Supplementary Material

The supplementary material contains:

(a) A complete discussion on the regularity conditions used to establish the main results in Sections 3 and 5.
(b) The detailed list of supporting lemma along with the proofs.

A Assumptions

Assumptions about the kernel function:

(K) $K(\cdot)$ is a symmetric and non-negative density function on $[-1, 1]$.

Assumptions about the regression function:

(R1) For every $j = 1, \cdots, L$, $\theta_j(\cdot)$ has p derivatives.

(R2) For every $j = 1, \cdots, L$, $\theta_j^{(p)}(\cdot)$ satisfies a Lipschitz condition of degree $q \in (0, 1]$ in a neighborhood of z.

Assumptions about the bandwidth:

(H1) $h + (nh)^{-1} \to 0$ as $n \to \infty$.

Assumptions about distributions of (X, Z):

(D1) The densities f_i, f_{ij}, f_{ijk} and $f_{ijk\ell}$, $(1 \le i < j < k < \ell \le n)$ as defined earlier are bounded uniformly in large n in neighborhoods of all combinations of arguments. We also assume for some $\epsilon > 0$,

$$\max_{i_1} \sup_{|u_{i_1}|<\epsilon} |f_{i_1}(u_{i_1} + z)| < \infty, \quad \max_{i_s : s=1,2} \sup_{|u_{i_s}|<\epsilon} |f_{i_1 i_2}(u_{i_1} + z, u_{i_2} + z)| < \infty.$$

(D2) For $\epsilon > 0$, define

$$c(z, \epsilon) = \max_{i,j,k} \sup_{|w|\le\epsilon} |m_{i,jk}(w + z) - m_{i,jk}(z)|.$$

Assume that for some $\epsilon > 0$,

$$\limsup_{n\to\infty} c(z, \epsilon) < \infty. \tag{A.1}$$

(D3) Let $g_{i,\ell k}(z) = E(X_{i\ell}^2 X_{ik}^2 | Z_i = z)$ if $\ell \ne k$ or $E(X_{i\ell}^4 | Z_i = z)$ if $\ell = k$. Assume that $n^{-1} \sum_{i=1}^n g_{i,\ell k}(z) f_i(z)$ converges as $n \to \infty$ and for some $\epsilon > 0$,

$$\limsup_{n\to\infty} \sup_{|u|\le\epsilon} |n^{-1} \sum_{i=1}^n g_{i,\ell k}(u + z) f_i(u + z) - n^{-1} \sum_{i=1}^n g_{i,\ell k}(z) f_i(z)| < \infty \tag{A.2}$$

This assumption is needed to facilitate computation for the asymptotic covariance of the estimates. Specifically, our strategy of proofs includes applying WLLN to show that the term $n^{-1} \sum_{i=1}^n E(X_{i\ell} X_{ik} | Z_i = z)$ converges in probability to a fixed quantity (See Lemma 1). The complication arises because (X_i, Z_i), $i = 1, \cdots, n$ are not assumed to be iid, not even independent. This is a much complicated situation than the standard VCM where those are assumed to be iid in which case the standard and simpler assumption is $E(X_\ell X_k | Z = z)$ exists and is finite. Since we are considering a more general situation, we need the assumption (D3) to facilitate the technical argument.

It is also worth mentioning that in the simpler setting of iid VCM a typical assumption on X is $EX_j^{2s} < \infty$ for $s > 2, j = 1, \cdots, l$ (Zhang and Lee (2000)).

(D4) Define

$$\rho_{ij,\ell k}(z_1,z_2)=\Big[E(X_{i\ell}X_{ik}X_{j\ell}X_{jk}|Z_i=z_1,Z_j=z_2)f_{ij}(z_1,z_2)$$
$$-E(X_{i\ell}X_{ik}|Z_i=z_1)E(X_{j\ell}X_{jk}|Z_j=z_2)f_i(z_1)f_j(z_2)\Big].$$

Assume for some $\epsilon>0$,

$$\lim_{n\to\infty} n^{-2}\sum_{i\neq j=1}^{n}\sup_{|u_1|\leq\epsilon,|u_2|\leq\epsilon}|\rho_{ij,\ell k}(z_1+u_1,z_2+u_2)|=0. \tag{A.3}$$

This particular assumption is needed in different steps of our proof in lieu of relaxing the independence condition of $(X_i,Z_i), i=1,\cdots,n$ which is assumed to be true in the standard iid VCM literature. Note that $\rho_{ij,lk}(z_1,z_2)=0$ under the standard assumption in a typical VCM setup where X_i and Z_i are independent and hence the assumption in (D4) is trivially satisfied.

