Economics Bulletin

Volume 31, Issue 3

A simple method for variance shift detection at unknown time points

Loredana Ureche-Rangau
Université de Picardie Jules Verne, CRIISEA

Franck Speeg
MSC Student, EM-Lyon

Abstract

Financial literature considers volatility as a good proxy for the risk level and thus the crucial parameter in many financial techniques and strategies. As such, the aim of this paper is to analyse the evolution of time series volatility and detect significant long-term variance changes. Building up on the variance ratio detection technique introduced by Tsay (1988), our paper extends it in two ways: first, we propose the computation of a moving variance ratio implemented on a selected part of the series, thus reducing the amount of calculus and increasing the reliability and second, as in reality permanent variance changes are almost inexistent, we proceed to an adjustment on a specified part of the series only after the detected variance change. Our moving variance ratio technique proves its efficiency in detecting variance changes and removing them from the series, both on simulated and real financial data. More specifically, two significant variance changes are detected within the series of the Hang Seng daily log-returns between 1994 and 2007: the first one on August 15, 1997 and can be linked to the Asian financial crisis, and the second one on July 27, 2001 corresponding to the beginning of a high volatility regime in emerging markets following the Internet bubble crash along with the first signs of the financial crisis in Argentina.

Citation: Loredana Ureche-Rangau and Franck Speeg, (2011) "A simple method for variance shift detection at unknown time points", *Economics Bulletin*, Vol. 31 no.3 pp. 2204-2218.
Submitted: Jul 05 2011. **Published:** July 28, 2011.

1. Introduction

Structure changes and outliers are very common in financial time series data. Most of these extraordinary observations can easily be attributed to dramatic events such as wars, crashes or political changes. The presence of level shifts and outliers can mislead time series analysis and thus lead to erroneous conclusions due to important model misspecification. If the type and date of the disturbances are known, their effects can easily be controlled. However, in practice, the date and magnitude of such events are seldom known *ex ante*. Therefore, detecting and handling outliers, level shifts and variance changes in time series becomes crucial. Several approaches were proposed to identify these abnormal patterns.

Abraham and Box (1979) use a Bayesian method, whereas Martin and Yohai (1986) consider outliers as anomalies generated by a deterministic probability distribution. Fox (1972) establishes two parametric models which have been adopted by Chang (1982), who develops an iterative procedure for outlier detection. This iterative procedure is widely applied, especially by Chang and Tiao (1983), Hillmer *et al.* (1983) and Tsay (1986). Tsay (1986) uses both Chang and Tiao (1983) and the Extended Sample Autocorrelation Function (ESACF) model identification method developed by Tsay and Tiao (1984). This modelling enables the distinction between an additive outlier, i.e. a “gross error” model that affects a single observation, and an innovational outlier, i.e. a disturbance that affects the process at the time it occurs but also the following observations. It is an iterative procedure based on a simple computation using basic linear regression techniques that enables outlier detection and classification along with model specification that includes the identified outliers.

Outlier detection highlights non-linearity in most financial time series. Generalized Autoregressive Conditionally Heteroscedastic (GARCH) variance models, introduced and developed by Engle (1982) and Bollerslev (1986) are also able to capture this non-linearity, but they do not always provide a satisfactory explanation for the non-Gaussian behaviour. Although model misspecifications could be the source of this non-normality, the variety of models that exhibit such a departure from normality in the residuals supports the idea that excess skewness and kurtosis are caused by the presence of outliers in financial time series. Hence, a recent strand of the financial literature focuses on the detection and correction qualities of different outlier identification methods within the GARCH framework (see for example Franses and Ghijsels, 1999; Zhang and King, 2005; Bali and Guirguis, 2006 or Ané *et al.*, 2007 among others).

The detection of outliers is a significant step towards level shift and variance change detection. An outlier can be defined as a single, big and unpredictable value. Outliers are usually disparate within the series, even though sometimes they may cluster, especially over periods of consecutive macroeconomic events. In contrast, level shifts and variance changes are considered as long-term changes. The detection of level shifts in time series starts when the impact of a specified intervention has to be measured. Early evidence on level shifts focuses on environmental (Box and Tiao, 1975) or social issues (Harvey and Durbin, 1986; Hsu, 1977). Remarkable contributions in detecting structural changes with unknown break dates are due to Bai and Perron (1998, 2003a, 2003b), synthesized in Perron (2006). Their different papers on multiple structural changes introduce and confront different issues for univariate regressions, namely the selection of the number of breaks, the algorithms to compute the estimates, the consistency of the break dates and the confidence intervals for these dates. Alternative approaches and extensions of this framework, including multivariate analysis, can be found in Liu *et al.* (1997), Bai *et al.* (1998), Bai (2000), Perron and Qu (2006), Kejriwal and Perron (2006a,b), Qu and Perron (2007a), among others.

Despite this wide amount of research, testing for structural changes in the variance of the regression error is rather scarce. Tsay (1988) proposes a unified method for detecting and handling outliers, level shifts and variance changes for univariate time series based on variance ratios. Alternative, ad-hoc approaches consist in applying standard sup-Wald type tests (Bai and Perron, 1998) to identify changes in the mean of the absolute value of the estimated residuals, or extensions of the CUSUM of squares test (introduced by Brown e*t al.*, 1975), as in Deng and Perron (2008)[1], following Inclan and Tiao (1994). However, even though these tests provide accurate break point

[1] See also Horvath (1993), Davis *et al.* (1995), Qu and Perron (2007a), Perron and Zhou (2008) and Zhou and Perron (2008).

detections, they do not provide any indication on the impact induced by these breaks on the variance, namely in terms of permanent or transient change.

In this paper we aim at analysing the evolution of time series volatility and detecting significant long-term variance changes. Given that volatility is considered a good proxy for the risk level, identifying the existence of different levels of variance could then be used by many financial techniques and strategies, like derivatives pricing and hedging. Volatility fluctuates, never diverges, and sometimes undergoes instant changes that could be transient or permanent. Permanent, i.e. definitive, changes are very rare, even in very volatile markets. It is noticeable that due to extraordinary events such as wars, crisis and political changes, most financial series experience a change in the mean level very often accompanied by a change in variance. This abnormal level of volatility may bias the estimation of risk. Detecting and removing such anomalies would provide a better picture of the evolution of past volatility. Thus, a more accurate pricing of derivatives could be obtained. If no adjustments were made, the level of volatility could over/understate the level of risk with direct consequences for the price and efficiency of the hedge. For example, the derivative would be over/under-priced and arbitrage strategies could become potentially profitable.

We develop the variance ratio detection technique introduced by Tsay (1988) in two ways. First, Tsay (1988) uses a variance ratio that takes into account all the residuals after and before a date "t". We implement this computation only on a selected part of the series. This will thus reduce the amount of calculus and increase the reliability. Second, in reality, permanent variance changes are extremely rare. The change in variance does not last infinitely and its amplitude reduces gradually. This means that only a specified part of the series after the detected variance change should be adjusted in contrast with Tsay (1988) procedure, in which the standardization is applied to all the observations following the detected change.

This paper is organized as follows. The next section explains, in a simulation framework, the variance change detection technique that we propose. Then we apply this technique on real data, namely the Hang Seng index from 1994 to 2007. A comparison with the results provided by a standard Bai and Perron (1998) test for break points identification is also provided. The last section concludes and indicates potential further developments.

