Evaluating exponential GARCH models

Hans Malmsten
Department of Economic Statistics
Stockholm School of Economics
Box 6501, SE-113 83 Stockholm, Sweden

SSE/EFI Working Paper Series in Economics and Finance
No 564
August 2004

Abstract

In this paper, a unified framework for testing the adequancy of an estimated EGARCH model is presented. The tests are Lagrange multiplier or Lagrange multiplier type tests and include testing an EGARCH model against a higher-order one and testing parameter constancy. Furthermore, various existing ways of testing the EGARCH model against GARCH one are investigated as another check of model adequancy. This is done by by size and power simulations. Small-sample properties of the other tests are also investigated by simulations.

Keywords. Evaluation of volatility models; modelling volatility; parameter constancy; GARCH

JEL Classification Codes: C22, C52

Acknowledgements. This research has been supported by the Swedish Research Council for Humanities and Social Sciences, and by Jan Wallander's and Tom Hedelius's Foundation, Grant No. J99/37. The author also wishes to thank Timo Teräsvirta for useful comments.

1 Introduction

Model evaluation is an important part of modelling not only for the conditional mean models but for the conditional variance specifications as well. It is useful, one could argue even necessary, to carry out an in-sample evaluation of a volatility model before it is used for forecasting. Engle & Ng (1993), Li & Mak (1994) and Chu (1995), to name a few examples, derived misspecification tests for generalized autoregressive heteroskedasticity (GARCH) models. Recently, Lundbergh & Teräsvirta (2002) presented a unified framework for testing the adequacy of an estimated GARCH model. Their framework covers, among other things, testing the null of no ARCH in the standardized errors, testing symmetry against a smooth transition GARCH (STGARCH) model and a test of parameter constancy against smoothly changing parameters.

It appears that less work has been done for the evaluation of exponential GARCH (EGARCH) model by Nelson (1991). In fact, Nelson already suggested several tests based on orthogonality conditions that the errors of the model satisfy under the null hypothesis, but not much has happened since, nor have Nelson's tests been regularly applied in empirical work. In this paper we continue Nelson's work and consider a number of misspecification tests for the EGARCH model. They are Lagrange multiplier or Lagrange multiplier type tests and include testing an EGARCH model against a higher-order one and testing parameter constancy. Furthermore, we investigate various ways of testing the EGARCH model against GARCH ones as another check of model adequacy. The literature on testing non-nested hypotheses for volatility models includes Chen & Kuan (2002), Kim, Shephard & Chib (1998) and Lee & Brorsen (1997); see also Engle & Ng (1993) and Ling & McAleer (2000). Their tests are considered in the present framework, and the small-sample properties of the tests are investigated by simulation.

The plan of the paper is as follows. The model is defined in Section 2 and the estimation of parameters is discussed briefly in Section 3. Section 4 considers testing an EGARCH model against a higher order one and testing parameter constancy. In Section 5 non-nested tests for testing EGARCH and GARCH models against each other are discussed. Section 6 contains results of a simulation experiment and Section 7 an empirical example. Finally, conclusion appear in Section 8.

2 The model

Let

$$y_t = f(\mathbf{w}_t; \boldsymbol{\varphi}) + \varepsilon_t, t = 1, ..., T \tag{1}$$

where f is at least twice continuously differentiable function of $\boldsymbol{\varphi}$, with $\mathbf{w}_t = (1, y_{t-1}, ..., y_{t-n}, x_{1t}, ..., x_{kt})'$. The error process is parameterized as

$$\varepsilon_t = z_t h_t^{1/2}, t = 1, ..., T \tag{2}$$

where $\{z_t\}$ is a sequence of independent identically distributed random variables with zero mean and unit variance. A family of EGARCH(p,q) models may be defined as a combination of (2) and

$$\ln h_t = \alpha_0 + \sum_{j=1}^{q} g_j(z_{t-j}) + \sum_{j=1}^{p} \beta_j \ln h_{t-j}. \tag{3}$$

The conditional variance is constrained to be non-negative by the assumption that the logarithm of h_t is a function of past z_t's. Equations (1) and (3) define a class of EGARCH(p,q) models. Setting

$$g_j(z_{t-j}) = \alpha_j z_{t-j} + \psi_j(|z_{t-j}| - E\,|z_t|), j = 1, ..., q, \tag{4}$$

in (3) yields the EGARCH(p,q) model proposed by Nelson (1991). The overwhelmingly most popular EGARCH model in application has been (4) with $p = q = 1$. When $g_j(z_{t-j}) = \alpha_j \ln z_{t-j}^2, j = 1, ..., q$, (2) and (3) define the logarithmic GARCH (LGARCH) model that Geweke (1986) and Pantula (1986) proposed. The specification in (4) amounts to $g(z_t)$ being a function of both the magnitude and sign of z_t. This enables h_t to respond asymmetrically to positive and negative values of ε_t, which is believed to be important for example in modelling the behaviour of stock returns. As to the distribution of z_t, we assume it to be symmetric around zero, which implies $\mathsf{E}z_t^3 = 0$. This assumption together with $\mathsf{E}z_t^3 = 0$ guarantees block diagonality of the information matrix of the log-likelihood function. Block diagonality turn allows us to concentrate of the conditional variance function (3) without simultaneously considering (1).

3 Estimation of parameters

Before considering misspecification tests we briefly discuss parameter estimation. If we complete the previous assumptions about z_t by assuming normality, the log-likelihood function of the EGARCH(p, q) model is

$$L_t = c - (1/2)\sum_{t=1}^{T} \ln h_t - (1/2)\sum_{t=1}^{T}(\varepsilon_t^2/h_t) \tag{5}$$

with

$$\ln h_t = \alpha_0 + \sum_{j=1}^{q}\{\alpha_j z_{t-j} + \psi_j(|z_{t-j}| - E\,|z_t|)\} + \sum_{j=1}^{p} \beta_j \ln h_{t-j}. \tag{6}$$

Let $\boldsymbol{\beta} = (\alpha_0, \alpha_1, ..., \alpha_q, \psi_1, ..., \psi_q, \beta_1, ..., \beta_p)'$. Nelson (1991) discussed maximum likelihood estimation under the assumption that the errors have a generalized error distribution, but we do not follow his path here. The first partial derivatives with respect to the EGARCH parameters are

$$\sum_{t=1}^{T} \frac{\partial l_t}{\partial \boldsymbol{\beta}} = (1/2)\sum_{t=1}^{T}(\frac{\varepsilon_t^2}{h_t} - 1)\frac{\partial \ln h_t}{\partial \boldsymbol{\beta}} \tag{7}$$

where

$$\frac{\partial \ln h_t}{\partial \boldsymbol{\beta}} = \mathbf{x}_{\beta t} - (1/2) \sum_{j=1}^{q} \{\alpha_j z_{t-j} + \psi_j |z_{t-j}|\} \frac{\partial \ln h_{t-j}}{\partial \boldsymbol{\beta}} + \sum_{j=1}^{p} \beta_j \frac{\partial \ln h_{t-j}}{\partial \boldsymbol{\beta}} \tag{8}$$

where $\mathbf{x}_{\beta t} = (1, z_{t-1}, ..., z_{t-q}, |z_{t-1}| - E|z_t|, ..., |z_{t-q}| - E|z_t|, \ln h_{t-1}, ..., \ln h_{t-p})'$.

The parameters of (1) with EGARCH errors (6) may be estimated jointly by maximum likelihood. The normality assumption guarantees block diagonality of the information matrix such that the off-diagonal blocks involving partial derivatives with respect to both mean and variance parameters are null matrices. Thus the parameters of the conditional mean defined by (1) can be estimated separately without asymptotic loss of efficiency. This implies that maximum likelihood estimates for the parameters in (6) can be obtained numerically from the first-order conditions defined by setting (7) equal to zero.

