Supplementary Material: Confidence intervals for multiple change points in linear models with heteroscedastic errors

Lajos Horváth*, Gregory Rice†, Yuqian Zhao‡

Appendix A. Proof of the main results

Lemma A.1. *If Assumption 2.5 holds, then for all* $1 \le \ell \le M+1$, $\mathbf{x}_i\epsilon_i, \mathbf{x}_i\mathbf{x}_i^\top, m_{\ell-1} < i \le m_\ell$ *are Bernoulli shifts and with some* $c > 0$

$$\left(E\left\|\mathbf{x}_i\epsilon_i - \mathbf{x}^*_{i,j}\epsilon^*_{i,j}\right\|^{\nu/2}\right)^{2/\nu} \le cj^{-\vartheta},$$

$$\left(E\left\|\mathbf{x}_i\mathbf{x}_i^\top - \mathbf{x}^*_{i,j}(\mathbf{x}^*_{i,j})^\top\right\|^{\nu/2}\right)^{2/\nu} \le cj^{-\vartheta},$$

where ν *and* ϑ *are defined in Assumption 2.5 and* $\mathbf{x}^*_{i,j}, \epsilon^*_{i,j}$ *are the first* d *and the last coordinates of* $\mathbf{z}^*_{i,j}$.
Also, for all $-\infty < m < \infty$

$$\left(E\left\|\mathbf{x}_i\mathbf{x}_i^\top\epsilon_{i+m} - \mathbf{x}^*_{i,j}(\mathbf{x}^*_{i,j})^\top\epsilon^*_{i+m,j}\right\|^{\nu/3}\right)^{3/\nu} \le cj^{-\vartheta}$$

and

$$\left(E\left\|\mathbf{x}_i\mathbf{x}_i^\top\mathbf{x}_{i+m}\mathbf{x}_{i+m}^\top - \mathbf{x}^*_{i,j}(\mathbf{x}^*_{i,j})^\top\mathbf{x}^*_{i+m,j}(\mathbf{x}^*_{i+m,j})^\top\right\|^{\nu/4}\right)^{4/\nu} \le cj^{-\vartheta}.$$

Proof. The Bernoulli shift property is an immediate consequence of Assumption 2.5. It is easy to see that

$$\mathbf{x}_i\epsilon_i - \mathbf{x}^*_{i,j}\epsilon^*_{i,j} = (\mathbf{x}_i - \mathbf{x}^*_{i,j})\epsilon_i + \mathbf{x}^*_{i,j}(\epsilon_i - \epsilon^*_{i,j})$$

and therefore by the Cauchy–Schwartz inequality we have

$$\begin{aligned}
E\left\|\mathbf{x}_i\epsilon_i - \mathbf{x}^*_{i,j}\epsilon^*_{i,j}\right\|^{\nu/2} &\le 2^{\nu/2}\left(E\left[\|\mathbf{x}_i - \mathbf{x}^*_{i,j}\||\epsilon_i|\right]^{\nu/2} + E\left[\|\mathbf{x}^*_{i,j}\||\epsilon_i - \epsilon^*_{i,j}|\right]^{\nu/2}\right)\\
&\le 2^{\nu/2}\left(\left[E\|\mathbf{x}_i - \mathbf{x}^*_{i,j}\|^\nu\right]^{1/2}\left[E|\epsilon_i|^\nu\right]^{1/2} + \left[E\|\mathbf{x}^*_{i,j}\|^\nu\right]^{1/2}\left[E|\epsilon_i - \epsilon^*_{i,j}|^\nu\right]^{1/2}\right)\\
&\le c_1 j^{-\vartheta\nu/2},
\end{aligned}$$

where in the last step we used Assumption 2.5.
The proofs of the remaining parts are the same and hence they are omitted. □

We start with the behavior of $\hat{\boldsymbol{\beta}}_N$ under the exactly one change in the regression parameter hypothesis. We show in the next lemma that $\hat{\boldsymbol{\beta}}_N$ is close to

$$\bar{\boldsymbol{\beta}}_N = \theta\boldsymbol{\beta}_1 + (1-\theta)\boldsymbol{\beta}_2.$$

*Department of Mathematics, University of Utah, USA.
†Department of Statistics and Actuarial Science, University of Waterloo, Canada.
‡University of Sussex Business School, University of Sussex, UK.

Lemma A.2. *If Assumptions 2.1–2.7 hold and $R=1$, then*

$$\left\|\hat{\boldsymbol{\beta}}_N-\bar{\boldsymbol{\beta}}_N\right\|=O_P\left(N^{-1/2}\right).$$

Proof. Using Theorem S2.1 in Aue et al. (2014) we obtain

$$\left\|\frac{1}{N}\mathbf{X}_N^\top\mathbf{X}_N-\mathbf{A}\right\|=O_P\left(N^{-1/2}\right).$$

Under Assumption 2.6, $\mathbf{A}$ is a non-singular matrix and therefore

$$\left\|\left(\frac{1}{N}\mathbf{X}_N^\top\mathbf{X}_N\right)^{-1}-\mathbf{A}^{-1}\right\|=O_P\left(N^{-1/2}\right).$$

Since

$$\mathbf{X}_N^\top\mathbf{Y}_N=\sum_{i=1}^{k^*}\mathbf{x}_i\mathbf{x}_i^\top\boldsymbol{\beta}_1+\sum_{i=k^*+1}^{N}\mathbf{x}_i\mathbf{x}_i^\top\boldsymbol{\beta}_2+\sum_{j=1}^{M+1}\sum_{i=m_{j-1}+1}^{m_j}\mathbf{x}_i\epsilon_i,$$

applying again Theorem S2.1 of Aue et al. (2014) (cf. also Berkes et al., 2014) on the intervals of stationarity, we conclude

$$\left\|\sum_{i=m_{j-1}+1}^{m_j}\mathbf{x}_i\epsilon_i\right\|=O_P(N^{1/2}),\quad 1\le j\le M+1,$$

which results in

$$\left\|\frac{1}{N}\mathbf{X}_N^\top\mathbf{Y}_N-\mathbf{A}\bar{\boldsymbol{\beta}}_N\right\|=O_P\left(N^{-1/2}\right).$$

The proof is now complete. □

Lemma A.3. *If Assumptions 2.1–2.7 hold and $R=1$, then*

$$\left|\frac{\hat{k}_N}{N}-\theta\right|=o_P(1).$$

Proof. First we consider the case when $0\le\kappa<1/2$. Using the definition of $\hat{\epsilon}_i$ we get

$$\begin{aligned}\sum_{i=1}^{k}\mathbf{x}_i\hat{\epsilon}_i&=\sum_{i=1}^{k}\mathbf{x}_i\hat{\epsilon}_i-\frac{k}{N}\sum_{i=1}^{N}\mathbf{x}_i\hat{\epsilon}_i\\&=\sum_{i=1}^{k}\mathbf{x}_i\epsilon_i-\frac{k}{N}\sum_{i=1}^{N}\mathbf{x}_i\epsilon_i-\left(\sum_{i=1}^{k}\mathbf{x}_i\mathbf{x}_i^\top-\frac{k}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i^\top\right)(\hat{\boldsymbol{\beta}}_N-\bar{\boldsymbol{\beta}}_N)+\mathbf{G}_N(k),\end{aligned}$$

where

$$\mathbf{G}_N(k) = \begin{cases} \left((1-\theta)\left[\sum_{i=1}^{k}\mathbf{x}_i\mathbf{x}_i^\top - \frac{k}{N}\sum_{i=1}^{k^*}\mathbf{x}_i\mathbf{x}_i^\top\right] + \theta\frac{k}{N}\sum_{i=k^*+1}^{N}\mathbf{x}_i\mathbf{x}_i^\top\right)\mathbf{\Delta}_N, & \text{if} \quad 1 \le k \le k^*, \\ \left((1-\theta)\left[1-\frac{k}{N}\right]\sum_{i=1}^{k^*}\mathbf{x}_i\mathbf{x}_i^\top - \theta\left[\sum_{i=k^*+1}^{k}\mathbf{x}_i\mathbf{x}_i^\top - \frac{k}{N}\sum_{i=k^*+1}^{N}\mathbf{x}_i\mathbf{x}_i^\top\right]\right)\mathbf{\Delta}_N, & \\ \qquad\qquad \text{if} \quad k^*+1 \le k \le N. \end{cases}$$

It is shown in Horváth et al. (2023, 2025) that

$$\max_{1\le k<N} \frac{N^{2\kappa-1/2}}{(k(N-k))^\kappa}\left\|\sum_{i=1}^{k}\mathbf{x}_i\epsilon_i - \frac{k}{N}\sum_{i=1}^{N}\mathbf{x}_i\epsilon_i\right\| = O_P(1) \tag{A.1}$$

and

$$\max_{1\le k<N} \frac{N^{2\kappa-1/2}}{(k(N-k))^\kappa}\left\|\sum_{i=1}^{k}\mathbf{x}_i\mathbf{x}_i^\top - \frac{k}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i^\top\right\| = O_P(1). \tag{A.2}$$

Hence

$$\max_{1\le k<N} \frac{N^{2\kappa-1/2}}{(k(N-k))^\kappa}\left\|\sum_{i=1}^{k}\mathbf{x}_i\hat{\epsilon}_i - \mathbf{G}_N(k)\right\| = o_P(1).$$

Let

$$\bar{\mathbf{G}}_N(k) = \begin{cases} \frac{k(N-k^*)}{N}\mathbf{A}\mathbf{\Delta}_N, & \text{if} \quad 1 \le k \le k^* \\ \frac{k^*(N-k)}{N}\mathbf{A}\mathbf{\Delta}_N, & \text{if} \quad k^*+1 \le k \le N. \end{cases}$$

It follows from (A.2)

$$\max_{1\le k<N} \frac{N^{-1/2+2\kappa}}{(k(N-k))^\kappa}\|\mathbf{G}_N(k) - \bar{\mathbf{G}}_N(k)\| = O_P(\|\mathbf{\Delta}_N\|).$$

It is easy to see that

$$\max_{1\le k<N} \frac{N^{2\kappa-1/2}}{(k(N-k))^\kappa}\left\|\bar{\mathbf{G}}_N(k)\right\| = N(\theta(1-\theta))^{1-\kappa}\left\|\mathbf{A}\mathbf{\Delta}_N\right\|(1+o(1)) \to \infty,$$

so this is the dominating term. We observe that for every $\varepsilon > 0$

$$\max_{|k^*-k|\le N\varepsilon} \frac{N^{2\kappa-1/2}}{(k(N-k))^\kappa}\left\|\bar{\mathbf{G}}_N(k)\right\| = N^{1/2}(\theta)(1-\theta)^{1-\kappa}\|\mathbf{A}\mathbf{\Delta}_N\| \to \infty.$$

Since

$$\lim_{N\to\infty} \frac{1}{N^{1/2}(\theta)(1-\theta)^{1-\kappa}\|\mathbf{A}\mathbf{\Delta}_N\|}\max_{|k^*-k|\ge N\varepsilon} \frac{N^{2\kappa-1/2}}{(k(N-k))^\kappa}\left\|\bar{\mathbf{G}}_N(k)\right\| < 1,$$

the lemma is proven when $0 \leq \kappa < 1/2$. If $\kappa = 1/2$ we just need to replace (A.1) and (A.2) with

$$\max_{1\leq k<N} \frac{N^{1/2}}{(k(N-k))^{1/2}} \left\| \sum_{i=1}^{k} \mathbf{x}_i\epsilon_i - \frac{k}{N}\sum_{i=1}^{k}\mathbf{x}_i\epsilon_i \right\| = O_P\left((\log\log N)^{1/2}\right)$$

and

$$\max_{1\leq k<N} \frac{N^{1/2}}{(k(N-k))^{1/2}} \left\| \sum_{i=1}^{k} \mathbf{x}_i\mathbf{x}_i^\top - \frac{k}{N}\sum_{i=1}^{k}\mathbf{x}_i\mathbf{x}_i^\top \right\| = \left((\log\log N)^{1/2}\right).$$

□

Some technical results we use in the proofs are collected in the following lemma.

Lemma A.4. *If Assumptions 2.1–2.7 hold and $R = 1$, then*

$$\left\|\mathbf{X}_N^\top\mathbf{X}_N - N\mathbf{A}\right\| = O_P\left(N^{1/2}\right),$$

$$\max_{1\leq k\leq N}\left\|\sum_{i=1}^{k}\mathbf{x}_i\epsilon_i\right\| = O_P\left(N^{1/2}\right)$$

and

$$a^{1/2}\max_{1\leq k\leq k^*-a}\frac{1}{k^*-k}\left\|\sum_{i=k+1}^{k^*}\mathbf{x}_i\epsilon_i\right\| = O_P(1)$$

if $a = a(N), a(N)\to\infty$ and $a(N)/N \to 0$.

Proof. The results in the lemma are taken from Lemma 5 of Horváth et al. (2023). □

The next result provides the rate of convergence in Lemma A.3.

Lemma A.5. *If Assumptions 2.1–2.7 hold and $R = 1$, then*

$$\|\mathbf{\Delta}_N\|^2\left|\frac{\hat{k}_N}{N} - \theta\right| = O_P(1).$$

Proof. We use the decomposition

$$\sum_{i=1}^{k}\mathbf{x}_i\hat{\epsilon}_i - \frac{k}{N}\sum_{i=1}^{N}\mathbf{x}_i\hat{\epsilon}_i = \mathbf{Q}_1(k) + \ldots + \mathbf{Q}_4(k),$$

where

$$\mathbf{Q}_1(k) = \sum_{i=1}^{k}\mathbf{x}_i\epsilon_i - \frac{k}{N}\sum_{i=1}^{N}\mathbf{x}_i\epsilon_i,$$

$$\mathbf{Q}_2(k) = -\left(\sum_{i=1}^{k}\mathbf{x}_i\epsilon_i - \frac{k}{N}\sum_{i=1}^{N}\mathbf{x}_i\epsilon_i\right)(\hat{\boldsymbol{\beta}}_N - \bar{\boldsymbol{\beta}}_N),$$

$$\mathbf{Q}_3(k) = (1-\theta)\sum_{i=1}^{k}\mathbf{x}_i\mathbf{x}_i^\top(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2),$$

$$\mathbf{Q}_4(k) = -\frac{k}{N}\left[(1-\theta)\sum_{i=1}^{k^*}\mathbf{x}_i\mathbf{x}_i^\top - \theta\sum_{i=k^*+1}^{N}\mathbf{x}_i\mathbf{x}_i^\top\right](\boldsymbol{\beta}_1-\boldsymbol{\beta}_2).$$

Using the notation above we have

$$\begin{aligned}\left(\frac{N}{k(N-k)}\right)^{2\kappa}&\left\|\sum_{i=1}^{k}\mathbf{x}_i\hat{\epsilon}_i - \frac{k}{N}\sum_{i=1}^{N}\mathbf{x}_i\hat{\epsilon}_i\right\|^2 - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\left\|\sum_{i=1}^{k^*}\mathbf{x}_i\hat{\epsilon}_i - \frac{k^*}{N}\sum_{i=1}^{N}\mathbf{x}_i\hat{\epsilon}_i\right\|^2\\ &= \left(\frac{N}{k(N-k)}\right)^{2\kappa}\sum_{i,j=1}^{4}\mathbf{Q}_i^\top(k)\mathbf{Q}_j(k) - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\sum_{i,j=1}^{4}\mathbf{Q}_i^\top(k^*)\mathbf{Q}_j(k^*).\end{aligned}$$

Elementary arguments yield for all $0<\alpha<\theta$

$$\max_{\lfloor N\alpha\rfloor\le k<k^*}\frac{1}{k^*-k}\left|\left(\frac{N}{k(N-k)}\right)^{2\kappa} - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\right| = O\left(N^{-1-2\kappa}\right) \tag{A.3}$$

and therefore by Lemma A.4

$$\begin{aligned}\max_{\lfloor N\alpha\rfloor\le k\le k^*-a}&\frac{1}{N^{1-2\kappa}\|\boldsymbol{\Delta}_N\|^2(k^*-k)}\left|\left(\frac{N}{k(N-k)}\right)^{2\kappa} - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\right|\|\mathbf{Q}_1(k)\|^2\\ &= O(1)\frac{1}{N\|\boldsymbol{\Delta}_N\|^2}\max_{1\le k\le N}\left\|N^{-1/2}\mathbf{Q}_1(k)\right\|^2\\ &= O_P\left(\frac{1}{N\|\boldsymbol{\Delta}_N\|^2}\right)\\ &= o_P(1),\end{aligned}$$

where

$$a = C/\|\boldsymbol{\Delta}_N\|^2.$$

Next we write

$$\left|\|\mathbf{Q}_1(k)\|^2 - \|\mathbf{Q}_1(k^*)\|^2\right| \le \left|\|\mathbf{Q}_1(k)\| - \|\mathbf{Q}_1(k^*)\|\right|\left(\|\mathbf{Q}_1(k)\| + \|\mathbf{Q}_1(k^*)\|\right).$$

Using again Lemma A.4 we conclude

$$\begin{aligned}&\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\max_{\lfloor N\alpha\rfloor\le k\le k^*-a}\frac{1}{N^{1-2\kappa}\|\boldsymbol{\Delta}_N\|^2(k^*-k)}\left|\|\mathbf{Q}_1(k)\|^2 - \|\mathbf{Q}_1(k^*)\|^2\right|\\ &\le\left[a^{1/2}\max_{1\le k\le k^*-a}\frac{1}{k^*-k}\left|\|\mathbf{Q}_1(k)\| - \|\mathbf{Q}_1(k^*)\|\right|\right]\left[N^{-1/2}\left(\max_{1\le k\le N}\|\mathbf{Q}_1(k)\| + \|\mathbf{Q}_1(k^*)\|\right)\right]\\ &\qquad\times\frac{N^{1/2}a^{1/2}}{N\|\boldsymbol{\Delta}_N\|^2} = o_P(1).\end{aligned}$$

Similar arguments give

$$\max_{\lfloor N\alpha\rfloor\le k\le k^*-a}\frac{N^{2\kappa-1}}{(k^*-k)\|\Delta_N\|^2}\left|\left(\frac{N}{k(N-k)}\right)^{2\kappa}\mathbf{Q}_i^\top(k)\mathbf{Q}_j(k) - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\mathbf{Q}_i^\top(k^*)\mathbf{Q}_j(k^*)\right|$$

$$= o_P(1),$$

for $i, j = 1, 2$. Lemma A.4 implies

$$\left\|(1-\theta)\sum_{i=1}^{k^*}\mathbf{x}_i\mathbf{x}_i^\top - \theta\sum_{i=k^*+1}^{N}\mathbf{x}_i\mathbf{x}_i^\top\right\| \tag{A.4}$$
$$\leq N\left|\frac{k^*}{N}-\theta\right|\|\mathbf{A}\| + \left\|\sum_{i=1}^{k^*}[\mathbf{x}_i\mathbf{x}_i^\top - \mathbf{A}]\right\| + \left\|\sum_{i=k^*+1}^{N}[\mathbf{x}_i\mathbf{x}_i^\top - \mathbf{A}]\right\|$$
$$= O_P\left(N^{1/2}\right).$$

Elementary algebra yields

$$\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}\left|\left(\frac{N}{k(N-k)}\right)^{2\kappa}\mathbf{Q}_4^\top(k)\mathbf{Q}_4(k) - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\mathbf{Q}_4^\top(k^*)\mathbf{Q}_4(k^*)\right|$$
$$\leq \max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}\left|\left(\frac{N}{k(N-k)}\right)^{2\kappa} - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\right|\|\mathbf{Q}_4(k)\|^2$$
$$+ \max_{1\leq k\leq k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\left|\mathbf{Q}_4^\top(k)\mathbf{Q}_4(k) - \mathbf{Q}_4^\top(k^*)\mathbf{Q}_4(k^*)\right|$$

and by (A.3) and (A.4)

$$\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}\left|\left(\frac{N}{k(N-k)}\right)^{2\kappa} - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\right|\|\mathbf{Q}_4(k)\|^2$$
$$= O_P(1)\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{(k^*-k)N^{-1-2\kappa}}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}\max_{1\leq k\leq k^*}\|\mathbf{Q}_4(k)\|^2$$
$$= O_P(1)\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{(k^*-k)N^{-1-2\kappa}}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}N\|\mathbf{\Delta}_N\|^2$$
$$= o_P(1).$$

Also, by (A.4)

$$\max_{1\leq k\leq k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\left|\mathbf{Q}_4^\top(k)\mathbf{Q}_4(k) - \mathbf{Q}_4^\top(k^*)\mathbf{Q}_4(k^*)\right|$$
$$\leq \max_{1\leq k\leq k^*-a}\frac{1}{N(k^*-k)\|\mathbf{\Delta}_N\|^2}\|\mathbf{Q}_4(k)-\mathbf{Q}_4(k^*)\|\max_{1\leq k\leq k^*}[\|\mathbf{Q}_4(k)\| + \|\mathbf{Q}_4(k^*)\|]$$
$$= O_P(1)\max_{1\leq k\leq k^*-a}\frac{1}{N^{1/2}(k^*-k)\|\mathbf{\Delta}_N\|}\|\mathbf{Q}_4(k)-\mathbf{Q}_4(k^*)\|$$
$$= o_P(1).$$

Repeating our previous arguments one can verify

$$\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}\Bigg|\left(\frac{N}{k(N-k)}\right)^{2\kappa}\mathbf{Q}_4^\top(k)\mathbf{Q}_i^\top(k)$$
$$-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\mathbf{Q}_4^\top(k^*)\mathbf{Q}_i^\top(k^*)\Bigg|=o_P(1),\quad i=1,2,3$$

and

$$\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}\Bigg|\left(\frac{N}{k(N-k)}\right)^{2\kappa}\mathbf{Q}_3^\top(k)\mathbf{Q}_2^\top(k)$$
$$-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\mathbf{Q}_3^\top(k^*)\mathbf{Q}_2^\top(k^*)\Bigg|=o_P(1).$$

Using again (A.3) we obtain

$$\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}\left|\left(\frac{N}{k(N-k)}\right)^{2\kappa}-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\right|\|\mathbf{Q}_1(k)\|\|\mathbf{Q}_3(k)\|$$
$$=O_P\left(\frac{1}{N^{1/2}\|\mathbf{\Delta}_N\|}\right)=o_P(1).$$

Next we note

$$\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}\left|\mathbf{Q}_1^\top(k)\mathbf{Q}_3(k)-\mathbf{Q}_1^\top(k^*)\mathbf{Q}_3(k^*)\right|$$
$$\leq\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{N(k^*-k)\|\mathbf{\Delta}_N\|^2}\|\mathbf{Q}_1(k)-\mathbf{Q}_1(k^*)\|\max_{1\leq k\leq k^*}\|\mathbf{Q}_3(k)\|$$
$$+\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{N(k^*-k)\|\mathbf{\Delta}_N\|^2}\|\mathbf{Q}_3(k)-\mathbf{Q}_3(k^*)\|\|\mathbf{Q}_1(k^*)\|$$
$$=O_P(1)\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{(k^*-k)\|\mathbf{\Delta}_N\|}\|\mathbf{Q}_1(k)-\mathbf{Q}_1(k^*)\|$$
$$+O_P(1)\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{N^{1/2}(k^*-k)\|\mathbf{\Delta}_N\|^2}\|\mathbf{Q}_3(k)-\mathbf{Q}_3(k^*)\|.$$

By the definition of $\mathbf{Q}_1(k)$ we have

$$\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{(k^*-k)\|\mathbf{\Delta}_N\|}\|\mathbf{Q}_1(k)-\mathbf{Q}_1(k^*)\|$$
$$\leq\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{(k^*-k)\|\mathbf{\Delta}_N\|}\frac{k^*-k}{N}\left\|\sum_{i=1}^N\mathbf{x}_i\epsilon_i\right\|+\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{(k^*-k)\|\mathbf{\Delta}_N\|}\left\|\sum_{i=k+1}^{k^*}\mathbf{x}_i\epsilon_i\right\|$$

and

$$\max_{\lfloor N\alpha\rfloor\leq k\leq k^*-a}\frac{1}{(k^*-k)\|\mathbf{\Delta}_N\|}\frac{k^*-k}{N}\left\|\sum_{i=1}^N\mathbf{x}_i\epsilon_i\right\|=O_P\left(\frac{1}{N^{1/2}\|\mathbf{\Delta}_N\|}\right)=o_P(1).$$