(D5) There exist finite valued functions $\lambda_{s_1s_2}(z)$ and $\psi_{s_1s_2}(z)$ such that with $\beta_{ij}=cov(e_i,e_j)$,

$$n^{-1}\sum_{i=1}^{n} m^*_{i,s_1s_2}(z)\sim\lambda_{s_1s_2}(z) \text{ as } n\to\infty,$$
$$\frac{n^{-2}\sum_{i\neq j=1}^{n}\beta_{ij}m^*_{ij,s_1s_2}(z)}{n^{-2}\sum_{i\neq j=1}^{n}\beta_{ij}}\sim\psi_{s_1s_2}(z) \text{ as } n\to\infty.$$

Define the $L\times L$ matrix $\lambda^*(z)$ such that its (j,k)-th element is $\lambda_{jk}(z)$ for $j,k=1,\ldots,L$ and define the matrix

$$\Lambda=\begin{bmatrix}\nu_0\lambda^*(z) & \nu_1\lambda^*(z) & \ldots & \nu_p\lambda^*(z)\\ \nu_1\lambda^*(z) & \nu_2\lambda^*(z) & \ldots & \nu_{p+1}\lambda^*(z)\\ \vdots & \vdots & \ddots & \vdots\\ \nu_p\lambda^*(z) & \nu_{p+1}\lambda^*(z) & \ldots & \nu_{2p}\lambda^*(z)\end{bmatrix}, \tag{A.4}$$

where $\nu_{\ell_1+\ell_2}, \ell_1,\ell_2=0,\cdots,p$ is as defined in (3.5). Similarly, let us define the $L\times L$ matrix Ψ with $\psi_{jk}(z)$ and $\widetilde{\kappa}(\ell_1,\ell_2)=\kappa_{\ell_1}\kappa_{\ell_2}(\ell_1!\ell_2!)^{-1}$ with κ_r as defined in (3.4) replacing $\lambda_{jk}(z)$ and $\nu_{\ell_1+\ell_2}$.

It is interesting to note that in standard VCM setup where X_i and Z_i are independent (See, e.g., Fan and Zhang (1999)) a typical assumption is $E(X_\ell X_k|Z=z)$ exists and finite for all z (in addition to the assumption that $f_i(z)$ exists). This is essentially similar to the assumption we have made through defining $\lambda_{s_1,s_2}(z)$. In our case since we are also allowing for heteroscedastic models where σ^2 is a function of both (X_i,Z_i), we need to account for it in the definitions of $m^*_{i,s_1s_2}(z)$ and hence $\lambda_{s_1,s_2}(z)$.

Regarding the expression for ψ, we need this definition to cope up with the dependence structure within the errors. Note that, if the errors are independent as in standard VCMs this definition of ψ is not needed.

(D6) Let us define the following quantities for any $m,n,k,q\in\{1,2,\cdots\}$,

$$\rho_{i_1^m}(z_{i_1})=E(\sigma^m_{i_1}(X_{i_1},Z_{i_1})X_{i_1\ell_1}|Z_{i_1}=z_{i_1}),$$
$$\rho_{i_1^m i_2^n}(z_{i_1},z_{i_2})=E(\sigma^m_{i_1}(X_{i_1},Z_{i_1})\sigma^n_{i_2}(X_{i_2},Z_{i_2})X_{i_1\ell_1}X_{i_2\ell_2}|Z_{i_1}=z_{i_1},Z_{i_2}=z_{i_2}),$$
$$\rho_{i_1^m i_2^n i_3^k}(z_{i_1},z_{i_2},z_{i_3})=E(\sigma^m_{i_1}(X_{i_1},Z_{i_1})\sigma^n_{i_2}(X_{i_2},Z_{i_2})\sigma^k_{i_3}(X_{i_3},Z_{i_3})$$
$$X_{i_1\ell_1}X_{i_2\ell_2}X_{i_3\ell_3}|Z_{i_1}=z_{i_1},Z_{i_2}=z_{i_2},Z_{i_3}=z_{i_3}),$$
$$\rho_{i_1^m i_2^n i_3^k i_4^q}(z_{i_1},z_{i_2},z_{i_3},z_{i_q})=E(\sigma^m_{i_1}(X_{i_1},Z_{i_1})\sigma^n_{i_2}(X_{i_2},Z_{i_2})\sigma^k_{i_3}(X_{i_3},Z_{i_3})\sigma^q_{i_4}(X_{i_4},Z_{i_4})$$
$$X_{i_1\ell_1}X_{i_2\ell_2}X_{i_3\ell_3}X_{i_4\ell_4}|Z_{i_1}=z_{i_1},Z_{i_2}=z_{i_2},Z_{i_3}=z_{i_3},Z_{i_4}=z_{i_4}).$$

Note that the ρ's as defined above also depend on the indices of variable X. However to make the notations simple we suppress those indices while defining ρ.

