On the Sensitivity of Kernel-based Tests of Conditional Moment Restrictions

Eduardo Fé-Rodríguez
Health Economics Research at Manchester,
University of Manchester,
e.fe-rodriguez@manchester.ac.uk

Chris D. Orme
Economics, School of Social Sciences,
University of Manchester,
chris.orme@manchester.ac.uk

October 5, 2009

Abstract

These notes were written primarily for the benefit of the authors in to construct a detailed proof of the main Theorem in the above paper.

1 Assumptions & Preliminary Central Limit Theorem

Let $\{W_i\}_{i=1}^n$ be a simple random sample drawn on a random variable W from an unknown Data Generation Process (DGP). For simplicity of exposition it is assumed that $X \in \mathbb{R}^k$, a subvector of W, is a continuous random variable with probability density $f(.)$. The unknown DGP is characterized by a $(p \times 1)$ parameter vector partitioned as $\varphi = (\theta', \gamma')' \in \Theta \times \Gamma \subset \mathbb{R}^p$, with true value $\varphi_0 = (\theta_0', \gamma_0')'$, and the corresponding estimation criterion has the general form $Q_n(\varphi)$, with $\tilde{\varphi} \equiv \arg\max_\varphi Q_n(\varphi)$. Throughout, unless stated otherwise, expectations are taken with respect the true DGP. It is assumed that standard regularity conditions support the following high level assumptions:

Assumption A

1. $\tilde{\varphi} - \varphi_0 \xrightarrow{p} 0$, and φ_0 lies in the interior of the compact and convex parameter space $\Theta \times \Gamma \subset \mathbb{R}^p$
2. $\sqrt{n}\partial Q_n(\varphi_0)/\partial\varphi = O_p(1)$.
3. $\partial^2 Q_n(\varphi)/\partial\varphi\partial\varphi' - J(\varphi) \xrightarrow{p} 0$, uniformly in φ, with $J(\varphi) = O(1)$, continuous in φ, and $J(\varphi_0)$ is negative definite.

Definition L^α, $\alpha > 0$, is the class of functions $l(.) : \mathbb{R}^k \to \mathbb{R}$ satisfying the following: $\exists\ \delta > 0$ such that for all $x \in \mathbb{R}^k$, $\sup_{\|d\| \leq \delta} |l(x+d) - l(x)| / \|d\| \leq L(x)$ and $l(.)$ and $L(.)$ have finite moments of order α (or are bounded if $\alpha = +\infty$).

Assumption B

1. $K(.)$ is even, bounded, integrates to 1 and $\lim_{\|u\| \to \infty} \|u\|^k |K(u)| = 0$.
2. $h \to 0$ as $n \to \infty$, such that $nh^k \to \infty$.
3. $f(.) \in L^\infty$; i.e., $\sup_x f(x) \leq \Delta < \infty$.
4. $\varepsilon_r(W; \varphi_0)$, $r = 1, ..., m$, satisfies the following:

(a) $E\left[\varepsilon_r(W;\varphi_0)|X\right]=0$.

(b) $c_{rs}(X)=E\left[\varepsilon_r(W;\varphi_0)\varepsilon_s(W;\theta_0)|X\right]\in L^4$.

(c) $\kappa_r(X)=E\left[|\varepsilon_r(W;\varphi_0)|^4|X\right]\in L^2$.

It is assumed, for simplicity of exposition, that $K(.)$ is a symmetric density function. Assumption B is sufficient for the following multivariate generalisation of Hall's (1984) Central Limit Theorem for a second order degenerate U-Statistic (in order to support the following result and the proof of the main theorem, we exploit some Technical Lemmata, which are detailed in Section 3.

Proposition 1 *Under Assumption B*

$$\Omega_0^{-1/2}nh^{k/2}T_n(\varphi_0)\xrightarrow{d}N(0,I_m), \tag{1}$$

where

$$\Omega_0=2E\left[\{C(X;\varphi_0)\odot C(X;\varphi_0)\}f(X)\right]\int K^2(u)\,du$$

has typical element $2E\left[c_{rs}^2(X;\varphi_0)f(X)\right]\int K^2(u)\,du$, $r,s=1,...,m$, *and is finite and positive definite.*

Proof. *The proof proceeds in two stages.*

1. *Consider, first, just a typical element of* T_n, *denoted*

$$T_{nr}=\frac{1}{n(n-1)h^k}\sum_{i=1}^{n}\sum_{j\neq i}\varepsilon_r(W_i;\varphi_0)K_{ij}\varepsilon_r(W_j;\varphi_0).$$

For this part of the proof, drop the subscript r *on* ε *and let* $\varepsilon_i\equiv\varepsilon(W_i;\varphi_0)$. *Then,* $nh^{k/2}T_{nr}$ *is a degenerate U-Statistic because* $E\left[\varepsilon_i\varepsilon_jK_{ij}|W_i\right]=0$. *Need to verify Hall's (1984) conditions. Let* $H_n(W_i,W_j)=h^{-k/2}\varepsilon_iK_{ij}\varepsilon_j$, $U_n=\frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j\neq i}^n H_n(W_i,W_j)$ *and define*

$$G_n(W_1,W_2)=E\left[H_n(W_3,W_1)H_n(W_3,W_2)|W_1,W_2\right]$$

then, provided

$$\lim_{n\to\infty}\frac{E\left[G_n^2(W_1,W_2)\right]+n^{-1}E\left[H_n^4(W_1,W_2)\right]}{\left\{E\left[H_n^2(W_1,W_2)\right]\right\}^2}=0$$

we get

$$\frac{nU_n}{\sqrt{2E\left[H_n(W_1,W_2)^2\right]}}=\frac{nh^{k/2}T_{nr}}{\sqrt{2E\left[H_n(W_1,W_2)^2\right]}}\xrightarrow{d}N(0,1).$$

From Lemma 2, (2),

$$\begin{aligned}E\left[H_n^2(W_1,W_2)\right]&=h^{-k}E\left[\varepsilon_1^2\varepsilon_2^2K_{12}^2\right]=E\left[c^2(X)f(X)\right]\int K^2(u)\,du+o(1)=O(1)\\ E\left[H_n^4(W_1,W_2)\right]&=h^{-2k}E\left[\varepsilon_1^4\varepsilon_2^4K_{12}^4\right]=h^{-k}E\left[\kappa^2(X)f(X)\right]\int K^4(u)\,du+o(h^{-k})=O(h^{-k}).\end{aligned}$$

For the first we need: $c(X)=E\left[|\varepsilon|^2|X\right]\in G^2$ *(we need the bound,* ***and*** *the Lipschitz condition on* $c(X)$*).*

For the second we need: $\kappa(X)=E\left[|\varepsilon|^4|X\right]\in G^2$ *(again, we need the bound* ***and*** *the*

Lipschitz condition on $\kappa(X)$).
By (5), [1]

$$\begin{aligned} E\left[G_n^2(W_1,W_2)\right] &= h^{-2p}E\left[\varepsilon_1^2\varepsilon_2^2E\left[\varepsilon_3^2K_{31}K_{32}|X_1,X_2\right]^2\right] \\ &= h^kE\left[c^4(X)f^3(X)\right]\int\left\{\int K(u)\,K(u-v)du\right\}^2dv+o(h^k)=O(h^k) \end{aligned}$$

thus

$$\frac{E\left[G_n^2(W_1,W_2)\right]+n^{-1}E\left[H_n^4(W_1,W_2)\right]}{\left\{E\left[H_n^2(W_1,W_2)\right]\right\}^2}=O(h^k)+O(n^{-1}h^{-k})=o(1)$$

since $h^k\to 0$ and $\left(nh^k\right)^{-1}\to 0$. Note that $2E\left[H_n(W_1,W_2)^2\right]=\Sigma_0+o(1)$, where $\Sigma_0=2E\left[f(X)c^2(X)\right]\int K^2(u)\,du$, It is then immediate that $nh^{k/2}T_{nr}=nU_n\xrightarrow{d}N(0,\Sigma_0)$.

2. *Second, we apply a Cramer-Wold device. Now use a Cramer-Wold device, and define $Z_n=\xi' T_n(\varphi_0)$, for any $\|\xi\|=1$, and $\zeta_{ij}=\sum_{r=1}^m\xi_r\varepsilon_{ir}\varepsilon_{jr}=\zeta_{ji}$, $\varepsilon_{ir}=\varepsilon_r(W_i;\theta_0)$,*

$$\begin{aligned} nh^{k/2}Z_n &= nU_n \\ U_n &= \frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j\neq i}^n h^{-k/2}\zeta_{ij}K_{ij} \end{aligned}$$

which is degenerate, because $E\left[\zeta_{ij}K_{ij}|W_i\right]=0$. In this case, define $H_n(W_i,W_j)=h^{-k/2}\zeta_{ij}K_{ij}$, so that by Lemma 2, (2),

$$\begin{aligned} E\left[H_n(W_1,W_2)^2\right] &= h^{-p}E\left[\zeta_{12}^2K_{12}^2\right] \\ &= h^{-p}E\left[\left\{\sum_{r=1}^m\xi_r\varepsilon_{1r}\varepsilon_{2r}\right\}^2K_{12}^2\right] \\ &= h^{-p}E\left[\sum_{r=1}^m\xi_r^2\left(\varepsilon_{1r}^2K_{12}^2\varepsilon_{2r}^2\right)+\sum_{r=1}\sum_{s\neq r}\xi_r\xi_s\left(\varepsilon_{1r}\varepsilon_{1s}K_{ij}^2\varepsilon_{2r}\varepsilon_{2s}\right)\right] \\ &= E\left[\sum_{r=1}^m\xi_rc_{rr}^2(X)+\sum_{r=1}\sum_{s\neq r}\xi_r\xi_sc_{rs}^2(X)\right]\int K^2(u)\,du+o(1) \\ &= E\left[\left(\xi'\left[C(X)\odot C(X)\right]\xi\right)f(X)\right]\int K^2(u)\,du+o(1) \\ &= \frac{1}{2}\xi'\Omega_0\xi+o(1)=O(1), \end{aligned}$$

using Assumption C4b, where

$$\Omega_0=2E\left[\left[C(X)\odot C(X)\right]f(X)\right]\int K^2(u)\,du.$$

Further,

$$\begin{aligned} G_n(W_1,W_2) &= E\left[H_n(W_3,W_1)H_n(W_3,W_2)|W_1,W_2\right] \\ &= h^{-k}E\left[\zeta_{31}\zeta_{32}K_{31}K_{32}|W_1,W_2\right] \\ E\left[H_n^4(W_1,W_2)\right] &= h^{-2k}E\left[\zeta_{12}^4K_{12}^4\right] \end{aligned}$$

