Finite Sample and Optimal Inference in Possibly Nonstationary ARCH Models with Gaussian and Heavy-Tailed Errors*

JEAN MARIE DUFOUR[†] and EMMA M. IGLESIAS[‡]
Université de Montréal University of Alicante

First version: April 27th, 2004

Abstract

Most of the literature on testing ARCH models focuses on the null hypothesis of no-ARCH effects. In this paper, we consider the general problem of testing any possible set of coefficient values in ARCH models, which may be non-stationary, with Gaussian and non-Gaussian errors, as well as with any number exogenous regressors in the mean equation. Both Engle-type and point-optimal tests are studied. Special problems considered include the hypothesis of no-ARCH effects and IARCH structure. We propose exact inference based on pivotal Monte Carlo tests [as in Dufour and Kiviet (1996, 1998) and Dufour, Khalaf, Bernard and Genest (2004)] and maximised Monte Carlo tests [Dufour (2004))], depending on whether nuisance parameters are present. This will allow the introduction of dynamics in the mean equation as well. We show that the method suggested provides provably valid tests in both finite and large samples, in cases where standard asymptotic and bootstrap methods may fail in the presence of heavy-tailed errors [as shown by Hall and Yao (2003)]. The performance of the proposed procedures with both Gaussian and non-Gaussian

*The second author thanks the hospitality of the Department of Economics at the University of Montreal while visiting them, and the financial support from an ESRC grant (Award number: T026 27 1238).

†Canada Research Chair Holder (Econometrics). CIRANO, CIREQ, and Département de sciences économiques, Université de Montréal. Mailing address: Département de sciences économiques, Université de Montréal, C. P. 6128 succursale Centre-ville, Montréal, Québec, Canada H3C 3J7. TEL: 1 514 343 2400; FAX: 1 514 343 5831; e-mail: jean.marie.dufour@umontreal.ca. Web page: http://www.fas.umontreal.ca/SCECO/Dufour.

‡Dpt. Fundamentos del Análisis Económico. University of Alicante. Campus de San Vicente, 03080, Alicante, Spain. E-mail address: emiglesias@merlin.fae.ua.es.

errors is analyzed in a simulation experiment. Our results show that the proposed procedures work well from the viewpoints of size and power. The powers gains provided by the point optimal procedures are in many cases spectacular. The tests also exhibit good behaviour outside the stationarity region [following the work of Jensen and Rahbek (2004)]. Finally, the technique is applied to the US inflation.

1 Introduction

In this paper we focus on developing finite sample inference procedures as well as optimal tests for ARCH(p) models. Specially, we concentrate on testing any value of the conditional heteroscedastic coefficients. This also allows us to test for the presence of ARCH effects and the unit root case (with the IARCH process), and we compare our results with other available procedures. There are already available several tests for accounting for ARCH and GARCH effects: see for example Engle (1982), Lee (1991) and Lee and King (1993). More recently, Dufour, Khalaf, Bernard and Genest (2004) have proposed to improve the properties of the inference by exploiting the use of Monte Carlo (MC) techniques. In this paper we propose also the use of MC techniques, although we go further, and we present new procedures for testing any value of the conditional heteroscedastic coefficients. Among others, we propose a point optimal test. We also use the Maximised Monte Carlo (MMC) technique to deal with nuisance parameters (see Dufour (2004) for more details).

In the case of testing the null of IGARCH(1,1), Lumsdaine (1995) gets results that Wald tests seem to have the best size, although the standard Lagrange multiplier statistic is badly oversized. At the same time, versions of the LM that are robust to possible nonnormality of the data perform only marginally better. In any case, Lumsdaine reports that in general, from her simulations, the Lagrange multiplier, likelihood ratio and Wald do not behave very well in small samples. Our framework also allows us to test for this hypothesis and with the MC technique we will be able to control for the size.

We also show through simulation that the tests developed in this paper can be applied both in the framework of gaussian and non-gaussian errors; including errors following a t-distribution with very low degrees of freedom (where asymptotic theory breaks down). Hall and Yao (2003) report problems with the conservatism of their subsampling technique when the tails of the errors are very light and problems with the anticonservatism for heavy-tails. Due to the fact that our procedure controls the size and the exactness of the test, this makes our proposal more attractive than the subsampling technique of Hall and Yao (2003). Our simulations also show that our optimal procedure has very good power. That implies that the procedures developed in this paper can be used by practicioners in any type of scenarios: including gaussian and non-gaussian errors without being worried about the existence of moments. Besides, Hall and Yao (2003) only show results for sample sizes 1000 and 500 and even already in those cases their procedure has size problems. We will show the good perfomance of our procedure even for sample sizes of 50.

Jensen and Rahbek (2004) have shown recently that the QMLE is always asymptotically normal provided that the fourth moment of the innovation process is finite in an ARCH(1) model. However, although this shows what happens asymptotically, it is still unknown in the literature which is the effect of being in a non-stationarity region in finite samples. In this paper we also consider a case where we are in a non-stationary region, and we will provide the behaviour of tests in this framework. Again in this case, the existence of moments is crucial for the result of Jensen and Rahbek (2004), while our tests (through the MC) can work both in the non-stationary region and in those cases where moments do not exist (with very fat tails).

Finally, making use of the Monte Carlo technique, we can afford the introduction of any number of exogenous variables in a very straightforward way. With asymptotic approximations, this would change the framework of the test, while the Monte Carlo technique takes that into account directly. We also make operational our procedures for testing the null of a sub-group of the ARCH coefficients equal to a value by making use of the Maximised Monte Carlo (MMC) technique. We go even further, and with the MMC we can allow for the existence of dynamics in the mean equation.

The inmediate applications of the results in this paper are several. Among them, first, from our inference procedures we can construct confidence intervals for conditional volatility. Second, we can retrieve from there predictions of volatility through the confidence intervals and point predictions. Third, from our method we can get predictions of the underlying variables.

The plan of the paper is as follows. We first present three different tests for any value of the ARCH coefficients, among which we consider a point optimal test. Later we carried out a simulation study to find out about the size and power properties of the proposed test procedures. In Section 3 we show the usefulness of our test in practice, when we re-visit the analysis of the US implicit price deflator for GNP. Finally, in Section 4 we conclude.

2 Alternative test statistics for any value of the ARCH coefficients

We proceed now to propose alternative tests for any value of the ARCH coefficients. We consider the next model:

$$y_t = x_t^{\prime}\beta + \sum_{i=1}^{q} \phi_i y_{t-i} + \varepsilon_t \tag{1}$$

$$\varepsilon_t = \sqrt{h_t}\eta_t, t = 1, ...T, \; h_t = E\left(\varepsilon_t^2 / I_{t-1}\right) = \theta_0 + \theta_1\varepsilon_{t-1}^2 + \theta_2\varepsilon_{t-2}^2 + ... + \theta_p\varepsilon_{t-p}^2 \tag{2}$$

where $x_t = (x_{t1}, x_{t2}, ..., x_{tk})'$, $X \equiv [x_1, ..., x_T]'$ is a full-column rank $T \times k$ matrix, $\beta = (\beta_1, ..., \beta_k)'$ is a $k \times 1$ vector of unknown coefficients, $\sqrt{h_1}, ..., \sqrt{h_T}$ are (possibly random) scale parameters, and $\eta_t = (\eta_1, ..., \eta_T)'$ is a random vector. The case of a conditional gaussian distribution is a special case, although we can allow for possibly any other distribution.

Let's consider first the case where we only deal with exogenous in the mean equation. We are interested in the problem of testing any possible set of values for the ARCH coefficients:

$$H_0 : h_t = \overline{h}_t = \overline{\theta}_0 + \overline{\theta}_1 \varepsilon_{t-1}^2 + ... + \overline{\theta}_p \varepsilon_{t-p}^2 \tag{3}$$

We stress the fact that our procedures allow as well to test the null hypothesis of no-ARCH and the integrated ARCH cases. Our scenario using the MC technique allows us as well the introduction of exogenous variables in the mean equation, and we also show later in the simulation study that we can allow for the presence of normal or non normal errors (including those cases where asymptotic theory even breaks down (see Hall and Yao (2003)). Due to the fact that we are using residual-based tests (see e.g. Dufour, Khalaf, Bernard and Genest (2004) for more details), we can justify that the tests are invariant to the choice of the intercept in the conditional variance and in the mean equation, and to the parameters of any number of exogenous variables that are included in the mean equation. We justify this in the following proposition and corollary:

Proposition 1 *(Pivotality of a statistic).*
Under (1) and (2), let $S(y, X) = (S_1(y, X), S_2(y, X), ..., S_m(y, X))'$ *be any vector of real-valued statistics such that*

$$S(cy + Xd, X) = S(y, X), \forall c, d \in R^k.$$

Then, for any positive constant $\sqrt{h_0} > 0$, *we can write*

$$S(y, X) = S\left(\varepsilon / \sqrt{h_0}, X\right)$$

and the conditional distribution of $S(y, X)$, *given* X, *is completely determined by the matrix* X *and the conditional distribution of* $\varepsilon / \sqrt{h_0} = \Delta\eta / \sqrt{h_0}$ *given* X, *where* $\Delta = diag\ (\sqrt{h_t}, t = 1, ..., T)$. *In particular, under* H_0 *in (3), we have:*

$$S(y, X) = S(\eta, X)$$

where $\eta = \varepsilon / \sqrt{\overline{h}_t}$, *and the conditional distribution of* $S(y, X)$, *given* X, *is completely determined by the matrix* X *and the conditional distribution of* η *given* X.

