Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
March-April 2022 Page No. 1188-1196
Article History: Received: 08 February 2022, Revised: 10 March 2022, Accepted: 02 April 2022, Publication: 30 April 2022

Estimation of Polytropic Points of Digital Characteristic Variance in Flat Panel Data

Guangjun Liu[1,*], Xiaoping Xu[1], Feng Wang[2], Long Liu[3]
[1]School of Sciences, Xi'an University of Technology, Xi'an, P.R. China
[2]School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, P.R. China
[3]School of Automation and Information Engineering, Xi'an University of Technology, Xi'an, P.R. China
[*]Corresponding Author.

Abstract:

Flat panel data model has many advantages, such as increasing the degree of freedom between variables, reducing multicollinearity, obtaining more effective estimators and so on. The change of variance in its data means the change of some risk, so studying the variance change point of flat panel data can effectively control the risk. In this paper, the quasi maximum likelihood estimate and cumulative sum estimate are used to reasonably estimate the variance change point in panel data, and further combined with the binary segmentation method, it is extended to the polytropic point problem. Theoretical reasoning and numerical simulation show that the proposed method is reasonable.

Keywords: *data, digital features, polytropic point, estimate*

I. INTRODUCTION

Studying the structural changes of flat panel data can objectively reveal the law of economic operation, timely grasp the changes of the market and effectively prevent risks [1-4]. From the perspective of probability and statistics, the estimation of structural change of panel data can be reduced to the estimation of structural change point of panel data model in essence.

The change point problem was originally raised by Page from the quality control problem [5]. It is that people spot check products from the production line to detect whether the product quality fluctuates significantly. Chen proposed that the change point refers to the point where one or some quantities in the model change suddenly [6]. In change point analysis, variance change point is an important type of change point. Analyzing variance change point plays an important role in controlling and reducing financial risks. On the research of variance change point, Bai discusses the estimation of variance change point in panel data by using the quasi maximum likelihood method, proves the consistency of change point estimator, and gives the limit distribution of change point estimator [7]; Li et al studied the test of variance change point in panel data by cumulative sum method [8]; Hou further discussed the estimation of variance change point [9]; Liu studied the estimation of variance change point in panel data based on the least square estimator, proved the consistency of the estimator, and gave the convergence speed of the estimator [10]; However, in reality, there may be more than one change point. For example, there may be multiple change points in financial data for a long time. Therefore, it is necessary to study the change points in practical problems.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
March-April 2022 Page No. 1188-1196
Article History: Received: 08 February 2022, Revised: 10 March 2022, Accepted: 02 April 2022, Publication: 30 April 2022

In recent years, many scholars have used different methods to study the structural variability of panel data model, and achieved a series of results. Under the assumption that there are multiple common structural change point panel data models, Qian and Su proposed the penalty least square method and the penalty generalized moment method to estimate the structural change parameters respectively based on the two cases with or without endogenous variables [11]; Ding et al composed the meteorological data of different cities into panel data, and detected and estimated the meteorological variability based on quasi maximum likelihood estimation statistics and binary segmentation method [12]. Among many methods, because the binary segmentation method can detect all change points in the local sequence, can reduce the limitations of global statistics, has the advantages of good detection effect and high accuracy. Therefore, the binary segmentation method is a commonly used method in dealing with the problem of multiple points of panel data [13].

After constructing the quasi maximum likelihood estimate and cumulative sum estimate for the variance change point of the flat panel data model, this paper gives the consistency proof of the change point estimator, and gives the algorithm steps of estimating the variance change point combined with the binary segmentation method. Finally, numerical simulation shows that the proposed method is effective.

II. MODEL AND ASSUMPTIONS

Consider the following flat panel data variance polytropic points model:

$$F_{js} = \begin{cases} \mu_j + \sigma_{j1}\xi_{js}, & s = 1,2,\cdots,l_1, \\ \mu_j + \sigma_{j2}\xi_{js}, & s = l_1+1, l_1+2,\cdots,l_2, \\ \quad\cdots & \\ \mu_j + \sigma_{jn}\xi_{js}, & s = l_{n-1}+1,_{n-1}+2,\cdots,l_n, \\ \mu_j + \sigma_{j,n+1}\xi_{js}, & s = l_n+1, l_n+2,\cdots,L, \end{cases} \tag{1}$$

where, $j = 1,2,L\ ,P$, F_{js} is the observed value, μ_j is the mean value of the jth panel, ξ_{js} is the error term, P is the number of flat panel data, L is the sample size on each panel. If $\sigma_{js} \neq \sigma_{j,s+1}$, $s = 1,2,\cdots,q$, it is called q change points.

