ECM-Cointegration test with GARCH(1,1) Errors.

Panagiotis Mantalos
Department of Science and Health
Blekinge Institute of Technology
Sweden

ABSTRACT

By using Monte Carlo experiment, we study the robustness of the ECM-Cointegration tests when the first difference of the series follow a GARCH(1,1) process and the behaviour of the ARCH-tests of ECM residuals for variation of GARCH parameters. We found that if two series are individually I(1) and follow a GARCH (1,1), then ECM-Cointegration test tend to overreject for small samples but the problem is not very serious for large sample except when the errors’s GARCH process is nearly integrated and the volatility parameter is not small. Finally the Wild Bootstrap is robust on GARCH(1,1) without lost of power.

Keywords: ARCH, Bootstrap, Cointegration, GARCH, Wild Bootstrap.

Communications should be sent to:
Panagiotis Mantalos, Department of Science and Health, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden.
E-mail: panagiotis.mantalos@bth.se
1. Introduction


At the same time it is also known that there is some ARCH/GARCH process that the different stocks follow, and it is important to investigate what effects this has on the ECM-Cointegration tests.

Pantula (1989) first investigated nonstationary univariate AR model with a regular unit root, and a first-order ARCH error, and derived the asymptotic distribution of the Least Squares Estimator, LSE of the unit root, which is the same as that given by Dickey and Fuller, (1979) However Kim and Schmidt,(1993) by using Monte Carlo simulation show that the Dickey-Fuller tests tend to over-reject in the presence of GARCH errors but the problem is not very serious except when the error’s GARCH process is nearly integrated and the volatility parameter is not small.

Ling and Li, 1998a investigated an ARIMA model with GARCH (p, q) errors, and derived the asymptotic distribution of the Maximum Likelihood Estimation, MLE. The asymptotic distributions of the MLE of various unit roots involve a series of bivariate Brownian motions, and the MLE of unit roots is more efficient than the corresponding LSE when GARCH innovations are present. Using these asymptotic distributions, Ling and Li, 1998b constructed
some new unit root tests, and show by using Monte Carlo simulation, that the new tests can have a better performance than the Dickey-Fuller tests.

In the cointegration case Hansen and Rahbek (1999) based on an operational drift criterion from Markov chain theory, showed that Cointegration analysis of vector autoregressive models is robust with respect to ARCH innovations.

Li, Ling and Wong (1998) by studying an m-dimensional autoregressive process with GARCH(p,q) innovations, showed that the full rank and the reduced rank MLE are more efficient than the LSE.

However, as the ECM procedure is a special case of Johansen’s procedure for a system in which the Cointegrating vectors appear only in the equation of interest, KED suggested analysing only the equation of interest as a conditional equation. And Mantalos (1999) in an empirical investigation of the Cointegration relationship between French and German stock markets, under period 1.11.1997- 31.3.1999 found that even when the series individually follow a GARCH(1,1) (nearly integrated GARCH process with small volatility parameter ), all the ARCH test showed that there is no ARCH in variance of the ”ECM-Cointegration tests regressions” residuals.

That is, there is a need for study of the ECM-Cointegration tests in the presence of conditional heteroskedasticity of the GARCH(1,1) form.

The purpose of this paper is to study the robustness of the ECM-Cointegration tests when the first difference of the series follow a GARCH(1,1) process and the behaviour of the ARCH-tests of ECM residuals nearly integrated GARCH process with small volatility parameter.
The paper is arranged as follows. In the next section we present the design of our Model and the Monte Carlo experiment. In Section 3 we describe the results concerning the size of the various tests while power is analysed in Section 4. Finally, a brief summary and conclusions are presented in Section 5.

2. The Model and the Monte Carlo Experiment.

Consider the data-generation process (DGP) consists of a linear first-order vector autoregression equation with at least one unit root and Granger causality in only one direction. The DGP is given by a conditional error-correction-based model, (ECM) as in equation (1) below, and a marginal unit-root process (2),

\[
\begin{align*}
\Delta y_t &= \alpha \Delta x_t + \beta (y - x)_{t-1} + \varepsilon_{1t}, \quad t = 1, \ldots, T \\
\Delta x_t &= \varepsilon_{2t}, \quad t = 1, \ldots, T
\end{align*}
\]

where \(\Delta\) denotes the first-difference and \(T\) is the number of observations in the sample.

