Testing for misspecification in GARCH-type models

Thomas Chuffart*[a,b], Emmanuel Flachaire[a] and Anne Péguin-Feissolle[a]

[a]Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS, France
[b]Université de Lille, France

Abstract

In this article, a misspecification test for GARCH-type models is presented. We propose a Lagrange Multiplier (LM) type test based on a Taylor expansion to distinguish between (G)ARCH models and unknown GARCH-type models. This new test can be seen as a general misspecification test of a large set of GARCH-type univariate models. We investigate the size and the power of this test through Monte Carlo experiments and we compare it to two other standard LM tests, which are more restrictive. We show the usefulness of our test with an illustrative empirical example based on daily exchange rate returns.

Index terms— Conditional heteroskedasticity ; Mnisspecication test ; Nonlinear time series ; Lagrange Multiplier test ; Exchange rates modeling
JEL classification: C12; C14; C22; C50; C52; C58

*Corresponding author: Thomas Chuffart, Centre de la vieille Charité, 2 rue de la Charité, 13002 Marseille.
E-Mail: thomaschuffart@gmail.fr, Tel: +33-4-91140727

1 Introduction

Υ The Autoregressive Conditional Heteroskedastic (ARCH) model developed by Engle (1982) provides a fruitfulness framework to analyze volatility in financial time series. This line of research has quickly become one of the most active research topics in financial econometrics. In 1986, Bollerslev proposed the Generalized Autoregressive Conditional Heteroskedastic (GARCH) model where the volatility is a linear function of past volatility and squared residual past shocks. However, shortly after their introduction, empirical studies revealed that these models do not always adequately fit the data over a long period of time. For example, Lamoureux and Lastrapes (1990) showed that neglected nonlinearities such as structural changes may result in an upward bias in GARCH estimates of persistence in conditional variance. To circumvent this problem, researchers introduced nonlinearity in the specification of conditional variance. For example, the seminal paper by Engle and Ng (1993) presented the Nonlinear-GARCH model, followed by a rapidly increasing body of literature on theoretical and empirical derivatives of nonlinear GARCH-type modeling (see, for example, Teräsvirta (2012) and Haas and Paolella (2012)).

Given this range of conditional heteroskedastic models, practitioners studying empirical data need to choose the most appropriate type of model, particularly to forecast conditional volatility. A misspecification may lead to misinterpretations. For example, Chuffart (2015) performs some Monte-Carlo simulations to detect misspecification using information criteria. He shows that they can lead to the choice of a Logistic Smooth Transition GARCH model instead of the Markov Switching GARCH in some cases.

When choosing the best specification of conditional variance, the first possible approach relies on selection criteria or loss functions. For example, Patton

(2011) investigates the robustness of loss functions when they are computed with an imperfect proxy of volatility. The second possible approach is based on inference. Lundbergh and Teräsvirta (2002) and Halunga and Orme (2009) develop Lagrange Multiplier (LM) tests to choose between linear GARCH and other GARCH-type models.

In this article, we follow the second approach. We introduce a new misspecification test where the alternative hypothesis is represented by a general nonlinear model of conditional variance, characterized by a function G of lagged conditional variances and lagged residual shocks. This test has the advantage of being robust to a misspecification of the model under the alternative hypothesis which can be very general. As an illustration, let us consider a GARCH-type model of the form

$$\varepsilon_t = h_t^{1/2}\eta_t, \qquad \eta_t \sim \mathrm{IID}(0,1)$$

where h_t is a positive process (volatility) and η_t an identically and independently distributed random variable with zero mean and unit variance. Moreover, let us assume that the true model is

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_2 h_{t-2}$$

and we consider the test of the following hypotheses:

$$H_0 : h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \quad vs. \quad H_1 : h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 h_{t-1}.$$

It is entirely possible that a test statistic would not reject the null hypothesis ($\gamma_1 = 0$), even though the model under this hypothesis is not the appropriate one. On the contrary, a test statistic based on the following alternative

$$H_1' : h_t = \alpha_0 + G\left(\varepsilon_{t-1}, h_{t-1}, h_{t-2}, \rho\right),$$

where $G(\varepsilon_{t-1}, h_{t-1}, h_{t-2}, \rho)$ is a function depending on a vector of unknown parameters ρ, would likely reject the null hypothesis model. The basic idea of the test we introduce is thus to define the alternative hypothesis by a general model of conditional volatility. The above illustrates why this test is a useful tool to check whether a given parametric specification is appropriate, helping to overcome the conditional variance specification problem.

Our new misspecification test called LM is compared here with the Lundbergh and Teräsvirta (2002) and the Halunga and Orme (2009) nonlinear tests. The three tests are LM-tests and have the advantage to be less time-consuming, only requiring estimation of GARCH model under the null hypothesis. The new test is based on a Taylor expansion of the unknown general function defining the conditional variance around an arbitrary fixed point in the sample space. Taylor expansion is a way to linearize the testing problem by approximating the true relationship. This approach has been used to test, for example, causality (Péguin-Feissolle et al. (2013)), heteroskedasticity (Lebreton and Péguin-Feissolle (2007)) or constancy of conditional correlation in multivariate GARCH-type models (Péguin-Feissolle and Sanhaji (2015)). It presents some fundamental advantages. First, it requires little knowledge of the functional relationship which determines conditional volatility. Secondly, it can be generalized and is easy to implement. Finally, our small-sample experiments show that it works better than existing tests.

The paper is organized as follows. Section 2 details the new misspecification test. Section 3 reports results of a simulation study: we use Monte-Carlo experiments to investigate both test size and power and robustness to nonnormality and jumps. In Section 4, we illustrate the advantages of the test in the modeling of the exchange rate returns of four currencies during the subprime crisis. Section 5 concludes.

