Online Supplement to "Limit Theory for Locally Flat Functional Coefficient Regression"*

Peter C. B. Phillips[a,b,c,d] and Ying Wang[e]

[a]University of Auckland, [b]Yale University,
[c]Singapore Management University, [d]University of Southampton
[e]Renmin University of China

May 19, 2022

This Online Supplement contains three parts. Each of these addresses challenges that arise from the fact that the flatness degree parameter L is typically unknown in practical work. The first part proves that the oracle estimator $L_n^{\dagger}$ is consistent for L. The second part shows that the plausible estimator $\hat{L}$ defined in (3.12) of the main paper could be consistent for L with appropriately selected bandwidth. That approach is not practically feasible because knowledge of L is needed to ensure consistency. Some simulations are included to verify the findings. The third part demonstrates that the adaptive bias estimator $\hat{B}(z)$ that is designed for inference in the absence of knowledge of L is not consistent for the true bias $h^{L^*}\mathcal{B}_L(z)$ in either the stationary or the nonstationary case. Some further complications with the adaptive approach to inference are also discussed.

1 Proof that $L_n^{\dagger} \to_p L$

We use the Taylor series representation $\beta(z_t) - \beta(z) = \frac{\beta^{(L)}(\tilde{z}_t)}{L!}(z_t - z)^L$ where $\tilde{z}_t$ lies on the line segment between z_t and z and $\beta^{(L)}(z) \neq 0$ by assumption. Taking sample averages of the magnitudes of the weighted differentials $|\beta(z_t) - \beta(z)|\, K_{tz}$ we have

$$\frac{1}{n}\sum_{t=1}^{n} |\beta(z_t) - \beta(z)|\, K_{tz} = \frac{1}{n}\sum_{t=1}^{n} \left|\frac{\beta^{(L)}(\tilde{z}_t)}{L!}\right| |z_t - z|^L K_{tz}$$
$$\sim_a \int \left|\frac{\beta^{(L)}(\tilde{z}_t)}{L!}\right| |z_t - z|^L K\left(\frac{z_t - z}{h}\right) f(z_t)\, dz_t$$

*Phillips acknowledges research support from the NSF under Grant No. SES 18-50860 at Yale University and a Kelly Fellowship at the University of Auckland. Wang acknowledges support from the National Natural Science Foundation of China (Grant 72103197).

$$= h\int\left|\frac{\beta^{(L)}\left(z+\tilde{s}h\right)}{L!}\right|\left|s\right|^{L}h^{L}K\left(s\right)f\left(z+sh\right)ds\sim_{a}h^{L+1}f\left(z\right)\left|\frac{\beta^{(L)}\left(z\right)}{L!}\right|\int\left|s\right|^{L}K\left(s\right)ds,$$

with $\tilde{s}$ on the line segment between s and 0. Further,

$$\frac{1}{n}\sum_{t=1}^{n}K_{tz}\sim_{a}\int K\left(\frac{z_{t}-z}{h}\right)f\left(z_{t}\right)dz_{t}=h\int K\left(s\right)f\left(z+sh\right)ds\sim_{a}hf\left(z\right).$$

Then

$$\begin{aligned}
&\log\left(\frac{1}{n}\sum_{t=1}^{n}\left|\beta(z_{t})-\beta(z)\right|w_{tz}\right)=\log\left(\frac{1}{n}\sum_{t=1}^{n}\left|\beta(z_{t})-\beta(z)\right|K_{tz}\right)-\log\left(\frac{1}{n}\sum_{t=1}^{n}K_{tz}\right)\\
&\sim_{a}\log\left(h^{L+1}f\left(z\right)\left|\frac{\beta^{(L)}\left(z\right)}{L!}\right|\int\left|s\right|^{L}K\left(s\right)ds\right)-\log\left(hf\left(z\right)\right)\\
&=\left(L+1\right)\log\left(h\right)+\log\left(\left|\frac{\beta^{(L)}\left(z\right)}{L!}\right|\int\left|s\right|^{L}K\left(s\right)ds\right)-\log\left(h\right)\\
&=L\log\left(h\right)+\log\left(\left|\frac{\beta^{(L)}\left(z\right)}{L!}\right|\int\left|s\right|^{L}K\left(s\right)ds\right).
\end{aligned}$$

It follows that as $n\to\infty$ and $h\to 0$

$$L_{n}^{\dagger}=\frac{1}{\log\left(h\right)}\log\left(\sum_{t=1}^{n}\left|\beta(z_{t})-\beta(z)\right|w_{tz}\right)\sim_{a}L+\frac{1}{\log\left(h\right)}\log\left(\left|\frac{\beta^{(L)}\left(z\right)}{L!}\right|\int\left|s\right|^{L}K\left(s\right)ds\right)\to_{p}L, \tag{1.1}$$

since $\beta^{L}\left(z\right)\neq 0$, and thus

$$\log(h)(L_{n}^{\dagger}-L)=\log\left(\left|\frac{\beta^{(L)}\left(z\right)}{L!}\right|\right)+\log\left(\int\left|s\right|^{L}K\left(s\right)ds\right)+o_{p}(1), \tag{1.2}$$

so that the rate of convergence of $L_{n}^{\dagger}$ is $O(log(h))$ but with a deterministic bias function as given on the right side of (1.2). ■

2 Discussion of the properties of $\hat{L}$

This section studies the properties of the estimator $\hat{L}$ given in (3.12) of the main paper. We begin with the simple case where $L=1$. Note that

$$\frac{1}{n}\sum_{t=1}^{n}\left|\hat{\beta}(z_{t})-\hat{\beta}(z)\right|K_{tz}\sim_{a}\mathbb{E}\left[\left|\hat{\beta}(z_{t})-\hat{\beta}(z)\right|K_{tz}\right]$$

$$= \int \left|\hat{\beta}(z_t) - \hat{\beta}(z)\right| K_{tz} f(z_t) dz_t$$

$$= h \int \left|\hat{\beta}(z + ph) - \hat{\beta}(z)\right| K(p) f(z + ph) dp \tag{2.1}$$

$$= h \int \left| g^{(1)}(z) ph + \frac{g^{(2)}(z)}{2} p^2 h^2 + ... \right| K(p) f(z + ph) dp. \tag{2.2}$$

To avoid notational confusion, we denote $g(z) = \hat{\beta}(z)$, $g^{(1)}(z) = \partial \hat{\beta}(z)/\partial z$, the first derivative of the estimator $\hat{\beta}(z)$, and similarly, $g^{(2)}(z) = \partial^2 \hat{\beta}(z)/\partial z^2$.

To simplify notation we take the case where x_t is a scalar and stationary process, and is independent of the process z_t. Then $g(z) = \sum x_t y_t K_{tz} / \sum x_t^2 K_{tz}$. Let $g_1(z) = \sum x_t y_t K_{tz}$ and $g_2(z) = \sum x_t^2 K_{tz}$. It follows that $g^{(1)}(z) = [g_1^{(1)}(z) g_2(z) - g_1(z) g_2^{(1)}(z)]/g_2^2(z)$.

We first examine the property of $g^{(1)}(z)$ in (2.2). We start with $g_1^{(1)}(z)$ and $g_2^{(1)}(z)$. Note that $\partial K_{tz}/\partial z = \partial K(\frac{z_t - z}{h})/\partial(\frac{z_t - z}{h}) \times \partial(\frac{z_t - z}{h})/\partial z = (-\frac{1}{h}) G_{tz}$, where $G_{tz} = G((z_t - z)/h)$ and $G(u) = \partial K(u)/\partial u$. It is not hard to verify that $G(u)$ satisfies: $\int G(u) du = 0, \int u G(u) = \mu_1(G) \neq 0, \int u^2 G(u) = 0, \int u^3 G(u) = \mu_3(G) \neq 0, \int G^2(u) du = \nu_0(G) \neq 0$. For both the second order Epanechnikov kernel and the Gaussian kernel it is easy to verify that $\mu_1(G) = -1$. With the help of $G(u)$ we have

$$g_1^{(1)}(z) = \partial g_1(z)/\partial z = \sum x_t y_t \partial K_{tz}/\partial z = -\frac{1}{h} \sum x_t y_t G_{tz} = -\frac{1}{h} \left(\sum x_t^2 \beta(z_t) G_{tz} + \sum x_t u_t G_{tz} \right),$$

and

$$g_2^{(1)}(z) = \partial g_2(z)/\partial z = \sum x_t^2 \partial K_{tz}/\partial z = -\frac{1}{h} \sum x_t^2 G_{tz}.$$

It follows that

$$\begin{aligned}
g^{(1)}(z) &= \frac{-\frac{1}{h}\left(\sum x_t^2 \beta(z_t) G_{tz} + \sum x_t u_t G_{tz}\right) \sum x_t^2 K_{tz} + \left(\sum x_t^2 \beta(z_t) K_{tz} + \sum x_t u_t K_{tz}\right) \frac{1}{h} \sum x_t^2 G_{tz}}{\left(\sum x_t^2 K_{tz}\right)^2} \\
&= \frac{-\frac{1}{h} \sum x_t^2 \beta(z_t) G_{tz} \sum x_t^2 K_{tz} + \frac{1}{h} \sum x_t^2 \beta(z_t) K_{tz} \sum x_t^2 G_{tz}}{\left(\sum x_t^2 K_{tz}\right)^2} \\
&+ \frac{-\frac{1}{h} \sum x_t u_t G_{tz} \sum x_t^2 K_{tz} + \frac{1}{h} \sum x_t u_t K_{tz} \sum x_t^2 G_{tz}}{\left(\sum x_t^2 K_{tz}\right)^2} \\
&\equiv \Pi_1(z) + \Pi_2(z), \qquad (2.3)
\end{aligned}$$

defining $\Pi_1(z)$ and $\Pi_2(z)$. It is easy to see that $\Pi_2(z)$ contributes to the asymptotic distribution and $\Pi_1(z)$ contributes bias and possibly to the asymptotic distribution.