2. Variance changes detection using iterative procedures

The aim of this section is to present and test the moving variance ratio technique proposed in this paper on simulated data. More specifically, we will focus on variance shift detection. As mentioned previously, our framework builds on the method introduced and developed by Tsay (1988), i.e. a variance ratio method using an iterative procedure decomposed into 4 consecutive steps: Gaussian Autoregressive Moving Average (ARMA) model specification, variance ratio computation using the residuals of the ARMA model from the first step, type and date of the variance change identification (the two extreme values of the variance ratio are identified, say $\lambda_{V,\max}$ the highest value and $\lambda_{V,\min}$ the lowest; then, $\lambda_V = Max(\lambda_{V,\max}, \lambda_{V,\min}^{-1})$ and is compared with a pre-specified critical value C)[2], variance change effect removal through a standardization process followed by ARMA model reestimation. We start by introducing d denoting the distance, i.e. length of the selected part of the series used for the calculation of the variance ratio. We test the moving variance ratio technique on simulated series containing a single variance change, and then extend it on series exhibiting several simulated variance changes.

The first series is a vector whose components are independent, normally distributed random variables with zero mean and known variances, including a change in variance at time d_0. For each d_i, the variance ratio is not computed on the whole series but only on a part of it. Hence, d is defined as the distance from d_i on which we compute the variance ratio $\hat{r}_i$. We compute the ratio of the variance of the d-data following d_i on the variance of the d-data preceding d_i and repeat it for each d_i. When an upward (resp. downward) movement in variance occurs, we notice a corresponding rise (resp. drop) in the value of this ratio. If the change in variance is significant, a small value for d

[2] If $\lambda_V > C$, a variance change occurs at the date at which λ_V is detected, say d_0, with C= 3.5

may be enough to efficiently detect the exact changing point. On the contrary, when the change in variance is small, a larger value of d may be needed. Hence, the moving variance ratio $\hat{r}_i$ can be defined as follows:

$$\forall\, i \in \{d+1; N-d\},\ \hat{r}_i = \frac{\sum_{t=i}^{i+d} (Z_t - \overline{Z}'_{i \le t \le i+d})^2}{\sum_{t=i-d}^{i-1} (Z_t - \overline{Z}'_{i-d \le t \le i-1})^2} \qquad (1)$$

where N is the sample size and $\overline{Z}'$ the mean of the d-values on which the variance is computed.

The first simulation that we implement consists in generating 10,000 data-points with a change in variance at time $d^* = 5001$. The 56 generated series are the combination of eight 5,000x1 vectors of independent, normally distributed random variables described by $N(0,\sigma)$ with $\sigma = \{1,2,3,4,6,8,15,20\}$.

To detect the exact date of the variance change, we first need to define $\lambda_{V,\max}$ denoting the largest $\hat{r}_i$ and $\lambda_{V,\min}$ for the smallest $\hat{r}_i$. Thus, the detected change d_0 occurs at $\lambda_{Vd_0} = Max(\lambda_{V,\max}, \lambda_{V,\min}^{-1})$.

The following table shows the impact of different values of d. We can notice that when the standard deviation of the first part of the series is much lower than the standard deviation of the second part, a d equal to 20 is enough to almost perfectly detect the variance change at the exact time point $d^* = 5001$. Moreover, when the standard deviations are quite close, d has to be higher to allow a significant detection. 39 of these 56 series only need d equal to 20 to detect the exact date with less than 0.2% divergence. The exact changing point of these 56 series has been detected with an error that is less than 2.5%.

Table 1. Variance Shifts Detection when N=10,000 and d*=5001. The series follow a Gaussian distribution described by N(0,σ)

N = 10,000	d=20	d=80	Part 1							
d* = 5001	d=60	d>=200	N(0,1)	N(0,2)	N(0,3)	N(0,4)	N(0,6)	N(0,8)	N(0,15)	N(0,20)
Part 2	N(0,1)	Max ri		2,669	4,026	5,797	3,797	6,175	5,558	7,210
		Min ri		0,263	0,139	0,057	0,023	0,019	0,007	0,003
		λv		3,809	7,188	17,502	43,397	52,467	148,464	324,358
		Detected do		**5005**	**5003**	**4998**	**5005**	**5005**	**5005**	**5006**
	N(0,2)	Max ri	3,767		2,116	2,545	3,930	6,384	6,384	7,210
		Min ri	0,310		0,436	0,243	0,103	0,060	0,021	0,010
		λv	3,767		2,294	4,116	9,743	16,632	47,104	98,895
		Detected do	**4977**		**4977**	**4998**	**5001**	**5002**	**5004**	**5004**
	N(0,3)	Max ri	11,915	4,833		1,738	2,322	2,483	6,377	7,210
		Min ri	0,200	0,245		0,477	0,275	0,232	0,089	0,040
		λv	11,915	4,833		2,096	3,631	4,310	11,242	24,818
		Detected do	**5001**	**5000**		**5050**	**5019**	**5002**	**5000**	**5007**
	N(0,4)	Max ri	17,486	7,141	4,478		1,501	2,549	4,087	7,210
		Min ri	0,200	0,245	0,238		0,446	0,308	0,126	0,091
		λv	17,486	7,141	4,478		2,241	3,249	7,944	11,037
		Detected do	**5000**	**4998**	**4994**		**5126**	**5000**	**4999**	**5001**
	N(0,6)	Max ri	34,612	16,136	9,343	2,786		1,446	3,216	7,210
		Min ri	0,200	0,245	0,238	0,365		0,581	0,188	0,128
		λv	34,612	16,136	9,343	2,786		1,722	5,323	7,790
		Detected do	**5004**	**5006**	**5006**	**5010**		**5014**	**5000**	**5001**
	N(0,8)	Max ri	33,242	13,784	7,513	4,087	1,753		2,642	7,210
		Min ri	0,200	0,245	0,238	0,294	0,676		0,226	0,126
		λv	33,242	13,784	7,513	4,087	1,753		4,420	7,956
		Detected do	**5000**	**4998**	**4994**	**4991**	**4937**		**5000**	**5005**
	N(0,15)	Max ri	161,156	67,138	32,702	12,504	8,107	4,596		1,746
		Min ri	0,123	0,145	0,200	0,126	0,200	0,311		0,372
		λv	161,156	67,138	32,702	12,504	8,107	4,596		2,685
		Detected do	**5002**	**5002**	**4994**	**5004**	**5004**	**4998**		**5010**
	N(0,20)	Max ri	225,948	80,788	47,160	14,953	7,366	5,362	1,698	
		Min ri	0,064	0,064	0,064	0,096	0,144	0,214	0,677	
		λv	225,948	80,788	47,160	14,953	7,366	5,362	1,698	
		Detected do	**5001**	**5001**	**4998**	**4990**	**4999**	**5010**	**4998**	

This first illustration shows that in the case of a single variance change, the moving variance ratio technique is able to accurately detect the shift in variance at the exact point. The upper triangle represents the down-changes while the lower triangle represents the up-changes. Black cells correspond to cases where the variance does not experience any change. We can also notice an asymmetry between the upper and lower parts of the table. Indeed, 11 of the 17 series which need a d higher than 20 are situated in the upper part of the table corresponding to down-changes in variance. Hence, down-changes are harder to properly detect than up-changes which in turn may come from the fact that volatility increases have larger impact than volatility decreases, i.e. volatility asymmetry.

