STATISTICA, anno LXIII, n. 3, 2003

THE COINTEGRATION ANALYSIS IN HYPOTHESIS OF HETEROSCHEDASTICITY: THE WILD BOOTSTRAP TEST

M. Gerolimetto, I. Procidano

1. INTRODUCTION

The importance of the concept of cointegration (Engle, Granger, 1987) has increased in the recent years and it received much attention in the literature. Some of the seminal works are those of Johansen (1988, 1996), Phillips and Durlauf (1986), Phillips and Hansen (1990), Xiao and Phillips (1999).

The cointegration analysis is usually carried out for macroeconomic and financial time series that are often affected by heteroschedasticity. This aspect is extremely relevant considering that homoschedasticity is one of the conditions underlying the applicability of the cointegration analysis procedures frequently used. So it is important to evaluate the robustness as regards heteroschedasticity of the procedures for the cointegration analysis and this is also the aim of this work. We consider and compare some traditional procedures and also an interesting proposal of a bootstrap test for cointegration. This test is originally a test for unit root (Procidano, Rigatti Luchini, 1999, 2000) which employs the bootstrap method called Wild Bootstrap.

Most of the results about the estimators for non-stationary series are asymptotic and their goodness has been evaluated by Montecarlo experiments, for example Schwert (1989), Diebold and Rudenbush (1991), Dejong *et al.* (1992). If the samples are finite, it seemed that the most frequently used tests exhibit serious level of size distortion and different powers.

These results oriented the most recent researches about this theme to the development of new procedures of estimation and testing by using the bootstrap method. This is the context, where it is inserted the proposal of the Wild Bootstrap test for unit root, that we want here to extend to the cointegration analysis.

In this work we want to compare, working by simulations, the robustness as regards heteroschedasticity of different procedures for cointegration analysis and, in particular, to verify if the Wild Bootstrap test is robust also in the context of the cointegration analysis. We consider the Dickey-Fuller test (DF: Dickey and Fuller, 1979), the Sargan-Bhargava test (DW: Sargan, Bhargava, 1983; Bhargava, 1986), the Johansen tests (λ_{trace} and λ_{max}: Johansen, 1988) and the Wild Bootstrap test (B) in hypothesis of GARCH errors. This typology of heteroschedasticity has been chosen because it is considered general enough to represent the major part of situations of non-stationarity in variance, typical of several economical and financial time series. In this context, like in the unit root analysis (Procidano and Rigatti Luchini, 1999, 2000, 2001), the simulations show a robustness of the Wild Bootstrap test as regards heteroschedasticity, particularly for small samples.

The article is organized as follows, section 2 introduces the notions of cointegration, section 3 explains the algorithm for the implementation of the wild Bootstrap Test, section 4 describes the simulation experiment, finally, section 5 concludes.

2. THE CONCEPT OF COINTEGRATION

The idea of a cointegration relationship between two or, more in general $n>2$, time series was intuitively proposed by Granger (1983) and then successively formalized in the following definition: given two time series $x_{1,t}$ and $x_{2,t}$, both I(1), without drift and trend, if it exists a linear combination

$$z_t = \beta_1 x_{1,t} - \beta_2 x_{2t} \tag{1}$$

that is I(0), then $x_{1,t}$ and $x_{2,t}$ are cointegrated and $\boldsymbol{\beta}=(\beta_1,\beta_2)$ is called vector the cointegration.

Moreover, if $x_{1,t}$ and $x_{2,t}$ are both I(1) and cointegrated, they are always generated by the model ECM (Error Correction Mechanism) whose matricial formulation is the following:

$$\nabla \mathbf{X}_t = \mathbf{\Gamma}_1 \nabla \mathbf{X}_{t-1} + \boldsymbol{\alpha}(\boldsymbol{\beta}' \mathbf{X}_{t-1}) + \boldsymbol{\varepsilon}_t \tag{2}$$

where $\mathbf{X}_t=(x_{1,t}, x_{2,t})$, $\boldsymbol{\varepsilon}_t=(\varepsilon_{1,t},\varepsilon_{2,t})$ is a white noise vector, $\mathbf{\Gamma_1}$ is a (2x2) matrix, $\mathbf{\Pi}=\boldsymbol{\alpha\beta}'$ is a (2x2) matrix of rank 1, $\boldsymbol{\alpha}=(\alpha_1,\alpha_2)$ is called speed of adjustment and $\boldsymbol{\beta}=(\beta_1,\beta_2)$ is the cointegration vector.

