

# Potential roles of social distancing in preventing the spread of coronavirus disease 2019 (COVID-19) in South Korea

Sang Woo Park<sup>1,\*</sup> Kaiyuan Sun<sup>2</sup> Cécile Viboud<sup>2</sup> Jonathan Dushoff<sup>3,4,5</sup> Bryan T. Grenfell<sup>1,2,6</sup>

**1** Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA

**2** Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

**3** Department of Mathematics and Statistics, McMaster University, Hamilton, ON, Canada

**4** M. G. DeGroote Institute for Infectious Disease Research, McMaster University, Hamilton, ON, Canada

**5** Department of Biology, McMaster University, Hamilton, ON, Canada

**6** Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ, USA

\*Corresponding author: swp2@princeton.edu

Since its first appearance in Wuhan, China, in December 2019, the novel coronavirus disease (COVID-19) has spread internationally, including to South Korea. The first COVID-19 case in South Korea was confirmed on January 20, 2020, from a traveling resident of Wuhan, China [1]. The epidemic then took off in mid-February when COVID-19 began to spread within a church from a city of Daegu — as of March 18, 2020, 8,413 cases have been confirmed, of which 60% are related to the church [1]. While its success in controlling the epidemic has been widely attributed to its extensive testing [2], other factors, such as social distancing, are also likely to have played important roles. Here, we describe potential roles of social distancing in mitigating the spread of COVID-19 by using metro traffic data to compare epidemics in two geographically separated major cities in South Korea.

## 1 Data description

We analyzed epidemiological data, collected between January 20–March 16, 2020, describing the COVID-19 outbreak in Korea. Daily number of reported cases in each geographic region was transcribed from press releases by Korea Centers for Disease Control and Prevention (KCDC) [1]. Partial line lists were transcribed from press releases from KCDC and various local and provincial governments. All data and original reports are stored in a publicly available GitHub repository: <https://github.com/parksw3/COVID19-Korea>.

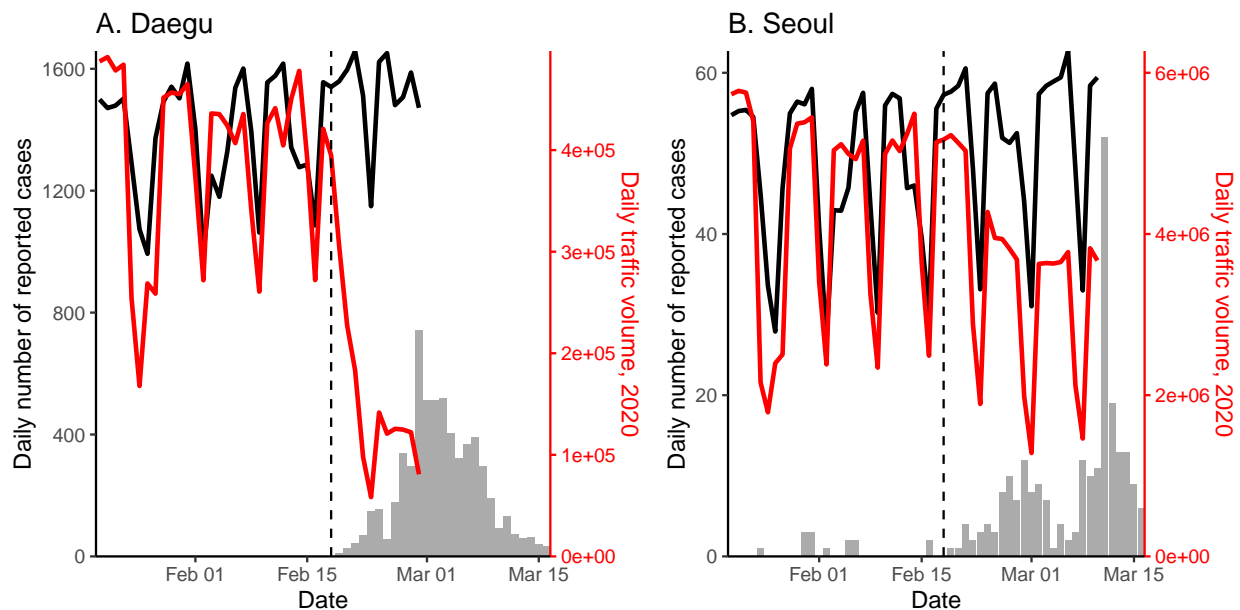


Figure 1: **Comparison of epidemiological and traffic data from Daegu and Seoul.** Solid lines represent the daily metro traffic volume in 2020 (red) and mean daily traffic volume between 2017–2019 (black). Daily traffic from previous years have been shifted by 1–3 days to align day of the week. Vertical lines indicate Feb 18, 2020, when the first case was confirmed in Daegu.

