

SUPPLEMENT TO “TIME SERIES ANALYSIS OF COVID-19 INFECTION CURVE: A CHANGE-POINT PERSPECTIVE”

BY FEIYU JIANG^{†*}, ZIFENG ZHAO[§] AND XIAOFENG SHAO^{¶†}

*Tsinghua University[‡], University of Notre Dame[§] and University of Illinois at Urbana
Champaign[¶]*

This supplement consists of three parts. Appendix A contains technical proofs. Appendix B extends the piecewise linear structure of model (1.1) to piecewise polynomial and presents an analysis for cumulative confirmed cases in 8 representative countries using a piecewise quadratic model. Appendix C provides the lag-1 to lag-30 (P)ACF plots of the residuals for cumulative confirmed cases in the 8 countries presented in Section 4.2.

APPENDIX A: PROOFS

In what follows, we denote \Rightarrow as the weak convergence on $D[\epsilon, 1]$, the space of functions on $[\epsilon, 1]$ which are right continuous and have left limits, endowed with Skorohod metric. Let $X_n \in \mathbb{R}^d$ with dimension $d > 0$ be a set of random vector defined in a probability space $(\Omega, \mathbb{P}, \mathcal{F})$. For a corresponding set of constants a_n , we say $X_n = O_p^s(a_n)$ if for any $\epsilon > 0$, there exists a finite $M > 0$ and a finite $N > 0$ such that for $n > N$,

$$\mathbb{P}(\|X_n/a_n\|_d > M) + \mathbb{P}(\|X_n/a_n\|_d < 1/M) < \epsilon,$$

where $\|\cdot\|_d$ denotes the L_d norm.

PROOF OF THEOREM 2.1 (i) It is a direct application of Theorem 3.1 in Rho and Shao (2015) and continuous mapping theorem. In particular, the result of (i) in Theorem 3.1 in

*Jiang is supported by China Scholarship Council (No.201906210093), National Natural Science Foundation of China (No.11571348 and No.11771239) and acknowledges that the work was carried out during the visit at Department of Statistics, University of Illinois at Urbana-Champaign.

†Shao is supported in part by NSF-DMS 1807032.

Rho and Shao (2015) corresponds to the case of (i) in Assumption 2.1 for linear processes while the result of (ii) in Theorem 3.1 in Rho and Shao (2015) corresponds to the case of (ii) in Assumption 2.1 for nonlinear processes.

(ii) On one hand, note that the continuous mapping theorem indicates that

$$L_{n,\delta}(1, \boldsymbol{\tau}, n) \Rightarrow \Gamma^2 L_\delta(\kappa), \quad \text{and} \quad R_{n,\delta}(1, \boldsymbol{\tau}, n) \Rightarrow \Gamma^2 R_\delta(\kappa).$$

and it follows that $V_{n,\delta}(1, \boldsymbol{\tau}, n) \Rightarrow \Gamma^2 V_\delta(\kappa)$.

On the other hand,

$$D_n(1, \boldsymbol{\tau}, n) = \kappa(1 - \kappa)\sqrt{n}(\widehat{\boldsymbol{\beta}}_{1,\boldsymbol{\tau}} - \widehat{\boldsymbol{\beta}}_{\boldsymbol{\tau}+1,n} + \mathbf{b}) - \kappa(1 - \kappa)\sqrt{n}\mathbf{b},$$

and it is clear that

$$\begin{aligned} & \kappa(1 - \kappa)\sqrt{n}(\widehat{\boldsymbol{\beta}}_{1,\boldsymbol{\tau}} - \widehat{\boldsymbol{\beta}}_{\boldsymbol{\tau}+1,n} + \mathbf{b}) \\ \Rightarrow & \kappa(1 - \kappa)\Gamma Q(\kappa)^{-1} B_F(\kappa) - \Gamma[Q(1) - Q(\kappa)]^{-1}[B_F(1) - B_F(\kappa)] = \Gamma D(\kappa). \end{aligned}$$

Then the continuous mapping theorem indicates that

$$\begin{aligned} \text{(A.1)} \quad & (n\|\mathbf{b}\|_2^2)^{-1} D_n(1, \boldsymbol{\tau}, n)^\top V_{n,\delta}(1, \boldsymbol{\tau}, n)^{-1} D_n(1, \boldsymbol{\tau}, n) \\ & \Rightarrow \kappa^2(1 - \kappa)^2 (\|\mathbf{b}\|_2^{-1} \mathbf{b})^\top V_\delta(\kappa)^{-1} (\|\mathbf{b}\|_2^{-1} \mathbf{b}) = O_p^s(1). \end{aligned}$$

Here the last equality uses the fact that RHS of (A.1) is greater than 0 with probability 1, or equivalently, $L_\delta(\kappa)$ and $R_\delta(\kappa)$ is positive definite with probability 1, which will hold by similar arguments in Lemma A.1 using Cauchy–Schwarz inequality.