[1] This is where we need $c(X)\in G^4$.

and we need to show that $E\left[G_n^2(W_1,W_2)\right]+n^{-1}E\left[H_n(W_1,W_2)^4\right]=o(1)$. *First, by the* C_r *Inequality, and Assumption C4c, Lemma 2, (2), gives*

$$\begin{aligned}
E\left[H_n^4(W_1,W_2)\right] &= h^{-2k}E\left[\left\{\sum_{r=1}^m \xi_r\varepsilon_{1r}\varepsilon_{2r}K_{12}\right\}^4\right] \\
&\leq m^3h^{-k}\sum_{r=1}^m E\left[h^{-p}\xi_r^4\varepsilon_{1r}^4\varepsilon_{2r}^4K_{12}^4\right] \\
&= m^3h^{-k}\sum_{r=1}^m\left\{E\left[\xi_r^4\kappa_r^2(X)f(X)\right]\int K^4(u)\,du+o(1)\right\} \\
&= O(h^{-p}).
\end{aligned}$$

Second,

$$\begin{aligned}
E\left[G_n^2(W_1,W_2)\right] &= h^{-2k}E\left[\left\{E\left[\zeta_{31}\zeta_{32}K_{31}K_{32}|W_1,W_2\right]\right\}^2\right] \\
&= h^{-2k}E\left[E\left[\sum_{r=1}^m\xi_r\varepsilon_{3r}\varepsilon_{1r}K_{31}\sum_{r=1}^m\xi_r\varepsilon_{3r}\varepsilon_{2r}K_{32}|W_1,W_2\right]^2\right] \\
&= h^{-2k}E\left[\left\{\sum_{r=1}^m\xi_r^2\varepsilon_{1r}\varepsilon_{2r}E\left[c_{rr}(X_3)K_{31}|X_1,X_{32}\right]\right.\right. \\
&\qquad \left.\left.\sum_{r=1}^m\sum_{s\neq r}\xi_r\xi_s\varepsilon_{1r}\varepsilon_{2s}E\left[c_{rs}(X_3)K_{31}K_{32}|X_1,X_2\right]\right\}^2\right] \\
&= h^{-2k}E\left[\left\{\sum_{r=1}^m(A_r+B_r)\right\}^2\right] \\
&\leq 2h^{-2k}E\left[\left\{\sum_{r=1}^m A_r\right\}^2+\left\{\sum_{r=1}^m B_r\right\}^2\right] \\
&\leq 2mh^{-2k}E\left[\sum_{r=1}^m\left(A_r^2+B_r^2\right)\right]
\end{aligned}$$

using the C_r *Inequality and where*

$$\begin{aligned}
A_r &= \xi_r^2\varepsilon_{1r}\varepsilon_{2r}E\left[c_{rr}(X_3)K_{31}K_{32}|X_1,X_2\right] \\
B_r &= \xi_r\varepsilon_{1r}\sum_{s\neq r}\xi_s\varepsilon_{2s}E\left[c_{rs}(X_3)K_{31}K_{32}|X_1,X_2\right].
\end{aligned}$$

Now, by Assumption C4b, Lemma 2, (5), yields

$$\begin{aligned}
h^{-3p}E\left[A_r^2\right] &= h^{-3p}E\left[\xi_r^4\varepsilon_{1r}^2\varepsilon_{2r}^2\left\{E\left[c_{rr}(X_3)K_{31}K_{32}|X_1,X_2\right]\right\}^2\right] \\
&= h^{-3p}E\left[\xi_r^4c_{rr}(X_1)c_{rr}(X_2)\left\{E\left[c_{rr}(X_3)K_{31}K_{32}|X_1,X_2\right]\right\}^2\right] \\
&= \xi_r^4E\left[c_{rr}^4(X)f^3(X)\right]\int\left\{\int K(u)\,K(u-v)du\right\}^2dv+o(1)
\end{aligned}$$

thus $h^{-2k}E\left[A_r^2\right]=O(h^p)$. *Similarly, by Assumption C4b (and using the* C_r *Inequality* *again)*

$$\begin{aligned}
h^{-3k}E\left[B_r^2\right] &\leq (m-1)h^{-3k}\sum_{s\neq r}E\left[\xi_r^2\xi_s^2\varepsilon_{1r}^2\varepsilon_{2s}^2\left\{E\left[c_{rs}(X_3)K_{31}K_{32}|X_1,X_2\right]\right\}^2\right] \\
&= (m-1)\,h^{-3k}\sum_{s\neq r}E\left[\xi_r^2\xi_s^2c_{rr}(X_1)c_{ss}(X_2)\left\{E\left[c_{rs}(X_3)K_{31}K_{32}|X_1,X_2\right]\right\}^2\right] \\
&= (m-1)\sum_{s\neq r}\xi_r^2\xi_s^2E\left[c_{rr}(X)c_{ss}(X)c_{rs}^2(X)f^3(X)\right] \\
&\quad\times\int\left\{\int K(u)\,K(u-v)du\right\}^2dv+o(1)
\end{aligned}$$

so that $h^{-2k}E\left[B_r^2\right]=O(h^k)$*, noting that (by repeated application of Cauchy-Schwartz)*

$$\begin{aligned}
E\left[c_{rr}(X)c_{ss}(X)c_{rs}^2(X)f^3(X)\right] &\leq B^3E\left[c_{rr}(X)c_{ss}(X)c_{rs}^2(X)\right] \\
&\leq B^3\left\{E\left[c_{rr}^2(X)c_{rs}^2(X)\right]E\left[c_{ss}^2(X)c_{rs}^2(X)\right]\right\}^{1/2} \\
&= O(1)
\end{aligned}$$

since $E\left[c_{rr}^2(X)c_{rs}^2(X)\right]\leq\left\{E\left[c_{rr}^4(X)\right]E\left[c_{rs}^4(X)\right]\right\}^{1/2}<\infty$*. Thus,*

$$\frac{E\left[G_n(W_1,W_2)^2\right]+n^{-1}E\left[H_n^4(W_1,W_2)\right]}{\left\{E\left[H_n^2(W_1,W_2)\right]\right\}^2}=O(h^k)+O(n^{-1}h^{-k})=o(1)$$

since $h^k\rightarrow 0$ *and* $\left(nh^k\right)^{-1}\rightarrow 0$*. Therefore,* $\dfrac{nU_n}{\sqrt{2E\left[H_n(W_1,W_2)^2\right]}}\xrightarrow{d}N(0,1)$, $U_n=\dfrac{1}{n(n-1)}\sum_{i=1}^n\sum_{j\neq i}^n h^{-k/2}\zeta_{ij}K_{ij}$*, by part (a), implying* $nh^{k/2}Z_n=nh^{k/2}\xi'T_n(\theta_0)\xrightarrow{d}N(0,\xi'\Omega_0\xi)$*, for all* $\|\xi\|=1$*. Thus by the Cramer-Wold device,* $nh^{k/2}T_n(\theta_0)\xrightarrow{d}N(0,\Omega_0)$*.*

This completes the proof. ■

In addition, the following assumptions are sufficient to justify the asymptotic expansions employed to obtain the limit distribution of the *CCM* test indicator:

Assumption C

For all $r=1,...,m$, and each w, $\varepsilon_r(w;\theta)$ is thrice differentiable in φ, and let $g_r(W;\varphi)=\partial\varepsilon_r(W;\varphi)\partial\varphi,\,(p\times 1)$, with typical element $\{g_{rt}(W;\varphi)\}$, $t=1,...,p$, $G_r(W;\varphi)=\partial^2\varepsilon_r(W;\varphi)/\partial\varphi\partial\varphi'$, $(p\times p)$, $F_r(W;\varphi)=\partial vecG_r(W;\varphi)/\partial\varphi$, $\left(p^2\times p\right)$, satisfying:

1. $E\left[|\varepsilon_r(W;\varphi_0)|\,|X\right]\in L^2$.
2. For all $t=1,...,p$: (i) $E\left[|g_{rt}(W;\varphi_0)|^{8/3}\right]<\infty$; (ii) $E\left[g_{rt}(W;\varphi_0)|X\right]\in L^{8/3}$; (iii) $E[|g_{rt}(W;\varphi_0)|\,|X]\in L^2$; and, (iv) $E\left[|g_{rt}(W;\varphi_0)|^2\,|X\right]\in L^2$
3. (i) $E\left[\|G_r(W;\varphi_0)\|^{8/3}\right]<\infty$; and, (ii) $\sup_\varphi\|G_r(W;\varphi)\|<M(W)$, for all r, $E[M(W)]<\infty$ with $\lambda(X)=E\left[M(W)|X\right]\in L^2$.
4. (i) $F_r(W;\varphi)$ is continous in φ, for each w; and, (ii) $\sup_\varphi\|F_r(W;\varphi)\|<P(W)$, for all r, $E[P(W)]<\infty$ with $E\left[P^2(W)|X\right]<\infty$.

2 Theorem and Proof

Theorem 1 *Define the following:* $J(\varphi_0) = \begin{bmatrix} J_{\theta\theta} & J_{\theta\gamma} \\ J_{\gamma\theta} & J_{\gamma\gamma} \end{bmatrix}$, $\xi = \begin{bmatrix} J_{\theta\theta}^{-1} J_{\theta\gamma}\delta \\ -\delta \end{bmatrix}$, $d(X;\varphi) = E\left[\left.\frac{\partial\varepsilon(W;\varphi)}{\partial\varphi'}\right| X\right]$ *and* $\mu(\varphi_0) = E\left[\|d(X;\varphi_0)\xi\|^2 f(X)\right]$.
Under the true DGP characterised by $\gamma_0 = \delta/\sqrt{nh^{k/2}}$, $0 \leq \|\delta\| < \infty$, *and Assumptions A-C*

$$nh^{k/2}V_n(\hat{\varphi}) \xrightarrow{d} N(\mu_0, \Sigma_0),$$

where $\mu_0 = \lim_{n\to\infty}\mu(\varphi_0)$ *and* $\Sigma_0 = \lim_{n\to\infty}\iota'\Omega_0\iota$.

