Online Appendix: Nonparametric estimation and forecasting of time-varying parameter models

The online appendix is organized as follows. In Section A, we discuss the implementation details on how to apply our methods when loss function is not smooth. Section B presents the proofs of Lemmas B1-B4. Section C provides proofs of main theorems 1-3. Section D presents some additional simulation results. Notice that, we write $\theta_u = \theta(u)$ for $u \in (0, 1]$ and $\tilde{\theta}_T = \tilde{\theta}(1)$, $\tilde{\theta}_T^{(1)} = \tilde{\theta}^{(1)}(1)$.

A Quantile predictive regressions

Consider the quantile predictive regression model with time-varying parameters:

$$\mathbb{Q}_\tau(y_{t+h}|x_t) = \alpha_{\tau,t} + \beta_{\tau,t} x_t,$$

where $\mathbb{Q}_\tau(y_{t+h}|x_t) := \inf\{y : F(y|x_t) \geqslant \tau\}$, $F(\cdot|x_t)$ is the conditional c.d.f. of Y_{t+h} given $X_t = x_t$, with density $f(\cdot|x_t)$.

As in Giglio et al. (2016), forecast evaluation at a given quantile τ is based on the check loss:

$$\rho_\tau(\hat{\varepsilon}_{T+h}) = \hat{\varepsilon}_{T+h}[\tau - \mathbb{1}_{\hat{\varepsilon}_{T+h}<0}],$$

where $\hat{\varepsilon}_{T+h} = y_{T+h} - \hat{\alpha}_{\tau,T} - \hat{\beta}_{\tau,T} x_T$. Let $\theta_\tau(1) = (\alpha_{\tau,T}, \beta'_{\tau,T})'$. As the check loss is not differentiable with respect to $\theta_\tau(1)$, our methods in Section 3 can not be directly applied.

Following Fernandes et al. (2021), we consider to apply kernel smoothing to the loss function, yielding

$$\tilde{\rho}_\tau(\varepsilon_{T+h}) = (\rho_\tau * \overline{K}_{\overline{b}})(\varepsilon_{T+h}) = \int_{-\infty}^{\infty} \rho_\tau(v)\overline{K}_{\overline{b}}(v - \varepsilon_{T+h})dv,$$

where $*$ is the convolution operator and $\overline{K}_{\overline{b}}(\cdot)$ is another kernel weighting function with bandwidth parameter $\overline{b} > 0$. Now the smoothed loss function $\tilde{\rho}_\tau(\varepsilon_{T+h})$ is twice continuous differentiable with respect to $\theta(1)$, by the arguments in (7)-(10), we obtain

$$R_T(K,b) = \Big(\hat{\theta}_{K,b,T} - \theta(1)\Big)' \, E_T\Big(\frac{\partial^2 \tilde{\rho}_\tau(\varepsilon_{T+h})}{\partial\theta\partial\theta'}\Big)\Big(\hat{\theta}_{K,\bar{b},T} - \theta(1)\Big).$$

We can replace $E_T\Big(\frac{\partial^2 \tilde{\rho}_\tau(\varepsilon_{T+h})}{\partial\theta\partial\theta'}\Big)$ by the second order derivatives of the in-sample smoothed loss function:

$$\tilde{L}_{\overline{b}}(\theta) = \frac{1}{T}\sum_{t=1}^{T-h} \tilde{\rho}_\tau(\varepsilon_{t+h}) = \int_{-\infty}^{\infty} \rho_\tau(v)\overline{K}_{\overline{b}}(v - \varepsilon_{t+h})dv.$$

It can be easily shown that second order derivatives of $\tilde{L}_{\overline{b}}(\theta)$ is given by

$$\frac{\partial^2 \tilde{L}_{\overline{b}}(\theta)}{\partial\theta\partial\theta'} = \frac{1}{T}\sum_{t=1}^{T-h} x_t x_t' \overline{K}_{\overline{b}}(-\varepsilon_{t+h}(\theta)).$$

B Proof of Lemmas

B.1 Proof of Lemma B1

For the local estimator $\hat{\theta}_{K,b,T}$, it is assumed that $\theta(t/T) \approx \theta_1$. Then θ_1 is obtained via M-estimation minimizing the sample loss function:

$$\hat{\theta}_{K,b,T} = \arg\min_{\theta_1 \in \Theta} \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\ell_{t,T}(\theta_1), \tag{1}$$

where $\ell_{t,T}(\theta_1) = \ell(y_{t,T}, \hat{y}_{t,T|t-1,T}(\theta_1))$. Let $L_T(\theta_1) = \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\ell_{t,T}(\theta_1)$.

<u>*Proof of (i)*</u>: Write $L_T(\theta_1|1) = \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\ell_{1,t}(\theta_1)$, where $\ell_{1,t}(\cdot)$ is the stationary approximation

of $\ell_{t,T}$ at the time point 1. By Definition A1, we have

$$\sup_{\theta_1\in\Theta} |L_T(\theta_1) - L_T(\theta_1|1)| \leqslant \sup_{\theta_1\in\Theta} \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\left|\ell_{t,T}(\theta_1) - \ell_{1,t}(\theta_1)\right|$$
$$\leqslant \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}(T^{-1} + \rho^t) = O(T^{-1}) + O((Tb^{-1/2})) = o(1), \qquad (2)$$

where order of the second term follows from Cauchy–Schwarz inequality:

$$\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\rho^t \leqslant \sqrt{\frac{1}{(Tb)^2}\sum_{t=1}^{T} k_{tT}^2}\sqrt{\sum_{t=1}^{T}\rho^{2t}} = O((Tb^{-1/2})).$$

This implies that (1) can be viewed as

$$\hat{\theta}_{K,b,T} = \arg\min_{\theta_1\in\Theta}\ L_T(\theta_1|1).$$

In view of Theorem 2.1 in Newey and MacFadden (1994), it is sufficient to verify that

(i) $E[\ell_{1,0}(\theta)]$ is uniquely minimized at θ_1 (assumed in Assumption A3(i));

(ii) Θ is compact (assumed in Assumption A1);

(iii) $L_T(\theta_1|1)$ is continuous (implied by Assumption A2(i));

(iv) Uniform weak law of large numbers (UWLLN):

$$\sup_{\theta_1\in\Theta}\left|\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\ell_{1,t}(\theta_1) - E[\ell_{1,0}(\theta)]\right| = o_p(1).$$

What remains is to show (iv). The ergodicity assumed in Assumption A3(i) implied that

$$\left|\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\ell_{1,t}(\theta_1) - E[\ell_{1,0}(\theta)]\right| = o_p(1).$$

Then, uniform consistency result follows if we could show that $L_T(\theta_1|1)$ is stochastic equicontinuous, which follows from the fact that $\ell_{u,t}(\theta)$ is L_1 continuous.