Proof. Taking $c = 1/\sqrt{h_0}$ and $d = -\beta/\sqrt{h_0}$, then:

$$cy + Xd = (X\beta + \varepsilon)/\sqrt{h_0} - X\beta/\sqrt{h_0} = \varepsilon/\sqrt{h_0}.$$

From (2) then, and under (3), we have that $\varepsilon = \Delta\eta = \sqrt{h_t}\eta$, and then, taking $\sqrt{h_0} = \sqrt{h_t}$, we get finally $\varepsilon/\sqrt{h_0} = \eta$ and $S(y, X) = S(\eta, X)$. ■

It is also necessary to prove that the pivotality characteristic also holds because in our case our statistics are scale-invariant functions of OLS residuals. We prove this in the next Corollary:

Corollary 1 *(Pivotal property of residual-based statistics).*

Under (1) and (2), let $S(y, X) = (S_1(y, X), S_2(y, X), ..., S_m(y, X))'$ *be any vector of real-valued statistics such that*

$$S(y, X) = \overline{S}(A(X)y, X)$$

where $A(X)$ *is any* $n \times k$ *matrix* $(n \geq 1)$ *such that*

$$A(X)X = 0$$

and $\overline{S}(A(X)y, X)$ *satisfies the scale invariance condition*

$$\overline{S}(cA(X)y, X) = \overline{S}(A(X)y, X), forallc > 0.$$

Then for any positive constant $\sqrt{h_0} > 0$, *we can write*

$$S(y, X) = \overline{S}\left(A(X)\varepsilon/\sqrt{h_0}, X\right)$$

and the conditional distribution of $S(y, X)$, *given* X, *is completely determined by the matrix* X *jointly with the conditional distribution of* $A(X)\varepsilon/\sqrt{h_0}$ *given* X.

It is straightforward to show that in all the statistics that we present in this paper $A(X) = I_T - X(X'X)^{-1}X'$ (to apply in Corollary 1).

In the case we wanted to test sub-vector coefficients of the ARCH process, we could still preserve the exactness of the test by dealing with the nuisance parameters through the MMC technique. The same type of methodology would be used in case we would need to incorporate dynamics in the mean equation.

2.1 Extension of the Engle test under a null hypothesis different from the one of no-ARCH effects.

A first alternative we consider is what we name to be an extension of "an Engle-type test". The original Engle (1982) test was designed as an LM test:

$$h_t = E\left(\varepsilon_t^2/I_{t-1}\right) = \theta_0 + \theta_1\varepsilon_{t-1}^2 + \theta_2\varepsilon_{t-2}^2 + ... + \theta_p\varepsilon_{t-p}^2$$

where the null hypothesis is $H_0 : \theta_1 = \theta_2 = ... = \theta_p = 0$, h_t is the conditional variance, and the test is formulated by TR^2 where T is the sample size and R^2 is the determination coefficient of a regression of OLS residuals $\widehat{\varepsilon}_t^2$ on a constant and $\widehat{\varepsilon}_{t-i}^2$ for $i = 1, ..., p$. The test is distributed as a χ^2 with p degrees of freedom.

In this paper we propose the next extension:

Theorem 1 *Suppose that* $y_t = x_t\acute{}\beta + \sum_{i=1}^{q}\phi_i y_{t-i} + \varepsilon_t$ *and* $h_t = E\left(\varepsilon_t^2/I_{t-1}\right) = \theta_0 + \theta_1\varepsilon_{t-1}^2 + \theta_2\varepsilon_{t-2}^2 + ... + \theta_p\varepsilon_{t-p}^2$*, then, an extension of the Engle (1982) test to any possible set of values of the ARCH coefficients* $H_0 : \theta_1 = \overline{\theta}_1, \theta_2 = \overline{\theta}_2, ..., \theta_p = \overline{\theta}_p$*, is given by* $TR^2 \sim \chi_p^2$*, where* $\overline{h}_t = \overline{\theta}_0 + \overline{\theta}_1\varepsilon_{t-1}^2 + ... + \overline{\theta}_p\varepsilon_{t-p}^2$ *and* R^2 *is the determination coefficient coming from the regression:*

$$\varepsilon_t^2 = 2\overline{h}_t^2\left[\gamma_0 + \gamma_1\varepsilon_{t-1}^2 + \gamma_2\varepsilon_{t-2}^2 + ... + \gamma_p\varepsilon_{t-p}^2\right] + \overline{h}_t + v_t.$$

In practice, the test would imply to take the residuals $\widehat{\varepsilon}_t^2$, to compute the dependent variable $\left(\frac{\widehat{\varepsilon}_t^2}{2\left(\overline{\theta}_0 + \overline{\theta}_1\widehat{\varepsilon}_{t-1}^2 + ... + \overline{\theta}_p\widehat{\varepsilon}_{t-p}^2\right)^2} - \frac{1}{2\left(\overline{\theta}_0 + \overline{\theta}_1\widehat{\varepsilon}_{t-1}^2 + ... + \overline{\theta}_p\widehat{\varepsilon}_{t-p}^2\right)}\right)$, and to regress this depend variable on a constant and $\widehat{\varepsilon}_{t-i}^2$ for $i = 1, ..., p$. The test is distributed as a χ^2 with p degrees of freedom.

We present now how we obtain the previous expression. The construction of (what we denote) an Engle test for the null of any possible set of values for the ARCH coefficients, implies the following. Let's suppose an ARCH(p) model:

$$y_t = x_t\acute{}\beta + \sum_{i=1}^{q}\phi_i y_{t-i} + \varepsilon_t$$

$$h_t = E\left(\varepsilon_t^2/I_{t-1}\right) = \theta_0 + \theta_1\varepsilon_{t-1}^2 + \theta_2\varepsilon_{t-2}^2 + ... + \theta_p\varepsilon_{t-p}^2$$

The log-likelihood:

$$L \propto -\frac{1}{2}\sum_{t=1}^{T}\log h_t - \frac{1}{2}\sum_{t=1}^{T}\frac{\varepsilon_t^2}{h_t}$$

So the gradient (grad) and the hessian (hes) are:

$$grad = \left[\frac{\varepsilon_t^2}{2h_t^2} - \frac{1}{2h_t}\right]Z$$

$$hes = \left[\frac{1}{2h_t^2} - \frac{\varepsilon_t^2}{h_t^3} \right] ZZ'$$

where Z is the vector $Z = \left(1, \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, ..., \varepsilon_{t-p}^2\right)'$.

Under any null hypothesis, we have to get the R^2 coming from regressing a column of ones on the derivatives of the log likelihood function computed at the restricted estimator.

Following Engle (1982), we can denote f^0 to be the first part of the gradient evaluated at the restricted estimator under the null:

$$f^0 = \left[\frac{\varepsilon_t^2}{2h_t^2} - \frac{1}{2h_t} \right]_0$$

So, the Engle test comes from TR^2, where R^2 comes from regressing f^0 on Z.

In the special case where the null is no-ARCH effects, since (following Engle (1982)) adding a constant and multiplying by a scalar won't change the R^2 of a regression, this will be the R^2 coming from regressing ε_t^2 on an intercept and p lagged values of ε_t^2.

