SUPPLEMENT

Supplement to "Modeling the COVID-19 infection trajectory: a piecewise linear quantile trend model"

Feiyu Jiang[1] | Zifeng Zhao[2] | Xiaofeng Shao[3]

The supplement contains all technical proofs and additional simulation and real data analysis results. Specifically, Section S1 presents theoretical results on single change-point estimation and includes several important lemmas to be used for deriving theoretical results on multiple change-point estimation. Section S2 gathers proofs for the main proposition and theorem. Section S3 contains auxiliary lemmas and proofs. Section S4 proposes the multi-scanning M-GOALS. Section S5 conducts numerical experiments to examine the finite sample performance of (M-)GOALS and multi-scanning M-GOALS. Section S6 presents additional results for real data analysis.

We first introduce some notations. Let $(\Omega, \mathcal{F}, P)$ be the underlying probability space. Let $\Rightarrow$ denote process convergence (or weak convergence) in some suitable function space, $\to_D$ denote convergence in distribution and $\to_p$ denote convergence in probability. For any compact set $\mathcal{T} \in \mathbb{R}^d$, $C(T)$ denotes the space of real-valued continuous functions on $\mathcal{T}$ endowed with the uniform topology; $l^\infty(\mathcal{T})$ is the space of real-valued bounded functions on $\mathcal{T}$ endowed with the uniform topology; $C(\mathbb{R}^d \times \mathcal{T})$ is the space of real-valued continuous functions on $\mathbb{R}^d \times \mathcal{T}$ endowed with the topology of locally uniform convergence. In addition, we assume C is a generic constant that may vary from line to line.

Without loss of generality, for GOALS, we can assume $Q_\tau(\varepsilon) = 0$ (otherwise, we can let $\widetilde{\beta}_t(\tau) = \theta_t + Q_\tau(\varepsilon_t)\gamma_t$ and $\widetilde{\varepsilon}_t = \varepsilon_t - Q_\tau(\varepsilon_t)$ so that $Y_t = X_t^\top\theta_t + [X_t^\top\gamma_t]\varepsilon_t = X_t^\top\widetilde{\beta}_t(\tau) + [X_t^\top\gamma_t]\widetilde{\varepsilon}_t$) and we drop "$(\tau)$" in $\beta_t = \beta_t(\tau)$ and $\widehat{\beta}_{\lfloor nr_1\rfloor,\lfloor nr_2\rfloor} = \widehat{\beta}_{\lfloor nr_1\rfloor,\lfloor nr_2\rfloor}(\tau)$ for $0 \le r_1 \le r_2 \le 1$.

S1 | THEORY FOR SINGLE CHANGE-POINT ESTIMATION

For simplicity, we consider a fixed quantile level $\tau \in \tau^M$. Under the model (5) with one change-point, i.e. $Y_t = X_t^\top\theta_t + (X_t^\top\gamma_t)\varepsilon_t$, and $Q_\tau(Y_t) = X_t^\top[\theta_t + Q_\tau(\varepsilon)\gamma_t] := X_t^\top\beta_t$, we have $\beta^{(i)} = \theta^{(i)} + \gamma^{(i)}Q_\tau(\varepsilon)$, $i = 1, 2$ such that

$$\beta_t = \begin{cases} \beta^{(1)}, & 1 \le t \le k_1 \\ \beta^{(2)}, & k_1 + 1 \le t \le n, \end{cases}$$

for some $k_1 = \lfloor nq_1\rfloor$, $q_1 \in (0,1)$.

It is natural to estimate the change-point location by

$$\widehat{k} = \arg\max_{k\in[h,n-h]} T_{n,\delta}(k),$$

where $T_{n,\delta}(k) = D_n(1,k,n)^\top V_{n,\delta}(1,k,n)^{-1} D_n(1,k,n)$. The next proposition presents the consistency of the change-point location estimator under the single change-point setting.

Proposition 1 *Suppose the model (5) admits a single change-point and Assumptions 1-3 hold. Let* $\mathbf{b} = \boldsymbol{\beta}^{(2)} - \boldsymbol{\beta}^{(1)}$ *and* $\kappa = \|\mathbf{b}\|$ *satisfies* $\log(n)^{-2} n\kappa^2 \to \infty$ *and* $\kappa \to 0$ *as* $n \to \infty$. *Let* $\widehat{q} = \widehat{k}/n$, *we have* $\widehat{q} \to_p q_1$.

The proof of Proposition 1 is deferred to Section S2. Below we present several key lemmas. In particular, Lemma 2 gives the uniform Bahadur representation of $\widehat{\boldsymbol{\beta}}_{\lfloor nr_1 \rfloor+1, \lfloor nr_2 \rfloor}$ when there is no change-point, see also Zhou and Shao (2013); Lemma 3 derives the asymptotic behavior of $T_{n,\delta}(k_1)$ under the single change-point setting; Lemma 4 obtains the uniform distributional limits for $\widehat{\boldsymbol{\beta}}_{\lfloor nr_1 \rfloor+1, \lfloor nr_2 \rfloor}$ after suitable centering and standardization under the single change-point setting; Lemma 5 generalizes Lemma 4 to the multiple change-point setting.

To simplify the notation, we define $\varphi_t = [X_t^\top \gamma_t]\varepsilon_t$, $g_t(\boldsymbol{\beta}) = \rho_\tau(Y_t - X_t^\top \boldsymbol{\beta}) - \rho_\tau(\varphi_t)$, where $\rho_\tau(u) = u\psi_\tau(u)$ is the quantile check function. Let $\omega(a, b; \eta) = \{(r_1, r_2) | a \le r_1 < r_2 \le b, r_2 - r_1 \ge \eta\}$ for some $\eta > 0$ and $0 \le a < b \le 1$. In addition, when change-points are present, recall $c_i \in \mathbb{R}^2/\{(0,0)^\top\}$ in Assumption 4 denotes the slope difference normalized by κ, we let $\boldsymbol{\alpha}_i = \sum_{\ell=1}^{i} c_\ell$ for $i = 1, \cdots, m_0$ and $\boldsymbol{\alpha}_0 = (0,0)^\top$.

Lemma 2 *Suppose Assumptions 1-3 hold. Then for all* $\eta \in (0,1)$, *if there is no change-point, then (i)*

$$\sup_{(r_1,r_2)\in\omega(0,1;\eta)} \left| \sqrt{n}(\widehat{\boldsymbol{\beta}}_{\lfloor nr_1 \rfloor+1, \lfloor nr_2 \rfloor} - \boldsymbol{\beta}) - \frac{1}{\sqrt{n}} [\Sigma_1(r_2) - \Sigma_1(r_1)]^{-1} f(0)^{-1} \sum_{t=\lfloor nr_1 \rfloor+1}^{\lfloor nr_2 \rfloor} X_t \psi_\tau(\varepsilon_t) \right| = o_p(1);$$

(ii) on a richer probability space, there exist i.i.d. standard normal random variables $V_1, V_2, \cdots, V_n$ *such that*

$$\max_{1\le i\le j\le n} \left| \sum_{t=i}^{j} X_t f(0)^{-1} \psi_\tau(\varepsilon_t) - \Gamma_\varepsilon \sum_{t=i}^{j} X_t V_t \right| = o_p(1).$$

The proof of Lemma 2 follows by using the same arguments in the proofs of Lemma 4 and Proposition 3 in Zhou and Shao (2013), hence omitted.

Lemma 3 *Under the conditions of Proposition 1,*

$$(n\kappa^2)^{-1} T_{n,\delta}(k_1) \to_D q_1^2 (1-q_1)^2 (\kappa^{-1}\mathbf{b})^\top V_\delta(q_1)^{-1} (\kappa^{-1}\mathbf{b}), \tag{S1}$$

where $V_\delta(s) = L_\delta(s) + R_\delta(s)$ *and*

$$L_\delta(s) = \int_\delta^{s-\delta} \frac{r^2(s-r)^2}{s^2} \left\{ \Sigma_1(r)^{-1} B_X(r) - [\Sigma_1(s) - \Sigma_1(r)]^{-1} [B_X(s) - B_X(r)] \right\}^{\otimes 2} dr,$$
$$R_\delta(s) = \int_{s+\delta}^{1-\delta} \frac{(r-s)^2(1-r)^2}{(1-s)^2} \left\{ [\Sigma_2(1) - \Sigma_2(r)]^{-1} [B_X(1) - B_X(r)] - [\Sigma_2(r) - \Sigma_2(s)]^{-1} [B_X(r) - B_X(s)] \right\}^{\otimes 2} dr.$$

The proof of Lemma 3 is deferred to Section S3.

Lemma 4 *Define* $\mathbf{r} = (r_1, r_2) \in [0, q_1 - \eta] \times [q_1 + \eta, 1] := \mathcal{T} \subset [0,1]^2$ *for some* $\eta \in (0, \delta)$. *For observations* $\{Y_t\}_{t=1}^n$ *under the single change-point alternative, let*

$$\widehat{\mathbf{c}}_n(\mathbf{r}) = \arg\min_{\mathbf{c}\in\mathbb{R}^2} Z_n(\mathbf{c}, \mathbf{r}), \tag{S2}$$

where

$$Z_n(\mathbf{c},\mathbf{r}) = \frac{1}{n\kappa^2}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} g_t(\boldsymbol{\beta}) + \frac{(nr_2-\lfloor nr_2\rfloor)}{n\kappa^2} g_{\lfloor nr_2\rfloor+1}(\boldsymbol{\beta}) + \frac{(\lfloor nr_1\rfloor+1-nr_1)}{n\kappa^2} g_{\lfloor nr_1\rfloor}(\boldsymbol{\beta}) \tag{S3}$$

is the objective function based on partial observations $\{Y_t\}_{t=\lfloor nr_1\rfloor}^{\lfloor nr_2\rfloor+1}$ *satisfying* $\mathbf{c} = \kappa^{-1}(\boldsymbol{\beta}-\boldsymbol{\beta}^{(1)})$*. Furthermore, for each* $\mathbf{r}\in\mathcal{T}$*, let* $\mathbf{c}(\mathbf{r}) = \arg\min_{\mathbf{c}\in\mathbb{R}^2} Z_\infty(\mathbf{c},\mathbf{r})$*, such that*

$$Z_\infty(\mathbf{c},\mathbf{r}) = \frac{f(0)}{2}\Big\{\int_{r_1}^{q_1}\frac{[(1,x)\mathbf{c}]^2}{(1,x)\gamma^{(1)}}dx + \int_{q_1}^{r_2}\frac{[(1,x)(\mathbf{c}-\boldsymbol{\alpha})]^2}{(1,x)\gamma^{(2)}}dx\Big\}, \tag{S4}$$

where $\boldsymbol{\alpha} = \kappa^{-1}\mathbf{b}\in\mathbb{R}^2$*. Then, under the conditions of Proposition 1,*

$$\widehat{\mathbf{c}}_n(\cdot) \Rightarrow \mathbf{c}(\cdot),$$

where $\mathbf{c}(\mathbf{r}) = [\Sigma_1(q_1)-\Sigma_1(r_1)+\Sigma_2(r_2)-\Sigma_2(q_1)]^{-1}[\Sigma_2(r_2)-\Sigma_2(q_1)]\boldsymbol{\alpha}$ *is unique.*

The proof of Lemma 4 is deferred to Section S3.

Remark When $0<\underline{c}<\kappa<\infty$, i.e. the change size in the coefficients is not diminishing, we have

$$\begin{aligned} Z_\infty(\mathbf{c},\mathbf{r}) =& \kappa^{-1}\Big\{\int_{r_1}^{q_1}\int_0^{(1,x)\mathbf{c}}[F([X(s)^\top\gamma^{(1)}]^{-1}\kappa s)-F(0)]dsdx \\ &+\kappa^{-1}\int_{q_1}^{r_2}\int_0^{(1,x)(\mathbf{c}-\boldsymbol{\alpha})}[F([X(s)^\top\gamma^{(2)}]^{-1}\kappa s)-F(0)]dsdx\Big\}. \end{aligned}$$

In this case, $\mathbf{c}(\mathbf{r}) = \arg\min_{\mathbf{c}\in\mathbb{R}^2} Z_\infty(\mathbf{c},\mathbf{r})$ depends on the exact form of the distribution function $F(\cdot)$. Similarly to the diminishing case, we have $\{\kappa^{-1}(\widehat{\boldsymbol{\beta}}_{\lfloor nr_1\rfloor,\lfloor nr_2\rfloor}-\boldsymbol{\beta}^{(1)})\}_{\mathbf{r}\in\mathcal{T}} \Rightarrow \{\mathbf{c}(\mathbf{r})\}_{\mathbf{r}\in\mathcal{T}}$, and in particular, $\widehat{\boldsymbol{\beta}}_{1,n}\to_p \kappa\mathbf{c}(0,1)+\boldsymbol{\beta}^{(1)}$.

Lemma 5 *Let* $\boldsymbol{q} = (q_1,\cdots,q_{m_0})$ *be the change-point locations. If* $m_0\ge 1$*, and for pairs of* $\mathbf{r}=(r_1,r_2)\in[q_{i-1},q_i-\eta]\times[q_j+\eta,q_{j+1}] := \mathcal{T}\subset[0,1]^2$ *for some* $1\le i\le j\le m_0$ *and* $\eta\in(0,\delta)$*, (i.e.* (r_1,r_2) *contains* $j-i+1$ *change-points).*

Let

$\widehat{\mathbf{c}}_n(\mathbf{r}) = \arg\min_{\mathbf{c}\in\mathbb{R}^2} Z_n(\mathbf{c},\mathbf{r})$*, where* $Z_n(\mathbf{c},\mathbf{r})$ *is defined in (S3) satisfying* $\mathbf{c}=\kappa^{-1}(\boldsymbol{\beta}-\boldsymbol{\beta}^{(1)})$*. Furthermore, for each* $\mathbf{r}\in\mathcal{T}$*, let* $\mathbf{c}(\mathbf{r}) = \arg\min_{\mathbf{c}\in\mathbb{R}^2} Z_\infty(\mathbf{c},\mathbf{r})$ *such that*

$$\begin{aligned} Z_\infty(\mathbf{c},\mathbf{r}) =& \frac{f(0)}{2}\Big[\int_{r_1}^{q_i}\frac{[(1,x)(\mathbf{c}-\boldsymbol{\alpha}_{i-1})]^2}{(1,x)\gamma^{(i)}}dx + \sum_{\ell=1}^{j-i}\int_{q_{i+\ell-1}}^{q_{i+\ell}}\frac{[(1,x)(\mathbf{c}-\boldsymbol{\alpha}_{i+\ell-1})]^2}{(1,x)\gamma^{(i+\ell)}}dx \\ &+\int_{q_j}^{r_2}\frac{[(1,x)(\mathbf{c}-\boldsymbol{\alpha}_j)]^2}{(1,x)\gamma^{(j+1)}}dx\Big]. \end{aligned}$$

Then under the conditions of Theorem 1 in the case of change-points, $\widehat{\mathbf{c}}_n(\cdot)\Rightarrow\mathbf{c}(\cdot)$*, where*

$$\begin{aligned} \mathbf{c}(\mathbf{r}) =& \Big\{\Sigma_i(q_i)-\Sigma_i(r_1)+\sum_{\ell=1}^{j-i}[\Sigma_{i+\ell}(q_{i+\ell})-\Sigma_{i+\ell}(q_{i+\ell-1})]+\Sigma_{j+1}(r_2)-\Sigma_{j+1}(q_j)\Big\}^{-1} \\ &\times\Big\{[\Sigma_i(q_i)-\Sigma_i(r_1)]\boldsymbol{\alpha}_{i-1}+\sum_{\ell=1}^{j-i}[\Sigma_{i+\ell}(q_{i+\ell})-\Sigma_{i+\ell}(q_{i+\ell-1})]\boldsymbol{\alpha}_{i+\ell-1}+[\Sigma_{j+1}(r_2)-\Sigma_{j+1}(q_j)]\boldsymbol{\alpha}_j\Big\} \end{aligned}$$

is unique with $\Sigma_i(r) = \int_0^r [X(s)^\top \gamma^{(i)}]^{-1} X(s) X(s)^\top ds$.

The proof of Lemma 5 is similar to that of Lemma 4, hence omitted.

S2 | PROOF OF MAIN THEOREMS

S2.1 | Proof of Proposition 1

By Lemma 3, we have that with probability tending to one, $(n\kappa^2)^{-1} T_{n,\delta}(k_1) > 0$. That is, at the true change-point, the statistic $T_{n,\delta}(k_1)$ diverges at the rate $n\kappa^2$. Letting $M_{n,\eta} = \{k : |\frac{k}{n} - q_1| > \eta\}$ for any $\eta \in (0, \delta)$, then it suffices to show that

$$(n\kappa^2)^{-1} \max_{k \in [h, n-h] \cap M_{n,\eta}} D_n(1,k,n)^\top V_{n,\delta}(1,k,n)^{-1} D_n(1,k,n) = o_p(1). \tag{S5}$$

That is, we show that for points far away from k_1, the statistic is of smaller order. By symmetry, it suffices to consider the points $k \in M_{n,\eta}^{(1)} := \{k : \frac{k}{n} < q_1 - \eta\}$.

Note that $\sup_{q \in (0, q_1 - \eta)} |\sqrt{n}(\widehat{\beta}_{1,\lfloor nq \rfloor} - \beta^{(1)})| = O_p(1)$ by Lemma 2, hence using $n\kappa^2 \to \infty$ as $n \to \infty$, we have $\kappa^{-1} \sup_{q \in (0, q_1 - \eta)} |\widehat{\beta}_{1,\lfloor nq \rfloor} - \beta^{(1)}| \to_p 0$. By Lemma 4, we can show that $\{\kappa^{-1}(\widehat{\beta}_{\lfloor nq \rfloor + 1, n} - \beta^{(1)})\}_{q \in (0, q_1 - \eta)} \Rightarrow \{\mathbf{c}(q,1)\}_{q \in (0, q_1 - \eta)}$, hence

$$\left\{(n\kappa^2)^{-1/2} D_n(1, \lfloor nq \rfloor, n)\right\}_{q \in (0, q_1 - \eta)} \Rightarrow \left\{-q(1-q)\mathbf{c}(q,1)\right\}_{q \in (0, q_1 - \eta)}, \tag{S6}$$

where $\mathbf{c}(q,1)$ is defined in Lemma 4. Next, for each $k < k_1$, we decompose $R_{n,\delta}(1,k,n)$ by

$$\begin{aligned} R_{n,\delta}(1,k,n) =& \Big[\sum_{i=k+\lfloor n\delta \rfloor}^{k_1 + \lfloor n\delta \rfloor - 1} + \sum_{i=k_1+\lfloor n\delta \rfloor}^{n - \lfloor n\delta \rfloor}\Big] \frac{(i-1-k)^2 (n-i+1)^2}{n^2 (n-k)^2} (\widehat{\beta}_{i,n} - \widehat{\beta}_{k+1,i-1})^{\otimes 2} \\ :=& R_{n,\delta,1}(1,k,n) + R_{n,\delta,2}(1,k,n). \end{aligned}$$

where we have

$$R_{n,\delta,2}(1,k,n) = \sum_{i=k_1+\lfloor n\delta \rfloor}^{n-\lfloor n\delta \rfloor} \frac{(i-1-k)^2 (n-i+1)^2}{n^2(n-k)^2} (\widehat{\beta}_{i,n} - \widehat{\beta}_{k+1,i-1})^{\otimes 2}.$$

By Lemma 4, we obtain that

$$\begin{aligned} &\left\{\kappa^{-1}\left(\widehat{\beta}_{\lfloor nr \rfloor, n} - \beta^{(2)}\right)\right\}_{r \in (q_1, 1)} \Rightarrow 0 \\ &\left\{\kappa^{-1}\left(\widehat{\beta}_{\lfloor nq \rfloor + 1, \lfloor nr \rfloor - 1} - \beta^{(1)}\right)\right\}_{(q,r) \in (0, q_1 - \eta) \times (q_1, 1)} \Rightarrow \{\mathbf{c}(q,r)\}_{(q,r) \in (0, q_1 - \eta) \times (q_1, 1)}. \end{aligned}$$

Hence, using the continuous mapping theorem, we obtain that

$$\left\{(n\kappa^2)^{-1} R_{n,\delta,2}(1, \lfloor nq \rfloor, n)\right\}_{q \in (0, q_1 - \eta)} \Rightarrow \left\{\overline{R}_{\delta,2}(q)\right\}_{q \in (0, q_1 - \eta)}, \tag{S7}$$

where $\overline{R}_{\delta,2}(q) = \int_{q_1+\delta}^{1-\delta} \frac{(r-q)^2(1-r)^2}{(1-q)^2}[\boldsymbol{\alpha} - \mathbf{c}(q,r)]^{\otimes 2} dr$.

