2 \frac{\psi_{\alpha,k_1}(\alpha_{k_1}^*)'}{v_{\psi,k_1}^*} H_{F_{n,k_1}}^{-2} \frac{\psi_{\alpha,k_1}(\alpha_{k_1}^*)}{v_{\psi,k_1}^*} \\
\times \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)'}{v_{\psi,k_2}^*} \tilde{H}_{k_2,n}^{-1} \tilde{D}_{k_1,k_2,n} \tilde{H}_{k_1,n}^{-2} \tilde{D}_{k_1,k_2,n} \tilde{H}_{k_2,n}^{-1} \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}^*} \\
\leq \left\| \tilde{H}_{k_1,n} - H_{F_{n,k_1}} \right\| ^2 \rho_{\max}(\tilde{D}_{k_1,k_2,n} \tilde{D}_{k_1,k_2,n}) \\
\times \frac{\rho_{\min}(H_{F_{n,k_1}})^2}{\rho_{\min}(\tilde{H}_{k_1,n})^2 \rho_{\min}(\tilde{H}_{k_2,n})} \frac{\psi_{\alpha,k_1}(\alpha_{k_1}^*)'}{v_{\psi,k_1}^*} \psi_{\alpha,k_1}(\alpha_{k_1}^*) \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}^*} \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}^*} = o_p(1) \quad (E.27)
\]

where the last equality is by Assumptions 6.2(a), 6.2(e), (E.13), (E.14) and (E.24). Using the same arguments in showing (E.27), but replacing (E.24) with (E.23), we have

\[
\left| \frac{\psi_{\alpha,k_1}(\alpha_{k_1}^*)'}{v_{\psi,k_1}^*} H_{F_{n,k_1}}^{-1} D_{F_{n,k_1,k_2}} H_{F_{n,k_1}}^{-1} (\tilde{H}_{k_1,n} - H_{F_{n,k_1}}) \tilde{H}_{k_1,n}^{-1} \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}^*} \right|^2 = o_p(1). \quad (E.28)
\]

By the Cauchy-Schwarz inequality,

\[
\left| \frac{\psi_{\alpha,k_1}(\alpha_{k_1}^*)'}{v_{\psi,k_1}^*} H_{F_{n,k_1}}^{-1} (\tilde{D}_{k_1,k_2,n} - D_{F_{n,k_1,k_2}}) \tilde{H}_{k_2,n}^{-1} \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}^*} \right|^2 \\
\leq \left\| \tilde{D}_{k_1,k_2,n} - D_{F_{n,k_1,k_2}} \right\|^2 \frac{\psi_{\alpha,k_1}(\alpha_{k_1}^*)'}{v_{\psi,k_1}^*} \psi_{\alpha,k_1}(\alpha_{k_1}^*) \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}^*} \frac{\psi_{\alpha,k_2}(\alpha_{k_2}^*)}{v_{\psi,k_2}^*} = o_p(1), \quad (E.29)
\]
where the last equality is by Assumptions 4.1(f), 6.2(a), 6.2(e), (E.14) and (E.21). Collecting the results in (E.22), (E.25)-(E.29), we immediately prove (E.20).

(c) By the consistency of $\hat{\omega}_{j,n}^2$ and $\hat{\sigma}_n$, and the fact that $|\rho_{0,j,n}| \leq 1$, it is sufficient to show that

$$
\hat{\rho}_{0,j,n} = \frac{\psi_{\alpha,k_j}^\prime(\hat{\alpha}_{k_j}) \hat{H}_{k_j,n}^{-1} \sum_{i=1}^n \ell_{\alpha,k_j}(Z_i; \hat{\alpha}_{k_j}) \ell(Z_i; \hat{\alpha}_n)}{n v_{\psi,k_j}^* \sigma_{F_n,n}} = \rho_{0,j,n} + o_p(1).
$$

(E.30)

By definition, we can write

$$
\hat{\rho}_{0,j,n} - \rho_{0,j,n} = \frac{\psi_{\alpha,k_j}^\prime(\alpha_{k_j}^*) H_{F_{n,k_j}}^{-1} \sum_{i=1}^n \ell_{\alpha,k_j}(Z_i; \alpha_{k_j}^*) \ell(Z_i; \alpha_n^*)}{n v_{\psi,k_j}^* \sigma_{F_n,n}} - \frac{\psi_{\alpha,k_j}^\prime(\hat{\alpha}_{k_j}) \hat{H}_{k_j,n}^{-1} \sum_{i=1}^n \ell_{\alpha,k_j}(Z_i; \hat{\alpha}_{k_j}) \ell(Z_i; \hat{\alpha}_n)}{n v_{\psi,k_j}^* \sigma_{F_n,n}} + \frac{\psi_{\alpha,k_j}(\hat{\alpha}_{k_j}) \hat{H}_{k_j,n}^{-1} \sum_{i=1}^n \ell_{\alpha,k_j}(Z_i; \hat{\alpha}_{k_j}) \ell(Z_i; \hat{\alpha}_n)}{n v_{\psi,k_j}^* \sigma_{F_n,n}}.
$$