In the general case of any other null except the no-ARCH effects one, h_t evaluated at a restricted estimator is going to contain lagged ε_t^2 and the values we are testing. The test in this case can be interpreted as regressing under the null:

$$\left[\frac{\varepsilon_t^2}{2h_t^2} - \frac{1}{2h_t} \right]_0 = \gamma_0 + \gamma_1 \varepsilon_{t-1}^2 + \gamma_2 \varepsilon_{t-2}^2 + ... + \gamma_p \varepsilon_{t-p}^2 + v_t \tag{4}$$

Or:

$$\varepsilon_t^2 = 2h_t^2 \left[\gamma_0 + \gamma_1 \varepsilon_{t-1}^2 + \gamma_2 \varepsilon_{t-2}^2 + ... + \gamma_p \varepsilon_{t-p}^2 \right] + h_t + v_t$$

So running (4), can be equivalent to run as well a regression of ε_t^2 under the null, on $2h_t^2$ times the usual stuff in the Engle test plus the h_t (everything evaluated under the null). Implicitly, because h_t has ε_{t-i}^2 inside, in order to construct the confidence sets, this may have some equivalent representation to running ε_t^2 under the null on a constant, as many lags of ε_t as the order of the ARCH(p) we are testing, raised to six, to four and to two (so, terms of the form $\varepsilon_{t-i}^6, \varepsilon_{t-i}^4, \varepsilon_{t-i}^2, \forall i = 1, ..., p$), and cross products of as many lags of ε_t as the order of the ARCH(p) we are testing raised to four and two (so terms of the form $\varepsilon_{t-i}^4 \varepsilon_{t-j}^2, \forall i \neq j$), plus cross products of as many lags of ε_t as order of the ARCH(p) we are testing raised to two (so terms of the form $\varepsilon_{t-i}^2 \varepsilon_{t-j}^2, \forall i \neq j$).

But with the "problem" of multiplying in each case the previous terms by the values we are testing under the null.

2.2 A Point Optimal Test

To find out the most powerful test for ARCH processes, we develop now a point optimal test.

Theorem 2 *Suppose that* $y_t = x_t'\beta + \sum_{i=1}^{q}\phi_i y_{t-i} + \varepsilon_t$ *and* $h_t = E\left(\varepsilon_t^2/I_{t-1}\right) = \theta_0 + \theta_1\varepsilon_{t-1}^2 + \theta_2\varepsilon_{t-2}^2 + ... + \theta_p\varepsilon_{t-p}^2$. *Then a point optimal test at* $\theta_1 = \theta_1^1, ..., \theta_p = \theta_p^1$ *for the null of* $H_0 : \theta_1 = \overline{\theta}_1, \theta_2 = \overline{\theta}_2, ..., \theta_p = \overline{\theta}_p$, *under conditional normality is given by:*

$$LR\left(\overline{\theta}_1, ..., \overline{\theta}_p, \theta_1^1, ..., \theta_p^1\right) = \sum_{t=1}^{T}\ln\left(\overline{h}_t\right) + \sum_{t=1}^{T}\left(\frac{1}{\overline{h}_t} - 1\right)\frac{\varepsilon_t^2}{\left(1 + \overline{\theta}_1\varepsilon_{t-1}^2 + ... + \overline{\theta}_p\varepsilon_{t-p}^2\right)}$$

where $\overline{h}_t = \left(\frac{\left(1+\theta_1^1 y_{t-1}^2+...+\theta_p^1 y_{t-p}^2\right)}{\left(1+\overline{\theta}_1 y_{t-1}^2+...+\overline{\theta}_p y_{t-p}^2\right)}\right)$.

In practice, the test implies the following. Take the residuals $\widehat{\varepsilon}_t^2$, and compute the test statistic:

$$\sum_{t=1}^{T}\ln\left(\frac{\left(1+\theta_1^1\widehat{\varepsilon}_{t-1}^2+...+\theta_p^1\widehat{\varepsilon}_{t-p}^2\right)}{\left(1+\overline{\theta}_1\widehat{\varepsilon}_{t-1}^2+...+\overline{\theta}_p\widehat{\varepsilon}_{t-p}^2\right)}\right)+\sum_{t=1}^{T}\left(\frac{\left(1+\overline{\theta}_1\widehat{\varepsilon}_{t-1}^2+...+\overline{\theta}_p\widehat{\varepsilon}_{t-p}^2\right)}{\left(1+\theta_1^1\widehat{\varepsilon}_{t-1}^2+...+\theta_p^1\widehat{\varepsilon}_{t-p}^2\right)} - 1\right)\frac{\widehat{\varepsilon}_t^2}{\left(1+\overline{\theta}_1\widehat{\varepsilon}_{t-1}^2+...+\overline{\theta}_p\widehat{\varepsilon}_{t-p}^2\right)}$$

Due to the non-standard distribution of the statistic, in this paper we propose to retrieve the critical values through the MC technique (see Dufour and Kiviet (1996, 1998), Dufour, Khalaf, Bernard and Genest (2004) for more details).

We will provide now with an example of the ARCH(1) process how the point optimal test is obtained. Let's suppose a simple ARCH(1) process:

$$y_t = \left(1 + \theta_1 y_{t-1}^2\right)^{1/2} v_t$$

Under H_0: $\theta_1 = \overline{\theta}_1$,

$$v_t = \frac{y_t}{\left(1 + \overline{\theta}_1 y_{t-1}^2\right)^{1/2}} \equiv y_t\left(\overline{\theta}_1\right)$$

Under H_1: $\theta_1 = \theta_1^1$,

$$v_t = \frac{y_t}{\left(1 + \theta_1^1 y_{t-1}^2\right)^{1/2}}$$

and

$$y_t\left(\overline{\theta}_1\right) = \frac{\left(1 + \theta_1^1 y_{t-1}^2\right)^{1/2}}{\left(1 + \overline{\theta}_1 y_{t-1}^2\right)^{1/2}} v_t = \overline{h}_t\left(\overline{\theta}_1, \theta_1^1\right)^{1/2} v_t$$

where:

$$\overline{h}_t\left(\overline{\theta}_1, \theta_1^1\right) = \frac{\left(1 + \theta_1^1 y_{t-1}^2\right)}{\left(1 + \overline{\theta}_1 y_{t-1}^2\right)} \equiv \overline{h}_t$$

Let

$$y_t\left(\overline{\theta}_1\right)=\left(y_1\left(\overline{\theta}_1\right),...,y_T\left(\overline{\theta}_1\right)\right)$$

Then

$$l_T\left(y\left(\overline{\theta}_1\right),\overline{\theta}_1\right)=-\frac{T}{2}\ln 2\pi-\frac{1}{2}\sum_{t=1}^{T}y_t\left(\overline{\theta}_1\right)^2$$

$$l_T\left(y\left(\overline{\theta}_1\right),\theta_1^1\right)=-\frac{T}{2}\ln 2\pi-\frac{1}{2}\sum_{t=1}^{T}\ln\left(\overline{h}_t\right)-\frac{1}{2}\sum_{t=1}^{T}\frac{y_t\left(\overline{\theta}_1\right)^2}{\overline{h}_t}$$

So:

$LR\left(\overline{\theta}_1,\theta_1^1\right)=-2\left[l_t\left(v\left(\overline{\theta}_1\right),\theta_1^1\right)-l_t\left(v\left(\overline{\theta}_1\right),\overline{\theta}_1\right)\right]=$
$-2\left[-\frac{1}{2}\sum_{t=1}^{T}\ln\left(\overline{h}_t\right)-\frac{1}{2}\sum_{t=1}^{T}\left(\frac{1}{\overline{h}_t}-1\right)y_t\left(\overline{\theta}_1\right)^2\right]=\sum_{t=1}^{T}\left[\ln\left(\overline{h}_t\right)+\left(\frac{1}{\overline{h}_t}-1\right)y_t\left(\overline{\theta}_1\right)^2\right]=$
$\sum_{t=1}^{T}\left[\ln\left(\overline{h}_t\right)+\left(\frac{1}{\overline{h}_t}-1\right)\frac{y_t^2}{\left(1+\overline{\theta}_1y_{t-1}^2\right)}\right]=\sum_{t=1}^{T}\left[\ln\left(\overline{h}_t\right)+\left(\frac{1}{1+\theta_1^1y_{t-1}^2}-\frac{1}{1+\overline{\theta}_1y_{t-1}^2}\right)y_t^2\right]=$
$\sum_{t=1}^{T}\left[\ln\left(\overline{h}_t\right)+\frac{\left(\overline{\theta}_1-\theta_1^1\right)y_{t-1}^2y_t^2}{\left(1+\theta_1^1y_{t-1}^2\right)\left(1+\overline{\theta}_1y_{t-1}^2\right)}\right]$

To summarize, it is necessary to consider:

$$\sum_{t=1}^{T}\ln\left(\overline{h}_t\right)+\sum_{t=1}^{T}\left(\frac{1}{\overline{h}_t}-1\right)\frac{y_t^2}{\left(1+\overline{\theta}_1y_{t-1}^2\right)}$$

where:

$$\overline{h}_t=\left(\frac{\left(1+\theta_1^1y_{t-1}^2\right)}{\left(1+\overline{\theta}_1y_{t-1}^2\right)}\right)$$

θ_1^1 =point where the test maximizes power.
$\overline{\theta}_1$ =the null hypothesis.