Proof of (ii) and (iii): Let us first define the score and the Hessian:

$$H_{1,T}(\theta) = \frac{\partial^2 L_T(\theta)}{\partial\theta\partial\theta'} = \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial^2 \ell_{t,T}(\theta)}{\partial\theta\partial\theta'}, \quad S_T(\theta) = \frac{\partial L_T(\theta)}{\partial\theta} = \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial \ell_{t,T}(\theta)}{\partial\theta}.$$

By a Taylor series expansion of $\frac{\partial L_T(\hat{\theta}_{K,b,T})}{\partial\theta} = 0$ around the true value θ_1, we have

$$\frac{\partial L_T(\theta_1)}{\partial\theta} + \frac{\partial^2 L_T(\overline{\theta}_1)}{\partial\theta\partial\theta'}(\hat{\theta}_{K,b,T} - \theta_1) = 0,$$

where $\overline{\theta}_1$ lies between θ_1 and $\hat{\theta}_{K,b,T}$. By rearranging terms, we have

$$\begin{aligned}
\hat{\theta}_{K,b,T} - \theta_1 &= -\Big(\frac{\partial^2 L_T(\overline{\theta}_1)}{\partial\theta\partial\theta'}\Big)^{-1}\Big(\frac{\partial L_T(\theta_1)}{\partial\theta}\Big) \\
&= -\Big(\frac{\partial^2 L_T(\theta_1)}{\partial\theta\partial\theta'}\Big)^{-1}\Big(\frac{\partial L_T(\theta_1)}{\partial\theta}\Big) + \left[\Big(\frac{\partial^2 L_T(\theta_1)}{\partial\theta\partial\theta'}\Big)^{-1} - \Big(\frac{\partial^2 L_T(\overline{\theta}_1)}{\partial\theta\partial\theta'}\Big)^{-1}\right]\frac{\partial L_T(\theta_1)}{\partial\theta} \\
&= -\Big(\frac{\partial^2 L_T(\theta_1)}{\partial\theta\partial\theta'}\Big)^{-1}\Big(\frac{\partial L_T(\theta_1)}{\partial\theta}\Big) + \Big(\frac{\partial^2 L_T(\theta_1)}{\partial\theta\partial\theta'}\Big)^{-1}\left[\frac{\partial^2 L_T(\overline{\theta}_1)}{\partial\theta\partial\theta'} - \frac{\partial^2 L_T(\theta_1)}{\partial\theta\partial\theta'}\right] \\
&\qquad\qquad \times \Big(\frac{\partial^2 L_T(\overline{\theta}_1)}{\partial\theta\partial\theta'}\Big)^{-1}\frac{\partial L_T(\theta_1)}{\partial\theta}, \\
&= -H_{1,T}^{-1}(\theta_1)S_T(\theta_1) + H_{1,T}^{-1}(\theta_1)\Big[H_{1,T}(\overline{\theta}_1) - H_{1,T}(\theta_1)\Big]H_{1,T}^{-1}(\overline{\theta}_1)S_T(\theta_1)
\end{aligned} \tag{3}$$

We will show that

$$\left\|S_T(\theta_1)\right\| = O_p((Tb)^{-1/2} + b^{\gamma}), \tag{4}$$

$$\left\|H_{1,T}^{-1}(\theta_1)\right\| = O_p(1), \tag{5}$$

$$\left\|H_{1,T}(\overline{\theta}_1) - H_{1,T}(\theta_1)\right\| = o_p(1). \tag{6}$$

These bounds together with (3) implies the consistency rate in B1(i).

Proof of (4). We have that

$$
\begin{aligned}
S_T(\theta_1) = \frac{\partial L_T(\theta_1)}{\partial \theta} &= \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial \ell_{t,T}(\theta_1)}{\partial \theta} \\
&= \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial \ell_{t,T}(\theta(t/T))}{\partial \theta} + \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial^2 \ell_{t,T}(\bar{\theta}_1)}{\partial \theta \partial \theta'}\left(\theta_1 - \theta(t/T)\right) \\
&= S_{1,T} + B_{2,T},
\end{aligned}
$$

where the second line follows from Taylor series expansion and $\bar{\theta}(1)$ lies between θ_1 and $\theta(t/T)$. We see that the score term is decomposed into a variance term $S_{1,T}$ and a bias term $B_{2,T}$. Using the similar argument as in (2), we have

$$
\left\| S_{1,T} - S^*_{1,T} \right\| = o(1).
$$

where $S^*_{1,T} = \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial \ell_{1,t}\left(\theta(t/T)\right)}{\partial \theta}$. A further Taylor series expansion around θ_1 gives

$$
\begin{aligned}
S^*_{1,T} &= \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial \ell_{1,t}(\theta_1)}{\partial \theta} + \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial^2 \ell_{1,t}(\bar{\theta}_1)}{\partial \theta \partial \theta'}\left(\theta_1 - \theta(t/T)\right) \\
&= S^*_{1,1,T} + S^*_{1,2,T}.
\end{aligned}
$$

By Assumption B3(ii), we have $\left\|S^*_{1,1,T}\right\| = O_p(\frac{1}{\sqrt{Tb}})$. For $S^*_{1,2,T}$, together with Assumption A1(i) and A3(iii), we have

$$
\left\|S^*_{1,2,T}\right\| \leqslant C\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\left\|\frac{\partial^2 \ell_{1,t}(\theta_1)}{\partial \theta \partial \theta'}\right\|\Big(\frac{|t-T|}{T}\Big)^{\gamma} \leqslant C\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\Big(\frac{|t-T|}{T}\Big)^{\gamma} \sim b^{\gamma}\int_{\mathcal{B}} K(u)u^{\gamma}du,
$$

which implies that $\left\|S^*_{1,2,T}\right\| = O_p(b^{\gamma}) = o_p(1)$. This further implies that $\left\|S_{1,T}\right\| \leqslant \left\|S^*_{1,1,T}\right\| + \left\|S^*_{1,2,T}\right\| = O_p(\frac{1}{\sqrt{Tb}})$.

Let us move on to analyze the bias term $B_{2,T}$. Let $\overline{\theta}(1) \to \theta(1)$. Then, we have that

$$\left\|B_{2,T}\right\| \leqslant C\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial^2 \ell_{t,T}(\theta_1)}{\partial\theta\partial\theta'}\Big(\frac{|t-T|}{T}\Big)^{\gamma} = B_{2,T,1}.$$

For $B_{2,T,1}$, following (2), we have

$$\sup_{\theta_1\in\Theta}\left\|B_{2,T,1} - B^*_{2,T,1}\right\| = O_p(b^{\gamma}) = o_p(1),$$

where $B^*_{2,T,1} = \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial^2 \ell_{1,t}(\theta_1)}{\partial\theta\partial\theta'}\Big(\frac{|t-T|}{T}\Big)^{\gamma}$. Since

$$\begin{aligned}
B^*_{2,T,1} &= \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\Big[\frac{\partial^2 \ell_{1,t}(\theta_1)}{\partial\theta\partial\theta'} - E\Big[\frac{\partial^2 \ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big] + E\Big[\frac{\partial^2 \ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big]\Big]\Big(\frac{|t-T|}{T}\Big)^{\gamma} \\
&= \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\Big[\frac{\partial^2 \ell_{1,t}(\theta_1)}{\partial\theta\partial\theta'} - E\Big[\frac{\partial^2 \ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big]\Big]\Big(\frac{|t-T|}{T}\Big)^{\gamma} + \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}E\Big[\frac{\partial^2 \ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big]\Big(\frac{|t-T|}{T}\Big)^{\gamma} \\
&= B^*_{2,T,1,1} + B^*_{2,T,1,2}.
\end{aligned}$$