In order to reasonably estimate the variance change point, the following assumptions are given here:

Assumption 1: The expectation of ξ_{js} is 0 and the variance is 1, for each panel j, $\xi_{js} = \sum_{s=0}^{\infty} b_{js}\zeta_{j,s-j}$, ζ_{js} follows a normal distribution with the mean of 0 and the variance of 1, $\sum_{s=1}^{\infty} s\left|b_{js}\right| < \infty$, $D(\xi_{js}) = \sum_{j=0}^{\infty} b_{js}^2 = 1$.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
March-April 2022 Page No. 1188-1196
Article History: Received: 08 February 2022, Revised: 10 March 2022, Accepted: 02 April 2022, Publication: 30 April 2022

Assumption 2: There is a positive number Z such that $E(\zeta_{js}^4) \le Z$.

Assumption 3: $P\log(\log(L))/L$ tends to zero.

Assumption 4: When P goes to infinity, $\sum_{j=1}^{P} g(\sigma_{js}^2/\sigma_{j,s+1}^2)$ equals 1, $s=1,2,\cdots,q$, $g(y)=y-1-\log(y)$, y is a positive number.

Assumption 5: There is a positive number $\upsilon_j^0 \in (0,1)$, $\zeta > 0$ such that $l_j^0 = [L\upsilon_j^0]$ and $\upsilon_{j+1}^0 - \upsilon_j^0 > \zeta$, $j=0,1,\cdots,q$, where $[\cdot]$ is the rounding function.

Assumption 1 ensures that the error term satisfies the stationarity, Assumption 2 requires that the fourth-order moment of the error term is limited, Assumption 3 allows the sample size to be much larger than the number of panels, Assumption 4 ensures the consistent estimation of variance change points, and Assumption 5 ensures that there are enough samples between each two change points of the model, which is the basic condition for the establishment of the law of large numbers and the functional central limit theorem, ε usually adopts smaller numbers such as 0.05 and 0.01 [14].

III. CHANGE POINT ESTIMATION

3.1 Polytropic Points Estimation Based on Quasi Maximum Likelihood Estimate

Consider the following case of a change point:

$$\begin{aligned} F_{js} &= \mu_{j1} + \sigma_{j1}\xi_{js}, \quad s=1,2,\cdots,l_0\ , j=1,2,\cdots,P, \\ F_{js} &= \mu_{j2} + \sigma_{j2}\xi_{js}, \quad s=l_0+1,\cdots,L, j=1,2,\cdots,P. \end{aligned} \tag{2}$$

Let

$$\sigma_{j1}^2(l) = (\sum_{s=1}^{l}\left(F_{js}-\tilde{F}_{j1}\right)^2)/l\,,\quad \sigma_{j2}^2(l) = (\sum_{s=l+1}^{L}\left(F_{js}-\tilde{F}_{j2}\right)^2)/(L-l)\,,$$

where, $\tilde{F}_{j1} = (\sum_{s=1}^{l} F_{js})/l$, $\tilde{F}_{j2} = (\sum_{s=l+1}^{L} F_{js})/(L-l)$, then the objective function of the ith panel is

$$l\sum_{j=1}^{P}\log\sigma_{j1}^2(l) + (L-l)\sum_{j=1}^{P}\log\sigma_{j2}^2(l)\,,\quad j=1,2,\cdots,P\,.$$

In this way, for P panels, the objective function becomes

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
March-April 2022 Page No. 1188-1196
Article History: Received: 08 February 2022, Revised: 10 March 2022, Accepted: 02 April 2022, Publication: 30 April 2022

$$W_{PL}(l) = l\sum_{j=1}^{P}\log\sigma_{j1}^{2}(l) + (L-l)\sum_{j=1}^{P}\log\sigma_{j2}^{2}(l)$$

Thus, the quasi maximum likelihood estimator of the change point is

$$\breve{l} = \arg\min_{1<l<L} W_{PL}(l) \tag{3}$$

Theorem 1. If Assumptions 1-5 hold, then when P and L tend to infinity, $p(\breve{l} = l_0)$ tend to one.