Under sufficient regularity conditions, the maximum likelihood estimators can be expected to be consistent and asymptotically normal. It appears, however, that these conditions have not yet been verified in the present situation. Verifying them in the GARCH case has been a demanding task, and things do not appear to be any easier in the case of EGARCH models. In what follows, it is assumed that the maximum likelihood estimators are consistent and asymptotically normal.

It is seen from (8) that parameter estimation implies a number of recursions, and starting-values for parameters are therefore necessary. Nelson (1991) discussed the role of starting-values and concluded that in his simulations the use of other starting-values than the unconditional mean of $\ln h_t$ very rapidly led to values of h_t obtained by starting from the estimate of $\mathsf{E} \ln h_t$.

4 Evaluation of EGARCH models

4.1 Testing against a higher-order EGARCH model

In this section our starting-point is that the parameters of the EGARCH(p,q) model have been estimated by maximum likelihood, assuming that the errors are standard normal and independent. If, in addition to independence, it is only assumed that $\mathsf{E}z_t = 0$, $\mathsf{E}z_t^2 = 1$ and $\mathsf{E}z_t^3 = 0$, the estimators are quasi maximum likelihood estimators. First we consider testing an EGARCH(p,q) model against a higher-order model, either an EGARCH($p+r$,q) or EGARCH(p,$q+r$), $r > 0$. This is analogous to Bollerslev's test of GARCH(p,q) against GARCH($p+r$,q) or GARCH(p,$q+r$), $r > 0$. Consider now an augmented version of model (2),

$$\varepsilon_t = z_t h_t^{1/2} g_t^{1/2} \tag{9}$$

where

$$\ln g_t = \sum_{j=1}^{r} \{\alpha_{q+j} z_{t-q-j} + \psi_{q+j}(|z_{t-q-j}| - E|z_t|)\} \tag{10}$$

The null hypothesis $H_0 : (\alpha_j, \psi_j) = (0,0), j = q+1, ..., q+r$. Under this hypothesis, $g_t \equiv 1$, and the model collapses into a EGARCH(p,q). Assume now that the alternative H_1 : at least one $\psi_j \neq 0, j = q+1, ..., q+r$. The log-likelihood function of the model is

$$L_T = c - (1/2)\sum_{t=1}^{T}(\ln h_t + \ln g_t) - (1/2)\sum_{t=1}^{T}[\varepsilon_t^2/(h_t g_t)]. \tag{11}$$

Let $\boldsymbol{\beta}_r = (\alpha_{q+1}, ..., \alpha_{q+r}, \psi_{q+1}, ..., \psi_{q+r})'$. The block of the score vector containing the partial derivatives with respect to $\boldsymbol{\beta}_r$ has the form

$$\frac{\partial L_T}{\partial \boldsymbol{\beta}_r} = (1/2)\sum_{t=1}^{T}(\frac{\varepsilon_t^2}{h_t g_t} - 1)\frac{\partial \ln g_t}{\partial \boldsymbol{\beta}_r} \tag{12}$$

where

$$\frac{\partial \ln g_t}{\partial \boldsymbol{\beta}_r} = \mathbf{x}_{\beta_r t} - (1/2)\sum_{j=1}^{r}\{\alpha_j z_{t-q-j} + \psi_j(|z_{t-q-j}| - E\,|z_t|)\}\frac{\partial \ln g_{t-q-j}}{\partial \boldsymbol{\beta}_r} \tag{13}$$

with $\mathbf{x}_{\beta_r t} = (z_{t-q-1}, |z_{t-q-1}| - E\,|z_t|\,, ..., z_{t-q-r}, |z_{t-q-r}| - E\,|z_t|)'$. Let $\widehat{h}_t$ and $\partial \ln \widehat{h}_t/\partial\boldsymbol{\beta}$ be the conditional variance h_t and $\partial \ln h_t/\partial\boldsymbol{\beta}$, respectively, estimated under H_0, and let

$$\widehat{\mathbf{v}}_t = (|\widehat{\varepsilon}_{t-q-1}|\,/\widehat{h}_{t-q-1}, ..., |\widehat{\varepsilon}_{t-q-r}|\,/\widehat{h}_{t-q-r}, \widehat{\varepsilon}_{t-q-1}/\widehat{h}_{t-q-1}, ..., \widehat{\varepsilon}_{t-q-r}/\widehat{h}_{t-q-r})'.$$

Assume, furthermore, that 1-$\sum_{j=1}^{p}\beta_j L^j$ has its roots outside the unit circle. This, together with the assumption of normality for z_t, guarantees that ε_t has all moments see Nelson (1991) and He, Teräsvirta & Malmsten (2002). Thus the moment conditions required for the asymptotic distribution theory of the LM test statistic are satisfied. The LM test can be carried out in the TR^2 form as follows:

1. Estimate the parameters of the EGARCH(p,q) model and compute the squared standardized residuals $\widehat{\varepsilon}_t^2/\widehat{h}_t - 1$, $t = 1, ..., T$, and the "residual sum of squares" $SSR_0^* = \sum_{t=1}^{T}(\widehat{\varepsilon}_t^2/\widehat{h}_t - 1)^2$.

2. Regress $\widehat{\varepsilon}_t^2/\widehat{h}_t - 1$ on $\partial \ln \widehat{h}_t/\partial\boldsymbol{\beta}$ and $\widehat{\mathbf{v}}_t$ and compute the sum of squared residuals, SSR_1^*.

3. Compute the value of the test statistic

$$LM_{addEGARCH} = T\frac{SSR_0^* - SSR_1^*}{SSR_0^*} \tag{14}$$

   that has an asymptotic χ^2 distribution with $2r$ degrees of freedom under the null hypothesis.

If the normality assumption does not hold, this distribution theory is not valid. Nevertheless, it is possible to robustify the test against non-normal errors following Wooldridge (1991). Assuming that $\mathsf{E}\,|\varepsilon_t|^3 < \infty$, the robust version of the test is carried out as follows:

1. Regress $\widehat{\mathbf{v}}_t$ on $\partial \ln \widehat{h}_t / \partial \boldsymbol{\beta}$, and compute the $(2r \times 1)$ residual vectors $\mathbf{r}_t$, $t = 1, ..., T$.

2. Regress 1 on $(\widehat{\varepsilon}_t^2 / \widehat{h}_t - 1)\mathbf{r}_t$ and compute the residual sum of squares SSR^* from that regression. The test statistic is

$$LM_{addEGARCH-R} = T - SSR^* \tag{15}$$

and has an asymptotic χ^2 distribution with $2r$ degrees of freedom under the null hypothesis.

When $p = q = 0$, the test collapses into a test of no EARCH against EARCH(r). The test against H_1 : "at least one $\boldsymbol{\beta}_j \neq 0, j = q+1, ..., q+r$", is constructed analogously by redefining vector $\widehat{\mathbf{v}}_t$. Note that when $p = q = 0$, the corresponding test would be meaningless.

4.2 Testing parameter constancy

Testing parameter constancy is important in its own right but also because nonconstancy signals an apparent lack of covariance stationarity. Here 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. This test may be viewed as the EGARCH counterpart of the test for parameter constancy against smooth continuous change in parameters for the GARCH model in Lundbergh & Teräsvirta (2002). Lin & Teräsvirta (1994) applied the same idea to testing parameter constancy in the conditional mean. Consider now the augmented model (9) where

$$\ln g_t = (\pi_0 + \sum_{j=1}^{q} \{\pi_{1j} z_{t-j} + \pi_{2j} |z_{t-j}|\} + \sum_{j=1}^{p} \pi_{3j} \ln h_{t-j}) G_n(t; \gamma, \mathbf{c}) \tag{16}$$

with the transition function

$$G_n(t; \gamma, \mathbf{c}) = \left[1 + \exp(-\gamma \prod_{i=1}^{n} (t - c_i))\right]^{-1}, \gamma > 0, c_1 \leq \ ... \ \leq c_n. \tag{17}$$

In (17) γ is a slope parameter, and $\mathbf{c} = (c_1, ..., c_n)$ a location vector. Conditions $\gamma > 0$ and $c_1 \leq ... \leq c_n$ are identifying conditions. When $\gamma = 0$, $G_n(t; \gamma, \mathbf{c}) \equiv 1/2$. Typically in practice, $n = 1$ or $n = 2$. The former choice yields a standard logistic function. When the slope parameter $\gamma \to \infty$, (17) with $n = 1$ becomes a step function whose value equals one for $t > c_1$ and zero otherwise. This special case represents a single structural break in the model at $t = c_1$. When $n = 2$,

(17) is symmetric about $(c_1 + c_2)/2$, and its minimum value, achieved at this point, lies between zero and 1/2. The value of the function approaches unity as $t \to \pm\infty$. When $\gamma \to \infty$, function (17) becomes a "double step" function that obtains value zero for $c_1 \leq t \leq c_2$ and unity otherwise.