(E.31)

Under the i.i.d assumption, Assumptions 4.1(f), 4.3(b) and 6.2(c),

$$
\text{Var} \left[ \frac{\psi_{\alpha,k_j}(\alpha_{k_j}^*) H_{F_{n,k_j}}^{-1}}{n v_{\psi,k_j}^* \sigma_{F_n,n}} \sum_{i=1}^n \ell_{\alpha,k_j}(Z_i; \alpha_{k_j}^*) \ell(Z_i; \alpha_n^*) - E_{F_n} \left[ \ell_{\alpha,k_j}(Z; \alpha_{k_j}^*) \ell(Z; \alpha_n^*) \right] \right]
$$

$$
= \frac{\psi_{\alpha,k_j}(\alpha_{k_j}^*) H_{F_{n,k_j}}^{-1}}{n^2 v_{\psi,k_j}^* \sigma_{F_n,n}} E_{F_n} \left[ \ell^2(Z; \alpha_n^*) \ell_{\alpha,k_j}(Z; \alpha_{k_j}^*) \ell_{\alpha,k_j}(Z; \alpha_{k_j}^*) \right] \frac{H_{F_{n,k_j}}^{-1} \psi_{\alpha,k_j}(\alpha_{k_j}^*)}{n^2 v_{\psi,k_j}^* \sigma_{F_n,n}},
$$

$$
\leq \frac{\rho_{\text{max}}(D_{\ell,k_j,n})}{\rho_{\text{min}}((H_{F_{n,k_j}})^2)} \frac{\psi_{\alpha,k_j}(\alpha_{k_j}^*) \psi_{\alpha,k_j}(\alpha_{k_j}^*)}{v_{\psi,k_j}^2} \frac{1}{n \sigma_{F_n,n}^2} = o_p(1),
$$

(E.32)

which together with the Markov inequality implies that the first summand in (E.31) is $o_p(1)$. By the triangle inequality

$$
\left\| \frac{\psi_{\alpha,k_j}(\alpha_{k_j}^*) H_{F_{n,k_j}}^{-1}}{n v_{\psi,k_j}^* \sigma_{F_n,n}} \sum_{i=1}^n \ell_{\alpha,k_j}(Z_i; \alpha_{k_j}^*) \ell(Z_i; \alpha_n^*) \right\|^2
$$

$$
\leq \left[ \frac{\psi_{\alpha,k_j}(\alpha_{k_j}^*)}{v_{\psi,k_j}^2} - \frac{\psi_{\alpha,k_j}(\hat{\alpha}_{k_j})}{v_{\psi,k_j}^2} \right] \hat{H}_{k_j,n}^{-2} \left[ \frac{\psi_{\alpha,k_j}(\alpha_{k_j}^*)}{v_{\psi,k_j}^2} - \frac{\psi_{\alpha,k_j}(\hat{\alpha}_{k_j})}{v_{\psi,k_j}^2} \right]
$$

$$
+ \frac{\psi_{\alpha,k_j}(\alpha_{k_j}^*) H_{F_{n,k_j}}^{-1} \hat{H}_{k_j,n} - H_{F_{n,k_j}} \hat{H}_{k_j,n} \hat{H}_{k_j,n} H_{F_{n,k_j}}^{-1} \psi_{\alpha,k_j}(\alpha_{k_j}^*)}{v_{\psi,k_j}^2}
$$

41
where the equality is by Assumption 6.2(b). Also note that

\[
\ell(Z_i; \alpha^*_n, \alpha^*_k) - E_{F_n, n} \left[ \ell(Z_i; \alpha^*_n) \ell_{\alpha, k_j}(Z_i; \alpha^*_k) \right] \leq E_{F_n, n} \left[ \ell(Z_i; \alpha^*_n) \ell_{\alpha, k_j}(Z_i; \alpha^*_k) \right] = o_p(k_j n^{-1}),
\]

(E.34)

where the equality is by Assumption 6.2(b). Also note that

\[
\left\| \ell_{\alpha, k_j}(Z_i; \alpha^*_n) \ell(Z_i; \alpha^*_k) \right\|^2 \leq E_{F_n, n} \left[ (\ell(Z_i; \alpha^*_n))^2 \right] E_{F_n, n} \left[ \ell_{\alpha, k_j}(Z_i; \alpha^*_n) \ell_{\alpha, k_j}(Z_i; \alpha^*_k) \right] = o_p(k_j \sigma_{F_n, n}^2)
\]

The above two displays together with the triangle inequality imply that

\[
\left\| n^{-1} \sum_{i=1}^n \ell(Z_i; \alpha^*_n) \ell_{\alpha, k_j}(Z_i; \alpha^*_k) \right\|^2 \leq o_p(k_j n^{-1}) + o_p(k_j \sigma_{F_n, n}^2).
\]

(E.35)

Equations (E.33) and (E.35) imply that the second summand in (E.31) is bounded from above by

\[
o_p(|k|^{-1/2})(O(k_j^{1/2} n^{-1/2} \sigma_{F_n, n}^{-1}) + o_p(k_j^{1/2} \sigma_{F_n, n} \sigma_{F_n, n}^{-1})) = o_p(1),
\]

where the equality is by Assumption 4.3(b). Thus, the second summand in (E.31) is also \(o_p(1)\).

