Proceedings of the International Conference on Industrial Engineering and Operations Management
Dubai, UAE, March 10-12, 2020

Identification of the Automotive Manufacturing Change Point Approaching Panel Data

Karim Atashgar and Naser Rafiee
Industrial Engineering Faculty
Malek Ashtar University of Technology
Tehran, Iran
Atashgar@iust.ac.ir, Naserrafiee67@gmail.com

Abstract

Change point refers to the time when a process really shifts from a normal condition to an unexpected condition. Identification of the change point leads one to an effectiveness root cause analysis of the process. In the case that there exists a multi-dimensional data sets i.e. panel data, an expert can analyze a variable of a cross-section on a determined time series at the same time. This research attempts to detect and analyze the change point based on the number of cars manufactured by different companies around the world in the last 19 years. This paper first identifies the change point of each country separately and then addresses the change point of all countries simultaneously using panel data approach. The comparative report of this paper helps one to analyze factors performed the change of the production volume.

Keywords

Change point, Panel data, Automotive Company.

1. Introduction

Change point analysis of panel data well addresses a valuable time to identify the factor(s) influenced the process. The change point refers to the time when the process departures from a natural condition to an unexpected performance. Identification of the change point helps experts and managers to identify the root cause of the new condition.

Bai (2010) used the least square error (LSE) method to estimate the common change point in means of a panel data and also used the quasi-maximum likelihood (QML) method to estimate the change point in mean, variance, and both. Horváth & Hušková (2012) investigated the statistic of the change point estimator in the panel data means generated by the Bai (2010) and proposed approximately the distribution of the statistic. Horváth & Hušková (2012) used their proposed method to identify change in the Gini coefficients of 33 countries, including European countries, Australia, the United States, South America, China and Taiwan. De Wachter & Tzavalis (2012) proposed a test to detecting the change point in the structure of a dynamic linear panel data model and used it to identify the effects of the 1997 Asian crisis on investment decisions of Asian companies. Enomoto & Nagata (2016) extended the Mahalanobis – Taguchi (MT) method proposed by Taguchi (2002) using Bayes inference to identify the change point in a panel data. Enomoto & Nagata (2016) used their proposed method to detecting the change point in the annual consumption of four beverages in states of Japan for the years 2001 - 2013. They identified the change point in 2010 for Kagoshima state and Shiman state, and in 2011 for Tokyo. Peštová & Pešta (2015) propose a test to identify a common change point in means of the panel data. Their proposed statistic is based on the LSE method and they use the Bootstrap method to determine the test criteria. Peštová & Pešta (2017) proposed a common change point estimator in the panel data means. Their proposed method is also derived from the LSE method. In the case that the last observation is addressed as the change point of the panel, it indicates the lack of the change point, otherwise, it is used to detection of the change point using the bootstrap method proposed by Pestova & Pesta (2015). Cho (2016) proposed a method for identifying the change point in panel data using CUSUM method. Cho

(2016) also used his method to identify change points in financial data sets from stock prices of S & P 100 index components over a one year period.

This paper attempts to investigate the presence or the lack of the change point in automobile production of different countries of the world from 2000 to 2018. For this purpose, we use a change point identification method in panel data and a change point identification method in time series data.

This paper is organized as follows: In the next section, detection methods proposed for identifying the change point in panel data is described. Section 3 is allocated to identifying the change point in time series. Section 4, investigates the presence of the change point in the volume of cars manufactured by different countries of the world. Finally, the last section of this paper is allocated to conclusion and remarks.

2. Detection method of the change point in the panel data

Assume the panel data model is as follows:

$$Y_{it} = \mu_i + \delta_i|\{t > \tau\} + e_{it}\,,\ \ 1 \le i \le N\,, 1 \le t \le T \tag{1}$$

Where N denotes the number of cross-sectional units and T represents the length of time series for each cross-sectional unit of a panel data, and τ addresses the change point. In equation 1, μ_i denotes the mean of the cross-section i under the in-control condition, e_{it} shows the process error in cross-section i and time t, and $\delta_i \neq 0$ case indicates the size of the shift in μ_i after an unknown time of $\tau \in [1, T)$.