The second simulation consists in the detection of two variance changes. In the case of multiple changes, an adjustment of the series has to be conducted to enable the next variance change detection. This adjustment can be assimilated to a standardization of the series and will be detailed at step 3. The stepwise procedure thus becomes:

Step 1. Compute the moving variance ratio $\hat{r}_i$ on the series.

Step 2. Find the two extreme values $\lambda_{V,\max}$ and $\lambda_{V,\min}$. and then compute $\lambda_{Vd_0} = Max(\lambda_{V,\max}, \lambda_{V,\min}^{-1})$.

$\lambda_{Vd_0}^{*}$ denotes the selected value among $\lambda_{V,\max}$ and $\lambda_{V,\min}$.

Step 3. Adjust the series Z_i by a standardization process Z_i^* as follows:

$$Z_i = \begin{cases} Z & \text{if } t<d_0 \\ \overline{Z} + \lambda_{Vd_0}^{*-1/2}(Z_i - \overline{Z}) & \text{if } t\geq d_0 \end{cases} \tag{2}$$

where $\overline{Z}$ is the sample mean of Z_i and go back to Step 1.

Repeat this process until no other significant variance changes are detected.

The aim of the standardization process is to remove the impact of the variance shift, hence allowing the detection of the following change. With no adjustment, the biggest variance change would always be detected. Once all the variance changes are detected, only small variations may still be observed on the lessened series. The modification we propose with respect to Tsay's (1988) procedure is introduced at Step 3. Indeed, we use the initial extreme value $\lambda_{Vd_0}^{*}$ instead of the modified maximum λ_{Vd_0} as done by Tsay. In the case of an up-change $\lambda_{Vd_0}^{*} = \lambda_{Vd_0} = \lambda_{V,\max}$, whereas in the case of a down-change $\lambda_{Vd_0}^{*} = \lambda_{Vd_0}^{-1} = \lambda_{V,\min}$. This means that Tsay's procedure can detect a down-change but the standardization process Z_i^* does not perform the right correction.

In order to study a case with two variance changes, we first simulate a series of 10,000 data points with a change at each third of the series. Hence, the first change occurs at $i = 3334$ and the second one at $i = 6667$. The three parts of the series contain centred Gaussian data with three different standard deviations equal to 2, 4 and 8 respectively. We take d equal to 50 for all variance shift detections since this value seems to be enough for most of the detections.

Table 2. Results of the 2 detections when σ = 2, 4 and 8 respectively.

	1st Detection	2nd Detection
Max $\hat{r}_i$	3.593392207	3.145228444
Min $\hat{r}_i$	0.43531926	0.43531926
λ_{Vd_0}	3.593392207	3.145228444
Detected d_o	3334	6667
Observation nb: 1st change	3334	3334
Observation nb: 2nd change	6667	6667

In this multiple variance shifts example, the moving variance ratio technique accurately detects both variance changes at the exact dates $i = 3334$ and $i = 6667$. The two following figures show the different steps of the detection and standardization process.

Figure 1. Evolution of the residuals after each standardization when σ = 2, 4 and 8 respectively.

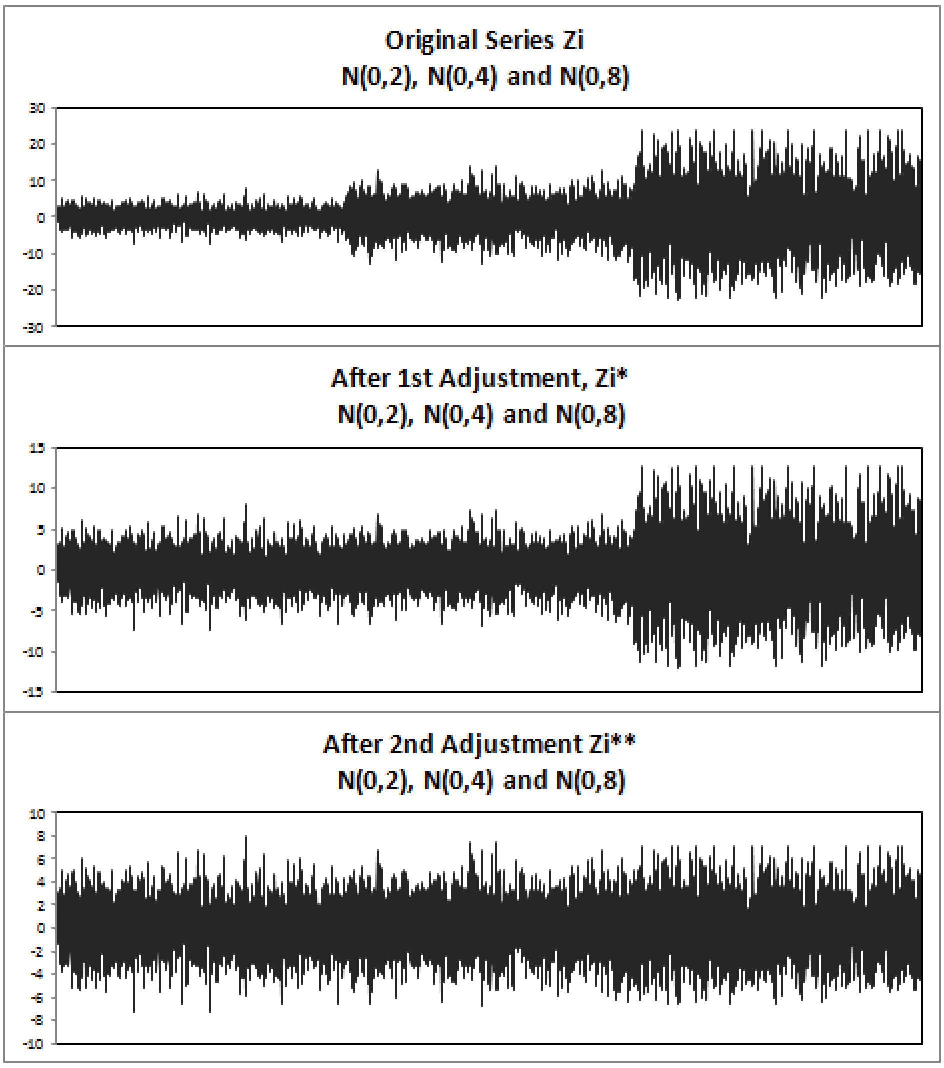
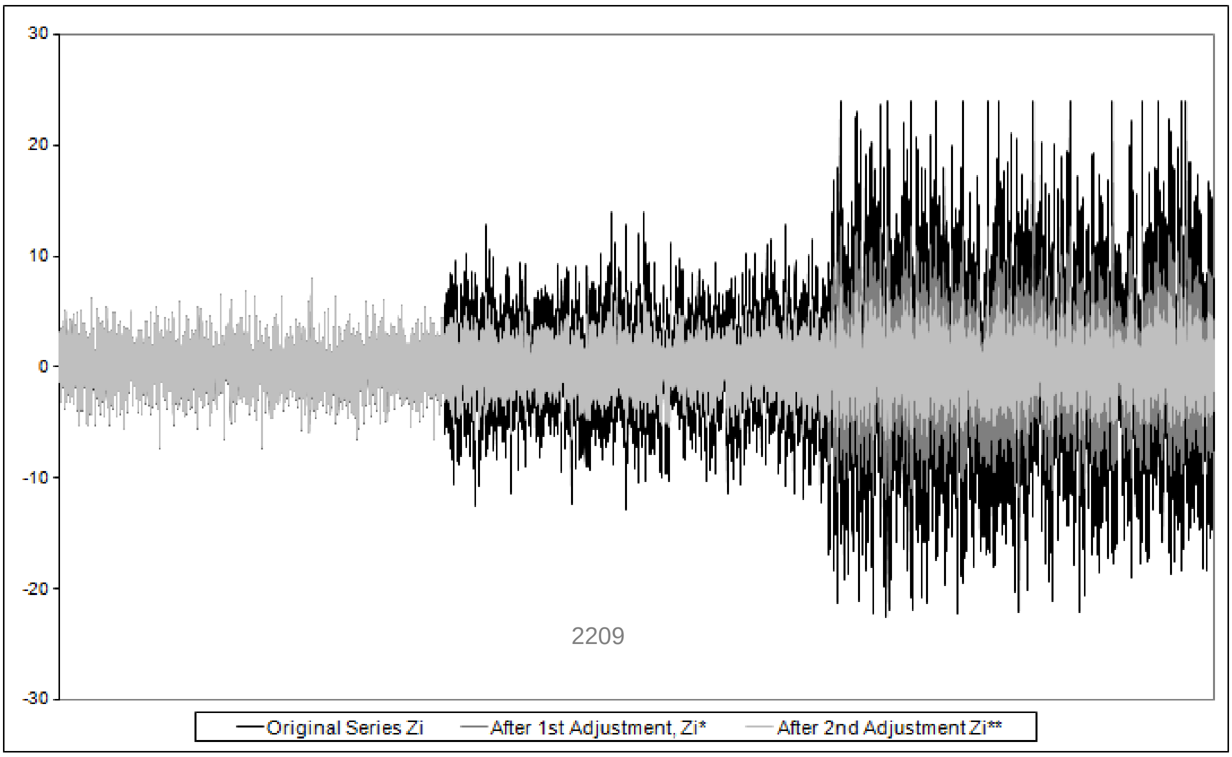