2 A general misspecification test

2.1 The null GARCH(p,q) model

Consider $\{y_t\}$ a stochastic process such that:

$$y_t = E\left[y_t|\mathscr{F}_{t-1}\right] + \varepsilon_{0t}, \qquad \text{and} \qquad Var\left[y_t|\mathscr{F}_{t-1}\right] = h_{0t}, \tag{1}$$

for $t = 1, ..., T$ where T is the number of observations; $E[y_t|\mathscr{F}_{t-1}]$ and $Var[y_t|\mathscr{F}_{t-1}]$ are respectively the conditional expectation and the conditional variance of y_t with respect to $\mathscr{F}_{t-1}$, the sigma-field generated by all the information until time $t-1$. Moreover, we assume a general specification for the conditional expectation:

$$E[y_t|\mathscr{F}_{t-1}] = m(x_t, \lambda_0), \tag{2}$$

where m is a function at least twice continuously differentiable with respect to the $k_{\lambda_0} \times 1$ vector of parameters $\lambda_0 \in \Lambda$, for all $k_x \times 1$ vectors x_t of exogenous variables. ε_{0t} is defined as follows:

$$\varepsilon_{0t} = h_{0t}^{1/2}\eta_t, \qquad \eta_t \sim \text{IID}(0,1). \tag{3}$$

The conditional variance is specified as

$$h_{0t} = \alpha_{00} + \sum_{j=1}^{q} \alpha_{0j}\varepsilon_{0t-j}^2 + \sum_{j=1}^{p} \gamma_{0j} h_{0t-j}. \tag{4}$$

The process above is defined for the true set of parameters $\theta_0 = (\lambda_0', \phi_0')'$ with $\phi_0 = (\alpha_{00}, \alpha_{01}, \ldots, \alpha_{0q}, \gamma_{01}, \ldots, \gamma_{0p})'$. We denote the corresponding model with unknown parameters $\theta = (\lambda', \phi')'$. To ensure identifiability of this process, we make the following assumption:

Assumption 1

i the elements of (y_t, x_t) *are strictly stationary and ergodic,*

ii the parameter space Θ *is compact and* θ_0 *lies in the interior of* Θ.

This model has been widely studied in the literature. Ling and McAleer (2003) show that $E(\varepsilon_{0t}^6) < \infty$ is needed to ensure asymptotic normality of the Quasi Maximum Likelihood (QML) estimator. Chan and McAleer (2002), Berkes et al. (2003) and Francq and Zakoian (2007) show that under weaker moment assumptions, the QML estimator of GARCH and ARMA-GARCH models is consistent and asymptotically normal. The quasi-loglikelihood function of such models is given by

$$L_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} l_t(\theta), \quad l_t(\theta) = -\frac{1}{2}\left[ln(h_t) + \frac{\varepsilon_t^2}{h_t}\right]. \tag{5}$$

Given Assumption 1 and the hypotheses formed on $m(x_t, \lambda)$ in (2), the score vector and the information matrix, respectively defined as $\frac{\partial l_t(\theta)}{\partial \theta}$ and $I(\theta) = -E[\frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'}]$, are both finite and $I(\theta)$ is positive definite.

2.2 Presentation of the test

In this paper, we are interested in testing the null hypothesis characterized by a standard GARCH(p,q) model presented in the previous section: $H_0 : h_t = h_{0t}$ where h_{0t} is given in (4). The alternative hypothesis of the test is assumed to be a general conditional heteroskedastic model where conditional variance is defined by

$$H_1 : \quad h_t = \alpha_0 + G(\varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_{t-q_a}, h_{t-1}, h_{t-2}, ..., h_{t-p_a}, \rho), \tag{6}$$

where ρ is a $k_\rho \times 1$ parameter vector and h_t satisfies the usual regularity conditions. The lags under the alternative, respectively q_a and p_a, can be different from the lags under the null. The functional form of G is unknown; we assume that it adequately represents the exact specification of conditional variance h_t and is at least twice continuously differentiable for all ρ everywhere in the sample space.

The test is based on a finite-order Taylor expansion of G. We thus linearize G in (6) by expanding the function into a kth-order Taylor series around an arbitrary fixed point in the sample space. After approximating G, merging terms and reparametrizing in order to make the GARCH(p, q) model recognizable from the first three terms, we obtain:

$$
\begin{aligned}
h_t = \quad & \alpha_0 + \textstyle\sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \gamma_j h_{t-j} \\
& + \textstyle\sum_{j=q+1}^{q_a} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=p+1}^{p_a} \gamma_j h_{t-j} + \sum_{j=1}^{q_a} \psi_j \varepsilon_{t-j} \\
& + \textstyle\sum_{j_1=1}^{q_a-1} \sum_{j_2=j_1+1}^{q_a} \phi_{j_1 j_2} \varepsilon_{t-j_1} \varepsilon_{t-j_2} \\
& + \textstyle\sum_{j_1=1}^{q_a} \sum_{j_2=1}^{p_a} \delta_{j_1 j_2} \varepsilon_{t-j_1} h_{t-j_2} \\
& + \textstyle\sum_{j_1=1}^{p_a} \sum\limits_{j_2=j_1}^{p_a} \gamma_{j_1 j_2} h_{t-j_1} h_{t-j_2} + ... \\
& + \textstyle\sum_{j_1=1}^{q_a} \sum_{j_2=j_1}^{q_a} ... \sum_{j_k=j_{k-1}}^{q_a} \phi_{j_1 ... j_k} \varepsilon_{t-j_1} ... \varepsilon_{t-j_k} + ... \\
& + \textstyle\sum_{j_1=1}^{p_a} \sum_{j_2=j_1}^{p_a} ... \sum_{j_k=j_{k-1}}^{p_a} \gamma_{j_1 ... j_k} h_{t-j_1} ... h_{t-j_k},
\end{aligned} \tag{7}
$$

where $k \geq 2$, $p \leq p_a \leq k$ and $q \leq q_a \leq k$ for notational convenience. Expansion (7) contains all possible combinations of lagged values of ε_t and h_t up to order k. Therefore, the null hypothesis is that all parameters, except α_0, α_j for

$j = 1, ..., q$ and γ_j for $j = 1, .., p$ are equal to 0, i.e.

