Evaluating GARCH models

Stefan Lundbergh and Timo Teräsvirta

Department of Economic Statistics,

Stockholm School of Economics,

P.O.Box 6501, SE-113 83 Stockholm, Sweden

February 5, 1999

Abstract

This paper suggests a unified framework for testing the adequacy of an estimated GARCH model. Nothing more complicated than standard asymptotic theory is required. Parametric tests of no ARCH in standardized errors, symmetry, and parameter constancy are suggested. Estimating the alternative when the null hypothesis is rejected may give useful ideas of how to improve the specification. It is also shown that the recent portmanteau test of Li and Mak (1994) is asymptotically equivalent to our test of no ARCH in the standardized error process.

Keywords: Conditional heteroskedasticity, model misspecification test, nonlinear time series, parameter constancy, smooth transition GARCH.

JEL Classification Code: C22, C52.

Acknowledgements: The first author acknowledges financial support from the Tore Browaldh's Foundation. The research of the second author has been supported by the Swedish Council for Research in Humanities and Social Sciences. He would also like to thank the Tinbergen Institute for hospitality during his visit there. Material from the paper has been presented at the EC^2 Meeting, Stockholm, December 1998. We wish to thank Changli He for useful comments and James Chu for sending us his GAUSS code for performing his parameter constancy test. Any errors and shortcomings in the paper remain our own responsibility.

Note: This paper is also available as Working Paper No. 292 in the SSE/EFI Working Paper Series in Economics and Finance. Stockholm School of Economics.

1. Introduction

When modelling the conditional mean, at least when it is a linear function of parameters, the estimated model is regularly subjected to a battery of misspecification tests to check its adequacy. The hypothesis of no (conditional) heteroskedasticity, no error autocorrelation, linearity, and parameter constancy, to name a few, are tested using various methods. In models of conditional variance, such as the popular GARCH model, testing the adequacy of the estimated model has been much less common in practice. But then, misspecification tests do exist in the literature also for GARCH models. For example, Bollerslev (1986) already suggested a score or Lagrange Multiplier (LM) test for testing a GARCH model of a given order against a higher-order GARCH model. Li and Mak (1994) derived a portmanteau type test for testing the adequacy of a GARCH model. Engle and Ng (1993) considered testing the GARCH specification against asymmetry using the so-called sign-bias test. Chu (1995) derived a test of parameter constancy against a single structural break. This test has a nonstandard asymptotic null distribution, but Chu provided tables for critical values.

In this paper we provide a unified framework for misspecification testing in GARCH models. The framework covers the most common alternative hypotheses. The idea is to make misspecification testing easy without sacrificing power. We suggest tests for testing the null of no ARCH in the standardized errors, a general test for misspecification of the functional form, testing symmetry against a smooth transition GARCH, and a test of parameter constancy against smoothly changing parameters. A single structural break is nested in the alternative hypothesis of the parameter constancy test. A two-regime asymmetric GARCH model such as the so-called GJR model (Glosten, Jagannathan and Runkle, 1993) is nested in the alternative of smooth transition GARCH. Note that the test of Bollerslev (1986) fits well into our framework. Furthermore, we show that the portmanteau test of Li and Mak (1994) is asymptotically equivalent to our test of no remaining ARCH. All our tests are LM-tests and require only standard asymptotic distribution theory. They may be obtained from the same "root" by merely changing the definitions of the elements of the score vector corresponding to the alternative hypothesis. This makes testing easy as the sample counterparts of the analytical first and second order derivatives of the logarithmic likelihood function may be computed without difficulty using the results in Fiorentini, Calzolari and Panattoni (1996). Our Monte Carlo simulations show that the tests we propose have reasonable power, that is, they compare favourably with the tests currently available in the literature.

The plan of the paper is as follows. In section 2 we define the model. In section 3 we discuss testing the null of no ARCH in the standardized errors and compare our LM-test with the portmanteau test of Li and Mak (1994). Section 4 considers testing the functional form, symmetry and parameter constancy. Section 5 contains results of a simulation experiment in which our tests are compared with other tests proposed in the literature and Section 6 concludes.

2. The model

Consider a conditionally heteroskedastic model where the conditional mean has the following structure:

$$y_t = f(\mathbf{w}_t; \boldsymbol{\varphi}) + u_t \tag{2.1}$$

where f at least is twice continuously differentiable with respect to $\boldsymbol{\varphi} \in \Phi$, for all $\mathbf{w}_t \in \Re^k$ everywhere in Φ. The conditional variance is parameterized as:

$$u_t = \xi_t \sqrt{h(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta})} \tag{2.2}$$

where $\{\xi_t\}$ is a sequence of independent standard normal variables. The normality assumption is made just for the purpose of defining the likelihood function but is not needed for the asymptotic results. Existence of a number of moments has to be assumed, however, for each of the cases considered below. We assume that $h_t = \boldsymbol{\eta}'\mathbf{z}_t$, that is, a linear function of the parameters $\boldsymbol{\eta}$. The standard GARCH(p,q) model where $\boldsymbol{\eta} = (\alpha_0, \alpha_1, ..., \alpha_q, \beta_1, ..., \beta_p)'$ and the observation vector $\mathbf{z}_t = (1, u_{t-1}^2, ..., u_{t-q}^2, h_{t-1}, ..., h_{t-p})'$ constitutes an example. Furthermore, $u_t = y_t - f(\mathbf{w}_t; \boldsymbol{\varphi})$ and $\boldsymbol{\varphi}$ is assumed not to depend on $\boldsymbol{\eta}$. This guarantees that $E[u_t] = 0$ and $E[u_t u_{t-j}] = 0$, $j \neq 0$. We assume that the maximum likelihood estimators of the parameters of the GARCH process are consistent and asymptotically normal. Lumsdaine (1996) gave the required assumptions for the GARCH(1,1) process. For more general GARCH processes, see Bollerslev and Wooldridge (1992). Since the standard GARCH(p,q) model is symmetric and satisfies the usual regularity conditions, see Engle (1982), the information matrix is block-diagonal in $\boldsymbol{\varphi}$ and $\boldsymbol{\eta}$. The restrictions $\alpha_0 > 0, \alpha_i \geqslant 0, i = 1, ..., q-1, \alpha_q > 0, \beta_i \geqslant 0,$ ensure nonnegative conditional variance with probability one but they can be relaxed as in Nelson and Cao (1992); see also He and Teräsvirta (1999c). In what follows we shall mainly focus on the conditional variance and do not consider the conditional mean. This we do for simplicity, and in cases where the information matrix of the log-likelihood is block diagonal this approach is justified. Joint modelling of the conditional mean and the conditional variance is discussed, for example, in Lee and Li (1998) and Lundbergh and Teräsvirta (1998).

3. Testing the null of no ARCH in standardized errors

3.1. LM-Test

We consider the situation where we have estimated a GARCH(p,q) model under the assumption that the standardized errors $\xi_t = u_t h_t^{-1/2}$ of this model are independent normal. We want to test this hypothesis against the alternative that these errors follow an ARCH(m) process. Consider (2.2) but assume that

$$\xi_t = \varepsilon_t \sqrt{g(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi})} \tag{3.1}$$

where $\{\varepsilon_t\}$ is a sequence of independent standard normal variables. The alternative of higher-order dependence in (3.1) is parametrized as $g_t = 1 + \boldsymbol{\pi}'\mathbf{v}_t$ where $\boldsymbol{\pi} = (\pi_1, ..., \pi_m)'$ and $\mathbf{v}_t = (\xi_{t-1}^2, ..., \xi_{t-m}^2)'$ so that $E[\xi_t^2 \xi_{t-j}^2] \neq 0$, $j \neq 0$. This implies that $\{\xi_t\}$ follows an ARCH(m) model. The null hypothesis of no ARCH in $\{\xi_t\}$ is $\mathrm{H}_0 : \boldsymbol{\pi} = \mathbf{0}$ which in turn is equivalent to $g_t \equiv 1$. Under the alternative $\boldsymbol{\pi} \neq \mathbf{0}$ the standard GARCH(p,q) model is misspecified because $\{\xi_t\}$ is no longer a sequence of independent variables. For simplicity, rewrite (2.2) as

$$u_t = \varepsilon_t \sqrt{h(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi})} \tag{3.2}$$

