The Social Sciences 13 (3): 766-768, 2018
ISSN: 1818-5800

Structural Breaks in Real Time: Evidence from the Firm Level of US Firms

[1]Nur Syazwani Mazlan and [2]George Bulkley
[1]Department of Economics, Faculty of Economics and Management,
Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[2]Department of Accounting and Finance, School of Economics, Finance and Management,
University of Bristol, BS8 1TN, Bristol, United Kingdom

Abstract: We study the process of learning about breaks in the fundamentals of stock prices of the US firms in real time without the benefit of hindsight. The identification of structural breaks takes several years which in turn affects the speed of post-break revision of the new parameters of data generating process. We look at different procedures for break detection, investigate their success and also examine how quickly and efficiently errors get corrected in real time.

Key words: Structural break, learning, real time, error, investigate, fundamentals

INTRODUCTION

The investment in equity or stock is subject to a daunting array of uncertainties. One of the biggest uncertainties is about the parameters of the data generating process. Relatively short samples mean sampling errors will always be a significant issue. This difficulty is compounded by the risk of structural breaks.

Forecasts are typically based on the assumption that model parameters are constant but in practice this may simply not be the case. The investors are aware that structural breaks may occur. However, the problem arises when investors observe an outlier from what they believe is the return or dividend generating process. They have to decide whether this observation is indeed an outlier in a model that is unchanged or whether it is a sign that a structural break has occurred. Like everyone else, investors do not have the benefit of hindsight in the present tense.

Mazlan and Bulkley (2015) identify structural breaks or instability in the dividend of US firms at the firm level by utilising the method developed by Bai and Perron (2003) in which the multiple structural change(s) models are estimated based on deterministic econometric approach. The main ingredient of the Bai and Perron (2003) method is the dynamic programming algorithm. The method of break detection, global optimization estimates the breaks according to the number of breaks pre-specified in the program. The results from their event study support the idea that abnormal returns should reflect the market's attempt to infer the existence of breaks. In this study, we study the learning process of structural breaks by utilising the same firm-level dividend dataset of the US firms in real time. We examine the performance of the techniques used for the identification of optimal breaks from Bai and Perron (2003) in terms of their efficiency and reliability in the midst of errors.

MATERIALS AND METHODS

Structural break analysis: We utilize the Bai and Perron (2003) program that allows for the construction of estimates of the parameters in models with multiple structural breaks. The algorithm of this program is based on the principle of dynamic programming and information criteria and sequential hypothesis testing give the optimal number of breaks. Besides that it is also designed to construct confidence intervals and test for structural change. We can also estimate either pure or partial structural change models and choose the options whether to allow for heterogeneity and/or serial correlation in the data and the errors across segments or not. The multiple linear regression models with m breaks (m +1 regimes) are described as:

$$\begin{aligned} y_t &= x_t^{'}\beta + z_t^{'} d_1 + u_t,\ t = 1, \dots, T_1 \\ y_t &= x_t^{'}\beta + z_t^{'} d_2 + u_t,\ t = T_1 + 1, \dots, T_2 \\ &\vdots \\ y_t &= x_t^{'}\beta + z_t^{'} d_{m+1} + u_t,\ t = T_{m+1} + 1, \dots, T \end{aligned} \quad (1)$$

Corresponding Author: Nur Syazwani Mazlan, Department of Economics, Faculty of Economics and Management, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

where, y_t is the observed dependent or response variable at time t; $x_t(p\times1)$ is the vector of variable(s), fixed throughout the analysis; $z_t(q\times1)$ is the vector of variable(s) subject to structural breaks at time t, a and a_j (j = 1, ..., m+1) are the vectors of coefficients of x_t and z_t, respectively u_t is the error or disturbance at time t. The maximum number of breakpoints is given by m. For the purpose of our structural break analysis we consider two different (general) structural break models as the following. Trend-stationary break model (Model 1):

$$y_t = f(t) + u_t \tag{2}$$

Where:

t = Time
f = A deterministic (linear) function in which f(t) = α
u_t = The disturbance at time t

The variable(s) subject to breaks is given by z_t = {f(t)} whereas xt = {}. It is a trend stationary break model when {f(t)} = α +βt. Autoregressive break model (Model 2):

$$y_t = a + y_{t-1} + u_t \tag{3}$$

Where:

t = Time
α = Drift
y_t-1 = The lag of dependent variable or unit root term
u_t = the disturbance at time t

The variable(s) subject to breaks is given by z_t = {a, y_t-1} whereas x_t = {}.