We assume for some $\epsilon > 0$,

$$\max_{i_1} \sup_{|u_{i_1}|<\epsilon} |\rho_{i_1^2}(u_{i_1}+z)| < \infty \ , \max_{i_1} \sup_{|u_{i_1}|<\epsilon} |\rho_{i_1^4}(u_{i_1}+z)| < \infty,$$

$$\max_{i_s:s=1,2} \sup_{|u_{i_s}|<\epsilon} |\rho_{i_1^1 i_2^1}(u_{i_1}+z, u_{i_2}+z)| < \infty \ , \max_{i_s:s=1,2} \sup_{|u_{i_s}|<\epsilon} |\rho_{i_1^2 i_2^2}(u_{i_1}+z, u_{i_2}+z)| < \infty$$

$$\max_{i_s:s=1,2} \sup_{|u_{i_s}|<\epsilon} |\rho_{i_1^3 i_2^1}(u_{i_1}+z, u_{i_2}+z)| < \infty \ , \max_{i_s:s=1,2,3} \sup_{|u_{i_s}|<\epsilon} |\rho_{i_1^2 i_2^1 i_3^1}(u_{i_1}+z, u_{i_2}+z, u_{i_3}+z)| < \infty.$$

(D7) Let us define $t_n = |n^{-2}\sum_{i\neq j=1}^n \beta_{ij}|$, $s_n = (nh)^{-1}$, and the scaling sequence

$$c_n = \begin{cases} nh & \text{if } \ t_n/s_n \to c \in [0,\infty), \\ n^2 \Big/ \left| \sum_{i\neq j=1}^n \beta_{ij} \right| & \text{if } \ t_n/s_n \to \infty, \end{cases}$$

For any $\epsilon > 0$ let us define the following quantities:

$$\begin{aligned} &\tilde{\rho}_{i_1 i_2}(z,z,\epsilon) \\ &= \sup_{|u_{i_s}|<\epsilon; s=1,2} |\rho_{i_1^2 i_2^2}(u_{i_1}+z, u_{i_2}+z) f_{i_1 i_2}(u_{i_1}+z, u_{i_2}+z) \\ &\qquad -\rho_{i_1^2}(u_{i_1}+z)\rho_{i_2^2}(u_{i_2}+z) f_{i_1}(u_{i_1}+z) f_{i_2}(u_{i_2}+z)|, \end{aligned}$$

$$\begin{aligned} &\tilde{\rho}_{i_1 i_2 i_3}(z,z,z,\epsilon) \\ &= \sup_{|u_{i_s}|<\epsilon; s=1,2,3} |\rho_{i_1^2 i_2^1 i_3^1}(u_{i_1}+z, u_{i_2}+z, u_{i_3}+z) f_{i_1 i_2 i_3}(u_{i_1}+z, u_{i_2}+z, u_{i_3}+z) \\ &\qquad -\rho_{i_1^2}(u_{i_1}+z)\rho_{i_2^1 i_3^1}(u_{i_2}+z, u_{i_3}+z) f_{i_1}(u_{i_1}+z) f_{i_2 i_3}(u_{i_2}+z, u_{i_3}+z)|, \\ &\tilde{\rho}_{i_1 i_2 i_3 i_4}(z,z,z,z,\epsilon) \\ &= \sup_{|u_{i_s}|<\epsilon; s=1,2,3,4} |\rho_{i_1^1 i_2^1 i_3^1 i_4^1}(u_{i_1}+z, u_{i_2}+z, u_{i_3}+z, u_{i_4}+z) f_{i_1 i_2 i_3 i_4}(u_{i_1}+z, u_{i_2}+z, u_{i_3}+z, \\ &\qquad u_{i_4}+z) - \rho_{i_1^1 i_2^1}(u_{i_1}+z, u_{i_2}+z)\rho_{i_3^1 i_4^1}(u_{i_3}+z, u_{i_4}+z) f_{i_1 i_2}(u_{i_1}+z, u_{i_2}+z) \\ &\qquad f_{i_3 i_4}(u_{i_3}+z, u_{i_4}+z)|. \end{aligned}$$