$$H_0 : \begin{cases} \alpha_j = 0 \quad \text{for } j = q+1, ..., q_a \\ \gamma_j = 0 \quad \text{for } j = p+1, ..., p_a \\ \psi_j = 0 \quad \text{for } j = 1, ..., q_a \\ \phi_{j_1 j_2} = 0 \quad \text{for } j_1 = 1, ..., q_a - 1,\ j_2 = j_1 + 1, ..., q_a \\ \delta_{j_1 j_2} = 0 \quad \text{for } j_1 = 1, ..., q_{,a}\ \ j_2 = 1, ..., p_a \\ \gamma_{j_1 j_2} = 0 \quad \text{for } j_1 = 1, ..., p_a,\ j_2 = j_1, ..., p_a \\ \vdots \\ \gamma_{j_1 \ldots j_k} = 0 \quad \text{for } j_1 = 1, ..., p_a,\ j_2 = j_1, ..., p_a, ...,\ j_k = j_{k-1}, ..., p_a. \end{cases} \tag{8}$$

The number of parameters to be tested under the null hypothesis is:

$$N^* = \sum_{j=1}^{k} \begin{pmatrix} q_a + p_a + j - 1 \\ j \end{pmatrix} - (p+q). \tag{9}$$

The conditional quasi-loglikelihood is given in (5) where h_t is defined in (7), $\varepsilon_t = y_t - m(x_t, \lambda)$ and θ is the vector of all the parameters of the model: $\theta = (\lambda', \phi', B')'$ with B the $N^* \times 1$ vector of the parameters that are equal to zero under H_0. The LM statistic for testing H_0 represented by (8) is

$$LM = \frac{1}{T} \left(\sum_{t=1}^{T} \frac{\partial l_t(\theta)}{\partial \theta} \right)' I(\theta)^{-1} \left(\sum_{t=1}^{T} \frac{\partial l_t(\theta)}{\partial \theta} \right), \tag{10}$$

where $I(\theta)$ is given by $-E[\frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'}]$. LM is evaluated in $\widehat{\theta}_R = (\widehat{\lambda}', \widehat{\phi}', 0')'$ the QML estimator of θ under the null hypothesis. We demonstrate in Appendix 1 that LM can be written as follows (hereafter, for the LM statistic, the lags of the alternative q_a, p_a and order k of the Taylor expansion are given):

$$LM_{(p_a,q_a,k)} = \frac{1}{4}\left[\sum_{t=1}^{T}\left(\frac{\varepsilon_t^2}{h_t}-1\right)v_t\right]' V^{-1}\left[\sum_{t=1}^{T}\left(\frac{\varepsilon_t^2}{h_t}-1\right)v_t\right] \tag{11}$$

with

$$\begin{aligned} V =& \frac{1}{2}\sum_{t=1}^{T} v_t v_t' - \frac{1}{2}\sum_{t=1}^{T}(z_t v_t')'\sum_{t=1}^{T}(z_t z_t')^{-1}\sum_{t=1}^{T}(z_t v_t') + \\ & \sum_{t=1}^{T}(c_t v_t')'\sum_{t=1}^{T}\left(\frac{1}{h_t}f_t f_t' + \frac{1}{2}c_t c_t'\right)^{-1}\sum_{t=1}^{T}(c_t v_t'), \end{aligned}$$

and

$$f_t = \frac{\partial m(x_t,\lambda)}{\partial\lambda}, \qquad c_t = \frac{1}{h_t}\frac{\partial h_t}{\partial\lambda}, \qquad z_t = \frac{1}{h_t}\frac{\partial h_t}{\partial\phi}, \qquad v_t = \frac{1}{h_t}\frac{\partial h_t}{\partial B}. \tag{12}$$

$\partial h_t/\partial\lambda$, $\partial h_t/\partial\phi$ and $\partial h_t/\partial B$ are respectively $k\times 1$, $(p+q+1)\times 1$ and $N^*\times 1$ vectors and are given for the GARCH(1,1) by:

$$\frac{\partial h_t}{\partial\lambda} = -2\alpha_1\sum_{j=1}^{t}\gamma^{j-1}\varepsilon_{t-j}f_{t-j}, \tag{13}$$

$$\frac{\partial h_t}{\partial\phi} = \begin{pmatrix}1\\ \varepsilon_{t-1}^2\\ h_{t-1}\end{pmatrix} + \gamma_1\frac{\partial h_{t-1}}{\partial\phi} = \sum_{i=1}^{t}\gamma_1^{i-1}\begin{pmatrix}1\\ \varepsilon_{t-i}^2\\ h_{t-i}\end{pmatrix}, \tag{14}$$

$$\frac{\partial h_t}{\partial B} = \mathbf{s}_{t-1}' + \gamma_1\frac{\partial h_{t-1}}{\partial B} = \sum_{i=1}^{t}\gamma_1^{i-1}\mathbf{s}_{t-i}' \tag{15}$$

assuming $h_0 = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2$. $\mathbf{s_t}$ is a vector which contains all the derivatives with respect to the B parameter vector in the Taylor expansion (7). All the terms are evaluated in $\widehat{\theta}_R$.

As Lee et al. (1993), Lebreton and Péguin-Feissolle (2007) and Péguin-

Feissolle et al. (2013) point out, some matrices used to build such test statistics may suffer from collinearity problems when k, p_a or q_a are large. Thus we conduct the test using the main principal components of the $N^* \times T$ matrix where each column is $\partial h_t/\partial B$ for $t = 1, \ldots, T$. To avoid components with insufficient explanatory power regarding variation in the matrix, we use the largest principal components that together explain at least 90% of this variation, the number π^* of principal components being determined automatically. The null hypothesis will be that the parameters linked to the main principal components are equal to zero. Therefore, we obtain a new $\pi^* \times T$ matrix that we call $\mathbf{P}$, whose vectors P_t, for $t = 1, \ldots, T$, are the main principal components of the $N^* \times T$ matrix. We therefore consider the following model under the alternative:

$$h_t = \alpha_0 + \sum_{j=1}^{q_a} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p_a} \gamma_j h_{t-j} + \sum_{j=1}^{\pi^*} \delta_j^* P_{tj} \tag{16}$$

and we test $\delta^* = 0$ with $\delta^* = (\delta_1^*, \ldots, \delta_{\pi^*}^*)$. We have in this case $\partial h_t/\partial B = \sum_{i=1}^{t} \gamma_1^{i-1} \mathbf{P}_{t-i}$. Then we compute the LM statistic given by (11).