Now, provided that $\overline{R}_{\delta,2}(q)$ is invertible when $k \in M_{n,\eta}^{(1)}$ (which is shown in Lemma 10), we have

$$
\begin{aligned}
&D_n(1,k,n)^\top V_{n,\delta}^{-1}(1,k,n) D_n(1,k,n)\\
\le &D_n(1,k,n)^\top R_{n,\delta,2}^{-1}(1,k,n) D_n(1,k,n)\\
= &[(n\kappa^2)^{-1/2} D_n(1,k,n)]^\top [(n\kappa^2)^{-1} R_{n,\delta,2}(1,k,n)]^{-1} [(n\kappa^2)^{-1/2} D_n(1,k,n)],
\end{aligned}
$$

where in the first inequality we use the fact that $V_{n,\delta}(1,k,n)^{-1} \le R_{n,\delta}(1,k,n)^{-1} \le R_{n,\delta,2}(1,k,n)^{-1}$ with the convention that $A \le B$ implies $B - A$ is semi-positive definite when A and B are square matrices.

Then, by (S6) and (S7), we see that

$$
\begin{aligned}
&\max_{k \in M_{n,\eta}^{(1)}} [(n\kappa^2)^{-1/2} D_n(1,k,n)]^\top [(n\kappa^2)^{-1} R_{n,\delta,2}(1,k,n)]^{-1} [(n\kappa^2)^{-1/2} D_n(1,k,n)]\\
\to_D &\sup_{q \in (0, q_1 - \eta)} [q^2(1-q)^2 \mathbf{c}(q,1)^\top \overline{R}_{\delta,2}^{-1}(q) \mathbf{c}(q,1)].
\end{aligned}
$$

Therefore, (S5) follows as $n\kappa^2 \to \infty$. ∎

S2.2 | Proof of Theorem 1

Case 1: First, we derive the asymptotic distributions for GOALS and M-GOALS when there is no change-point, and both methods will not deliver any estimates with probability tending to one.

(i) For GOALS, when there is no change-point, using the similar arguments as in Lemma 2 and Lemma 6 in Zhou and Shao (2013), we have

$$
\begin{aligned}
&\left\{\frac{(q-u_1)(u_2-q)}{(u_2-u_1)^{3/2}} \sqrt{n}(\widehat{\boldsymbol{\beta}}_{\lfloor nu_1 \rfloor, \lfloor nq \rfloor} - \widehat{\boldsymbol{\beta}}_{\lfloor nq \rfloor + 1, \lfloor nu_1 \rfloor})\right\}_{q \in (\epsilon, 1-\epsilon), (u_1,u_2) \in G_\epsilon(q)} \Rightarrow \left\{\Gamma_\varepsilon D(u_1,q,u_2)\right\}_{q \in (\epsilon, 1-\epsilon), (u_1,u_2) \in G_\epsilon(q)},\\
&\left\{L_{n,\delta}(\lfloor nu_1 \rfloor, \lfloor nq \rfloor, \lfloor nu_2 \rfloor)\right\}_{q \in (\epsilon, 1-\epsilon), (u_1,u_2) \in G_\epsilon(q)} \Rightarrow \Gamma_\varepsilon^2 \left\{L_\delta(u_1,q,u_2)\right\}_{q \in (\epsilon, 1-\epsilon), (u_1,u_2) \in G_\epsilon(q)},\\
\text{and } &\left\{R_{n,\delta}(\lfloor nu_1 \rfloor, \lfloor nq \rfloor, \lfloor nu_2 \rfloor)\right\}_{q \in (\epsilon, 1-\epsilon), (u_1,u_2) \in G_\epsilon(q)} \Rightarrow \Gamma_\varepsilon^2 \left\{R_\delta(u_1,q,u_2)\right\}_{q \in (\epsilon, 1-\epsilon), (u_1,u_2) \in G_\epsilon(q)},
\end{aligned}
$$

where $D(u_1,q,u_2)$, $L_\delta(u_1,q,u_2)$, $R_\delta(u_1,q,u_2)$ are defined in Theorem 1.

Therefore

$$
\left\{T_{n,\epsilon,\delta}(\lfloor nq \rfloor)\right\}_{q \in (\epsilon, 1-\epsilon)} \Rightarrow \left\{\max_{(u_1,u_2) \in G_\epsilon(q)} D(u_1,q,u_2)^\top V_\delta(u_1,q,u_2)^{-1} D(u_1,q,u_2)\right\}_{q \in (\epsilon, 1-\epsilon)}.
$$

Hence by continuous mapping theorem, it follows

$$
\max_{k=h,\cdots,n-h} T_{n,\epsilon,\delta}(k) \to_D \sup_{q \in (\epsilon, 1-\epsilon)} \max_{(u_1,u_2) \in G_\epsilon(q)} D(u_1,q,u_2)^\top V_\delta(u_1,q,u_2)^{-1} D(u_1,q,u_2).
$$

(ii) For M-GOALS, using results in (i) for GOALS, we can obtain that for any $\tau \in [\tau_L, \tau_U]$, for all $\eta \in (0,1)$,

$$\sup_{(r_1,r_2)\in\omega(0,1;\eta)} \left|\sqrt{n}\left\{\begin{pmatrix}\widehat{\beta}_{0;\lfloor r_1 n\rfloor+1,\lfloor r_2 n\rfloor}(\tau)\\ \widehat{\beta}_{1;\lfloor r_1 n\rfloor+1,\lfloor r_2 n\rfloor}(\tau)\end{pmatrix} - \begin{pmatrix}\beta_0\\ \beta_1\end{pmatrix}\right\} - \frac{1}{\sqrt{n}}[\Sigma_1(r_2)-\Sigma_1(r_1)]^{-1} f(Q_\tau(\varepsilon))^{-1}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} X_t\psi_\tau(\varepsilon_t)\right| = o_p(1).$$

Recall $e_2 = (0,1)^\top$, then

$$\sup_{(r_1,r_2)\in\omega(0,1;\eta)} \left|\sqrt{n}\left\{\widehat{\beta}_{1;\lfloor nr_1\rfloor+1,\lfloor nr_2\rfloor}(\tau) - \beta_1\right\} - \frac{1}{\sqrt{n}f(Q_\tau(\varepsilon))} e_2^\top[\Sigma_1(r_2)-\Sigma_1(r_1)]^{-1}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} X_t\psi_\tau(\varepsilon_t)\right| = o_p(1).$$

Then, we stack the above equation for each $\tau \in \boldsymbol{\tau}^M$, since M is assumed to be finite, we can obtain

$$\sup_{(r_1,r_2)\in\omega(0,1;\eta)} \left|\sqrt{n}\left\{\widehat{\beta}^M_{\lfloor nr_1\rfloor+1,\lfloor nr_2\rfloor} - \beta_1^M\right\} - \frac{1}{\sqrt{n}}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} [e_2^\top[\Sigma_1(r_2)-\Sigma_1(r_1)]^{-1}X_t]\upsilon_t^M\right| = o_p(1),$$

with $\upsilon_t^M = (\frac{\psi_{\tau_1}(\varepsilon_t)}{f(Q_{\tau_1}(\epsilon))}, \cdots, \frac{\psi_{\tau_M}(\varepsilon_t)}{f(Q_{\tau_M}(\epsilon))})^\top$. Similar to Lemma 4 and Lemma 6 in Zhou and Shao (2013), we can show that for all $\eta \in (0,1)$

$$\begin{aligned}&\left\{\frac{1}{\sqrt{n}}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor}[e_2^\top[\Sigma_1(r_2)-\Sigma_1(r_1)]^{-1}X_t]\upsilon_t^M\right\}_{0\le r_1\le r_2-\eta\le 1-\eta}\\ \Rightarrow&\Gamma^M\left\{\int_{r_1}^{r_2}[e_2^\top[\Sigma_1(r_2)-\Sigma_1(r_1)]^{-1}X(s)]dB^M(s)\right\}_{0\le r_1\le r_2-\eta\le 1-\eta}\end{aligned}$$

where $\Gamma^M = \lim_{n\to\infty}\mathrm{Var}\Big(\frac{1}{\sqrt{n}}\sum_{t=1}^n \upsilon_t^M\Big)$.

Hence,

$$\left\{D_n^M(1,\lfloor nq\rfloor,n)\right\}_{q\in(\epsilon,1-\epsilon)} \Rightarrow \Gamma^M\left\{D^M(q)\right\}_{q\in(\epsilon,1-\epsilon)},\quad \left\{V^M_{n,\delta}(1,\lfloor nq\rfloor,n)\right\}_{q\in(\epsilon,1-\epsilon)} \Rightarrow \left\{\Gamma^M V^M_\delta(q)\Gamma^M\right\}_{q\in(\epsilon,1-\epsilon)},$$

the result follows by continuous mapping theorem.

Since $\zeta_n \to \infty$ as $n \to \infty$, combined with results in (i) and (ii),

$$P(\widehat{m}=0) = P(\max_{k=h,\cdots,n-h} T_{n,\epsilon,\delta}(k) < \zeta_n) \to 1. \tag{S8}$$

Case 2: Now we turn to the case with change-points. For simplicity, we only show the consistency for GOALS. The same argument applies to M-GOALS, as all SN-based test statistics involved in M-GOALS have the same stochastic orders as their counterparts in GOALS. We first show the following two intermediate results:

(1) at change-point locations,

$$P(\text{for all } k_i, T_{n,\epsilon,\delta}(k_i) > \zeta_n) \to 1; \tag{S9}$$

(2) at non-change-point locations, we show that for all $\eta > 0$,

$$P(\max_{k \in M_{n,\eta}} \max_{(t_1,t_2) \in G_n(k)} T_{n,\epsilon,\delta}(k) > \zeta_n) \to 0, \tag{S10}$$

where $M_{n,\eta} = \{k : |k - k_i| > n\eta, \forall k_i\} \cap \{k : h \le k \le n - h\}$.

(1) At each change-point k_i, we have

$$(n\kappa^2)^{-1} T_{n,\epsilon,\delta}(k_i) \ge (n\kappa^2)^{-1} D_n(k_i - h + 1, k_i, k_i + h)^\top V_{n,\delta}^{-1}(k_i - h + 1, k_i, k_i + h) D_n(k_i - h + 1, k_i, k_i + h). \tag{S11}$$

Note that we assume that $\min_{1 \le i \le m_0+1}(k_i - k_{i-1}) > h$, so there are no change-points between $[k_i - h, k_i + h]$ in the sample. Hence, by Lemma 2,

$$V_{n,\delta}(k_i - h + 1, k_i, k_i + h) \to_D \Gamma_\epsilon^2 V_\delta(q_i - \epsilon, q_i, q_i + \epsilon),$$

and

$$D_n(k_i - h + 1, k_i, k_i + h) = \frac{h^2}{(2h)^{3/2}}(\widehat{\boldsymbol{\beta}}_{k_i-h+1,k_i} - \boldsymbol{\beta}^{(i)} - \widehat{\boldsymbol{\beta}}_{k_i,k_i+h} + \boldsymbol{\beta}^{(i+1)}) + \frac{h^2}{(2h)^{3/2}}(\boldsymbol{\beta}^{(i)} - \boldsymbol{\beta}^{(i+1)}).$$

Therefore, by Lemma 4, we obtain that

$$(\sqrt{n}\kappa)^{-1} D_n(k_i - h + 1, k_i, k_i + h) \to_p -\frac{\epsilon^{1/2}}{\sqrt{8}} c_i.$$

This implies that

$$(n\kappa^2)^{-1} T_{n,\delta}(k_i - h + 1, k_i, k_i + h) \to_D \frac{\epsilon \Gamma_\epsilon^{-2}}{8} c_i^\top V_\delta^{-1}(q_i - \epsilon, q_i, q_i + \epsilon) c_i = O_p(1), \tag{S12}$$

i.e. $T_{n,\delta}(k_i - h + 1, k_i, k_i + h) = O_p(n\kappa^2)$.

Since the number of change-points m_0 is finite, we have

$$P\Big(\bigcap_{i=1}^{m_0} \{T_{n,\epsilon,\delta}(k_i) > \zeta_n\}\Big) \to 1,$$

i.e. (S9) holds.

(2) At non-change-point locations, we recall the definition of (6) and decompose $G_\epsilon(u)$ as $G_\epsilon(u) = G^{(1)}(u) \cup G^{(2)}(u)$, where

$$G^{(1)}(u) = \{(u_1, u_2) | \exists q_i,\ q_i \in (u_1, u_2)\} \cap G_\epsilon(u), \quad G^{(2)}(u) = \{(u_1, u_2) | \forall q_i,\ q_i \le u_1 \text{ or } q_i \ge u_2)\} \cap G_\epsilon(u).$$

The above sets define the neighborhood grid that contain at least one change-point and contain no change-points, respectively. Similarly, we define their rescaled counterparts,

$$G_n^{(1)}(k) = \{(t_1, t_2) | \exists k_i,\ k_i \in (t_1, t_2)\} \cap G_n(k), \quad G_n^{(2)}(k) = \{(t_1, t_2) | \forall k_i,\ k_i \le t_1 \text{ or } k_i \ge t_2)\} \cap G_n(k).$$

Then, for $i = 1, 2$ we denote

$$T^{(i)}_{n,\epsilon,\delta}(k) = \max_{(t_1,t_2)\in G^{(i)}_n(k)} D_n(t_1,k,t_2)^\top V_{n,\delta}(t_1,k,t_2)^{-1} D_n(t_1,k,t_2).$$

We only need to consider $T^{(1)}_{n,\epsilon,\delta}(k)$ since by similar arguments used in the case (i) with no change-point, we can show that

$$P(\max_{k\in M_{n,\eta}} T^{(2)}_{n,\epsilon,\delta}(k) < \zeta_n) \to 1. \tag{S13}$$

By Lemma 5 and we note that u_1 and u_2 are linear functions of q if $(u_1, u_2) \in G_\epsilon(q)$, hence

$$\begin{aligned}
&\left\{\kappa^{-1}(\widehat{\boldsymbol\beta}_{\lfloor nu_1\rfloor,\lfloor nq\rfloor} - \boldsymbol\beta^{(1)})\right\}_{q\in(\epsilon,1-\epsilon),(u_1,u_2)\in G^{(1)}(q)} \Rightarrow \left\{\mathbf{c}(u_1,q)\right\}_{q\in(\epsilon,1-\epsilon),(u_1,u_2)\in G^{(1)}(q)},\\
&\left\{\kappa^{-1}(\widehat{\boldsymbol\beta}_{\lfloor nq\rfloor+1,\lfloor nu_2\rfloor} - \boldsymbol\beta^{(1)})\right\}_{q\in(\epsilon,1-\epsilon),(u_1,u_2)\in G^{(1)}(q)} \Rightarrow \left\{\mathbf{c}(q,u_2)\right\}_{q\in(\epsilon,1-\epsilon),(u_1,u_2)\in G^{(1)}(q)},
\end{aligned}$$

which implies that

$$\begin{aligned}
&\left\{n^{-1/2}\kappa^{-1}D_n(\lfloor nu_1\rfloor,\lfloor nq\rfloor,\lfloor nu_2\rfloor)\right\}_{q\in(\epsilon,1-\epsilon),(u_1,u_2)\in G^{(1)}(q)}\\
\Rightarrow&\left\{\frac{(q-u_1)(u_2-q)}{(u_2-u_1)^{3/2}}(\mathbf{c}(u_1,q)-\mathbf{c}(q,u_2))\right\}_{q\in(\epsilon,1-\epsilon),(u_1,u_2)\in G^{(1)}(q)},
\end{aligned} \tag{S14}$$

where $\mathbf{c}(u_1, q)$ and $\mathbf{c}(q, u_2)$ are defined in Lemma 5.

Next, we analyze the behavior of $V^{-1}_{n,\delta}(t_1, k, t_2)$. Note that when $k \in M_{n,\eta}$, and $(t_1, t_2) \in G^{(1)}_n(k)$, we must have (t_1, k) or (k, t_2) contains at least one change-point (notice that k is not a change-point). Without loss of generality, we assume (t_1, k) contains at least one change-point and let $M_\eta = \{q : |q - q_i| > \eta, \forall q_i\} \cap (\epsilon, 1 - \epsilon)$.

Recall that

$$L_{n,\delta}(t_1,k,t_2) = \sum_{i=t_1+\lfloor n\delta\rfloor+1}^{k-\lfloor n\delta\rfloor} \frac{(i-t_1)^2(k-i)^2}{(t_2-t_1+1)^2(k-t_1+1)^2}(\widehat{\boldsymbol\beta}_{t_1,i} - \widehat{\boldsymbol\beta}_{i+1,k})^{\otimes 2}.$$

By Lemma 5, we obtain that

$$\left\{(n\kappa^2)^{-1}L_{n,\delta}(\lfloor nu_1\rfloor,\lfloor nq\rfloor,\lfloor nu_2\rfloor)\right\}_{q\in(\epsilon,1-\epsilon),(u_1,u_2)\in G^{(1)}(q)} \Rightarrow \left\{\overline{L}_\delta(u_1,q,u_2)\right\}_{q\in(\epsilon,1-\epsilon),(u_1,u_2)\in G^{(1)}(q)},$$

where

$$\overline{L}_\delta(u_1,q,u_2) = \int_{u_1+\delta}^{q-\delta} \frac{(r-u_1)^2(q-r)^2}{(u_2-u_1)^2(q-u_1)^2}(\mathbf{c}(u_1,r)-\mathbf{c}(r,q))^{\otimes 2}dr, \tag{S15}$$

with $\mathbf{c}(u_1, r)$ and $\mathbf{c}(r, q)$ defined in Lemma 5.

Therefore, since $\overline{L}_\delta(u_1, q, u_2)$ is invertible by similar arguments in Lemma 10 when $q \in M_\eta$, $(u_1, u_2) \in G^{(1)}(q)$, then using the fact that $V_{n,\delta}(t_1, k, t_2)^{-1} - L_{n,\delta}(t_1, k, t_2)^{-1} \le 0$, we have

$$\max_{k\in M_{n,\eta}} T^{(1)}_{n,\epsilon,\delta}(k)$$

$$\le \max_{k\in M_{n,\eta}} \max_{(t_1,t_2)\in G^{(1)}(k)} (n\kappa^2)^{-1/2} D_n(t_1,k,t_2)^\top [(n\kappa^2)^{-1} L_{n,\delta}(t_1,k,t_2)]^{-1} (n\kappa^2)^{-1/2} D_n(t_1,k,t_2)$$

$$\Rightarrow \sup_{q\in M_\eta} \max_{(u_1,u_2)\in G^{(1)}(q)} \left\{ \frac{q^2(1-q)^2}{(u_2-u_1)^3} \Big(\mathbf{c}(u_1,q)-\mathbf{c}(q,u_2)\Big)^\top \overline{L}_\delta(u_1,q,u_2)^{-1} \Big(\mathbf{c}(u_1,q)-\mathbf{c}(q,u_2)\Big) \right\} < \infty,$$

i.e.

$$P(\max_{k\in M_{n,\eta}} T^{(1)}_{n,\epsilon,\delta}(k) < \zeta_n) \to 1.$$

This and (S13) complete the proof for (S10).

To finish the proof, note that by (S10), the points above the threshold ζ_n can only be in the neighborhood of change-point locations with probability tending to 1, i.e., $\forall \eta > 0$,

$$P(\{k : T_{n,\epsilon,\delta}(k) > \zeta_n\} \subset M^c_{n,\eta}) \to 1. \tag{S16}$$

On one hand, for each change-point location k_i, since the algorithm only allows for one change-point k_i itself in the interval $(k_i - h, k_i + h)$, by choosing $\eta < (\epsilon_o - \epsilon)/2$, we can ensure that $P(\{k : T_{n,\epsilon,\delta}(k) > \zeta_n\} \cap (k_i - h, k_i + h) \subset \{k : |k - k_i| < n\eta\}) \to 1$. That is to say, if some points in the interval $(k_i - h, k_i + h)$ are above the threshold ζ_n, they have to be very close to k_i, i.e., they are all in the neighborhood $\{k : |k - k_i| < n\eta\}$. By (S9), these points are nonempty. Therefore, the local maximizer, say $\widehat{k}_i$ in $(k_i - h, k_i + h)$ should also be in the $\{k : |k - k_i| < n\eta\}$ such that $T_{n,\epsilon,\delta}(\widehat{k}_i) \ge T_{n,\epsilon,\delta}(k_i) > \zeta_n$. Clearly, we have $|\widehat{q}_i - q_i| = |\frac{\widehat{k}_i}{n} - \frac{k_i}{n}| < \eta$.