Real-time analysis: In general, following Clements and Galvao (2013) we have access to the "vintage" t-values of the observations on y up to time period T-1 where "vintage" is defined as the information set that one has available in hand at a given or specific date and the compilation of such vintage is the "real-time data set" (Croushore and Stark, 2003). The T-vintage which can be written as {y_t T} t = 1, 2, T-1. This is also called the latest available T-vintage whereas the previous vintages, for example, the T-j vintage is {y_t T-j} for j = 1, 2, 3 and where t = 1, 2, T-j-1. When we have the full data set with hindsight, I have the T-vintage in which the true breaks are detected as in the previous chapter. The regression model for T-vintage with m breaks (m+1 regimes) of interest is:

$$\begin{aligned} y_t^T &= x_t'^T ß + z_t^T d_1 + e_t^T,\ t = 1, \dots, T_1 \\ y_t^T &= x_t'^T ß + z_t^T d_2 + e_t^T,\ t = T_1 + 1, \dots, T_2 \\ &\vdots \\ y_t^T &= x_t'^T ß + z_t^T d_{m+1} + e_t^T,\ t = T_{m+1} + 1, \dots, T-1 \end{aligned} \tag{4}$$

The true set of breaks is given by {T_k} where k = 1, 2, …, m where m is the maximum number of break allowed in the empirical exercise. For the real time analysis we carry out the structural breaks analysis of the Bai and Perron (2003) program by using all the previous vintages that I have, i.e., {y_t T-j} for j = 1, 2, 3, …, and where t = 1, 2, …, T-j-1:

$$\begin{aligned} y_t^{T-j} &= x_t'^{T-j} ß + z_t^{T-j} d_1 + e_t^{T-j}\ t = 1, \dots, T_1 \\ y_t^{T-j} &= x_t'^{T-j} ß + z_t^{T-j} d_2 + e_t^{T-j}\ t = T_1 + 1, \dots, T_2 \\ &\vdots \\ y_t^{t-j} &= x_t'^{t-j} β + z_t'^{t-j} δ_{m+1} + e_t^{t-j},\ t = T_{m+1} + 1, \dots, T-j-1 \end{aligned} \tag{5}$$

With the benefit of hindsight that a break had occurred at T_k at 5% significance level we would expect to find the same break at $T_k$as more data arrive. For instance we would expect to detect a break at a past date, T_k at T_k+1 by using T_k+1-vintage, i.e., {$y^T{}_t$}t = 1, 2, …, T-1. Similarly, we would always expect to detect the same break in the next periods as more data become available. However, there are times that this happens not to be the case. GThe error in judgement in real time may present in the form of type 1 and 2 error.

Type 1 error: This happens in the case of a rejection of the null hypothesis of no break when it is actually true, i.e., a break was identified when there was no break.

Type 2 error: This happens in the case of a failure to reject the null hypothesis when it is actually not true, i.e., a break was not identified when there was a break. In the context of our structural break analysis in real time if we were to explain judgement error in the form of type 1 and 2 error as how it would naturally have been thought of this would lead us to some confusion which can further lead to misleading analysis. For the detection of structural breaks in real time, the following would have been our set of hypotheses:

- Null hypothesis; there is no (true) break(s) at data point t
- Alternative hypothesis; there is a (true) break(s) at data point t

Essentially, we investigate the following:

- How long does it take for correction to happen in real time
- How soon do we learn about the breaks in real time

RESULTS AND DISCUSSION

How long does it take for correction to happen in real time: When the breaks found are not the true breaks we investigate the correction speed of realising that mistakes have been made in the selection of optimal breaks in real time. Table 1 reports the results for the correction time taken by each procedure and break model in real time. For the break model of trend stationary (Model 1) we find that sequential and repartition on average, take around three quarters or less than a year to correct the judgement error or mistakes made in real time. BIC and LWZ are seen to be slightly slower and the time recorded for the correction to happen is around five quarters on average. As for autoregressive break model (Model 2), the correction time is slightly shorter in which it is around two quarters for the correction to take place for sequential and repartition. For LWZ, this is around four quarters or 1 year and for BIC, it is around 5 quarters which is the longest compared to other techniques for optimal breaks selection considered in this study.

