Supplementary Materials: Intuitions and Proofs

Appendix A: Some Intuitions

In this appendix, we will provide some intuitions for some key assumptions and results in the main text. We will first provide a concrete example to illustrate that $f(u|x)$ may have a cusp at $x=\pi$ when there is imprecise control on x, and then provide some intuitions for the n-consistency of $\widehat{\pi}$, and finally discuss two equivalence results on the asymptotic distribution of $\widehat{\pi}$ in Theorem 4.

Illustration of the Nonsmoothness of $f(u|x)$ at $x=\pi$

Suppose x is related to U by $x(U)=Z+e(U)$, where Z can be precisely controlled by individuals so its support is on $[\pi,\infty)$, while e is a random error independent of Z. Because individuals exert different efforts to manage x for x greater than π and x less than π, the density of e is not continuous at $e=0$. To be more specific, suppose $\pi=0$, the density of Z, $f_Z(z)$, is standard exponential, the density of U, $f_U(u)$, is uniform on $[0,1]$, and $e|U$ has a density as follows:[15]

$$f(e|U)=\begin{cases}\frac{1}{2(2+U)}\exp\left\{\frac{e}{2+U}\right\}, & \text{if } e<0;\\ \frac{1}{2(1+U)}\exp\left\{-\frac{e}{1+U}\right\}, & \text{if } e\geq 0.\end{cases}$$

Then x's density conditional on U is

$$\begin{aligned} f(x|U) &= \int f_Z(x-e)f(e|U)de=\int_{-\infty}^{x}\exp\{e-x\}f(e|U)de \\ &= \begin{cases}\frac{1}{2(3+U)}\exp\left\{\frac{1}{2+U}x\right\}; & \text{if } x<0;\\ \frac{1}{2U}\exp\left\{-\frac{1}{1+U}x\right\}-\frac{3}{2U(3+U)}\exp\{-x\}; & \text{if } x\geq 0;\end{cases}\end{aligned}$$

which is continuous but not smooth at $x=0$. Consequently, $f(u|x)=\frac{f(x|U)f_U(u)}{\int f(x|U=u)f_U(u)du}$ has a cusp at $x=0$. The densities $f(e)$, $f(e|U)$, $f_Z(z)$, $f(x|U)$, $f(x)$ and $f(U|x)$ are shown in Figure 6.

n-Consistency of $\widehat{\pi}$

First, we discuss how the convergence rate is determined for a general estimator defined by maximizing an objective function. Suppose the parameter $\theta\in\Theta$ is estimated by

$$\widehat{\theta}=\arg\max_{\theta\in\Theta}Q_n(\theta)=\arg\max_{\theta\in\Theta}\left[Q_n(\theta)-Q_n(\theta_0)\right],$$

where $Q_n(\cdot)$ is the objective function. Then because $\widehat{\theta}$ is the maximizer of $Q_n(\theta)-Q_n(\theta_0)$ on Θ, and $\theta_0\in\Theta$,

$$0\leq Q_n(\widehat{\theta})-Q_n(\theta_0)=\left[Q(\widehat{\theta})-Q(\theta_0)\right]+\left[\left(Q_n(\widehat{\theta})-Q(\widehat{\theta})\right)-\left(Q_n(\theta_0)-Q(\theta_0)\right)\right],$$

where the first term on the right-hand side is the limit process and less than zero since $\theta_0=\arg\max Q(\theta)$, and the second term is the modulus of continuity of the empirical process and greater than zero. We must

[15] If U is interpreted as ability in the scholarship example of van der Klaauw (2002), then this density means that a student with higher ability (smaller U) can exert larger power in managing her score.

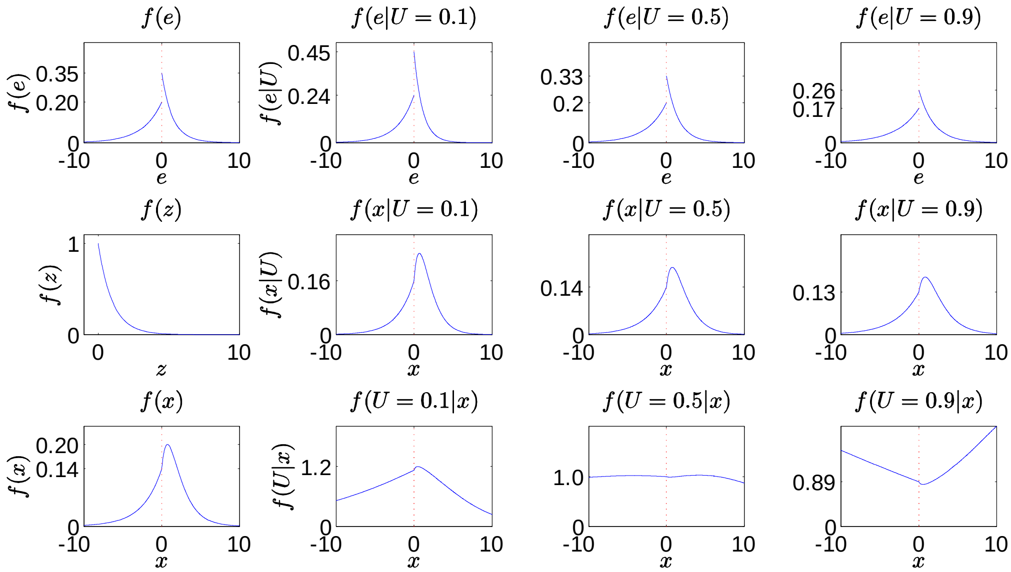

Figure 6: $f(e)$, $f(e|U)$, $f_Z(z)$, $f(x|U)$, $f(x)$ and $f(U|x)$ When Selection is Present

balance $Q(\widehat{\theta}) - Q(\theta_0)$ and

$$\frac{\phi_n(\delta)}{\sqrt{n}} \equiv \sup_{|\theta-\theta_0|\leq\delta} \left[(Q_n(\theta) - Q(\theta)) - (Q_n(\theta_0) - Q(\theta_0))\right]$$

such that their sum is greater than zero.

For a regular parameter μ, say, the mean of a random variable, $Q(\mu)$ is smooth, so for $|\mu - \mu_0| \leq \delta$, $Q(\mu) - Q(\mu_0) = O\left(\delta^2\right)$. And by the empirical process technique, $\frac{\phi_n(\delta)}{\sqrt{n}} = O_p\left(\frac{\delta}{\sqrt{n}}\right)$. Suppose $\widehat{\mu} - \mu_0 = O_p\left(r_n^{-1}\right)$, and let $\left(\frac{1}{r_n}\right)^2 \approx \frac{r_n^{-1}}{\sqrt{n}}$, we get $r_n = \sqrt{n}$, the usual convergence rate for a regular parameter.

The balancing in estimating π is different. Consider a simple model with the fixed design:

$$y_i = \alpha 1\left(x_i \geq \pi\right) + \varepsilon_i, \tag{13}$$

where ε_i's are i.i.d. and follow $N(0,1)$, $x_i = i/n$, $i = 1, \cdots, n$, and $\pi_0 = 1/2$. WLOG, suppose $\alpha > 0$. In the parametric estimation of π, $\widehat{\pi}$ maximizes $\widehat{\alpha}(\pi) - \widehat{\alpha}(\pi_0)$. For $|\pi - \pi_0| \leq \delta$, and $\pi > \pi_0$,

$$\begin{aligned} & \widehat{\alpha}(\pi) - \widehat{\alpha}(\pi_0) \\ = & \left[\frac{1}{n(1-\pi)} \sum_{i=n\pi+1}^{n} y_i - \frac{1}{n\pi}\sum_{i=1}^{n\pi} y_i\right] - \left[\frac{1}{n(1-\pi_0)} \sum_{i=n\pi_0+1}^{n} y_i - \frac{1}{n\pi_0}\sum_{i=1}^{n\pi_0} y_i\right] \\ = & \left[\alpha + \frac{1}{n(1-\pi)} \sum_{i=n\pi+1}^{n} \varepsilon_i - \frac{n\left(\pi-\pi_0\right)}{n\pi}\alpha - \frac{1}{n\pi}\sum_{i=1}^{n\pi} \varepsilon_i\right] - \left[\alpha + \frac{1}{n(1-\pi_0)} \sum_{i=n\pi_0+1}^{n} \varepsilon_i - \frac{1}{n\pi_0}\sum_{i=1}^{n\pi_0} \varepsilon_i\right] \\ \approx & -\alpha\left(\frac{\pi-\pi_0}{\pi}\right) - \left(\frac{1}{n\pi} + \frac{1}{n(1-\pi_0)}\right)\sum_{i=n\pi_0+1}^{n\pi} \varepsilon_i = O\left(\pi-\pi_0\right) + O\left(n^{-1}\right) O_p\left(\sqrt{n\left(\pi-\pi_0\right)}\right) = O(\delta) + O_p\left(\sqrt{\frac{\delta}{n}}\right), \end{aligned} \tag{14}$$

so $\phi_n(\delta) = \sqrt{\delta}$ and $Q(\pi) - Q(\pi_0) = O(\delta)$. Similar results hold for $\pi < \pi_0$. Suppose $\widehat{\pi} - \pi_0 = O_p\left(r_n^{-1}\right)$,

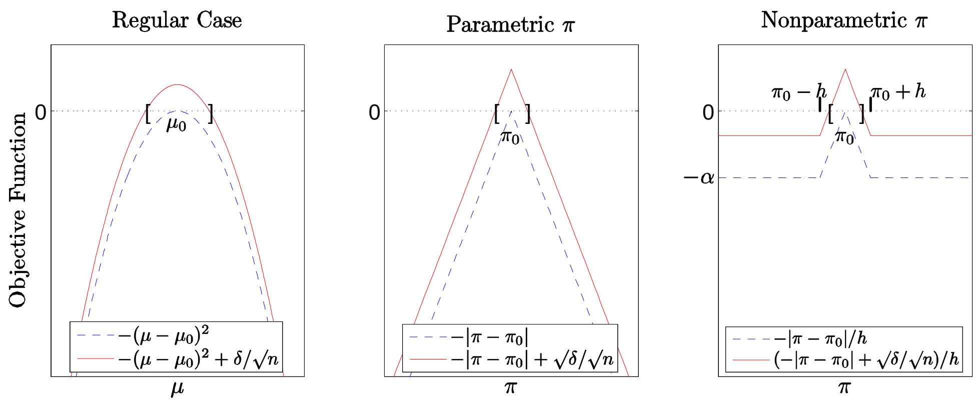

Figure 7: Balancing $Q(\theta)-Q(\theta_0)$ and $\frac{\phi_n(\delta)}{\sqrt{n}}$ in the Regular Case and Parametric and Nonparametric Estimation of π

and let $\frac{1}{r_n} \approx \frac{\sqrt{1/r_n}}{\sqrt{n}}$, we get $r_n = n$. In the nonparametric estimation of π, the scale changes to h. For $\pi = \pi_0 + ah$ with $a > 0$ small enough,

$$\begin{aligned}
&\widehat{\alpha}(\pi)-\widehat{\alpha}(\pi_0)\\
=\;&\left[\frac{1}{nh}\sum_{i=n\pi+1}^{n(\pi+h)} y_i-\frac{1}{nh}\sum_{i=n(\pi-h)}^{n\pi} y_i\right]-\left[\frac{1}{nh}\sum_{i=n\pi_0+1}^{n(\pi_0+h)} y_i-\frac{1}{nh}\sum_{i=n(\pi_0-h)}^{n\pi_0} y_i\right]\\
=\;&\left[\alpha+\frac{1}{nh}\sum_{i=n\pi+1}^{n(\pi+h)}\varepsilon_i-\frac{n(\pi-\pi_0)}{nh}\alpha-\frac{1}{nh}\sum_{i=n(\pi-h)}^{n\pi}\varepsilon_i\right]-\left[\alpha+\frac{1}{nh}\sum_{i=n\pi_0+1}^{n(\pi_0+h)}\varepsilon_i-\frac{1}{nh}\sum_{i=n(\pi_0-h)}^{n\pi_0}\varepsilon_i\right]\\
\approx\;&-\alpha\left(\frac{\pi-\pi_0}{h}\right)-\frac{2}{nh}\sum_{i=n\pi_0+1}^{n\pi}\varepsilon_i=O\left(\frac{\pi-\pi_0}{h}\right)-O\left(\frac{1}{nh}\right)O_p\left(\sqrt{n(\pi-\pi_0)}\right)=O\left(\frac{\delta}{h}\right)+O_p\left(\frac{1}{h}\sqrt{\frac{\delta}{n}}\right).
\end{aligned}$$

So by solving $\frac{r_n^{-1}}{h} \approx \frac{1}{h}\frac{\sqrt{1/r_n}}{\sqrt{n}}$, we get $r_n = n$.

