Journal of Econometrics 232 (2023) 1–17

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

# Time series analysis of COVID-19 infection curve: A change-point perspective

Feiyu Jiang [a,1], Zifeng Zhao [b,2], Xiaofeng Shao [c,*,3]

[a] Center for Statistical Science and Department of Industrial Engineering, Tsinghua University, Beijing 100084, China
[b] Department of Information Technology, Analytics, and Operations, Mendoza College of Business, University of Notre Dame, Notre Dame, IN 46556, USA
[c] Department of Statistics, University of Illinois at Urbana Champaign, Champaign, IL 61820, USA

ARTICLE INFO

*Article history:*
Received 31 May 2020
Received in revised form 13 June 2020
Accepted 6 July 2020
Available online 30 July 2020

ABSTRACT

In this paper, we model the trajectory of the cumulative confirmed cases and deaths of COVID-19 (in log scale) via a piecewise linear trend model. The model naturally captures the phase transitions of the epidemic growth rate via change-points and further enjoys great interpretability due to its semiparametric nature. On the methodological front, we advance the nascent self-normalization (SN) technique (Shao, 2010) to testing and estimation of a single change-point in the linear trend of a nonstationary time series. We further combine the SN-based change-point test with the NOT algorithm (Baranowski et al., 2019) to achieve multiple change-point estimation. Using the proposed method, we analyze the trajectory of the cumulative COVID-19 cases and deaths for 30 major countries and discover interesting patterns with potentially relevant implications for effectiveness of the pandemic responses by different countries. Furthermore, based on the change-point detection algorithm and a flexible extrapolation function, we design a simple two-stage forecasting scheme for COVID-19 and demonstrate its promising performance in predicting cumulative deaths in the U.S.



## 1. Introduction

Since the initial outbreak of the novel coronavirus in Wuhan, China in early January 2020, the COVID-19 pandemic has rapidly spread across the world. Due to the high infectivity of the virus and the lack of immunity in the human population, the epidemic grows exponentially without intervention, and thus can greatly stress the public health system and bring enormous disruption to economy and society. Thus, a crucial task facing every country is to reduce the transmission rate and flatten the (infection) curve. Various emergency measures, such as regional lockdown and mass testing, have been taken by different countries and a natural question is whether (and to what degree) these interventions are effective in slowing down the pandemic. Additionally, each country is at a different stage of the epidemic and it is essential for countries to understand its own pattern of virus growth, as such information is critical for important policy decisions

* Corresponding author.
*E-mail addresses:* jfy16@mails.tsinghua.edu.cn (F. Jiang), zzhao2@nd.edu (Z. Zhao), xshao@illinois.edu (X. Shao).

1 Jiang is supported by the National Key Research and Development Program of China (No.2018YFC1603105), China Scholarship Council (No. 201906210093), National Natural Science Foundation of China (No. 71973077 and No. 11771239) and acknowledges that the work was carried out during the visit at Department of Statistics, University of Illinois at Urbana-Champaign.

2 Zhao is supported in part by NSF-DMS 2014053.

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

https://doi.org/10.1016/j.jeconom.2020.07.039
0304-4076/

such as extending lockdown or reopening. To (at least partially) answer these questions, a natural step is to analyze the trajectory of the infection curve of COVID-19 since the initial outbreak in each country.

In this paper, we propose to model the time series of cumulative confirmed cases and deaths (in log scale) of each country via a piecewise linear trend model (see formal definition later). In other words, we model the mean of the logarithm of cumulative infection as a linear trend with an unknown number of potential changes in the intercept and slope, as it is natural to expect that the spread of COVID-19 may experience several phases, where the initial growth is typically rapid due to absence of immunity and lack of preparation, and the spread may then evolve into phases with slower growth depending on government intervention and public health responses (i.e. flattening the curve). The estimation of such a model can be formulated as a change-point detection problem.

In recent years, change-point analysis has become an increasingly active research area in statistics and econometrics thanks to its applications across a wide range of fields, including bioinformatics (Fan and Mackey, 2017), climate science (Gromenko et al., 2017), economics (Bai, 1994, 1997; Cho and Fryzlewicz, 2015), finance (Fryzlewicz, 2014), medical science (Chen and Gupta, 2011), and signal processing (Chen and Gu, 2018); see Perron (2006), Aue and Horváth (2013) and Truong et al. (2020) for some recent reviews. However, most existing change-point literature operates under the piecewise stationarity assumption, where it is assumed that the time series of interest is (potentially) non-stationary but can be partitioned into piecewise stationary segments such that observations within each segment are stationary and share a common parameter of interest such as mean or variance. While the piecewise stationarity assumption is proven to be reasonable and fruitful for many applications, methods developed under this framework cannot handle time series with intrinsic non-stationarity, such as the cumulative infection curve of COVID-19.

A simple but important class of time series with intrinsic non-stationarity is the piecewise linear trend model, which has the following mathematical formulation. Let the time series $\{Y_t\}_{t=1}^n$ admit

$$\begin{aligned} &Y_t = a_t + b_t(t/n) + u_t, \ \ t = 1, \ldots, n, \\ &(a_t, b_t)^\top = \boldsymbol{\beta}^{(i)} = (\beta_0^{(i)}, \beta_1^{(i)})^\top, \ \tau_{i-1} + 1 \le t \le \tau_i, \ \text{ for } i = 1, \ldots, m+1, \end{aligned} \tag{1.1}$$

where $(a_t, b_t)^\top$ is the linear trend (intercept and slope) of $\mathbb{E}(Y_t)$ at time $t$, $\{u_t\}$ is a weakly dependent stationary error process, $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_m)$ denotes the $m \ge 0$ change-points with the convention that $\tau_0 = 0$ and $\tau_{m+1} = n$, and we require $\boldsymbol{\beta}^{(i)} \neq \boldsymbol{\beta}^{(i+1)}$, $i = 1, \ldots, m$. In this paper, we set $\{Y_t\}_{t=1}^n$ to be the time series of daily cumulative confirmed cases or deaths (in log scale) of COVID-19. Due to the log transformation, the slope $b_t$ naturally measures the growth rate of the virus at day $t$.

The piecewise linear trend model is intuitive, interpretable and is useful for tracking the dynamics of a pandemic as it naturally segments the spread process into phases with (approximately) the same growth rate. The slope of the last segment can shed light on the current status of the pandemic and provide short-term forecast, while the estimated change-points can be compared with dates when emergency measures such as lockdown were introduced to help assess the effectiveness of different policies. Also, the semiparametric nature of (1.1) helps to achieve model flexibility while maintaining simplicity, which is advantageous for modeling the cumulative cases at the early stage of a pandemic as the time series is relatively short, curbing the use of sophisticated fully nonparametric methods.

An important part in estimation of (1.1) is to recover the unknown number $m$ and location $\boldsymbol{\tau}$ of the change-points. As discussed above, such a problem has mostly been ignored in the change-point literature with only a few exceptions. A CUSUM based detection algorithm is proposed in Baranowski et al. (2019), and a model selection based procedure is derived in Maidstone and Letchford (2019). However, both methods assume temporal independence of $\{u_t\}$, which can be restrictive as serial dependence is commonly found in time series data. Although Baranowski et al. (2019) briefly discussed possible extensions to temporally dependent series, potentially important issues such as choice of tuning parameters seem not carefully addressed. Bai and Perron (1998) can detect structural breaks in the linear trend model under serial dependence. However, numerical study (see Section 4) suggests that their method is relatively sensitive to positive temporal dependence, which is indeed exhibited by the COVID-19 data, and may give less favorable estimation performance under small sample size.

Based on the self-normalization (SN) idea in Shao (2010), we propose a novel SN-based change-point detection procedure for the estimation of (1.1) that is robust to temporal dependence both in asymptotic theory and in finite sample. The essential idea of SN is using an inconsistent variance estimator to absorb the unknown serial dependence in the data. See a brief review of SN in Section 2.1 and Shao (2015) for a comprehensive overview of recent developments of SN for low dimensional time series.

Using the proposed SN method and the piecewise linear trend model, we analyze the time series of cumulative confirmed cases and deaths of COVID-19 (in log scale) in 30 major countries. We find that the spread of coronavirus in each country can typically be segmented into several phases with distinct growth rates and countries with geographical proximity share similar spread patterns, which is particularly evident for continental European countries and developing countries in Latin America. In addition, the transition date from rapid growth phases to moderate growth phases is typically associated with the initiation of emergency measures such as lockdown and mass testing with contact tracing, which partially provides evidence that strict social distancing rules help slow down the virus growth and flatten the curve. Moreover, our analysis further indicates that compared to developed countries, most developing countries are still in the early stages of the pandemic and are generally less efficient in terms of controlling the spread of coronavirus, thus may need more international aids to help contain the epidemic.

Combining the SN-based change-point detection algorithm with a flexible extrapolation function, we further design a simple two-stage forecasting scheme for COVID-19. The proposed method is used to forecast the cumulative deaths in the U.S. and is found to deliver accurate prediction valuable to data-driven public health decision-making.

## 2. Methodology

In this section, we propose a novel SN-based method for change-point detection in model (1.1) that is robust against a wide range of temporal dependence. Specifically, an SN-based test statistic is first proposed for testing a single change-point alternative and then modified to consistently estimate the change-point. A multiple change-point estimation procedure is further developed by combining the proposed SN test with the NOT algorithm in Baranowski et al. (2019).

### 2.1. Testing for a single change-point

We start with a change-point testing problem where for model (1.1) we want to test the null hypothesis $H_0$ of no change-point against the alternative $H_a$ of one change-point:

$$H_0 : \boldsymbol{\beta}_1 = \cdots = \boldsymbol{\beta}_n = \boldsymbol{\beta} \quad \text{v.s.} \quad H_a : \boldsymbol{\beta}_t = \begin{cases} \boldsymbol{\beta}^{(1)}, & 1 \le t \le \tau \\ \boldsymbol{\beta}^{(2)}, & \tau + 1 \le t \le n, \end{cases} \quad \text{such that } \boldsymbol{\beta}^{(1)} \ne \boldsymbol{\beta}^{(2)},$$

where $\boldsymbol{\beta}_t = (a_t, b_t)^\top$, $\tau = \lfloor \kappa n \rfloor$ is an unknown change-point satisfying $\epsilon < \kappa < 1 - \epsilon$ for some $0 < \epsilon < 1/2$ and $\epsilon$ is the commonly used trimming parameter in the change-point analysis (see e.g. Andrews (1993)).

Throughout this paper, we operate under the following mild assumption of $\{u_t\}$, which covers a wide range of weakly dependent error process and is weaker than most existing literature where independence of $\{u_t\}$ is assumed.

**Assumption 2.1.** The error process $\{u_t\}$ is strictly stationary such that $\mathbb{E}(u_t) = 0$, $\mathbb{E}(u_t^4) < \infty$ and the long-run variance satisfies $\Gamma^2 = \lim_{n\to\infty} \mathrm{Var}(n^{-1/2} \sum_{t=1}^n u_t) \in (0, \infty)$. Denote $\{e_t\}$ as a sequence of i.i.d. random variables with zero mean and unit variance, we further assume that $\{u_t\}$ admits one of the following two representations:

(i). $u_t = \sum_{j=0}^\infty c_j e_{t-j}$ and $\sum_{j=0}^\infty |j c_j| < \infty$.