Observe that $\max_k T_{n,\delta}(k) \geq D_n(1, \boldsymbol{\tau}, n)^\top V_{n,\delta}(1, \boldsymbol{\tau}, n)^{-1} D_n(1, \boldsymbol{\tau}, n) = O_p^s(n\|\mathbf{b}\|_2^2)$. The result follows by noting $n\|\mathbf{b}\|_2^2 \rightarrow L$ and $L \rightarrow \infty$, \square

PROOF OF THEOREM 2.2 Note that by (A.1), we have shown that with probability tending to one, $(n\|\mathbf{b}\|_2^2)^{-1} T_{n,\delta}(\boldsymbol{\tau}) \geq \kappa^2(1 - \kappa)^2 (\|\mathbf{b}\|_2^{-1} \mathbf{b})^\top V_\delta(\kappa)^{-1} (\|\mathbf{b}\|_2^{-1} \mathbf{b}) = O_p^s(1)$.

Then, let $M_{n,\eta} = \{k : |\frac{k}{n} - \kappa| > \eta\}$, it suffices to show that

$$(n\|\mathbf{b}\|_2^2)^{-1} \max_{k \in [h, n-h] \cap M_{n,\eta}} D_n(1, k, n)^\top V_{n,\delta}(1, k, n)^{-1} D_n(1, k, n) = o_p(1).$$

By symmetricity, we can consider $M_{n,\eta}^{(1)} = \{k : \frac{k}{n} < \kappa - \eta\}$, and on $\{k \in M_{n,\eta}^{(1)}\}$, we have

$$D_n(1, k, n) = \frac{k(n-k)}{n^{3/2}} \left\{ [Q_n(1) - Q_n(\frac{k}{n})]^{-1} [Q_n(1) - Q_n(\frac{\tau}{n})] (\boldsymbol{\beta}^{(1)} - \boldsymbol{\beta}^{(2)}) \right. \\ \left. + n^{-1/2} Q_n(\frac{k}{n})^{-1} B_{n,F}(\frac{k}{n}) - n^{-1/2} [Q_n(1) - Q_n(\frac{k}{n})]^{-1} [B_{n,F}(1) - B_{n,F}(\frac{k}{n})] \right\}.$$

Let $\nu = \lim_{n \rightarrow \infty} \frac{k}{n} := \lim_{n \rightarrow \infty} \frac{k(n)}{n}$, then by similar arguments in (i) of Theorem 2.1, we have

(A.2)

$$n^{-1/2} \|\mathbf{b}\|_2^{-1} D_n(1, k, n) = \nu(1-\nu)[Q(1) - Q(\nu)]^{-1} [Q(1) - Q(\kappa)] \|\mathbf{b}\|_2^{-1} \mathbf{b} + O_p(n^{-1/2} \|\mathbf{b}\|_2^{-1}).$$

Next, since $k < \tau - n\eta$, we decompose $R_{n,\delta}(1, k, n)$ by

$$R_{n,\delta}(1, k, n) = \left[\sum_{i=k+3+[n\delta]}^{\tau+[n\delta]-1} + \sum_{i=\tau+[n\delta]}^{n-1-[n\delta]} \right] \frac{(i-1-k)^2(n-i+1)^2}{n^2(n-k)^2} (\widehat{\boldsymbol{\beta}}_{i,n} - \widehat{\boldsymbol{\beta}}_{k+1,i-1})^{\otimes 2} \\ := R_{n,\delta,1}(1, k, n) + R_{n,\delta,2}(1, k, n).$$

It follows easily that $V_{n,\delta}(1, k, n)^{-1} \leq R_{n,\delta}(1, k, n)^{-1} \leq R_{n,\delta,2}(1, k, n)^{-1}$ where for semi-positive definite matrices A and B , $A \leq B$ indicates $B - A$ is semi-positive definite.