$\Pi_1(z)$ and $\Pi_2(z)$ are now considered separately, starting with $\Pi_1(z)$. We need the following preliminary results, whose proofs are given at the end of the section:

$$\begin{aligned}\mathbb{E}\beta(z_t)G_{tz} &= -h^2\left[\beta(z)f^{(1)}(z)+\beta^{(1)}(z)f(z)\right]\\ &-\frac{3}{5}h^4\left[\beta^{(1)}(z)f^{(2)}(z)/2+\beta^{(2)}(z)f^{(1)}(z)/2+\beta^{(3)}(z)f(z)/6+\beta(z)f^{(3)}(z)/6\right]+o(h^4)\\ &\equiv \mu_1(G)h^2\Delta_1(z)+\mu_3(G)h^4\Delta_3(z)+o(h^4),\end{aligned} \tag{2.4}$$

$$\mathbb{E}\beta^2(z_t)G_{tz}^2 = h\beta(z)f(z)\int G^2(u)du+O(h^3)=\nu_0(G)h\beta(z)f(z)+O(h^3)=O(h), \tag{2.5}$$

$$\mathbb{E}G_{tz}=\mu_1(G)h^2f^{(1)}(z)+\mu_3(G)h^4f^{(3)}(z)/6+o(h^4), \tag{2.6}$$

and

$$\mathbb{E}G_{tz}^2=\nu_0(G)hf(z)+o(h)=O(h). \tag{2.7}$$

Combining (2.4) and (2.5) gives

$$\sum x_t^2\beta(z_t)G_{tz}\sim_a n\mathbb{E}x_t^2\mathbb{E}\beta(z_t)G_{tz}+O_p(\sqrt{nh})\sim_a \mu_1(G)nh^2\Delta_1(z)\mathbb{E}x_t^2+O_p(\sqrt{nh}). \tag{2.8}$$

Combining (2.6) and (2.7) we have

$$\sum x_t^2G_{tz}\sim_a n\mathbb{E}x_t^2\mathbb{E}G_{tz}+O_p(\sqrt{nh})\sim_a \mu_1(G)nh^2f^{(1)}(z)\mathbb{E}x_t^2+O_p(\sqrt{nh}). \tag{2.9}$$

The following results are standard and readily obtained:

$$\mathbb{E}K_{tz}=hf(z)+h^3f^{(2)}(z)\mu_2(K)/2+o(h^3), \tag{2.10}$$

$$\mathbb{E}K_{tz}^2=hf(z)\nu_0(K)+o(h)=O(h), \tag{2.11}$$

$$\sum x_t^2K_{tz}\sim_a nhf(z)\mathbb{E}x_t^2, \tag{2.12}$$

$$\begin{aligned}\mathbb{E}\beta(z_t)K_{tz} &= h\beta(z)f(z)+h^3\left[\beta(z)f^{(2)}(z)/2+\beta^{(1)}(z)f^{(1)}(z)+\beta^{(2)}(z)f(z)/2\right]\mu_2(K)\\ &\equiv h\beta(z)f(z)+h^3\Delta_2(z)\mu_2(K),\end{aligned} \tag{2.13}$$

$$\mathbb{E}\beta^2(z_t)K_{tz}^2=h\beta^2(z)f(z)\nu_0(K)+o(h)=O(h), \tag{2.14}$$

$$\sum x_t^2\beta(z_t)K_{tz}\sim_a n\mathbb{E}x_t^2\mathbb{E}\beta(z_t)K_{tz}+O_p(\sqrt{nh})\sim_a nh\beta(z)f(z). \tag{2.15}$$

Combining (2.8), (2.9), (2.12), and (2.15) we have

$$\Pi_1(z)=\frac{-\frac{1}{h}\sum x_t^2\beta(z_t)G_{tz}\sum x_t^2K_{tz}+\frac{1}{h}\sum x_t^2\beta(z_t)K_{tz}\sum x_t^2G_{tz}}{\left(\sum x_t^2K_{tz}\right)^2}$$

$$
\begin{aligned}
&\sim_a \frac{-\frac{1}{h}\left[\mu_1(G)nh^2\Delta_1(z)\mathbb{E}x_t^2 + O_p(\sqrt{nh})\right](nhf(z)\mathbb{E}x_t^2) + \frac{1}{h}nh\beta(z)f(z)\left[\mu_1(G)nh^2 f^{(1)}(z)\mathbb{E}x_t^2 + O_p(\sqrt{nh})\right]}{\left[nhf(z)\mathbb{E}x_t^2\right]^2} \\
&= \frac{-\mu_1(G)nh\Delta_1(z)\mathbb{E}x_t^2(nhf(z)\mathbb{E}x_t^2) + \mu_1(G)n\beta(z)f(z)nh^2 f^{(1)}(z)\mathbb{E}x_t^2}{\left[nhf(z)\mathbb{E}x_t^2\right]^2} \\
&+ \frac{-\frac{1}{h}O_p(\sqrt{nh})nhf(z)\mathbb{E}x_t^2 + n\beta(z)f(z)O_p(\sqrt{nh})}{\left[nhf(z)\mathbb{E}x_t^2\right]^2} \\
&= -\mu_1(G)\Delta_1(z)/f(z) + \mu_1(G)\beta(z)f^{(1)}(z)/f(z) + O_p(1/\sqrt{nh^3}) \\
&= -\mu_1(G)\beta^{(1)}(z) + O_p(1/\sqrt{nh^3}),
\end{aligned} \tag{2.16}
$$

and

$$
\begin{aligned}
\Pi_2(z) &= \frac{-\frac{1}{h}\sum x_t u_t G_{tz}\sum x_t^2 K_{tz} + \frac{1}{h}\sum x_t u_t K_{tz}\sum x_t^2 G_{tz}}{\left(\sum x_t^2 K_{tz}\right)^2} \\
&= \frac{-\frac{1}{h}O_p(\sqrt{nh})nhf(z)\mathbb{E}x_t^2 + \frac{1}{h}O_p(\sqrt{nh})\left[\mu_1(G)nh^2 f^{(1)}(z)\mathbb{E}x_t^2 + O_p(\sqrt{nh})\right]}{\left[nhf(z)\mathbb{E}x_t^2\right]^2} \\
&= O_p\left(1/\sqrt{nh^3}\right).
\end{aligned} \tag{2.17}
$$

In view of (2.3), (2.16), and (2.17), we obtain

$$
g^{(1)}(z) = -\mu_1(G)\beta^{(1)}(z) + O_p\left(1/\sqrt{nh^3}\right). \tag{2.18}
$$

Note that $\mu_1(G) = -1$ for both second order Epanechnikov and Gaussian kernels. Result (2.18) suggests that when $L = 1$ and $\beta^{(1)}(z) \neq 0$, $g^{(1)}(z) = O_p(1)$ as $nh^3 \to c \in (0, \infty]$. Following (2.2) we have

$$
\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}\left|\hat{\beta}(z_t) - \hat{\beta}(z)\right| K_{tz} &\sim_a h\int \left|g^{(1)}(z)ph\right| K(p)f(z+ph)dp \\
&\sim_a h^2\left|g^{(1)}(z)\right| f(z)\int |p|K(p)dp,
\end{aligned} \tag{2.19}
$$

and then

$$
\begin{aligned}
\hat{L} &= \frac{1}{\log(h)}\log\left(\frac{1}{n}\sum_{t=1}^{n}\left|\hat{\beta}(z_t) - \hat{\beta}(z)\right| K_{tz} \Big/ \frac{1}{n}\sum_{t=1}^{n} K_{tz}\right) \\
&\sim_a \frac{1}{\log(h)}\log\left(h^2\left|g^{(1)}(z)\right| f(z)\int |p|K(p)dp \Big/ hf(z)\right) \\
&\sim_a \frac{1}{\log(h)}\left[\log(h) + \log\left(\left|g^{(1)}(z)\right|\right) + \log\left(\int |p|K(p)dp\right)\right]
\end{aligned}
$$

$$\sim_a 1 + \log\left(\left|g^{(1)}(z)\right|\right)\Big/\log(h) + \log\left(\int |p|K(p)dp\right)/\log(h) \xrightarrow{p} 1.$$

This proves that $\hat{L}$ is consistent when $L = 1$ under the condition $nh^3 \to c \in (0, \infty]$.

When $L > 1$ and $\beta^{(1)}(z) = 0$, result (2.18) is insufficient. For this situation we need to consider higher order bias terms of $g^{(1)}(z)$ to determine whether $\hat{L}$ is consistent. For simplicity, we consider the case where $L = 2$. The function $\beta(\cdot)$ appears only in the numerator of $\Pi_1(z)$. So the result for $\Pi_2(z)$ given in (2.17) still applies in the case $L = 2$. We therefore focus on the numerator of $\Pi_1(z)$, and in particular $\sum x_t^2\beta(z_t)G_{tz}$ and $\sum x_t^2\beta(z_t)K_{tz}$. When $\beta^{(1)}(z) = 0$, result (2.4) becomes

$$\begin{aligned}\mathbb{E}\beta(z_t)G_{tz} &= \mu_1(G)h^2\beta(z)f^{(1)}(z) + \mu_3(G)h^4\left[\beta^{(2)}(z)f^{(1)}(z)/2 + \beta^{(3)}(z)f(z)/6 + \beta(z)f^{(3)}(z)/6\right] + \ldots \\ &\equiv \mu_1(G)h^2\beta(z)f^{(1)}(z) + \mu_3(G)h^4\Delta_3^*(z),\end{aligned} \tag{2.20}$$

and (2.13) becomes

$$\mathbb{E}\beta(z_t)K_{tz} = h\beta(z)f(z) + h^3\left[\beta(z)f^{(2)}(z)/2 + \beta^{(2)}(z)f(z)/2\right]\mu_2(K) + \ldots \equiv h\beta(z)f(z) + h^3\Delta_2^*(z)\mu_2(K). \tag{2.21}$$

For the numerator of $\Pi_1(z)$ we then have

$$\begin{aligned}&-\frac{1}{h}\sum x_t^2\beta(z_t)G_{tz}\sum x_t^2K_{tz} + \frac{1}{h}\sum x_t^2\beta(z_t)K_{tz}\sum x_t^2G_{tz} \\ &\sim_a -\frac{1}{h}\left[n\left(\mu_1(G)h^2\beta(z)f^{(1)}(z) + \mu_3(G)h^4\Delta_3^*(z)\right)\mathbb{E}x_t^2 + O_p(\sqrt{nh})\right]\left[n\left(hf(z) + h^3f^{(2)}(z)\mu_2(K)/2\right)\mathbb{E}x_t^2\right] \\ &+\frac{1}{h}\left[n\left(h\beta(z)f(z) + h^3\Delta_2^*(z)\mu_2(K)\right)\mathbb{E}x_t^2\right]\left[n\left(\mu_1(G)h^2f^{(1)}(z) + \mu_3(G)h^4f^{(3)}(z)/6\right)\mathbb{E}x_t^2 + O_p(\sqrt{nh})\right] \\ &\sim_a n^2h^4\left[\left(\mu_1(G)\mu_2(K) - \mu_3(G)\right)f(z)f^{(1)}(z)\beta^{(2)}(z)/2 - \mu_3(G)f^2(z)\beta^{(3)}(z)/6\right](\mathbb{E}x_t^2)^2 + O_p\left(n\sqrt{nh}\right).\end{aligned}$$