2.3 An alternative test

Another possible test is constructed by exploiting the use of pivotal properties of ARCH processes:

Theorem 3 *Suppose that* $y_t=x_t'\beta+\sum_{i=1}^{q}\phi_iy_{t-i}+\varepsilon_t$ *and* $h_t=E\left(\varepsilon_t^2/I_{t-1}\right)=\theta_0+\theta_1\varepsilon_{t-1}^2+\theta_2\varepsilon_{t-2}^2+...+\theta_p\varepsilon_{t-p}^2$*, then, a possible test for the null of* $H_0:\theta_1=\overline{\theta}_1,\theta_2=\overline{\theta}_2,...,\theta_p=\overline{\theta}_p$*, is given as an F-type test from:*

$$\frac{\varepsilon_t^2}{\left(1+\overline{\theta}_1\varepsilon_{t-1}^2+...+\overline{\theta}_p\varepsilon_{t-p}^2\right)}-1=\theta_1^*\frac{\varepsilon_{t-1}^2}{\left(1+\overline{\theta}_1\varepsilon_{t-1}^2+...+\overline{\theta}_p\varepsilon_{t-p}^2\right)}+...+\theta_p^*\frac{\varepsilon_{t-p}^2}{\left(1+\overline{\theta}_1\varepsilon_{t-1}^2+...+\overline{\theta}_p\varepsilon_{t-p}^2\right)}+w_t$$

This implies:

$$\varepsilon_t^2 - \left(1 + \overline{\theta}_1 \varepsilon_{t-1}^2 + ... + \overline{\theta}_p \varepsilon_{t-p}^2\right) = \theta_1^* \varepsilon_{t-1}^2 + ... + \theta_p^* \varepsilon_{t-p}^2 + w_t$$

where $\theta_i^* = \left(\theta_i - \overline{\theta}_i\right), \forall i = 1, .., p.$

In practice, the test implies the following. Take the residuals $\widehat{\varepsilon}_t^2$, and regress $\frac{\widehat{\varepsilon}_t^2}{\left(1+\overline{\theta}_1 \widehat{\varepsilon}_{t-1}^2 + ... + \overline{\theta}_p \widehat{\varepsilon}_{t-p}^2\right)}$ on $\frac{\widehat{\varepsilon}_{t-i}^2}{\left(1+\overline{\theta}_1 \widehat{\varepsilon}_{t-1}^2 + ... + \overline{\theta}_p \widehat{\varepsilon}_{t-p}^2\right)}$ for $i = 1, ..., p$. The critical values can be obtain from the asymptotic theory from an F-distribution or through the MC technique.

The previous test can be explained in detail now. Let us consider the expression:

$$v_t\left(\overline{\theta}_1\right) = \frac{\varepsilon_t}{\left(\theta_0 + \overline{\theta}_1 \varepsilon_{t-1}^2 + ... + \overline{\theta}_p \varepsilon_{t-p}^2\right)^{1/2}}$$

where, for the case of the ARCH(1):

$$y_t = \varepsilon_t$$

$$h_t = E\left(\varepsilon_t^2 / I_{t-1}\right) = \theta_0 + \theta_1 \varepsilon_{t-1}^2$$

$$E\left(\ v_t^2\left(\overline{\theta}_1\right)\right) = \frac{E\left(y_t / I_{t-1}\right)}{\left(1 + \overline{\theta}_1 y_{t-1}^2\right)} = \frac{\left(\theta_0 + \theta_1 y_{t-1}^2\right)}{\left(1 + \overline{\theta}_1 y_{t-1}^2\right)}$$

So:

$$E\left(\ v_t^2\left(\overline{\theta}_1\right) - \frac{1}{\left(1 + \overline{\theta}_1 y_{t-1}^2\right)} / I_{t-1}\right) = \frac{\theta_1 y_{t-1}^2}{\left(1 + \overline{\theta}_1 y_{t-1}^2\right)}$$

$$E\left(\ v_t^2\left(\overline{\theta}_1\right) - 1\right) = \left(\theta_1 - \overline{\theta}_1\right) \frac{y_{t-1}^2}{\left(1 + \overline{\theta}_1 y_{t-1}^2\right)}$$

So finally in this case we will regress:

$$\frac{y_t^2}{\left(1 + \overline{\theta}_1 y_{t-1}^2\right)} - 1 = \theta_1^* \frac{y_{t-1}^2}{\left(1 + \overline{\theta}_1 y_{t-1}^2\right)} + w_t$$

In the general case of an ARCH(p) model:

$$E\left(v_t^2\left(\overline{\theta}_1,...,\overline{\theta}_p\right)-1\right)=\left(\theta_1-\overline{\theta}_1\right)\frac{y_{t-1}^2}{\left(1+\overline{\theta}_1 y_{t-1}^2+...+\overline{\theta}_p y_{t-p}^2\right)}+...$$

$$+\left(\theta_p-\overline{\theta}_p\right)\frac{y_{t-p}^2}{\left(1+\overline{\theta}_1 y_{t-1}^2+...+\overline{\theta}_p y_{t-p}^2\right)}$$

So the testing would imply to regress:

$$\frac{y_t^2}{\left(1+\overline{\theta}_1 y_{t-1}^2+...+\overline{\theta}_p y_{t-p}^2\right)}-1=\theta_1^*\frac{y_{t-1}^2}{\left(1+\overline{\theta}_1 y_{t-1}^2+...+\overline{\theta}_p y_{t-p}^2\right)}+...+\theta_p^*\frac{y_{t-p}^2}{\left(1+\overline{\theta}_1 y_{t-1}^2+...+\overline{\theta}_p y_{t-p}^2\right)}+w_t$$

Finally, it is important to stress that we can apply these tests not only in the presence of pivotality when we test for the null of the full ARCH coefficients vector (using the MC technique) equal to a value; but also, using the Maximised Monte Carlo (MMC) technique, we can handle those cases where we loose the pivotal property when we test only for a sub-vector of those ARCH coefficients equal to a value (see Dufour (2004) for more details)). The MMC also allows for the presence of autoregressions in the mean equation.

3 Comparison by simulation of the proposed test procedures

We proceed now to compare our different procedures in the context of an ARCH(2) process under normal, t(5) and t(3) errors. This model was already analysed in Hall and Yao (2003) and it was chosen for comparative purposes. The model is then given by:

$y_t = \varepsilon_t$, where $E\left(\varepsilon_t^2/I_{t-1}\right)=\theta_1+\theta_2\varepsilon_{t-1}^2+\theta_3\varepsilon_{t-2}^2$. Assuming $\theta_1 = 0.81$.

We consider seven different types of null hypothesis. The results are given in the Appendix. We will present the results both using asymptotic theory and the MC technique (see Dufour and Kiviet (1996, 1998), Dufour, Khalaf, Bernard and Genest (2004) for more details). While in Hall and Yao (2003) they only reported results for sample sizes 500 and 1000, we will report results for sample sizes 500, 200 and 50, to show the good finite sample properties of our procedures. We also consider a much more variety of null and alternative hypothesis than those given in Hall and Yao (2003). In this last paper, they report that with their proposed subsampling technique, light tails in the distribution of the errors tend to produce relatively conservative confidence intervals. On the other hand, for extreme heavy-tailed-errors (t-3) the anticonservatism becames a problem. Our procedure through the MC technique allows to control for the size, and besides, we will provide clear results about the power properties.

The results for the Monte Carlo technique are based on 40000 replications, and N=99. The results for asymptotics where carried out with 40000 replications. The results for the simulated annealing (for the Maximised Monte Carlo technique) have been based on 1000 replications and N=99.

We cover 7 different types of nulls in the simulation results. All tests are very conservative in relation to the size distorsion.We proceed now to comment in detail results in each of the null hypothesis:

1) the first one, is the case of the IARCH(2) process. Here specially, the advantages of the point optimal test are huge. The point optimal test evaluated in the middle of the parameter space seems the best alternative (following Elliot, Rothenberg, and Stock (1996)). We also check that the procedure does not loose important power when the residuals are not gaussian. In this case, for alternatives with low values of the coefficients, to set the point optimal test to the middle value seems not to have good power properties. In this case, it would be necessary either to maximise power against a weighted average of possible alternatives (Andrews and Ploberger (1994)); or to use a small fraction of the sample (at the beginning) to estimate a plausible value for the alternative, and the rest to perform the test (taking as given the observations used to estimate the alternative). Given the Markovian structure of ARCH processes this will preserve the exactness of the test. Test 2 also has a good performance.

