ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 216, 2008

*Wojciech Sarnowski**

ON DETECTION OF HOMOGENEOUS SEGMENTS OF OBSERVATIONS IN FINANCIAL TIME SERIES

ABSTRACT. The aim of this article is to present financial data modelling in presence of stochastic disorders. Change-point analysis is applied. We adapt universal method of change-point detection for disorder in parameters of GARCH processes. A comparison of the model fitted to whole sample with models built on homogenous data subset is made.

Key words: detection of change-points, minimum contrast estimator, GARCH models, stochastic volatility.

I. INTRODUCTION

The disorder of probabilistic mechanism driving the data is common in financial data analysis. It is known that markets generate clusters of different stochastic volatility violating data homogeneity. This phenomenon can be interpreted as variance disorder. Analysts, when modelling financial data set, are very often faced with volatility effect. Unfortunately, in some cases it is really hard to find model well-fitted to the whole sample. The aim of this paper is to present financial data modelling supported by change-point analysis which help us to solve the problem of disordered data. Taking into account the type of analyzed disorder we use a method assuming changes in parameters of marginal distributions of data, in particular changes in variances. Such a method was proposed in Lavielle (1999). We adapt it to GARCH process case. Finally, we fit a GARCH model to series of returns of dollar on DM exchange rate taking into account data disorders.

* MSc, Wroclaw University of Technology, Institute of Mathematics and Informatics; Wybrzeże Wyspiańskiego 27, Wrocław, Poland.

II. PROBLEM STATEMENT AND METHOD DESCRIPTION

Suppose that process $\{X_t\}_{t=1}^{n}$ is observed. $K-1$ changes occur at unknown times $t_1^*,\ldots,t_{K-1}^*$ in marginal distributions F_t of X_t's, $t=1,2,\ldots,n$. They affect parameter $\theta, \theta\in\Theta\subseteq R^d$. For the first t_1^* observations we have $\theta=\theta_1^*$. Next, between instants t_1^*+1 and t_2^*, $\theta=\theta_2^*$. Finally, marginal distributions of $X_{t_{K-1+1}^*},\ldots,X_n$ depend on $\theta=\theta_K^*$. Thus we can consider a vector $\underline{\theta}^*=\left(\theta_1^*,\ldots,\theta_K^*\right)$, $\theta_j^*\in\Theta\subseteq R^d$, $j=1,\ldots,K$ and the following model

$$P\left(X_t\in A\right)=\sum_{j=1}^{K}F_t(A;\theta_j^*)1_{\left(t_{j-1}^*,t_j^*\right]}(t)$$

for any Borel subset A of R^p (X_t's can be p-dimension vectors) and with $t_0^*=0$. Detection of the change-point set $\{t_1^*,\ldots,t_{K-1}^*\}$ bases on minimizing a contrast function. For any subsequence $X_t,\ldots,X_{t'}$ we assume that there exist functions $\phi:\Theta\to R^p, \psi:\Theta\to R^m$ (twice continuously derivable functions) and $\xi:R^p\to R^m$ for which the contrast function W_n satisfies:

$$\forall\theta\in\Theta\ W_n\left(X_t,\ldots,X_{t'};\theta\right)=\frac{1}{n}\sum_{i=t}^{t'}\left(\phi(\theta)+\left\langle\psi(\theta),\xi(X_i)\right\rangle\right), \tag{1}$$

where $1\le t\le t'\le n$ and $\langle .,.\rangle$ stands for the inner product. Define:

$$J_n\left(\underline{\tau},\underline{\theta}\right)=\sum_{k=1}^{K}W_n\left(\underline{X}_k,\theta_k\right). \tag{2}$$