On the other hand, for any estimated change-point location $\widehat{q}_i \in \widehat{\mathbf{q}}$, by (S16), it has to be in the neighborhood of a true change-point. That is, there is at least one true change-point, say k_i in $(\widehat{k}_i - n\eta, \widehat{k}_i + n\eta)$. Note that by assumption, there could only be one-change-point in such neighborhood as the length of this interval is $2n\eta < \lfloor n\epsilon_o \rfloor$, hence such change-point is unique.

Therefore, we have established one-to-one mapping of $\widehat{\mathbf{q}}$ and $\mathbf{q}$. The consistency follows by the assumption that the number of change-points is finite. ∎

S3 | AUXILIARY LEMMAS AND PROOFS

Proof of Lemma 3: Before the change-point location, by Lemma 2 (i), we obtain that for all $\eta \in (0, q_1)$

$$\sup_{(r_1,r_2)\in\omega(0,q_1;\eta)} \left| \sqrt{n}(\widehat{\beta}_{\lfloor r_1 n\rfloor+1,\lfloor nr_2\rfloor} - \beta^{(1)}) - \frac{1}{\sqrt{n}} [\Sigma_1(r_2) - \Sigma_1(r_1)]^{-1} f(0)^{-1} \sum_{t=\lfloor r_1 n\rfloor+1}^{\lfloor r_2 n\rfloor} X_t \psi_\tau(\varepsilon_t) \right| = o_p(1).$$

By Lemma 2 (ii) and Lemma 6 in Zhou and Shao (2013),

$$\left\{ \frac{1}{\sqrt{n}} [\Sigma_1(r_2) - \Sigma_1(r_1)]^{-1} f(0)^{-1} \sum_{t=\lfloor r_1 n\rfloor+1}^{\lfloor r_2 n\rfloor} X_t \psi_\tau(\varepsilon_t) \right\}_{0\le r_1\le r_2-\eta\le q_1-\eta}$$
$$\Rightarrow \Gamma_\varepsilon \left\{ [\Sigma_1(r_2) - \Sigma_1(r_1)]^{-1} \int_{r_1}^{r_2} X(s)dB(s) \right\}_{0\le r_1\le r_2-\eta\le q_1-\eta}.$$

Therefore, by continuous mapping theorem, we obtain

$$\left\{\sqrt{n}(\widehat{\beta}_{1,\lfloor nr\rfloor}-\widehat{\beta}_{\lfloor nr\rfloor+1,\lfloor nq_1\rfloor})\right\}_{r\in(\delta,q_1-\delta)}\Rightarrow\Gamma_\varepsilon\left\{\Sigma_1(r)^{-1}B_X(r)-[\Sigma_1(q_1)-\Sigma_1(r)]^{-1}[B_X(q_1)-B_X(r)]\right\}_{r\in(\delta,q_1-\delta)}.$$

Similarly, after the change-point location, we also have

$$\begin{aligned}&\left\{\sqrt{n}(\widehat{\beta}_{\lfloor nr\rfloor+1,n}-\widehat{\beta}_{\lfloor nq_1\rfloor+1,\lfloor nr\rfloor})\right\}_{r\in(q_1+\delta,1-\delta)}\\ \Rightarrow&\Gamma_\varepsilon\left\{[\Sigma_2(1)-\Sigma_2(r)]^{-1}[B_X(1)-B_X(r)]-[\Sigma_2(r)-\Sigma_2(q_1)]^{-1}[B_X(r)-B_X(q_1)]\right\}_{r\in(q_1+\delta,1-\delta)}.\end{aligned}$$

Then the continuous mapping theorem indicates that

$$\begin{aligned}L_{n,\delta}(1,k_1,n)&\to_D\Gamma_\varepsilon^2\int_\delta^{q_1-\delta}\frac{r^2(q_1-r)^2}{q_1^2}\left\{\Lambda_1(0,r)-\Lambda_1(r,q_1)\right\}^{\otimes 2}dr=\Gamma_\varepsilon^2L_\delta(q_1),\\ R_{n,\delta}(1,k_1,n)&\to_D\Gamma_\varepsilon^2\int_{q_1+\delta}^{1-\delta}\frac{(r-q_1)^2(1-r)^2}{(1-q_1)^2}\left\{\Lambda_2(r,1)-\Lambda_2(q_1,r)\right\}^{\otimes 2}dr=\Gamma_\varepsilon^2R_\delta(q_1).\end{aligned}$$

In addition, we have that

$$D_n(1,k_1,n)=\frac{k_1(n-k_1)}{n^2}\sqrt{n}(\widehat{\beta}_{1,k_1}-\widehat{\beta}_{k_1+1,n}),$$

where it is not hard to see that

$$\sqrt{n}\left[(\widehat{\beta}_{1,k_1}-\widehat{\beta}_{k_1+1,n})+\mathbf{b}\right]\Rightarrow\Gamma_\varepsilon\Lambda_1(0,q_1)-\Gamma_\varepsilon\Lambda_2(q_1,1),$$

hence we can see that

$$n^{-1/2}\kappa^{-1}D_n(1,k_1,n)=(q_1)(1-q_1)(-\kappa^{-1}\mathbf{b})+O_p(n^{-1/2}\kappa^{-1}).$$

Then the continuous mapping theorem indicates that (S1) holds. ∎

Proof of Lemma 4: Note that Lemma 6 shows that $Z_n(\mathbf{c},(a,b))$ converges in probability to $Z_\infty(\mathbf{c},(a,b))\in\mathbb{R}$, which then implies the finite dimensional convergence in probability, i.e.

$$\Big(Z_n(\mathbf{c},(a_1,b_1)),Z_n(\mathbf{c},(a_2,b_2)),\cdots,Z_n(\mathbf{c},(a_k,b_k))\Big)\to_p\Big(Z_\infty(\mathbf{c},(a_1,b_1)),Z_\infty(\mathbf{c},(a_2,b_2)),\cdots,Z_\infty(\mathbf{c},(a_k,b_k))\Big).$$

Further, Lemma 7 implies the stochastic equicontinuity of $Z_n(\mathbf{c},\cdot)$. Therefore, we have shown that

$$Z_n(\mathbf{c},\cdot)\Rightarrow(\text{or }\to_p)Z_\infty(\mathbf{c},\cdot). \tag{S17}$$

Since $Z_\infty(\mathbf{c},\cdot)\in C(\mathcal{T})$ is a nonrandom function, and the convergence is in probability, hence we have

$$\Big(Z_n(\mathbf{c}_1,\cdot),Z_n(\mathbf{c}_2,\cdot),\cdots,Z_n(\mathbf{c}_k,\cdot)\Big)\Rightarrow\Big(Z_\infty(\mathbf{c}_1,\cdot),Z_\infty(\mathbf{c}_2,\cdot),\cdots,Z_\infty(\mathbf{c}_k,\cdot)\Big). \tag{S18}$$

Finally, the result of Lemma 4 will follow by Theorem 1 in Kato (2009) once the following three conditions are satisfied:

(i) $Z_n(\mathbf{c},\mathbf{r})$ and $Z_\infty(\mathbf{c},\mathbf{r})$ are convex in $\mathbf{c}$ for each $\mathbf{r}$ and continuous in $\mathbf{r}$ for each $\mathbf{c}$;

(ii) $\mathbf{c}(\mathbf{r})$ is the unique minimum point of $Z_\infty(\mathbf{c},\mathbf{r})$ for each $\mathbf{r}$;

(iii) $\widehat{\mathbf{c}}_n(\mathbf{r}) \in l^\infty(\mathcal{T})$ is bounded.

It suffices to consider (iii). In fact, by (S17) and detailed investigation of proof in Theorem 1 in Kato (2009), it can be shown that there exists a sequence of bounded stochastic processes $\widehat{\mathbf{c}}_n(\cdot)$ uniformly converging in probability to $\mathbf{c}(\cdot)$ such that with probability approaching one (see discussions after the proof of Theorem 1 in Kato (2009)). ∎

Lemma 6 *Under the conditions of Lemma 4, for any fixed* $(a,b)\in\mathcal{T}$,

$$Z_n(\mathbf{c},(a,b)) \to_p Z_\infty(\mathbf{c},(a,b)). \tag{S19}$$

Proof of Lemma 6: By the identity in Knight (1998), i.e. $\rho_\tau(u-v)-\rho_\tau(u) = -v\psi_\tau(u)+\int_0^v[1(u\le s)-1(u\le 0)]ds$, we have

$$\begin{aligned}
&Z_n(\mathbf{c},(a,b))\\
=&\frac{1}{n\kappa^2}\sum_{t=\lfloor na\rfloor+1}^{k_1}[\rho_\tau(\varphi_t-\kappa X_t^\top\mathbf{c})-\rho_\tau(\varphi_t)]+\frac{1}{n\kappa^2}\sum_{t=k_1+1}^{\lfloor nb\rfloor}[\rho_\tau(\varphi_t-\kappa X_t^\top(\mathbf{c}-\boldsymbol{\alpha}))-\rho_\tau(\varphi_t)]\\
&+\frac{(nb-\lfloor nb\rfloor)}{n\kappa^2}g_{\lfloor nb\rfloor+1}(\boldsymbol{\beta})+\frac{(\lfloor na\rfloor+1-na)}{n\kappa^2}g_{\lfloor na\rfloor}(\boldsymbol{\beta})\\
=&\frac{1}{n\kappa}\sum_{t=\lfloor na\rfloor+1}^{k_1}-X_t^\top\mathbf{c}\psi_\tau(\varphi_t)+\frac{1}{n\kappa^2}\sum_{t=\lfloor na\rfloor+1}^{k_1}\int_0^{\kappa X_t^\top\mathbf{c}}[1(\varphi_t\le s)-1(\varphi_t\le 0)]ds\\
&+\frac{1}{n\kappa}\sum_{t=k_1+1}^{\lfloor nb\rfloor}-X_t^\top(\mathbf{c}-\boldsymbol{\alpha})\psi_\tau(\varphi_t)+\frac{1}{n\kappa^2}\sum_{t=\lfloor na\rfloor+1}^{k_1}\int_0^{\kappa X_t^\top(\mathbf{c}-\boldsymbol{\alpha})}[1(\varphi_t\le s)-1(\varphi_t\le 0)]ds\\
&+\frac{(nb-\lfloor nb\rfloor)}{n\kappa^2}g_{\lfloor nb\rfloor+1}(\boldsymbol{\beta})+\frac{(\lfloor na\rfloor+1-na)}{n\kappa^2}g_{\lfloor na\rfloor}(\boldsymbol{\beta})\\
=:&\sum_{i=1}^{6}Z_{ni}(\mathbf{c},(a,b)).
\end{aligned}$$

Since $n\kappa^2\to\infty$ as $n\to\infty$, it is easy to see that $Z_{n5}(\mathbf{c},(a,b))+Z_{n6}(\mathbf{c},(a,b))=o_p(1)$. Under Assumptions 1-3 and the fact that $n\kappa^2\to\infty$, from Lemma 8 (i), it follows that $Z_{n1}(\mathbf{c},(a,b))+Z_{n3}(\mathbf{c},(a,b))=o_p(1)$.

By a change-of-variable, we get

$$Z_{n2}(\mathbf{c},(a,b)) = \frac{1}{n^{1/2}\kappa}\frac{1}{\sqrt{n}}\sum_{t=\lfloor na\rfloor+1}^{k_1}\int_0^{X_t^\top\mathbf{c}}[1(\varphi_t\le\kappa s)-1(\varphi_t\le 0)]ds.$$

Then, by Lemma 8 (ii), we have $\mathrm{Var}(Z_{n2}(\mathbf{c},(a,b)))=o(1)$ and similarly $\mathrm{Var}(Z_{n4}(\mathbf{c},(a,b)))=o(1)$. Hence, by Chebyshev's inequality, we have

$$Z_n(\mathbf{c},(a,b)) = EZ_{n2}(\mathbf{c},(a,b))+EZ_{n4}(\mathbf{c},(a,b))+o_p(1).$$

Note that $P(\varphi_t\le s)=P(\varepsilon_t\le[X_t^\top\gamma_t]^{-1}s)$, then by change-of-variable, the fact that $|X_t^\top\mathbf{c}|\le\|\mathbf{c}\|_2$, $X_t^\top\gamma_t>0$

and mean value theorem, we have

$$
\begin{aligned}
EZ_{n2}(\mathbf{c},(a,b)) =&\frac{1}{n\kappa^2}\sum_{t=\lfloor na\rfloor}^{k_1}\int_0^{\kappa X_t^\top \mathbf{c}}[F([X_t^\top \gamma_t]^{-1}s)-F(0)]ds=\frac{1}{n\kappa}\sum_{t=\lfloor na\rfloor}^{k_1}\int_0^{X_t^\top \mathbf{c}}[F([X_t^\top \gamma_t]^{-1}\kappa s)-F(0)]ds \\
=&\frac{1}{n}\sum_{t=\lfloor na\rfloor}^{k_1}\int_0^{X_t^\top \mathbf{c}}\Big([X_t^\top\gamma_t]^{-1}f(0)s+o(1)\Big)ds=\frac{f(0)}{2n}\sum_{t=\lfloor na\rfloor}^{k_1}\frac{[X_t^\top \mathbf{c}]^2}{[X_t^\top\gamma^{(1)}]}+o(1).
\end{aligned} \tag{S20}
$$

Similarly, we have

$$
EZ_{n4}(\mathbf{c},(a,b))=\frac{f(0)}{2n}\sum_{t=k_1+1}^{\lfloor nb\rfloor}\frac{[X_t^\top(\mathbf{c}-\boldsymbol{\alpha})]^2}{[X_t^\top\gamma^{(2)}]}+o(1).
$$

Hence,

$$
\begin{aligned}
Z_n(\mathbf{c},(a,b)) =&EZ_{n2}(\mathbf{c},(a,b))+EZ_{n4}(\mathbf{c},(a,b))+o_p(1) \\
=&\frac{f(0)}{2}\Big\{\int_a^{q_1}\frac{[(1,x)\mathbf{c}]^2}{(1,x)\gamma^{(1)}}dx+\int_{q_1}^{b}\frac{[(1,x)(\mathbf{c}-\boldsymbol{\alpha})]^2}{(1,x)\gamma^{(2)}}dx\Big\}+o_p(1),
\end{aligned}
$$

i.e. (S19) is proved. ∎

Lemma 7 *Under conditions of Lemma 4, for any $x>0$, and $(a_1,b_1),(a_2,b_2)\in\mathcal{T}$,*

$$
\lim_{\Delta\downarrow 0}\limsup_{n\to\infty}P\Big(\sup_{|a_1-a_2|\le\Delta,|b_1-b_2|\le\Delta}|Z_n(\mathbf{c},(a_1,b_1))-Z_n(\mathbf{c},(a_2,b_2))|>x\Big)=0. \tag{S21}
$$

Proof of Lemma 7: It suffices to show that for $\forall\eta>0$ and $\forall x>0$, there exists a $\Delta>0$ such that for n large enough,

$$
P\Big(\sup_{|a_1-a_2|\le\Delta,|b_1-b_2|\le\Delta}|Z_n(\mathbf{c},(a_1,b_1))-Z_n(\mathbf{c},(a_2,b_2))|>x\Big)<\eta. \tag{S22}
$$

Note that $\max\{a_1,a_2\}<\min\{b_1,b_2\}$, there are only four types of configurations for (a_1,a_2,b_1,b_2). We only consider the case $a_1<a_2<b_1<b_2$ since other cases are similar.

Let $T_n(\mathbf{c},(a,b))=\frac{1}{n\kappa^2}\sum_{t=\lfloor na\rfloor+1}^{k_1}[\rho_\tau(\varphi_t-\kappa X_t^\top\mathbf{c})-\rho_\tau(\varphi_t)]+\frac{1}{n\kappa^2}\sum_{t=k_1+1}^{\lfloor nb\rfloor}[\rho_\tau(\varphi_t-\kappa X_t^\top(\mathbf{c}-\boldsymbol{\alpha}))-\rho_\tau(\varphi_t)]$. Using the boundedness of $\rho_\tau(\cdot)$, it is easy to see that

$$
\begin{aligned}
&\sup_{|a_1-a_2|\le\Delta,|b_1-b_2|\le\Delta}|Z_n(\mathbf{c},(a_1,b_1))-Z_n(\mathbf{c},(a_2,b_2))| \\
\le&\sup_{|a_1-a_2|\le\Delta,|b_1-b_2|\le\Delta}|T_n(\mathbf{c},(a_1,b_1))-T_n(\mathbf{c},(a_2,b_2))|+\frac{4}{n\kappa^2} \\
\le&\sup_{|a_1-a_2|\le\Delta,b_1}|T_n(\mathbf{c},(a_1,b_1))-T_n(\mathbf{c},(a_2,b_1))|+\sup_{|b_1-b_2|\le\Delta,a_2}|T_n(\mathbf{c},(a_2,b_1))-T_n(\mathbf{c},(a_2,b_2))|+\frac{4}{n\kappa^2} \\
=&\sup_{|a_1-a_2|\le\Delta}|T_n(\mathbf{c},(a_1,q_1))-T_n(\mathbf{c},(a_2,q_1))|+\sup_{|b_1-b_2|\le\Delta}|T_n(\mathbf{c},(q_1,b_1))-T_n(\mathbf{c},(q_1,b_2))|+\frac{4}{n\kappa^2}.
\end{aligned}
$$

Note that $(n\kappa^2)^{-1}\to 0$, hence it suffices to show the tightness of $T_n(\mathbf{c},(\cdot,q_1))$ and $T_n(\mathbf{c},(q_1,\cdot))$. Since they are similar, we only prove it for $T_n(\mathbf{c},(q_1,\cdot))$.

Denote $T_n^*(\mathbf{c},(q_1,b_1) = T_n(\mathbf{c},(q_1,b_1)) - ET_n(\mathbf{c},(q_1,b_1))$ as the demeaned version. Then

$$
\begin{aligned}
&\sup_{|b_2-b_1|\le\Delta} |T_n(\mathbf{c},(q_1,b_1)) - T_n(\mathbf{c},(q_1,b_2)| \\
\le &\sup_{|b_2-b_1|\le\Delta} |ET_n(\mathbf{c},(q_1,b_1)) - ET_n(\mathbf{c},(q_1,b_2)| + \sup_{|b_2-b_1|\le\Delta} |T_n^*(\mathbf{c},(q_1,b_1)) - T_n^*(\mathbf{c},(q_1,b_2)|.
\end{aligned}
$$

First, using Knight's identity, and the fact that $\int_0^x F(cs) - F(0)ds \ge 0$ for all x and $c > 0$, we obtain

$$
\sup_{|b_2-b_1|\le\Delta} |ET_n(\mathbf{c},(q_1,b_1)) - ET_n(\mathbf{c},(q_1,b_2)| \le \sup_{b_1} \frac{1}{n\kappa^2} \sum_{t=\lfloor nb_1\rfloor+1}^{\lfloor n(b_1+\Delta)\rfloor} \int_0^{\kappa X_t^\top(\mathbf{c}-\boldsymbol{\alpha})} [F([X_t^\top\gamma_t]^{-1}s) - F(0)]ds \le C\Delta
$$

by similar arguments used in (S20).