2) the second one is where θ_2=0.98 and θ_3=0.01; namely, it is similar to an ARCH(1) process (because θ_3 is very small), and there, any point optimal test seems to be doing it ok, what means that to set the point to optimize the test equal to (0.49, 0.49) seems to be a good suggestion (in an ARCH(1), then for nulls of the type near the unit root it would be a good idea to set the point optimal test to maximize in 0.49). Test 2 also provides a good performance.

3) The third case is when the null is quite close to that recommendation of setting the point to the middle of the parameter space. Here the recommendation to the researcher is to follow again the route of maximing power through a weighted average of possible alternatives (Andrews and Ploberger (1994)) or to use a small fraction of the sample to estimate a plausible value for the alternative. Test 2 has a bad performance power in this case.

4) The fourth null covers the case of very low values of θ_2 and θ_3, and there the best point to optimize the test is again the recommendation of Elliot, Rothenberg, and Stock (1996): (0.49, 0.49) for the ARCH(2) (and for the ARCH(1), it would be in 0.49).

5) The fifth case covers the case of testing for ARCH effects. Here, the Engle test has a much better performance than for the null of any other value. Even so, our point optimal test has superior properties to the Engle (1982) test, specially for low value alternatives and low degrees of freedom as well. Our point optimal test also has very good power when the distribution of the innovation process has very fat tails.

6) The sixth null relates to the case where we would be interested in testing for

subsets of the whole set of ARCH coefficients. we see how the MMC technique also makes our tests operational in case we want to test for a null hypothesis different from the one of the whole coefficient ARCH vector equal to a value. The MMC technique was carried out using the simulated annealing algorithm. The results show that even for sample sizes of 50, the point optimal test although it decreases the power for alternative hypothesis quite close to the null, still has good power, after optimising the p-value out of all possible values for the nuisance parameters.

7) In the seventh case, we consider a non-stationary case where we show the behaviour of our point optimal case against the asymptotic alternatives (in the context of Jensen and Rahbek (2004)). They have proved that in the setting of a non-stationary region, the QMLE is still asymptotically normal. However, there are still no results in the literature that show what happens in small samples. We provide results in this context both when the fourth moment of the innovation process exists and when it does not exist. We show that the behaviour of our point optimal test in this setting has very good power regardless if the errors have very fat tails or not (at the same time of controlling for the size). The extension of the Engle test and teh other test we propose have very low power in some of the alternatives. This possibility of very good behaviour of point optimal tests outside the stationarity region was already showed in Dufour and King (1991).

So, from the simulation results, we advice the practitioners always to use as a rule of thumb the value equal to the middle of the parameter space (e.g. 0.49 for an ARCH(1) process); and when it is possible, it is always better to maximise power against a weighted alternative as in Andrews and Ploberger (1994), or to split the sample size, and to use a small fraction of the sample (at the beginning) to estimate a plausible value for the alternative. We also show the behaviour of the point optimal test under non-normal-t(3) errors, and the findings indicate good power in these cases (allowing also for control of the size and then, improving on Hall and Yao (2003)).

4 Application of an ARCH model for US Inflation

In this section we re-visit the analysis of the Implicit price deflator for GNP done by Engle and Kraft (1983) and also reported by Bollerslev (1986). The series is also analysed in Greene (2000, page 809). The data corresponds to quaterly observations on the implicit price deflator for GNP from 1948.II to 1983.IV. We have obtained the data from the U.S. Department of Commerce: Bureau of Economic Analysis.

For this data, Engle and Kraft (1983) selected an AR(4)-ARCH(8) model such as:

$$\widehat{\pi}_t = 0.138 + 0.423\pi_{t-1} + 0.222\pi_{t-2} + 0.377\pi_{t-3} - 0.175\pi_{t-4}$$

$$(0.059)\ (0.081) \qquad (0.108) \qquad (0.078) \qquad (0.104)$$

$$\widehat{h}_t = 0.058 + 0.808 \sum_{i=1}^{8} \left(\frac{9-i}{36} \right) \varepsilon_{t-i}^2 \tag{5}$$

$$(0.033) \qquad\qquad (0.265)$$

where:

$$\pi_t = 100 \ln \frac{P_t}{P_{t-1}}$$

Asymptotic standard errors are given in brackets. Acording to the previous study, the values in that model decline linearly from 0.179 to 0.022. The normalised residuals from this model show no evidence of autocorrelation, nor do their squares.

We first carried out the analysis of the series for the same time period 1948.II to 1983.IV using an AR(4) model in the mean equation and using asymptotic Newey-West HAC (1987) standard errors. The results were:

$$\widehat{\pi}_t = 0.182 + 0.595\pi_{t-1} + 0.147\pi_{t-2} + 0.144\pi_{t-3} - 0.075\pi_{t-4}$$

$$(0.090)\ (0.106) \qquad (0.110) \qquad (0.131) \qquad (0.110)$$

The same that happens in the study of Engle and Kraft (1983), when we test for ARCH effects at lags 1, 4 and 8, we reject the null hypothesis both using the asymptotic LM test of Engle (1982) and when we apply our point optimal test (in all the cases giving very small p-values of 0.000). To use our point optimal test in this framework, we have applied the MMC technique where we have kept the AR coefficients in the mean equation as nuisance parameters, in order to obtain provably valid results. We have used the point to optimise equal to the middle of the parameter space. Then we proceed as the previous study to fit an unrestricted ARCH(8) model where we have used asymptotic Bollerslev-Wooldridge (1992) robust standard errors, and we obtain:

$$\widehat{\pi}_t = 0.136 + 0.544\pi_{t-1} + 0.171\pi_{t-2} + 0.158\pi_{t-3} - 0.011\pi_{t-4}$$

$$(0.040)\ (0.084) \qquad (0.094) \qquad (0.076) \qquad (0.081)$$

$$\widehat{h}_t = 0.064 + 0.166\varepsilon^2_{t-1} + 0.109\varepsilon^2_{t-2} - 0.021\varepsilon^2_{t-3} + 0.120\varepsilon^2_{t-4} + 0.033\varepsilon^2_{t-5}$$

$$(0.015) \quad (0.206) \quad (0.109) \qquad (0.095) \qquad (0.111) \quad (0.104)$$

$$+0.189\varepsilon^2_{t-6} + 0.051\varepsilon^2_{t-7} - 0.050\varepsilon^2_{t-8}$$

$$(0.114) \quad (0.123) \quad (0.048)$$

As Greene (2000) says, the linear restriction of the linear lag model given in (5) on the unrestricted ARCH(8) model appears not to be statistically significant. We indeed tested for the existence of the linear restriction in (5) to see if it holds in our model, and both using an asymptotic test and our point optimal test we reject the linear restriction with a p-value of 0.000 in both cases. However, one interesting puzzle is why although with the Engle-test and with our point optimal test we reject the null of no-ARCH effects of order 8, when we fit the previous model we obtain that all the coefficients in the ARCH(8) are not individually statistically significant according to the asymptotic results. We proceed then to apply our point optimal test using the MMC technique to test if the ARCH coefficients in the previous model where individually statistically significant, and we found in the 8 cases a p-value around 0.000 rejecting the null hypothesis that each of the coefficients are individually equal to zero. In this case, it is proved that asymptotic results give pretty bad inference for the individual statistically significance for this example.

5 Conclusions

In this paper we are interested in finding the best procedure for practitioners to test for any value of ARCH coefficients. We have developed three tests, including a point optimal test, and we have provided evidence of its good performance both in the case of normal errors, very fat tails (to compare with Hall and Yao (2003)), and/or in a non-stationary region (to cover as well the framework of Jensen and Rahbek (2004)). We have shown mainly that our point optimal test can have a good performance in all these cases. We have also shown that our tests can be made operational not only in the case where pivotality is guaranteed, but also when we test for the null of a sub-vector of the whole coefficient vector in the ARCH structure or when an AR process is introduced in the mean equation without loosing the exactness of the test. We have also reported the usefulness of our test in the case of an ARCH model applied to the US inflation.