The error components \((\varepsilon_{1t}, \varepsilon_{2t})\) in (1) and (2) are generated by GARCH(1,1) models, i.e.,

\[
\begin{align*}
\varepsilon_{it} &= h_{it} \upsilon_{it}, \quad i = 1, 2 \\
\upsilon_{it} &\text{i.i.d., } E(\upsilon_{it}) = 0, E(\upsilon_{it}^2) = 1 \\
h_{it}^2 &= \gamma_i + \phi_i h_{i,t-1}^2 + \phi_i \varepsilon_{i,t-1}^2 \\
\text{and } \text{Cov}(\varepsilon_{1t}, \varepsilon_{2t}) &= 0.
\end{align*}
\]

The variables \(y_t\) and \(x_t\) are integrated of order one, I(1), and are possibly cointegrated. Without loss of generality, the cointegrating vector for \((y_t, x_t)\) is \((1, -1)\) if \(y_t\) and \(x_t\) are cointegrated and, for simplicity, the "hypothesised" cointegrating vector is assumed known. The parameter \(\beta\) is the error-correction coefficient in the conditional model. The parameter space is restricted to \(\{0 \leq \alpha \leq 1, \ -1 < \beta \leq 0\}\). The variables \(y_t\) and \(x_t\) are cointegrated if
\( \beta < 0 \), and non-cointegrated if \( \beta = 0 \). Thus, in the ECM approach as in equation (1), the t-ratio based upon the Least Squares (OLS) estimate \( \hat{\beta} \), denoted \( \hat{\beta}_{OLS} \), is the ECM statistic (denoted \( T_{ECM} \)). That is, the estimated equation (1) is:

\[
\Delta y_t = \hat{\alpha} \Delta x_t + \hat{\beta} \left( y_{t-1} - x_{t-1} \right) + \hat{\varepsilon}_t.
\]

The ECM statistic is \( T_{ECM} = \hat{\beta} / \text{ese}(\hat{\beta}) \), where \( \text{ese}(.) \) denotes the estimated standard error of its argument, to be used in testing the null hypothesis that \( \beta = 0 \), i.e., that \( y_t \) and \( x_t \) are not cointegrated.

The condition for finite variance is \( \phi_i + \varphi_i < 1 \) and the condition for finite fourth moment is \( 3\phi_i^2 + 2\phi_i\varphi_i + \varphi_i^2 < 1 \). And if \( \gamma_i > 0 \) and \( \phi_i + \varphi_i < 1 \), then the unconditional variance of the \( \varepsilon_i \), exist and equals \( \sigma_{\varepsilon_i}^2 = (\gamma_i - \phi_i - \varphi_i) \).

Note that when \( \phi = \varphi = 0 \), the \( \varepsilon_i \) reduced to iid white noises.

In the case when the error components \( (\varepsilon_{1t}, \varepsilon_{2t}) \) in (1) and (2) are iid as bivariate normal with \( E(\varepsilon_{1t}) = E(\varepsilon_{2t}) = 0 \) and \( \text{Var}(\varepsilon_{1t}) = \sigma_{\varepsilon_1}^2 \), \( \text{Var}(\varepsilon_{2t}) = \sigma_{\varepsilon_2}^2 \) and \( \text{Cov}(\varepsilon_{1t}, \varepsilon_{2t}) = 0 \), \( \forall t \), studied by Kremers, Ericsson and Dolado (1992), KED.