In order to consider the testing problem, let $\overline{G}_n = G_n - 1/2$. This transformation simplifies notation in deriving the test but does not effect the generality of the arguments. The smooth transition alternative poses an identification problem. The null hypothesis can be expressed as $\mathrm{H}_0 : \gamma = 0$ in $\overline{G}_n$. It can be seen from (16) and (17) that when the null hypothesis holds $\pi_0, \pi_{1j}, \pi_{2j}, j = 1, ..., q$, and $\pi_{3j}, j = 1, ..., p$, in (16) and $c_1, ..., c_n$ in (17) are unidetified nuisance parameters. The standard asymptotic distribution theory is thus not available in this situation, for a general discussion see Hansen (1996).

We circumvent the identification problem by following Luukkonen, Saikkonen & Teräsvirta (1988), see also Lundbergh & Teräsvirta (2002). This is done by expanding the transition function $\overline{G}_n$ into a first-order Taylor series around $\gamma = 0$, replacing the transition function (17) with this Taylor approximation in (16) and rearranging terms. This results in

$$\ln g_t = \sum_{i=1}^{n} \boldsymbol{\delta}_i' \mathbf{v}_{it} + R \tag{18}$$

where $\boldsymbol{\delta}_i = \gamma \widetilde{\boldsymbol{\delta}}_i$, $\widetilde{\boldsymbol{\delta}}_i \neq \mathbf{0}$, $\mathbf{v}_{it} = t^i \mathbf{x}_{\beta t}$, $i = 1, ..., n$, and R is the remainder. The new null hypothesis based on (18) is equals $\boldsymbol{\delta}_1 = ... = \boldsymbol{\delta}_n = \mathbf{0}$. Note that under $\mathrm{H}_0 : R = 0$ so that the remainder does not affect the asymptotic distribution theory. The test can be carried out in TR^2 form via an auxiliary regression exactly as in the previous section. Vectors $\mathbf{v}_{it}$ now contains the additional variables that appear in the auxiliary regression such that $\widehat{\mathbf{v}}_t = (\mathbf{v}_{1t}', ..., \mathbf{v}_{nt}')'$ with $\widehat{\mathbf{v}}_{it} = t^i \widehat{\mathbf{x}}_{\beta t}$, $i = 1, ..., n$. The test can easily be modified to concern only a subset of parameters. A number of terms in the auxiliary equation now contains trending variables. Nevertheless, applying the results of Lin & Teräsvirta (1994), it can be shown that the asymptotic null distribution even in this case is a chi-squared one. The number of degrees of freedom in the test statistic equals $n(p + q + 1)$. The test can be robustified against non-normality in the same way as the previous one.

It is also possible to construct a test against a single structural break by adapting the test of Chu (1995) to the EGARCH case (see Hansen (1996) for obtaining critical values), but that has not been done here.

5 Testing EGARCH against GARCH

As the GARCH model of Bollerslev (1986) and Taylor (1986) is a very popular alternative to the EGARCH model, it would be useful in practice to also compare the estimated EGARCH model with its GARCH counterpart in order to see if one is to be preferred to the other. In this section we discuss three non-nested tests for testing EGARCH and GARCH models against each other. The question we pose is whether or not the GARCH model characterizes some features in the data that the EGARCH model is unable to capture. The tests can thus been seen as misspecification tests of the EGARCH model against the GARCH model or vice versa, depending on which one of the models is the null model. In the GARCH(p,q) model, the conditional variance is

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}. \tag{19}$$

A sufficient condition for the conditional variance to be positive is $\alpha_0 > 0$, $\alpha_j \geq 0,\ j = 1, ..., q, \beta_j \geq 0, j = 1, ..., p$. The necessary and sufficient conditions for positivity of the conditional variance in higher-order GARCH models are complicated; see Nelson & Cao (1992).

The standard GARCH model has been extended to characterize asymmetric responses to shocks. The GJR-GARCH model (Glosten, Jagannathan & Runkle (1993)) is obtained by adding $\sum_{j=1}^{q} \omega_j I(\varepsilon_{t-j})\varepsilon_{t-j}^2$ to the GARCH specification (19) where $I(\varepsilon_{t-1}) = 1$ if $\varepsilon_{t-1} < 0$, and $I(\varepsilon_{t-1}) = 0$ otherwise. A useful nonlinear version of the GJR-GARCH model is obtained by making the transition between regimes smooth. A smooth transition GARCH (STGARCH) model may be defined as (2) with

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_{1j} \varepsilon_{t-j}^2 + \sum_{j=1}^{q} \alpha_{2j} \varepsilon_{t-j}^2 G_n(\varepsilon_{t-j}; \gamma, \mathbf{c}) + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{20}$$

where ε_{t-j} is the transition variable. When $n = 1$, G_1 is the logistic function that controls the change of the coefficient of ε_{t-j}^2 from α_j to $\alpha_j + \omega_j$ as a function of ε_{t-j}. In that case, letting $\gamma \to \infty$ yields the GJR-GARCH model. For discussions of the STGARCH model, see Hagerud (1997), Gonzalez-Rivera (1998), Anderson, Nam & Vahid (1999), Lanne & Saikkonen (2002) and Lundbergh & Teräsvirta (2002). The EGARCH model does not nest these models, and next we shall present three nonnested tests for testing EGARCH against GARCH. In particular, we are intrested in the case where the alternative is a GJR-GARCH model.

5.1 The encompassing test

In this subsection we consider an LM test suggested in Engle & Ng (1993). It is based on a minimal nesting model; see Mizon & Richard (1986). The idea is to construct model that encompasses both alternatives. Thus, decomposing $\ln h_t$ into two components

$$\ln h_t = \ln k_t + \ln g_t \tag{21}$$

where $k_t = \boldsymbol{\theta}'\mathbf{z}_t$ and $g_t = \exp(\boldsymbol{\phi}'\mathbf{x}_t)$ both k_t and g_t are functions of lags of h_t, $\mathbf{z}_t$ is a $k \times 1$ and $\mathbf{x}_t$ is a $m \times 1$ vector of explanatory variables, and $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ are parameter vectors, yields such a model. Equation (21) shows that the model is another special case of the augmented EGARCH model (9). This model is the smallest model which encompasses both the EGARCH and GARCH models and can be used for testing EGARCH and GARCH models against each other. For example, if $\ln k_t = \alpha^* z_{t-1} + \psi^* |z_{t-1}| + \beta^* \ln h_{t-1}$ and $\ln g_t = \ln(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \omega_1 I(\varepsilon_{t-1})\varepsilon_{t-1}^2 + \beta_1 h_{t-1})$, then the model encompassing the EGARCH(1,1) and GJR-GARCH(1,1) ones is

$$\begin{aligned} \ln h_t &= \alpha^* z_{t-1} + \psi^* |z_{t-1}| + \beta^* \ln h_{t-1} \\ &\quad + \ln(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \omega_1 I(\varepsilon_{t-1})\varepsilon_{t-1}^2 + \beta_1 h_{t-1}) \end{aligned} \tag{22}$$

