Additional Proofs for:
Crump, R., V. J. Hotz, G. Imbens and O. Mitnik, "Nonparametric Tests for Treatment Effect Heterogeneity"

Final Draft: December 2007

1 Additional Proofs

Proof of Lemma A.1 We will generalize the proof in Imbens, Newey and Ridder (2006). For (i) we will show

$$\mathbb{E}\left[\left\|\hat{\Omega}_{w,K}-\Omega_{w,K}\right\|^2\right]\leq C\cdot\zeta(K)^2K/N$$

so that the result follows by Markov's inequality. Note first that $\mathbb{E}[\hat{\Omega}_{w,K}]=\Omega_{w,K}$ so that,

$$\begin{aligned}
&\mathbb{E}\left[\left\|\hat{\Omega}_{w,K}-\Omega_{w,K}\right\|^2\right]\\
&=\mathbb{E}\left[\operatorname{tr}\left(\hat{\Omega}_{w,K}^2\right)+\operatorname{tr}\left(\Omega_{w,K}^2\right)-2\operatorname{tr}\left(\hat{\Omega}_{w,K}\Omega_{w,K}\right)\right]\\
&=\operatorname{tr}\left(\mathbb{E}\left[\hat{\Omega}_{w,K}^2\right]\right)+\operatorname{tr}\left(\Omega_{w,K}^2\right)-2\operatorname{tr}\left(\mathbb{E}\left[\hat{\Omega}_{w,K}\right]\Omega_{w,K}\right)\\
&=\operatorname{tr}\left(\mathbb{E}\left[\hat{\Omega}_{w,K}^2\right]\right)-\operatorname{tr}\left(\Omega_{w,K}^2\right).
\end{aligned}\tag{B.1}$$

The first term of equation (B.1) is

$$\begin{aligned}
&\operatorname{tr}\left(\mathbb{E}\left[\hat{\Omega}_{w,K}^2\right]\right)\\
&=\frac{1}{N_w^2}\mathbb{E}\left[\sum_{k=1}^K\sum_{l=1}^K\left(\sum_{i=1}^N\mathbf{1}_w\left(W_i\right)R_{kK}(X_i)R_{lK}(X_i)\right)^2\right]\\
&=\frac{1}{N_w^2}\sum_{k=1}^K\sum_{l=1}^K\sum_{i=1}^N\sum_{j=1}^N\mathbb{E}\left[\mathbf{1}_w\left(W_i\right)\mathbf{1}_w\left(W_j\right)R_{kK}(X_i)R_{lK}(X_i)R_{kK}(X_j)R_{lK}(X_j)\right].
\end{aligned}$$

We may partition this expression into terms with $i=j$,

$$\frac{1}{N_w^2}\sum_{k=1}^K\sum_{l=1}^K\sum_{i=1}^N\mathbb{E}\left[\mathbf{1}_w\left(W_i\right)R_{kK}(X_i)^2R_{lK}(X_i)^2\right]\tag{B.2}$$

and with terms $i\neq j$,

$$\frac{1}{N_w^2}\sum_{k=1}^K\sum_{l=1}^K\sum_{i\neq j}\mathbb{E}\left[\mathbf{1}_w\left(W_i\right)R_{kK}(X_i)R_{lK}(X_i)\right]\mathbb{E}\left[\mathbf{1}_w\left(W_j\right)R_{kK}(X_j)R_{lK}(X_j)\right].\tag{B.3}$$

For a random variable U with $E\left|U\right|<\infty$ and an event G with $\Pr\left(\mathcal{G}\right)>0$, then

$$\mathbb{E}\left[U|\,\mathcal{G}\right]=\frac{\mathbb{E}\left[U\cdot 1_{\mathcal{G}}\right]}{\Pr\left(\mathcal{G}\right)}.$$

Using this we may rewrite equations (B.2) and (B.3) as,

$$\operatorname{tr}\left(\mathbb{E}\left[\hat{\Omega}_{w,K}^2\right]\right)=\pi_w\frac{N}{N_w^2}\sum_{k=1}^K\sum_{l=1}^K\mathbb{E}\left[R_{kK}(X)^2R_{lK}(X)^2\middle|\,W=w\right]+\pi_w^2\frac{N\left(N-1\right)}{N_w^2}\operatorname{tr}(\Omega_{w,K}^2).\tag{B.4}$$

To deal with the first term of equation (B.4) consider,

$$\sum_{k=1}^{K}\sum_{l=1}^{K}\mathbb{E}\left[R_{kK}(X)^2R_{lK}(X)^2\middle| W=w\right]=\mathbb{E}\left[\sum_{k=1}^{K}R_{kK}(X)^2\sum_{l=1}^{K}R_{lK}(X)^2\middle| W=w\right]$$

$$\leq \zeta(K)^2\sum_{l=1}^{K}\mathbb{E}\left[R_{lK}(X)^2\middle| W=w\right] \tag{B.5}$$

$$=\zeta(K)^2\,\mathrm{tr}\left(\Omega_{w,K}\right)$$

$$\leq \lambda_{\max}\left(\Omega_{w,K}\right)\cdot\zeta(K)^2K$$

$$\leq C\cdot\zeta(K)^2K. \tag{B.6}$$

Equation (B.5) follows by,

$$\zeta(K)=\sup_x\|R_K(x)\|=\sup_x\left(\sum_{k=1}^{K}R_{kK}^2(x)\right)^{\frac{1}{2}}$$

which then implies that

$$\sum_{k=1}^{K}R_{kK}^2(x)\leq\zeta(K)^2.$$

Equation (B.6) follows since the maximum eigenvalue of $\Omega_{w,K}$ is $O\left(1\right)$ (see below). Thus, the first term of equation (B.4) is

$$\pi_w\frac{N}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\mathbb{E}\left[R_{kK}(X)^2R_{lK}(X)^2\middle| W=w\right]\leq C\cdot\zeta(K)^2KN^{-1}. \tag{B.7}$$

To deal with second term of equation (B.4) and the second term of equation (B.1) we have,

$$\mathrm{tr}(\Omega_{w,K}^2)\leq\lambda_{\max}\left(\Omega_{w,K}^2\right)\cdot K=\lambda_{\max}\left(\Omega_{w,K}\right)^2\cdot K\leq C\cdot K.$$

Thus,

$$\left(\pi_w^2\frac{N\left(N-1\right)}{N_w^2}-1\right)\mathrm{tr}(\Omega_{w,K}^2)=\left(-\frac{1}{N}+o\left(1\right)\right)O\left(K\right)=O\left(KN^{-1}\right). \tag{B.8}$$

Combining the results from equations (B.7) and (B.8) yields,

$$\mathbb{E}\left[\left\|\hat{\Omega}_{w,K}-\Omega_{w,K}\right\|^2\right]=O\left(\zeta(K)^2KN^{-1}\right)+O\left(KN^{-1}\right)=O\left(\zeta(K)^2KN^{-1}\right).$$

For (ii), first note that for any two positive semi-definite matrices A and B, and conformable vectors a and b, if $A\geq B$ in a positive semi-definite sense, then for

$$\lambda_{\min}(A)=\min_{a'a=1}a'Aa=\underline{a}'A\underline{a},\quad \lambda_{\min}(B)=\min_{b'b=1}b'Bb=\underline{b}'B\underline{b},$$

and

$$\lambda_{\max}(A)=\max_{a'a=1}a'Aa=\bar{a}'A\bar{a},\quad \lambda_{\max}(B)=\max_{b'b=1}b'Bb=\bar{b}'B\bar{b},$$

we have that,

$$\lambda_{\min}(A)\geq\lambda_{\min}(B) \tag{B.9}$$

and

$$\lambda_{\max}(A) \geq \lambda_{\max}(B). \tag{B.10}$$

Now, let $f_w(x) = f_{X|W}(x|W = w)$ and define

$$q(x) = f_0(x)/f_1(x)$$

and note that by Assumptions 2.3 and 3.1 we have that

$$0 < \underline{q} \leq q(x) \leq \bar{q} < \infty.$$

Thus we may define $q(x) \equiv \underline{q} + \tilde{q}(x)$ so that,

$$\begin{aligned}
\Omega_{0,K} &= \mathbb{E}\left[R_K(x)R_K(x)'|W = 0\right] \\
&= \int R_K(x)R_K(x)' f_0(x)dx \\
&= \int R_K(x)R_K(x)' q(x) f_1(x)dx \\
&= \int R_K(x)R_K(x)' \left(\underline{q} + \tilde{q}(x)\right) f_1(x)dx \\
&= \underline{q} \int R_K(x)R_K(x)' f_1(x)dx + \int R_K(x)R_K(x)'\tilde{q}(x) f_1(x)dx \\
&= \underline{q} \cdot \Omega_{1,K} + \int R_K(x)R_K(x)'\tilde{q}(x) f_1(x)dx \\
&= \underline{q} \cdot \Omega_{1,K} + \tilde{Q}
\end{aligned}$$

$\tilde{Q}$ is a positive semi-definite matrix, which implies that $\Omega_{0,K} \geq \underline{q} \cdot \Omega_{1,K}$ in a positive semi-definite sense. Thus by equation (B.9)

$$\lambda_{min}\left(\Omega_{0,K}\right) \geq \underline{q} \cdot \lambda_{min}\left(\Omega_{1,K}\right) = \underline{q}$$

and the minimum eigenvalue of $\Omega_{0,K}$ is bounded away from zero. Next, observe that

$$0 \leq \tilde{q}(x) \leq \bar{q} - \underline{q} < \infty$$

and so by the above we have that

$$\begin{aligned}
\Omega_{0,K} &= \underline{q} \cdot \Omega_{1,K} + \int R_K(x)R_K(x)'\tilde{q}(x) f_1(x)dx \\
&\leq \underline{q} \cdot \Omega_{1,K} + \left(\bar{q} - \underline{q}\right) \int R_K(x)R_K(x)' f_1(x)dx \\
&= \bar{q} \cdot \Omega_{1,K},
\end{aligned}$$

in a positive semi-definite sense. Now by equation (B.10) we have

$$\lambda_{\max}\left(\Omega_{0,K}\right) \leq \lambda_{\max}\left(\bar{q} \cdot \Omega_{1,K}\right) = \bar{q}$$

and the maximum eigenvalue of $\Omega_{0,K}$ is bounded. Both the minimum and maximum eigenvalue of $\Omega_{1,K}$ are bounded away from zero and bounded, respectively, by construction. For (iii) consider the minimum eigenvalue of $\hat{\Omega}_{w,K}$,

$$\begin{aligned}
\lambda_{\min}\left(\hat{\Omega}_{w,K}\right) &= \min_{d'd=1} d'\left(\hat{\Omega}_{w,K}\right) d \\
&= \min_{d'd=1}\left(d'\left(\Omega_{w,K}\right) d + d'\left(\hat{\Omega}_{w,K}-\Omega_{w,K}\right) d\right) \\
&\geq \min_{d_1'd_1=1} d_1'\left(\Omega_{w,K}\right) d_1 + \min_{d_2'd_2=1} d_2'\left(\hat{\Omega}_{w,K}-\Omega_{w,K}\right) d_2 \\
&= \lambda_{\min}\left(\Omega_{w,K}\right) + \lambda_{\min}\left(\hat{\Omega}_{w,K}-\Omega_{w,K}\right) \\
&\geq \lambda_{\min}\left(\Omega_{w,K}\right) - \left\|\hat{\Omega}_{w,K}-\Omega_{w,K}\right\| && \text{(B.11)} \\
&= \lambda_{\min}\left(\Omega_{w,K}\right) - O_p\left(\zeta(K)K^{1/2}N^{-1/2}\right) && \text{(B.12)}
\end{aligned}$$

Where (B.11) follows since for a symmetric matrix A

$$\|A\|^2 = \operatorname{tr}\left(A^2\right) \geq \lambda_{\min}(A)^2,$$

and since the norm is nonnegative

$$\|A\| \geq -\lambda_{min}(A)$$

and

$$\|A\| \geq \lambda_{min}(A)$$

for all values of $\lambda_{min}(A)$. Finally, (B.12) follows by part (i). Next, consider the maximum eigenvalue of $\hat{\Omega}_{w,K}$.