We assume for some $\epsilon > 0$,

$$\sum_{\substack{i_s=1 \\ s=1,2}}^n \tilde{\rho}_{i_1 i_2}(z,z,\epsilon) = o_p(n^2), \tag{A.5}$$

$$\sum_{\substack{i_s=1 \\ s=1,2,3}}^n |\beta_{i_2 i_3}| \tilde{\rho}_{i_1 i_2 i_3}(z,z,z,\epsilon) = o_p(n^2 t_n), \tag{A.6}$$

$$\sum_{\substack{i_s=1 \\ s=1,\cdots,4}}^n |\beta_{i_1 i_2}\beta_{i_3 i_4}| \tilde{\rho}_{i_1 i_2 i_3 i_4}(z,z,z,z,\epsilon) = o_p(n^4 t_n^2), \tag{A.7}$$

with the understanding that the indices i_s in the above sums are not equal. Note that under the standard VCM assumptions (i.e., X_i and Z_i are independent and the errors are homoscedastic) the assumptions in (D6) essentially implies that $E(X_{i\ell}|Z_i = z)$ is bounded in a local neighborhood of z. This is a reasonable assumption in standard VCM literature.

As mentioned earlier, in this work we are working under a more general framework and it is easy to see that the quantities $\tilde{\rho}$ defined in (D7) are zero when we have the usual VCMs. Hence the assumptions as in (A.5)-(A.7) as defined in (D7) are practically redundant. In fact, the last two assumptions in (D7) are trivially satisfied when the errors are independent. However because of the fact that we are considering a more general case with heteroscedastic error variance as well as dependence among X_i and Z_i, we are compelled to make more complicated assumptions as in (D6) and (D7).

Assumptions about the weights $\{\alpha_{ij}\}$*:*

(W1) Define

$$b_n = \sup_j \Big(\sum_{i=1}^{n} |\alpha_{ij}|\Big)^2 / \sum_{j=-\infty}^{\infty} \Big(\sum_{i=1}^{n} |\alpha_{ij}|\Big)^2.$$

Assume that $\limsup_{n\to\infty} \sup_j \sum_{i=1}^{n} \alpha_{ij}^2 \leq \infty$ and $b_n \to 0$ as $n \to 0$.

(W2) the covariances $\beta_{ij} = cov(e_i, e_j),\ 1 \leq i, j \leq n,\ n \geq 1$ satisfy

$$\lim_{n\to\infty} n^{-2} \sum_{i\neq j} |\beta_{ij}| = 0, \tag{A.8}$$

$$\sup_n \sup_j \sum_{i=1}^{\infty} |\alpha_{ij}| + \sup_n \max_{i=1,\cdots,n} \sum_{j=-\infty}^{\infty} |\alpha_{ij}| < \infty. \tag{A.9}$$

Under our assumption that ϵ_i has finite variance

$$\beta_{ij} = \sum_{k=-\infty}^{\infty} \alpha_{ik}\alpha_{jk}.$$

So, under (A.9),

$$\limsup_{n\to\infty} n^{-1} \sum_{i,j=1}^{n} |\beta_{ij}| \leq \infty.$$

(W3) If $s_n/t_n \to 0$, then $\sum_{i,j,k=1}^{n} \beta_{ij}\beta_{ik} = o(n^3 t_n)$ as $n \to \infty$, where s_n and t_n are as defined in (D7).

B Supporting Lemmas

Lemma 1 *Under the assumptions (K), (H1), (D2) - (D4),*

$$B_{rjk} - \kappa_r (r!)^{-1} n^{-1} \sum_{i=1}^{n} m_{i,jk}(z) \to^p 0 \ for\ r, j, k \geq 0.$$

Proof We start by noting that for $r, j, k \geq 0$,

$$\begin{aligned}
&E[(nh)^{-1} \sum_{i=1}^{n} K\{(Z_i - z)/h\}(Z_i - z)^r X_{ij} X_{ik}/h^r r!] \\
&\quad = E[(nh)^{-1} \sum_{i=1}^{n} E(X_{ij}X_{ik}|Z_i) K\{(Z_i - z)/h\}(Z_i - z)^r/h^r r!] \\
&\quad = n^{-1}(r!)^{-1} \sum_{i=1}^{n} \int m_{i,jk}(wh + z) K(w) w^r dw \\
&\quad = \kappa_r (r!)^{-1} n^{-1} \sum_{i=1}^{n} m_{i,jk}(z) + n^{-1}(r!)^{-1} \sum_{i=1}^{n} \int K(w) w^r \{m_{i,jk}(wh + z) - m_{i,jk}(z)\} dw.
\end{aligned}$$