Second, denote $h_t(\mathbf{c}) = [\rho_t(\varphi_t - \kappa X_t^\top(\mathbf{c}-\boldsymbol{\alpha})) - \rho_\tau(\varphi_t)] - [E\rho_t(\varphi_t - \kappa X_t^\top(\mathbf{c}-\boldsymbol{\alpha})) - E\rho_\tau(\varphi_t)]$. For any $q_1 < s_1 < s_2$, using Knight's identity and Lemma 8, we have

$$
\begin{aligned}
&E[T_n^*(\mathbf{c},(q_1,s_1)) - T_n^*(\mathbf{c},(q_1,s_2))]^2 = \frac{1}{n^2\kappa^4} E\Big\{\sum_{t=\lfloor ns_1\rfloor+1}^{\lfloor ns_2\rfloor} h_t(\mathbf{c})\Big\}^2 \\
\le &\frac{2}{n^2\kappa^2} E\Big\{\sum_{t=\lfloor ns_1\rfloor+1}^{\lfloor ns_2\rfloor} X_t^\top(\mathbf{c}-\boldsymbol{\alpha})\psi_\tau(\varphi_t)\Big\}^2 \\
&+ \frac{2}{n^2\kappa^2} E\Big\{\sum_{t=\lfloor ns_1\rfloor+1}^{\lfloor ns_2\rfloor} \int_0^{\kappa X_t^\top(\mathbf{c}-\boldsymbol{\alpha})} [1(\varphi_t \le s) - 1(\varphi_t \le 0)] - [F([X_t^\top\gamma_t]^{-1}s) - F(0)]ds\Big\}^2 \\
\le &C\frac{(s_2-s_1)}{n\kappa^2}.
\end{aligned}
\tag{S23}
$$

Then, we see that

$$
\begin{aligned}
&P(\sup_{q_1<\min\{b_1,b_2\},|b_2-b_1|\le\Delta} |T_n^*(\mathbf{c},(q_1,b_1)) - T_n^*(\mathbf{c},(q_1,b_2)| > x) \\
\le &\sum_{i=0}^{\lfloor\Delta^{-1}(1-q_1)\rfloor} P(\sup_{q_1+i\Delta\le s\le q_1+(i+1)\Delta} |T_n^*(\mathbf{c},(q_1,s)) - T_n^*(\mathbf{c},(q_1,q_1+i\Delta))| > \frac{x}{3}) \\
\le &x^{-2}C\sum_{i=0}^{\lfloor\Delta^{-1}(1-q_1)\rfloor} E\Big\{\sup_{q_1+i\Delta\le s\le q_1+(i+1)\Delta} |T_n^*(\mathbf{c},(q_1,s)) - T_n^*(\mathbf{c},(q_1,q_1+i\Delta))|\Big\}^2 \\
= &x^{-2}C\sum_{i=0}^{\lfloor\Delta^{-1}(1-q_1)\rfloor} E\Big\{\max_{1\le j\le\lfloor n\Delta\rfloor} |\frac{1}{n\kappa^2}\sum_{t=\lfloor n(q_1+i\Delta)\rfloor}^{j+\lfloor n(q_1+i\Delta)\rfloor} h_t(\mathbf{c})|\Big\}^2 \\
\le &x^{-2}C\Delta^{-1}(1-q_1)[\log_2(4n\Delta)]^2\frac{\Delta}{n\kappa^2} = C\frac{\log^2(n)}{n\kappa^2} \to 0.
\end{aligned}
$$

In the above expressions, the first inequality holds by equation (8.6) in the proof of Theorem 8.3 in Billingsley (1968), the second by Chebyshev's inequality and the last by Proposition 1 in Wu (2007) and (S23). ∎

Lemma 8 *Under Assumptions 1-3, for any $0 \le r_1 < r_2 \le 1$ and for any $\mathbf{c} \in \mathbb{R}^2$, we have for some constant $C > 0$,*

(i)

$$\text{Var}\Big(\frac{1}{\sqrt{n}}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} X_t^\top \mathbf{c}\psi_\tau(\varphi_t)\Big) \le C(r_2-r_1).$$

(ii)

$$\text{Var}\Big(\frac{1}{\sqrt{n}}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} \int_0^{X_t^\top \mathbf{c}} [1(\varphi_t \le s) - 1(\varphi_t \le 0)]ds\Big) \le C(r_2-r_1).$$

Proof of Lemma 8 (i) Observe that $\psi_\tau(\varphi_t) = \tau - 1(\varphi_t < 0) = \tau - 1([X_t^\top \gamma_t]\varepsilon_t < 0)$. Note that $X_t^\top \gamma_t > 0$, hence we obtain $\psi_\tau(\varphi_t) = \psi_\tau(\varepsilon_t)$. Then

$$\begin{aligned}\text{Var}\Big(\frac{1}{\sqrt{n}}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} X_t^\top \mathbf{c}\psi_\tau(\varphi_t)\Big) =&\text{Var}\Big(\frac{1}{\sqrt{n}}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} X_t^\top \mathbf{c}\psi_\tau(\varepsilon_t)\Big)\\ \le&\frac{1}{n}\sum_{t,t'=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} X_t^\top \mathbf{c} X_{t'}^\top \mathbf{c}\text{Cov}(\psi_\tau(\varepsilon_t),\psi_\tau(\varepsilon_{t'}))\\ \le&\frac{\|\mathbf{c}\|_2^2}{n}\sum_{t,t'=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor}\Big|\text{Cov}(1(\varepsilon_t\le 0),1(\varepsilon_{t'}\le 0))\Big|.\end{aligned}$$

By Lemma 9, we obtain that

$$\text{Var}\Big(\frac{1}{\sqrt{n}}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} X_t^\top \mathbf{c}\psi_\tau(\varepsilon_t)\Big) \le \frac{1}{n}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor}\sum_{j=0}^{\infty} C\delta_G^*(j,p)^{p/(2p+2)}.$$

Note that by Assumption 3, we have $\delta_G(k,4)$ decays with rate faster than $O(k^{-4})$, hence $\delta_G^*(k,4)$ decays with rate faster than $O(k^{-3})$. Then choosing $p=4$ implies $\sum_{j=0}^{\infty}\delta_G^*(j,p)^{p/(2p+2)} < \infty$, hence (i) follows.

(ii) Similar to (i), we have that

$$\begin{aligned}&\text{Var}\Big(\frac{1}{\sqrt{n}}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} \int_0^{X_t^\top \mathbf{c}} [1(\varphi_t \le s) - 1(\varphi_t \le 0)]ds\Big)\\ =&\text{Var}\Big(\frac{1}{\sqrt{n}}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} \int_0^{X_t^\top \mathbf{c}} [1(\varepsilon_t \le [X_t^\top\gamma_t]^{-1}s) - 1(\varepsilon_t \le 0)]ds\Big)\\ \le&\frac{1}{n}\sum_{t,t'=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor}\int_0^{X_t^\top \mathbf{c}}\int_0^{X_{t'}^\top \mathbf{c}}|\text{Cov}(d_{\varepsilon_t}([X_t^\top\gamma_t]^{-1}s),d_{\varepsilon_{t'}}([X_{t'}^\top\gamma_{t'}]^{-1}s'))|dsds',\end{aligned}$$

where $d_{\varepsilon_t}(s) = [1(\varepsilon_t \le s) - 1(\varepsilon_t \le 0)]$.

Note that $X_t^\top\gamma_t$ is a bounded sequence, and

$$\begin{aligned}\text{Cov}(d_{\varepsilon_t}(s),d_{\varepsilon_{t'}}(s')) =&\text{Cov}(1(\varepsilon_t\le s),1(\varepsilon_{t'}\le s')) + \text{Cov}(1(\varepsilon_t\le 0),1(\varepsilon_{t'}\le 0))\\ &-\text{Cov}(1(\varepsilon_t\le s),1(\varepsilon_{t'}\le 0)) - \text{Cov}(1(\varepsilon_t\le 0),1(\varepsilon_{t'}\le s')),\end{aligned}$$

hence by Lemma 9, (ii) can be proved based on a similar argument used in (i). ∎

Lemma 9 *For any $s, s' \in \mathbb{R}$,*

$$|\text{Cov}(1(\varepsilon_0 \le s), 1(\varepsilon_t \le s'))| \le C\delta_G^*(t,p)^{p/(2p+2)}$$

for some constant $C > 0$ independent of s and s'.

Proof of Lemma 9 Note that for any $s, s' \in \mathbb{R}$,

$$\begin{aligned}\text{Cov}(1(\varepsilon_0 \le s), 1(\varepsilon_t \le s')) =& E[1(\varepsilon_0 \le s)1(\varepsilon_t \le s')] - E[1(\varepsilon_0 \le s)]E[1(\varepsilon_t \le s')] \\ =& E\{[1(\varepsilon_t \le s') - 1(\varepsilon_t^* \le s')]1(\varepsilon_0 \le s)\},\end{aligned}$$

where we use the fact that $\varepsilon_t^* = G(\mathcal{F}_i^*)$ is independent of ε_0. Then, by the Cauchy-Schwarz inequality, we obtain that $|\text{Cov}(1(\varepsilon_0 \le s), 1(\varepsilon_t \le s'))| \le \|1(\varepsilon_t \le s') - 1(\varepsilon_t^* \le s')]\|_2$.

Furthermore, we can show that for any $s \in \mathbf{R}$ and any $\eta > 0$,

$$\begin{aligned}&E|[1(\varepsilon_t \le s') - 1(\varepsilon_t^* \le s')]1(|\varepsilon_t - \varepsilon_t^*| \le \eta)|^2 \\ =& P(\varepsilon_t \le s, \varepsilon_t^* > s, |\varepsilon_t - \varepsilon_t^*| \le \eta) + P(\varepsilon_t > s, \varepsilon_t^* \le s, |\varepsilon_t - \varepsilon_t^*| \le \eta) \le 2P(s < \varepsilon_t \le s + \eta) \le 2C\eta,\end{aligned}$$

by Assumption 1. For some $p \ge 1$

$$E|[1(\varepsilon_t \le s') - 1(\varepsilon_t^* \le s')]1(|\varepsilon_t - \varepsilon_t^*| > \eta)|^2 \le P(|\varepsilon_t - \varepsilon_t^*| > \eta) \le \eta^{-p}E|\varepsilon_t - \varepsilon_t^*|^p = \delta_G^*(t,p).$$

By choosing $\eta = [\delta_G^*(t,p)^p/(2C)]^{1/(p+1)}$, we obtain $\|1(\varepsilon_t \le s') - 1(\varepsilon_t^* \le s')\|_2 \le 2(2C)^{p/(p+1)}\delta_G^*(t,p)^{p/(2p+2)}$.
∎

Lemma 10 *$\overline{R}_{\delta,2}(q)$ defined in (S7) is invertible for $q < q_1$ when $\|\boldsymbol{\alpha}\| \ne 0$.*

Proof of Lemma 10: Recall that for $q < q_1 < r$, we have $\mathbf{c}(q,r) = [\Sigma_1(q_1) - \Sigma_1(q) + \Sigma_2(r) - \Sigma_2(q_1)]^{-1}[\Sigma_2(r) - \Sigma_2(q_1)]\boldsymbol{\alpha}$. Therefore,

$$\overline{R}_{\delta,2}(q) = \int_{q_1+\delta}^{1-\delta} \frac{(r-q)^2(1-r)^2}{(1-q)^2}\Big\{[\Sigma_1(q_1) - \Sigma_1(q) + \Sigma_2(r) - \Sigma_2(q_1)]^{-1}[\Sigma_1(q_1) - \Sigma_1(q)]\boldsymbol{\alpha}\Big\}^{\otimes 2} dr.$$

Denote

$$\frac{(r-q)(1-r)}{(1-q)}\Big\{[\Sigma_1(q_1) - \Sigma_1(q) + \Sigma_2(r) - \Sigma_2(q_1)]^{-1}[\Sigma_1(q_1) - \Sigma_1(q)]\boldsymbol{\alpha}\Big\} := \Big(g_1(r), g_2(r)\Big)^\top,$$

then the invertibility of $\overline{R}_{\delta,2}(q)$ is equivalent to $\det(\overline{R}_{\delta,2}(q)) > 0$ since $\overline{R}_{\delta,2}(q)$ is semi-positive definite, i.e.

$$\int_{q_1+\delta}^{1-\delta} g_1(r)^2 dr \int_{q_1+\delta}^{1-\delta} g_2(r)^2 dr - [\int_{q_1+\delta}^{1-\delta} g_1(r)g_2(r)dr]^2 > 0.$$

Then, by Cauchy-Schwarz inequality for integrals, the invertibility of $\overline{R}_{\delta,2}(q)$ is equivalent to the following statement:

$$\frac{g_1(r,q,q_1,\delta,\alpha_1,\alpha_2)}{g_2(r,q,q_1,\delta,\alpha_1,\alpha_2)} \neq c \quad \text{for any constant } c \text{ uniformly in } r \in (q_1+\delta, 1-\delta).$$

We will prove by contradiction. Denote $[\Sigma_1(q_1)-\Sigma_1(q)-\Sigma_2(q_1)] = \{A_{ij}\}_{i,j=1,2}$, $[\Sigma_1(q_1)-\Sigma_1(q)]\alpha := (w_1,w_2)^\top$, then notice that when $\gamma_0^{(i)} \neq 0$, we have

$$\Sigma_2(r) = \begin{pmatrix} \frac{[\ln|\gamma_0^{(2)}+\gamma_1^{(2)}r|-\ln|\gamma_0^{(2)}|]}{\gamma_1^{(2)}} & \frac{\gamma_1^{(2)}r-\gamma_0^{(2)}[\ln|\gamma_0^{(2)}+\gamma_1^{(2)}r|-\ln|\gamma_0^{(2)}|]}{\gamma_1^{(2)2}} \\ \frac{\gamma_1^{(2)}r-\gamma_0^{(2)}[\ln|\gamma_0^{(2)}+\gamma_1^{(2)}r|-\ln|\gamma_0^{(2)}|]}{\gamma_1^{(2)2}} & \frac{\gamma_0^{(2)2}[\ln|\gamma_0^{(2)}+\gamma_1^{(2)}r|-\ln|\gamma_0^{(2)}|]-\gamma_1^{(2)}\gamma_0^{(2)}r}{\gamma_1^{(2)3}} + \frac{r^2}{2\gamma_1^{(2)}} \end{pmatrix} := \{B_{ij}(r)\}_{i,j=1,2}$$

Now, suppose $\frac{g_1(r)}{g_2(r)} = c$ for some constant c uniformly in r, note that $(g_1,g_2)^\top = [A+B(r)]^{-1}(w_1,w_2)^\top$, we have

$$\frac{[B_{22}(r)+A_{22}]w_1 - [B_{12}(r)+A_{12}]w_2}{-[B_{21}(r)+A_{21}]w_1 - [B_{11}(r)+A_{11}]w_2} = c.$$

When we compare the coefficients of r^2 in both denominator and numerator of the above equation, we can conclude the $w_1 B_{22}(r) = 0$, i.e. $w_1 = 0$. Then we compare the coefficients of r and determine that $w_2 = 0$. This implies that $\alpha_1 = \alpha_2 = 0$, which contradicts the model setting. ∎

Proposition 11 *Suppose $X = (X_1,\cdots,X_d)^\top \in \mathbb{R}^d$ follows an elliptical distribution, i.e. there exist a vector $\mu = (\mu_1,\cdots,\mu_d)^\top \in \mathbb{R}^d$, a positive semi-definite matrix M and a function $\Psi : \mathbb{R}^+ \to \mathbb{R}$ such that the characteristic function $\varphi_{X-\mu}(t)$ of $X-\mu$ takes the form $\varphi_{X-\mu}(t) = \Psi(t^\top M t)$, $t \in \mathbb{R}^d$. For $\tau \in (0.5,1)$, we have $Q_\tau(\sum_{i=1}^d X_i) \le \sum_{i=1}^d Q_\tau(X_i)$, and for $\tau \in (0,0.5)$, we have $Q_\tau(\sum_{i=1}^d X_i) \ge \sum_{i=1}^d Q_\tau(X_i)$.*

Proof of Proposition 11: Let $e_i \in \mathbb{R}^d$ be the vector with i-th entry 0 and other entries 1, then by the definition of X, we have for any $t \in \mathbb{R}$

$$E\Big(\exp(itM_{ii}^{-1/2}e_i^\top[X-\mu])\Big) = \Psi(t^2 M_{ii}^{-1}e_i^\top M e_i) = \Psi(t^2)$$

where M_{ij} denotes the (i,j)th entry of M. That is,

$$M_{ii}^{-1/2}e_i^\top(X-\mu) = \frac{X_i-\mu_i}{\sqrt{M_{ii}}} \overset{d}{=} Z$$

where Z is a random variable with the characteristic function $\varphi_Z(t) = \Psi(t^2)$. Hence, we obtain that $Q_\tau(X_i) = \mu_i + M_{ii}^{1/2}Q_\tau(Z)$.

In addition, since

$$E\Big(\exp(it[\sum_{i=1}^d e_i]^\top[X-\mu])\Big) = \Psi(t^2\sum_{i=1}^d\sum_{j=1}^d M_{ij}),$$

we obtain that

$$Q_\tau(\sum_{i=1}^d X_i) = \sum_{i=1}^d \mu_i + \Big(\sum_{i=1}^d \sum_{j=1}^d M_{ij}\Big)^{1/2} Q_\tau(Z).$$

Note that $Q_\tau(Z) < 0$ for $\tau \in (0, 0.5)$ and $Q_\tau(Z) > 0$ for $\tau \in (0.5, 1)$ since it is straightforward to see that Z is symmetric around 0. The result follows by observing that $|M_{ij}| \le M_{ii}^{1/2} M_{jj}^{1/2}$ when M is positive semi-definite. ∎

S4 | MULTI-SCANNING (M-)GOALS

As discussed in Section 2 of the main text, in practice, the optimal choice of (ϵ, δ) for (M-)GOALS may depend on the specific real data application, and different choices of (ϵ, δ) could lead to different segmentation results by (M-)GOALS. Thus, it is desirable to have a fully data-driven procedure that automatically selects a suitable (ϵ, δ) based on the observations.

To this end, we propose *multi-scanning* (M-)GOALS, which further augments (M-)GOALS with a model selection based post-processing step and automatically consolidates estimated change-points from (M-)GOALS with different trimming parameters (ϵ, δ) via minimizing a quantile regression BIC function. To conserve space, we focus on the presentation of multi-scanning M-GOALS. The procedure for multi-scanning GOALS can be derived accordingly in a straightforward manner.

The key idea of multi-scanning M-GOALS is a post-processing procedure based on a modified quantile regression BIC. Denote $\widehat{\mathbf{k}} = (\widehat{k}_1, \cdots, \widehat{k}_{\widehat{m}})$ as the estimated change-points based on $\{Y_t\}_{t=1}^n$, and further define $\widehat{k}_0 := 0$ and $\widehat{k}_{\widehat{m}+1} := n$. Given the multiple quantile levels $\boldsymbol{\tau}^M = (\tau_1, \cdots, \tau_M)$ of interest, the modified quantile regression BIC function is defined as

$$\mathrm{BIC}(\widehat{\mathbf{k}}, \boldsymbol{\tau}^M) = \sum_{j=1}^M \left[\log\left(\sum_{i=0}^{\widehat{m}} \sum_{t=\widehat{k}_i+1}^{\widehat{k}_{i+1}} \rho_{\tau_j}\left(Y_t - X_t^\top \widehat{\boldsymbol{\beta}}_{\widehat{k}_i+1, \widehat{k}_{i+1}}(\tau_j) \right) \right) + (3\widehat{m}+2)\frac{\log n}{2n} \right], \tag{S24}$$

where $\rho_\tau(u) = u(\tau - \mathbb{I}(u < 0))$ is the check loss at quantile level τ, $X_t = (1, t/n)$ is the deterministic regressor and $\widehat{\boldsymbol{\beta}}_{\widehat{k}_i+1,\widehat{k}_{i+1}}(\tau_j) = (\widehat{\beta}_{0;\widehat{k}_i+1,\widehat{k}_{i+1}}(\tau_j), \widehat{\beta}_{1;\widehat{k}_i+1,\widehat{k}_{i+1}}(\tau_j))^\top$ is the estimated linear trend parameters based on the $(i+1)$th estimated segment $\{Y_t\}_{t=\widehat{k}_i+1}^{\widehat{k}_{i+1}}$ via quantile regression at the quantile level τ_j, for $j = 1, \cdots, M$ and $i = 0, 1, \cdots, \widehat{m}$. The factor $3\widehat{m}+2$ is used to account for the number of parameters in the $\widehat{m}+1$ quantile regressions and $\widehat{m}$ change-point locations.