KED examine the reason for the low power of the Dickey-Fuller test, (DF) and by using Monte Carlo methods studied the small sample properties of two different versions of the cointegration test, DF and ECM. KED found also by defining a "signal-to-noise" ratio:

\[ q = - (\alpha - 1)s \]

where \( s \) denote the ratio \( \sigma_{\varepsilon_2} / \sigma_{\varepsilon_1} \) (assumed strictly positive), and \( q^2 \) is the variance of \( (\alpha - 1)\Delta x \), relative to \( \varepsilon_{1t} \) (see KED Section 3, p. 330) and using three different sets of asymptotic critical values, that the size of the ECM tests is relative to \( q \). That is, the size of the ECM test in the KED study shows that there is a need for a set of robust critical values.
One way to solve this problem and produce a set of robust critical values for ECM test is to use bootstrap technique. Mantalos and Shukur (1998), MS, by using the same DGP as in KED, improved the critical values of the ECM test statistic by employing residual based bootstrap technique and they showed with Monte Carlo experiment that the size of the test approaches its nominal value without loss of power. However here we use the Wild Bootstrap, Wu (1986) that is robust against heteroskedasticity.

Before we describe our Monte Carlo Experiment we have to explain why we study the near Integrated GARCH model that the residual from ECM model (3) follows.

Because the unconditional variance equal \( \left( \gamma, 1 - \phi - \varphi \right) \), one might expect that \( \gamma \) is close to zero when \( \phi + \varphi \) is close to one, and vice versa, so that the unconditional variance is not too large or too small.

Moreover as Kim and Schmidt (1993) point out in their study that Nelson (1990, 1992) considers GARCH model as an approximation to an underlying continuous-time diffusion process. Letting \( L \to 0 \) if the parameters follow the sequence:

\[
\gamma = \omega L, \quad \phi = a \sqrt{L}, \quad \text{and} \quad \varphi = 1 - \phi,
\]

with “\( \omega \)” and “\( a \)” fixed. According to this result we might expected GARCH model applied to high-frequency data to be nearly integrated and exhibit small values of \( \gamma \) and \( \phi \) and finally \( \varphi \) near unity.

Furthermore the ECM is an extension to Dickey-Fuller test for unit roots in system of variables, and is interesting to compare the Kim and Schmidt (1993) results with our study. And also because the GARCH model applied to stock market data, in Mantalos (1999) study, was nearly integrated and exhibit small values of \( \gamma \) and \( \phi \) and \( \varphi \) near unity.

The Monte Carlo experiment has been performed by generating data according to the model defined by (1) and (2), we simulate also three near integrated GARCH with \( L \) equal to 0.25,
0.09 and 0.0064 and three GARCH with a) high persistence (0.01,0.09,0.9), b) medium persistence (0.05, 0.05, 0.9) and c) low persistence (0.20,0.05,0.75).

We chose the parameter values, by fixing $\sigma_{\epsilon t}^2 = 1$, and leaving the parameters $(\sigma_{\epsilon t}^2, \alpha, \beta, \gamma, \phi, \varphi)$ and the sample size $T$ as experimental design variables.

About the significance levels to be used when judging the properties of the tests, different authors have put forward reasons for using both larger and smaller significance levels. Maddala (1992) suggests using significance levels of as much as 25% in diagnostic testing, while MacKinnon (1992) suggest going in the other direction.

To reduce this problem, in this study, we use mainly graphical methods that may provide more information about the size and the power of the tests. We use simple graphical methods that developed and illustrated by Davidson and MacKinnon (1997) and are easy to interpret, the “P value plot” to study the size and the “Size-Power curves” to study the power of the tests.

Furthermore to judge the reasonability of the results we use a 95% confidence interval for the actual size ($\pi$) as : $\pi_0 \pm 2 \sqrt{\frac{\pi_0(1-\pi_0)}{N}}$, where $N$ is the number of replications. Results that lie between these bounds will be considered satisfactory.

For each time series 20 presample values are generated with zero initial conditions, taking net sample sizes of $T = 100, 500$.