Figure 7 illustrates the intuition above. Roughly speaking, in estimating π, $Q(\pi)-Q(\pi_0)$ is a nonsmooth function of π in the neighborhood of π_0 such that π_0 can be identified more easily than μ_0. In the nonparametric case, we use a smaller scale h, and focus on the discussion in a h neighborhood of π_0.

Two Equivalence Results

Given the n-consistency of $\widehat{\pi}$, we continue to find the weak limit of the localized objective function. Here, we need the Lipschitz continuity of $k(\cdot)$, and the uniform kernel above does not work. To simplify the discussion, let π_0 be the closest $\frac{i}{n}$ to $\frac{1}{2}$. Assume further that both v and nh are positive integers. Then for $\widehat{\alpha}(\pi)$ based on a kernel function $k(\cdot)$,

$$\begin{aligned}
&nh\left(\widehat{\alpha}\left(\pi_0+\frac{v}{n}\right)-\widehat{\alpha}(\pi_0)\right)\\
=\;&\left[\sum_{i=1}^{nh}k\left(\frac{i}{nh}\right)y_{n\pi_0+v+i}-\sum_{i=-nh}^{0}k\left(\frac{i}{nh}\right)y_{n\pi_0+v+i}\right]-\left[\sum_{i=1}^{nh}k\left(\frac{i}{nh}\right)y_{n\pi_0+i}-\sum_{i=-nh}^{0}k\left(\frac{i}{nh}\right)y_{n\pi_0+i}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=-nh}^{-nh+v-1} k\left(\frac{i}{nh}\right)\varepsilon_{n\pi_0+i} - \sum_{i=-nh+v}^{0}\left[k\left(\frac{i-v}{nh}\right)-k\left(\frac{i}{nh}\right)\right]\varepsilon_{n\pi_0+i} \\
&\quad -\sum_{i=1}^{v}\left\{k\left(\frac{i-v}{nh}\right)\alpha+\left[k\left(\frac{i}{nh}\right)+k\left(\frac{i-v}{nh}\right)\right]\varepsilon_{n\pi_0+i}\right\} \\
&\quad +\sum_{i=1}^{nh-v}\left[k\left(\frac{i}{nh}\right)-k\left(\frac{i+v}{nh}\right)\right]\varepsilon_{n\pi_0+v+i}+\sum_{i=nh-v+1}^{nh}k\left(\frac{i}{nh}\right)\varepsilon_{n\pi_0+v+i} \\
&\equiv I+II+III+IV+V.
\end{aligned}
$$

Figure 8 shows the effect of $k(\cdot)$ on the five terms. When $k(\cdot)$ is Lipschitz, all terms except III are $o_p(1)$. Especially, $k\left(\frac{i}{nh}\right)$ in I and V is close to zero. But if $k(\cdot)$ is uniform, I and V will not disappear. The summands in III correspond to the jumps in $D(v)$. Note that $\frac{i-v}{nh}$ and $\frac{i}{nh}$ are close to zero when i and v are finite, so the summand in III is approximately $-k(0)\left(\alpha+2\varepsilon_{n\pi_0+i}\right)$.[16] Similar results hold for $v<0$.

Now, we can state two equivalence results for the asymptotic distribution of $\widehat{\pi}$. To simplify our discussion, we continue using the model with the fixed design. First, the asymptotic distribution of $\widehat{\pi}$ is the same as the parametric estimator based on (14). In other words, although RDDs are nonparametrically formulated, we can estimate π_0 as if the model is parametric as long as $m(\cdot)$ is smooth on $\Pi\backslash\{\pi_0\}$. Given that $\widehat{\pi}$ is n-consistent, only the data in a n^{-1} neighborhood of π_0 (that is, **finite** data points) are informative to π_0; see the derivation above. Because $nh\to\infty$, there are **infinite** data points in the h neighborhood of kernel smoothing. Also, the kernel estimator treats $m(\cdot)$ as a constant in any h neighborhood. So the kernel estimator is smooth enough to identify π_0 as if $m(\cdot)$ were constant in big enough left and right neighborhoods of π_0. Following such an argument, it is not surprising that π_0 can be estimated as if in a parametric model.

Second, the asymptotic distribution of $\widehat{\pi}$ is the same as that of the least squares estimator (LSE) in (1).[17] Recall that the objective function of the LSE is

$$
\sum_{i=1}^{n}\left[y_i-\alpha_\pi d_i(\pi)-\sum_{j=1}^{n}w_j^i\left(y_j-\alpha_\pi d_j(\pi)\right)\right]^2, \tag{15}
$$

where $w_j^i=\frac{K_{h,ij}}{\sum_{l=1}^{n}K_{h,il}}$, and $\sum_{j=1}^{n}w_j^i\left(y_j-\alpha_\pi d_j(\pi)\right)$ can be treated as an estimator of $m_\pi(x)$ at x_i. This is exactly the objective function of the partially linear estimator in Porter (2003). Note that (15) can be written as

$$
\left\|(I-W)\,Y-\alpha_\pi\,(I-W)\,D_\pi\right\|^2, \tag{16}
$$

where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^n$, I is an $n\times n$ identity matrix, $W=\left(w_j^i\right)_{i,j=1,\cdots,n}$ is symmetric, and $D_\pi=\left(d_i(\pi)\right)_{i=1,\cdots,n}$. So the estimator of α_π given π is

$$
\widehat{\alpha}(\pi)=\frac{D_\pi'(I-W)^2Y}{D_\pi'(I-W)^2D_\pi},
$$

[16] Since our objective function is $\widehat{\alpha}^2(\pi)$, $nh\left(\widehat{\alpha}^2(\pi_0+\frac{v}{n})-\widehat{\alpha}^2(\pi_0)\right) = nh\left(\widehat{\alpha}\left(\pi_0+\frac{v}{n}\right)-\widehat{\alpha}(\pi_0)\right)\left(\widehat{\alpha}\left(\pi_0+\frac{v}{n}\right)+\widehat{\alpha}(\pi_0)\right)$. $\widehat{\alpha}\left(\pi_0+\frac{v}{n}\right)+\widehat{\alpha}(\pi_0)$ converges to $2\alpha_{\pi_0}$, so the jumps in $D(v)$ should be $-k(0)\left(2\alpha_{\pi_0}\right)\left(\alpha_{\pi_0}+2\varepsilon_{n\pi_0+i}\right)$, which are different from the jumps of $D(v)$ in Theorem 4 only by a constant $2k(0)$. Of course, a constant will not affect the output of the argmax operator.

[17] See Section 4.1 of Yu (2008) for the asymptotic distribution of the least squares estimator in the parametric case, but as argued above, it should be the same as in the nonparametric case.

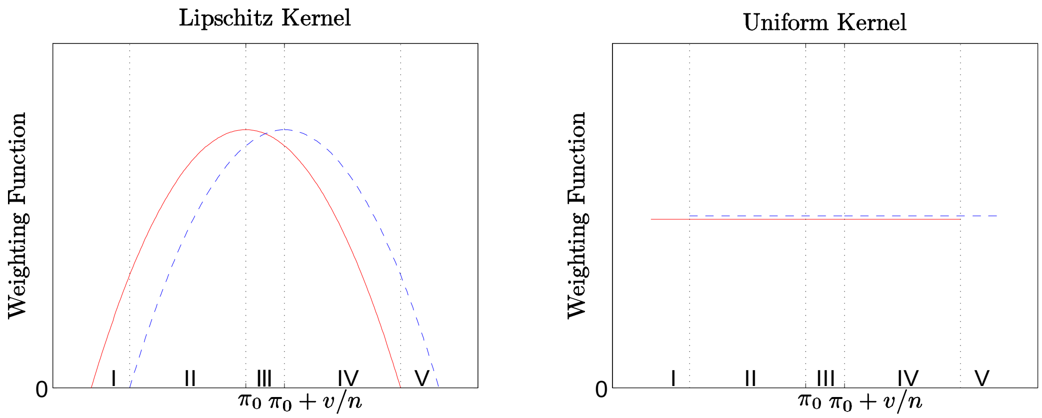

Figure 8: Weighting Functions in Localized Objective Function Using Different Kernels

which is the coefficient in the least squares projection of $(I-W)\,Y$ on $(I-W)\,D_\pi$. Now, (16) becomes

$$\|(I-W)\,Y\|^2 - \left\|(I-W)\,D_\pi \frac{D'_\pi\,(I-W)^2\,Y}{D'_\pi\,(I-W)^2\,D_\pi}\right\|^2 = \|(I-W)\,Y\|^2 - \|P_\pi\,(I-W)\,Y\|^2\,,$$

where $P_\pi = \frac{(I-W)D_\pi D'_\pi(I-W)}{D'_\pi(I-W)^2 D_\pi}$ is a projection matrix. Because the first term $\|(I-W)\,Y\|^2$ does not involve π, minimizing the objective function with respect to π is equivalent to maximizing $\|P_\pi\,(I-W)\,Y\|^2$ with respect to π. Note that $\|P_\pi\,(I-W)\,Y\|^2 = \widehat{\alpha}^2\,(\pi)\,D'_\pi\,(I-W)^2\,D_\pi$. From Theorem 2 of Porter (2003),

$$D'_\pi\,(I-W)^2\,D_\pi \xrightarrow{p} 2\int_0^1 \left(\int_u^1 k(v)dv\right)^2 du$$

independent of π, so maximizing $\|P_\pi\,(I-W)\,Y\|^2$ is equivalent to maximizing $\widehat{\alpha}^2\,(\pi)$.

Appendix B: Proofs

Throughout the proofs, H_0 indicates both $H_0^{(1)}$ and $H_0^{(2)}$, and H_1 indicates both $H_1^{(1)}$ and $H_1^{(2)}$.
Proof of Theorem 1 and 2. First, decompose I_n by using (8):

$$\begin{aligned} I_n &= \frac{nh^{1/2}}{n\,(n-1)} \sum_i \sum_{j\neq i} 1_i^\Pi 1_j^\Pi \left\{(m_i-\widehat{m}_i)\,(m_j-\widehat{m}_j) + \varepsilon_i\varepsilon_j + \widehat{\varepsilon}_i\widehat{\varepsilon}_j \right.\\ &\quad \left. +2\varepsilon_i\,(m_j-\widehat{m}_j) - 2\widehat{\varepsilon}_i\,(m_j-\widehat{m}_j) - 2\varepsilon_i\widehat{\varepsilon}_j\right\} K_{h,ij} \\ &\equiv I_{1n} + I_{2n} + I_{3n} + 2I_{4n} - 2I_{5n} - 2I_{6n}. \end{aligned}$$

We shall complete the proof by examining $I_{1n}, \cdots, I_{6n}$, respectively, and showing that $v_n^2 = \Sigma + o_p\,(1)$ under both H_0 and H_1. Throughout this proof, $z_i = (x_i, \varepsilon_i)'$ and $E_i\,[\cdot] = E\,[\cdot|x_i]$.

First, I_{2n}, I_{3n} and I_{6n} are invariant under H_0 and H_1. Propositions 3 and 6 show that I_{3n} and I_{6n} are both $o_p(1)$. Proposition 2 shows that $I_{2n} \xrightarrow{d} N\,(0, \Sigma)$.

Under H_0, Proposition 1 shows that $I_{1n} = o_{P_m}\,(1)$, and Propositions 4 and 5 show that I_{4n} and I_{5n} are both $o_{P_m}\,(1)$ uniformly in $m(\cdot) \in H_0$.

Under H_1, Propositions 4 and 5 show that I_{4n} and I_{5n} are dominated by I_{1n}, and Proposition 1 shows that $I_{1n} = O_p\left(nh^{1/2}b^3\right)$ under $H_1^{(1)}$ and $I_{1n} = O_p\left(nh^{1/2}b\right)$ under $H_1^{(2)}$. The local power can be easily obtained from the proof of Proposition 1.

At last, Proposition 7 shows that $v_n^2 = \Sigma + o_p(1)$. So the proof is complete. ■

Proof of Theorem 3. This proof is similar but more tedious than that of Theorem 1 and 2. Note that $\Phi(z)$ is a continuous function. By Polya's theorem, it suffices to show for any fixed value of $z \in \mathbb{R}$, $|P(T_n^* \leq z|\mathcal{F}_n) - \Phi(z)| = o_p(1)$.