3 Monte Carlo experiments

In this section, we study the finite sample performance of the new misspecification test compared to the tests of Lundbergh and Teräsvirta (2002) and Halunga and Orme (2009), respectively called $L\&T$ and $H\&0$ hereafter.

3.1 Lundbergh and Teräsvirta (2002) and Halunga and Orme (2009) tests

These two tests based on the LM principle distinguish between linear GARCH models and nonlinear GARCH-type models in the case of univariate models. They have similar null hypotheses but more restrictive alternative hypotheses than the test we propose. Lundbergh and Teräsvirta (2002) consider the standard GARCH(p,q) model under the null hypothesis. The alternative hypothesis is characterized by the following augmented model

$$\varepsilon_t = \eta_t (h_t + g_t)^{\frac{1}{2}}, \tag{17}$$

where h_t is a GARCH(1,1) and g_t is a nonlinear function which depends on parameters such that under the null hypothesis the function g_t is null. They consider the following logistic function:

$$F_n(\varepsilon_t; \psi, c) = \Big(1 + \exp\{-\psi\, \Pi_{l=1}^{n} (\varepsilon_t - c_l)\}\Big)^{-1} \tag{18}$$

where ψ is a strictly positive slope parameter, ε_t the transition variable and $c = (c_1, \ldots, c_n)$ a local vector with $c_1 \leq \cdots \leq c_n$. In practice, $n = 1$ or $n = 2$. If $n = 1$, when $\psi \to \infty$, F_n becomes a step function which is equal to one if $\varepsilon_t > c_1$ and zero otherwise. If $n = 2$, F_n is symmetric around $(c_1 + c_2)/2$ and becomes a double-step function when $\psi \to \infty$. Therefore, the alternative hypothesis can be written as (17) with

$$g_t = \sum_{j=1}^{q} \alpha_{0j} F_n(\varepsilon_{t-j}; \psi, c) + \sum_{j=1}^{q} \alpha_{1j} F_n(\varepsilon_{t-j}; \psi, c) \varepsilon_{t-j}^2. \tag{19}$$

The LM statistic follows a χ^2 distribution under the null hypothesis with $m = (n+1)q$ degrees of freedom.

Halunga and Orme (2009) introduce a first-order asymptotic theory framework to correct Lundbergh and Teräsvirta (2002)'s test. Their approach is based on the fact that if the GARCH specification is the true generating model, it follows from (3) that

$$E\left[(\eta_t^2 - 1) | \mathscr{F}_{t-1}\right] = 0$$

with $\eta_t = \varepsilon_t / h_t^{1/2}$. Therefore, a misspecification test of GARCH models can be constructed as a test of the following moment conditions

$$E\left[(\eta_t^2 - 1)\pi | \mathscr{F}_{t-1}\right] = 0, \tag{20}$$

where π is a measurable function of $\mathscr{F}_{t-1}$. The intuition is that the squared standardized residuals should be serially uncorrelated with any past information if the GARCH model is the generating model. The alternative hypothesis is thus:

$$H_1: \quad h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 h_{t-1} + g(\nu_t; B), \tag{21}$$

where $g(\nu_t; B) = B'\nu_t$ is a nonlinear function of ε_{t-1}, for example the logistic similar as (19), with ν_t being a vector of omitted variables. The null hypothesis is $H_0 : B = 0$ which implies $g_t = 0$. Their statistic is also based on a score principle and follows asymptotically a χ^2 distribution under the null hypothesis.

3.2 Simulation design

We test the null hypothesis of a GARCH(1,1) model, with the standard $L\&T$ and $H\&0$ statistics and the new test statistic. We consider several $LM_{(p_a,q_a,k)}$ statistics (11) where the order k of the Taylor expansion is 2 or 3[1]. The p_a and q_a orders of the GARCH(p,q) under the alternative are 1 or 2. When N^* given in (9) is large, i.e. $k \geq 3$, $p_a \geq 1$ and $q_a \geq 1$, we select only the largest principal components that together explain 90% of the variation in the corresponding matrix, the number of principal components being thus determined automatically ; we note $LM_{P,(p_a,q_a,k)}$ if we apply the principal component analysis (PCA). We generate 2000 more observations than required to eliminate initialization effects. We choose the true value of parameters to start the estimation procedure via

[1]Simulations, not given here but available from the authors upon request, show that $k = 2$ or $k = 3$ is sufficient in practice.

QML and the unconditional variance for h_0 and ε_0^2. The number of replications is $R = 1000$.

To study size, we generate data under a *true* null hypothesis. Therefore, we use the following GARCH(1,1) model:

$$y_t = \varepsilon_t \quad \text{where} \quad \varepsilon_t = h_t^{1/2}\eta_t, \qquad \eta_t \sim \text{NID}(0,1), \tag{22}$$

for $t = 1, \ldots, T$, and

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma h_{t-1}. \tag{23}$$

Size is investigated with different parameter values. To study power, we need to generate data from a *false* null hypothesis. More precisely, we generate data from a GARCH-type model with regime switches (GJR-GARCH) and with Markov switching regimes (MS-GARCH), which departs progressively from a benchmark model corresponding to the GARCH(1,1) as defined in (23) with $\alpha_0 = 0.01$, $\alpha_1 = 0.1$ and $\gamma = 0.7$.

Asymptotic p-values are computed from a χ^2 distribution. Bootstrap p-values are computed as follows:

1. Estimate a GARCH(1,1) from the original sample $\{y_1, \ldots, y_T\}$. Save the parameter estimates $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\gamma}$ and the residuals $\{\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_T\}$. Compute the test statistic τ_0 following (11);

2. Generate a bootstrap sample $\{y_1^\star, \ldots, y_T^\star\}$, where $y_t^\star = h_t^{\star 1/2}\eta_t^\star$ with $h_t^\star = \hat{\alpha}_0 + \hat{\alpha}_1 \varepsilon_{t-1}^{\star 2} + \hat{\gamma} h_{t-1}^\star$ and $\eta_t^\star$ is a random sampling with replacement in $\{\hat{\varepsilon}_1 \hat{h}_1^{-1/2}, \ldots, \hat{\varepsilon}_T \hat{h}_T^{-1/2}\}$;[2]

3. Estimate a GARCH(1,1) from the bootstrap sample. Compute the test statistic $\tau_1^\star$ similar to τ_0;

[2]To simulate the GARCH bootstrap, we use the initial conditions detailed above.