Table 1 shows the different parameters of our Monte Carlo design. The number of replications per model is 10,000, for the size and 1000 for the power of the tests. The calculations were performed using GAUSS 3.2.
Table 1: Monte Carlo Parameter

<table>
<thead>
<tr>
<th>Size</th>
<th>α</th>
<th>b</th>
<th>γ</th>
<th>φ</th>
<th>(1−γ−φ)</th>
<th>σ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near Integrated</td>
<td>1/0.5</td>
<td>0.02</td>
<td>0.005</td>
<td>0.3√L</td>
<td>(1−γ−φ)</td>
<td>1/16</td>
</tr>
<tr>
<td>High Persistence</td>
<td>1/0.5</td>
<td>0.02</td>
<td>0.01</td>
<td>0.09</td>
<td>0.90</td>
<td>1/16</td>
</tr>
<tr>
<td>Medium</td>
<td>1/0.5</td>
<td>0.02</td>
<td>0.05</td>
<td>0.05</td>
<td>0.90</td>
<td>1/16</td>
</tr>
<tr>
<td>Low</td>
<td>1/0.5</td>
<td>0.02</td>
<td>0.20</td>
<td>0.05</td>
<td>0.75</td>
<td>1/16</td>
</tr>
</tbody>
</table>

Power

<table>
<thead>
<tr>
<th>Size</th>
<th>α</th>
<th>b</th>
<th>γ</th>
<th>φ</th>
<th>(1−γ−φ)</th>
<th>σ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near Integrated</td>
<td>1/0.5</td>
<td>0.02</td>
<td>0.005</td>
<td>0.3√L</td>
<td>(1−γ−φ)</td>
<td>1/16</td>
</tr>
<tr>
<td>High Persistence</td>
<td>1/0.5</td>
<td>0.02</td>
<td>0.01</td>
<td>0.09</td>
<td>0.90</td>
<td>1/16</td>
</tr>
<tr>
<td>Medium</td>
<td>1/0.5</td>
<td>0.02</td>
<td>0.05</td>
<td>0.05</td>
<td>0.90</td>
<td>1/16</td>
</tr>
<tr>
<td>Low</td>
<td>1/0.5</td>
<td>0.02</td>
<td>0.20</td>
<td>0.05</td>
<td>0.75</td>
<td>1/16</td>
</tr>
</tbody>
</table>

Where $L$ is 0.25, 0.09, 0.0064

The estimated equation (1):

$$\Delta y_t = \alpha \Delta x_t + \beta (y_{t-1} - x_{t-1}) + \epsilon_{yt} \quad (3)$$

and the ECM statistic $T_{ECM}$ is used in testing the null hypothesis that $\beta = 0$, i.e., that $y_t$ and $x_t$ are not cointegrated.

The ARCH LM procedure test for autoregressive conditional heteroskedasticity (ARCH) is used. And particularly the test is based on the regression of squared residuals from equation (3) on one lagged, squared residuals and a constant.

The output from the test is a TR2 statistic, distributed as $\chi^2$, with degrees of freedom equal to the number of lagged, squared residuals, that is in our case one. Each statistic provides a test of the hypothesis that the coefficients of the lagged squared residuals are all zero that is no ARCH.

About our Bootstrap procedure we use the wild Bootstrap technique. A direct wild Bootstrap gives:

$$\Delta y_t^* = \alpha \Delta x_t + \epsilon_{yt} \quad (4)$$

to test if the variables are cointegrated. Where $u_t^*$ are i.i.d observations drawn from some distribution with c.d.f F, defined so as to satisfy:
\[ E(u_t^* \mid F) = 0, \quad E(u_t^{2} \mid F) = 1 \quad \text{and} \quad E(u_t^{3} \mid F) = 1. \] (5)

Let us denote by \( T_s \) the ECM test statistics, and by \( T_s^* \) the ”Bootstrap test statistics”.

A Bootstrap estimate of the P-value for testing is \( P^* \{ T_s^* \geq T_s \} \) and this approach we use to examine the size of the Bootstrap test with the help of the ”P-value plot”.

The number of the Bootstrap sample used to estimate Bootstrap critical values and P-values in our study is \( N_b = 199 \).

3. Results of the Size of the Tests

In this section we present the results of our Monte Carlo experiment concerning the sizes of the tests.

Now for the P value plots, we have that if the distribution used to compute the \( p_s \) is correct, each of the \( p_s \) should be distributed as uniform (0,1). Therefore the resulting graph should be close to the 45° line i.e. as Figure 1a shows.