The quantile regression BIC is originally proposed in Lee et al. (2014) for *i.i.d.* data. The modified quantile regression BIC proposed in (S24) is adapted from equation (2.3) in Lee et al. (2014) to handle the piecewise linear quantile trend model. In the literature, another commonly used quantile regression BIC is based on the sum of check losses without the logarithm function (see Wu and Zen, 1999), which is later adopted in the quantile regression change-point literature by Oka and Qu (2011) and Aue et al. (2014). However, as argued by Lee et al. (2014), the BIC in (S24) is preferable as it is invariant to the scale of $\{Y_t\}_{t=1}^n$. We refer to Section 2.1 of Lee et al. (2014) for more discussions.

We proceed by introducing notations necessary for the presentation of multi-scanning M-GOALS. Given a significance level $\alpha \in (0,1)$, denote $\widehat{\mathbf{k}}(\epsilon, \delta)$ as the set of change-points estimated by M-GOALS with trimming parameters (ϵ, δ) and threshold $\zeta_n^M(\epsilon, \delta)$, where $\zeta_n^M(\epsilon, \delta)$ is the $(1-\alpha)\times 100\%$ quantile of the pivotal limiting distribution $\mathcal{T}^M(\epsilon, \delta)$ in Theorem 1. Denote $\mathcal{C}$ as the set of all specifications of (ϵ, δ). Note that $\mathcal{C}$ should cover a reasonably wide range of (ϵ, δ). In numerical studies, we set C as

$\{(0.06, 0.01), (0.08, 0.01), (0.08, 0.02), (0.1, 0.01), (0.1, 0.02), (0.12, 0.01), (0.12, 0.02), (0.15, 0.01), (0.15, 0.02)\}$.

As discussed above, M-GOALS based on different trimming parameters $(\epsilon, \delta) \in C$ could potentially return different estimated change-points $\widehat{\mathbf{k}}(\epsilon, \delta)$. A natural solution is to select the change-point estimate (and thus the trimming parameter) that minimizes the modified quantile regression BIC in (S24). In other words, define

$$(\epsilon^*, \delta^*) = \underset{(\epsilon,\delta)\in C}{\arg\min}\, \mathrm{BIC}(\widehat{\mathbf{k}}(\epsilon, \delta), \boldsymbol{\tau}^M), \tag{S25}$$

we set the final change-point estimator as $\widehat{\mathbf{k}}(\epsilon^*, \delta^*)$ and refer it as the best individual M-GOALS by BIC. Simulation studies in Section S5 indicate that $\widehat{\mathbf{k}}(\epsilon^*, \delta^*)$ performs well in practice.

One potential drawback of $\widehat{\mathbf{k}}(\epsilon^*, \delta^*)$ is that it treats estimation results from different trimming parameters separately and chooses a single (ϵ^*, δ^*) for M-GOALS. However, when there are multiple change-points, there could be cases where the best estimator for different true change-point is given by M-GOALS with different trimming parameters. Intuitively, if we can consolidate the estimated change-points across M-GOALS with different $(\epsilon, \delta) \in C$ in a sensible way, the final estimate can be improved. This motivates the development of multi-scanning M-GOALS.

Denote $\widehat{\mathbf{K}} = \bigcup_{(\epsilon,\delta)\in C} \widehat{\mathbf{k}}(\epsilon, \delta)$ as the collection of estimated change-points pooled from M-GOALS with different (ϵ, δ) in C. Multi-scanning M-GOALS consolidates $\widehat{\mathbf{K}}$ by searching through all subsets of $\widehat{\mathbf{K}}$ and defines the final change-point estimate as

$$\widehat{\mathbf{k}}_* = \underset{\widehat{\mathbf{k}}\subset\widehat{\mathbf{K}}}{\arg\min}\, \mathrm{BIC}(\widehat{\mathbf{k}}, \boldsymbol{\tau}^M). \tag{S26}$$

Unlike $\widehat{\mathbf{k}}(\epsilon^*, \delta^*)$, the estimated change-points in $\widehat{\mathbf{k}}_*$ can come from M-GOALS with different trimming parameters, and it is easy to see that $\mathrm{BIC}(\widehat{\mathbf{k}}_*, \boldsymbol{\tau}^M) \le \mathrm{BIC}(\widehat{\mathbf{k}}(\epsilon^*, \delta^*), \boldsymbol{\tau}^M)$. Thus, via the quantile regression BIC, multi-scanning M-GOALS implicitly integrates different change-point estimates from different (ϵ, δ). The optimization of (S26) can be done via exhaustive search, as the cardinality of $\widehat{\mathbf{K}}$ is typically small (less than 10 in the numerical studies), otherwise, a forward selection procedure can be used for heuristic optimization. In Section S5, multi-scanning M-GOALS is seen to provide the most favorable performance in the simulation studies.

Remark 2: Multi-scanning M-GOALS is a hybrid of M-GOALS and the model selection based approach. The estimated change-points by M-GOALS with different (ϵ, δ) create a promising candidate pool $\widehat{\mathbf{K}}$ of potential change-points, which is then refined by the quantile regression BIC in (S26) to form the final change-point estimate $\widehat{\mathbf{k}}_*$. Multi-scanning M-GOALS capitalizes on both steps, where the candidate pool generated by M-GOALS provides robustness to serial dependence and heteroscedasticity thanks to the use of SN, and the model selection step helps implicitly select the best trimming parameter and consolidates the change-point estimation by integrating results obtained across M-GOALS with different (ϵ, δ). Similar hybrid strategies are previously investigated in Niu and Zhang (2012) and Yau and Zhao (2016) under the context of change-point estimation in univariate mean and autoregressive models, respectively.

Remark 3: It is natural to consider a fully model selection based approach, where we estimate change-points by directly minimizing (S26) without the M-GOALS step. In other words, we set $\widehat{\mathbf{K}} = \{1, 2, \cdots, n\}$ in (S26). However, there are several drawbacks for such an approach. First, an exact optimization of (S26) is impossible as the search space is of cardinality 2^n and it is easy to see that dynamic programming cannot be used to optimize (S26). Second and more importantly, without the pre-screening from M-GOALS, the estimation can be more sensitive to unknown serial dependence and heteroscedasticity, as the modified quantile regression BIC in (S26) does not explicitly account for such effects.

S5 | SIMULATION STUDIES

In this section, we conduct numerical experiments to investigate the finite sample performance of GOALS, M-GOALS and multi-scanning M-GOALS for multiple change-point detection, and further compare with Oka and Qu (2011). Section S5.1 focuses on GOALS and M-GOALS, and Section S5.2 examines multi-scanning M-GOALS and further conducts a sensitivity analysis.

We consider the following three data-generating processes (DGPs) with $n = 210$, which is the average length of the COVID-19 infection curves across the 35 countries used in real data analysis.

$$\text{DGP1:} \quad Y_t = \begin{cases} 5.8 + 26(t/n) + \varphi_t, & 1 \le t \le \lfloor 0.1 * n \rfloor, \\ 8.4 + 2(t/n - 0.1) + \varphi_t, & \lfloor 0.1 * n \rfloor + 1 \le t \le \lfloor 0.3 * n \rfloor, \\ 8.8 - 7(t/n - 0.3) + \varphi_t, & \lfloor 0.3 * n \rfloor + 1 \le t \le \lfloor 0.55 * n \rfloor, \\ 7.05 + 4(t/n - 0.55) + \varphi_t, & \lfloor 0.55 * n \rfloor + 1 \le t \le \lfloor 0.8 * n \rfloor, \\ 8.05 + 11(t/n - 0.8) + \varphi_t, & \lfloor 0.8 * n \rfloor + 1 \le t \le n. \end{cases}$$

$$\text{DGP2:} \quad Y_t = \begin{cases} 5.5 + 45(t/n) + \varphi_t, & 1 \le t \le \lfloor 0.1 * n \rfloor, \\ 10 - 1.3(t/n - 0.1) + \varphi_t, & \lfloor 0.1 * n \rfloor + 1 \le t \le \lfloor 0.4 * n \rfloor, \\ 9.61 + 7.5(t/n - 0.4) + \varphi_t, & \lfloor 0.4 * n \rfloor + 1 \le t \le \lfloor 0.6 * n \rfloor, \\ 11.11 - 4(t/n - 0.6) + \varphi_t, & \lfloor 0.6 * n \rfloor + 1 \le t \le \lfloor 0.8 * n \rfloor, \\ 10.31 + 6(t/n - 0.8) + \varphi_t, & \lfloor 0.8 * n \rfloor + 1 \le t \le n. \end{cases}$$

$$\text{DGP3:} \quad Y_t = \begin{cases} 6 - 23(t/n) + \varphi_t, & 1 \le t \le \lfloor 0.15 * n \rfloor, \\ 2.55 - 3(t/n - 0.1) + \varphi_t, & \lfloor 0.1 * n \rfloor + 1 \le t \le \lfloor 0.35 * n \rfloor, \\ 1.8 + 19(t/n - 0.35) + \varphi_t, & \lfloor 0.35 * n \rfloor + 1 \le t \le \lfloor 0.6 * n \rfloor, \\ 6.55 - 14(t/n - 0.6) + \varphi_t, & \lfloor 0.6 * n \rfloor + 1 \le t \le \lfloor 0.8 * n \rfloor, \\ 3.75 - 3(t/n - 0.8) + \varphi_t, & \lfloor 0.8 * n \rfloor + 1 \le t \le n \end{cases}$$

The error process $\{\varphi_t\}$ takes the form $\varphi_t = [1 + \gamma_1(t/n)]\varepsilon_t$ and ε_t is generated via an AR(1) process where $\varepsilon_t = \rho\varepsilon_{t-1} + e_t$, $e_t \overset{i.i.d.}{\sim} \mathcal{N}(0, (1-\rho^2)\sigma^2)$. The DGP 1-3 are designed to mimic the trajectory of daily coronavirus new cases (in log scale) in the U.K, the U.S. and Australia respectively, see Figure S1 for typical realizations of the three DGPs with $\rho = 0.3$, $\gamma_1 = 0.3$ and $\sigma = 0.2$.

We vary ρ to control the temporal dependence of the error process and vary σ to control the variance of the error process. Furthermore, we set $\gamma_1 = 0$ to generate homogeneous errors and set $\gamma_1 = 0.3$ to generate heteroscedastic errors. For each simulation setting considered in the following, we repeat the experiments 500 times.

For comparison, we consider the multiple change-point estimation procedure for regression quantiles proposed in Oka and Qu (2011). The estimation method proposed in Aue et al. (2014) shares similar ideas with Oka and Qu (2011) by minimizing an MDL-based objective function but can only perform change-point detection at a single quantile level. Thus, we focus on the comparison with Oka and Qu (2011) (hereafter, OQ).

To assess the accuracy of change-point estimation, we use the Hausdorff distance between two sets. Denote the

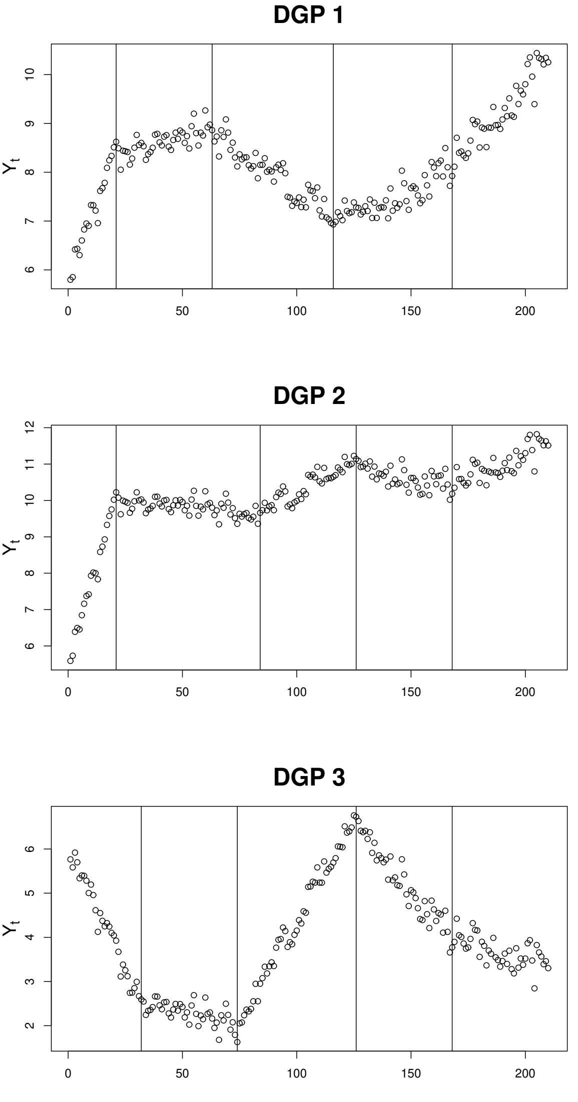

FIGURE S1 Typical realizations of DGP 1-3 with $\rho = 0.3$, $\gamma_1 = 0.3$ and $\sigma = 0.2$. Vertical lines mark the true change-points.

set of true change-points as $\mathbf{q}_0$ and the set of estimated change-points as $\widehat{\mathbf{q}}$, we define $d_1(\mathbf{q}_0, \widehat{\mathbf{q}}) = \max_{q_1 \in \widehat{\mathbf{q}}} \min_{q_2 \in \mathbf{q}_0} |q_1 - q_2|$ and $d_2(\mathbf{q}_0, \widehat{\mathbf{q}}) = \max_{q_1 \in \mathbf{q}_0} \min_{q_2 \in \widehat{\mathbf{q}}} |q_1 - q_2|$, where d_1 measures the over-segmentation error while d_2 measures the under-segmentation error of $\widehat{\mathbf{q}}$. In addition, we report the adjusted Rand index (ARI) which measures the similarity between two partitions of the same observations. Higher ARI (with the maximum value of 1) corresponds to more accurate change-point estimation. For the definition and detailed discussions of ARI, we refer to Hubert and Arabie (1985).

S5.1 | GOALS and M-GOALS

In this section, we conduct extensive numerical experiments to examine the performance of GOALS and M-GOALS across various simulation settings of DGP 1-3. Specifically, we vary $\rho = 0, \pm 0.3, \pm 0.5$ to generate a wide range of temporal dependence and vary $\sigma = 0.1, 0.2$ to generate low and high volatility. We set $\gamma_1 = 0$ to generate homogeneous errors and set $\gamma_1 = 0.3$ to generate heteroscedastic errors.

For each simulated time series $\{Y_t\}_{t=1}^n$, GOALS and OQ are employed for change-point detection at a single quantile level τ with $\tau = 0.5$ or 0.9, and M-GOALS and OQ are employed for change-point detection across multiple quantile levels with $\boldsymbol{\tau}^M = (0.1, 0.5, 0.9)$. In this subsection, we fix the trimming parameters for GOALS and M-GOALS at $(\epsilon, \delta) = (0.1, 0.02)$, which is the same as the one used in real data analysis in Section 4 of the main text. Both (M-)GOALS and OQ require a significance level α in their implementation, and we set $\alpha = 0.1$ for both methods, same as in Section 4 of the main text.

Table S1 (homogeneous case) and Table S2 (heteroscedastic case) summarize the performance of GOALS and OQ for change-point detection at a single quantile $\tau = 0.5$ or 0.9, where we report the average ARI, the number of estimated change-point $\widehat{m}$, and Hausdorff distances (d_1, d_2) across the 500 experiments. It can be seen that both methods deliver accurate change-point estimation for the quantile level $\tau = 0.5$. In particular, GOALS tends to have better ARI and smaller d_2, while OQ tends to have smaller d_1. However, for the quantile level $\tau = 0.9$, OQ severely underestimates the number of change-points in DGP1 and DGP2, especially under high variance ($\sigma = 0.2$). In comparison, GOALS remains effective under all settings. This suggests that GOALS is more robust when applied to detecting changes in extreme quantiles.

Table S3 summarizes the performance of M-GOALS and OQ for change-point detection across multiple quantile levels $\boldsymbol{\tau}^M = (0.1, 0.5, 0.9)$. Compared with the results based on a single quantile level, it can be seen that the performance of both methods improves when multiple quantile levels are used simultaneously. The two methods have similar performance under lower variance ($\sigma = 0.1$) and homogeneous ($\gamma_1 = 0$) errors. However, M-GOALS performs more favorably under high volatility ($\sigma = 0.2$) and heterogeneous ($\gamma_1 = 0.3$) errors, especially in terms of ARI.

For both change-point estimation at a single quantile level and across multiple quantile levels, we find that increasing ρ (temporal dependence) or σ (variance) negatively affects the accuracy of change-point estimation for both methods. In addition, estimation is more accurate in the homogeneous case ($\gamma_1 = 0$) than in heterogeneous case ($\gamma_1 = 0.3$).

TABLE S1 Change-point estimation at a single quantile level $\tau = 0.5$ or 0.9 (homogeneous case)

		GOALS						OQ					
		$\tau = 0.5$			$\tau = 0.9$			$\tau = 0.5$			$\tau = 0.9$		
σ	ρ	ARI	$\widehat{m}$	d_1/d_2(%)	ARI	$\widehat{m}$	d_1/d_2(%)	ARI	$\widehat{m}$	d_1/d_2(%)	ARI	$\widehat{m}$	d_1/d_2(%)
DGP 1													
	-0.5	0.913	4.00	1.72 /1.74	0.876	4.02	2.60 /2.67	0.898	4.00	1.90 /1.90	0.731	3.24	12.02 /3.63
	-0.3	0.911	4.00	1.80 /1.80	0.874	4.01	2.67 /2.62	0.895	4.00	1.99 /1.99	0.723	3.24	11.91 /3.82
0.1	0	0.902	4.00	1.99 /2.01	0.870	4.02	2.76 /2.79	0.887	4.00	2.21 /2.22	0.712	3.20	12.00 /3.98
	0.3	0.890	4.01	2.34 /2.38	0.863	4.03	2.93 /3.05	0.873	4.02	2.55 /2.64	0.701	3.16	12.94 /4.15
	0.5	0.875	4.01	2.78 /2.73	0.849	4.04	3.34 /3.46	0.856	4.10	2.87 /3.31	0.683	3.14	13.44 /4.50
	-0.5	0.861	4.01	2.89 /2.85	0.784	3.88	7.06 /4.23	0.821	3.96	3.84 /3.34	0.472	2.14	24.41 /6.29
	-0.3	0.851	4.00	3.39 /3.11	0.775	3.90	7.11 /4.45	0.815	3.98	3.94 /3.60	0.471	2.13	25.20 /6.10
0.2	0	0.837	3.99	3.95 /3.37	0.755	3.83	8.72 /4.59	0.796	3.97	4.45 /4.04	0.483	2.18	23.31 /6.09
	0.3	0.797	3.93	6.18 /4.11	0.726	3.73	10.07 /4.93	0.772	3.99	5.14 /4.74	0.483	2.20	23.24 /6.11
	0.5	0.759	3.86	8.08 /4.82	0.708	3.71	10.91 /5.31	0.746	4.04	5.78 /5.44	0.499	2.30	23.04 /6.28
DGP 2													
	-0.5	0.930	4.00	1.42 /1.47	0.899	4.03	1.98 /2.35	0.915	4.00	1.71 /1.71	0.582	2.54	26.03 /4.14
	-0.3	0.927	4.01	1.48 /1.57	0.897	4.03	2.04 /2.33	0.913	4.00	1.76 /1.76	0.569	2.49	26.75 /4.37
0.1	0	0.920	4.01	1.71 /1.85	0.894	4.03	2.14 /2.44	0.907	4.00	1.95 /1.95	0.564	2.50	26.62 /4.49
	0.3	0.906	4.02	2.01 /2.20	0.887	4.04	2.40 /2.70	0.891	4.03	2.29 /2.43	0.567	2.51	25.95 /4.58
	0.5	0.892	4.05	2.28 /2.78	0.877	4.03	2.71 /2.94	0.878	4.08	2.58 /3.02	0.566	2.54	25.65 /4.65
	-0.5	0.887	4.01	2.51 /2.53	0.809	3.89	6.29 /3.99	0.830	3.90	4.18 /3.44	0.280	1.40	42.35 /4.32
	-0.3	0.881	4.01	2.78 /2.75	0.797	3.82	7.32 /4.05	0.827	3.90	4.40 /3.54	0.278	1.37	41.83 /4.45
0.2	0	0.863	3.98	3.67 /3.11	0.787	3.79	7.73 /4.00	0.821	3.95	4.11 /3.81	0.290	1.42	38.95 /4.97
	0.3	0.821	3.91	5.78 /3.90	0.756	3.68	9.87 /4.24	0.794	3.94	5.09 /4.60	0.302	1.47	36.26 /5.22
	0.5	0.776	3.80	8.28 /4.42	0.734	3.66	10.60 /4.72	0.779	3.99	5.27 /5.18	0.312	1.53	35.06 /5.46
DGP 3													
	-0.5	0.936	4.00	1.21 /1.21	0.914	4.00	1.56 /1.57	0.935	4.00	1.27 /1.30	0.877	3.80	4.64 /1.93
	-0.3	0.933	4.00	1.26 /1.26	0.910	4.00	1.65 /1.65	0.933	4.00	1.33 /1.33	0.880	3.84	4.19 /2.00
0.1	0	0.932	4.00	1.32 /1.32	0.905	4.00	1.80 /1.81	0.931	4.00	1.38 /1.40	0.866	3.82	4.84 /2.21
	0.3	0.923	4.00	1.57 /1.53	0.901	4.00	1.91 /1.94	0.921	4.05	1.55 /1.80	0.861	3.86	4.90 /2.43
	0.5	0.914	4.00	1.68 /1.70	0.898	4.00	2.01 /2.01	0.910	4.11	1.80 /2.36	0.871	3.89	4.17 /2.38
	-0.5	0.905	4.00	1.94 /1.91	0.861	3.93	4.00 /2.54	0.888	4.00	2.22 /2.22	0.774	3.51	9.47 /3.06
	-0.3	0.903	3.99	2.18 /1.91	0.853	3.90	4.48 /2.64	0.884	4.00	2.35 /2.35	0.768	3.49	9.76 /3.10
0.2	0	0.892	3.97	2.68 /2.16	0.842	3.87	5.30 /2.75	0.874	4.00	2.56 /2.58	0.758	3.46	9.92 /3.28
	0.3	0.865	3.91	4.28 /2.51	0.826	3.81	6.33 /2.92	0.856	4.02	3.01 /3.05	0.747	3.47	10.10 /3.59
	0.5	0.839	3.87	5.61 /3.00	0.820	3.85	6.18 /3.27	0.843	4.07	3.37 /3.63	0.737	3.52	10.49 /3.99