In addition, we have

$$R_{n,\delta,2}(1, k, n) = \sum_{i=\tau+[n\delta]}^{n-1-[n\delta]} \frac{(i-1-k)^2(n-i+1)^2}{n^2(n-k)^2} (\widehat{\boldsymbol{\beta}}_{i,n} - \widehat{\boldsymbol{\beta}}_{k+1,i-1})^{\otimes 2},$$

where for $r \in (\kappa, 1)$ uniformly, we can show

$$\sqrt{n} \left(\widehat{\boldsymbol{\beta}}_{[nr],n} - \widehat{\boldsymbol{\beta}}_{[n\nu]+1,[nr]-1} - [Q(r) - Q(\nu)]^{-1} [Q(\kappa) - Q(\nu)] \mathbf{b} \right) \\ \Rightarrow \Gamma [Q(1) - Q(r)]^{-1} [B_F(1) - B_F(r)] - \Gamma [Q(r) - Q(\nu)]^{-1} [B_F(r) - B_F(\nu)] = O_p^s(1).$$

Therefore, it follows that

$$(A.3) \quad (n \|\mathbf{b}\|_2^2)^{-1} R_{n,\delta,2}(1, k, n) \\ \Rightarrow \int_{\kappa+\delta}^{1-\delta} \frac{(r-\nu)^2(1-r)^2}{(1-\nu)^2} \left\{ \|\mathbf{b}\|_2^{-1} [Q(r) - Q(\nu)]^{-1} [Q(\kappa) - Q(\nu)] \mathbf{b} \right\}^{\otimes 2} dr := \overline{R}_{\delta,2}(\nu).$$

By Lemma A.1, when $\nu < \kappa$, $\bar{R}_{\delta,2}(\nu)$ is invertible, hence

(A.4)

$$\begin{aligned} & (n\|\mathbf{b}\|_2^2)^{-1}D_n(1, k, n)^\top V_{n,\delta}(1, k, n)^{-1}D_n(1, k, n) \\ &= (n\|\mathbf{b}\|_2^2)^{-1}[n^{-1/2}\|\mathbf{b}\|_2^{-1}D_n(1, k, n)]^\top [(n\|\mathbf{b}\|_2^2)^{-1}R_{n,\delta,2}(1, k, n)]^{-1}[n^{-1/2}\|\mathbf{b}\|_2^{-1}D_n(1, k, n)] \\ &\Rightarrow (n\|\mathbf{b}\|_2^2)^{-1}\left\{\nu(1-\nu)[Q(1)-Q(\nu)]^{-1}[Q(1)-Q(\kappa)]\|\mathbf{b}\|_2^{-1}\mathbf{b}\right\}^\top \bar{R}_{\delta,2}(\nu) \\ &\quad \times \left\{\nu(1-\nu)[Q(1)-Q(\nu)]^{-1}[Q(1)-Q(\kappa)]\|\mathbf{b}\|_2^{-1}\mathbf{b}\right\} \Rightarrow 0 \end{aligned}$$

by (A.2) and (A.3). \square

LEMMA A.1. $\bar{R}_{\delta,2}(\nu)$, defined in (A.3), is invertible for $\nu < \kappa$ and $\|\mathbf{b}\|_2 \neq 0$.

PROOF OF LEMMA A.1

Note that

$$\begin{aligned} [Q(\kappa) - Q(\nu)] &= (\kappa - \nu) \begin{pmatrix} 1 & \frac{\kappa + \nu}{2} \\ \frac{\kappa + \nu}{2} & \frac{\nu^2 + \kappa^2 + \kappa\nu}{3} \end{pmatrix}, \\ [Q(r) - Q(\nu)]^{-1} &= 12(r - \nu)^{-3} \begin{pmatrix} \frac{r^2 + \nu^2 + r\nu}{3} & -\frac{r + \nu}{2} \\ -\frac{r + \nu}{2} & 1 \end{pmatrix}. \end{aligned}$$

We first let $\mathbf{b} = (b_1, b_2)^\top$, then

$$[Q(\kappa) - Q(\nu)]\|\mathbf{b}\|_2^{-1}\mathbf{b} = (\kappa - \nu)\|\mathbf{b}\|_2^{-1} \begin{pmatrix} b_1 + \frac{\kappa + \nu}{2}b_2 \\ \frac{\kappa + \nu}{2}b_1 + \frac{\nu^2 + \kappa^2 + \kappa\nu}{3}b_2 \end{pmatrix} := (w_1, w_2)^\top.$$