4. Repeat steps 2 to 3 a large number of times in order to obtain N_b bootstrap statistics, $\tau_b^\star$, $b = 1, \dots, N_b$.

A bootstrap p-value is computed by

$$p_{boot} = \frac{1}{N_b} \sum_{b=1}^{N_b} \mathbb{1}(\tau_b^\star > \tau_0)$$

where $\mathbb{1}(.)$ is the indicator function, equal to 1 if its argument is true, and to 0 otherwise. The number N_b of bootstrap replications is 499.[3] We compute the rejection probability, or rejection frequency, as the proportion of bootstrap p-values below a nominal level equal to 0.05.

3.3 Size

We generate data from several GARCH(1,1) models, as defined in (22) and (23):

- Model 1: $h_t = 0.01 + 0.09\,\varepsilon_{t-1}^2 + 0.9\,h_{t-1}$.
- Model 2: $h_t = 0.05 + 0.05\,\varepsilon_{t-1}^2 + 0.9\,h_{t-1}$.
- Model 3: $h_t = 0.2 + 0.05\,\varepsilon_{t-1}^2 + 0.75\,h_{t-1}$.
- Model 4: $h_t = 0.01 + 0.1\,\varepsilon_{t-1}^2 + 0.7\,h_{t-1}$.

The first three models are chosen to replicate the results of Halunga and Orme (2009). Without loss of generality, the unconditional variance of ε_t is equal to one. The model 4 is considered as a benchmark in the power experiments.

Table 1 shows rejection frequencies of asymptotic and bootstrap tests, for the $LM_{(p_a,q_a,k)}$, $L\&T$ and $H\&0$ statistics, at nominal level $\alpha = 0.05$, with sample size $T = 1000$. Our results suggest that bootstrap p-values outperform asymptotic p-values. Therefore bootstrap inference is reliable, since it provides rejection frequencies closer to the nominal level in many cases.

[3] see Davidson and MacKinnon (2000) for the choice of the number of bootstrap replications.

Table 2 shows the usefulness of the PCA. The size distorsion increases with the values of parameters p_a, q_a and k. However, the same test with principal components works well in all cases, and although the bootstrap analysis seems to correct these size distortions too, the results may be biased due to the collinearity in the Taylor expansion. We thus suggest using principal component analysis when $k \geq 3$, $p_a > 1$ and $q_a > 1$.

3.4 Power

The alternative hypothesis of the LM statistic is more general than the alternative of the $L\&T$ and $H\&0$ statistics. These last two statistics should therefore be more powerful than the LM when the true DGP corresponds to the alternative hypotheses of $L\&T$ and $H\&O$ and the opposite otherwise. This is the well-known trade-off between robustness and efficiency in statistical inference.

First, we consider a case where the true DGP is encompassed by the alternative hypothesis of the $L\&T$ and $H\&0$ statistics. The GJR-GARCH model is defined by Glosten et al. (1993) as

$$h_t = \alpha_0 + \Big(\alpha_1 + \omega\, \mathbb{1}(\varepsilon_{t-1} < 0)\Big)\varepsilon_{t-1}^2 + \gamma h_{t-1}, \tag{24}$$

where $\mathbb{1}(.)$ is the indicator function. This model is a special case of the alternative hypothesis defined in (19) of the $L\&T$ and $H\&0$ statistics, where $p = q = 1$, $\alpha_{01} = c_1 = 0$, $n = 1$, $\alpha_{11} = \omega$ and $\psi \to -\infty$ in (18), since the logistic function becomes a double step function like the indicator function in the GJR-GARCH, i.e. $F_1(\varepsilon_{t-1}; \psi, c) = \mathbb{1}(\varepsilon_{t-1} < 0)$. This model is able to capture some asymmetric effects in financial time series. Indeed, empirical evidence has shown that the increase in volatility can be larger when returns are negative than when they are positive;[4] this characteristic is known as the "leverage-effect".

[4] See Black (1976), Ding et al. (1993), Franses and Van Dijk (1996), Loudon et al. (2000).

Data are generated from the GJR-GARCH model (24), with the parameter values of our benchmark model ($\alpha_0 = 0.01, \alpha_1 = 0.1, \gamma = 0.7$) and with varying parameter ω. We tested in the last section the case where $\omega = 0$, i.e. we tested the true null hypothesis of a GARCH(1,1) model (size). In this section we test the case where $\omega \neq 0$, i.e. we test the false null hypothesis of a GARCH(1,1) model (power). A reliable test gains power the further the ω moves away from 0 (and the further the GJR-GARCH model moves away from a GARCH model). Figure 1 shows rejection frequencies of bootstrap tests corresponding to different values of the parameter ω, for the $LM_{(p_a,q_a,k)}$, $L\&T$ and $H\&0$ test statistics, at nominal level $\alpha = 0.05$, with sample size $T = 1000$. From this figure, we can see that the $LM_{P,(2,2,3)}$ performs better than the others.

We now turn to a case where the true DGP is not encompassed by the alternative hypothesis of the $L\&T$ and $H\&0$ statistics. Let us consider a Markov switching model of Haas et al. (2004), or MS-GARCH model, with two regimes:

$$\begin{cases} \text{Regime 1:} \quad h_{1,t} = \alpha_0(1) + \alpha_1(1)\varepsilon_{t-1}^2 + \gamma(1)h_{1,t-1} \\ \text{Regime 2:} \quad h_{2,t} = \alpha_0(2) + \alpha_1(2)\varepsilon_{t-1}^2 + \gamma(2)h_{2,t-1} \end{cases} \tag{25}$$

with probabilities $1 - \pi$ and π of being respectively in Regime 1 and Regime 2. In this model, the volatility is driven by two independent regimes which depend on a variable $\{\Delta_t\} = \{1, 2\}$ which follows a Markov chain. The transition matrix is $P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$ with $p_{ij} = P(\Delta_t = j|\Delta_{t-1} = i)$. Therefore, $\pi = (1 - p_{11})/(2 - p_{11} - p_{22})$. In the simulation, we only adjust p_{22} (from 0.05 to 0.9) when $\pi < 0.5$ and p_{11} otherwise (from 0.9 to 0.05).[5]. Both models, GJR-GARCH and MS-GARCH, have regime switches, deterministic for the first and stochastic for the second. The class of Markov Switching models is studied

[5]Note that $p_{ii} = 1 - p_{ij}$ When we change the value of p_{ii}, p_{jj} is fixed to 0.9 to keep a persistent regime in each case.

under the assumption that there is more than one state in the economy. For example, Maheu and McCurdy (2000) provide empirical evidence that volatility is higher in Bull markets than in Bear markets.

