DYNAMIC ECONOMETRIC MODELS
Vol. 6 – Nicolaus Copernicus University – Toruń – 2004

Joanna Górka, Joanna Stempińska
Nicolaus Copernicus University in Toruń
University of Warmia and Mazury in Olsztyn

Heteroskedastic Cointegration

1. Introduction

Conditional heteroskedasticity has been commonly found in macroeconomic and financial time series (Bollerslev, Chou, Kroner (1992), as well as Bera, Higgins (1993)). It is known that many financial prices and economic indicators are volatile and the constant conditional variance assumption is impractical. Studies conducted in the eighties proved that many macromodels based on time series have less stable variance of the innovations than it was assumed (Welfe (1998)). It often appeared in the case of financial models, and it could be explained through the economic theory. The problem of volatile conditional variance can be solved by means of the ARCH model or its extensions.

The notion of cointegration introduced by Granger (1981) and developed by Engle and Granger (1987) became popular due to two seemingly contradictory facts: (1) economic time series typically appear to posses unit roots (2) the economic theory often suggests existence of equilibrium or a long run relationship between variables. In the classical model of cointegration (model E-G) it is assumed that the processes $y_t, x_t \sim I(1)$, while the errors from the long run relationship u_t are $I(0)$. Hansen (1992) noticed that the model of cointegration formulated by Engle and Granger is sufficiently general to cover all nonstationary economic models. The errors of cointegration regression differ stochastically from the regressors, which have a fixed mean and a bounded variance. One might expect that as the regressors increase, the residual variance would also increase. We might also expect that the variance of the error process changes over time due to changing factors. Summing up we can expect that the error variance will not be stationary over time.

The notion of the heteroskedastic cointegration model (HCI) was proposed by Hansen (1992). He reduced assumptions in the classical cointegration model E-G. Here we have only assumptions concerning the cointegration equation: (1) $x_t \sim I(1)$, (2) error (w_t) of the equation is a bi-integrated process (BI). Assumption (2) does not contradict the classical assumption, because a BI process is a $I(0)$ process.

Hansen in his papers has given two justifications for the HCI model:

1. Both x_t and w_t are characterized by variance which grows at the same rate.
2. In economic research the time-varying parameter (TVP) model are used more and more frequently. For instance, consider the linear regression model with a time-varying parameter

$$y_t = \beta_t x_t + u_t \tag{1}$$

$$\beta_t = \beta + v_t \tag{2}$$

The model described by equations (1)-(2) can be written in an equivalent form, as:

$$y_t = \beta x_t + v_t x_t + u_t \tag{3}$$

If time series $x_t \sim I(1)$, then model (3) becomes HCI model, where $v_t x_t$ is a BI process. Model (3) suggests that the residuals will be proportional to the regressors i.e. big disturbances will appear more often when the process x_t achieves big values.

The aim of this paper is to present the properties of the BI process and to show the heteroskedastic cointegration model.

2. The bi-integrated process

Hansen (1992) introduced the notion of the bi-integrated process.
The process w_t generated by $w_t = \sigma_t u_t$, where $\sigma_t \sim I(1)$ and $u_t \sim I(0)$ is called a bi-integrated (BI) process.

The process u_t is understood as a „stationary part" of process w_t, whereas the process σ_t (more precisely σ_t^2) – as a „variance part".

Let

$$w_t = \sigma_t u_t, \tag{4}$$

where $\sigma_t = \sigma_{t-1} + v_t$, and $u_t \sim NID(0, \delta_1)$, $v_t \sim NID(0, \delta_2)$. Note that

$$E(w_t) = 0 \tag{5}$$

$$\mathrm{var}(w_t) = E(\sigma_t^2 u_t^2) = \delta_1^2 \delta_2^2 t, \text{ if } \mathrm{cov}(u_t, v_t) = 0 \tag{6}$$

$$\mathrm{cov}(w_t w_{t-s}) = E(\sigma_t u_t \sigma_{t-s} u_{t-s}) = 0 \tag{7}$$

Thus the variance of the BI process, similarly to the variance process σ_t is a function of time and autocorrelation does not exist. Moreover, the process w_t frequently tends to cross its mean. Individual graphs of BI process realization resemble the GARCH process. Analysis of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the BI process as well as various GARCH processes led us to the following conclusions:

1. the values of the ACF and PACF coefficients for the BI process are statistically insignificant,
2. the values of the ACF coefficients for the squares are statistically significant. Moreover, they disappear very slowly (slower than for the type of GARCH process, for which the ACF dies out after several/a few dozen lags depending on the distance of the sum of the estimated parameters to one). This shows that the processes have a long memory variance process. Thus BI processes resemble IGARCH processes,
3. the values of the PACF coefficients for the squares of observations are statistically significant. Their disappearing occurs although we can not affirm that all the coefficients values after several/a few dozen lags are statistically insignificant,
4. in the boundary cases distinguishing BI and GARCH processes with the sum of the parameter close to one by means of ACF and PACF functions is not possible.