$$\begin{aligned}
\lambda_{\max}\left(\hat{\Omega}_{w,K}\right) &= \max_{d'd=1} d'\left(\hat{\Omega}_{w,K}\right) d \\
&= \max_{d'd=1}\left(d'\left(\Omega_{w,K}\right) d + d'\left(\hat{\Omega}_{w,K}-\Omega_{w,K}\right) d\right) \\
&\leq \max_{d_1'd_1=1} d_1'\left(\Omega_{w,K}\right) d_1 + \max_{d_2'd_2=1} d_2'\left(\hat{\Omega}_{w,K}-\Omega_{w,K}\right) d_2 \\
&= \lambda_{\max}\left(\Omega_{w,K}\right) + \lambda_{\max}\left(\hat{\Omega}_{w,K}-\Omega_{w,K}\right) \\
&\leq \lambda_{\max}\left(\Omega_{w,K}\right) + \left\|\hat{\Omega}_{w,K}-\Omega_{w,K}\right\| && \text{(B.13)} \\
&= \lambda_{\max}\left(\Omega_{w,K}\right) + O_p\left(\zeta(K)K^{1/2}N^{-1/2}\right) && \text{(B.14)}
\end{aligned}$$

Where (B.13) follows by similar arguments as above and (B.14) follows by part (i). For (iv) let us first define

$$\tilde{\Sigma}_{w,K} = \frac{R_{w,K}'\tilde{D}_{w,K}R_{w,K}}{N_w}, \qquad \tilde{D}_{w,K} = \operatorname{diag}\left\{\mathbf{1}_w\left(W_i\right)\varepsilon_{w,i}^2; i=1,\ldots,N\right\}.$$

Next recall that for matrices A and B we have that

$$\|A+B\|^2 \leq 2\|A\|^2 + 2\|B\|^2.$$

Thus,

$$\begin{aligned}\mathbb{E}\left\|\hat{\Sigma}_{w,K}-\Sigma_{w,K}\right\|^2 &= \mathbb{E}\left\|\hat{\Sigma}_{w,K}-\tilde{\Sigma}_{w,K}+\tilde{\Sigma}_{w,K}-\Sigma_{w,K}\right\|^2 \\ &\leq 2\cdot\mathbb{E}\left\|\hat{\Sigma}_{w,K}-\tilde{\Sigma}_{w,K}\right\|^2+2\cdot\mathbb{E}\left\|\tilde{\Sigma}_{w,K}-\Sigma_{w,K}\right\|^2 \\ &= 2\cdot\mathbb{E}\left\|\frac{R'_{w,K}\left(\hat{D}_{w,K}-\tilde{D}_{w,K}\right)R_{w,K}}{N_w}\right\|^2 \qquad \text{(B.15)} \\ &\quad +2\cdot\mathbb{E}\left\|\tilde{\Sigma}_{w,K}-\Sigma_{w,K}\right\|^2 \qquad \text{(B.16)}\end{aligned}$$

Before we deal with equations (B.15) and (B.16), we need to establish conditions for consistency of the estimated errors. Note that,

$$\left(\hat{\varepsilon}_{w,i}-\varepsilon_{w,i}\right)(X_i)=\left(Y_i-\hat{\mu}_w(X_i)\right)-\left(Y_i-\mu_w(X_i)\right)=\mu_w(X_i)-\hat{\mu}_w(X_i)$$

and so by Lemma A.6 (v)

$$\sup_x\left|\left(\hat{\varepsilon}_{w,i}-\varepsilon_{w,i}\right)(x)\right|=O_p\left(\zeta(K)^2KN^{-1}\right)+O\left(\zeta(K)K^{-s/d}\right).$$

Moreover,

$$\hat{\varepsilon}^2_{w,i}-\varepsilon^2_{w,i}=2\varepsilon_{w,i}\left(\hat{\varepsilon}_{w,i}-\varepsilon_{w,i}\right)+O_p\left(\zeta(K)^2KN^{-1}\right)+O\left(\zeta(K)K^{-s/d}\right).$$

and so, for $M\in\mathbb{R}_{++}$

$$\begin{aligned}&\Pr\left(\left|\hat{\varepsilon}^2_{w,i}-\varepsilon^2_{w,i}\right|>M\right)\\ &\leq\frac{\mathbb{E}\left|\hat{\varepsilon}^2_{w,i}-\varepsilon^2_{w,i}\right|}{M}\\ &=\frac{\mathbb{E}\left|2\varepsilon_{w,i}\left(\hat{\varepsilon}_{w,i}-\varepsilon_{w,i}\right)+O_p\left(\zeta(K)^2KN^{-1}\right)+O\left(\zeta(K)K^{-s/d}\right)\right|}{M}\\ &\leq\frac{\mathbb{E}\left|2\varepsilon_{w,i}\left(\hat{\varepsilon}_{w,i}-\varepsilon_{w,i}\right)\right|}{M}+\frac{O_p\left(\zeta(K)^2KN^{-1}\right)+O\left(\zeta(K)K^{-s/d}\right)}{M}\\ &\leq C\cdot\sup_x\left|\left(\hat{\varepsilon}_{w,i}-\varepsilon_{w,i}\right)(x)\right|\frac{\mathbb{E}\left|\varepsilon_{w,i}\right|}{M}+O_p\left(\zeta(K)^2KN^{-1}\right)+O\left(\zeta(K)K^{-s/d}\right)\\ &=O_p\left(\zeta(K)^2KN^{-1}\right)+O\left(\zeta(K)K^{-s/d}\right),\end{aligned}$$

and so

$$\hat{\varepsilon}^2_{w,i}-\varepsilon^2_{w,i}=O_p\left(\zeta(K)^2KN^{-1}+\zeta(K)K^{-s/d}\right). \qquad \text{(B.17)}$$

We begin with equation (B.16) first. First note that $E\left[\tilde{\Sigma}_{w,K}\right]=\Sigma_{w,K}$ and so,

$$\begin{aligned}\mathbb{E}\left\|\tilde{\Sigma}_{w,K}-\Sigma_{w,K}\right\|^2&=\mathbb{E}\left[\operatorname{tr}\left(\tilde{\Sigma}^2_{w,K}\right)-2\operatorname{tr}\left(\tilde{\Sigma}_{w,K}\Sigma_{w,K}\right)+\operatorname{tr}\left(\Sigma^2_{w,K}\right)\right]\\ &=\operatorname{tr}\left(\mathbb{E}\left[\tilde{\Sigma}^2_{w,K}\right]\right)-2\operatorname{tr}\left(\mathbb{E}\left[\tilde{\Sigma}_{w,K}\right]\Sigma_{w,K}\right)+\operatorname{tr}\left(\Sigma^2_{w,K}\right)\\ &=\operatorname{tr}\left(\mathbb{E}\left[\tilde{\Sigma}^2_{w,K}\right]\right)-\operatorname{tr}\left(\Sigma^2_{w,K}\right). \qquad \text{(B.18)}\end{aligned}$$

The first term of equation (B.18) is

$$\begin{aligned}
\operatorname{tr}\left(\mathbb{E}\left[\tilde{\Sigma}_{w,K}^{2}\right]\right) &= \frac{1}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left[\mathbf{1}_w\left(W_i\right)\mathbf{1}_w\left(W_j\right)\varepsilon_i^2\varepsilon_j^2 R_{kK}(X_i)R_{lK}(X_i)R_{kK}(X_j)R_{lK}(X_j)\right] \\
&= \frac{1}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{i=1}^{N}\mathbb{E}\left[\mathbf{1}_w\left(W_i\right)\varepsilon_{w,i}^4 R_{kK}(X_i)^2 R_{lK}(X_i)^2\right] \\
&+ \frac{1}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{i\neq j}\mathbb{E}\left[\mathbf{1}_w\left(W_i\right)\varepsilon_{w,i}^2 R_{kK}(X_i)R_{lK}(X_i)\right]\mathbb{E}\left[1_{\{W_j=w\}}\varepsilon_{w,j}^2 R_{kK}(X_j)R_{lK}(X_j)\right] \\
&= \frac{N}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\mathbb{E}\left[\mathbf{1}_w\left(W\right)\varepsilon_w^4 R_{kK}(X)^2 R_{lK}(X)^2\right] \qquad \text{(B.19)} \\
&+ \frac{N\left(N-1\right)}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\left(\mathbb{E}\left[\mathbf{1}_w\left(W\right)\varepsilon_w^2 R_{kK}(X)R_{lK}(X)\right]\right)^2. \qquad \text{(B.20)}
\end{aligned}$$

Equation (B.19) may be rewritten as,

$$\begin{aligned}
&\frac{N}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\mathbb{E}\left[\mathbf{1}_w\left(W\right)\varepsilon_w^4 R_{kK}(X)^2 R_{lK}(X)^2\right] \\
&= \frac{N}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\mathbb{E}\left[\mathbf{1}_w\left(W\right)\mathbb{E}\left[\left.\varepsilon_w^4\right| X,\mathbf{1}_w\left(W\right)\right] R_{kK}(X)^2 R_{lK}(X)^2\right] \\
&= \frac{N}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\mathbb{E}\left[\mathbf{1}_w\left(W\right)\mathbb{E}\left[\left.\varepsilon_w^4\right| X\right] R_{kK}(X)^2 R_{lK}(X)^2\right] \\
&\leq C\cdot\frac{N}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\mathbb{E}\left[\mathbf{1}_w\left(W\right)R_{kK}(X)^2 R_{lK}(X)^2\right] \\
&\leq C\cdot\frac{N}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\mathbb{E}\left[\left.R_{kK}(X)^2 R_{lK}(X)^2\right| W=w\right].
\end{aligned}$$

Thus by equation (B.7) we have that equation (B.19) is,

$$\frac{N}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\mathbb{E}\left[\mathbf{1}_w\left(W\right)\varepsilon_w^4 R_{kK}(X)^2 R_{lK}(X)^2\right] \leq C\cdot\zeta(K)^2 K N^{-1}. \qquad \text{(B.21)}$$

Equation (B.20) and the second term of equation (B.18) are,

$$\begin{aligned}
&\frac{N\left(N-1\right)}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\left(\mathbb{E}\left[\mathbf{1}_w\left(W\right)\varepsilon_w^2 R_{kK}(X)R_{lK}(X)\right]\right)^2 - \operatorname{tr}\left(\Sigma_{w,K}^2\right) \qquad \text{(B.22)} \\
&= \pi_w^2\frac{N\left(N-1\right)}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\left(\mathbb{E}\left[\left.\sigma_w^2\left(X\right)R_{kK}(X)R_{lK}(X)\right| W=w\right]\right)^2 - \operatorname{tr}\left(\Sigma_{w,K}^2\right) \\
&= \left(\pi_w^2\frac{N\left(N-1\right)}{N_w^2}-1\right)\operatorname{tr}\left(\Sigma_{w,K}^2\right).
\end{aligned}$$

The first factor is,

$$\pi_w^2\frac{N\left(N-1\right)}{N_w^2}-1 = -\frac{1}{N}+o\left(1\right).$$

By Assumption 3.2 the second factor is,

$$\operatorname{tr}\left(\Sigma_{w,K}^{2}\right) \leq \lambda_{\max}\left(\Sigma_{w,K}^{2}\right) \cdot K \leq \bar{\sigma}^{4} \cdot \lambda_{\max}\left(\Omega_{w,K}\right)^{2} \cdot K \leq C \cdot K \tag{B.23}$$

and so equation (B.22) is,

$$\frac{N\left(N-1\right)}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\left(\mathbb{E}\left[\mathbf{1}_w\left(W\right)\varepsilon_w^2 R_{kK}(X)R_{lK}(X)\right]\right)^2 - \operatorname{tr}\left(\Sigma_{w,K}^2\right) = O\left(KN^{-1}\right).$$

Thus, by equations (B.21) and (B.23) we have,

$$\mathbb{E}\left\|\tilde{\Sigma}_{w,K}-\Sigma_{w,K}\right\|^2 = O\left(\zeta(K)^2 KN^{-1}\right) + O\left(KN^{-1}\right) = O\left(\zeta(K)^2 KN^{-1}\right). \tag{B.24}$$

Now consider equation (B.15).

