Statistics and Its Interface Volume 14 (2021) 23–24

Discussion on "The timing and effectiveness of implementing mild interventions of COVID-19 in large industrial regions via a synthetic control method" by Tian et al.

Song Xi Chen* and Xianyu Zheng

We congratulate the authors for a well thought and timely work that contributes to both the studies on COVID-19 and the synthetic control (1) method (SCM). The treatment effects were estimated by the SCM, while the impacts of the timing for control measures in Shenzhen are evaluated via the SIHR (Susceptible, Infectious, Hospitalized, Removed) model. The results show that if there were no interventions implemented in Shenzhen, the expected cumulative confirmed cases would be 5.75 times of the observed numbers at the end of the studied period (February 3, 2020). Besides, the effects of the implementation of the mild interventions are shown to be effective as compared with a delayed one, which suggests the importance of a quicker action.

The authors select a 16-day period from January 19, 2020, when the first confirmed case was reported in Shenzhen to February 3, 2020, which includes a 4-day of pre-intervention period and 12 days of post-intervention period. As for the potential control units, the authors selected 68 counties in the United States during the early stage of the epidemic from March 1 to March 16 when no interventions were adopted. The outcome is the cumulative numbers of confirmed cases per 10^5 persons. The covariates include the population density and the latitude, which are found to be associated with the transmission of the COVID-19. In constructing the synthetic control, the authors do not match the pre-intervention outcomes directly. Instead, they perform the principal component analysis (PCA) first and then match the first two components which explain over 95% of the total variation.

We note that the length of the pre-intervention period is quite short at 4 days. It means that only 4 pre-intervention outcomes need to be matched, which is relatively small in comparison to the number of potential controls (68). It appears that the necessity of using PCA is not quite clear. The authors could compare the results between matching the PCA and matching the original covariates, respectively. Another comment is that although some previous studies show that latitude is correlated with the transmission of COVID-19, linear combinations of latitude seems not very natural. And Table 1 suggests the latitude of the synthetic Shenzhen is not very close to the latitude of Shenzhen.

The authors also perform a placebo test to determine the significance of the intervention effects estimated by the SCM. By exchanging Shenzhen and each of the 68 counties in the United States, the authors obtained 68 placebo gaps. The results show that the gap between Shenzhen and the "synthetic Shenzhen" was the negatively largest among the negative gaps, which suggests that the mild interventions had effectively reduced the number of cumulative cases. However, Figure 3 shows two of the placebo gaps had rather unusually large positive gaps as compared to the rest. One possible reason is that the pretreatment matching is poor, which is shown in Figure 3. It is worthy to investigate the cause of the poor match.

Except from using the SCM to estimate the effects of interventions, the authors employed the SHIR (2) model to study the effect of timing for taking the intervention. The SIHR model considers the pre-symptomatic transmission pattern during the incubation period and divides the population into four classes: Susceptible, Infectious, Hospitalized, Removed. The SIHR model employed time-varying contact rate $\beta(t)$, constant parameters L_{in} (the mean length of the incubation), and constant γ (the removing rate). The contact rate was modeled by a logistic function as the following

$$\beta(t) = \frac{\beta_0}{1 + \exp\left(\lambda_{m_2}\left(t - m_1 - m_2/2\right)\right)}.$$

The parameter m_1 reflect the timing of the intervention. The authors studied the effects of delayed interventions by varying the values of m_1 and applying the delay operator $\mathcal{D}_h : m_1 \to m_1 + h$ for $h = 0, \cdots, 5$, which seems to be equivalent to changing the infectious rate $\beta(t) \to \beta(t-h)$. As the logistic specification of $\beta(t)$ is very crucial in this part of the analysis, we wonder if there is a way to calibrate the specification empirically.

In this paper, the authors used both the causal inference method (SCM) and the epidemic compartment model (SIHR) to study the effects of interventions in Shenzhen. We

*Corresponding author.

wonder if what are the connections between the two sets of analyses, especially in the empirical studies. Is the SIHR model used to validate the SCM?

PS: We have a comment on the title for the authors to consider. We would think "The timing and effectiveness of implementing mild interventions of COVID-19 in **Shenzhen** via a synthetic control method" may be more adequate for the study as the SCM is only applied to Shenzhen.

Received 8 October 2020

REFERENCES

[1] Abadie, A., Diamond, A. and Hainmueller, J. (2010), 'Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program', *Journal of the American Statistical Association* **105**(490), 493–505. MR2759929

[2] Tian, T., Zhang, J., Lin, S., Jiang, Y., Tan, J., Li, Z. and Wang, X. (2020), 'Data-driven analysis on the simulations of the spread of covid-19 under different interventions of China', *medRxiv*. MR3884849

Song Xi Chen
Guanghua School of Management
Peking University
Beijing 100871
P.R. China
Center for Statistical Science
Peking University
Beijing 100871
P.R. China
School of Mathematical Science
Peking University
Beijing 100871
P.R. China
E-mail address: csx@gsm.pku.edu.cn

Xianyu Zheng
Guanghua School of Management
Peking University
Beijing 100871
P.R. China
E-mail address: zhengxiangyu@pku.edu.cn