By Assumption A3(iii), we have $\left\|B^*_{2,T,1,1}\right\| = O_p(b^{\gamma})$. For $B^*_{2,T,1,2}$, we again have

$$\left\|B^*_{2,T,1,2}\right\| \leqslant C\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\Big(\frac{|t-T|}{T}\Big)^{\gamma} \sim b^{\gamma}\int_C K(u)u^{\gamma}du,$$

which also implies that $\left\|B^*_{2,T,1,2}\right\| = O_p(b^{\gamma})$. Then, (4) follows again from triangular inequality.

Proof of (5). It follows again similarly from (2) that

$$\sup_{\theta_1\in\Theta}\left\|H_{1,T}(\theta_1) - H^*_{1,T}(\theta_1)\right\| = o_p(1),$$

where $H^*_{1,T}(\theta_1) = \frac{1}{Tb}\sum_{t=1}^T k_{tT}\frac{\partial^2\ell_{1,t}(\theta_1)}{\partial\theta\partial\theta'}$. Write

$$\begin{aligned} H^*_{1,T}(\theta_1) &= \frac{1}{Tb}\sum_{t=1}^T k_{tT}E\Big[\frac{\partial^2\ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big] + \frac{1}{Tb}\sum_{t=1}^T k_{tT}\Big(\frac{\partial^2\ell_{1,t}(\theta_1)}{\partial\theta\partial\theta'} - E\Big[\frac{\partial^2\ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big]\Big) \\ &= H^*_{1,T,1} + H^*_{1,T,2} = H^*_{1,T,1}(I_k + \tilde{\Delta}^*_T), \end{aligned} \tag{7}$$

where $\tilde{\Delta}^*_T = (H^*_{1,T,1})^{-1}(H^*_{1,T} - H^*_{1,T,1})$. By Assumption A3(iii), there exists $v > 0$ such that for all $t \geqslant 1$,

$$a'E\Big[\frac{\partial^2\ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big]a \geqslant 1/v > 0.$$

Thus, we have, for any $k \times 1$ vector $a = (a_1, \cdots, a_k)'$ such that $\|a\|^2 = 1$

$$\min_{\|a\|=1} a'H^*_{1,T,1}a = \min_{\|a\|=1} \Big(\frac{1}{Tb}\sum_{t=1}^T k_{tT}a'E\Big[\frac{\partial^2\ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big]a\Big) \geqslant \frac{1}{v}\Big(\frac{1}{Tb}\sum_{t=1}^T k_{tT}\Big) > 0.$$

This means that the smallest eigenvalue of $H^*_{1,T,1}$ is not smaller than $1/v > 0$, which further implies that

$$\Big\|(H^*_{1,T,1})^{-1}\Big\|_{sp} = O_p(1).$$

In addition, by Assumption A3(iii), we have

$$\Big\|H^*_{1,T} - H^*_{1,T,1}\Big\|_{sp} = o_p(1).$$

Then,

$$\Big\|H^{*-1}_{1,T}(\theta_1)\Big\|_{sp} \leqslant \Big\|(H^*_{1,T,1})^{-1}\Big\|_{sp}(1 - \Big\|H^*_{1,T} - H^*_{1,T,1}\Big\|_{sp})^{-1} = O_p(1),$$

which implies that $\Big\|H^{-1}_{1,T}(\theta_1)\Big\|_{sp} = O_p(1)$.

Proof of (6). This follow immediately by the consistency: $\hat{\theta}_{K,b,T} \xrightarrow{p} \theta_1$.

Back to (3), we have

$$\sqrt{Tb}\left(\hat{\theta}_{K,b,T}-\theta_1\right)=-H^{*-1}_{1,T}(\theta_1)\ \sqrt{Tb}(S_{1,T}+B_{2,T})$$

Since $\left\|\sqrt{Tb}S_{1,T}\right\|=O_p(1)$, $\left\|\sqrt{Tb}S_{2,T}\right\|=O_p(T^{1/2}b^{1/2+\gamma})$, under the condition $T^{1/2}b^{1/2+\gamma}\to 0$, the dominating term is the first one, by applying CLT on $\sqrt{Tb}S_{1,T}$, together with Slutsky's theorem, we obtain

$$\sqrt{Tb}\left(\hat{\theta}_{K,b,T}-\theta_1\right)\xrightarrow{d}\mathcal{N}(0,\phi_{0,K}\Sigma_1),$$

where $\Sigma_1=H_1^{-1}\Lambda_1H^{-1}$, $H_1=E\left[\frac{\partial^2\ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\right]$ and $\Lambda_1=Var\left(\frac{\partial\ell_{1,0}(\theta_1)}{\partial\theta'}\right)$.

B.2 Proof of Lemma B2

The objective function is given by

$$L_T(\theta_1,\theta_1^{(1)})=\frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\ell_{t,T}\Big(\theta_1+\theta_1^{(1)}(t/T-1)\Big).$$

Define $\beta_1=\theta_1+\theta_1^{(1)}(t/T-1)$. Similarly as in (3), we have that

$$\begin{pmatrix}\tilde{\theta}_T-\theta_1\\ \tilde{\theta}_T^{(1)}-\theta_1^{(1)}\end{pmatrix}=-\Big(\frac{\partial L_T^2(\beta_1)}{\partial\beta_1\partial\beta_1'}\Big)^{-1}\frac{\partial L_T(\beta_1)}{\partial\beta_1}+o_p(1). \tag{8}$$

Notice that

$$\frac{\partial L_T^2(\beta_1)}{\partial\beta_1\partial\beta_1'}=\begin{bmatrix}\frac{1}{T\tilde{b}}\sum_{t=1}^T\tilde{k}_{tT}\frac{\partial^2\ell_{t,T}\left(\beta_1\right)}{\partial\theta_1\partial\theta_1'} & \frac{1}{T\tilde{b}}\sum_{t=1}^T\tilde{k}_{tT}\frac{\partial^2\ell_{t,T}\left(\beta_1\right)}{\partial\theta_1\partial\theta_1'^{(1)}}\left(\frac{t-T}{T}\right)\\ \frac{1}{T\tilde{b}}\sum_{t=1}^T\tilde{k}_{tT}\frac{\partial^2\ell_{t,T}\left(\beta_1\right)}{\partial\theta_1^{(1)}\partial\theta_1'}\left(\frac{t-T}{T}\right) & \frac{1}{T\tilde{b}}\sum_{t=1}^T\tilde{k}_{tT}\frac{\partial^2\ell_{t,T}\left(\beta_1\right)}{\partial\theta_1^{(1)}\partial\theta_1'^{(1)}}\left(\frac{t-T}{T}\right)^2\end{bmatrix}.$$