Figure 2. Evolution of the residuals after each standardization when σ = 2, 4 and 8 respectively.



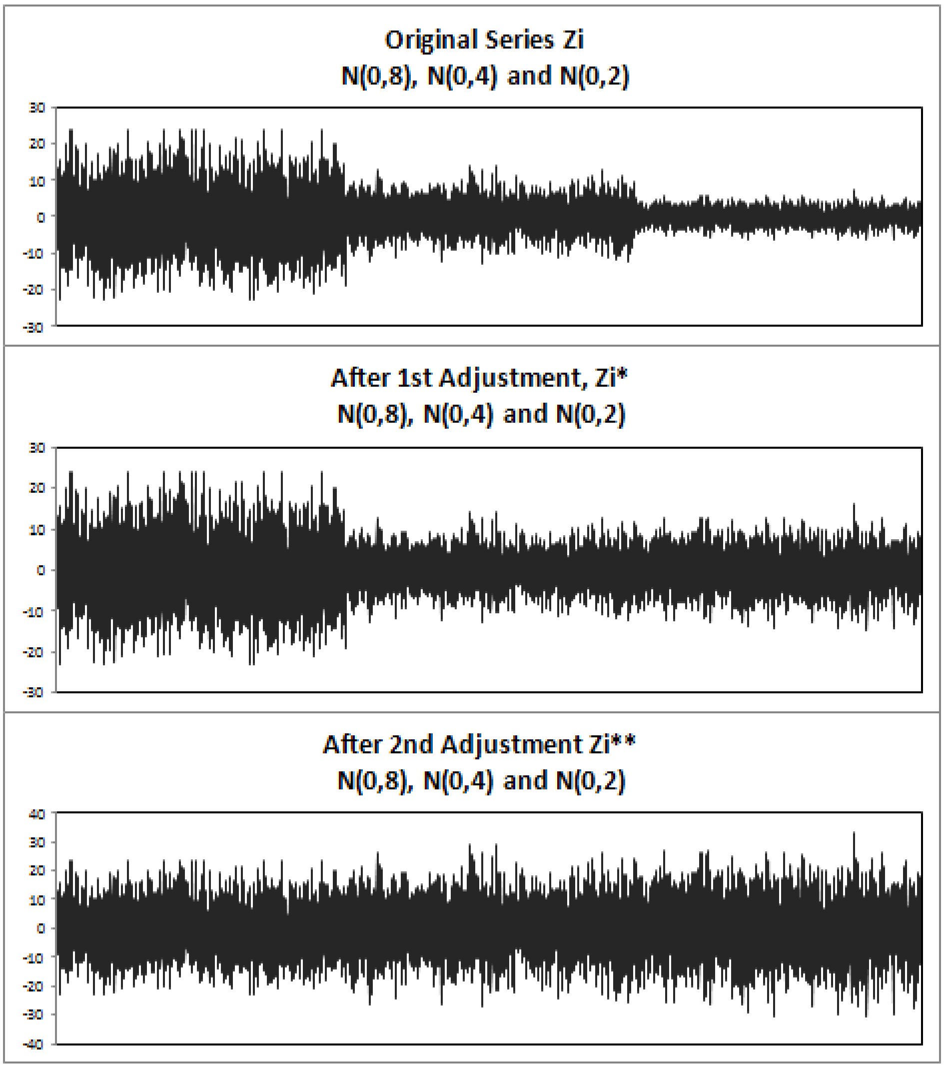
In a second example, we simulate 10,000 data points with a change at each third of the series and three different standard deviations of 8, 4 and 2 respectively.

Table 3. Results of the 2 detections when σ = 8, 4 and 2 respectively.

	1st Detection	2nd Detection
Max $\hat{r}_i$	2.721904899	2.721904899
Min $\hat{r}_i$	0.204068686	0.226943962
λ_{Vd_0}	4.900310874	4.406374113
Detected d_o	6667	3329
Observation nb: 1st change	3334	3334
Observation nb: 2nd change	6667	6667

In this second example, the two variance shifts are down-changes. The identified d_0 corresponds to the exact date for the first down-change, and is situated only 5 observations before the exact second down-change. We can thus state that the accuracy of this technique is also shown for down-changes. The fact that the detection is not as perfect as in the case of up-changes may come from the presence of the asymmetry in volatility as stated previously.

Figure 3. Evolution of the residuals after each standardization when σ = 8, 4 and 2 respectively.