Using again Lemma A.4 and the definition of a we conclude

$$\lim_{C\to\infty}\limsup_{N\to\infty} P\left\{\frac{1}{\|\mathbf{\Delta}_N\|}\max_{1\le k\le k^*-a}\frac{1}{k^*-k}\left\|\sum_{i=k+1}^{k^*}\mathbf{x}_i\epsilon_i\right\|>x\right\}=0,$$

for all $x>0$. It is easy to see that

$$\max_{\lfloor N\alpha\rfloor\le k\le k^*-a}\frac{1}{N^{1/2}(k^*-k)\|\mathbf{\Delta}_N\|^2}\|\mathbf{Q}_3(k)-\mathbf{Q}_3(k^*)\|=O_P\left(\frac{1}{N^{1/2}\|\mathbf{\Delta}_N\|}\right)=o_P(1).$$

Applying Lemma A.4 and (A.3) yields

$$\max_{\lfloor N\alpha\rfloor\le k\le k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathbf{\Delta}_N\|^2}\Bigg|\left(\frac{N}{k(N-k)}\right)^{2\kappa}\|\mathbf{Q}_3(k)\|^2-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\|\mathbf{Q}_3(k^*)\|^2$$
$$-\left(\left(\frac{N}{k(N-k)}\right)^{2\kappa}k^2-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}(k^*)^2\right)\mathbf{\Delta}_N^\top\mathbf{A}^2\mathbf{\Delta}_N\Bigg|=o_P(1).$$

Putting together our estimates we conclude

$$\max_{\lfloor N\alpha\rfloor\le k\le k^*-a}\left(U_N(k)+\frac{1}{2}V_N(k)\right)\xrightarrow{P}-\infty,$$

with

$$U_N(k)=\left(\frac{N}{k(N-k)}\right)^{2\kappa}\sum_{(i,j)\in\mathscr{T}}\mathbf{Q}_i^\top(k)\mathbf{Q}_j(k)-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\sum_{(i,j)\in\mathscr{T}}\mathbf{Q}_i^\top(k^*)\mathbf{Q}_j(k^*)\tag{A.5}$$

and

$$V_N(k)=\left(\frac{N}{k(N-k)}\right)^{2\kappa}\mathbf{Q}_3^\top(k)\mathbf{Q}_3(k)-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\mathbf{Q}_3^\top(k^*)\mathbf{Q}_3(k^*),$$

where

$$\mathscr{T}=\{(i,j):\ 1\le i,j\le 4\}\setminus\{(1,3),(3,1),(3,3)\}.$$

For the remaining term $\mathbf{Q}_1(k)\mathbf{Q}_3(k)$ we have

$$\lim_{C\to\infty}\limsup_{N\to\infty}P\Bigg\{\max_{\lfloor N\alpha\rfloor\le k\le k^*-C/\|\mathbf{\Delta}_N\|^2}\Bigg[2\Bigg(\left(\frac{N}{k(N-k)}\right)^{2\kappa}\mathbf{Q}_1^\top(k)\mathbf{Q}_3(k)$$
$$-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\mathbf{Q}_1^\top(k^*)\mathbf{Q}_3(k^*)\Bigg)+\frac{1}{2}V_N(k)\Bigg]>x\Bigg\}=0,$$

for all $x<0$. Thus we showed that

$$\lim_{C\to\infty}\limsup_{N\to\infty}P\left\{\hat{k}_N\le k^*-C/\|\mathbf{\Delta}_N\|^2\right\}=0$$

and by symmetry

$$\lim_{C\to\infty}\limsup_{N\to\infty}P\left\{\hat{k}_N\ge k^*+C/\|\mathbf{\Delta}_N\|^2\right\}=0.$$

□

We extend the definition of $\mathbf{Q}_3(k)$ for $k > k^*$

$$\mathbf{Q}_3(k) = \theta \sum_{i=k+1}^{N} \mathbf{x}_i \mathbf{x}_i^\top (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2), \quad \text{if } k > k^*.$$

We introduce further notations

$$R_{N,1}(k) = \begin{cases} N^{-(1-2\kappa)} \left(\left(\frac{N}{k(N-k)} \right)^{2\kappa} \sum_{i=1}^{k} \mathbf{x}_i^\top \epsilon_i \mathbf{Q}_3(k) - \left(\frac{N}{k^*(N-k^*)} \right)^{2\kappa} \sum_{i=1}^{k^*} \mathbf{x}_i^\top \epsilon_i \mathbf{Q}_3(k^*) \right), \\ \qquad \text{if } 1 \le k \le k^*, \\ N^{-(1-2\kappa)} \left(\left(\frac{N}{k(N-k)} \right)^{2\kappa} \left(- \sum_{i=k+1}^{N} \mathbf{x}_i^\top \epsilon_i \right) \mathbf{Q}_3(k) \right. \\ \qquad \left. - \left(\frac{N}{k^*(N-k^*)} \right)^{2\kappa} \left(- \sum_{i=k^*+1}^{N} \mathbf{x}_i^\top \epsilon_i \right) \mathbf{Q}_3(k^*) \right), \quad \text{if } k^*+1 \le k \le N, \end{cases}$$

$$R_{N,2}(k) = \begin{cases} N^{-(1-2\kappa)} \left(\left(\frac{N}{k(N-k)} \right)^{2\kappa} (1-\theta)^2 k^2 (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2)^\top \mathbf{A}^2 (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2) \right. \\ \qquad \left. - \left(\frac{N}{k^*(N-k^*)} \right)^{2\kappa} (1-\theta)^2 (k^*)^2 (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2)^\top \mathbf{A}^2 (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2) \right), \quad \text{if } 1 \le k \le k^*, \\ N^{-(1-2\kappa)} \left(\left(\frac{N}{k(N-k)} \right)^{2\kappa} \theta^2 (N-k)^2 (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2)^\top \mathbf{A}^2 (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2) \right. \\ \qquad \left. - \left(\frac{N}{k^*(N-k^*)} \right)^{2\kappa} \theta^2 (N-k^*)^2 (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2)^\top \mathbf{A}^2 (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2) \right), \quad \text{if } k^*+1 \le k \le N, \end{cases}$$

$$R_{N,3}(k) = \begin{cases} N^{-(1-2\kappa)} \left(\frac{N}{k^*(N-k^*)} \right)^{2\kappa} (1-\theta) \left(- \sum_{i=k+1}^{k^*} \mathbf{x}_i^\top \epsilon_i \right) \sum_{i=1}^{k^*} \mathbf{x}_i \mathbf{x}_i^\top (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2), \quad \text{if } 1 \le k \le k^*, \\ N^{-(1-2\kappa)} \left(\frac{N}{k^*(N-k^*)} \right)^{2\kappa} \theta \left(\sum_{i=k^*}^{k} \mathbf{x}_i^\top \epsilon_i \right) \sum_{i=k^*+1}^{N} \mathbf{x}_i \mathbf{x}_i^\top (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2), \quad \text{if } k^*+1 \le k \le N. \end{cases}$$

and

$$R_{N,4}(k) = \begin{cases} (\theta(1-\theta))^{1-2\kappa} \left(- \sum_{i=k+1}^{k^*} \mathbf{x}_i^\top \epsilon_i \right) \mathbf{A} (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2), \quad \text{if } 1 \le k \le k^*, \\ (\theta(1-\theta)^{1-2\kappa} \left(\sum_{i=k^*+1}^{k} \mathbf{x}_i^\top \epsilon_i \right) \mathbf{A} (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2), \quad \text{if } k^*+1 \le k \le N. \end{cases}$$

Lemma A.6. *If Assumptions 2.1–2.7 hold and $R = 1$, then we have for all $C > 0$*

$$\max_{|k^*-k|\le C/\|\boldsymbol{\Delta}_N\|^2}\left|N^{-(1-2\kappa)}\left[\left(\frac{N}{k(N-k)}\right)^{2\kappa}\mathbf{Q}_1^\top(k)\mathbf{Q}_3(k)-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\mathbf{Q}_1^\top(k^*)\mathbf{Q}_3(k^*)\right]\right.$$
$$\left.-R_{N,1}(k)\right| = o_P(1),$$

$$N^{-(1-2\kappa)}\max_{|k^*-k|\le C/\|\boldsymbol{\Delta}_N\|^2}\left|N^{-(1-2\kappa)}\left[\left(\frac{N}{k(N-k)}\right)^{2\kappa}\mathbf{Q}_3^\top(k)\mathbf{Q}_3(k)\right.\right.$$
$$\left.\left.-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\mathbf{Q}_3^\top(k^*)\mathbf{Q}_3(k^*)\right]-R_{N,2}(k)\right| = o_P(1),$$

$$\sup_{|k-k^*|\le C/\|\boldsymbol{\Delta}_N\|^2}|R_{N,1}(k)-R_{N,3}(k)| = o_P(1),$$
$$\sup_{|k-k^*|\le C/\|\boldsymbol{\Delta}_N\|^2}|R_{N,3}(k)-R_{N,4}(k)| = o_P(1),$$

and

$$N^{-(1-2\kappa)}\max_{|k-k^*|\le C/\|\boldsymbol{\Delta}_N\|^2}|U_N(k)| = o_P(1),$$

where $U_N(k)$ is defined in (A.5).

Proof. It goes along the lines of the proof of Lemma A.5 and therefore it is omitted. □

Lemma A.7. *If Assumptions 2.1–2.7 hold and $R = 1$, then*

$$\sup_{|s|\le C}\left|R_{N,2}(k^*+s\gamma/\|\boldsymbol{\Delta}_N\|^2)+2\gamma(\theta(1-\theta))^{1-2\kappa}|s|m_\kappa(s)\|\mathbf{A}\mathbf{d}\|^2\right| = o(1).$$

Proof. Elementary calculations yield the result. □

Let

$$Z_{N,1}(s) = \|\boldsymbol{\Delta}_N\|\left(-\sum_{i=k^*-s/\|\boldsymbol{\Delta}_N\|^2}^{k^*}\mathbf{x}_i^\top\epsilon_i\right)\mathbf{A}\frac{\boldsymbol{\Delta}_N}{\|\boldsymbol{\Delta}_N\|},\qquad s\ge 0, \tag{A.6}$$

and

$$Z_{N,2}(s) = \|\boldsymbol{\Delta}_N\|\left(\sum_{i=k^*+1}^{k^*+s/\|\boldsymbol{\Delta}_N\|^2}\mathbf{x}_i^\top\epsilon_i\right)\mathbf{A}\frac{\boldsymbol{\Delta}_N}{\|\boldsymbol{\Delta}_N\|},\qquad s\ge 0. \tag{A.7}$$

We recall the definition of $\sigma^2(i)$ from equation (2.6).

Lemma A.8. *We assume that Assumptions 2.1–2.7 hold and $R = 1$.*
(i) If $\tau_{i-1} < \theta < \tau_i$, then

$$\{Z_{N,1}(s), Z_{N,2}(s), 0 \le s \le C\} \xrightarrow{\mathscr{D}^2[0,C]} \{\sigma(i)W^{(1)}(s), \sigma(i)W^{(2)}(s), 0 \le s \le C\}.$$

(ii) If $\theta = \tau_i$, then

$$\{Z_{N,1}(s), Z_{N,2}(s), 0 \le s \le C\} \xrightarrow{\mathscr{D}^2[0,C]} \{\sigma(i)W^{(1)}(s), \sigma(i+1)W^{(2)}(s), 0 \le s \le C\},$$

where $\{W^{(1)}(s), s \ge 0\}$ and $\{W^{(2)}(s), s \ge 0\}$ are independent Wiener processes.

Proof. Due to Assumption 2.5, $\mathbf{x}_i\epsilon_i$ is stationary on the intervals $(m_{k-1}, m_k], 1 \le k \le M+1$. Hence, the asymptotic independence as well as the weak convergence follows from Theorem S2.1 of Aue et al. (2014). □

Proof of Theorem 2.1.

Proof. According to Lemmas A.5 and A.6 it is enough to consider

$$Z_N(s) = 2R_{N,4}(k^* + \varsigma s/\|\mathbf{\Delta}_N\|^2) + R_{N,2}(k^* + \varsigma s/\|\mathbf{\Delta}_N\|^2),\ -C \le s \le C,$$

for all $C > 0$. Assume that $\tau_{i-1} < \theta < \tau_i$. It follows from Lemmas A.7 and A.8 that

$$Z_N(s) \xrightarrow{\mathscr{D}[-C,C]} 2(\theta(1-\theta))^{1-2\kappa}\left(\varsigma^{1/2}\sigma(i)W(s) - \varsigma|s|m_\kappa(s)\|\mathbf{Ad}\|^2\right). \tag{A.8}$$

Choosing

$$\varsigma = \left(\frac{a(\theta)+a(\theta+)}{2}\right)^2 \frac{1}{\|\mathbf{Ad}\|^4}$$

we get

$$\left(\varsigma^{1/2}\sigma(i)W(s) - \varsigma|s|m_\kappa(s)\|\mathbf{Ad}\|^2\right) = \left(\frac{a(\theta)+a(\theta+)}{2}\right)^2 \frac{1}{\|\mathbf{Ad}\|^2}(\sigma(i)W(s) - |s|m_\kappa(s)). \tag{A.9}$$

We can choose C as large as we wish, the proof is complete when $\tau_{i-1} < \theta < \tau_i$.
If $\theta = \tau_i$, then (A.8) and (A.9) change to

$$Z_N(s) \xrightarrow{\mathscr{D}[-C,C]} 2(\theta(1-\theta))^{1-2\kappa}\Bigg(\varsigma^{1/2}\left[\sigma(i)W(s)\mathbb{1}\{s \le 0\} + \sigma(i+1)W(s)\mathbb{1}\{s > 0\}\right] - \varsigma|s|m_\kappa(s)\|\mathbf{Ad}\|^2\Bigg).$$

and

$$\left(\varsigma^{1/2}\sigma(i)W(s) - \varsigma|s|m_\kappa(s)\|\mathbf{Ad}\|^2\right) = \left(\frac{a(\theta)+a(\theta+)}{2}\right)^2 \frac{1}{\|\mathbf{Ad}\|^2}\Bigg(\sigma(i)W(s)\mathbb{1}\{s \le 0\} \tag{A.10}$$

$$+ \sigma(i+1)W(s)\{s>0\} - |s|m_\kappa(s)\Bigg).$$

□

Proof of Theorem 2.3.

Lemma A.9. *If Assumptions 2.1–2.6 hold, then*

$$\lim_{\varepsilon\to 0}\limsup_{N\to\infty} P\left\{ N^{\kappa-1/2}\max_{|k-k_i|\le N\varepsilon}\max_{k_i-N\varepsilon\le j<k}\frac{1}{(k-j)^\kappa}\left\|\sum_{\ell=1}^{k}\mathbf{x}_\ell\epsilon_\ell - \sum_{\ell=1}^{j}\mathbf{x}_\ell\epsilon_\ell\right\| > x\right\} = 0,$$

for all $x>0$ and

$$\max_{|k-k_i|\le N\varepsilon}\max_{k_i-N\varepsilon\le j<k}\frac{1}{(k-j)^{1/2}}\left\|\sum_{\ell=1}^{k}\mathbf{x}_\ell\epsilon_\ell - \sum_{\ell=1}^{j}\mathbf{x}_\ell\epsilon_\ell\right\| = o_P\left(N^{1/\nu}\right).$$

Proof. Assume that $0\le\kappa<1/2$. We prove that for all $x>0$ and $1\le\hbar\le d$

$$\lim_{\varepsilon\to 0}\limsup_{N\to\infty} P\left\{ N^{\kappa-1/2}\max_{|k-k_i|\le N\varepsilon}\max_{k_i-N\varepsilon\le j<k}\frac{1}{(k-j)^\kappa}\left|\sum_{\ell=1}^{k}x_{\ell,\hbar}\epsilon_\ell - \sum_{\ell=1}^{j}x_{\ell,\hbar}\epsilon_\ell\right| > x\right\} = 0.$$

Consider first when $0\le\kappa<1/2$. It follows from Theorem 1 of Berkes et al. (2011) that for all $0<\epsilon<\min_{1\le i\le R}(\theta_i-\theta_{i+1})$

$$\begin{aligned} &N^{\kappa-1/2}\max_{|k-k_i|\le N\epsilon}\max_{k_i-N\epsilon\le j<k}\frac{1}{(k-j)^\kappa}\left|\sum_{\ell=1}^{k}x_{\ell,\hbar}\epsilon_\ell - \sum_{\ell=1}^{j}x_{\ell,\hbar}\epsilon_\ell\right| \\ &\qquad \xrightarrow{\mathscr{D}} \sup_{|t-\theta_i|\le\epsilon}\sup_{\theta_i-\epsilon\le s<t}\frac{\sigma}{(t-s)^\kappa}|W(t)-W(s)|. \end{aligned}$$

The random variable on the right hand side of the above is well defined by Garsia's modulus of continuity result (cf. Csörgő and Horváth, 1997 and Mörters and Peres, 2010). When $\kappa=1/2$, for every $\epsilon>0$, $0<\delta<1/2$,

$$\begin{aligned} &\max_{|k-k_i|\le N\epsilon}\max_{k_i-N\epsilon\le j<k}\frac{1}{(k-j)^{1/2}}\left|\sum_{\ell=1}^{k}x_\ell\epsilon_\ell - \sum_{\ell=1}^{j}x_{\ell,\hbar}\epsilon_\ell\right| \\ &\quad\le \max_{|k-k_i|\le N\epsilon}\max_{k_i-N\epsilon\le j<k}(k-j)^{-\delta}\frac{1}{(k-j)^{1/2-\delta}}\left|\sum_{\ell=1}^{k}x_{\ell,\hbar}\epsilon_\ell - \sum_{\ell=1}^{j}x_{\ell,\hbar}\epsilon_\ell\right| \\ &\quad\le \max_{|k-k_i|\le N\epsilon}\max_{k_i-N\epsilon j<k}\frac{1}{(k-j)^{1/2-\delta}}\left|\sum_{\ell=1}^{k}x_{\ell,\hbar}\epsilon_\ell - \sum_{\ell=1}^{j}x_{\ell,\hbar}\epsilon_\ell\right| \\ &\quad= O_P\left(N^{1/2-(1/2-\delta)}\right) \\ &\quad= O_P\left(N^{\delta}\right), \end{aligned}$$

where we have applied the first part of the lemma. Taking $\delta=1/\nu$ finishes the proof. □

We define the event in Assumption 2.3

$$\mathscr{A}_N = \{|\hat{\theta}_r - \theta_r| \le \varepsilon \ \text{ for all } \ 1 \le r \le R\} \cap \{\hat{R}_N = R\}.$$

Let

$$\mathbf{d}_i = \lim_{N\to\infty} \frac{\mathbf{\Delta}_{i,N}}{\|\mathbf{\Delta}_{i,N}\|}.$$

We introduce

$$\bar{a}(\theta_i) = \mathbf{d}_i^\top \mathbf{A}\mathbf{D}_j\mathbf{A}\mathbf{d}_i, \text{ if } \tau_j < \theta_i \le \tau_{j+1}, j \in \{1,\dots,M\}, i \in \{1,\dots,R\}.$$

Appendix B. Proof of Theorem 2.2 ($R > 1$)

We recall the notation describing the disjoint (topological) union space of $\mathbb{R}^k$,

$$\mathbb{U} \ := \ \bigcup_{k=0}^{\infty} \mathbb{R}^k \ = \ \{(k,\mathbf{x}) : k \in \mathbb{N}\cup\{0\}, \ \mathbf{x}\in\mathbb{R}^k\},$$

equipped with the metric

$$d_{\mathbb{U}}((k_1,\mathbf{x}_1),(k_2,\mathbf{x}_2)) = \mathbb{1}\{k_1 \ne k_2\} + \frac{\|\mathbf{x}_1-\mathbf{x}_2\|}{1+\|\mathbf{x}_1-\mathbf{x}_2\|}\mathbb{1}\{k_1 = k_2\}.$$

$\mathbb{U}$ is a Polish space; see Chapter 1 of Kechris (1995). The random element $(\hat{\mathfrak{k}}_{N,1},\dots,\hat{\mathfrak{k}}_{N,\hat{R}_N}) := (\hat{R}_N, (\hat{\mathfrak{k}}_{N,1},\dots,\hat{\mathfrak{k}}_{N,\hat{R}_N})^\top)$ takes value in $\mathbb{U}$. We first begin by noting an elementary property of weak convergence in $\mathbb{U}$.

Lemma B.1. *Consider an array of random variables $X_{i,N}$ such that for a fixed positive integer R, (a) $(X_{1,N},...,X_{R,N}) \xrightarrow{\mathscr{D}(\mathbb{R}^R)} (X_1,...,X_R)$, and (b) $\hat{R}_N \xrightarrow{P} R$, as $N\to\infty$. Then*

$$(\hat{R}_N, X_{1,N},...,X_{\hat{R},N}) \xrightarrow{\mathscr{D}(\mathbb{U})} (R, X_{1,N},...,X_{R,N}).$$

Proof. Let

$$Z_N \ := \ \big(\hat{R}_N, \ (X_{1,N},\dots,X_{\hat{R}_N,N})\big) \in \mathbb{U}$$

and $Z = (R,(X_1,\dots,X_R))$. For any bounded continuous $f : \mathbb{U} \to \mathbb{R}$, by the triangle inequality

$$\Big|Ef(Z_N)-Ef(Z)\Big| \le \Big|Ef(Z_N)-Ef\big(R,(X_{1,N},\dots,X_{R,N})\big)\Big|+\Big|Ef(Z)-Ef\big(R,(X_{1,N},\dots,X_{R,N})\big)\Big|.$$

For the first term, $|Ef(Z_N) - Ef\big(R,(X_{1,N},\dots,X_{R,N})\big)| \ \le \ 2\|f\|_\infty \mathbb{P}(\hat{R}_N \ne R) \ \to \ 0.$ as $N\to\infty$. The second term vanishes as $N\to\infty$ by hypothesis (a). This completes the proof. □

Lemma B.1 in effect means that in order to prove Theorem 2.2 under Assumption 2.3, we only need to consider the case when $\hat{R}_N = R$, and establish the weak convergence of fixed dimensional vector $(\hat{\mathfrak{k}}_{N,1}, \ldots, \hat{\mathfrak{k}}_{N,R})$.

Henceforth we thus assume $\hat{R}_N = R$. Let for $i \in \{1, ..., R\}$

$$\zeta_i = \frac{k_i - \hat{k}_{i-1}}{\hat{k}_{i+1} - \hat{k}_{i-1}}, \quad \iota_{i,1} = \frac{k_{i-1} - \hat{k}_{i-1}}{\hat{k}_{i+1} - \hat{k}_{i-1}} \mathbb{1}\{\hat{k}_{i-1} < k_{i-1}\}, \quad \iota_{i,2} = \frac{k_{i+1} - \hat{k}_{i+1}}{\hat{k}_{i+1} - \hat{k}_{i-1}} \mathbb{1}\{k_{i+1} < \hat{k}_{i+1}\}$$

and

$$\bar{\boldsymbol{\beta}}_{i,N} = \zeta_i \boldsymbol{\beta}_i + (1 - \zeta_i)\boldsymbol{\beta}_{i+1} - \iota_{i,1}\boldsymbol{\Delta}_{N,i-1} + \iota_{i,2}\boldsymbol{\Delta}_{N,i+1}.$$

We recall that $\hat{\boldsymbol{\beta}}_{i,N} = \hat{\boldsymbol{\beta}}(\hat{k}_{i_1}, \hat{k}_{i-1})$ denotes the least squares estimator computed from $y_\ell, \mathbf{x}_\ell, \hat{k}_{i-1} < \ell \leq \hat{k}_{i+1}, 1 \leq i \leq M$.