where $h(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) = (\boldsymbol{\eta}'\mathbf{z}_t) g(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi})$. Let $\boldsymbol{\omega} = (\boldsymbol{\varphi}, \boldsymbol{\eta}')'$ denote the parameters of the standard GARCH model with the conditional mean specified according to (2.1). In that case, $h(\mathbf{z}_t; \boldsymbol{\omega}, \mathbf{a}) = (\boldsymbol{\eta}'\mathbf{z}_t)(1 + \boldsymbol{\pi}'\mathbf{v}_t)$. The Lagrange multiplier (or score) test statistic is defined as

$$\mathrm{LM}_{\boldsymbol{\pi}} = T \begin{pmatrix} 0 \\ \frac{1}{T}\sum_{t=1}^{T} \frac{\partial l_t}{\partial \boldsymbol{\pi}}|_{\boldsymbol{\pi}=\mathbf{0}} \end{pmatrix}' I(\widehat{\boldsymbol{\omega}}, \boldsymbol{\pi}|_{\boldsymbol{\pi}=\mathbf{0}})^{-1} \begin{pmatrix} 0 \\ \frac{1}{T}\sum_{t=1}^{T} \frac{\partial l_t}{\partial \boldsymbol{\pi}}|_{\boldsymbol{\pi}=\mathbf{0}} \end{pmatrix} \tag{3.3}$$

where T is the number of observations. The information matrix $I(\widehat{\boldsymbol{\omega}}, \boldsymbol{\pi}|_{\boldsymbol{\pi}=\mathbf{0}})$ is estimated by the estimated negative expectation of the Hessian. Estimation of the GARCH model using analytical derivatives, Fiorentini et al. (1996), yields as a by-product numerically reliable estimates for the elements of the inverse of the information matrix and can therefore be recommended.

The first derivative of the log-likelihood of observation t with respect to $\boldsymbol{\pi}$ evaluated under H_0 has the form

$$\begin{aligned} \frac{\partial l_t}{\partial \boldsymbol{\pi}'}|_{\boldsymbol{\pi}=\mathbf{0}} &= \frac{1}{2}\left(\frac{u_t^2}{h_t} - 1\right)\frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\pi}'}|_{\boldsymbol{\pi}=0} \\ &= \frac{1}{2}\left(\frac{u_t^2}{h_t} - 1\right)\mathbf{v}_t \end{aligned}$$

where $\mathbf{v}_t = (\xi_{t-1}^2, ..., \xi_{t-m}^2)' = (\frac{u_{t-1}^2}{h_{t-1}}, ..., \frac{u_{t-m}^2}{h_{t-m}})' = \frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\pi}'}|_{\boldsymbol{\pi}=0}$. Under the null hypothesis, the information matrix is block-diagonal in $\boldsymbol{\varphi}$ and $\boldsymbol{\eta}$. The negative (conditional) expectations of the relevant second order derivatives are

$$-\mathsf{E}\frac{\partial^2 l_t}{\partial \boldsymbol{\eta}\partial \boldsymbol{\eta}'}|_{\boldsymbol{\pi}=0} = \frac{1}{2}\mathbf{x}_t\mathbf{x}_t' \quad -\mathsf{E}\frac{\partial^2 l_t}{\partial \boldsymbol{\pi}\partial \boldsymbol{\eta}'}|_{\boldsymbol{\pi}=0} = \frac{1}{2}\mathbf{v}_t\mathbf{x}_t' \quad -\mathsf{E}\frac{\partial^2 l_t}{\partial \boldsymbol{\pi}\partial \boldsymbol{\pi}'}|_{\boldsymbol{\pi}=0} = \frac{1}{2}\mathbf{v}_t\mathbf{v}_t'$$

where $\mathbf{x}_t = \frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\eta}'}|_{\boldsymbol{\pi}=0}$. The test statistic (3.3) may then be written as

$$\mathrm{LM}_{\boldsymbol{\pi}} = \frac{1}{4T}\sum\left((\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)\widehat{\mathbf{v}}_t'\right) V_{\mathrm{LM}}(\widehat{\boldsymbol{\eta}})^{-1}\sum\left(\widehat{\mathbf{v}}_t(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)\right) \tag{3.4}$$

where $V_{\mathrm{LM}}(\widehat{\boldsymbol{\eta}}) = \frac{1}{2T}\left(\sum \widehat{\mathbf{v}}_t\widehat{\mathbf{v}}_t' - \sum \widehat{\mathbf{v}}_t\widehat{\mathbf{x}}_t'(\sum \widehat{\mathbf{x}}_t\widehat{\mathbf{x}}_t')^{-1}\sum \widehat{\mathbf{x}}_t\widehat{\mathbf{v}}_t'\right)$. Furthermore, $\widehat{\mathbf{v}}_t$ and $\widehat{\mathbf{x}}_t$ are the sample counterparts of the corresponding derivatives under the null hypothesis and they may be computed iteratively; see Fiorentini et al. (1996) for details. Assuming that the relevant moments, including the fourth moment of ξ_t, exist, (3.4) has an asymptotic χ^2 distribution with m degrees of freedom when H_0 holds.

This test statistic may also be computed by using an artificial regression. The F-version of the test is then carried out as follows.

1. Estimate the parameters of the conditional variance model under the null, compute $\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1$ and the sum of squared residuals, $SSR_0 = \sum_{t=1}^{T}(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)^2$

2. Regress $(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)$ on $\widehat{\mathbf{x}}_t'$, $\widehat{\mathbf{v}}_t'$ and compute the sum of squared residuals, $SSR_1 = \sum_{t=1}^{T}\widehat{\varepsilon}_t^2$.

3. Compute the F-version of the test statistic as $F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T - \dim(\boldsymbol{\omega}) - m)}$.

For the sample sizes relevant in GARCH modelling, there is no essential difference between the properties of the F-test and its asymptotically correct χ^2 counterpart $\chi^2 = T(SSR_0 - SSR_1)/SSR_0$.

3.2. Comparison with a portmanteau test

The test in the previous section was explicitly derived as an LM-test. Li and Mak (1994) recently introduced a portmanteau statistic for testing the adequacy of the standard GARCH(p,q) model. The null hypothesis is that the squared and standardized error process is not autocorrelated. In practice, one tests this hypothesis for the first m autocorrelations. Let $\mathbf{r} = (r_1, ..., r_m)'$ be the $m \times 1$ vector of the first m autocorrelations so that $\mathrm{H}_0 : \mathbf{r} = \mathbf{0}$. Li and Mak (1994) showed that under this hypothesis $\sqrt{T}\widehat{\mathbf{r}}$ is asymptotically normally distributed where T is the number of observations. The vector of autocorrelations is estimated by

$$\widehat{\widehat{\mathbf{r}}} = \frac{1}{T}\sum_{t=1}^{T}\left((\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)\widehat{\mathbf{v}}_t^*\right) / \left(\frac{1}{T}\sum_{t=1}^{T}(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)^2\right)$$

where $\widehat{u}_t = y_t - f(\mathbf{w}_t; \widehat{\boldsymbol{\varphi}})$, $\widehat{h}_t = h_t(\mathbf{w}_t; \widehat{\boldsymbol{\omega}})$ and $\widehat{\mathbf{v}}_t^* = \widehat{\mathbf{v}}_t - \mathbf{1}_m = (\frac{\widehat{u}_{t-1}^2}{\widehat{h}_{t-1}} - 1, ..., \frac{\widehat{u}_{t-m}^2}{\widehat{h}_{t-m}} - 1)'$ with $\mathbf{1}_m$= (1, ...,1)$'$ an $m \times 1$ vector. Note that under normality $\frac{1}{T}\sum_{t=1}^{T}(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)^2 \to 2$ in probability as $T \to \infty$. Thus

$$\widehat{\mathbf{r}} = \frac{1}{2T}\sum(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)\widehat{\mathbf{v}}_t^*$$

is asymptotically equivalent to $\widehat{\widehat{\mathbf{r}}}$. Under the null hypothesis, the asymptotic covariance matrix of $\sqrt{T}\widehat{\mathbf{r}}$ is block diagonal in $\boldsymbol{\varphi}$ and $\boldsymbol{\eta}$ and therefore estimated by

$$\mathbf{V}_{\mathbf{r}}(\widehat{\boldsymbol{\eta}}) = \mathbf{I}_m - \mathbf{X}_{\mathbf{r}}(\widehat{\boldsymbol{\eta}})'\mathbf{G}^{-1}(\widehat{\boldsymbol{\eta}})\mathbf{X}_{\mathbf{r}}(\widehat{\boldsymbol{\eta}})$$

since $\frac{1}{T}\sum \widehat{\mathbf{v}}_t^* \widehat{\mathbf{v}}_t^{*\prime} \to 2\mathbf{I}_m$ under H_0 as $T \to \infty$ and

$$\mathbf{X}_{\mathbf{r}}(\widehat{\boldsymbol{\eta}}) = -\frac{1}{2T}\sum_{t=1}^{T}\left(\frac{1}{\widehat{h}_t}\frac{\partial \widehat{h}_t}{\partial \widehat{\boldsymbol{\eta}}}\widehat{\mathbf{v}}_t^{*\prime}\right) = -\frac{1}{2T}\sum \widehat{\mathbf{x}}_t \widehat{\mathbf{v}}_t^{*\prime}.$$