Denote $m_i^* = \widehat{y}_i$ and define $\widehat{m}_i^*$ and $\widehat{\varepsilon}_i^*$ by

$$\widehat{m}_i^* = \frac{1}{n-1}\sum\nolimits_{j\neq i} m_j^* L_{b,ij} \Big/ \widehat{f}_i,$$

and

$$\widehat{\varepsilon}_i^* = \frac{1}{n-1}\sum\nolimits_{j\neq i} \varepsilon_j^* L_{b,ij} \Big/ \widehat{f}_i.$$

Then using $\widehat{e}_i^* = y_i^* - \widehat{y}_i^* = m_i^* + \varepsilon_i^* - (\widehat{m}_i^* + \widehat{\varepsilon}_i^*)$, we get

$$\begin{aligned} I_n^* &= \frac{nh^{1/2}}{n(n-1)}\sum_i\sum_{j\neq i} 1_i^\Pi 1_j^\Pi \left\{(m_i^* - \widehat{m}_i^*)\left(m_j^* - \widehat{m}_j^*\right) + \varepsilon_i^*\varepsilon_j^* + \widehat{\varepsilon}_i^*\widehat{\varepsilon}_j^* \right. \\ &\quad \left. + 2\varepsilon_i^*\left(m_j^* - \widehat{m}_j^*\right) - 2\widehat{\varepsilon}_i^*\left(m_j^* - \widehat{m}_j^*\right) - 2\varepsilon_i^*\widehat{\varepsilon}_j^*\right\} K_{h,ij} \\ &\equiv I_{1n}^* + I_{2n}^* + I_{3n}^* + 2I_{4n}^* - 2I_{5n}^* - 2I_{6n}^*. \end{aligned}$$

The theorem will be proved if we can show that $I_{in}^*|\mathcal{F}_n = o_p(1)$ for $i = 1, 3, 4, 5, 6$ and $I_{2n}^*/v_n^*|\mathcal{F}_n \to N(0,1)$ in probability. The first part is similar to those of Proposition 1, 3, 4, 5 and 6 under H_0. Only note that $m^*(x)|\mathcal{F}_n$ defined as above satisfies H_0 even if $m(x)$ is from H_1; see Gu et al. (2007) for a similar analysis in testing omitted variables. But there is some difference to show the second part.

First, because $\varepsilon_i^*|\mathcal{F}_n$ are mean zero and mutually independent and have variance $\widehat{e}_i^2$,

$$\frac{nh^{1/2}}{n(n-1)}\sum_i\sum_{j\neq i} 1_i^\Pi 1_j^\Pi \varepsilon_i^*\varepsilon_j^* K_{h,ij} = \frac{2nh^{1/2}}{n(n-1)}\sum_i\sum_{j>i} 1_i^\Pi 1_j^\Pi \varepsilon_i^*\varepsilon_j^* K_{h,ij} \equiv \sum_i\sum_{j>i} U_{n,ij}^*$$

is a second order degenerate U-statistic with conditional variance

$$\frac{2h}{n(n-1)}\sum_i\sum_{j\neq i} 1_i^\Pi 1_j^\Pi \widehat{e}_i^2\widehat{e}_j^2 K_{h,ij}^2 = v_n^2.$$

Because $U_{n,ij}^*$ depends on i and j, we must use the central limit theorem of de Jong (1987) for generalized quadratic forms rather than Hall (1984) to find the asymptotic distribution of I_{2n}^*. From his Proposition 3.2, we know $I_{2n}^*/v_n|\mathcal{F}_n \to N(0,1)$ in probability as long as

$$\begin{aligned} G_I^* &= \sum_i\sum_{j>i} E^*\left[U_{n,ij}^{*4}\right] = o_p\left(v_n^4\right), \\ G_{II}^* &= \sum_i\sum_{j>i}\sum_{l>j>i} E^*\left[U_{n,ij}^{*2}U_{n,il}^{*2} + U_{n,ji}^{*2}U_{n,jl}^{*2} + U_{n,li}^{*2}U_{n,lj}^{*2}\right] = o_p\left(v_n^4\right), \\ G_{IV}^* &= \sum_i\sum_{j>i}\sum_{k>j>i}\sum_{l>k>j>i} E^*\left[U_{n,ij}^*U_{n,ik}^*U_{n,lj}^*U_{n,lk}^* + U_{n,ij}^*U_{n,il}^*U_{n,kj}^*U_{n,kl}^* + U_{n,ik}^*U_{n,il}^*U_{n,jk}^*U_{n,jl}^*\right] = o_p\left(v_n^4\right). \end{aligned}$$

It is straightforward to show that

$$G_I^* = O_p\left(\left(n^2h\right)^{-1}\right), G_{II}^* = O_p(n^{-1}), G_{IV}^* = O_p(h),$$

so the result follows by $v_n^4 = O_p(1)$. Next, it is easy to check that $E^*\left[v_n^{*2}\right] = v_n^2 + o_p(1)$, and $Var^*\left(v_n^{*2}\right) = o_p(1)$. Thus $I_{2n}^*/v_n^*|\mathcal{F}_n \to N(0,1)$ in probability. ■

Proof of Theorem 4. We assume $m_+(\pi)$ and $m_-(\pi)$ are estimated by local constant estimators since by Fan and Gijbels (1996), the p-th order local polynomial estimator is asymptotically equivalent to the local constant estimator with a $(p+1)$-th order boundary kernel. More importantly, the contribution made to the local polynomial fit by terms of degree j is of order $O(h^j)$, so the local constant has the dominating effect.

Proposition 8 proves $\widehat{\pi}$ is consistent, and Proposition 9 proves $\widehat{\pi} - \pi_0 = O_p\left(n^{-1}\right)$. Now, we show $n\left(\widehat{\pi} - \pi_0\right)$ has the asymptotic distribution as stated in the theorem by applying the argmax continuous mapping theorem. To achieve this goal, we first analyze the localized objective function:

$$nh\left(\widehat{\alpha}^2(\pi_0 + \frac{v}{n}) - \widehat{\alpha}^2(\pi_0)\right) = \left(\widehat{\alpha}\left(\pi_0 + \frac{v}{n}\right) + \widehat{\alpha}(\pi_0)\right) nh\left(\widehat{\alpha}(\pi_0 + \frac{v}{n}) - \widehat{\alpha}(\pi_0)\right).$$

Since $\left(\widehat{\alpha}(\pi_0 + \frac{v}{n}) + \widehat{\alpha}(\pi_0)\right) \xrightarrow{p} 2\alpha_{\pi_0}$, we need only analyze $nh\left(\widehat{\alpha}(\pi_0 + \frac{v}{n}) - \widehat{\alpha}(\pi_0)\right)$. Proposition 10 shows that for v on any compact set in $\mathbb{R}$,

$$\begin{aligned} & nh\left(\widehat{\alpha}(\pi_0 + \frac{v}{n}) - \widehat{\alpha}(\pi_0)\right) \qquad (17) \\ = & -\frac{k(0)}{f(\pi_0)}\sum\nolimits_{j=1}^{n}\left\{(\alpha_{\pi_0} - 2\varepsilon_j)1\left(\pi_0 + \frac{v}{n} \le x_j < \pi_0\right) + (\alpha_{\pi_0} + 2\varepsilon_j)1\left(\pi_0 \le x_j < \pi_0 + \frac{v}{n}\right)\right\} + o_p(1), \end{aligned}$$

so

$$\begin{aligned} & nh\left(\widehat{\alpha}^2(\pi_0 + \frac{v}{n}) - \widehat{\alpha}^2(\pi_0)\right) \\ = & -C\sum\nolimits_{j=1}^{n}\left\{\left(\alpha_{\pi_0}^2 - 2\alpha_{\pi_0}\varepsilon_j\right)1\left(\pi_0 + \frac{v}{n} \le x_j < \pi_0\right) + \left(\alpha_{\pi_0}^2 + 2\alpha_{\pi_0}\varepsilon_j\right)1\left(\pi_0 \le x_j < \pi_0 + \frac{v}{n}\right)\right\} + o_p(1), \end{aligned}$$

where $C = 2k(0)/f(\pi_0)$. Now, by a straightforward application of the proof idea of Theorem 1 and 2 in Yu (2012), we can get the asymptotic distribution of $\widehat{\pi}$. The only difference here is that $nh\left(\widehat{\alpha}^2(\pi_0 + \frac{v}{n}) - \widehat{\alpha}^2(\pi_0)\right)$ is a caglad instead of cadlag process. Note also that ε_i^- in Theorem 1 of Yu (2012) is defined as the limiting conditional distribution of ε_i given $\pi_0 + \Delta \le x_i < \pi_0$, $\Delta < 0$ with $\Delta \uparrow 0$, and ε_i^+ is defined as the limiting conditional distribution of ε_i given $\pi_0 \le x_i < \pi_0 + \Delta$, $\Delta > 0$ with $\Delta \downarrow 0$. Given Assumption E, ε_i^- has the conditional density $f(\varepsilon|x = \pi_0-)$, and ε_i^+ has the conditional density $f(\varepsilon|x = \pi_0+)$. ■

Proof of Theorem 5. As argued in Theorem 4, we need only prove the result for the local constant estimator. Since $\widehat{\pi} - \pi_0 = O_p\left(n^{-1}\right)$, $\widehat{\pi}$ will fall into $\left(\pi_0 - \frac{C}{n}, \pi_0 + \frac{C}{n}\right)$ for some positive C with any large probability when n is large enough. Combining this fact and (17), we need only show that

$$\sup_{v\in(-C,C)}\sum\nolimits_{j=1}^{n}\left\{(\alpha_{\pi_0} - 2\varepsilon_j)1\left(\pi_0 + \frac{v}{n} \le x_j < \pi_0\right) + (\alpha_{\pi_0} + 2\varepsilon_j)1\left(\pi_0 \le x_j < \pi_0 + \frac{v}{n}\right)\right\} = O_p(1),$$

for any $C > 0$ because $nh \to \infty$. Due to the similarity between $v < 0$ and $v > 0$, we need only show

$$\sup_{v\in(0,C)}\sum\nolimits_{j=1}^{n}(\alpha_{\pi_0} + 2\varepsilon_j)1\left(\pi_0 \le x_j < \pi_0 + \frac{v}{n}\right) = O_p(1).$$

First,

$$\sup_{v\in(0,C)} \sum\nolimits_{j=1}^{n} \alpha_{\pi_0} 1\left(\pi_0 \le x_j < \pi_0 + \frac{v}{n}\right) = O_p(1),$$

because $\sup_{v\in(0,C)} \sum_{j=1}^{n} 1\left(\pi_0 \le x_j < \pi_0 + \frac{v}{n}\right)$ is no larger than $\sum_{j=1}^{n} 1\left(\pi_0 \le x_j < \pi_0 + \frac{C}{n}\right)$ which is $O_p(1)$ by Assumption F. Second,

$$\sup_{v\in(0,C)} \sum\nolimits_{j=1}^{n} \varepsilon_j 1\left(\pi_0 \le x_j < \pi_0 + \frac{v}{n}\right) = O_p(1).$$

This is because $\left\{1\left(\pi_0 \le x_j < \pi_0 + \frac{v}{n}\right), v \in (0,C)\right\}$ is a VC-class with envelope $1\left(\pi_0 \le x_j < \pi_0 + \frac{C}{n}\right)$, and so

$$E\left[\left|\sup_{v\in(0,C)} \sum\nolimits_{j=1}^{n} \varepsilon_j 1\left(\pi_0 \le x_j < \pi_0 + \frac{v}{n}\right)\right|\right] \le \sqrt{n}C' \sqrt{E\left[\varepsilon_j^2 1\left(\pi_0 \le x_j < \pi_0 + \frac{C}{n}\right)\right]} = O(1),$$

where the inequality is from, e.g., Theorem 2.14.2 of Van der Vaart and Wellner (1996), C' is some positive constant, and the equality is from Assumption E and F. As to the asymptotic independence between $\widehat{\pi}$ and $\widehat{\alpha}_\pi$, see the proof of Theorem 1 and 2 in Yu (2012) where the characteristic function is used to show this result. ■

Proof of Corollary 1. Take $\widehat{\pi}_2$ as an example. Following the proof of Proposition 8 and 9, it is easy to see that $\widehat{\pi}_1 - \pi_{10} = O_p\left(n^{-1}\right)$. Because $nh \to \infty$, π_{10} will stay in the h neighborhood of $\widehat{\pi}_1$ w.p.a.1. Because we exclude a $2h$ neighborhood of $\widehat{\pi}_1$, w.p.a.1., the estimation of π_2 and α_{π_2} will not use any data in the h neighborhood of $\widehat{\pi}_1$. As a result, $\widehat{\pi}_2$ and $\widehat{\alpha}_{\pi_2}$ are asymptotically independent of $\widehat{\pi}_1$ and $\widehat{\alpha}_{\pi_1}$. The asymptotic independence between $\widehat{\pi}_2$ and $\widehat{\alpha}_{\pi_2}$ can be similarly shown as in Theorem 5. ■

Appendix C: Propositions

When we evaluate the order of some terms, if there is no confusion, we will use n, $(n-1)$, and $(n-2)$ interchangeably. By Lemma 3, we can assume $f(\cdot) = 1$ on Π_ϵ throughout the following proof, so $f(\cdot)$ is depressed unless necessary. WLOG, suppose $M = 1$, and the constant L in $\mathcal{M}_\Pi(L, s)$ is 1.