Lemma B.2. *If Assumptions 2.1–2.8 hold, then we have*

$$\left\| \hat{\boldsymbol{\beta}}_{i,N} - \bar{\boldsymbol{\beta}}_{i,N} \right\| = O_P\left(N^{-1/2}\right).$$

Proof. By definition

$$\hat{\boldsymbol{\beta}}_{i,N} = \left(\mathbf{X}_{i,N}^\top \mathbf{X}_{i,N}\right)^{-1} \mathbf{X}_{i,N}\mathbf{Y}_{i,N},$$

where

$$\mathbf{X}_{i,N}^\top \mathbf{X}_{i,N} = \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top \quad \text{and} \quad \mathbf{Y}_{i,N} = \left(y_{\hat{k}_{i-1}+1}, \ldots, y_{\hat{k}_{i+1}}\right)^\top.$$

By Assumption 2.5 there is $c_0 > 0$ such that for all $c \leq c_0$

$$\lim_{N\to\infty} P\left\{ \left\| \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} (\mathbf{x}_\ell \mathbf{x}_\ell^\top - \mathbf{A}) \right\| \leq \max_{k_{i-1}-Nc \leq m < n \leq k_{i+1}+Nc} \left\| \sum_{\ell=m+1}^{n} (\mathbf{x}_\ell^\top \mathbf{x}_\ell - \mathbf{A}) \right\| \right\} = 1.$$

The sequence $\mathbf{x}_\ell \epsilon_\ell$ is stationary on intervals, so by Theorem S2.1 of Aue et al. (2014) we get

$$\frac{1}{N} \max_{k_{i-1}-Nc \leq m < n \leq k_{i+1}+Nc} \left\| \sum_{\ell=m+1}^{n} (\mathbf{x}_\ell^\top \mathbf{x}_\ell - \mathbf{A}) \right\| = O_P\left(N^{-1/2}\right).$$

Assumption 2.1 implies

$$\frac{N}{\hat{k}_{i+1} - \hat{k}_{i-1}} = O_P(1),$$

and therefore

$$\left\| \frac{1}{\hat{k}_{i+1} - \hat{k}_{i-1}} \mathbf{X}_{i,N}^\top \mathbf{X}_{i,N} - \mathbf{A} \right\| = O_P\left(N^{-1/2}\right),$$

hence

$$\left\| \left(\frac{1}{\hat{k}_{i+1} - \hat{k}_{i-1}} \mathbf{X}_{i,N}^\top \mathbf{X}_{i,N} \right)^{-1} - \mathbf{A}^{-1} \right\| = O_P\left(N^{-1/2}\right).$$

We note that

$$\begin{aligned}\mathbf{X}_{i,N}^\top \mathbf{Y}_{i,N} = &\sum_{\ell=\hat{k}_{i-1}+1}^{k_i} \mathbf{x}_\ell \mathbf{x}_\ell^\top \boldsymbol{\beta}_i - \left(\sum_{\ell=\hat{k}_{i-1}+1}^{k_i-1} \mathbf{x}_\ell \mathbf{x}_\ell^\top \boldsymbol{\Delta}_{i-1} \right) \mathbb{1}\{\hat{k}_{i-1} < k_{i-1}\} \\ &+ \sum_{\ell=k_i+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top \boldsymbol{\beta}_{i+1} + \left(\sum_{\ell=k_{i+1}+1}^{\hat{k}_{i+1}-1} \mathbf{x}_\ell \mathbf{x}_\ell^\top \boldsymbol{\Delta}_{i+1} \right) \mathbb{1}\{k_{i+1} < \hat{k}_{i+1}\} \\ &+ \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \epsilon_\ell .\end{aligned}$$

Using again Theorem S2.1 of Aue et al. (2014) for the piecewise stationary sequence $\{\mathbf{x}_\ell \epsilon_\ell, -\infty < \ell < \infty\}$, we get

$$\left\| \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \epsilon_\ell \right\| \leq \max_{0 \leq m < n \leq N} \left\| \sum_{\ell=m+1}^{n} \mathbf{x}_\ell \epsilon_\ell \right\| = O_P\left(N^{1/2}\right).$$

Since

$$\left\| \sum_{\ell=\hat{k}_{i-1}+1}^{k_i} (\mathbf{x}_\ell \mathbf{x}_\ell^\top - \mathbf{A}) \right\| = O_P\left(N^{1/2}\right) \quad \text{and} \quad \left\| \sum_{\ell=k_i+1}^{\hat{k}_{i+1}} (\mathbf{x}_\ell \mathbf{x}_\ell^\top - \mathbf{A}) \right\| = O_P\left(N^{1/2}\right),$$

the proof is complete. □

Remark B.1. If information is given about the rate of convergence of $\hat{k}_{i-1}/N$ and $\hat{k}_{i+1}/N$ to θ_{i-1} and θ_{i+1}, Assumption 2.8 can be replaced with weaker conditions. If, for example,

$$\left| \frac{\hat{k}_{i-1}}{N} - \theta_{i-1} \right| = O_P(r_{i-1}) \quad \text{and} \quad \left| \frac{\hat{k}_{i+1}}{N} - \theta_{i+1} \right| = O_P(r_{i+1}),$$

then

$$\frac{r_{i-1}\|\boldsymbol{\Delta}_{i-1,N}\|}{\|\boldsymbol{\Delta}_{i,N}\|} \to 0 \quad \text{and} \quad \frac{r_{i+1}\|\boldsymbol{\Delta}_{i+1,N}\|}{\|\boldsymbol{\Delta}_{i,N}\|} \to 0$$

are sufficient for Theorem 2.2 to hold. The typical rate of convergence is the fast rate

$$r_{i-1} = \frac{\|\boldsymbol{\Delta}_{i-1,N}\|^2}{N} \quad \text{and} \quad r_{i+1} = \frac{\|\boldsymbol{\Delta}_{i+1,N}\|^2}{N}.$$

Lemma B.3. *If Assumptions 2.1–2.8 hold, then we have*

$$\left| \frac{\hat{\mathfrak{k}}_{i,N}}{\hat{k}_{i+1} - \hat{k}_{i-1}} - \frac{\theta_i - \theta_{i-1}}{\theta_{i+1} - \theta_{i-1}} \right| = o_P(1).$$

Proof. First we assume that $\kappa < 1/2$. According to Assumption 2.1

$$\lim_{N \to \infty} P\left\{ \hat{k}_{i-1} < k_i < \hat{k}_{i+1} \right\} = 1.$$

By the definition of $\{\hat{\epsilon}_{i,\ell}, \hat{k}_{i-1} < \ell \leq \hat{k}_{i+1}\}$ we have

$$\hat{\epsilon}_{i,\ell} = \epsilon_\ell + \begin{cases} \mathbf{x}_\ell^\top(\boldsymbol{\beta}_{i-2} - \hat{\boldsymbol{\beta}}_{i,N}), & \hat{k}_{i-1} < \ell < \max(k_{i-1}, \hat{k}_{i-1}), \\ \mathbf{x}_\ell^\top(\boldsymbol{\beta}_{i-1} - \hat{\boldsymbol{\beta}}_{i,N}), & \max(k_{i-1}, \hat{k}_{i-1}) \leq \ell < k_i, \\ \mathbf{x}_\ell^\top(\boldsymbol{\beta}_{i} - \hat{\boldsymbol{\beta}}_{i,N}), & k_i \leq \ell < \min(k_{i+1}, \hat{k}_{i+1}), \\ \mathbf{x}_\ell^\top(\boldsymbol{\beta}_{i+1} - \hat{\boldsymbol{\beta}}_{i,N}), & \min(k_{i+1}, \hat{k}_{i+1}) \leq \ell < \hat{k}_{i+1}. \end{cases} \tag{B.1}$$

There are four possible cases, depending on if the first and the last lines in (B.1) are empty. We consider only the most difficult case, the proof of the other cases are the same but somewhat simpler. If $\hat{k}_{i-1} < k_{i-1}, k_{i+1} < \hat{k}_{i+1}$ we have

$$\begin{aligned} \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \hat{\epsilon}_{i,\ell} = & \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \epsilon_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \epsilon_\ell \\ & - \left(\sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top \right) \left(\hat{\boldsymbol{\beta}}_{i,N} - \bar{\boldsymbol{\beta}}_{i,N} \right) \\ & + \mathbf{G}_{i,N}(k) + \mathbf{G}_{i,N}^{(1)}(k) + \mathbf{G}_{i,N}^{(2)}(k), \end{aligned}$$

with

$$\mathbf{G}_{i,N}(k) = \begin{cases} \left((1-\zeta_i) \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \mathbf{x}_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \left[(1-\zeta_i) \sum_{\ell=\hat{k}_{i-1}+1}^{k_i} \mathbf{x}_\ell \mathbf{x}_\ell^\top \right.\right. \\ \qquad \left.\left. -\zeta_i \sum_{\ell=k_i+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top \right] \right) \boldsymbol{\Delta}_{N,i}, \quad \text{if} \quad \hat{k}_{i-1} < k \leq k_i, \\ \left(-\zeta_i \sum_{\ell=k+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top + \frac{\hat{k}_{i+1}-k}{\hat{k}_{i+1}-\hat{k}_{i-1}} \left[(1-\zeta_i) \sum_{\ell=\hat{k}_{i-1}+1}^{k_i} \mathbf{x}_\ell \mathbf{x}_\ell^\top \right.\right. \\ \qquad \left.\left. -\zeta_i \sum_{\ell=k_i+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top \right] \right) \boldsymbol{\Delta}_{N,i}, \quad \text{if} \quad k_i + 1 \leq k \leq \hat{k}_{i+1}, \end{cases}$$

$$\mathbf{G}_{i,N}^{(1)}(k) = \begin{cases} \sum_{\ell=\hat{k}_{i-1}}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top (-1 + \iota_{i,1}) \boldsymbol{\Delta}_{N,i-1}, & \hat{k}_{i-1} < k < k_{i-1}, \\ \sum_{\ell=\hat{k}_{i-1}}^{k_{i-1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top (-\boldsymbol{\Delta}_{N,i-1}) + \sum_{i=\hat{k}_{i-1}}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top \iota_{i,1} \boldsymbol{\Delta}_{N,i-1}, & k_{i-1} \leq k < \hat{k}_{i+1}, \end{cases}$$

and

$$\mathbf{G}_{i,N}^{(2)}(k) = \begin{cases} \displaystyle\sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top (-\iota_{i,2}) \boldsymbol{\Delta}_{N,i+1}, & \hat{k}_{i-1} < k \le k_{i+1}, \\ \displaystyle\sum_{\ell=k_{i+1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell \iota_{i,2} \boldsymbol{\Delta}_{N,i+1}, & k_{i+1} < k < \hat{k}_{i+1}. \end{cases}$$

Using Assumptions 2.1–2.5 we obtain

$$\begin{aligned} N^{-1+2\kappa} \max_{\hat{k}_{i-1}<k<\hat{k}_{i+1}} & \frac{1}{(k-\hat{k}_{i-1})^\kappa} \frac{1}{(\hat{k}_{i+1}-k)^\kappa} \left(\|\mathbf{G}_{i,N}^{(1)}(k)\| + \|\mathbf{G}_{i,N}^{(2)}(k)\| \right) \\ &= O_P(1) \left(\left| \frac{\hat{k}_{i-1}}{N} - \frac{k_{i-1}}{N} \right| + \left| \frac{\hat{k}_{i+1}}{N} - \frac{k_{i+1}}{N} \right| \right) (\|\boldsymbol{\Delta}_{N,i-1}\| + \|\boldsymbol{\Delta}_{N,i+1}\|) \\ &= o_P(1) (\|\boldsymbol{\Delta}_{N,i-1}\| + \|\boldsymbol{\Delta}_{N,i+1}\|) . \end{aligned} \tag{B.2}$$

We note

$$\begin{aligned} N^{-1/1+2\kappa} & \max_{\hat{k}_{i-1}<k<\hat{k}_{i+1}} \frac{1}{(k-\hat{k}_{i-1})^\kappa} \frac{1}{(\hat{k}_{i+1}-k)^\kappa} \left\| \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \epsilon_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}+1} \mathbf{x}_\ell \epsilon_\ell \right\| \\ &\le N^{-1/2+2\kappa} \max_{\hat{k}_{i-1}<k\le k_i} \frac{1}{(k-\hat{k}_{i-1})^\kappa} \frac{1}{(\hat{k}_{i+1}-k)^\kappa} \left\| \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \epsilon_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \epsilon_\ell \right\| \\ &\quad + N^{-1/2+2\kappa} \max_{k_i<k<\hat{k}_{i+1}} \frac{1}{(k-\hat{k}_{i-1})^\kappa} \frac{1}{(\hat{k}_{i+1}-k)^\kappa} \left\| \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \epsilon_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \epsilon_\ell \right\| . \end{aligned}$$

By Assumptions 2.1–2.5 we get

$$\begin{aligned} N^{-1/2+2\kappa} & \max_{\hat{k}_{i-1}<k\le k_i} \frac{1}{(k-\hat{k}_{i-1})^\kappa} \frac{1}{(\hat{k}_{i+1}-k)^\kappa} \left\| \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \epsilon_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \epsilon_\ell \right\| \\ &= O_P(1) N^{-1/2+\kappa} \max_{\hat{k}_{i-1}<k\le k_i} \frac{1}{(k-\hat{k}_{i-1})^\kappa} \left\| \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \epsilon_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \epsilon_\ell \right\| , \end{aligned}$$

and

$$\begin{aligned} N^{-1/2+\kappa} & \max_{\hat{k}_{i-1}<k\le k_i} \frac{1}{(k-\hat{k}_{i-1})^\kappa} \left\| \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \epsilon_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \epsilon_\ell \right\| \\ \le N^{-1/2+\kappa} & \max_{\hat{k}_{i-1}<k\le k_i} \frac{1}{(k-\hat{k}_{i-1})^\kappa} \left\| \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \epsilon_\ell \right\| + N^{-1/2+\kappa} \max_{\hat{k}_{i-1}<k\le k_i} \frac{(k-\hat{k}_{i-1})^{1-\kappa}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \left\| \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \epsilon_\ell \right\| . \end{aligned}$$

Assumption 2.1 yields that for all $c > 0$

$$P\left\{\max_{\hat{k}_{i-1}<k\le k_i}\frac{1}{(k-\hat{k}_{i-1})^{\kappa}}\left\|\sum_{\ell=\hat{k}_{i-1}+1}^{k}\mathbf{x}_\ell\epsilon_\ell\right\| \le \max_{k_{i-1}-Nc\le m\le k_{i-1}+Nc}\max_{m<k\le k_i}\frac{1}{(k-m)^{\kappa}}\left\|\sum_{\ell=m+1}^{k}\mathbf{x}_\ell\epsilon_\ell\right\|\right\} \to 1,$$

and using Proposition 2 of Berkes et al. (2011) coordinatewise, we get

$$N^{-1/2+\kappa}\max_{k_{i-1}-Nc\le m\le k_{i-1}+Nc}\max_{m<k\le k_i}\frac{1}{(k-m)^{\kappa}}\left\|\sum_{\ell=m+1}^{k}\mathbf{x}_\ell\epsilon_\ell\right\| = O_P(1).$$

Theorem S2.1 of Aue et al. (2014) implies

$$N^{-1/2}\max_{0\le m<k\le N}\left\|\sum_{\ell=m+1}^{k}\mathbf{x}_\ell\epsilon_\ell\right\| = O_P(1),$$

and thus we conclude

$$N^{-1/2+\kappa}\max_{\hat{k}_{i-1}<k\le k_i}\frac{1}{(k-\hat{k}_{i-1})^{\kappa}}\left\|\sum_{\ell=\hat{k}_{i-1}+1}^{k}\mathbf{x}_\ell\epsilon_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}}\sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}}\mathbf{x}_\ell\epsilon_\ell\right\| = O_P(1). \tag{B.3}$$

Similar arguments yield

$$N^{-1/2+\kappa}\max_{k_i<k<\hat{k}_{i+1}}\frac{1}{(k-\hat{k}_{i-1})^{\kappa}}\left\|\sum_{\ell=\hat{k}_{i-1}+1}^{k}\mathbf{x}_\ell\epsilon_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}}\sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}}\mathbf{x}_\ell\epsilon_\ell\right\| = O_P(1) \tag{B.4}$$

and therefore

$$N^{-1/2+2\kappa}\max_{\hat{k}_{i-1}<k\le k_i}\frac{1}{(k-\hat{k}_{i-1})^{\kappa}}\frac{1}{(\hat{k}_{i+1}-k)^{\kappa}}\left\|\sum_{\ell=\hat{k}_{i-1}+1}^{k}\mathbf{x}_\ell\hat{\epsilon}_{i,\ell} - \mathbf{G}_{i,N}(k)\right\| = O_P(1).$$

Following the proof of (B.2) one can show that

$$N^{-1/2+2\kappa}\max_{\hat{k}_{i-1}<k\le k_i}\frac{1}{(k-\hat{k}_{i-1})^{\kappa}}\frac{1}{(\hat{k}_{i+1}-k)^{\kappa}}\Bigg\|(1-\zeta_i)\sum_{\ell=\hat{k}_{i-1}+1}^{k}[\mathbf{x}_\ell\mathbf{x}_\ell - \mathbf{A}] - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}}$$
$$\times\left[(1-\zeta_i)\sum_{\ell=\hat{k}_{i-1}+1}^{k_i}[\mathbf{x}_\ell\mathbf{x}_\ell^\top - \mathbf{A}] + \zeta_i\sum_{\ell=k_i+1}^{\hat{k}_{i+1}}[\mathbf{x}_\ell\mathbf{x}_\ell^\top - \mathbf{A}]\right]\Bigg\| = O_P(1)$$

and

$$N^{-1/2+2\kappa}\max_{k_i<k<\hat{k}_{i+1}}\frac{1}{(k-\hat{k}_{i-1})^{\kappa}}\frac{1}{(\hat{k}_{i+1}-k)^{\kappa}}\Bigg\|-\zeta_i\sum_{\ell=k+1}^{\hat{k}_{i+1}}[\mathbf{x}_\ell\mathbf{x}_\ell^\top - \mathbf{A}] + \frac{\hat{k}_{i+1}-k}{\hat{k}_{i+1}-\hat{k}_i}$$

$$\times\left[(1-\zeta_i)\sum_{\ell=\hat{k}_{i-1}+1}^{k_i}[\mathbf{x}_\ell\mathbf{x}_\ell^\top-\mathbf{A}]-\zeta_i\sum_{\ell=k_i+1}^{\hat{k}_{i+1}}[\mathbf{x}_\ell\mathbf{x}_\ell-\mathbf{A}]\right]\Bigg\|=O_P(1).$$

We also define

$$\bar{\mathbf{G}}_N(k)=N^{-1+2\kappa}\frac{1}{(k-\hat{k}_{i-1})^\kappa}\frac{1}{(\hat{k}_{i+1}-k)^\kappa}\times\begin{cases}(1-\zeta_i)(k-\hat{k}_{i-1})-\dfrac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}}\Big[(k_i-\hat{k}_{i-1})\\ \qquad\qquad -\zeta_i(\hat{k}_{i+1}-k_i)\Big], \quad \hat{k}_{i-1}<k\le k_i,\\ -\zeta_i(\hat{k}_{i+1}-k)+\dfrac{\hat{k}_{i+1}-k}{\hat{k}_{i+1}-\hat{k}_{i-1}}\Big[(1-\zeta_i)(k_i-\hat{k}_{i-1})\\ \qquad\qquad -\zeta_i(\hat{k}_{i+1}-k_i)\Big], \quad k_i<k<\hat{k}_{i+1}.\end{cases}$$

We note

$$\bar{\mathbf{G}}_N(k)=N^{-1+2\kappa}\frac{1}{(k-\hat{k}_{i-1})^\kappa}\frac{1}{(\hat{k}_{i+1}-k)^\kappa}\times\begin{cases}\dfrac{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)}{\hat{k}_{i+1}-\hat{k}_{i-1}}, & \hat{k}_{i-1}<k\le k_i,\\ -\dfrac{(\hat{k}_{i+1}-k)(k_i-\hat{k}_{i-1})}{\hat{k}_{i+1}-\hat{k}_{i-1}} & k_i<k<\hat{k}_{i+1}.\end{cases}$$

The rest of the proof is the same as the last arguments used in the proof of Lemma A.2. If $\kappa=1/2$, then we can repeat the arguments above but the upper bounds $O_P(1)$ in (B.3) and (B.4) are replaced with $O_P((\log N)^{1/2})$. Also, (B.1) is modified to

$$N^{-1+2\kappa}\max_{\hat{k}_{-1}<k<\hat{k}_{i+1}}\frac{1}{(k-\hat{k}_{i-1})^\kappa}\frac{1}{(\hat{k}_{i+1}-k)^\kappa}\left(\|\mathbf{G}_{i,N}^{(1)}(k)\|+\|\mathbf{G}_{i,N}^{(2)}(k)\|\right)$$
$$=o_P(1)(\log N)^{1/2}\left(\|\mathbf{\Delta}_{N,i-1}\|+\|\mathbf{\Delta}_{N,i+1}\|\right).$$

□

Lemma B.4. *If Assumptions 2.1–2.8 hold, then we have*

$$\frac{1}{\|\mathbf{\Delta}_{N,i}\|^2}\left|\frac{\hat{\mathfrak{k}}_i}{N}-\theta_i\right|=O_P(1).$$

Proof. We modify the proof of Lemma A.4. We consider the case when $\hat{k}_{i-1}<k_{i-1}, k_{i+1}<\hat{k}_{i+1}$. Let

$$\sum_{\ell=\hat{k}_{i-1}+1}^{k}\mathbf{x}_\ell\hat{\epsilon}_{i,\ell}-\frac{k}{\hat{k}_{i+1}-\hat{k}_{i-1}}\sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}}\mathbf{x}_\ell\hat{\epsilon}_{i,\ell}=\mathbf{Q}_{1,i}(k)+\ldots+\mathbf{Q}_{8,i}(k),$$

where

$$\mathbf{Q}_{1,i}(k) = \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \epsilon_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \epsilon_\ell,$$

$$\mathbf{Q}_{2,i}(k) = -\left(\sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \epsilon_\ell - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \epsilon_\ell \right) \left(\hat{\boldsymbol{\beta}}_{i,N} - \bar{\boldsymbol{\beta}}_{i,N} \right),$$

$$\mathbf{Q}_{3,i}(k) = \left(1 - \frac{k_i-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \right) \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2),$$

$$\mathbf{Q}_{4,i}(k) = -\frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \left((1-\zeta_i) \sum_{\ell=\hat{k}_{i-1}+1}^{k_i} \mathbf{x}_\ell \mathbf{x}_\ell^\top - \zeta_i \sum_{\ell=k_i+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top \right) (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2),$$

$$\begin{aligned}\mathbf{Q}_{5,i}(k) = &\left(\sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top \right) (1+\iota_{i,1}) \mathbb{1}\{\hat{k}_{i-1} < k < k_{i-1}\} \boldsymbol{\Delta}_{N,i-1} \\ &+ \left(\sum_{\ell=\hat{k}_{i-1}+1}^{k_{i-1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top (1+\iota_{i,1}) + \sum_{\ell=k_{i-1}+1}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top \iota_{i,1} \right) \mathbb{1}\{\hat{k}_{i-1} < k < \hat{k}_{i+1}\} \boldsymbol{\Delta}_{N,i-1},\end{aligned}$$

$$\mathbf{Q}_{6,i}(k) = -\frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \left((1+\iota_{i,1}) \sum_{\ell=\hat{k}_{i-1}+1}^{k_{i-1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top + \iota_{i,1} \sum_{\ell=k_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top \right) \boldsymbol{\Delta}_{N,i-1},$$

$$\begin{aligned}\mathbf{Q}_{7,i}(k) = &\sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top (-\iota_{i,2}) \mathbb{1}\{\hat{k}_{i-1} < k \le k_{i+1}\} \boldsymbol{\Delta}_{N,i+1} \\ &+ \left(\sum_{\ell=\hat{k}_{i-1}+1}^{k_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top (-\iota_{i,2}) + \sum_{\ell=k_{i+1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top (1-\iota_{i,2}) \right) \mathbb{1}\{k_{i+1} < k < \hat{k}_{i+1}\} \boldsymbol{\Delta}_{N,i+1},\end{aligned}$$

$$\mathbf{Q}_{8,i}(k) = -\frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \left((1-\iota_{i,2}) \sum_{\ell=k_{i+1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top - \iota_{i,2} \sum_{\ell=\hat{k}_{i-1}+1}^{k_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top \right) \boldsymbol{\Delta}_{N,i+1}.$$