Therefore we obtain

$$\begin{aligned} & \frac{(r - \nu)(1 - r)}{(1 - \nu)} [Q(r) - Q(\nu)]^{-1} [Q(\kappa) - Q(\nu)]\|\mathbf{b}\|_2^{-1}\mathbf{b} = \frac{12(1 - r)}{(r - \nu)^2(1 - \nu)} \begin{pmatrix} \frac{r^2 + \nu^2 + r\nu}{3}w_1 - \frac{r + \nu}{2}w_2 \\ -\frac{r + \nu}{2}w_1 + w_2 \end{pmatrix} \\ & := \left(g_1(r, \nu, \kappa, b_1, b_2), g_2(r, \nu, \kappa, b_1, b_2) \right)^\top. \end{aligned}$$

Then, since

$$\bar{R}_{\delta,2}(\nu) = \int_{\kappa + \delta}^{1 - \delta} \left((g_1(r, \nu, \kappa, b_1, b_2), g_2(r, \nu, \kappa, b_1, b_2)) \right)^\top \left((g_1(r, \nu, \kappa, b_1, b_2), g_2(r, \nu, \kappa, b_1, b_2)) \right) dr,$$

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

$$\int_{\kappa+\delta}^{1-\delta} g_1(r, \nu, \kappa, b_1, b_2)^2 dr \int_{\kappa+\delta}^{1-\delta} g_2(r, \nu, \kappa, b_1, b_2)^2 dr - \left[\int_{\kappa+\delta}^{1-\delta} g_1(r, \nu, \kappa, b_1, b_2) g_2(r, \nu, \kappa, b_1, b_2) dr \right]^2 > 0,$$

which is implied by Cauchy–Schwarz inequality as long as

$$(A.5) \quad \frac{g_1(r, \nu, \kappa, b_1, b_2)}{g_2(r, \nu, \kappa, b_1, b_2)} = \frac{2(r^2 + \nu^2 + r\nu)w_1 - 3(r + \nu)w_2}{-6(r + \nu)w_1 + 12w_2}$$

is not a constant for all $r \geq \kappa$.

To see this, suppose $\overline{R}_{\delta,2}(\nu)$ is not invertible, then (A.5) is a constant for all $r \geq \kappa$. Note that the numerator and the denominator of RHS of (A.5) can be written in a quadratic form of r as

$$(A.6) \quad 2w_1r^2 + (2\nu w_1 - 3w_2)r + (2\nu^2 w_1 - 3\nu w_2),$$

$$(A.7) \quad 0r^2 - 6w_1r + (-6\nu w_1 + 12w_2),$$

respectively.

Therefore, comparing coefficients of the quadratic functions (A.6) and (A.7) w.r.t r , it must hold that $w_1 = 0$, and hence $w_2 = 0$, i.e.

$$b_1 + \frac{\kappa + \nu}{2}b_2 = 0, \quad \text{and} \quad \frac{\kappa + \nu}{2}b_1 + \frac{\nu^2 + \kappa^2 + \kappa\nu}{3}b_2 = 0.$$

Solving these equations for b_1 and b_2 we obtain that $b_1 = b_2 = 0$, contradiction.

Hence, $\overline{R}_{\delta,2}(\nu)$ is invertible. □

APPENDIX B: PIECEWISE POLYNOMIAL TREND MODEL

In this section, we extend the piecewise linear structure in model (1.1) of the main text to a piecewise polynomial structure. We further apply a piecewise quadratic trend model to analyze the cumulative confirmed cases in 8 representative countries as in Section 4.2.

B.1. Model formulation and inference. We extend the piecewise linear trend model (1.1) by allowing higher order polynomial terms. Specifically, let the time series $\{Y_t\}_{t=1}^n$ admit

$$(B.1) \quad Y_t = \boldsymbol{\beta}_t^\top F_t^{(p)} + u_t = \beta_{0,t} + \beta_{1,t}(t/n) + \cdots + \beta_{p,t}(t/n)^p + u_t, \quad t = 1, \dots, n,$$

$$(\beta_{0,t}, \dots, \beta_{p,t})^\top = \boldsymbol{\beta}^{(i)} = (\beta_0^{(i)}, \dots, \beta_p^{(i)})^\top, \tau_{i-1} + 1 \leq t \leq \tau_i, \text{ for } i = 1, \dots, m+1,$$

where $F_t^{(p)} = (1, t/n, \dots, (t/n)^p)^\top$ and $\boldsymbol{\beta}_t = (\beta_{0,t}, \dots, \beta_{p,t})^\top$ are the coefficients at time t with fixed $p \geq 1$. Same as in model (1.1), $\{u_t\}$ is a weakly dependent stationary error process, $\boldsymbol{\tau} = (\tau_1, \dots, \tau_m)$ denotes the $m \geq 0$ change-points with the convention that $\tau_0 = 0$ and $\tau_{m+1} = n$, and we require $\boldsymbol{\beta}^{(i)} \neq \boldsymbol{\beta}^{(i+1)}, i = 1, \dots, m$. Model (B.1) extends the piecewise linear model by allowing for polynomial trends and provides more flexibility of modeling observations in each segment.