Proof. It is shown in the main paper that $\sqrt{nh^{k/2}}(\hat{\varphi}_n - \varphi_0) = O_p(1)$
It will be useful to define:
(i) $\hat{\varepsilon}_{ir} \equiv \varepsilon_r(W_i;\hat{\varphi})$, $\varepsilon_{ir}^0 \equiv \varepsilon_r(W_i;\varphi_0)$, $g_{ir}(\varphi) \equiv g_r(W_i;\varphi)$, $G_{ir}(\varphi) \equiv G_r(W_i;\varphi)$, $F_{ir}(\varphi) \equiv F_r(W_i;\varphi)$,
with
(ii) $d(X;\varphi)$, $(m \times p)$ having typical element $d_{rt}(X;)$, $r = 1,...m$, $t = 1,...,p$, and $d_r(X;\varphi)$, $(p \times 1)$, having typical element $d_{rt}(X;\varphi)$, $t = 1,...,p$.
Then, if $T_n(\varphi)$ has typical element $T_{nr}(\varphi)$, we have

$$\begin{aligned} nh^{k/2}T_{nr}(\hat{\varphi}) &= nh^{k/2}T_{nr}(\varphi_0) \\ &\quad +2\frac{nh^{k/2}}{n(n-1)h^k}\sum_{i=1}^n\sum_{j\neq i}\varepsilon_{ir}^0 K_{ij}(\hat{\varepsilon}_{jr} - \varepsilon_{jr}^0) + \frac{nh^{k/2}}{n(n-1)h^k}\sum_{i=1}^n\sum_{j\neq i}(\hat{\varepsilon}_{ir} - \varepsilon_{ir}^0)K_{ij}(\hat{\varepsilon}_{jr} - \varepsilon_{jr}^0). \\ &= nh^{k/2}T_{nr}(\varphi_0) + 2T_{1nr} + T_{2nr}, \quad \text{say.} \end{aligned}$$

It suffices to show that $T_{1nr} = o_p(1)$, whilst $T_{2nr} = E\left[|d_r(X;\varphi_0)'\xi|^2 f(X)\right] + o_p(1)$:

1. T_{1nr} : write

$$\begin{aligned} \hat{\varepsilon}_{jr} - \varepsilon_{jr}^0 &= g_{jr}(\varphi_0)'(\hat{\varphi} - \varphi_0) + \frac{1}{2}(\hat{\varphi} - \varphi_0)'G_{jr}(\varphi_0)(\hat{\varphi} - \varphi_0) \\ &\quad +\frac{1}{6}vec((\hat{\varphi} - \varphi_0)(\hat{\varphi} - \varphi_0)')'F_{jr}(\bar{\varphi}^{(r)})(\hat{\varphi} - \varphi_0) \end{aligned}$$

where $\bar{\varphi}^{(r)}$ is a "mean value" such that $\left\|\bar{\varphi}^{(r)} - \varphi_0\right\| \leq \|\hat{\varphi} - \varphi_0\| = O_p(n^{-1/2}h^{-k/4})$. Substituting this expression into T_{1nr} yields

$$\begin{aligned} T_{1nr} &= h^{k/4}\left\{\frac{\sqrt{n}}{n(n-1)h^k}\sum_{i=1}^n\sum_{j\neq i}\varepsilon_{ir}^0 K_{ij}g_{jr}(\varphi_0)'\right\}\xi_n \\ &\quad +\frac{1}{2}\xi_n'\left\{\frac{1}{n(n-1)h^k}\sum_{i=1}^n\sum_{j\neq i}\varepsilon_{ir}^0 K_{ij}G_{jr}(\varphi_0)\right\}\xi_n \\ &\quad +\frac{1}{6}vec(\xi_n\xi_n')'\left\{\frac{1}{n(n-1)h^k}\sum_{i=1}^n\sum_{j\neq i}\varepsilon_{ir}^0 K_{ij}F_{jr}(\bar{\varphi}^{(r)})\right\}(\hat{\varphi} - \varphi_0), \\ &= h^{k/4}S_{1n}'\xi_n + \frac{1}{2}\xi_n'S_{2n}\xi_n + \frac{1}{6}vec(\xi_n\xi_n')'S_{3n}(\hat{\varphi}_n - \varphi_0) \end{aligned}$$

where (here) $\xi_n = \sqrt{nh^{k/2}}(\hat{\varphi} - \varphi_0) = \xi + o_p(1) = O_p(1)$.

(a) $S_{1n} = O_p(1)$
Appeal to Corollary 1 and consider a typical element of S_{1n}. Write $S_{1n} = \{S_{1nt}\}$, $t = 1, ..., p$ where

$$\begin{aligned} S_{1nt} &= \frac{\sqrt{n}}{n(n-1)h^k} \sum_{i=1}^{n} \sum_{j \neq i} \varepsilon_{ir}^0 K_{ij} g_{rt}(W_j; \varphi_0) \\ &= \frac{\sqrt{n}}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i}^{n} H_n(W_i, W_j) \end{aligned}$$

where (dropping the subscriprs r and t on H_n, for notational convenience)

$$H_n(W_1, W_2) = \frac{1}{2h^k} \left\{ \varepsilon_{1r}^0 K_{12} g_{rt}(W_2; \varphi_0) + \varepsilon_{2r}^0 K_{21} g_{rt}(W_1; \varphi_0) \right\}.$$

Now, since $E|A+B|^2 \leq 2\left\{E|A|^2 + E|B|^2\right\}$ we can write, by Lemma 2, result (2),

$$\begin{aligned} E|H_n(W_1, W_2)|^2 &\leq \frac{1}{h^{2k}} E\left|\varepsilon_{1r}^0 K_{12} g_{rt}(W_2; \varphi_0)\right|^2 \\ &= \frac{1}{h^{2k}} E\left[\left|\varepsilon_{1r}^0\right|^2 K_{12}^2 \left|g_{rt}(W_2; \varphi_0)\right|^2\right] \\ &= h^{-k}\left\{E\left[|d_{rt}(W; \varphi_0)|^2 c_{rr}(X) f(X)\right] \int |K(u)|^2 du + o(1)\right\} \\ &= O(h^{-k}) = O\left(n/\left(nh^k\right)\right) = o(n) \end{aligned}$$

because $E\left[|g_{rt}(W; \varphi_0)|^2 c_{rr}(X) f(X)\right] = O(1)$, by Assumption B4b and C2. More precisely,

$$\begin{aligned} \frac{1}{h^k} E\left[\left|\varepsilon_{1r}^0\right|^2 K_{12} \left|g_{rt}(W_2; \varphi_0)\right|^2\right] &= E\left[|g_r(W; \varphi_0)|^2 c_{rr}(X) f(X)\right] + o(1) \\ &\leq \Delta\left(E|g_{rt}(W; \varphi_0)|^{8/3}\right)^{3/4} \left(E|c_{rr}(X)|^4\right)^{1/4} + o(1) \\ &= O(1) \end{aligned}$$

by Lemma 2, equation (2), with $|g_{rt}(W; \varphi_0)|^2 \equiv z(W)$, $|\varepsilon_r(W; \varphi_0)|^2 \equiv m(W)$, and $s = 4$.
Note: here we have just used Assumption B4b and $E|g_{rt}(W; \varphi_0)|^{8/3} < \infty$; we have not used $E[g_{rt}(W; \varphi_0)|X] \in L^{8/3}$, yet; but we have used the Lipschitz condition on $c_{rr}(X)$.
Thus, the conditions of Lemma 1 are satisfied, with $\bar{\mu}_{nt}(W_i) = \frac{1}{2}\varepsilon_{ir}^0 E\left[h^{-k} K_{ij} g_{rt}(W_j; \varphi_0)|X_i\right]$ and $\mu_{nt} = E\left[\bar{\mu}_{nt}(W_i)\right] = E\left[H_n(W_i, W_j)\right] = 0$, and we can write

$$S_{1nt} = \frac{2}{\sqrt{n}} \sum_{i=1}^{n} \bar{\mu}_{nt}(W_i) + o_p(1).$$

Apply Lemma 2, equation (3), with $|g_{rt}(W; \varphi_0)| \equiv z(W)$ and $|\varepsilon_r(W; \varphi_0| = m(W)$, with $s = 4$. That is,

$$\begin{aligned} |\bar{\mu}_{nt}(W_1)| &= \frac{1}{2}\left|\varepsilon_{1r}^0 E\left[h^{-k} K_{12} g_{rt}(W_2; \varphi_0)|X_1\right]\right| \\ &\leq \frac{1}{2}\left|\varepsilon_{1r}^0\right| E\left[h^{-k} K_{12} \left|g_{rt}(W_2; \varphi_0)\right| |X_1\right] \end{aligned}$$

so that

$$
\begin{aligned}
E\left|\bar{\mu}_{nt}(W_1)\right|^2 &\leq \frac{1}{4}E\left\{\left|\varepsilon_{1r}^0\right| E\left[h^{-k}K_{12}\left|g_{rt}(W_2;\varphi_0)\right| |X_1\right]\right\}^2 \\
&= \frac{1}{4}E\left[c_{rr}(X)\gamma_{rt}^2(X)f^2(X)\right]\left\{\int |K(u)|\, du\right\}^2 + o(1) \\
&= O(1)
\end{aligned}
$$

because $E\left[c_{rr}^2(X)\right] < \infty$ and $\gamma_{rt}(X) = E\left[g_{rt}(W;\varphi_0)|X\right] \in L^{8/3}$ (to apply (3), we now need the Lipschitz condition on $\gamma_{rt}(X)$ as well). Thus, by Corollary 1, $S_{1n} = O_p(1)$ and we are done.