Using the processes above we have the following decomposition:

$$\begin{aligned}&\left(\frac{k_{i+1}-k_{i-1}}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)} \right)^{2\kappa} \left\| \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \hat{\epsilon}_{i,\ell} - \frac{k-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \hat{\epsilon}_{i,\ell} \right\|^2 \\ &\quad - \left(\frac{k_{i+1}-k_{i-1}}{(k_i-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)} \right)^{2\kappa} \left\| \sum_{\ell=\hat{k}_{i-1}+1}^{k_i} \mathbf{x}_\ell \hat{\epsilon}_{i,\ell} - \frac{k_i-\hat{k}_{i-1}}{\hat{k}_{i+1}-\hat{k}_{i-1}} \sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \hat{\epsilon}_{i,\ell} \right\|^2\end{aligned}$$

$$= \left(\frac{k_{i+1}-k_{i-1}}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa} \sum_{m,n=1}^{8} \mathbf{Q}_{m,i}^{\top}(k)\mathbf{Q}_{n,i}(k)$$

$$- \left(\frac{k_{i+1}-k_{i-1}}{(k_i-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)}\right)^{2\kappa} \sum_{m,n=1}^{8} \mathbf{Q}_{m,i}^{\top}(k_i)\mathbf{Q}_{n,i}(k_i). \tag{B.5}$$

Let $0 < \alpha < \theta_i/\theta_{i-1} - 1$. Elementary arguments yield

$$\max_{k_{i-1}(1+\alpha)\le k<k_i} \frac{1}{k_i-k}\left|\left(\frac{k_{i+1}-k_{i-1}}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa} - \left(\frac{k_{i+1}-k_{i-1}}{(k_i-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)}\right)^{2\kappa}\right| = O_P\left(N^{-1-2\kappa}\right).$$

It follows from Theorem S2.1 in Aue et al. (2014) that

$$\max_{k_{i-1}(1-\delta)\le k\le k_{i+1}(1+\delta)} \|\mathbf{Q}_{1,i}(k)\| = O_P(N^{1/2})$$

and for all $\delta > 0$

$$P\left\{\max_{\hat{k}_{i-1}<k<\hat{k}_{i+1}} \|\mathbf{Q}_{1,i}(k)\| \le \max_{k_{i-1}(1-\delta)\le k\le k_{i+1}(1+\delta)} \|\mathbf{Q}_{1,i}(k)\|\right\} \to 1.$$

Let

$$a = \frac{C}{\|\mathbf{\Delta}_{N,i}\|^2}, \qquad \text{where} \quad C > 0.$$

Thus we get

$$\begin{aligned}
&\max_{k_{i-1}(1+\alpha)\le k\le k_i-a} \frac{N^{-1+2\kappa}}{\|\mathbf{\Delta}_{N,i}\|^2(k_i-k)}\left|\left(\frac{k_{i+1}-k_{i-1}}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa} - \left(\frac{k_{i+1}-k_{i-1}}{(k_i-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)}\right)^{2\kappa}\right| \\
&\qquad \times \|\mathbf{Q}_{1,i}(k)\|^2 \\
&= O_P(1)\frac{1}{N\|\mathbf{\Delta}_{N,i}\|^2} \max_{k_{i-1}(1+\alpha)\le k\le k_i-a} \left\|N^{-1/2}\mathbf{Q}_{1,i}(k)\right\|^2 \\
&= O_P\left(\frac{1}{N\|\mathbf{\Delta}_{N,i}\|^2}\right) \\
&= o_P(1).
\end{aligned}$$

We use

$$\left|\|\mathbf{Q}_{1,i}(k)\|^2 - \|\mathbf{Q}_{1,i}(k_i)\|^2\right| \le \left|\|\mathbf{Q}_{1,i}(k)\| - \|\mathbf{Q}_{1,i}(k_i)\|\right| \left(\|\mathbf{Q}_{1,i}(k)\| + \|\mathbf{Q}_{1,i}(k_i)\|\right)$$

and

$$\|\mathbf{Q}_{1,i}(k) - \mathbf{Q}_{1,i}(k_i)\| \le \left\|\sum_{\ell=k+1}^{k_i} \mathbf{x}_\ell\epsilon_\ell\right\| + \frac{k_i-k}{\hat{k}_{i+1}-\hat{k}_{i-1}}\left\|\sum_{\ell=\hat{k}_{i-1}+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell\epsilon_\ell\right\|.$$

Lemma A.3 yields that

$$a^{1/2} \max_{1 \le k \le k_i - a} \frac{1}{k_i - k} \left\| \sum_{\ell = k+1}^{k_i} \mathbf{x}_\ell \epsilon_\ell \right\| = O_P(1)$$

and

$$\max_{\hat{k}_{i-1} < k \le \hat{k}_{i+1}} \left\| \sum_{\ell = \hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell \epsilon_\ell \right\| = O_P\left(N^{1/2}\right).$$

Thus we conclude

$$\begin{aligned}
&\left(\frac{N}{(k_i - \hat{k}_{i-1})(\hat{k}_{i+1} - k_i)}\right)^{2\kappa} \max_{k_{i-1}(1+\alpha) \le k \le k_i - a} \frac{N^{-1+2\kappa}}{\|\boldsymbol{\Delta}_{N,i}\|^2 (k_i - k)} \left| \|\mathbf{Q}_{1,i}(k)\|^2 - \|\mathbf{Q}_{1,i}(k_i)\|^2 \right| \\
&\quad \le \left[a^{1/2} \max_{k_{i-1}(1+\alpha) \le k \le k_i - a} \frac{1}{k_i - k} \left| \|\mathbf{Q}_{1,i}(k)\| - \|\mathbf{Q}_{1,i}(k_i)\| \right| \right] \\
&\qquad\qquad \times \left[N^{-1/2} \left(\max_{k_{i-1}(1+\alpha) \le k \le k_i} \|\mathbf{Q}_{1,i}(k)\| + \|\mathbf{Q}_{1,i}(k_i)\| \right) \right] \frac{N^{1/2} a}{N \|\boldsymbol{\Delta}_{N,i}\|^2} \\
&\quad = o_P(1).
\end{aligned}$$

We can modify the arguments of Lemma A.4, as in case the terms $\|\mathbf{Q}_{1,i}(k)\|^2 - \|\mathbf{Q}_{1,i}(k_i)\|^2$, to get

$$\begin{aligned}
\text{(B.6)} \qquad \max_{k_{i-1}(1+\alpha) \le k \le k_i - a} \frac{N^{-1+2\kappa}}{(k_i - k)\|\boldsymbol{\Delta}_{N,i}\|^2} &\left| \left(\frac{N}{(k - \hat{k}_{i-1})(\hat{k}_{i+1} - k)}\right)^{2\kappa} \mathbf{Q}_{m,i}^\top(k) \mathbf{Q}_{n,i}(k) \right. \\
&\left. - \left(\frac{N}{(k_i - \hat{k}_{i-1})(\hat{k}_{i+1} - k_i)}\right)^{2\kappa} \mathbf{Q}_{m,i}^\top(k_i) \mathbf{Q}_{n,i}(k_i) \right| = o_P(1),
\end{aligned}$$

if $(m, n) = (1,1), (1,2), (2,1), (2,2), (2,3), (3,2), (4,1), (4,2), (4,3), (4,4), (3,2), (2,3)$.
Now we consider terms containing $\mathbf{Q}_{5,i}, \ldots, \mathbf{Q}_{8,i}$. By definition, if $k_{i-1}(1+\alpha) \le k \le k_i - a$

$$\begin{aligned}
\|\mathbf{Q}_{5,i}(k)\|^2 &= \left\| \sum_{\ell = \hat{k}_{i-1}+1}^{k_{i-1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top (1 + \iota_{i,1}) + \sum_{\ell = k_{i-1}+1}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top \iota_{i,1} \right\|^2 \|\boldsymbol{\Delta}_{N,i-1}\|^2 \\
&= \left[\left\| \sum_{\ell = \hat{k}_{i-1}+1}^{k_{i-1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top (1 + \iota_{i,1}) \right\|^2 + \left(\sum_{\ell = \hat{k}_{i-1}+1}^{k_{i-1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top (1 + \iota_{i,1}) \right)^\top \left(\sum_{\ell = k_{i-1}+1}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top \iota_{i,1} \right) \right. \\
&\qquad \left. + \left(\sum_{\ell = k_{i-1}+1}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top \iota_{i,1} \right)^\top \left(\sum_{\ell = \hat{k}_{i-1}+1}^{k_{i-1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top (1 + \iota_{i,1}) \right) + \left\| \sum_{\ell = k_{i-1}+1}^{k} \mathbf{x}_\ell \mathbf{x}_\ell^\top \iota_{i,1} \right\|^2 \right] \|\boldsymbol{\Delta}_{N,i-1}\|^2
\end{aligned}$$

and therefore

$$\left|\|\mathbf{Q}_{5,i}(k)\|^2-\|\mathbf{Q}_{5,i}(k_i)\|^2\right|\leq 2\left\|\sum_{\ell=\hat{k}_{i-1}+1}^{k_{i-1}}\mathbf{x}_\ell\mathbf{x}_\ell^\top(1+\iota_{i,1})\right\|\left\|\sum_{\ell=k+1}^{k_i}\mathbf{x}_\ell\mathbf{x}_\ell^\top\iota_{i,1}\right\|\|\mathbf{\Delta}_{N,i-1}\|^2$$
$$+\left|\left\|\sum_{\ell=k_{i-1}+1}^{k}\mathbf{x}_\ell\mathbf{x}_\ell^\top\iota_{i,1}\right\|^2-\left\|\sum_{\ell=k_{i-1}+1}^{k_i}\mathbf{x}_\ell\mathbf{x}_\ell^\top\iota_{i,1}\right\|^2\right|\|\mathbf{\Delta}_{N,i-1}\|^2.$$

It follows from Assumption 2.5

$$\max_{k_{i-1}(1+\alpha)\leq k\leq k_i-a}\|\mathbf{Q}_{5,i}(k)\|^2=O_P(N\|\mathbf{\Delta}_{N,i-1}\|^2)$$

and thus we get

$$\max_{k_{i-1}(1+\alpha)\leq k\leq k_i-a}\frac{N^{-1+2\kappa}}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}\left|\left(\frac{N}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}-\left(\frac{N}{(k_i-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)}\right)^{2\kappa}\right|\|\mathbf{Q}_{5,i}(k)\|^2$$
$$=O_P\left(\frac{1}{N}\right)\frac{\|\mathbf{\Delta}_{N,i-1}\|^2}{\|\mathbf{\Delta}_{N,i}\|^2}=o_P(1).$$

Now

$$\max_{k_{i-1}(1+\alpha)\leq k\leq k_i-a}\left(\frac{N}{(k_i-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)}\right)^{2\kappa}\frac{N^{-1+2\kappa}}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}\left\|\sum_{\ell=\hat{k}_{i-1}+1}^{k_{i-1}}[\mathbf{x}_\ell\mathbf{x}_\ell^\top-\mathbf{A})]\right\|$$
$$\times\left\|\sum_{\ell=k+1}^{k_i}[\mathbf{x}_\ell\mathbf{x}_\ell^\top-\mathbf{A}]\iota_{i,1}\right\|\|\mathbf{\Delta}_{N,i-1}\|^2$$
$$=o_P(1)\frac{\|\mathbf{\Delta}_{N,i-1}\|^2}{N^{1/2}\|\mathbf{\Delta}_{N,i}\|}=o_P(1), \tag{B.7}$$

since by Assumption 2.1

$$\left\|\sum_{\ell=\hat{k}_{i-1}+1}^{k_{i-1}}[\mathbf{x}_\ell\mathbf{x}_\ell^\top-\mathbf{A}]\right\|=O_P\left(N^{1/2}\right)$$

and

$$\max_{k_{i-1}(1+\alpha)\leq k\leq k_i-a}\frac{1}{k_i-k}\left\|\sum_{\ell=k+1}^{k_i}[\mathbf{x}_\ell\mathbf{x}_\ell^\top-\mathbf{A}]\iota_{i,1}\right\|=O_P\left(\|\mathbf{\Delta}_{N,i}\|\iota_{i,1}\right).$$

We note

$$\max_{k_{i-1}(1+\alpha)\leq k\leq k_i-a}\left(\frac{N}{(k_i-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)}\right)^{2\kappa}\frac{N^{-1+2\kappa}}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}$$
$$\times|k_{i-1}-\hat{k}_{i-1}|(k_i-k)\iota_{i,1}\|\mathbf{A}\|\|\mathbf{\Delta}_{N,i-1}\|^2=o_P(1). \tag{B.8}$$

Utilizing

$$\left|\left\|\sum_{\ell=k_{i-1}+1}^{k}\mathbf{x}_\ell\mathbf{x}_\ell^\top\iota_{i,1}\right\|^2-\left\|\sum_{\ell=k_{i-1}+1}^{k_i}\mathbf{x}_\ell\mathbf{x}_\ell^\top\iota_{i,1}\right\|^2\right|$$
$$\leq \iota_{i,1}\left\|\sum_{\ell=k+1}^{k_i}\mathbf{x}_\ell\mathbf{x}_\ell^\top\right\|\left(\left\|\sum_{\ell=k_{i-1}+1}^{k}\mathbf{x}_\ell\mathbf{x}_\ell^\top\right\|+\left\|\sum_{\ell=k_{i-1}+1}^{k_i}\mathbf{x}_\ell\mathbf{x}_\ell^\top\right\|\right)$$
$$\leq \iota_{i,1}\left\{\left\|\sum_{\ell=k+1}^{k_i}[\mathbf{x}_\ell\mathbf{x}_\ell^\top-\mathbf{A}]\right\|+(k_i-k)\|\mathbf{A}\|\right\}\left\{\left\|\sum_{\ell=k_{i-1}+1}^{k}\mathbf{x}_\ell\mathbf{x}_\ell^\top\right\|+\left\|\sum_{\ell=k_{i-1}+1}^{k_i}\mathbf{x}_\ell\mathbf{x}_\ell^\top\right\|\right\},$$

we get, as in the proof of (B.7) and (B.8), that

$$\max_{k_{i-1}(1+\alpha)\leq k\leq k_i-a}\left(\frac{N}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}\frac{N^{-1+2\kappa}}{(k_i-k)\|\boldsymbol{\Delta}_{N,i}\|^2}$$
$$\times\left|\left\|\sum_{\ell=k_{i-1}+1}^{k}\mathbf{x}_\ell\mathbf{x}_\ell^\top\iota_{i,1}\right\|^2-\left\|\sum_{\ell=k_{i-1}+1}^{k_i}\mathbf{x}_\ell\mathbf{x}_\ell^\top\iota_{i,1}\right\|^2\right|\|\boldsymbol{\Delta}_{N,i-1}\|^2=o_P(1).$$

The same arguments show that (B.6) holds, when $m,n=5,6,7,8$. Next we show that (B.6) holds when $m=1$ and $n=5$, We observe

$$\max_{k_{i-1}(1+\alpha)\leq k\leq k_i-a}\frac{N^{-1+2\kappa}}{(k_i-k)\|\boldsymbol{\Delta}_{N,i}\|^2}\left|\left(\frac{N^{-1+2\kappa}}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}-\left(\frac{N^{-1+2\kappa}}{(k_i-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)}\right)^{2\kappa}\right|$$
$$\times\|\mathbf{Q}_{1,i}(k_i)\|\|\mathbf{Q}_{5,i}(k_i)\|$$
$$=O_P(1)\frac{\|\boldsymbol{\Delta}_{N,i-1}\|}{N\|\boldsymbol{\Delta}_{N,i}\|^2}=o_P(1),$$

since

$$\|\mathbf{Q}_{1,i}(k_i)\|=O_P\left(N^{1/2}\right)\qquad\text{and}\qquad\|\mathbf{Q}_{5,i}(k_i)\|=O_P\left(N^{1/2}\|\boldsymbol{\Delta}_{N,i-1}\|\right).$$

Also,

$$\max_{k_{i-1}(1+\alpha)\leq k\leq k_i-a}\left(\frac{N}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}\frac{N^{-1+2\kappa}}{(k_i-k)\|\boldsymbol{\Delta}_{N,i}\|^2}\|\mathbf{Q}_{1,i}(k_i)-\mathbf{Q}_{1,i}(k)\|\|\mathbf{Q}_{5,i}(k)\|$$
$$=O_P(1)\frac{\|\boldsymbol{\Delta}_{N,i-1}\|}{N^{1/2}\|\boldsymbol{\Delta}_{N,i}\|}=o_P(1)$$

and

$$\max_{k_{i-1}(1+\alpha)\leq k\leq k_i-a}\left(\frac{N}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}\frac{N^{-1+2\kappa}}{(k_i-k)\|\boldsymbol{\Delta}_{N,i}\|^2}\|\mathbf{Q}_{5,i}(k_i)-\mathbf{Q}_{5,i}(k)\|\|\mathbf{Q}_{1,i}(k)\|$$

$$= O_P(1))\frac{\|\mathbf{\Delta}_{N,i-1}\|}{N^{1/2}\|\mathbf{\Delta}_{N,i}\|} = o_P(1).$$

The same argument shows that (B.6) holds when $m = 1, 2, 4$ and $n = 5, 6, 7, 8$.

Next we consider $\mathbf{Q}_{3,i}(k)\mathbf{Q}_{5,i}(k)$. We proceed as before,

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a} \frac{N^{-1+2\kappa}}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}\left|\left(\frac{N}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa} - \left(\frac{N}{(k_i-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)}\right)^{2\kappa}\right|$$

$$\times \|\mathbf{Q}_{3,i}(k)\|\|\mathbf{Q}_{5,i}(k)\| = O_P\left(\iota_{i,1}\frac{\|\mathbf{\Delta}_{N,i-1}\|}{\|\mathbf{\Delta}_{N,i}\|}\right) = o_P(1).$$

Also,

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\left(\frac{N}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}\frac{N^{-1+2\kappa}}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}\|\mathbf{Q}_{5,i}(k)-\mathbf{Q}_{5,i}(k_i)\|\|\mathbf{Q}_{3,i}(k)\| = o_P(1),$$

since

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\frac{1}{k_i-k}\|\mathbf{Q}_{5,i}(k)-\mathbf{Q}_{5,i}(k_i)\| = O_P(\iota_{i,1}\|\mathbf{\Delta}_{N,i-1}\|) = o_P(\|\mathbf{\Delta}_{N,i-1}\|)$$

and

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\|\mathbf{Q}_{3,i}(k)\| = O_P(N\|\mathbf{\Delta}_{N,i}\|).$$

Using similar arguments we have

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\left(\frac{N}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}\frac{N^{-1+2\kappa}}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}\|\mathbf{Q}_{3,i}(k)-\mathbf{Q}_{3,i}(k_i)\|\|\mathbf{Q}_{5,i}(k)\| = o_P(1),$$

since

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\frac{1}{k_i-k}\|\mathbf{Q}_{3,i}(k)-\mathbf{Q}_{3,i}(k_i)\|\|\mathbf{Q}_{5,i}(k)\| = O_P(\|\mathbf{\Delta}_{N,i}\|)$$

and

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\|\mathbf{Q}_{5,i}(k)\| = O_P(N\iota_{i,1}\|\mathbf{\Delta}_{N,i-1}\|) = o_P(N\|\mathbf{\Delta}_{N,i-1}\|).$$

The same arguments imply that (B.6) holds when $m = 3$ and $n = 5, 6, 7, 8$. If $\mathscr{F} = \{(m,n) : 1 \le n, m \le 8\} \setminus \{(3,1),(1,3),(3,3)\}$, then

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\left(\frac{N}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}\frac{N^{-1+2\kappa}}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}|U_{N,i}(k)| = o_P(1)$$

with

$$U_{N,i}(k) = \sum_{(m,n)\in\mathscr{F}}\left(\frac{k_{i+1}-k_{i-1}}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}\mathbf{Q}_{m,i}^\top(k)\mathbf{Q}_{n,i}(k)$$

$$-\left(\frac{k_{i+1}-k_{i-1}}{(k_i-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)}\right)^{2\kappa}\mathbf{Q}_{m,i}^\top(k_i)\mathbf{Q}_{n,i}(k_i).$$

Next we consider $\mathbf{Q}_{3,i}^\top(k)\mathbf{Q}_{3,i}(k)$ and $\mathbf{Q}_{1,i}^\top(k)\mathbf{Q}_{3,i}(k)$. We note

$$\begin{aligned}
&\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\frac{1}{N^{1/2}(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}\|\mathbf{Q}_{3,i}(k)-\mathbf{Q}_{3,i}(k_i)\|\\
&\quad\le \max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\frac{1}{N^{1/2}(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}\left\|\sum_{\ell=k+1}^{k_i}\mathbf{x}_\ell\mathbf{x}_\ell^\top\right\|\|\mathbf{\Delta}_{N,i}\|\\
&\quad= O_P\left(\frac{1}{N^{1/2}\|\mathbf{\Delta}_{N,i}\|}\right)=o_P(1)
\end{aligned}$$

and thus we get

$$\begin{aligned}
&\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\frac{N^{-1+2\kappa}}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}\\
&\times\Bigg|\left(\frac{N}{(k-\hat{k}_{i-1})(k-\hat{k}_{i+1}}\right)^{2\kappa}\|\mathbf{Q}_{3,i}(k)\|^2-\left(\frac{N}{(k_i-\hat{k}_{i-1})(k_i-\hat{k}_{i+1}}\right)^{2\kappa}\|\mathbf{Q}_{3,i}(k_i)\|^2\\
&-\left(\left(\frac{N}{(k-\hat{k}_{i-1})(k-\hat{k}_{i+1}}\right)^{2\kappa}(k-\hat{k}_{i-1})^2-\left(\frac{N}{(k_i-\hat{k}_{i-1})(k_i-\hat{k}_{i+1}}\right)^{2\kappa}(k_i-\hat{k}_{i-1})^2\right)\mathbf{\Delta}_{N,i}^\top\mathbf{A}\mathbf{\Delta}_{N,i}\Bigg|\\
&=o_P(1).
\end{aligned}$$

We define

$$V_{N,i}(k)=\left(\frac{N}{(k-\hat{k}_{i-1})(k-\hat{k}_{i+1}}\right)^{2\kappa}\|\mathbf{Q}_{3,i}(k)\|^2-\left(\frac{N}{(k_i-\hat{k}_{i-1})(k_i-\hat{k}_{i+1}}\right)^{2\kappa}\|\mathbf{Q}_{3,i}(k_i)\|^2.$$

Our estimates yield

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\left(U_{N,i}(k)++\frac{1}{2}V_{N,i}(k)\right)\stackrel{P}{\to}-\infty.$$

We showed that

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\|\mathbf{Q}_{1,i}(k)\|=O_P\left(N^{1/2}\right)\quad\text{and}\quad\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\|\mathbf{Q}_{3,i}(k)\|=O_P\left(N\|\mathbf{\Delta}_{N,i}\|\right)$$

and therefore

$$\begin{aligned}
&\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\frac{N^{-1+2\kappa}}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}\left|\left(\frac{N}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}-\left(\frac{N}{(k_i-\hat{k}_{i-1})(\hat{k}_{i+1}-k_i)}\right)^{2\kappa}\right|\\
&\qquad\times\|\mathbf{Q}_{1,i}(k)\|\|\mathbf{Q}_{3,i}(k)\|=O_P\left(\frac{1}{N^{1/2}\|\mathbf{\Delta}_{N,i}\|}\right)=o_P(1).
\end{aligned}$$

Now

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a} \frac{N^{-1+2\kappa}}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}\left(\frac{N}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}\|\mathbf{Q}_{1,i}(k)-\mathbf{Q}_{1,i}(k_i)\|\|\mathbf{Q}_{3,i}(k)\|$$
$$=O_P(1)\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\frac{1}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|}\left\|\sum_{\ell=k+1}^{k_i}\mathbf{x}_\ell\epsilon_\ell\right\|.$$

It is shown in Horváth et al. (2025) that

$$\lim_{C\to\infty}\limsup_{N\to\infty}P\left\{\max_{1\le k\le k_i-C/\|\mathbf{\Delta}_{N,i}\|^2}\frac{1}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|}\left\|\sum_{\ell=k+1}^{k_i}\mathbf{x}_\ell\epsilon_\ell\right\|>x\right\}=0,\quad \text{for all}\quad x>0.$$

Also,

$$\max_{k_{i-1}(1+\alpha)\le k\le k_i-a} \frac{N^{-1+2\kappa}}{(k_i-k)\|\mathbf{\Delta}_{N,i}\|^2}\left(\frac{N}{(k-\hat{k}_{i-1})(\hat{k}_{i+1}-k)}\right)^{2\kappa}\|\mathbf{Q}_{3,i}(k)-\mathbf{Q}_{3,i}(k_i)\|\|\mathbf{Q}_{1,i}(k)\|$$
$$=O_P\left(\frac{1}{N^{1/2}\|\mathbf{\Delta}_{N,i}\|}\right)=o_P(1).$$