To show that the last summand in (E.31) is \(o_p(1)\), let

\[
\hat{A}_n = \frac{\psi_{\alpha, k_j}(\alpha_k)}{\psi_{\alpha, k_j} \sigma_{F_n, n}}, \quad A_n = \frac{\psi_{\alpha, k_j}(\alpha_k) H_{F_n, k_j}}{\psi_{\alpha, k_j} \sigma_{F_n, n}} \quad \hat{B}_n = n^{-1} \sum_{i=1}^n [\ell_{\alpha, k_j}(Z_i; \alpha_k^*) H_{\ell, k_j} \bar{\ell}_{\alpha, n}(\alpha_k^*)]
\]

\[
B_n = E_{F_n, n} [\ell_{\alpha, k_j}(Z_i; \alpha_k^*) H_{\ell, k_j} \bar{\ell}_{\alpha, n}(\alpha_k^*)] \quad \hat{C}_n = n^{-1} \sum_{i=1}^n \ell(Z_i; \alpha_k^*) H_{\ell, k_j} \bar{\ell}_{\alpha, n}(\alpha_k^*) \quad \text{and} \quad C_n = H_{\ell, k_j} H_{\ell, k_j} \bar{\ell}_{\alpha, n}(\alpha_k^*),
\]

(E.36)
Using Assumption 6.2(c), we can write the last summand in (E.31) as
\[
\tilde{A}_n(B_n + \tilde{C}_n + o_p(\sigma_{F_n,n})).
\]

Next, note that (E.33) implies that
\[
\hat{A}_n = A_n + \tilde{A}_n - A_n = \frac{\psi_{\alpha,k_j}(\alpha_{k_j}^*)' H_{F_n,k_j}^{-1}}{v_{\psi,k_j}^* \sigma_{F_n,n}} + o_p(|k|^{-1/2} \sigma_{F_n,n}^{-1}) = O_p(\sigma_{F_n,n}^{-1}),
\]
where the last equality is by Assumptions 4.1(f) and 6.2(e). Under Assumptions 4.1(a) and (f)
\[
E_{F_n} \left[ \tilde{\ell}_{\alpha,n}(\alpha_k^*') H_{F_n,k}^{-2} \tilde{\ell}_{\alpha,n}(\alpha_k^*) \right] = \frac{1}{n} \text{tr} \left( H_{F_n,k}^{-2} D_{F_n,k} \right) = O(|k| n^{-1}),
\]
which together with the Markov inequality implies that
\[
\tilde{\ell}_{\alpha,n}(\alpha_k^*') H_{F_n,k}^{-2} \tilde{\ell}_{\alpha,n}(\alpha_k^*) = O_p(|k| n^{-1}).
\]
By the Cauchy-Schwarz inequality,
\[
\| \hat{B}_n - B_n \| \leq \left\| \frac{1}{n} \sum_{i=1}^{n} \ell_{\alpha,k_j}(Z_i; \alpha_{k_j}^*) \ell_{\alpha,k_j}'(Z_i; \alpha_{k_j}^*) - E_{F_n} \left[ \ell_{\alpha,k_j}(Z; \alpha_{k_j}^*) \ell_{\alpha,k_j}'(Z; \alpha_{k_j}^*) \right] \right\| \sqrt{E_{F_n} \left[ \tilde{\ell}_{\alpha,n}(\alpha_k^*) H_{F_n,k}^{-2} \tilde{\ell}_{\alpha,n}(\alpha_k^*) \right]}
\]
\[
= o_p(n^{-\frac{1}{2}}),
\]
where the equality is by (E.38) and Assumption 6.2(f). Note that Lemma F.2 combined with Assumptions 4.1(f) implies that
\[
\rho_{\max} \left( E_{F_n} \left[ \ell_{\alpha,k_j}(Z; \alpha_{k_j}^*) \ell_{\alpha,k_j}'(Z; \alpha_{k_j}^*) \right] E_{F_n} \left[ \ell_{\alpha,k}(Z; \alpha_k^*) \ell_{\alpha,k}'(Z; \alpha_k^*) \right] \right) \leq C.
\]
Under the i.i.d. assumption,
\[
E_{F_n} \left[ \| B_n \|^2 \right] = \text{tr} \left( E_{F_n} \left[ \ell_{\alpha,k_j}(Z; \alpha_{k_j}^*) \ell_{\alpha,k_j}'(Z; \alpha_{k_j}^*) \right] H_{F_n,k}^{-1} D_{F_n,k} H_{F_n,k}^{-1} \right) E_{F_n} \left[ \ell_{\alpha,k}(Z; \alpha_k^*) \ell_{\alpha,k}'(Z; \alpha_k^*) \right]
\]
\[
\leq \frac{\rho_{\max}(D_{F_n,k})}{n \rho_{\min}((H_{F_n,k})^2)} \text{tr} \left( E_{F_n} \left[ \ell_{\alpha,k_j}(Z; \alpha_{k_j}^*) \ell_{\alpha,k_j}'(Z; \alpha_{k_j}^*) \right] E_{F_n} \left[ \ell_{\alpha,k}(Z; \alpha_k^*) \ell_{\alpha,k}'(Z; \alpha_k^*) \right] \right)
\]
\[
= O(|k| n^{-1})
\]
where the last equality is by Assumptions 4.1(f) and (E.40). (E.41) combined with the Markov
inequality yields