$$\begin{aligned}
&\mathbb{E}\left\|\frac{R'_{w,K}\left(\hat{D}_{w,K}-\tilde{D}_{w,K}\right)R_{w,K}}{N_w}\right\|^2 \\
&= \frac{1}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left[\mathbf{1}_w\left(W_i\right)\mathbf{1}_w\left(W_j\right)\left(\hat{\varepsilon}_i^2-\varepsilon_i^2\right)\left(\hat{\varepsilon}_j^2-\varepsilon_j^2\right)R_{kK}(X_i)R_{lK}(X_i)R_{kK}(X_j)R_{lK}(X_j)\right].
\end{aligned}$$

Then, by equation (B.17) we have,

$$\begin{aligned}
&\frac{1}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left[\mathbf{1}_w\left(W_i\right)\mathbf{1}_w\left(W_j\right)\left(\hat{\varepsilon}_i^2-\varepsilon_i^2\right)\left(\hat{\varepsilon}_j^2-\varepsilon_j^2\right)R_{kK}(X_i)R_{lK}(X_i)R_{kK}(X_j)R_{lK}(X_j)\right] \\
&= \left[O_p\left(\zeta\left(K\right)^2 KN^{-1} + \zeta\left(K\right)K^{-s/d}\right)\right]^2 \\
&\qquad \times \frac{1}{N_w^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left[\mathbf{1}_w\left(W_i\right)\mathbf{1}_w\left(W_j\right)R_{kK}(X_i)R_{lK}(X_i)R_{kK}(X_j)R_{lK}(X_j)\right] \\
&= \left[O_p\left(\zeta\left(K\right)^2 KN^{-1}\right) + O\left(\zeta\left(K\right)K^{-s/d}\right)\right]^2 \cdot \operatorname{tr}\left(\mathbb{E}\left[\hat{\Omega}_{w,K}^2\right]\right).
\end{aligned}$$

From the proof of (i), we have that

$$\operatorname{tr}\left(\mathbb{E}\left[\hat{\Omega}_{w,K}^2\right]\right) = O\left(\zeta\left(K\right)^2 KN^{-1}\right) + O\left(KN^{-1}\right) + O\left(K\right) = O\left(K\right),$$

and so we have,

$$\mathbb{E}\left\|\frac{R'_{w,K}\left(D_{w,K}-\tilde{D}_{w,K}\right)R_{w,K}}{N_w}\right\|^2 = \left[O_p\left(\zeta\left(K\right)^2 KN^{-1}\right) + O\left(\zeta\left(K\right)K^{-s/d}\right)\right]^2 O\left(K\right). \tag{B.25}$$

Combining equations (B.24) and (B.25) yields,

$$\begin{aligned}
\mathbb{E}\left\|\hat{\Sigma}_{w,K}-\Sigma_{w,K}\right\|^2 &\leq 2\cdot\mathbb{E}\left\|\tilde{\Sigma}_{w,K}-\Sigma_{w,K}\right\|^2+2\cdot\mathbb{E}\left\|\frac{R'_{w,K}\left(D_{w,K}-\tilde{D}_{w,K}\right)R_{w,K}}{N_w}\right\|^2 \\
&= O\left(\zeta(K)^2KN^{-1}\right)+\left[O_p\left(\zeta\left(K\right)^2KN^{-1}\right)+O\left(\zeta\left(K\right)K^{-s/d}\right)\right]^2 O\left(K\right) \\
&= O\left(\zeta(K)^2KN^{-1}\right) \\
&+\left[O_p\left(\zeta\left(K\right)^4K^2N^{-2}\right)+O\left(\zeta\left(K\right)^2K^{-2s/d}\right)+O_p\left(\zeta\left(K\right)^3KK^{-s/d}N^{-1}\right)\right]O\left(K\right) \\
&= O\left(\zeta(K)^2KN^{-1}\right) \\
&+O_p\left(\zeta\left(K\right)^4K^3N^{-2}\right)+O\left(\zeta\left(K\right)^2KK^{-2s/d}\right)+O_p\left(\zeta\left(K\right)^3K^2K^{-s/d}N^{-1}\right) \\
&= O_p\left(\zeta\left(K\right)^4K^3N^{-2}\right)+O\left(\zeta\left(K\right)^2KK^{-2s/d}\right),
\end{aligned}$$

and the result follows. For (v) note that,

$$\underline{\sigma}^2\cdot\Omega_{w,K}\leq\Sigma_{w,K}\leq\bar{\sigma}^2\cdot\Omega_{w,K}$$

in a positive semi-definite sense. Thus,

$$\lambda_{\min}\left(\Sigma_{w,K}\right)\geq\lambda_{\min}\left(\underline{\sigma}^2\cdot\Omega_{w,K}\right)=\underline{\sigma}^2\cdot\lambda_{\min}\left(\Omega_{w,K}\right)\geq\underline{\sigma}^2\cdot\min\left(\underline{q},1\right).$$

Similarly,

$$\lambda_{\max}\left(\Sigma_{w,K}\right)\leq\lambda_{\max}\left(\bar{\sigma}^2\cdot\Omega_{w,K}\right)\leq\bar{\sigma}^2\cdot\lambda_{\max}\left(\Omega_{w,K}\right)\leq\bar{\sigma}^2\cdot\max\left(\bar{q},1\right).$$

For (vi) consider,

$$\begin{aligned}
\lambda_{\min}\left(\hat{\Sigma}_{w,K}\right) &= \min_{d'd=1}d'\left(\hat{\Sigma}_{w,K}\right)d \\
&= \min_{d'd=1}\left[d'\left(\Sigma_{w,K}\right)d+d'\left(\Sigma_{w,K}-\hat{\Sigma}_{w,K}\right)d\right] \\
&\geq \min_{d_1'd_1=1}d_1'\left(\Sigma_{w,K}\right)d_1+\min_{d_2'd_2=1}d_2'\left(\Sigma_{w,K}-\hat{\Sigma}_{w,K}\right)d_2 \\
&= \lambda_{\min}\left(\Sigma_{w,K}\right)+\lambda_{\min}\left(\Sigma_{w,K}-\hat{\Sigma}_{w,K}\right) \\
&\geq \lambda_{\min}\left(\Sigma_{w,K}\right)-\left\|\Sigma_{w,K}-\hat{\Sigma}_{w,K}\right\| \\
&= \lambda_{\min}\left(\Sigma_{w,K}\right)-O_p\left(\zeta\left(K\right)^2K^{3/2}N^{-1}\right)-O_p\left(\zeta\left(K\right)K^{1/2}K^{-s/d}\right).
\end{aligned}$$

Next,

$$\begin{aligned}
\lambda_{\max}\left(\hat{\Sigma}_{w,K}\right) &= \max_{d'd=1}d'\left(\hat{\Sigma}_{w,K}\right)d \\
&= \max_{d'd=1}\left[d'\left(\Sigma_{w,K}\right)d+d'\left(\Sigma_{w,K}-\hat{\Sigma}_{w,K}\right)d\right] \\
&\leq \max_{d_1'd_1=1}d_1'\left(\Sigma_{w,K}\right)d_1+\max_{d_2'd_2=1}d_2'\left(\Sigma_{w,K}-\hat{\Sigma}_{w,K}\right)d_2 \\
&= \lambda_{\max}\left(\Sigma_{w,K}\right)+\lambda_{\max}\left(\Sigma_{w,K}-\hat{\Sigma}_{w,K}\right) \\
&\leq \lambda_{\max}\left(\Sigma_{w,K}\right)+\left\|\Sigma_{w,K}-\hat{\Sigma}_{w,K}\right\| \\
&= \lambda_{\max}\left(\Sigma_{w,K}\right)+O_p\left(\zeta\left(K\right)^2K^{3/2}N^{-1}\right)+O_p\left(\zeta\left(K\right)K^{1/2}K^{-s/d}\right).
\end{aligned}$$

■

Proof of Lemma A.2 For this proof we need two results. Let A be a symmetric positive definite matrix and B a conformable matrix, then

$$\lambda_{\min}(B'AB) \geq \lambda_{\min}(A) \cdot \lambda_{\min}(B'B), \qquad \lambda_{\max}(B'AB) \leq \lambda_{\max}(A) \cdot \lambda_{\max}(B'B).$$

Using the above result we have,

$$\begin{aligned}
&\lambda_{\min}\left(\Omega_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right) \\
&\geq \lambda_{\min}(\Sigma_{w,K}) \cdot \lambda_{\min}\left(\Omega_{w,K}^{-2}\right) \\
&\geq \underline{\sigma}^2 \cdot \min\left(\underline{q}, 1\right) \cdot \left[\lambda_{\max}\left(\Omega_{w,K}^{2}\right)\right]^{-1} \\
&= \underline{\sigma}^2 \cdot \min\left(\underline{q}, 1\right) \cdot \left[\left(\lambda_{\max}(\Omega_{w,K})\right)^2\right]^{-1} \\
&\geq \underline{\sigma}^2 \cdot \min\left(\underline{q}, 1\right) \cdot \left[\max(\bar{q}, 1)\right]^{-2},
\end{aligned} \tag{B.26}$$

and,

$$\begin{aligned}
&\lambda_{\max}\left(\Omega_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right) \\
&\leq \lambda_{\max}(\Sigma_{w,K}) \cdot \lambda_{\max}\left(\Omega_{w,K}^{-2}\right) \\
&\leq \bar{\sigma}^2 \cdot \max(\bar{q}, 1) \cdot \left[\lambda_{\min}\left(\Omega_{w,K}^{2}\right)\right]^{-1} \\
&= \bar{\sigma}^2 \cdot \max(\bar{q}, 1) \cdot \left[\left(\lambda_{\min}(\Omega_{w,K})\right)^2\right]^{-1} \\
&\leq \bar{\sigma}^2 \cdot \max(\bar{q}, 1) \cdot \left[\min\left(\underline{q}, 1\right)\right]^{-2}.
\end{aligned} \tag{B.27}$$

Now consider,

$$\begin{aligned}
\lambda_{\min}(N \cdot V) &= \min_{d'd=1} d'\left(\frac{1}{\pi_0}\Omega_{0,K}^{-1}\Sigma_{0,K}\Omega_{0,K}^{-1} + \frac{1}{\pi_1}\Omega_{1,K}^{-1}\Sigma_{1,K}\Omega_{1,K}^{-1}\right)d \\
&\geq \frac{1}{\pi_0}\min_{d'd=1} d'\Omega_{0,K}^{-1}\Sigma_{0,K}\Omega_{0,K}^{-1}d + \frac{1}{\pi_1}\min_{d'd=1} d'\left(\Omega_{1,K}^{-1}\Sigma_{1,K}\Omega_{1,K}^{-1}\right)d \\
&= \frac{1}{\pi_0}\lambda_{\min}\left(\Omega_{0,K}^{-1}\Sigma_{0,K}\Omega_{0,K}^{-1}\right) + \frac{1}{\pi_1}\lambda_{\min}\left(\Omega_{1,K}^{-1}\Sigma_{1,K}\Omega_{1,K}^{-1}\right),
\end{aligned}$$

which is bounded away from zero by equation (B.26) and Assumption 2.3. Finally, consider

$$\begin{aligned}
\lambda_{\max}(N \cdot V) &= \max_{d'd=1} d'\left(\frac{1}{\pi_0}\Omega_{0,K}^{-1}\Sigma_{0,K}\Omega_{0,K}^{-1} + \frac{1}{\pi_1}\Omega_{1,K}^{-1}\Sigma_{1,K}\Omega_{1,K}^{-1}\right)d \\
&\leq \frac{1}{\pi_0}\max_{d'd=1} d'\Omega_{0,K}^{-1}\Sigma_{0,K}\Omega_{0,K}^{-1}d + \frac{1}{\pi_1}\max_{d'd=1} d'\left(\Omega_{1,K}^{-1}\Sigma_{1,K}\Omega_{1,K}^{-1}\right)d \\
&= \frac{1}{\pi_0}\lambda_{\max}\left(\Omega_{0,K}^{-1}\Sigma_{0,K}\Omega_{0,K}^{-1}\right) + \frac{1}{\pi_1}\lambda_{\max}\left(\Omega_{1,K}^{-1}\Sigma_{1,K}\Omega_{1,K}^{-1}\right),
\end{aligned}$$

which is bounded by equation (B.27) and Assumption 2.3. For (ii) we have,

$$
\begin{aligned}
&\lambda_{\min}\left(\hat{\Omega}_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}\right)\\
&\geq \lambda_{\min}\left(\hat{\Sigma}_{w,K}\right)\cdot\lambda_{\min}\left(\hat{\Omega}_{w,K}^{-2}\right)\\
&\geq \left[\lambda_{\min}\left(\Sigma_{w,K}\right) - O_p\left(\zeta\left(K\right)^2 K^{3/2}N^{-1}\right) - O_p\left(\zeta\left(K\right)K^{1/2}K^{-s/d}\right)\right]\cdot\lambda_{\min}\left(\hat{\Omega}_{w,K}^{-1}\right)^2\\
&\geq \left[\lambda_{\min}\left(\Sigma_{w,K}\right) - O_p\left(\zeta\left(K\right)^2 K^{3/2}N^{-1}\right) - O_p\left(\zeta\left(K\right)K^{1/2}K^{-s/d}\right)\right]\\
&\quad\cdot\left[\lambda_{\min}\left(\Omega_{w,K}^{-1}\right) - O_p\left(\zeta\left(K\right)K^{1/2}N^{-1/2}\right)\right]^2\\
&= \lambda_{\min}\left(\Sigma_{w,K}\right)\left[\lambda_{\min}\left(\Omega_{w,K}^{-1}\right)\right]^2 + O_p\left(\zeta\left(K\right)^2 K^{3/2}N^{-1}\right) + O_p\left(\zeta\left(K\right)K^{1/2}K^{-s/d}\right).
\end{aligned}
\tag{B.28}
$$