It follows that

$$\begin{aligned}\Pi_1(z) &\sim_a \frac{n^2h^4\left[(\mu_1(G)\mu_2(K) - \mu_3(G))f(z)f^{(1)}(z)\beta^{(2)}(z)/2 - \mu_3(G)f^2(z)\beta^{(3)}(z)/6\right](\mathbb{E}x_t^2)^2 + O_p\left(n\sqrt{nh}\right)}{(nhf(z)\mathbb{E}x_t^2)^2} \\ &\sim_a h^2\left[\frac{\mu_1(G)\mu_2(K) - \mu_3(G)}{2}f^{-1}(z)f^{(1)}(z)\beta^{(2)}(z) - \frac{\mu_3(G)}{6}\beta^{(3)}(z)\right] + O_p\left(1/\sqrt{nh^3}\right).\end{aligned} \tag{2.22}$$

Note that $\Pi_2(z) = O_p(1/\sqrt{nh^3})$ in the case $L = 2$. We have

$$g^{(1)}(z) \sim_a h^2\left(\frac{\mu_1(G)\mu_2(K) - \mu_3(G)}{2}f^{-1}(z)f^{(1)}(z)\beta^{(2)}(z) - \frac{\mu_3(G)}{6}\beta^{(3)}(z)\right) + O_p\left(1/\sqrt{nh^3}\right). \tag{2.23}$$

Following (2.2), we have

$$\frac{1}{n}\sum_{t=1}^{n}\left|\hat{\beta}(z_t)-\hat{\beta}(z)\right|K_{tz}\sim_a h\int\left|g^{(1)}(z)ph+\frac{g^{(2)}(z)}{2}p^2h^2+...\right|K(p)f(z+ph)dp. \tag{2.24}$$

Given $g^{(1)}(z)=O_p(h^2+1/\sqrt{nh^3})$, to determine the order of the RHS of (2.24), we need to study the order of $g^{(2)}(z)$, which is complicated. Based on the result that $g^{(1)}(z)\to_p -\mu_1(G)\beta^{(1)}(z)$ in the case of $L=1$, we conjecture that $g^{(2)}(z)\to_p const.\times\beta^{(2)}(z)\neq 0$ in the case of $L=2$. Then $g^{(2)}(z)=O_p(1)$ when $L=2$ and from (2.24) we deduce

$$\frac{1}{n}\sum_{t=1}^{n}\left|\hat{\beta}(z_t)-\hat{\beta}(z)\right|K_{tz}\sim_a h\left[O_p\left(h(h^2+1/\sqrt{nh^3})\right)+O_p(h^2)\right]=O_p\left(h^3+\sqrt{h/n}\right). \tag{2.25}$$

Consequently,

$$\begin{aligned}\hat{L}&=\frac{1}{\log(h)}\log\left(\frac{1}{n}\sum_{t=1}^{n}\left|\hat{\beta}(z_t)-\hat{\beta}(z)\right|K_{tz}\Big/\frac{1}{n}\sum_{t=1}^{n}K_{tz}\right)\\&\sim_a\frac{1}{\log(h)}\log\left(O_p\left(h^3+\sqrt{h/n}\right)\Big/hf(z)\right)\\&\sim_a\frac{1}{\log(h)}\log\left(O_p\left(h^2+1/\sqrt{nh}\right)\right)\\&\to 2\text{ when }nh^5\to c\in(0,\infty].\end{aligned} \tag{2.26}$$

Result (2.26) implies that with an appropriately chosen bandwidth $\hat{L}$ can deliver a consistent estimate in the case $L=2$. We expect that for large L results parallel to (2.25) would be very involved and include n,h and L. Then, similar to (2.26), for carefully chosen h, we expect that $\hat{L}$ can deliver a consistent estimate of L. Based on our present findings, we conjecture that the rate condition $nh^{2L+1}\to c\in(0,\infty]$ is needed for $\hat{L}$ to be consistent. This condition is intuitive, indicating that larger bandwidth orders are needed to achieve consistency the flatter is the function (i.e., the true L is larger). This intuition is confirmed in the simulations reported below. But since L itself is unknown, it is impossible to determine the appropriate order of h before estimation. So $\hat{L}$ is not a feasible estimator in general.

To corroborate the above analysis we conduct simulations to check the performance of $\hat{L}$ under different bandwidth orders. The data generating process is: $z_t\sim iidN(0,1)$, and x_t follows stationary $AR(1)$ process with coefficient 0.5. A second order Epanechnikov kernel is used. Bandwidth h is determined by $h=\hat{\sigma}_z n^{\gamma}$. Two β functions are considered: $\beta_1(z)=z^2+z+1$ and $\beta_2(z)=z^3+1$. For $\beta_1(z)$, we have $L=1$ at point $z=0.5$ and $L=2$ at point $z=-0.5$. For $\beta_2(z)$ we have $L=3$ at point $z=0$. Tables 1-3 report the performance of $\hat{L}$ at those three points with different bandwidth order γ. For all three tables, the left panel uses 'too small' a bandwidth with $nh^{2L+1}\to 0$. In this case, $\hat{L}$ is not consistent according to our analysis.

Table 1: Performance of $\hat{L}$ at point $z = 0.5$ with $L = 1$

	$\gamma = -2/5$			$\gamma = -1/3$			$\gamma = -1/5$		
n	*mean*	*var*	*MSE*	*mean*	*var*	*MSE*	*mean*	*var*	*MSE*
200	1.0866	0.1171	0.1246	1.2254	0.1394	0.1902	1.3869	0.1170	0.2666
400	1.0525	0.0820	0.0848	1.1915	0.0960	0.1327	1.3061	0.0525	0.1462
800	1.0306	0.0593	0.0603	1.1724	0.0732	0.1029	1.2586	0.0276	0.0945
1600	1.0080	0.0460	0.0461	1.1516	0.0570	0.0800	1.2218	0.0149	0.0641
3200	0.9917	0.0355	0.0356	1.1396	0.0452	0.0646	1.1971	0.0083	0.0471
6400	0.9758	0.0297	0.0303	1.1261	0.0383	0.0542	1.1760	0.0048	0.0358

Table 2: Performance of $\hat{L}$ at point $z = -0.5$ with $L = 2$

	$\gamma = -1/3$			$\gamma = -1/5$			$\gamma = -1/7$		
n	*mean*	*var*	*MSE*	*mean*	*var*	*MSE*	*mean*	*var*	*MSE*
200	1.5198	0.1507	0.3812	2.8166	0.3491	1.0158	3.8599	0.5702	4.0294
400	1.4652	0.1041	0.3902	2.7092	0.2430	0.7460	3.7024	0.3812	3.2792
800	1.4177	0.0752	0.4142	2.6436	0.1839	0.5981	3.5665	0.2619	2.7157
1600	1.3790	0.0587	0.4443	2.5786	0.1466	0.4815	3.4400	0.1849	2.2586
3200	1.3444	0.0477	0.4776	2.5294	0.1259	0.4062	3.3368	0.1346	1.9216

Bandwidths in the middle and right panels satisfy the consistency condition obtained above. The middle panel has $nh^{2L+1} \to const$ and the right panel has $nh^{2L+1} \to \infty$. Results are obtained with 10,000 replications.

From findings in the three tables we draw the following conclusions: (i) with bandwidth too small as in the left panel of the three tables, $\hat{L}$ is evidently inconsistent. Especially in Table 2 and 3, the inconsistency is revealed from the increasing MSE as n becomes very large. In Table 1, although MSE is decreasing, we can expect that bias will keep growing with n and this will finally lead to inconsistency. In particular, the 'mean' column shows that too small a bandwidth eventually causes $\hat{L}$ to underestimate L. (ii) When condition $nh^{2L+1} \to c \in (0, \infty]$ is satisfied, the results in the middle and right panels of all three tables show that MSE and bias are both monotonically decreasing as n increases, corroborating consistency of $\hat{L}$. When $nh^{2L+1} \to \infty$ and a larger bandwidth is used, $\hat{L}$ has a tendency to overestimate L and suffer some efficiency loss compared to the middle panel where $nh^{2L+1} \to const$. The efficiency loss is more evident in Tables 2 and 3, which suggests that although asymptotic theory indicates that larger L requires larger bandwidths for consistency, too large a bandwidth will cause some overestimation and loss of efficiency.

In summary, the simulation findings confirm that consistency of $\hat{L}$ relies on appropriately selected bandwidth orders. But without prior information concerning L there is no way to be clear that the estimate $\hat{L}$ is consistent. To illustrate: suppose the estimate $\hat{L}$ is around 2.7 with $\gamma = -1/5$. This outcome could match either the middle panel in Table 2 or the left panel in Table 3, leading to uncertainty about the true value. Such uncertainty can be expected to be

Table 3: Performance of $\hat{L}$ at point $z=0$ with $L=3$

	$\gamma=-1/5$			$\gamma=-1/7$			$\gamma=-1/9$		
n	*mean*	*var*	*MSE*	*mean*	*var*	*MSE*	*mean*	*var*	*MSE*
200	2.9180	0.3513	0.3580	4.0793	0.6772	1.8420	4.9752	1.0137	4.9149
400	2.8166	0.2555	0.2891	3.9611	0.4726	1.3962	4.7619	0.6732	3.7775
800	2.7390	0.1952	0.2633	3.8539	0.3619	1.0911	4.5964	0.4908	3.0391
1600	2.6595	0.1496	0.2656	3.7660	0.2815	0.8682	4.4387	0.3648	2.4345
3200	2.6106	0.1253	0.2769	3.6969	0.2327	0.7184	4.3038	0.2785	1.9783

common in applications and hinders practical use of the estimate $\hat{L}$. Further research is needed to determine whether this approach can be improved to be better suited for practical work.