Cho (2016) proposed double CUSUM statistic for identifying the change point in panel. Let $\mathbb{Y}^{i}_{\tau\in[1,T)}$ denote CUSUM statistic for the cross-sectional unit i, as follows:

$$\mathbb{Y}^{i}_{\tau\in[1,T)} = \frac{1}{\sigma_i}\sqrt{\frac{\tau(T-\tau)}{T}}\left(\frac{1}{\tau}\sum_{t=1}^{\tau} Y_{it} - \frac{1}{(T-\tau)}\sum_{t=\tau+1}^{T} Y_{it}\right) \tag{2}$$

Where σ_i^2 is the variance of the observations at cross-section i. The double CUSUM statistic is defined as follows:

$$D_{m\in[1,N]}\left(\left\{\left|\mathbb{Y}^{(i)}_{\tau\in[1,T)}\right|\right\}_{i=1}^{N}\right) = \left\{\frac{m(2N-m)}{2N}\right\}^{\varphi}\left(\frac{1}{m}\sum_{i=1}^{m}\left|\mathbb{Y}^{(i)}_{\tau\in[1,T)}\right| - \frac{1}{2N-m}\sum_{i=m+1}^{N}\left|\mathbb{Y}^{(i)}_{\tau\in[1,T)}\right|\right) \tag{3}$$

Where D_m is denoted as double CUSUM statistic operator which takes the ordered CUSUM values $\left|\mathbb{Y}^{(1)}_{\tau\in[1,T)}\right| \ge \left|\mathbb{Y}^{(2)}_{\tau\in[1,T)}\right| \ge \cdots \ge \left|\mathbb{Y}^{(N)}_{\tau\in[1,T)}\right|$ at each τ, and $\varphi \in \{0, ½\}$. The test statistic proposed by Cho (2016) to detecting the presence of the change point over a given interval [1, T] is defined as follows:

$$P_{N,T} = \max_{\tau\in[1,T)}\max_{1\le m\le N} D_m\left(\left\{\left|\mathbb{Y}^{(i)}_{\tau\in[1,T)}\right|\right\}_{i=1}^{N}\right) \tag{4}$$

According to the properties of the general dynamic factor model, Cho (2016) proposed a bootstrap method to determine the test criterion considering both cross-sectional and within-series correlations (See Forni et al. (2000) for details of the general dynamic factor model). In this case the test statistic $P_{N,T}$ is compared with the test criterion π_{NT}. Where the panel experiences the case of $P_{NT} > \pi_{NT}$, the location of the change point is estimated as follows:

$$\hat{\tau}_N = \arg\max_{\tau\in[1,T)}\max_{1\le m\le N} D_m\left(\left\{\left|\mathbb{Y}^{(i)}_{\tau\in[1,T)}\right|\right\}_{i=1}^{N}\right) \tag{5}$$

The density of the number of cross-sectional units influenced the change point estimation is estimated as follows. This density addresses the number of countries affected by a non-random variability.

$$\widehat{m}_N = \arg\max_{1\le m\le N} D_m\left(\left\{\left|\mathbb{Y}^{(i)}_{\hat{\tau}_N}\right|\right\}_{i=1}^{N}\right) \tag{6}$$

3. Method of identifying the change point in time series data

Pettitt (1979) proposes a nonparametric test to identify the change point in the mean of time series. His method is suitable for short time series. Let y_t for t = 1, …, T denotes a random variable. Assume a change point occurs at time τ of time series. In this case, y_t follows a common distribution function F_1 for t = 1,..., τ and follows a common distribution function F_2 for t = τ+1, ..., T, where $F_1 \neq F_2$. It is possible to define a sequence U_t for t = 1, …, T as follows:

$$U_t = \sum_{i=1}^{t}\sum_{j=t+1}^{T} \operatorname{sgn}(y_j - y_i) \tag{7}$$

Where the function sgn(.) is defined as:

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \tag{8}$$

The test statistic proposed by Pettitt (1979) to identifying the presence of the change point over a given interval [1, T] is defined as follows:

$$P_T = 2 \times e^{\frac{-6k^2}{T^3+T^2}} \tag{9}$$

Where k is addressed as:

$$k = \max_{1\le t<T} |U_t| \tag{10}$$

In this study, we propose the following Monte Carlo simulation to obtain the test criteria:

1) Obtain the change point location estimation for the observations $y_1, y_2, \dots, y_T$ via Equation 11:

$$\hat{\tau} = \arg\max_{1\le t<T} |U_t| \tag{11}$$

2) Estimate the error values of the observations based on Equation 12:

$$\hat{e}_t = \begin{cases} y_t - \bar{y}_{\hat{\tau}} \ , & t \le \hat{\tau} \\ y_t - \tilde{y}_{\hat{\tau}} \ , & t > \hat{\tau} \end{cases} \tag{12}$$

Where $\bar{y}_{\hat{\tau}} = \frac{1}{\hat{\tau}}\sum_{t=1}^{\hat{\tau}} y_t$ and $\tilde{y}_{\hat{\tau}} = \frac{1}{T-\hat{\tau}}\sum_{t=\hat{\tau}+1}^{T} y_t$.