In addition,

$$
\begin{aligned}
&\lambda_{\max}\left(\hat{\Omega}_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}\right)\\
&\leq \lambda_{\max}\left(\hat{\Sigma}_{w,K}\right)\cdot\lambda_{\max}\left(\hat{\Omega}_{w,K}^{-2}\right)\\
&\leq \left[\lambda_{\max}\left(\Sigma_{w,K}\right) + O_p\left(\zeta\left(K\right)^2 K^{3/2}N^{-1}\right) + O_p\left(\zeta\left(K\right)K^{1/2}K^{-s/d}\right)\right]\cdot\lambda_{\max}\left(\hat{\Omega}_{w,K}^{-1}\right)^2\\
&\leq \left[\lambda_{\max}\left(\Sigma_{w,K}\right) + O_p\left(\zeta\left(K\right)^2 K^{3/2}N^{-1}\right) + O_p\left(\zeta\left(K\right)K^{1/2}K^{-s/d}\right)\right]\\
&\quad\cdot\left[\lambda_{\max}\left(\Omega_{w,K}^{-1}\right) + O_p\left(\zeta\left(K\right)K^{1/2}N^{-1/2}\right)\right]^2\\
&= \lambda_{\max}\left(\Sigma_{w,K}\right)\cdot\left[\lambda_{\max}\left(\Omega_{w,K}^{-1}\right)\right]^2 + O_p\left(\zeta\left(K\right)^2 K^{3/2}N^{-1}\right) + O_p\left(\zeta\left(K\right)K^{1/2}K^{-s/d}\right).
\end{aligned}
\tag{B.29}
$$

Thus,

$$
\begin{aligned}
\lambda_{\min}\left(N\cdot\hat{V}\right) &= \min_{d'd=1} d'\left(\frac{N}{N_0}\hat{\Omega}_{0,K}^{-1}\hat{\Sigma}_{0,K}\hat{\Omega}_{0,K}^{-1} + \frac{N}{N_1}\hat{\Omega}_{1,K}^{-1}\hat{\Sigma}_{1,K}\hat{\Omega}_{1,K}^{-1}\right)d\\
&\geq \frac{N}{N_0}\min_{d'd=1} d'\left(\hat{\Omega}_{0,K}^{-1}\hat{\Sigma}_{0,K}\hat{\Omega}_{0,K}^{-1}\right)d + \frac{N}{N_1}\min_{d'd=1} d'\left(\hat{\Omega}_{1,K}^{-1}\hat{\Sigma}_{1,K}\hat{\Omega}_{1,K}^{-1}\right)d\\
&= \frac{N}{N_0}\lambda_{\min}\left(\hat{\Omega}_{0,K}^{-1}\hat{\Sigma}_{0,K}\hat{\Omega}_{0,K}^{-1}\right) + \frac{N}{N_1}\lambda_{\min}\left(\hat{\Omega}_{1,K}^{-1}\hat{\Sigma}_{1,K}\hat{\Omega}_{1,K}^{-1}\right)
\end{aligned}
$$

is bounded away from zero in probability by equation (B.28), Assumption 2.3 and Assumptions 3.2 and 3.3. Finally,

$$
\begin{aligned}
\lambda_{\max}\left(N\cdot\hat{V}\right) &= \max_{d'd=1} d'\left(\frac{N}{N_0}\hat{\Omega}_{0,K}^{-1}\hat{\Sigma}_{0,K}\hat{\Omega}_{0,K}^{-1} + \frac{N}{N_1}\hat{\Omega}_{1,K}^{-1}\hat{\Sigma}_{1,K}\hat{\Omega}_{1,K}^{-1}\right)d\\
&\leq \frac{N}{N_0}\max_{d'd=1} d'\left(\hat{\Omega}_{0,K}^{-1}\hat{\Sigma}_{0,K}\hat{\Omega}_{0,K}^{-1}\right)d + \frac{N}{N_1}\max_{d'd=1} d'\left(\hat{\Omega}_{1,K}^{-1}\hat{\Sigma}_{1,K}\hat{\Omega}_{1,K}^{-1}\right)d\\
&= \frac{N}{N_0}\lambda_{\max}\left(\hat{\Omega}_{0,K}^{-1}\hat{\Sigma}_{0,K}\hat{\Omega}_{0,K}^{-1}\right) + \frac{N}{N_1}\lambda_{\max}\left(\hat{\Omega}_{1,K}^{-1}\hat{\Sigma}_{1,K}\hat{\Omega}_{1,K}^{-1}\right)
\end{aligned}
$$

which is bounded in probability by equation (B.29), Assumption 2.3 and Assumptions 3.2 and 3.3. ■

Proof of Lemma A.6 For (i) note that we would like to provide regressors up to a certain power, say n, including cross product terms. However, we would like to relate the number of covariates, d, and the uppermost power desired, n, to the number of terms in the series, K. We may do so by,

$$
K = \sum_{i=0}^{n}\binom{i+d-1}{i} \leq (n+1)\binom{n+d-1}{n}.
$$

Then note that,

$$\begin{aligned}\binom{n+d-1}{n} &= \frac{(n+d-1)!}{n!\,(d-1)!} \\ &= \frac{(n+d-1)\cdots(n+1)}{(d-1)\cdots 1} \\ &= \prod_{j=1}^{d-1}\left(\frac{n}{j}+1\right) \\ &\leq (n+1)^{d-1}.\end{aligned}$$

Thus, we have that $K = C_1 \cdot (n+1)^d$, or equivalently that $n = C_2 \cdot K^{1/d}$ (and so $n^{-s} = C_2^{-s} \cdot K^{-s/d}$) for some $C_1, C_2 \in R_{++}$. Finally, by Lorentz (1986, Theorem 8) we have that

$$\sup_x \left|\mu_w(x) - R_K(x)' \gamma^0_{w,K}\right| = O\left(n^{-s}\right) = O\left(K^{-s/d}\right),$$

as desired. The proofs of (*ii*) and (*iv*) may be found in Imbens, Newey and Ridder (2006). For (*iii*) consider,

$$\begin{aligned}\left\|\hat{\gamma}_K - \gamma^0_K\right\| &= \left\|\frac{1}{N_w}\hat{\Omega}^{-1}_{w,K} R'_{w,K} Y_w - \frac{1}{N_w}\hat{\Omega}^{-1}_{w,K} R'_{w,K} R_{w,K} \cdot \gamma^0_{w,K}\right\| \\ &= \left\|\frac{1}{N_w}\hat{\Omega}^{-1}_{w,K} R'_{w,K}\left(Y_w - R_{w,K}\cdot\gamma^0_{w,K}\right)\right\| \\ &\leq \lambda_{\max}\left(\hat{\Omega}^{-1/2}_{w,K}\right)\cdot\left\|\frac{1}{N_w}\hat{\Omega}^{-1/2}_{w,K} R'_{w,K}\left(Y_w - R_{w,K}\cdot\gamma^0_{w,K}\right)\right\|.\end{aligned}$$

By Lemma A.1, the first factor is,

$$\lambda_{\max}\left(\hat{\Omega}^{-1/2}_{w,K}\right) = \lambda_{\max}\left(\Omega^{-1/2}_{w,K}\right) + O_p\left(\zeta(K) K^{1/2} N^{-1/2}\right) = O(1) + O_p\left(\zeta(K) K^{1/2} N^{-1/2}\right). \quad \text{(B.30)}$$

The second factor may be broken up into,

$$\begin{aligned}&\left\|\frac{1}{N_w}\hat{\Omega}^{-1/2}_{w,K} R'_{w,K}\left(Y_w - R_{w,K}\cdot\gamma^0_{w,K}\right)\right\| \\ &= \left\|\frac{1}{N_w}\hat{\Omega}^{-1/2}_{w,K} R'_{w,K}\left(Y_w - \mu_w(\boldsymbol{X}) + \mu_w(\boldsymbol{X}) - R_{w,K}\cdot\gamma^0_{w,K}\right)\right\| \\ &\leq \left\|\frac{1}{N_w}\hat{\Omega}^{-1/2}_{w,K} R'_{w,K}\boldsymbol{\varepsilon}_w\right\| \qquad \text{(B.31)} \\ &+ \left\|\frac{1}{N_w}\hat{\Omega}^{-1/2}_{w,K} R'_{w,K}\left(\mu_w(\boldsymbol{X}) - R_{w,K}\cdot\gamma^0_{w,K}\right)\right\|. \qquad \text{(B.32)}\end{aligned}$$

For equation (B.31) we have,

$$\begin{aligned}
&\mathbb{E}\left\|\frac{1}{N_w}\hat{\Omega}_{w,K}^{-1/2}R'_{w,K}\boldsymbol{\varepsilon}_w\right\|^2\\
&=\mathbb{E}\left[\operatorname{tr}\left(\frac{1}{N_w^2}\boldsymbol{\varepsilon}'_w R_{w,K}\hat{\Omega}_{w,K}^{-1}R'_{w,K}\boldsymbol{\varepsilon}_w\right)\right]\\
&=\frac{1}{N_w}\mathbb{E}\left[\operatorname{tr}\left(R_{w,K}\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}\boldsymbol{\varepsilon}_w\boldsymbol{\varepsilon}'_w\right)\right]\\
&=\frac{1}{N_w}\operatorname{tr}\left(\mathbb{E}\left[R_{w,K}\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}\mathbb{E}\left[\boldsymbol{\varepsilon}_w\boldsymbol{\varepsilon}'_w\right|\boldsymbol{X}]\right]\right)\\
&\leq\bar{\sigma}^2\cdot\frac{1}{N_w}\mathbb{E}\left[\operatorname{tr}\left(R_{w,K}\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}\right)\right]\\
&=\bar{\sigma}^2\cdot\frac{1}{N_w}K\\
&\leq C\cdot KN^{-1},
\end{aligned}$$

and so

$$\left\|\frac{1}{N_w}\hat{\Omega}_{w,K}^{-1/2}R'_{w,K}\boldsymbol{\varepsilon}_w\right\|=O_p\left(K^{1/2}N^{-1/2}\right) \tag{B.33}$$

by Markov's inequality. For equation (B.32) we have,

$$\begin{aligned}
&\left\|\frac{1}{N_w}\hat{\Omega}_{w,K}^{-1/2}R'_{w,K}\left(\mu_w\left(\boldsymbol{X}\right)-R_{w,K}\cdot\gamma_{w,K}^0\right)\right\|^2\\
&=\frac{1}{N_w}\left(\mu_w\left(\boldsymbol{X}\right)-R_{w,K}\cdot\gamma_{w,K}^0\right)'R_{w,K}\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}\left(\mu_w\left(\boldsymbol{X}\right)-R_{w,K}\cdot\gamma_{w,K}^0\right)\\
&\leq\frac{1}{N_w}\left(\mu_w\left(\boldsymbol{X}\right)-R_{w,K}\cdot\gamma_{w,K}^0\right)'\left(\mu_w\left(\boldsymbol{X}\right)-R_{w,K}\cdot\gamma_{w,K}^0\right)\\
&\leq\sup_x\left|\mu_w\left(x\right)-R_K\left(x\right)'\gamma_K^0\right|^2\\
&=O\left(K^{-2s/d}\right),
\end{aligned}$$

by (i), and so

$$\left\|\frac{1}{N_w}\hat{\Omega}_{w,K}^{-1/2}R'_{w,K}\left(\mu_w\left(\boldsymbol{X}\right)-R_{w,K}\cdot\gamma_{w,K}^0\right)\right\|=O\left(K^{-s/d}\right) \tag{B.34}$$