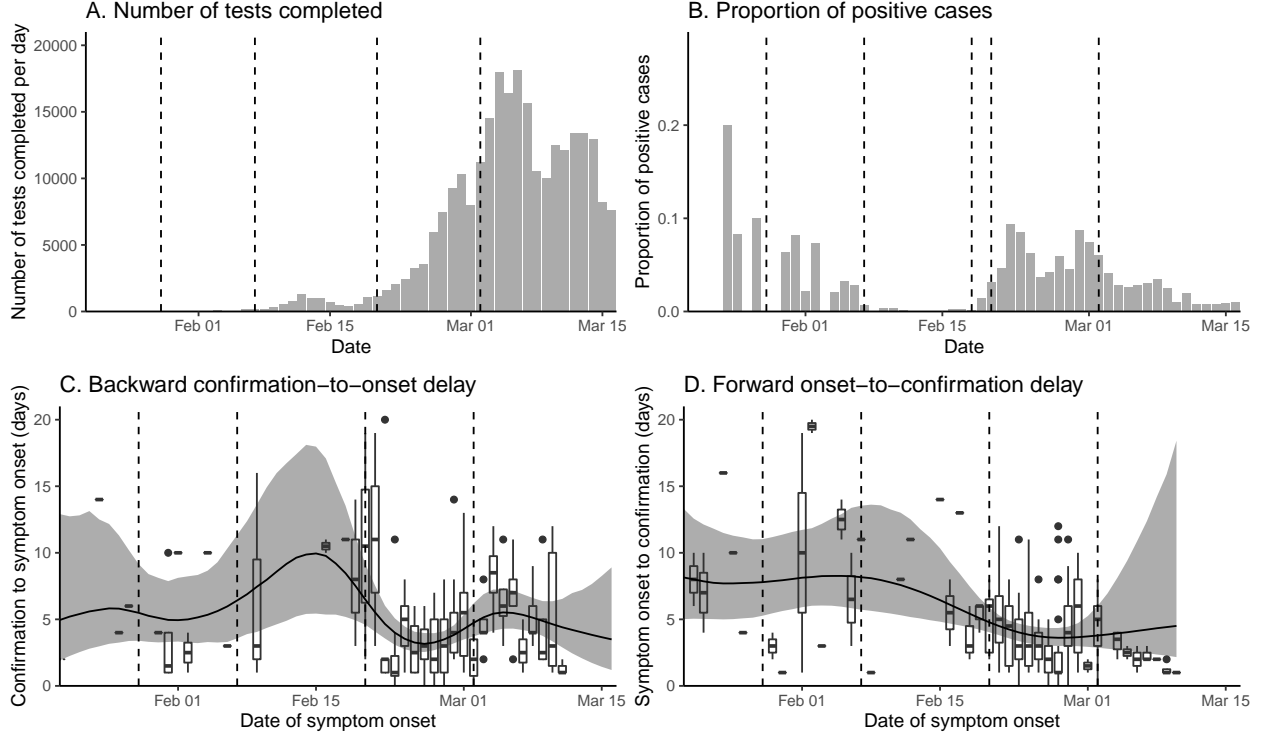


Figure 2: **Changes in the number of tests and delay distributions over time.** Verticle lines indicate the date on which testing criteria expanded.

We compared epidemiological dynamics of COVID-19 from two cities in which the biggest number of COVID-19 cases have been reported: Daegu and Seoul. Between January 20–March 16, 2020, 6,083 cases were reported from Daegu and 248 from Seoul. Unlike the epidemic in Daegu, which is characterized by a single, large peak followed by a gradual decrease, the epidemic in Seoul consists of several small outbreaks (Fig. 1).

Daily metro traffic in Daegu and Seoul between 2017–2020 was obtained from `data.go.kr` and `data.seoul.go.kr`, respectively. We compared the the daily number of individuals who got on the subway — including a monorail in Daegu — across all stations (Fig. 1). Soon after the first church-related case was confirmed in Daegu on Feb 18, 2020, the daily traffic volume decreased by about 80% and 50% compared to previous years in Daegu and Seoul, respectively.