\[ B_n = O_p(|k|^{\frac{1}{2}} n^{-\frac{1}{2}}). \]  

By the Cauchy-Schwarz inequality,

\[
\|\tilde{C}_n - C_n\| \leq \left\| \frac{1}{n} \sum_{i=1}^{n} \ell(Z_i; \alpha_{k_j}^*) \ell_{\alpha, k_j}(Z_i; \alpha_{k_j}^*) - H_{\ell, k_j,n} \right\| \sqrt{\ell_{\alpha, n}(\alpha_{k_j}^*) H_{F_{n, k_j}}^{-1} \ell_{\alpha, n}(\alpha_{k_j})}
\]

\[ = o_p(n^{-\frac{1}{2}}), \]  

(E.43)

where the equality is by (E.38) and Assumption 6.2(d). Under the i.i.d. assumption,

\[
E_{F_n} \left[ \|C_n\|^2 \right] = tr \left( n^{-1} H_{\ell, k_j,n} H_{F_{n, k}}^{-1} D_{F_{n, k}} H_{F_{n, k}}^{-1} H_{\ell, k_j,n} \right) = O(|k| n^{-1}),
\]

(E.44)

where the last equality is by Assumptions 4.1(f) and 6.2(b). Thus,

\[ C_n = O_p(|k| n^{-1}). \]  

(E.45)

Under the i.i.d. Assumption,

\[
E_{F_n} \left[ |A_n B_n|^2 \right] = \frac{\psi_{\alpha, k_j}(\alpha_{k_j}^*)'}{v_{\psi, k_j, F_{n, n}}^*} E_{F_n} \left[ \ell_{\alpha, k_j}(Z; \alpha_{k_j}^*) \ell_{\alpha, k_j}(Z; \alpha_{k_j}^*)' \right] \frac{H_{F_{n, k}}^{-1} D_{F_{n, k}} H_{F_{n, k}}^{-1}}{n} H_{\ell, k_j,n}^{-1} \psi_{\alpha, k_j}(\alpha_{k_j}^*)
\]

\[
\times E_{F_n} \left[ \ell_{\alpha, k_j}(Z; \alpha_{k_j}^*) \ell_{\alpha, k_j}(Z; \alpha_{k_j}^*) \right] \frac{H_{F_{n, k}}^{-1} \psi_{\alpha, k_j}(\alpha_{k_j}^*)}{v_{\psi, k_j, F_{n, n}}^*} \leq C \rho_{\max}(D_{F_{n, k}}) \frac{\psi_{\alpha, k_j}(\alpha_{k_j}^*)'}{v_{\psi, k_j, F_{n, n}}^*} \frac{\psi_{\alpha, k_j}(\alpha_{k_j}^*)}{v_{\psi, k_j, F_{n, n}}^*} \frac{1}{n \sigma_{F_{n, n}}^2} = o(1)
\]

(E.46)

where the inequality is by (E.40), the last equality is by Assumptions 4.1(f), 4.3(b) and 6.2(e). Similarly,

\[
E_{F_n} \left[ |A_n C_n|^2 \right] = \frac{\psi_{\alpha, k_j}(\alpha_{k_j}^*)'}{v_{\psi, k_j, F_{n, n}}^*} H_{\ell, k_j,n} H_{F_{n, k}, j}^{-1} \frac{H_{F_{n, k}}^{-1} D_{F_{n, k}} H_{F_{n, k}}^{-1} H_{\ell, k_j,n}^{-1} H_{F_{n, k}}^{-1} H_{\ell, k_j,n}^{-1}}{n} \frac{H_{F_{n, k}}^{-1} \psi_{\alpha, k_j}(\alpha_{k_j}^*)}{v_{\psi, k_j, F_{n, n}}^*} \frac{1}{n \sigma_{F_{n, n}}^2} = o(1),
\]

