

## Referee Report on JE-2020370

Motivated by the case of COVID-19, this paper considered detecting trends differences between pairs of several time series on various different intervals. Based on theory about Gaussian approximation for maximum statistic, they proposed a multi-scale test and find the suitable cutoff value through Monte Carlo simulation to control the FWER. Simulation study are provided to illustrate the performance. The method is then applied on COVID-19 data. In general the paper is well written and is very readable. I have some concerns mainly from the theoretical point of view.

1. Since the difference could be canceled out, why should one consider  $\sum_t (X_{it} - X_{jt}) \mathbf{1}_{t/T \in I_k}$ ? Isn't it more appropriate to use  $|X_{it} - X_{jt}|$  or  $|X_{it} - X_{jt}|^2$  to capture the distance?
2. Now the conclusion will be largely interfered by the choice of interval sets. I am wondering whether we can reach some unified result without the influence of such selection. That is whether we can aggregate the rejected intervals  $I_k$  and draw some meaningful conclusion?
3. The author mentioned this method can be used to identify locations of changes in the trends. But the detail is not very clear to me. For example consider a very simple case: if the two series  $i, j$  differ from time  $t_1$  to  $t_2$  and are the same before and after this interval, where  $t_1, t_2, t_2 - t_1$  are all unknown. Can we somehow able to identify this interval  $[t_1, t_2]$  using our method and how well can we estimate  $t_1$  and  $t_2$ ? If one takes difference of each pair  $(i, j)$ , and then the trends are zero except some unknown intervals. Then the task is to detect those unknown intervals. Such problem can be possibly solved by for example MOSUM. Can author comments about this?
4. Some theory question: 1) since the result in Chenozukov et al's Gaussian approximation(GA) does not require the series to be independent cross sectionally, I wonder does that mean the current result can be extended to data with cross-sectional dependence? 2) It would be better if the author can derive power under certain alternatives, so that one can get a better idea as how different the trends needs to be

in order to be detected. 3) The argument about no need for time dependent data is reasonable, just a short comment: there already exists result extending chenozukov et al's GA to time dependent case, maybe this paper can be further extended to time dependent data as well.

5. The specific allowance of  $p = |W|$ , which is essential in high dimensional analysis, is not mentioned until appendix. Please put them forward in the main context to provide some guidance in application. Also since the convergence speed of Gaussian approximation depends on  $T, p$ , it would be better to keep the bound in terms of those parameters, so that we know how large the sample size we need in order to obtain the desired accuracy.

## REFERENCES