Different procedures can be used to test for the existence of a cointegration relationship and they can be classified in two groups. The first group includes the "two steps procedures": in the first step the cointegration relationship is estimated by a regression between the variables, in the second step the null hypothesis of unit root in the residuals is verified by the usual tests for unit root. Among the "two steps procedures" there are the Dickey-Fuller Test (DF), the Sargan-Bhargava Test (DW) and the Wild Bootstrap Test (B).

More precisely, given the following model:

$$x_t = \alpha x_{t-1} + \varepsilon_t \quad t = 1,2,\ldots \mathrm{T} \tag{3}$$

where ε_t is a white noise with σ^2 variance, the statistic of the Dicey-Fuller test is:

$$\tau = (\hat{\alpha} - 1)S^{-1}\left(\sum_{t=1}^{T} x_{t-1}^2\right)^{1/2} \tag{4}$$

where $S^2 = (T-2)^{-1}\sum_{t=1}^{T}(x_t - \hat{\alpha}x_{t-1})^2$ and $\hat{\alpha} = \left(\sum_{t=1}^{T} x_t x_{t-1}\right)\left(\sum_{t=1}^{T} x_{t-1}^2\right)^{-1}$

The statistic of the Sargan-Bhargava test is:

$$R_1 = \sum_{t=1}^{T}(x_t - x_{t-1})^2 \Big/ \sum_{t=1}^{T}(x_t - \overline{x})^2 \tag{5}$$

The second group includes the formulations that work directly on the time series, for example the Johansen test proposed in 1988. In this procedure the hypothesis of no cointegration is tested analysing the rank of the matrix $\mathbf{\Pi}=\boldsymbol{\alpha\beta}'$ by maximum likelihood techniques.

Johansen proposed two statistics:

$$\lambda_{\text{trace}} = -T\sum_{i=r+1}^{n}\log(1-\hat{\lambda}_i) \tag{6}$$

and

$$\lambda_{\max} = -T\log(1-\hat{\lambda}_{r+1}) \tag{7}$$

where $\hat{\lambda}_i$ are the eigenvalues of the matrix $\mathbf{\Pi}=\boldsymbol{\alpha\beta}'$ of the ECM model, T is the sample size and r, $0 \leq r < n$ is the rank of the $\mathbf{\Pi}$ matrix.

From the several works with comparisons among different cointegration tests (for example Podivinsky, 1998; Haug, 1996) it results that the performances are quite different. These results imply for the applied researchers that more than one cointegration test should be applied. In general, most of the Monte Carlo studies reveal a trade off between power and size distortions of the cointegration tests, even though the least size distortion is exhibited (but only for high sample size) by the Johansen tests.

3. THE WILD BOOTSTRAP TEST

The B test proposed by Procidano and Rigatti Luchini (1999, 2000) to test for a unit root refers to the resample method proposed by Wu in 1986. This method is robust as regards heteroschedasticity and it consists of carrying out extractions from an external distribution with zero mean and unit variance whose choice represents a subjectivity element (Davidson and Flaichaire, 2000) that does not seem to significantly modify the results.

The following is the general explanation of algorithm for the implementation of the B test:

1 - We generate a time series x_t from the model (3) in the null hypothesis $\alpha=1$;
2 - We calculate the value of the τ statistics;
3 - We generate a replication of x_t with the wild bootstrap and then we calculate the value of the τ statistics (indicated by τ^*) on this replication;
4 - We repeat the step 3 J times: so we obtain the empirical distribution of τ^*;
5 - We calculate the percentile corresponding the level of the test (1-δ): we obtain so the extreme limit of the acceptance region with the wild bootstrap;
6 - We register if the statistics τ does not belong to the MacKinnon acceptance region or to the acceptance region calculated with the wild bootstrap following step 5;
7 - we repeat steps 1-2-3-4-5-6 N times: we obtain so the rejects proportion as regards the acceptance region calculated with the wild bootstrap.