Figure 1 shows the ARCH LM test for different “L”, and we see that the test has the right size for white noise residuals and has adequate power with ARCH parameter larger than 0.09.

Note also that the Figure 1a for ARCH parameter 0.024 shows that the power of the test is very low and is almost equal to the size. This result explains why the ARCH tests in Mantalos (1999) shows no ARCH effects on the ECM residuals. That is if two series are individually I(1) and follow a GARCH(1,1) process that is nearly Integrated with the same GARCH parameter and the small volatility parameter(<0.03), then the ARCH-tests show that there is no ARCH effect on the least squares residual by regress the first difference of the series.
Figure 1 ARCH test with different L

Figure 1a White Noise

Figure 1a L=0.0064

Figure 1a L=0.09

Figure 1a L=0.25

Note that the whole lines for the figures are the approximate 95% confidence interval for actual size.

Figure 2 shows some of the advantages of the P value plots, since they make easy to distinguish between tests that work badly, and tests that work well. We see that in almost all the cases, for large sample (500 observations), the ECM cointegration tests with least squares estimation seems to be robust for GARCH innovations. However ECM test over-rejects the null hypothesis for small sample. Note also in the near integrated case it is a tendency to over-rejects null hypothesis even for large sample when the ARCH parameter is not small.
Figure 2: P Value Plots for ECM-Cointegration (GARCH innovations).

Figure 2a: High Persistence GARCH

Figure 2b: Medium Persistence GARCH

Figure 2c: Low Persistence GARCH

Figure 2d: Near Integrated : L=0.25

Figure 2e: Near Integrated : L=0.09

Figure 2f: Near Integrated : L=0.0064

Note that the dot lines for the figures are the approximate 95% confidence interval for actual size.
Figure 3: P Value Plots for ECM-Cointegration (GARCH innovations), q Effect.

Now, Figure 3 shows the combination of the GARCH and q effect on the ECM-cointegration test. Remind the results from KED and MS studies; that for large q the ECM tests systematically under-reject the null hypothesis. However, it seems to be a small tendency that the over-reject from the GARCH effect to cancel the under-reject effect from large q, note also that we q is only 2 in our experiment. Moreover the q effects it is obviously that dominate over the GARCH effects.
By study the Figure 4, which shows the rejection of the tests on 5% nominal size, we can summarise the results about the size of the tests, in our experiment.

Without q effect ECM cointegration test rejects about the right proportion of the time the null hypothesis for large sample and for small ARCH parameter in near integrated case and over reject for small sample even when there is low persistence effects.

With q=2 there is no large effect from the GARCH effects on ECM tests it seems that the q effect dominates.

Figure 4. Comparison of The “q” and “GARCH” on 5% confidence interval

The Bootstrap test rejects about the right proportion of the time the null hypothesis, as Figure 5 shows, in the 95% confidence interval for both case without q effect and with q=2. Note that we use only 100 Bootstrap-resample in our experiment more Bootstrap-resample makes the line smooth without change the result.
Figure 5 Bootstrap Test

Figure 5a Bootstrap High

![Diagram of Figure 5a Bootstrap High]

Figure 5b Bootstrap Medium

![Diagram of Figure 5b Bootstrap Medium]

Figure 5c Bootstrap Low

![Diagram of Figure 5c Bootstrap Low]
Finally, Figure 6 for 100 observations summarise all the above results for high, medium and low persistence GARCH with or without q effects and results is that the wild Bootstrap is robust. The ECM-cointegration test tend to over-reject in the presence of GARCH errors but the problem is not very serious in large samples except when the error’s GARCH process is nearly integrated and the volatility parameter is not small. And this result is the same with Kim and Schmidt (1993) results.
4. Analysis of the Power of the Tests

In this section we discuss the most interesting results of our Monte Carlo experiment, designed to gather evidence concerning the power of the various test versions. We analysed the power of the ECM tests using sample sizes of, 100, and 500 observations and Bootstrap tests using sample sizes of, 100, observations. The power function is estimated by calculating the rejection frequencies in 1000 replications using values of the $\beta$ coefficients = 0,2. The estimated power functions of the tests have been compared only graphically as in the size case. We use the Size-Power Curves to compare the power of alternative test statistics. This has proved to be quite adequate, since the tests that gave reasonable results as regard size usually differed very little regarding power. We follow the same procedure as for the size investigation to evaluate the EDF’s denoted $\hat{F}(x_j)$, by using the same sequence of random numbers as the one that we used to estimate the size of the tests.