Proposition 1 *I_{1n} is $o_{P_m}(1)$ uniformly in m under H_0, and is $O_p\left(nh^{1/2}b^3\right)$ under $H_1^{(1)}$ and $O_p\left(nh^{1/2}b\right)$ under $H_1^{(2)}$.*

Proof. Because $k(\cdot)$ is nonnegative and symmetric,

$$\begin{aligned} I_{1n} &= \frac{nh^{1/2}}{n(n-1)} \sum_i \sum_{j\neq i} 1_i^\Pi 1_j^\Pi (m_i - \widehat{m}_i)(m_j - \widehat{m}_j) K_{h,ij} \\ &\le \frac{nh^{1/2}}{2n(n-1)} \sum_i \sum_{j\neq i} \left\{1_i^\Pi [(m_i - \widehat{m}_i)]^2 + 1_j^\Pi [(m_j - \widehat{m}_j)]^2\right\} K_{h,ij} \\ &= \frac{nh^{1/2}}{n(n-1)} \sum_i \sum_{j\neq i} 1_i^\Pi [(m_i - \widehat{m}_i)]^2 K_{h,ij} = h^{1/2} \sum_i 1_i^\Pi [(m_i - \widehat{m}_i)]^2 \widetilde{f}_i \\ &\le h^{1/2} \left(\sup_{x_i\in\Pi} \widetilde{f}_i\right) \sum_i 1_i^\Pi [(m_i - \widehat{m}_i)]^2, \end{aligned}$$

where $\widetilde{f}_i = \frac{1}{n-1}\sum_{j\neq i} K_{h,ij}$, $\sup_{x_i\in\Pi} \widetilde{f}_i = O_p(1)$ is a well-known result, and $\sum_i 1_i^\Pi [(m_i - \widehat{m}_i)]^2$ is $O_p\left(n(b^{2\eta} + (nb)^{-1})\right)$ under $H_0^{(1)}$ and is $O_p\left(n(b^{2\eta} + b^3 + (nb)^{-1})\right)$ under $H_0^{(2)}$ as shown in Lemma 1. So $I_{1n} \le h^{1/2} O_p(1) O_p\left(n\left(b^{2\eta} + (nb)^{-1}\right)\right) =$

$O_p\left(nh^{1/2}b^{2\eta}+h^{1/2}/b\right)=o_p(1)$ under Assumption B(a), and $I_{1n}\leq h^{1/2}O_p(1)O_p\left(n\left(b^{2\eta}+b^3+(nb)^{-1}\right)\right)=O_p\left(nh^{1/2}\left(b^{2\eta}+b^3\right)+h^{1/2}/b\right)=o_p(1)$ under Assumption B(b).

Under H_1, $\widehat{f}_i^{-1}=f_i^{-1}+o_p(1)$ uniformly over $x_i\in\Pi$, so

$$\begin{aligned}
I_{1n} &= \frac{nh^{1/2}}{n(n-1)}\sum_i\sum_{j\neq i}1_i^{\Pi}1_j^{\Pi}\left(m_i-\widehat{m}_i\right)\widehat{f}_i\left(m_j-\widehat{m}_j\right)\widehat{f}_jK_{h,ij}\left(\widehat{f}_i^{-1}\widehat{f}_j^{-1}\right)\\
&\approx \frac{nh^{1/2}}{n(n-1)^3}\sum_i\sum_{j\neq i}\sum_{l\neq i}\sum_{k\neq j}1_i^{\Pi}\left(m_i-m_l\right)L_{b,il}1_j^{\Pi}\left(m_j-m_k\right)L_{b,jk}K_{h,ij}f_i^{-1}f_j^{-1}.
\end{aligned}$$

It can be shown that the case where i,j,l,k are all different from each other dominates, so

$$I_{1n}\approx O_p(nh^{1/2}E\left[1_1^{\Pi}\left(m_1-m_2\right)L_{b,12}1_3^{\Pi}\left(m_3-m_4\right)L_{b,34}K_{h,13}f_1^{-1}f_3^{-1}\right]).$$

Because $h/b\to 0$, we can treat $x_1=x_3$. Specifically,

$$\begin{aligned}
& E\left[1_1^{\Pi}\left(m_1-m_2\right)L_{b,12}1_3^{\Pi}\left(m_3-m_4\right)L_{b,34}K_{h,13}f_1^{-1}f_3^{-1}\right]\\
=\ & E\left[1_1^{\Pi}\left(m_1-m_2\right)L_{b,12}f_1^{-1}\int 1\left(x_1+uh\in\Pi\right)\left(m(x_1+uh)-m_4\right)\frac{1}{b}l\left(\frac{x_1+uh-x_4}{b}\right)k(u)du\right]\\
\approx\ & E\left[1_1^{\Pi}\left(m_1-m_2\right)L_{b,12}\left(m_1-m_4\right)L_{b,14}f_1^{-1}\right]\\
=\ & E\left\{1_1^{\Pi}f_1^{-1}\left\{E_1\left[\left(m_1-m_2\right)L_{b,12}\right]\right\}^2\right\}=\int_{\underline{\pi}}^{\overline{\pi}}\left[\int\left(m(x_1)-m(x_2)\right)\frac{1}{b}l\left(\frac{x_2-x_1}{b}\right)f(x_2)dx_2\right]^2dx_1\\
\approx\ & O\left(b^{2\eta}\right)+\int_{\pi-b}^{\pi+b}\left[\int_{-1}^{1}\left(m(x_1)-m(x_1+ub)\right)l\left(u\right)f(x_1+ub)du\right]^2dx_1\\
\approx\ & O\left(b^{2\eta}\right)+f^2(\pi)\int_{\pi-b}^{\pi+b}\left[\int_{\frac{\pi-x_1}{b}}^{1}\left(m(x_1)-m(x_1+ub)\right)l\left(u\right)du+\int_{-1}^{\frac{\pi-x_1}{b}}\left(m(x_1)-m(x_1+ub)\right)l\left(u\right)du\right]^2dx_1
\end{aligned}$$

Under $H_1^{(1)}$,

$$\begin{aligned}
& \int_{\pi-b}^{\pi+b}\left[\int_{\frac{\pi-x_1}{b}}^{1}\left(m(x_1)-m(x_1+ub)\right)l\left(u\right)du+\int_{-1}^{\frac{\pi-x_1}{b}}\left(m(x_1)-m(x_1+ub)\right)l\left(u\right)du\right]^2dx_1\\
=\ & \int_{\pi}^{\pi+b}\left[\int_{\frac{\pi-x_1}{b}}^{1}\left(m(x_1)-m(x_1+ub)\right)l\left(u\right)du+\int_{-1}^{\frac{\pi-x_1}{b}}\left(m(x_1)-m(x_1+ub)\right)l\left(u\right)du\right]^2dx_1\\
& +\int_{\pi-b}^{\pi}\left[\int_{\frac{\pi-x_1}{b}}^{1}\left(m(x_1)-m(x_1+ub)\right)l\left(u\right)du+\int_{-1}^{\frac{\pi-x_1}{b}}\left(m(x_1)-m(x_1+ub)\right)l\left(u\right)du\right]^2dx_1\\
=\ & b\int_0^1\left[\int_{-v}^{1}\left[m(\pi+vb)-m(\pi+(v+u)b)\right]l\left(u\right)du+\int_{-1}^{-v}\left[m(\pi+vb)-m(\pi+(v+u)b)\right]l\left(u\right)du\right]^2dv\\
& +b\int_0^1\left[\int_{v}^{1}\left[m(\pi-vb)-m(\pi+(u-v)b)\right]l\left(u\right)du+\int_{-1}^{v}\left[m(\pi-vb)-m(\pi+(u-v)b)\right]l\left(u\right)du\right]^2dv\\
=\ & b^3\int_0^1\left[-\int_{-v}^{1}m_+'(\pi)ul\left(u\right)du+\int_{-1}^{-v}\left[m_+'(\pi)v-m_-'(\pi)(v+u)\right]l\left(u\right)du\right]^2dv\\
& +b^3\int_0^1\left[\int_{v}^{1}\left[-m_-'(\pi)v-m_+'(\pi)\left(u-v\right)\right]l\left(u\right)du-m_-'(\pi)\int_{-1}^{v}ul\left(u\right)du\right]^2dv
\end{aligned}$$

$$\begin{aligned}
&= b^3 \int_0^1 \left[-\left(m'_+(\pi) - m'_-(\pi)\right) \int_v^1 ul\left(u\right) du + \left(m'_+(\pi) - m'_-(\pi)\right) v \int_v^1 l\left(u\right) du \right]^2 dv \\
&\quad + b^3 \int_0^1 \left[\left(m'_+(\pi) - m'_-(\pi)\right) v \int_v^1 l\left(u\right) du - \left(m'_+(\pi) - m'_-(\pi)\right) \int_v^1 ul\left(u\right) du \right]^2 dv \\
&= 2\left(m'_+(\pi) - m'_-(\pi)\right)^2 b^3 \int_0^1 \left(v \int_v^1 l\left(u\right) du - \int_v^1 ul\left(u\right) du \right)^2 dv.
\end{aligned}$$

Under $H_1^{(2)}$,

$$\begin{aligned}
& \int_{\pi-b}^{\pi+b} \left[\int_{\frac{\pi-x_1}{b}}^1 \left(m(x_1) - m(x_1+ub)\right) l\left(u\right) du + \int_{-1}^{\frac{\pi-x_1}{b}} \left(m(x_1) - m(x_1+ub)\right) l\left(u\right) du \right]^2 dx_1 \\
\approx & \int_{\pi}^{\pi+b} \left[- \int_{\frac{\pi-x_1}{b}}^1 m'_+(\pi) ubl(u) du + \int_{-1}^{\frac{\pi-x_1}{b}} \left(\alpha_\pi + Cub\right) l(u) du \right]^2 dx_1 \\
& + \int_{\pi-b}^{\pi} \left[\int_{\frac{\pi-x_1}{b}}^1 \left(-\alpha_\pi + Cub\right) l(u) du - \int_{-1}^{\frac{\pi-x_1}{b}} m'_-(\pi) ubl(u) du \right]^2 dx_1 \\
\approx & \; b\alpha_\pi^2 \left[\int_0^1 \left(\int_{-1}^{-v} l(u) du \right)^2 dv + \int_0^1 \left(\int_v^1 l(u) du \right)^2 dv \right] = 2b\alpha_\pi^2 \int_0^1 \left(\int_v^1 l(u) du \right)^2 dv.
\end{aligned}$$

The result follows. ■

Proposition 2 $I_{2n} \xrightarrow{d} N\left(0, \Sigma\right)$.