How soon do we learn about the breaks in real time: Table 2 reports the period required for the first-time detection of breaks in real time. A "perfect" outcome here would be if we were to find this true breaks right after they happen. For the break model of trend stationary (Model 1) we observe that it takes on average, 36- 54 quarters or 9 to slightly >13 years for the techniques for optimal breaks selection considered in this study to learn about the (true) breaks in real time. Repartition is the quickest in detecting the breaks in real time followed by sequential, LWZ and BIC, respectively. For the unit root break model (Model 2), we observe that the learning time is just slightly quicker, although not much different on average. It takes on average, 35-52 quarters or 8.75-13 years for the techniques for the optimal breaks selection considered in this study to learn about the (true) breaks in real time. Again, Repartition is seen to be the quickest in doing so followed by sequential, LWZ and BIC, respectively.

Table 1: Correction in real time

	Descriptive statistics						
Process	N	Mean	Median	SD	Min.	Max.	Range
Model 1 (Trend stationary)							
Sequential	4722	3.12	3	1.72	1	8	7
Repartition	3957	2.86	3	1.64	1	8	7
BIC	4834	5.06	4	2.21	1	12	11
LWZ	3270	4.50	4	2.11	1	198	197
Model 2 (Autoregressive)							
Sequential	1863	2.41	2	1.49	1	7	6
Repartition	1863	2.41	2	1.49	1	7	6
BIC	2508	4.77	5	2.13	1	12	11
LWZ	1947	3.99	4	1.61	1	198	197

Table 2: Learning about breaks in real time

	Descriptive statistics						
Process	N	Mean	Median	SD	Min.	Max.	Range
Model 1 (Trend stationary)							
Sequential	787	39.55	43	10.48	17	56	39
Repartition	767	35.70	43	12.47	12	56	44
BIC	777	53.59	58	14.70	23	68	45
LWZ	654	48.35	55	10.42	30	55	25
Model 2 (Autoregressive)							
Sequential.	345	35.97	41	10.81	14	41	27
Repartition	345	34.85	43	12.57	12	47	35
BIC	425	51.79	58	14.23	27	64	37
LWZ	301	42.50	48	13.64	22	55	33

CONCLUSION

We study the dynamics of learning about the breaks in real time by using the dataset of the dividend process of the US firms at the firm level. The previous study, Mazlan and Bulkley (2015) present the evidence of breaks in the firm-level of US dividends which are then assumed to be the true breaks for the real-time analysis in this paper. In real time we show that mistakes are made in the identification of breaks and it takes some time to learn about the breaks and also for the correction to take place. The judgement error or mistakes can happen in several ways. For instance we observe that the true breaks are not found at some dates and I also observe that sometimes, the breaks that are not the true breaks are found at some dates.

We look at the speed of learning about the (true) breaks as well as the speed of correcting the mistakes in real time in which we find that the learning time recorded is around 9 to <14 years for sequential, repartition, BIC and LWZ. For the correction speed, this is around 3-5 quarters or 9-20 months. Repartition is seen to perform better than others as it records the shortest learning and correction time in my investigation. Overall, the autoregressive (Model 2) is seen to perform a bit better than the trend stationary model (Model 1).

REFERENCES

Bai, J. and P. Perron, 2003. Computation and analysis of multiple structural change models. J. Appl. Econ., 18: 1-22.

Clements, M.P. and A.B. Galvao, 2013. Real-time forecasting of inflation and output growth with autoregressive models in the presence of data revisions. J. Appl. Econ., 28: 458-477.

Croushore, D. and T. Stark, 2003. A real-time data set for macroeconomists: Does the data vintage matter?. Rev. Econ. Stat., 85: 605-617.

Mazlan, N.S. and G. Bulkley, 2015. How stable is the underlying process of stock prices? Empirical evidence of structural breaks in the firm-level dividend of the US firms. Int. J. Econ. Manage., 9: 342-355.