Then, the minimum contrast estimator

$$\left(\hat{\underline{\tau}}_n,\hat{\underline{\theta}}_n\right)=\left(\hat{t}_1/n,\ldots,\hat{t}_{K-1}/n,\hat{\theta}_1^*,\ldots,\hat{\theta}_K^*\right) \text{ for } \left(t_1^*/n,\ldots,t_{K-1}^*/n,\hat{\theta}_1^*,\ldots,\hat{\theta}_K^*\right)$$

is obtained as a solution of the following minimization problem:

$$\forall\left(\underline{\tau},\underline{\theta}\right)\in T_K\times\Theta_K\quad J_n\left(\hat{\underline{\tau}}_n,\hat{\underline{\theta}}_n\right)\le J_n\left(\underline{\tau},\underline{\theta}\right), \tag{3}$$

where:
$T_K = \{ \underline{\tau} = (\tau_0, \tau_1, \dots, \tau_K), 0 = \tau_0 < \tau_1 < \dots < \tau_{K-1} < \tau_K = 1 \}$, with $\tau_j = t_j / n$;
$\Theta_K = \{ \underline{\theta} = (\theta_1, \dots, \theta_K), \theta_j \in \Theta \}$;
$W_n(\underline{X}_k, \theta_k)$ is the contrast function calculated over segment $\underline{X}_k = (X_{t_{k-1}+1}, \dots, X_{t_k})$. Proposed estimator works under two assumption. The first one is imposed on the contrast function:

Assumption 1 *There exist a function* $w: \Theta \times \Theta \to R$ *such that*

$$\forall\, 1 \le j \le K\ \forall \theta \in \Theta\ \ w(\theta_j^*, \theta) = \phi(\theta) + \langle \psi(\theta), E\xi(X_i) \rangle, \tag{4}$$

where $t_{j-1}^* + 1 \le i \le t_j^*$ *and such that, for any*

$(\theta, \theta') \in \Theta \times \Theta$, $w(\theta, \theta) \le w(\theta, \theta')$ *with* $w(\theta, \theta) = w(\theta, \theta')$ *if and only if* $\theta = \theta'$. *Furthermore, for any* $1 \le j \le K$, *there exists a neighborhood* $U(\theta_j^*) \subset \Theta$ *of* θ_j^* *and a constant* $B > 0$ *such that*

$$\forall \theta \in U(\theta_j^*)\ \ w(\theta_j^*, \theta) - w(\theta_j^*, \theta_j^*) \ge B \| \theta_j^* - \theta \|_2^2$$

The second assumption is expressed in the terms of process $\{ \eta_t(\theta) \}_{t=1}^n$ defined below:

$$\forall \theta \in \Theta\ \ \eta_t(\theta) = \langle \psi(\theta), \xi(X_t) - E\xi(X_t) \rangle, 1 \le t \le n. \tag{5}$$

Assumption 2 *There exists* $h \in [1, 2)$, *such that*

$$E\left(\sum_{i=t}^{t+s} \eta_i(\theta) \right)^2 \le C(\theta) s^h, 1 \le t \le t + s \le n \tag{6}$$

for some constant $C(\theta)$.

The considered method can be summarized by the following theorem proved in Lavielle (1999):

Theorem 1 *Let* $\hat{\underline{\tau}}_n$ *be the estimate of the normalized change-points sequence* $\underline{t}^* / n$ *and* $\hat{\underline{\theta}}_n$ *be the estimate of the parameters in different segments,* *obtained as a solution of the following minimization problem:* $\forall(\underline{\tau},\underline{\theta}) \in T_K \times \Theta_K \quad J_n(\underline{\hat{\tau}}_n,\underline{\hat{\theta}}_n) \le J_n(\underline{\tau},\underline{\theta})$.

Then, under assumptions 1 and 2, $(\underline{\hat{\tau}}_n,\underline{\hat{\theta}}_n)$ *converges in P-probability to* $(\underline{\tau}^*,\underline{\theta}^*)$.

Notice that process η and contrast function depend on the type of disorder. Changes in mean, variance or some other parameter determine different formulas (1)–(6). The point of our interest are disorders in GARCH parameters. In the next section we show that changes in such parameters can be interpreted as disorders in variances of X_t's.