The proof of the theorem is shown in Reference [7].

In the above discussion, we use the quasi maximum likelihood estimator to estimate the single variation point of variance. Next, we discuss the estimation of variance variation point.

Assuming that the number of change points q is known, combined with the binary segmentation theory, the above method is extended to the case of change points. If the number of change points q is unknown, the recursive test method in Reference [14] can be used to estimate the number of change points. The specific steps of estimating the variance points in panel data model (1) are as follows:

Step 1: Use Equation (3) to estimate the first change point $\breve{l}_1$;

Step 2: Divide the whole panel data into $F_{j1}, F_{j2}, \cdots, F_{j\breve{l}_1}$ and $F_{j(\breve{l}_1+1)}, F_{j(\breve{l}_1+2)}, \cdots, F_{jL}$ ($j = 1, 2, \cdots, P$) two parts at $\breve{l}_1$, and estimate the change points in the two parts of the panel by using Equation (3), which is recorded as l_2^*, l_2^{**};

Step 2.1: Calculate WO and $W_{PL}^1(l_2^*)$ in the panel data in the previous section;

$$WO = \breve{l}_1\sum_{j=1}^{P}\log\sigma_{j1}^{2}(\breve{l}_1),$$

$$W_{PL}^{1}(l_2^*) = \min_{1<l<\breve{l}_1}\left\{l\sum_{j=1}^{P}\log\sigma_{j1}^{2}(l) + (\breve{l}_1 - l)\sum_{j=1}^{P}\log\sigma_{j2}^{2}(l)\right\}.$$

Step 2.2: Calculated WG and $W_{PL}^2(l_2^{**})$ in the later panel data;

$$WG = (L-\breve{l}_1)\sum_{j=1}^{P}\log\sigma_{j2}^{2}(\breve{l}_1),$$

$$W_{PL}^{2}(l_2^{**}) = \min_{\breve{l}_1<l<L}\left\{(l-\breve{l}_1)\sum_{j=1}^{P}\log\sigma_{j1}^{2}(l) + (L-l)\sum_{j=1}^{P}\log\sigma_{j2}^{2}(l)\right\}.$$

Step 2.3. Let $x_1 = WO - W_{PL}^1(l_2^*)$, $x_2 = WG - W_{PL}^2(l_2^{**})$. Compare the size of x_1 and x_2, if x_1 is greater than x_2, $\breve{l}_2$ is equal to l_2^*, otherwise, $\breve{l}_2$ is equal to l_2^{**}.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
March-April 2022 Page No. 1188-1196
Article History: Received: 08 February 2022, Revised: 10 March 2022, Accepted: 02 April 2022, Publication: 30 April 2022

Step 3: Sort the change point $\breve{l}_1$ estimated by Step 1 and the change point $\breve{l}_2$ estimated by Step 2. Based on these two change points, the whole panel data is divided into three parts, similar to Step 2 to estimate the third change point. This continues until q change points are estimated.

3.2 Polytropic Points Estimation Based on Cumulative Sum Estimate

Considering the model (2) given above, let

$$G_{PL}(l)=\frac{1}{\sqrt{P}}\sum_{j=1}^{P}\left\{\frac{1}{\breve{\psi}_j}\frac{1}{\sqrt{L}}\left(\sum_{s=1}^{l}\zeta_{js}^2-\frac{l}{L}\sum_{s=1}^{L}\zeta_{js}^2\right)\right\}$$

where, $1<l<L$, $\zeta_{js}=F_{js}-\tilde{F}_L(j)$, $\tilde{F}_L(j)=(\sum_{s=1}^{L}Y_{js})/L$, $\breve{\psi}_j^2=(\sum_{s=1}^{L}\zeta_{js}^4-\sum_{s=1}^{L}\zeta_{js}^2)/L$, $j=1,2,\cdots,P$. Thus, the cumulative sum of the change point is estimated to be

$$\breve{l}=\underset{1<l<L}{\arg\max}\left|G_{PL}(l)\right| \tag{4}$$

In order to prove the consistency of the change point estimate, the following assumptions need to be given:

Assumption 6: When $\min(P,L)$ tends to infinity, $L\Theta_{PL}/P$ tends to infinity, where, $\Theta_{PL}=\sum_{j=1}^{P}\Phi_j^2$, $\Phi_j^2=\sigma_{js}^2-\sigma_{j,s-1}^2$.