Setting $\alpha_1 = \omega_1 = \beta_1 = 0$ yields the EGARCH(1,1) model. On the other hand, $\alpha^* = \psi^* = \beta^* = 0$ corresponds to the GJR-GARCH(1,1) model. A test of the latter restrictions can also been seen as a misspecification test of the GJR-GARCH model against the EGARCH model. Considering the former case implies the null hypothesis $\mathrm{H}_0 : \alpha_1 = \omega_1 = \beta_1 = 0$. The relevant block of the score vector evaluated under H_0 has the form

$$\frac{\partial L_T}{\partial \boldsymbol{\beta}_r}|_{\mathrm{H}_0} = (1/2)\sum_{t=1}^{T}(\frac{\widehat{\varepsilon}_t^2}{\widehat{h}_t} - 1)\frac{\partial \ln g_t}{\partial \boldsymbol{\beta}_r}|_{\mathrm{H}_0} \tag{23}$$

where

$$\frac{\partial \ln g_t}{\partial \boldsymbol{\beta}_r}|_{\mathrm{H}_0} = \frac{1}{\widehat{h}_t}(\widehat{\varepsilon}_{t-1}^2, I(\widehat{\varepsilon}_{t-1})\widehat{\varepsilon}_{t-1}^2, \widehat{h}_{t-1})' \tag{24}$$

and $\boldsymbol{\beta}_r = (\alpha_1, \omega, \beta_1)'$. Using previous notation, $\widehat{\mathbf{v}}_t = (\widehat{\varepsilon}_{t-1}^2/\widehat{h}_t, I(\widehat{\varepsilon}_{t-1})\widehat{\varepsilon}_{t-1}^2/\widehat{h}_t, \widehat{h}_{t-1}/\widehat{h}_t)'$ in the auxiliary TR^2 regression. Test statistic (14) has an asymptotic χ^2 distribution with three degrees of freedom when the null hypothesis is valid. The test can be robustified against non-normal errors in the same way as the tests in Section 4. It should be noted that when the null hypothesis is rejected, the rejection is against the encompassing model and not the GJR-GARCH one.

5.2 The Pseudo-Score Test

In this subsection we briefly describe a competing test suggested by Chen & Kuan (2002). It is based on the finite sample counterpart of the pseudo-true score, the limit of the expected value of the score function from the alternative model, where the expectation is taken with respect to the null model. Suppose that the null model is a EGARCH model, whereas the alternative model is a GARCH model. Then the pseudo-true score function is

$$PS_{GARCH} = \lim_{T\to\infty} E_{EGARCH(\boldsymbol{\theta})} \frac{1}{2T}\sum_{t=1}^{T}(\frac{\varepsilon_t^2}{h_t} - 1)\frac{\partial \ln h_t}{\partial \boldsymbol{\beta}} \tag{25}$$

where $E_{EGARCH(\boldsymbol{\theta})}$ denotes the expectation taken with respect to the EGARCH model. When the GARCH model is the correct one, expectation (25) equals zero. The test can be constructed by checking if the estimate of (25) is sufficiently close to zero. As shown in Chen & Kuan (2002) the finite sample counterpart of the pseudo-true score equals

$$\widehat{PS}_{GARCH} = \frac{1}{2}\sum_{t=1}^{T}\frac{(\widehat{h}_{EGARCH,t} - \widehat{h}_t)}{\widehat{h}_t}\partial \ln \widehat{h}_t / \partial \boldsymbol{\beta} \tag{26}$$

where $\widehat{h}_{EGARCH,t}$ is the estimate of the conditional variance under the EGARCH model. The test statistic is

$$CK = T\widehat{PS}'_{GARCH}\widehat{\boldsymbol{\Omega}}^{-}\widehat{PS}_{GARCH} \tag{27}$$

where $\widehat{\boldsymbol{\Omega}}$ is a consistent estimator of the information matrix $\boldsymbol{\Omega}$, and $\widehat{\boldsymbol{\Omega}}^{-}$ is its generalized inverse. The test statistic is asymptotically distributed as chi-square with r degrees of freedom when the null hypothesis is true, where r is the rank of $\widehat{\boldsymbol{\Omega}}$. Note that $\boldsymbol{\Omega}$ is not of full rank. In our simulations we estimate $\boldsymbol{\Omega}$ using the estimate of $cov(\widehat{PS}_{GARCH})$ given in Chen & Kuan (2000).

5.3 Simulated likelihood ratio statistic

Our remaining test is the one proposed by Lee & Brorsen (1997) and Kim et al. (1998). The latter authors suggested it for testing the GARCH model against the autoregressive stochastic volatility model or vice versa. The test is based on the log likelihood ratio

$$LR = L_T(\widehat{\boldsymbol{\theta}}_{GARCH}) - L_T(\widehat{\boldsymbol{\theta}}_{EGARCH}) \tag{28}$$

where $L_T(\widehat{\boldsymbol{\theta}}_{GARCH})$ and $L_T(\widehat{\boldsymbol{\theta}}_{EGARCH})$ are the maximized log-likelihood function under the GARCH model and under the EGARCH model, respectively. The asymptotic distribution of LR under the hypothesis that the EGARCH model is the true model or under the hypothesis that the GARCH model is the true one is unknown and an empirical distribution is constructed by simulation. Under the assumption that the EGARCH model is true and that its parameter vector is $\widehat{\boldsymbol{\theta}}_{EGARCH}$, we generate N time series from the "true" model. For each simulated series we estimate the parameters of the GARCH and EGARCH models and record the value of LR^i, $i = 1, ..., N$. The resulting values of LR^i are a sample from the exact distribution of LR under the EGARCH model. This gives us the critical value LR_α of the test to which LR is compared. If $LR > LR_\alpha$ the null hypothesis is rejected. For a general discussion of Monte Carlo tests of this type; see Ripley (1987).

6 Simulation experiment

The above distribution theory is asymptotic, and 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)

$$\begin{aligned} y_t &= \varepsilon_t \\ \varepsilon_t &= z_t h_t^{1/2} \end{aligned} \tag{29}$$

where the definition of the conditional variance h_t depends on the test statistic to be simulated. Under the null hypothesis h_t is the conditional variance of the standard EGARCH(1,1) model (4) with $p = q = 1$. The random numbers, z_t, have been generated by the random number generator in GAUSS 3.2. The distribution for the random numbers sampled is either standard normal or a standardized (unit variance) generalized error $GED(\nu)$ distribution, see Nelson (1991). In the latter case, parameter ν is chosen such that the kurtosis $\mathsf{E}z_t^4 = 5$. The first 1000 observations of each generated series have been discarded to avoid initialization effects. Size experiments are performed with series of 1000 and 3000 observations. The empirical power of the tests is investigated using series of 1000 observations. We use 1000 replications in each experiment. Both the normality-based and the robust version of each test are considered.

6.1 Evaluation of EGARCH models

6.1.1 Testing against a higher-order EGARCH model

First, we consider the test against a higher-order EGARCH model. We define a DGP such that the conditional variance follows a symmetric EGARCH(2,2) process. Thus,

$$\begin{aligned} \ln h_t &= -0.00127 + 0.11605(|z_{t-1}| - E\,|z_{t-1}|)\} + 0.95 \ln h_{t-1} + \\ &\quad + \psi_2(|z_{t-2}| - E\,|z_{t-2}|) + \beta_2 \ln h_{t-2}. \end{aligned} \tag{30}$$

The values of ψ_2 and β_2 are chosen being varied in simulations. The moment structure of the EGARCH(p,q) model has been worked out in He (2000). For $\psi_2 = \beta_2 = 0$ the DGP reduces to a symmetric EGARCH(1,1) model. In the simulations these tests were all computed with a singel parameter in the alternative. That is, for different values of ψ_2 we choose $\widehat{\mathbf{v}}_t = (|\widehat{\varepsilon}_{t-2}| / \widehat{h}_{t-2})$. For different values of β_2 we choose $\widehat{\mathbf{v}}_t = \ln \widehat{h}_{t-2}$. The asymptotic null distribution is thus $\chi^2(1)$. The actual rejection frequencies based on the significant level 0.05 under the asymptotic distribution are reported.