These yield that for all $x>0$

$$\lim_{C\to\infty}\limsup_{N\to\infty}P\left\{\max_{k_{i-1}(1+\alpha)\le k\le k_i-a}\left[2\left(\left(\frac{N}{(k-\hat{k}_i)(\hat{k}_{i+1}-k)}\right)^{2\kappa}\mathbf{Q}_{1,i}^\top(k)\mathbf{Q}_{3,i}(k)\right.\right.\right.$$
$$\left.\left.\left.-\left(\frac{N}{(k_i-\hat{k}_i)(\hat{k}_{i+1}-k_i)}\right)^{2\kappa}\mathbf{Q}_{1,i}^\top(k_i)\mathbf{Q}_{3,i}(k_i)\right)+\frac{1}{2}V_{N,i}(k)\right]>x\right\}=0.$$

We conclude

$$\lim_{C\to\infty}\limsup_{N\to\infty}P\left\{\hat{\mathfrak{k}}_{N,i}<k_i-C/\|\mathbf{\Delta}_{N,i}\|^2\right\}=0.$$

By symmetry,

$$\lim_{C\to\infty}\limsup_{N\to\infty}P\left\{\hat{\mathfrak{k}}_{N,i}>k_i+C/\|\mathbf{\Delta}_{N,i}\|^2\right\}=0,$$

which concludes the proof when $0\le\kappa<1/2$. The proof when $\kappa=1/2$ requires some minor modifications only. Mainly,

$$\max_{1\le m<n\le N}\frac{1}{(n-m)^\kappa}\left\|\sum_{\ell=m+1}^{n}\mathbf{x}_\ell\epsilon_\ell\right\|=O_P\left(N^{1/2-\kappa}\right)$$

is replaced with

$$\max_{1\le m<n\le N}\frac{1}{(n-m)^{1/2}}\left\|\sum_{\ell=m+1}^{n}\mathbf{x}_\ell\epsilon_\ell\right\|=O_P\left((\log N)^{1/2}\right).$$

□

We defined $\mathbf{Q}_{3,i}(k)$ when $\hat{k}_{i-1} < k \le k_i$, and we extend this definition

$$\mathbf{Q}_{3,i}(k) = \zeta_i \sum_{\ell=k+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top \boldsymbol{\Delta}_{N,i}, \quad k_i < k \le \hat{k}_{i+1}.$$

We also introduce

$$\bar{\zeta}_i = \frac{\theta_i - \theta_{i-1}}{\theta_{i+1} - \theta_{i-1}},$$

$$R_{N,1,i}(k) = \begin{cases} (\hat{k}_{i+1} - \hat{k}_{i-1})^{-1+2\kappa} \left(\left(\frac{\hat{k}_{i+1} - \hat{k}_i}{(k - \hat{k}_{i-1})(\hat{k}_{i+1} - k)} \right)^{2\kappa} \sum_{\ell=\hat{k}_{i-1}+1}^{k} \mathbf{x}_\ell^\top \epsilon_\ell \mathbf{Q}_{3,i}^\top(k) \right. \\ \qquad \left. - \left(\frac{\hat{k}_{i+1} - \hat{k}_i}{(k_i - \hat{k}_{i-1})(\hat{k}_{i+1} - k_i)} \right)^{2\kappa} \sum_{\ell=\hat{k}_{i-1}+1}^{k_i} \mathbf{x}_\ell^\top \epsilon_\ell \mathbf{Q}_{3,i}^\top(k_i) \right), \quad \hat{k}_{i-1} < k \le k_i, \\ (\hat{k}_{i+1} - \hat{k}_{i-1})^{-1+2\kappa} \left(\left(\frac{\hat{k}_{i+1} - \hat{k}_i}{(k - \hat{k}_{i-1})(\hat{k}_{i+1} - k)} \right)^{2\kappa} \left(- \sum_{\ell=k+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell^\top \epsilon_\ell \right) \mathbf{Q}_{3,i}(k) \right. \\ \qquad \left. - \left(\frac{\hat{k}_{i+1} - \hat{k}_i}{(k_i - \hat{k}_{i-1})(\hat{k}_{i+1} - k_i)} \right)^{2\kappa} \left(- \sum_{\ell=k_i+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell^\top \epsilon_\ell \right) \mathbf{Q}_{3,i}(k) \right), \quad k_i < k \le \hat{k}_{i+1}, \end{cases}$$

$$R_{N,2,i} = \begin{cases} (k_{i+1} - k_i)^{-1+2\kappa} \left(\left(\frac{k_{i+1} - k_i}{(k - k_{i-1})(k_{i+1} - k)} \right)^{2\kappa} (1 - \bar{\zeta}_i)^2 (k - k_{i-1})^2 \boldsymbol{\Delta}_{N,i}^\top \mathbf{A} \boldsymbol{\Delta}_{N,i} \right. \\ \qquad \left. - \left(\frac{k_{i+1} - k_i}{(k_i - k_{i-1})(k_{i+1} - k_i)} \right)^{2\kappa} (1 - \bar{\zeta}_i)^2 (k_i - k_{i-1})^2 \boldsymbol{\Delta}_{N,i}^\top \mathbf{A} \boldsymbol{\Delta}_{N,i} \right), \quad \hat{k}_{i-1} < k \le k_i, \\ (k_{i+1} - k_i)^{-1+2\kappa} \left(\left(\frac{k_{i+1} - k_i}{(k - k_{i-1})(k_{i+1} - k)} \right)^{2\kappa} \bar{\zeta}_i^2 (k_{i+1} - k)^2 \boldsymbol{\Delta}_{N,i}^\top \mathbf{A} \boldsymbol{\Delta}_{N,i} \right. \\ \qquad \left. - \left(\frac{k_{i+1} - k_i}{(k_i - k_{i-1})(k_{i+1} - k_i)} \right)^{2\kappa} \bar{\zeta}_i^2 (k_{i+1} - k)^2 \boldsymbol{\Delta}_{N,i}^\top \mathbf{A} \boldsymbol{\Delta}_{N,i} \right), \quad k_i < k \le \hat{k}_{i+1}, \end{cases}$$

$$R_{N,3,i}(k) = \begin{cases} (\hat{k}_{i+1} - \hat{k}_{i-1})^{-1+2\kappa} \left(\dfrac{\hat{k}_{i+1} - \hat{k}_{i-1}}{(k_i - \hat{k}_{i-1})(\hat{k}_{i+1} - k_i)} \right)^{2\kappa} (1 - \bar{\zeta}_i) \left(-\sum_{\ell=k+1}^{k_i} \mathbf{x}_\ell^\top \epsilon_\ell \right) \\ \qquad \times \left(\sum_{\ell=\hat{k}_{i-1}+1}^{k_i} \mathbf{x}_\ell \mathbf{x}_\ell^\top \right) \boldsymbol{\Delta}_{N,i}, \quad \hat{k}_{i-1} < k \le k_i, \\ (\hat{k}_{i+1} - \hat{k}_{i-1})^{-1+2\kappa} \left(\dfrac{\hat{k}_{i+1} - \hat{k}_{i-1}}{(k_i - \hat{k}_{i-1})(\hat{k}_{i+1} - k_i)} \right)^{2\kappa} \bar{\zeta}_i \left(\sum_{\ell=k_i+1}^{k} \mathbf{x}_\ell^\top \epsilon_\ell \right) \left(\sum_{\ell=k_i+1}^{\hat{k}_{i+1}} \mathbf{x}_\ell \mathbf{x}_\ell^\top \right) \boldsymbol{\Delta}_{N,i}, \\ \qquad k_i < k \le \hat{k}_{i+1}, \end{cases}$$

and

$$R_{N,4,i}(k) = \begin{cases} (\bar{\zeta}_i(1 - \bar{\zeta}_i))^{1-2\kappa} \left(-\sum_{\ell=k+1}^{k_i} \mathbf{x}_\ell^\top \epsilon_\ell \right) \mathbf{A} \boldsymbol{\Delta}_{N,i}, & \hat{k}_{i-1} < k \le k_i, \\ (\bar{\zeta}_i(1 - \bar{\zeta}_i))^{1-2\kappa} \left(\sum_{\ell=k_i+1}^{k} \mathbf{x}_\ell^\top \epsilon_\ell \right) \mathbf{A} \boldsymbol{\Delta}_{N,i}, & k_i < k \le \hat{k}_{i+1}. \end{cases}$$

Lemma B.5. *If Assumptions 2.1–2.8 hold, then we have*

$$(k_{i+1} - k_i)^{-1+2\kappa} \max_{|k-k_i| \le C/\|\boldsymbol{\Delta}_{N,i}\|^2} |U_{N,i}(k)| = o_P(1), \tag{B.9}$$

$$\max_{|k-k_i| \le C/\|\boldsymbol{\Delta}_{N,i}\|^2} \Bigg| (k_{i+1} - k_{i-1})^{-1+2\kappa} \Bigg[\left(\frac{k_{i+1} - k_i}{(k_i - \hat{k}_{i-1})(\hat{k}_{i+1} - k)} \right)^{2\kappa} \mathbf{Q}_{1,i}^\top(k) \mathbf{Q}_{3,i}(k) - \left(\frac{k_{i+1} - k_{i-1}}{(k_i - \hat{k}_{i-1})(\hat{k}_{i+1} - k_i)} \right)^{2\kappa} \mathbf{Q}_{1,i}^\top(k_i) \mathbf{Q}_{3,i}(k_i) \Bigg] - R_{N,1,i}(k) \Bigg| = o_P(1), \tag{B.10}$$

$$\max_{|k-k_i| \le C/\|\boldsymbol{\Delta}_{N,i}\|^2} \Bigg| (k_{i+1} - k_{i-1})^{-1+2\kappa} \Bigg[\left(\frac{k_{i+1} - k_i}{(k_i - \hat{k}_{i-1})(\hat{k}_{i+1} - k)} \right)^{2\kappa} \mathbf{Q}_{3,i}^\top(k) \mathbf{Q}_{3,i}(k) - \left(\frac{k_{i+1} - k_{i-1}}{(k_i - \hat{k}_{i-1})(\hat{k}_{i+1} - k_i)} \right)^{2\kappa} \mathbf{Q}_{3,i}^\top(k_i) \mathbf{Q}_{3,i}(k_i) \Bigg] - R_{N,2,i}(k) \Bigg| = o_P(1), \tag{B.11}$$

$$\max_{|k-k_i| \le C/\|\boldsymbol{\Delta}_{N,i}\|^2} |R_{N,1,i}(k) - R_{N,3,i}(k)| = o_P(1) \tag{B.12}$$

and

$$\max_{|k-k_i|\le C/\|\boldsymbol{\Delta}_{N,i}\|^2} |R_{N,2,i}(k)-R_{N,4,i}(k)| = o_P(1). \tag{B.13}$$

Proof. The proofs of (B.9)–(B.13) go along the lines of Lemma B.4 so the details are omitted. □

Lemma B.5 means that in the decomposition of (B.5) the dominating terms contain $\mathbf{Q}_{3,i}^\top(k)\mathbf{Q}_{1,i}(k)$ and $\mathbf{Q}_{3,i}^\top(k)\mathbf{Q}_{3,i}(k)$. These terms are approximated with the sequences $R_{N,3,i}(k)$ and $R_{N,4,i}(k)$. Next we define the continuous version of $\mathbf{Q}_{1,i}^\top(k)\mathbf{Q}_{3,i}(k)$:

$$Z_{N,1,i}(s) = \|\boldsymbol{\Delta}_{N,i}\|\left(-\sum_{\ell=k_i-s/\|\boldsymbol{\Delta}_{N,i}\|^2}^{k_i}\mathbf{x}_\ell\epsilon_\ell\right)\mathbf{A}\frac{\boldsymbol{\Delta}_{N,i}}{\|\boldsymbol{\Delta}_{N,i}\|}, \quad s\ge 0, \tag{B.14}$$

and

$$Z_{N,2,i}(s) = \|\boldsymbol{\Delta}_{N,i}\|\left(\sum_{\ell=k_i+1}^{k_i-s/\|\boldsymbol{\Delta}_{N,i}\|^2}\mathbf{x}_\ell\epsilon_\ell\right)\mathbf{A}\frac{\boldsymbol{\Delta}_{N,i}}{\|\boldsymbol{\Delta}_{N,i}\|}, \quad s\ge 0, \tag{B.15}$$

We recall the definition of $\bar a(\theta_i)$ and $\bar a(\theta_i+)$ from (2.20).

Lemma B.6. *We assume that Assumptions 2.1–2.8 hold. Let $\{W_i^{(1)}(s), W_i^{(2)}(s), s\ge 0, 1\le \ell\le M+1\}$ be independent Wiener processes. The processes $\{Z_{N,1,i}(s), Z_{N,2,i}(s), s\ge 0, 1\le i\le K\}$ converge jointly in $\mathscr{D}^{2K}[0,C]$ to independent scaled Wiener processes.*
(i) If $\tau_{\ell-1}<\theta_i<\theta_\ell$, then

$$\{Z_{N,1,i}(s), Z_{N,2,i}(s), 0\le s\le C\} \stackrel{\mathscr{D}^2[0,C]}{\longrightarrow} \{\bar a(\theta_i)W_i^{(1)}(s), \bar a(\theta_i)W_i^{(2)}(s), 0\le s\le C\}.$$

(ii) If $\theta_i=\tau_\ell$, then

$$\{Z_{N,1,i}(s), Z_{N,2,i}(s), 0\le s\le C\} \stackrel{\mathscr{D}^2[0,C]}{\longrightarrow} \{\bar a(\theta_i)W_i^{(1)}(s), \bar a(\theta_i+)W_i^{(2)}(s), 0\le s\le C\}.$$

Proof. The process $\mathbf{x}_\ell\epsilon_\ell$ is stationary on the intervals $(m_{k-1}, m_k]$ so the result, including the independence of the limiting processes follows from Theorem S2.1 of Aue et al. (2014). □

We also use the following version of Lemma A.6.

Lemma B.7. *If Assumptions 2.1–2.8 hold, then we have*

$$\sup_{|s|\le C}\left|R_{N,2,i}(k_i+s\gamma/\|\boldsymbol{\Delta}_{N,i}\|^2)+2\gamma(\bar\zeta_i(1-\bar\zeta_i))^{1-2\kappa}|s|m_\kappa(s,\bar\zeta_i)\|\mathbf{A}\mathbf{d}_i\|^2\right| = o(1).$$

Proof of Theorem 2.2.

Proof. Based on Lemmas B.3–B.5, we need to consider the process

$$Z_{N,i}(s) = 2R_{N,4,i}(k_i+\gamma s/\|\boldsymbol{\Delta}_{N,i}\|^2)+R_{N,2,i}(k_i+\gamma s/\|\boldsymbol{\Delta}_{N,i}\|^2), \quad -C\le s\le C,$$

for all $C > 0$. If $\tau_{\ell-1} < \theta_i < \theta_\ell$, then by Lemmas B.6 and B.7 we have

$$Z_{N,i}(s) \stackrel{\mathscr{D}[-C,C]}{\longrightarrow} 2(\bar{\zeta}_i(1-\bar{\zeta}_i))^{1-2\kappa}\left(\gamma^{1/2}\bar{a}(\theta_i)W(s) - \gamma|s|m_\kappa(s,\bar{\zeta}_i)\|\mathbf{A}\mathbf{d}_i\|^2\right),$$

where $\{W(s), -\infty < s < \infty\}$ is a two sided Wiener process. Choosing

$$\gamma = \left(\frac{\bar{a}(\theta_i)+\bar{a}(\theta_i+)}{2}\right)^2 \frac{1}{\|\mathbf{A}\mathbf{d}_i\|^4}, \tag{B.16}$$

we get

$$\begin{aligned}
&2(\bar{\zeta}_i(1-\bar{\zeta}_i))^{1-2\kappa}\left(\gamma^{1/2}\bar{a}(\theta_i)W(s) - \gamma|s|m_\kappa(s,\bar{\zeta}_i)\|\mathbf{A}\mathbf{d}_i\|^2)\right)\\
&\qquad = 2(\bar{\zeta}_i(1-\bar{\zeta}_i))^{1-2\kappa}\left(\frac{\bar{a}(\theta_i)+\bar{a}(\theta_i+)}{2}\right)^2 \frac{1}{\|\mathbf{A}\mathbf{d}_i\|^2}\left(\frac{2\bar{\theta}_i}{\bar{a}(\theta_i)+\bar{a}(\theta_i+)}W(s) - |s|m_\kappa(s,\bar{\zeta}_i)\right).
\end{aligned}$$

If $\theta_i = \tau_\ell$, then

$$\begin{aligned}
Z_{N,i}(s) &\stackrel{\mathscr{D}[-C,C]}{\longrightarrow} 2(\bar{\zeta}_i(1-\bar{\zeta}_i))^{1-2\kappa}\\
&\quad\times\left(\gamma^{1/2}\bar{a}(\theta_i)W(s)\mathbb{1}\{s\le 0\} + \gamma^{1/2}\bar{a}(\theta_i+)W(s)\mathbb{1}\{s>0\} - \gamma|s|m_\kappa(s,\bar{\zeta}_i)\|\mathbf{A}\mathbf{d}_i\|^2\right).
\end{aligned}$$

Using again γ in (B.16) we get

$$\begin{aligned}
&2(\bar{\zeta}_i(1-\bar{\zeta}_i))^{1-2\kappa}\left(\gamma^{1/2}\bar{a}(\theta_i)W(s)\mathbb{1}\{s\le 0\} + \gamma^{1/2}\bar{a}(\theta_i+)W(s)\mathbb{1}\{s>0\} - \gamma|s|m_\kappa(s,\bar{\zeta}_i)\|\mathbf{A}\mathbf{d}_i\|^2\right)\\
&\quad = 2(\bar{\zeta}_i(1-\bar{\zeta}_i))^{1-2\kappa}\left(\frac{\bar{a}(\theta_i)+\bar{a}(\theta_i+)}{2}\right)^2 \frac{1}{\|\mathbf{A}\mathbf{d}_i\|^2}\left(\frac{2\bar{a}(\theta_i)}{\bar{a}(\theta_i)+\bar{a}(\theta_i+)}W(s)\mathbb{1}\{s\le 0\}\right.\\
&\qquad\qquad \left. + \frac{2\bar{a}(\theta_i+)}{\bar{a}(\theta_i)+\bar{a}(\theta_i+)}W(s)\mathbb{1}\{s>0\} - |s|m_\kappa(s,\bar{\zeta}_i)\right).
\end{aligned}$$

Since we can choose C is large as we wish, Theorem 2.2 is proven. □

Proof of Theorem 2.3.

Proof. Let $\mathscr{A}_N = \{|\hat{\theta}_r - \theta_r| \le \varepsilon \;\text{ for all }\; 1 \le r \le R\} \cap \{\hat{R}_N = R\}$. Then

$$P(\{k_1, ..., k_R\} \subseteq \mathscr{S}_N) = P(\{\{k_1, ..., k_R\} \subseteq \mathscr{S}_N\} \cap \mathscr{A}_N) + P(\{\{k_1, ..., k_R\} \subseteq \mathscr{S}_N\} \cap \mathscr{A}_N^c).$$

The second term on the right hand side above vanishes as N tends to infinity by Assumption 2.3. Further, under Assumption 2.3 $\|\hat{\mathbf{\Delta}}_{N,i} - \mathbf{\Delta}_{N,i}\| \stackrel{P}{\to} 0$, and $\max\{|\hat{a}(\theta_i) - \bar{a}(\theta_i)|, |\hat{a}(\theta_i+) - \bar{a}(\theta_i+)|\} \stackrel{P}{\to} 0$. Hence, by Theorem 2.2 and Slutsky's Theorem

$$P(\{\{k_1, ..., k_R\} \subseteq \mathscr{S}_N\} \cap \mathscr{A}_N) \to (1-\alpha)^R,$$

as needed. □

Appendix C. Estimation of the long run variances when $R = 1$

Assumption C.1. *$\mathscr{K}()$ is a symmetric, bounded function with a compact support, $\mathscr{K}(0) = 1$ and $\mathscr{K}()$ is Lipschitz continuous at 0.*

Assumption C.2. *(i) $\wp_{1,N} \to \infty, \wp_{2,N} \to \infty, \wp_{1,N}/N \to 0, \wp_{2,N}/N \to 0,$*
(ii) $h = h(N) \to \infty, h(N)/\wp_{1,N}^{1/2} \to 0, h(N)/\wp_{2,N}^{1/2} \to 0,$
(iii)

$$\|\boldsymbol{\Delta}_N\|^3 \wp_{1,N}^{1/2} \to 0 \quad \textit{and} \quad \|\boldsymbol{\Delta}_N\|^3 \wp_{2,N}^{1/2} \to 0. \tag{C.1}$$

We note that (C.1) in Assumption C.2 is equivalent with

$$\left|\frac{\hat{k}_N - k^*}{N}\right| \|\boldsymbol{\Delta}_N\| \wp_{1,N}^{1/2} \to 0.$$

Theorem C.1. *We assume that $R = 1$ and Assumptions 2.1–2.7 are satisfied, Assumption 2.5 holds with $\nu > 8$. If in addition Assumptions C.1 and C.2 also hold, then we have*
(i) if $\tau_{\ell-1} < \theta \le \tau_\ell$, then

$$\hat{\mathbf{D}}_{1,N} \xrightarrow{P} \mathbf{D}_\ell;$$

(ii) if $\tau_{\ell-1} < \theta < \tau_\ell$, then

$$\hat{\mathbf{D}}_{2,N} \xrightarrow{P} \mathbf{D}_\ell;$$

(iii) if $\theta = \tau_\ell$, then

$$\hat{\mathbf{D}}_{2,N} \xrightarrow{P} \mathbf{D}_{\ell+1}.$$

Proof. We recall that $\hat{\boldsymbol{\beta}}_1$ is the least square estimator computed from $y_\ell, \mathbf{x}_\ell, 1 \le \ell < \hat{k}_N$. The corresponding residuals satisfy

$$\bar{\epsilon}_{\ell,1} = y_\ell - \mathbf{x}_\ell^\top \hat{\boldsymbol{\beta}}_1 = \begin{cases} \epsilon_\ell - \mathbf{x}_\ell^\top(\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1), & \text{if} \quad 1 \le \ell < k^*, \\ \epsilon_\ell - \mathbf{x}_\ell^\top(\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_2), & \text{if} \quad k^* \le \ell < \hat{k}_N. \end{cases} \tag{C.2}$$

(If $\hat{k}_N < k^*$, then the last line in (C.2) is not needed.) Similarly, the estimator $\hat{\boldsymbol{\beta}}_1$ might contain observations which parameter $\boldsymbol{\beta}_2$, but not more than $(\hat{k}_N - k^*)\mathbb{1}\{k^* < \hat{k}_N\}$. Hence

$$\hat{\boldsymbol{\beta}}_1 = \left\{ \begin{array}{ll} \boldsymbol{\beta}_1, & \text{if} \ \hat{k}_N < k^*, \\ \dfrac{k^*}{\hat{k}_N}\boldsymbol{\beta}_1 + \left(1 - \dfrac{k^*}{\hat{k}_N}\right)\boldsymbol{\beta}_2, & \text{if} \quad k^* \le \hat{k}_N, \end{array} \right\} + \boldsymbol{R}_{N,1}$$

and

$$\|\boldsymbol{R}_{N,1}\| = O_P\left(N^{-1/2}\right).$$

We can assume that $\mathbf{x}_\ell, \epsilon_\ell, \ell < k^*$ are from the same volatility regime. If $\hat{k}_N \le k^*$, we have

$$\mathbf{x}_j \bar{\epsilon}_{j,1} \mathbf{x}_{\ell+j}^\top \bar{\epsilon}_{\ell+j,1} = \mathbf{x}_j \epsilon_j \mathbf{x}_{j+\ell}^\top \epsilon_{j+\ell} + \mathbf{U}_{j,\ell},$$

where

$$\mathbf{U}_{j,\ell} = -\mathbf{x}_j\epsilon_j\boldsymbol{R}_N^\top\mathbf{x}_{j+\ell}\mathbf{x}_{j+\ell}^\top - \mathbf{x}_j\mathbf{x}_j^\top\boldsymbol{R}_{N,1}\mathbf{x}_{j+\ell}^\top\epsilon_{j+\ell} + \mathbf{x}_\ell\mathbf{x}_\ell^\top\boldsymbol{R}_{N,1}\boldsymbol{R}_{N,1}^\top\mathbf{x}_{j+\ell}\mathbf{x}_{j+\ell}^\top.$$