(b) $S_{2n} = o_p(1)$:
Appeal to Lemma 1. Write $S_{2n} = \frac{1}{n(n-1)h^k}\sum_{i=1}^n \sum_{j\neq i} \varepsilon_{ir}^0 K_{ij} G_{jr}(\varphi_0)$ and by similar arguments to those put forward in the first part of (a) above (replace g_{rt} by G_r), Assumptions B1-4b , C3(i), and Lemma 1 imply that $\frac{1}{n(n-1)h^k}\sum_{i=1}^n \sum_{j\neq i} \varepsilon_{ir}^0 K_{ij} G_{jr}(\varphi_0) = o_p(1)$. Specfically, $E[S_{1n}] = E[h^{-k}\varepsilon_{ir}^0 K_{ij} G_{jr}(\varphi_0)] = 0$, for all n, since $E[\varepsilon_r(W;\varphi_0)|X] = 0$. Thus, in terms of Lemma 1, $\mu_n = 0$. Write $S_{2n} = \frac{1}{n(n-1)}\sum_{i=1}^n \sum_{j\equiv i} H_n(W_i, W_j)$, where

$$
H_n(W_1, W_2) = \frac{1}{2h^k}\left\{\varepsilon_{1r}^0 K_{12} G_r(W_2;\varphi_0) + \varepsilon_{1r}^0 K_{21} G_r(W_1;\varphi_0)\right\}
$$

and

$$
\begin{aligned}
E\left\|H_n(W_1,W_2)\right\|^2 &\leq \frac{1}{h^{2k}}E\left\|\varepsilon_{1r}^0 K_{12} G_r(W_2;\varphi_0)\right\|^2 \\
&= \frac{1}{h^{2k}}E\left[\left|\varepsilon_{1r}^0\right|^2 K_{12}^2 \left\|G_r(W_2;\varphi_0)\right\|^2\right] \\
&= h^{-k}\left\{E\left[\left\|G_r(W;\varphi_0)\right\|^2 c_{rr}(X)f(X)\right]\int |K(u)|^2\, du + o(1)\right\} \\
&= O(h^{-k}) = O\left(n/\left(nh^k\right)\right) = o(n)
\end{aligned}
$$

because $E\left[\left\|G_r(W;\varphi_0)\right\|^2 c_{rr}(X)f(X)\right] = O(1)$, by Assumption B4b and C3(i). More precisely,

$$
\begin{aligned}
\frac{1}{h^k}E\left[\left|\varepsilon_{1r}^0\right|^2 K_{12}\left\|G_r(W_2;\varphi_0)\right\|^2\right] &= E\left[\left\|G_r(W;\varphi_0)\right\|^2 c_{rr}(X)f(X)\right] + o(1) \\
&\leq \Delta\left(E\left\|G_r(W;\varphi_0)\right\|^{8/3}\right)^{3/4}\left(E\left|c_{rr}(X)\right|^4\right)^{1/4} + o(1) \\
&= O(1)
\end{aligned}
$$

by Lemma 2, equation (2), with $\left\|G_r(W;\varphi_0)\right\|^2 \equiv z(W)$, $\left|\varepsilon_r(W;\varphi_0)\right|^2 \equiv m(W)$, and $s = 4$. Thus, by Lemma 1, $S_{in} \xrightarrow{p} 0$.

(c) $S_{3n} = O_p(1)$
Write $S_{3n}(\varphi) = \frac{1}{n(n-1)h^k}\sum_{i=1}^n \sum_{j\neq i} \varepsilon_{ir}^0 K_{ij} F_{jr}(\varphi)$. Assumptions B1-3, C1 and C4 imply

$$
\begin{aligned}
E\left[\sup_{\varphi}\left\|\frac{1}{n(n-1)h^k}\sum_{i=1}^n\sum_{j\neq i}\varepsilon_{ir}^0 K_{ij} F_{jr}(\varphi)\right\|\right] &\leq \frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j\neq i} E[h^{-k}\left|\varepsilon_{ir}^0\right| K_{ij} P(W_j)] \\
&= O(1)
\end{aligned}
$$

so that $\frac{1}{n(n-1)h^k}\sum_{i=1}^n\sum_{j\neq i}\varepsilon_{ir}^0 K_{ij}F_{jr}(\bar{\varphi}^{(r)}) = O_p(1)$ by Markov's Inequality.
By Lemma 2, equation (2), we have $E\sup_\varphi \|S_{3n}(\varphi)\| = O(1)$, since $\mu(X) = E\left[|\varepsilon_r(W;\varphi_0)|\,|X\right] \in L^2$ by Assumption C1 (we need the Lipschitz condition) and by Assumption C4 $E\left[P^2(W)|X\right] < \infty$

2. T_{2nr} : write

$$\begin{aligned}
\hat{\varepsilon}_{ir} - \varepsilon_{ir}^0 &= g_{ir}(\varphi_0)'(\hat{\varphi}-\varphi_0) + \frac{1}{2}(\hat{\varphi}-\varphi_0)'G_{ir}(\tilde{\varphi}^{(r)})(\hat{\varphi}-\varphi_0)\\
&= g_{ir}(\varphi_0)'(\hat{\varphi}-\varphi_0) + \frac{1}{2}vec(\tilde{G}_{ir})'\left((\hat{\varphi}-\varphi_0)\otimes(\hat{\varphi}-\varphi_0)\right)
\end{aligned}$$

where $\tilde{\varphi}^{(r)}$ is a "mean value" such that $\left\|\tilde{\varphi}^{(r)}-\varphi_0\right\| \leq \|\hat{\varphi}-\varphi_0\| = O_p(n^{-1/2}h^{-k/4})$, and $\tilde{G}_{ir} = G_{ir}(\tilde{\varphi}^{(r)})$. Substituting this expression into T_{2nr} yields

$$\begin{aligned}
T_{2nr} &= \xi_n'\left\{\frac{1}{n(n-1)h^k}\sum_{i=1}^n\sum_{j\neq i} g_{ir}(\varphi_0)K_{ij}g_{jr}(\varphi_0)'\right\}\xi_n + R_n, \quad \text{say}\\
&= \xi_n' S_{1n}\xi_n + R_n
\end{aligned}$$

(a) $\xi_n' S_{1n}\xi_n = E\left[|d_r(X;\varphi_0)'\xi|^2 f(X)\right] + o_p(1)$.
Write $S_{1n} = \frac{1}{n(n-1)h^k}\sum_{i=1}^n\sum_{j\neq i} g_{ir}(\varphi_0)K_{ij}g_{jr}(\varphi_0)'$, and consider a typical element $S_{1nts} = \frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j\neq i} H_n(W_i,W_j)$, where

$$H_n(W_1,W_2) = h^{-k}g_{rt}(W_1;\varphi_0)K_{12}g_{rt}(W_2;\varphi_0)',\quad t,s = 1,...,p,$$

and which is symmetric in i and j. Note that (i) $E[H_n(W_1,W_2)] = E[f(X)d_{rt}(X;\varphi_0)d_{rs}(X;\varphi_0)'] + o(1)$, by Lemma 2, equation (2), which follows from Assumption C2(ii) - the moment bound is slightly stronger than actually required here - and (ii).$E\left|H_n(W_1,W_2)\right|^2 = o(n)$, by Lemma 2, equation (2), which follows from Assumption C2(iv). Thus, by Lemma 1, $S_{1n} = E[f(X)d_r(X;\varphi_0)d_r(X;\varphi_0)'] + o_p(1)$ and the result follows since $\xi_n = \xi + o_p(1) = O_p(1)$.

(b) $R_n = O_p(1)$.
We can express R_n as

$$\begin{aligned}
R_n &= (\hat{\varphi}-\varphi_0)'\left\{\frac{1}{n(n-1)h^k}\sum_{i=1}^n\sum_{j\neq i} g_{ir}(\varphi_0)K_{ij}vec(\tilde{G}_{ir})'\right\}(\xi_n\otimes\xi_n)\\
&\quad + \frac{1}{4nh^{k/2}}(\xi_n\otimes\xi_n)'\left\{\frac{1}{n(n-1)h^k}\sum_{i=1}^n\sum_{j\neq i} vec(\tilde{G}_{ir})K_{ij}vec(\tilde{G}_{ir})'\right\}(\xi_n\otimes\xi_n)\\
&= (\hat{\varphi}-\varphi_0)'R_{1n}(\xi_n\otimes\xi_n) + \frac{1}{4nh^{k/2}}(\xi_n\otimes\xi_n)'R_{2n}(\xi_n\otimes\xi_n)
\end{aligned}$$

and, since $nh^{k/2}\to\infty$, $(\hat{\varphi}-\varphi_0) = o_p(1)$ and $\xi_n = O_p(1)$, all need to show is that $R_{1n} = o_p(1)$ and $R_{2n} = o_p(1)$. This is done by Markov's Inequality. Consider first a typical row of R_{1n}, denoted R_{1nt}, so that by Assumption C3(ii) (with $\|vec(A)\| =$

$\|A\|$)

$$\|R_{1nt}\| \leq \frac{1}{n(n-1)h^k}\sum_{i=1}^{n}\sum_{j\neq i}|g_{rt}(W_i;\varphi_0)|\,K_{ij}\left\|vec(G_r(W_j;\tilde{\varphi}^{(r)}\right\|$$
$$\leq \frac{1}{n(n-1)h^k}\sum_{i=1}^{n}\sum_{j\neq i}|g_{rt}(W_i;\varphi_0)|\,K_{ij}M(W_j).$$

Second, again by Assumption C3(ii),

$$\|R_{2n}\| \leq \left\{\frac{1}{n(n-1)h^k}\sum_{i=1}^{n}\sum_{j\neq i}M(W_i)K_{ij}M(W_j)\right\}.$$

Markov's Inquality ensures that both $\frac{1}{n(n-1)h^k}\sum_{i=1}^{n}\sum_{j\neq i}|g_{rt}(W_i;\varphi_0)|\,K_{ij}M(W_j)$ and $\frac{1}{n(n-1)h^k}\sum_{i=1}^{n}\sum_{j\neq i}M(W_i)K_{ij}M(W_j)$ are $O_p(1)$, by Assumptions B1-3, C2(iii)&3(ii), since

$$E[h^{-k}|g_{rt}(W_1;\varphi_0)|\,K_{12}M(W_2)] = E[E[|g_{rt}(W;\varphi_0)|\,|X]\lambda(X)f(X)]+o(1)$$
$$\leq \Delta\sqrt{E[\{E[|g_{rt}(W;\varphi_0)|\,|X]\}^2]E[\lambda^2(X)]}+o(1)=O(1)$$

by Assumptions C2(iii) and since $E[\lambda^2(X)]<\infty$ (we just need the moment bolund here), and

$$E[h^{-k}M(W_1)K_{12}M(W_2)] = E[E(M^2(W)|X)f(X)]+o(1)$$

by Assumption C3, $\lambda(X)=E[M(W)|X]\in L^2$ and an application of Lemma 2, equation (2).