Using similar arguments for the proofs of (4)-(5), we have

$$\Big\|\frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial^2\ell_{t,T}\big(\beta_1\big)}{\partial\theta_1\partial\theta_1'}\Big\| = O_p(1), \quad \Big\|\frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial^2\ell_{t,T}\big(\beta_1\big)}{\partial\theta_1\partial\theta_1'^{(1)}}\Big(\frac{t-T}{T}\Big)\Big\| = O_p(\tilde{b})$$
$$\Big\|\frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial^2\ell_{t,T}\big(\beta_1\big)}{\partial\theta_1^{(1)}\partial\theta_1'}\Big(\frac{t-T}{T}\Big)\Big\| = O_p(\tilde{b}), \quad \Big\|\frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial^2\ell_{t,T}\big(\beta_1\big)}{\partial\theta_1^{(1)}\partial\theta_1'^{(1)}}\Big(\frac{t-T}{T}\Big)^2\Big\| = O_p(\tilde{b}^2)$$

Using the property of the inverse of the partitioned matrices (see, Abadir and Magnus (2005)), we have

$$\Big(\frac{\partial L_T^2(\beta_1)}{\partial\beta_1\partial\beta_1'}\Big)^{-1} = \begin{bmatrix} O_p(1) & O_p(\tilde{b}^{-1}) \\ O_p(\tilde{b}^{-1}) & O_p(\tilde{b}^{-2}) \end{bmatrix}.$$

Next, we have

$$\frac{\partial L_T(\beta_1)}{\partial\beta_1} = \begin{bmatrix} \frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial\ell_{t,T}\big(\beta_1\big)}{\partial\theta_1} \\ \frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial\ell_{t,T}\big(\beta_1\big)}{\partial\theta_1^{(1)}}\Big(\frac{t-T}{T}\Big) \end{bmatrix} = \begin{bmatrix} \tilde{S}_{1,T} \\ \tilde{S}_{2,T} \end{bmatrix}. \tag{9}$$

By Assumption A1(ii), we have that

$$\theta(t/T) \approx \theta_1 + \theta_1^{(1)}\Big(\frac{t-T}{T}\Big) + \frac{\theta_1^{(2)}}{2}\Big(\frac{t-T}{T}\Big)^2.$$

Then, for $\tilde{S}_{1,T}$, Taylor expansion around $\theta(t/T)$ gives

$$\begin{aligned}\tilde{S}_{1,T} &= \frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial\ell_{t,T}(\theta(t/T))}{\partial\theta_1} + \frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial^2\ell_{t,T}(\bar{\theta}_1)}{\partial\theta_1\partial\theta_1'}\Big(\theta_1 + \theta^{(1)}(1)\Big(\frac{t-T}{T}\Big) - \theta(t/T)\Big) \\ &= \frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial\ell_{t,T}(\theta(t/T))}{\partial\theta_1} + \frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial^2\ell_{t,T}(\bar{\theta}_1)}{\partial\theta_1\partial\theta_1'}\frac{\theta_1^{(2)}}{2}\Big(\frac{t-T}{T}\Big)^2,\end{aligned}$$

where $\bar{\theta}_1$ lies between $\theta(t/T)$ and β_1. Similarly, we have

$$\tilde{S}_{2,T} = \frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial\ell_{t,T}(\theta(t/T))}{\partial\theta^{(1)}}\Big(\frac{t-T}{T}\Big) + \frac{1}{T\tilde{b}}\sum_{t=1}^{T}\tilde{k}_{tT}\frac{\partial^2\ell_{t,T}(\bar{\theta}_1)}{\partial\theta^{(1)}\partial\theta'^{(1)}}\frac{\theta^{(2)}(1)}{2}\Big(\frac{t-T}{T}\Big)^3.$$

Now, back to (8), we have

$$\begin{pmatrix} \tilde{\theta}_T - \theta_1 \\ \tilde{\theta}_T^{(1)} - \theta_1^{(1)} \end{pmatrix} = -\Big(\frac{\partial L_T^2(\beta_1)}{\partial \beta_1 \partial \beta_1'}\Big)^{-1} \underbrace{\begin{bmatrix} \frac{1}{T\tilde{b}} \sum_{t=1}^T \tilde{k}_{tT} \frac{\partial \ell_{t,T}\big(\theta(t/T)\big)}{\partial \theta_1} \\ \frac{1}{T\tilde{b}} \sum_{t=1}^T \tilde{k}_{tT} \frac{\partial \ell_{t,T}\big(\bar{\theta}_1\big)}{\partial \beta_2}\Big(\frac{t-T}{T}\Big) \end{bmatrix}}_{Q_{1,T}}$$

$$- \Big(\frac{\partial L_T^2(\beta_1)}{\partial \beta_1 \partial \beta_1'}\Big)^{-1} \underbrace{\begin{bmatrix} \frac{1}{T\tilde{b}} \sum_{t=1}^T \tilde{k}_{tT} \frac{\partial^2 \ell_{t,T}\big(\bar{\theta}_1\big)}{\partial \theta_1 \partial \theta_1'} \frac{\theta^{(2)}(1)}{2}\Big(\frac{t-T}{T}\Big)^2 \\ \frac{1}{T\tilde{b}} \sum_{t=1}^T \tilde{k}_{tT} \frac{\partial^2 \ell_{t,T}\big(\bar{\theta}_1\big)}{\partial \theta_1^{(1)} \partial \theta_1'^{(1)}} \frac{\theta^{(2)}(1)}{2}\Big(\frac{t-T}{T}\Big)^3 \end{bmatrix}}_{Q_{2,T}}.$$

Following again the proofs of (4)-(5), we have

$$Q_{1,T} = \begin{bmatrix} O_p((T\tilde{b})^{-1/2}) \\ O_p((T\tilde{b})^{-1/2}\tilde{b}) \end{bmatrix}, \quad Q_{2,T} = \begin{bmatrix} O_p(\tilde{b}^2) \\ O_p(\tilde{b}^3) \end{bmatrix}.$$

Therefore, we obtain the consistency rate for $\tilde{\theta}_T$:

$$\big\|\tilde{\theta}_T - \theta_1\big\| = O_p((T\tilde{b})^{-1/2} + \tilde{b}^2).$$