The results of both size and power simulations can be found in Table 1. They indicate that both test is well sized for $T = 1000$. When the errors are normal, the nonrobust test is somewhat more powerful than the robust one. When the error distribution is a $GED(5)$ one, the robust test is more powerful than the nonrobust one. A tentative recommendation would be to always use the robustified test unless there is strong evidence of the errors being normally distributed.

6.1.2 Testing parameter constancy

We consider two cases of parameter nonconstancy for the symmetric EGARCH(1,1) model: the DGP is a EGARCH with either (a) a single or (b) double structural break in the intercept. Our test is computed using $n = 1$ in case (a), and $n = 2$ in case (b).

We consider the following symmetric models with a break in the constant term

$$\begin{aligned}
\ln h_t &= -0.00127 + 0.11605(|z_{t-1}| - E\,|z_{t-1}|) + 0.95 \ln h_{t-1}, \\
\text{(a) } t &< \eta T,\ \text{(b) } t < \eta_1 T,\ t > \eta_2 T, \\
\ln h_t &= -0.03593 + 0.11605(|z_{t-1}| - E\,|z_{t-1}|) + 0.95 \ln h_{t-1}, \\
\text{(a) } t &\geqslant \eta T,\ \text{(b) } \eta_1 T \leq t \leq \eta_2 T
\end{aligned} \tag{31}$$

where T is the sample size and $0 \leq \eta, \eta_1, \eta_2 \leq 1$. The parameters under the null hypothesis are chosen to mimic one of the sets of parameter values considered in Engle & Ng (1993), see below. In the simulation experiment, the unconditional variance is halved at ηT and $\eta_1 T$. Even if only the intercept changes in (31), we assume that under the alternative, the break affects all three parameters. The asymptotic null distribution is thus $\chi^2(3)$.

The results of both size and power simulations can be found in Table 2. They indicate that both tests are somewhat oversized for $T = 1000$ but well-sized for $T = 3000$. When the errors are normal, the nonrobust test is more powerful than the robust one for both a single and a double structural break in the intercept. When the error distribution is a $GED(5)$ one, the robust test is more powerful than the nonrobust one. A tentative recommendation would be similar to the previous one: use the robustified test unless there is strong evidence favouring normal errors.

6.2 Testing EGARCH against GARCH

In this section we consider the small-sample performance of the nonnested tests of testing EGARCH against GARCH. First we consider symmetric, then asymmetric models.

6.2.1 Symmetric models

We consider six pairs of parameter vectors for the GARCH(1,1) and the symmetric EGARCH(1,1) model. They can be found in Table 3. For GARCH, $\alpha_1 + \beta_1$ is the exponential decay rate of the autocorrelations of squared observations, which has been used as a measure for persistence in volatility. We choose three different values of the persistence. Engle & Ng (1993) used the same values in their simulation experiments. The three parameters, α_0,α_1 and β_1, are selected such that the unconditional variance $\mathsf{E}\varepsilon_t^2$ equals unity but the kurtosis equals either 6 or 12. These parameter values are obtained from the analytic expressions of the second moment, the kurtosis and the autocorrelation function of squared observations of a family of GARCH models with normal or t distributed errors that are available in He & Teräsvirta (1999). This family includes the GJR-GARCH model.

The parameter values for the EGARCH model are chosen to be as comparable with the ones for the GARCH models as possible. Thus, β is set equal to $\alpha_1 + \beta_1$ in the corresponding GARCH model, as β controls the decay of the autocorrelation function of the squared observations in the EGARCH model. Note, however, that while the decay rate of the autocorrelation of ε_t^2 in the GARCH(1,1) model equals $\alpha_1 + \beta_1$, it only approaches β from below with increasing lag length in the EGARCH(1,1) model. Parameters α_0 and ψ are chosen such that the unconditional variance and the kurtosis are the same in both models as well. This can be done using the analytic expressions for the relevant moments of the EGARCH(1,1) model in He et al. (2002). The parameters of the EGARCH model (31) under the null hypothesis are chosen to mimic one of the sets of GARCH parameter values considered in Engle & Ng (1993). They can be found in Table 3. In simulating the LR statistic we use 99 replications to construct the empirical null distribution.

A general result valid for all our simulation experiments is that the size of both the encompassing test and the simulated LR test is close to the nominal size already at $T = 1000$, see Tables 4 and 5. As to the pseudo-score statistic, it is oversized even for $T = 3000$. This is due to the estimated matrix $\widehat{\mathbf{\Omega}}$. In our experiment the rank of the (3×3) matrix $\mathbf{\Omega}$ equals two, but the estimated matrix $\widehat{\mathbf{\Omega}}$ is not seriously ill-conditioned. This causes the test statistic to be oversized. In fact, assuming rank equal to three, that is, using the $\chi^2(3)$ distribution instead of $\chi^2(2)$ as the null distribution, would not be such a bad idea for $T = 1000$. The test would be conservative (undersized), but not overly so.

The case of $\alpha_1 + \beta_1 = 1$ in the GARCH model has received attention in the literature. Engle & Bollerslev (1986) called the model with this restriction the integrated GARCH (IGARCH) model. The behaviour of $\widehat{\alpha}_1 + \widehat{\beta}_1$ when the true model is GARCH(1,1) with $\alpha_1 + \beta_1 < 1$ has also received attention. In that case there is a substantial probability of estimating this persistence parameter to be greater than one when T is small; see Shephard (1996). Figure 1 contains the estimated density of $\widehat{\alpha}_1 + \widehat{\beta}_1$ when the true model is a symmetric EGARCH(1,1). If we generate data from a stationary EGARCH(1,1) model with normal errors and fit a GARCH(1,1) model with normal errors to the observations, there is a large probability of 'finding IGARCH', that is, ending up with $\widehat{\alpha}_1 + \widehat{\beta}_1 \geq 1$. Furthermore, this probability increases with the sample size. Lamoureux & Lastrapes (1990) obtained a similar result when they generated data with a GARCH(1,1) model with a structural break, but here the DGP is a constant-parameter stationary EGARCH model. In simulating the LR statistic we only use the replications with $\widehat{\alpha}_1 + \widehat{\beta}_1 < 1$. We discard the rest and add new ones until there are 1000 replications in each experiment.

Another result valid for all our simulations is that the simulated LR test is more powerful than the encompassing test, see Table 6. Because of the size problems, the power of the pseudo-score test is not comparable. For all tests, the power is higher for models with low than with high persistence.

We use the EGARCH model (31) under the null hypothesis and the corresponding GARCH model in the simulation experiments of the robust version of the tests. They can be found in Table 3. Because of size problems, the behaviour of the pseudo-score test is not investigated. Our results indicate that the nonrobust version of the simulated LR test is undersized when the error distribution is a GED, see Table 7. The robust version of the encompassing test is undersized for $T = 1000$ but well-sized for $T = 3000$, see Table 8.

In Table 9 we report the result of the power simulations. When the errors are normal, the robust tests perform as well as the nonrobust ones. When the error distribution is a GED, the robust tests are always more powerful than the nonrobust ones. The simulated LR test is more powerful than the encompassing test.

6.2.2 Asymmetric models

We now turn to asymmetric models. We add the asymmetric component ϕz_{t-1} to four of the symmetric EGARCH models assuming $\phi = -0.04$. The asymmetry introduced through ϕ tends to increase kurtosis and, at the same time, reduce the the first-order autocorrelation of squared observations. We choose α_0 and ψ such that the unconditional variance equals unity and the kurtosis is 6 or 12 even here. The parameter values of the asymmetric EGARCH models are reported in Table 3. We use the reduction in the first-order autocorrelation of squared observations, in percentage, due to the asymmetry as an additional condition to identify the sets of parameters in the GJR-GARCH model. The corresponding GJR-GARCH models are also found in Table 3.