Thus, $T_{2nr}=E\left[|d_r(X;\varphi_0)'\xi|^2 f(X)\right]+o_p(1)$.

Finally, the Assumptions ensure that $\mu(\varphi_0)=E\left[\|d(X;\varphi_0)\xi\|^2 f(X)\right]=O(1)$, and continuous in φ_0, so that $\lim_{n\to\infty}\mu(\varphi_0)=\mu_0$ exists. Finally, by (1), $nh^{k/2}T_n(\varphi_0)'\iota/\sqrt{\iota'\Omega_0\iota}\xrightarrow{d}N(0,1)$ and due to the local alternatives note that $\lim_{n\to\infty}\iota'\Omega_0\iota=\Sigma_0$. The result then follows, noting that $\sum_{r=1}^{m}|d_r(X;\varphi_0)'\xi|^2=\|d(X;\varphi_0)'\xi\|^2$. ■

3 Technical lemmata

Lemma 1 (Powell, Stock and Stoker, 1989) *Let $U_n=\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}^{n}H_n(W_i,W_j)$ be a U-statistic with $H_n(W_i,W_j)=H_n(W_j,W_i)$ and W_i being independently and identically distributed. Assume $E\|H_n(W_i,W_j)\|^2$ exists for each n, but which may depend upon n, and define $\bar{\mu}_n(W_i)=E[H_n(W_i,W_j)|W_i]$, $\mu_n=E[\bar{\mu}_n(W_i)]=E[H_n(W_i,W_j)]$ and $\tilde{U}_n=\mu_n+\frac{2}{n}\sum_{i=1}^{n}[\bar{\mu}_n(W_i)-\mu_n]$. If $E\|H_n(W_i,W_j)\|^2=o(n)$, then*

$$U_n = \tilde{U}_n+o_p(n^{-1/2})$$
$$U_n = \mu_n+o_p(1).$$

Remark 1 *The above shows that $\sqrt{n}(U_n-\tilde{U}_n)=o_p(1)$, so that $\sqrt{n}(U_n-\mu_n)=\sqrt{n}(\tilde{U}_n-\mu_n)+o_p(1)$, where $\sqrt{n}(\tilde{U}_n-\mu_n)=\frac{2}{\sqrt{n}}\sum_{i=1}^{n}[\bar{\mu}_n(W_i)-\mu_n]$.*

We thus have the following Corollary:

Corollary 1 *If* U_n, $\bar{\mu}_n(W_i)$ *and* μ_n *are as in Lemma 1, and* $E\left\|\bar{\mu}_n(W_i)\right\|^2 = E\left\|E\left[H_n(W_i,W_j)|W_i\right]\right\|^2 = O(1)$*, then* $\sqrt{n}(U_n-\mu_n) = O_p(1)$*, with* $\sqrt{n}(U_n-\mu_n) \xrightarrow{d} N(0,V)$, $V = 4\lim_{n\to\infty} var\left(\bar{\mu}_n(W_i)-\mu_n\right)$.

Proof. *First,and by Chebyshev's Inequality, provided* $E\left\|\bar{\mu}_n(W_i)-\mu_n\right\|^2 = O(1)$*, then* $\sqrt{n}(U_n - \mu_n) = O_p(1)$*. Note that*

$$\begin{aligned} E\left\|\bar{\mu}_n(W_i)-\mu_n\right\|^2 &= E\left[(\bar{\mu}_n(W_i)-\mu_n)'(\bar{\mu}_n(W_i)-\mu_n)\right] \\ &= E\left\|\bar{\mu}_n(W_i)\right\|^2 - \left\|\mu_n\right\|^2 \\ &= O(1) \end{aligned}$$

since $E\left\|\bar{\mu}_n(W_i)\right\|^2 = O(1)$*, by assumption, and* $E\left\|\bar{\mu}_n(W_i)\right\|^2 \geq \mu_n^2$*, by Jensen's Inequality. (Note, also by Jensen's Inequality, that since* $E(\left\|H_n(W_i,W_j)\right\|^2|W_i) \geq \left\|E[H_n(W_i,W_j)|W_i]\right\|^2$*,*

$$o(n) = E\left\|H_n(W_i,W_j)\right\|^2 = E\left[E(\left\|H_n(W_i,W_j)\right\|^2|W_i)\right] \geq E\left\|\bar{\mu}_n(W_i)\right\|^2$$

so that, in fact, $E\left\|H_n(W_i,W_j)\right\|^2 = o(n)$ *is insufficient for* $E\left\|\bar{\mu}_n(W_i)\right\|^2 = O(1)$*, and we therefore need the extra condition.) With this condition,* $var\left(\bar{\mu}_n(W_i)-\mu_n\right) = O(1)$ *and the iid assumption ensures that the limit exists.* ■

Here we state and prove a fundamental Lemma. This exploits Lemma 3 which immediately follows it.

Lemma 2 *Let* $m(W)$ *and* $z(W)$ *be scalar random variables, and define for any* $r \geq 1$

$$Q_n(W_1,W_2) = h^{-k}m(W_1)K_{12}^r z(W_2).$$

In addition to Assumption A, assume the following conditions hold, where $s \geq 1$ *:*

1. $f(x) \in L^\infty$*, implying that* $\sup_x f(x) \leq \Delta < \infty$*.*
2. $\eta(X) = E\left[m(W)|X\right] \in L^s$*. (The condition of* $\eta(X) \in L^s$ *is not only important becasue of bounded moments, it also gives the required local Lipschitz (in x) condition on* $\eta(x)$*.)*
3. $E\left[|z(W)|^{s/(s-1)}\right] < \infty$ *or* $z(W)$ *is bounded if* $s=1$ *(note that* $s/(s-1) > 1$*).*
   Under Assumption A and Assumptions 1-3 above,

$$\begin{aligned} E\left[Q_n(W_1,W_2)\right] &= E\left\{z(W)\eta(X)f(X)\right\}\int K^r(u)\,du + o(1) = O(1) \qquad (2) \\ &= E\left\{\gamma(X)\eta(X)f(X)\right\}\int K^r(u)\,du + o(1) = O(1) \end{aligned}$$

where $\gamma(X) = E[z(W)|X]$*, which exists a.s. since* $E\left|z(W)\right| < \infty$*. In fact, the proof of this indicates that a slightly weaker condition is that* $E\left|\gamma(X)\right|^{s/(s-1)} < \infty$*. Since* $v = s/(s-1) > 1$*, Jensen's Inequality implies that*

$$E[|z(W)|^v] = E\left\{E[|z(W)|^v|X]\right\} \geq E\left\{E[|z(W)||X]\right\}^v \geq E\left\{|E[z(W)|X]|\right\}^v = E\left|\gamma(X)\right|^v$$

so that $E\left|z(W)\right|^{s/(s-1)} < \infty \implies E\left|\gamma(X)\right|^{s/(s-1)} < \infty$*.*

Remark 2 *If* $m(W) \equiv z(W)$*, then set* $s=2$ *and will require just* $\eta(X) = E\left[m(W)|X\right] \in L^2$*.*

4. *Suppose* $\tau(X) = E[m^2(W)|X] \in L^s$ *and* $\gamma(X) = E[z(W)|X] \in L^{2s/(s-1)}$.
Define $q_n(W_1) = E\left[Q_n(W_1, W_2)|W_1\right]$, *then under Assumption A and Assumptions 1-4 above,*
$$E\left|q_n(W_1)\right|^2 = E\left[\tau(X)\gamma^2(X)f^2(X)\right]\left\{\int K^r(u)du\right\}^2 + o(1). \tag{3}$$
Actually, for the above, all we need is $E\left|\tau(X)\right|^s < \infty$ *and* $\gamma(X) \in L^{2s/(s-1)}$
We need, though, both $\tau(X) = E[m^2(W)|X] \in L^s$ *and* $\gamma(X) \in L^{2s/(s-1)}$ *for the following:*
Define $P_n(W_1, W_2) = E\left[Q_n(W_3, W_1)Q_n(W_3, W_2)|W_1, W_2\right]$, *the under Assumption A and Assumptions 1-4 above*
$$E\left[P_n(W_1, W_2)\right] = E\left[\tau(X)\gamma^2(X)f^2(X)\right]\int\int K^r(u)\, K^r(u-v)dudv + o(1). \tag{4}$$

5. *And we need both* $\tau(X) = E[m^2(W)|X] \in L^{2s}$ *(* $\Longrightarrow$ $\tau(X) = E[m^2(W)|X] \in L^s$*) and* $\lambda(X) = E\left[z^2(W)|X\right] \in L^{2s/(s-1)}$ *for the following:*
If $P_n(W_1, W_2) = E\left[Q_n(W_3, W_1)Q_n(W_3, W_2)|W_1, W_2\right]$, *as above then under Assumption A and Assumptions 1-5 above*
$$h^k E\left|P_n(W_1, W_2)\right|^2 = E\left[\tau^2(X)\lambda^2(X)f^3(X)\right]\int\left\{\int K^r(u)\, K^r(u-v)du\right\}^2 dv + o(1). \tag{5}$$

6. *Let* $m_r(W)$ *and* $z_r(W)$ *be scalar random variables,* $r = 1, 2$, *satisfying* $\tau_{rs}(X) = E[m_r(W)m_s(W)|X] \in L^{2s}$ *and* $\lambda_{rs}(X) = E\left[z_r(W)z_s(W)|X\right] \in L^{2s/(s-1)}$.
Define $P_n(W_1, W_2) = E\left[Q_{1n}(W_3, W_1)Q_{2n}(W_3, W_2)|W_1, W_2\right]$, *where*
$$Q_{1n}(W_1, W_2) = h^{-k}m_1(W_1)K_{12}^r z_1(W_2)$$
and
$$Q_{2n}(W_1, W_2) = h^{-k}m_2(W_1)K_{12}^r z_2(W_2),$$
then under Assumption A and Assumptions 1-6 above
$$h^k E\left|P_n(W_1, W_2)\right|^2 = E\left[\tau_{12}^2(X)\lambda_{11}(X)\lambda_{22}(X)f^3(X)\right]\int\left\{\int K^r(u)\, K^r(u-v)du\right\}^2 dv + o(1). \tag{6}$$