Furthermore, $\mathbf{G}^{-1}(\widehat{\boldsymbol{\eta}})$ is some consistent estimator of the relevant block of the information matrix, evaluated at $\boldsymbol{\eta} = \widehat{\boldsymbol{\eta}}$. The portmanteau statistic becomes

$$Q(m) = T\widehat{\mathbf{r}}'\mathbf{V}_{\mathbf{r}}(\widehat{\boldsymbol{\eta}})^{-1}\widehat{\mathbf{r}} \tag{3.5}$$

which is asymptotically χ^2- distributed with m degrees of freedom under the null hypothesis. We may now also define

$$Q(m)^* = \frac{1}{4T}\left(\sum(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)\widehat{\mathbf{v}}_t^{*\prime}\right)\mathbf{V}_{\mathbf{r}}^*(\widehat{\boldsymbol{\eta}})^{-1}\left(\sum \widehat{\mathbf{v}}_t^*(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)\right) \tag{3.6}$$

where $\mathbf{V}_{\mathbf{r}}^*(\widehat{\boldsymbol{\eta}}) = \frac{1}{2T}\left(\sum \widehat{\mathbf{v}}_t^* \widehat{\mathbf{v}}_t^{*\prime} - \sum \widehat{\mathbf{v}}_t^* \widehat{\mathbf{x}}_t'(\sum \widehat{\mathbf{x}}_t \widehat{\mathbf{x}}_t')^{-1}\sum \widehat{\mathbf{x}}_t \widehat{\mathbf{v}}_t^{*\prime}\right)$. The only difference between (3.5) and (3.6) is the choice of the consistent estimator of the covariance matrix. $\mathbf{V}_{\mathbf{r}}^*(\widehat{\boldsymbol{\eta}})$ makes use of $\frac{1}{T}\sum \widehat{\mathbf{v}}_t^* \widehat{\mathbf{v}}_t^{*\prime}$ whereas its expectation $2\mathbf{I}_m$ appears in $\mathbf{V}_{\mathbf{r}}(\widehat{\boldsymbol{\eta}})$. The two covariance matrices are thus asymptotically equal. As (3.6) is identical to (3.4), the apparent difference in the expressions being due to centring, it follows that the Li and Mak (1994) portmanteau statistic and the LM-test (3.4) are asymptotically equivalent. This means that the portmanteau test of model adequacy is in fact an LM-test of no ARCH in the standardized errors against ARCH(m). This is analogous to the McLeod and Li (1983) portmanteau test being asymptotically equivalent to the classic LM-test of no ARCH of Engle (1982); see, for example, Luukkonen, Saikkonen and Teräsvirta (1988b) for a discussion. As a matter of fact, when $h_t \equiv \alpha_0$, our test and that of Li and Mak (1994) reduce to the Engle (1982) and McLeod and Li (1983) tests, respectively.

4. Misspecification of structure

In this section we present three different misspecification tests for an estimated conditional variance model. The first one can be interpreted as a test of the functional form. The second one is a test against nonlinearity or, in some cases, asymmetry. It is a modification of a test in Hagerud (1997). Finally we propose a test of parameter constancy against smooth continuous change in parameters. All three tests may be viewed as conditional variance counterparts of the tests for the nonlinear conditional mean in Eitrheim and Teräsvirta (1996). To describe the common features in these tests we first introduce a general structure and thereby consider each test separately.

4.1. General structure

Consider (2.2) and define

$$h(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi}) = \boldsymbol{\eta}'\mathbf{z}_t + G(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi}) \tag{4.1}$$

We assume that $h(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi})$ satisfies the regularity conditions mentioned in Section 2, and that $G(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi})$ is at least twice differentiable for all $\boldsymbol{\pi}$ everywhere in its sample space. Furthermore, let $G(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \mathbf{0}) \equiv \mathbf{0}$ which does not affect the generality of the argument. We also assume that the necessary moments of $\{u_t\}$ exists. Let $\boldsymbol{\omega} = (\boldsymbol{\varphi}', \boldsymbol{\eta}')'$ denote the parameters of the standard GARCH(p,q) model with the conditional mean specified according to (2.1). In that case, $h(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) = \boldsymbol{\eta}'\mathbf{z}_t$, and the null hypothesis of no additional structure in $h(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi})$ becomes $\mathrm{H}_0 : \boldsymbol{\pi} = 0$. The Lagrange multiplier (or score) test statistic is again (3.3) which is asymptotically χ^2-distributed with dim $(\boldsymbol{\pi})$ degrees of freedom under the null hypothesis and the required regularity conditions. Due to the fact that the information matrix is block diagonal under the null hypothesis it has the expression, (3.4) where $\mathbf{v}_t = \frac{\partial l_t}{\partial \boldsymbol{\pi}'}|_{\boldsymbol{\pi}=0} = \frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\pi}'}|_{\boldsymbol{\pi}=0}$ and $\mathbf{x}_t = \frac{\partial l_t}{\partial \boldsymbol{\eta}'}|_{\boldsymbol{\pi}=0} = \frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\eta}'}|_{\boldsymbol{\pi}=0}$. Furthermore $\widehat{\mathbf{v}}_t$ and $\widehat{\mathbf{x}}_t$ are the sample counterparts of the corresponding derivatives under H_0. Thus (3.4) may be computed by using an artificial regression as described in the previous section. The partial derivatives required to construct the Hessian needed for the estimation of the information matrix can be found in Appendix A.

4.2. Testing the functional form

Another way of testing the null hypothesis of no error autocorrelation in the squared residuals is to lag the linear combination $\boldsymbol{\eta}'\mathbf{z}_t$ and enter it in the conditional variance process, h_t, under the alternative. This may be viewed as a general but possibly parsimonious misspecification test along the lines in Bollerslev (1986). The test is obtained by defining $G(\mathbf{z}_t;\boldsymbol{\omega},\boldsymbol{\pi}) = \boldsymbol{\pi}'\mathbf{v}_t$ where $\boldsymbol{\pi} = (\pi_1, ..., \pi_r)'$ and $\mathbf{v}_t = (\boldsymbol{\eta}'\mathbf{z}_{t-1}, ..., \boldsymbol{\eta}'\mathbf{z}_{t-r})'$, in (4.1). The null hypothesis of no remaining serial dependence in the squared residuals or no model misspecification is $\mathrm{H}_0 : \boldsymbol{\pi} = 0$. The moment condition $Eu_t^4 < \infty$ must hold for the asymptotic theory to go through. Under the null hypothesis, the LM-statistic (3.3) is asymptotically χ^2- distributed with $\dim(\boldsymbol{\pi})$ degrees of freedom. On the other hand, Bollerslev (1986) suggested another test which is obtained by defining $G(\mathbf{z}_t;\boldsymbol{\omega},\boldsymbol{\pi}) = \boldsymbol{\eta}'\mathbf{z}_t^*$ where $\mathbf{z}_t^* = (u_{t-q-1}^2, ..., u_{t-q-m}^2)'$ or $\mathbf{z}_t^* = (h_{t-p-1}, ..., h_{1-p-n})'$. The alternative is thus a higher-order GARCH model. Note that it is not possible to test a GARCH(p,q) against a GARCH($p+n$,$q+m$) model using standard procedures when both m and n are assumed positive.