(E.47)

where the last equality is by Assumptions 4.1(f), 4.3(b), 6.2(b) and 6.2(e). Therefore

\[
\tilde{A}_n(\tilde{B}_n + \tilde{C}_n + o_p(\sigma_{F_{n, n}}))
\]

\[ = \tilde{A}_n(\tilde{B}_n + \tilde{C}_n) + o_p(1)
\]

44
\[
= \hat{A}_n(\hat{B}_n - B_n + \hat{C}_n - C_n) + A_nB_n + A_nC_n + (\hat{A}_n - A_n)(B_n + C_n) + o_p(1)
\]
\[
= O_p(\sigma_1^{-1} o_p(n^{-1/2}) + o_p(1) + o_p(|k|^{-1/2} \sigma_1^{-1} o_p(|k|^{1/2} n^{-1/2})) + o_p(1)
\]
\[
= o_p(1). \quad (E.48)
\]

Thus, the last summand in (E.31) is \(o_p(1)\). This completes the proof of (E.30) and hence the claim (c) of the theorem.

\section{F Auxiliary Lemmas}

\textbf{Lemma F.1} Let \(X\) be a scalar random variable with variance \(\sigma_X^2\) and \(Y = (Y_1, \ldots, Y_d)'\) be a \(d\)-dimensional random vector with variance-covariance matrix \(D_Y\). Let \(\rho\) be the vector correlation coefficient of \(X\) and \(Y\). That is, \(\rho = \sigma_X^{-1} (D_Y^{-1/2})^+ \text{Cov}(X,Y)\), where \(A^{1/2}\) is the unique symmetric matrix square root of the positive semi-definite matrix \(A\) and \(A^+\) is the Moore-Penrose inverse of \(A\). Then (a) \(\rho' \rho \leq 1\); (b) for any positive semi-definite matrix \(A\), \(\rho' A \rho \leq \lambda_{\text{max}}(A)\), where \(\lambda_{\text{max}}(A)\) is the maximum eigenvalue of \(A\); and (c) \((D_Y^{1/2})^+ \text{Cov}(X,Y)' D_Y^{1/2} = \text{Cov}(X,Y)'\).

\textbf{Proof of Lemma [F.1]} First we show part (a). If \(\sigma_X = 0\), then \(\rho = 0\) by definition and part (a) holds. Thus below, we consider the nontrivial case of \(\sigma_X > 0\). Without loss of generality, normalize \(\sigma_X = 1\).

Consider the simple case that \(D_Y\) is invertible first. Then,
\[
\begin{pmatrix}
1 & \rho'
\rho & I_d
\end{pmatrix}
= \begin{pmatrix}
1 & 0
0 & D_Y^{-1/2}
\end{pmatrix}
\Sigma_{(X,Y)'} \begin{pmatrix}
1 & 0
0 & D_Y^{-1/2}
\end{pmatrix},
\]
where \(\Sigma_{(X,Y)'}\) is the variance covariance matrix of the random vector \((X, Y')\). Due to the positive (semi-)definiteness of \(D_Y\) and \(\Sigma_{(X,Y)'}\), we have \(\begin{pmatrix}
1 & \rho'
\rho & I_d
\end{pmatrix}\) is positive semi-definite. That implies that
\[
\begin{pmatrix}
1 & \rho'
\rho & I_d
\end{pmatrix} \geq 0. \quad (F.2)
\]

Direct calculation shows that \(\begin{vmatrix}
1 & \rho'
\rho & I_d
\end{vmatrix} = 1 - \rho' \rho\). Thus, \(\rho' \rho \leq 1\).

Now consider the simple noninvertible case: \(D_Y\) is a singular diagonal matrix. Without loss of generality, suppose that the first \(J\) elements of \(D_Y\) are zeros for a positive integer \(J \leq d\) and the rest are strictly positive. If \(J = d\), then \(\rho = 0\) and part (a) holds trivially. If \(J < d\), we have
\( \rho_j = 0 \) for \( j = 1, \ldots, J \), and

\[
\rho_{(J+1):d} = D_{Y,(J+1):d}^{-1/2} \text{Cov}(X, Y_{(J+1):d}), \tag{F.3}
\]

where \( \rho_{(J+1):d} = (\rho_{J+1}, \ldots, \rho_d) \), \( Y_{(J+1):d} = (Y_{J+1}, \ldots, Y_d) \) and \( D_{Y,(J+1):d} \) is \( D_Y \) with the first \( J \) rows and columns removed. By the arguments for the invertible \( D_Y \) case, we have \( \rho_{(J+1):d} \rho_{(J+1):d} \leq 1 \). Thus \( \rho' \rho = \rho_{(J+1):d} \rho_{(J+1):d} \leq 1 \).