Proof. *To prove (2), we can write (by independence)*
$$\begin{aligned} E\left[Q_n(W_1, W_2)\right] &= h^{-k}E\left[m(W_1)K_{12}^r z(W_2)\right] \\ &= h^{-k}E\left\{z(W_2)K_{12}^r E\left[m(W_1)|X_1, W_2\right]\right\} \\ &= h^{-k}E\left\{z(W_2)K_{12}^r E\left[m(W_1)|X_1\right]\right\} \\ &= h^{-k}E\left\{z(W_2)K_{12}^r \eta(X_1)\right\} \\ &= h^{-k}E\left\{z(W_2)E\left[K_{12}^r \eta(X_1)|W_2\right]\right\} \\ &= h^{-k}E\left\{z(W_2)E\left[K_{12}^r \eta(X_1)|X_2\right]\right\} \\ &= h^{-k}E\left\{z(W)\int\left\{K\left(\frac{x-X}{h}\right)\right\}^r \eta(x)f(x)dx\right\}. \end{aligned}$$
Now, $\eta(X) = E\left[m(W)|X\right] \in L^s$, $s \geq 1$, *implies* $\int\left|\eta(x)\right| f(x)dx < \infty$. *Furthermore, for all* $x \in \mathbb{R}^p$
$$\begin{aligned} & \sup_{\|d\|<\delta}\left|\eta(x+d)f(x+d) - \eta(x)f(x)\right| / \|d\| \\ = & \sup_{\|d\|\leq\delta}\left|\left(\eta(x+d)-\eta(x)\right)f(x+d) + \left(f(x+d)-f(x)\right)\eta(x)\right| / \|d\| \\ \leq & \sup_{\|d\|\leq\delta}\frac{\left|\eta(x+d)-\eta(x)\right|}{\|d\|}f(x+d) + \sup_{\|d\|\leq\delta}\frac{\left|f(x+d)-f(x)\right|}{\|d\|}\left|\eta(x)\right| \\ \leq & BL(x) + B\left|\eta(x)\right| \end{aligned}$$

where $B = \max\left[\sup_x f(x), \sup_{\|d\|\le\delta} \dfrac{|f(x+d)-f(x)|}{\|d\|}\right]$. *Thus, by an application of Lemma 3, we have*

$$\begin{aligned}
&\left|h^{-k}\int\left\{K\left(\frac{x-X}{h}\right)\right\}^r \eta(x)f(x)dx - \eta(X)f(X)\int K^r(u)\,du\right| \\
\le\;& \delta B\left\{L(X)+|\eta(X)|\right\}\int_{\|u\|\le\delta/h}|K(u)|^r\,du \\
&+\sup_{\|u\|>\delta/h}\|u\|^p|K(u)|^r\frac{1}{\delta^k}\int|\eta(x)|f(x)dx + |\eta(X)|f(X)\int_{\|u\|>\delta/h}|K(u)|^r\,du.
\end{aligned}$$

Thus, since

$$\begin{aligned}
&\left|E\left[Q_n(W_1,W_2)\right] - E\left\{z(W)\eta(X)f(X)\right\}\int K^r(u)\,du\right| \\
=\;& \left|E\left[z(W)h^{-k}\int\left\{K\left(\frac{x-X}{h}\right)\right\}^r\eta(x)f(x)dx - z(W)\eta(X)f(X)\int K^r(u)\,du\right]\right| \\
\le\;& E\left|z(W)h^{-k}\int\left\{K\left(\frac{x-X}{h}\right)\right\}^r\eta(x)f(x)dx - z(W)\eta(X)f(X)\int K^r(u)\,du\right| \\
=\;& E\left[|z(W)|\left|h^{-k}\int\left\{K\left(\frac{x-X}{h}\right)\right\}^r\eta(x)f(x)dx - z(W)\eta(X)f(X)\int K^r(u)\,du\right|\right],
\end{aligned}$$

we obtain

$$\begin{aligned}
&\left|E\left[Q_n(W_1,W_2)\right] - E\left\{z(W)\eta(X)f(X)\right\}\int K^r(u)\,du\right| \\
\le\;& \delta BE\left\{|z(W)|L(X)+|z(W)\eta(X)|\right\}\int_{\|u\|\le\delta/h}|K(u)|^r\,du \\
&+\sup_{\|u\|>\delta/h}\|u\|^p|K(u)|\frac{1}{\delta^k}E\left[|z(W)|\right]E[|\eta(X)|] \\
&+E\left\{|z(W)\eta(X)|f(X)\right\}\int_{\|u\|>\delta/h}|K(u)|^r\,du.
\end{aligned}$$

Now, $E\left[|z(W)|\right] < \infty$ *and* $E[|\eta(X)|] < \infty$ *by Assumptions 2 & 3, and* $E\left[|z(W)\eta(X)|\right]$, $E\left[|z(W)|L(X)\right]$ *and* $E[|z(W)\eta(X)|f(X)]$ *are all bounded, by Hölder's Inequality. For example, by Assumptions 1-3,*

$$\begin{aligned}
E\left\{|z(W)\eta(X)|f(X)\right\} &\le B\times E|z(W)\eta(X)| \\
&\le B\times\left(E|z(W)|^{s/(s-1)}\right)^{(s-1)/s}\left(E|\eta(X)|^s\right)^{1/s} = O(1).
\end{aligned}$$

Therefore, by iterative expectations,

$$\begin{aligned}
E\left[Q_n(W_1,W_2)\right] &= E\left\{z(W)\eta(X)f(X)\right\}\int K^r(u)\,du + o(1) \\
&= E\left\{\gamma(X)\eta(X)f(X)\right\}\int K^r(u)\,du + o(1).
\end{aligned}$$

This, therefore, also establishes the existence of $q_n(W_1) = E\left[Q_n(W_1,W_2)|W_1\right]$, *a.s..*
The proof of (3) follows in a straightforward way. As notes, Assumptions 1-3 establish the

existence of $q_n(W_1) = E\left[Q_n(W_1, W_2)|W_1\right]$, *almost surely, and*

$$\begin{aligned}
E\left[Q_n(W_1, W_2)|W_1\right] &= m(W_1)E\left[h^{-k}K_{12}^r z(W_2)|W_1\right] \\
&= m(W_1)E\left[h^{-k}K_{12}^r z(W_2)|X_1\right] \\
&= m(W_1)E\left[h^{-k}K_{12}^r \left\{E\left[z(W_2)|X_1, X_2\right]\right\}|X_1\right] \\
&= m(W_1)E\left[h^{-k}K_{12}^r \left\{E\left[z(W_2)|X_2\right]\right\}|X_1\right] \\
&= m(W_1)E\left[h^{-k}K_{12}^r \gamma(X_2)|X_1\right] \\
&= m(W_1)h^{-k}\int\left\{K\left(\frac{X_1 - x}{h}\right)\right\}^r \gamma(x)f(x)dx. \\
&= m(W)h^{-k}\int\left\{K\left(\frac{X - x}{h}\right)\right\}^r \gamma(x)f(x)dx.
\end{aligned}$$

Substituting $u = \dfrac{X - x}{h}$, *for* x, *yields*

$$\begin{aligned}
h^{-k}\int\left\{K\left(\frac{X - x}{h}\right)\right\}^r \gamma(x)f(x)dx &= \int K^r(u)\gamma(X - uh)f(X - uh)du \\
&= \gamma(X)f(X)\int K^r(u)du + t_n(X), \quad say.
\end{aligned}$$

where $|t_n(X)|$, *by Lemma 3, is linear in* $L(X)$, $\gamma(X)$ *and* $\gamma(X)f(X)$. *Specifically,*

$$\begin{aligned}
|t_n(X)| &= \left|h^{-k}\int\left|K\left(\frac{X - x}{h}\right)\right|^r \gamma(x)f(x)dx - \gamma(X)f(X)\int\left|K(u)\right|^r du\right| \\
&\leq \delta B\left\{L(X) + |\gamma(X)|\right\}\int_{\|u\|\leq\delta/h}|K(u)|\,du \\
&\quad + \sup_{\|u\|>\delta/h}\|u\|^p\,|K(u)|^r\,\frac{1}{\delta^k}E\left[|\gamma(X)|\right] \\
&\quad + |\gamma(X)|\,f(X)\int_{\|u\|>\delta/h}|K(u)|^r\,du.
\end{aligned}$$

Thus,

$$\begin{aligned}
q_n^2(W) &= \left\{m(W)\gamma(X)f(X)\int K^r(u)du + m(W)t_n(X)\right\}^2 \\
&\leq |m(W)\gamma(X)f(X)|^2\left\{\int K^r(u)du\right\}^2 \\
&\quad + 2\left|m^2(W)\gamma(X)t_n(X)f(X)\right|\left\{\int K^r(u)du\right\} + |m(W)t_n(X)|^2
\end{aligned}$$

so that

$$\begin{aligned}
E\left|q_n(W)\right|^2 &= E\left|m(W)\gamma(X)f(X)\right|^2\left\{\int K^r(u)du\right\}^2 + o(1) \\
&= E\left[\tau(X)\gamma^2(X)f^2(X)\right]\left\{\int K^r(u)du\right\}^2 + o(1)
\end{aligned}$$

provided $E\left[m^2(W)L^2(X)\right]$, $E\left[m^2(W)\gamma^2(X)\right]$, $E\left[m^2(W)\gamma(X)L(X)\right]$ *are all bounded, since* $f(X)$ *is bounded. By Hölder's Inequality, and Assumption 4,*

$$E\left[m^2(W)\gamma(X)L(X)\right] \leq \left\{E\left|m^2(W)\right|^s\right\}^{1/s}\left\{E\left|\gamma(X)L(X)\right|^{s/(s-1)}\right\}^{(s-1)/s} < \infty$$

since $E\left|m^2(W)\right|^s = E\left|m(W)\right|^{2s} < \infty$, $E\left|\gamma(X)L(X)\right|^{s/(s-1)} < \infty$ *by Cauchy-Schwartz, and*

$$\begin{aligned} E\left[m^2(W)\gamma^2(X)\right] &\leq \left\{E\left|m^2(W)\right|^s\right\}^{1/s}\left\{E\left|\gamma(X)\right|^{2s/(s-1)}\right\}^{(s-1)/s} < \infty \\ E\left[m^2(W)L^2(X)\right] &\leq \left\{E\left|m^2(W)\right|^s\right\}^{1/s}\left\{E\left|L(X)\right|^{2s/(s-1)}\right\}^{(s-1)/s} < \infty. \end{aligned}$$