The estimation procedure of model (B.1) is essentially the same as the one for model (1.1). Given the grid parameter ϵ , we let $h = \lfloor \epsilon n \rfloor$. Define $F^{(p)}(s) = (1, s, \dots, s^p)^\top$. For $1 \leq i < j \leq n$, we denote $\widehat{\boldsymbol{\beta}}_{i,j} = \left[\sum_{t=i}^j F^{(p)}(t/n) F^{(p)}(t/n)^\top \right]^{-1} \sum_{t=i}^j F^{(p)}(t/n) Y_t$ as the OLS estimator of $\boldsymbol{\beta}$ based on $\{Y_t\}_{t=i}^j$. Let the trimming parameter satisfy $0 \leq \delta < \epsilon/2$. For any $1 \leq t_1 < k < t_2 \leq n$, given the subsample $\{Y_t\}_{t=t_1}^{t_2}$ and a potential change-point k , we define a contrast statistic $D_n^{(p)}(t_1, k, t_2)$, and the self-normalizer $V_{n,\delta}^{(p)}(t_1, k, t_2) = L_{n,\delta}^{(p)}(t_1, k, t_2) + R_{n,\delta}^{(p)}(t_1, k, t_2)$ in the same spirit as (2.1), (2.2) and (2.3) by:

$$D_n^{(p)}(t_1, k, t_2) = \frac{(k - t_1 + 1)(t_2 - k)}{(t_2 - t_1 + 1)^{3/2}} (\widehat{\boldsymbol{\beta}}_{t_1,k} - \widehat{\boldsymbol{\beta}}_{k+1,t_2}),$$

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

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

Then, the test statistic targeting against the one change-point alternative is defined as:

$$G_n^{(p)} = \max_{k \in \{h, \dots, n-h\}} T_{n,\delta}^{(p)}(k), \quad T_{n,\delta}^{(p)}(k) = D_n^{(p)}(1, k, n)^\top V_{n,\delta}^{(p)}(1, k, n)^{-1} D_n^{(p)}(1, k, n).$$

Define $Q^{(p)}(r) = \int_0^r F^{(p)}(s)F^{(p)}(s)^\top ds$ and $B_F^{(p)}(r) = \int_0^r F^{(p)}(s)dB(s)$ where $B(\cdot)$ is a standard Brownian motion. The following theorem extends Theorem 2.1 in the main text.

THEOREM B.1. *Suppose Assumption 2.1 holds. Then,*

(i) *under H_0 , we have*

$$G_n^{(p)} \xrightarrow{\mathcal{D}} G^{(p)}(\epsilon, \delta) := \sup_{\eta \in (\epsilon, 1-\epsilon)} D^{(p)}(\eta)^\top V_\delta^{(p)}(\eta) D^{(p)}(\eta),$$

where $D^{(p)}(\eta)$ and $V_\delta^{(p)}(\eta) = L_\delta^{(p)}(\eta) + R_\delta^{(p)}(\eta)$ have the similar expression as given in Theorem 2.1, except $F(\cdot)$ and $Q(\cdot)$ are replaced by $F^{(p)}(\cdot)$ and $Q^{(p)}(\cdot)$ respectively.

(ii) *under H_a , given that $n\|\mathbf{b}\|_2^2 \rightarrow L$, we have*

$$\lim_{L \rightarrow \infty} \lim_{n \rightarrow \infty} G_n^{(p)} = \infty, \quad \text{in probability.}$$

The proof is a simple extension of Appendix A, hence omitted.

B.2. Analysis of cumulative confirmed cases in 8 representative countries.

We use the piecewise quadratic trend model, i.e. model (B.1) with $p = 2$ to re-analyze the cumulative confirmed cases in the 8 countries as in Section 4.2. Figure B.1 gives the estimated models for each country. As can be seen, compared to Figure 4.1 in the main text, the estimated number of change-points decreases for every country, which is intuitive as more flexibility is brought into the model. For most countries, a piecewise quadratic model with one or two change-points fits the data reasonably well.