The results in the asymmetric case are similar to the symmetric case. The probability of finding GJR-IGARCH model in which $\alpha_1+\beta_1+\omega_1/2 = 1$ when the observations have been generated by an EGARCH(1,1) model is quite large, see Figure 2. The size of the encompassing test is close to the nominal size already at $T = 1000$, see Table 10. As to the pseudo-score statistic, it is oversized in the case of asymmetric models for $T = 3000$; see Table 11. In Table 12 we report result of the power simulations. Note that the results appearing in the tables are not based on size-adjusted tests. For both tests, the power is higher for models with low than the ones with high persistence.

A conclusion of our simulations is that the simulated LR is more powerful than the encompassing test. The pseudo-score test in the form applied in this paper cannot be recommended because of the size problems pointed out.

7 Empirical example

In this section we apply our tests to daily return series of the 29 most actively traded stocks in Stockholm stock exchange. The list of stocks appears in Table 16 together with information about the length of the series. The period investigated ends April 24, 2001. The return series are continuously compounded returns calculated from the closing prices obtained from Datastream.

In Table 13 we report results of the test against a higher-order model. There is some evidence of a need for an EGARCH(1,2) model. The p-value of the test lies between 0.01 and 0.05 in 13 cases out of 29, but there is only one occasion in which it does not exceed 0.01. An EGARCH(2,1) model is not a likely alternative. As a whole it seems that the need for higher-order EGARCH models is not very strong. Table 14 containes results of the parameter constancy test. In almost about half of the cases, there is strong evidence of time-varying parameters. It seems that nonconstancy of the intercept is often a strong reason for rejection. This suggests that the unconditional variance of the series changes over time. The dynamic behaviour of the conditional variance may be less prone to change in time.

Turning to choosing between EGARCH and GARCH, Table 15 contains results based on the robust version of the encompassing test and the simulated LR test for testing GJR-GARCH(1,1) and EGARCH(1,1) models against each other. They indicate that both models fit the data more or less equally well. In most cases there is no clear difference between the models. The encompassing test does not reject either model in 16 cases. The simulated LR test does not reject either model in 9 cases. It is rare that both models are rejected simultaneously. For the encompassing test this happens only once. The EGARCH model is rejected more often than the GJR-GARCH model. Because of the size problems, the pseudo-score test is not applied to these series.

The main conclusion of the empirical example is that there is substantial evidence for parameter nonconstancy. Rejections, measured in p-values, are generally weaker for the other tests applied to the estimated models.

8 Conclusions

In this paper we consider misspecification tests for an EGARCH(p,q) model. We derive two new misspecification tests for an EGARCH model. Since both test statistics are asymptotically χ^2-distributed under the null hypothesis, possible misspecification of an EGARCH model can be detected at low computational cost. Because the tests of an EGARCH model against a higher-order EGARCH model and testing parameter constancy are parametric, the alternative may be estimated if the null hypothesis is rejected. This is useful for a model builder who wants to find out possible weakness of estimated specification. It may also give him/her useful ideas of how the model could be further improved. These tests may be viewed as the EGARCH counterpart of the tests for the GARCH model in Lundbergh & Teräsvirta (2002).

Furthermore, we investigate various way of testing the EGARCH model against GARCH ones as another check of model adequacy. Our simulations show that the simulated LR test is more powerful than the encompassing test and that the size of the test may be a problem in applying the pseudo-score test.

Finally, the simulation results indicate that in practice, the robust versions of our tests should be preferred to nonrobust ones. They can be recommended as standard tools when it comes to testing the adequacy of an estimated EGARCH (p,q) model.

References

Anderson, H. M., Nam, K. & Vahid, F. (1999), *Asymmetric Nonlinear Smooth Transition GARCH Models*, In Rothman, P (Ed.), Nonlinear time series analysis of economic and financial data. Kluwer, Boston.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics* **31**, 307–327.

Chen, Y.-T. & Kuan, C.-M. (2000), The pseudo-true score encompassing test for non-nested hypotheses. Working Paper Academia Sinica, Taipei, Taiwan.

Chen, Y.-T. & Kuan, C.-M. (2002), 'The pseudo-true score encompassing test for non-nested hypotheses', *Journal of Econometrics* **106**, 271–295.

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

Engle, R. F. & Bollerslev, T. (1986), 'Modelling the persistence of conditional variances', *Econometric Reviews* **5**, 1–50.

Engle, R. F. & Ng, V. K. (1993), 'Measuring and testing the impact of news on volatility', *Journal of Finance* **48**, 1749–1778.

Geweke, J. (1986), 'Modelling persistence of conditional variances: Comment', *Econometric Reviews* **5**, 57–61.

Glosten, L., Jagannathan, R. & 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.

Gonzalez-Rivera, G. (1998), 'Smooth transition GARCH models', *Studies in Non-linear Dynamics and Econometrics* **3**, 61–78.

Hagerud, G. E. (1997), *A New Non-Linear GARCH Model*, Stockholm: EFI Economic Research Institute.

He, C. (2000), Moments and the autocorrelation structure of the exponential GARCH(p,q) process. Working Paper Series in Economics and Finance, Stockholm School of Economics, No. 359.

He, C. & Teräsvirta, T. (1999), 'Properties of moments of a family of GARCH processes', *Journal of Econometrics* **92**, 173–192.

He, C., Teräsvirta, T. & Malmsten, H. (2002), 'Moment structure of a family of first-order exponential GARCH models', *Econometric Theory* **18**, 868–885.

Kim, S., Shephard, N. & Chib, S. (1998), 'Stochastic volatility: Likelihood inference and comparison with ARCH models', *Review of Economic Studies* **65**, 361–393.

Lamoureux, C. G. & Lastrapes, W. D. (1990), 'Persistence in variance, structural change and the GARCH model', *Journal of Business and Economic Statistics* **8:2**, 225–234.

Lanne, M. & Saikkonen, P. (2002), Nonlinear GARCH models for highly persistent volatility. Discussion Paper 20/2002 Humboldt-Universität zu Berlin.

Lee, J.-H. & Brorsen, B. W. (1997), 'A non-nested test of GARCH vs. EGARCH models', *Applied Economics Letters* **4**, 765–768.

Li, W. K. & 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. & Teräsvirta, T. (1994), 'Testing the constancy of regression parameters against continous structural change', *Journal of Econometrics* **62**, 211–228.

Ling, S. & McAleer, M. (2000), Testing GARCH versus E-GARCH, *in* C. W-S, W. K. Li & H. Tong, eds, 'Statistics and Finance: An Interface', Imperial College Press, London, pp. 226–242.

Lundbergh, S. & Teräsvirta, T. (2002), 'Evaluating GARCH models', *Journal of Econometrics* **110**, 417–435.

Luukkonen, R., Saikkonen, P. & Teräsvirta, T. (1988), 'Testing linearity against smooth transition autoregressive models', *Biometrika* **75**, 491–499.

Mizon, G. E. & Richard, J. F. (1986), 'The encompassing principle and its application to testing non-nested hypotheses', *Econometrica* **54**, 657–678.

Nelson, D. B. (1991), 'Conditional heteroskedasticity in asset returns: A new approach', *Econometrica* **59**, 347–370.

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

Pantula, S. G. (1986), 'Modelling persistence of conditional variances: Comment', *Econometric Reviews* **5**, 71–74.

Ripley, B. D. (1987), *Stochastic Simulation*, Wiley, New York.

Taylor, S. (1986), *Modelling Financial Time Series*, Wiley, New York.