B.3 Proof of Lemma B3

Write $\hat{\theta}_{\overline{K},b,T} = \hat{\theta}_{b,T}$. As in (3), the estimator can be decomposed as

$$\begin{aligned} \hat{\theta}_{b,T} - \theta_1 &= -H_{1,T} S_T + o_p(1) \\ &= -H_{1,T}(S_{1,T} + B_{2,T}) + o_p(1), \end{aligned} \tag{10}$$

where

$$
\begin{aligned}
H_{1,T} &= \left(\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial^2 \ell_{t,T}(\theta_1)}{\partial\theta\partial\theta'}\right)^{-1}, \\
S_{1,T} &= \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial \ell_{t,T}(\theta(t/T))}{\partial\theta}, \quad B_{2,T} = \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial^2 \ell_{t,T}(\bar{\theta}_1)}{\partial\theta\partial\theta'}(\theta_1 - \theta(t/T)),
\end{aligned}
$$

and $\bar{\theta}_1$ lies between θ_1 and $\theta(t/T)$. We will show that

$$
\sup_{b\in I_T}\left\|\frac{1}{T^{1/2}b^{1/2-\delta}}\sum_{t=1}^{T} k_{tT}\frac{\partial \ell_{t,T}(\theta_1)}{\partial\theta}\right\| = O_p(1), \quad \text{for } 0<\delta<1/2, \tag{11}
$$

$$
\sup_{b\in I_T}\left\|\left(\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial^2 \ell_{t,T}(\theta_1)}{\partial\theta\partial\theta'}\right)^{-1}\right\| = O_p(1), \tag{12}
$$

$$
\sup_{b\in I_T}\left\|\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial^2 \ell_{t,T}(\bar{\theta}_1)}{\partial\theta\partial\theta'}(\theta_1 - \theta(t/T))\right\| = O_p(b^{\gamma}) \quad \text{for } 0<\gamma\leqslant 1 \tag{13}
$$

These bounds together with (10) prove (B1).

Proof of (11). By Boole's inequality and Chebyshev's inequality, we have, for any $\varepsilon > 0$,

$$
\begin{aligned}
\mathbb{P}\left(\sup_{b\in I_T}\left\|\frac{1}{T^{1/2}b^{1/2-\delta}}\sum_{t=1}^{T} k_{tT}\frac{\partial \ell_{t,T}(\theta_1)}{\partial\theta}\right\| > \varepsilon\right) &\leqslant \sum_{b\in I_T}\mathbb{P}\left(\left\|\frac{1}{T^{1/2}b^{1/2-\delta}}\sum_{t=1}^{T} k_{tT}\frac{\partial \ell_{t,T}(\theta_1)}{\partial\theta}\right\| > \varepsilon\right) \\
&\leqslant |I_T|\times\sup_{b\in I_T}\ \mathbb{P}\left(\left\|\frac{1}{T^{1/2}b^{1/2-\delta}}\sum_{t=1}^{T} k_{tT}\frac{\partial \ell_{t,T}(\theta_1)}{\partial\theta}\right\| > \varepsilon\right) \\
&\leqslant |I_T|\times\sup_{b\in I_T}\ \frac{C}{b^{\delta}\varepsilon^2} = O(1),
\end{aligned}
$$

where the third inequality follows from (4).

Proof of (12). As in (5), consider $H^*_{1,T}$:

$$
\begin{aligned}
\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\frac{\partial^2 \ell_{1,t}(\theta_1)}{\partial\theta\partial\theta'} &= \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}E\Big[\frac{\partial^2 \ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big] + \frac{1}{Tb}\sum_{t=1}^{T} k_{tT}\left(\frac{\partial^2 \ell_{1,t}(\theta_1)}{\partial\theta\partial\theta'} - E\Big[\frac{\partial^2 \ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big]\right) \\
&= H^*_{1,T,1} + H^*_{1,T,2} = H^*_{1,T,1}(I_k + \tilde{\Delta}^*_T),
\end{aligned} \tag{14}
$$

where $\tilde{\Delta}_T^* = (H_{1,T,1}^*)^{-1}(H_{1,T}^* - H_{1,T,1}^*)$. First, (5) holds uniformly over b:

$$\sup_{b\in I_T} \left\|(H_{1,T,1}^*)^{-1}\right\|_{sp} = O_p(1). \tag{15}$$

For $\tilde{\Delta}_T^*$, let $\tilde{\Delta}_t^* = \frac{\partial^2 \ell_{1,t}(\theta_1)}{\partial\theta\partial\theta'} - E\Big[\frac{\partial^2 \ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big]$. Then, for any $\varepsilon > 0$, , by Boole's inequality and Chebyshev's inequality, we have

$$\begin{aligned}\mathbb{P}\left(\sup_{b\in I_T} \left\|\frac{1}{Tb}\sum_{t=1}^T k_{tT}\tilde{\Delta}_t^*\right\| > \varepsilon\right) &\leqslant \sum_{b\in I_T} \mathbb{P}\left(\left\|\frac{1}{Tb}\sum_{t=1}^T k_{tT}\tilde{\Delta}_t^*\right\| > \varepsilon\right)\\ &\leqslant |I_T| \times \sup_{b\in I_T} \mathbb{P}\left(\left\|\frac{1}{Tb}\sum_{t=1}^T k_{tT}\tilde{\Delta}_t^*\right\| > \varepsilon\right).\end{aligned}$$

Similarly as in the proof of (11), we have

$$\sup_{b\in I_T} \left\|\frac{1}{Tb}\sum_{t=1}^T k_{tT}\tilde{\Delta}_t^*\right\| = o_p(1). \tag{16}$$

Since the cardinality of the set $|I_T|$ must be $o(1)$, we have $\sup_{b\in I_T} \left\|\tilde{\Delta}_T^*\right\|_{sp} = o_p(1)$. To sum up, we continue from (7):

$$\sup_{b\in I_T} \left\|H_{1,T}^{*-1}\right\|_{sp} \leqslant \underbrace{\sup_{b\in I_T} \left\|(H_{1,T,1}^*)^{-1}\right\|_{sp}}_{O_p(1) \text{ by } (15)} \Big(1 - \underbrace{\sup_{b\in I_T} \left\|\tilde{\Delta}_T^*\right\|_{sp}}_{o_p(1) \text{ by } (16)}\Big)^{-1} = O_p(1).$$

This also implies (12).

Proof of (13). Let $\overline{\theta}_1 \to \theta_1$ and consider $B_{2,T}^*$:

$$\begin{aligned}B_{2,T}^* &= \frac{1}{Tb}\sum_{t=1}^T k_{tT}\Big(\frac{\partial^2 \ell_{1,t}(\theta_1)}{\partial\theta\partial\theta'} - E\Big[\frac{\partial^2 \ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big]\Big)(\theta_1 - \theta(t/T)) + \frac{1}{Tb}\sum_{t=1}^T k_{tT} E\Big[\frac{\partial^2 \ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big](\theta_1 - \theta(t/T))\\ &= B_{2,T,1}^* + B_{2,T,2}^*.\end{aligned}$$

For $B^*_{2,T,1}$, again, similarly as in (11), we have

$$\mathbb{P}\left(\sup_{b\in I_T} \left\|B^*_{2,T,1}\right\| > \varepsilon\right) \leqslant \sum_{b\in I_T} \mathbb{P}\left(\left\|B^*_{2,T,1}\right\| > \varepsilon\right)$$
$$\leqslant |I_T| \times \sup_{b\in I_T} \mathbb{P}\left(\left\|B^*_{2,T,1}\right\| > \varepsilon\right) = O(b^{\gamma+\delta}),$$

for some $0 < \delta < 1/2$. Moving to $B^*_{2,T,2}$, notice that

$$\left\|B^*_{2,T,2}\right\| \leqslant C\Big(\frac{1}{Tb}\sum_{t=1}^{T} k_{tT}(\frac{|t-T|}{T})^{\gamma}\Big) \approx b^{\gamma}\int_C u^{\gamma}K(u)du = O(b^{\gamma}),$$

which holds uniformly over b. Thus, we have

$$\sup_{b\in I_T}\left\|B^*_{2,T}\right\| \leqslant \sup_{b\in I_T}\left\|B^*_{2,T,1}\right\| + \sup_{b\in I_T}\left\|B^*_{2,T,2}\right\| = O_p(b^{\gamma}),$$

which implies (13).