4.3. Testing linearity (symmetry)

The above test of the functional form offers few hints about what may be wrong when the null hypothesis is rejected. Bollerslev's (1986) test is more explicit about the alternative which is a higher-order GARCH model. Nevertheless, it may also be useful to consider other parametric alternatives to the symmetric GARCH(p,q) model. In some cases we may expect the response to be a function not only of the size of the shock but also of its direction. Engle and Ng (1993), see also references therein, considered this possibility. We call such a shock response asymmetric and parameterize it by generalizing the GJR-GARCH model of Glosten, Jagannathan and Runkle (1993). This is done in three ways. First, we make the transition between the extreme regimes smooth. Second, we incorporate a nonlinear version of the quadratic GARCH model of Sentana (1995) in our alternative. Finally, while the GJR-GARCH model is asymmetric, the present generalization may also retain the symmetry although the model becomes nonlinear. For smooth transition GARCH, see also Hagerud (1997) and González-Rivera (1998). Let

$$H_n(s_t;\gamma,\mathbf{c}) = \left(1 + \exp(-\gamma\prod_{l=1}^{n}(s_t - c_l))\right)^{-1}, \gamma > 0, c_1 \leq ... \leq c_n \tag{4.2}$$

where s_t is the transition variable at time t, γ is a slope parameter, and $\mathbf{c}$ a location vector. When $\gamma = 0$, $H_n(s_t;\gamma,\mathbf{c}) \equiv 1/2$ and when $\gamma \to \infty$, $H_n(s_t;\gamma,\mathbf{c})$ becomes a step function. The logistic function (4.2) is used for parameterizing the maintained model, and we assume $n \leq 2$. The alternative may now be written as

$$\begin{aligned} h_t &= \alpha_0 + \sum_{j=1}^{q} \alpha_{0j} H_n(u_{t-j}; \gamma, \mathbf{c}) \\ &\quad + \sum_{j=1}^{q} \{\alpha_{1j} + \alpha_{2j} H_n(u_{t-j}; \gamma, \mathbf{c})\} u_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \end{aligned} \tag{4.3}$$

where $\alpha_0 + \sum_{j=1}^{q} \alpha_{0j} H_n > 0$ and $\alpha_{1j} + \alpha_{2j} H_n \geq 0$, $j = 1, ..., q$, for $0 \leq H_n \leq 1$. This implies

$$G(\mathbf{z}_t; \boldsymbol{\theta}) = \sum_{j=1}^{q} \alpha_{0j} H_n(u_{t-j}; \gamma, \mathbf{c}) + \sum_{j=1}^{q} \alpha_{2j} H_n(u_{t-j}; \gamma, \mathbf{c}) u_{t-j}^2 \tag{4.4}$$

where $n = 1$ or 2. Assuming that the first sum on the right-hand side of (4.4) is identically to zero and letting $\gamma \to \infty$ yields the GJR-GARCH model. The test of the standard GARCH model against nonlinear GARCH in Hagerud (1997) may be viewed as a special case of this specification with $G(\mathbf{z}_t; \boldsymbol{\theta}) = \sum_{j=1}^{q} \alpha_{2j} H_n(u_{t-j}; \gamma, \mathbf{c}) u_{t-j}^2$.

Another way of parameterizing the alternative is to assume that the transition variable has a fixed delay. This assumption results in the following conditional variance model

$$\begin{aligned} h_t &= \alpha_0 + \alpha_{0d} H_n(u_{t-d}; \gamma, \mathbf{c}) \\ &\quad + \sum_{j=1}^{q} (\alpha_{1j} + \alpha_{2j} H_n(u_{t-d}; \gamma, \mathbf{c})) u_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \end{aligned} \tag{4.5}$$

which gives $G(\mathbf{z}_t; \boldsymbol{\theta}) = \alpha_{1d} H_n(u_{t-d}; \gamma, \mathbf{c}) + \sum_{j=1}^{q} \alpha_{2j} H_n(u_{t-d}; \gamma, \mathbf{c}) u_{t-j}^2$. This specification is very much in the spirit of Teräsvirta (1994) for the STAR-type conditional mean. Probably the most common case in practice is $d = q = p = 1$ so that (4.3) and (4.5) are equal. We do not impose any nonlinear structure on the h_{t-j}, $j = 1, ..., p$, as the alternative structure is already very flexible even without such an extension.

These smooth transition alternatives pose an identification problem. The null hypothesis can be expressed as $\mathrm{H}_0 : \gamma = 0$ in (4.2). It is seen that the parameters in (4.3) or (4.5) assuming a logistic transition function (4.2) are only identified under the alternative. The classical test procedures thus do not work; see, for example, Hansen (1996) for discussion. We circumvent the identification problem by following

Luukkonen, Saikkonen and Teräsvirta (1988a). This is done by expanding the transition function into a Taylor series around $\gamma = 0$, replacing the transition function with this Taylor approximation in (4.4) or (4.5) and rearranging terms. This results in

$$h_t = \boldsymbol{\eta}'\mathbf{z}_t + \boldsymbol{\pi}'\mathbf{v}_t + R_1\left(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}\right)$$

where it can be seen that $\boldsymbol{\pi} = \gamma\widetilde{\boldsymbol{\pi}}$ with $\widetilde{\boldsymbol{\pi}} \neq \mathbf{0}$. This being the case, the new null hypothesis $\mathrm{H}_0' : \boldsymbol{\pi} = 0$. Thus $G(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) = \boldsymbol{\pi}'\mathbf{v}_t + R_1\left(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}\right)$ in (4.1). Note that under H_0 we have $R_1(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) \equiv 0$ so that the remainder does not affect the distribution theory. For the smooth GJR-like alternative (4.3) we have $\boldsymbol{\pi} = (\pi_1, ..., \pi_{(1+n)q})'$ and $\mathbf{v}_t = (u_{t-1}, u_{t-1}^3, ..., u_{t-1}^{n+2} ..., u_{t-q}, u_{t-q}^3, ..., u_{t-q}^{n+2})'$. For the alternative with a fixed delay (4.5) there are two possibilities depending on the delay. If the delay is to be found within q then $\mathbf{a} = (a_1, ..., a_{1+nq})'$ and $\mathbf{v}_t = (u_{t-d}, u_{t-d}u_{t-1}^2, ..., u_{t-d}^n u_{t-1}^2 ..., u_{t-d}u_{t-q}^2, ..., u_{t-d}^n u_{t-q}^2)'$, otherwise $\boldsymbol{\pi} = (\pi_1, ..., \pi_{3+nq})'$ and $\mathbf{v}_t = (u_{t-d}, u_{t-d}^2, u_{t-d}^3, u_{t-d}u_{t-1}^2, ..., u_{t-d}^n u_{t-q}^2)'$. For both models the moment condition of $Eu_t^{2n+2} < \infty$ must hold for the asymptotic theory to go through. Under the null hypothesis the LM-statistic (3.3) is asymptotically χ^2- distributed with $\dim(\boldsymbol{\pi})$ degrees of freedom. If d in (4.5) is assumed unknown a priori, the test may be generalized to that situation along the lines in Luukkonen et al. (1988a). Note that if $n = 2$ and $c_1 = -c_2$ the alternative is symmetric in that a positive shock and a negative one of the same magnitude still have the same (mirror) effect on the conditional variance. The response to the shock, however, is a nonlinear function of lags of u_t^2.