Wooldridge, J. (1991), 'On the application of robust-based diagnostics to model of conditional means and conditional variances', *Journal of Econometrics* **47**, 5–46.

Tables

Table 1: Empirical p-values of the nonrobust and robustified test against higher-order model when the observations have been generated from model (30) using normal and GED(5) errors, for 1000 and 3000 observations, based on 1000 replications. The nominal significance level equals 0.05.

Error distribution			Normal(0,1)		GED(5)	
ψ_2	$-\beta_2$	T	Nonrobust	Robustified	Nonrobust	Robustified
0	0	1000	4.3	4.2	4.6	4.9
0	0	3000	5.2	4.7	5.0	4.7
0.1	0	1000	5.5	5.4	6.9	18.8
0.2	0	1000	13.5	10.6	11.9	25.8
0	0.1	1000	6.2	6.1	5.3	18.3

Table 2: Empirical p-values of the nonrobust and robustified parameter constancy test when the observations have been generated from model (31) using normal and GED(5) errors, for 1000 and 3000 observations, based on 1000 replications. The nominal significance level equals 0.05.

Error distribution				Normal(0,1)		GED(5)	
η	η_1	η_2	T	Nonrobust	Robustified	Nonrobust	Robustified
1	–	–	1000	8.2	7.4	6.8	6.2
1	–	–	3000	6.6	5.3	6.6	5.5
0.25	–	–	1000	55.8	50.8	32.7	36.0
0.5	–	–	1000	80.8	75.9	54.3	59.2
–	0.333	0.667	1000	31.2	23.5	16.2	25.7
–	0.25	0.75	1000	63.4	57.4	32.1	45.2

Table 3: This table reports the DGP's in the simulations. The first six pairs of parameters are the symmetric models in the simulation experiment. The next four are the asymmetric models. The last pair is for the simulation experiments of the robust version of the tests.

	GARCH					EGARCH			
Model	α_0	α_1	β	ω	Model	$-\alpha_0^*$	ψ	β^*	$-\phi$
G1	0.01	0.0705	0.9195	-	E1	0.0033	0.1876	0.99	-
G2	0.01	0.0864	0.9036	-	E2	0.0065	0.2614	0.99	-
G3	0.05	0.1561	0.7939	-	E3	0.0155	0.3974	0.95	-
G4	0.05	0.1912	0.7588	-	E4	0.0299	0.5453	0.95	-
G5	0.2	0.3	0.5	-	E5	0.0565	0.7095	0.8	-
G6	0.2	0.3674	0.4326	-	E6	0.1064	0.9555	0.8	-
GJR1	0.01	0.0422	0.9251	0.0453	asE1	0.0032	0.1731	0.99	0.04
GJR2	0.01	0.0562	0.9112	0.0495	asE2	0.0064	0.2504	0.99	0.04
GJR3	0.05	0.1202	0.7986	0.0624	asE3	0.0154	0.3891	0.95	0.04
GJR4	0.05	0.1515	0.8036	0.0698	asE4	0.0297	0.5383	0.95	0.04
RG	0.05	0.05	0.9	-	RE	0.0013	0.1161	0.95	-

Table 4: Empirical p-values from size simulations of three tests of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Simulated LR		Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%	5%	10%
G1	5.7	11.2	5.0	9.6	6.1	13.0
G2	5.6	10.6	4.8	9.7	6.4	13.0
G3	5.7	11.4	4.4	9.6	8.4	15.1
G4	5.9	11.7	4.9	9.5	8.1	16.1
G5	5.6	11.1	4.8	9.4	7.4	17.1
G6	5.1	10.7	4.5	8.8	8.4	16.4
E1	5.5	11.1	5.4	10.5	6.5	13.1
E2	4.8	9.7	5.7	11.0	7.0	14.2
E3	5.9	12.0	4.8	10.0	9.4	15.4
E4	5.4	10.5	5.0	10.1	9.8	16.0
E5	5.1	9.9	5.3	10.1	11.0	18.3
E6	5.0	9.7	4.9	9.6	11.0	17.7

Table 5: Empirical p-values from size simulations of two tests of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 3000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%
G1	5.2	9.8	5.7	11.4
G2	4.9	10.0	5.2	11.7
G3	5.1	9.7	5.5	11.6
G4	5.3	10.1	6.2	11.7
G5	5.1	10.2	7.4	15.5
G6	4.1	10.0	7.8	15.6
E1	5.2	10.4	6.0	13.0
E2	5.1	10.3	7.7	15.5
E3	4.5	9.6	6.6	13.8
E4	5.8	10.3	7.0	13.9
E5	4.4	9.3	5.9	12.6
E6	4.7	10.0	5.3	13.2

Table 6: Empirical p-values from power simulations of three tests of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Simulated LR		Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%	5%	10%
G1	42.3	53.8	13.2	22.2	21.9	33.2
G2	45.8	57.2	12.4	20.7	26.0	39.3
G3	56.7	68.4	22.2	34.0	64.5	76.2
G4	65.8	74.2	25.0	37.9	89.3	93.8
G5	87.2	97.3	57.3	71.4	21.9	33.2
G6	94.1	99.1	74.7	84.9	26.0	39.3
E1	48.0	61.4	17.4	26.0	20.0	29.8
E2	39.0	51.8	20.7	32.8	24.9	38.6
E3	69.9	80.7	26.1	35.0	44.4	57.8
E4	71.7	79.1	30.3	40.7	61.1	70.8
E5	89.4	99.1	43.2	53.4	64.5	76.2
E6	97.8	99.4	51.0	61.1	89.3	93.8

Table 7: Empirical p-values from size simulations of two tests, nonrobust and robust version, of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05.

	Normal(0,1)				t(7) or GED(5)			
	Engle-Ng		Simulated LR		Engle-Ng		Simulated LR	
True model	Nonr.	Robust	Nonr.	Robust	Nonr.	Robust	Nonr.	Robust
RG	4.8	4.6	5.5	5.9	3.3	3.1	2.3	5.4
RE	4.4	4.5	5.8	6.0	3.6	3.3	2.1	5.9

Table 8: Empirical p-values from size simulations of two tests, nonrobust and robust version, of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 3000 observations, based on 1000 replications. The nominal significance level equals 0.05.

	Normal(0,1)		t(7) or GED(5)	
	Engle-Ng		Engle-Ng	
True model	Nonr.	Robust	Nonr.	Robust
RG	4.9	4.6	4.1	5.0
RE	4.2	4.2	3.2	4.3

Table 9: Empirical p-values from power simulations of two tests, nonrobust and robust version, of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05.

	Normal(0,1)				t(7) or GED(5)			
	Engle-Ng		Simulated LR		Engle-Ng		Simulated LR	
True model	Nonr.	Robust.	Nonr.	Robust.	Nonr.	Robust.	Nonr.	Robust
RG	43.9	37.9	50.1	53.1	19.6	14.8	13.6	40.2
RE	20.2	25.1	45.3	43.2	17.0	23.4	23.8	48.8

Table 10: Empirical p-values from size simulations of two tests of testing EGARCH(1,1) against GJR-GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%
GJR1	5.8	11.4	11.3	19.0
GJR2	5.7	10.8	11.9	19.0
GJR3	5.8	11.3	10.8	19.3
GJR4	6.4	11.7	12.9	21.8
asE1	5.8	11.8	7.8	16.6
asE2	5.7	11.0	7.6	14.9
asE3	5.6	11.2	10.4	18.3
asE4	5.2	10.7	10.5	18.4

Table 11: Empirical p-values from size simulations of two tests of testing EGARCH(1,1) against GJR-GARCH(1,1) and vice versa, 3000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%
GJR1	5.7	10.7	5.8	12.4
GJR2	5.5	10.4	7.7	14.0
GJR3	5.6	10.4	7.4	13.6
GJR4	5.7	10.6	8.4	16.3
asE1	5.6	11.2	4.3	10.0
asE2	5.4	10.6	5.7	12.2
asE3	5.4	10.7	5.7	12.1
asE4	5.0	10.2	5.8	12.4

Table 12: Empirical p-values from power simulations of two tests of testing EGARCH(1,1) against GJR-GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%
GJR1	23.2	34.7	19.8	31.1
GJR2	20.5	32.0	26.0	37.2
GJR3	30.2	45.0	43.1	55.2
GJR4	34.9	50.4	62.5	73.0
asE1	20.1	36.1	11.8	21.1
asE2	24.1	39.4	14.1	22.4
asE3	32.7	44.4	19.1	29.6
asE4	41.3	53.1	22.1	34.2

Table 13: Rejection frequencies of the robust version of the test against a higher-order model.