In summary, a general rule to advice to practitioners (supported by the simulation results) is to use our point optimal test setting the point to optimise equal to the middle value of the parameter space (following the recommendation of Elliot, Rothenberg, and Stock, (1996)). This would be a good rule of thumb, although the best recommendation, if possible, is to maximise power against a weighted average of alternatives (Andrews and Ploberger (1994)) or to split the sample using a small fraction at the beginning to estimate a plausible value for the alternative. Our simulations show that our point optimal test has good properties in finite samples regardless if the errors are gaussian or not (including if the errors have very fat-tails) and/or if the process is in the stationarity region or not.

6 Appendix

Null 1: $H_0 : \theta_2 = 0.81; \theta_3 = 0.19$

	Engle-normal		Engle-t(5)		Engle-t(3)		T2-normal		T2-t(5)	
	AS	MC	AS	MC	AS	MC	AS	MC	AS	MC
Size: simulating under the null										
T=500	0.002	0.050	0.001	0.048		0.051	0.018	0.047	0.020	0.050
T=200	0.002	0.050	0.001	0.047		0.047	0.017	0.050	0.021	0.049
T=50	0.002	0.047	0.001	0.050		0.048	0.014	0.050	0.018	0.051
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.16$										
T=500	0.004	0.780	0.001	0.565		0.700	0.963	1.000	0.438	0.845
T=200	0.003	0.560	0.001	0.420		0.555	0.690	1.000	0.192	0.460
T=50	0.006	0.145	0.001	0.130		0.095	0.101	0.315	0.050	0.145
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.4$										
T=500	0.001	0.395	0.001	0.145		0.115	0.824	1.000	0.554	1.000
T=200	0.003	0.280	0.001	0.140		0.075	0.410	1.000	0.298	0.955
T=50	0.003	0.095	0.002	0.065		0.055	0.095	0.245	0.093	0.255
Power: simulating under $H_A : \theta_2 = 0.01; \theta_3 = 0.98$										
T=500	0.011	1.000	0.001	0.370		0.220	1.000	1.000	0.997	1.000
T=200	0.003	1.000	0.001	0.095		0.210	0.999	1.000	0.967	1.000
T=50	0.009	0.170	0.001	0.085		0.095	0.582	1.000	0.485	1.000

	T2-t(3)		PO (0.01,0.98)		PO (0.49,0.49)		PO (0.16,0.16)	
	AS	MC	MC-t(3)	MC-normal	MC-t(3)	MC-normal	MC-t(3)	MC-normal
Size: simulating under the null								
T=500		0.050	0.048	0.051	0.050	0.051	0.050	0.051
T=200		0.049	0.050	0.049	0.049	0.049	0.054	0.049
T=50		0.050	0.047	0.050	0.050	0.049	0.050	0.048
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.16$								
T=500		0.330	1.000	1.000	0.060	1.000	1.000	1.000
T=200		0.230	1.000	1.000	0.050	0.755	1.000	1.000
T=50		0.095	0.500	0.580	0.040	0.090	1.000	1.000
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.4$								
T=500		0.795	0.535	1.000	0.325	1.000	0.600	1.000
T=200		0.625	0.335	1.000	0.210	1.000	0.270	0.495
T=50		0.175	0.190	0.320	0.140	0.485	0.110	0.150
Power: simulating under $H_A : \theta_2 = 0.01; \theta_3 = 0.98$								
T=500		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=200		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=50		1.000	1.000	1.000	1.000	1.000	0.485	0.705

AS: Asymptotic. MC: Monte Carlo. PO(θ_2^1,θ_3^1): Point optimal. T2: Test in Theorem 3.

Null 2: $H_0 : \theta_2 = 0.98; \theta_3 = 0.01$

	Engle-normal		Engle-t(5)		Engle-t(3)		T2-normal		T2-t(5)	
	AS	MC	AS	MC	AS	MC	AS	MC	AS	MC
Size: simulating under the null										
T=500	0.003	0.054	0.001	0.047		0.049	0.023	0.049	0.022	0.049
T=200	0.001	0.054	0.001	0.049		0.055	0.026	0.052	0.024	0.053
T=50	0.001	0.053	0.001	0.054		0.055	0.025	0.055	0.027	0.050
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.16$										
T=500	0.456	1.000	0.089	1.000		1.000	0.999	1.000	0.885	1.000
T=200	0.171	1.000	0.021	1.000		1.000	0.909	1.000	0.625	1.000
T=50	0.012	0.395	0.003	0.465		0.410	0.261	0.560	0.232	0.540
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.4$										
T=500	0.029	0.255	0.001	0.110		0.085	0.999	1.000	0.988	1.000
T=200	0.008	0.195	0.001	0.145		0.050	0.905	1.000	0.861	1.000
T=50	0.002	0.110	0.001	0.100		0.070	0.395	1.000	0.431	1.000
Power: simulating under $H_A : \theta_2 = 0.01; \theta_3 = 0.98$										
T=500	0.003	0.135	0.001	0.135		0.150	1.000	1.000	0.999	1.000
T=200	0.005	0.125	0.001	0.085		0.095	0.995	1.000	0.977	1.000
T=50	0.002	0.130	0.001	0.110		0.145	0.662	1.000	0.613	1.000

	T2-t(3)		PO (0.01,0.98)		PO (0.49,0.49)		PO (0.16,0.16)	
	AS	MC	MC-t(3)	MC-normal	MC-t(3)	MC-normal	MC-t(3)	MC-normal
Size: simulating under the null								
T=500		0.051	0.055	0.049	0.051	0.050	0.051	0.051
T=200		0.051	0.047	0.050	0.049	0.049	0.048	0.049
T=50		0.049	0.051	0.050	0.051	0.048	0.051	0.052
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.16$								
T=500		1.000	1.000	1.000	0.860	1.000	1.000	1.000
T=200		1.000	1.000	1.000	0.440	1.000	1.000	1.000
T=50		0.545	0.765	1.000	0.250	0.785	1.000	1.000
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.4$								
T=500		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=200		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=50		0.980	0.930	1.000	1.000	1.000	1.000	1.000
Power: simulating under $H_A : \theta_2 = 0.01; \theta_3 = 0.98$								
T=500		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=200		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=50		1.000	1.000	1.000	1.000	1.000	1.000	1.000

Null 3: $H_0 : \theta_2 = 0.5; \theta_3 = 0.4$

	Engle-normal		Engle-t(5)		Engle-t(3)		T2-normal		T2-t(5)	
	AS	MC	AS	MC	AS	MC	AS	MC	AS	MC
Size: simulating under the null										
T=500	0.078	0.050	0.001	0.051		0.052	0.016	0.044	0.019	0.047
T=200	0.035	0.047	0.001	0.055		0.050	0.016	0.051	0.017	0.053
T=50	0.013	0.048	0.002	0.052		0.051	0.012	0.049	0.012	0.047
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.16$										
T=500	0.607	1.000	0.137	1.000		1.000	0.784	1.000	0.015	0.045
T=200	0.365	0.895	0.049	1.000		1.000	0.351	1.000	0.010	0.035
T=50	0.070	0.310	0.001	0.780		0.305	0.032	0.135	0.008	0.040
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.4$										
T=500	0.393	0.480	0.023	0.600		0.570	0.762	1.000	0.214	0.400
T=200	0.177	0.410	0.012	0.545		0.450	0.293	1.000	0.101	0.320
T=50	0.041	0.165	0.007	0.145		0.175	0.040	0.135	0.028	0.105
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.64$										
T=500	0.094	0.040	0.001	0.110		0.085	0.981	1.000	0.804	1.000
T=200	0.041	0.085	0.001	0.105		0.115	0.656	1.000	0.654	1.000
T=50	0.019	0.045	0.001	0.080		0.095	0.118	0.400	0.093	0.435
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.16$										
T=500	0.237	0.215	0.004	0.240		0.210	0.364	0.515	0.071	0.255
T=200	0.037	0.135	0.003	0.225		0.185	0.094	0.305	0.040	0.150
T=50	0.024	0.209	0.004	0.095		0.100	0.015	0.060	0.011	0.040
Power: simulating under $H_A : \theta_2 = 0.9; \theta_3 = 0.09$										
T=500	0.005	0.015	0.001	0.015		0.025	0.995	1.000	0.824	1.000
T=200	0.004	0.010	0.001	0.005		0.035	0.711	1.000	0.449	1.000
T=50	0.005	0.020	0.001	0.025		0.060	0.092	0.315	0.061	0.220
Power: simulating under $H_A : \theta_2 = 0.09; \theta_3 = 0.9$										
T=500	0.007	0.020	0.001	0.010		0.015	1.000	1.000	0.984	1.000
T=200	0.007	0.005	0.001	0.010		0.030	0.964	1.000	0.843	1.000
T=50	0.008	0.040	0.002	0.030		0.060	0.324	1.000	0.241	0.860