4.4. Testing parameter constancy

Testing parameter constancy is important in its own right but also because nonconstancy may manifest itself as apparent lack of weak stationarity (IGARCH); see, for example, Lamoureux and Lastrapes (1990). In this paper we assume that the alternative to constant parameters in the conditional variance is that the parameters, or a subset of them, change smoothly over time. Lin and Teräsvirta (1994) applied this idea to testing parameter constancy in the conditional mean. We postulate

$$h_t = \boldsymbol{\eta}(t)'\mathbf{z}_t \tag{4.6}$$

where the time-varying parameter is $\boldsymbol{\eta}(t) = \boldsymbol{\eta}^* + \boldsymbol{\lambda}H_n(t;\gamma,\mathbf{c})$. If the null hypothesis only concerns a subset of parameters then only the corresponding elements in $\boldsymbol{\lambda}$ are assumed to be nonzero a priori. The transition function $H_n(t;\gamma,\mathbf{c})$ is assumed to be a logistic function of order n defined in (4.2) with $s_t \equiv t$. If $\gamma \to \infty$, $H_1(t;\gamma,\mathbf{c})$ becomes a step-function and characterizes a single structural break in the model. Chu (1995) discussed testing parameter constancy against this alternative. The null hypothesis of parameter constancy becomes $\mathrm{H}_0 : \gamma = 0$ against $\mathrm{H}_1 : \gamma > 0$. We can again circumvent the lack of identification under the null hypothesis by a Taylor approximation of the transition function. A first-order Taylor-expansion of $H_n(t;\gamma,\mathbf{c})$ around $\gamma = 0$ yields, after a reparameterization,

$$h_t = \boldsymbol{\beta}_0'\mathbf{z}_t + \boldsymbol{\pi}'\mathbf{v}_t + R_2(\mathbf{z}_t;\boldsymbol{\omega},\boldsymbol{\pi})$$

where $\boldsymbol{\pi} = (\boldsymbol{\beta}_1',...,\boldsymbol{\beta}_n')' = \gamma\widetilde{\boldsymbol{\pi}}(\boldsymbol{\eta},\gamma,\mathbf{c})$ and $\mathbf{v}_t = ((\mathbf{z}_t t)',...,(\mathbf{z}_t t^n)')'$. Thus, $G(\mathbf{z}_t;\boldsymbol{\omega},\boldsymbol{\pi}) = \boldsymbol{\pi}'\mathbf{v}_t + R_2(\mathbf{z}_t;\boldsymbol{\omega},\boldsymbol{\pi})$. Note, however, that $R_2(\mathbf{z}_t;\boldsymbol{\omega},\boldsymbol{\pi}) \equiv 0$ under H_0 so that it does not affect the distribution theory. Our null hypothesis is $\mathrm{H}_0 : \boldsymbol{\pi} = \mathbf{0}$. We note that components of $\mathbf{w}_t$ are trending but modifying a corresponding proof in Lin and Teräsvirta (1994) the asymptotic null distribution of (3.4) can be shown to be a χ^2 distribution if the fourth moment of u_t exists.

An advantage of a parametric alternative such as (4.6) to parameter constancy is that if the null hypothesis is rejected we can estimate the parameters of the alternative model. This helps us find out where in the sample the parameters under test seem to be changing and how rapid the change is. This is useful information if respecification of the model to achieve parameter constancy is attempted.

5. Simulation experiments

As the above theory is asymptotic we have to find out how our tests behave in finite samples. This is done by simulation. For all simulations we used the following data generating process (DGP)

$$\begin{aligned} y_t &= u_t \\ u_t &= \varepsilon_t \sqrt{h_t}. \end{aligned} \tag{5.1}$$

The conditional variance h_t varies with the alternative hypothesis we shall be testing against. Under the null hypothesis h_t is the conditional variance of the standard GARCH(1,1) model. The random numbers, ε_t, were generated by the random number generator in GAUSS 3.2.31. The random numbers sampled were all assumed to be normally distributed with expectation zero and unit variance. The first 200 observations of each generated series were always discarded to avoid initialization effects. For the tests against remaining structure, series of 2000 observations were generated. Series of 1000 observations were used both for the linearity test and parameter constancy test. For each design a total of 1000 replications were performed.

5.1. Test of the functional form

First we consider the test of no ARCH in the standardized errors and the functional misspecification test of Section 4.2. We define a DGP is such that the conditional variance either (a) follows a GARCH$(1,2)$ process or (b) follows a GARCH$(2,1)$ process. Thus,

$$h_t = 0.5 + 0.05u_{t-1}^2 + \alpha u_{t-2}^2 + 0.9h_{t-1} \tag{5.2a}$$

$$h_t = 0.5 + 0.05u_{t-1}^2 + 0.9h_{t-1} + \beta h_{t-2} \tag{5.2b}$$

In the experiment the value of α and β vary within limits such that the conditional variance of the process remains positive with probability 1, see Nelson and Cao (1992), and the condition for covariance stationarity holds. For (a) this is the case when $-0.045 < \alpha < 0.05$ and for (b) when $-0.2025 < \beta < 0.1$. For $\alpha = 0$ in (a) and $\beta = 0$ in (b) the DGP reduces to a standard GARCH$(1,1)$ model. The results for the test of functional misspecification and for the test of no ARCH in the standardized errors for (a) and (b), are reported in Figures B.1 and B.2. The nominal significance level equals 0.1, and we use the test of Bollerslev (1986) as a benchmark. In the simulations these tests were all computed with a single parameter in the alternative, that is, with either $\widehat{\mathbf{v}}_t = \widehat{u}_{t-1}^2/h_{t-1}$ (no ARCH) or with $\widehat{\mathbf{v}}_t = \widehat{\boldsymbol{\eta}}\mathbf{z}_{t-1}$ where $\mathbf{z}_{t-1} = (1, \widehat{u}_{t-1}^2, \widehat{h}_{t-1})$ (functional form). We also used Bollerslev's test such that the alternative was in the DGP. In practice one does not know if the unconditional fourth moment of $\{u_t\}$ exists, although the estimated model does contain information about this. It is therefore of interest to investigate how the tests behave when the unconditional fourth moment does not exist; the existence condition is given in He and Teräsvirta (1999a). This is the case for parameter values on the right hand side of the dashed vertical line in Figures B.1 and B.2.

In the experiment where the conditional variance of the DGP is (5.2a) the power of the functional form test is lower than that of Bollerslev's test. On the other hand, for positive values of β when the DGP is (5.2b) the test of the functional form performs better than Bollerslev's test. The latter test has no power against the GARCH(2,1) alternative when the unconditional fourth moment exists. When the existence condition for the unconditional variance is violated the power of the functional form test first sharply increases and then decreases as a function of β. As to Bollerslev's test, its power increases as a function of β without dropping again. The power of our test against remaining ARCH(1), which is asymptotically equivalent to the test of Li and Mak (1994), is found to be equal to that of the Bollerslev's test in both cases (a) and (b).

It may be pointed out already that these tests have no power when the DGP is a nonlinear GJR-GARCH(1,1) model (5.3) or contains a structural break (5.4). Simulations of other tests against such alternatives will be discussed in the next two sections.

5.2. Testing linearity (symmetry)

In this section we consider the small-sample performance of the linearity test. The test can be expected to be powerful against smooth transition alternatives for which it is designed. Considering its performance against, say, the GJR-GARCH model would constitute a tougher trial for our test and make it possible to compare the performance of the tests with that of the sign-bias test of Engle and Ng (1993). The DGP is (2.2) with

$$h_t = 0.005 + 0.28 \left[|u_{t-1}| + \omega u_{t-1}\right]^2 + 0.7 h_{t-1} \quad (5.3)$$

where u_t is assumed conditionally normal. For (5.3), $|\omega| < 0.267$ is required for covariance stationarity. In this special case the unconditional fourth moment also exists under this restriction; the relevant moment condition appeared in He and Teräsvirta (1999b). The joint sign-bias test of Engle and Ng (1993) mentioned above is designed for detecting asymmetry in the conditional variance. We compute the values of the sign-bias test statistic by using (3.3); the difference between this test and the TR^2 version, suggested by Engle and Ng (1993), is negligible at our sample size. Engle and Ng (1993) used (2.2) and (5.3) with $\omega = -0.23$ as the DGP in their evaluation of the sign-bias test. Note that for $\omega = 0$ the DGP reduces to a standard GARCH(1,1) model. The power curves are reported in Figure B.3 for $\omega \leq 0$. For positive values of ω the results are similar and therefore not reported. Our smooth transition GARCH test was computed by assuming $n = 3$ in (4.2). The power of the test compares very favourably with that of the sign-bias test. This accords with the results in Hagerud (1997)

5.3. Testing parameter constancy

We consider two cases of parameter nonconstancy: the DGP is a GARCH(1,1) model with either (a) a single or (b) a double structural break. We did not simulate smooth parameter change because our test can be expected to perform well against such an alternative. Our choice of alternative also gives us an opportunity to compare our test against that of Chu (1995) which is a test against a single structural break. If T is the total number of observations, the single structural break parametrization corresponding to alternative (a) is assumed to have a change at time ηT where η lies between 0 to 1. The double structural break parametrization corresponding to (b) first postulates a change at time $\eta_1 T$ and a return to the original parameters at $\eta_2 T$, $0 \leq \eta_1 < \eta_2 \leq 1$. The test of Chu (1995) should outperform ours in case (a). We use the version of Chu's test that assumes normal errors. Our test is computed with $n = 1$, case (a), and $n = 2$, case (b), where n is the order of the logistic function in (4.6).