To illustrate the hypothetical realization of the BI, GARCH and ARCH processes are shown in the Fig. 1. The data generating process (DGP) for each case has the following form:

BI process

$$w_t = \sigma_t u_t , \tag{8}$$

$$\sigma_t = \sigma_{t-1} + v_t , \quad u_t , v_t \sim NID(0, 1), \tag{9}$$

GARCH process

$$\varepsilon_t \,|\, y_{t-1}, y_{t-2} \ldots \sim N(0, h_t), \tag{10}$$

$$h_t = 0.1 + 0.89 h_{t-1} + 0.1\varepsilon_{t-1}^2 , \tag{11}$$

ARCH process

$$\varepsilon_t \,|\, y_{t-1}, y_{t-2} \ldots \sim N(0, h_t), \tag{12}$$

$$h_t = 0.1 + 0.4\varepsilon_{t-1}^2 + 0.3\varepsilon_{t-2}^2 + 0.2\varepsilon_{t-3}^2 + 0.05\varepsilon_{t-4}^2 . \tag{13}$$

In Fig. 1 histograms and the values of the ACF and PACF for the squares, respectively are shown.

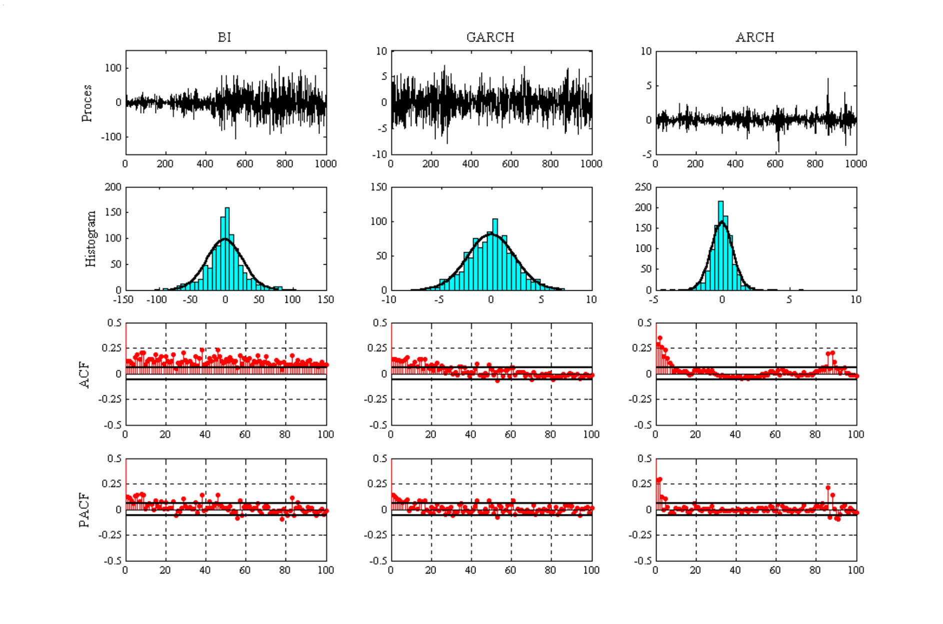

Fig. 1. Hypothetical realizations of BI, GARCH, ARCH processes and their profiles

The BI process is an explosive process. It is due to the variance of the process which is characterized by a stable partial autocorrelation function that does not die out. The distribution of this process is not close to the normal distribution.

In order to describe statistical properties of the BI process the simulation experiment has been made. It consisted in generating processes: BI, ARCH, GARCH described by the equations (8)–(13). All simulations were based on 100 and 300 replications, where the data generating process had 1000 and 1300 observations. Table 1. contains the results of simulations for the time series whose length are 1000 observations and 300 replications. The conclusions from the simulations based on the data generating series which have 1300 observations and 100 replications do not differ from the results shown in Table 1.