B.4 Proof of Lemma B4

Let us first expand $A(b)$:

$$
\begin{aligned}
A(b) &= (\hat{\theta}_{b,T} - \tilde{\theta}_T)' \, \omega_T(\tilde{\theta}_T) \, (\hat{\theta}_{b,T} - \tilde{\theta}_T) \\
&= (\hat{\theta}_{b,T} - \theta_1 + \theta_1 + \tilde{\theta}_T)' \Bigg(\omega_T(\theta_1) + \underbrace{\Big[\frac{\partial \omega_T(\theta_1)}{\partial \theta_1} (\tilde{\theta}_T - \theta_1)) \ \ldots \ \frac{\partial \omega_T(\theta_1)}{\partial \theta_p} (\tilde{\theta}_T - \theta_1) \Big]}_{\tilde{\omega}_T(\theta_1)_{p \times p}} \Bigg) \\
&\qquad\qquad \times (\hat{\theta}_{b,T} - \theta_1 + \theta_1 + \tilde{\theta}_T) \\
&= L(b) - 2(\hat{\theta}_{b,T} - \theta_1)' \omega_T(\theta_1)(\tilde{\theta}_T - \theta_1) + (\tilde{\theta}_T - \theta_1)' \omega_T(\theta_1)(\tilde{\theta}_T - \theta_1) \\
&\quad + (\hat{\theta}_{b,T} - \theta_1)' \tilde{\omega}_T(\theta_1)(\hat{\theta}_{b,T} - \theta_1) - 2(\hat{\theta}_{b,T} - \theta_1)' \tilde{\omega}_T(\theta_1)(\tilde{\theta}_T - \theta_1) \\
&\quad + (\tilde{\theta}_T - \theta_1)' \tilde{\omega}_T(\theta_1)(\tilde{\theta}_T - \theta_1) \\
&= L(b) - 2D_1(b) + D_1' + D_2(b) - 2D_3(b) + D_2',
\end{aligned}
$$

where

$$
\begin{aligned}
&D_1(b) = (\hat{\theta}_{b,T} - \theta_1)' \omega_T(\theta_1)(\tilde{\theta}_T - \theta_1), \quad D_1' = (\tilde{\theta}_T - \theta_1)' \omega_T(\theta_1)(\tilde{\theta}_T - \theta_1), \\
&D_2(b) = (\hat{\theta}_{b,T} - \theta_1)' \tilde{\omega}_T(\theta_1)(\hat{\theta}_{b,T} - \theta_1), \quad D_3(b) = (\hat{\theta}_{b,T} - \theta_1)' \tilde{\omega}_T(\theta_1)(\tilde{\theta}_T - \theta_1), \\
&D_2' = (\tilde{\theta}_T - \theta_1)' \tilde{\omega}_T(\theta_1)(\tilde{\theta}_T - \theta_1).
\end{aligned}
$$

Then, we have

$$
\frac{L(b) - A(b)}{L(b)} = \frac{2D_1(b)}{L(b)} - \frac{D_1'}{L(b)} - \frac{D_2(b)}{L(b)} + \frac{D_3(b)}{L(b)} - \frac{D_2'}{L(b)}.
$$

By Lemma B2 and Assumption A5(i), we have

$$
\left\| \tilde{\theta}_T - \theta_1 \right\| = O_p((T\tilde{b})^{-1/2}). \tag{17}
$$

We will show that

$$\sup_{b\in I_T}\left|\frac{D_1(b)}{L(b)}\right| = o_p(1),\quad \sup_{b\in I_T}\left|\frac{D_2(b)}{L(b)}\right| = o_p(1),\quad \sup_{b\in I_T}\left|\frac{D_3(b)}{L(b)}\right| = o_p(1), \tag{18}$$

$$\sup_{b\in I_T}\left|\frac{D_1'}{L(b)}\right| = o_p(1),\quad \sup_{b\in I_T}\left|\frac{D_2'}{L(b)}\right| = o_p(1). \tag{19}$$

These bounds together with triangular inequality imply (B2).

Proof of (18). First, by Lemma B3, we have

$$\sup_{b\in I_T}|L(b)| \leqslant \sup_{b\in I_T}\left\|\hat{\theta}_{b,T}-\theta_1\right\| \|\omega_T(\theta_1)\|_{sp} \sup_{b\in I_T}\left\|\hat{\theta}_{b,T}-\theta_1\right\| = O_p(r^2_{T,b,\delta,\gamma}), \tag{20}$$

for some $0<\delta<1/2$ and $0<\gamma\leqslant 1$. Write $\tilde{r}_{T,\tilde{b}} = (T\tilde{b})^{-1/2}$, we also have

$$\begin{aligned}
\sup_{b\in I_T}|D_1(b)| &\leqslant \sup_{b\in I_T}\left\|\hat{\theta}_{b,T}-\theta_1\right\| \|\omega_T(\theta_1)\|_{sp} \left\|\tilde{\theta}_T-\theta_1\right\| = O_p(r_{T,b,\delta,\gamma}\tilde{r}_{T,\tilde{b}}),\\
\sup_{b\in I_T}|D_2(b)| &\leqslant \sup_{b\in I_T}\left\|\hat{\theta}_{b,T}-\theta_1\right\| \|\tilde{\omega}_T(\theta_1)\|_{sp} \sup_{b\in I_T}\left\|\hat{\theta}_{b,T}-\theta_1\right\| = O_p(r^2_{T,b,\delta,\gamma}\tilde{r}_{T,\tilde{b}}),\\
\sup_{b\in I_T}|D_3(b)| &\leqslant \sup_{b\in I_T}\left\|\hat{\theta}_{b,T}-\theta_1\right\| \|\tilde{\omega}_T(\theta_1)\|_{sp} \left\|\tilde{\theta}_T-\theta_1\right\| = O_p(r_{T,b,\delta,\gamma}\tilde{r}^2_{T,\tilde{b}}).
\end{aligned}$$

These bounds imply that

$$\sup_{b\in I_T}\left|\frac{D_1(b)}{L(b)}\right| = O_p\Big(\frac{\tilde{r}_{T,\tilde{b}}}{r_{T,b,\delta,\gamma}}\Big) = o_p(1),$$

where $\frac{\tilde{r}_{T,\tilde{b}}}{r_{T,b,\delta,\gamma}} \to 0$ is guaranteed by Assumption A5. Similarly, we have

$$\sup_{b\in I_T}\left|\frac{D_2(b)}{L(b)}\right| = O_p(\tilde{r}_{T,\tilde{b}}) = o_p(1),$$

as $T\tilde{b}\to\infty$. Finally, we have

$$\sup_{b\in I_T}\left|\frac{D_3(b)}{L(b)}\right| = O_p\Big(\frac{\tilde{r}^2_{T,\tilde{b}}}{r_{T,b,\delta,\gamma}}\Big) = o_p(1),$$

where $\frac{\tilde{r}^2_{T,\tilde{b}}}{r_{T,b,\delta,\gamma}} \to 0$ is again guaranteed by Assumption A5.