(ii). $u_t = G(\mathcal{F}_t)$ for some measurable function $G$ and $\mathcal{F}_t = (e_t, e_{t-1}, \ldots)$. For some $\chi \in (0, 1)$, $\|G(\mathcal{F}_k) - G(\{\mathcal{F}_{-1}, e_0', e_1, \ldots, e_k\})\|_4 = O(\chi^k)$ if $k \ge 0$ and 0 otherwise. Here $e_0'$ is an i.i.d. copy of $e_0$ and $\|X\|_4 = (\mathbb{E}(X^4))^{1/4}$ for a random variable $X$.

Assumption 2.1(i) is popular in the linear process literature to ensure the central limit theorem and the invariance principle. Assumption 2.1(ii) is basically equivalent to the geometric moment contracting condition for the nonlinear causal process in Wu and Shao (2004) and Wu (2005), which implies invariance principle.

Earlier works on this testing problem include Andrews (1993) and Bai and Perron (1998) where Lagrangian multiplier, Wald, likelihood ratio and $F$ statistics are considered. These tests typically require an estimator of the long-run variance (LRV) $\Gamma$ due to the unknown temporal dependence of the error process $\{u_t\}$. However, as pointed out in Shao and Zhang (2010), the size and power performance of these tests may depend crucially on the selection of various tuning parameters. In particular, if a data-driven bandwidth parameter is used for the estimation of LRV, an undesirable non-monotonic power phenomenon may occur; see Crainiceanu and Vogelsang (2007) and Shao and Zhang (2010). To avoid the bandwidth selection involved in the estimation of LRV, we instead adapt the idea of self-normalization in Shao (2010), which was originally proposed for inference of stationary time series and was generalized to change-point testing for piecewise stationary time series in Shao and Zhang (2010) and Zhang and Lavitas (2018). See Shao (2015) for a review of SN.

To proceed, we first introduce some notations. Given $\epsilon$, denote $h = \lfloor \epsilon n \rfloor$. For a vector $x$, denote the $l_2$ norm as $\|x\|_2$ and denote $x^{\otimes 2} = xx^\top$. Define $F(s) = (1, s)^\top$, for $1 \le i < j \le n$, we denote $\widehat{\boldsymbol{\beta}}_{i,j} = \left[\sum_{t=i}^j F(t/n)F(t/n)^\top\right]^{-1} \sum_{t=i}^j F(t/n)Y_t$ as the OLS estimator of $\boldsymbol{\beta}$ based on $\{Y_t\}_{t=i}^j$. For any $1 \le t_1 < k < t_2 \le n$, given the subsample $\{Y_t\}_{t=t_1}^{t_2}$ and a potential change-point $k$, we define a contrast statistic $D_n$ where

$$D_n(t_1, k, t_2) = \frac{(k - t_1 + 1)(t_2 - k)}{(t_2 - t_1 + 1)^{3/2}} (\widehat{\boldsymbol{\beta}}_{t_1,k} - \widehat{\boldsymbol{\beta}}_{k+1,t_2}). \tag{2.1}$$

Note that $D_n(t_1, k, t_2)$ is a normalized difference between the OLS estimates of $\boldsymbol{\beta}$ with pre-$k$ samples $\{Y_t\}_{t=t_1}^k$ and post-$k$ samples $\{Y_t\}_{t=k+1}^{t_2}$. Intuitively, a large $\max_{h \le k \le n-h} \|D_n(1, k, n)\|_2$ leads to the rejection of $H_0$. However, the asymptotic distribution of $D_n(1, k, n)$ depends on the unknown LRV of $\{u_t\}$, and as discussed before the accurate estimation of LRV is rather challenging and problematic in practice.

To bypass the problematic estimation of LRV, we utilize the self-normalization technique. Define $0 < \delta < \epsilon/2$ as a local trimming parameter, we define the self-normalizer $V_{n,\delta}(t_1, k, t_2) = L_{n,\delta}(t_1, k, t_2) + R_{n,\delta}(t_1, k, t_2)$ where

$$L_{n,\delta}(t_1, k, t_2) = \sum_{i=t_1+1+\lfloor n\delta \rfloor}^{k-2-\lfloor n\delta \rfloor} \frac{(i - t_1 + 1)^2 (k - i)^2}{(k - t_1 + 1)^2 (t_2 - t_1 + 1)^2} (\widehat{\boldsymbol{\beta}}_{t_1,i} - \widehat{\boldsymbol{\beta}}_{i+1,k})^{\otimes 2}, \tag{2.2}$$

**Table 2.1**
Simulated quantiles of $G$.

| $\epsilon$ | $\delta$ | $1-\alpha$ | | | | |
|---|---|---|---|---|---|---|
| | | 90% | 95% | 99% | 99.5% | 99.9% |
| 0.1 | 0.01 | 14.963 | 19.284 | 32.168 | 36.145 | 45.354 |
| | 0.02 | 24.959 | 32.727 | 53.645 | 64.898 | 92.982 |
| | 0.03 | 38.277 | 50.872 | 83.713 | 107.062 | 137.433 |
| | 0.04 | 54.569 | 76.244 | 116.497 | 144.437 | 182.786 |
| 0.2 | 0.01 | 4.656 | 5.905 | 9.691 | 12.037 | 14.148 |
| | 0.02 | 7.217 | 9.404 | 15.486 | 18.389 | 24.079 |
| | 0.03 | 10.526 | 13.767 | 23.060 | 26.758 | 36.388 |
| | 0.04 | 14.439 | 19.075 | 33.049 | 37.426 | 49.495 |

$$R_{n,\delta}(t_1,k,t_2)=\sum_{i=k+3+\lfloor n\delta\rfloor}^{t_2-1-\lfloor n\delta\rfloor}\frac{(i-1-k)^2(t_2-i+1)^2}{(t_2-t_1+1)^2(t_2-k)^2}(\widehat{\boldsymbol{\beta}}_{i,t_2}-\widehat{\boldsymbol{\beta}}_{k+1,i-1})^{\otimes 2}. \tag{2.3}$$

The local trimming parameter $\delta$ is introduced to make sure all the subsample estimates of $\boldsymbol{\beta}$ in the self-normalizer $V_{n,\delta}(t_1,k,t_2)$ are constructed with a subsample of size being a positive fraction of n, which is a technical condition necessary in our theoretical analysis. We later discuss the implication of the trimming parameters $(\epsilon,\delta)$.

Based on the contrast statistic $D_n(1,k,n)$ and the self-normalizer $V_{n,\delta}(1,k,n)$, we propose an SN-based test statistic $G_n$ for testing the single change-point alternative where

$$G_n=\max_{k\in\{h,\ldots,n-h\}}T_{n,\delta}(k),\quad T_{n,\delta}(k)=D_n(1,k,n)^\top V_{n,\delta}(1,k,n)^{-1}D_n(1,k,n). \tag{2.4}$$

Intuitively, due to the presence of the self-normalizer, the LRVs in $D_n(1,k,n)$ and $V_{n,\delta}(1,k,n)$ cancel out with each other, leading to a test statistic $G_n$ that is invariant to LRV. This phenomenon is made formal in Theorem 2.1.

Denote $\xrightarrow{\mathcal{D}}$ as convergence in distribution and $\mathbf{b}=\boldsymbol{\beta}^{(2)}-\boldsymbol{\beta}^{(1)}$. Define $Q(r)=\int_0^r F(s)F(s)^\top ds$ and $B_F(r)=\int_0^r F(s)dB(s)$ where $B(\cdot)$ is a standard Brownian motion. Theorem 2.1 states the asymptotic behavior of the SN test statistic $G_n$ under $H_0$ and $H_a$ respectively.

**Theorem 2.1.** *Suppose Assumption 2.1 holds. Let $G_n$ be defined in (2.4), we have*
*(i) under $H_0$, we have*

$$G_n\xrightarrow{\mathcal{D}}G(\epsilon,\delta):=\sup_{\eta\in(\epsilon,1-\epsilon)}D(\eta)^\top V_\delta(\eta)D(\eta), \tag{2.5}$$

*where* $D(\eta)=\eta(1-\eta)\Big\{Q(\eta)^{-1}B_F(\eta)-[Q(1)-Q(\eta)]^{-1}[B_F(1)-B_F(\eta)]\Big\}$ *and* $V_\delta(\eta)=L_\delta(\eta)+R_\delta(\eta)$ *with* $L_\delta(\eta)=\int_\delta^{\eta-\delta}\frac{r^2(\eta-r)^2}{\eta^2}\big\{Q(r)^{-1}B_F(r)-[Q(\eta)-Q(r)]^{-1}[B_F(\eta)-B_F(r)]\big\}^{\otimes 2}dr$, $R_\delta(\eta)=\int_{\eta+\delta}^{1-\delta}\frac{(r-\eta)^2(1-r)^2}{(1-\eta)^2}\times\big\{[Q(1)-Q(r)]^{-1}[B_F(1)-B_F(r)]-[Q(r)-Q(\eta)]^{-1}[B_F(r)-B_F(\eta)]\big\}^{\otimes 2}dr$.
*(ii) under $H_a$, given that $n\|\mathbf{b}\|_2^2\to L$, we have*

$$\lim_{L\to\infty}\lim_{n\to\infty}G_n=\infty,\quad \textit{in probability}.$$

Due to self-normalization, the limiting distribution $G(\epsilon,\delta)$ in (2.5) is pivotal and invariant to the LRV. The corresponding critical values can be easily obtained via simulation. Table 2.1 gives the $1-\alpha$ quantiles of $G(\epsilon,\delta)$ for some combinations of $(\epsilon,\delta)$ (based on 10000 replications). Note that the limiting null distribution $G(\epsilon,\delta)$ explicitly depends on the choice of $(\epsilon,\delta)$, thus the impact of trimming parameters $(\epsilon,\delta)$ is accounted for at the first order, in the same spirit of the fixed-$b$ asymptotics (Kiefer and Vogelsang, 2005). See also Zhou and Shao (2013). Throughout the paper, we set $(\epsilon,\delta)=(0.1,0.02)$.

Give that the null hypothesis $H_0$ is rejected, we estimate the change-point $\tau$ by $\widehat{\tau}=\arg\max_{k\in\{h,\ldots,n-h\}}T_{n,\delta}(k)$. The following theorem gives the consistency result of $\widehat{\kappa}=n^{-1}\widehat{\tau}$.

**Theorem 2.2.** *Under $H_a$, suppose Assumption 2.1 holds, and $n\|\mathbf{b}\|_2^2\to\infty$ as $n\to\infty$. Then, we have that for any $\eta>0$,*

$$\lim_{n\to\infty}\mathbb{P}(|\widehat{\kappa}-\kappa|<\eta)=1.$$

Theorem 2.2 allows a diminishing change size $\|\mathbf{b}\|_2$ with the sample size $n$ as long as $n\|\mathbf{b}\|_2^2\to\infty$. Note that no consistency result is provided in Shao and Zhang (2010) for the change-point location estimation, and our result seems to be the first formal attempt based on the SN technique. However, it is challenging to obtain an explicit rate of convergence for $\widehat{\tau}$ due to the complicated nature of the self-normalizer $V_{n,\delta}$ and we leave it for future investigation.