Proofs of the preliminary results (2.4), (2.6) and (2.7):

Proof of (2.4):

$$\begin{aligned}
\mathbb{E}\beta(z_t)G_{tz} &= \int \beta(z_t)G\left(\frac{z_t-z}{h}\right)f(z_t)dz_t \\
&= h\int \beta(z+hu)G(u)f(z+hu)du \\
&= h\int \left[\beta(z)+\beta^{(1)}(z)hu+...\right]G(u)\left[f(z)+f^{(1)}(z)hu+...\right]du \\
&= h^2\left[\beta(z)f^{(1)}(z)+\beta^{(1)}(z)f(z)\right]\int uG(u)du \\
&+ h^4\left[\beta^{(1)}(z)f^{(2)}(z)/2+\beta^{(2)}(z)f^{(1)}(z)/2+\beta^{(3)}(z)f(z)/6+\beta(z)f^{(3)}(z)/6\right]\int u^3G(u)du+... \\
&= \mu_1(G)h^2\left[\beta(z)f^{(1)}(z)+\beta^{(1)}(z)f(z)\right] \\
&+ \mu_3(G)h^4\left[\beta^{(1)}(z)f^{(2)}(z)/2+\beta^{(2)}(z)f^{(1)}(z)/2+\beta^{(3)}(z)f(z)/6+\beta(z)f^{(3)}(z)/6\right]+... \\
&\equiv \mu_1(G)h^2\Delta_1(z)+\mu_3(G)h^4\Delta_3(z)+o(h^4). \qquad (2.27)
\end{aligned}$$

Proof of (2.6):

$$\begin{aligned}
\mathbb{E}G_{tz} &= \int G\left(\frac{z_t-z}{h}\right)f(z_t)dz_t = h\int G(u)f(z+hu)du \\
&= h^2 f^{(1)}(z)\int uG(u)du + h^4\frac{f^{(3)}(z)}{6}\int u^3G(u)du+... \\
&= \mu_1(G)h^2f^{(1)}(z)+\mu_3(G)h^4f^{(3)}(z)/6+o(h^4). \qquad (2.28)
\end{aligned}$$

Proof of (2.7):

$$\mathbb{E}G_{tz}^2 = \int G^2\left(\frac{z_t-z}{h}\right)f(z_t)dz_t = h\int G^2(u)f(z+hu)du$$

$$= hf(z)\int G^2(u)du + ... = \nu_0(G)hf(z) + o(h). \tag{2.29}$$

Difficulty in the estimation of h^L

By Taylor expansion, we have

$$h(L_n^\dagger) - h(L) = h^{(1)}(\tilde{L})(L_n^\dagger - L) = h^{\tilde{L}}\log(h)(L^\dagger - L) = h^L \log\left(\left|\frac{\beta^{(L)}(z)}{L!}\right|\int |s|^L K(s)ds\right) + o_p(h^L) \tag{2.30}$$

since $\tilde{L}$ is on the line segment between $L_n^\dagger$ and L and $L_n^\dagger \to_p L$. Then

$$h(L_n^\dagger) = h(L) \times \left(1 + \log\left(\left|\frac{\beta^{(L)}(z)}{L!}\right|\int |s|^L K(s)ds\right)\right) + o_p(h^L),$$

and $h(L_n^\dagger) = h^{L_n^\dagger} \sim_a d_L h^L$, where $d_L = \left(1 + \log\left(\left|\frac{\beta^{(L)}(z)}{L!}\right|\int |s|^L K(s)ds\right)\right)$, so that $h^{L_n^\dagger}$ is inconsistent. Thus, the slow rate of convergence of $L_n^\dagger$ interferes with the consistent estimation of the factor h^L needed for bias correction.

3 Failure of the adaptive approach

We provide some details to show that the adaptive approach does not work in cases where $L > 1$. More specifically, we will show the adaptive bias estimator $\hat{B}(z)$, which is

$$\hat{B}(z) = A_n(z)^{-1}\left(\sum_{t=1}^n x_t x_t'\right)\left(\frac{1}{n}\sum_{s=1}^n [\hat{\beta}(z_s) - \hat{\beta}(z)]K\left(\frac{z_s - z}{h}\right)\right), \tag{3.1}$$

is not consistent for the true bias $h^{L^*}\mathcal{B}_L(z) = h^{L^*}\frac{\mu_{L^*}(K)}{f(z)}C_L(z)$ when $L > 1$. As before we examine the stationary and nonstationary x_t cases separately.

(i) Stationary x_t In an attempt to show $\hat{B}(z) \sim_a h^{L^*}\frac{\mu_{L^*}(K)}{f(z)}C_L(z) + o(h^{L^*})$, the critical step is to obtain the following asymptotic representation

$$\frac{1}{n}\sum_{t=1}^n [\hat{\beta}(z_t) - \hat{\beta}(z)]K\left(\frac{z_t - z}{h}\right) \sim_a \mathbb{E}\xi_{\beta t} = h^{L^*+1}\mu_{L^*}(K)C_L(z) + o(h^{L^*+1}) \tag{3.2}$$

with $C_L(z) = \frac{f(z)\beta^{(L)}(z)}{L!}\mathbf{1}_{\{L=\text{even}\}} + \left[\frac{\beta^{(L)}(z)}{L!}f^{(1)}(z) + \frac{\beta^{(L+1)}(z)}{(L+1)!}f(z)\right]\mathbf{1}_{\{L=\text{odd}\}}$, and $L^* = L\mathbf{1}_{\{L=\text{even}\}} + (L+1)\mathbf{1}_{\{L=\text{odd}\}}$. However, the asymptotic representation (3.2) does not hold, as we now show. To begin, we write

$$\frac{1}{n}\sum_{t=1}^n \left[\hat{\beta}(z_t) - \hat{\beta}(z)\right]K\left(\frac{z_t - z}{h}\right)$$

$$= \frac{1}{n}\sum_{t=1}^{n}\left\{\beta(z_t)-\beta(z)+\hat\beta(z_t)-\beta(z_t)-[\hat\beta(z)-\beta(z)]\right\}K\left(\frac{z_t-z}{h}\right)$$
$$= \frac{1}{n}\sum_{t=1}^{n}\{\beta(z_t)-\beta(z)\}K\left(\frac{z_t-z}{h}\right)+\frac{1}{n}\sum_{t=1}^{n}\left\{\hat\beta(z_t)-\beta(z_t)-[\hat\beta(z)-\beta(z)]\right\}K\left(\frac{z_t-z}{h}\right)$$
$$=: \Pi_{1n}+\Pi_{2n}. \tag{3.3}$$

By Assumption 1(i) and standard manipulations, $\Pi_{1n} = \frac{1}{n}\sum_{t=1}^{n}\{\beta(z_t)-\beta(z)\}K\left(\frac{z_t-z}{h}\right) \sim_a \mathbb{E}\xi_{\beta t}$.

Now consider Π_{2n}. To simplify the derivations, we consider the case where $\{x_t\}$ and $\{z_t\}$ are independent. Note that

$$\hat\beta(z)-\beta(z) = \left(\sum_{t=1}^{n}x_tx_t'K_{tz}\right)^{-1}\left\{\sum_{t=1}^{n}x_tx_t'[\beta(z_t)-\beta(z)]K_{tz}+\sum_{t=1}^{n}x_tu_tK_{tz}\right\}$$
$$\sim_a (nh)^{-1}\Sigma_{xx}^{-1}f^{-1}(z)\left\{\sum_{s=1}^{n}x_sx_s'[\beta(z_s)-\beta(z)]K_{sz}+\sum_{s=1}^{n}x_su_sK_{sz}\right\}. \tag{3.4}$$

Then

$$\hat\beta(z_t)-\hat\beta(z) = \hat\beta(z_t)-\beta(z_t)-[\hat\beta(z)-\beta(z)]$$
$$\sim_a (nh)^{-1}\Sigma_{xx}^{-1}\left\{f^{-1}(z_t)\left(\sum_{s=1}^{n}x_sx_s'[\beta(z_s)-\beta(z_t)]K_{st}+\sum_{s=1}^{n}x_su_sK_{st}\right)\right.$$
$$\left.-f^{-1}(z)\left(\sum_{s=1}^{n}x_sx_s'[\beta(z_s)-\beta(z)]K_{sz}+\sum_{s=1}^{n}x_su_sK_{sz}\right)\right\}$$
$$= (nh)^{-1}\Sigma_{xx}^{-1}\left\{f^{-1}(z_t)\sum_{s=1}^{n}x_sx_s'[\beta(z_s)-\beta(z_t)]K_{st}-f^{-1}(z)\sum_{s=1}^{n}x_sx_s'[\beta(z_s)-\beta(z)]K_{sz}\right\}$$
$$+(nh)^{-1}\Sigma_{xx}^{-1}\left\{f^{-1}(z_t)\sum_{s=1}^{n}x_su_sK_{st}-f^{-1}(z)\sum_{s=1}^{n}x_su_sK_{sz}\right\}. \tag{3.5}$$