Table 1. Statistic properties of processes: GARCH, ARCH, BI (300 replications)

	Mean			Standard deviation		
	min	max	mean	min	max	mean
GARCH	-0.2840	0.4303	-0.0007	1.9665	6.3936	3.0704
ARCH	-0.2463	0.2772	-0.0015	0.6074	6.3549	1.1127
BI	-3.4503	2.4904	-0.0068	5.4636	72.7990	23.4574
	Skewness			Kurtosis		
	min	max	mean	min	max	mean
GARCH	-1.0083	0.7226	0.0059	2.8884	14.5002	4.8383
ARCH	-11.7519	12.7833	-0.0073	4.2094	359.1711	23.3872
BI	-0.7728	0.7752	0.0010	2.8709	14.6710	5.8697

The simulation results showed that the mean for all the processes tends to zero while the variance of the BI process has high values (keeps the higher value than variance for GARCH and ARCH processes). The BI process turned out to be more leptokurtic than the GARCH process, however less leptokurtic than the ARCH process (heightened kurtosis). One can observe the occurrence of havy tails (see: Fig. 1).

Beside the results shown above, we tested the presence of the autocorrelation by means of the Box-Ljung test, the ARCH effect by means of the Engle test, the unit roots by means of the DF test and linearity by means of the McLeod-Li test for examined processes. The results showed that it was not possible to distinguish these processes.

Analyzing the statistical properties, the question concerning the comparison of bi-integrated models with bilinear models (BL) arises. The latter are similar to GARCH models which are characterized by concentration of variance and in some cases (a subdiagonal process) have higher kurtosis with relation to GARCH models (compare Bruzda (2003)). Since the autocorrelation research for squares observations[1] for BL and GARCH models do not give us any possibility to distinguish between them (compare Bruzda (2003)), similarly to BI and GARCH models, we conclude that also for BL and BI models the research autocorrelation for squares observations we are not able to distinguish among these processes.

The basic characteristic distinguishing the BI process from the GARCH or BL process is the value of variance and, to a small extent, analysis of the ACF for squares. The BI model has a very big variance (bigger than GARCH or BL) and frequent changeability. It can be used to describe relatively calm processes, which at some moment „explode” (with their variance considerably rising; the variance depends on the process σ_t which is $I(1)$).

3. Heteroskedastic cointegration

Hansen (1992) gave the following definition of the heteroskedastic cointegration model:

Consider the linear regression model:

$$y_t = \beta_0 + \beta_1^{'} x_t + w_t \tag{12}$$

where $x_t \sim I(1)$ is a regressor vector $n \times 1$

$$x_t = x_{t-1} + u_{3t} \tag{13}$$

The error w_t is a term bi-integrated process defined earlier

$$w_t = \sigma_t u_{1t},$$

where $\sigma_t \sim I(1)$ is a scale process (14)

[1] For example Engle or McLeod-Li tests.

$$\sigma_t = \sigma_{t-1} + u_{2t} \tag{15}$$

The initial values x_0 and σ_0 are random variables with a finite absolute expectation.
The model given by the equation (12)–(15) is called a model of hetroskedastic cointegration – HCI

Note that

$$\mathrm{var}(x_t) \approx C_1 t, \qquad 0 < C_1 < \infty,$$
$$\mathrm{var}(w_t) \approx C_2 t, \qquad 0 < C_2 < \infty.$$

Thus x_t, w_t have the same stochastic order. This model differs from the standard E-G cointegration model, as the variance of the regression errors is time-varying. Although they have variances of the same stochastic order, the series behave entirely different. Process x_t, shows no tendency to return either to mean or to any particular value, while process w_t is running around its mean (crosses its mean) which is equal to zero (equation (5)). For visual distinction between CI and HCI processes, the individual realization of these processes has been shown in Fig. 2.