III. ADAPTATION TO THE GARCH CASE

Let us consider a GARCH(p,q) process:

$$X_t = \sqrt{H_t} Z_t, t \in N, \tag{7}$$

where $\{Z_t\}$ is a sequence of i.i.d. random variables such that $EZ_t = 0$ and $EZ_t^2 = 1$. Moreover $\{H_t\}$ follows the equation:

$$H_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j X_{t-j}^2 + \sum_{j=1}^{p} \beta_j H_{t-j}. \tag{8}$$

We have:

$$\begin{aligned} Var(X_t) = E(X_t^2) = E\big(E(X_t^2 \mid F_{t-1})\big) = E(H_t) = \\ = \alpha_0 + \sum_{j=1}^{q} \alpha_j E(X_{t-j}^2) + \sum_{j=1}^{p} \beta_j E(H_{t-j}), \end{aligned} \tag{9}$$

where: $F_{t-1} = \sigma(X_1,\dots X_{t-1})$. We infer that $\sigma^2(t) := Var(X_t)$ depends on the vector of parameters $(\alpha_0,\alpha_1,\dots,\alpha_q,\beta_1,\dots,\beta_p)$. Thus, a change in θ can be considered as a change in variance $\sigma^2(t)$. This means the method used for the detection of variance disorders could be applied here. For that kind of disorders the following function $J_n(\underline{\tau},\underline{\theta})$ is proposed (see Lavielle (1999)):

$$J_n(\underline{\tau},\underline{\theta})=\frac{1}{n}\sum_{k=1}^{K}\left(\frac{\sum_{i=1}^{n_k}\left(X_{t_{k-1}+i}-\mu\right)^2}{\sigma_k^2}+n_k\log\sigma_k^2\right), \tag{10}$$

where:

$X_{t_{k-1}+i}$ – i-th observation from segment $\left(X_{t_{k-1}+1},\ldots,X_{t_k}\right)$;

n_k – length of k-th segment; $\sigma_k^2=\sigma^2(t_{k-1}+i)$ for $1\le i\le n_k$; $1\le k\le K$; $\mu=E(X_t)$; $1\le t\le n$. Moreover, in this case, for $\theta=\sigma^2$,

$$\eta_t(\theta)=\frac{1}{\sigma^2}\left[\left(X_t-E(X_t)\right)^2-Var(X_t)\right]. \tag{11}$$

Given (10) and (11) we are able to verify both assumptions for GARCH series. Assumption 1 for such function J_n is satisfied as J_n bases on the Gaussian likelihood function and $w(\theta',\theta)-w(\theta,\theta)$ is the Kullback-Liebler distance (see Lavielle (1999)). On the other hand there are several sufficient conditions stated in Lavielle (1999) under which assumption 2 holds. If we impose on $\{\eta_t\}_{t\in N}$ the following covariance structure

$$Cov(\eta_t,\eta_{t+s})=O(s^{-a}),\ t,s\in N$$

for some $a>0$, then (6) is satisfied with $h=\max\{2-a,1\}$.

We are going to use this fact to show that assumption 2 is fulfilled for GARCH processes in case of variance disorder. First, applying formula (11), let us compute covariance function for $\{\eta_t\}_{t\in N}$:

$$\begin{aligned}Cov(\eta_t,\eta_{t+s})&=\frac{1}{\sigma^4}Cov\left(X_t^2-E(X_t^2),X_{t+s}^2-E(X_{t+s}^2)\right)=\\&=\frac{1}{\sigma^4}Cov\left(X_t^2,X_{t+s}^2\right).\end{aligned}$$

Notice that $Cov(\eta_t,\eta_{t+s})$ requires knowledge about autocovariance function for $\{X_t^2\}$. To tackle this problem we will use well known property of squared