When the B test is inserted in a context of cointegration analysis, the algorithm above described is applied to the residuals of the regression among the variables with the aim of establishing if the same residuals contain a unit root i.e. if the variables are cointegrated.

4. THE SIMULATION EXPERIMENT

In this section the performance of the B test is compared to that of the other cointegration tests. The comparison has been carried out by the simulation of 5000 time series of length T = 25, 50, 100, 200 for each of the two following schemes, partially referring to what Lee and Tse did in their work in 1996:

1) In the experiment for examining the size of the tests we generate non cointegrated systems with GARCH (1,1) errors. Let $\mathbf{X_t}=(x_{1t},x_{2t}, \ldots,x_{Nt})'$ be an Nx1 vector of integrated series with

$$\nabla \mathbf{X_t} = \boldsymbol{\varepsilon}_t \tag{8}$$

The error vector $\varepsilon_t=(e_{1t},e_{2t},\ldots,e_{Nt})'$ is assumed to follow an N-variate conditional normal distribution with $E(e_{it} \mid F_{t-1})=0$ and $E(e_{it}^2 \mid F_{t-1})=\sigma_{it}^2$, where $i = 1,2,\ldots N$, F_{t-1} is the σ-field generated by all the information available at time t-1 and

$$\sigma_{it} = \phi_{i0} + \phi_{i1} e_{i,t-1}^2 + \phi_{i,2}\sigma_{i,t-1} \tag{9}$$

2) In the experiment for examinig the power of the tests we generate bivariate cointegrated systems with GARCH (1,1) errors. The system generated is

$$\begin{aligned} \nabla x_{1,t} &= -0.2(x_{1,t-1} - x_{2,t-1}) + \varepsilon_{1,t} \\ \nabla x_{2,t} &= \varepsilon_{2,t} \end{aligned} \tag{10}$$

In the first case, the simulations permit to evaluate the size distortion of the different tests of cointegration expressed as proportion of rejects of the null hypothesis, in the second case the power of the different tests is estimated.

The parameters of the GARCH model have been chosen like in Procidano and Rigatti Luchini (1999, 2000, 2001) *i.e.* considering quasi-integrated ($\varphi_1+\varphi_2 \cong 1$) and quasi-degenerete ($\varphi_0 \cong 0$) structures.

5. RESULTS

In the following tables, for each choice of the GARCH parameters the size distortions of the tests are reported.

TABLE 5.1

Proportion of rejects (φ_0=0.1 φ_1=0.3 φ_2=0.6)
Proportions of rejects out of 5000 tests, nominal level 0.95
500 replications, normal error distribution

Test	T=25	T=50	T=100	T=200
		φ_0=0.1 φ_1=0.3 φ_2=0.6		
DF	0.0668	0.0628	0.0621	0.0593
DW	0.0715	0.0678	0.0631	0.0618
λ_{Trace}	0.0701	0.0691	0.0591	0.0572
B	*0.0598*	*0.0583*	*0.0569*	*0.0521*
λ_{max}	0.0737	0.0652	0.0636	0.0609

TABLE 5.2

Proportion of rejects (φ_0=0.05 φ_1=0.3 φ_2=0.65)
Proportions of rejects out of 5000 tests, nominal level 0.95
500 replications, normal error distribution

Test	T=25	T=50	T=100	T=200
		φ_0=0.05 φ_1=0.3 φ_2=0.65		
DF	0.0684	0.0678	0.0659	0.0612
DW	0.0721	0.0682	0.0628	0.0624
λ_{Trace}	0.0798	0.0672	0.0622	0.0604
B	*0.0619*	*0.0614*	*0.0593*	*0.0554*
λ_{max}	0.0816	0.0705	0.0624	0.0607

TABLE 5.3

Proportion of rejects (φ_0=0.001 φ_1=0.3 φ_2=0.699)
Proportions of rejects out of 5000 tests, nominal level 0.95
500 replications, normal error distribution

Test	T=25	T=50	T=100	T=200
		φ_0=0.001 φ_1=0.3 φ_2=0.699		
DF	0.0692	0.0689	0.0678	0.0627
DW	0.0734	0.0712	0.0695	0.0653
λ_{Trace}	0.0786	0.0688	0.0646	0.0617
B	*0.0563*	*0.0638*	*0.0621*	*0.0612*
λ_{max}	0.0833	0.0677	0.0636	0.0621

From the analysis of tables 5.1, 5.2, 5.3, it is possible to see how the size distortion of all the tests tends to approach the nominal level as the sample size increases. Moreover, the size distortion exhibited by the B test is always the lowest for every sample size and for every parametric choice. This result is particularly important for small sample sizes (*e.g.* $T = 25$) where the test B permits to get proportions of rejects closer to the nominal level than those got by the other tests. As the sample size increases the difference becomes less evident because the size distortion tends to approach the nominal level.