We consider $\mathbf{U}_{j,\ell}$ coordinate wise. If $\mathbf{x}_j = (c_1(j),\ldots,c_d(j))^\top$, $\boldsymbol{R}_N = (r_1,\ldots,r_d)^\top$ and $\mathbf{x}_{j+\ell}\mathbf{x}_{j+\ell}^\top = \{a_{m,n}(j+\ell), 1\le m,n\le d\}$. The (m,n)th coordinate of $\mathbf{x}_j\epsilon_j\boldsymbol{R}_N^\top\mathbf{x}_{j+\ell}\mathbf{x}_{j+\ell}^\top$ is

$$d(m,n) = \sum_{v=1}^d r_v[a_{v,n}(j+\ell)c_m(j)\epsilon_j].$$

It follows from Assumption 2.5 that

$$\left|\sum_{\ell=0}^\infty \frac{1}{\mathscr{P}_{1,N}}\sum_{\ell=\hat{k}_N-\mathscr{P}_{1,N}}^{\hat{k}_N-\ell} E[a_{v,n}(j+\ell)c_m(j)\epsilon_j]\right| < \infty$$

and therefore

$$\left|\sum_{\ell=0}^\infty \frac{1}{\mathscr{P}_{1,N}}\sum_{j=\hat{k}_N-\mathscr{P}_{1,N}}^{\hat{k}_N-\ell} r_v E[a_{v,n}(j+\ell)c_m(j)\epsilon_j]\right| < \infty = O_P\left(N^{-1/2}\right).$$

Using again Assumption 2.5 we get

$$\begin{aligned} E&\left|\sum_{j=\hat{k}_N-\mathscr{P}_{1,N}}^{\hat{k}_N-\ell}\{a_{v,n}(j+\ell)c_m(j)\epsilon_j - E[a_{v,n}(j+\ell)c_m(j)\epsilon_j]\}\right| \\ &\le E\left(\max_{\mathscr{P}_{1,N}<z\le N}\left|\sum_{j=z-\mathscr{P}_{1,N}}^{z-\ell}\{a_{v,n}(j+\ell)c_m(j)\epsilon_j - E[a_{v,n}(j+\ell)c_m(j)\epsilon_j]\}\right|\right) \\ &\le C\mathscr{P}_{1,N}^{1/2}, \end{aligned} \tag{C.3}$$

where C does not depend on ℓ. Thus we get

$$\begin{aligned} &\left|\sum_{\ell=0}^\infty \frac{1}{\mathscr{P}_{1,N}}\sum_{j=\hat{k}_N-\mathscr{P}_{1,N}}^{\hat{k}_N-\ell} r_v\{a_{v,n}(j+\ell)c_m(j)\epsilon_j - E[a_{v,n}(j+\ell)c_m(j)\epsilon_j]\}\right| \\ &= O_P\left(\left(\frac{\mathscr{P}_{1,N}}{N}\right)^{1/2}\right) = o_P(1). \end{aligned}$$

We now conclude

$$\left\|\sum_{\ell=0}^\infty \mathscr{K}\left(\frac{\ell}{h}\right)\frac{1}{\mathscr{P}_{1,N}-\ell}\sum_{j=\hat{k}_N-\mathscr{P}_{1,N}}^{\hat{k}_N-\ell}\mathbf{x}_j\epsilon_j\boldsymbol{R}_N^\top\mathbf{x}_{j+\ell}\mathbf{x}_{j+\ell}^\top\right\| = o_P(1). \tag{C.4}$$

By symmetry,

$$\left\| \sum_{\ell=0}^{\infty} \mathscr{K}\left(\frac{\ell}{h}\right) \frac{1}{\not{p}_{1,N} - \ell} \sum_{j=\hat{k}_N - \not{p}_{1,N}}^{\hat{k}_N - \ell} \mathbf{x}_j \mathbf{x}_j^\top \boldsymbol{R}_{N,1} \mathbf{x}_{j+\ell}^\top \epsilon_{j+\ell} \right\| = o_P(1). \tag{C.5}$$

$$\begin{aligned} &\left\| \sum_{\ell=0}^{\infty} \mathscr{K}\left(\frac{\ell}{h}\right) \frac{1}{\not{p}_{1,N} - \ell} \sum_{j=\hat{k}_N - \not{p}_{1,N}}^{\hat{k}_N - \ell} \mathbf{x}_\ell \mathbf{x}_\ell^\top \boldsymbol{R}_{N,1} \boldsymbol{R}_{N,1}^\top \mathbf{x}_{j+\ell} \mathbf{x}_{j+\ell}^\top \right\| \\ &\le \sum_{\ell=0}^{\infty} \mathscr{K}\left(\frac{\ell}{h}\right) \frac{1}{\not{p}_{1,N} - \ell} \sum_{j=\hat{k}_N - \not{p}_{1,N}}^{\hat{k}_N - \ell} \|\mathbf{x}_\ell\|^2 \|\boldsymbol{R}_{N,1}\|^2 \|\mathbf{x}_{j+\ell}\|^2 \\ &= O_P\left(\frac{\not{p}_{1,N}}{N}\right) = o_P(1). \end{aligned} \tag{C.6}$$

Combining (C.4)–(C.6) we obtain

$$\left\| \sum_{\ell=0}^{\infty} \mathscr{K}\left(\frac{\ell}{h}\right) \frac{1}{\not{p}_{1,N} - \ell} \sum_{j=\hat{k}_N - \not{p}_{1,N}}^{\hat{k}_N - \ell} \left[\mathbf{x}_j \bar{\epsilon}_{j,1} \mathbf{x}_{j+\ell}^\top \bar{\epsilon}_{j+\ell,1} - \mathbf{x}_j \epsilon_j \mathbf{x}_{j+\ell}^\top \epsilon_{j+\ell} \right] \right\| = o_P(1). \tag{C.7}$$

Next, we consider when $k^* < \hat{k}_N$. Assume $j, j+\ell \le k^*$ we have

$$\mathbf{x}_j \bar{\epsilon}_{j,1} \mathbf{x}_{\ell+j}^\top \bar{\epsilon}_{\ell+j,1} = \mathbf{x}_j \epsilon_j \mathbf{x}_{j+\ell}^\top \epsilon_{j+\ell} + \mathbf{U}_{j,\ell},$$

where

$$\begin{aligned} \mathbf{U}_{j,\ell} = &-\mathbf{x}_j \epsilon_j (\boldsymbol{R}_N + (1 - k^*/\hat{k}_N)\boldsymbol{\Delta}_N)^\top \mathbf{x}_{j+\ell} \mathbf{x}_{j+\ell}^\top - \mathbf{x}_j \mathbf{x}_j^\top (\boldsymbol{R}_{N,1} + (1 - k^*/\hat{k}_N)\boldsymbol{\Delta}_N) \mathbf{x}_{j+\ell}^\top \epsilon_{j+\ell} \\ &+ \mathbf{x}_\ell \mathbf{x}_\ell^\top (\boldsymbol{R}_{N,1} + (1 - k^*/\hat{k}_N)\boldsymbol{\Delta}_N)(\boldsymbol{R}_{N,1} + (1 - k^*/\hat{k}_N)\boldsymbol{\Delta}_N)^\top \mathbf{x}_{j+\ell} \mathbf{x}_{j+\ell}^\top. \end{aligned}$$

Now this is essentially the same, except that we have an additional term to $\boldsymbol{R}_{N,1}$, $(\hat{k}_N - k^*)\boldsymbol{\Delta}_N$. Hence, by the arguments used to prove (C.4)–(C.6) we get under Assumption C.2 that (C.7) holds in the present case. The remaining cases can be handled similarly.

Next, we show that

$$\left\| \sum_{\ell=0}^{\infty} \mathscr{K}\left(\frac{\ell}{h}\right) \frac{1}{\not{p}_{1,N} - \ell} \left[\sum_{j=\hat{k}_N - \not{p}_{1,N}}^{\hat{k}_N - \ell} \mathbf{x}_j \epsilon_j \mathbf{x}_{j+\ell}^\top \epsilon_{j+\ell} - \sum_{j=k^* - \not{p}_{1,N}}^{k^* - \ell} \mathbf{x}_j \epsilon_j \mathbf{x}_{j+\ell}^\top \epsilon_{j+\ell} \right] \right\| = o_P(1). \tag{C.8}$$

Assumption 2.5 yields (cf. (C.3))

$$E \left\| \sum_{j=\hat{k}_N - \not{p}_{1,N}}^{\hat{k}_N - \ell} \left\{ \mathbf{x}_j \epsilon_j \mathbf{x}_{j+\ell}^\top \epsilon_{j+\ell} - E[\mathbf{x}_j \epsilon_j \mathbf{x}_{j+\ell}^\top \epsilon_{j+\ell}] \right\} \right\| \le C \not{p}_{1,N}^{1/2}$$

and

$$E\left\|\sum_{j=k^*-\mathscr{p}_{1,N}}^{k^*-\ell}\left\{\mathbf{x}_j\epsilon_j\mathbf{x}_{j+\ell}^\top\epsilon_{j+\ell}-E[\mathbf{x}_j\epsilon_j\mathbf{x}_{j+\ell}^\top\epsilon_{j+\ell}]\right\}\right\|\le C\mathscr{p}_{1,N}^{1/2},$$

where C does not depend on ℓ. Since $h/\mathscr{p}_{1,N}^{1/2}\to 0$, (C.7) is proven.

Now we consider long run covariance matrix estimators based on $\mathbf{x}_j\epsilon_j$. If $\theta_{\ell-1}<\theta<\tau_\ell$, then $\mathbf{x}_j\epsilon_j, k^*-\mathscr{p}_{1,N}\le j\le k^*+\mathscr{p}_{2,N}$ is a stationary sequence. Standard results on long run covariance matrices can be used for the kernel estimator based on stationary observations (see e.g. Andrews (1991)). Hence, we conclude

$$\hat{\mathbf{D}}_{N,1}\xrightarrow{P}\mathbf{D}_\ell \quad\text{and}\quad \hat{\mathbf{D}}_{N,2}\xrightarrow{P}\mathbf{D}_\ell.$$

If $\theta=\tau_\ell$, then $\mathbf{x}_j\epsilon_j, k^*-\mathscr{p}_{1,N}\le j<k^*$ and also $\mathbf{x}_j\epsilon_j, k^*\le j\le k^*+\mathscr{p}_{2,N}$ are stationary variables with different long run covariance matrices. Hence

$$\hat{\mathbf{D}}_{N,1}\xrightarrow{P}\mathbf{D}_\ell \quad\text{and}\quad \hat{\mathbf{D}}_{N,2}\xrightarrow{P}\mathbf{D}_{\ell+1}.$$

□

Appendix D. Estimation of the long run variances when $R>1$

We need the following modification of Assumption C.2 for the $R>1$ case.

Assumption D.1. *For all $1\le i\le R$ we assume*
(i) $\mathscr{p}_{1,N,i}\to\infty, \mathscr{p}_{2,N,i}\to\infty, \mathscr{p}_{1,N,i}/N\to 0, \mathscr{p}_{2,N,i}/N\to 0,$
(ii) $h=h(N)\to\infty, h(N)/\mathscr{p}_{1,N,i}^{1/2}\to 0, h(N)/\mathscr{p}_{2,N,i}^{1/2}\to 0,$
(iii) $\|\boldsymbol{\Delta}_{N,i}\|^3\mathscr{p}_{1,N,i}^{1/2}\to 0$ *and* $\|\boldsymbol{\Delta}_{N,i}\|^3\mathscr{p}_{2,N,i}^{1/2}\to 0.$

Theorem D.1. *We assume that $R\ge 1$ and Assumptions 2.1–2.7 are satisfied, Assumption 2.5 holds with $\nu>8$. If in addition Assumptions C.1 and D.1 also hold, then we have for all $1\le i\le R$*
(i) if $\tau_{\ell-1}<\theta_i\le\tau_\ell$, then

$$\hat{\mathbf{D}}_{1,N,i}\xrightarrow{P}\mathbf{D}_\ell;$$

(ii) if $\tau_{\ell-1}<\theta_i<\tau_\ell$, then

$$\hat{\mathbf{D}}_{2,N,i}\xrightarrow{P}\mathbf{D}_\ell;$$

(iii) if $\theta=\tau_\ell$, then

$$\hat{\mathbf{D}}_{2,N,i}\xrightarrow{P}\mathbf{D}_{\ell+1}.$$

Proof. There is little difference between the estimation of long run covariance matrices in the $R=1$ and $R>1$ cases. In the construction of the estimator of the long run covariance matrices immediately before and after k_i, we use the observations $\mathbf{x}_j\hat{\epsilon}_{j,i.1}, \hat{k}-\mathscr{p}_{1,N,i}\le j<\hat{k}_i$ and $\mathbf{x}_j\hat{\epsilon}_{j,i.1}, \hat{k}_i\le j\le\hat{k}+\mathscr{p}_{2,N,i}$. We also note that $\mathscr{p}_{1,N,i}/\hat{k}_{i-1}\to 0$ and $\mathscr{p}_{2,N,i}/\hat{k}_{i+1}\to 0$ in probability. Hence if $\tau_{\ell-1} < \theta_i < \tau_\ell$, then

$$\hat{\mathbf{D}}_{N,i,1} \xrightarrow{P} \mathbf{D}_\ell \quad \text{and} \quad \hat{\mathbf{D}}_{N,i,2} \xrightarrow{P} \mathbf{D}_\ell.$$

If $\theta_i = \tau_\ell$, then

$$\hat{\mathbf{D}}_{N,i,1} \xrightarrow{P} \mathbf{D}_\ell \quad \text{and} \quad \hat{\mathbf{D}}_{N,i,2} \xrightarrow{P} \mathbf{D}_{\ell+1}.$$

□

Appendix E. Proof of Theorem 3.1

Lemma E.1. *If Assumptions 2.1–2.7(i), 3.1 and 3.2 with* $\kappa = \min(\kappa_1, \kappa_2)$ *hold, then*

$$\|\mathbf{\Delta}_N\|^2 \left| \frac{\hat{k}_N}{N} - \theta \right| = O_P(1).$$

Proof. The proof goes along the lines of Lemma A.5. □

We modify the definitions of $Z_{N,1}(s)$ and $Z_{N,2}(s)$ of (A.6) and (A.7) as

$$\hat{Z}_{N,1}(s) = \|\mathbf{\Delta}_N\| \left(- \sum_{j=k^*-s/\|\mathbf{\Delta}_N\|^2}^{k^*} g(j/N)\mathbf{x}_j^\top \epsilon_j \right) \mathbf{A} \frac{\mathbf{\Delta}_N}{\|\mathbf{\Delta}_N\|}, \quad s \geq 0,$$

and

$$\hat{Z}_{N,2}(s) = \|\mathbf{\Delta}_N\| \left(\sum_{j=k^*}^{k^*+s/\|\mathbf{\Delta}_N\|^2} g(j/N)\mathbf{x}_j^\top \epsilon_j \right) \mathbf{A} \frac{\mathbf{\Delta}_N}{\|\mathbf{\Delta}_N\|}, \quad s > 0.$$

Next we define

$$\bar{\sigma}_L^2(i) = g^2(\theta-)\mathbf{d}^\top \mathbf{A}\mathbf{D}_i\mathbf{A}\mathbf{d} \;\text{ and }\; \bar{\sigma}_R^2(i) = g^2(\theta+)\mathbf{d}^\top \mathbf{A}\mathbf{D}_i\mathbf{A}\mathbf{d}, \;\text{ if }\; \tau_{i-1} < \theta \leq \tau_i, 1 \leq i \leq M+1,$$

where $g(u)$ and $g(u+)$ are the limits of g at u from the left and from the right, respectively and $\mathbf{d}$ is from (2.6).

Lemma E.2. *We assume that Assumptions 2.1–2.7(i), 3.1 and 3.2 with* $\kappa = \min(\kappa_1, \kappa_2)$ *hold.*

(i) If $\tau_{i-1} < \theta < \tau_i$, *then*

$$\left\{ \hat{Z}_{N,1}(s), \hat{Z}_{N,2}(s), 0 \leq s \leq C \right\} \xrightarrow{\mathscr{D}^2[0,C]} \left\{ \bar{\sigma}_L(i)W^{(1)}(s), \bar{\sigma}_R(i)W^{(2)}(s), 0 \leq s \leq C \right\}. \tag{E.1}$$

(ii) If $\theta = \tau_i$, *then*

$$\left\{ \hat{Z}_{N,1}(s), \hat{Z}_{N,2}(s), 0 \leq s \leq C \right\} \xrightarrow{\mathscr{D}^2[0,C]} \left\{ \bar{\sigma}_L(i)W^{(1)}(s), \bar{\sigma}_R(i+1)W^{(2)}(s), 0 \leq s \leq C \right\}, \tag{E.2}$$

where $\{W^{(1)}(u), u \geq 0\}$ *and* $\{W^{(2)}(u), u \geq 0\}$ *are independent Wiener processes.*

Proof. Due to Assumption 2.5, the sequence $\mathbf{x}_\ell \epsilon_\ell$ is stationary on $(m_{i-1}, m_i]$, and $[k^* - C/\|\mathbf{\Delta}_N\|^2, k^* + C/\|\mathbf{\Delta}_N\|^2] \subset (m_{i-1}, m_i]$. Hence Aue et al. (2014) yields that

$$\left\{ \|\mathbf{\Delta}_N\| \sum_{j=k^*-s/\|\mathbf{\Delta}\|^2}^{k^*} \mathbf{x}_j \epsilon_j, \;\; \|\mathbf{\Delta}_N\| \sum_{j=k^*}^{k^*+s/\|\mathbf{\Delta}\|^2} \mathbf{x}_j \epsilon_j, 0 \le s \le C \right\} \xrightarrow{\mathscr{D}^{2d}[0,C]} \left\{ \mathbf{W}^{(1)}(s), \mathbf{W}^{(2)}(s), 0 \le s \le C \right\}, \tag{E.3}$$

where $\mathbf{W}^{(1)}(s)$ and $\mathbf{W}^{(2)}(s)$ are independent, identically distributed Gaussian processes with $E\mathbf{W}^{(1)}(s) = E\mathbf{W}^{(2)}(s) = \mathbf{0}$ and

$$E\mathbf{W}^{(1)}(s) \left(\mathbf{W}^{(1)}(t)\right)^\top = E\mathbf{W}^{(2)}(s) \left(\mathbf{W}^{(2)}(t)\right)^\top = \min(t, s)\mathbf{D}_i.$$

We write

$$\|\mathbf{\Delta}_N\| \sum_{j=k^*-s/\|\mathbf{\Delta}\|^2}^{k^*} g(j/N)\mathbf{x}_j \epsilon_j = \int_{-s}^{0} g\left(x\|\mathbf{\Delta}_N\|^2/N + k^*/N\right) d \left(\|\mathbf{\Delta}_N\| \sum_{j=k^*+x/\|\mathbf{\Delta}_N\|^2}^{k^*} \mathbf{x}_j \epsilon_j \right)$$

and

$$\|\mathbf{\Delta}_N\| \sum_{j=k^*}^{k^*+s/\|\mathbf{\Delta}\|^2} g(j/N)\mathbf{x}_j \epsilon_j = \int_{0}^{s} g\left(x\|\mathbf{\Delta}_N\|^2/N + k^*/N\right) d \left(\|\mathbf{\Delta}_N\| \sum_{j=k^*}^{k^*+x/\|\mathbf{\Delta}_N\|^2} \mathbf{x}_j \epsilon_j \right).$$

Since

$$\sup_{-C \le x < 0} \left| g\left(x\|\mathbf{\Delta}_N\|^2/N + k^*/N\right) - g(\theta) \right| \to 0$$

and

$$\sup_{0 < x \le C} \left| g\left(x\|\mathbf{\Delta}_N\|^2/N + k^*/N\right) - g(\theta+) \right| \to 0,$$

it follows from (E.3) and Theorem 2.7 Kurtz and Protter (1991) that

$$\left\{ \|\mathbf{\Delta}_N\| \sum_{j=k^*-s/\|\mathbf{\Delta}\|^2}^{k^*} g(j/N)\mathbf{x}_j \epsilon_j, \;\; \|\mathbf{\Delta}_N\| \sum_{j=k^*}^{k^*+s/\|\mathbf{\Delta}\|^2} g(j/N)\mathbf{x}_j \epsilon_j, 0 \le s \le C \right\} \xrightarrow{\mathscr{D}^{2d}[0,C]} \left\{ g(\theta)\mathbf{W}^{(1)}(s), g(\theta+)\mathbf{W}^{(2)}(s), 0 \le s \le C \right\}.$$

Thus we conclude

$$\left\{ \hat{Z}_{N,1}(s), \hat{Z}_{N,2}(s), 0 \le s \le C \right\} \xrightarrow{\mathscr{D}^{2}[0,C]} \left\{ g(\theta)\left(\mathbf{W}^{(1)}(s)\right)^\top \mathbf{Ad}, g(\theta+)\left(\mathbf{W}^{(2)}(s)\right)^\top \mathbf{Ad}, 0 \le s \le C \right\}.$$

We observe that

$$Eg(\theta)\left(\mathbf{W}^{(1)}(s)\right)^\top \mathbf{Ad} = Eg(\theta+)\left(\mathbf{W}^{(2)}(s)\right)^\top \mathbf{Ad} = 0$$

$$E\left[g(\theta)\left(\mathbf{W}^{(1)}(s)\right)^{\top}\mathbf{Ad}g(\theta)\left(\mathbf{W}^{(1)}(t)\right)^{\top}\mathbf{Ad}\right]=\sigma_L^2(i)\min(t,s)$$

and

$$E\left[g(\theta+)\left(\mathbf{W}^{(1)}(s)\right)^{\top}\mathbf{Ad}g(\theta+)\left(\mathbf{W}^{(1)}(t)\right)^{\top}\mathbf{Ad}\right]=\sigma_R^2(i)\min(t,s),$$

hence, the lemma is proven. □

Proof of Theorem 3.1

We follow the proof of Theorem 2.1. We recall the definition of $Z_N(s)$ in (A.8). Putting together Lemmas A.6 and (E.2) we get

$$Z_N(s)\overset{\mathscr{D}[-C,C]}{\longrightarrow}2(\theta(1-\theta))^{1-2\kappa}\left(\varsigma^{1/2}\bar{W}(s)-\varsigma|s|m_\kappa(s)\|\mathbf{Ad}\|^2\right),$$

where

$$\hat{W}(s)=\begin{cases}g(\theta)a(\theta)W_1(-s), & \text{if } s\le 0,\\ g(\theta+)a(\theta+)W_2(s), & \text{if } s>0,\end{cases}$$

$\{W_1(s), s\geq 0\}$ and $\{W_2(s), s\geq 0\}$ are independent Wiener processes. Choosing $\varsigma = 1/\|\mathbf{Ad}\|^4$, we get the result. □

Appendix F. Algorithm: Confidence Intervals with Multiple Change Points

This section details the procedure for constructing confidence intervals for multiple change points. Algorithm 1 consists of five steps, with Steps 1-3 describing the estimation of up to R refined change points. Step 1 implements the binary segmentation method, as illustrated in Example 1 of Figure F.1. Once the first break $\hat{k}_1$ is identified, the sample is segmented to estimate $\hat{k}_2'$ and $\hat{k}_2$ in the subsamples before and after $\hat{k}_1$, respectively. We then compute the residual sum of squares (RSS) of model (1.1) for two candidate sets of break points: $\{\hat{k}_1,\hat{k}_2'\}$ and $\{\hat{k}_1,\hat{k}_2\}$. If the RSS of the model with $\{\hat{k}_1,\hat{k}_2\}$ is smaller, $\hat{k}_2$ is retained as the second break, while $\hat{k}_2'$ remains a candidate for the next iteration.

Given the second break $\hat{k}_2$, the sample after $\hat{k}_1$ is further divided into segments $y_{\{\hat{k}_1+1,\dots,\hat{k}_2\}}$, $\mathbf{x}_{\{\hat{k}_1+1,\dots,\hat{k}_2\}}$, and $y_{\{\hat{k}_2+1,\dots,N\}}$, $\mathbf{x}_{\{\hat{k}_2+1,\dots,N\}}$, and the next potential break points $\hat{k}_3$ and $\hat{k}_3'$ are estimated. We again compare the RSS from models (1.1) with candidate break sets $\{\hat{k}_1,\hat{k}_2,\hat{k}_2'\}$, $\{\hat{k}_1,\hat{k}_2,\hat{k}_3'\}$, and $\{\hat{k}_1,\hat{k}_2,\hat{k}_3\}$. If the model with $\{\hat{k}_1,\hat{k}_2,\hat{k}_3\}$ yields the smallest RSS, $\hat{k}_3$ is selected as the third break, while $\hat{k}_2'$ and $\hat{k}_3'$ are retained as candidate break points for the next iteration. The estimation procedure terminates once the number of identified breaks reaches the maximum value R, or when a subsample contains fewer than 20 observations, in which case the long-run variance cannot be reliably estimated.