Note that the third line of the penultimate display follows since $R_{w,K}\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}$ is a projection matrix. Combining equations (B.30), (B.33) and (B.34) yields,

$$\begin{aligned}
\left\|\hat{\gamma}_K-\gamma_K^0\right\|&=\left[O\left(1\right)+O_p\left(\zeta\left(K\right)K^{1/2}N^{-1/2}\right)\right]\left[O_p\left(K^{1/2}N^{-1/2}\right)+O\left(K^{-s/d}\right)\right]\\
&=O_p\left(K^{1/2}N^{-1/2}\right)+O\left(K^{-s/d}\right)+O_p\left(\zeta\left(K\right)KN^{-1}\right)+O_p\left(\zeta\left(K\right)K^{1/2}K^{-s/d}N^{-1/2}\right)\\
&=O\left(K^{-s/d}\right)+O_p\left(\zeta\left(K\right)KN^{-1}\right)+O_p\left(\zeta\left(K\right)K^{1/2}K^{-s/d}N^{-1/2}\right).
\end{aligned}$$

However, by Assumption 3.3, we have that,

$$O_p\left(\zeta\left(K\right)K^{1/2}K^{-s/d}N^{-1/2}\right)=o_p\left(K^{-s/d}\right)=o_p\left(1\right)$$

and so

$$\left\|\hat{\gamma}_K-\gamma_K^0\right\|=O_p\left(\zeta\left(K\right)KN^{-1}\right)+O\left(K^{-s/d}\right),$$

as desired. Finally, for (v) we have,

$$\sup_x |\mu_w(x) - \hat{\mu}_{w,K}(x)| \leq \sup_x \left|\mu_w(x) - \mu^0_{w,K}(x)\right| + \sup_x \left|\mu^0_w(x) - \hat{\mu}_{w,K}(x)\right|.$$

The first term is $O\left(K^{-s/d}\right)$ by (i). For the second term we have,

$$\begin{aligned}
&\sup_x \left|\mu^0_w(x) - \hat{\mu}_{w,K}(x)\right| \\
&= \sup_x \left|R_K(x)'\left(\gamma^0_{w,K} - \hat{\gamma}_{w,K}\right)\right| \\
&\leq \sup_x \|R_K(x)\| \left\|\gamma^0_{w,K} - \hat{\gamma}_{w,K}\right\| \\
&= \zeta(K) \cdot \left[O_p\left(\zeta(K) K N^{-1}\right) + O\left(K^{-s/d}\right)\right] \\
&= O_p\left(\zeta(K)^2 K N^{-1}\right) + O\left(\zeta(K) K^{-s/d}\right),
\end{aligned}$$

where we use the definition of $\zeta(K)$ and the result from (iv). Thus,

$$\begin{aligned}
\sup_x |\mu_w(x) - \hat{\mu}_{w,K}(x)| &= O\left(K^{-s/d}\right) + O_p\left(\zeta(K)^2 K N^{-1}\right) + O\left(\zeta(K) K^{-s/d}\right) \\
&= O_p\left(\zeta(K)^2 K N^{-1}\right) + O\left(\zeta(K) K^{-s/d}\right),
\end{aligned}$$

as desired. ■

Proof of Lemma A.7 (Additional Details) Equation (A.4) is

$$\begin{aligned}
&N_w^{-1/2}\left\|\left[\hat{\Omega}^{-1}_{w,K}\hat{\Sigma}_{w,K}\hat{\Omega}^{-1}_{w,K}\right]^{-1/2}\hat{\Omega}^{-1}_{w,K}R'_{w,K}\boldsymbol{\varepsilon}^*_{w,K} - \left[\hat{\Omega}^{-1}_{w,K}\hat{\Sigma}_{w,K}\hat{\Omega}^{-1}_{w,K}\right]^{-1/2}\hat{\Omega}^{-1}_{w,K}R'_{w,K}\cdot\boldsymbol{\varepsilon}_w\right\| \\
&\leq \left\|\left[\hat{\Sigma}_{w,K}\hat{\Omega}^{-1}_{w,K}\right]^{-1/2}\right\| \left\|\hat{\Omega}^{-1/2}_{w,K}R'_{w,K}\left(\boldsymbol{\varepsilon}^*_{w,K} - \boldsymbol{\varepsilon}_w\right)\Big/ N_w^{1/2}\right\|
\end{aligned}$$

The first factor is,

$$\begin{aligned}
&\left\|\left[\hat{\Sigma}_{w,K}\hat{\Omega}^{-1}_{w,K}\right]^{-1/2}\right\| \\
&= \left\|\hat{\Omega}^{1/2}_{w,K}\hat{\Sigma}^{-1/2}_{w,K}\right\| \\
&\leq \lambda_{\max}\left(\hat{\Sigma}^{-1/2}_{w,K}\right)\left\|\hat{\Omega}^{1/2}_{w,K}\right\| \\
&\leq \lambda_{\max}\left(\hat{\Sigma}^{-1/2}_{w,K}\right)\lambda_{\max}\left(\hat{\Omega}^{1/2}_{w,K}\right)K^{1/2} \\
&= \left[\lambda_{\max}\left(\Sigma^{-1/2}_{w,K}\right) + O_p(\zeta(K)^2 K^{3/2} N^{-1}) + O_p\left(\zeta(K) K^{1/2} K^{-s/d}\right)\right] \\
&\qquad \times \left[\lambda_{\max}\left(\Omega^{1/2}_{w,K}\right) + O_p\left(\zeta(K) K^{1/2} N^{-1/2}\right)\right] K^{1/2} \\
&= \lambda_{\max}\left(\Sigma^{-1/2}_{w,K}\right)\lambda_{\max}\left(\Omega^{1/2}_{w,K}\right)K^{1/2} + O_p(\zeta(K)^2 K^2 N^{-1}) + O\left(\zeta(K) K K^{-s/d}\right) \\
&= O\left(K^{1/2}\right) + O_p(\zeta(K)^2 K^2 N^{-1}) + O_p\left(\zeta(K) K K^{-s/d}\right).
\end{aligned}$$

For the second factor we have,

$$
\begin{aligned}
&\mathbb{E}\left\|\hat{\Omega}_{w,K}^{-\frac{1}{2}}R'_{w,K}\left(\boldsymbol{\varepsilon}^*_{w,K}-\boldsymbol{\varepsilon}_w\right)/\sqrt{N_w}\right\|^2\\
&=\mathbb{E}\left[\frac{1}{N_w}\mathrm{tr}\left(\left(\boldsymbol{\varepsilon}^*_{w,K}-\boldsymbol{\varepsilon}_w\right)'R_{w,K}\hat{\Omega}_{w,K}^{-1}R'_{w,K}\left(\boldsymbol{\varepsilon}^*_{w,K}-\boldsymbol{\varepsilon}_w\right)\right)\right]\\
&=\mathbb{E}\left[\left(\left(\boldsymbol{\varepsilon}^*_{w,K}-\boldsymbol{\varepsilon}_w\right)'R_{w,K}\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}\left(\boldsymbol{\varepsilon}^*_{w,K}-\boldsymbol{\varepsilon}_w\right)\right)\right]\\
&\leq\mathbb{E}\left[\left(\boldsymbol{\varepsilon}^*_{w,K}-\boldsymbol{\varepsilon}_w\right)'\left(\boldsymbol{\varepsilon}^*_{w,K}-\boldsymbol{\varepsilon}_w\right)\right] &&\text{(B.35)}\\
&=\mathbb{E}\left[\left(\mu_w(\mathbf{X})-R_{w,K}\gamma^*_{w,K}\right)'\left(\mu_w(\mathbf{X})-R_{w,K}\gamma^*_{w,K}\right)\right]\\
&\leq N_w\cdot\sup_x\left|\mu_w(x)-R_K(x)'\gamma^*_{w,K}\right|^2\\
&\leq N_w\cdot\sup_x\left(\left|\mu_w(x)-R_K(x)'\gamma^0_{w,K}\right|+\left|R_K(x)'\gamma^0_{w,K}-R_K(x)'\gamma^*_{w,K}\right|\right)^2\\
&=N_w\left(O\left(K^{-\frac{s}{d}}\right)+O\left(\zeta(K)K^{\frac{1}{2}}K^{-\frac{s}{d}}\right)\right)^2 &&\text{(B.36)}\\
&=O\left(N\right)\cdot\left(O\left(\zeta(K)K^{\frac{1}{2}}K^{-\frac{s}{d}}\right)\right)^2
\end{aligned}
$$

Thus,

$$
\left\|\hat{\Omega}_{w,K}^{-1/2}R'_{w,K}\left(\boldsymbol{\varepsilon}^*_{w,K}-\boldsymbol{\varepsilon}_w\right)\Big/N_w^{1/2}\right\|=O_p\left(\zeta(K)K^{1/2}K^{-s/d}N^{1/2}\right)
$$

by Markov's inequality. Equation (B.35) follows by the fact that $R_{w,K}(R'_{w,K}R_{w,K})^{-1}R'_{w,K}$ is a projection matrix and equation (B.36) follows from Lemma A.6 (i) and (iv). Then equation (A.4) is

$$
\begin{aligned}
&N_w^{-1/2}\left\|\left[\hat{\Omega}_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}\right]^{-1/2}\hat{\Omega}_{w,K}^{-1}R'_{w,K}\boldsymbol{\varepsilon}^*_{w,K}-\left[\hat{\Omega}_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}\right]^{-1/2}\hat{\Omega}_{w,K}^{-1}R'_{w,K}\cdot\boldsymbol{\varepsilon}_w\right\|\\
&=\left[O\left(K^{1/2}\right)+O_p(\zeta\left(K\right)^2K^2N^{-1})+O_p\left(\zeta\left(K\right)KK^{-s/d}\right)\right]\left[O_p\left(\zeta(K)K^{1/2}K^{-s/d}N^{1/2}\right)\right]\\
&=O_p\left(\zeta(K)KK^{-s/d}N^{1/2}\right)+O_p\left(\zeta(K)^3K^{5/2}K^{-s/d}N^{-1/2}\right)+O_p\left(\zeta(K)^2K^{3/2}K^{-2s/d}N^{1/2}\right)\\
&=O_p\left(\zeta(K)^2K^{3/2}K^{-2s/d}N^{1/2}\right),
\end{aligned}
$$

by Assumption 3.3. Next, equation (A.5) is

$$
\begin{aligned}
&N_w^{-1/2}\left\|\left[\hat{\Omega}_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}\right]^{-1/2}\hat{\Omega}_{w,K}^{-1}R'_{w,K}\cdot\boldsymbol{\varepsilon}_w-\left[\hat{\Omega}_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right]^{-1/2}\hat{\Omega}_{w,K}^{-1}R'_{w,K}\cdot\boldsymbol{\varepsilon}_w\right\|\\
&\leq\left\|\left[\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}\right]^{-1/2}-\left[\Sigma_{w,K}\Omega_{w,K}^{-1}\right]^{-1/2}\right\|\left\|N_w^{-1/2}\hat{\Omega}_{w,K}^{-1/2}R'_{w,K}\cdot\varepsilon_w\right\|
\end{aligned}
$$