TABLE S2 Change-point estimation at a single quantile level $\tau = 0.5$ or 0.9 (heteroscedastic case)

		GOALS						OQ					
		$\tau = 0.5$			$\tau = 0.9$			$\tau = 0.5$			$\tau = 0.9$		
σ	ρ	ARI	$\widehat{m}$	d_1/d_2(%)	ARI	$\widehat{m}$	d_1/d_2(%)	ARI	$\widehat{m}$	d_1/d_2(%)	ARI	$\widehat{m}$	d_1/d_2(%)
						DGP 1							
0.1	-0.5	0.904	4.00	1.95/1.97	0.905	4.00	1.83/1.81	0.885	4.00	2.21/2.21	0.853	3.74	13.01 /4.25
	-0.3	0.900	4.00	2.06/2.06	0.903	4.00	1.86/1.86	0.881	4.00	2.31/2.31	0.846	3.72	12.15 /4.17
	0	0.892	4.00	2.25/2.26	0.898	4.00	2.06/1.95	0.871	4.00	2.57/2.57	0.845	3.76	13.18 /4.57
	0.3	0.877	4.00	2.71/2.70	0.892	4.00	2.26/2.20	0.858	4.02	2.91/2.99	0.846	3.79	13.74 /4.67
	0.5	0.859	4.01	3.23/3.13	0.885	4.01	2.40/2.35	0.837	4.11	3.31/3.70	0.843	3.88	13.83 /4.92
0.2	-0.5	0.841	3.97	4.02/3.24	0.832	3.81	6.30/2.74	0.770	3.78	7.02/4.01	0.737	3.30	26.09 /6.10
	-0.3	0.830	3.96	4.48/3.43	0.820	3.77	7.13/2.82	0.763	3.83	6.62/4.37	0.739	3.36	27.00 /6.08
	0	0.809	3.92	5.77/3.74	0.810	3.74	7.88/3.00	0.753	3.85	6.83/4.72	0.718	3.25	26.44 /6.03
	0.3	0.764	3.82	8.53/4.47	0.793	3.67	8.84/3.14	0.728	3.85	7.82/5.37	0.702	3.28	26.74 /6.27
	0.5	0.721	3.70	10.63/5.14	0.788	3.71	8.73/3.53	0.705	3.89	8.35/6.10	0.715	3.33	26.47 /6.07
						DGP 2							
0.1	-0.5	0.923	4.01	1.61/1.67	0.887	4.03	2.36/2.71	0.903	4.00	1.98/1.98	0.532	2.35	29.20 /4.46
	-0.3	0.918	4.00	1.77/1.79	0.886	4.04	2.35/2.72	0.902	4.00	2.01/2.01	0.534	2.36	28.56 /4.69
	0	0.913	4.02	1.88/2.04	0.880	4.02	2.64/2.75	0.894	4.00	2.24/2.24	0.528	2.34	28.32 /4.66
	0.3	0.893	4.02	2.41/2.57	0.872	4.04	2.92/3.14	0.879	4.01	2.63/2.70	0.517	2.28	29.00 /4.77
	0.5	0.879	4.05	2.68/3.08	0.856	4.04	3.41/3.55	0.862	4.08	2.93/3.34	0.502	2.27	28.72 /4.81
0.2	-0.5	0.864	3.95	3.88/2.85	0.751	3.61	10.63/4.07	0.747	3.55	9.52/4.48	0.180	0.95	46.97 /3.84
	-0.3	0.849	3.92	4.87/3.21	0.738	3.57	11.55/4.21	0.768	3.69	7.24/4.62	0.187	0.96	48.43 /3.65
	0	0.823	3.85	6.37/3.52	0.710	3.44	13.38/4.14	0.754	3.70	7.90/4.87	0.190	0.98	45.00 /4.23
	0.3	0.770	3.69	9.42/4.08	0.680	3.35	15.35/4.47	0.735	3.67	8.59/5.46	0.215	1.05	41.06 /4.88
	0.5	0.726	3.56	12.03/4.52	0.673	3.38	14.84/4.89	0.727	3.80	8.10/6.05	0.248	1.21	35.42 /5.73
						DGP 3							
0.1	-0.5	0.931	4.00	1.31/1.31	0.861	4.02	3.17/3.10	0.928	4.00	1.44/1.44	0.674	3.07	5.88 /2.23
	-0.3	0.929	4.00	1.35/1.35	0.860	4.02	3.15/3.05	0.925	4.00	1.49/1.51	0.687	3.15	6.30 /2.34
	0	0.925	4.00	1.48/1.46	0.854	4.02	3.27/3.24	0.921	4.00	1.64/1.64	0.663	3.06	5.93 /2.50
	0.3	0.913	3.99	1.89/1.76	0.842	4.04	3.77/3.57	0.910	4.03	1.92/2.06	0.654	3.04	5.80 /2.57
	0.5	0.905	4.00	2.01/1.93	0.827	4.05	4.32/4.08	0.902	4.08	2.03/2.46	0.648	3.03	5.44 /2.78
0.2	-0.5	0.887	3.95	3.18/2.07	0.748	3.75	9.50/4.46	0.873	4.00	2.56/2.58	0.455	2.01	11.60 /3.26
	-0.3	0.880	3.93	3.60/2.27	0.731	3.70	10.33/4.70	0.870	4.00	2.69/2.65	0.445	1.93	11.59 /3.29
	0	0.862	3.87	4.90/2.38	0.707	3.63	12.13/4.86	0.857	4.01	3.08/3.06	0.453	1.96	12.46 /3.60
	0.3	0.830	3.76	7.12/2.65	0.687	3.56	12.99/5.29	0.840	4.02	3.48/3.51	0.450	1.99	12.61 /4.00
	0.5	0.809	3.74	7.99/3.14	0.673	3.56	13.07/5.58	0.823	4.06	4.01/4.09	0.465	2.05	12.21 /3.95

TABLE S3 Change-point estimation across multiple quantile levels $\boldsymbol{\tau}^M = (0.1, 0.5, 0.9)$

		Homogeneous case						Heterogeneous case					
		M-GOALS			OQ			M-GOALS			OQ		
σ	ρ	ARI	$\widehat{m}$	d_1/d_2(%)	ARI	$\widehat{m}$	d_1/d_2(%)	ARI	$\widehat{m}$	d_1/d_2(%)	ARI	$\widehat{m}$	d_1/d_2(%)
							DGP 1						
	-0.5	0.924	4.00	1.47/1.49	0.912	4.00	1.66/1.66	0.917	4.00	1.64/1.66	0.902	4.00	1.86/1.86
	-0.3	0.919	4.00	1.61/1.64	0.905	4.00	1.82/1.82	0.909	4.01	1.84/1.92	0.894	4.00	2.06/2.06
	0	0.904	4.01	2.06/2.04	0.890	4.00	2.19/2.19	0.894	4.01	2.26/2.30	0.877	4.00	2.48/2.49
	0.3	0.876	4.02	2.85/2.76	0.875	4.02	2.53/2.65	0.857	4.00	3.66/3.18	0.860	4.03	2.89/3.04
	0.5	0.847	4.05	3.94/3.78	0.858	4.10	2.84/3.29	0.826	4.06	4.81/4.21	0.839	4.11	3.33/3.74
	-0.5	0.875	4.02	2.52/2.64	0.848	3.98	3.16/2.94	0.858	4.02	3.25/3.05	0.792	3.79	6.58/3.55
	-0.3	0.861	4.02	3.01/3.04	0.835	3.98	3.52/3.23	0.840	4.01	4.33/3.54	0.785	3.82	6.40/3.86
	0	0.828	3.98	4.72/3.69	0.808	3.97	4.33/3.92	0.795	3.91	7.06/4.07	0.764	3.85	6.74/4.56
	0.3	0.783	3.97	6.94/4.67	0.778	4.00	4.96/4.63	0.749	3.88	9.04/5.19	0.733	3.87	7.46/5.25
	0.5	0.738	3.98	8.54/5.74	0.749	4.06	5.59/5.43	0.706	3.94	10.11/6.28	0.711	3.92	7.79/5.98
							DGP 2						
	-0.5	0.937	4.04	1.18/1.62	0.924	4.00	1.53/1.53	0.931	4.05	1.29/1.85	0.915	4.00	1.73/1.73
	-0.3	0.935	4.04	1.25/1.66	0.918	4.00	1.68/1.70	0.928	4.04	1.41/1.86	0.909	4.00	1.89/1.89
	0	0.921	4.05	1.54/2.12	0.909	4.00	1.88/1.88	0.914	4.05	1.72/2.27	0.897	4.00	2.20/2.20
	0.3	0.902	4.03	2.19/2.44	0.894	4.02	2.26/2.40	0.892	4.04	2.42/2.70	0.879	4.01	2.65/2.73
	0.5	0.879	4.04	2.97/3.03	0.878	4.09	2.56/3.02	0.864	4.01	3.74/3.32	0.862	4.07	2.98/3.32
	-0.5	0.905	4.06	1.91/2.54	0.851	3.91	3.74/3.04	0.895	4.06	2.20/2.84	0.767	3.56	8.88/4.14
	-0.3	0.897	4.07	2.09/2.72	0.842	3.91	4.03/3.24	0.883	4.08	2.50/3.17	0.783	3.70	7.05/4.35
	0	0.872	4.04	3.04/3.22	0.830	3.95	3.96/3.71	0.850	4.03	3.94/3.75	0.762	3.70	7.80/4.70
	0.3	0.821	4.06	4.98/4.63	0.797	3.95	5.04/4.53	0.795	4.01	6.28/5.09	0.739	3.69	8.32/5.35
	0.5	0.776	4.00	6.83/5.46	0.778	4.00	5.25/5.21	0.748	3.92	8.33/5.71	0.725	3.80	8.18/6.16
							DGP3						
	-0.5	0.939	4.01	1.12/1.25	0.940	4.00	1.17/1.19	0.935	4.01	1.21/1.31	0.934	4.00	1.30/1.32
	-0.3	0.936	4.00	1.20/1.24	0.936	4.00	1.25/1.25	0.933	4.01	1.26/1.32	0.929	4.00	1.42/1.44
	0	0.928	4.01	1.36/1.44	0.931	4.00	1.39/1.41	0.923	4.01	1.44/1.56	0.922	4.00	1.60/1.60
	0.3	0.917	4.00	1.70/1.67	0.922	4.04	1.58/1.83	0.908	3.99	2.15/1.88	0.911	4.03	1.86/2.02
	0.5	0.908	4.00	2.08/1.94	0.912	4.11	1.75/2.36	0.893	3.98	2.85/2.25	0.899	4.12	2.06/2.63
	-0.5	0.916	4.00	1.63/1.65	0.902	4.00	1.90/1.90	0.910	4.00	1.76/1.78	0.888	4.00	2.23/2.25
	-0.3	0.913	4.00	1.75/1.79	0.892	4.00	2.18/2.18	0.901	4.01	2.08/2.10	0.878	4.00	2.55/2.50
	0	0.896	4.00	2.20/2.15	0.875	4.00	2.54/2.55	0.881	3.98	3.01/2.41	0.861	4.01	2.93/2.96
	0.3	0.866	4.00	3.36/2.90	0.857	4.02	2.97/3.11	0.852	3.96	4.14/3.05	0.841	4.01	3.48/3.50
	0.5	0.846	3.98	4.42/3.26	0.843	4.08	3.32/3.65	0.824	3.93	5.52/3.58	0.823	4.07	3.97/4.10

S5.2 | Multi-scanning M-GOALS and sensitivity analysis

In this section, we examine the finite sample performance of the multi-scanning M-GOALS proposed in Section S4, which is a hybrid of M-GOALS and the modified quantile regression BIC. We consider change-point estimation for DGP 1-3 and vary $\sigma = 0.1, 0.2$ to generate low and high volatility. We set $\gamma_1 = 0$ to generate homogeneous errors and set $\gamma_1 = 0.3$ to generate heteroscedastic errors. To conserve space, we present the result with temporal dependence $\rho = 0.3$, which is most relevant to real data analysis of COVID-19 infection curves. Results for $\rho = 0, -0.3, \pm 0.5$ are similar and thus omitted.

For each simulated time series $\{Y_t\}_{t=1}^n$, we consider change-point detection across multiple quantile levels with $\tau^M = (0.1, 0.5, 0.9)$ using four different methods: OQ, M-GOALS, best individual M-GOALS by BIC (BI-GOALS) in (S25) and multi-scanning M-GOALS (MS-GOALS) in (S26).

For OQ and M-GOALS, their implementation is the same as in Section S5.1. Recall for M-GOALS, we fix its trimming parameter at $(\epsilon, \delta) = (0.1, 0.02)$. For BI-GOALS and MS-GOALS, we fix the set C as

$$\{(0.06, 0.01), (0.08, 0.01), (0.08, 0.02), (0.1, 0.01), (0.1, 0.02), (0.12, 0.01), (0.12, 0.02), (0.15, 0.01), (0.15, 0.02)\},$$

which covers a wide range of trimming parameters (ϵ, δ). All four methods require a significance level α in their implementation, and we vary $\alpha = 0.1, 0.05$ to examine the sensitivity of the estimation results w.r.t. the significance level.

Table S4 summarizes the performance of the four methods, where we report the average ARI, the number of estimated change-point $\widehat{m}$, and Hausdorff distances (d_1, d_2) across the 500 experiments. As can be seen clearly, across all simulation settings, BI-GOALS and MS-GOALS deliver more favorable performance than M-GOALS, which suggests the benefit of combining estimation across M-GOALS with different trimming parameters, compared to using a single trimming parameter. The improvement is more notable under the case of low signal-to-noise ratio (SNR), where the error exhibits high variance ($\sigma = 0.2$) and heterogeneity ($\gamma_1 = 0.3$). Overall, MS-GOALS gives the best performance, BI-GOALS comes second, and then M-GOALS and OQ, though the performance difference is minimal under the high SNR scenario with low variance and homogeneous error. As for the sensitivity w.r.t. the significance level α, all methods are quite robust to the choice of $\alpha = 0.1$ and 0.05, though M-GOALS and OQ seem to be slightly more sensitive compared to MS-GOALS.

Table S5 further reports the performance of M-GOALS with different trimming parameters $(\epsilon, \delta) \in C$, together with the performance of MS-GOALS. To conserve space, we only report the result for the significance level $\alpha = 0.1$, as the result for $\alpha = 0.05$ is essentially the same. Note that for DGP1 and DGP2, the minimum spacing between change-points is 0.1 and for DGP3, it is 0.15. Thus, by Theorem 1, we know that M-GOALS with $\epsilon > 0.1$ may not work well for DGPs 1 and 2.

As can be seen from Table S5, under the case of high SNR with low variance and homogeneous error, M-GOALS delivers accurate and robust performance across $\epsilon \in \{0.08, 0.10, 0.12\}$, and its performance is further robust to the local trimming parameter $\delta \in \{0.01, 0.02\}$. However, visible variations of the performance can be seen under weak SNR with high variance and heterogeneous error. M-GOALS with $\epsilon = 0.06$ suffers from over-estimation (false positives) as the estimation error of the subsample SN is difficult to control for very small ϵ with the moderate sample size $n = 210$, while M-GOALS with $\epsilon = 0.15$ suffers from under-estimation for DGP 1 and 2 as ϵ exceeds the minimum spacing 0.1 by a large margin. Across all simulation settings, M-GOALS with $(\epsilon, \delta) = (0.1, 0.02)$ and $(0.1, 0.01)$ seem to provide the best performance, which is then further improved by multi-scanning M-GOALS, which combines estimation results given by different M-GOALS with $(\epsilon, \delta) \in C$.

TABLE S4 Change-point estimation across multiple quantile levels $\tau^M = (0.1, 0.5, 0.9)$. BI-GOALS stands for best individual M-GOALS by BIC (i.e. $\widehat{\mathbf{k}}(\epsilon^*, \delta^*)$ in (S25)). MS-GOALS stands for multi-scanning M-GOALS (i.e. $\widehat{\mathbf{k}}_*$ in (S26)).