This suffices.
To prove (4), we have

$$\begin{aligned} P_n(W_1,W_2) &= E\left[Q_n(W_3,W_1)Q_n(W_3,W_2)|W_1,W_2\right] \\ &= h^{-2k}z(W_1)z(W_2)E\left[m^2(W_3)K_{31}^rK_{32}^r|X_1,X_2\right] \end{aligned}$$

so that

$$\begin{aligned} E\left[P_n(W_1,W_2)\right] &= h^{-2k}E\left[\gamma(X_1)\gamma(X_2)E\left[\tau(X_3)K_{31}^rK_{32}^r|X_1,X_2\right]\right] \\ &= h^{-2k}E\left[\gamma(X_1)E\left[\gamma(X_2)E\left[\tau(X_3)K_{31}^rK_{32}^r|X_1,X_2\right]|X_1\right]\right] \\ &= h^{-2k}E\left[\gamma(X_1)\int\int\gamma(z)\left\{K\left(\frac{x-X_1}{h}\right)K\left(\frac{x-z}{h}\right)\right\}^r\tau(x)f(x)f(z)dxdz\right] \\ &= h^{-2k}E\left[\gamma(X)\int\int\gamma(z)f(z)\tau(x)f(x)\left\{K\left(\frac{x-X}{h}\right)K\left(\frac{x-z}{h}\right)\right\}^r dxdz\right] \end{aligned}$$

Thus, by an application of Lemma 3, part 2, we have

$$E\left[P_n(W_1,W_2)\right] = E\left[\gamma^2(X)\tau(X)f^2(X)\right]\int\left\{\int K^r(u)\,K^r(u-v)du\right\}dv + o(1)$$

provided $E\left[\gamma^2(X)\tau(X)\right] < \infty$*, which it is by Hölder's Inequality:*

$$E[\gamma^2(X)\tau(X)] \leq \left\{E\left|\tau(X)\right|^s\right\}^{1/s}\left\{E\left|\gamma(X)\right|^{2s/(s-1)}\right\}^{(s-1)/s} < \infty.$$

To prove (5), we have

$$\begin{aligned} P_n(W_1,W_2) &= E\left[Q_n(W_3,W_1)Q_n(W_3,W_2)|W_1,W_2\right] \\ &= h^{-2k}z(W_1)z(W_2)E\left[m^2(W_3)K_{31}^rK_{32}^r|X_1,X_2\right] \end{aligned}$$

so that

$$\begin{aligned} h^kE\left|P_n(W_1,W_2)\right|^2 &= h^{-3k}E\left[\lambda(X_1)\lambda(X_2)\left\{E\left[\tau(X_3)K_{31}^rK_{32}^r|X_1,X_2\right]\right\}^2\right] \\ &= h^{-3k}E\left[\lambda(X_1)E\left(\lambda(X_2)\left\{E\left[\tau(X_3)K_{31}^rK_{32}^r|X_1,X_2\right]\right\}^2|X_1\right)\right] \\ &= h^{-3k}E\left[\lambda(X_1)\int\lambda(z)\left\{\int\left\{K\left(\frac{x-X_1}{h}\right)K\left(\frac{x-z}{h}\right)\right\}^r\tau(x)f(x)dx\right\}^2 f(z)dz\right] \\ &= h^{-3k}E\left[\lambda(X)\int\lambda(z)f(z)\left\{\int\tau(x)f(x)\left\{K\left(\frac{x-X}{h}\right)K\left(\frac{x-z}{h}\right)\right\}^r dx\right\}^2 dz\right]. \end{aligned}$$

Thus, by an application of Lemma 3, part 2, we have

$$h^kE\left|P_n(W_1,W_2)\right|^2 = E\left[\lambda^2(X)\tau^2(X)f^3(X)\right]\int\left\{\int K^r(u)\,K^r(u-v)du\right\}^2 dv + o(1)$$

provided $E\left[\lambda^2(X)\tau^2(X)\right] < \infty$*, which it is by Hölder's Inequality :*

$$E[\lambda^2(X)\tau^2(X)] \leq \left\{E\left|\tau(X)\right|^{2s}\right\}^{1/s}\left\{E\left|\lambda(X)\right|^{2s/(s-1)}\right\}^{(s-1)/s} < \infty.$$

The result in (6) is a straightforward extension of the previous result. We have

$$\begin{aligned} P_n(W_1, W_2) &= E\left[Q_{1n}(W_3, W_1)Q_{2n}(W_3, W_2)|W_1, W_2\right] \\ &= h^{-2k} z_1(W_1) z_2(W_2) E\left[m_1(W_3)m_2(W_3)K_{31}^r K_{32}^r|X_1, X_2\right] \end{aligned}$$

so that

$$\begin{aligned} h^k E\left|P_n(W_1, W_2)\right|^2 &= h^{-3k} E\left[\lambda_1(X_1)\lambda_2(X_2)\left\{E\left[\tau_{12}(X_3)K_{31}^r K_{32}^r|X_1, X_2\right]\right\}^2\right] \\ &= h^{-3k} E\left[\lambda_1(X_1) E\left(\lambda_2(X_2)\left\{E\left[\tau_{12}(X_3)K_{31}^r K_{32}^r|X_1, X_2\right]\right\}^2 |X_1\right)\right] \\ &= h^{-3k} E\left[\lambda_1(X_1)\int \lambda_2(z)\left\{\int\left\{K\left(\frac{x - X_1}{h}\right)K\left(\frac{x - z}{h}\right)\right\}^r \tau_{12}(x)f(x)dx\right\}^2 f(z)dz\right] \\ &= h^{-3k} E\left[\lambda_1(X)\int \lambda_2(z)f(z)\left\{\int \tau_{12}(x)f(x)\left\{K\left(\frac{x - X}{h}\right)K\left(\frac{x - z}{h}\right)\right\}^r dx\right\}^2 dz\right] \end{aligned}$$

Thus, by an application of Lemma 3, part 2, we have

$$h^k E\left|P_n(W_1, W_2)\right|^2 = E\left[\lambda_1(X)\lambda_2(X)\tau_{12}^2(X)f^3(X)\right]\int\left\{\int K^r(u)\,K^r(u - v)du\right\}^2 dv + o(1)$$

provided $E\left[\lambda_1(X)\lambda_2(X)\tau_{12}^2(X)\right] < \infty$, *which it is by Hölder's Inequality :*

$$E[\lambda_1(X)\lambda_2(X)\tau_{12}^2(X)] \le \left\{E\left|\tau_{12}(X)\right|^{2s}\right\}^{1/s}\left\{E\left|\lambda_1(X)\lambda_2(X)\right|^{s/(s-1)}\right\}^{(s-1)/s} < \infty$$

and, by Cauchy-Schwartz, $E\left|\lambda_1(X)\lambda_2(X)\right|^{s/(s-1)} \le \left\{E\left|\lambda_1(X)\right|^{2s/(s-1)} E\left|\lambda_2(X)\right|^{2s/(s-1)}\right\}^{1/2} < \infty$. ■

Lemma 3 *Define the class of functions* S *to be* $s(.) : \mathbb{R}^k \to \mathbb{R}$ *such that for all* $x \in \mathbb{R}^k$, $\sup_{\|d\|\le\delta}|s(x + d) - s(x)|/\|d\| \le S < \infty$, *for some* $\delta > 0$. *Note that, since this Lipschitz condition holds for all* $x \in \mathbb{R}^p$, *this implies* $\int|s(x)|\,dx < \infty$, *and that that* $s(x)$ *is bounded:* $\sup_x|s(x)| \le B < \infty$.

1. *If* $g \in S$, *then, for all* $r \ge 1$,

$$h^{-k}\int\left\{K\left(\frac{x - x_0}{h}\right)\right\}^r g(x)dx = g(x_0)\int K^r(u)\,du + o(1),$$

where $K^r(u) \equiv \{K(u)\}^r$; *or, more precisely, for any* $h > 0$ *and* $\delta > 0$

$$\begin{aligned} \left|h^{-k}\int\left\{K\left(\frac{x - x_0}{h}\right)\right\}^r g(x)dx - g(x_0)\int K^r(u)\,du\right| &\le \delta S\int_{\|u\|\le\delta/h}|K(u)|^r du \\ &\quad + \sup_{\|u\|>\delta/h}\|u\|^p|K(u)|^r\frac{1}{\delta^k}\int|g(x)|\,dx \\ &\quad + |g(x_0)|\int_{\|u\|>\delta/h}|K(u)|^r du. \end{aligned}$$

2. *If both* $a \in S$ *and* $g \in S$, *then for all* $r \ge 1$

$$\begin{aligned} & h^{-2p}\int a(z)\int\left\{K\left(\frac{x - x_0}{h}\right)K\left(\frac{x - z}{h}\right)\right\}^r g(x)dxdz \\ &= a(x_0)g(x_0)\int\int K^r(u)\,K^r(u - v)dudv + o(1). \\ & h^{-3p}\int a(z)\left\{\int\left\{K\left(\frac{x - x_0}{h}\right)K\left(\frac{x - z}{h}\right)\right\}^r g(x)dx\right\}^2 dz \\ &= a(x_0)g^2(x_0)\int\left\{\int K^r(u)\,K^r(u - v)du\right\}^2 dv + o(1). \end{aligned}$$

Proof.