Proof of (19). First, it is straightforward to show that

$$\left|D_1'\right| = O_p(\tilde{r}^2_{T,\tilde{b}}), \ \ \left|D_2'\right| = O_p(\tilde{r}^3_{T,\tilde{b}}).$$

Together with (20) and following the same reasoning above, we have

$$\sup_{b\in I_T}\left|\frac{D_1'}{L(b)}\right| = O_p\Big(\frac{\tilde{r}^2_{T,\tilde{b}}}{r^2_{T,b,\delta,\gamma}}\Big) = o_p(1), \ \ \sup_{b\in I_T}\left|\frac{D_2'}{L(b)}\right| = O_p\Big(\frac{\tilde{r}^3_{T,\tilde{b}}}{r^2_{T,b,\delta,\gamma}}\Big) = o_p(1).$$

C Proofs of the theorems

C.1 Proof of Theorem 1

Write $\hat{\theta}_{\overline{K},b,T} = \hat{\theta}_{b,T}$ and $\omega_T(\theta_1) = E_T\Big(\frac{\partial^2 \ell_{T+h}(\theta_1)}{\partial\theta\partial\theta'}\Big)$. It follows from Lemma B1 that, the infeasible objective function can be written as

$$(\hat{\theta}_{b,T} - \theta_1)'\, \omega_T(\theta_1)\, (\hat{\theta}_{b,T} - \theta_1) = r_{T,b}q_T,$$

where q_T is a scalar $O_p(1)$ random variable and $r_{T,b,\gamma} = (Tb)^{-1/2} + b^\gamma$ for some $0 < \gamma \leqslant 1$. The first-order condition of $r_{T,b,\gamma}$ with respect to b gives $\hat{b} = O_p(T^{-\frac{1}{2\gamma+1}})$. Since the second order derivative of $r_{T,b,\gamma}$ is always positive, the optimal bandwidth minimize the objective function.

C.2 Proof of Theorem 2

Write $\hat{\theta}_{\overline{K},b,T} = \hat{\theta}_{b,T}$ and $\omega_T(\theta_1) = E_T\Big(\frac{\partial^2 \ell_{T+h}(\theta_1)}{\partial\theta\partial\theta'}\Big)$. Let

$$\hat{b} := \arg\min_{b\in I_T} (\hat{\theta}_{b,T} - \tilde{\theta}(1))'\, \omega_T(\tilde{\theta}(1))\, (\hat{\theta}_{b,T} - \tilde{\theta}(1))$$

be the bandwidth selected according to the feasible criterion. As in the proof of Lemma B4, the decomposition of $A(b)$ implies that

$$A(\hat{b}) = L(\hat{b}) - 2D_1(\hat{b}) + D_1' + D_2(\hat{b}) - 2D_3(\hat{b}) + D_2'.$$

Then, we have

$$\begin{aligned}\frac{A(\hat{b})}{\inf_{b\in I_T} L(b)} &= \frac{L(\hat{b})}{\inf_{b\in I_T} L(b)} - \frac{2D_1(\hat{b})}{\inf_{b\in I_T} L(b)} + \frac{D_2(\hat{b})}{\inf_{b\in I_T} L(b)} - \frac{2D_3(\hat{b})}{\inf_{b\in I_T} L(b)} + \frac{D_1'}{\inf_{b\in I_T} L(b)} + \frac{D_2'}{\inf_{b\in I_T} L(b)} \\ &= I_1(\hat{b}) + I_2(\hat{b}) + I_3(\hat{b}) + I_4(\hat{b}) + I_5 + I_6.\end{aligned}$$

Following (18) and (19), we have

$$I_2(\hat{b}) = o_p(1), \;\; I_3(\hat{b}) = o_p(1), \;\; I_4(\hat{b}) = o_p(1), \;\; I_5 = o_p(1), \;\; I_6 = o_p(1).$$

What remains is to show that

$$I_1(\hat{b}) \xrightarrow{p} 1,$$

which is equivalent to verify that, for any $b, b' \in I_T$,

$$\sup_{b,b'\in I_T} \left| \frac{L(b) - L(b') - (A(b) - A(b'))}{L(b) + L(b')} \right| \xrightarrow{p} 0.$$

This follows immediately from Lemma B4:

$$\sup_{b,b'\in I_T} \left| \frac{L(b) - L(b') - (A(b) - A(b'))}{L(b) + L(b')} \right| \leqslant \sup_{b\in I_T} \left| \frac{L(b) - A(b)}{L(b)} \right| + \sup_{b'\in I_T} \left| \frac{L(b') - A(b')}{L(b')} \right| = o_p(1).$$

C.3 Proof of Theorem 3

In Lemma B1, we show that the local estimator obeys the following expansion:

$$\hat{\theta}_{K,b,T} - \theta_1 = -H_{1,T}^{-1} S_{1,T} - H_{1,T}^{-1} B_{2,T},$$

where

$$H_{1,T} = \frac{1}{Tb} \sum_{t=1}^{T} k_{tT} \frac{\partial^2 \ell_{t,T}(\theta_1)}{\partial\theta\partial\theta'}, \quad S_{1,T} = \frac{1}{Tb} \sum_{t=1}^{T} k_{tT} \frac{\partial \ell_{t,T}(\theta(t/T))}{\partial\theta}, \quad B_{2,T} = \frac{1}{Tb} \sum_{t=1}^{T} k_{tT} \frac{\partial^2 \ell_t(\bar{\theta}_1)}{\partial\theta\partial\theta'} (\theta_1 - \theta(t/T)),$$

and $\bar{\theta}_1$ lies between $\hat{\theta}_{K,b,T}$ and θ_1. Following the proof of Lemma B1, $\left\|H_{1,T}^{-1}\right\| = O_p(1)$, $\left\|S_{1,T}\right\| = O_p((Tb)^{-1/2})$, $\left\|B_{2,T}\right\| = O_p(b^{\gamma})$, when $T^{1/2} b^{1/2+\gamma} \to 0$, the dominating term is $S_{1,T}$. CLT in LemmaB1(ii) holds. Theorem 3(i) follows immediately from continuous mapping theorem.