### 2.2. Multiple change-point estimation

To extend single change-point testing to multiple change-point estimation, the classical idea is to combine the change-point test with binary segmentation (BS). Although conceptually and computationally simple, it is well known that BS can cause severe power loss for detecting non-monotonic changes (Olshen et al., 2004), which is common in real data. Several variants of BS have been proposed to address this drawback, such as wild binary segmentation (WBS) (Fryzlewicz, 2014) and Narrowest-Over-Threshold (NOT) (Baranowski et al., 2019). Since NOT is shown to be superior to WBS, we combine the SN-based test with the NOT algorithm to estimate multiple change-points and name our algorithm SN-NOT.

The essential idea of SN-NOT is to compute the SN test on a large collection of random subsamples of $\{Y_t\}_{t=1}^n$ instead of the entire sample $\{Y_t\}_{t=1}^n$. With high probability, some subsamples will only contain a *single* change-point, where the SN test statistics are expected to exhibit large values, leading to the discovery of a change-point.

Denote $F_n^M = \{(s_i, e_i) : i = 1, \ldots, M\}$ as the set of $M$ random intervals such that each pair of integers $(s_i, e_i)$ are drawn uniformly from $\{1, \ldots, n\}$ and satisfy $1 \le s_i < e_i \le n$ and $e_i - s_i + 1 \ge 2h$. For each random interval $(s, e) \in F_n^M$, we calculate the SN test

$$G_{n,\delta}(s, e) = \max_{k \in \{s+h-1, \ldots, e-h\}} T_{n,\delta}(s, k, e), \quad T_{n,\delta}(s, k, e) = D_n(s, k, e) V_{n,\delta}(s, k, e)^{-1} D_n(s, k, e)^\top.$$

SN-NOT finds the narrowest interval $(s, e) \in F_n^M$ where the test statistic $G_{n,\delta}(s, e)$ exceeds a given threshold $\zeta_n$ and estimates the change-point as $\widehat{\tau} = \arg\max_{k \in \{s+h-1, \ldots, e-h\}} T_{n,\delta}(s, k, e)$. Note that for large $M$, with high probability there is only one change-point in this narrowest interval, which thus remedies the drawback of BS in detecting non-monotonic changes. Once a change-point $\widehat{\tau}$ is identified, SN-NOT then divides the sample into two subsamples accordingly and apply the same procedure on each of them. The process is implemented recursively until no change-point is detected. In addition to the advantage of detecting non-monotonic changes, SN-NOT broadens the applicability of the NOT algorithm itself by allowing for temporal dependence in the error process thanks to the self normalization technique.

The detailed implementation of SN-NOT is given in Algorithm 1. We propose to select the threshold $\zeta_n$ as follows. Generate $B$ sequences of i.i.d $\mathcal{N}(0, 1)$ random variables $\{\varepsilon_t^b\}_{t=1}^n$, $b = 1, \ldots, B$; for the $b$th sample, we calculate

$$\zeta_n^b = \arg\max_{i=1,\ldots,M} G_{n,\delta}(s_i, e_i), \quad b = 1, \ldots, B.$$

The threshold $\zeta_n$ is set as the 95% sample quantile of $\{\zeta_n^b\}_{b=1}^B$. Since the SN test statistic is asymptotically pivotal, this threshold is expected to well approximate the 95% quantile of the finite sample distribution of the maximum SN test statistic on the $M$ random intervals under null. Throughout this paper, we set $B = 1000$, $M = 300$.

**Algorithm 1:** SN-NOT

**Input**: Data $\{Y_t\}_{t=1}^n$, threshold $\zeta_n$, trimming size $d = \lfloor \delta n \rfloor$ and $h = \lfloor \epsilon n \rfloor$, random intervals $F_n^M$.
**Output**: Estimated number of change-points $\widehat{m}$ and estimated change-points set $\widehat{\boldsymbol{\tau}}$
**Initialization**: SN-NOT$(1, n, \zeta_n)$
**Procedure**: SN-NOT$(s, e, \zeta_n)$

**if** $e - s + 1 < 2h$ **then**
  Stop
**else**
  $\mathcal{M}_{(s,e)} := \{i : [s_i, e_i] \in F_n^M, [s_i, e_i] \subset [s, e], e_i - s_i + 1 \ge 2h\}$ ;
  **if** $\mathcal{M}_{(s,e)} = \varnothing$ **then**
    Stop
  **else**
    $\mathcal{O}_{(s,e)} := \left\{ i \in \mathcal{M}_{(s,e)} : G_{n,\delta}(s_i, e_i) > \zeta_n \right\}$;
    **if** $\mathcal{O}_{(s,e)} = \varnothing$ **then**
      Stop
    **else**
      $i^* = \arg\min_{i \in \mathcal{O}_{(s,e)}} |e_i - s_i + 1|$;
      $\tau^* = \arg\max_{k \in \{s_i^* + h - 1, \cdots, e_i^* - h\}} T_{n,\delta}(s_{i^*}, k, e_{i^*})$ ;
      $\widehat{\boldsymbol{\tau}} = \widehat{\boldsymbol{\tau}} \cup \tau^*$, $\widehat{m} = \widehat{m} + 1$;
      SN-NOT$(s, \tau^*, \zeta_n)$;
      SN-NOT$(\tau^* + 1, e, \zeta_n)$;
    **end**
  **end**
**end**

**Table 3.1**
Size and size-adjusted power for SN test and supLM test.

| $\alpha$ | $\rho$ | SN | | | | | supLM | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | −0.5 | −0.2 | 0 | 0.2 | 0.5 | −0.5 | −0.2 | 0 | 0.2 | 0.5 |
| | | Size | | | | | | | | | |
| 5% | $n = 100$ | 0.003 | 0.012 | 0.026 | 0.042 | 0.093 | 0.043 | 0.023 | 0.016 | 0.012 | 0 |
| 10% | | 0.008 | 0.028 | 0.053 | 0.091 | 0.160 | 0.091 | 0.064 | 0.047 | 0.035 | 0.018 |
| 5% | $n = 500$ | 0.022 | 0.033 | 0.036 | 0.045 | 0.057 | 0.049 | 0.042 | 0.032 | 0.030 | 0.020 |
| 10% | | 0.051 | 0.064 | 0.074 | 0.085 | 0.105 | 0.101 | 0.089 | 0.082 | 0.078 | 0.064 |
| 5% | $n = 1000$ | 0.040 | 0.045 | 0.045 | 0.045 | 0.049 | 0.042 | 0.037 | 0.036 | 0.034 | 0.025 |
| 10% | | 0.086 | 0.086 | 0.089 | 0.092 | 0.096 | 0.108 | 0.096 | 0.090 | 0.079 | 0.067 |
| | | Power | | | | | | | | | |
| 5% | $n = 100$ | 1 | 0.990 | 0.909 | 0.654 | 0.269 | 1 | 0.999 | 0.965 | 0.846 | 0.438 |
| 10% | | 1 | 1 | 0.983 | 0.879 | 0.531 | 1 | 1 | 0.996 | 0.925 | 0.587 |
| 5% | $n = 500$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.989 | 0.277 |
| 10% | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.998 | 0.568 |
| 5% | $n = 1000$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.906 |
| 10% | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.989 |

## 3. Simulation

In this section, we study the finite sample performance of the SN test in testing single change-point and the SN-NOT algorithm in detecting multiple change-points through numerical experiments. All results are reported based on 1000 replications.

### 3.1. Testing size and power

We generate the data from model (1.1) with sample size $n = 100, 500$ and 1000 respectively. For the size performance, we let $\boldsymbol{\beta} = (3, 0.05n)$ while for the power performance, we let $\boldsymbol{\beta}^{(1)} = (3, 0.06n)$ and $\boldsymbol{\beta}^{(2)} = (3 + 0.015n, 0.03n)$ with the change-point $\tau = n/2$. The error process $\{u_t\}$ is generated via an AR(1) model where $u_t = \rho u_{t-1} + e_t$, $e_t \overset{i.i.d.}{\sim} \mathcal{N}(0, (1-\rho^2)\sigma^2)$ with $\rho = 0, \pm 0.2, \pm 0.5$ and $\sigma = 0.15$.

For comparison, we also implement the supLM test defined in Andrews (1993) (using function `sctest` of the R package `strucchange`) with the same trimming parameter $\epsilon = 0.1$. The results are summarized in Table 3.1 at significance levels $\alpha = 5\%$ and 10%. It can be seen that when $n$ is small, both methods have distorted sizes. In particular, SN is prone to be conservative when $\rho$ is negative and oversized when $\rho$ is positive while supLM is undersized in all cases. As $n$ increases, we find that both tests tend to have more accurate sizes. For $n = 100$, supLM test has slightly higher power than SN test while for $n = 500$ and $n = 1000$, SN test beats supLM test under positive $\rho$. Note that both tests are more powerful under negative $\rho$.

### 3.2. Multiple change-point estimation

We examine the numerical performance of SN-NOT by considering the following DGP with $n = 100$:

$$Y_t = \begin{cases} 3 + 3.2(t/n) + u_t, & 1 \le t \le 20, \\ 5.8 + 1.8(t/n) + u_t, & 21 \le t \le 40, \\ 9.8 + 0.8(t/n) + u_t, & 41 \le t \le 70, \\ 15.05 + 0.05(t/n) + u_t, & 71 \le t \le 100. \end{cases}$$

The error process $\{u_t\}$ is generated via an AR(1) model where $u_t = \rho u_{t-1} + e_t$, $e_t \overset{i.i.d.}{\sim} \mathcal{N}(0, (1-\rho^2)\sigma^2)$ with $\rho = 0, \pm 0.2, \pm 0.5$ and $\sigma = 0.15$. For comparison, we also implement the multiple change-point detection procedure proposed in Bai and Perron (1998) (denoted as BP hereafter), which is the most widely used detection algorithm allowing for temporal dependence in the error term of model (1.1). BP is implemented using function `breakpoints` of the R package `strucchange`.