Once all candidate change points are identified, Step 2 permutes these points to select the model with the appropriate number of breaks using the BIC criterion. The selected change points are then refined by re-estimating a single break within each corresponding segment, as illustrated in the lower panel of Figure F.1. Steps 4–5 outline the construction of confidence intervals for these refined change points. We implement the estimator $\hat{k}_N^{HET}(1/2)$ for each refined change point. Given the several parameters involved, we concisely summarize them below.

Summary of parameter estimation:

Recall Equation (2.11):

$$\left[\hat{k}_N(\kappa) - \left(\frac{\hat{\xi}_{1-\alpha/2}(\kappa,\hat{\theta})\hat{\sigma}^*}{\|\hat{\mathbf{\Delta}}_N\|^2}\right),\ \hat{k}_N(\kappa) - \left(\frac{\hat{\xi}_{\alpha/2}(\kappa,\hat{\theta})\hat{\sigma}^*}{\|\hat{\mathbf{\Delta}}_N\|^2}\right)\right].$$

The change point estimate $\hat{k}_N(\kappa)$ and break fraction estimate $\hat{\theta}$ depend on the tuning weight parameter κ, which is set to $\kappa = 0.5$ by default. The change magnitude $\hat{\mathbf{\Delta}}_N$ is computed by comparing the model coefficients before and after $\hat{k}_N(\kappa)$.

The term $\hat{\sigma}^*$ is then estimated using Equation (2.7), as recalled below

$$\hat{\sigma}^* = \left(\frac{\hat{a}(\theta)+\hat{a}(\theta+)}{2}\right)^2 \frac{1}{\|\hat{\mathbf{A}}\hat{\mathbf{d}}\|^4}.$$

In particular, the sandwich variance object $\hat{a}(\theta)$ (or $\hat{a}(\theta+)$) is estimated as $\hat{a}^2(\theta) = \hat{\mathbf{d}}^\top\hat{\mathbf{A}}\hat{\mathbf{D}}_{1,N}\hat{\mathbf{A}}\hat{\mathbf{d}}$ (replacing $\hat{\mathbf{D}}_{1,N}$ with $\hat{\mathbf{D}}_{2,N}$ for $\hat{a}(\theta+)$). The vector $\hat{\mathbf{d}}$ represents the direction of change and is defined as $\hat{\mathbf{d}} = \hat{\mathbf{\Delta}}_N/\|\hat{\mathbf{\Delta}}_N\|$. The matrix $\hat{\mathbf{A}}_N$ denotes the covariance of the covariates, estimated using the sample covariance. The long-run variance parameters $\mathbf{D}_{1,N}$ ($\mathbf{D}_{2,N}$) are estimated with tuning parameter $\wp_{1,N}$ ($\wp_{2,N}$). By default, we employ the Bartlett kernel with the Andrews (1991) optimal bandwidth selection, using observations from $\{\hat{k}_N - \wp_{1,N} : \hat{k}_N\}$ for $\hat{\mathbf{D}}_{1,N}$ (from $\{\hat{k}_N + 1 : \hat{k}_N + \wp_{2,N}\}$ for $\hat{\mathbf{D}}_{2,N}$). We set $\wp_{1,N} = \wp_{2,N} = \lfloor\sqrt{N}\rfloor$.

The sample-specific limit distribution $\hat{\xi}(\kappa,\hat{\theta})$ is generated by (2.8)–(2.10) with $\kappa = 0.5$, using the estimated values of $\hat{a}(\theta)$ and $\hat{a}(\theta+)$.

Algorithm 1

Input: Dependent variable y_i with covariates $\mathbf{x}_i$, $1 \le i \le N$, and a maximum number of structural breaks R.

Step 1: Given data $y_{\{1,\dots,N\}}$ and $\mathbf{x}_{\{1,\dots,N\}}$, estimate a single change point $\hat{k}_1$ using $\hat{\mathbf{Z}}_N$ in (2.1), with the weight parameter $\kappa = 1/2$. Split the sample at $\hat{k}_1$ and estimate $\hat{k}_2$ in the subsample $\{y_{\{1,\dots,\hat{k}_1\}}, \mathbf{x}_{\{1,\dots,\hat{k}_1\}}\}$ and $\hat{k}_3$ in $\{y_{\{\hat{k}_1+1,\dots,N\}}, \mathbf{x}_{\{\hat{k}_1+1,\dots,N\}}\}$. Repeat this binary segmentation procedure until R change points are identified.

Step 2: Select the number of change points using the BIC criterion in (4.2), as discussed in Section 4. For each selected change point $\hat{k}_r$, refine its estimated location by re-estimating a single change point within the corresponding segment, denoted as $\hat{\mathfrak{k}}_r$.

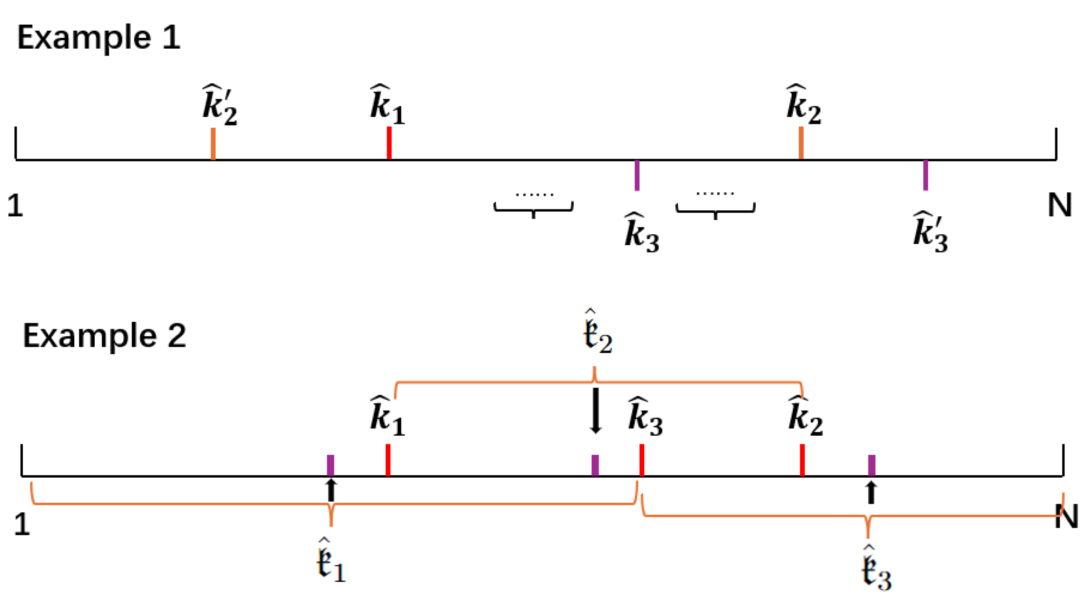
Step 3: For the first refined change point $\hat{\mathfrak{k}}_1$, estimate $a(\theta)$ and $a(\theta+)$ according to (2.14) and (2.15), with $\wp_{1,N} = \wp_{2,N} = \lfloor \sqrt{N} \rfloor$. Simulate the limit distribution of $\xi(\kappa)$ from (2.8), using a Wiener process generated with the estimated $\hat{a}(\theta)$ and $\hat{a}(\theta+)$.

Step 4: Estimate the magnitude of each change, $\hat{\Delta}_{1,N}$, by fitting least squares regressions before and after $\hat{\mathfrak{k}}_1$ up to $\hat{\mathfrak{k}}_2$. Obtain $\hat{\mathbf{A}}_N$ from the full sample and compute σ^* using (2.7). Construct the confidence interval for $\hat{\mathfrak{k}}_1$ following (2.11), denoted as the estimator $k_N^{HET}(1/2)$.

Step 5: Repeat Steps 3–4 to construct confidence intervals for the remaining refined change points.

Output: Confidence intervals for selected and refined R change points.

FIGURE F.1. Illustrative examples of estimating multiple and refined change points. The upper panel of Example 1 demonstrates the procedure for estimating multiple change points up to R changes, while the lower panel of Example 2 illustrates the estimation of refined change points.



Appendix G. further simulation results

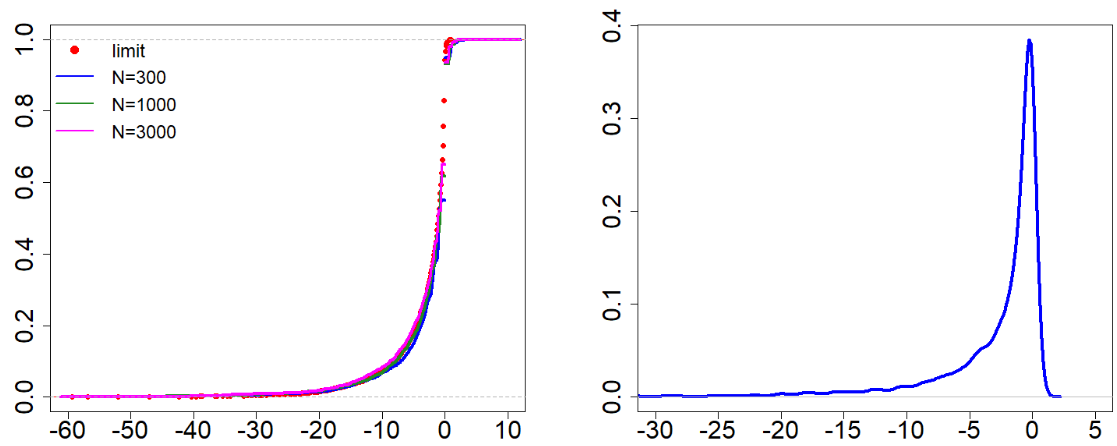
Figure G.1 illustrates the approximation of the limiting distribution in Theorem 2.1 when a single change occurs at $k_1 = \lfloor 0.5N \rfloor$ with $\mathbf{\Delta}_N = 3N^{-1/5}$. The change point estimator $\hat{k}_N(\kappa)$ is computed via (2.4), using $\kappa = 1/2$. The results show good approximation, with the limiting density displaying asymmetry due to heteroscedasticity before and after the change point.

We then assess the empirical coverage rates for $N = 300$ under a single change point at the midpoint of the sample, $k_1 = \lfloor 0.5N \rfloor$, with either a variance change at $m_1 = \lfloor 0.2N \rfloor$ or at the identical location $m_1 = \lfloor 0.5N \rfloor$. In addition to $\hat{k}_N^{HET}(\kappa)$ (with $\not{p}_{1,N} = \not{p}_{2,N} = \lfloor N^{1/2} \rfloor$), we also consider $\hat{k}_N^{Hx}(\kappa)$ (with $\not{p}_{1,N} + \not{p}_{2,N} = N$), which represents the case of choosing larger values for $\not{p}_{1,N}$ and $\not{p}_{2,N}$. The data generating processes remain the same as in Section 5. Figures G.2 and G.3 present the empirical coverage rates and confidence interval lengths for models with homoscedastic and heteroscedastic errors, respectively. Compared with Figures 5.1 and 5.2 in Section 5, the empirical coverage is closer to the nominal level when the change is smaller and the confidence intervals are also narrower in general. The simulation indicates that the estimators' performances improve when the change point occurs near the center of the sample. In particular, to compare $\hat{k}_N^{HET}(1/2)$ and $\hat{k}_N^{Hx}(1/2)$, the two estimators perform similarly when $m_1 = k_1$. However, $\hat{k}_N^{HET}(1/2)$ outperforms $\hat{k}_N^{Hx}(1/2)$ when $m_1 \neq k_1$, as the latter may be affected by estimating the long-run variance from a subsample that includes a variance change.

When the sample size is $N = 100$ and $k_1 = \lfloor 0.2N \rfloor$ with $m_1 = \lfloor 0.2N \rfloor$ or $\lfloor 0.5N \rfloor$ when $N = 100$. Figures G.4 and G.5 show that compared to Figures 5.1 and 5.2, the empirical coverage rates decrease and the lengths of the confidence intervals increase. The simulation indicates that estimator performance deteriorates in smaller samples. In both simulations, the relative performance of the methods remains consistent with the main results in Section 5: the confidence intervals $\hat{k}_N^{Hx}(1/2)$ and $\hat{k}_N^{HET}(1/2)$ generally outperform the alternatives, with $\hat{k}_N^{HET}(1/2)$ performing best when coefficient shifts k_1 are not aligned with changes in error variance m_1.

Following the simulation results, we now focus on the estimator $\hat{k}_N^{HET}(1/2)$. The asymptotic behavior of $\hat{k}_N^{HET}(1/2)$ depends on the tuning parameters $\not{p}_{1,N}$ and $\not{p}_{2,N}$, for which we previously set $\not{p}_{1,N} = \not{p}_{2,N} = \lfloor N^{1/2} \rfloor$. We next evaluate alternative choices of: Size 1 – $\not{p}_{1,N} = \not{p}_{2,N} = \lfloor N^{2/5} \rfloor$, Size 2 – $\not{p}_{1,N} = \not{p}_{2,N} = \lfloor N^{1/3} \rfloor$, Size 3 – $\not{p}_{1,N} = \not{p}_{2,N} = \lfloor \log(N) \rfloor$, and Size 4 – $\not{p}_{1,N} = \not{p}_{2,N} = \lfloor N^{1/4} \rfloor$. Figure G.6 reports the results under homoscedastic errors (**IID**, **AR**) and heteroscedastic errors (**HeteIID**, **HeteAR**, **HeteSmooth**). Across all scenarios, selecting larger values of $\not{p}_{1,N}$ and $\not{p}_{2,N}$ yields slightly higher empirical coverage rates and wider confidence intervals. Note that increasing $\not{p}_{1,N}$ and $\not{p}_{2,N}$ enhances the estimator by incorporating more data in the long-run variance estimation, and this is the case when

FIGURE G.1. Approximation of Limiting Distributions in Theorems 2.1. The left subplot corresponds to the model with **HeteIID** heteroscedastic errors and illustrates the approximation of $\frac{\|\mathbf{\Delta}_N\|^2}{\sigma^*(\theta)}(\hat{k}_N(1/2)-k^*) \xrightarrow{\mathscr{D}} \xi(1/2)$, where σ^* and $\xi(1/2)$ are defined in (2.7) and (2.8), respectively. The right subplot corresponds to the density of $\xi(1/2)$. The statistics are calculated with sample sizes $N = 300$, 1000, and 3000.



(a) Theorem 2.1 (b) Density

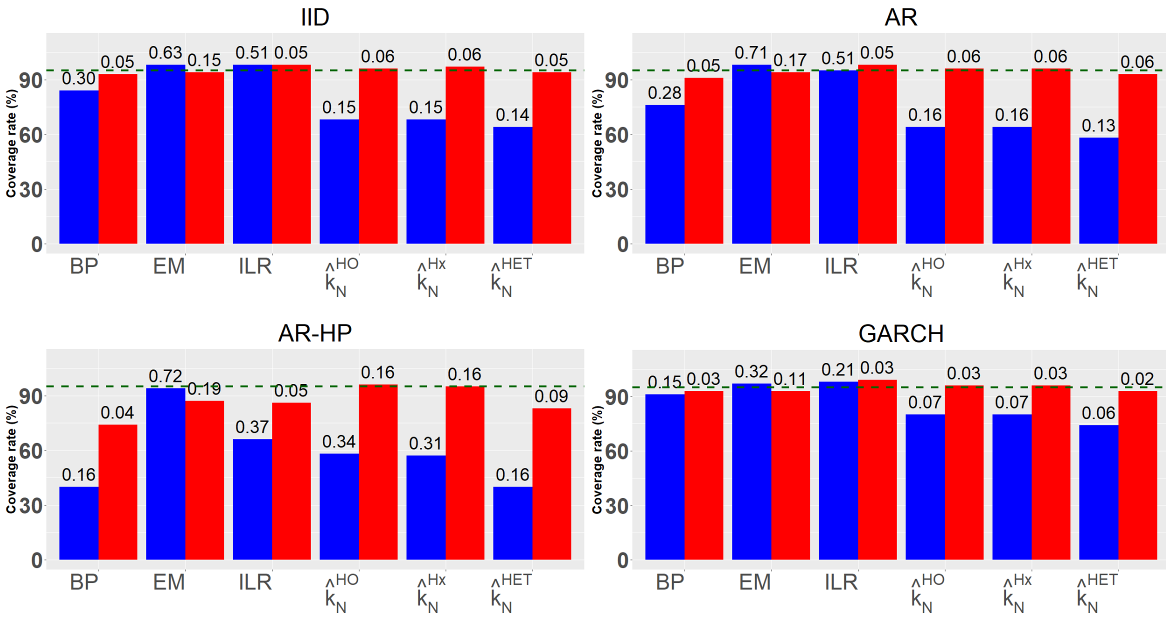
there is no error variance change occuring over these observations. Otherwise, changes in error variance can distort the performance of $\hat{k}_N^{HET}(1/2)$, analogous to the distortion observed when applying $\hat{k}_N^{HET}(1/2)$ under $k_1 \neq m_1$.

Next, we evaluate the effect of the choice of the weight parameter κ. Figure G.7 shows the empirical coverage and the lengths of confidence intervals of the estimator $\hat{k}_N^{HET}(\kappa)$ when there is an early or a middle change.

By setting the sample size $N = 300$, we find that increasing κ from 0 to 0.45 leads to higher empirical coverage rates and narrower confidence intervals when the change occurs early, i.e., $k^* = \lfloor 0.2N \rfloor$. In contrast, the opposite pattern emerges when the change occurs in the middle of the sample. This effect is slightly weaker in heteroscedastic models when the structural change occurs near the center of the sample. Considering the overall performance in contrast with the results in the main text, we suggest choosing $\kappa = 1/2$ for general superiority regardless of the location of the change points.

We also examine the impact of signal-to-noise ratios on the performance of the estimator $\hat{k}_N^{HET}(1/2)$ in homoscedastic and heteroscedastic models. Here we consider a parameter r adjusting the signal-to-noise ratios. By fixing the noise at the same level in all heteroscedastic DGPs, we adjust the signal-to-noise ratio by increasing the size of the change, $\Delta_N = r3N^{-1/5}(1,1)^\top\zeta_{SNR}$, where $r = 0.5$, 1.5, and 2. Recall that in the simulations displayed above, the parameter $r = 1$. By setting $r = 0.5$, 1.5 and 2, we examine the performances of

FIGURE G.2. Empirical coverage rates of BP, EM, ILR, $\hat{k}_N^{HO}(1/2)$, $\hat{k}_N^{Hx}(1/2)$ and $\hat{k}_N^{HET}(1/2)$. The DGPs include models with homoscedastic **IID**, **AR**, **AR-HP** and **GARCH** errors. For each estimator, the left-side blue bars and right-side red bars present the results when the size of change $\Delta_{N,1} = N^{-1/5}(1,1)^\top \zeta_{SNR}$ and $\Delta_{N,2} = 3N^{-1/5}(1,1)^\top \zeta_{SNR}$, respectively. The values displayed on the top of the bar charts document the lengths of confidence intervals, in proportion to the sample size. The change point occurs at the middle of the sample $k_1 = \lfloor 0.5N \rfloor$, with the sample size $N = 300$.



$\hat{k}_N^{HET}(1/2)$ across scenarios ranging from low to high signal-to-noise ratios. Figure G.8 displays the improvement in performance as the signal-to-noise ratio increases when the model experiences a moderate change with a sample size of $N = 300$.

FIGURE G.3. Empirical coverage rates of BP, EM, ILR, $\hat{k}_N^{HO}(1/2)$, $\hat{k}_N^{Hx}(1/2)$ and $\hat{k}_N^{HET}(1/2)$. The DGPs include models with homoscedastic **IID**, **AR**, **AR-HP**, **GARCH**, and heteroscedastic **HeteIID**, **HeteAR**(**-HP**), **HeteGARCH** and **HeteSmooth** errors. For each estimator, the left-side blue bars and right-side red bars present the results when the size of change $\Delta_{N,1} = N^{-1/5}(1,1)^\top \zeta_{SNR}$ and $\Delta_{N,2} = 3N^{-1/5}(1,1)^\top \zeta_{SNR}$, respectively. The values displayed on the top of the bar charts document the average length of confidence intervals, in proportion to the sample size. The change point is set at the beginning of the sample, $k_1 = \lfloor 0.5N \rfloor$, with the error variance change either coinciding at $m_1 = \lfloor 0.5N \rfloor$ or differing at $m_1 = \lfloor 0.2N \rfloor$, under a sample size of $N = 300$.

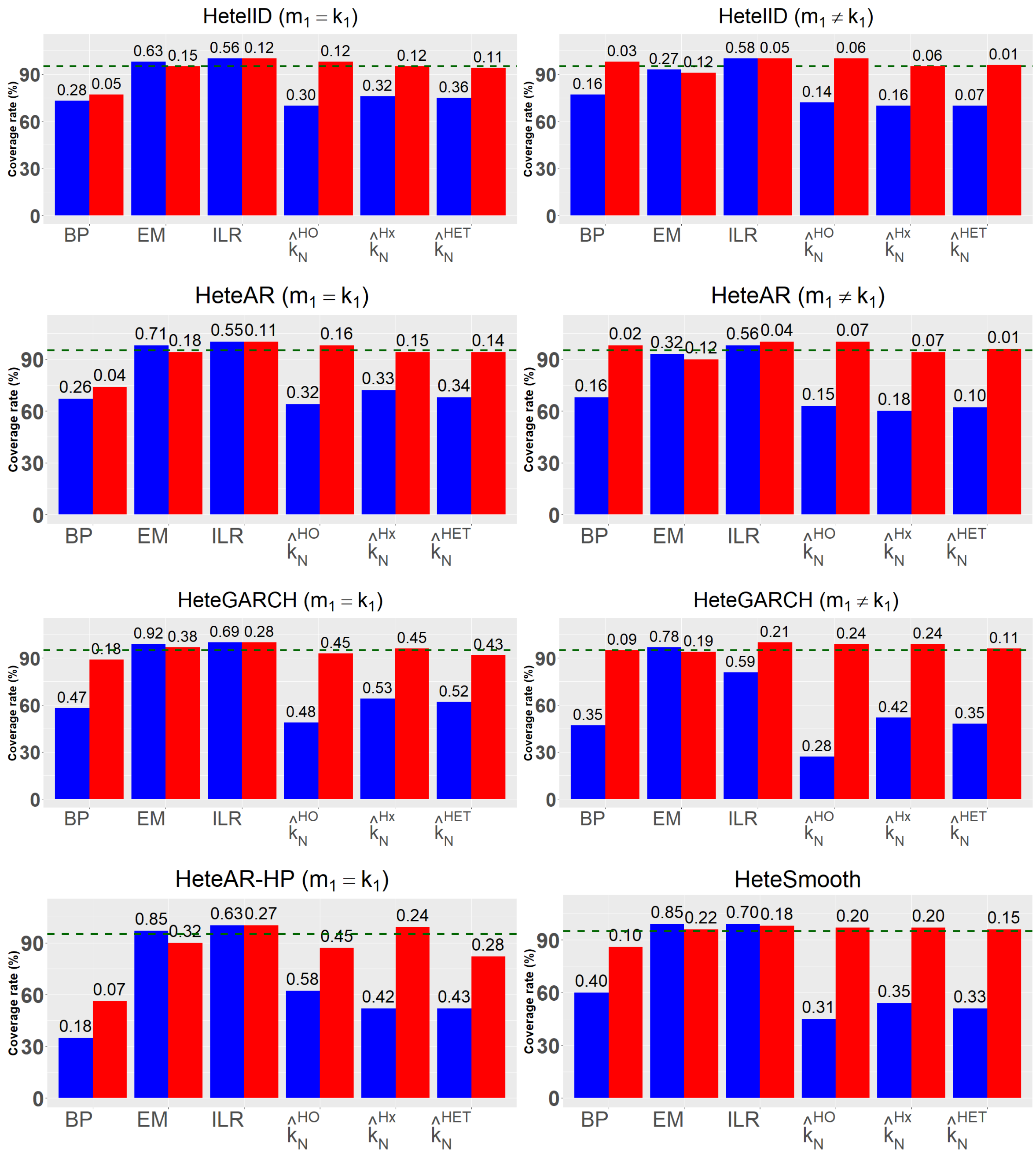


FIGURE G.4. Empirical coverage rates of BP, EM, ILR, $\hat{k}_N^{HO}(1/2)$, $\hat{k}_N^{Hx}(1/2)$ and $\hat{k}_N^{HET}(1/2)$. The DGPs include models with homoscedastic **IID**, **AR**, **AR-HP** and **GARCH** errors. For each estimator, the left-side blue bars and right-side red bars present the results when the size of change $\Delta_{N,1} = N^{-1/5}(1,1)^\top \zeta_{SNR}$ and $\Delta_{N,2} = 3N^{-1/5}(1,1)^\top \zeta_{SNR}$, respectively. The values displayed on the top of the bar charts document the lengths of confidence intervals, in proportion to the sample size. The change point occurs at the middle of the sample $k_1 = \lfloor 0.2N \rfloor$, with the sample size $N = 100$.