When $T^{1/2} b^{1/2+\gamma} \to \infty$, the dominating term is $S_{2,T}$. By similar analysis as in the proof of Lemma B1, we have

$$b^{-\gamma} S_{2,T} \xrightarrow{p} \mu_{\gamma,K} E\Big[\frac{\partial^2 \ell_{1,0}(\theta_1)}{\partial\theta\partial\theta'}\Big] C,$$

where $C = (c_1, \cdots, c_{\bar{k}})'$ is a collection of constant given in Assumption A1(i). Theorem 3(ii) follows again from continuous mapping theorem. Theorem 3(iii) follows immediately by combining the results obtained in (i) and (ii).

D Additional simulation results

This section presents additional simulation results using different specifications for $(a_t, b_t)'$ in (17). In particular, we consider the case when the parameters are constant over time $(a_t, b_t)' = (a, b)$ for all t, as well as cases when $(a_t, b_t)'$ have a one-time break point. We have 7 different specifications for $(a_t, b_t)'$. For DGP D1, $(a_t, b_t)' = (0.9, 1)'$ for all t. For DGPs D2-D7, they are generated according to

(1) D2-D4: $a_t = 0.9 - \frac{1}{T^{0.2}}\mathbb{1}(t \geqslant \pi T + 1)$, $b_t = 1 + \frac{1}{T^{0.2}}\mathbb{1}(t \geqslant \pi T + 1)$, where $\pi = 0.25, 0.5, 0.75$, respectively;

(2) D5-D7: $a_t = 0.9 - \frac{1}{T^{0.5}}\mathbb{1}(t \geqslant \pi T + 1)$, $b_t = 1 + \frac{1}{T^{0.5}}\mathbb{1}(t \geqslant \pi T + 1)$, where $\pi = 0.25, 0.5, 0.75$, respectively.

Notice that, the main differences between DGPs D5-D7 and D2-D4 are that the break size is relatively larger in the previous group.

Table D1 presents the out-of-sample forecasting performance from the simulated dataset. Overall, for 1-step ahead forecast, the local estimator with optimal bandwidths selection is quite useful when break size is relatively small. For 12-step ahead forecast, all local estimators perform better than non-local estimator even if the true parameters are constant over time. Overall, rolling window forecasts with window size equal to 60 are more likely to be the best, but the performance from using $K_2(u)$ with optimal bandwidth selection improves as sample size increases, particularly for 12-step ahead forecast.

References

Abadir, K. M. and J. R. Magnus (2005). *Matrix algebra*, Volume 1. Cambridge University Press.

Fernandes, M., E. Guerre, and E. Horta (2021). Smoothing quantile regressions. *Journal of Business & Economic Statistics 39*(1), 338–357.

Giglio, S., B. Kelly, and S. Pruitt (2016). Systemic risk and the macroeconomy: An empirical evaluation. *Journal of Financial Economics 119*(3), 457–471.

Newey, W. and D. MacFadden (1994). Large sample estimation and hypothesis testing, chapter 36. *Handbook of Econometrics Vol 4*.

Table D1: Forecasting performance from simulated dataset

DGP	Fixed1	Fixed2	Opt-K_1	Opt-K_2	Opt-K_3	Fixed1	Fixed2	Opt-K_1	Opt-K_2	Opt-K_3
			$h = 1$					$h = 12$		
					T=150					
D1	1.035	1.022	1.049	1.019	1.058	0.889	0.871	0.917	0.887	0.926
D2	0.919	0.909	0.944	0.927	0.947	0.909	0.902	0.942	0.917	0.947
D3	0.829	0.820	0.852	0.849	0.856	0.710	0.703	0.738	0.723	0.741
D4	0.838	0.899	0.834	0.874	0.835	0.710	0.718	0.719	0.718	0.720
D5	1.026	1.012	1.043	1.013	1.053	0.882	0.868	0.913	0.882	0.926
D6	1.017	1.005	1.035	1.006	1.045	0.829	0.820	0.860	0.838	0.870
D7	1.007	1.002	1.020	0.998	1.027	0.830	0.825	0.850	0.832	0.858
					T=300					
D1	1.034	1.023	1.030	1.015	1.036	0.914	0.894	0.916	0.900	0.926
D2	0.923	0.912	0.928	0.918	0.929	0.907	0.895	0.916	0.901	0.920
D3	0.826	0.816	0.831	0.826	0.834	0.723	0.716	0.732	0.725	0.734
D4	0.830	0.822	0.829	0.833	0.830	0.721	0.710	0.727	0.722	0.730
D5	1.032	1.019	1.027	1.012	1.031	0.907	0.888	0.912	0.896	0.920
D6	1.022	1.006	1.018	1.003	1.024	0.857	0.847	0.865	0.853	0.870
D7	1.021	1.012	1.020	1.004	1.022	0.852	0.839	0.856	0.846	0.863
					T=450					
D1	1.032	1.023	1.025	1.012	1.027	0.906	0.887	0.895	0.887	0.901
D2	0.911	0.902	0.908	0.902	0.907	0.882	0.869	0.878	0.871	0.881
D3	0.845	0.837	0.842	0.838	0.844	0.728	0.718	0.728	0.723	0.731
D4	0.845	0.837	0.840	0.838	0.841	0.725	0.716	0.722	0.716	0.724
D5	1.031	1.023	1.024	1.012	1.026	0.901	0.883	0.894	0.883	0.897
D6	1.029	1.015	1.017	1.007	1.021	0.867	0.849	0.860	0.853	0.865
D7	1.022	1.013	1.013	1.003	1.017	0.863	0.843	0.854	0.846	0.860
					T=600					
D1	1.037	1.025	1.023	1.013	1.025	0.921	0.899	0.901	0.892	0.907
D2	0.898	0.889	0.891	0.888	0.889	0.877	0.866	0.871	0.864	0.871
D3	0.864	0.852	0.857	0.854	0.858	0.730	0.720	0.725	0.721	0.727
D4	0.855	0.847	0.847	0.842	0.848	0.727	0.714	0.717	0.713	0.719
D5	1.030	1.020	1.018	1.009	1.020	0.903	0.882	0.886	0.877	0.891
D6	1.024	1.016	1.013	1.005	1.015	0.869	0.848	0.856	0.849	0.858
D7	1.019	1.010	1.009	1.000	1.009	0.876	0.858	0.864	0.858	0.866

Notes: Fixed1: rolling window estimator with window size equal to 40; Fixed2: rolling window estimator with window size equal to 60; Opt-K_i: local estimator with optimal bandwidth selection, where $K_1(u) = \mathbb{1}_{\{-1<u<0\}}$, $K_2(u) = \frac{2}{\sqrt{2\pi}} \exp\big(-\frac{u^2}{2}\big)\mathbb{1}_{\{u<0\}}$ and $K_3(u) = \frac{3}{2}(1-u^2)\mathbb{1}_{\{-1<u<0\}}$.