As far as the power is concerned, in the following tables the powers of the test are reported:

TABLE 5.4

Estimation of the power (φ_0=0.1 φ=0.3 φ_2=0.6)
Estimation of the power of the tests, nominal level 0.95
500 replications, normal error distribution

Test	T=25	T=50	T=100	T=200
		φ_0=0.01 φ=0.3 φ_2=0.6		
DF	0.175	0.492	0.628	0.915
DW	0.194	0.538	0.702	0.965
λ_{Trace}	0.190	0.412	0.850	0.926
B	*0.201*	*0.498*	*0.711*	*0.971*
λ_{max}	0.219	0.505	0.916	0.992

TABLE 5.5

Estimation of the power (φ_0=0.05 φ=0.3 φ_2=0.65)
Estimation of the power of the tests, nominal level 0.95
500 replications, normal error distribution

Test	T=25	T=50	T=100	T=200
		φ_0=0.05 φ_1=0.3 φ_2=0.65		
DF	0.171	0.499	0.622	0.907
DW	0.168	0.532	0.697	0.968
λ_{Trace}	0.193	0.390	0.849	0.918
B	*0.204*	*0.513*	*0.709*	*0.978*
λ_{max}	0.200	0.466	0.936	0.984

TABLE 5.6

Estimation of the power (φ_0=0.001 φ_1=0.3 φ_2=0.699)
Estimation of the power of the tests, nominal level 0.95
500 replications, normal error distribution

Test	T=25	T=50	T=100	T=200
		φ_0=0.001 φ_1=0.3 φ_2=0.699		
DF	0.179	0.502	0.620	0.909
DW	0.170	0.490	0.692	0.965
λ_{Trace}	0.184	0.489	0.859	0.952
B	*0.191*	*0.507*	*0.698*	*0.970*
λ_{max}	0.190	0.517	0.927	0.989

Even though the power of the B test is not always the best, anyway it turns out to be high and compared to the DF test both compared to the other test in general. For every test the power increases as the sample size increases.

These results seem to confirm the robustness of the Wild Bootstrap test with respect to heteroschedasticity also in the interesting context of the cointegration analysis. In our opinion, these are very important results, particularly because they could be profitably employed considering other heteroschedastic structures. Moreover, these results can be employed also in the recent and challenging context of fractional cointegration.

Dipartimento di Scienze Satistiche MARGHERITA GEROLIMETTO
Università di Padova

Dipartimento di Statistica ISABELLA PROCIDANO
Università Ca' Foscari, Venezia

REFERENCES

A. BHARGAVA, (1986), *On the theory of testing for unit roots in observed time series*, "Review of Economic Studies", 53, pp. 369-384.

R. DAVIDSON, E. FLANCHAIRE (2000), *The wild bootstrap tamed at last*, Manoscritto.

D.N. DEJONG. et al. (1992), *The power problems of unit root tests in time series with autoregressive errors*, "Journal of Econometrics", 53, pp. 324-343.

D.A. DICKEY, W.A. FULLER (1979), *Distribution of the estimators for autoregressive time series with unit root*, "Journal of the American Statistical Association", 74, pp. 427-431.

F. DIEBOLD, G RUDEBUSCH (1991), *On the power of dickey fuller tests against fractional alternatives*, "Economic Letters", 35, pp. 155-160.

R.F. ENGLE, C.W.J. GRANGER (1987), *Co-integration and error correction representation, estimation and testing*, "Econometrica", 55, pp. 251-276.

C.W.J. GRANGER (1983), *Cointegrated variables and error correcting modelling*, Working paper n. 83-13, University of California.

C.W.J. GRANGER, R.F. ENGLE (1985), *Dynamic model specification with equilibrium constraints*, Mimeo, (University of California, San Diego, CA).