GARCH(p,q) process: if $\{X_t\}_{t\in N}$ is GARCH(p,q) satisfying (7) and (8) then $\{X_t^2\}_{t\in N}$ has ARMA representation:

$$X_t^2 = \alpha_0 + \sum_{j=1}^{r}(\alpha_j + \beta_j)X_{t-j}^2 - \sum_{j=1}^{p}\beta_j \nu_{t-j} + \nu_t,$$

where $r = \max\{p,q\}$ and $\nu_t = X_t^2 - H_t$. The innovation process $\{\nu_t\}$ is white noise with finite second moment. Thus, if $\{X_t^2\}$ can be rewritten as the ARMA process then its autocovariance function is geometrically bounded: $Cov(X_t^2, X_{t+s}^2) \le C r^s$, with $r \in (0,1)$ and C – some positive constant (see chapter 13 of the book Brockwell, Davies (1991)). Using this fact we obtain that $Cov(\eta_t, \eta_{t+s}) \le \tilde{C} r^s$ for some constant opositive $\tilde{C}$. Of course the condition: $r^s = O(s^{-a})$, $a \in [1,2)$ is met. Hence, usage of presented method for detection of changes in parameters of GARCH models has found a justification. Practical application of the method for $K = 2$ is presented in the next section.

IV. REAL DATA EXAMPLE – DETECTION OF CHANGE-POINT

In practical part of the paper we model financial time series using change-point analysis. We study series of returns of Dollar on DM exchange rate from 18 May 1971 to 18 April 1975 (961 observations). The series of daily returns is displayed in the left panel of figure 1. We decided to model our data set using family of GARCH(p,q) processes with $p = q = 1$. The empirical studies in the field of financial time series reveal that $p = q = 1$ is by far the most common model order for GARCH series.

We can observe two different regimes in the pattern of the variance. The first interval refers to low market volatility, the second interval (followed by outlier) corresponds to high volatility.

This reasonable preliminary analysis suggests two homogeneous segments of data. The estimated change-point confirms our observation - see right panel of figure 1. Applying the method described in sections 2 and 3, we obtained minimum of estimation procedure at point 430 what corresponds to 14 February 1973. Thus we register a disorder point at this day. The change-point analysis splits data set into two homogenous segments. The next step is to compare the quality of the model fitted to the whole sample with models fitted to each regime separately.

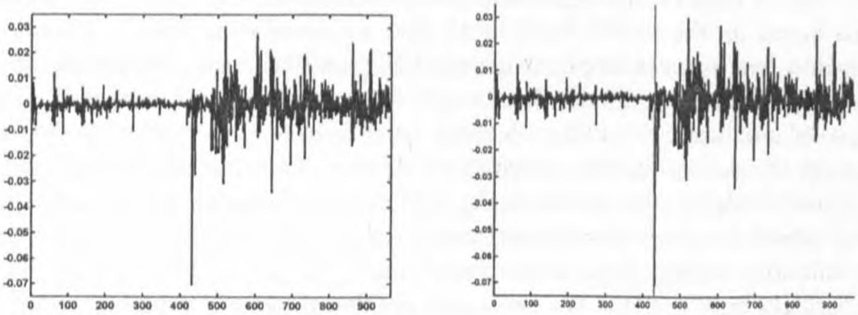

Figure 1. Data plot. Left panel: data without the line expressing change-point, right panel: data with the line of change-point.

V. REAL DATA EXAMPLE – COMPARISON OF THE MODELS

According to carried out change-point analysis we divided the series in two segment. The first contains observations from 1 to 430 (18/May/1971 – 15/Feb/1973). The second one in turn – from 431 to 961 (15/Feb/1973 – 18/Apr/1975). In table 1 we present results of parameters estimation. It is notable that parameter values differ significantly from subset to subset. The detected disorder stands behind this effect. As a consequence of nonhomogeneous regimes, estimation conducted on the whole sample results in compromise between stochastically different segments.

Table 1

Parameters of fitted models

Data set	α_0	α_1	β_1
whole sample	$0.378 \cdot 10^{-5}$	0.1592	0.7550
first segment	$0.764 \cdot 10^{-6}$	0.4398	0.5051
second segment	$0.545 \cdot 10^{-5}$	0.2290	0.6754

Source: own calculations.