Data are generated from the MS-GARCH model (25), with parameter values of the benchmark model for Regime 1 ($\alpha_0 = 0.01, \alpha_1 = 0.1, \gamma = 0.7$); the parameters for Regime 2 are: $\alpha_0 = 0,001, \alpha_1 = 0.2, \gamma = 0.3$. When $\pi = 0$ or $\pi = 1$, we test the true null hypothesis of a GARCH(1,1) model (size). When π is different from 0 or 1, we test the false null hypothesis of a GARCH(1,1) model (power). A reliable test increasingly rejects the null hypothesis the further π moves away from 0 and 1 (and the further the MS-GARCH model moves away from a GARCH model). Figure 2 shows rejection frequencies of bootstrap tests for different values of π, for the $LM_{(p_a,q_a,k)}$, $L\&T$ and $H\&0$ test statistics, at nominal level $\alpha = 0.05$, with sample size $T = 1000$. From figure 2, we can see that the LM test for high values of parameters ($LM_{P,(2,2,3)}$) is the most powerful. Clearly, the $L\&T$ and $H\&0$ tests fail to detect an MS-GARCH model. Therefore, as expected, the LM test is more powerful than the L&T and $H\&O$ tests when the true DGP is not a model corresponding to the alternative hypothesis of the $L\&T$ and $H\&O$ tests, proving the usefulness of the new test in this case. Indeed, the LM test is based on a general alternative hypothesis given by (6), i.e. an unspecified unknown model that we linearize by a Taylor expansion. Moreover, an MS(k)-GARCH(p,q) process could be represented by a GARCH(p^*,q^*) process with $p^* > p$ and $q^* > q$, following the results of Krolzig (1997) and Cavicchioli (2014). Indeed, Krolzig (1997) derives the ARMA(p^*,q^*) representation of an MS(k)-AR(p) with p^* greater than p. Cavicchioli (2014) develops the same method to show that a K-state switching M-dimensional Autoregressive model AR(p) can be represented by an VARMA(p^*,q^*) with $p^* \leq k + Mp - 1$ and $q^* \leq k + (M-1)p - 1$.

3.5 Robustness to non-normality

In this subsection, we study the behavior of the different tests when the error term follows a Student distribution with five degrees of freedom. Table 3 presents the size in the first column and the power in the others. Non-normality leads returns to exhibit a fat-tailed distribution which is a frequent feature of financial data (see Laurent et al. (2015)). The simulation experiments show that the tests provide rejection frequencies quite close to the nominal level in the size simulation i.e. in the GARCH-t column. The power is studied for two processes, a GJR-GARCH and an MS-GARCH given respectively by:

$$h_t = 0.1 + [0.1 + 0.2I(\varepsilon_{t-1} < 0)]\varepsilon_{t-1}^2 + 0.7h_{t-1} \tag{26}$$

and

$$h_t = \begin{pmatrix} 0.1 \\ 0.01 \end{pmatrix} + \begin{pmatrix} 0.1 \\ 0.2 \end{pmatrix} \varepsilon_{t-1}^2 + \begin{pmatrix} 0.7 & 0 \\ 0 & 0.3 \end{pmatrix} h_{t-1} \tag{27}$$

with a transition matrix $P = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}$, i.e $\pi = 0.5$. The results for the normal and the Student distributions are shown in Table 3. In the two Student distribution cases (GJR-t and MS-t), there is a loss of power even if the results are close to the results obtained when η_t follows a Gaussian distribution (GJR-N, MS-N).

4 Application to daily exchange rate returns

We illustrate the usefulness of the LM test in modeling daily exchange rate returns of four currencies against the US dollar: the Japanese yen (USD-JPY), the Chinese yuan (USD-RMB), the Euro (EUR-USD) and the Chilean Peso

(USD-CLP). This latter choice was based on the fact that Chile's monetary policy differs from those of other Latin American countries. They switched to a floating exchange rate in 1999 and conducted a monetary policy of inflation targeting. The dataset starts before the subprime crisis and ends before the second European public debt crisis,[6] i.e. from January 24, 2006 to July 1, 2013 ($T = 1940$ observations). Modeling exchange rate returns is a real challenge for econometricians, as shown by the extensive literature in this area ; among others, Erdemlioglu et al. (2013) review recent developments in modeling exchange rate volatility and jumps. The volatility of nominal exchange rate returns has many sources. The monetary policies of both countries have a non-negligible impact on volatility, as shown in Bonser-Neal and Tanner (1996), Dominguez (1998), Beine et al. (2003). Galì and Monacelli (2005) investigate the effect of different monetary policies in a small open economy on nominal and real exchange rate volatility. Otherwise, some papers have shown that the price of commodities, for example oil, also has a strong impact on exchange rates (see among others Zhang et al. (2015) and Ferraro et al. (2015)). Here, we contribute to the existing literature on exchange rate returns by giving a strategy to specify the most appropriate type of models. Indeed, based on economic intuitions and stylized facts, we try to achieve a better fit for the conditional volatility of these four exchange rate series.

The series correspond to closing prices of the exchange rates obtained from the Federal Reserve Economic Data. We compute returns by using the usual formula $y_t = 100 \times \log(\frac{P_t}{P_{t-1}})$, where P_t represents the daily closing price at time t, $t = 2, \ldots, T$. Table 4 presents some statistics on the exchange rate returns. As expected, they exhibit a zero mean process with fat tails. This is especially true of the USD-JPY, USD-RMB and USD-CLP exchange rates, which have

[6] The period of the daily time series is chosen for its interesting features, i.e. the financial crisis and the multiple interventions of the central banks during this period.

large excess kurtosis. Results from Table 4 suggest that $y_t = \varepsilon_t$ is probably not normally distributed. We also compute standard tests to detect possible autocorrelation across residuals and squared residuals and the LM-ARCH test of Engle (1982) to detect ARCH effects. To test autocorrelation, we apply a heteroskedasticity consistent serial correlation test ; it does not reject the null hypothesis of no autocorrelation on residuals. Moreover, applied to the squared residuals, it rejects the null hypothesis in all cases. There is thus a clear evidence of ARCH effects in the exchange rate returns. The four exchange rate series and their returns are shown in Figure 3.