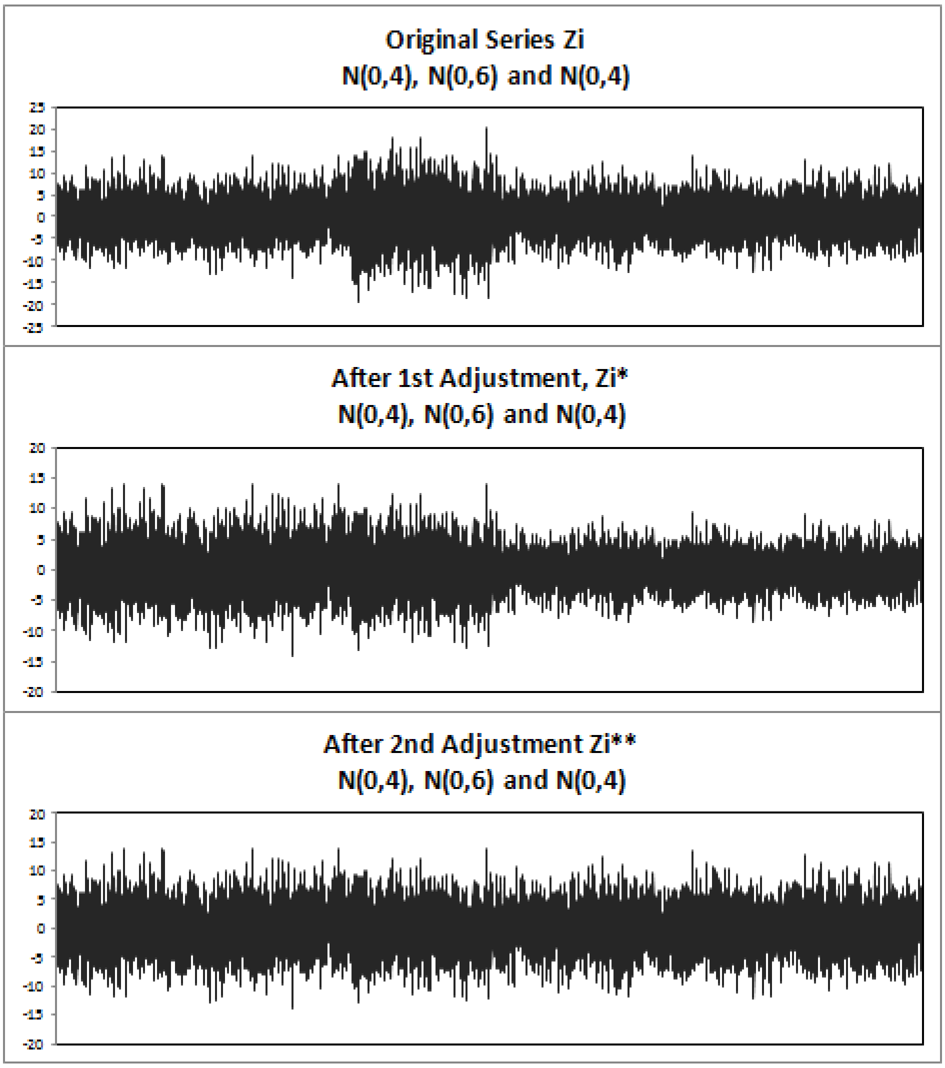


For the last example, we simulate 10,000 observations with the first variance change at $i = 3334$ and the second at $i = 5001$. The three different standard deviations are 4, 6 and 4 respectively. In this example, we are in the case of two small variance changes, upward and then downward. Because of the small variance shift, we set d equal to 150.

Table 4. Results of the 2 detections when σ = 4, 6, 4 respectively.

	1st Detection	2nd Detection
Max $\hat{r}_i$	2.166149627	1.965881614
Min $\hat{r}_i$	0.484617473	0.484617473
λ_{Vd_0}	2.166149627	2.06348317
Detected d_o	3346	5096
Observation nb: 1st change	3334	3334
Observation nb: 2nd change	5001	5001

Figure 4. Evolution of the residuals after each standardization when σ = 4, 6, 4 respectively.



Our technique is accurate for this last example too. The two detected d_0 remain quite close to the real changing dates. We can again notice that the detection of the down-change is less accurate. The adjusted series does not seem to contain any other major variance change. These three examples confirm that this simple moving variance ratio technique is efficient in detecting variance changes and the standardization process uniformly lessened the series[3]. This technique can now be applied for variance shift detections on real financial time series.

3. The moving variance ratio technique applied to financial time series

In this section we test the technique developed in the previous part on the Hong Kong stock market index. We decided to focus on the Hang Seng index in our empirical investigations as Hong Kong provides a good example of a very active medium size stock market. It is a very liquid, but also sometimes extremely volatile market, as for example in 1997, during the Asian financial crisis following the devaluation of the Thai baht.

Step 1: ARMA model specification

We use an ARMA(p,q) model as follows:

$$u_t = \Phi_1 u_{t-1} + \Phi_2 u_{t-2} + \cdots + \Phi_p u_{t-p} + \xi_t + \theta_1 \xi_{t-1} + \theta_2 \xi_{t-2} + \cdots + \theta_q \xi_{t-q} \tag{3}$$

Selecting the order of lags p and q is the result of an arbitrage between efficiency and accuracy. The use of a reduced number of lags may bias the detection of serial correlations at high orders whereas on the contrary, significant low order correlations may be affected by insignificant correlations at higher orders and a loss of degrees of freedom.

Step2: Moving Variance Ratio

The residuals of the ARMA(p,q) process become our Z_i series. Then, we compute the moving variance ratio $\hat{r}_i$ on these residuals as in "(1)". The two extreme values $\lambda_{V,\max}$ and $\lambda_{V,\min}$ can now be identified and $\lambda_{Vd_0} = Max(\lambda_{V,\max}, \lambda_{V,\min}^{-1})$ computed. λ_{Vd_0} is then compared to a pre-specified critical value, say C. This critical value C sets a level of significance for the variance change. If λ_{Vd_0} > C, a variance change occurs at d_i. We set C equal to 4 for the variance change detection on real financial time series[4].

Step3: Limited adjustment

We are now in the real financial series case and permanent variance changes are very rare. The change in variance does not last infinitely and its amplitude decreases gradually. This means that only a specified part of the series after the detected variance change has to be adjusted. We then compute a slightly different ratio $\hat{r}_i'$ where the denominator, i.e. the variance of the d-data[5] preceding the detected variance change at d_0, remains constant.

$$\forall\, i \in \{d_0 ; N-d\}, \hat{r}_i' = \frac{\sum_{t=i}^{i+d} (Z_t - \overline{Z}'_{i \le t \le i+d})^2}{\sum_{t=d_0-d}^{d_0-1} (Z_t - \overline{Z}'_{d_0-d \le t \le d_0-1})^2} \tag{4}$$

One should bear in mind that in the case of a constant variance, this variance ratio is equal to 1. Thus, this ratio converges to 1 when the variance comes back to the level preceding the detected variance change. Therefore, k denotes the number of consecutive variance ratios $\hat{r}_i'$ that are above or below 1 respectively, in the case of an up or down-change. Hence, the adjustment is operated on the k-data following the detected d_0 as follows:

$$Z_i^* = \begin{cases} Z_i & \text{if } i<d_0 \\ \overline{Z}'_{d_0 \le i \le d_0+k} + \lambda_{Vd_0}^{*-1/2} (Z_i - \overline{Z}'_{d_0 \le i \le d_0+k}) & \text{if } d_0<i\le d_0+k \\ Z_i & \text{if } i > d_0+k \end{cases} \tag{5}$$

where $\lambda_{Vd_0}^*$ is either $\lambda_{V,\max}$ or $\lambda_{V,\min}$.

[3] Results for all the simulations we performed are available upon request.