As a consequence of (A.1), given any $\delta > 0$, there exists large enough n, such that the absolute value of the second term is bounded by

$$(r!)^{-1}c(z,h)\int K(w)|w^r|dw < \delta.$$

Hence we have for all large enough n,

$$\Big|E[(nh)^{-1}\sum_{i=1}^{n} X_{ij}X_{ik}K\{(Z_i-z)/h\}(Z_i-z)^r/h^r r!] - \kappa_r (r!)^{-1}n^{-1}\sum_{i=1}^{n} m_{ijk}(z)\Big| < \delta.$$

In addition,

$$\begin{aligned}
&var[(nh)^{-1}\sum_{i=1}^{n} X_{i\ell}X_{ik}K\{(Z_i-z)/h\}(Z_i-z)^r/h^r r!]\\
&= (nh)^{-2}\sum_{i=1}^{n} var[X_{i\ell}X_{ik}K\{(Z_i-z)/h\}(Z_i-z)^r/h^r r!]\\
&\quad +(nh)^{-2}\sum_{i\neq j=1}^{n} cov[X_{i\ell}X_{ik}K\{(Z_i-z)/h\}(Z_i-z)^r, X_{j\ell}X_{jk}K\{(Z_j-z)/h\}(Z_j-z)^r]/h^{2r}(r!)^2.
\end{aligned} \tag{B.10}$$

The first term is bounded by

$$\begin{aligned}
&(nh)^{-2}\sum_{i=1}^{n} E[E(X_{i\ell}^2 X_{ik}^2|Z_i)K^2\{(Z_i-z)/h\}(Z_i-z)^{2r}/h^{2r}(r!)^2]\\
&= (r!)^{-2}(nh)^{-1}\int K^2(w)w^{2r}\Big[n^{-1}\sum_{i=1}^{n} g_{i,\ell k}(wh+z)f_i(wh+z)\Big]dw \to 0,
\end{aligned}$$

where $g_{i,\ell k}(z) = E(X_{i\ell}^2 X_{ik}^2|Z_i = z)$ if $\ell \neq k$ or $E(X_{i\ell}^4|Z_i = z)$ if $\ell = k$ as defined in (D3). The last limit holds according to our assumption in (A.2).

Also, for the covariance term in (B.10), we have

$$(r!)^{-2}\int K(w_1)w_1^r K(w_2)w_2^r\Big[n^{-2}\sum_{i\neq j=1}^{n} \rho_{ij,\ell k}(w_1h+z, w_2h+z)\Big]dw_1dw_2 \to 0,$$

by assumption (A.3). Hence, the proof is completed.

Lemma 2 *Under the assumptions (D1), (D6), (D7), and (W3),*

$$c_n^2 E(||T - E(T)||^2) \to^p 0.$$

Proof To start with, let us first note that

$$E\{T-E(T)\}\{T-E(T)\}' = \sum_{j,\ell=1}^{N}[E(W_jW_j'W_\ell W_\ell') - E(W_jW_j')E(W_\ell W_\ell')] = \sum_{j,\ell=1}^{N} S_{j\ell}.$$

A typical element of $S_{j\ell}$ is given by

$$\begin{aligned}
&(nh)^{-4}\sum_{i_s=1,s=1,\cdots,4}^{n} \alpha_{i_1j}\alpha_{i_2j}\alpha_{i_3j}\alpha_{i_4j}/h^{r_s}\Big[E[\prod_{s=1}^{4} K\{(Z_{i_s}-z)/h\}\sigma_{i_s}(X_{i_s},Z_{i_s})(Z_{i_s}-z)^{r_s}X_{i_s\ell_s}]\\
&-E[\prod_{s=1}^{2} K\{(Z_{i_s}-z)/h\}\sigma_{i_s}(X_{i_s},Z_{i_s})(Z_{i_s}-z)^{r_s}X_{i_s\ell_s}]\\
&\qquad\times E[\prod_{s=3}^{4} K\{(Z_{i_s}-z)/h\}\sigma_{i_s}(X_{i_s},Z_{i_s})(Z_{i_s}-z)^{r_s}X_{i_s\ell_s}]\Big].
\end{aligned}$$