Hence

$$\Pi_{2n} = \frac{1}{n}\sum_{t=1}^{n}\left\{\hat\beta(z_t)-\beta(z_t)-[\hat\beta(z)-\beta(z)]\right\}K\left(\frac{z_t-z}{h}\right)$$
$$\sim_a (nh)^{-1}\Sigma_{xx}^{-1}\frac{1}{n}\sum_{t=1}^{n}\left\{f^{-1}(z_t)\sum_{s=1}^{n}x_sx_s'[\beta(z_s)-\beta(z_t)]K_{st}-f^{-1}(z)\sum_{s=1}^{n}x_sx_s'[\beta(z_s)-\beta(z)]K_{sz}\right\}K_{tz}$$
$$+(nh)^{-1}\Sigma_{xx}^{-1}\frac{1}{n}\sum_{t=1}^{n}\left\{f^{-1}(z_t)\sum_{s=1}^{n}x_su_sK_{st}-f^{-1}(z)\sum_{s=1}^{n}x_su_sK_{sz}\right\}K_{tz}$$
$$= \Sigma_{xx}^{-1}\frac{1}{nh}\sum_{t=1}^{n}f^{-1}(z_t)K_{tz}\frac{1}{n}\sum_{s=1}^{n}x_sx_s'[\beta(z_s)-\beta(z_t)]K_{st}-\Sigma_{xx}^{-1}f^{-1}(z)\frac{1}{nh}\sum_{t}K_{tz}\frac{1}{n}\sum_{s}x_sx_s'[\beta(z_s)-\beta(z)]K_{sz}$$

$$+ h\Sigma_{xx}^{-1}\frac{1}{nh}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\frac{1}{nh}\sum_{s=1}^{n} x_s u_s K_{st} - h\Sigma_{xx}^{-1} f^{-1}(z)\frac{1}{nh}\sum_{t=1}^{n} K_{tz}\frac{1}{nh}\sum_{s=1}^{n} x_s u_s K_{sz}$$

$$=: \Pi_{2n1} - \Pi_{2n2} + \Pi_{2n3} - \Pi_{2n4}. \tag{3.6}$$

Consider the terms Π_{2ni}, $i = 1, 2, 3, 4$ in turn. Start with Π_{2n1} and to simplify derivations in what follows assume that the z_t are independent and identically distributed. The weakly dependent case involves much longer calculations but can be handled under Assumption 1 along lines such as those used in the proof of Lemma B.2. Following a similar line of argument as that leading to (B.5) we have, using conditional expectations $\mathbb{E}_{z\_t}$ given z_t,

$$\Pi_{2n1} = \Sigma_{xx}^{-1}\frac{1}{nh}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\frac{1}{n}\sum_{s=1}^{n} x_s x_s'[\beta(z_s) - \beta(z_t)]K_{st}$$

$$= \Sigma_{xx}^{-1}\frac{1}{nh}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\frac{1}{n}\sum_{s\neq t}^{n} x_s x_s'[\beta(z_s) - \beta(z_t)]K_{st}$$

$$\sim_a \Sigma_{xx}^{-1}\frac{1}{nh}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\Sigma_{xx}\mathbb{E}_{z\_t}[\beta(z_s) - \beta(z_t)]K_{st}$$

$$\sim_a \frac{1}{nh}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\mathbb{E}_{z\_t}[\beta(z_s) - \beta(z_t)]K_{st}$$

$$\sim_a \frac{1}{nh}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\int_{-1}^{1}[\beta(z_s) - \beta(z_t)]K\left(\frac{z_s - z_t}{h}\right) f(z_s)dz_s$$

$$\sim_a \frac{1}{n}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\int_{-1}^{1}[\beta(z_t + ph) - \beta(z_t)]K(p)f(z_t + ph)dp$$

$$\sim_a \frac{1}{n}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\int_{-1}^{1}\sum_{j=1}^{L+1}\frac{1}{j!}\beta^{(j)}(z_t)(ph)^j K(p)[f(z_t) + f^{(1)}(z_t)ph]dp$$

$$\sim_a \frac{1}{n}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\sum_{j=1}^{L+1}\frac{1}{j!}\beta^{(j)}(z_t)h^j[\int p^j K(p)dpf(z_t) + \int p^{j+1}K(p)dpf^{(1)}(z_t)h]$$

$$\sim_a \frac{1}{n}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\sum_{j=1}^{L+1}\frac{1}{j!}\beta^{(j)}(z_t)h^j[\mu_j(K)f(z_t) + \mu_{j+1}(K)f^{(1)}(z_t)h]$$

$$\sim_a \int K_{tz}\sum_{j=1}^{L+1}\frac{1}{j!}\beta^{(j)}(z_t)h^j[\mu_j(K)f(z_t) + \mu_{j+1}(K)f^{(1)}(z_t)h]dz_t$$

$$\sim_a \sum_{j=1}^{L+1}\frac{1}{j!}h^j\mu_j(K)\int \beta^{(j)}(z_t)K_{tz}f(z_t)dz_t + \sum_{j=1}^{L+1}\frac{1}{j!}h^{j+1}\mu_{j+1}(K)\int \beta^{(j)}(z_t)K_{tz}f^{(1)}(z_t)dz_t$$

$$\sim_a \sum_{j=1}^{L+1}\frac{1}{j!}h^{j+1}\mu_j(K)\int \beta^{(j)}(z + hu)K(u)f(z + hu)du + \sum_{j=1}^{L+1}\frac{1}{j!}h^{j+2}\mu_{j+1}(K)\int \beta^{(j)}(z + hu)K(u)f^{(1)}(z + hu)du$$

$$
\begin{aligned}
&\sim_a \sum_{j=1}^{L+1} \frac{1}{j!} h^{j+1} \mu_j(K) \int \sum_{k=0}^{L-j+1} \frac{1}{k!} \beta^{(j+k)}(z)(hu)^k K(u) f(z+hu) du \\
&+ \sum_{j=1}^{L+1} \frac{1}{j!} h^{j+2} \mu_{j+1}(K) \int \sum_{k=0}^{L+1-j} \frac{1}{k!} \beta^{(j+k)}(z)(hu)^k K(u) f^{(1)}(z+hu) du \\
&\sim_a \sum_{j=1}^{L} \frac{1}{j!} h^{j+1} \mu_j(K) \frac{1}{(L-j)!} \beta^{(L)}(z) \int (hu)^{L-j} K(u) f(z+hu) du \\
&+ \sum_{j=1}^{L+1} \frac{1}{j!} h^{j+1} \mu_j(K) \frac{1}{(L+1-j)!} \beta^{(L+1)}(z) \int (hu)^{L+1-j} K(u) f(z+hu) du \\
&+ \sum_{j=1}^{L} \frac{1}{j!} h^{j+2} \mu_{j+1}(K) \frac{1}{(L-j)!} \beta^{(L)}(z) \int (hu)^{L-j} K(u) f^{(1)}(z+hu) du \\
&\sim_a h^{L+1} \sum_{j=1}^{L} \frac{1}{j!} \mu_j(K) \frac{1}{(L-j)!} \beta^{(L)}(z) \mu_{L-j}(K) f(z) + h^{L+2} \sum_{j=1}^{L} \frac{1}{j!} \mu_j(K) \frac{1}{(L-j)!} \beta^{(L)}(z) \mu_{L+1-j}(K) f^{(1)}(z) \\
&+ h^{L+2} \sum_{j=1}^{L+1} \frac{1}{j!} \mu_j(K) \frac{1}{(L+1-j)!} \beta^{(L+1)}(z) \mu_{L+1-j}(K) f(z) + h^{L+2} \sum_{j=1}^{L} \frac{1}{j!} \mu_{j+1}(K) \frac{1}{(L-j)!} \beta^{(L)}(z) \mu_{L-j}(K) f^{(1)}(z) \\
&\sim_a h^{L+1} \left\{ f(z) \beta^{(L)}(z) \sum_{j=1}^{L} \frac{1}{j!} \mu_j(K) \frac{1}{(L-j)!} \mu_{L-j}(K) \right\} \times 1_{\{L=even\}} \\
&+ h^{L+2} \left\{ \beta^{(L)}(z) f^{(1)}(z) \sum_{j=1}^{L} \frac{1}{j!(L-j)!} [\mu_j(K)\mu_{L+1-j}(K) + \mu_{j+1}(K)\mu_{L-j}(K)] \right. \\
&\left. + \beta^{(L+1)}(z) f(z) \sum_{j=1}^{L+1} \frac{1}{j!(L+1-j)!} \mu_j(K) \mu_{L+1-j}(K) \right\} \times 1_{\{L=odd\}} \\
&=: h^{L^*+1} D_L(z).
\end{aligned}
$$

The second to last equation is due to the fact that when $L = even$, $\mu_j(K)\mu_{L+1-j}(K) = 0$ and when $L = odd$, $\mu_j(K)\mu_{L-j}(K) = 0$.

For the term Π_{2n2} we have

$$
\begin{aligned}
\Pi_{2n2} &= \Sigma_{xx}^{-1} f^{-1}(z) \frac{1}{nh} \sum_t K_{tz} \frac{1}{n} \sum_s x_s x_s' [\beta(z_s) - \beta(z)] K_{sz} \\
&\sim_a \Sigma_{xx}^{-1} f^{-1}(z) f(z) \mathbb{E}(x_s x_s') \mathbb{E}[\beta(z_s) - \beta(z)] K_{sz} \sim_a \mathbb{E}\xi_{\beta t}. \qquad (3.7)
\end{aligned}
$$

So the leading terms of Π_{1n} and Π_{2n2} are both asymptotically equivalent to $\mathbb{E}\xi_{\beta t}$. Below we will see $\Pi_{2n3} - \Pi_{2n4}$ has zero mean and does not contribute to the bias centering expression. It follows that the leading term of $\Pi_{1n} + \Pi_{2n}$ is determined by that of Π_{2n1}, which is very complicated and involves the unknown value of L. Hence, (3.2) is unlikely to hold. As a result,

$\hat{B}(z)$ is not a consistent estimate of the bias $h^{L^*}\mathcal{B}_L(z)$; and, due to the dependence of the limit of Π_{2n1} on unknown L, it is impossible to scale adjust $\hat{B}(z)$ to achieve a consistent bias estimate.