Proof.

$$\begin{aligned}
I_{2n} &= \frac{nh^{1/2}}{n\left(n-1\right)} \sum_i \sum_{j \neq i} 1_i^\Pi 1_j^\Pi \varepsilon_i \varepsilon_j K_{h,ij} \\
&\equiv \frac{nh^{1/2}}{n\left(n-1\right)} \sum_i \sum_{j \neq i} H_n(z_i, z_j) \equiv nh^{1/2} U_n,
\end{aligned}$$

where U_n is a second order degenerate U-statistic with kernel function H_n. We can apply theorem 1 of Hall (1984) to find its asymptotic distribution. Two conditions should be checked: (i) $E[H_n^2(z_1, z_2)] < \infty$; (ii)

$$\frac{E\left[G_n^2(z_1, z_2)\right] + n^{-1} E[H_n^4(z_1, z_2)]}{E^2[H_n^2(z_1, z_2)]} \to 0 \text{ as } n \to \infty,$$

where $G_n(z_1, z_2) = E[H_n(z_3, z_1) H_n(z_3, z_2) | z_1, z_2]$. This checking is very similar to that in lemma 3.3a of Zheng (1996), so omitted here. The conclusion is that

$$nU_n / \sqrt{2E[H_n^2(z_1, z_2)]} \xrightarrow{d} N(0, 1).$$

It is easy to check that

$$\begin{aligned} E[H_n^2(z_1,z_2)] &= E\left[1_1^{\Pi}1_2^{\Pi}K_{h,12}^2E[\varepsilon_1^2\varepsilon_2^2|x_1,x_2]\right] \\ &= \int_{\underline{\pi}}^{\overline{\pi}}\int_{\underline{\pi}}^{\overline{\pi}}\frac{1}{h^2}k^2\left(\frac{x_1-x_2}{h}\right)\sigma^2(x_1)\sigma^2(x_2)f(x_1)f(x_2)dx_1dx_2 \\ &= \int_{\underline{\pi}}^{\overline{\pi}}\int_{\frac{x-\overline{\pi}}{h}}^{\frac{x-\underline{\pi}}{h}}\frac{1}{h}k^2(u)\,\sigma^2(x)\sigma^2(x-hu)f(x)f(x-hu)dudx \\ &= \frac{1}{h}\int k^2(u)\,du\int_{\underline{\pi}}^{\overline{\pi}}\sigma^4(x)f^2(x)dx+o\left(\frac{1}{h}\right)\approx\frac{1}{h}\frac{\Sigma}{2}, \end{aligned}$$

so the result follows. ■

Proposition 3 $I_{3n}=o_p(1)$.

Proof.

$$\begin{aligned} I_{3n} &= \frac{nh^{1/2}}{n(n-1)}\sum_i\sum_{j\neq i}1_i^{\Pi}1_j^{\Pi}\widehat{\varepsilon}_i\widehat{\varepsilon}_jK_{h,ij}\leq\frac{nh^{1/2}}{2n(n-1)}\sum_i\sum_{j\neq i}\left\{1_i^{\Pi}\widehat{\varepsilon}_i^2+1_j^{\Pi}\widehat{\varepsilon}_j^2\right\}K_{h,ij} \\ &= \frac{nh^{1/2}}{n(n-1)}\sum_i\sum_{j\neq i}1_i^{\Pi}\widehat{\varepsilon}_i^2K_{h,ij}=h^{1/2}\sum_i1_i^{\Pi}\widehat{\varepsilon}_i^2\widetilde{f}_i\leq h^{1/2}\left(\sup_{x_i\in\Pi}\widetilde{f}_i\right)\sum_i1_i^{\Pi}\widehat{\varepsilon}_i^2 \\ &= h^{1/2}O_p(1)O_p\left(b^{-1}\right)=o_p(1), \end{aligned}$$

where $\sum_i 1_i^{\Pi}\widehat{\varepsilon}_i^2=O_p\left(b^{-1}\right)$ is shown in Lemma 2. ■

Proposition 4 *I_{4n} is $o_{P_m}(1)$ uniformly in m under H_0, and is $o_p(I_{1n})$ under H_1.*

Proof.

$$\begin{aligned} I_{4n} &= \frac{nh^{1/2}}{n(n-1)}\sum_i\sum_{j\neq i}1_i^{\Pi}1_j^{\Pi}\varepsilon_i(m_j-\widehat{m}_j)\widehat{f}_jK_{h,ij}\widehat{f}_j^{-1} \\ &\approx \frac{nh^{1/2}}{n(n-1)}\sum_i\sum_{j\neq i}\sum_{k\neq j}1_i^{\Pi}1_j^{\Pi}\varepsilon_i(m_j-m_k)L_{b,jk}K_{h,ij}, \end{aligned}$$

because $\sup_{x_j\in\Pi}\widehat{f}_j^{-1}=O_p(1)$. Note that $E[I_{4n}]=0$ and we need only calculate the second moment of I_{4n}. By a similar tedious analysis as in Proposition A.4 of Fan and Li (1996), we can show

$$E\left[I_{4n}^2\right]=\begin{cases}O(nhb^{2\eta}), & \text{under } H_0^{(1)},\\ O(nh\left(b^{2\eta}+b^3\right)), & \text{under } H_0^{(2)},\end{cases}$$

which is $o(1)$ under Assumption B(a) and B(b), respectively.

Under H_1, we can show

$$E\left[I_{4n}^2\right]=\begin{cases}O(nh\left(b^{2\eta}+b^3\right)), & \text{under } H_1^{(1)},\\ O(nh\left(b^{2\eta}+b\right)), & \text{under } H_1^{(2)},\end{cases}$$

so

$$I_{4n}=\begin{cases}O_p(n^{1/2}h^{1/2}b^{3/2})=o_p(I_{1n}), & \text{under } H_1^{(1)},\\ O_p\left(n^{1/2}h^{1/2}b^{1/2}\right)=o_p(I_{1n}), & \text{under } H_1^{(2)}.\end{cases}$$

■

Proposition 5 *I_{5n} is $o_{P_m}(1)$ uniformly in m under H_0, and is $o_p(I_{1n})$ under H_1.*

Proof. Under H_0,

$$\begin{aligned} I_{5n} &= \frac{nh^{1/2}}{n(n-1)} \sum_i \sum_{j \neq i} 1_i^\Pi 1_j^\Pi \widehat{\varepsilon}_i (m_j - \widehat{m}_j) K_{h,ij} \\ &\leq \frac{nh^{1/2}}{2n(n-1)} \sum_i \sum_{j \neq i} \left[1_i^\Pi \widehat{\varepsilon}_i^2 + 1_j^\Pi (m_j - \widehat{m}_j)^2 \right] K_{h,ij} \\ &= o_p(1), \end{aligned}$$

by Propositions 1 and 3. Under H_1, a similar analysis as in Proposition 4 can show $I_{5n} = o_p(I_{1n})$. ■

Proposition 6 *$I_{6n} = o_p(1)$.*

Proof.

$$\begin{aligned} I_{6n} &= \frac{nh^{1/2}}{n(n-1)} \sum_i \sum_{j \neq i} 1_i^\Pi 1_j^\Pi \varepsilon_i \widehat{\varepsilon}_j K_{h,ij} = \frac{nh^{1/2}}{n(n-1)} \sum_i \sum_{j \neq i} 1_i^\Pi 1_j^\Pi \varepsilon_i \widehat{\varepsilon}_j \widehat{f}_j K_{h,ij} \widehat{f}_j^{-1} \\ &= \frac{nh^{1/2}}{n(n-1)} \sum_i \sum_{j \neq i} 1_i^\Pi 1_j^\Pi \varepsilon_i \widehat{f}_i \widehat{\varepsilon}_j \widehat{f}_j K_{h,ij} \widehat{f}_j^{-1} \widehat{f}_i^{-1} \\ &\approx \frac{nh^{1/2}}{n(n-1)} \sum_i \sum_{j \neq i} 1_i^\Pi 1_j^\Pi \varepsilon_i \widehat{f}_i \widehat{\varepsilon}_j \widehat{f}_j K_{h,ij}, \end{aligned}$$

because $\sup_{x_j \in \Pi} \widehat{f}_j^{-1} \sup_{x_i \in \Pi} \widehat{f}_i^{-1} = O_p(1)$. By Proposition A.6 of Fan and Li (1996), $I_{6n} = o_p(1)$; note only that b (h) plays the same role as a^{q_1} (h^d) there. ■

Proposition 7 *$v_n^2 = \Sigma + o_p(1)$.*

Proof. It takes some algebra but is straightforward to show that

$$\begin{aligned} &\frac{h}{n(n-1)} \sum_i \sum_{j \neq i} 1_i^\Pi 1_j^\Pi K_{h,ij}^2 \widehat{e}_i^2 \widehat{e}_j^2 \\ = \; &\frac{h}{n(n-1)} \sum_i \sum_{j \neq i} 1_i^\Pi 1_j^\Pi K_{h,ij}^2 (\varepsilon_i + m_i - \overline{m}_i)^2 (\varepsilon_j + m_j - \overline{m}_j)^2 + o_p(1) \\ = \; &E\left[1_1^\Pi (\varepsilon_1 + m_1 - \overline{m}_1)^2 f_1 E_1 \left[(\varepsilon_1 + m_1 - \overline{m}_1)^2 \right] \right] \int k^2(u) du + o_p(1) \\ = \; &E\left[1_1^\Pi \varepsilon_1^2 f_1 E_1 \left[\varepsilon_1^2 \right] \right] \int k^2(u) du + E\left[1_1^\Pi (m_1 - \overline{m}_1)^2 f_1 E_1 \left[(m_1 - \overline{m}_1)^2 \right] \right] \int k^2(u) du + o_p(1) \\ = \; &E\left[1_x^\Pi f(x) \sigma^4(x) \right] \int k^2(u) du + E\left[1_1^\Pi (m_1 - \overline{m}_1)^4 f_1 \right] \int k^2(u) du + o_p(1) \end{aligned}$$

where $\overline{m}_i = E_i[\widehat{m}_i]$, and the second equality is because $h/b \to 0$ so that we can treat $x_i = x_j$ for x_j in a h neighborhood of x_i. By a similar proof as in Lemma 1, $E\left[1_1^\Pi (m_1 - \overline{m}_1)^4 f_1 \right]$ is at most $O(b)$ under H_0 and H_1, so the results follow. ■

Proposition 8 *$\widehat{\pi} - \pi_0 = O_p(h)$.*

Proof. We will apply Lemma 4 to prove this result. WLOG, assume $\alpha_{\pi_0} > 0$. In this case, $\theta = \pi, \theta_0 = \pi_0, \theta_n = \pi_n, \mathcal{N}_n = [\pi_0 - h, \pi_0 + h]$, $Q_n(\theta) = \overline{\alpha}_n^2(\pi), \widehat{Q}_n(\theta) = \widehat{\alpha}^2(\pi)$, where

$$\begin{aligned} \overline{\alpha}_n(\pi) &= \overline{m}_+(\pi) - \overline{m}_-(\pi), \pi_n = \arg\max_{\pi\in\Pi} \overline{\alpha}_n^2(\pi), \\ \overline{m}_+(\pi) &= \frac{\int_0^1 k(u)m(\pi+uh)f(\pi+uh)du}{\int_0^1 k(u)f(\pi+uh)du}, \overline{m}_-(\pi) = \frac{\int_{-1}^0 k(u)m(\pi+uh)f(\pi+uh)du}{\int_{-1}^0 k(u)f(\pi+uh)du}. \end{aligned}$$

We first check condition (iii). If we use the local constant estimator,

$$\begin{aligned} \widehat{m}_+(\pi) &= \frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\, y_j d_j(\pi)}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\, d_j(\pi)}, \\ \widehat{m}_-(\pi) &= \frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\, y_j (1-d_j(\pi))}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)(1-d_j(\pi))}. \end{aligned}$$

We can multiply the numerator and denominator by 2 and redefine $k(\cdot)$ such that $\int_0^1 k(u)du = \int_{-1}^0 k(u)du = 1$. By a similar argument as in Lemma B.1 of Newey (1994), we can show

$$\begin{aligned} \sup_{\pi\in\Pi}\left|\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\, y_j d_j(\pi) - \int_0^1 k(u)m(\pi+uh)f(\pi+uh)du\right| &= O_p\left(\sqrt{\frac{\ln n}{nh}}\right), \\ \sup_{\pi\in\Pi}\left|\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\, y_j (1-d_j(\pi)) - \int_{-1}^0 k(u)m(\pi+uh)f(\pi+uh)du\right| &= O_p\left(\sqrt{\frac{\ln n}{nh}}\right), \end{aligned}$$

$$\begin{aligned} \sup_{\pi\in\Pi}\left|\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\, d_j(\pi) - \int_0^1 k(u)f(\pi+uh)du\right| &= O_p\left(\sqrt{\frac{\ln n}{nh}}\right), \\ \sup_{\pi\in\Pi}\left|\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)(1-d_j(\pi)) - \int_{-1}^0 k(u)f(\pi+uh)du\right| &= O_p\left(\sqrt{\frac{\ln n}{nh}}\right). \end{aligned}$$

So

$$\sup_{\pi\in\Pi}\left|\widehat{\alpha}^2(\pi) - \overline{\alpha}_n^2(\pi)\right| \le \sup_{\pi\in\Pi}|\widehat{\alpha}(\pi) - \overline{\alpha}_n(\pi)| \sup_{\pi\in\Pi}|\widehat{\alpha}(\pi) + \overline{\alpha}_n(\pi)| = o_p(1)O_p(1) = o_p(1).$$