However, compared to the piecewise linear trend model, the quadratic model loses its interpretability as the parameters of each segment cannot be naturally linked to growth rate. Thus the meaning of ‘‘change-point’’ needs a more delicate definition. Moreover, within each segment, the growth rate of the virus still changes from day to day, making it difficult to interpret the behavior of the estimated segments. For example, we find that most estimated change-points can hardly be associated with the initiations of emergency public health measures, as the intervention effect may have been absorbed into the quadratic function. Therefore, we prefer the piecewise linear trend model for the analysis.

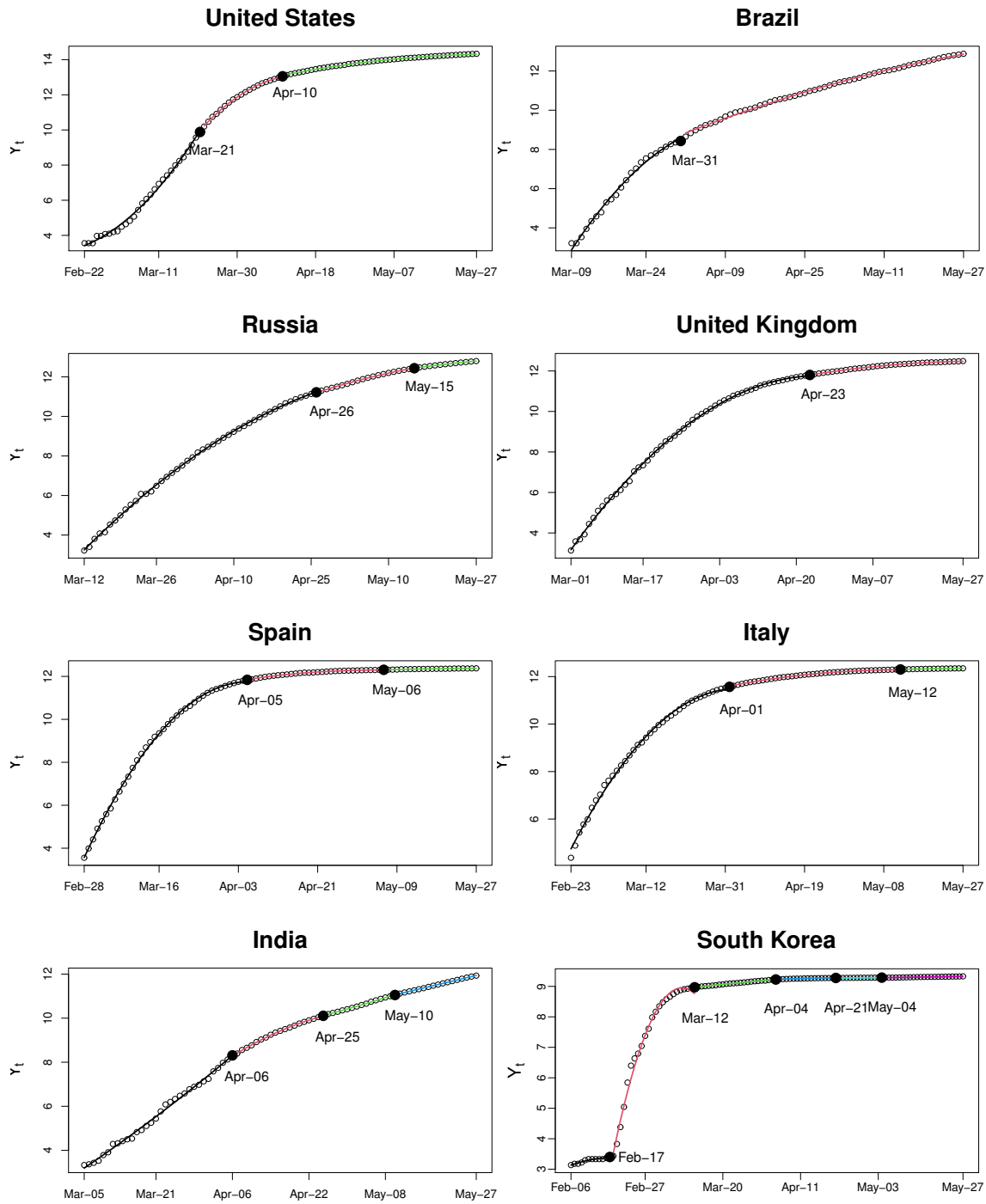


FIG B.1. Estimated piecewise quadratic trend for 8 representative countries

APPENDIX C: ACF AND PACF PLOTS OF RESIDUALS (AFTER FITTING
PIECEWISE LINEAR TREND MODEL) FOR CUMULATIVE
CONFIRMED CASES IN 8 COUNTRIES