Next, consider the remaining two terms of (3.6) involving the difference

$$\Pi_{2n3}-\Pi_{2n4}=h\Sigma_{xx}^{-1}\frac{1}{nh}\sum_{t=1}^{n}f^{-1}(z_t)K_{tz}\frac{1}{nh}\sum_{s=1}^{n}x_su_sK_{st}-h\Sigma_{xx}^{-1}f^{-1}(z)\frac{1}{nh}\sum_{t=1}^{n}K_{tz}\frac{1}{nh}\sum_{s=1}^{n}x_su_sK_{sz}, \tag{3.8}$$

which has zero mean and therefore does not contribute to the bias centering expression. But, as in Phillips and Wang (2021), we need to analyze this term's contribution to the variance of the limit distribution of the statistic. For ease of presentation and calculation, set $g=f^{-1}$, take the scalar x_t case and let $u_t\sim_{iid}\left(0,\sigma_u^2\right)$ and z_t be iid, as above. Write (3.8) as

$$\Pi_{2n3}-\Pi_{2n4}=\frac{1}{n}\Sigma_{xx}^{-1}\frac{1}{\sqrt{nh}}\sum_{s=1}^{n}x_su_s\left(\frac{1}{\sqrt{nh}}\sum_{t=1}^{n}K_{tz}\left[K_{st}g\left(z_t\right)-K_{sz}g\left(z\right)\right]\right), \tag{3.9}$$

and since x_su_s is a martingale difference, this expression has mean zero and variance

$$\begin{aligned}
&\frac{\sigma_u^2\Sigma_{xx}^{-1}}{n^2h}\mathbb{E}\left(\frac{1}{\sqrt{nh}}\sum_{t=1}^{n}K_{tz}\left[K_{st}g\left(z_t\right)-K_{sz}g\left(z\right)\right]\right)^2\\
&=\frac{\sigma_u^2\Sigma_{xx}^{-1}}{n^3h^2}\mathbb{E}\left(\sum_{t,r=1}^{n}K_{tz}K_{rz}\left[K_{st}g\left(z_t\right)-K_{sz}g\left(z\right)\right]\left[K_{sr}g\left(z_r\right)-K_{sz}g\left(z\right)\right]\right)\\
&=\frac{\sigma_u^2\Sigma_{xx}^{-1}}{n^2h^2}\mathbb{E}\left(K_{tz}^2\left[K_{st}g\left(z_t\right)-K_{sz}g\left(z\right)\right]^2\right)+\frac{\sigma_u^2\Sigma_{xx}^{-1}}{n^3h^2}\sum_{t\neq r}^{n}\mathbb{E}\left(K_{tz}K_{rz}\left[K_{st}g\left(z_t\right)-K_{sz}g\left(z\right)\right]\left[K_{sr}g\left(z_r\right)-K_{sz}g\left(z\right)\right]\right)\\
&=:D_1+D_2.
\end{aligned} \tag{3.10}$$

Noting that the particular case where $s=t$ in the dual summation (3.9) is of smaller order and can therefore be neglected in (3.9), we have

$$\begin{aligned}
D_1&=\frac{\sigma_u^2\Sigma_{xx}^{-1}}{n^2h^2}\mathbb{E}\left(K_{tz}^2\left[K_{st}g\left(z_t\right)-K_{sz}g\left(z\right)\right]^2\right)\\
&=\frac{\sigma_u^2\Sigma_{xx}^{-1}}{n^2h^2}\int_{-1}^{1}\int_{-1}^{1}K\left(\frac{z_t-z}{h}\right)^2\left[K\left(\frac{z_s-z_t}{h}\right)g\left(z_t\right)-K\left(\frac{z_s-z}{h}\right)g\left(z\right)\right]^2f\left(z_s\right)f\left(z_t\right)dz_sdz_t\\
&=\frac{\sigma_u^2\Sigma_{xx}^{-1}}{n^2}\int_{-1}^{1}\int_{-1}^{1}K\left(p\right)^2\left[K\left(p-q\right)g\left(z+ph\right)-K\left(q\right)g\left(z\right)\right]^2f\left(z+qh\right)f\left(z+ph\right)dpdq\\
&\sim_a\frac{\sigma_u^2\Sigma_{xx}^{-1}g\left(z\right)^2f\left(z\right)^2}{n^2}\int_{-1}^{1}\int_{-1}^{1}K\left(p\right)^2\left[K\left(p-q\right)-K\left(q\right)\right]^2dpdq\\
&=\frac{\sigma_u^2\Sigma_{xx}^{-1}}{n^2}\int_{-1}^{1}\int_{-1}^{1}K\left(p\right)^2\left[K\left(p-q\right)-K\left(q\right)\right]^2dpdq.
\end{aligned} \tag{3.11}$$

Next, again noting that the case where $s = t$ is of smaller order and can be neglected, we have

$$\begin{aligned}
D_2 &= \frac{\sigma_u^2 \Sigma_{xx}^{-1}}{n^3 h^2} \sum_{t \neq r}^{n} \mathbb{E}\left(K_{tz} K_{rz} \left[K_{st} g\left(z_t\right) - K_{sz} g\left(z\right)\right]\left[K_{sr} g\left(z_r\right) - K_{sz} g\left(z\right)\right]\right) \\
&\sim_a \frac{\sigma_u^2 \Sigma_{xx}^{-1}}{n h^2} \mathbb{E}\left(K_{tz} K_{rz} \left[K_{st} g\left(z_t\right) - K_{sz} g\left(z\right)\right]\left[K_{sr} g\left(z_r\right) - K_{sz} g\left(z\right)\right]\right) \\
&\sim_a \frac{\sigma_u^2 \Sigma_{xx}^{-1}}{n h^2} \int\int\int K\left(\frac{z_t - z}{h}\right) K\left(\frac{z_r - z}{h}\right)\left[K\left(\frac{z_s - z_t}{h}\right) g\left(z_t\right) - K\left(\frac{z_s - z}{h}\right) g\left(z\right)\right] \\
&\times \left[K\left(\frac{z_s - z_r}{h}\right) g\left(z_r\right) - K\left(\frac{z_s - z}{h}\right) g\left(z\right)\right] f\left(z_s\right) f\left(z_t\right) f(z_r) dz_s dz_t dz_r \\
&= h \frac{\sigma_u^2 \Sigma_{xx}^{-1}}{n} \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} K\left(p\right) K(w)\left[K\left(p - q\right) g\left(z + ph\right) - K\left(q\right) g\left(z\right)\right] \\
&\times \left[K(q - w) g(z + hw) - K(q) g(z)\right] f\left(z + qh\right) f\left(z + ph\right) f(z + wh) dp dq dw \\
&\sim_a h \frac{\sigma_u^2 \Sigma_{xx}^{-1}}{n} f(z) \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} K\left(p\right) K(w)\left[K\left(p - q\right) - K\left(q\right)\right]\left[K(q - w) - K(q)\right] dp dq dw.
\end{aligned} \tag{3.12}$$

It follows that $\Pi_{2n3} - \Pi_{2n4}$ has zero mean and asymptotic variance $D_1 + D_2 = O(h/n)$ in view of (3.10) - (3.12), so that $\Pi_{2n3} - \Pi_{2n4} = O_p(\sqrt{h/n})$. As a result, $A_n(z)^{-1}(\sum_t x_t x_t')(\Pi_{2n3} - \Pi_{2n4}) = O_p(\frac{1}{nh} n \sqrt{\frac{h}{n}}) = O_p(1/\sqrt{nh})$, which is of the same order with the asymptotic variance term in the stationary case since the convergence rate is $\sqrt{nh}$. Consequently,

$$\begin{aligned}
\hat{B}(z) &= A_n(z)^{-1}\left(\sum_t x_t x_t'\right)\left[\Pi_{1n} + \Pi_{2n1} - \Pi_{2n2} + \Pi_{2n3} - \Pi_{2n4}\right] \\
&\sim_a \frac{1}{h f(z)}\left(h^{L^*+1} D_L(z) + o_p(h^{L^*+1})\right) + A_n(z)^{-1}\left(\sum_t x_t x_t'\right)\left(\Pi_{2n3} - \Pi_{2n4}\right) \\
&\sim_a h^{L^*} f^{-1}(z) D_L(z) + o_p(h^{L^*}) + O_p\left(1/\sqrt{nh}\right).
\end{aligned} \tag{3.13}$$

This analysis reveals that $\hat{B}(z)$ is not consistent for the true bias $h^{L^*}\mathcal{B}_L(z)$. And also, it retains a random element that is $O_p(1)$ after standardization by the convergence rate $\sqrt{nh}$ in the stationary case. So the bias term adjustment $\hat{B}(z)$ affects the limit distribution of the bias corrected estimation error $\hat{\beta}(z) - \beta(z) - \hat{B}(z)$. In effect, the adaptive bias adjustment $\hat{B}(z)$ introduces an estimation error through the presence of the element $\hat{\beta}(z_t) - \hat{\beta}(z)$. This estimation error contributes a random term $A_n(z)^{-1}(\sum_t x_t x_t')(\Pi_{2n3} - \Pi_{2n4})$ to the limit distribution that has the same order $O_p(\frac{1}{\sqrt{nh}})$ as the asymptotic variance term (the last term in (A.8) in the main paper).

In view of these difficulites, we do not consider the adaptive variance estimators $\hat{\Omega}_n(z)$ (in the stationary case) and $\hat{\Omega}_n^*(z)$ (in the nonstationary case). In short, without an adaptive consistent bias estimator $\hat{B}(z)$ to adjust estimation error $\hat{\beta}(z) - \beta(z)$ even a correctly adjusted variance matrix estimator would be unable to produce an asymptotically valid test statistic.

(ii) Nonstationary x_t The adaptive approach to constructing a test statistic does not work in this case either. First, as shown below, the adaptive bias estimator is not consistent for the same reason as in the stationary case. Moreover, the adaptive bias estimator introduces additional variation in the limit distribution and this additional variance takes a more complicated form in the nonstationary case than in the stationary case and depends on the unknown flatness parameter L and the derivative $\beta^{(L)}(z)$. Attempts to estimate these variances in an adaptive way introduces further bias terms in the adaptive variance estimator. These interactions among bias and variance effects make it difficult to formulate a successful adaptive approach to inference that is not reliant on knowledge or consistent estimation of the parameter L in the nonstationary case. Thus, in spite of the apparent simplicity of the adaptive statistic and its formulation independent of L, the high technical complexity of its asymptotics and interactive bias and variance effects are serious challenges that are left for subsequent research on adaptation in the presence of flat function behavior.