Plotting the estimated power functions against the nominal size, we have the Size-Power Curves. Note that the Size-Power Curves on a correct size-adjusted basis i.e plotting the estimated power functions against the true size, that is $\hat{F}(x_j)$ against $\hat{F}(x_j)$, do not show much different than the standard Size-Power Curves and we do not show here.
Figure 7 Power –Size Plots of the ECM Test

Figure 7a Power of High Persistence

Figure 7b Power of Medium Persistence

Figure 7c Power of Low Persistence

Figure 7d Power with L=0.25

Figure 7e Power with L=0.09

Figure 7f Power with L=0.0064
Figure 7 presents the Size-Power Curves for the tests for different GARCH models in all cases we see that it is a sample effect, higher power with large sample. We do not notice any GARCH effect (solid line is the White noise model).

However Figure 8 shows, the already known from KED MS studies, that the larger q is, the higher power of the test. Even with the Size-Power Curves on a correct size-adjusted basis (Figure 8c) the result does not change.

Finally the Bootstrap test as Figure shows is the same as the “asymptotic” tests. For q=1 and a little higher with q=2. That is the Wild Bootstrap has the right size without lost of power.

Figure 8: Power size Plots of the ECM q effects

<table>
<thead>
<tr>
<th>Figure 8a Power of High Persistence</th>
<th>Figure 8c Power of Low Persistence</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="#" alt="Power of High Persistence" /></td>
<td><img src="#" alt="Power of Low Persistence" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 8b Power of Medium Persistence</th>
<th>Figure 8a Power of correct size-adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="#" alt="Power of Medium Persistence" /></td>
<td><img src="#" alt="Power of correct size-adjusted" /></td>
</tr>
</tbody>
</table>
5. Summary and conclusions

By using Monte Carlo experiment, we study the robustness of the ECM-Cointegration tests when the first difference of the series follow a GARCH(1,1) process and the behaviour of the ARCH-tests of ECM residuals for variation of GARCH parameters. We defining a ”signal-to-noise” ratio: $q = - (\alpha - 1) s$, where $s$ denote the ratio $\sigma_{\varepsilon_2} / \sigma_{\varepsilon_1}$ (assumed strictly positive), and $q^2$ is the variance of $(\alpha - 1) \Delta x_t$ relative to $\varepsilon_{1t}$ (see Section 2).

We use simple graphical methods that developed and illustrated by Davidson and MacKinnon (1997) and are easy to interpret, the ”P value plot” to study the size and the ”Size-Power curves” to study the power of the tests.

Without $q$ effect ECM cointegration test rejects about the right proportion of the time the null hypothesis for large sample and for small ARCH parameter in near integrated case and over reject for small sample even when there is low persistence effects.
With $q=2$ there is no large effect from the GARCH effects on ECM tests it seems that the $q$ effect dominates.

The Bootstrap test rejects about the right proportion of the time the null hypothesis, as Figure 5 shows, in the 95% confidence interval for both case without $q$ effect and with $q=2$.

About the power of the tests the Size-Power Curves for different GARCH models in all cases shows that there is a sample effect, higher power with large sample. We do not notice any GARCH effect.

However Fig the Size-Power Curves show, the already known from KED MS studies, that the larger $q$ is, the higher power of the test. Even with the Size-Power Curves on a correct size-adjusted basis (Figure 8c) the result does not change.

Finally the Bootstrap test as Figure shows is the same as the “asymptotic” tests. For $q=1$ and a little higher with $q=2$.

That is the Wild Bootstrap has the right size without lost of power.
References


Wu, C.F.J. (1986); “Jackknife, bootstrap and other resampling methods in regression analysis, *Ann.Statist.*, 14, 1261-1350