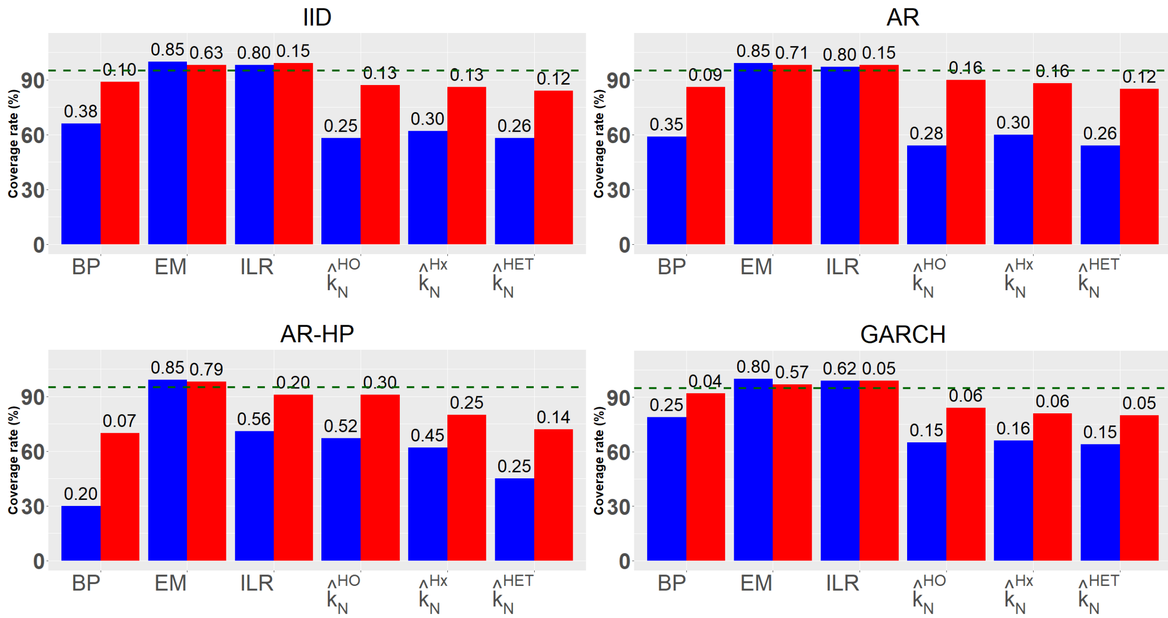


FIGURE G.5. Empirical coverage rates of BP, EM, ILR, $\hat{k}_N^{HO}(1/2)$, $\hat{k}_N^{Hx}(1/2)$ and $\hat{k}_N^{HET}(1/2)$. The DGPs include models with homoscedastic **IID**, **AR**, **AR-HP**, **GARCH**, and heteroscedastic **HeteIID**, **HeteAR**(**-HP**), **HeteGARCH** and and **HeteSmooth** errors. For each estimator, the left-side blue bars and right-side red bars present the results when the size of change $\Delta_{N,1} = N^{-1/5}(1,1)^\top \zeta_{SNR}$ and $\Delta_{N,2} = 3N^{-1/5}(1,1)^\top \zeta_{SNR}$, respectively. The values displayed on the top of the bar charts document the average length of confidence intervals, in proportion to the sample size. The change point is set at the beginning of the sample, $k_1 = \lfloor 0.5N \rfloor$, with the error variance change either coinciding at $m_1 = \lfloor 0.2N \rfloor$ or differing at $m_1 = \lfloor 0.5N \rfloor$, under a sample size of $N = 100$.

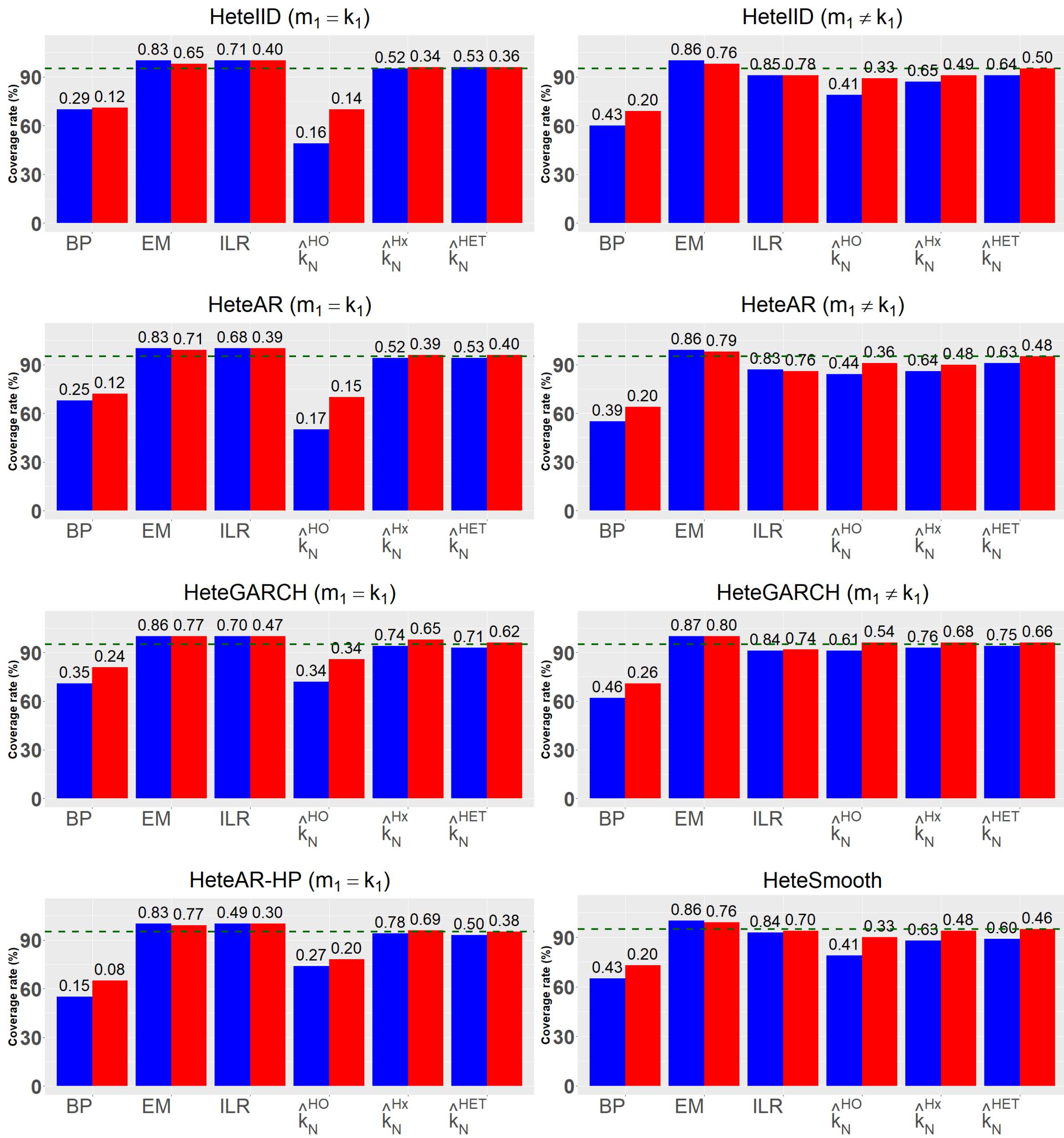


FIGURE G.6. Empirical coverage rates of $\hat{k}_N^{HET}(1/2)$. The DGPs include models with homoscedastic **IID**, **AR**, and heteroscedastic **HeteIID**, **HeteAR**, and **HeteSmooth** errors. The bandwidth parameter $\wp_{1,N}$ and $\wp_{2,N}$ are chosen in four different scenarios: Size 1 – $\wp_{1,N} = \wp_{2,N} = \lfloor N^{2/5} \rfloor$, Size 2 – $\wp_{1,N} = \wp_{2,N} = \lfloor N^{1/3} \rfloor$, Size 3 – $\wp_{1,N} = \wp_{2,N} = \lfloor \log(N) \rfloor$, and Size 4 – $\wp_{1,N} = \wp_{2,N} = \lfloor N^{1/4} \rfloor$. The size of change $\Delta_N = 3N^{-1/5}(1,1)^\top \zeta_{SNR}$, with the sample size $N = 300$. An early structural change occurs at $k_1 = \lfloor 0.2N \rfloor$, with $m_1 \neq k_1$, under heteroscedastic **HeteIID** and **HeteAR** error specifications.

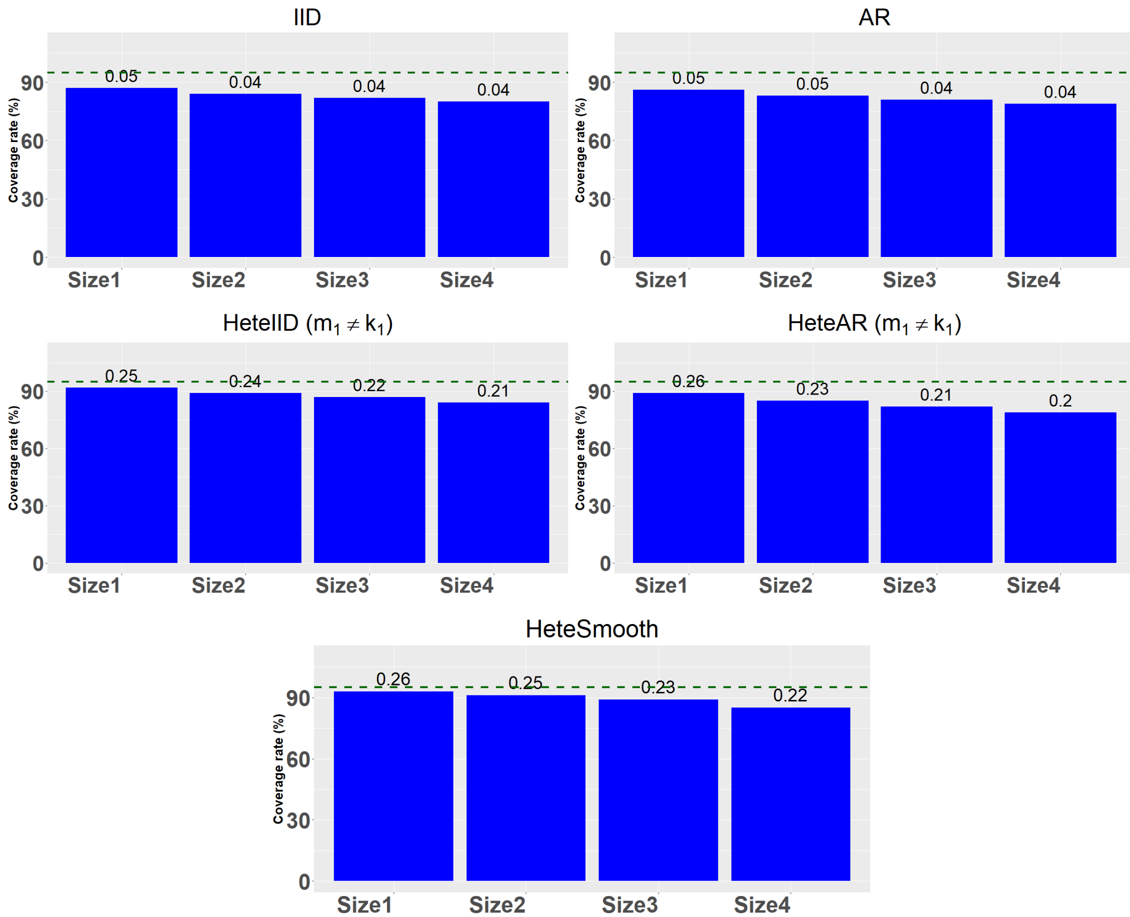


FIGURE G.7. Empirical coverage rates of $\hat{k}_N^{HET}(\kappa)$. The DGPs include models with homoscedastic **IID**, **AR**, and heteroscedastic **HeteIID** and **HeteAR** errors. For each DGP, the bars are arranged from left to right to display the results for the weight parameter $\kappa = 0$, 0.15, 0.3, and 0.45. The left sub–figures show the outcomes when an early change point occurs at $k^* = \lfloor 0.2N \rfloor$, while the right sub–figures illustrate the scenarios with a middle change at $k_1 = \lfloor 0.5N \rfloor$, with m_1 coinciding with k_1. The size of change $\Delta_N = 3N^{-1/5}(1,1)^\top \zeta_{SNR}$, with the sample size $N = 300$. The values displayed on the top of the bar charts document the lengths of confidence intervals, in proportion to the sample size.

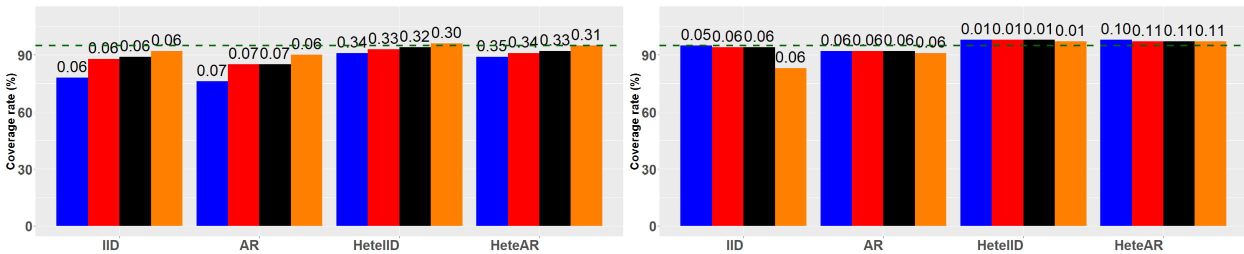
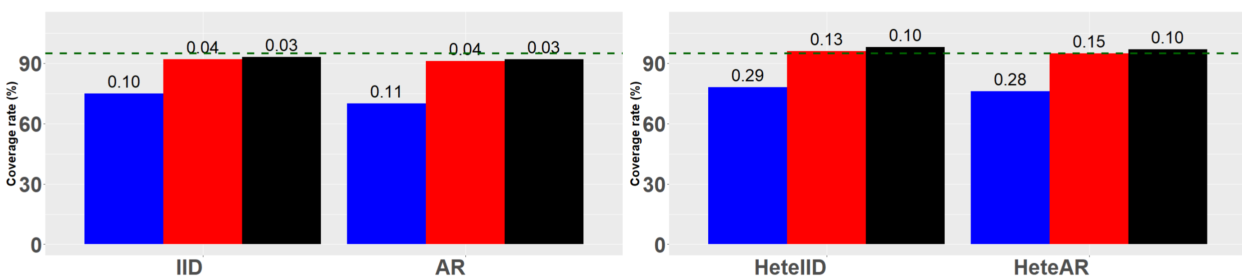


FIGURE G.8. Empirical coverage rates of $\hat{k}_N^{HET}(\kappa)$. The DGPs include models with homoscedastic errors **IID** and **AR** (left sub-plot), and heteroscedastic **HeteIID** and **HeteAR** errors (right sub-plot). For each estimator, the bars from the left to the right present the ratio $r = 0.5$, 1.5 and 2. The sample size $N = 300$ and the size of change $\Delta_N = r3N^{-1/5}(1,1)^\top \zeta_{SNR}$. The change point occurs at the middle of the sample $k_1 = \lfloor 0.2N \rfloor$ with m_1 identical with k_1. The values displayed on the top of the bar charts document the lengths of confidence intervals, in proportion to the sample size.



Appendix H. Data application: changes in common risk factors in Bitcoin

In the second data application, we apply the estimator and confidence interval $\hat{k}_N^{HET}(1/2)$ to estimate changes in the relationship between Bitcoin returns and common risk factors in the cryptocurrency market. Cryptocurrency, as a decentralized digital currency, has experienced rapid growth in the recent decade. Analyzing the cryptocurrency market provides important implications for empirical asset pricing, as it helps to understand the asset pricing common risk factors in traditional financial assets and these innovative virtual assets. The recent literature of Sockin and Xiong (2020) identifies a cryptocurrency momentum premium led by the investor overreaction effect. Borri and Shakhnov (2022) find that the Bitcoin prices can be dispersed on different trading exchanges and propose a risk–based explanation for the pricing differences. Similarly to the equity market, the cross–sections of cryptocurrency show price co–movements. In order to capture the systematic risk in the cryptocurrency market, Liu et al. (2022) consider all the digital currencies with market capitalizations above one million dollars and their weekly returns from January 2014 to July 2020, and they explore three common risk factors – the cryptocurrency market factor (Cmkt), cryptocurrency size factor (Csmb), and a cryptocurrency momentum factor (Cmom) to account for the cross–sectional cryptocurrency excess returns. In the present paper, we check if these three common risk factors consistently account for the Bitcoin excess returns.

The weekly common risk factors data is collected from Yukun Liu's personal website[4]. We also collect the weekly data on Bitcoin price and three–month Treasury bill rate from Investing.com and Yahoo Finance, respectively. The data ranges from 20 January 2014 (week 4) to 20 July 2020 (week 30). We then use the following three–factor model for Bitcoin excess returns:

$$R_i^B - R_i^f = \alpha + \beta_1 Cmkt_i + \beta_2 Csize_i + \beta_3 Cmom_i + \varepsilon_i, \quad 1 \le i \le N, \tag{H.1}$$

where R_i^B is the log return of Bitcoin price, and for risk free rate R_i^f we use the three–month Treasury bill rate.

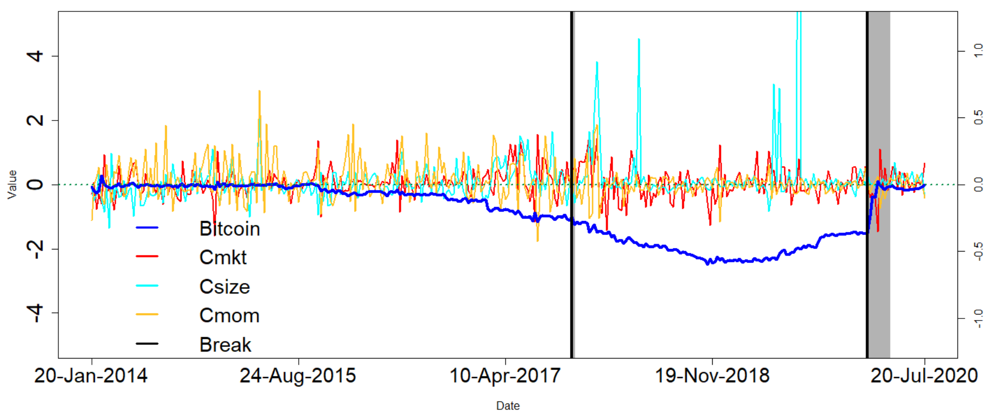
Figure H.1 illustrates the evolution of Bitcoin excess returns and the series of three cryptocurrency common risk factors during the sample period. We observe a glide of Bitcoin returns during 2017–2020, which may be attributed to the burst of the cryptocurrency bubble, and then the whole market rallies in 2020. By using the estimator $\hat{k}_N^{HET}(1/2)$, we identify two change points in week 43, 2017 and week 7, 2020.[5] In unreported results, the estimators BP and ILR only detect the first change in 2017. In Table H.1, we document the estimates in model (H.1) in three sub–samples that are split by the two estimated change points. The

[4]https://www.yukunliu.com/research/

[5]The confidence intervals are, respectively, from week 42, 2017 to week 44, 2017, and from week 7, 2020 to week 16, 2020.

last column $SD(\varepsilon)$ presents the estimated standard deviation of the residuals, indicating some heteroscedasticity in the model. The results show that three cryptocurrency risk factors significantly account for the Bitcoin excess return during the sample period from week 4, 2014 to week 43, 2017. However, these common risk factors become ineffective in explaining the Bitcoin excess return during the burst of cryptocurrency bubbles between week 44, 2017 and week 7, 2020. This finding is similar to the performances of Fama-French three factors in the equity market during the market crashes (Linnainmaa and Roberts, 2018). The second change point is detected at the end of the burst of the cryptocurrency bubble. In the sample from week 8, 2020 to week 30, 2020, the risk factors still cannot explain the Bitcoin excess returns though this may be improved when more observations are available.

FIGURE H.1. Plots of Bitcoin return and the series of three factors Cmkt, Csize, and Cmom, with the estimated change point and corresponding confidence intervals via $\hat{k}_N^{HET}(1/2)$ in the shaded areas. The left-sided y-axis is for Bitcoin return and the right-sided y-axis is for three cryptocurrency factors.



References

[1] Aue, A., Hörmann, S., Horváth, L. and Hušková, M.: Dependent functional linear models with applications to monitoring structural change. *Statistica Sinica* **24**(2014), 1043–1073.

[2] Berkes, I., Hörmann, S. and Schauer, J. (2011). Split invariance principles for stationary processes.

[3] Berkes, I., Liu, W. and Wu, W.B.: Komlós-Major-Tusnády approximation under dependence. *Annals of Probability* **42**(2014), 794–817.

[4] Borri, N., and Shakhnov, K.: The cross-section of cryptocurrency returns. *The Review of Asset Pricing Studies* (2022), raac007.

[5] Csörgő, M. and Horváth, L.: *Limit Theorems in Change–Point Analysis.* Wiley, New York, 1997.

[6] Horváth, L., Rice, G. and Zhao, Y.: Testing for changes in linear models using weighted residuals. *Journal of Multivariate Analysis* **198**(2023), 105210.

[7] Kechris, A. S.: *Classical Descriptive Set Theory.* Graduate Texts in Mathematics, Vol. 156. Springer–Verlag, New York (1995).

TABLE H.1. The estimates of Model (H.1) in the three sub–samples with the last column $SD(\varepsilon)$ representing the standard deviation of the model residuals. The superscript of coefficient ***, ** and * indicate statistical significance at levels 1%, 5% and 10%, respectively.

	20-January-2014 (week 4) : 23-October-2017 (week 43)				
Parameter	α	β_1	β_2	β_3	$SD(\varepsilon)$
Estimate	-0.28^{***}	-0.70^{***}	-1.09^{***}	0.32^{**}	0.31
SD	0.02	0.21	0.20	0.14	
	30-October-2017 (week 44) : 10-February-2020 (week 7)				
Estimate	-1.92^{***}	0.42	0.01	0.07	0.36
SD	0.03	0.27	0.08	0.36	
	17-February-2020 (week 8) : 20-July-2020 (week 30)				
Estimate	-0.16^{***}	0.48	-0.19	0.79	0.20
SD	0.05	0.41	0.91	0.80	

[8] Kurtz, T.G., and Protter, P.: Weak limit theorems for stochastic integrals and stochastic differential equations. *Annals of Probability* **19**(1991), 1035–1070.

[9] Linnainmaa, J. T. and Roberts, M.R.: The history of the cross-section of stock returns. *The Review of Financial Studies*, **31**(2018), 2606–2649.

[10] Liu, Y., Tsyvinski, A. and Wu, X.: Common risk factors in cryptocurrency. *The Journal of Finance*, **77**(2022), 1133–1177.

[11] Mörters, P. and Peres, Y.: *Brownian Motion.* Cambridge University Press, Cambridge, UK, 2010.

[12] Sockin, M. and Xiong, W.: A model of cryptocurrencies (No. w26816). *National Bureau of Economic Research.* (2020).