To assess the accuracy of change-point estimation, we define the Hausdorff distance between two sets. Denote the set of true change-points as $\boldsymbol{\tau}_o$ and the set of estimated change-points as $\widehat{\boldsymbol{\tau}}$, we define $d_1(\boldsymbol{\tau}_o, \widehat{\boldsymbol{\tau}}) = \max_{\tau_1 \in \widehat{\boldsymbol{\tau}}} \min_{\tau_2 \in \boldsymbol{\tau}_o} |\tau_1 - \tau_2|$ and $d_2(\boldsymbol{\tau}_o, \widehat{\boldsymbol{\tau}}) = \max_{\tau_1 \in \boldsymbol{\tau}_o} \min_{\tau_2 \in \widehat{\boldsymbol{\tau}}} |\tau_1 - \tau_2|$, where $d_1$ measures the over-segmentation error of $\widehat{\boldsymbol{\tau}}$ and $d_2$ measures the under-segmentation error of $\widehat{\boldsymbol{\tau}}$. The Hausdorff distance is then defined as $d_H(\boldsymbol{\tau}_o, \widehat{\boldsymbol{\tau}}) = \max(d_1(\boldsymbol{\tau}_o, \widehat{\boldsymbol{\tau}}), d_2(\boldsymbol{\tau}_o, \widehat{\boldsymbol{\tau}}))$. In addition,

**Table 3.2**
Estimation results for SN-NOT and BP.

| $\rho$ | SN-NOT | | | | | BP | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | −0.5 | −0.2 | 0 | 0.2 | 0.5 | −0.5 | −0.2 | 0 | 0.2 | 0.5 |
| ARI | 0.844 | 0.852 | 0.849 | 0.828 | 0.784 | 0.863 | 0.852 | 0.840 | 0.805 | 0.714 |
| $d_1$ | 4.817 | 3.846 | 3.953 | 4.765 | 6.049 | 2.854 | 3.176 | 3.379 | 3.970 | 4.837 |
| $d_2$ | 2.949 | 3.170 | 3.574 | 3.964 | 6.032 | 2.854 | 3.252 | 3.915 | 5.605 | 10.457 |
| $d_H$ | 4.830 | 3.877 | 4.141 | 4.960 | 7.152 | 2.854 | 3.252 | 3.915 | 5.605 | 10.457 |
| $\widehat{m}=3$ | 0.902 | 0.955 | 0.950 | 0.922 | 0.808 | 1 | 0.989 | 0.930 | 0.775 | 0.337 |
| $\lvert\widehat{m}-3\rvert=1$ | 0.098 | 0.045 | 0.050 | 0.078 | 0.186 | 0 | 0.011 | 0.069 | 0.198 | 0.402 |
| $\lvert\widehat{m}-3\rvert>1$ | 0 | 0 | 0 | 0 | 0.006 | 0 | 0 | 0.001 | 0.027 | 0.261 |

**Table 4.1**
Summary of estimated models (1.1) for cumulative confirmed cases in 8 representative countries.

| Country | Start | $n$ | No.CP | 1st CP ($S_1$) | 2nd CP ($S_2$) | Latest CP ($S_{\widehat{m}+1}$) | $\widehat{\rho}$ |
|---|---|---|---|---|---|---|---|
| United States | Feb-22 | 96 | 5 | Mar-04 (0.113) | Mar-24 (0.292) | May-09 (0.015) | 0.492 |
| Brazil | Mar-09 | 80 | 2 | Mar-25 (0.301) | Apr-12 (0.129) | Apr-12 (0.066) | 0.438 |
| Russia | Mar-12 | 77 | 4 | Apr-05 (0.218) | Apr-21 (0.146) | May-17 (0.028) | 0.573 |
| United Kingdom | Mar-01 | 88 | 5 | Mar-20 (0.254) | Mar-29 (0.181) | May-12 (0.011) | 0.575 |
| Spain | Feb-28 | 90 | 5 | Mar-14 (0.359) | Mar-27 (0.176) | May-01 (0.004) | 0.611 |
| Italy | Feb-23 | 95 | 6 | Mar-09 (0.289) | Mar-22 (0.151) | May-18 (0.003) | 0.616 |
| India | Mar-05 | 83 | 5 | Mar-24 (0.159) | Apr-02 (0.142) | May-09 (0.052) | 0.375 |
| South Korea | Feb-06 | 112 | 6 | Feb-18 (0.022) | Mar-03 (0.360) | May-08 (0.002) | 0.749 |

we report the adjusted Rand index (ARI) which measures the similarity between two partitions of the same observations. Roughly speaking, a higher ARI (with the maximum value of 1) means more accurate change-point estimation. For the definition and detailed discussions of ARI, we refer to Hubert and Arabie (1985).

Table 3.2 summarizes the numerical result where we report ARI, $d_1$, $d_2$, $d_H$ and the frequency of $|\widehat{m}-m_o|$ for SN-NOT and BP. It can be seen that SN-NOT is overall better than BP in terms of ARI, $d_H$ and the estimated number of change-points when $\rho \geq 0$. This finding suggests using SN-NOT could be more advantageous for analyzing COVID-19 data, which exhibit positive temporal dependence (see the last column of Table 4.1). For applications where negatively correlated error is expected, BP could be a better choice.

## 4. Analysis for cumulative confirmed cases and deaths of COVID-19

In this section, based on the proposed SN-NOT algorithm, we provide detailed in-sample analysis of the cumulative confirmed cases (Sections 4.2–4.3) and deaths (Section 4.4) of COVID-19 (in log scale) in 30 major countries.

### 4.1. Data and method

We focus on G20 (with 19 sovereign countries[4]) and 11 other countries leading the total infected cases as of May 27, 2020, including Australia (AUS), Argentina (ARG), Belgium (BEL), Brazil (BRA), Canada (CAN), Chile (CHI), China (CHN), France (FRA), Germany (GER), India (IND), Indonesia(INA), Iran (IRI), Italy (ITA), Japan (JPN), Mexico (MEX), Netherlands (NED), Pakistan (PAK), Peru (PER), Portugal (POR), Qatar(QAT), Russia (RUS), Saudi Arabia (KSA), Spain (ESP), South Africa (RSA), South Korea (ROK), Sweden (SWE), Switzerland (SUI), Turkey (TUR), United Kingdom (GBR), United States (USA).

We obtain the data from https://ourworldindata.org/coronavirus-source-data maintained by "Our World in Data", where cumulative measures such as confirmed cases and deaths are updated daily for each nation. For each country, the logarithm of cumulative confirmed cases (or deaths) $\{Y_t\}$ starts on the date when the cumulative cases (or deaths) exceeded 20 and ends on May 27.

We study the cumulative confirmed cases and deaths (in log scale) of each country via the piecewise linear trend model (1.1), where given $\{Y_t\}$, the change-points $(\tau_1, \ldots, \tau_{\widehat{m}})$ are estimated by the SN-NOT algorithm. An OLS is then used to recover the linear model for the $i$th estimated segment $\{Y_t\}_{t=\widehat{\tau}_{i-1}+1}^{\widehat{\tau}_i}$, $i = 1, 2, \ldots, \widehat{m}+1$. With a slight abuse of notation, denote $\widehat{b}_i$ as the estimated slope for the $i$th segment. We define the normalized slope $S_i = \widehat{b}_i/n$ for each segment. As can be seen from (1.1), the normalized slope $S_i$ measures $\mathbb{E}[Y_{t+1} - Y_t]$ for the $i$th segment, which can be interpreted as the "log-return" and measures the daily growth rate of the cumulative confirmed cases (or deaths) in the original scale.

Methodologically speaking, for cumulative confirmed cases, the piecewise linearity allows us to assess the growth rate of the coronavirus at any given time and further facilitates short-term forecast. In particular, the estimated slope $S_i$ of each segment indicates the pace of the growth rate during the corresponding period. Moreover, by comparing the slope before and after each change-point, we can quantitatively assess the changes in growth rate, which partially measure the effectiveness of policies taken by the government.

[4] G20 is an international forum for the governments and central bank governors from 19 countries and the European Union. We will view members of the European Union as individual countries because the responses to COVID-19 usually come from the national level.

### 4.2. Detailed analysis of cumulative confirmed cases in 8 representative countries

We first conduct a detailed case study for eight representative countries that either lead confirmed cases (the U.S., Brazil, Russia, and India) in the corresponding continent or receive most media attention (the U.K., Spain, Italy, and South Korea).

Table 4.1 summarizes the detailed estimation result for each country (in descending order of the cumulative confirmed cases), where we report the starting date of the series, length of the series $n$, the estimated number of change-points, dates of the first, second and latest estimated change-point. The first ($S_1$), the second ($S_2$) and the current normalized slope ($S_{\widehat{m}+1}$) are also presented. In addition, we report the lag-1 sample autocorrelation $\widehat{\rho}$ of the error process. From the table, we can see all of these countries have been affected by the coronavirus for more than two months. The average length of segments between two adjacent change-points is around 13–20 days, indicating that the spread rate can be relatively steady for a window of 2–3 weeks. The latest change-point for most countries appeared in May except for Brazil. We also note that the current normalized slopes (i.e. growth rate) vary considerably across countries with comparably large values in Brazil and India. Meanwhile, the lag-1 sample autocorrelation $\widehat{\rho}$ are all positive, which suggests the use of SN-NOT instead of BP as discussed in Section 3.2. In Figures C.1 and C.2 of the supplementary material, we further plot the lag-1 to lag-30 ACF and PACF of the residuals, which rules out the scenario of long memory and supports the validity of Assumption 2.1.

Fig. 4.1 visualizes the estimated piecewise linear models for the eight countries, which gives a more direct perception of how the growth rate changes over time. Note that the U.S. and South Korea are the only two countries that witnessed an increase in the slope after the first change-point. For the U.S., the first change-point is March 4, one day after the first confirmed case appeared in New York. Since then, the pandemic underwent an outbreak in the New York state, which has been the leading state in the U.S. in terms of infected cases. The second change-point appeared on March 24, after which the slope began to drop. This is also noteworthy as on March 20, the U.S. began barring entry of foreign nationals who had traveled to 28 European countries within the past 14 days. While in South Korea, after February 18, the infected cases increased drastically, and the slope dropped after March 3. We find that the first change-point is the day when the first super-spreader in South Korea was diagnosed.[5] The second change-point, March 3, is when the drive-through testing was made widely available to Korean citizens.

The growth rate decreased after the first change-point in other countries. For the U.K., the first and second change-points are quite close. In particular, we find the U.K. governments gradually increased the restrictions on freedom of movement for the general public between these two change-points (March 20 and March 29). This could help explain why both change-points are associated with significant drops in the virus growth rate. In addition, we find that Italy extended the quarantine lockdown from region-focused to nationwide on March 10, one day after the first estimated change-point. For Spain, the first change-point is estimated as March 14, which is one day after Spain declared the nationwide state of emergency. Similar to Italy, the slopes dropped drastically after the first change-point. Generally speaking, the first or second change-point of these countries are closely associated with the date when local or nationwide interventions from the governments were initiated. These countries typically transition from a rapid growth phase to a moderate growth phase after the first or second change-point. This may serve as evidence that government intervention such as lockdown and massive testing could effectively slow down the spread of the coronavirus.

From Fig. 4.1, we also find the situations in Brazil, Russia and India rather somber, as of May 27. Russia is still transitioning from the rapid growth phase to the moderate growth phase, while the fast growing trend in Brazil has not changed since April 12. Even though Brazil managed to bring down the slope by a significant amount at the first change-point on March 25, it seemed the right-wing government took few follow-up effective measures. The situation in India is also grim where the decreases of growth rate at the first and second change-points are quite small and the current growth rate is still high, suggesting that stricter measures to be taken. In summary, these three countries still have a long way to go in terms of slowing down the spread of COVID-19.