Then, we compute the misspecification tests on exchange rate returns in order to determine whether the conditional variance of these series can be represented by a GARCH(1,1) ; results are given in Table 5. The null hypothesis of a GARCH(1,1) model is rejected by all the tests in the case of the Japanese Yen, by the $LM_{P,(2,2,3)}$ for the Chinese Yuan and by $LM_{P,(1,1,3)}$ and $LM_{P,(2,2,3)}$ for the Chilean Peso at 5%. For the Euro, all tests do not reject the null. This illustration shows that tests can lead to different conclusions according to the series studied. Following our results, the conditional volatility process for the USD-JPY cannot be fitted by a GARCH process. The decision for the USD-CLP and USD-RMB are less conclusive, since one or two tests out of five reject the null hypothesis. Finally, the EUR-USD conditional volatility can be fitted by a GARCH process. However, certain features in these results stand out: p-values differ greatly across the USD-RMB and the USD-CLP. For example, the p-value of $H\&0$ for the USD-RMB is equal to 0.537, whereas that of $LM_{P,(2,2,3)}$ is equal to 0.016. In the same vein, for the Chilean Peso, the p-value of $H\&0$ is equal to 0.793 and less than 0.1 for the three LM tests.

To go beyond a simple test of misspecification, we estimate for each series a GARCH, a GJR and an MS-GARCH model. For the first two models, we estimate both Gaussian and Student versions. Since the distribution assumption may lead to inaccurate results, we estimate only a Student version of the MS-GARCH. In this way, we will see whether a GARCH(1,1) process can adequately represent the conditional volatility process for all series, or whether a nonlinear model is more appropriate. Our selection method is based on Information Criteria (IC), and deeper analysis of out-sample forecasting and in-sample properties is required to confirm that the model selected by IC is the most appropriate. Other conditional volatility models would also need to be tested. However, this is beyond the scope of the paper. Our immediate objective is to illustrate the usefulness of all information yielded by the statistical tests.

Tables 6, 7, 8 and 9 give the estimation results for the four series, respectively for the USD-JPY, the USD-RMB, the EUR-USD and the USD-CLP exchange rate returns.

First, for the USD-JPY returns (Table 6), the GJR-t model outperforms all the others, with both AIC and BIC lower for this model. The MS-GARCH model estimation shows two regimes (the first with time varying volatility, the second with constant volatility). GJR-t model selection means that there is a leverage effect such that negative returns (i.e. when the US dollar depreciates against the Yen) have a stronger impact on conditional variance than positive shocks. This results can be counter-intuitive since the foreign exchange market is two-sided: for bilateral exchange rates, positive returns for one currency correspond to negative returns for the other, i.e. we can not disentangle "bad news" and "good news". This implies that exchange rate volatility should have symmetric responses to positive and negative returns. However, we can argue for the presence of asymmetry in the foreign exchange rates market and give two reasons. First, some currencies have greater importance from an economic point of view. Secondly, central bank interventions lead to higher volatility, for example, to depreciate its currency, a central bank would buy its domestic currency and sell USD. Wang and Yang (2009) find a similar result for a longer period between 1996 and 2004 using realized volatility measure. They explain this phenomena by the position of the Bank of Japan. After 2008, there was an appreciation of the JPY, partially caused by the decreasing amount of carry trades. Thus, to compensate, the Bank of Japan sold JPY and bought USD, which may explain partially the volatility asymmetry of the USD-JPY.

Secondly, for the USD-RMB (Table 7), the MS-GARCH model with a Student distribution better fits the data than the other possible models according to the Information Criteria. This result is explained by the monetary policy of the Chinese government during this period. On the July 21, 2005 the People's Bank of China announced a number of measures to reform the Renminbi exchange rate regime. The main measure was that the exchange rate regime would move immediately into a managed floating exchange rate based on market supply and demand. Furthermore, they scrapped the currency peg to the US dollar, shifting to a basket of main currencies to determine the value of the Chinese currency. These measures led the RMB appreciating by about 21% against the dollar. As a result of the 2008 Global Financial Crisis, China temporarily halted the Yuan floating exchange rate at the end of the year, in response to a slump in worldwide demand for Chinese products. The floating exchange rate was reintroduced in June 19, 2010. During the period of suspension of the floating exchange rate, the Chinese monetary authorities introduced a soft peg. The USD-RMB exchange rate had on average been 6.831, with a very low volatility. As shown by Figure 4a, the MS-GARCH model detects this period of low volatility, suggesting its overall ability to predict the USD-RMB exchange rate returns. Of course, another structural change model of conditional volatility could be used, for example Cai et al. (2012) propose a structural conditional mean and volatility with a dummy variable representing the political switch in both equations.

For the EUR-USD (Table 8), GARCH-t and GJR-t perform quite similarly. The conditional volatility is linear according to missspecification tests. Both central banks let their currency float without any intervention. Thus, either an appreciation or a depreciation in the dollar would have the same impact on volatility which is linearly time varying.

Finally, the MS-GARCH Student model is selected according to the AIC to model the USD-CLP (Table 9) exchange rate returns. The smoothed probability of the turbulent regime is given in Figure 4b. It shows that the exchange rate returns were in a low volatility regime period from the beginning of 2009 to mid 2011. By the end of 2008, a dramatic fall in copper prices negatively impacted the Chilean economy, which is very dependent on the price of this commodity price. To counter the global financial crisis and this price drop, the Central Bank of Chile launched a monetary intervention: they purchased a large amount of dollars. Thus, the 2008 intervention by the Central Bank of Chile may be one determinant of this low volatility regime, in contrast to the 2011 intervention which seems not to have any impact on exchange rate returns. However, there are a lot of other determinants which could explain this volatility.