The first factor is

$$
\left\|\left[\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}\right]^{-1/2}-\left[\Sigma_{w,K}\Omega_{w,K}^{-1}\right]^{-1/2}\right\|=O_p\left(\zeta\left(K\right)^2K^2N^{-1}\right)+O_p\left(\zeta\left(K\right)KN^{-1/2}\right)+O_p\left(\zeta\left(K\right)KK^{-s/d}\right)
$$

by Lemma B.2. The second factor is,

$$
\begin{aligned}
&\mathbb{E}\left\|\hat{\Omega}_{w,K}^{-\frac{1}{2}}\cdot R'_{w,K}\boldsymbol{\varepsilon}_w/\sqrt{N_w}\right\|^2\\
&=\mathbb{E}\left[\frac{1}{N_w}\text{tr}\left(\boldsymbol{\varepsilon}'_w R_{w,K}\hat{\Omega}_{w,K}^{-1}R'_{w,K}\boldsymbol{\varepsilon}_w\right)\right]\\
&=\mathbb{E}\left[\text{tr}\left(\boldsymbol{\varepsilon}'_w R_{w,K}\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}\boldsymbol{\varepsilon}_w\right)\right]\\
&=\mathbb{E}\left[\text{tr}\left(R_{w,K}\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}\boldsymbol{\varepsilon}_w\boldsymbol{\varepsilon}'_w\right)\right]\\
&=\text{tr}\left(\mathbb{E}\left[R_{w,K}\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}\mathbb{E}\left[\boldsymbol{\varepsilon}_w\boldsymbol{\varepsilon}'|\mathbf{X}\right]\right]\right)\\
&\leq\bar{\sigma}_w^2\cdot\text{tr}\left(\mathbb{E}\left[R_{w,K}\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}\right]\right)\\
&=\bar{\sigma}_w^2\cdot\mathbb{E}\left[\text{tr}\left(R_{w,K}\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}\right)\right]\\
&=\bar{\sigma}_w^2\cdot\mathbb{E}\left[\text{tr}\left(\left(R'_{w,K}R_{w,K}\right)^{-1}R'_{w,K}R_{w,K}\right)\right]\\
&=\bar{\sigma}_w^2\cdot\text{tr}\left(I_K\right)\\
&=\bar{\sigma}_w^2\cdot K.
\end{aligned}
$$

Thus, the the second factor is $O_p\left(K^{1/2}\right)$ by Markov's inequality. Putting this together yields,

$$
\begin{aligned}
&N_w^{-1/2}\left\|\left[\hat{\Omega}_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}\right]^{-1/2}\hat{\Omega}_{w,K}^{-1}R'_{w,K}\cdot\boldsymbol{\varepsilon}_w-\left[\hat{\Omega}_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right]^{-1/2}\hat{\Omega}_{w,K}^{-1}R'_{w,K}\cdot\boldsymbol{\varepsilon}_w\right\|\\
&=\left[O_p\left(\zeta\left(K\right)^2K^2N^{-1}\right)+O_p\left(\zeta\left(K\right)KN^{-1/2}\right)+O_p\left(\zeta\left(K\right)KK^{-s/d}\right)\right]O_p\left(K^{1/2}\right)\\
&=O_p\left(\zeta\left(K\right)K^{3/2}N^{-1/2}\right)+O_p\left(\zeta\left(K\right)K^{3/2}K^{-s/d}\right).
\end{aligned}
$$

Finally, equation (A.6) is

$$
\begin{aligned}
&N_w^{-1/2}\left\|\left[\hat{\Omega}_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right]^{-1/2}\hat{\Omega}_{w,K}^{-1}R'_{w,K}\cdot\boldsymbol{\varepsilon}_w-\left[\Omega_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right]^{-1/2}\Omega_{w,K}^{-1}R'_{w,K}\cdot\boldsymbol{\varepsilon}_w\right\|\\
&\leq\left\|\left[\Sigma_{w,K}\Omega_{w,K}^{-1}\right]^{-1/2}\right\|\left\|\hat{\Omega}_{w,K}^{-1/2}-\Omega_{w,K}^{-1/2}\right\|\left\|N_w^{-1/2}R'_{w,K}\cdot\boldsymbol{\varepsilon}_w\right\|
\end{aligned}
$$

The first factor is,

$$
\left\|\left[\Sigma_{w,K}\Omega_{w,K}^{-1}\right]^{-1/2}\right\|\leq C\cdot\|I\|=O\left(K^{1/2}\right).
$$

The second factor is $O_p\left(\zeta(K) K^{1/2} N^{-1/2}\right)$ by Lemma A.1. For the third factor consider,

$$
\begin{aligned}
&\mathbb{E}\left\|R_{w,K}'\boldsymbol{\varepsilon}_w/\sqrt{N_w}\right\|^2 \\
&= \mathbb{E}\left[\frac{1}{N_w}\operatorname{tr}\left(\boldsymbol{\varepsilon}_w' R_{w,K} R_{w,K}'\boldsymbol{\varepsilon}_w\right)\right] \\
&= \mathbb{E}\left[\frac{1}{N_w}\operatorname{tr}\left(R_{w,K}'\boldsymbol{\varepsilon}_w\boldsymbol{\varepsilon}_w' R_{w,K}\right)\right] \\
&= \operatorname{tr}\left(\frac{1}{N_w}\mathbb{E}\left[R_{w,K}'\mathbb{E}\left[\boldsymbol{\varepsilon}_w\boldsymbol{\varepsilon}_w'|\mathbf{X}\right] R_{w,K}\right]\right) \\
&\leq \bar{\sigma}_w^2 \cdot \operatorname{tr}\left(\mathbb{E}\left[R_{w,K}' R_{w,K}/N_w\right]\right) \\
&= \bar{\sigma}_w^2 \cdot \operatorname{tr}\left(\Omega_{w,K}\right) \\
&\leq \bar{\sigma}_w^2 \cdot K \cdot \lambda_{max}\left(\Omega_{w,K}\right) \\
&\leq C \cdot K
\end{aligned}
$$

Thus, the third factor is $O_p\left(K^{1/2}\right)$ by Assumption 3.2, Lemma A.1 (ii) and Markov's inequality. Putting this together yields,

$$
\begin{aligned}
&N_w^{-1/2}\left\|\left[\hat{\Omega}_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right]^{-1/2}\hat{\Omega}_{w,K}^{-1}R_{w,K}' \cdot \boldsymbol{\varepsilon}_w - \left[\Omega_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right]^{-1/2}\Omega_{w,K}^{-1}R_{w,K}' \cdot \boldsymbol{\varepsilon}_w\right\| \\
&= O_p\left(\zeta(K) K^{3/2} N^{-1/2}\right).
\end{aligned}
$$

■

Before proving Theorem 3/3. we need the following lemma.

Lemma B.1 *We may partition $\hat{V}$ and V analogously to the partition of $\hat{V}_P$ and V_P in Section 3.4.*

$$
\hat{V} = \begin{pmatrix} \hat{V}_{00} & \hat{V}_{01} \\ \hat{V}_{10} & \hat{V}_{11} \end{pmatrix}
$$

and

$$
V = \begin{pmatrix} V_{00} & V_{01} \\ V_{10} & V_{11} \end{pmatrix}
$$

where $\hat{V}_{00}$ and V_{00} are scalars, $\hat{V}_{01}$ and V_{01} are $1 \times (K-1)$ vectors, $\hat{V}_{10}$ and V_{10} are $(K-1) \times 1$ vectors and $\hat{V}_{11}$ and V_{11} are $(K-1) \times (K-1)$ matrices. Then,

$$
\lambda_{\min}\left([N \cdot V]^{-1}\right) \leq \lambda_{\min}\left([N \cdot V_{11}]^{-1}\right), \qquad \lambda_{\max}\left([N \cdot V]^{-1}\right) \geq \lambda_{\max}\left([N \cdot V_{11}]^{-1}\right)
$$

and if $O_p\left(\zeta(K)^2 K^{3/2} N^{-1}\right) + O_p\left(\zeta(K) K^{1/2} K^{-s/d}\right) = o_p(1)$,

$$
\lambda_{\min}\left(\left[N \cdot \hat{V}\right]^{-1}\right) \leq \lambda_{\min}\left(\left[N \cdot \hat{V}_{11}\right]^{-1}\right), \qquad \lambda_{\max}\left(\left[N \cdot \hat{V}\right]^{-1}\right) \geq \lambda_{\max}\left(\left[N \cdot \hat{V}_{11}\right]^{-1}\right)
$$

with probability approaching one.

Proof The proof follows by the interlacing theorem, (see, for example, Li and Mathias (2002)): *If A is an $n \times n$ positive semi-definite Hermitian matrix with eigenvalues $\lambda_1 \geq ... \geq \lambda_n$, B is a $k \times k$ principal submatrix of A with eigenvalues $\tilde{\lambda}_1 \geq ... \geq \tilde{\lambda}_k$, then*

$$
\lambda_i \geq \tilde{\lambda}_i \geq \lambda_{i+n-k}, \quad i = 1, ..., k.
$$

In our case, $N \cdot \hat{V}$ and $N \cdot V$ are positive semi-definite, symmetric and thus positive semi-definite, Hermitian. So then, by the interlacing theorem

$$\lambda_{\min}(N \cdot V) \leq \lambda_{\min}(N \cdot V_{11}) \Longrightarrow \lambda_{\max}\left([N \cdot V]^{-1}\right) \geq \lambda_{\max}\left([N \cdot V_{11}]^{-1}\right)$$

and

$$\lambda_{\max}(N \cdot V) \geq \lambda_{\max}(N \cdot V_{11}) \Longrightarrow \lambda_{\min}\left([N \cdot V]^{-1}\right) \leq \lambda_{\min}\left([N \cdot V_{11}]^{-1}\right).$$

Moreover, by Lemma A.2, $N \cdot \hat{V}$ is nonsingular with probability approaching one so that again by the interlacing theorem we obtain

$$\lambda_{\min}\left(N \cdot \hat{V}\right) \leq \lambda_{\min}\left(N \cdot \hat{V}_{11}\right) \Longrightarrow \lambda_{\max}\left(\left[N \cdot \hat{V}\right]^{-1}\right) \geq \lambda_{\max}\left(\left[N \cdot \hat{V}_{11}\right]^{-1}\right)$$

and

$$\lambda_{\max}\left(N \cdot \hat{V}\right) \geq \lambda_{\max}\left(N \cdot \hat{V}_{11}\right) \Longrightarrow \lambda_{\min}\left(\left[N \cdot \hat{V}\right]^{-1}\right) \leq \lambda_{\min}\left(\left[N \cdot \hat{V}_{11}\right]^{-1}\right).$$

■

Proof of Theorem 3.3 When the conditional average treatment effect is constant we may choose the two approximating sequences, $\gamma_{0,K}^{0}$ and $\gamma_{1,K}^{0}$, to differ only by way of the first element (the coefficient of the constant term in the approximating sequence). To simplify notation, define

$$\hat{\delta}_{1K} = \hat{\gamma}_{11,K} - \hat{\gamma}_{01,K}, \qquad \delta_{1K}^{*} = \gamma_{11,K}^{*} - \gamma_{01,K}^{*}.$$

We may again follow the logic of Lemmas (A.3), (A.4), and (A.5) to conclude that

$$T^{*\prime} \equiv \left(\left(\hat{\delta}_{1K} - \delta_{1K}^{*}\right)' \cdot \hat{V}_{11}^{-1} \cdot \left(\hat{\delta}_{1K} - \delta_{1K}^{*}\right) - (K-1)\right) \Big/ \sqrt{2(K-1)} \xrightarrow{d} \mathcal{N}(0,1)$$