There could be the following cases: (1) $i_1 = i_2 = i_3 = i_4$, (2) $i_1 = i_2 = i_3 \neq i_4$, (3) $i_1 = i_2 \neq i_3 \neq i_4$, (4) $i_1 = i_3 \neq i_2 \neq i_4$, (5) $i_1 = i_3 \neq i_2 = i_4$, (6) $i_1 = i_2 \neq i_3 = i_4$ and (7) $i_1 \neq i_2 \neq i_3 \neq i_4$.

To prove this lemma we will follow the same approach as in Robinson (2011) and we will bound each of this seven terms separately. Let us first consider the first case where all i_s are different, i.e., $i_1 \neq i_2 \neq i_3 \neq i_4$. In this case it can be shown that for some constant C the term is bounded by

$$Cn^{-4} \sum_{\substack{i_s=1\\ s=1,\cdots,4}}^{n} |\beta_{i_1 i_2}\beta_{i_3 i_4}|\tilde{\rho}_{i_1 i_2 i_3 i_4}(z,z,z,z,\epsilon)$$

for some $\epsilon > 0$. According to the assumption in (A.7) this term is $o_p(t_n^2)$.

Next consider the case $i_1 = i_2 \neq i_3 \neq i_4$. In this case for some constant C the term is bounded by

$$Cn^{-3}h^{-1} \sum_{\substack{i_s=1\\ s=1,2,3}}^{n} |\beta_{i_2 i_3}|\tilde{\rho}_{i_1 i_2 i_3}(z,z,z,\epsilon).$$

This is $o_p(s_n t_n)$ using (A.6), which is $o_p(s_n^2)$ if $t_n = O_p(s_n)$ and $o_p(t_n^2)$ if $s_n = o_p(t_n)$.

For the case $i_1 = i_2 \neq i_3 = i_4$, the term is bounded by the term

$$Cn^{-2}h^{-2}\tilde{\rho}_{i_1 i_2}(z,z,\epsilon) = o_p(s_n^2).$$

Next we consider the case when $i_1 = i_3 \neq i_2 \neq i_4$. The first term of the sum can be bounded by,

$$Cn^{-4}h^{-1} \sum_{\substack{i_s=1\\ s=1,2,3}}^{n} |\beta_{i_1 i_2}||\beta_{i_1 i_3}| \sup_{\substack{|u_{i_s}|<\epsilon\\ s=1,2,3}} |\rho_{i_1^2 i_2^1 i_3^1}(u_{i_1}+z, u_{i_2}+z, u_{i_3}+z)|$$
$$\times \sup_{\substack{|u_{i_s}|<\epsilon\\ s=1,2,3}} f_{i_1 i_2 i_3}(u_{i_1}+z, u_{i_2}+z, u_{i_3}+z).$$

The second term is bounded by

$$Cn^{-4} \sum_{\substack{i_s=1\\ s=1,2,3}}^{n} |\beta_{i_1 i_2}||\beta_{i_1 i_3}| \sup_{\substack{|u_{i_s}|<\epsilon\\ s=1,2}} |\rho_{i_1^1 i_2^1}(u_{i_1}+z, u_{i_2}+z)| \sup_{\substack{|u_{i_s}|<\epsilon\\ s=1,2}} f_{i_1 i_2}(u_{i_1}+z, u_{i_2}+z)$$
$$\sup_{\substack{|u_{i_s}|<\epsilon\\ s=1,3}} |\rho_{i_1^1 i_3^1}(u_{i_1}+z, u_{i_3}+z)| \sup_{\substack{|u_{i_s}|<\epsilon\\ s=1,3}} f_{i_1 i_3}(u_{i_1}+z, u_{i_3}+z).$$

According to our assumptions in (D1), (D6) and (W3), each of these terms are $o_p(s_n t_n)$. Now consider the case when $i_1 = i_3 \neq i_2 = i_4$. In this case the first term is bounded by

$$Cn^{-4}h^{-2} \sum_{\substack{i_s=1\\ s=1,2}}^{n} |\beta_{i_1 i_2}^2| \sup_{\substack{|u_{i_s}|<\epsilon\\ s=1,2}} |\rho_{i_1^2 i_2^2}(u_{i_1}+z, u_{i_2}+z)| \sup_{\substack{|u_{i_s}|<\epsilon\\ s=1,2}} f_{i_1 i_2}(u_{i_1}+z, u_{i_2}+z).$$