## 2 Trends in time-dependent reproduction number and traffic volume

In order to estimate the time-dependent reproduction number  $R_t$  (i.e., the expected number of secondary cases caused by an individual infected at time  $t$  [3]), we first accounted for changes in testing criteria (Fig. 2B) by multiplying the daily number of reported cases by

the relative detection rate for each criterion (i.e., a proportion of cases tested positive based on a criterion divided by the mean detection rate). Likewise, we accounted for changes in the number of tests completed on each day (Fig. 2A) by dividing the scaled number of reported cases by the relative number of tests completed on each day; this step was performed separately for each testing criterion because widening of criteria necessarily increases the number of tests. A sensitivity analysis showed that qualitative patterns of inferred  $R_t$  are robust to these adjustments (Supplementary Materials).

We then estimated time-dependent *backward* confirmation-to-onset delay distributions from the partial line list using a negative-binomial regression with `brms` package [4] (Fig. 2C). We combined the estimated posterior distribution of delay distributions with previously estimated incubation period distribution [5] to obtain posterior samples for date of infection for each confirmed case. Likewise, we estimated time-dependent *forward* onset-to-confirmation delay distribution using a negative-binomial regression while accounting for right-censoring (Fig. 2D). This allowed us to calculate the degree of censoring (i.e., median probability that a case infected on a given day will be reported after March 16, 2020). Details are provided in the Supplementary Materials.

For each posterior sample of the reconstructed incidence time series (i.e., the number of infected individuals on day  $t$ ), we divided incidence by  $1 - \text{degree of censoring}$  and estimated the time-dependent reproduction number using the renewal equation with a 14-day sliding window [3]:

$$\mathcal{R}_t = \frac{I_t}{\sum_{k=1}^{14} I_{t-k} w_k}, \quad (1)$$

where  $I_t$  is the censoring-adjusted incidence time series and  $w_k$  is the assumed generation-interval distribution. We restricted the calculation of the time-dependent reproduction number between February 2, 2020 (14 days after the first confirmed case was imported) and March 10, 2020 (because the degree of censoring would be too large after this date).

Fig. 3 compares the estimates of  $\mathcal{R}_t$  in Daegu and Seoul. Estimates of  $\mathcal{R}_t$  gradually decrease in Daegu and eventually drop below 1 about a week after the introduction of its first case, coinciding with the decrease in the metro traffic volume (Fig. 3A); our estimates of  $\mathcal{R}_t$  are consistent with [6], although their  $\mathcal{R}_t$  estimates drop below 1 slightly later because they rely on number of symptomatic cases instead. On the other hand, estimates of  $\mathcal{R}_t$  remain around 1 in Seoul (Fig. 3B), suggesting that the degree of social distancing may have been insufficient to prevent the spread in Seoul; stronger distancing or other control measures will be required to decrease  $\mathcal{R}_t$  below 1. Similar patterns in the estimates of  $\mathcal{R}_t$  are found in adjacent provinces, providing support for the robustness of our analysis (Supplementary Materials).

### 3 Discussion

The ongoing COVID-19 outbreak in South Korea provides a unique perspective to understanding and controlling the pandemic. Its experience provides evidence that the epidemic can be contained without draconian measures as in China. Our analysis provides indirect,