		Homogeneous case						Heterogeneous case					
		$\sigma = 0.1$			$\sigma = 0.2$			$\sigma = 0.1$			$\sigma = 0.2$		
α	Model	ARI	$\widehat{m}$	$d_1/d_2(\%)$	ARI	$\widehat{m}$	$d_1/d_2(\%)$	ARI	$\widehat{m}$	$d_1/d_2(\%)$	ARI	$\widehat{m}$	$d_1/d_2(\%)$
		DGP 1											
0.1	OQ	0.875	4.02	2.53/2.65	0.778	4.00	4.96/4.63	0.860	4.03	2.89/3.04	0.733	3.87	7.46/5.25
	M-GOALS	0.876	4.02	2.85/2.76	0.783	3.97	6.94/4.67	0.857	4.00	3.66/3.18	0.749	3.88	9.04/5.19
	BI-GOALS	0.888	4.08	2.14/2.70	0.803	4.19	3.90/4.71	0.872	4.10	2.45/3.08	0.773	4.21	5.13/5.55
	MS-GOALS	0.889	4.05	2.23/2.47	0.812	4.06	3.92/4.15	0.874	4.05	2.58/2.85	0.791	4.05	4.80/4.67
0.05	OQ	0.876	4.01	2.53/2.59	0.767	3.92	5.74/4.69	0.863	4.01	2.88/2.92	0.714	3.73	8.93/5.33
	M-GOALS	0.878	4.02	2.72/2.73	0.771	3.89	8.01/4.62	0.858	4.02	3.53/3.21	0.732	3.75	11.04/4.96
	BI-GOALS	0.888	4.08	2.13/2.69	0.800	4.20	4.05/4.72	0.872	4.10	2.46/3.08	0.773	4.21	5.13/5.55
	MS-GOALS	0.889	4.05	2.23/2.46	0.811	4.06	4.00/4.21	0.874	4.05	2.57/2.85	0.791	4.05	4.80/4.67
		DGP 2											
0.1	OQ	0.894	4.02	2.26/2.40	0.797	3.95	5.04/4.53	0.879	4.01	2.65/2.73	0.739	3.69	8.32/5.35
	M-GOALS	0.902	4.03	2.19/2.44	0.821	4.06	4.98/4.63	0.892	4.04	2.42/2.70	0.795	4.01	6.28/5.09
	BI-GOALS	0.908	4.08	1.77/2.40	0.839	4.20	3.22/4.48	0.898	4.08	2.00/2.68	0.815	4.23	3.85/5.19
	MS-GOALS	0.909	4.04	1.92/2.22	0.849	4.05	3.39/3.62	0.899	4.04	2.16/2.41	0.833	4.05	3.84/4.01
0.05	OQ	0.896	4.00	2.26/2.26	0.779	3.82	6.58/4.64	0.879	4.00	2.68/2.67	0.683	3.40	13.01/5.45
	M-GOALS	0.905	4.05	2.06/2.43	0.819	4.02	5.28/4.42	0.893	4.04	2.35/2.67	0.781	3.89	7.64/4.78
	BI-GOALS	0.908	4.08	1.77/2.40	0.838	4.19	3.25/4.51	0.898	4.08	2.01/2.69	0.810	4.25	4.01/5.36
	MS-GOALS	0.909	4.04	1.92/2.23	0.848	4.06	3.39/3.67	0.899	4.04	2.17/2.41	0.831	4.05	3.86/4.03
		DGP 3											
0.1	OQ	0.922	4.04	1.58/1.83	0.857	4.02	2.97/3.11	0.911	4.03	1.86/2.02	0.841	4.01	3.48/3.50
	M-GOALS	0.917	4.00	1.70/1.67	0.866	4.00	3.36/2.90	0.908	3.99	2.15/1.88	0.852	3.96	4.14/3.05
	BI-GOALS	0.929	4.02	1.39/1.53	0.882	4.04	2.38/2.60	0.920	4.02	1.59/1.78	0.869	4.05	2.65/2.89
	MS-GOALS	0.928	4.02	1.50/1.59	0.879	4.03	2.52/2.67	0.920	4.02	1.68/1.80	0.870	4.02	2.71/2.83
0.05	OQ	0.924	4.01	1.57/1.64	0.859	4.01	2.96/3.00	0.913	4.01	1.87/1.93	0.841	4.00	3.54/3.44
	M-GOALS	0.918	4.01	1.67/1.68	0.868	3.98	3.45/2.73	0.910	4.00	1.98/1.88	0.848	3.92	4.67/2.92
	BI-GOALS	0.929	4.02	1.39/1.52	0.882	4.04	2.38/2.59	0.920	4.02	1.58/1.78	0.868	4.05	2.70/2.91
	MS-GOALS	0.928	4.02	1.50/1.59	0.879	4.03	2.52/2.67	0.920	4.02	1.68/1.80	0.869	4.02	2.78/2.86

TABLE S5 Change-point estimation across multiple quantile levels $\boldsymbol{\tau}^M = (0.1, 0.5, 0.9)$ by M-GOALS with different trimming parameters $(\epsilon, \delta) \in \mathcal{C}$ and by MS-GOALS.

	Homogeneous case						Heterogeneous case					
	$\sigma = 0.1$			$\sigma = 0.2$			$\sigma = 0.1$			$\sigma = 0.2$		
(ϵ, δ)	ARI	$\widehat{m}$	$d_1/d_2(\%)$	ARI	$\widehat{m}$	$d_1/d_2(\%)$	ARI	$\widehat{m}$	$d_1/d_2(\%)$	ARI	$\widehat{m}$	$d_1/d_2(\%)$
DGP 1												
MS-GOALS	0.889	4.05	2.23/2.47	0.812	4.06	3.92/4.15	0.874	4.05	2.58/2.85	0.791	4.05	4.80/4.67
(0.06,0.01)	0.739	5.86	2.70/10.24	0.652	5.86	6.31/10.12	0.726	5.84	3.39/10.11	0.633	5.81	7.30/10.13
(0.08,0.01)	0.844	4.41	2.64/5.63	0.732	4.17	8.69/6.35	0.828	4.35	3.43/5.51	0.695	4.02	11.36/6.57
(0.08,0.02)	0.794	4.77	3.33/7.49	0.696	5.06	5.92/9.00	0.774	4.84	3.83/7.88	0.678	5.09	6.47/9.21
(0.10,0.01)	0.868	3.99	3.74/3.01	0.747	3.76	10.03/4.83	0.845	3.96	4.78/3.39	0.715	3.70	11.78/5.32
(0.10,0.02)	0.876	4.02	2.85/2.76	0.783	3.97	6.94/4.67	0.857	4.00	3.66/3.18	0.749	3.88	9.04/5.19
(0.12,0.01)	0.782	3.41	12.04/3.72	0.685	3.36	13.67/4.40	0.758	3.40	12.00/4.04	0.657	3.29	14.81/4.73
(0.12,0.02)	0.821	3.71	7.72/3.23	0.737	3.65	9.63/4.19	0.807	3.73	7.48/3.43	0.704	3.52	12.06/4.37
(0.15,0.01)	0.727	2.88	18.69/3.62	0.589	2.52	19.80/5.57	0.687	2.75	19.41/4.02	0.559	2.41	20.27/5.89
(0.15,0.02)	0.761	2.98	18.75/3.05	0.646	2.75	18.85/4.86	0.735	2.92	19.00/3.38	0.617	2.64	19.38/5.15
DGP 2												
MS-GOALS	0.909	4.04	1.92/2.22	0.849	4.05	3.39/3.62	0.899	4.04	2.16/2.41	0.833	4.05	3.84/4.01
(0.06,0.01)	0.722	5.98	2.44/10.85	0.638	6.16	5.41/11.89	0.706	6.04	2.85/10.87	0.617	6.00	6.76/11.83
(0.08,0.01)	0.835	4.37	2.77/6.03	0.728	4.48	6.52/9.36	0.814	4.41	3.26/6.69	0.694	4.37	8.82/9.58
(0.08,0.02)	0.811	4.62	2.99/7.33	0.706	5.06	5.40/9.95	0.791	4.69	3.48/7.86	0.690	5.06	6.06/10.05
(0.10,0.01)	0.909	4.09	2.01/2.78	0.800	3.97	6.67/4.71	0.895	4.07	2.65/2.95	0.770	3.89	8.21/5.14
(0.10,0.02)	0.902	4.03	2.19/2.44	0.821	4.06	4.98/4.63	0.892	4.04	2.42/2.70	0.795	4.01	6.28/5.09
(0.12,0.01)	0.765	3.99	6.54/6.08	0.698	3.63	10.99/5.45	0.749	3.94	7.28/6.13	0.682	3.50	12.21/5.28
(0.12,0.02)	0.808	4.06	4.31/4.82	0.745	3.83	7.85/5.01	0.795	4.06	4.59/5.03	0.722	3.70	9.55/4.98
(0.15,0.01)	0.735	3.76	10.29/7.32	0.595	3.12	16.31/6.84	0.710	3.65	11.38/7.30	0.559	2.95	17.93/6.75
(0.15,0.02)	0.753	3.87	9.12/7.37	0.637	3.31	14.5/6.78	0.736	3.80	9.75/7.35	0.608	3.18	15.74/6.74
DGP 3												
MS-GOALS	0.928	4.02	1.50/1.59	0.879	4.03	2.52/2.67	0.920	4.02	1.68/1.80	0.870	4.02	2.71/2.83
(0.06,0.01)	0.780	5.83	1.79/9.96	0.744	5.68	3.70/9.70	0.775	5.79	2.05/9.90	0.734	5.57	4.65/9.49
(0.08,0.01)	0.877	4.43	1.91/5.32	0.832	4.26	4.17/5.07	0.864	4.40	2.51/5.34	0.809	4.12	6.08/4.92
(0.08,0.02)	0.864	4.41	2.28/5.34	0.804	4.62	3.65/6.74	0.853	4.45	2.53/5.56	0.788	4.64	4.25/6.98
(0.10,0.01)	0.927	4.02	1.62/1.67	0.871	3.90	4.25/2.32	0.917	4.01	2.13/1.91	0.848	3.85	5.54/2.58
(0.10,0.02)	0.917	4.00	1.70/1.67	0.866	4.00	3.36/2.90	0.908	3.99	2.15/1.88	0.852	3.96	4.14/3.05
(0.12,0.01)	0.866	3.76	6.31/1.74	0.795	3.56	10.26/2.51	0.850	3.71	7.34/1.89	0.783	3.54	10.83/2.67
(0.12,0.02)	0.897	3.93	3.01/1.72	0.837	3.82	5.94/2.60	0.882	3.89	3.93/1.92	0.819	3.76	7.12/2.79
(0.15,0.01)	0.896	3.81	4.90/1.33	0.770	3.23	13.74/1.84	0.870	3.69	7.23/1.37	0.737	3.09	15.79/2.01
(0.15,0.02)	0.905	3.89	3.49/1.47	0.816	3.51	9.75/1.92	0.889	3.83	4.70/1.54	0.785	3.37	11.89/2.05

S6 | ADDITIONAL RESULTS FOR REAL DATA ANALYSIS

This section provides additional results for real data analysis. Table S6 provides detailed information of the 35 countries analyzed in the real data analysis. Table S7 and Table S8 provide the detailed results of the one-week (5-day) and two-week (12-day) ahead forecasts given by M-GOALS and CDC Ensemble.

TABLE S6 Detailed information on 35 major countries used in the real data analysis

Full name	Abbreviation	Start date	n	Full name	Abbreviation	Start date	n
United States	USA	Mar-11	242	Iraq	IRQ	Apr-07	211
India	IND	Mar-30	223	Ukraine	UKR	Apr-04	210
Brazil	BRA	Mar-22	229	Indonesia	IDN	Mar-28	225
Russia	RUS	Mar-28	225	Bangladesh	BGD	Apr-15	207
France	FRA	Mar-09	227	Czech Republic	CZE	Mar-23	230
Spain	ESP	Mar-08	206	Netherlands	NLD	Mar-16	237
Argentina	ARG	Apr-02	208	Philippines	PHL	Mar-29	219
United Kingdom	GBR	Mar-13	240	Turkey	TUR	Mar-23	230
Colombia	COL	Apr-02	219	Saudi Arabia	SAU	Mar-27	225
Mexico	MEX	Mar-24	218	Pakistan	PAK	Mar-26	216
Peru	PER	Apr-01	251	Israel	ISR	Mar-23	228
Italy	ITA	Mar-01	251	Romania	ROU	Mar-27	225
South Africa	ZAF	Mar-28	224	Australia	AUS	Mar-22	230
Iran	IRN	Mar-03	247	Canada	CAN	Mar-22	227
Germany	DEU	Mar-10	243	China	CHN	Jan-25	285
Chile	CHL	Mar-26	227	Japan	JPN	Mar-21	226
Belgium	BEL	Mar-13	240	South Korea	KOR	Feb-26	250
Poland	POL	Mar-26	226				

Figure S2 and Figure S3 plot the estimated piecewise linear quantile trend models based on M-GOALS and the corresponding SN test statistics $\{T^M_{n,\epsilon,\delta}(k)\}_{k=1}^n$ for nine representative countries, including United States (1st globally/ 1st North America), India (2nd globally/ 1st Asia), Brazil (3rd globally/ 1st South America), France (4th globally), Russia (5th globally), Spain (6th globally), United Kingdom (7th globally), South Africa (1st Africa) and Australia (1st Oceania). For comparison, Figure S4 plots the estimated quantile curves across $\tau^M = (0.1, 0.5, 0.9)$ by the L_1 quantile trend filtering (TF) in Brantley et al. (2020), which performs notably worse than M-GOALS, especially for U.S., France and South Africa.

Figure S5 visualizes the dissimilarity matrix D calculated at the quantile level $\tau = 0.1$ and $\tau = 0.9$ based on M-GOALS, which are consistent with the pattern exhibited by the dissimilarity matrix D at $\tau = 0.5$ in Figure 2 of the main text. Figure S6 plots the clustering results based on multi-scanning M-GOALS, which closely matches patterns in Figure 2 and Figure S5, where continental European countries form a cluster and developing countries in Asia and Latin America tend to cluster together. This result helps further confirm the robustness of our analysis in the main text.

Figure S7 visualizes the M-GOALS based two-stage forecast scheme at four representative dates. For comparison, we further mark the in-sample change-points estimated by both M-GOALS and multi-scanning M-GOALS in Figure S7. The quantile regression BIC selects quadratic trend extrapolation on Nov-09 and Sep-21 and selects linear trend extrapolation on Oct-26 and Oct-05. A notable phenomenon in Figure S7 is the robustness of the estimated change-points by M-GOALS. For example, the change-points around Apr-04, Jun-16, Jul-20 and Sep-05 are consistently detected by both M-GOALS and multi-scanning M-GOALS even though the data lengths vary significantly across the four plots.

Figure S8 visualizes the forecast results given by multi-scanning M-GOALS and CDC Ensemble from Aug-03 to Nov-09. Note that the forecasts and prediction intervals given by multi-scanning M-GOALS are almost identical as the ones reported in Figure 3 by M-GOALS, and can be seen as further support for the robustness of our forecast analysis.

TABLE S7 One-week ahead forecast results by M-GOALS and CDC Ensemble

Date	Target	True	$M\text{-}G_{50\%}$	$M\text{-}G_{10\%}$	$M\text{-}G_{90\%}$	Covered	CDC	$CDC_{2.5\%}$	$CDC_{97.5\%}$	Covered
Nov-09	Nov-14	1000069	871137	842248	995037	0	790658	638931	947117	0
			-12.89%				-20.94%			
Nov-02	Nov-07	692118	583181	524650	641921	0	568680	493193	703080	1
			-15.74%				-17.83%			
Oct-26	Oct-31	553758	484375	424909	542660	0	472736	388545	544569	0
			-12.53%				-14.63%			
Oct-19	Oct-24	443528	406576	340737	487420	1	388452	312392	431219	0
			-8.33%				-12.42%			
Oct-12	Oct-17	385465	348801	327500	388603	1	328053	270335	380219	0
			-9.51%				-14.89%			
Oct-05	Oct-10	332391	324114	273873	375097	1	292071	232163	349045	1
			-2.49%				-12.13%			
Sep-28	Oct-03	298855	328112	274198	372623	1	302407	236640	335481	1
			9.79%				1.19%			
Sep-21	Sep-26	308763	280495	211615	336088	1	265469	210378	319854	1
			-9.16%				-14.02%			
Sep-14	Sep-19	279379	238552	196152	277262	0	240244	196505	292978	1
			-14.61%				-14.01%			
Sep-07	Sep-12	243562	254203	221821	293040	1	272867	227553	330970	1
			4.37%				12.03%			
Aug-31	Sep-05	284287	274667	252220	314902	1	273327	223480	334145	1
			-3.38%				-3.86%			
Aug-24	Aug-29	293712	285282	243747	311345	1	276838	237779	342803	1
			-2.87%				-5.75%			
Aug-17	Aug-22	310647	317664	295344	341963	1	333170	298148	405935	1
			2.26%				7.25%			
Aug-10	Aug-15	371284	372175	300024	406560	1	375174	308971	461211	1
			0.24%				1.05%			
Aug-03	Aug-08	379759	396751	330540	475710	1	434142	367977	528801	1
			4.47%				14.32%			

TABLE S8 Two-week ahead forecast results by M-GOALS and CDC Ensemble

Date	Target	True	$M\text{-}G_{50\%}$	$M\text{-}G_{10\%}$	$M\text{-}G_{90\%}$	Covered	CDC	$CDC_{2.5\%}$	$CDC_{97.5\%}$	Covered
Nov-09	Nov-21	2174399	2019046	1973337	2354446	1	1710205	1245752	2153942	0
			-7.14%				-21.35%			
Nov-02	Nov-14	1692187	1241133	1105565	1373880	0	1182692	975098	1479338	0
			-26.66%				-30.11%			
Oct-26	Nov-07	1245876	1005683	861745	1140638	0	980921	751895	1143782	0
			-19.28%				-21.27%			
Oct-19	Oct-31	997286	885732	715127	1098614	1	777290	591161	906026	0
			-11.19%				-22.06%			
Oct-12	Oct-24	828993	719134	684153	805284	0	662367	513274	775630	0
			-13.25%				-20.1%			
Oct-05	Oct-17	717856	680019	561827	793689	1	586936	444293	710447	0
			-5.27%				-18.24%			
Sep-28	Oct-10	631246	718933	578490	805452	1	599714	444284	679848	1
			13.89%				-5%			
Sep-21	Oct-03	607618	566564	393321	707163	1	525370	384559	667686	1
			-6.76%				-13.54%			
Sep-14	Sep-26	588142	452071	351938	543236	0	477286	372333	602009	1
			-23.14%				-18.85%			
Sep-07	Sep-19	522941	488394	412815	579727	1	549325	423477	672472	1
			-6.61%				5.05%			
Aug-31	Sep-12	527849	528992	479230	623420	1	540705	418297	680673	1
			0.22%				2.44%			
Aug-24	Sep-05	577999	551185	454447	610436	1	525749	437125	699347	1
			-4.64%				-9.04%			
Aug-17	Aug-29	604359	610720	568088	664063	1	635506	549572	821836	1
			1.05%				5.15%			
Aug-10	Aug-22	681931	730587	546471	809676	1	737726	577902	945571	1
			7.14%				8.18%			
Aug-03	Aug-15	751043	786146	611605	987338	1	843348	702270	1075070	1
			4.67%				12.29%			

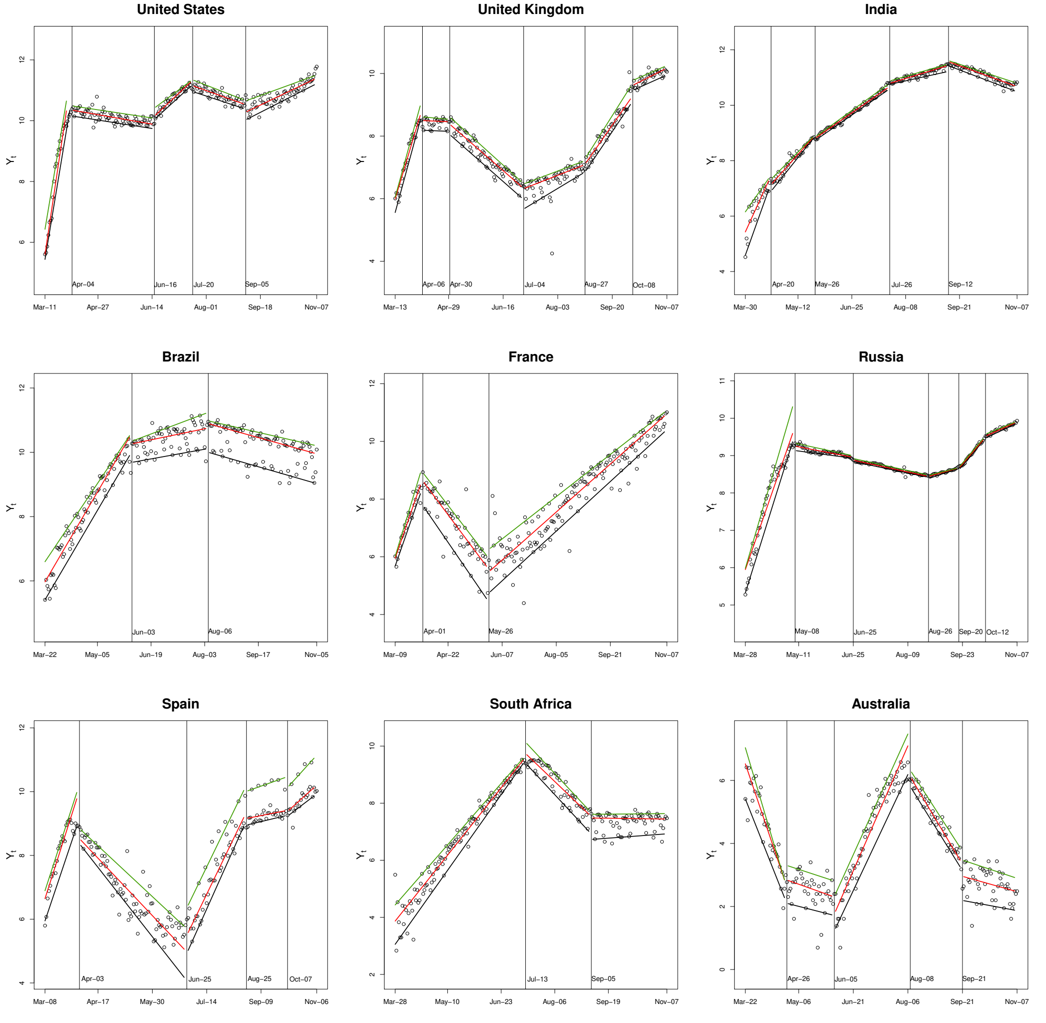

FIGURE S2 Estimated piecewise linear quantile trend models by M-GOALS for nine representative countries. The estimated change-points are marked by solid vertical lines.