The second term is bounded by

$$Cn^{-4} \sum_{\substack{i_s=1\\ s=1,2}}^{n} |\beta_{i_1 i_2}^2| \sup_{\substack{|u_{i_s}|<\epsilon\\ s=1,2}} |\rho_{i_1^1 i_2^1}^2(u_{i_1}+z, u_{i_2}+z)| \sup_{\substack{|u_{i_s}|<\epsilon\\ s=1,2}} f_{i_1 i_2}^2(u_{i_1}+z, u_{i_2}+z).$$

Again, according to our assumptions in (D1), (D6) and (W3), the first term is $O_p(s_n^2 t_n)$ and the second term is $o_p(s_n^2 t_n)$.

For the case $i_1 = i_2 = i_3 \neq i_4$ it can be shown that the first term is bounded by

$$Cn^{-4}h^{-2} \sum_{\substack{i_s=1 \\ s=1,2}}^{n} |\beta_{i_1 i_2}| \sup_{\substack{|u_{i_s}|<\epsilon \\ s=1,2}} |\rho_{i_1^3 i_2^1}(u_{i_1}+z, u_{i_2}+z)| \sup_{\substack{|u_{i_s}|<\epsilon \\ s=1,2}} f_{i_1 i_2}(u_{i_1}+z, u_{i_2}+z),$$

and the second term is bounded by

$$Cn^{-4}h^{-1} \sum_{\substack{i_s=1 \\ s=1,2}}^{n} |\beta_{i_1 i_2}| \sup_{|u_{i_1}|<\epsilon} |\rho_{i_1^2}(u_{i_1}+z)| \sup_{|u_{i_1}|<\epsilon} f_{i_1}(u_{i_1}+z)$$
$$\times \sup_{\substack{|u_{i_s}|<\epsilon \\ s=1,2}} |\rho_{i_1^1 i_2^1}(u_{i_1}+z, u_{i_2}+z)| \sup_{\substack{|u_{i_s}|<\epsilon \\ s=1,2}} f_{i_1 i_2}(u_{i_1}+z, u_{i_2}+z).$$

Then according to (D1), (D6) and (W3), both of the above terms are $O_p(s_n^2 t_n)$ and $o_p(s_n^2 t_n)$ respectively.

Lastly, let us consider the case when all i_s are equal. In this case, it can be shown that the first and the second term are bounded by

$$Cn^{-4}h^{-3} \sum_{i_1=1}^{n} \sup_{|u_{i_1}|<\epsilon} |\rho_{i_1^4}(u_{i_1}+z)| \sup_{|u_{i_1}|<\epsilon} f_{i_1}(u_{i_1}+z), \quad \text{and}$$
$$Cn^{-4}h^{-2} \sum_{i_1=1}^{n} \sup_{|u_{i_1}|<\epsilon} |\rho_{i_1^2}^2(u_{i_1}+z)| \sup_{|u_{i_1}|<\epsilon} f_{i_1}^2(u_{i_1}+z),$$

respectively. Again, according to our assumptions as in (D1) and (D6) these terms are $O_p(s_n^3)$ and $o_p(s_n^3)$ respectively.

Lemma 3 *Under assumptions (K), (H1), (D1) - (D6), and (W3),*

$$c_n E(T) \to^p \Sigma.$$

Proof We start the proof by first observing that there can be four types of terms in $E(T)$. A typical element of $E(T)$ is, up to a term of $o_p(c_n^{-1})$, for $r_1 \neq r_2$ and $s_1 \neq s_2$,