Theorem 2. if Assumptions 1-4 and Assumption 6 hold, then when P,L tend to infinity, $p(\breve{l}=l_0)$ tends to one.

The proof of Theorem 2 is similar to the proof of Theorem 2.1 in Reference [15].

The cumulative sum estimator method is used to estimate the single change point of the square difference. Similar to Section 3.1, assuming that the number of change points q is known, combined with the binary segmentation theory, the above method is extended to the case of multiple points. The specific steps for estimating the variance multiple points in the panel data model (1) are as follows:

Step 1: Use Equation (4) to estimate the first change point $\breve{l}_1$;

Step 2: Divide the whole panel data into $F_{j1},F_{j2},\cdots,F_{j\breve{l}_1}$ and $F_{j(\breve{l}_1+1)},F_{j(\breve{l}_1+2)},\cdots,F_{jL}$ ($j=1,2,\cdots,P$) two parts at $\breve{l}_1$, and estimate the change points in the two parts of the panel by using Equation (4), which is

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
March-April 2022 Page No. 1188-1196
Article History: Received: 08 February 2022, Revised: 10 March 2022, Accepted: 02 April 2022, Publication: 30 April 2022

recorded as l_2^*, l_2^{**};

Step 2.1: Calculate $G_{PL}^1(l_2^*)$ in the panel data in the previous section;

$$G_{PL}^1(l_2^*) = \max_{1<l<\breve{l}_1}\left|\frac{1}{\sqrt{P}}\sum_{j=1}^{P}\left\{\frac{1}{\breve{\psi}_j}\frac{1}{\sqrt{\breve{l}_1}}\left(\sum_{s=1}^{l}\zeta_{js}^2 - \frac{l}{\breve{l}_1}\sum_{s=1}^{\breve{l}_1}\zeta_{js}^2\right)\right\}\right|.$$

Step 2.2: Calculated $G_{PL}^2(l_2^{**})$ in the later panel data;

$$G_{PL}^2(l_2^{**}) = \max_{\breve{l}_1<l<L}\left|\frac{1}{\sqrt{P}}\sum_{j=1}^{P}\left\{\frac{1}{\breve{\psi}_j}\frac{1}{\sqrt{L-\breve{l}_1}}\left(\sum_{s=\breve{l}_1+1}^{l}\zeta_{js}^2 - \frac{l-\breve{l}_1}{L-\breve{l}_1}\sum_{s=\breve{l}_1+1}^{L}\zeta_{js}^2\right)\right\}\right|.$$

Step 2.3. Compare the size of $G_{PL}^1(l_2^*)$ and $G_{PL}^2(l_2^{**})$, if $G_{PL}^1(l_2^*)$ is greater than $G_{PL}^2(l_2^{**})$, $\breve{l}_2$ is equal to l_2^*, otherwise, $\breve{l}_2$ is equal to l_2^{**}.

Step 3: Sort the change point $\breve{l}_1$ estimated by Step 1 and the change point $\breve{l}_2$ estimated by Step 2. Based on these two change points, the whole panel data is divided into three parts, similar to Step 2 to estimate the third change point. This continues until q change points are estimated.