We consider the following model for a change in the constant term:

$$\begin{aligned} h_t &= 0.5 + 0.1u_{t-1}^2 + 0.8h_{t-1},\ \text{(a) } t < \eta T,\ \text{(b) } t < \eta_1 T, t > \eta_2 T \\ h_t &= 0.5(1+\triangle) + 0.1u_{t-1}^2 + 0.8h_{t-1},\ \text{(a) } t \geq \eta T,\ \text{(b) } \eta_1 T \leq t \leq \eta_2 T \end{aligned} \tag{5.4}$$

where $\triangle = 0.4, 0.8$. Chu (1995) used (a) in (5.4) as the DGP in his own simulation experiments. The power curves for the DGPs in (5.4) with a single structural break at η for $\triangle = 0.4$ and 0.8 appear in Figures B.4 and B.5. The values $\eta = 0, 1$ correspond to the null hypothesis.

For $0.4 < \eta < 0.8$ our test has the same as or higher power than the test of Chu (1995), but otherwise the relationship is the opposite. This occurs for both $\triangle = 0.4$ and $\triangle = 0.8$ which includes these DGPs. Thus the Chu test does not dominate ours as one might expect. We can see that a small change, $\triangle = 0.4$, is difficult to detect. As $\triangle$ doubles to 0.8, the change is detected more easily. For comparison, we also simulated another version of the test in which we assumed that only the constant term is time-varying under the alternative. In that case our test outperforms that of Chu (1995) for almost all η. In yet another experiment we allowed the coefficient of u_{t-1}^2 to change once within the sample period. The behaviour of the tests was similar to the previous case and the details are not reported here.

We turn to the case of a double structural break. The DGP for the experiment is such that $\eta_2 = \eta_1 + 0.3$ where η_1 is varied from 0 to 0.7. For $\eta_1 = 0$ and $\eta_1 = 0.7$ the DGP thus has only a single structural break. The power curves of this experiment for $\triangle = 0.4$ and 0.8 can be found in Figures B.6 and B.7. In this case the test of Chu (1995) cannot be expected to be very powerful because the design of the experiment does not favour it, and our test does have superior power for all double break points considered. The power of the Chu test is high for η_1 close to zero and 0.7 because the test is designed for detecting a single structural break.

6. Conclusions

In this paper we have derived a unified framework for testing the adequacy of an estimated GARCH model. Our selection contains a number of new tests while some existing ones fit into this framework as well. Nothing more complicated than standard asymptotic distribution theory is required. As a result, misspecification of a GARCH model may be detected quite easily at low computational cost. Because the test of symmetry and parameter constancy are parametric, the alternative may be estimated if the null hypothesis is rejected. This helps the model builder find out what the weaknesses of his/her estimated specification are and may give useful ideas of how the current specification could be improved.

We also show that our test of no ARCH in the standardized error process is asymptotically equivalent to a portmanteau test of Li and Mak (1994). This links the work of these authors to our framework and indicates that the null hypothesis of no remaining ARCH can be tested in different ways while the asymptotic theory remains the same. This article does not contain any applications of these new tests. Empirical examples can be found instead in the companion paper Lundbergh and Teräsvirta (1998).

References

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31**: 307–327.

Bollerslev, T. and Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances., *Econometric Reviews* **11**: 143–172.

Chu, C.-S. J. (1995). Detecting parameter shift in GARCH models, *Econometric Reviews* **14**: 241–266.

Eitrheim, Ø. and Teräsvirta, T. (1996). Testing the adequacy of smooth transition autoregressive models, *Journal of Econometrics* **74**: 59–75.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50**: 987–1007.

Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility, *Journal of Finance* **48**: 1749–1777.

Fiorentini, G., Calzolari, G. and Panattoni, L. (1996). Analytic derivatives and the computation of GARCH estimates, *Journal of Applied Econometrics* **11**: 399–417.

Glosten, L., Jagannathan, R. and Runkle, D. (1993). On the relation between expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* **48**: 1779–1801.

González-Rivera, G. (1998). Smooth transition GARCH models, *Studies in Nonlinear Dynamics and Econometrics* (forthcoming) .

Hagerud, G. E. (1997). *A New Non-Linear GARCH Model*, EFI, Stockholm School of Economics.

Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica* **64**: 413–430.

He, C. and Teräsvirta, T. (1999a). Fourth moment structure of the GARCH(p,q) process, *Econometric Theory* (forthcoming) .

He, C. and Teräsvirta, T. (1999b). Properties of moments of a family of GARCH processes, *Journal of Econometrics* (forthcoming) .

He, C. and Teräsvirta, T. (1999c). Properties of the autocorrelation function of squared observations for second order GARCH processes under two sets of parameter constraints, *Journal of Time Series Analysis* (forthcoming) .

Lamoureux, C. G. and Lastrapes, W. G. (1990). Persistence in variance, structural change and the GARCH model, *Journal of Business and Economic Statistics* **8**: 225–234.

Lee, Y. N. and Li, W. K. (1998). On smooth transition double threshold models. Department of Statistics, University of Hong Kong, Research Report No. 198.

Li, W. K. and Mak, T. K. (1994). On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity, *Journal of Time Series Analysis* **15**: 627–636.

Lin, C.-F. and Teräsvirta, T. (1994). Testing the constancy of regression parameters against continuous structural change, *Journal of Econometrics* **62**: 211–228.

Lumsdaine, R. L. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models, *Econometrica* **64**: 575–596.

Lundbergh, S. and Teräsvirta, T. (1998). Modelling economic high-frequency time series with STAR-STGARCH models. Stockholm School of Economics, Working Paper Series in Economics and Finance, No. 291.

Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988a). Testing linearity against smooth transition autoregressive models, *Biometrika* **75**: 491–499.

Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988b). Testing linearity in univariate time series models, *Scandinavian Journal of Statistics* **15**: 161–175.

McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared residual autocorrelations, *Journal of Time Series Analysis* **4**: 269–273.

Nelson, D. B. and Cao, C. Q. (1992). Inequality constraints in the univariate GARCH model, *Journal of Business and Economic Statistics* **10**: 229–235.

Sentana, E. (1995). Quadratic ARCH models, *Review of Economic Studies* **62**: 639–661.

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association* **89**: 208–218.

A. Information matrix

Consider the model defined by (2.1) and (2.2):

$$\begin{aligned} y_t &= \boldsymbol{\varphi}'\mathbf{w}_t + u_t \\ u_t &= \varepsilon_t \sqrt{h(z_t;\boldsymbol{\eta},\boldsymbol{\pi})} \end{aligned}$$

where $h_t = h(z_t;\boldsymbol{\varphi},\boldsymbol{\eta},\boldsymbol{\pi})$ is the parametrization of the conditional variance including the alternative. The assumed null hypothesis is $\mathrm{H}_0 : \boldsymbol{\pi} = \mathbf{0}$. If we assume that $\{\varepsilon_t\}$ is a sequence of independent standard normal errors, the log-likelihood function at time t is given by:

$$l_t = \text{const} - \frac{1}{2}\ln h_t - \frac{1}{2}\frac{u_t^2}{h_t}$$

where $u_t = y_t - \boldsymbol{\varphi}'\mathbf{w}_t$ as $\boldsymbol{\varphi}$ is not assumed to depend on $\boldsymbol{\eta}$ or $\boldsymbol{\pi}$.