Finally, consider the case where \( D_Y \) is singular but not diagonal. Because \( D_Y \) is a variance-covariance matrix and thus is positive semi-definite, \( D_Y \) have the following eigenvalue decomposition:

\[
D_Y = Q_Y \Lambda_Y Q_Y', \tag{F.4}
\]

where \( Q_Y \) is an orthonormal matrix and \( \Lambda_Y \) is a diagonal matrix whose diagonal elements are eigenvalues of \( D_Y \). Using this decomposition, we have

\[
\rho = Q_Y (\Lambda^{1/2})^+ Q_Y' \text{Cov}(X, Y), \tag{F.5}
\]

Thus,

\[
Q_Y' \rho = (\Lambda^{1/2})^+ \text{Cov}(X, Q_Y Y). \tag{F.6}
\]

Then by the arguments for the singular diagonal \( D_Y \), we have \( \rho' Q_Y Q_Y' \rho \leq 1 \). But because \( Q_Y Q_Y' = I_d \), we have \( \rho' \rho = \rho' Q_Y Q_Y' \rho \leq 1 \). This concludes the proof of part (a).

Now we show part (b). The matrix \( A \) has the following decomposition:

\[
A = Q_A \Lambda_A Q_A', \tag{F.7}
\]

where \( Q_A \) is a orthonormal matrix, and \( \Lambda_A \) is a diagonal matrix of eigenvalues of \( A \). Thus,

\[
\rho' A \rho = \rho' Q_A \Lambda_A Q_A' \rho = \rho' Q_A \lambda_{\max}(A) I_d Q_A' \rho + \rho' Q_A (\lambda_{\max}(A) I_d - \Lambda_A) Q_A' \rho \\
\leq \rho' Q_A \lambda_{\max}(A) I_d Q_A' \rho = \lambda_{\max}(A) \rho' \rho \leq \lambda_{\max}(A). \tag{F.8}
\]

This shows part (b).

Finally we show part (c). Without loss of generality, suppose that \( D_Y \) has \( J \) many zero eigenvalues which corresponds to the first \( J \) diagonal elements in \( \Lambda \). The claim of the lemma holds trivially when \( J = 0 \) or \( d \). Hence we only need to consider the case that \( 0 < J < d \). By \( (F.4) \), \( D_Y^{1/2} = Q_Y \Lambda_Y^{1/2} Q_Y' \) and \( (D_Y^{1/2})^+ = Q_Y (\Lambda^{1/2})^+ Q_Y' \) where \( (\Lambda^{1/2})^+ \) is a symmetric matrix. Let \( Y^* = Q_Y Y \), then the first \( J \) element of \( Y^* \) is zero almost surely. Let \( Y_{d-J} \) denotes the last \( d - J \)
elements of $Y^*$. Then

$$((D_{Y}^{1/2})^*\text{Cov}(X,Y))^\prime D_{Y}^{1/2}Q_{Y} = \text{Cov}(X,Y)^\prime (D_{Y}^{1/2})^* D_{Y}^{1/2}Q_{Y}$$

$$= \text{Cov}(X,Y^*)^\prime \begin{pmatrix} 0_{J \times J} & 0_{J \times (d-J)} \\ 0_{(d-J) \times J} & I_{d-J} \end{pmatrix}$$

$$= \begin{pmatrix} 0_{1 \times J} \text{Cov}(X,Y_{d-J}^*)^\prime \\ \text{Cov}(X,Y_{d-J}^*)^\prime \end{pmatrix}. \quad \text{(F.9)}$$

By definition $\text{Cov}(X,Y)^\prime Q_{Y} = \text{Cov}(X,Y^*)^\prime = \begin{pmatrix} 0_{1 \times J} \text{Cov}(X,Y_{d-J}^*)^\prime \\ \text{Cov}(X,Y_{d-J}^*)^\prime \end{pmatrix}$ which together with the above equation implies that

$$((D_{Y}^{1/2})^*\text{Cov}(X,Y))^\prime D_{Y}^{1/2}Q_{Y} = \text{Cov}(X,Y)^\prime Q_{Y}. \quad \text{(F.10)}$$

As $Q_{Y}$ is a non-singular matrix, the claim of the lemma follows immediately by (F.10). ■

**Lemma F.2** Let $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, where $A_{12} = A_{21}$, $A_{11}$ and $A_{22}$ are $k_1 \times k_1$ and $k_2 \times k_2$ symmetric matrices respectively. Then $\rho_{\text{max}}(A) \leq C$ implies that:

(a) $\rho_{\text{max}}(A_{11}^2 + A_{12}A_{21}) \leq C^2$ and $\rho_{\text{max}}(A_{22}^2 + A_{21}A_{12}) \leq C^2$;

(b) $\rho_{\text{max}}(A_{12}A_{21}) \leq C^2$ and $\rho_{\text{max}}(A_{21}A_{12}) \leq C^2$.