We first show the bias estimator $\hat{B}(z)$ is not consistent. The analysis follows lines similar to those used above while allowing for limiting moment behavior of the nonstationary x_t. We have the same decomposition as in (3.3), viz.,

$$\frac{1}{n}\sum_{t=1}^{n}[\hat{\beta}(z_t)-\hat{\beta}(z)]K\left(\frac{z_t-z}{h}\right)=\Pi_{1n}^{*}+\Pi_{2n}^{*}. \tag{3.14}$$

As before, we have $\Pi_{1n}^{*}=\frac{1}{n}\sum_{t=1}^{n}\left\{\beta(z_t)-\beta(z)\right\}K\left(\frac{z_t-z}{h}\right)\sim_a \mathbb{E}\xi_{\beta t}$. Proceeding in the nonstationary case, we have

$$\begin{aligned}
\Pi_{2n}^{*} &= \frac{1}{n}\sum_{t=1}^{n}\left\{\hat{\beta}(z_t)-\beta(z_t)-[\hat{\beta}(z)-\beta(z)]\right\}K\left(\frac{z_t-z}{h}\right) \\
&\sim_a \frac{1}{n^2h}\left(\int B_xB_x'\right)^{-1}\frac{1}{n}\sum_{t=1}^{n}\left\{f^{-1}(z_t)\sum_{s=1}^{n}x_sx_s'[\beta(z_s)-\beta(z_t)]K_{st}-f^{-1}(z)\sum_{s=1}^{n}x_sx_s'[\beta(z_s)-\beta(z)]K_{sz}\right\}K_{tz} \\
&+\frac{1}{n^2h}\left(\int B_xB_x'\right)^{-1}\frac{1}{n}\sum_{t=1}^{n}\left\{f^{-1}(z_t)\sum_{s=1}^{n}x_su_sK_{st}-f^{-1}(z)\sum_{s=1}^{n}x_su_sK_{sz}\right\}K_{tz} \\
&=\left(\int B_xB_x'\right)^{-1}\frac{1}{n^2h}\sum_{t=1}^{n}f^{-1}(z_t)K_{tz}\frac{1}{n}\sum_{s=1}^{n}x_sx_s'[\beta(z_s)-\beta(z_t)]K_{st} \\
&-\left(\int B_xB_x'\right)^{-1}f^{-1}(z)\frac{1}{n^2h}\sum_{t}K_{tz}\frac{1}{n}\sum_{s}x_sx_s'[\beta(z_s)-\beta(z)]K_{sz} \\
&+h\left(\int B_xB_x'\right)^{-1}\frac{1}{n^2h}\sum_{t=1}^{n}f^{-1}(z_t)K_{tz}\frac{1}{nh}\sum_{s=1}^{n}x_su_sK_{st} \\
&-h\left(\int B_xB_x'\right)^{-1}f^{-1}(z)\frac{1}{n^2h}\sum_{t=1}^{n}K_{tz}\frac{1}{nh}\sum_{s=1}^{n}x_su_sK_{sz}
\end{aligned}$$

$$=: \Pi^*_{2n1} - \Pi^*_{2n2} + \Pi^*_{2n3} - \Pi^*_{2n4}. \tag{3.15}$$

In analyzing these terms we employ similar arguments to those in the stationary case. For the first term, using conditional expectations $\mathbb{E}_{z\_t}$ given z_t, we have

$$\begin{aligned}
\Pi^*_{2n1} &= \left(\int B_x B_x'\right)^{-1} \frac{1}{n}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\frac{1}{n^2h}\sum_{s\neq t}^{n} x_s x_s'[\beta(z_s)-\beta(z_t)]K_{st} \\
&= \left(\int B_x B_x'\right)^{-1} \frac{1}{n}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\frac{1}{n^2h}\sum_{s\neq t}^{n} x_s x_s'\mathbb{E}_{z_t}\xi_{\beta st} \\
&+ \left(\int B_x B_x'\right)^{-1} \frac{1}{n}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\frac{1}{n^2h}\sum_{s\neq t}^{n} x_s x_s'[\xi_{\beta st}-\mathbb{E}_{z_t}\xi_{\beta st}] \\
&=: \Pi^{*a}_{2n1} + \Pi^{*b}_{2n1}
\end{aligned} \tag{3.16}$$

where $\xi_{\beta st} := [\beta(z_s)-\beta(z_t)]K_{st}$. For the first term Π^{*a}_{2n1} we have $\Pi^{*a}_{2n1} \sim_a h^{L^*+1}D_L(z)$ following the same lines as that of Π_{2n1} in the stationary case. For the second term Π^{*b}_{2n1}, first as in Lemma B.2 we can verify that $\mathbb{E}_{z\_t}\xi_{\beta st} = O(h^{L^*+1})$ and $\mathbb{E}_{z\_t}\xi^2_{\beta st} = O(h^{2L+1})$. As a result, $\sum_{s\neq t}^{n} x_s x_s'[\xi_{\beta st}-\mathbb{E}_{z\_t}\xi_{\beta st}] = O_p(n\sqrt{nh^{2L+1}}) \sim_a n\sqrt{nh^{2L+1}}\int B_x B_x' dB_{\xi,z_t}$ where B_{ξ,z_t} is dependent on z_t. Then we have

$$\begin{aligned}
\Pi^{*b}_{2n1} &= \left(\int B_x B_x'\right)^{-1} \frac{1}{n}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\frac{1}{n^2h}\sum_{s\neq t}^{n} x_s x_s'[\xi_{\beta st}-\mathbb{E}_{z_t}\xi_{\beta st}] \\
&\sim_a \left(\int B_x B_x'\right)^{-1} h\frac{1}{nh}\sum_{t=1}^{n} f^{-1}(z_t)K_{tz}\frac{1}{n^2h}n\sqrt{nh^{2L+1}}\int B_x B_x' dB_{\xi,z_t} \\
&= h\frac{1}{n^2h}n\sqrt{nh^{2L+1}}\left(\int B_x B_x'\right)^{-1} f^{-1}(z)f(z)\int B_x B_x' dB_{\xi,z} \\
&= \sqrt{\frac{h^{2L+1}}{n}}\left(\int B_x B_x'\right)^{-1}\int B_x B_x' dB_{\xi,z} \qquad (3.17)\\
&= O_p\left(\sqrt{\frac{h^{2L+1}}{n}}\right), \qquad (3.18)
\end{aligned}$$

where $B_{\xi,z}$ means localize z_t at z in B_{ξ,z_t}. Note that $A_n(z)^{-1}(\sum_t x_t x_t')\Pi^{*b}_{2n1} = O_p\left(\frac{1}{n^2h}n^2\sqrt{\frac{h^{2L+1}}{n}}\right) = O_p\left(\sqrt{\frac{h^{2L-1}}{n}}\right)$, which has the same order as the first term in (A.14) in the main paper. Therefore it may contribute to the limit distribution and should be retained in the bias corrected expression $\hat{\beta}(z) - \beta(z) - \hat{B}(z)$.

For Π_{2n2}^* we have

$$\begin{aligned}
\Pi_{2n2}^* &= \left(\int B_x B_x'\right)^{-1} f^{-1}(z) \frac{1}{nh} \sum_{t=1}^{n} K_{tz} \frac{1}{n^2} \sum_s x_s x_s' [\beta(z_s) - \beta(z)] K_{sz} \\
&= \left(\int B_x B_x'\right)^{-1} f^{-1}(z) \frac{1}{nh} \sum_{t=1}^{n} K_{tz} \frac{1}{n^2} \sum_s x_s x_s' \mathbb{E}\xi_{\beta s} \\
&+ \left(\int B_x B_x'\right)^{-1} f^{-1}(z) \frac{1}{nh} \sum_{t=1}^{n} K_{tz} \frac{1}{n^2} \sum_s x_s x_s' (\xi_{\beta s} - \mathbb{E}\xi_{\beta s}) \\
&=: \Pi_{2n2}^{*a} + \Pi_{2n2}^{*b},
\end{aligned}$$

where $\xi_{\beta s} := [\beta(z_s) - \beta(z)] K_{sz}$. For the first term Π_{2n2}^{*a} we have

$$\begin{aligned}
\Pi_{2n2}^{*a} &= \left(\int B_x B_x'\right)^{-1} f^{-1}(z) \frac{1}{nh} \sum_{t=1}^{n} K_{tz} \frac{1}{n^2} \sum_s x_s x_s' \mathbb{E}\xi_{\beta s} \\
&\sim_a f^{-1}(z) f(z) \mathbb{E}[\beta(z_s) - \beta(z)] K_{sz} \sim_a \mathbb{E}\xi_{\beta t}. \qquad (3.19)
\end{aligned}$$

For the second term Π_{2n2}^{*b}, first following Lemma B.2 $\mathbb{E}\xi_{\beta s} = O(h^{L^*+1})$ and $\mathbb{E}\xi_{\beta s}^2 = O(h^{2L+1})$. Therefore $\sum_s x_s x_s' (\xi_{\beta s} - \mathbb{E}\xi_{\beta s}) = O_p(n\sqrt{nh^{2L+1}}) \sim_a n\sqrt{nh^{2L+1}} \int B_x B_x' dB_\xi$. Then we have

$$\begin{aligned}
\Pi_{2n2}^{*b} &= \left(\int B_x B_x'\right)^{-1} f^{-1}(z) \frac{1}{nh} \sum_{t=1}^{n} K_{tz} \frac{1}{n^2} \sum_s x_s x_s' (\xi_{\beta s} - \mathbb{E}\xi_{\beta s}) \\
&\sim_a \left(\int B_x B_x'\right)^{-1} f^{-1}(z) f(z) \frac{1}{n^2} n\sqrt{nh^{2L+1}} \int B_x B_x' dB_\xi \\
&= \sqrt{\frac{h^{2L+1}}{n}} \left(\int B_x B_x'\right)^{-1} \int B_x B_x' dB_\xi \qquad (3.20) \\
&= O_p\left(\sqrt{\frac{h^{2L+1}}{n}}\right). \qquad (3.21)
\end{aligned}$$

Note that the leading variation term of Π_{2n1}^* and Π_{2n2}^*, namely Π_{2n1}^{*b} and Π_{2n2}^{*b}, cannot be cancelled although they share the same order. So they both should be retained in the bias corrected expression $\hat{\beta}(z) - \beta(z) - \hat{B}(z)$.