We then check condition (ii). By Assumptions M and F, $\sup_{\pi\in\Pi\backslash\mathcal{N}_n}|\overline{m}_+(\pi) - m(\pi)| = O(h)$, and $\sup_{\pi\in\Pi\backslash\mathcal{N}_n}|\overline{m}_-(\pi) - m(\pi)| = O(h)$, so $\sup_{\pi\in\Pi\backslash\mathcal{N}_n}|\overline{m}_+(\pi) - \overline{m}_-(\pi)| = O(h)$. On the contrary, for $\pi \in \mathcal{N}_n$, WLOG, let $\pi = \pi_0 - ah$, $a \in (0,1)$,

$$\begin{aligned} \overline{\alpha}_n(\pi) &= \frac{\int_0^1 k(u)m_{\pi_0}(\pi+uh)f(\pi+uh)du + \alpha_{\pi_0}\int_a^1 k(u)f(\pi+uh)du}{\int_0^1 k(u)f(\pi+uh)du} - \frac{\int_{-1}^0 k(u)m(\pi+uh)f(\pi+uh)du}{\int_{-1}^0 k(u)f(\pi+uh)du} \\ &= \alpha_{\pi_0}\frac{\int_a^1 k(u)f(\pi+uh)du}{\int_0^1 k(u)f(\pi+uh)du} + O(h) = \frac{\int_a^1 k(u)du}{\int_0^1 k(u)du}\alpha_{\pi_0} + O(h). \end{aligned}$$

Because $k(0) > 0$, $\frac{\int_a^1 k(u)du}{\int_0^1 k(u)du}\alpha_{\pi_0} < \alpha_{\pi_0}$. As a result, π_n converges to π_0, and there is a δ, say $\frac{\alpha_{\pi_0}^2}{2}$, such that $\sup_{\pi\in\Pi\backslash\mathcal{N}_n}\overline{\alpha}_n^2(\pi) < \overline{\alpha}_n^2(\pi_n) - \delta$. The proof is complete. ■

Remark 1 *In this proof, the limit objective function $\overline{\alpha}_n(\pi)$ depends on h and so on n. If its limit $\overline{\alpha}(\pi)$ as h converges to zero is taken as the limit objective function, then it is zero when $\pi \neq \pi_0$, and α_{π_0} when $\pi = \pi_0$.*

Because $\widehat{\alpha}(\pi)$ is a continuous function, the convergence from $\widehat{\alpha}(\pi)$ to $\overline{\alpha}(\pi)$ cannot be uniform. In this proof, we swell $\overline{\alpha}(\pi)$ to a continuous function $\overline{\alpha}_n(\pi)$ on $\mathcal{N}_n$ to make the convergence uniform.

Proposition 9 $\widehat{\pi}-\pi_0 = O_p\left(n^{-1}\right).$

Proof. WLOG, we use the same setup and notations as in the last proposition. For each n, the parameter space can be partitioned into the "shells" $S_{j,n}=\left\{\pi : 2^{j-1}<n\left|\pi-\pi_0\right|\le 2^j\right\}$ with j ranging over the integers. If $n\left|\widehat{\pi}-\pi_0\right|$ is larger than 2^J for a given integer J, then $\widehat{\pi}$ is in one of the shells $S_{j,n}$ with $j\ge J$. In that case the supremum of the map $\pi\mapsto\widehat{\alpha}^2(\pi)-\widehat{\alpha}^2(\pi_0)$ over this shell is nonnegative by the property of $\widehat{\pi}$.

$$
\begin{aligned}
& P\left(n\left|\widehat{\pi}-\pi_0\right|>2^J\right)\\
\le\ & P\left(\sup_{2^J<n|\pi-\pi_0|\le nh}\left(\widehat{\alpha}^2(\pi)-\widehat{\alpha}^2(\pi_0)\right)\ge 0\right)+P\left(\left|\widehat{\pi}-\pi_0\right|\ge h\right)\\
\le\ & \sum_{j=J}^{\log_2(nh)} P\left(\sup_{S_{j,n}}\widehat{\alpha}^2(\pi)\ge\widehat{\alpha}^2(\pi_0)\right)+P\left(\left|\widehat{\pi}-\pi_0\right|\ge h\right)\\
\le\ & \sum_{j=J}^{\log_2(nh)} P\left(\sup_{S_{j,n}}\widehat{\alpha}(\pi)-\widehat{\alpha}(\pi_0)>0\right)+\sum_{j=J}^{\log_2(nh)} P\left(\sup_{S_{j,n}}\widehat{\alpha}(\pi)+\widehat{\alpha}(\pi_0)<0\right)+P\left(\left|\widehat{\pi}-\pi_0\right|\ge h\right)
\end{aligned}
$$

We show the first sum converges to zero, the second is easier because we assume $\alpha_{\pi_0}>0$, and the third converges to zero by the consistency of $\widehat{\pi}$. First, suppose $\pi_0-h\le\pi<\pi_0$; then by the analysis in the last proposition,

$$
\begin{aligned}
& \overline{\alpha}_n(\pi)-\overline{\alpha}_n(\pi_0)\\
=\ & \frac{\int_0^1 k(u)m_{\pi_0}(\pi+uh)f(\pi+uh)du+\alpha_{\pi_0}\int_{\frac{\pi_0-\pi}{h}}^1 k(u)f(\pi+uh)du}{\int_0^1 k(u)f(\pi+uh)du}-\frac{\int_{-1}^0 k(u)m(\pi+uh)f(\pi+uh)du}{\int_{-1}^0 k(u)f(\pi+uh)du}\\
& -\frac{\int_0^1 k(u)m_{\pi_0}(\pi_0+uh)f(\pi_0+uh)du+\alpha_{\pi_0}\int_0^1 k(u)f(\pi_0+uh)du}{\int_0^1 k(u)f(\pi_0+uh)du}+\frac{\int_{-1}^0 k(u)m(\pi_0+uh)f(\pi_0+uh)du}{\int_{-1}^0 k(u)f(\pi_0+uh)du}\\
=\ & -\alpha_{\pi_0}\frac{\int_0^{\frac{\pi_0-\pi}{h}} k(u)f(\pi+uh)du}{\int_0^1 k(u)f(\pi+uh)du}+O\left(\pi-\pi_0\right)\\
\le\ & -C\left|\pi-\pi_0\right|/h,
\end{aligned}
$$

where the last inequality is because $k(0)>0$. This result holds when $\pi_0<\pi\le\pi_0+h$ by a similar argument. For $\delta\le h$,

$$
\begin{aligned}
& P\left(\sup_{|\pi-\pi_0|<\delta}\widehat{\alpha}(\pi)-\widehat{\alpha}(\pi_0)>0\right)\\
\le\ & P\left(\sup_{|\pi-\pi_0|<\delta}\left[\left(\widehat{\alpha}(\pi)-\overline{\alpha}_n(\pi)\right)-\left(\widehat{\alpha}(\pi_0)-\overline{\alpha}_n(\pi_0)\right)\right]>C\left|\pi-\pi_0\right|/h\right),
\end{aligned}
$$

so we need to analyze the process $(\widehat{\alpha}(\pi)-\overline{\alpha}_n(\pi))-(\widehat{\alpha}(\pi_0)-\overline{\alpha}_n(\pi_0))$, where $\overline{\alpha}_n(\pi)-\overline{\alpha}_n(\pi_0)$ is the centering process of $\widehat{\alpha}(\pi)-\widehat{\alpha}(\pi_0)$. Second, for $|\pi-\pi_0|\le h$, WLOG, suppose $\pi_0<\pi\le\pi_0+h$,

$$
\begin{aligned}
& \widehat{\alpha}(\pi)-\widehat{\alpha}(\pi_0)-(\overline{\alpha}_n(\pi)-\overline{\alpha}_n(\pi_0)) \\
= & \widehat{m}_+(\pi)-\widehat{m}_-(\pi)-(\widehat{m}_+(\pi_0)-\widehat{m}_-(\pi_0))-(\overline{\alpha}_n(\pi)-\overline{\alpha}_n(\pi_0)) \\
= & \frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,y_j d_j(\pi)}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,d_j(\pi)}-\frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,y_j(1-d_j(\pi))}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)(1-d_j(\pi))} \\
& -\left(\frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,y_j d_j(\pi_0)}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,d_j(\pi_0)}-\frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,y_j(1-d_j(\pi_0))}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)(1-d_j(\pi_0))}\right)-(\overline{\alpha}_n(\pi)-\overline{\alpha}_n(\pi_0)) \\
= & \frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,(m_{\pi_0}(x_j)+\alpha_{\pi_0})\,d_j(\pi)}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,d_j(\pi)}-\frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,m(x_j)(1-d_j(\pi))}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)(1-d_j(\pi))} \\
& -\frac{\frac{\alpha_{\pi_0}}{n}\sum_{j=1}^n k_h(x_j-\pi)\,1(\pi_0\le x_j<\pi)}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)(1-d_j(\pi))}-\left(\frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,(m_{\pi_0}(x_j)+\alpha_{\pi_0})\,d_j(\pi_0)}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,d_j(\pi_0)}\right. \\
& \left.-\frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,m(x_j)(1-d_j(\pi_0))}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)(1-d_j(\pi_0))}\right)-(\overline{\alpha}_n(\pi)-\overline{\alpha}_n(\pi_0)) \\
& +\frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,\varepsilon_j d_j(\pi)}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,d_j(\pi)}-\frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,\varepsilon_j(1-d_j(\pi))}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)(1-d_j(\pi))} \\
& -\left(\frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,\varepsilon_j d_j(\pi_0)}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,d_j(\pi_0)}-\frac{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,\varepsilon_j(1-d_j(\pi_0))}{\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)(1-d_j(\pi_0))}\right) \\
= & o_p\left(\frac{|\pi-\pi_0|}{h}\right)+f(\pi_0)^{-1}\left\{\left[\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,\varepsilon_j d_j(\pi)-\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,\varepsilon_j(1-d_j(\pi))\right]\right. \\
& \left.-\left[\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,\varepsilon_j d_j(\pi_0)-\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,\varepsilon_j(1-d_j(\pi_0))\right]\right\},
\end{aligned}
$$

where the last equality is from a tedious but straightforward analysis. Note that

$$
\begin{aligned}
& \frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,\varepsilon_j d_j(\pi)-\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,\varepsilon_j d_j(\pi_0) \\
& -\left(\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi)\,\varepsilon_j(1-d_j(\pi))-\frac{1}{n}\sum_{j=1}^n k_h(x_j-\pi_0)\,\varepsilon_j(1-d_j(\pi_0))\right) \\
= & \frac{1}{n}\sum_{j=1}^n (k_h(x_j-\pi)-k_h(x_j-\pi_0))\,\varepsilon_j d_j(\pi)-\left(\frac{1}{n}\sum_{j=1}^n (k_h(x_j-\pi)-k_h(x_j-\pi_0))\,\varepsilon_j(1-d_j(\pi_0))\right) \\
& -\frac{1}{n}\sum_{j=1}^n (k_h(x_j-\pi)+k_h(x_j-\pi_0))\,\varepsilon_j 1(\pi_0\le x_j<\pi).
\end{aligned}
$$