IV. NUMERICAL SIMULATION

Here, consider the following situations:

One change point: $F_{js} = \begin{cases} \mu_j + \sigma_{j1}\xi_{js},\ s = 1,2,\cdots,l_1, \\ \mu_j + \sigma_{j2}\xi_{js},\ s = l_1+1, l_1+2,\cdots,L \end{cases}$

Three change points: $F_{js} = \begin{cases} \mu_j + \sigma_{j1}\xi_{js},\ s = 1,2,\cdots,l_1, \\ \mu_j + \sigma_{j2}\xi_{js},\ s = l_1+1, l_1+2,\cdots,l_2, \\ \mu_j + \sigma_{j3}\xi_{js},\ s = l_2+1, l_2+2,\cdots,l_3, \\ \mu_j + \sigma_{j4}\xi_{js},\ s = l_3+1, l_3+2,\cdots,L, \end{cases}$

where, $j = 1,2,L\ ,P$, for simplicity, ξ_{js} follows a normal distribution with a mean of 0 and a variance of 1, $\mu_j = 1$, $(j = 1,2,\cdots,P)$. For the case of a change point, σ_{j1} and σ_{j2} take 0.4 and 0.8, respectively, $l_1 = L/5$, $l_1 = 2L/5$ and $l_1 = 4L/5$, $P, L = 40, 80$. For the case of three change points, σ_{j1}, σ_{j2}, σ_{j3} and σ_{j4} take 0.4, 0.8, 1.2 and 1.6, respectively, $l_1 = L/5$, $l_1 = 2L/5$ and $l_1 = 4L/5$, $P, L = 40, 80, 180$. The number of Monte Carlo simulation tests is 2000.

When there is a change point in the panel data, Table 1 shows the median, mean and standard deviation (std) of change point estimator based on quasi maximum likelihood estimate and cumulative sum estimate.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
March-April 2022 Page No. 1188-1196
Article History: Received: 08 February 2022, Revised: 10 March 2022, Accepted: 02 April 2022, Publication: 30 April 2022

It can be seen from Table 1 that for different sample sizes and change point positions, the change point estimators obtained based on quasi maximum likelihood estimates and cumulative sum estimates are close to the real change point, and the standard deviation is relatively small.

TABLE I. Estimation Results of the Two Methods For A Change Point

$\breve{l}_1$	*P/L*	quasi maximum likelihood			cumulative sum method		
		median	mean	std	median	mean	std
$l_1 = L/5$	40/40	8	8.01	0.05	8	8.11	0.54
	80/40	8	7.98	0.10	8	8.05	0.37
	40/80	16	16.00	0.06	16	16.27	0.65
	80/80	16	16.20	0.02	16	16.12	0.54
$l_1 = 2L/5$	40/40	16	16.00	0.06	16	15.98	0.20
	8/40	16	16.00	0.05	16	16.00	0.01
	40/80	32	32.01	0.03	32	32.02	0.05
	80/80	32	32.00	0.02	32	32.00	0.01
$l_1 = 4L/5$	40/40	32	32.00	0.04	32	32.00	0.00
	80/40	32	32.00	0.01	32	32.00	0.02
	40/80	64	64.00	0.02	64	64.00	0.03
	80/80	64	64.00	0.01	64	64.00	0.01

TABLE II. Estimation Results of the Two Methods for Three Change Points

$\breve{l}_1$	*P/L*	quasi maximum likelihood			cumulative sum method		
		median	mean	std	median	mean	std
$l_1 = L/5$	40/40	8	8.01	0.51	8	7.99	0.09
	80/80	16	16.00	0.02	16	16.00	0.03
	180/80	16	16.02	0.01	16	16.00	0.01
	80/180	36	36.00	0.03	36	36.00	0.00
$l_1 = 2L/5$	40/80	32	32.11	1.70	32	32.31	0.78
	80/80	32	32.00	0.64	32	32.20	0.51
	180/80	32	32.03	0.07	32	32.02	0.21
	80/180	72	71.99	0.30	72	72.09	0.43
$l_1 = 4L/5$	40/80	64	64.77	3.35	64	64.03	0.55
	80/80	64	63.88	0.83	64	64.05	0.17
	180/80	64	64.02	0.41	64	64.02	0.07
	80/180	144	144.10	0.27	144	143.96	0.16

Similarly, When there are three change points in the panel data, Table 2 shows the median, mean and standard deviation (std) of change point estimator based on quasi maximum likelihood estimate and cumulative sum estimate. It can be seen from Table 2 that for different sample sizes and change point positions, the change point estimators obtained based on quasi maximum likelihood estimates and cumulative sum estimates are also close to the real change points, and the standard deviation is relatively small.