We need only show that $|T^{*\prime} - T'| = o_p(1)$ to complete the proof. First, note that

$$\begin{aligned} \left\|\gamma_{w1,K}^{*} - \gamma_{w1,K}^{0}\right\|^{2} &= \sum_{i=2}^{K}\left(\gamma_{w1,K,i}^{*} - \gamma_{w1,K,i}^{0}\right)^{2} \\ &\leq \sum_{i=2}^{K}\left(\gamma_{w1,K,i}^{*} - \gamma_{w1,K,i}^{0}\right)^{2} + \left(\gamma_{w0,K}^{*} - \gamma_{w0,K}^{0}\right)^{2} \\ &= \left\|\gamma_{w,K}^{*} - \gamma_{w,K}^{0}\right\|^{2} \\ &= O\left(KK^{-2s/d}\right), \end{aligned} \tag{B.37}$$

by Lemma A.6 (*ii*) and

$$\begin{aligned} \left\|\hat{\gamma}_{w1,K} - \gamma_{w1,K}^{0}\right\|^{2} &= \sum_{i=2}^{K}\left(\hat{\gamma}_{w1,K,i} - \gamma_{w1,K,i}^{0}\right)^{2} \\ &\leq \sum_{i=2}^{K}\left(\hat{\gamma}_{w1,K,i} - \gamma_{w1,K,i}^{0}\right)^{2} + \left(\hat{\gamma}_{w0,K} - \gamma_{w0,K}^{0}\right)^{2} \\ &= \left\|\hat{\gamma}_{w,K} - \gamma_{w,K}^{0}\right\|^{2} \\ &= \left[O_p\left(\zeta(K)KN^{-1}\right) + O\left(K^{-s/d}\right)\right]^{2}. \end{aligned} \tag{B.38}$$

by Lemma A.6 (*iii*). We may choose the last $(K-1)$ elements of the approximating sequence to be equal, $\gamma^0_{11,K} = \gamma^0_{01,K}$. This allows us to bound,

$$\begin{aligned}
\left\|\hat{\delta}_{1K}\right\| &= \left\|\hat{\gamma}_{11,K} - \hat{\gamma}_{01,K}\right\| \\
&= \left\|\hat{\gamma}_{11,K} - \gamma^0_{11,K} + \gamma^0_{01,K} - \hat{\gamma}_{01,K}\right\| \\
&\leq \left\|\hat{\gamma}_{11,K} - \gamma^0_{11,K}\right\| + \left\|\gamma^0_{01,K} - \hat{\gamma}_{01,K}\right\| \\
&= O_p\left(\zeta(K) K N^{-1}\right) + O\left(K^{-s/d}\right) \qquad \text{(B.39)}
\end{aligned}$$

by equation (B.38). Also,

$$\begin{aligned}
\left\|\delta^*_{1K}\right\| &= \left\|\gamma^*_{11,K} - \gamma^*_{01,K}\right\| \\
&= \left\|\gamma^*_{11,K} - \gamma^0_{11,K} + \gamma^0_{01,K} - \gamma^*_{01,K}\right\| \\
&\leq \left\|\gamma^*_{11,K} - \gamma^0_{11,K}\right\| + \left\|\gamma^0_{01,K} - \gamma^*_{01,K}\right\| \\
&= O\left(K^{1/2} K^{-s/d}\right) \qquad \text{(B.40)}
\end{aligned}$$

by equation (B.37). Next note that,

$$\begin{aligned}
\left|T^{*\prime} - T'\right| &= \left(\left(\hat{\delta}_{1K} - \delta^*_{1K}\right)' \hat{V}_{11}^{-1}\left(\hat{\delta}_{1K} - \delta^*_{1K}\right) - \hat{\delta}'_{1K}\hat{V}_{11}^{-1}\hat{\delta}_{1K}\right) \Big/ \sqrt{2(K-1)} \\
&= \left(\delta^*_{1K}{}'\hat{V}_{11}^{-1}\delta^*_{1K} - 2\cdot\hat{\delta}'_{1K}\hat{V}_{11}^{-1}\delta^*_{1K}\right) \Big/ \sqrt{2(K-1)} \qquad \text{(B.41)}
\end{aligned}$$

First consider,

$$\begin{aligned}
& \left|\delta^*_{1K}{}'\hat{V}_{11}^{-1}\delta^*_{1K}\right| \\
=\ & \left|\mathrm{tr}\left(\delta^*_{1K}{}'\hat{V}_{11}^{-1}\delta^*_{1K}\right)\right| \\
=\ & N\cdot\left|\mathrm{tr}\left(\delta^*_{1K}{}'\left[N\cdot\hat{V}_{11}\right]^{-1}\delta^*_{1K}\right)\right| \\
\leq\ & N\cdot\lambda_{\max}\left(\left[N\cdot\hat{V}_{11}\right]^{-1}\right)\cdot\left\|\delta^*_{1K}\right\|^2 \\
\leq\ & N\cdot\lambda_{\max}(\left(\left[N\cdot\hat{V}\right]^{-1}\right)\cdot\left\|\delta^*_{1K}\right\|^2 \qquad \text{(B.42)} \\
=\ & N\cdot\left[O(1) + O_p\left(\zeta(K) K^{3/2} N^{-1/2}\right) + O_p\left(\zeta(K) K^{3/2} K^{-s/d}\right)\right] O\left(K K^{-2s/d}\right) \\
=\ & O_p\left(\zeta(K) K^{5/2} K^{-2s/d} N^{1/2}\right), \qquad \text{(B.43)}
\end{aligned}$$

where equation (B.42) follows by Lemma B.1 and equation (B.43) follows from Lemma B.2 and the proof of Theorem 3.1. Now consider,

$$
\begin{aligned}
& \left|\hat{\delta}_{1K}' \hat{V}_{11}^{-1} \delta_{1K}^{*}\right| \\
= & \left|\operatorname{tr}\left(\hat{\delta}_{1K}' \hat{V}_{11}^{-1} \delta_{1K}^{*}\right)\right| \\
= & N \cdot \left|\operatorname{tr}\left(\hat{\delta}_{1K}' \left[N \cdot \hat{V}_{11}\right]^{-1} \delta_{1K}^{*}\right)\right| \\
\leq & N \cdot \lambda_{\max}\left(\left[N \cdot \hat{V}_{11}\right]^{-1}\right) \left\|\hat{\delta}_{1K}\right\| \left\|\delta_{1K}^{*}\right\| \\
\leq & N \cdot \lambda_{\max}\left(\left[N \cdot \hat{V}\right]^{-1}\right) \left\|\hat{\delta}_{1K}\right\| \left\|\delta_{1K}^{*}\right\| && \text{(B.44)} \\
= & N \cdot \left[O(1) + O_p\left(\zeta(K) K^{3/2} N^{-1/2}\right) + O_p\left(\zeta(K) K^{3/2} K^{-s/d}\right)\right] \\
& \times \left[O_p\left(\zeta(K) K N^{-1}\right) + O\left(K^{-s/d}\right)\right]\left[O\left(K^{1/2} K^{-s/d}\right)\right] \\
= & O\left(\zeta(K) K^{2} K^{-3s/d} N\right) && \text{(B.45)}
\end{aligned}
$$

where equation (B.44) follows by Lemma B.1 and equation (B.45) follows from Lemma B.2 and the proof of Theorem 3.1. Thus, we have that equation (B.41) is,

$$
\begin{aligned}
\left|T^{*\prime} - T'\right| &= O\left(K^{-1/2}\right)\left[O_p\left(\zeta(K) K^{5/2} K^{-2s/d} N^{1/2}\right) + O_p\left(\zeta(K) K^{2} K^{-3s/d} N\right)\right] \\
&= O_p\left(\zeta(K) K^{2} K^{-2s/d} N^{1/2}\right) + O_p\left(\zeta(K) K^{3/2} K^{-3s/d} N\right)
\end{aligned}
$$

which is $o_p(1)$ under Assumptions 3.2 and 3.3. ■

Proof of Theorem 3.4 First, note that we may partition $R_K(x)$ as

$$
R_K(x) = \begin{pmatrix} R_1 \\ R_{K-1}(x) \end{pmatrix}.
$$

Next, consider

$$
\begin{aligned}
\rho_N \cdot \sup_{x \in \mathbb{X}} |\Delta(x)| &= \sup_{x} |\mu_1(x) - \mu_0(x) - \tau| \\
&\leq \sup_{x \in \mathbb{X}} \left|R_K(x)'\gamma_{1,K}^0 - \mu_1(x)\right| + \sup_{x \in \mathbb{X}} \left|R_K(x)'\gamma_{0,K}^0 - \mu_0(x)\right| \\
&\quad + \sup_{x \in \mathbb{X}} \left|R_K(x)'\hat{\gamma}_{0,K} - R_K(x)'\gamma_{0,K}^0\right| + \sup_{x \in \mathbb{X}} \left|R_K(x)'\hat{\gamma}_{1,K} - R_K(x)'\gamma_{1,K}^0\right| \\
&\quad + \sup_{x \in \mathbb{X}} \left|R_{K-1}(x)'\hat{\gamma}_{11,K} - R_{K-1}(x)'\hat{\gamma}_{01,K}\right| + \left|R_1\hat{\gamma}_{10,K} - R_1\hat{\gamma}_{00,K} - \tau\right| \\
&\leq \sup_{x \in \mathbb{X}} \left|R_K(x)'\gamma_{0,K}^0 - \mu_0(x)\right| + \sup_{x \in \mathbb{X}} \left|R_K(x)'\gamma_{1,K}^0 - \mu_1(x)\right| \\
&\quad + \sup_{x \in \mathbb{X}} \left\|R_K(x)\right\| \cdot \left\|\hat{\gamma}_{0,K} - \gamma_{0,K}^0\right\| + \sup_{x \in \mathbb{X}} \left\|R_K(x)\right\| \cdot \left\|\hat{\gamma}_{1,K} - \gamma_{1,K}^0\right\| \\
&\quad + \sup_{x \in \mathbb{X}} \left\|R_{K-1}(x)\right\| \cdot \left\|\hat{\gamma}_{11,K} - \hat{\gamma}_{01,K}\right\| + \left|R_1\hat{\gamma}_{10,K} - R_1\hat{\gamma}_{00,K} - \tau\right| \\
&\leq \sup_{x \in \mathbb{X}} \left|R_K(x)'\gamma_{0,K}^0 - \mu_0(x)\right| + \sup_{x \in \mathbb{X}} \left|R_K(x)'\gamma_{1,K}^0 - \mu_1(x)\right| \\
&\quad + \zeta(K) \cdot \left\|\hat{\gamma}_{0,K} - \gamma_{0,K}^0\right\| + \zeta(K) \cdot \left\|\hat{\gamma}_{1,K} - \gamma_{1,K}^0\right\| + \zeta(K) \cdot \left\|\hat{\gamma}_{11,K} - \hat{\gamma}_{01,K}\right\| \\
&\quad + \left|R_1\hat{\gamma}_{10,K} - R_1\hat{\gamma}_{00,K} - \tau\right|.
\end{aligned}
$$

Thus,

$$\begin{aligned}
\|\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\| \quad \geq \quad & \zeta(K)^{-1}\cdot\rho_N\cdot\sup_{x\in\mathbb{X}}|\Delta(x)|-\zeta(K)^{-1}\cdot\sup_{x\in\mathbb{X}}\left|R_K(x)'\gamma^0_{0,K}-\mu_0(x)\right| \\
& -\zeta(K)^{-1}\cdot\sup_{x\in\mathbb{X}}\left|R_K(x)'\gamma^0_{1,K}-\mu_1(x)\right|-\left\|\hat{\gamma}_{0,K}-\gamma^0_{0,K}\right\|-\left\|\hat{\gamma}_{1,K}-\gamma^0_{1,K}\right\| \\
& -\zeta(K)^{-1}\cdot\left|R_1\hat{\gamma}_{10,K}-R_1\hat{\gamma}_{00,K}-\tau\right|.
\end{aligned}$$

We may follow the steps of the proof of Theorem 3.2 to obtain, for any M',

$$\Pr\left(N^{1/2}\zeta\left(K\right)^{-1/2}K^{-1/2}\left\|\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right\|>M'\right)\longrightarrow 1. \tag{B.46}$$

Next, we show that this implies that

$$\Pr\left(\frac{\tilde{C}\cdot\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)'V_{11}^{-1}\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)-(K-1)}{\sqrt{2\,(K-1)}}>M\right)\longrightarrow 1, \tag{B.47}$$

for an arbitrary constant $\tilde{C}\in\mathbb{R}_{++}$. Denote $\lambda_{\min}([N\cdot V_{11}]^{-1})$ and $\lambda_{\max}([N\cdot V_{11}]^{-1})$ by $\underline{\lambda}_{11}$ and $\bar{\lambda}_{11}$, respectively and note that by Lemma A.2 and Lemma B.1 it follows that $\underline{\lambda}_{11}$ is bounded away from zero and $\bar{\lambda}_{11}$ is bounded.