### 4.3. Analysis of cumulative confirmed cases in 30 countries

We further extend the scope of analysis to 30 countries to obtain a relatively complete picture of the pandemic situations around the world. Specifically, we conduct a comparative study based on two important quantities: the maximum normalized slope and the current normalized slope, which are estimated by $S_{max} = \max_{1\le i\le \widehat{m}+1} n^{-1}\widehat{b}_i$ and $S_{cur} = n^{-1}\widehat{b}_{\widehat{m}+1}$ respectively. Combined together, the two measures allow us to obtain an overall picture of the phase when the virus transmitted fastest and the current situation in each country. In particular, $S_{max}$ provides information on the growth rate at the early stage of the pandemic for a particular country. In this phase, often no government regulations are imposed so it depicts the worst scenario if no emergency measure is taken. $S_{cur}$ gives the ongoing epidemic growth rate and could help make predictions in the short run.

In Fig. 4.2, we plot $S_{max}$ against $S_{cur}$ for each country. Note that by their relative positions in Fig. 4.2, the 30 countries can be roughly grouped into three clusters: East Asian countries and Australia, European and North American countries

[5] A member of the Shincheonji religious organization was diagnosed as 31st case in Daegu, see https://foreignpolicy.com/2020/02/27/coronavirus-south-korea-cults-conservatives-china/.

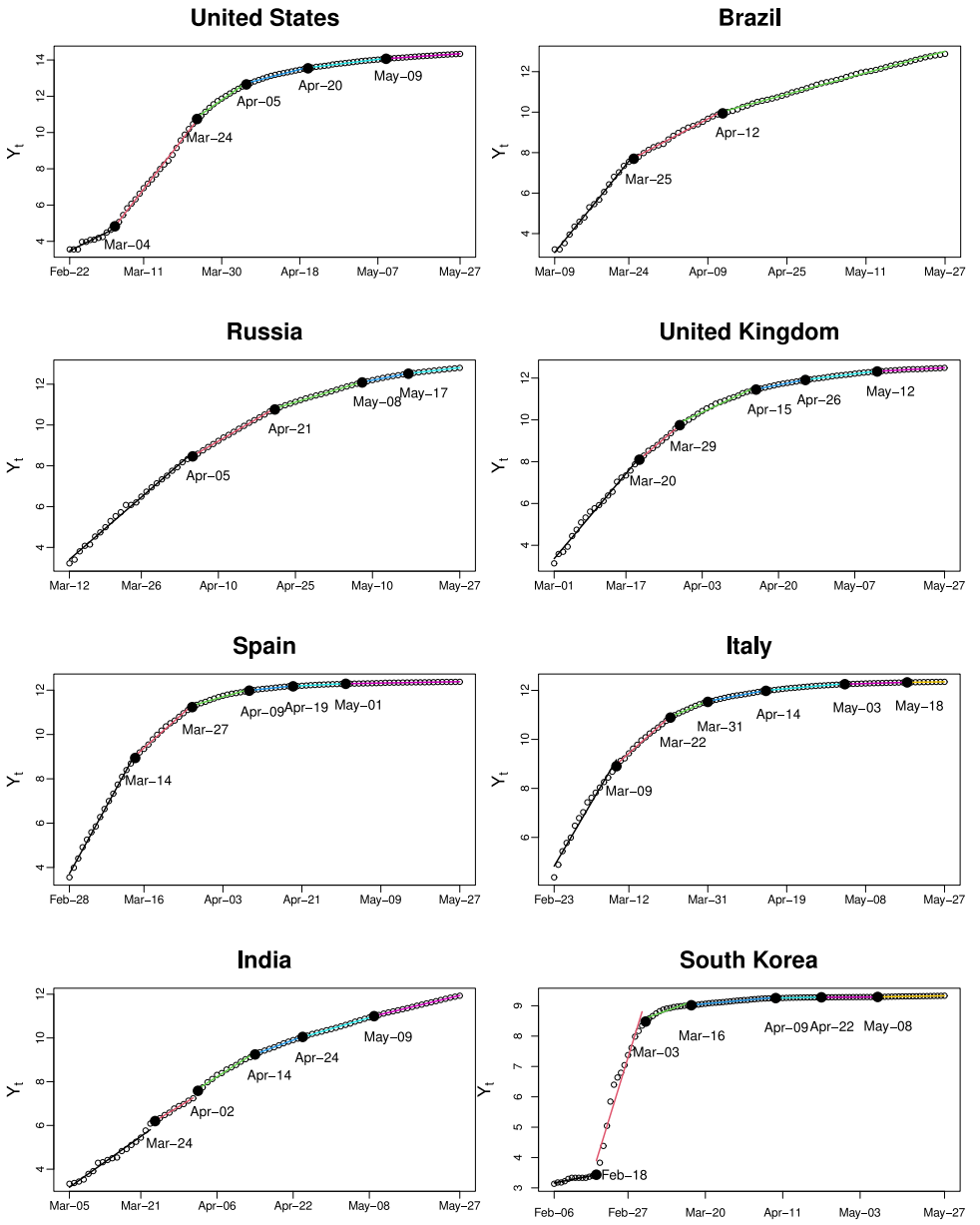


**Fig. 4.1.** Estimated piecewise linear trend for cumulative confirmed cases in 8 representative countries.

and other developing countries. We find that countries within the same cluster tend to have similar current growth rate. China, South Korea, and Australia are among the best with $S_{cur}$ close to zero. Most European and North American countries are in the second tier while countries in continental Europe generally have slower ongoing virus growth than the U.K., the U.S. and Canada. The only exceptions are Sweden and Russia. In fact, Sweden adopted a different strategy than other countries in that no lockdown has been imposed by the government and large parts of its society remain open. Note that Fig. 4.2 does not take the time effect into account, thus the cluster along the horizontal direction may also be attributed to the cluster of similar eruption time of the virus. This could help explain why Russia is closer to developing countries and why Latin American countries have the largest $S_{cur}$.

To take the time factor into consideration, in Fig. 4.3, we plot the ratio $S_{cur}/S_{max}$ against the days in between (i.e. $\tau_{cur} - \tau_{max}$ with $\tau_{max}$ as the start date for the segment with the largest slope and $\tau_{cur} = \tau_{\widehat{m}}$ as the latest change-point), which allows us to further understand how the growth rate changes from its peak to the current status with time. Horizontally speaking, for the same ratio $S_{cur}/S_{max}$, if country A is to the left of country B, then A acts faster than B in bringing down the virus growth from its peak value. Vertically speaking, for the same time length $\tau_{cur} - \tau_{max}$, if A is below B, then A is more effective than B in reducing the growth rate.

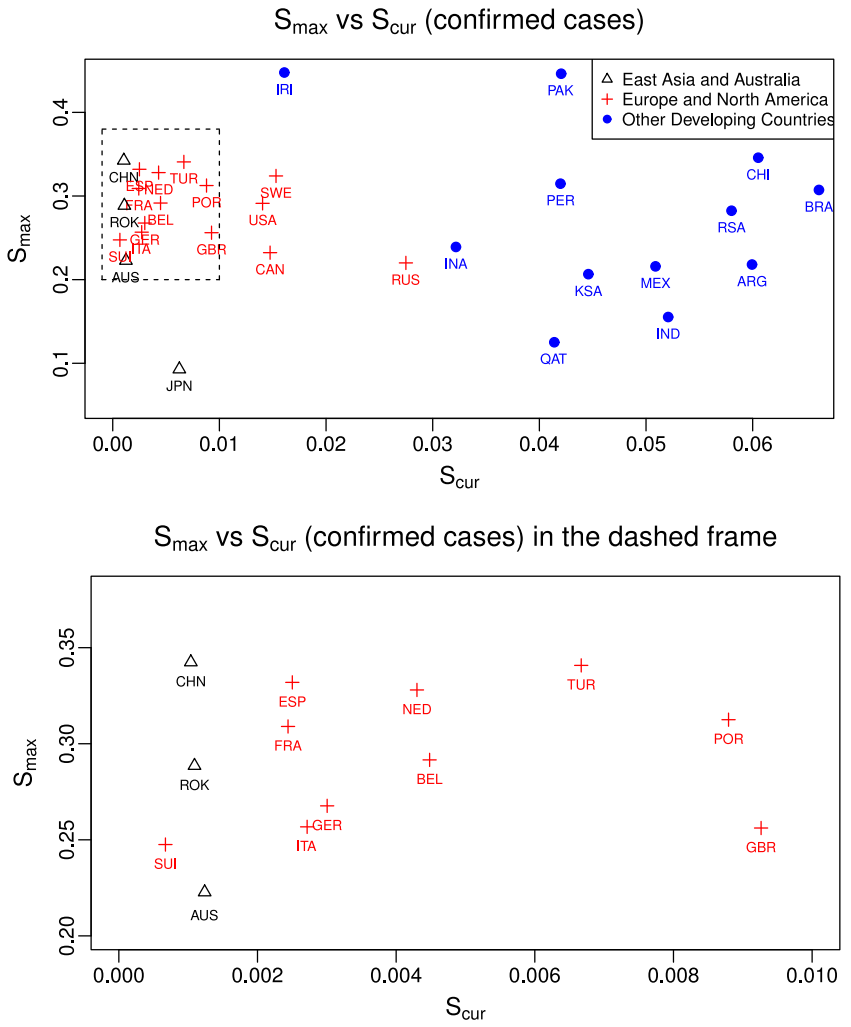


**Fig. 4.2.** Plot of maximum normalized slope $S_{max}$ and normalized slope after the latest change-point $S_{cur}$ for cumulative confirmed cases of each country. Black △: East Asian Countries and Australia; red +: European and North American Countries; blue •: Other developing countries.

We again find that most European and North American countries tend to share similar characteristics. The growth rates in the current phases for these countries are less than one-tenth of their peak value, and it took them about two to three months to achieve that. From the lower panel in Fig. 4.3, we find that South Korea, China and Australia outperform other countries as the ratios were brought to near zero in around 65 days. Again, we find that continental European countries (except Russia and Sweden) perform better than U.S, Canada and U.K.

Most developing countries are on the top-left of the plot, suggesting that they are still in the relatively early stage of the pandemic and the situation has not improved much since the beginning of the outbreak. In addition, we find Latin American countries, such as Mexico, Brazil, Chile, and Peru, tend to cluster. Given their geographical proximity, this is not a surprise. We note that developing countries tend to be less efficient in slowing the spread of COVID-19. For example, with roughly the same amount of time, the ratios in India and Argentina are three times larger than developed countries. In summary, more caution and attention should be given to the epidemic in developing countries as they may need more international aids compared to the developed countries.