A.1. Partial derivative of l_t

The first-order partial derivative (the gradient) of the log likelihood function at time t is

$$G_t = \left(\frac{\partial l_t}{\partial \boldsymbol{\varphi}'} \; \frac{\partial l_t}{\partial \boldsymbol{\eta}'} \; \frac{\partial l_t}{\partial \boldsymbol{\pi}'} \right)$$

where the corresponding elements are

$$\begin{aligned} \frac{\partial l_t}{\partial \boldsymbol{\varphi}'} &= \frac{u_t}{h_t}\mathbf{w}_t' + \frac{1}{2h_t}\left(\frac{u_t^2}{h_t} - 1\right)\frac{\partial h_t}{\partial \boldsymbol{\varphi}'} \\ \frac{\partial l_t}{\partial \boldsymbol{\eta}'} &= \frac{1}{2h_t}\left(\frac{u_t^2}{h_t} - 1\right)\frac{\partial h_t}{\partial \boldsymbol{\eta}'} \\ \frac{\partial l_t}{\partial \boldsymbol{\pi}'} &= \frac{1}{2h_t}\left(\frac{u_t^2}{h_t} - 1\right)\frac{\partial h_t}{\partial \boldsymbol{\pi}'} \end{aligned}$$

The second-order partial derivative (the Hessian) of the log-likelihood function at time t is

$$H_t = \begin{pmatrix} \frac{\partial^2 l_t}{\partial \boldsymbol{\varphi}\,\partial \boldsymbol{\varphi}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\varphi}\,\partial \boldsymbol{\eta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\varphi}\,\partial \boldsymbol{\pi}'} \\ \frac{\partial^2 l_t}{\partial \boldsymbol{\eta}\,\partial \boldsymbol{\varphi}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\eta}\,\partial \boldsymbol{\eta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\eta}\,\partial \boldsymbol{\pi}'} \\ \frac{\partial^2 l_t}{\partial \boldsymbol{\pi}\,\partial \boldsymbol{\varphi}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\pi}\,\partial \boldsymbol{\eta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\pi}\,\partial \boldsymbol{\pi}'} \end{pmatrix} \tag{A.1}$$

If H_0 holds, the expectations of the elements in (A.1) are

$$E\left[\frac{\partial^2 l_t}{\partial\varphi\partial\varphi'}|_{\pi=\mathbf{0}}\right] = -E\left[\frac{1}{h_t}\mathbf{w}_t\mathbf{w}_t' + \frac{1}{2h_t^2}\frac{\partial h_t}{\partial\varphi}\frac{\partial h_t}{\partial\varphi'}\right] \qquad E\left[\frac{\partial^2 l_t}{\partial\varphi\partial\eta'}|_{\pi=\mathbf{0}}\right] = -\frac{1}{2}E\left[\frac{1}{h_t^2}\frac{\partial h_t}{\partial\varphi}\frac{\partial h_t}{\partial\eta'}\right]$$

$$E\left[\frac{\partial^2 l_t}{\partial\eta\partial\eta'}|_{\pi=\mathbf{0}}\right] = -\frac{1}{2}E\left[\frac{1}{h_t^2}\frac{\partial h_t}{\partial\eta}\frac{\partial h_t}{\partial\eta'}\right] \qquad E\left[\frac{\partial^2 l_t}{\partial\varphi\partial\pi'}|_{\pi=\mathbf{0}}\right] = -\frac{1}{2}E\left[\frac{1}{h_t^2}\frac{\partial h_t}{\partial\varphi}\frac{\partial h_t}{\partial\pi'}\right]$$

$$E\left[\frac{\partial^2 l_t}{\partial\pi\partial\pi'}|_{\pi=\mathbf{0}}\right] = -\frac{1}{2}E\left[\frac{1}{h_t^2}\frac{\partial h_t}{\partial\pi}\frac{\partial h_t}{\partial\pi'}\right] \qquad E\left[\frac{\partial^2 l_t}{\partial\pi\partial\eta'}|_{\pi=\mathbf{0}}\right] = -\frac{1}{2}E\left[\frac{1}{h_t^2}\frac{\partial h_t}{\partial\pi}\frac{\partial h_t}{\partial\eta'}\right]$$

A.2. Partial derivative of the multiplicative conditional variance h_t.

Assume that the conditional variance under the alternative hypothesis is parameterized as in (3.2). The conditional variance is then

$$h_t = (\boldsymbol{\eta}'\mathbf{z}_t)(1+\boldsymbol{\pi}'\mathbf{v}_t) = (\alpha_0 + \sum_{i=1}^{q}\alpha_i u_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j})(1+\boldsymbol{\pi}'\mathbf{v}_t)$$

which reduces into the standard GARCH model under H_0. To initialize the iterative computation of h_t under null hypothesis the conditional variance is estimated with the unconditional variance (sample variance) in the pre-sample case. This is done for all $t \leq 0$ by setting $h_t = u_t^2 = \frac{1}{T}\sum_{s=1}^{T}u_s^2$ where $u_s = y_s - \boldsymbol{\varphi}'\mathbf{w}_s$.

To compute the test statistic (3.4) we need the first-order partial derivatives of the conditional variance h_t under the null hypothesis.

First-order derivative	Pre-sample values ($t \leq 0$)
$\frac{\partial h_t}{\partial\varphi'}\|_{\pi=\mathbf{0}} = -2\sum_{i=1}^{q}\alpha_i u_{t-i}\mathbf{w}_{t-i}' + \sum_{j=1}^{p}\beta_j\frac{\partial h_{t-j}}{\partial\varphi'}\|_{\pi=\mathbf{0}}$	$-2u_t\mathbf{w}_t'\|_{\pi=\mathbf{0}} = \frac{\partial h_t}{\partial\varphi'}\|_{\pi=\mathbf{0}} = -\frac{2}{T}\sum_{s=1}^{T}u_s\mathbf{w}_t'$
$\frac{\partial h_t}{\partial\eta'}\|_{\pi=\mathbf{0}} = \mathbf{z}_t' + \sum_{j=1}^{p}\beta_j\frac{\partial h_{t-j}}{\partial\eta'}\|_{\pi=\mathbf{0}}$	$\frac{\partial h_t}{\partial\eta'}\|_{\pi=\mathbf{0}} = 0$
$\frac{\partial h_t}{\partial\pi'}\|_{\pi=\mathbf{0}} = (\boldsymbol{\eta}'\mathbf{z}_t)\mathbf{v}_t'$	$\frac{\partial h_t}{\partial\pi'}\|_{\pi=\mathbf{0}} = 0$

A.3. Partial derivative of the additive conditional variance h_t.

Assume that the conditional variance under the alternative hypothesis is parameterized as in (4.1) The conditional variance is then

$$h_t = \boldsymbol{\eta}'\mathbf{z}_t + \boldsymbol{\pi}'\mathbf{v}_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i u_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j} + \boldsymbol{\pi}'\mathbf{v}_t$$

which reduces to the standard GARCH model under H_0. The iterative computation of h_t under null hypothesis is initialized in the same way as for the multiplicative conditional variance model. The first-order derivatives of the conditional variance h_t under the null hypothesis are required to compute the test statistic (3.3). These derivatives are given as follows:

First-order derivative	Pre-sample values ($t \leq 0$)
$\frac{\partial h_t}{\partial \boldsymbol{\varphi}'}\vert_{\boldsymbol{\pi}=\mathbf{0}} = -2\sum_{i=1}^{q}\alpha_i u_{t-i}\mathbf{w}_{t-i}^{'} + \sum_{j=1}^{p}\beta_j \frac{\partial h_{t-j}}{\partial \boldsymbol{\varphi}'}\vert_{\boldsymbol{\pi}=\mathbf{0}}$	$-2u_t\mathbf{w}_t'\vert_{\boldsymbol{\pi}=\mathbf{0}} = \frac{\partial h_t}{\partial \boldsymbol{\varphi}'}\vert_{\boldsymbol{\pi}=\mathbf{0}} = -\frac{2}{T}\sum_{s=1}^{T}u_s\mathbf{w}_t'$
$\frac{\partial h_t}{\partial \boldsymbol{\eta}'}\vert_{\boldsymbol{\pi}=\mathbf{0}} = \mathbf{z}_t' + \sum_{j=1}^{p}\beta_j \frac{\partial h_{t-j}}{\partial \boldsymbol{\eta}'}\vert_{\boldsymbol{\pi}=\mathbf{0}}$	$\frac{\partial h_t}{\partial \boldsymbol{\eta}'}\vert_{\boldsymbol{\pi}=\mathbf{0}} = \mathbf{0}$
$\frac{\partial h_t}{\partial \boldsymbol{\pi}'}\vert_{\boldsymbol{\pi}=\mathbf{0}} = \mathbf{v}_t' + \sum_{j=1}^{p}\beta_j \frac{\partial h_{t-j}}{\partial \boldsymbol{\pi}'}\vert_{\boldsymbol{\pi}=\mathbf{0}}$	$\frac{\partial h_t}{\partial \boldsymbol{\pi}'}\vert_{\boldsymbol{\pi}=\mathbf{0}} = \mathbf{0}$