$$\begin{aligned}
& \Pr\left(\frac{\tilde{C}\cdot\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)'V_{11}^{-1}\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)-(K-1)}{\sqrt{2\,(K-1)}}>M\right) \\
= \; & \Pr\left(\frac{\tilde{C}\cdot N\cdot\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)'\left[N\cdot V_{11}\right]^{-1}\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)-(K-1)}{\sqrt{2\,(K-1)}}>M\right) \\
= \; & \Pr\left(\tilde{C}\cdot N\cdot\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)'\left[N\cdot V_{11}\right]^{-1}\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)>M\sqrt{2\,(K-1)}+K-1\right) \\
\geq \; & \Pr\left(\underline{\lambda}_{11}\tilde{C}\cdot N\cdot\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)'\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)>M\sqrt{2\,(K-1)}+K-1\right) \\
= \; & \Pr\left(N\zeta\left(K\right)^{-1}K^{-1}\cdot\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)'\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)>\left(\underline{\lambda}_{11}\tilde{C}\right)^{-1}\zeta\left(K\right)^{-1}K^{-1}\left[M\sqrt{2}\,(K-1)^{1/2}+K-1\right]\right) \\
= \; & \Pr\left(N^{1/2}\zeta\left(K\right)^{-1/2}K^{-1/2}\left\|\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right\|>\left(\underline{\lambda}_{11}\tilde{C}\right)^{-1/2}\zeta\left(K\right)^{-1/2}\left(M\sqrt{2}K^{-1}\left(K-1\right)^{1/2}+1-K^{-1}\right)^{1/2}\right)
\end{aligned}$$

Since for any M, for large enough N, we have

$$\left(\underline{\lambda}_{11}\tilde{C}\right)^{-1/2}\zeta\left(K\right)^{-1/2}\left(M\sqrt{2}K^{-1}\left(K-1\right)^{1/2}+1-K^{-1}\right)^{1/2}<2\left(\underline{\lambda}\tilde{C}\right)^{-1/2}$$

it follows that this probability is for large N bounded from below by the probability

$$\Pr\left(N^{1/2}\zeta\left(K\right)^{-1/2}K^{-1/2}\left\|\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right\|>2\left(\underline{\lambda}\tilde{C}\right)^{-1/2}\right)$$

which goes to one by (B.46). To conclude we must show that this implies that

$$\Pr(T'>M)=\Pr\left(\frac{\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)'\hat{V}_{11}^{-1}\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)-(K-1)}{\sqrt{2\,(K-1)}}>M\right)\longrightarrow 1.$$

Let $\hat{\underline{\lambda}}_{11}=\lambda_{\min}([N\cdot\hat{V}_{11}]^{-1})$ for simplicity of notation. Let A_1 denote the event that $\hat{\underline{\lambda}}_{11}>\underline{\lambda}_{11}/2$ which satisfies $\Pr\left(A_1\right)\to 1$ as $N\to\infty$ by Lemmas A.1, A.2, and B.1 along with Assumptions 3.2 and 3.3. Also define the event A_2,

$$\frac{\left(\underline{\lambda}_{11}/2\right)N\cdot\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)'\left(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K}\right)-(K-1)}{\sqrt{2\,(K-1)}}>M.$$

Note that

$$\begin{aligned}
& \Pr\left(\frac{\tilde{C}\cdot(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})' V_{11}^{-1}(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})-(K-1)}{\sqrt{2(K-1)}}>M\right) \\
= & \Pr\left(\frac{\tilde{C}\cdot N\cdot(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})' [N\cdot V_{11}]^{-1}(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})-(K-1)}{\sqrt{2(K-1)}}>M\right) \\
\leq & \Pr\left(\frac{\bar{\lambda}_{11}\tilde{C}\cdot N\cdot(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})'(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})-(K-1)}{\sqrt{2(K-1)}}>M\right)
\end{aligned}$$

which goes to one as $N\to\infty$ by equation (B.47). Since $\tilde{C}$ was arbitrary we may choose $\tilde{C}=(\underline{\lambda}_{11}/2)\cdot\bar{\lambda}_{11}^{-1}$ and so $\Pr(A_2)\to 1$ as $N\to\infty$. Thus, $\Pr(A_1\cap A_2)\to 1$ as $N\to\infty$. Finally, note that the event $A_1\cap A_2$ implies that

$$\begin{aligned}
T' &= \frac{(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})' \hat{V}_{11}^{-1}(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})-(K-1)}{\sqrt{2(K-1)}} \\
&= \frac{N\cdot(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})' \left[N\cdot\hat{V}_{11}\right]^{-1}(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})-(K-1)}{\sqrt{2(K-1)}} \\
&\geq \frac{\hat{\underline{\lambda}}_{11}N\cdot(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})'(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})-(K-1)}{\sqrt{2(K-1)}} \\
&> \frac{(\underline{\lambda}_{11}/2)\,N\cdot(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})'(\hat{\gamma}_{11,K}-\hat{\gamma}_{01,K})-(K-1)}{\sqrt{2(K-1)}} \\
&> M.
\end{aligned}$$

Hence $\Pr(T'>M))\longrightarrow 1$. ■

Lemma B.2 *Suppose Assumptions 2.1-2.3 and 3.1-3.2 hold. Then* (i),

$$\left\|\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}-\Sigma_{w,K}\Omega_{w,K}^{-1}\right\| = O_p\left(\zeta(K)^2 K^2 N^{-1}\right)+O_p\left(\zeta(K)KN^{-1/2}\right)+O_p\left(\zeta(K)KK^{-s/d}\right),$$

and (ii)

$$\left\|\hat{\Omega}_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}-\Omega_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right\| = O_p\left(\zeta(K)K^{3/2}N^{-1/2}\right)+O_p\left(\zeta(K)K^{3/2}K^{-s/d}\right).$$

Proof For (i) note that,

$$\begin{aligned}
\left\|\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}-\Sigma_{w,K}\Omega_{w,K}^{-1}\right\| &= \left\|\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}-\Sigma_{w,K}\hat{\Omega}_{w,K}^{-1}+\Sigma_{w,K}\hat{\Omega}_{w,K}^{-1}-\Sigma_{w,K}\Omega_{w,K}^{-1}\right\| \\
&\leq \left\|\hat{\Sigma}_{w,K}-\Sigma_{w,K}\right\|\left\|\hat{\Omega}_{w,K}^{-1}\right\| \qquad \text{(B.48)} \\
&+ \left\|\Sigma_{w,K}\right\|\left\|\hat{\Omega}_{w,K}^{-1}-\Omega_{w,K}^{-1}\right\| \qquad \text{(B.49)}
\end{aligned}$$

First, consider equation (B.48),

$$\left\|\hat{\Sigma}_{w,K}-\Sigma_{w,K}\right\| = O_p\left(\zeta(K)^2 K^{3/2}N^{-1}\right)+O_p\left(\zeta(K)K^{1/2}K^{-s/d}\right)$$

by Lemma A.1. Next,

$$\left\|\hat{\Omega}_{w,K}^{-1}\right\| \leq K^{1/2}\cdot\lambda_{\max}\left(\hat{\Omega}_{w,K}^{-1}\right) = O\left(K^{1/2}\right)+O_p\left(\zeta(K)KN^{-1/2}\right).$$

For equation (B.49) we have that,

$$\|\Sigma_{w,K}\| \leq K^{1/2} \cdot \lambda_{\max}\left(\Sigma_{w,K}\right) = O\left(K^{1/2}\right).$$

Also, we have

$$\left\|\hat{\Omega}_{w,K}^{-1} - \Omega_{w,K}^{-1}\right\| = O_p\left(\zeta\left(K\right) K^{1/2} N^{-1/2}\right) \tag{B.50}$$

by Lemma A.1. Thus,

$$\begin{aligned}
&\left\|\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1} - \Sigma_{w,K}\Omega_{w,K}^{-1}\right\| \\
&= \left[O_p\left(\zeta\left(K\right)^2 K^{3/2} N^{-1}\right) + O_p\left(\zeta\left(K\right) K^{1/2} K^{-s/d}\right)\right]\left[O\left(K^{1/2}\right) + O_p\left(\zeta\left(K\right) K N^{-1/2}\right)\right] \\
&+ \left[O\left(K^{1/2}\right)\right]\left[O_p\left(\zeta\left(K\right) K^{1/2} N^{-1/2}\right)\right] \\
&= O_p\left(\zeta\left(K\right)^2 K^2 N^{-1}\right) + O_p\left(\zeta\left(K\right) K N^{-1/2}\right) + O_p\left(\zeta\left(K\right) K K^{-s/d}\right).
\end{aligned}$$

For (ii) note that,

$$\begin{aligned}
&\left\|\hat{\Omega}_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1} - \Omega_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right\| \\
&= \left\|\hat{\Omega}_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1} - \Omega_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1} + \Omega_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1} - \Omega_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right\| \\
&\leq \left\|\hat{\Omega}_{w,K}^{-1} - \Omega_{w,K}^{-1}\right\| \left\|\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}\right\| && \text{(B.51)} \\
&+ \left\|\Omega_{w,K}^{-1}\right\| \left\|\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1} - \Sigma_{w,K}\Omega_{w,K}^{-1}\right\| && \text{(B.52)}
\end{aligned}$$

For equation (B.51) the first factor is $O_p\left(\zeta\left(K\right) K^{1/2} N^{-1/2}\right)$ by equation (B.50) and the second factor is

$$\begin{aligned}
&\left\|\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}\right\| \\
&\leq \lambda_{\max}\left(\hat{\Sigma}_{w,K}\right)\left\|\hat{\Omega}_{w,K}\right\| \\
&\leq \lambda_{\max}\left(\hat{\Sigma}_{w,K}\right)\lambda_{\max}\left(\hat{\Omega}_{w,K}\right) K^{1/2} \\
&= \lambda_{\max}\left(\Sigma_{w,K}\right)\lambda_{\max}\left(\Omega_{w,K}\right) K^{1/2} + O_p\left(\zeta\left(K\right)^2 K^2 N^{-1}\right) + O_p\left(\zeta\left(K\right) K K^{-s/d}\right) \\
&= O\left(K^{1/2}\right) + O_p\left(\zeta\left(K\right)^2 K^2 N^{-1}\right) + O_p\left(\zeta\left(K\right) K K^{-s/d}\right).
\end{aligned}$$

Thus, equation (B.51) is

$$\left\|\hat{\Omega}_{w,K}^{-1} - \Omega_{w,K}^{-1}\right\| \left\|\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}\right\| = O_p\left(\zeta\left(K\right) K N^{-1/2}\right) + O_p\left(\zeta\left(K\right)^2 K^{3/2} K^{-s/d} N^{-1/2}\right).$$

For equation (B.52) the first factor is,

$$\left\|\Omega_{w,K}^{-1}\right\| \leq \lambda_{\max}\left(\Omega_{w,K}^{-1}\right)\|I\| = O\left(K^{1/2}\right).$$

The second factor is

$$\left\|\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1} - \Sigma_{w,K}\Omega_{w,K}^{-1}\right\| = O_p\left(\zeta\left(K\right)^2 K^2 N^{-1}\right) + O_p\left(\zeta\left(K\right) K N^{-1/2}\right) + O_p\left(\zeta\left(K\right) K K^{-s/d}\right)$$

by (i). Thus, equation (B.52) is

$$\left\|\Omega_{w,K}^{-1}\right\| \left\|\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1} - \Sigma_{w,K}\Omega_{w,K}^{-1}\right\| = O_p\left(\zeta\left(K\right) K^{3/2} N^{-1/2}\right) + O_p\left(\zeta\left(K\right) K^{3/2} K^{-s/d}\right)$$

Finally,

$$
\begin{aligned}
&\left\|\hat{\Omega}_{w,K}^{-1}\hat{\Sigma}_{w,K}\hat{\Omega}_{w,K}^{-1}-\Omega_{w,K}^{-1}\Sigma_{w,K}\Omega_{w,K}^{-1}\right\| \\
&=O_p\left(\zeta\left(K\right)KN^{-1/2}\right)+O_p\left(\zeta\left(K\right)^2K^{3/2}K^{-s/d}N^{-1/2}\right) \\
&+O_p\left(\zeta\left(K\right)K^{3/2}N^{-1/2}\right)+O_p\left(\zeta\left(K\right)K^{3/2}K^{-s/d}\right) \\
&=O_p\left(\zeta\left(K\right)K^{3/2}N^{-1/2}\right)+O_p\left(\zeta\left(K\right)K^{3/2}K^{-s/d}\right).
\end{aligned}
$$

■

Additional References

Imbens, Guido, Whitney Newey and Geert Ridder, "Mean-squared-error Calculations for Average Treatment Effects," unpublished manuscript, Department of Economics, Harvard University (2006).

Li, Chi-Kwong, and Roy Mathias, "Interlacing Inequalities for Totally Nonnegative Matrices," *Linear Algebra and its Applications* 341 (January 2002), 35-44.

Lorentz, George, *Approximation of Functions* New York: Chelsea Publishing Company, 1986).