	T2-t(3)		PO (0.01,0.98)		PO (0.49,0.49)		PO (0.16,0.16)	
Size: simulating under the null								
	AS	MC	MC-t(3)	MC-normal	MC-t(3)	MC-normal	MC-t(3)	MC-normal
T=500		0.048	0.051	0.050	0.051	0.051	0.050	0.049
T=200		0.048	0.051	0.048	0.050	0.051	0.051	0.048
T=50		0.053	0.046	0.051	0.047	0.051	0.048	0.052
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.16$								
T=500		0.015	1.000	1.000	0.015	0.010	1.000	1.000
T=200		0.010	1.000	1.000	0.010	0.006	1.000	1.000
T=50		0.020	0.165	0.070	0.010	0.005	1.000	0.770
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.4$								
T=500		0.255	1.000	1.000	0.030	0.025	1.000	1.000
T=200		0.210	1.000	1.000	0.035	0.015	0.710	1.000
T=50		0.095	0.250	0.275	0.030	0.015	0.375	0.310
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.64$								
T=500		1.000	1.000	1.000	0.175	1.000	0.250	0.315
T=200		1.000	1.000	1.000	0.150	1.000	0.160	0.175
T=50		0.260	0.370	0.815	0.120	0.380	0.100	0.085
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.16$								
T=500		0.160	0.380	0.070	0.015	0.010	1.000	1.000
T=200		0.095	0.145	0.030	0.015	0.005	0.690	0.555
T=50		0.040	0.035	0.015	0.020	0.001	0.185	0.130
Power: simulating under $H_A : \theta_2 = 0.9; \theta_3 = 0.09$								
T=500		1.000	0.010	0.005	0.025	0.015	0.010	0.025
T=200		0.970	0.010	0.002	0.025	0.011	0.015	0.022
T=50		0.150	0.010	0.001	0.025	0.010	0.030	0.020
Power: simulating under $H_A : \theta_2 = 0.09; \theta_3 = 0.9$								
T=500		1.000	1.000	1.000	0.960	1.000	0.010	0.040
T=200		1.000	1.000	1.000	0.585	1.000	0.015	0.036
T=50		0.790	0.860	1.000	0.290	1.000	0.040	0.035

Null 4: $H_0 : \theta_2 = 0.16; \theta_3 = 0.25$

	Engle-normal		Engle-t(5)		Engle-t(3)		T2-normal		T2-t(5)	
	AS	MC	AS	MC	AS	MC	AS	MC	AS	MC
Size: simulating under the null										
T=500	0.077	0.049	0.016	0.051		0.045	0.017	0.049	0.022	0.047
T=200	0.002	0.048	0.001	0.049		0.056	0.018	0.047	0.028	0.051
T=50	0.003	0.046	0.001	0.053		0.045	0.014	0.041	0.019	0.052
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.16$										
T=500	0.019	0.020	0.001	0.005		0.005	0.919	1.000	0.685	1.000
T=200	0.001	0.005	0.001	0.005		0.020	0.526	1.000	0.399	1.000
T=50	0.001	0.020	0.001	0.025		0.015	0.109	0.380	0.102	0.355
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.4$										
T=500	0.003	0.020	0.001	0.001		0.001	0.863	1.000	0.602	1.000
T=200	0.001	0.015	0.001	0.001		0.001	0.466	1.000	0.313	0.910
T=50	0.001	0.030	0.001	0.001		0.001	0.090	0.220	0.076	0.195
Power: simulating under $H_A : \theta_2 = 0.9; \theta_3 = 0.09$										
T=500	0.001	0.001	0.001	0.001		0.001	1.000	1.000	0.985	1.000
T=200	0.001	0.001	0.001	0.001		0.001	0.970	1.000	0.846	1.000
T=50	0.001	0.005	0.001	0.020		0.010	0.361	1.000	0.261	0.690
Power: simulating under $H_A : \theta_2 = 0.09; \theta_3 = 0.9$										
T=500	0.005	0.015	0.001	0.001		0.001	0.999	1.000	0.965	1.000
T=200	0.001	0.001	0.001	0.005		0.001	0.912	1.000	0.729	1.000
T=50	0.001	0.005	0.001	0.020		0.010	0.254	0.995	0.172	0.400

	T2-t(3)		PO (0.01,0.98)		PO (0.49,0.49)		PO (0.16,0.16)	
	AS	MC	MC-t(3)	MC-normal	MC-t(3)	MC-normal	MC-t(3)	MC-normal
Size: simulating under the null								
T=500		0.045	0.051	0.051	0.051	0.051	0.051	0.049
T=200		0.051	0.046	0.050	0.051	0.050	0.048	0.049
T=50		0.045	0.051	0.050	0.048	0.049	0.053	0.051
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.16$								
T=500		0.975	0.035	0.030	0.375	1.000	0.010	0.060
T=200		0.630	0.035	0.025	0.320	1.000	0.015	0.055
T=50		0.180	0.045	0.005	0.250	0.950	0.025	0.050
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.4$								
T=500		0.495	0.045	0.525	1.000	1.000	0.003	0.015
T=200		0.345	0.065	0.480	0.915	1.000	0.001	0.011
T=50		0.185	0.080	0.335	0.425	1.000	0.004	0.010
Power: simulating under $H_A : \theta_2 = 0.9; \theta_3 = 0.09$								
T=500		1.000	0.030	0.350	1.000	1.000	0.010	0.035
T=200		1.000	0.045	0.325	1.000	1.000	0.005	0.030
T=50		0.465	0.070	0.320	0.870	1.000	0.005	0.025
Power: simulating under $H_A : \theta_2 = 0.09; \theta_3 = 0.9$								
T=500		1.000	1.000	1.000	1.000	1.000	0.002	0.010
T=200		1.000	1.000	1.000	1.000	1.000	0.001	0.005
T=50		0.370	0.645	1.000	0.460	1.000	0.001	0.002

Null 5: $H_0 : \theta_2 = 0; \theta_3 = 0$

	Engle-normal		Engle-t(5)		Engle-t(3)		T2-normal		T2-t(5)	
	AS	MC	AS	MC	AS	MC	AS	MC	AS	MC
Size: simulating under the null										
T=500	0.045	0.052	0.045	0.048		0.047	0.023	0.049	0.016	0.046
T=200	0.043	0.051	0.041	0.053		0.050	0.023	0.048	0.028	0.048
T=50	0.039	0.050	0.033	0.048		0.050	0.019	0.047	0.025	0.050
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.16$										
T=500	0.957	1.000	0.984	1.000		1.000	0.923	1.000	0.873	1.000
T=200	0.654	1.000	0.815	1.000		1.000	0.620	1.000	0.650	1.000
T=50	0.191	0.375	0.288	0.665		0.960	0.170	0.575	0.241	0.700
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.4$										
T=500	1.000	1.000	0.995	1.000		1.000	0.999	1.000	0.900	1.000
T=200	0.991	1.000	0.976	1.000		1.000	0.961	1.000	0.834	1.000
T=50	0.528	1.000	0.594	1.000		1.000	0.541	1.000	0.488	1.000
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.16$										
T=500	1.000	1.000	0.995	1.000		1.000	0.998	1.000	0.980	1.000
T=200	0.698	1.000	0.962	1.000		1.000	0.926	1.000	0.886	1.000
T=50	0.433	1.000	0.512	1.000		1.000	0.387	1.000	0.420	1.000
Power: simulating under $H_A : \theta_2 = 0.09; \theta_3 = 0.9$										
T=500	0.999	1.000	0.976	1.000		1.000	0.999	1.000	0.976	1.000
T=200	0.984	1.000	0.949	1.000		1.000	0.509	1.000	0.888	1.000
T=50	0.568	1.000	0.543	1.000		1.000	0.468	1.000	0.443	1.000
Power: simulating under $H_A : \theta_2 = 0.9; \theta_3 = 0.09$										
T=500	1.000	1.000	0.990	1.000		1.000	0.998	1.000	0.981	1.000
T=200	0.990	1.000	0.961	1.000		1.000	0.957	1.000	0.906	1.000
T=50	0.558	1.000	0.580	1.000		1.000	0.490	1.000	0.481	1.000

	T2-t(3)		PO (0.01,0.98)		PO (0.49,0.49)		PO (0.16,0.16)	
	AS	MC	MC-t(3)	MC-normal	MC-t(3)	MC-normal	MC-t(3)	MC-normal
Size: simulating under the null								
T=500		0.051	0.051	0.050	0.051	0.051	0.052	0.049
T=200		0.046	0.051	0.053	0.053	0.050	0.050	0.048
T=50		0.055	0.046	0.048	0.051	0.049	0.051	0.051
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.16$								
T=500		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=200		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=50		0.790	1.000	1.000	1.000	1.000	1.000	1.000
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.4$								
T=500		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=200		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=50		1.000	1.000	1.000	1.000	1.000	1.000	1.000
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.16$								
T=500		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=200		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=50		1.000	1.000	1.000	1.000	1.000	1.000	1.000
Power: simulating under $H_A : \theta_2 = 0.09; \theta_3 = 0.9$								
T=500		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=200		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=50		1.000	1.000	1.000	1.000	1.000	1.000	1.000
Power: simulating under $H_A : \theta_2 = 0.9; \theta_3 = 0.09$								
T=500		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=200		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=50		1.000	1.000	1.000	1.000	1.000	1.000	1.000

Null 6: $H_0 : \theta_2 = 0$
Point optimal test set to (0.49,0.49) using MMC with t(3) and normal.