Since the three terms in the last equality are independent, the variance is bounded by

$$
\begin{aligned}
& \frac{1}{(nh)^2}\sum\nolimits_{j=1}^{n} E\left[\left(k\left(\frac{x_j-\pi}{h}\right)-k\left(\frac{x_j-\pi_0}{h}\right)\right)^2 \varepsilon_j^2 1\left(\pi \le x_j \le \pi+h\right)\right] \\
& +\frac{1}{(nh)^2}\sum\nolimits_{j=1}^{n} E\left[\left(k\left(\frac{x_j-\pi}{h}\right)-k\left(\frac{x_j-\pi_0}{h}\right)\right)^2 \varepsilon_j^2 1\left(\pi_0-h \le x_j < \pi_0\right)\right] \\
& +\frac{1}{(nh)^2}\sum\nolimits_{j=1}^{n} E\left[\left(k\left(\frac{x_j-\pi}{h}\right)+k\left(\frac{x_j-\pi_0}{h}\right)\right)^2 \varepsilon_j^2 1\left(\pi_0 \le x_j < \pi\right)\right] \\
\le\ & \frac{C}{(nh)^2}\left[nh\frac{(\pi-\pi_0)^2}{h^2}+nh\frac{(\pi-\pi_0)^2}{h^2}+n\left|\pi-\pi_0\right|\right] \le C\frac{n\left|\pi-\pi_0\right|}{(nh)^2}
\end{aligned}
$$

uniformly for $|\pi-\pi_0| \le h$. In consequence,

$$
\begin{aligned}
& P\left(\sup_{|\pi-\pi_0|<\delta} \widehat{\alpha}(\pi)-\widehat{\alpha}(\pi_0) > 0\right) \\
\le\ & C\, E\left[\left(\sup_{|\pi-\pi_0|<\delta}\left[(\widehat{\alpha}(\pi)-\overline{\alpha}_n(\pi))-(\widehat{\alpha}(\pi_0)-\overline{\alpha}_n(\pi_0))\right]\right)^2\right] \Big/ \frac{(\pi-\pi_0)^2}{h^2} \\
\le\ & C\frac{n\left|\pi-\pi_0\right|}{(nh)^2} \Big/ \frac{(\pi-\pi_0)^2}{h^2} \le \frac{C}{n\left|\pi-\pi_0\right|}
\end{aligned}
$$

by Markov's inequality. So

$$
\begin{aligned}
& \sum_{j=J}^{\log_2(nh)} P\left(\sup_{S_{j,n}} \widehat{\alpha}(\pi)-\widehat{\alpha}(\pi_0) > 0\right) \\
\le\ & \sum_{j\ge J}\frac{C}{n\cdot 2^j/n} = C\sum_{j\ge J} 2^{-j} \to 0
\end{aligned}
$$

as $J \to \infty$. The proof is complete. ■

Remark 2 *This proof extends Theorem 3.2.5 of Van der Vaart and Wellner (1996) by concentrating on a* h *neighborhood of* π_0*. The key point here is that the centering process* $\overline{\alpha}_n(\pi)-\overline{\alpha}_n(\pi_0)$ *is not differentiable at* π_0 *such that* π_0 *is easier to identify than in the regular case as shown intuitively in Section 4.2.*

Proposition 10 *For* v *in any compact set of* $\mathbb{R}$*,*

$$
\begin{aligned}
& nh\left(\widehat{\alpha}\left(\pi_0+\frac{v}{n}\right)-\widehat{\alpha}(\pi_0)\right) \\
=\ & -\frac{k(0)}{f(\pi_0)}\sum\nolimits_{j=1}^{n}\left\{(\alpha_{\pi_0}-2\varepsilon_j)\,1\left(\pi_0+\frac{v}{n}\le x_j<\pi_0\right)+(\alpha_{\pi_0}+2\varepsilon_j)\,1\left(\pi_0\le x_j<\pi_0+\frac{v}{n}\right)\right\}+o_p(1).
\end{aligned}
$$

Proof. Let $\pi = \pi_0 + \frac{v}{n}$, $v > 0$; then

$$\begin{aligned}
& nh\left(\widehat{\alpha}(\pi) - \widehat{\alpha}(\pi_0)\right) \\
= \quad & \frac{\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) y_j d_j(\pi)}{\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) d_j(\pi)} - \frac{\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) y_j\left(1-d_j(\pi)\right)}{\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right)\left(1-d_j(\pi)\right)} \\
& -\left(\frac{\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) y_j d_j(\pi_0)}{\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)} - \frac{\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) y_j\left(1-d_j(\pi_0)\right)}{\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right)\left(1-d_j(\pi_0)\right)}\right) \\
= \quad & 1\bigg/\left[\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) d_j(\pi)\cdot\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)\right] \\
& \cdot\left[\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right)\left(m_{\pi_0}(x_j)+\alpha_{\pi_0}+\varepsilon_j\right) d_j(\pi)\right. \\
& \left.-\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) d_j(\pi)\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right)\left(m_{\pi_0}(x_j)+\alpha_{\pi_0}+\varepsilon_j\right) d_j(\pi_0)\right] \\
& +1\bigg/\left[\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right)\left(1-d_j(\pi)\right)\cdot\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right)\left(1-d_j(\pi_0)\right)\right] \\
& \cdot\left[\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right)\left(1-d_j(\pi)\right)\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right)\left(m(x_j)+\varepsilon_j\right)\left(1-d_j(\pi_0)\right)\right. \\
& \left.-\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right)\left(1-d_j(\pi_0)\right)\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right)\left(m(x_j)+\varepsilon_j\right)\left(1-d_j(\pi_0)\right)\right] \\
& -\frac{\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right)\left(1-d_j(\pi_0)\right)\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right)\left(m_{\pi_0}(x_j)+\alpha_{\pi_0}+\varepsilon_j\right) 1\left(\pi_0 \le x_j < \pi\right)}{\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right)\left(1-d_j(\pi)\right)\cdot\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right)\left(1-d_j(\pi_0)\right)}, \\
\equiv \quad & T_1 + T_2 - T_3.
\end{aligned}$$

Let's analyze each term in turn. First,

$$\begin{aligned}
T_1 \quad = \quad & \frac{\frac{1}{nh}\sum\limits_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)\sum\limits_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) m_{\pi_0}(x_j) d_j(\pi) - \frac{1}{nh}\sum\limits_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) d_j(\pi)\sum\limits_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) m_{\pi_0}(x_j) d_j(\pi_0)}{\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) d_j(\pi)\cdot\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)} \\
& +\frac{\frac{1}{nh}\sum\limits_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)\sum\limits_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) d_j(\pi)\varepsilon_j - \frac{1}{nh}\sum\limits_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) d_j(\pi)\sum\limits_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)\varepsilon_j}{\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) d_j(\pi)\cdot\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)} \\
= \quad & 1\bigg/\left[\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) d_j(\pi)\cdot\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)\right] \\
& \cdot\left[\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)\sum_{j=1}^n\left(k\left(\frac{x_j-\pi}{h}\right)-k\left(\frac{x_j-\pi_0}{h}\right)\right)\left(m_{\pi_0}(x_j)-m_{\pi_0}(\pi_0)\right) d_j(\pi)\right. \\
& -\sum_{j=1}^n\left(k\left(\frac{x_j-\pi}{h}\right)-k\left(\frac{x_j-\pi_0}{h}\right)\right) d_j(\pi_0)\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right)\left(m_{\pi_0}(x_j)-m_{\pi_0}(\pi_0)\right) d_j(\pi_0) \\
& +\sum_{j=1}^n k\left(\frac{x_j-\pi}{h}\right) 1\left(\pi_0 \le x_j < \pi\right)\frac{1}{nh}\sum_{j=1}^n k\left(\frac{x_j-\pi_0}{h}\right)\left(m_{\pi_0}(x_j)-m_{\pi_0}(\pi_0)\right) d_j(\pi_0)
\end{aligned}$$

$$
\begin{aligned}
& -\frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0) \sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right)\left(m_{\pi_0}(x_j)-m_{\pi_0}(\pi_0)\right) 1(\pi_0 \le x_j < \pi)\Bigg] \\
& +1 \Bigg/ \left[\frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right) d_j(\pi) \cdot \frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)\right] \\
& \cdot\Bigg[\frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0) \sum\nolimits_{j=1}^{n}\left(k\left(\frac{x_j-\pi}{h}\right)-k\left(\frac{x_j-\pi_0}{h}\right)\right) d_j(\pi_0)\,\varepsilon_j \\
& -\sum\nolimits_{j=1}^{n}\left(k\left(\frac{x_j-\pi}{h}\right)-k\left(\frac{x_j-\pi_0}{h}\right)\right) d_j(\pi_0) \frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)\,\varepsilon_j \\
& +\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right) 1(\pi_0 \le x_j < \pi) \frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0)\,\varepsilon_j \\
& -\frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right) d_j(\pi_0) \sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right) 1(\pi_0 \le x_j < \pi)\,\varepsilon_j\Bigg] \\
\equiv\ & T_{11}+T_{12}.
\end{aligned}
$$

By the Lipschitz continuity of $m_{\pi_0}(\cdot)$ and $k(\cdot)$, $T_{11}=o_p(1)$. It is easy to show the first three terms of T_{12} are $o_p(1)$, but the last term of T_{12} is $O_p(1)$. Second,

$$
\begin{aligned}
& T_2-T_3 = 1 \Bigg/ \left[\frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right)(1-d_j(\pi)) \cdot \frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right)(1-d_j(\pi_0))\right] \\
& \cdot\Bigg[\frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right)(1-d_j(\pi)) \sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right) m_{\pi_0}(x_j)(1-d_j(\pi_0)) \\
& -\frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right)(1-d_j(\pi_0)) \sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right) m_{\pi_0}(x_j)(1-d_j(\pi))\Bigg] \\
& +1 \Bigg/ \left[\frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right)(1-d_j(\pi)) \cdot \frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right)(1-d_j(\pi_0))\right] \\
& \Bigg[\frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right)(1-d_j(\pi)) \sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right)(1-d_j(\pi_0))\,\varepsilon_j \\
& -\frac{1}{nh}\sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right)(1-d_j(\pi_0)) \sum\nolimits_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right)(1-d_j(\pi_0))\,\varepsilon_j\Bigg] \\
& -\frac{\frac{1}{nh}\sum_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right)(1-d_j(\pi_0)) \sum_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right)(\alpha_{\pi_0}+\varepsilon_j) 1(\pi_0 \le x_j < \pi)}{\frac{1}{nh}\sum_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right)(1-d_j(\pi)) \cdot \frac{1}{nh}\sum_{j=1}^{n} k\left(\frac{x_j-\pi_0}{h}\right)(1-d_j(\pi_0))}.
\end{aligned}
$$

By a similar analysis as T_1, we can show the first two terms of T_2-T_3 are $o_p(1)$, but the last term is $O_p(1)$. In summary, when $\pi=\pi_0+\frac{v}{n}$, $v>0$,

$$
\begin{aligned}
& nh\left(\widehat{\alpha}(\pi)-\widehat{\alpha}(\pi_0)\right) \\
=\ & -\frac{\sum_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right) 1(\pi_0 \le x_j < \pi)\,\varepsilon_j}{\frac{1}{nh}\sum_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right) d_j(\pi)} - \frac{\sum_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right)(\alpha_{\pi_0}+\varepsilon_j) 1(\pi_0 \le x_j < \pi)}{\frac{1}{nh}\sum_{j=1}^{n} k\left(\frac{x_j-\pi}{h}\right)(1-d_j(\pi))} + o_p(1) \\
=\ & -\frac{k(0)}{f(\pi_0)}\sum\nolimits_{j=1}^{n}(\alpha_{\pi_0}+2\varepsilon_j) 1(\pi_0 \le x_j < \pi) + o_p(1).
\end{aligned}
$$

Similarly, when $v<0$,

$$
nh\left(\widehat{\alpha}(\pi)-\widehat{\alpha}(\pi_0)\right) = -\frac{k(0)}{f(\pi_0)}\sum\nolimits_{j=1}^{n}(\alpha_{\pi_0}-2\varepsilon_j) 1(\pi_0 \le x_j < \pi)
$$

■

Appendix D: Lemmas

Lemma 1 *$\frac{1}{n}\sum_i 1_i^{\Pi}\left[(m_i-\widehat{m}_i)\right]^2$ is $O_p\left(b^{2\eta}+(nb)^{-1}\right)$ under $H_0^{(1)}$ and is $O_p\left(b^{2\eta}+b^3+(nb)^{-1}\right)$ under $H_0^{(2)}$.*

Proof. Because $H_0^{(2)}$ covers $H_0^{(1)}$, we need only prove the result under $H_0^{(2)}$.

$$\begin{aligned}\frac{1}{n}\sum_i 1_i^{\Pi}\left[(m_i-\widehat{m}_i)\right]^2 &= \frac{1}{n}\sum_i 1_i^{\Pi}\left[(m_i-\widehat{m}_i)\widehat{f}_i\right]^2\widehat{f}_i^{-2}\\ &\le \left(\sup_{x_i\in\Pi}\widehat{f}_i^{-1}\right)^2\frac{1}{n}\sum_i 1_i^{\Pi}\left[(m_i-\widehat{m}_i)\widehat{f}_i\right]^2\approx\frac{1}{n}\sum_i 1_i^{\Pi}\left[(m_i-\widehat{m}_i)\widehat{f}_i\right]^2\end{aligned}$$

because $\sup_{x_i\in\Pi}\widehat{f}_i^{-1}=O_p(1)$, which is a standard result. By identity of distribution,

$$E\left\{\frac{1}{n}\sum_i 1_i^{\Pi}\left[(m_i-\widehat{m}_i)\widehat{f}_i\right]^2\right\}=\frac{1}{n}\sum_i E\left[1_i^{\Pi}\left(\frac{1}{n-1}\sum_{j\neq i}(m_i-m_j)L_{b,ij}\right)^2\right]=\frac{1}{(n-1)^2}E\left[1_1^{\Pi}T^2\right],$$

where $T=\sum_{i\neq 1}t_i$, $t_i=(m_1-m_i)L_{b,1i}$. Note that $E\left[1_1^{\Pi}T^2\right]\le 2E\left\{1_1^{\Pi}\left[\sum_{i\neq 1}(t_i-t)\right]^2\right\}+2n^2E\left[1_1^{\Pi}t^2\right]$, where $t=E_1[t_i]$. Conditional on x_1, the t_i-t are independent with mean 0, so $E\left\{1_1^{\Pi}\left[\sum_{i\neq 1}(t_i-t)\right]^2\right\}=(n-1)E\left\{1_1^{\Pi}E_1\left[(t_i-t)^2\right]\right\}$. It remains to bound $E\left[1_1^{\Pi}t^2\right]$ and $E\left\{1_1^{\Pi}E_1\left[(t_i-t)^2\right]\right\}$.