### *4.4. Analysis of cumulative deaths in 30 countries*

Based on the same methodology, we analyze cumulative deaths in the 30 countries. Note that unlike confirmed cases, public health interventions naturally have a longer lagged effect on coronavirus-related deaths, as severe symptoms may not develop immediately upon infection. Thus, we believe a change-point analysis on cumulative confirmed cases should be preferred in terms of quantifying the effectiveness of emergency policies. Additionally, the criteria for certifying deaths

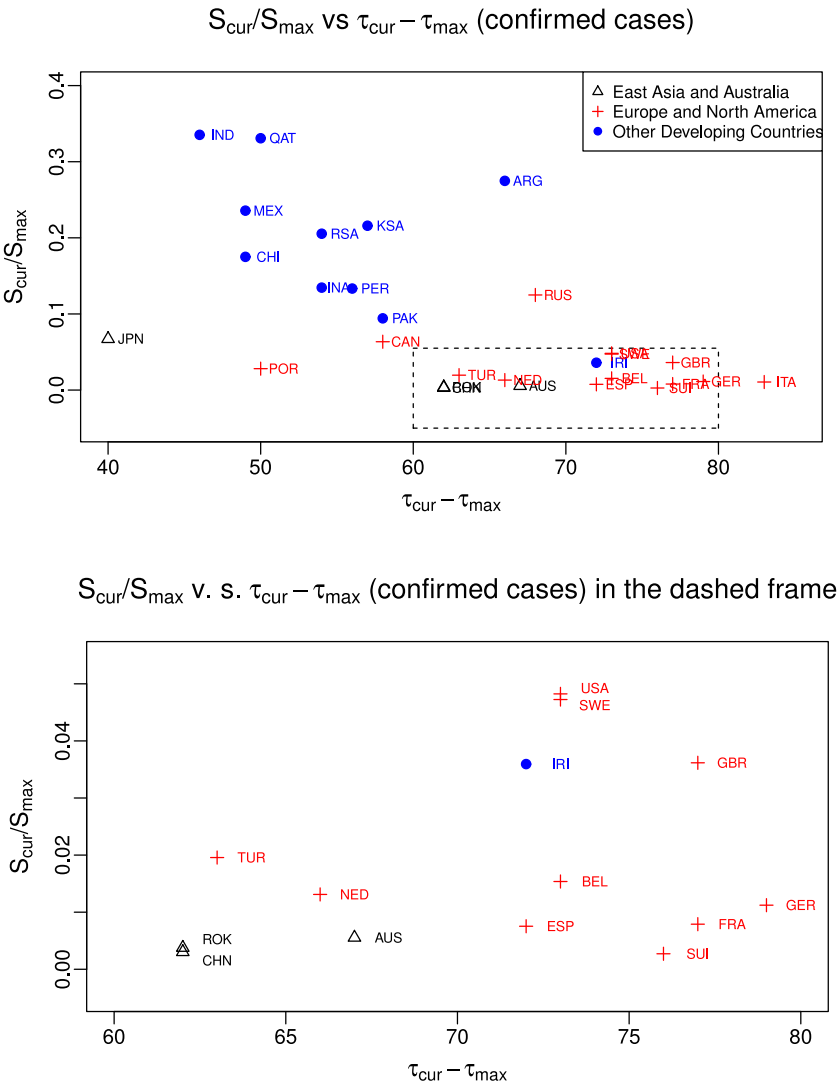


**Fig. 4.3.** Plot of ratio between the current normalized slope $S_{cur}$ and maximum normalized slope $S_{max}$ against days from the start date for the segment with the largest slope to the start date for the latest segment for cumulative confirmed cases of each country. Black △: East Asian Countries and Australia; red +: European and North American Countries; blue •: Other developing countries.

due to COVID-19 vary from nation to nation, thus comparative analysis across countries should be interpreted with caution.

Table 4.2 summarizes the detailed estimation result for cumulative deaths in the eight representative countries. Notably, for each country, the estimated number of change-points for deaths is smaller than or equal to that for cumulative confirmed cases in Table 4.1. This is intuitive as the history of cumulative deaths is shorter and number of deaths largely depend on infections (with a lag). Note that the duration between the starting date and the first change-point for cumulative deaths is around 2–3 weeks, which is consistent with that for confirmed cases in Table 4.1. The same phenomenon also applies to the duration between the first and second change-points. This consistency in part confirms the validity of the change-point estimation results and indicates a 2–3 weeks response lag between changes in growth rate of infections and changes in growth rate of deaths. We note that Italy and Spain have the highest growth rate of cumulative deaths before the first change-point, which highlights the extreme importance of "flattening the curve", as it is known that the exponential surge of coronavirus cases exhausted the public health system in the two countries at the early stage of the pandemic.

Fig. 4.4 further plots the estimated piecewise linear models for cumulative deaths in the eight countries. The pattern exhibited by each country is largely consistent with its pattern in Fig. 4.1, except for South Korea. Note that the start date of the cumulative death curve in South Korea is almost 30 days later than the start date of the cumulative confirmed cases, which partially explains the different pattern around its first change-point.

**Table 4.2**
Summary of estimated models (1.1) for cumulative deaths in 8 representative countries.

| Country | Start | $n$ | No.CP | 1st CP ($S_1$) | 2nd CP ($S_2$) | Latest CP ($S_{\widehat{m}+1}$) | $\widehat{\rho}$ |
|---|---|---|---|---|---|---|---|
| United States | Mar-09 | 80 | 5 | Mar-26 (0.229) | Apr-09 (0.195) | May-15 (0.012) | 0.556 |
| Brazil | Mar-23 | 66 | 2 | Apr-11 (0.195) | May-01 (0.086) | May-01 (0.056) | 0.696 |
| Russia | Apr-02 | 56 | 3 | Apr-22 (0.149) | May-03 (0.093) | May-11 (0.043) | 0.366 |
| United Kingdom | Mar-15 | 74 | 4 | Apr-03 (0.254) | Apr-19 (0.099) | May-15 (0.008) | 0.657 |
| Spain | Mar-10 | 79 | 5 | Mar-27 (0.307) | Apr-05 (0.121) | May-15 (−0.000[a]) | 0.507 |
| Italy | Feb-29 | 89 | 6 | Mar-14 (0.305) | Mar-22 (0.167) | May-08 (0.005) | 0.287 |
| India | Mar-29 | 60 | 3 | Apr-13 (0.179) | May-06 (0.070) | May-20 (0.039) | −0.012 |
| South Korea | Mar-02 | 87 | 4 | Mar-13 (0.100) | Mar-30 (0.0518) | May-07 (0.003) | 0.363 |

[a]Spain revised its death toll downwards on May 25, see https://english.elpais.com/society/2020-05-26/spanish-health-ministry-lowers-coronavirus-death-toll-by-nearly-2000.html.