[4] Tsay (1988) sets this critical value at 3.5. In our case, as the computation of the moving variance ratio on a restricted section of the series induces larger values for the ratio we increase the value of C.

[5] We take d=50 for all the detections as specified in the previous part.

The detection process then starts again from Step 1. The modified series Z_i^* is now used as the initial series on which a new ARMA model is applied. Then, the moving variance ratio detection technique is applied to the residuals which become the new series Z_i.

First variance change detection

We use the daily closing price of the Hang Seng index from January 1994 to April 2007 which accounts for 3462 observations. Our data comes from Datastream Thomson Financial. The first step consists in computing the daily logarithmic returns R_i. Figure 5 depicts the evolution of the daily Hang Seng Index returns and Table 5 provides the basic descriptive statistics of this return series.

Figure 5. Hang Seng Index log-returns between 1994 and 2007.

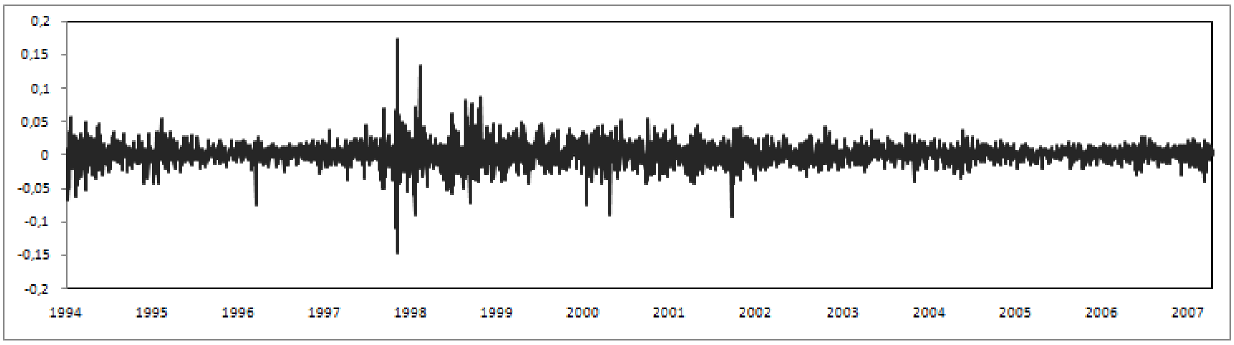


Table 5. Descriptive statistics of the raw Hang Seng Index returns.

Mean (%)	0.0150
Standard deviation (%)	1.5958
Skewness	-0.0935*
Kurtosis	5.9552*
Jarque-Bera	1264.1330 (p-value=0.0000)

* denotes significance at the 5% confidence level.

Our returns series shows an empirical distribution with heavy tails relative to the Gaussian distribution. The series also appears asymmetric. Finally, the Jarque-Bera statistic strongly rejects normality at the standard 5% confidence level.

The best model[6] fitting our returns series is an ARMA(1,1) model (detailed in Table 6) as follows:

$$R_i = 0.00023 - 0.56170 R_{i-1} + 0.62097 \xi_{t-1} \qquad (6)$$

Table 6. The ARMA(1,1) on the raw returns series coefficients and statistics

Variable	Coefficient	p-value
Constant	0.00023	0.57691
AR(1) term	-0.56170*	0.00025
MA(1) term	0.62097*	0.00002

* indicates significance at the conventional 5% risk level.

We then compute the moving variance ratio $\hat{r}_i$ on the residuals of the ARMA(1,1) model specified by "(1)". We thus obtain our series of $\hat{r}_i$, on which we look for the two extreme values $\lambda_{V,\max}$ and $\lambda_{V,\min}$, and finally compute λ_{Vd_0} = Max ($\lambda_{V,\max}$, $\lambda_{V,\min}^{-1}$). If λ_{Vd_0} is significant, a change in variance is

[6] We based our choice on the AIC and Schwarz criteria. The detailed results are available upon request.

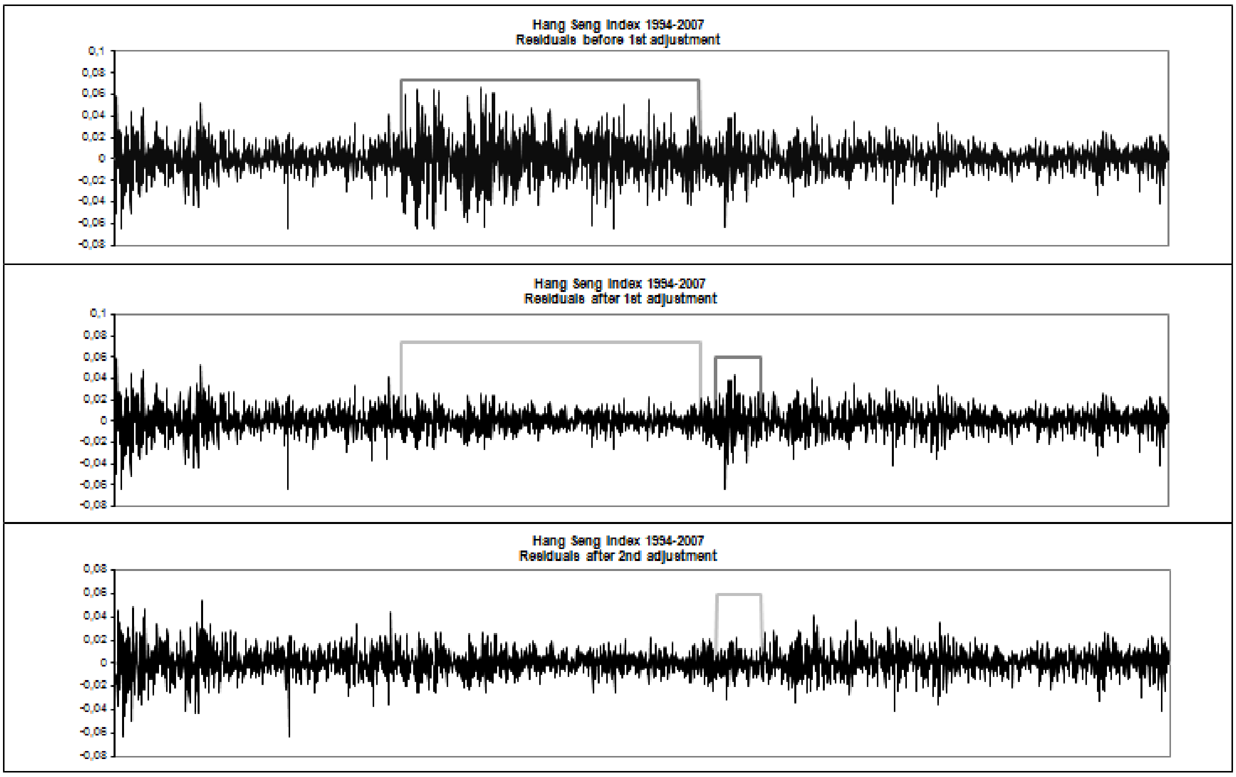
detected and an adjustment has to be made. This first variance change d_0 is detected at $i = 945$ which corresponds to August 15, 1997. This date corresponds to the beginning of a high volatility period on the Hong Kong market due to the spillover effects of the financial crisis from Thailand.