B. Figures

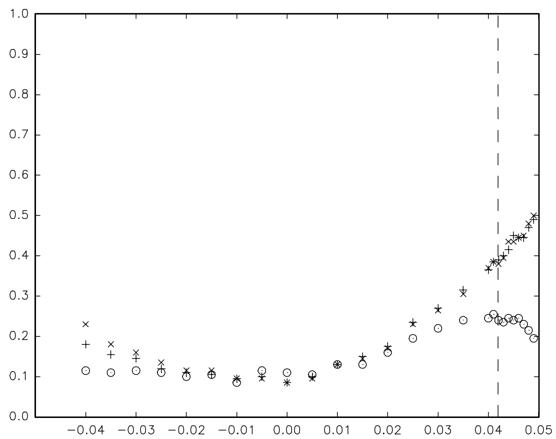

Figure B.1: Power curves at significance level 0.1 for the no ARCH test (cross), the functional form test (circle) and Bollerslev's test (plus). The DGP is a GARCH(1,2). The value of α in (5.2a) is given on the x-axis. The null hypothesis is the GARCH(1,1) model.

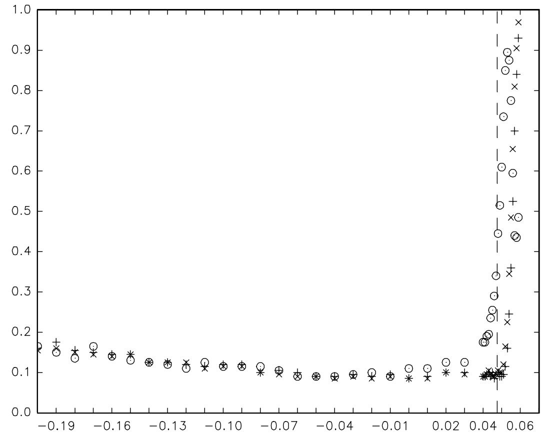

Figure B.2: Power curves at significance level 0.1 for the no ARCH test (cross), the functional form test (circle) and Bollerslev's test (plus). The DGP is a GARCH(2,1). The value of β in (5.2b) is given on the x-axis. The null hypothesis is the GARCH(1,1) model.

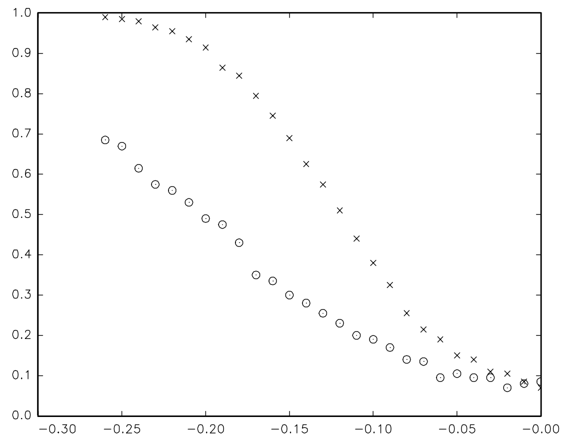

Figure B.3: Power curves at significance level 0.1 for the linearity (symmetry) test (cross) and the sign-bias test (circle). The DGP is a GJR-GARCH. The value of ω in (5.3) is given on the x-axis. The null hypothesis is the GARCH(1,1) model.

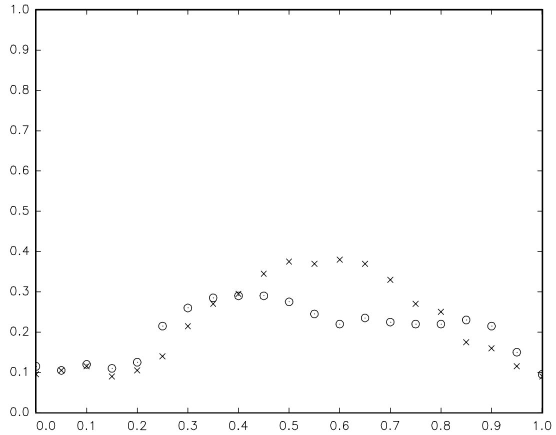

Figure B.4: Power curves at significance level 0.1 for the parameter constancy test (cross) and the test of Chu (circle). The DGP is a GARCH(1,1) with a single structural shift of size $\Delta = 0.4$ at η. In the figure η is given on the x-axis. The null hypothesis is a constant parameter GARCH(1,1) model.

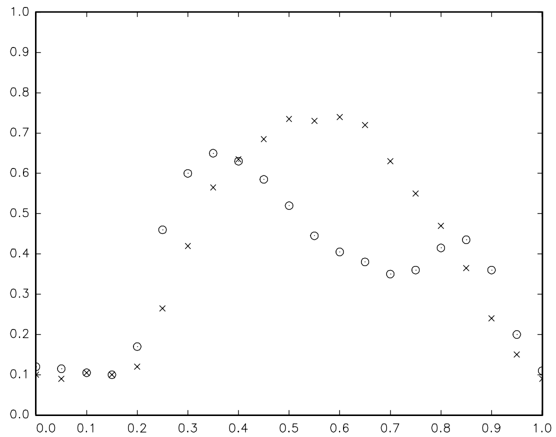

Figure B.5: Power curves at significance level 0.1 for the parameter constancy test (cross) and the test of Chu (circle). The DGP is a GARCH(1,1) with a single structural shift of size $\Delta = 0.8$ at η. In the figure η is given on the x-axis. The null hypothesis is a constant parameter GARCH(1,1) model.

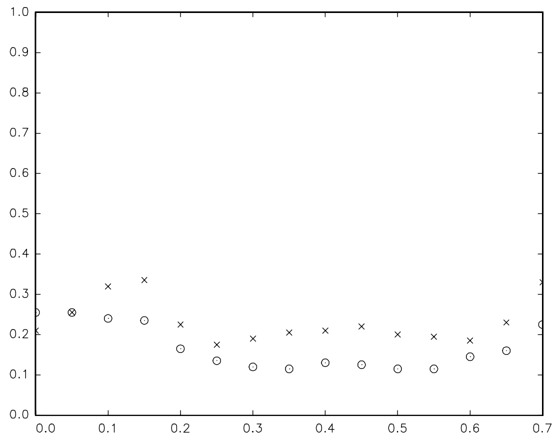

Figure B.6: Power curves at significance level 0.1 for our parameter constancy test (cross) and the test of Chu (circle). The DGP is a GARCH(1,1) process with a double structural shift in the constant of size $\Delta = 0.4$. The first shift appearing begin at η_1 and the return to original parameter values occuring at η_2. In the figure the value of η_1 is given on the x-axis and defined such that $\eta_2 = \eta_1 + 0.3$. The null hypothesis is a constant parameter GARCH(1,1) model.

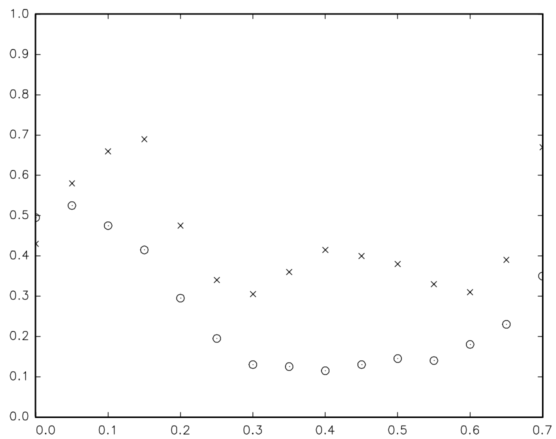

Figure B.7: Power curves at significance level 0.1 for our parameter constancy test (cross) and the test of Chu (circle). The DGP is a GARCH(1,1) process with a double structural shift in the constant of size $\Delta = 0.8$. The first shift appearing begin at η_1 and the return to original parameter values occuring at η_2. In the figure the value of η_1 is given on the x-axis and defined such that $\eta_2 = \eta_1 + 0.3$. The null hypothesis is a constant parameter GARCH(1,1) model.