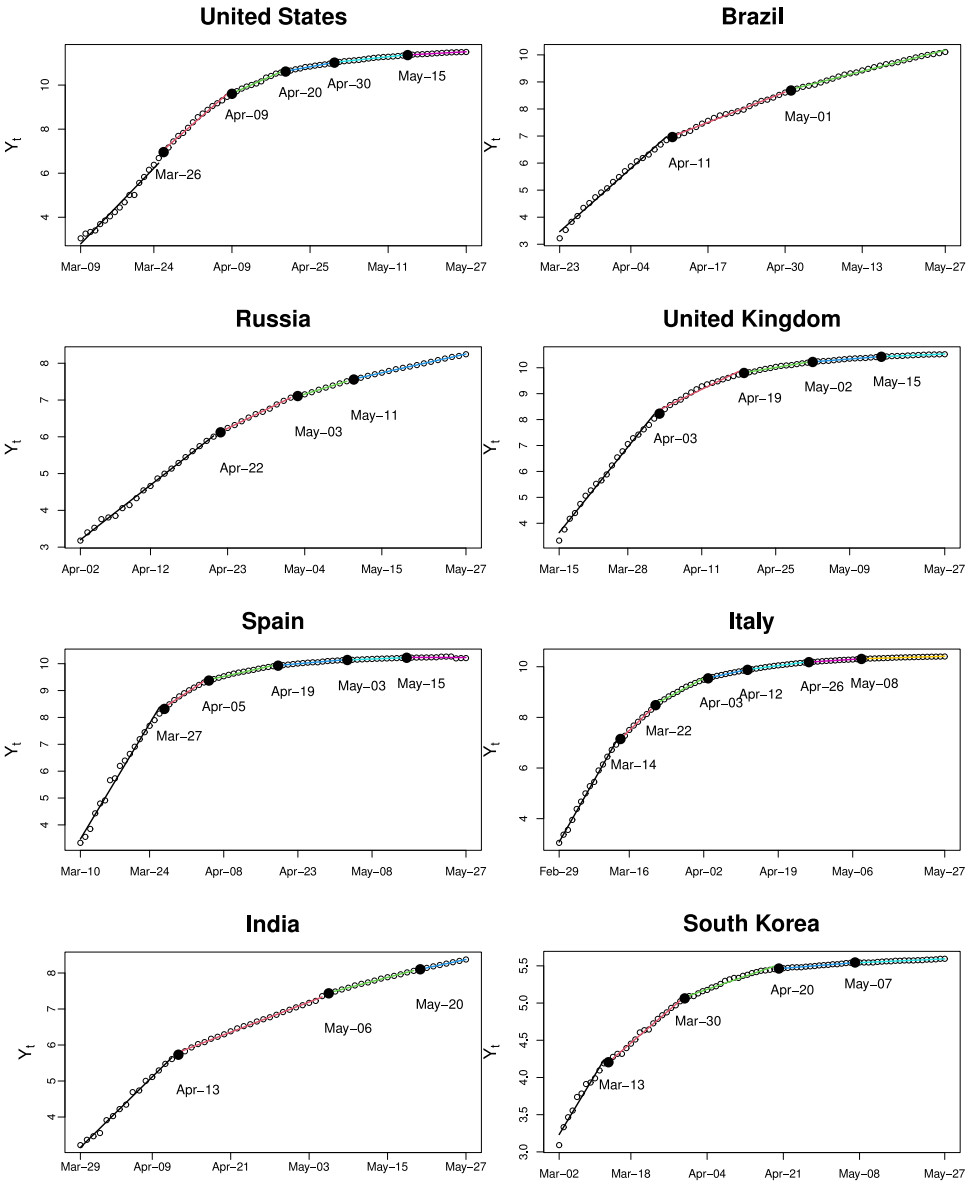


**Fig. 4.4.** Estimated piecewise linear trend for cumulative deaths in 8 representative countries.

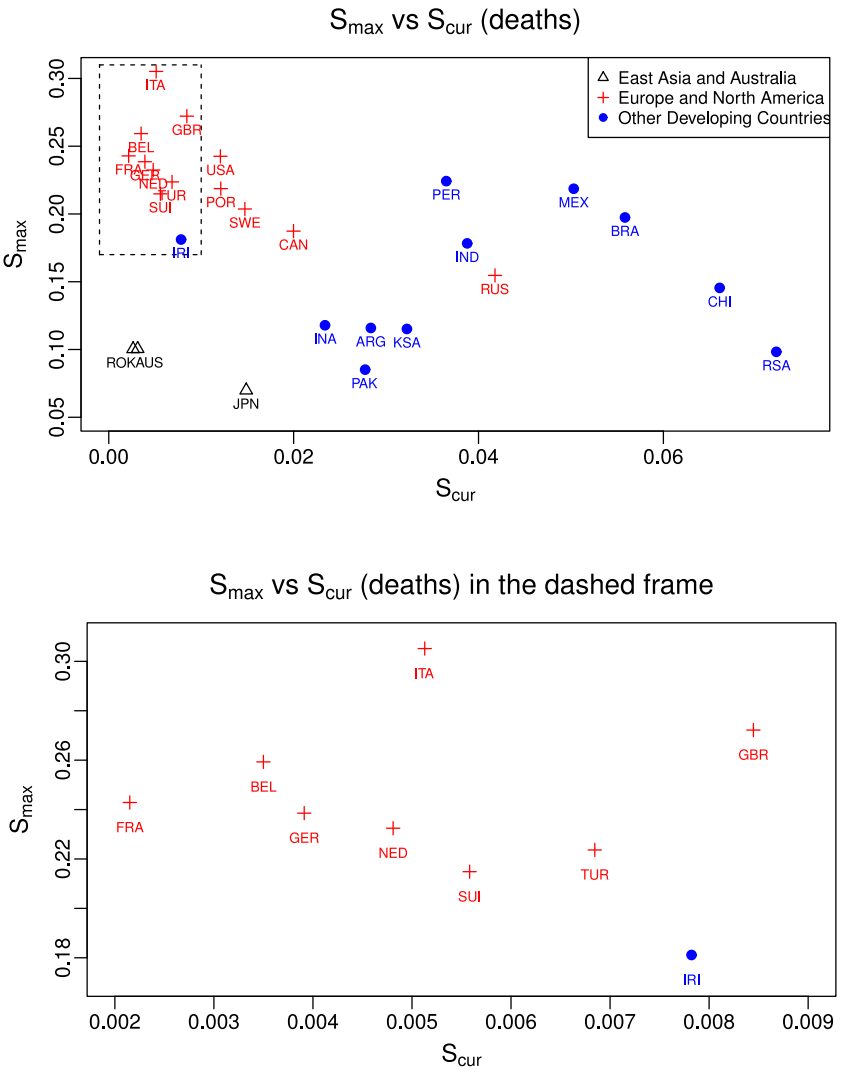


**Fig. 4.5.** Plot of maximum normalized slope $S_{max}$ and normalized slope after the latest change-point $S_{cur}$ for cumulative deaths of each country. Black △: East Asian Countries and Australia; red +: European and North American Countries; blue •: Other developing countries.

We further conduct a comparative analysis for cumulative deaths in 30 countries. We exclude China, Spain and Qatar in the analysis as the death tolls were either revised or unavailable.[6] Fig. 4.5 plots $S_{max}$ against $S_{cur}$ for each country. Similar to the results for confirmed cases in Fig. 4.2, European and North American countries tend to cluster while developing countries generally have higher ongoing growth rates $S_{cur}$.

Note that South Korea and Australia deliver the best responses with small $S_{max}$ and near-zero $S_{cur}$ for cumulative deaths. However, it is unexpected to see that western developed countries, such as Italy and the U.K., experience the largest maximum growth rate. Since the maximum growth rate always takes place in the first segment of the cumulative death curve, it indicates that the coronavirus may take these countries by surprise and the health systems may not be well prepared for the flood of coronavirus patients in the early stage of the pandemic. Another notable pattern is that Latin American countries tend to have larger values in both maximum and current growth rates than other developing countries, signaling the possibility of Latin America becoming the next epicenter of the COVID-19 pandemic.

Fig. 4.6 plots $S_{cur}/S_{max}$ against $\tau_{cur} - \tau_{\max}$ for cumulative deaths in each country, where the observed patterns are similar to the ones for cumulative confirmed cases in Fig. 4.6. Specifically, developing countries again tend to be less efficient in slowing the spread of COVID-19, where with roughly the same amount of time, the ratios $S_{cur}/S_{max}$ in developing countries are noticeably larger than developed countries.

6 China revised its death toll upwards on April 17, see https://www.nytimes.com/2020/04/17/world/asia/china-wuhan-coronavirus-death-toll.html. The death toll is not available for Qatar.

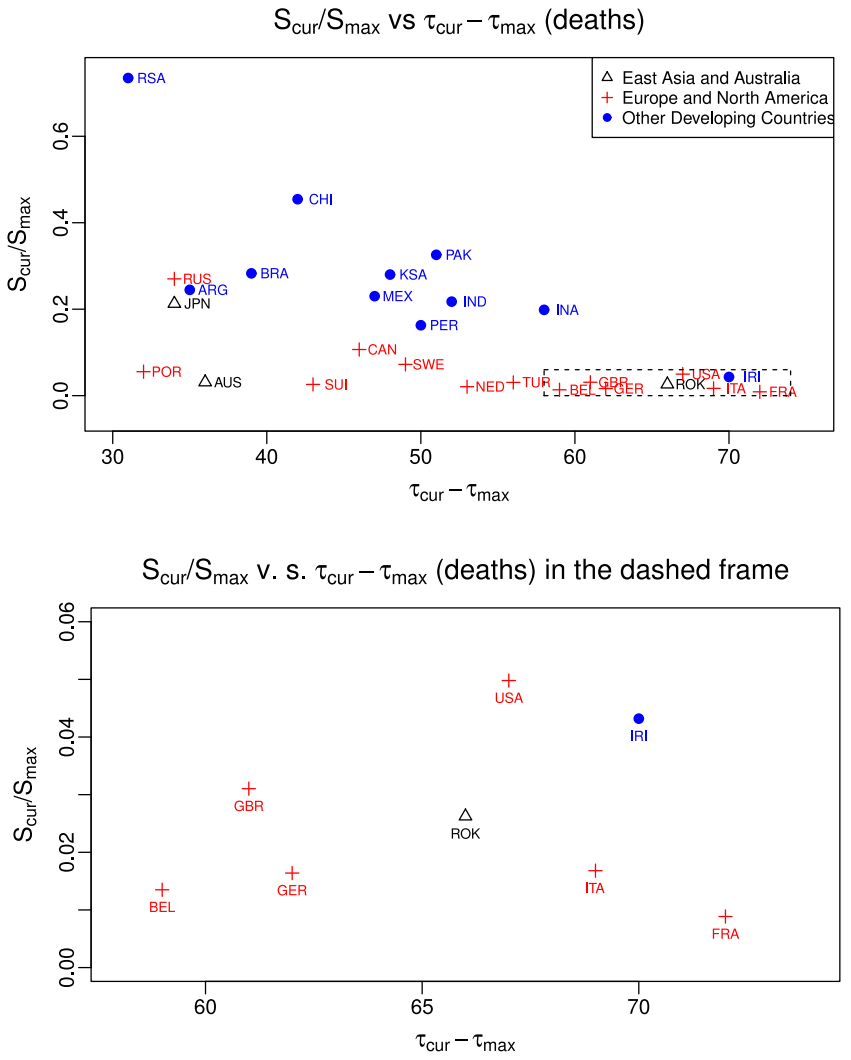


**Fig. 4.6.** Plot of ratio between the current normalized slope $S_{cur}$ and maximum normalized slope $S_{max}$ against days from the start date for the segment with the largest slope to the start date for the latest segment for cumulative deaths of each country. Black △: East Asian Countries and Australia; red +: European and North American Countries; blue •: Other developing countries.

## 5. SN-NOT based forecast for cumulative deaths

As stated by the Centers for Disease Control and Prevention (CDC),[7] accurate forecast of COVID-19 deaths is critical for public health decision-making, as it projects the likely impact of coronavirus to health systems in coming weeks and helps government officials develop data-driven public health policies for controlling the pandemic.

In Section 5.1, we propose a simple and intuitive forecasting scheme for cumulative deaths due to COVID-19 by combining SN-NOT with a flexible extrapolation function. In Section 5.2, we further demonstrate its promising performance in predicting cumulative deaths in the U.S.

### 5.1. Method

As suggested by the analysis in Section 4, the spread of coronavirus typically experiences several different stages due to external interventions. While a sophisticated epidemiology model based on differential equations may manage to take into account information about interventions and characterize the entire cumulative death curve, a more natural (and simpler) solution from the change-point aspect is to first segment the time series into periods with relatively stable behavior and

[7] https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html#why-forecasting-critical.

then generate forecast based on observations in the last segment, see for example, Pesaran and Timmermann (2002) and Bauwens et al. (2015).

Following this idea, we propose an SN-NOT based two-stage approach for cumulative deaths prediction. Specifically, in the first stage, given the cumulative deaths (in log scale) $\{Y_t\}_{t=1}^n$, a piecewise linear trend model is estimated via SN-NOT with change-points $\widehat{\boldsymbol{\tau}}$. In the second stage, a flexible function $f(t)$ is fitted on the last segment $\{Y_t\}_{t=\widehat{\tau}_{\widehat{m}}+1}^n$ with the assumption that $\mathbb{E}(Y_t) = f(t)$ and the $k$-day ahead forecast for cumulative deaths can be readily made via extrapolation of $\widehat{f}(t)$.

Note that the purpose of the first stage (in-sample) change-point analysis is to identify the most recent segment where $\{Y_t\}_{t=1}^n$ exhibits relatively stable behavior and thus facilitates the second stage (out-of-sample) forecast. As demonstrated in Section 4, the piecewise linear trend model with SN-NOT is sufficient for this task. However, as for prediction in the second stage, any flexible extrapolation function $f(t)$ can be considered, as it is expected that a linear function may only provide a reasonable forecast for short horizons due to its limited flexibility.

In the following, we consider three commonly used extrapolation functions (in the order of increasing flexibility) in the literature, including the linear function $f(t) = a + b(t/n)$, the quadratic function $f(t) = c + d(t/n) + e(t/n)^2$ and the logistic function $f(t) = \dfrac{L}{1 + \exp\big(-\alpha(t/n - t_0)\big)}$.

Based on $\{Y_t\}_{t=\widehat{\tau}_{\widehat{m}}+1}^n$, a standard OLS can be used to estimate the linear and quadratic functions and a standard nonlinear least square can be used to estimate the logistic function. The $k$-day ahead forecast for $Y_{n+k}$ is formulated respectively as

$$\begin{aligned}
&\text{SN-NOT + Linear [SNL]:} && \widehat{Y}_{n+k} = \widehat{a} + \widehat{b}(1 + k/n),\\
&\text{SN-NOT + Quadratic [SNQ]:} && \widehat{Y}_{n+k} = \widehat{c} + \widehat{d}(1 + k/n) + \widehat{e}(1 + k/n)^2,\\
&\text{SN-NOT + Logistic [SNLG]:} && \widehat{Y}_{n+k} = \frac{\widehat{L}}{1 + \exp\big(-\widehat{\alpha}(1 + k/n - \widehat{t}_0)\big)}.
\end{aligned}$$

The prediction for cumulative deaths on day $n + k$ is $\widehat{\text{Death}}_{n+k} = \exp(\widehat{Y}_{n+k})$.

### 5.2. Data and prediction results

We apply the SN-NOT based prediction method to forecast cumulative deaths in the U.S. and compare its performance with other forecasting models listed on the CDC website.[8] Specifically, following the CDC website, the forecast is generated on five dates, April-27, May-04, May-11, May-18 and May-25, and the forecast horizon is 5-day (one-week) ahead and 12-day (two-week) ahead.