Finally, we need to determine k through the computation of the following variance ratio as in "(4)". k represents the number of consecutive variance ratios $\hat{r}_i'$ above 1. For this first variance change detection k is equal to 980. Hence, the adjustment is operated on the 980 data following the detected d_0 at time $i = 945$ as follows:

$$Z_i^* = \begin{cases} Z_i & \text{if } i<945 \\ \overline{Z}'_{d_0 \le i \le d_0+k} + \lambda_{Vd_0}^{*-1/2}(Z_i - \overline{Z}'_{d_0 \le i \le d_0+k}) & \text{if } 945 \le i \le 945+980 \\ Z_i & \text{if } i>945+980 \end{cases} \quad (7)$$

The upper part of Figure 6 shows the series Z_i and the region on which the adjustment will be operated while the middle part of Figure 6 shows the adjusted series Z_i^*. One can notice that after the adjustment, the high-volatility period has mainly been removed from the series.

Figure 6. Hang Seng index: residuals after the first and second standardizations



Second variance change detection

The best model fitting the new, adjusted returns series is now an ARMA(2,2) process (detailed in Table 7) written as follows:

$$R_i = 0.0000274 + 0.1576 R_{i-1} - 0.6337 R_{i-2} - 0.1862 \xi_{t-1} + 0.6640 \xi_{t-2} \quad (8)$$

Table 7. ARMA(2,2) coefficients and statistics, second change detection

Variable	Coefficient	p-value
Constant	2.7355E-05	0.92097
AR(1)	0.15756*	0.04750
AR(2)	-0.63366*	4.9228E-18
MA(1)	-0.18620*	0.01649
MA(2)	0.66404*	2.8185E-20

* indicates significance at the conventional 5% risk level.

The moving variance ratio is now applied on the residuals of this ARMA(2,2) model. We obtain a series of $\hat{r}_i$, and then isolate the two extreme values $\lambda_{V,\max}$ and $\lambda_{V,\min}$ to compute λ_{Vd_0} = Max ($\lambda_{V,\max}$, $\lambda_{V,\min}^{-1}$). As soon as λ_{Vd_0} is significant, a change in variance is detected and an adjustment has to be made. This second change d_0 is detected at the date $i = 1975$ which corresponds to July 27, 2001. This date historically coincides with the beginning of a highly volatile period for emerging markets following the crash of the Internet bubble and the financial crisis in Argentina.

The adjustment has to be made on 148 data points. The middle part of Figure 6 depicts the evolution of the time series residuals before the second standardization with the first rectangle indicating the previous adjustment and the second rectangle the period on which the second adjustment will be operated. The lower part of Figure 6 shows the new residual series Z_i^* after the second adjustment, while Table 8 summarizes the descriptive statistics of the two obtained series following the first and the second adjustments.

Table 8: Comparative descriptive statistics of the residual series after the first and second adjustments

	After the first adjustment	After the second adjustment
Mean (%)	0.0000	0.0000
Standard deviation (%)	1.1078	1.0521
Skewness	-0.1917 *	-0.0954*
Kurtosis	6.0972*	5.6083*
Jarque-Bera	1404.149 (p-value=0.0000)	985.5049 (p-value=0.0000)

* denotes significance at the 5% confidence level.

According to these descriptive statistics, the newly adjusted series Z_i^* exhibits less excess kurtosis than the one obtained after the first adjustment. This means that the moving variance ratio technique manages to treat the different extreme values.

Restarting the stepwise algorithm, an ARMA(1,1) seems now the most accurate model and is written as follows:

$$R_i = 0.00000694 + 0.64134 R_{i-1} - 0.63588 \xi_{t-1} \tag{9}$$

Table 9. ARMA(1,1) coefficients and statistics, third change detection

Variable	**Coefficient**	**p-value.**
Constant	6.942E-06	0.91502
AR(1)	0.64134*	0.00161
MA(1)	-0.63588*	0.00192

* indicates significance at the conventional 5% risk level.

Then, the moving variance ratio is computed on the residual series from the preceding ARMA(1,1) model. The two extreme values $\lambda_{V,\max}$ and $\lambda_{V,\min}$ are isolated so that λ_{Vd_0} = Max ($\lambda_{V,\max}$, $\lambda_{V,\min}^{-1}$) can be computed. Here λ_{Vd_0} is equal to 3.8954 which is below the critical value set equal to 4. Thus, the series does not contain any new variance change.

Finally, we compare our results with those obtained by applying the Bai and Perron (1998, 2003) test that allows detecting the number and location of the structural breaks in the time paths of a time series. We use the squared residuals of the ARMA estimation on the returns as a proxy for the unconditional variance. The breakpoint selection procedure is based on the Bayesian Information Criteria (BIC) and the maximum number of breaks is initially set to be 3. Table 10 summarizes the results of the BP test. They are consistent with what was previously reported with our variance shift detection method, i.e. two break points[7]. More precisely, the first variance change point is exactly the same one, i.e. observation number 945 which corresponds to August 15, 1997. The second break point is not exactly the same, i.e. 1675 versus 1975 for our method. However, the variance change point identified by our method is within the 95% confidence interval of the Bai-Perron test. The difference in the exact location of this second break might be explained by the difference in the length of the segments on which the computations are done in the BP test and the lengths on which we apply the adjustments. Consequently, the BP test provides rather similar results when compared to our method. However, our approach is easier to implement; moreover, it also provides an adjustment procedure after each variance shift detection[8].

Table 10. Results of the Bai-Perron test for multiple structural break points

Number of breakpoints	Estimated break date	95% confidence interval (observation number)
2	945	[488 - 954]
	1675	[1675 - 2047]

4. Conclusion

The moving variance ratio technique has proven its efficiency in detecting variance changes and removing them from the series. The detection technique developed in this paper is different with respect to the one proposed by Tsay (1988) in two ways. First, through the introduction of d which refers to the number of observations used to compute the two variances of the ratio. The first advantage of this innovation is the reduction in the number of computations needed. Moreover, having a numerator and denominator computed on the same amount of observations d gives to each ratio the same weight and significance. The second development that we propose concerns the area on which the adjustment is made. Tsay (1988) standardizes all the observations following the detected variance change. We actually chose to introduce and define k as the number of observations on which the adjustment is operated. A variance change never lasts indefinitely, which is why we decide to adjust only the period experiencing a high volatility. We determine this period trough the computation of a moving variance ratio with a fixed denominator. This ratio converges to 1 when the volatility comes back to the level of the period preceding the detected level shift. The value k refers to the number of consecutive ratios, computed after the detected variance change, that are above or below 1 in the case of an up or down-change, respectively.

Two significant variance changes have been identified within the series of the Hang Seng daily log-returns stretching from 1994 to 2007. The first one is detected on August 15, 1997, which coincides with the beginning of the period of high volatility in Hong Kong due to the financial turmoil experienced by Thailand after the Thai Baht devaluation. This period of high volatility lasts 980 trading days, i.e. almost four years. The second variance change is detected on July 27, 2001 which is the beginning of a period of high volatility on emerging markets following the crash of the dot.com bubble, along with the first signs of the financial crisis in Argentina and exacerbated by the September 11, 2001 attacks.

Potential developments could consist in using an adjustment factor in the standardization that overweights the residuals close to the detected change and optimizing the level of the critical value C.