$$\begin{aligned} D &= (nh)^{-2}(r_1!r_2!)^{-1} \sum_{i=1}^{n}\sum_{k=1}^{n} \beta_{ik} E\Big[K\{(Z_i-z)/h\}K\{(Z_k-z)/h\}(Z_i-z)^{r_1}(Z_k-z)^{r_2} \\ &\qquad \times m^*_{ik,s_1 s_2}(Z_i,Z_k)/h^{r_1+r_2}\Big] \\ &= n^{-2}(r_1!r_2!)^{-1} \sum_{i\neq k=1}^{n} \beta_{ik} \int K(w_1)K(w_2)w_1^{r_1}w_2^{r_2} m^*_{ik,s_1 s_2}(w_1 h+z, w_2 h+z)dw_1 dw_2 \\ &\quad + n^{-2}h^{-1}(r_1!r_2!)^{-1} \sum_{i=1}^{n} \int K^2(w)w^{r_1+r_2} m^*_{i,s_1 s_2}(wh+z)dw \\ &= \kappa_{r_1}\kappa_{r_2}(r_1!r_2!)^{-1}n^{-2} \sum_{i\neq k=1}^{n} \beta_{ik} m^*_{ik,s_1 s_2}(z,z) + \nu_{r_1+r_2}(nh)^{-1}\Big[n^{-1}\sum_{i=1}^{n} m^*_{i,s_1 s_2}(z)\Big] \\ &= t_n \psi_{s_1 s_2}(z) + s_n \lambda_{s_1 s_2}(z). \end{aligned}$$

Now we observe that

$$D \sim \begin{cases} s_n \lambda_{s_1 s_2}(z) & \text{when } t_n/s_n \to 0; \\ s_n(c\psi_{s_1 s_2}(z) + \lambda_{s_1 s_2}(z)) & \text{when } t_n/s_n \to c \in (0,\infty); \\ t_n \psi_{s_1 s_2}(z) & \text{when } t_n/s_n \to \infty. \end{cases}$$

The result now follows by definition of Λ, Ψ and Σ as defined in Section A.

Lemma 4 *Under assumptions (K), (H1), (D1), (D7), and (W1),*

$$E[\sum_{j=1}^{N} w_j^4] = o(1).$$

Proof Proceeding in the similar way as in Lemma 2, we can prove this result. To maintain brevity we omit the details.

Lemma 5 *Let $T = \sum_{j=1}^{N} W_j W_j'$ where W_j is as defined in Eqn. (5.11). Define the matrix P such that $PP' = T$. Then under assumptions (K), (H1), (D1), (D7), (W1) and conditional on Z,*

$$P^{-1} \sum_{j=1}^{N} W_j \epsilon_j \to^d N(0, I). \tag{B.11}$$

Proof To prove (B.11), we need to show that for any unit vector v, conditional on Z,

$$v' P^{-1} \sum_{j=1}^{N} W_j \epsilon_j = \sum_{j=1}^{N} w_j \epsilon_j \to^d N(0, 1),$$

where $w_j = v' P^{-1} W_j$. Let $\mathcal{F}_{n,j} = \sigma(X_{1n}, \cdots, X_{jn}, Z_{1n}, \cdots, Z_{jn}),\ j = 1, \cdots, N$ be the filtration generated by the process (X, Z), for positive integer $N = N_n$, increasing with n. Then $w_j \epsilon_j$ is a martingale with respect to $\mathcal{F}_{n,j}$ with $\sum_{j=1}^{N} w_j \epsilon_j$ being a martingale with respect to $\mathcal{F}_{n,N}$.
Then for any $\eta > 0$,

$$\begin{aligned}
&\sum_{j=1}^{N} E[w_j^2 \epsilon_j^2 1(|w_j \epsilon_j| > \eta) | X_1, \cdots, X_n, Z_1, \ldots, Z_n] \\
&\quad \leq \sum_{j=1}^{N} w_j^2 E[\epsilon_j^2 1(\epsilon_j^2 > \eta/\delta) | X_1, \cdots, X_n, Z_1, \ldots, Z_n] + \sum_{j=1}^{N} w_j^2 1(w_j^2 > \eta\delta) \\
&\quad \leq \max_j E[\epsilon_j^2 1(\epsilon_j^2 > \eta/\delta)] + (\eta\delta)^{-1} \sum_{j=1}^{N} w_j^4.
\end{aligned}$$

This implies

$$\sum_{j=1}^{N} E[w_j^2 \epsilon_j^2 1(|w_j \epsilon_j| > \eta)] \leq \max_j E[\epsilon_j^2 1(\epsilon_j^2 > \eta/\delta)] + (\eta\delta)^{-1} \sum_{j=1}^{N} E(w_j^4). \tag{B.12}$$

In (B.12), the inequality is still valid since we have positive random variables at both sides. Hence using Lemma 4 and the fact that the first term in (B.12) can be made arbitrarily small, the asymptotic normality of $\sum_{j=1}^{N} w_j \epsilon_j$ is confirmed (e.g., see Scott (1973)). Therefore, using Cramer-Wold device, the convergence of $P^{-1} \sum_{j=1}^{N} W_j \epsilon_j$ is obtained.