	$p \leq 0.01$	$0.01 < p \leq 0.05$	$0.05 < p \leq 0.1$	$p > 0.1$
EGARCH(2,1)	0	0	2	27
EGARCH(1,2)	1	12	10	6

Table 14: Rejection frequencies of the robust version of the parameter constancy test with n=2.

	$p \leq 0.001$	$0.001 < p \leq 0.01$	$0.01 < p \leq 0.1$	$p > 0.1$
Intercept parameter	9	3	7	10
All four parameters	12	2	7	8

Table 15: Rejection frequencies of the EGARCH(1,1) and GJR-GARCH(1,1) models from the encompassing test and the simulated LR (between brackets) test for the 29 series. 1 simulation of the LR test do not converge.

		H_0:EGARCH			
		$p \leq 0.01$	$0.01<p\leq 0.05$	$p > 0.05$	$\sum$
H_0:GARCH	$p \leq 0.01$	1(0)	1(0)	2(2)	4(2)
	$0.01<p\leq 0.05$	0(0)	0(1)	0(4)	0(5)
	$p > 0.05$	6(6)	3(6)	16(9)	25(21)
	$\sum$	7(6)	4(7)	18(15)	29(28)

Table 16: This table lists the 29 stocks investigated. The column labeled "T", reports the number of observations.

y	T
ABB	3717
Assa A.	1617
Assi D.	1769
Astra	3591
Atlas C.	2915
Autoliv	1690
Electrolux	4577
Ericsson	4576
FSB	1470
Gambro	2454
Holmen	4568
Industriv.	2061
Investor	4146
Nokia	2907
OMG	2084
Pharmacia	1370
Sandvik	4576
Scania	1268
Securitas	2461
Skandia	4566
SEB	2984
Skanska	4337
SKF	4578
SSAB	2963
Stora	3263
SCA	4576
SHB	2612
Sw. Match	1239
VOLVO	5324

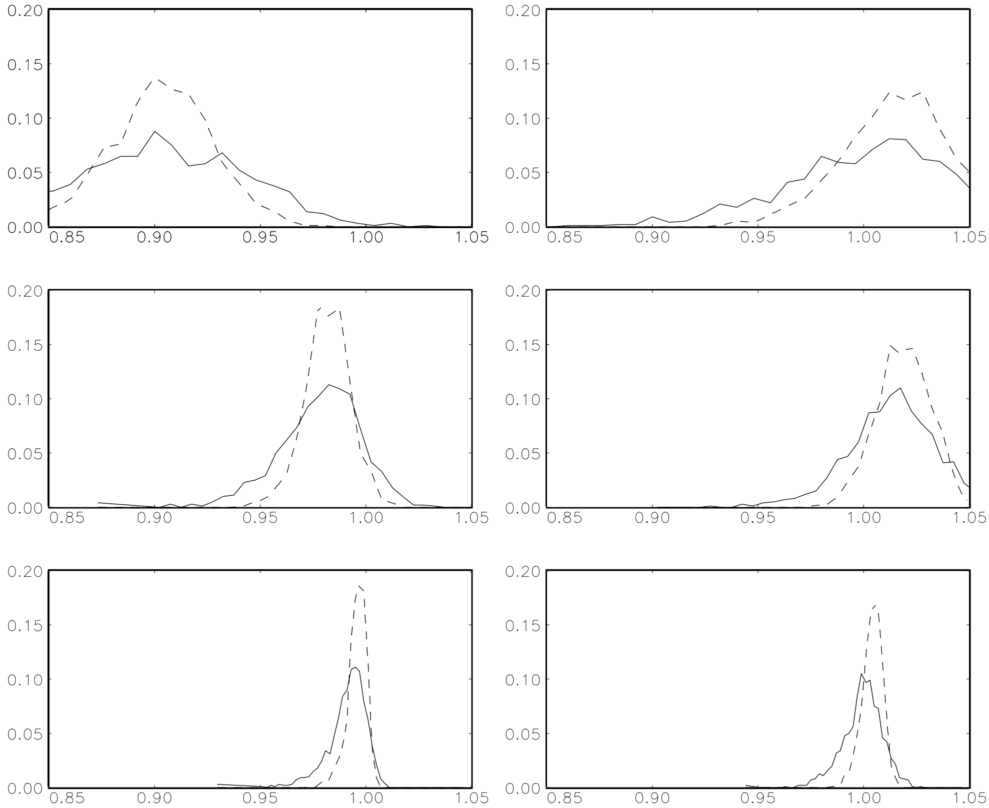

Figure 1: Estimated density of $\widehat{\alpha}_1 + \widehat{\beta}_1$ for the GARCH model when DGP is the symmetric EGARCH model. Upper panel, left: E5, upper panel, right: E6, middle panel, left: E3, middle panel, right: E4, lower panel, left: E1, lower panel, right: E2. Number of observations 1000 (solid) and 3000 (dashed).

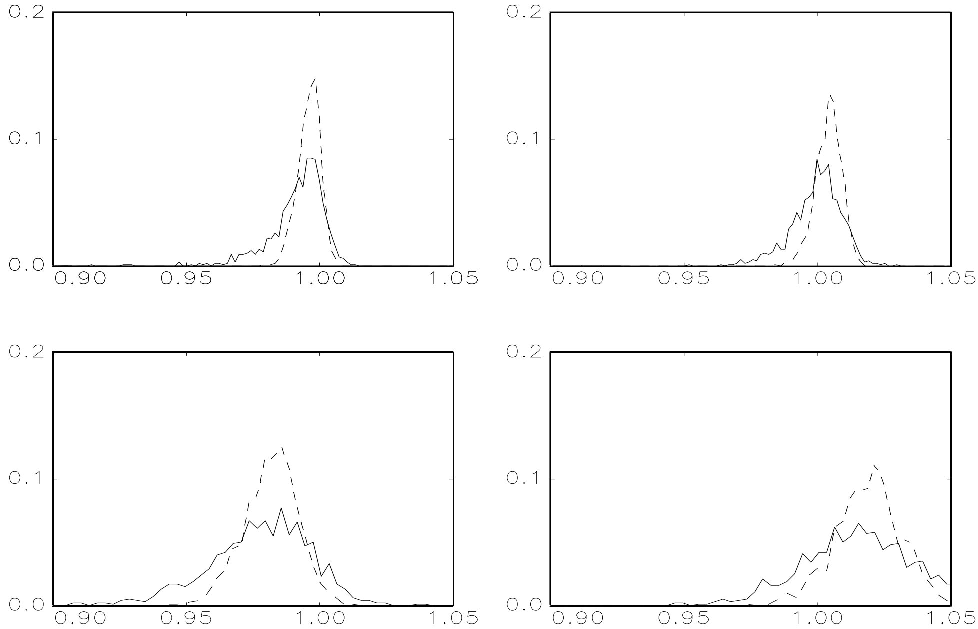

Figure 2: Estimated density of $\widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\omega}/2$ for the GJR-GARCH model when DGP is the EGARCH model. Upper panel, left: asE3, upper panel, right: asE4, lower panel, left: asE1, lower panel, right: asE2. Number of observations 1000 (solid) and 3000 (dashed).