There are multiple potential ways of using this moving variance ratio technique. Indeed, d can be set at different levels depending of the type of anomalies that one is willing to detect. If d is set at a very high level, large and permanent variance changes will be detected. On the contrary, when d is

[7] We also applied a CUSUM of squares test that points out the presence of two break points too.

[8] We also performed the moving variance procedure and the Bai and Perron test on 3 other Asian index returns, namely the Kospi Index, the Singapore Straits Times Index and the Nikkei 225 Index over the same time period. For space reasons, the detailed results are not provided here. Two conclusions emerge: 1/ the two approaches provide very similar results (the detected breaks are either exactly the same or within the confidence interval of the Bai and Perron test); 2/ the variance changes of all these Asian return series arise during the. Asian turmoil of 1997 and the dot.com bubble bursting coupled with the beginning the crisis in Argentina.

set at around 50, permanent and transient variance changes can be detected. Moreover, setting the value of d below 10 could be a way to capture outliers. Indeed, the presence of an abnormally high value will produce a huge increase in the variance ratio.

References

Abraham, B. and G.E.P. Box (1979) "Bayesian Analysis of Some Outlier Problems in Time Series" *Biometrika* **66**, 229-236.

Ané, T., Ureche-Rangau, L., Gambet, J.B. and J. Bouverot (2007) "Robust Outlier Detection for Asia-Pacific Stock Index Returns" *Journal of International Financial Markets, Institutions and Money* **18**, 326-343.

Bai, J. (2000) "Vector Autoregressive Models with Structural Changes in Regression Coefficients and in Variance-covariance Matrices" *Annals of Economics and Finance* **1**, 303-339.

Bai, J., Lumsdaine, R.L. and J.H. Stock (1998) "Testing for and Dating Breaks in Multivariate Time Series" *Review of Economic Studies* **65**, 395-432.

Bai, J. and P. Perron (1998) "Estimating and Testing Linear Models with Multiple Structural Changes" *Econometrica* **66**, 47-78.

Bai, J. and P. Perron (2003a) "Computation and Analysis of Multiple Structural Change Models" *Journal of Applied Econometrics* **18**, 1-22.

Bai, J. and P. Perron (2003b) "Critical Values for Multiple Structural Change Tests" *Econometrics Journal* **6**, 72-78.

Brown, R.L., Durbin, J. and J.M. Evans (1975) "Techniques for Testing the Constancy of Regression Relationships Over Time" *Journal of the Royal Statistical Society B* **37**, 149-163.

Bali, R. and H. Guirguis (2007) "Extreme Observations and Non-normality in ARCH and GARCH" *International Review of Economics and Finance* **16**, 332-346.

Black, F. and M. Scholes (1973) "The Pricing of Options and Corporate Liabilities" *Journal of Political Economy* **81**, 637–654.

Bollerslev, T. (1986) "Generalized Autoregressive Conditional Heteroskedasticity" *Journal of Econometrics* **31**, 307-327.

Box, G.E.P. and G.C. Tiao (1975) "Intervention Analysis with Applications to Economic and Environmental Problems" *Journal of the American Statistical Association* **70**, 70-79.

Chang, I. (1982) "*Outliers in Time Series*" Unpublished Ph.D. Thesis, University of Wisconsin-Madison, Dept. of Statistics.

Chang, I. and G.C. Tiao (1983) "*Estimation of Time Series Parameters in the Presence of Outliers*" Technical Report 8, University of Chicago, Statistics Research Center.

Davis, R.A., Huang, D. and Y.C. Yao (1995) "Testing for a Change in the Parameter Values and Order of an Autoregressive Model. *Annals of Statistics* **23**, 282-304.

Deng, A. and P. Perron (2008) "The Limit Distribution of the CUSUM of Squares Test under General Mixing Conditions" *Econometric Theory* **24**, 809-822.

Engle, R.F. (1982) "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation" *Econometrica* **50**, 987-1008.

Fox, A. J. (1972) "Outliers in Time Series" *Journal of the Royal Statistical Society, Series B* **34**, 350-363.

Franses, P.H. and H. Ghijsels (1999) "Additive Outliers, GARCH and Forecasting Volatility" *International Journal of Forecasting* **15**, 1-9.

Harvey, A.C. and J. Durbin (1986) "The Effects of Seat Belt Legislation on British Road Casualties: A Case Study in Structural Time Series Modelling" *Journal of the Royal Statistical Society Series A (General)* **149**, 187-227.

Hillmer, S.C., Bell, W. and G.C. Tiao (1983) "Modeling Considerations in the Seasonal Adjustment of Economic Time Series" in *Applied time series analysis of economic data* by A. Zellner, Ed., U.S. Department of Commerce, Bureau of the Census, Washington, 74-100.

Horváth, L. (1993) "The Maximum Likelihood Method for Testing Changes in the Parameters of Normal Observations" *Annals of Statistics* **21**, 671-680.

Hsu, D. A. (1977) "Tests for Variance Shift at an Unknown Time Point" *Applied Statistics* **26**, 279-284.

Inclán, C. and G.C. Tiao (1994) "Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance" *Journal of the American Statistical Association* **89**, 913-923.

Kejriwal, M. and P. Perron (2006a) "Testing for Multiple Structural Changes in Cointegrated Regression Models" Unpublished Manuscript, Department of Economics, Boston University.

Kejriwal, M. and P. Perron (2006b) "The Limit Distribution of the Estimates in Cointegrated Regression Models with Multiple Structural Changes" *Forthcoming in Journal of Econometrics*.

Liu, J., Wu, S. and J.V. Zidek (1997) "On Segmented Multivariate Regressions" *Statistica Sinica* **7**, 497-525.

Martin, D. and V.J. Yohai (1986) "Influence Functionals for Time Series" *Annals of Statistics* **14**, 781-818.

Perron, P. (2006) "Dealing with Structural Breaks" in *Palgrave Handbook of Econometrics* by K. Patterson and T.C. Mills, Eds., Vol. 1: Econometric Theory, Palgrave Macmillan, 278-352.

Perron, P. and Z. Qu (2006) "Estimating Restricted Structural Change Models" *Journal of Econometrics* **134**, 373-399.

Perron, P. and J. Zhou (2008) "*Testing Jointly for Structural Changes in the Error Variance and Coefficients of a Linear Regression Model"* Unpublished Manuscript, Department of Economics, Boston University.

Qu, Z. and P. Perron (2007a) "Estimating and Testing Multiple Structural Changes in Multivariate Regressions" *Econometrica* **75**, 459-502.

Tsay, R.S. (1986) "Time Series Model Specification in the Presence of Outliers" *Journal of the American Statistical Association* **81**, 132-141.

Tsay, R.S. (1988) "Outliers, Level Shifts, and Variance Change in Time Series" *International Journal of Forecasting* **7**, 1-20.

Tsay, R. S. and G.C. Tiao (1984) "Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models" *Journal of the American Statistical Association* **79**, 84-96.

Zhang, X. and M. King (2005) "Influence Diagnostics in Generalized Autoregressive Conditional Heteroscedasticity Processes" *Journal of Business and Economic Statistics* **23**, 118-129.

Zhou, J. and P. Perron (2008) "*Testing for Breaks in Coefficients and Error Variance: Simulations and Applications*" Unpublished Manuscript, Department of Economics, Boston University.