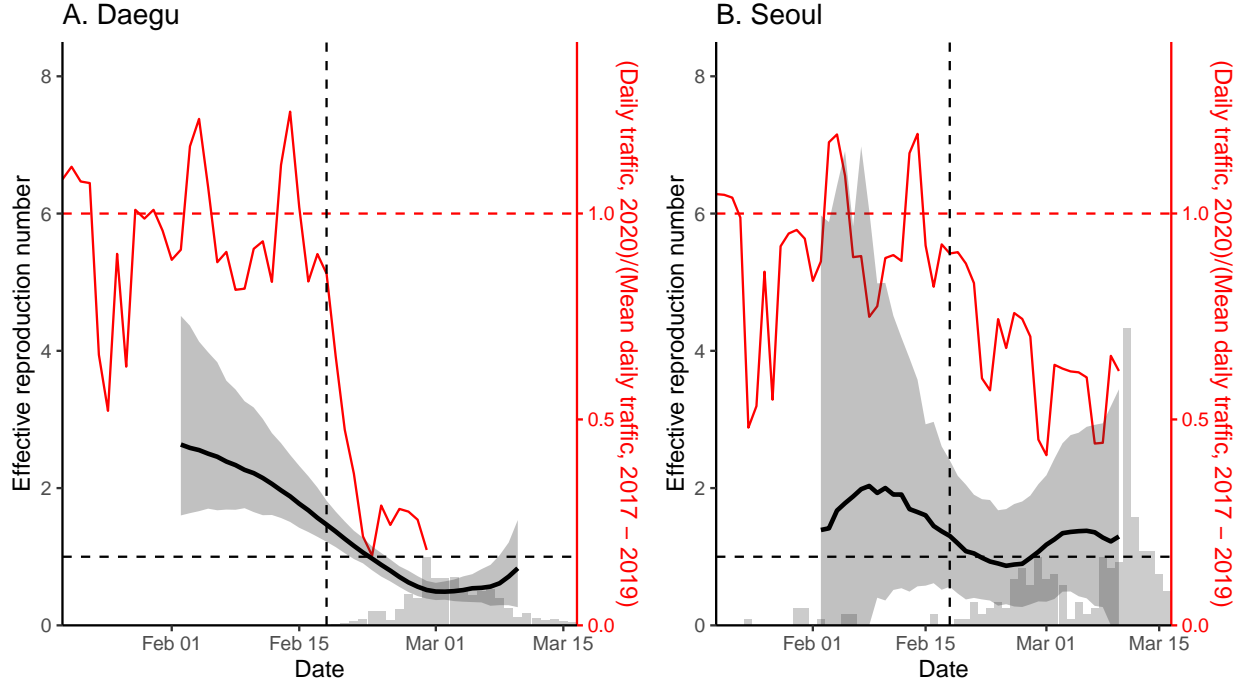


Figure 3: **Comparison of time-dependent reproduction number and normalized traffic in Daegu and Seoul.** Black lines represent the median estimates of  $R_t$ . Gray ribbons represent the corresponding 95% credible intervals. Red lines represent the normalized traffic volume. Vertical lines indicate Feb 18, 2020, when the first case was confirmed in Daegu.

but clear, evidence that social distancing is likely to have assisted in mitigating the epidemic in South Korea. Even though social distancing alone may not be able to fully prevent the spread of the disease, its ability to flatten the epidemic curve (cf. Fig. 3B) reduces burden for healthcare system and provides time to plan for the future [7].

There are many other factors that allowed South Korea to control the COVID-19 outbreak. In particular, the large number of tests as well as the intensity of the testing are likely to have played major roles.

Caveats:

- Only two cities...although two coupled provinces show strikingly similar patterns
- Correlation not causation
- Did not account for geographic heterogeneity in testing or etc. Estimation of delay distribution etc all based on national line list.

## Contribution

Data collection: SWP; analysis: SWP. All authors contributed to the writing and approval of the final report.

## References

- [1] Korea Centers for Disease Control and Prevention (KCDC). Press release (in Korean). 2020. <https://www.cdc.go.kr/board/board.es?mid=a20501000000&bid=0015#>. Accessed Jan 20 – Mar 16, 2020.
- [2] Dennis Normile. Coronavirus cases have dropped sharply in South Korea. What’s the secret to its success? *Science*, 2020. <https://www.sciencemag.org/news/2020/03/coronavirus-cases-have-dropped-sharply-south-korea-whats-secret-its-success>. Accessed Mar 21, 2020.
- [3] Christophe Fraser. Estimating individual and household reproduction numbers in an emerging epidemic. *PloS one*, 2(8), 2007.
- [4] Paul-Christian Bürkner et al. brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, 80(1):1–28, 2017.
- [5] Jantien A Backer, Don Klinkenberg, and Jacco Wallinga. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance*, 25(5), 2020.
- [6] Sam Abbott, Joel Hellewell, James D Munday, June Young Chun, Robin N Thompson, Nikos I Bosse, Yung-Wai Desmond Chan, Timothy W Russell, Christopher I Jarvis, CMMID nCov working group, Stefan Flasche, Adam J Kucharski, Rosalind Eggo, and Sebastian Funk. Temporal variation in transmission during the COVID-19 outbreak. *Science*, 2020. <https://cmmid.github.io/topics/covid19/current-patterns-transmission/global-time-varying-transmission.html>. Accessed Mar 21, 2020.
- [7] Roy M Anderson, Hans Heesterbeek, Don Klinkenberg, and T Déirdre Hollingsworth. How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet*, 2020.