The remaining two terms of (3.15) involve the difference

$$\begin{aligned}
\Pi_{2n3}^* - \Pi_{2n4}^* &= h \left(\int B_x B_x'\right)^{-1} \frac{1}{n^2 h} \sum_{t=1}^{n} f^{-1}(z_t) K_{tz} \frac{1}{nh} \sum_{s=1}^{n} x_s u_s K_{st} \\
&- h \left(\int B_x B_x'\right)^{-1} f^{-1}(z) \frac{1}{n^2 h} \sum_{t=1}^{n} K_{tz} \frac{1}{nh} \sum_{s=1}^{n} x_s u_s K_{sz}, \qquad (3.22)
\end{aligned}$$

which has zero mean and does not contribute to bias centering. As in the stationary x_t case, we need to examine this term's contribution to the variance and limit distribution. For ease of presentation and calculation, we again set $g = f^{-1}$, take the scalar x_t case and let $u_t \sim_{iid} \left(0, \sigma_u^2\right)$ and z_t be *iid*. Write (3.22) as

$$\Pi_{2n3}^* - \Pi_{2n4}^* = \frac{1}{n^2}\left(\int B_x^2\right)^{-1} \frac{1}{\sqrt{nh}} \sum_{s=1}^{n} x_s u_s \left(\frac{1}{\sqrt{nh}} \sum_{t=1}^{n} K_{tz}\left[K_{st} g\left(z_t\right) - K_{sz} g\left(z\right)\right]\right), \tag{3.23}$$

and since $x_s u_s$ is a martingale difference, this expression has mean zero and conditional variance given $\mathcal{F}_x$

$$\begin{aligned}
&\frac{\sigma_u^2}{n^4}\left(\int B_x^2\right)^{-2} \frac{1}{nh} \sum_{s=1}^{n} x_s^2\, \mathbb{E}\left(\frac{1}{\sqrt{nh}} \sum_{t=1}^{n} K_{tz}\left[K_{st} g\left(z_t\right) - K_{sz} g\left(z\right)\right]\right)^2 \\
&\sim_a \frac{\sigma_u^2\left(\int B_x^2\right)^{-2}}{n^3 h}\left(\int B_x^2\right) \mathbb{E}\left(\frac{1}{\sqrt{nh}} \sum_{t=1}^{n} K_{tz}\left[K_{st} g\left(z_t\right) - K_{sz} g\left(z\right)\right]\right)^2 \\
&= \frac{\sigma_u^2\left(\int B_x^2\right)^{-1}}{n^4 h^2} \mathbb{E}\left(\sum_{t,r=1}^{n} K_{tz} K_{rz}\left[K_{st} g\left(z_t\right) - K_{sz} g\left(z\right)\right]\left[K_{sr} g\left(z_t\right) - K_{sz} g\left(z\right)\right]\right) \\
&\sim_a \frac{\sigma_u^2\left(\int B_x^2\right)^{-1}}{n^3 h^2} \mathbb{E}\left(K_{tz}^2\left[K_{st} g\left(z_t\right) - K_{sz} g\left(z\right)\right]^2\right) \\
&+ \frac{\sigma_u^2\left(\int B_x^2\right)^{-1}}{n^2 h^2} \mathbb{E}\left\{K_{tz} K_{rz}\left[K_{st} g\left(z_t\right) - K_{sz} g\left(z\right)\right]\left[K_{sr} g\left(z_t\right) - K_{sz} g\left(z\right)\right]\right\} \\
&=: D_1^* + D_2^*.
\end{aligned} \tag{3.24}$$

Proceeding in the same way as the argument leading to (3.11) and (3.12), we find that

$$D_1^* = \frac{\sigma_u^2\left(\int B_x^2\right)^{-1}}{n^3} \int_{-1}^{1} \int_{-1}^{1} K\left(p\right)^2\left[K\left(p-q\right) - K\left(q\right)\right]^2 dp dq, \tag{3.25}$$

and

$$\begin{aligned}
D_2^* &= \frac{\sigma_u^2\left(\int B_x^2\right)^{-1}}{n^2 h^2} \mathbb{E}\left(K_{tz} K_{rz}\left[K_{st} g\left(z_t\right) - K_{sz} g\left(z\right)\right]\left[K_{sr} g\left(z_r\right) - K_{sz} g\left(z\right)\right]\right) \\
&\sim_a \frac{\sigma_u^2\left(\int B_x^2\right)^{-1}}{n^2 h^2} \int\int\int K\left(\frac{z_t - z}{h}\right) K\left(\frac{z_r - z}{h}\right)\left[K\left(\frac{z_s - z_t}{h}\right) g\left(z_t\right) - K\left(\frac{z_s - z}{h}\right) g\left(z\right)\right] \\
&\times\left[K\left(\frac{z_s - z_r}{h}\right) g\left(z_r\right) - K\left(\frac{z_s - z}{h}\right) g\left(z\right)\right] f\left(z_s\right) f\left(z_t\right) f(z_r) dz_s dz_t dz_r \\
&= h\frac{\sigma_u^2\left(\int B_x^2\right)^{-1}}{n^2} \int_{-1}^{1} \int_{-1}^{1} \int_{-1}^{1} K\left(p\right) K(w)\left[K\left(p-q\right) g\left(z+ph\right) - K\left(q\right) g\left(z\right)\right] \\
&\times\left[K(q-w) g(z+hw) - K(q) g(z)\right] f\left(z+qh\right) f\left(z+ph\right) f(z+wh) dp dq dw
\end{aligned}$$

$$\sim_a h\frac{\sigma_u^2\left(\int B_x^2\right)^{-1}}{n^2}f(z)\int_{-1}^1\int_{-1}^1\int_{-1}^1 K(p)\,K(w)\,[K(p-q)-K(q)]\,[K(q-w)-K(q)]dpdqdw. \tag{3.26}$$

It follows that $\Pi^*_{2n3}-\Pi^*_{2n4}$ has zero mean and conditional asymptotic variance $D_1^*+D_2^*=O_p(\frac{h}{n^2})$ in view of (3.24) - (3.26), so that $\Pi^*_{2n3}-\Pi^*_{2n4}=O_p(\frac{h^{1/2}}{n})$. The implication for the bias is that $A_n(z)^{-1}(\sum_t x_tx_t')(\Pi_{2n3}-\Pi_{2n4})=O_p(\frac{1}{n^2h}n^2\sqrt{\frac{h}{n^2}})=O_p(1/\sqrt{n^2h})$, which is of the same order as the asymptotic variance term in the nonstationary case. It therefore affects the limit distribution and should be retained in the bias corrected formula $\hat\beta(z)-\beta(z)-\hat B(z)$.

Letting $A^*_{n,h}:=\frac{1}{n^2h}\sum_{t=1}^n x_tx_t'K_{tz}$ we have the bias corrected estimation error

$$\begin{aligned}
&\hat\beta(z)-\beta(z)-\hat B(z)=A_n^{-1}(z)\sum_{t=1}^n x_tx_t'[\beta(z_t)-\beta(z)]K_{tz}+A_n^{-1}(z)\sum_{t=1}^n x_tu_tK_{tz}-\hat B(z)\\
&=A_n^{-1}(z)\sum_{t=1}^n x_tx_t'[\beta(z_t)-\beta(z)]K_{tz}+A_n^{-1}(z)\sum_{t=1}^n x_tu_tK_{tz}-A_n^{-1}(z)\left(\sum_{t=1}^n x_tx_t'\right)\left(\frac{1}{n}\sum_{s=1}^n[\hat\beta(z_s)-\hat\beta(z)]K_{sz}\right)\\
&=A^{*-1}_{n,h}\frac{1}{n^2h}\sum_{t=1}^n x_tx_t'[\beta(z_t)-\beta(z)]K_{tz}-A^{*-1}_{n,h}\frac{1}{n^2}\sum_{t=1}^n x_tx_t'\left(\frac{1}{nh}\sum_{s=1}^n[\hat\beta(z_s)-\hat\beta(z)]K_{sz}\right)+A_n^{-1}(z)\sum_{t=1}^n x_tu_tK_{tz}\\
&=A^{*-1}_{n,h}\frac{1}{n^2h}\sum_{t=1}^n x_tx_t'\mathbb{E}\xi_{\beta t}+A^{*-1}_{n,h}\frac{1}{n^2h}\sum_{t=1}^n x_tx_t'\eta_{\beta t}-A^{*-1}_{n,h}\frac{1}{n^2}\sum_{t=1}^n x_tx_t'\left(\frac{1}{nh}\sum_{t=1}^n\{\beta(z_t)-\beta(z)\}K_{tz}\right)\\
&-A^{*-1}_{n,h}\frac{1}{n^2}\sum_{t=1}^n x_tx_t'\left(\frac{1}{nh}\sum_{s=1}^n\left\{\hat\beta(z_t)-\beta(z_t)-[\hat\beta(z)-\beta(z)]\right\}K_{tz}\right)+A_n^{-1}(z)\sum_{t=1}^n x_tu_tK_{tz}\\
&=o(h^{L^*})+A^{*-1}_{n,h}\frac{1}{n^2h}\sum_{t=1}^n x_tx_t'\eta_{\beta t}-A^{*-1}_{n,h}\frac{1}{n^2}\sum_{t=1}^n x_tx_t'\left(\frac{1}{h}\Pi^*_{2n}\right)+A^{*-1}_{n,h}\frac{1}{n^2h}\sum_{t=1}^n x_tu_tK_{tz}\\
&=o(h^{L^*})+A^{*-1}_{n,h}\frac{1}{n^2h}\sum_{t=1}^n x_tx_t'\eta_{\beta t}-A^{*-1}_{n,h}\frac{1}{n^2}\sum_{t=1}^n x_tx_t'\frac{1}{h}\left(\Pi^{*b}_{2n1}-\Pi^{*b}_{2n2}\right)\\
&-A^{*-1}_{n,h}\frac{1}{n^2}\sum_{t=1}^n x_tx_t'\frac{1}{h}\left(\Pi^*_{2n3}-\Pi^*_{2n4}\right)+A^{*-1}_{n,h}\frac{1}{n^2h}\sum_{t=1}^n x_tu_tK_{tz},
\end{aligned}\tag{3.27}$$

where $\xi_{\beta t}=[\beta(z_t)-\beta(z)]K_{tz}$ and $\eta_{\beta t}=\xi_{\beta t}-\mathbb{E}\xi_{\beta t}$.

In view of the above analysis, we find that the bias estimator $\hat B(z)$ is not consistent. Furthermore, it introduces additional variation that affects the final limit distribution through $\Pi^{*b}_{2n1}-\Pi^{*b}_{2n2}$ and $\Pi^*_{2n3}-\Pi^*_{2n4}$. In particular, note that the variance of Π^{*b}_{2n1} and Π^{*b}_{2n2} depends on the unknown value of L and $\beta^{(L)}(z)$, like that of $B_{\eta,L}(\cdot)$ in Lemma B.2 (a). So the asymptotic variance of $\hat\beta(z)-\beta(z)-\hat B(z)$ includes two parts: one part involves $\beta^{(L)}(z)$ and L through the first two terms of (3.27), and the other involves the variation entering through the last two terms of (3.27). Therefore, use of the inconsistent estimator $\hat B(z)$ produces both bias and variance complications that lead to a non-pivotal limit theory for the adaptive statistic. ■

References

Phillips, P. C. B. and Wang, Y. (2021). When bias contributes to variance: True limit theory in functional coefficient cointegrating regression. *Journal of Econometrics, https://doi.org/10.1016/j.jeconom.2021.09.007.*