3) Consider $n_1 = \min_{1\le t<T} \hat{e}_t$ and $n_2 = \max_{1\le t<T} \hat{e}_t$ then, we perform the Monte Carlo simulation in order to obtain the empirical distribution from P_T test statistic using random sampling of the uniform distribution U (n_1,n_2).

4) We obtained the test criterion by the standard error α= 0.05. If the test criterion is equal to π_T and $P_T < \pi_T$, then the null hypothesis of no-change is rejected and the achieves the change point estimation via Equation 11.

4. Analyzing the change point in the world car production

The analyzed data in this study includes annual car productions in all countries of the world in the time interval of 2000 to 2018, which is accessible at the internet address http://www.oica.net/production-statistics/. This research investigates the presence of the change point in mean of car productions volume for each country separately using the method described in the previous section. The results of the analysis are shown in Table 1. In Table 1, the value of the test statistic is compared with the test criteria obtained by the Monte Carlo method. In the case that the value of the test statistic is smaller than the test criteria, one is led to the presence of the change point. Table 1 also addresses the estimated location of the change point if the analysis indicates the existence of the change point. The last column of Table 1 addresses the increase or decrease of the mean value of the production after the detected change point. The change point analysis reported in Table 1 for the annual car production in separate time series are obtained based on 5% error and with 1000 time iterations. As shown in Table 1, most countries experienced a change in mean of their car production volume between 2006 and 2009.

As shown in Figure 1 the panel data investigation of this research for car production volume corresponding to the interval 2000-2018 addresses the change point in the year 2008 (2008 is addressed by b=9 in Figure 1). In addition, using Equation 6 the density of cross-sectional units number for the estimated change point is addressed by $\hat{m}_N = 22$. The estimated $\hat{m}_N$ indicates that 22 countries lead the calculation to identifying the common change point. In this investigation R software is used.

As shown in Figure 1 the test statistic value is P_{NT}=15.94188 and the test threshold is π_{NT} =9.3776. Since the statistic value is larger than the threshold, the presence of the change point in the panel data is confirmed and the location of this change point is the 9th time (As b=9 addressed in Figure 1), which is related to 2008. The result is also approved by the economic analysis reports.

Table 1. The change point in the mean of car productions for each country separately

Country	Test statistic	Test criteria	Change point location	Increase or decrease of production
1. Argentina	0.021	0.12	2005	increase
2. Austria	no change point			
3. Belgium	0.002	0.10	2008	decrease
4. Brazil	0.020	0.12	2006	increase
5. Canada	0.002	0.10	2008	decrease
6. China	0.002	0.12	2008	increase
7. Czech Rep.	0.002	0.12	2008	increase
8. Egypt	0.013	0.10	2011	decrease
9. Finland	0.034	0.12	2013	increase
10. France	0.002	0.10	2008	decrease
11. Germany	no change point			
12. Hungary	0.006	0.12	2006	increase
13. India	0.002	0.15	2008	increase
14. Indonesia	0.002	0.10	2009	increase
15. Iran	0.021	0.12	2007	increase
16. Italy	0.004	0.10	2007	decrease
17. Japan	0.021	0.10	2008	decrease
18. Malaysia	0.003	0.12	2007	increase
19. Mexico	0.002	0.12	2009	increase
20. Others	no change point			
21. Poland	no change point			
22. Portugal	0.043	0.12	2006	decrease
23. Romania	0.002	0.10	2008	increase
24. Russia	0.021	0.10	2009	increase
25. Serbia	0.013	0.15	2012	increase
26. Slovakia	0.003	0.12	2010	increase
27. Slovenia	no change point			
28. South Africa	no change point			
29. South Korea	0.003	0.12	2009	increase
30. Spain	no change point			
31. Taiwan	no change point			
32. Thailand	0.002	0.10	2009	increase
33. Turkey	0.007	0.12	2006	increase
34. UK	no change point			
35. Ukraine	0.034	0.10	2012	decrease
36. USA	0.016	0.12	2006	decrease
37. Uzbekistan	0.007	0.10	2006	increase
Total	0.002	0.10	2009	increase

```
                  [,1]
level_index   1.00000
child_index   1.00000
s             1.00000
b             9.00000
e            19.00000
test_stat    15.94188

$ecp
b
9

$thr
    95%
9.34776

attr(,"class")
[1] "bin.tree"
```

Figure 1. The result report of the software

The result of this research is confirmed by the economic report. The crisis of automotive industries in 2008-2010 was reported as a part of the financial crisis of 2007-2008 leading to a great recession (Chakraborty Das 2008, Associated Press 2008). The crisis of automotive primarily reported for American industries and then the crisis affected European, Canadian, and Asian industries.