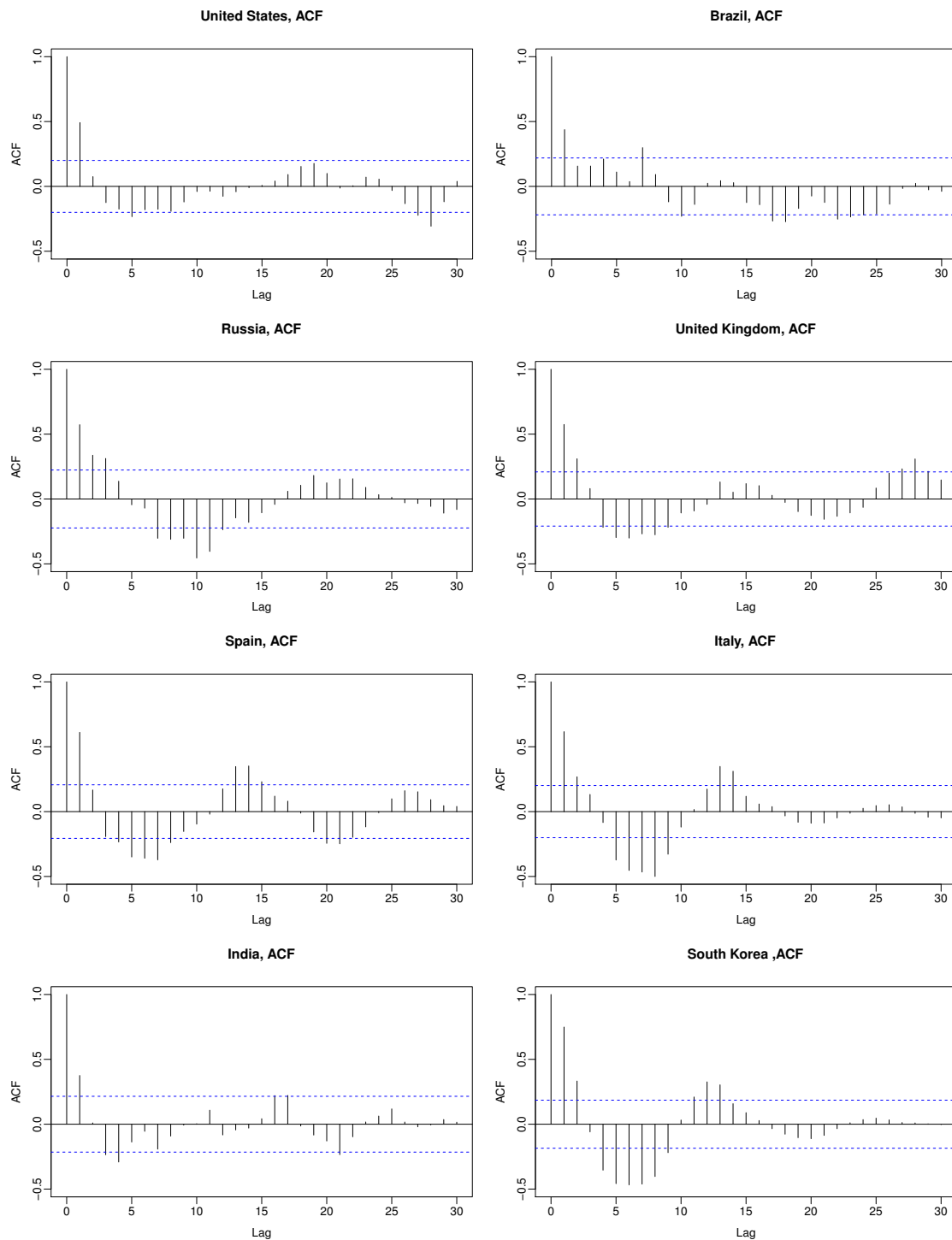


FIG C.1. ACF plot of residuals for 8 representative countries

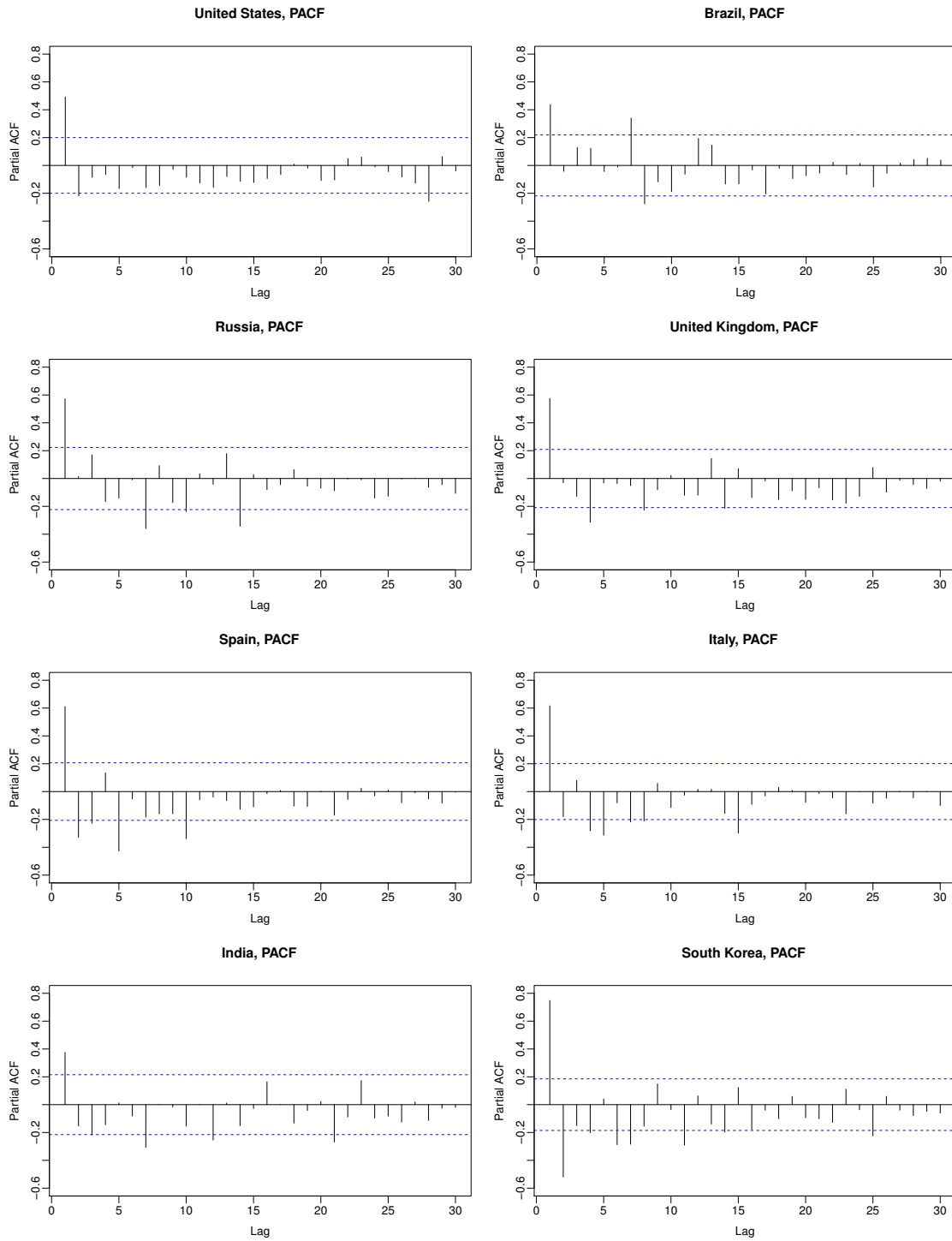


FIG C.2. PACF plot of residuals for 8 representative countries

REFERENCES

Rho, Y. and X. Shao (2015). Inference for time series regression models with weakly dependent and heteroscedastic errors. *Journal of Business & Economic Statistics* 33(3), 444–457.

CENTER FOR STATISTICAL SCIENCE

AND DEPARTMENT OF INDUSTRIAL ENGINEERING

TSINGHUA UNIVERSITY

BEIJING 100084, CHINA

E-MAIL: jfy16@mails.tsinghua.edu.cn

DEPARTMENT OF INFORMATION TECHNOLOGY, ANALYTICS, AND OPERATIONS

MENDOZA COLLEGE OF BUSINESS

UNIVERSITY OF NOTRE DAME

NOTRE DAME, INDIANA 46556, U.S.A.

E-MAIL: zzhao2@nd.edu

DEPARTMENT OF STATISTICS

UNIVERSITY OF ILLINOIS AT URBANA CHAMPAIGN

CHAMPAIGN, ILLINOIS 61820, U.S.A.

E-MAIL: xshao@illinois.edu