However when we look at table 2 we realize that it could be a bad compromise. The table presents results of diagnostic checking of residuals. Assuming that the fitted model is correct we should get i.i.d. residuals. We collected results of two test of randomness applied to analyzed models. We can suspect that residuals based on the model fitted to all data are correlated, because p-value of Ljung-Box test is quite small. Other models pass both tests. Diagnostic of residuals reveals that the model fitted to all data could be poorly adjusted. Comparison of stochastic volatility obtained on disjoin subsets with volatility obtained on the whole sample provides us another argument against an idea of fitting one model to the stochastically different data segments. In figure 2 we display absolute values of analyzed series and estimated volatilities. Left panel plots volatility coming from single model and right panel – volatility generated by separately built models. We can easily see that stochastic volatility generated by one model is overestimated and too smooth on the first segment. The second segment volatility looks reasonably in both cases.

Table 2

Diagnostic checking – tests of residuals randomness

Data set	Test	Test statistics	p-value
whole sample	Ljung-Box	32.632	0.037
	Turning points	618	0.102
first segment	Ljung-Box	19.731	0.475
	Turning points	273	0.157
Second segment	Ljung-Box	26.579	0.148
	Turning points	349	0.705

Source: own calculations.

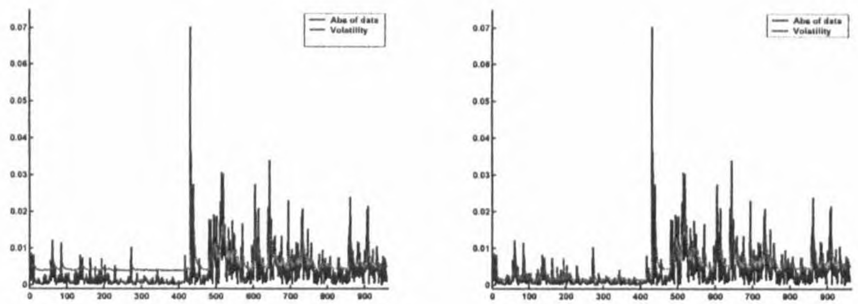

Figure 2. Absolute value of data versus estimated volatility. Left panel: volatility generated by the model fitted to whole sample, right panel: volatility combined from models built on separate subsets

VI. FINAL REMARKS

Statistical analysis of data set exhibiting strong non-homogeneity provides us conclusions that stochastic modelling should be followed by change-point analysis. Disorder detection and – after that – building separate models on homogeneous segments plays a crucial role. Single model is not enough to capture sometimes very different data segments and can result in strong overestimation (underestimation). Final inferring may be very uncertain then.

REFERENCES

Brockwell P.J, R. A. Davies, (1991), *Time series: theory and methods*, Springer-Verlag

Gouriéroux Ch, (1997), *ARCH models and financial applications*, Springer-Verlag, New York

Lavielle M, (1999), *Detection of multiple changes in a sequence of dependent variables*, Stochastic Processes and their Applications, 83. 79–102

Sarnowski W., (2003), *Wykrywanie skupień o jednorodnej wariancji w szeregach czasowych*, Master Thesis, Wroclaw University of Technology.

Wojciech Sarnowski

WYKRYWANIE JEDNORODNYCH SEGMENTÓW OBSERWACJI W FINANSOWYCH SZEREGACH CZASOWYCH

Praca podejmuje zagadnienie modelowania finansowych szeregów czasowych w obecności rozregulowań struktury probabilistycznej. Zmiany wykrywane są za pomocą uniwersalnej metody detekcji zaadaptowanej do wykrywania rozregulowań w parametrach procesów typu GARCH. Przeprowadzona została statystyczna analiza jakości modeli uwzględniających wykryte zaburzenia z modelami, które zakładają iż ciąg danych ma jednorodną strukturę probabilistyczną.