**Proof of Lemma F.2.** By definition,

$$A^2 = \begin{pmatrix} A_{11}^2 + A_{12}A_{21} & A_{11}A_{12} + A_{12}A_{22} \\ A_{21}A_{11} + A_{22}A_{21} & A_{22}^2 + A_{21}A_{12} \end{pmatrix}. $$

Then it is clear that

$$\rho_{\text{max}}(A_{11}^2 + A_{12}A_{21}) \leq \rho_{\text{max}}(A^2) = \rho_{\text{max}}^2(A),$$

which proves the first claim in (a). The second claim in (a) can be proved by the same way. Let $\gamma_{k_1} \in R^{k_1}$ be the eigenvector of the largest eigenvalue of $A_{12}A_{21}$. Then the above inequality implies that

$$\gamma_{k_1}' A_{11}^2 \gamma_{k_1} + \rho_{\text{max}}(A_{12}A_{21}) \leq \rho_{\text{max}}^2(A)$$

which immediately shows that $\rho_{\text{max}}(A_{12}A_{21}) \leq \rho_{\text{max}}^2(A) \leq C^2$. The second result in (b) can be proved similarly. ■
G Extra Simulation Studies

G.1 Simulation design 1

Consider two linear regression models

\[ M_1 : Y = \beta_{1,0} + X_1 \beta_{1,1} + X_2 \beta_2 + u_1, \]
\[ M_2 : Y = \beta_{2,0} + X_1 \beta_{2,1} + \sum_{k=1}^{K} X_{3,k} \beta_{3,k} + u_2. \]

The latent DGP (denoted as S1) is

\[ Y = 0.5 + 0.5X_1 + X_2a + \sum_{k=1}^{K} X_{3,k}b + \varepsilon \] (G.1)

where \((X_1, X_2, X_{3,1}, \ldots, X_{3,K}, \varepsilon)\) is a standard normal random vector. Thus,

\[-2f(M_1, F_0) = E_{F_0}[u_1^2] = Kb^2 + 1; \]
\[-2f(M_2, F_0) = E_{F_0}[u_2^2] = a^2 + 1. \] (G.2)

Therefore, the null hypothesis holds if and only if \(a^2 = Kb^2\), and when \(a^2 > Kb^2\), \(f(M_1, F_0) > f(M_2, F_0)\). When \(a^2 = Kb^2 = 0\), \(u_1 = u_2\) and hence \(\omega_{F_0,*}^2 = 0\). Otherwise, \(\omega_{F_0,*}^2 > 0\). There are 31 DGPs considered in this design, which are determined by different combinations of \(a\) and \(b\):

\[ \begin{pmatrix} a \\ b \end{pmatrix}_{j=1,...,31} = \begin{cases} \begin{pmatrix} a_0 + \frac{16-s}{50} \\ b_0 \end{pmatrix}_{s=1,...,15} & \begin{pmatrix} a_0 \\ b_0 + \frac{s}{50} \end{pmatrix}_{s=1,...,15} \end{cases} \] (G.3)

where \(a_0^2 = Kb_0^2\) and we consider two possible values for \(a_0\) in this design: \(a_0 = 0\) or 0.5. The null hypothesis \(H_0\) holds when \((a, b) = (a_0, b_0)\). Model \(M_1\) is better than model \(M_2\) under the first 15 DGPs in (G.3), while model the model \(M_2\) is better under the last 15 DGPs. The finite sample rejection rates of the tests are calculated using 10000 simulated samples.

Figure 6 presents the finite sample rejection rates of the tests when \(K = 2\). When \(a_0 = 0.5\), \(\omega_{F_0,*}^2 > 0\) and our nondegenerate test statistic has asymptotic standard normal distribution regardless the value of \(K\). From graphs (a) and (b), we can see that the finite sample rejection rates of our test (the red solid line) and the parametric test proposed in Shi (2015b) (the blue dashed line) are very close to the nominal level 5% (\(H_0\) holds at \(j = 16\)). When the sample size is 500, both tests have good power which is further improved with larger sample size 1000. When
Figure 6: Finite Sample Rejection Rates of the Tests ($K = 2$)

(a) $a_0 = 0.5$ and $n = 500$

(b) $a_0 = 0.5$ and $n = 1000$

(c) $a_0 = 0$ and $n = 500$

(d) $a_0 = 0$ and $n = 1000$

Legend:
- Nonp_QLR
- Para_QLR
- $\alpha = 0.05$
Figure 7: Finite Sample Rejection Rates of the Tests ($K = 8$)

(a) $a_0 = 0.5$ and $n = 500$

(b) $a_0 = 0.5$ and $n = 1000$

(c) $a_0 = 0$ and $n = 500$

(d) $a_0 = 0$ and $n = 1000$

Legend:
- Nonp_QLR
- Para_QLR
- $\alpha = 0.05$
$a_0 = 0, \omega^2_{F_0,*} = 0$ and the standard normal distribution may not be a good approximation of the finite sample distribution of our test statistic since $K$ is small here. In this case (graphs (c) and (d)), both our test and the parametric test proposed in [Shi (2015b)] under-reject. The finite sample rejection rates of both tests are close to zero under the null. The power of the tests becomes worse for both tests when $a_0 = 0$, although our nonparametric test has better power. Increasing the sample size from 500 to 1000 improves the power of both tests.