To sum up, when the binary segmentation method based on quasi maximum likelihood estimate and the binary segmentation method based on cumulative sum estimate are used to estimate the variance polymorphs of panel data model, there are three main influencing factors: sample size, change point position and error term distribution. When the sample size increases, the estimation results of the two methods are better. In other words, the method proposed in this paper is effective.

V. CONCLUSION

In this paper, the variance change point of panel data model is estimated based on quasi maximum likelihood estimate and cumulative sum estimate respectively. Combined with binary segmentation method, it is extended to the case of multiple points, and the specific algorithm steps are given. Monte Carlo simulation experiments show that the binary segmentation method based on quasi maximum likelihood estimate and cumulative sum estimate can effectively estimate the change point position. In the future, this method will be used in practical problems and is expected to play a better role.

ACKNOWLEDGMENTS

This work is supported by the Innovation Capability Support Program of Shaanxi Province of China (Grants No. 2020PT-023), National Natural Science Foundation of China (Grants No. 11801438), and Natural Science Basic Research Plan in Shaanxi Province of China (Grants No. 2018JQ1089) and the Innovative Training Item for College Student of Shaanxi Province.

REFERENCES

[1] Farooq, S., Ozturk, I., Majeed, M.T., and Akram, R. (2022), 'Globalization and CO emissions in the presence of EFC: a global panel data analysis', Gondwana Research, 106, pp. 367-378.
[2] Liu, C., and Nie, G.H. (2021), 'Identifying the driving factors of food nitrigen footprint in China: econometric analysis of provincial spatial panel data by STIRPAT model', Sustainability, 13, pp. 1-23.
[3] Bao, A.N. (2015), 'Application of panel data model in credit asset securitization', Journal of Inner Mongolia Nornal University, 44, pp. 316-321.
[4] Yu, Y., and Li, F. (2010), 'Analysis of the impact of agrometeorological disasters on China's grain output based on panel data', Economics and Management, 24, pp. 5-8.
[5] Page, E.S. (1954), 'Continuous inspection schemes', Biometrika, 41, pp. 100-115.
[6] Chen, X.R. (1991), 'Introduction to change point statistical analysis', Mathematical Statistics and Management, 10, pp. 52-59.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
March-April 2022 Page No. 1188-1196
Article History: Received: 08 February 2022, Revised: 10 March 2022, Accepted: 02 April 2022, Publication: 30 April 2022

[7] Bai, J. (2010), 'Common breaks in means and variances for panel data', Journal of Econometrics, 157, pp. 78-92.
[8] Wu, Y. (2020), 'Estimation of common change point and isolation of changed panels after sequential detection', Sequential Analysis, 39, pp. 52-64.
[9] Hou, B.S. (2016), 'Breakpoint analysis of mean and variance in panel data', Hangzhou: Zhejiang University.
[10] Liu, X., Zhao, W.Z., and Xue, Y.X. (2017), 'Least squares estimation of variance gradient change points in panel data', Journal of Basic Science of Textile University, 30, pp. 357-363.
[11] Lu, S.X., Xu, Q.F., Jiang, C.X., Liu, Y.Z., and Kusiak, A. (2022), 'Probabilistic load forecasting with a non-crossing sparse-group lasso-quantile regression deep neural network', Energy, 242, pp. 77-85.
[12] Ding, M.Y., Tan, C.C., and Zhou, J.Z. (2019), 'Research on meteorological variability based on panel data analysis', Journal of Hefei University of Technology, 42, pp. 570-576.
[13] Li, Y.P. (2018), 'Multivariate detection based on binary segmentation and its application in finance', Harbin: Harbin Institute of Technology.
[14] Han, S.E., Tian, Z., and Wu, X.Q. (2006), 'Multivariate test of a class of stock market volatility prediction model', System Engineering Theory and Practice, 36, pp. 94-101.
[15] Zhu, H.J., Sarafidis, V., and Silvapulle, M.J. (2020), 'A new structural break test for panels with common factors panel data models with multiple time-varying individual effects', Econometrics Journal, 23, pp. 137-155.