The change point addressed in this research is about the time that a substantial increase in the automotive fuel price experienced and automotive industries decided to change their strategies. Three big American companies, General Motors, Ford, and Chrysler, encouraged to focus on vehicles with high profit margin. The Chinese government reduced automotive taxes in order to spur flagging sales, in 2008.

In fact, the results of this paper addresses the value time in a scientific method. The change point can be focused on experts to analyze the source of the unnatural condition. In addition to, the method of used by this paper is capable of applying for monthly, weekly or daily observations.

It should be noted that this crisis affected on each country differently, for example, China experienced an increase in average car production after 2008, while this is the reverse for Canada. Thus identifying the change point in time series observations of the sum or mean of the world car production may mislead us into the accurate estimation of the change point. As can be seen in Table 1, the change point of the total of the world car production is estimated the 2009. In other words, the influence of the various external factors, at one time or at different times, may vary from one cross - sectional unit to another. The panel data provides the opportunity of the considering the effects of the cross - sectional units.

5. Conclusion

This paper analyzed car productions by the world countries in terms of the presence of a common change point in the mean. Change point identification provides the opportunity of analyzing the sources(s) manifested itself to the process. Change point addresses the time that really an assignable cause(s) has been experienced by the process. This paper focused on the annual automotive productions of the world and provided a panel data. The change point analysis of this panel using two approach addressed a valuable time i.e. the change point. The identified change point of this paper is confirmed by the economic reports. The method used in this paper can be approached by professional industrial management. This professional approach leads managers to eliminate the special cause of the process as soon as possible.

References

Bai, J., Common breaks in means and variances for panel data, *Journal of Econometrics*, vol. 157, no, 1, pp. 78-92, 2010.

Horváth, L., and Hušková, M., Change-point detection in panel data, *Time Series Anallysis*, vol. 33, no. 4, pp. 631-648, 2012.

De Wachter, S., and Tzavalis, E., Detection of structural breaks in linear dynamic panel data models, *Computational Statistics and Data Analysis*, vol. 56, no. 11, pp. 3020-3034, 2012.

Enomoto, T., and Nagata, Y., Detection of change points in panel data based on bayesian MT method, *Total Quality Science*, vol. 2, no. 1, pp. 36-47, 2016.

Taguchi, G., *Technological development in the MT system*, Japan Standards Association, 2002. (In Japanese)

Peštová, B. and Pešta, M., Testing structural changes in panel data with small fixed panel size and bootstrap, *Metrika*, vol. 78, no. 6, pp. 665-689, 2015.

Peštová, B. and Pešta, M., Change point estimation in panel data without boundary issue, *Risks*, vol. 5, no. 1, pp. 7, 2017.

Cho, H., Change-point detection in panel data via double CUSUM statistic, *Electronic Journal of Statistics*, vol. 10, no. 2, pp. 2000-2038, 2016.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L., The generalized dynamic factor model: identification and estimation, *The Review of Economics and Statistics*, vol. 82, no. 4, pp. 540-554, 2000.

Pettitt, A.N., A non-parametric approach to the change-point problem, *Applied Statistics*, vol. 28, no. 2, pp. 126-135, 1979.

Chakraborty Das, J., Uncertainty in U.S. auto industry puts pressure on suppliers, Available: https://web.archive.org/web/20080924054405/http://www.iht.com/articles/2008/09/18/business/deal19.php, 2008.

Associated Press, Gas prices put Detroit Big Three in crisis mode, Available: https://web.archive.org/web/20121103110602/http://www.msnbc.msn.com/id/24896359/, 2008.

Biographies

Karim Atashgar is an associate professor at Industrial Engineering Faculty. He received his PhD in Industrial Engineering from Iran University of Science and Technology. His primary research interests include Statistical Process Control (SPC), Profile Monitoring, Quality Management, Applied Statistical Analysis, Artificial Neural Network, Project Control, Project Management, Know-how management, and Transfer of Technology. He experienced different positions of management in different work areas such as Automotive Industry, Project based organizations and Planning and programming for holding companies.

Email: atashgar@iust.ac.ir

Naser Rafiee is a researcher at Iran University of Science and Technology. He received his BS in Applied Mathematics from University of Payame Noor and his MS in Industrial Engineering from Malek Ashtar University of Technology. His primary research interests is applied Statistical Analysis.

Email: naserrafiee67@gmail.com