1. *Making the substitution* $u = (x - x_0)/h$ *we have*

$$I_n(x_0) \equiv h^{-k} \int \left\{ K\left(\frac{x - x_0}{h} \right) \right\}^r g(x)dx = \int K^r(u) g(x_0 + uh) du$$

so that

$$\begin{aligned}
\left| I_n(x_0) - g(x_0) \int K(u)\, du \right| &= \left| \int K^r(u)(g(x_0 + uh) - g(x_0))du \right| \\
&\leq \int |K(u)|^r \, |g(x_0 + uh) - g(x_0)| \, du \\
&= \int_{\|u\| \leq \delta/h} |K(u)|^r \, |g(x_0 + uh) - g(x_0)| \, du \\
&\quad + \int_{\|u\| > \delta/h} |K(u)|^r \, |g(x_0 + uh) - g(x_0)| \, du \\
&= T_{1n} + T_{2n}, \quad say.
\end{aligned}$$

We show that $T_{jn} = o(1)$, $j = 1, 2$. *Now,*

$$\begin{aligned}
T_{1n} &\leq \int_{\|u\| \leq \delta/h} \|uh\| \, |K(u)|^r \, \frac{|g(x_0 + uh) - g(x_0)|}{\|uh\|} du \\
&\leq \sup_{\|uh\| < \delta} \frac{|g(x_0 + uh) - g(x_0)|}{\|uh\|} \int_{\|u\| \leq \delta/h} \|uh\| \, |K(u)|^r \, du \\
&\leq \delta S \int_{\|u\| \leq \delta/h} |K(u)|^r \, du
\end{aligned}$$

which is bounded for any $\delta > 0$ *and* $h > 0$. *Turning now to* T_{2n}, *we have*

$$\begin{aligned}
T_{2n} &\leq \int_{\|u\| > \delta/h} |K(u)|^r \, |g(x_0 + uh)| \, du + |g(x_0)| \int_{\|u\| > \delta/h} |K(u)|^r \, du \\
&= \int_{\|u\| > \delta/h} \|u\|^p \, |K(u)|^r \left| \frac{g(x_0 + uh)}{\|u\|^p} \right| du + |g(x_0)| \int_{\|u\| > \delta/h} |K(u)|^r \, du \\
&\leq \sup_{\|u\| > \delta/h} \|u\|^p \, |K(u)|^r \int_{\|u\| > \delta/h} \left| \frac{g(x_0 + uh)}{\|u\|^p} \right| du + |g(x_0)| \int_{\|u\| > \delta/h} |K(u)|^r \, du \\
&\leq \sup_{\|u\| > \delta/h} \|u\|^p \, |K(u)|^r \, \frac{h^k}{\delta^k} \int_{\|u\| > \delta/h} |g(x_0 + uh)| \, du + |g(x_0)| \int_{\|u\| > \delta/h} |K(u)|^r \, du \\
&= \sup_{\|u\| > \delta/h} \|u\|^p \, |K(u)|^r \, \frac{1}{\delta^k} \int_{\|x - x_0\| > \delta} |g(x)| \, dx + |g(x_0)| \int_{\|u\| > \delta/h} |K(u)|^r \, du \\
&\leq \sup_{\|u\| > \delta/h} \|u\|^p \, |K(u)|^r \, \frac{1}{\delta^k} \int |g(x)| \, dx + |g(x_0)| \int_{\|u\| > \delta/h} |K(u)|^r \, du
\end{aligned}$$

where the third inequality follows since, for $p \geq 1$, $1/\|u\|^p < h^k/\delta^k$, *the second equality follows by making the substitution* $x = x_0 + uh$, *and the last inequality because*

$$\int_{\|x - x_0\| > \delta} |g(x)| \, dx \leq \int |g(x)| \, dx.$$

Putting all of this together, we have for any $h > 0$ and $\delta > 0$

$$\begin{aligned}
\left| h^{-k} \int \left| K\left(\frac{x-x_0}{h}\right)\right|^r g(x)dx - g(x_0)\int |K(u)|^r du \right| &\leq \delta S \int_{\|u\|\leq \delta/h} |K(u)|^r du \\
&+ \sup_{\|u\|>\delta/h} \|u\|^p |K(u)|^r \frac{1}{\delta^k} \int |g(x)|\, dx \\
&+ |g(x_0)| \int_{\|u\|>\delta/h} |K(u)|^r du.
\end{aligned}$$

Since this is true for any $\delta > 0$, we now just let $h \to 0$ (as $n \to \infty$): the first integral is bounded by Assumption A2; $\sup_{\|u\|>\delta/h} \|u\|^p |K(u)|^r \to 0$, by Assumption A3; and, $\int_{\|u\|>\delta/h} |K(u)|^r du \to 0$. Then letting $\delta \to 0$, gives the result, noting that $\int |g(x)|\, dx$ and $|g(x_0)|$ are both bounded.

2. *For the first result, making the substitution $u = (x - x_0)/h$ and then $v = (z - x_0)/h$ we obtain*

$$\begin{aligned}
I_n(x_0) &\equiv h^{-k} \int a(z) \left\{ h^{-k} \int \left\{ K\left(\frac{x-x_0}{h}\right) K\left(\frac{x-z}{h}\right)\right\}^r g(x)dx \right\} dz \\
&= \int a(x_0+vh) \left\{ \int K^r(u)K^r(u-v)g(x_0+uh)du \right\} dv.
\end{aligned}$$

Writing

$$\begin{aligned}
\int K^r(u)K^r(u-v)g(x_0+uh)du &= g(x_0)\int K^r(u)K^r(u-v)du \\
&+ \int K^r(u)K^r(u-v)\{g(x_0+uh)-g(x_0)\} du \\
&= g(x_0)J(v) + R_n(x_0,v), \quad \text{say,}
\end{aligned}$$

we obtain

$$\begin{aligned}
\left| I_n(x_0) - a(x_0)g(x_0)\int\int K^r(u) K^r(u-v)dudv \right| &= \left| \int a(x_0+vh)\{g(x_0)J(v)+R_n(x_0,v)\} dv \right. \\
&\left. - a(x_0)g(x_0)\int J(v)dv \right| \\
&\leq \left| g(x_0)\int \{a(x_0+vh)-a(x_0)\} J(v)dv \right| \\
&+ \left| \int a(x_0+vh)R_n(x_0,v)dv \right| \\
&\leq B\int |a(x_0+vh)-a(x_0)|\,|J(v)|\, dv \\
&+ B\int |R_n(x_0,v)|\, dv \\
&= T_{1n} + T_{2n}
\end{aligned}$$

where we have exploited the fact that both $g(.)$ and $a(.)$ are bounded. We now show that the $T_{jn} = o(1)$, $j = 1,2$.
Firstly, note that $J(v) = \int K^r(u)K^r(u-v)du$ satisfies the conditions placed on $K(u)$ in Assumptions 1-3 so that, since $a \in S$, $T_{1n} = o(1)$ which follows immediately from the proof of part 1 above.

Second, we show that $\sup_v |R_n(x_0,v)| = o(1)$*, which suffices for* $T_{2n} = o(1)$.

$$\begin{aligned} |R_n(x_0,v)| &\le \int |K^r(u)K^r(u-v)| \, |g(x_0+uh) - g(x_0)| \, du \\ &\le \Delta^r \int |K(u)|^r \, |g(x_0+uh) - g(x_0)| \, du \end{aligned}$$

which is independent of v *and where* $\Delta = \sup_z |K(z)|$. *Again, from part 1,*

$$\int |K(u)|^r \, |g(x_0+uh) - g(x_0)| \, du = o(1)$$

so that $|R_n(x_0,v)| = o(1)$ *uniformly in* v.
Similarly, for the second result, making the substitution $u = (x - x_0)/h$ *we have and then* $v = (z - x_0)/h$ *we obtain*

$$\begin{aligned} I_n(x_0) &\equiv h^{-k} \int a(z) \left\{ h^{-k} \int \left\{ K\left(\frac{x-x_0}{h}\right) K\left(\frac{x-z}{h}\right) \right\}^r g(x)dx \right\}^2 dz \\ &= \int a(x_0+vh) \left\{ \int K^r(u)K^r(u-v)g(x_0+uh)du \right\}^2 dv. \end{aligned}$$

Writing

$$\begin{aligned} \int K^r(u)K^r(u-v)g(x_0+uh)du &= g(x_0) \int K^r(u)K^r(u-v)du \\ &\quad + \int K^r(u)K^r(u-v) \left\{ g(x_0+uh) - g(x_0) \right\} du \\ &= g(x_0)J(v) + R_n(x_0,v), \quad \textit{say}, \end{aligned}$$

we obtain

$$\begin{aligned} &\left| I_n(x_0) - a(x_0)g^2(x_0) \int \left\{ \int K^r(u) K^r(u-v)du \right\}^2 dv \right| \\ &= \left| \int a(x_0+vh) \left\{ g(x_0)J(v) + R_n(x_0,v) \right\}^2 dv \, - a(x_0)g^2(x_0) \int J^2(v)dv \right| \\ &\le \left| g^2(x_0) \int \left\{ a(x_0+vh) - a(x_0) \right\} J^2(v)dv \right| \\ &\quad + \left| \int a(x_0+vh) \left[\left\{ g(x_0)J(v) + R_n(x_0,v) \right\}^2 - g^2(x_0)J^2(v) \right] dv \right| \\ &\le B \int |a(x_0+vh) - a(x_0)| \, |J(v)|^2 \, dv \\ &\quad + B \int \left| \left\{ g(x_0)J(v) + R_n(x_0,v) \right\}^2 - g^2(x_0)J^2(v) \right| dv \\ &= T_{1n} + T_{2n} \end{aligned}$$

where we have exploited the fact that both $g(.)$ *and* $a(.)$ *are bounded. We now show that the* $T_{jn} = o(1)$, $j = 1, 2$.
Firstly, note that $J(v) = \int K^r(u)K^r(u-v)du$ *satisfies the conditions placed on* $K(u)$ *in Assumptions 1-3 so that, since* $a \in S$, $T_{1n} = o(1)$ *which follows immediately from the proof of part 1 above.*
Second, we have shown that $\sup_v |R_n(x_0,v)| = o(1)$*, which suffices for* $T_{2n} = o(1)$.

This completes the proof. ■

References

[1] Delgado, M.A., M.A. Dominguez and P. Lavergne (2006). Consistent Tests of Conditional Moment Restrictions. *Annales d'Économie et de Statistique,* 81, 33-67.

[2] Hall, P. (1984). Central Limit Theorem for Integrated Square Error of Multivariate Nonparametric Density Estimators. *Journal of Multivariate Analysis*, 14, 1-16.

[3] Newey, W.K. and J.L Powell (2003). Instrumental Variable Estimation of Nonparametric Models. *Econometrica* 71, 1565–1578.

[4] Powell, J.L., J.H. Stock, and T.M. Stoker (1989). Semiparametric Estimation of Index Coefficients. *Econometrica* 57, 1403 1430.

[5] Zheng, J.X. (1996). A Consistent Test of Functional Form via Nonparametric Estimation Techniques. *Journal of Econometrics*, 75, 263-289.