We compare with five forecasting models[9] available on the CDC website: "LANL" by Los Alamos National Laboratory (2020), "Imperial" by Unwin et al. (2020), "UT" by University of Texas (2020), "YYG" by Gu (2020) and "MOBS" by Laboratory for the Modeling of Biological and Socio-technical Systems (2020). These forecasting methods are mainly ensembles of complex mechanistic models (such as SEIR and SEIS), known as compartmental models in epidemiology, which track the spread of infectious disease via a system of differential equations. To highlight the importance of the first-stage change-point analysis, we additionally report the forecast given by fitting a logistic function on the entire time series without segmentation (and name it "Logistic").

Table 5.1 reports the prediction results and the findings can be summarized as follows.

(1) SNL gives comparable performance to other methods for the 5-day ahead forecast, while it considerably overestimates deaths at the 12-day horizon. In other words, linear extrapolation can only be used for short-term forecasts. This is not surprising as the linear function essentially assumes a constant growth rate for the cumulative deaths. While such an approximation is reasonable for short-term, it may not be able to track the growth rate for a long period to make accurate predictions. SNQ generally performs better than SNL due to its increased flexibility, though it tends to underestimate at the 12-day horizon as the quadratic function may pass its peak for long-horizon extrapolation.

(2) SNLG is consistently a top performer among all models thanks to the flexibility of the logistic function, which ensures the fitted curve is non-decreasing and is capable of tracking both increasing and decreasing growth rate. Note that there is a drastic performance difference between the two-stage SNLG forecast and the pure Logistic forecast, which indicates the value of the first-stage change-point estimation for identifying the most recent segment where cumulative deaths exhibit relatively stable behavior.

In summary, the SN-NOT based two-stage prediction, in particular SNLG, provides decent forecasts for the cumulative deaths in the U.S. Considering that SNLG is solely based on the time series of cumulative deaths, this result is rather promising and further confirms the value and validity of the change-point analysis. Though by no means SNLG can replace the complex mechanistic models built on epidemiology principles, we believe it can serve as a meaningful addition to the existing set of forecasting models for tracking the COVID-19 pandemic.

[8] https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html#.

[9] Other models can be found on the CDC website. The five models are chosen as their predictions are available on all the aforementioned dates while other models only report on some of the recent dates.

**Table 5.1**
Prediction performance for cumulative deaths in the U.S. (the top 3 performers on each forecast date are highlighted in bold).

| Date | Target | | True | Imperial | LANL | MOBS | UT | YYG | SNLG | SNL | SNQ | Logistic |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| End-of-Week | | | | | | | | | | | | |
| Apr-27 | May-02 | Forecast | 66527 | 66837 | 69410 | 63029 | 58720 | 73317 | 65067 | 70376 | 63775 | 55480 |
| | | Rel.error | / | **0.47%** | 4.33% | −5.26% | −11.74% | 10.21% | **−2.19%** | 5.79% | **−4.14%** | −16.61% |
| May-04 | May-09 | Forecast | 78946 | 79511 | 78755 | 77035 | 70646 | 77522 | 77178 | 85775 | 75703 | 62930 |
| | | Rel.error | / | **0.72%** | **−0.24%** | −2.42% | −10.51% | **−1.80%** | −2.24% | 8.65% | −4.11% | −20.29% |
| May-11 | May-16 | Forecast | 88893 | 91528 | 87022 | 88922 | 87666 | 88767 | 88128 | 86331 | 84965 | 69702 |
| | | Rel.error | / | 2.96% | −2.10% | **0.03%** | −1.38% | **−0.14%** | **−0.76%** | 2.88% | −4.42% | −21.59% |
| May-18 | May-23 | Forecast | 97220 | 98076 | 96582 | 97252 | 96128 | 97625 | 97573 | 99432 | 97307 | 75659 |
| | | Rel.error | / | 0.88% | −0.66% | **0.03%** | −1.12% | 0.42% | **0.36%** | 2.28% | **0.09%** | −22.18% |
| May-25 | May-30 | Forecast | 103915 | 104671 | 104085 | 104241 | 104736 | 104436 | 103923 | 108080 | 103197 | 80887 |
| | | Rel.error | / | 0.73% | **0.16%** | **0.31%** | 0.79% | 0.50% | **0.01%** | 4.01% | −0.69% | −22.16% |
| Date | Target | | True | Imperial[a] | LANL | MOBS | UT | YYG | SNLG | SNL | SNQ | Logistic |
| Two-week | | | | | | | | | | | | |
| Apr-27 | May-09 | Forecast | 78946 | | 84837 | 70156 | 65903 | 77336 | 74244 | 93730 | 67565 | 58341 |
| | | Rel.error | / | | **7.46%** | −11.13% | −16.52% | **−2.04%** | **−5.96%** | 18.73% | −14.42% | −29.72% |
| May-04 | May-16 | Forecast | 88893 | | 90078 | 85827 | 78243 | 87608 | 85896 | 109953 | 79531 | 64625 |
| | | Rel.error | / | | **1.33%** | −3.45% | −11.98% | **−1.45%** | **−3.37%** | 23.69% | −10.53% | −29.21% |
| May-11 | May-23 | Forecast | 97220 | | 93997 | 97513 | 96232 | 98365 | 96136 | 124580 | 89205 | 70719 |
| | | Rel.error | / | | −3.32% | **0.30%** | **−1.02%** | 1.18% | **−1.11%** | 28.14% | −8.24% | −27.26% |
| May-18 | May-30 | Forecast | 103915 | | 103461 | 104443 | 101060 | 106432 | 105985 | 111822 | 104819 | 76269 |
| | | Rel.error | / | | **−0.44%** | **0.51%** | −2.75% | 2.42% | 1.99% | 7.61% | **0.87%** | −26.60% |
| May-25 | June-06 | Forecast | 109802 | | 110640 | 110285 | 111759 | 111799 | 109708 | 119971 | 106995 | 81251 |
| | | Rel.error | / | | **0.76%** | **0.44%** | 1.79% | 1.82% | **−0.09%** | 9.26% | −2.56% | −26.00% |

[a]Imperial only gives one-week ahead forecast.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2020.07.039.

## References


Andrews, D.W., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61 (4), 821–856.
Aue, A., Horváth, L., 2013. Structural breaks in time series. J. Time Series Anal. 34 (1), 1–16.
Bai, J., 1994. Least squares estimation of a shift in linear processes. J. Time Series Anal. 15 (5), 453–472.
Bai, J., 1997. Estimation of a change point in multiple regression models. Rev. Econ. Stat. 79 (4), 551–563.
Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66 (1), 47–78.
Baranowski, R., Chen, Y., Fryzlewicz, P., 2019. Narrowest-over-threshold detection of multiple change points and change-point-like features. J. R. Stat. Soc. Ser. B Stat. Methodol. 81 (3), 649–672.
Bauwens, L., Koop, G., Korobilis, D., Rombouts, J.V., 2015. The contribution of structural break models to forecasting macroeconomic series. J. Appl. Econometrics (30), 596–620.
Chen, Y.J.Y., Gu, Y., 2018. Subspace change-point detection: A new model and solution. IEEE J. Sel. Top. Sign. Proces. 12(6), 1224–1239.
Chen, J., Gupta, A.K., 2011. Parametric Statistical Change Point Analysis: with Applications to Genetics, Medicine, and Finance. Springer Science & Business Media.
Cho, H., Fryzlewicz, P., 2015. Multiple change-point detection for high-dimensional time series via sparsified binary segmentation. J. R. Stat. Soc. Ser. B Stat. Methodol. 77 (2), 475–507.
Crainiceanu, C., Vogelsang, T., 2007. Nonmonotonic power for tests of mean shift in a time series. J. Stat. Comput. Simul. 77(6), 457–476.
Fan, Z., Mackey, L., 2017. An empirical Bayesian analysis of simultaneous changepoints in multiple data sequences. Ann. Appl. Stat. 11(4), 2200–2221.
Fryzlewicz, P., 2014. Wild binary segmentation for multiple change-point detection. Ann. Statist. 42 (6), 2243–2281.
Gromenko, O., Kokoszka, P., Reimherr, M., 2017. Detection of change in the spatiotemporal mean function. J. R. Stat. Soc. Ser. B Stat. Methodol. 79 (1), 29–50.
Gu, Y., 2020. YYG model. https://covid19-projections.com/.
Hubert, L., Arabie, P., 1985. Comparing partitions. J. Classification 2 (1), 193–218.
Kiefer, N., Vogelsang, T., 2005. A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. Econom. Theory 21, 1130–1164.
Laboratory for the Modeling of Biological and Socio-technical Systems, 2020. Modeling of COVID-19 epidemic in the United States. https://uploads-ssl.webflow.com/58e6558acc00ee8e4536c1f5/5e8bab44f5baae4c1c2a75d2_GLEAM_web.pdf/.
Los Alamos National Laboratory, 2020. LANL model. https://covid-19.bsvgateway.org/.
Maidstone, P.F.R., Letchford, A., 2019. Detecting changes in slope with an $L_0$ penalty. J. Comput. Graph. Statist. 28 (2), 265–275.
Olshen, A.B., Venkatraman, S., Lucito, R., Wigler, M., 2004. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 5 (4), 557–572.
Perron, P., 2006. Dealing with structural breaks. Palgrave Handb. Econom. 1 (2), 278–352.
Pesaran, M.H., Timmermann, A., 2002. Market timing and return prediction under model instability. J. Empir. Financ. 9 (5), 495–510.
Shao, X., 2010. A self-normalized approach to confidence interval construction in time series. J. R. Stat. Soc. Ser. B Stat. Methodol. 72 (3), 343–366.
Shao, X., 2015. Self-normalization for time series: A review of recent developments. J. Amer. Statist. Assoc. 110 (512), 1797–1817.
Shao, X., Zhang, X., 2010. Testing for change points in time series. J. Amer. Statist. Assoc. 105 (491), 1228–1240.
Truong, C., Oudre, L., Vayatis, N., 2020. Selective review of offline change point detection methods. Signal Process. 167.
University of Texas, 2020. The university of texas COVID-19 modeling consortium. https://covid-19.tacc.utexas.edu/projections/.

Unwin, H.J.T., Mishra, S., Bradley, V.C., et al., 2020. Report 23 - state-level tracking of COVID-19 in the United States. http://dx.doi.org/10.25561/79231.
Wu, W., 2005. Nonlinear system theory: Another look at dependence. Proc. Natl. Acad. Sci. USA 102(40), 14150–14154.
Wu, W., Shao, X., 2004. Limit theorems for iterated random functions. J. Appl. Probab. 41, 425–436.
Zhang, T., Lavitas, L., 2018. Unsupervised self-normalized change-point testing for time series. J. Amer. Statist. Assoc. 113 (522), 637–648.
Zhou, Z., Shao, X., 2013. Inference for linear models with dependent errors. J. R. Stat. Soc. Ser. B Stat. Methodol. 75 (2), 323–343.