First, analyze t. For $x_1\notin[\pi-b,\pi+b]$,

$$\begin{aligned}|t| &= \left|E\left[(m(x_2)-m(x_1))L_{b,21}|x_1\right]\right|\\ &= \left|\int(m(x_2)-m(x_1))f(x_2)\frac{1}{b}l\left(\frac{x_2-x_1}{b}\right)dx_2\right|\\ &= \left|\int\left(Q_m(x_2,x_1)+R_m(x_2,x_1)\right)\left(f(x_1)+Q_f(x_2,x_1)+R_f(x_2,x_1)\right)\frac{1}{b}l\left(\frac{x_2-x_1}{b}\right)dx_2\right|,\end{aligned}$$

where $Q_m(x_2,x_1)$ is $[s]$th-order Taylor expansion of $m(x_2)$ at $m(x_1)$, $R_m(x_2,x_1)$ is the remainder term, $Q_f(x_2,x_1)$ is $[\lambda]$th-order Taylor expansion of $f(x_2)$ at $f(x_1)$, and $R_f(x_2,x_1)$ is the remainder term. From Assumption K, $\int Q_m(x_2,x_1)\left(f(x_1)+Q_f(x_2,x_1)\right)\frac{1}{b}l\left(\frac{x_2-x_1}{b}\right)dx_2=0$, so $\left|E\left[(m(x_2)-m(x_1))L_{b,21}|x_1\right]\right|$ is bounded by

$$\begin{aligned}&\left|\int R_m(x_2,x_1)f(x_1)\frac{1}{b}l\left(\frac{x_2-x_1}{b}\right)dx_2\right|+\left|\int(m(x_2)-m(x_1))R_f(x_2,x_1)\frac{1}{b}l\left(\frac{x_2-x_1}{b}\right)dx_2\right|\\ \le\ & Cb^s+Cb^{\lambda+1}\le Cb^{\eta},\end{aligned}$$

where $\eta = \min(\lambda+1, s)$. When $x_1 \in [\pi - b, \pi + b]$, WLOG, assume $x_1 \leq \pi$; then

$$\begin{aligned}
& \left|E\left[\left(m(x_2)-m(x_1)\right)L_{b,21}|x_1\right]\right| \\
= & \left|\int_{x_1-b}^{\pi}\left(m(x_2)-m(x_1)\right)f(x_2)\frac{1}{b}l\left(\frac{x_2-x_1}{b}\right)dx_2\right| \\
& +\left|\int_{\pi}^{x_1+b}\left(m(x_2)-m(\pi)+m(\pi)-m(x_1)\right)f(x_2)\frac{1}{b}l\left(\frac{x_2-x_1}{b}\right)dx_2\right| \\
\leq & Cb.
\end{aligned}$$

So $E\left[t^2\right] \leq C\left(b^{2\eta}+b^3\right)$. Second, it is easy to show

$$E\left\{1_1^{\Pi}E_1\left[\left(t_i-t\right)^2\right]\right\} = O\left(b^{-1}\right)$$

by the boundedness of $m(\cdot)$ and $l(\cdot)$. Combining the analysis above, the result follows. ■

Lemma 2 $\frac{1}{n}\sum\limits_i 1_i^{\Pi}\widehat{\varepsilon}_i^2 = O_p\left((nb)^{-1}\right)$.

Proof. $\frac{1}{n}\sum\limits_i 1_i^{\Pi}\widehat{\varepsilon}_i^2 = \frac{1}{n}\sum\limits_i 1_i^{\Pi}\widehat{\varepsilon}_i^2\widehat{f}_i^2/\widehat{f}_i^2 \leq \sup\limits_{x_i\in\Pi}\widehat{f}_i^{-2}\frac{1}{n}\sum\limits_i 1_i^{\Pi}\widehat{\varepsilon}_i^2\widehat{f}_i^2 \approx \frac{1}{n}\sum\limits_i 1_i^{\Pi}\widehat{\varepsilon}_i^2\widehat{f}_i^2$. Note that

$$E\left\{\frac{1}{n}\sum_i 1_i^{\Pi}\widehat{\varepsilon}_i^2\widehat{f}_i^2\right\} = \frac{1}{n}\sum_i E\left[1_i^{\Pi}\left(\frac{1}{n-1}\sum_{j\neq i}\varepsilon_j L_{b,ij}\right)^2\right] = \frac{1}{(n-1)^2}E\left[1_1^{\Pi}T^2\right],$$

where $T = \sum\limits_{i\neq 1} t_i$, $t_i = \varepsilon_i L_{b,1i}$. Because $E\left[\varepsilon_1|x_1,\cdots,x_n\right] = 0$,

$$E\left[1_1^{\Pi}T^2\right] = (n-1)E\left[1_1^{\Pi}L_{b,1i}^2E\left[\varepsilon_i^2|x_i\right]\right] = (n-1)E\left[1_1^{\Pi}L_{b,1i}^2\sigma^2(x_i)\right] = O\left((n-1)/b\right),$$

and the result is proved. ■

Lemma 3 *Suppose $f(x)$ satisfies Assumption F; then*

$$\begin{aligned}
\underline{f}\inf_{\widetilde{m}(\cdot)\in\mathcal{C}(L,s)}\int_{\underline{\pi}}^{\overline{\pi}}\left(m(x)-\widetilde{m}(x)\right)^2dx & \leq \inf_{\widetilde{m}(\cdot)\in\mathcal{C}(L,s)}E\left[\left(m(x)-\widetilde{m}(x)\right)^2 1_x^{\Pi}\right] \\
& \leq \overline{f}\inf_{\widetilde{m}(\cdot)\in\mathcal{C}(L,s)}\int_{\underline{\pi}}^{\overline{\pi}}\left(m(x)-\widetilde{m}(x)\right)^2dx
\end{aligned}$$

for any $m(x) \in \mathcal{M}_{\Pi}(L,s)$.

Proof. Suppose

$$\overline{m}(\cdot) = \arg\inf_{\widetilde{m}(\cdot)\in\mathcal{C}(L,s)}E\left[\left(m(x)-\widetilde{m}(x)\right)^2 1_x^{\Pi}\right],$$

and

$$\overline{\overline{m}}(\cdot) = \arg\inf_{\widetilde{m}(\cdot)\in\mathcal{C}(L,s)}\int_{\underline{\pi}}^{\overline{\pi}}\left(m(x)-\widetilde{m}(x)\right)^2dx;$$

then

$$\inf_{\widetilde{m}(\cdot)\in\mathcal{C}(L,s)} E\left[(m(x)-\widetilde{m}(x))^2 1_x^{\Pi}\right] = E\left[(m(x)-\overline{m}(x))^2 1_x^{\Pi}\right] \le E\left[\left(m(x)-\overline{\overline{m}}(x)\right)^2 1_x^{\Pi}\right]$$
$$= \int_{\underline{\pi}}^{\overline{\pi}} \left(m(x)-\overline{\overline{m}}(x)\right)^2 f(x)dx \le \overline{f}\int_{\underline{\pi}}^{\overline{\pi}} \left(m(x)-\overline{\overline{m}}(x)\right)^2 dx = \overline{f}\inf_{\widetilde{m}(\cdot)\in\mathcal{C}(L,s)}\int_{\underline{\pi}}^{\overline{\pi}} (m(x)-\widetilde{m}(x))^2 dx.$$

Similarly,

$$\inf_{\widetilde{m}(\cdot)\in\mathcal{C}(L,s)}\int_{\underline{\pi}}^{\overline{\pi}} (m(x)-\widetilde{m}(x))^2 dx = \int_{\underline{\pi}}^{\overline{\pi}} \left(m(x)-\overline{\overline{m}}(x)\right)^2 dx \le \int_{\underline{\pi}}^{\overline{\pi}} (m(x)-\overline{m}(x))^2 dx$$
$$\le \int_{\underline{\pi}}^{\overline{\pi}} (m(x)-\overline{m}(x))^2 \frac{f(x)}{\underline{f}}dx = \frac{1}{\underline{f}}\int_{\underline{\pi}}^{\overline{\pi}} (m(x)-\overline{m}(x))^2 f(x)dx = \frac{1}{\underline{f}}\inf_{\widetilde{m}(\cdot)\in\mathcal{C}(L,s)} E\left[(m(x)-\widetilde{m}(x))^2 1_x^{\Pi}\right].$$

■

The following lemma is an extension of Theorem 2.1 of Newey and McFadden (1994).

Lemma 4 *If there is a deterministic function $Q_n(\theta)$ which depends on n such that (i) $Q_n(\theta)$ is minimized at θ_n which converges to a fixed point θ_0; (ii) for an open subset $\mathcal{N}_n$ containing θ_n, there exists a fixed positive number δ such that $\sup_{\theta\in\Theta\backslash\mathcal{N}_n} Q_n(\theta) < Q_n(\theta_n)-\delta$ for n large enough; (iii) $\sup_{\theta\in\Theta}\left|\widehat{Q}_n(\theta)-Q_n(\theta)\right| \xrightarrow{p} 0$, then $P\left(\widehat{\theta}\in\mathcal{N}_n\right)\to 1$.*

Proof. For any $\varepsilon>0$ we have w.p.a.1 (a) $\widehat{Q}_n\left(\widehat{\theta}\right) > \widehat{Q}_n(\theta_n)-\varepsilon/3$ by the fact that $\widehat{\theta}$ is the maximizer of $\widehat{Q}_n(\theta)$; (b) $Q_n\left(\widehat{\theta}\right) > \widehat{Q}_n\left(\widehat{\theta}\right)-\varepsilon/3$ by (iii); (c) $\widehat{Q}_n(\theta_n) > Q_n(\theta_n)-\varepsilon/3$ by (iii). Therefor, w.p.a.1,

$$Q_n\left(\widehat{\theta}\right) \overset{(b)}{>} \widehat{Q}_n\left(\widehat{\theta}\right)-\varepsilon/3 \overset{(a)}{>} \widehat{Q}_n(\theta_n)-2\varepsilon/3 \overset{(c)}{>} Q_n(\theta_n)-\varepsilon.$$

Thus, for any $\varepsilon>0$, $Q_n\left(\widehat{\theta}\right) > Q_n(\theta_n)-\varepsilon$ w.p.a.1. From (ii), $\sup_{\theta\in\Theta\backslash\mathcal{N}_n} Q_n(\theta) < Q_n(\theta_n)-\delta$ for n large enough. Choosing $\varepsilon=\delta<Q_n(\theta_n)-\sup_{\theta\in\Theta\backslash\mathcal{N}_n} Q_n(\theta)$, it follows that w.p.a.1 $Q_n\left(\widehat{\theta}\right) > Q_n(\theta_n)-\delta > \sup_{\theta\in\Theta\backslash\mathcal{N}_n} Q_n(\theta)$, hence $\widehat{\theta}\in\mathcal{N}_n$. ■

Additional References

De Jong, P., 1987, A Central Limit Theorem for Generalized Quadratic Forms, *Probability Theory and Related Fields*, 75, 261-277.

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

Newey, W.K., 1994, Kernel Estimation of Partial Means and a General Variance Estimator, *Econometric Theory*, 10, 233-253.

Newey, W.K. and D.L. McFadden, 1994, Large Sample Estimation and Hypothesis Testing, *Handbook of Econometrics*, Vol. 4, R.F. Eagle and D.L. McFadden, eds., Elsevier Science B.V., Ch. 36, 2111-2245.

Van der Vaart, A.W. and J.A. Wellner, 1996, *Weak Convergence and Empirical Processes: With Applications to Statistics*, New York : Springer.