The finite sample rejection rates of the tests when $K = 8$ are included in Figure 7. In graphs (a) and (b), we observe similar properties of both tests to the case with small $K$ ($K = 2$), which is expected since in both cases $\omega^2_{F_0,*} > 0$ and our nondegenerate test statistic has asymptotic standard normal distribution. When $a_0 = 0, \omega^2_{F_0,*} = 0$ and both tests are still under-rejection under the null. However the null rejection rates of both tests are closer to the nominal level with larger $K$. In graphs (c) and (d), we see that our nondegenerate test has larger power particularly when model $\mathcal{M}_1$ is better than model $\mathcal{M}_2$ (i.e., $j = 1, \ldots, 15$). The power of these tests are similar when model $\mathcal{M}_2$ is better than model $\mathcal{M}_1$ (i.e., $j = 17, \ldots, 31$). Increasing the sample size from 500 to 1000 improves the power of both tests.

### G.2 Simulation design 2

In the second simulation design, the latent DGP is

$$Y = 0.5 + 0.25 \sum_{k=1}^{K} X_{1,k} + X_2a + X_3b + \varepsilon,$$  \hspace{1cm} (G.4)

where $(X_{1,1}, \ldots, X_{1,K}, X_2, X_3, \varepsilon)$ is a standard normal random vector. There are two linear regression models

\[ \mathcal{M}_1 : \quad Y = \beta_{1,0} + \sum_{k=1}^{K} X_1 \beta_{1,1,k} + X_2 \beta_2 + u_1, \]
\[ \mathcal{M}_2 : \quad Y = \beta_{2,0} + \sum_{k=1}^{K} X_1 \beta_{2,1,k} + X_3 \beta_3 + u_2. \]

Thus

\[-2f(\mathcal{M}_1, F_0) = E_{F_0}[u_1^2] = b^2 + 1; \]
\[-2f(\mathcal{M}_2, F_0) = E_{F_0}[u_2^2] = a^2 + 1. \hspace{1cm} (G.5)\]
Therefore, the null hypothesis holds if and only if $a^2 = b^2$, and when $a^2 > b^2$, $f(M_1, F_0) > f(M_2, F_0)$. When $a^2 = b^2 = 0$, $u_1 = u_2$ and hence $\omega_{F_0,*}^2 = 0$. Otherwise, $\omega_{F_0,*}^2 > 0$. There are 31 DGP’s considered in this design, which are determined by different combinations of $a$ and $b$:

$$
\begin{bmatrix}
  a \\
  b
\end{bmatrix}_{j=1, \ldots, 31} = \begin{cases}
  \left( \begin{array}c
  a_0 + \frac{16-s}{50} \\
  b_0
\end{array} \right)_{s=1, \ldots, 15}, & \left( \begin{array}c
  a_0 \\
  b_0 + s_{50}
\end{array} \right)_{s=1, \ldots, 15}
\end{cases}
$$

(G.6)

where $a_0^2 = b_0^2$ and we consider two possible values for $a_0$ in this design: $a_0 = 0$ or 0.5. The null hypothesis $H_0$ holds when $(a, b) = (a_0, b_0)$. The model $M_1$ is better than the model $M_2$ under the first 15 DGP’s in (G.6), while model the model $M_2$ is better under the last 15 DGP’s. The finite sample rejection rates of the tests are calculated using 10000 simulated samples.

Figure 8 presents the finite sample rejection rates of the tests when $K = 10$. When $a_0 = 0.5$, $\omega_{F_0,*}^2 > 0$ and the null rejection rates of both tests are very close to the nominal level 5% ($H_0$ holds at $j = 16$) and their rejection rates are almost the same under the alternative, which is
similar to what we have observed in the first simulation design. When \( a_0 = 0, \omega^2_{F_0,*} = 0 \) and we see that both tests under-reject under the null. Since the two models \( \mathcal{M}_1 \) and \( \mathcal{M}_2 \) have the same dimensions, the rejections rates of both tests are symmetric around \( j = 16 \). Our nonparametric test has better power than the parametric test proposed in Shi (2015b) in both sample sizes 500 and 1000 (graphs (c) and (d)).

References