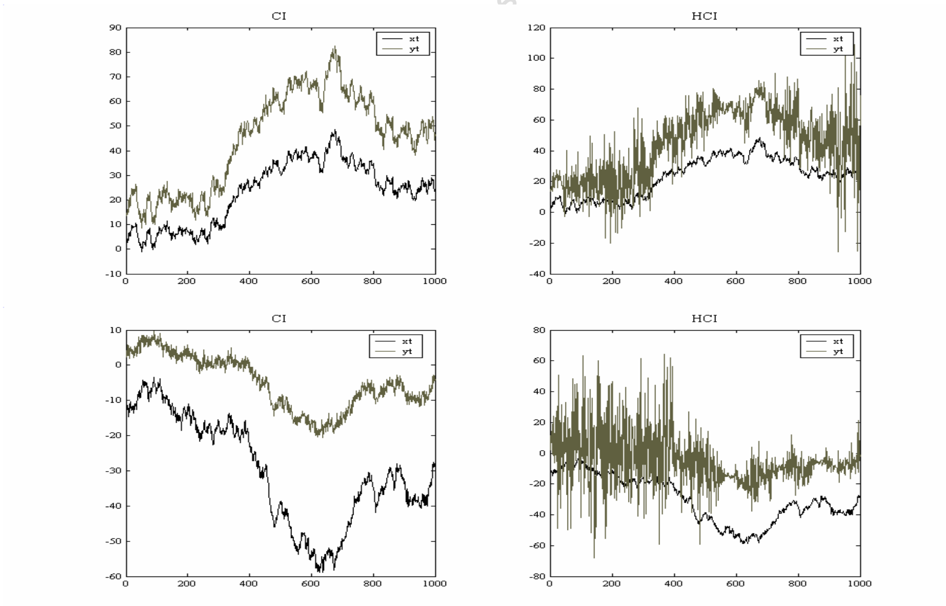

Fig.. 2. The individual realizations of CI and HCI processes

The CI processes have the same long development path, while HCI processes have the same feature, but with reference to the mean value. Weakening the assumption of the E-G model permits finding the long run relationship between two processes which have different size and frequency changeability.

In order to compare the CI and HCI models the Monte Carlo simulation was carried out. The generating models were the following:
HCI model

$$y_t = \beta_0 + \beta_1 x_t + w_t, \tag{16}$$

where

$$x_t = x_{t-1} + u_{3t},$$
$$w_t = \sigma_t u_{1t},$$
$$\sigma_t = \sigma_{t-1} + u_{2t},$$
$$u_{1t}, u_{2t}, u_{3t} \sim NID(0, 1),$$

CI model

$$y_t = \beta_0 + \beta_1 x_t + u_{1t}, \tag{17}$$

where

$$x_t = x_{t-1} + u_{2t},$$
$$u_{1t}, u_{2t} \sim NID(0, 1).$$

In each case 100 series of 1000 observations were generated. Then the parameters for the individual models by means of OLS[2] were estimated. In both cases estimated the values of all parameters are statistically significant.
The determination coefficient for CI model has high values while its size for the HCI model depends on estimated values of parameter β_1 (its value rises with a rise of β_1 – see: Table 2).

Table 2. The values of determination coefficient for HCI and CI models

		HCI			CI		
β_1	β_1	min	max	mean	min	max	mean
10	0.5	0.0003	0.6596	0.1253	0.8033	0.9944	0.9596
10	1.0	0.0092	0.9307	0.3167	0.9491	0.9999	0.9885
10	1.5	0.0205	0.9096	0.4281	0.9816	0.9993	0.9953

Residuals from a CI model have normal distribution, neither autocorrelation nor ARCH effect occur. Residuals from the HCI model, similarly to the CI model, are integrated of orders zero. However they are not normally distributed, the ARCH effect occurs and sometimes autocorrelation (of higher order).

[2] This estimation method for HCI model is used by Hansen (1992).

4. Conclusions

Modelling the economic processes using the cointegration is useful, because essential connections between nonstationary time series can be found. Hansen's work on the heteroskedastic cointegration showed that the statistical theory developed for the standard cointegration model can also be applied in the case of a wider class of models. In a standard cointegration model the regressor differs stochastically from residuals in two aspects: (1) the variance of regressors grows linearly in time, (2) the regressors have stochastic trend. In the hetroskedastic cointegration model, the regressors differ stochastically from residuals only with respect to the properties of the trend.

The HCI model allows to find the long run equilibrium between series with various changeability, with the variance growing in time and not necessarily, with the same properties of the trend function.

References

Bruzda, J. (2003), Procesy bi-linearne i GARCH w modelowaniu finansowych szeregów czasowych (Bilinear and GARCH Processes in Modeling of Financial Time Series), *Przegląd Statystyczny (Statistical Survey)*, 2.

Bera, A. K., Higgins, M. L. (1993), ARCH Models: Properties, Estimation and Testing, *Journal of Economic Surveys*, 7, 305-366.

Charemza, W. W., Deadman, D. F. (1997), *Nowa ekonometria (New Directions in Econometric Practice)*, PWE Warszawa.

Engle, R. F., Granger, C. W. J. (1987), Cointegration and Error Correction: Representation, Estimation and Testing, *Econometrica*, vol. 55, 251–276.

Hansen, B. E. (1992), Heteroskedastic cointegration, *Journal of Econometrics*, 54, 139–158.

Hyżyk, S., Osińska, M. (2000), Zmienność wariancji warunkowej w modelowaniu i prognozowaniu kursów walutowych (Variability of conditional variance in the modeling and forecasting of exchange rates), AUNC, XXX, UMK, Toruń.

Maddala, G. S., Kim, In-Moo. (2002), *Unit Roots, Cointegration, and Structural Change*, Cambridge.

Syczewska, E. M. (1999), *Analiza relacji dlugookresowych: estymacja i weryfikacja (Analysis of the long run relations: estimation and verification)*, SGH, Warszawa.

Welfe, A. (1998), *Ekonometria (Econometrics)*, PWE, Warszawa.