A.A. HAUG (1996), *Tests for cointegration a Monte Carlo comparison*, "Journal of Econometrics", 71, pp. 89-115.

S. JOHANSEN (1988), *Statistical analysis of cointegrating vectors*, "Journal of Economic Dynamics and Control", 12, pp. 231-254.

S. JOHANSEN (1992), *Determination of the cointegration rank in the presence of a linear trend*, "Oxford Bulletin of Economics and Statistics", 54, pp. 383-397.

S. JOHANSEN (1996), *Likelihood-based inference in cointegrated autoregressive models*, 2nd edn. Oxford University Press.

K. KIM, P. SCHMIDT (1993), *Unit root tests with conditional heteroscedasticity*, "Journal of Econometrics", 59, pp. 287-300.

T.H. LEE, Y. TSE (1996), *Cointegration tests with conditional heteroscedasticity*, "Journal of Econometrics", 73, pp. 401-410.

R.Y. LIU (1988), *Bootstrap Procedures under some non-i.i.d. models*, "Annals of Statistics", 16, pp. 1697-1708.

E. MAMMEM (1993), *Bootstrap and wild bootstrap for high dimensional linear models*, "Annals of Statistics", 21, pp. 255-285.

P.C.B. PHILLIPS, S. DURLAUF (1986), *Multiple time series regressions with integrated processes*, "Review of Economic Studies", 53, pp. 473-496.

P.C.B. PHILLIPS, B. HANSEN (1990), *Statistical inference in instrumental variables regression with I(1) processes*, "Review of Economic Studies", 57, pp. 99-125.

J.M. PODIVINSKY (1998), *Testing misspecified cointegrating relationships*, "Economic Letters", 60, pp. 1-9.

I. PROCIDANO, S. RIGATTI LUCHINI (1999), *Verifica di radici unitarie mediante bootstrap*, Convegno S.CO. 99, Venezia 27-29 settembre 1999, Atti, pp. 242-247.

I. PROCIDANO, S. RIGATTI LUCHINI (2000), *Il bootstrap esterno nella verifica di radici unitarie in condizioni di eteroschedasticità*, XL Riunione scientifica della Società italiana di Statistica, Firenze, 26-29 aprile 2000, Atti, pp. 303-306.

I. PROCIDANO, S. RIGATTI LUCHINI (2001), *Testing unit roots by bootstrap*, Metron, in corso di pubblicazione.

J.D. SARGAN, A. BHARGAVA (1983), *Testing residuals from least squares regression for being generated by the gaussian random walk*, "Econometrica", 51, pp. 153-174.

G.W. SCHWERT (1989), *Tests for unit roots: a Monte Carlo investigation*, "Journal of Business and Economic Statistics", 7, pp. 147-159.

C.F.J. WU (1986), *Jacknife, bootstrap and other resampling methods in regression analysis (with discussion)*, "Annals of Statistics", 14, pp. 1261-1295.

Z. XIAO, P.C.B. PHILLIPS (1999), *Efficient detrending in cointegrating regression*, "Econometric Theory", 15, pp. 519-548.

RIASSUNTO

La verifica di cointegrazione in ipotesi di eteroschedasticità: il test bootstrap esterno

In questo lavoro si vuole confrontare, tramite simulazioni, la robustezza rispetto all'eteroschedasticità di tipo GARCH (1,1) di alcune delle principali procedure proposte in letteratura per l'analisi di cointegrazione. In particolare, si considerano il test di Johansen e alcune procedure a due stadi, cioè il test Dickey-Fuller, il test Sargan-Bhargava ed un test bootstrap esterno. Quest'ultimo, confrontato con gli altri, ha fornito una buona performance specialmente alla bassa numerosità campionaria.

SUMMARY

The cointegration analysis in hypothesis of heteroschedasticity: the wild bootstrap test

We consider the problem of comparing, by simulations, the robustness as regards heteroschedasticity GARCH (1,1) of some of the most important procedures proposed in literature for the cointegration analysis. In particular, we consider the Johansen test and some "two steps" procedures", i.e. Dickey-Fuller test, Sargan-Bhargava test and an External Bootstrap test. The Bootstrap test performs very well, particularly for the lowest sample size.