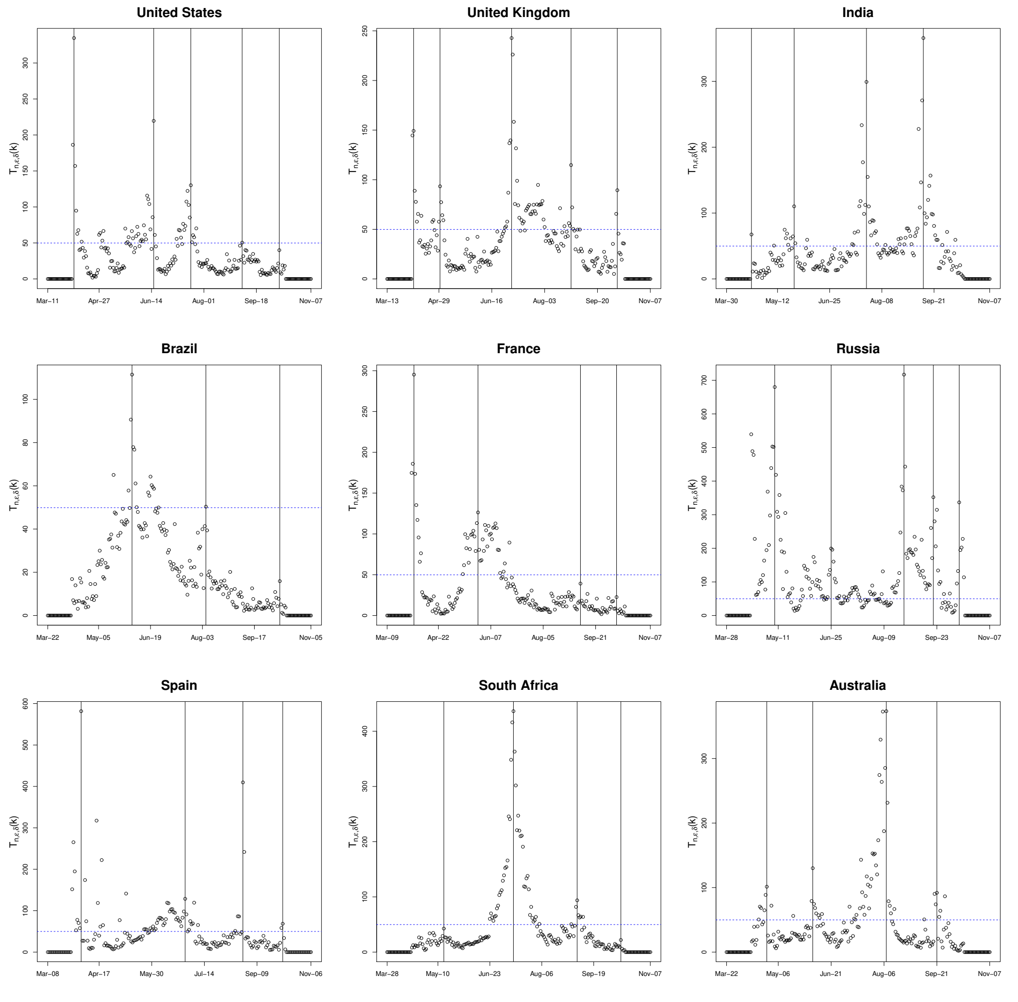

FIGURE S3 The computed SN test statistics $\{T^M_{n,\epsilon,\delta}(k)\}_{k=1}^n$ for nine representative countries. Vertical solid lines mark the h-local maximizer and horizontal dotted lines mark the threshold ζ_n^M of M-GOALS.

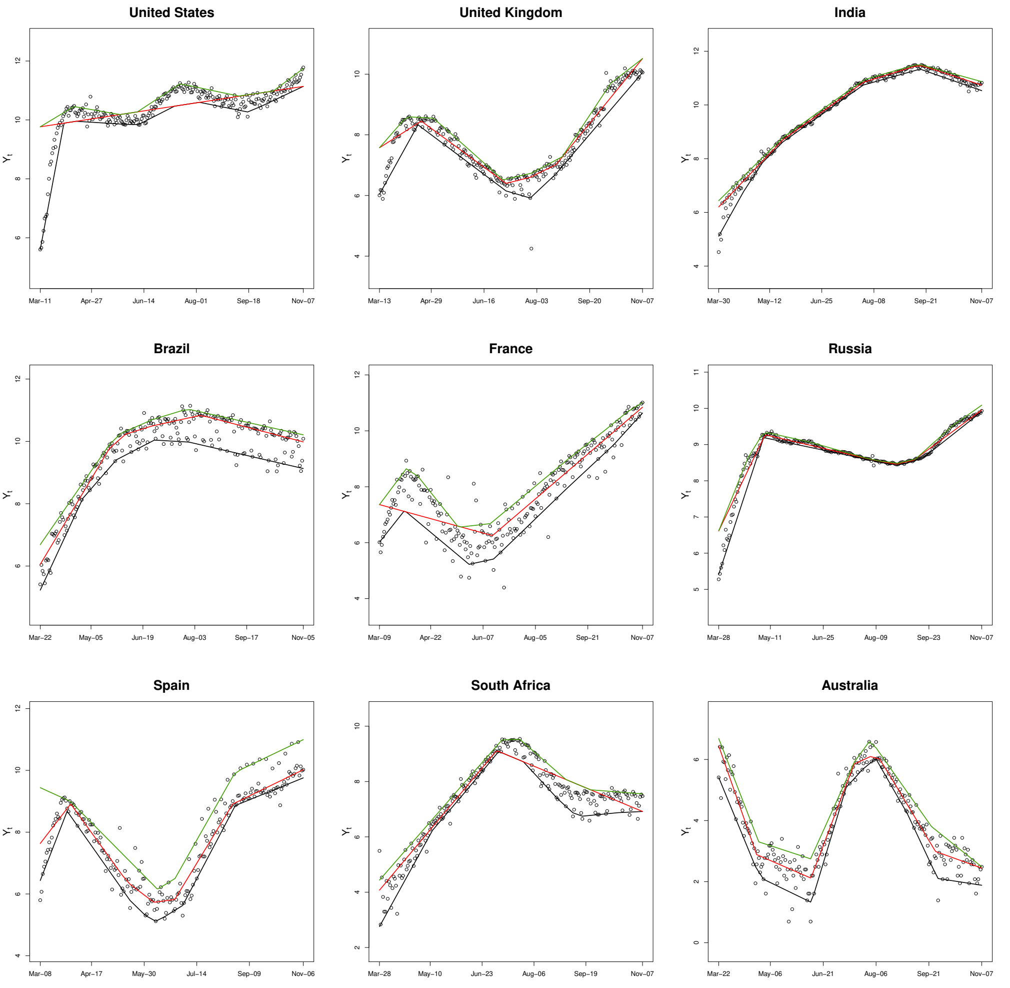

FIGURE S4 Estimated quantile curves at $\boldsymbol{\tau}^M = (0.1, 0.5, 0.9)$ by the L_1 quantile trend filtering in Brantley et al. (2020) for nine representative countries.

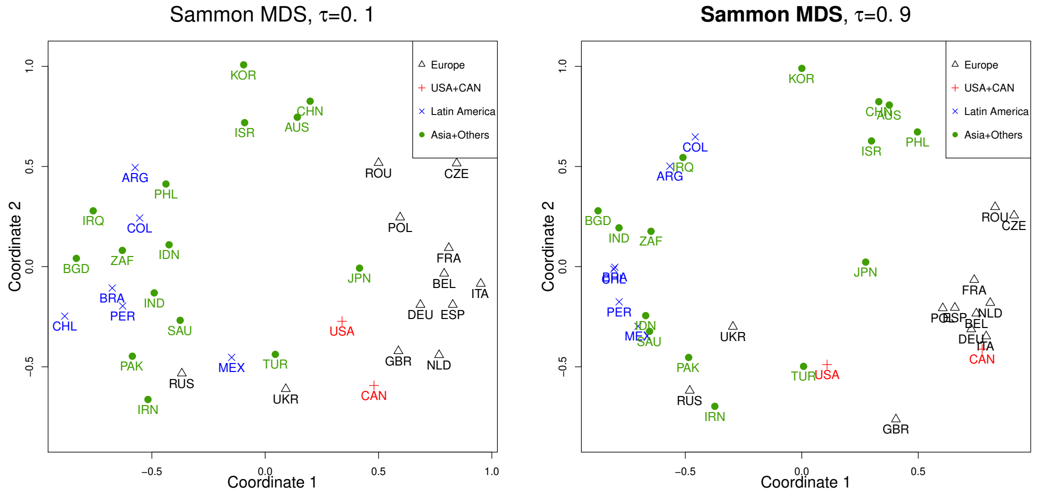

FIGURE S5 Visualization of the dissimilarity matrix D for $\tau = 0.1$ and $\tau = 0.9$ via the Sammon MDS.

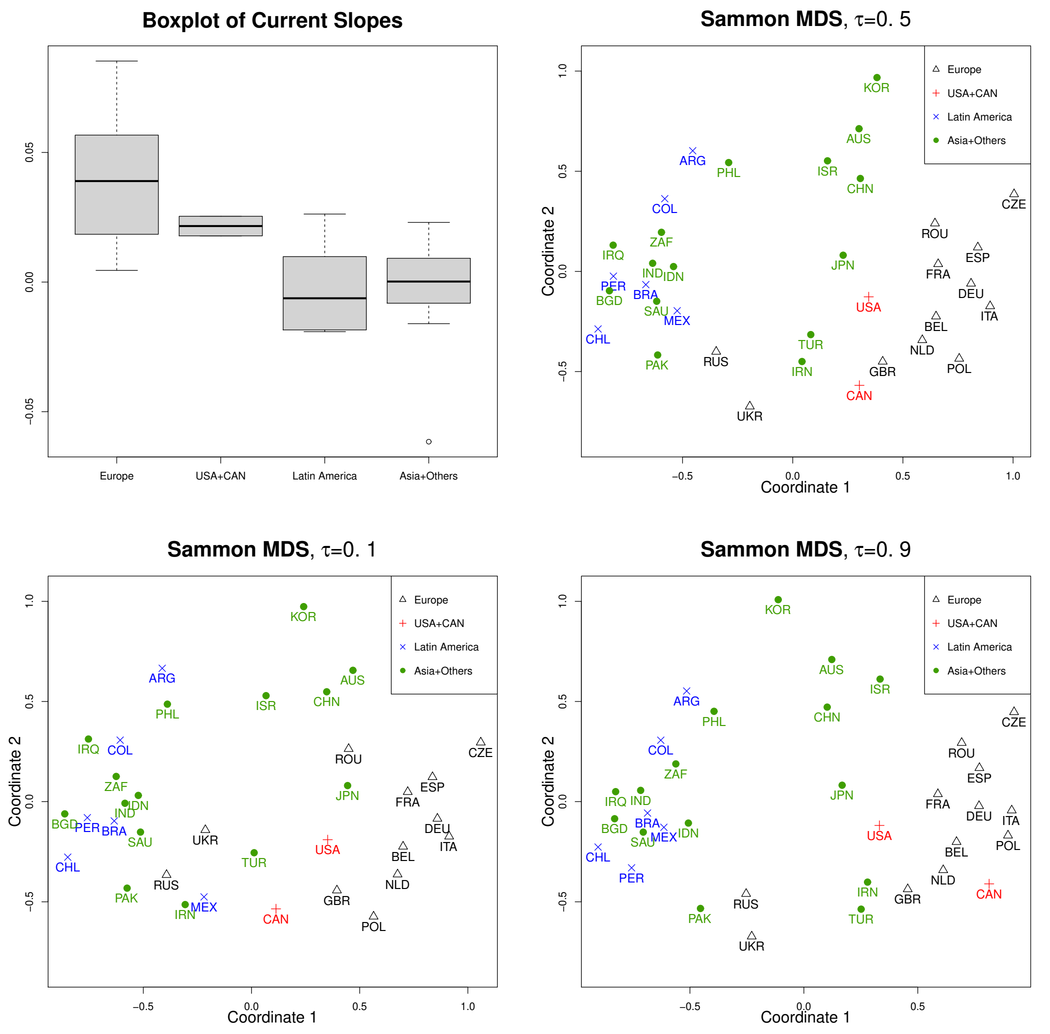

FIGURE S6 Boxplot of the current growth rate at $\tau = 0.5$ across four groups of countries, and visualization of the dissimilarity matrix D for $\tau = 0.1, 0.5, 0.9$ via the Sammon MDS. The result is based on multi-scanning M-GOALS.

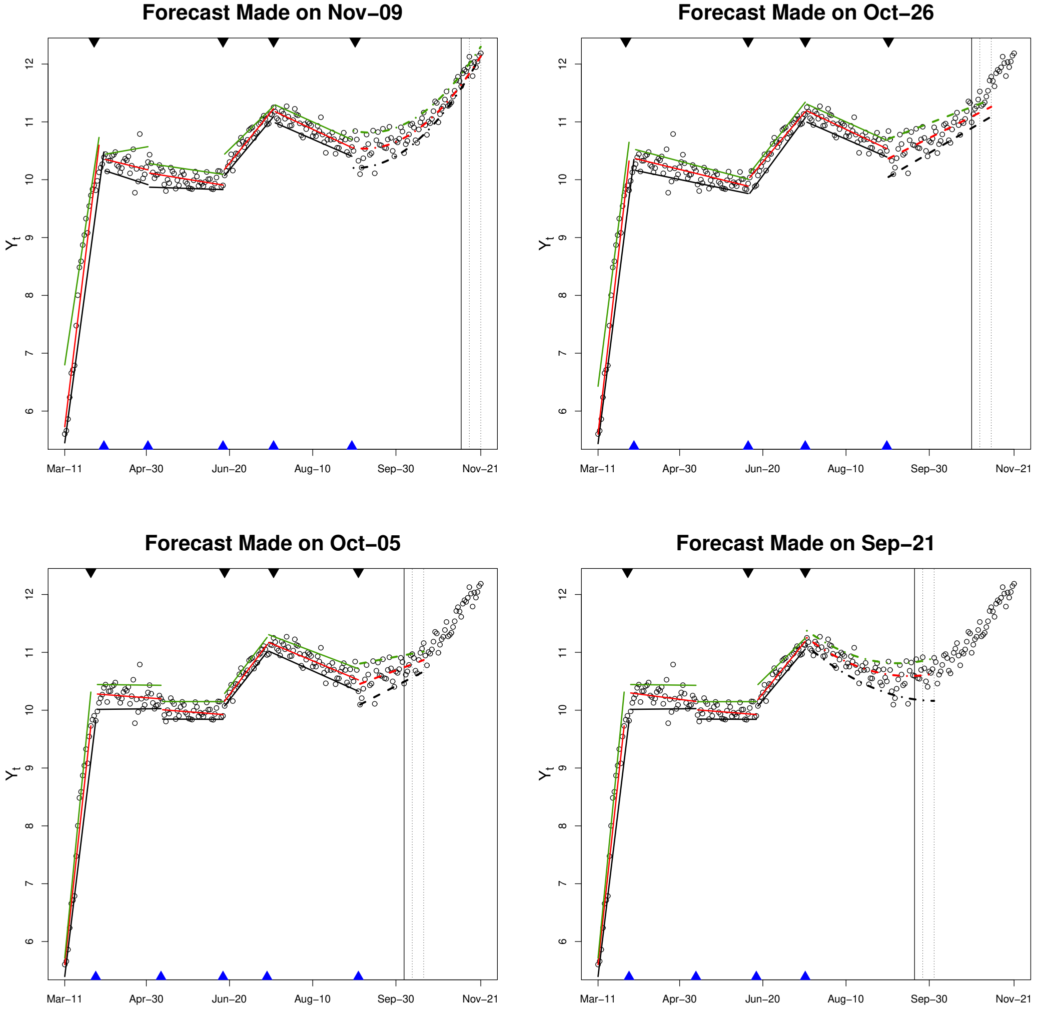

FIGURE S7 Out-of-sample forecasts for U.S. cumulative new cases on Sep-21, Oct-09, Oct-26 and Nov-09. Solid vertical lines mark the end of in-sample data used for prediction. Dotted vertical lines mark the one-week (5-day) and two-week (12-day) ahead target dates. The dashed lines mark the extrapolation function (estimated on last segment) selected by quantile regression BIC. Blue (black) triangles mark the in-sample change-points estimated by M-GOALS (multi-scanning M-GOALS).

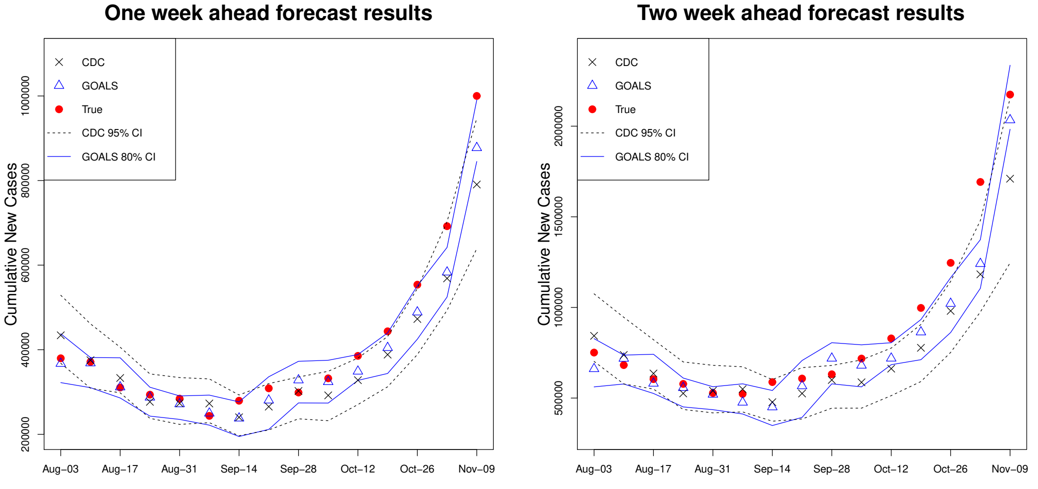

FIGURE S8 Forecast results for one-week and two-week ahead cumulative new cases from Aug-03 to Nov-09. The result is based on multi-scanning M-GOALS.

references

Aue, A., Cheung, R. C. Y., Lee, T. C. and Zhong, M. (2014) Segmented model selection in quantile regression using the minimum description length principle. *Journal of the American Statistical Association*, **109**, 1241–1256.

Billingsley, P. (1968) *Convergence of Probability Measures*. John Wiley & Sons.

Brantley, H. L., Guinness, J. and Chi, E. C. (2020) Baseline drift estimation for air quality data using quantile trend filtering. *Annals of Applied Statistics*, **14**, 585–604.

Hubert, L. and Arabie, P. (1985) Comparing partitions. *Journal of Classification*, **2**, 193–218.

Kato, K. (2009) Asymptotics for argmin processes: convexity arguments. *Journal of Multivariate Analysis*, **100**, 1816–1829.

Knight, K. (1998) Limiting distributions for l_1 regression estimators under general conditions. *Annals of Statistics*, **26**, 755–770.

Lee, E. R., Noh, H. and Park, B. U. (2014) Model selection via bayesian information criterion for quantile regression models. *Journal of the American Statistical Association*, **109**, 216–229.

Niu, S. Y. and Zhang, H. (2012) The screening and ranking algorithm to detect DNA copy number variations. *Annals of Applied Statistics*, **6**, 1306–1326.

Oka, T. and Qu, Z. (2011) Estimating structural changes in regression quantiles. *Journal of Econometrics*, **162**, 248–267.

Wu, W. B. (2007) Strong invariance principles for dependent random variables. *The Annals of Probability*, **35**, 2294–2320.

Wu, Y. and Zen, M. M. (1999) A strongly consistent information criterion for linear model selection based on m-estimation. *Probability Theory and Related Fields*, **113**, 599–625.

Yau, C. Y. and Zhao, Z. (2016) Inference for multiple change points in time series via likelihood ratio scan statistics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **78**, 895–916.

Zhou, Z. and Shao, X. (2013) Inference for linear models with dependent errors. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **75**, 323–343.