	Size		Power $\theta_2 = 0.9$		Power $\theta_2 = 0.5$		Power $\theta_2 = 0.16$	
	MMC-t(3)	MMC-nor	MMC-t(3)	MMC-nor	MMC-t(3)	MMC-nor	MMC-t(3)	MMC-nor
T=50	0.059	0.060	0.910	0.940	0.820	0.840	0.340	0.400

Null 7: $H_0 : \theta_2 = 0.81; \theta_3 = 0.49$

	Engle-normal		Engle-t(5)		Engle-t(3)		T2-normal		T2-t(5)	
	AS	MC	AS	MC	AS	MC	AS	MC	AS	MC
Size: simulating under the null										
T=500	0.003	0.049	0.002	0.051		0.051	0.020	0.050	0.017	0.052
T=200	0.002	0.049	0.001	0.052		0.052	0.016	0.049	0.015	0.053
T=50	0.002	0.051	0.001	0.051		0.048	0.014	0.052	0.013	0.049
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.16$										
T=500	0.007	1.000	0.002	1.000		1.000	0.956	1.000	0.505	0.950
T=200	0.007	0.720	0.002	0.695		1.000	0.723	1.000	0.137	0.535
T=50	0.009	0.180	0.002	0.185		0.225	0.111	0.390	0.018	0.065
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.4$										
T=500	0.003	0.505	0.001	0.180		0.655	0.363	0.565	0.018	0.070
T=200	0.001	0.385	0.001	0.150		0.400	0.108	0.355	0.016	0.050
T=50	0.005	0.120	0.001	0.080		0.115	0.022	0.110	0.012	0.040
Power: simulating under $H_A : \theta_2 = 0.01; \theta_3 = 0.98$										
T=500	0.002	0.600	0.001	0.430		0.975	1.000	1.000	0.989	1.000
T=200	0.001	0.360	0.001	0.360		0.695	0.992	1.000	0.910	1.000
T=50	0.004	0.090	0.001	0.100		0.125	0.417	1.000	0.311	1.000
Power: simulating under $H_A : \theta_2 = 0.49; \theta_3 = 0.81$										
T=500	0.001	0.080	0.002	0.045		0.040	0.927	1.000	0.724	1.000
T=200	0.001	0.090	0.001	0.075		0.050	0.513	1.000	0.348	1.000
T=50	0.001	0.060	0.001	0.040		0.085	0.099	0.455	0.077	0.265
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 1$										
T=500	0.003	0.160	0.002	0.175		0.275	1.000	1.000	0.984	1.000
T=200	0.003	0.125	0.001	0.185		0.240	0.970	1.000	0.832	1.000
T=50	0.003	0.085	0.001	0.095		0.100	0.323	1.000	0.231	0.840

	T2-t(3)		PO (0.01,0.98)		PO (0.49,0.49)		PO (0.16,0.16)	
	AS	MC	MC-t(3)	MC-normal	MC-t(3)	MC-normal	MC-t(3)	MC-normal
Size: simulating under the null								
T=500		0.050	0.052	0.049	0.051	0.053	0.049	0.047
T=200		0.051	0.053	0.052	0.049	0.052	0.049	0.052
T=50		0.049	0.048	0.051	0.048	0.050	0.052	0.049
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 0.16$								
T=500		0.050	1.000	1.000	1.000	1.000	1.000	1.000
T=200		0.075	1.000	1.000	1.000	1.000	1.000	1.000
T=50		0.035	1.000	1.000	1.000	1.000	1.000	1.000
Power: simulating under $H_A : \theta_2 = 0.5; \theta_3 = 0.4$								
T=500		0.055	1.000	1.000	1.000	1.000	1.000	1.000
T=200		0.075	1.000	1.000	1.000	1.000	1.000	1.000
T=50		0.040	0.280	0.220	0.370	0.285	0.390	0.355
Power: simulating under $H_A : \theta_2 = 0.01; \theta_3 = 0.98$								
T=500		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=200		1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=50		0.920	1.000	1.000	1.000	1.000	0.556	0.390
Power: simulating under $H_A : \theta_2 = 0.49; \theta_3 = 0.81$								
T=500		1.000	0.290	0.860	0.540	0.500	0.055	0.050
T=200		0.725	0.165	0.315	0.255	0.515	0.055	0.065
T=50		0.300	0.135	0.220	0.115	0.190	0.055	0.060
Power: simulating under $H_A : \theta_2 = 0.16; \theta_3 = 1$								
T=500		1.000	1.000	1.000	1.000	1.000	0.825	0.800
T=200		1.000	1.000	1.000	1.000	1.000	0.425	0.415
T=50		0.625	0.905	1.000	0.705	1.000	0.175	0.155

References

Andrews, D.W.K. and W. Ploberger (1994), Optimal Tests when a Nuisance Parameter is Present only under the Alternative, *Econometrica* 62, 6, 1383-1414.

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

Bollerslev, T. and J. M. Wooldridge (1992), Quasi Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances. *Econometric Reviews* 11, 143-172.

Dufour, J. M. (2004), Monte Carlo Tests with Nuisance Parameters: A General Approach to Finite-Sample Inference and Nonstandard Asymptotics in Econometrics, Mimeo.

Dufour, J. M., L. Khalaf, J.-T.- Bernard and I. Genest (2004), Simulation-Based Finite-Sample Tests for Heteroscedasticity and ARCH Effects, *Journal of Econometrics*, forthcoming.

Dufour, J. M. and M. King (1991), Optimal Invariant Tests for the Autocorrelation Coefficient in Linear Regressions with Stationary or Nonstationary AR(1) Errors, *Journal of Econometrics* 47, 115-143.

Dufour, J. M. and J.-F. Kiviet (1996), Exact Tests for Structural Change in First Order Dynamic Models, *Journal of Econometrics* 70, 39-68.

Dufour, J. M. and J.-F. Kiviet (1998), Exact Inference Methods for First-Order Autoregressive Distributed Lag Models, *Econometrica* 66, 79-104.

Elliot, G., T.J. Rothenberg, and J.H. Stock, (1996), Efficient Tests for an Autoregressive Unit Root, *Econometrica*, 64, 813-836.

Engle, R. F. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica* 50, 987-1007.

Engle, R. F. and D. Kraft (1983), Multiperiod Forecast Error Variances of Inflation Estimated from ARCH Models. In A. Zellner, ed., *Applied Time Series Analysis of Economic Data*. Washington D.C.: Bureau of the Census.

Greene, W. H. (2000), *Econometric Analysis*, fourth edition, Prentice-Hall.

Hall, P. and Q. Yao (2003), Inference in ARCH and GARCH Models with Heavy-Tailed Errors, *Econometrica* 71, 285-317.

Jensen, S. T. and A. Rahbek (2004), Asymptotic Normality of hte QMLE of ARCH in the Nonstationary Case, *Econometrica* 72, 2, 641-646.

Lee, J. H. H. (1991), A Lagrange Multiplier Test for GARCH Models, *Economics Letters* 37, 265-271.

Lee, J. H. and M. L. King (1993), A Locally Most Mean Powerful Based Score Test for ARCH and GARCH Regression Disturbances, *Journal of Business and Economic Statistics* 11, 17-27. Correction 12 (1994), 139.

Lumsdaine, R. (1995), Finite-Sample Properties of the Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models. A Monte Carlo Investigation, *Journal of Business and Economic Statistics* 13, 1-10.

Newey, W. K. and K. D. West (1987), A Simple, Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703-708.