# Supplementary Materials

## Estimating delay distributions between onset and reporting

We estimated both forward and backward delay distributions.

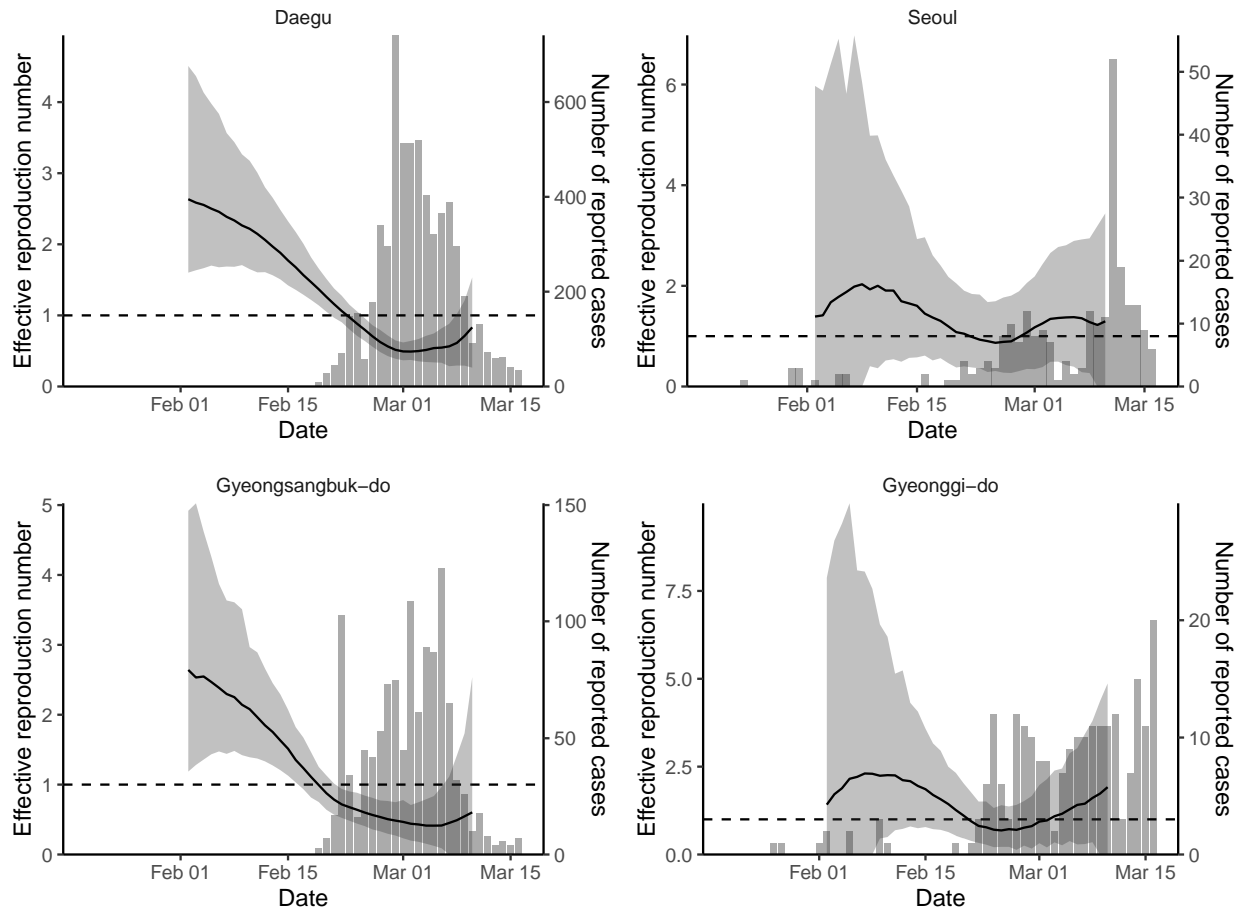


Figure S1: Comparison of effective reproduction number in Daegu, Seoul, Gyeongsangbuk-do, and Gyeonggi-do.