To conclude, conditional volatilities of these four exchange rate returns need to be carefully specified. The GARCH model is adequate for the EUR-USD and nonlinearity has to be introduced in the conditional volatility for the USD-JPY, the USD-RMB and the USD-CLP. However, only our test suggests it for the USD-RMB and the USD-CLP.

5 Conclusion

Dealing with misspecification is always difficult, since this may require the use of more complex models; estimation is thus generally more difficult. In this paper, we introduce a new test procedure to detect misspecification in conditional volatility based on a Taylor expansion of the unknown conditional volatility around a given point in a sample space. A feature of the new test is that the alternative hypothesis is represented by an unknown general relationship. Therefore, rejection of the null hypothesis of a GARCH(1,1) specification does not imply that the data have been generated from a model where the conditional volatility is specified as a known and fully specified function. Our test can thus be considered as a general misspecification test of very different univariate GARCH-type models. Our simulations show that in some cases it outperforms tests built with a more restrictive alternative hypothesis ; in other cases, it still performs well.

References

Beine, M., Laurent, S., and Lecourt, C. (2003). "Official central bank interventions and exchange rate volatility: Evidence from a regime-switching analysis". *European Economic Review* 47.5, pp. 891–911.

Berkes, I., Horváth, L., and Kokoszka, P. (2003). "GARCH processes: structure and estimation". *Bernoulli* 9.2, pp. 201–227.

Black, F. (1976). "Studies of stock price volatility changes". *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section*, pp. 177–181.

Bonser-Neal, C. and Tanner, G. (1996). "Central bank intervention and the volatility of foreign exchange rates: evidence from the options market ".

Journal of International Money and Finance 15.6, pp. 853–878. ISSN: 0261-5606.

Cai, Z., Chen, L., and Fang, Y. (2012). "A New Forecasting Model for USD/CNY Exchange Rate". *Studies in Nonlinear Dynamics & Econometrics* 16.3, pp. 1–20.

Cavicchioli, M. (2014). "Determining The Number Of Regimes In Markov Switching Var And Vma Models". *Journal of Time Series Analysis* 35.2, pp. 173–186.

Chan, F. and McAleer, M. (2002). "Maximum likelihood estimation of STAR and STAR-GARCH models: theory and Monte Carlo evidence". *Journal of Applied Econometrics* 17.5, pp. 509–534.

Chuffart, T. (2015). "Selection criteria in regime switching conditional volatility models". *Econometrics* 3.2, p. 289. ISSN: 2225-1146.

Davidson, R. and MacKinnon, J. (2000). "Bootstrap tests: how many bootstraps?" *Econometric Reviews* 19.1, pp. 55–68.

Ding, Z., Granger, C. W., and Engle, R. F. (1993). "A long memory property of stock market returns and a new model ". *Journal of Empirical Finance* 1.1, pp. 83–106. ISSN: 0927-5398.

Dominguez, K. M. (1998). "Central bank intervention and exchange rate volatility". *Journal of International Money and Finance* 17.1, pp. 161–190.

Engle, R. F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation". *Econometrica* 50.4, pp. 987–1007.

Engle, R. F. and Ng, V. K. (1993). "Measuring and Testing the Impact of News on Volatility". *Journal of Finance* 48.5, pp. 1749–78.

Erdemlioglu, D., Laurent, S., and Neely, C. J. (2013). "Econometric modeling of exchange rate volatility and jumps". *Handbook of Research Methods and Applications in Empirical Finance*. Edward Elgar. Chap. 16, pp. 373–427.

Ferraro, D., Rogoff, K., and Rossi, B. (2015). "Can oil prices forecast exchange rates? An empirical analysis of the relationship between commodity prices and exchange rates". *Journal of International Money and Finance* 54.C, pp. 116–141.

Francq, C. and Zakoian, J.-M. (2007). "Quasi-maximum likelihood estimation in GARCH processes when some coefficients are equal to zero". *Stochastic Processes and their Applications* 117.9, pp. 1265–1284.

Franses, P. H. and Van Dijk, D. (1996). "Forecasting stock market volatility using (non-linear) Garch models". *Journal of Forecasting* 15.3, pp. 229–235. ISSN: 1099-131X.

Galì, J. and Monacelli, T. (2005). "Monetary policy and exchange rate volatility in a small open economy". *Review of Economic Studies* 72.3, pp. 707–734.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks". *Journal of Finance* 48.5, pp. 1779–1801.

Haas, M., Mittnik, S., and Paolella, M. S. (2004). "A New Approach to Markov-Switching GARCH Models". *Journal of Financial Econometrics* 2.4, pp. 493–530.

Haas, M. and Paolella, M. S. (2012). "Mixture and Regime-Switching GARCH Models". *Handbook of Volatility Models and Their Applications*. John Wiley & Sons, Inc., pp. 71–102. ISBN: 9781118272039.

Halunga, A. G. and Orme, C. D. (2009). "First-Order Asymptotic Theory For Parametric Misspecification Tests Of Garch Models". *Econometric Theory* 25.02, pp. 364–410.

Krolzig, H.-M. (1997). *Markov-Switching Vector Autoregressions*. Red. by G. Fandel and W. Trockel. Vol. 454. Lecture Notes in Economics and Mathematical Systems. Berlin, Heidelberg: Springer Berlin Heidelberg.

Lamoureux, C. G. and Lastrapes, W. D. (1990). "Persistence in variance, structural change, and the GARCH model". *Journal of Business & Economic Statistics* 8.2, pp. 225–234.

Laurent, S., Lecourt, C., and Palm, F. C. (2015). "Testing for jumps in conditionally Gaussian ARMA-GARCH models, a robust approach". *Computational Statistics & Data Analysis*, forthcoming. ISSN: 0167-9473.

Lebreton, M. and Péguin-Feissolle, A. (2007). "Robust Tests for Heteroscedasticity in a general Framework". *Annals of Economics and Statistics* 85, pp. 159–187.

Lee, T.-H., White, H., and Granger, C. W. J. (1993). "Testing for neglected nonlinearity in time series models: A comparison of neural network methods and alternative tests". *Journal of Econometrics* 56.3, pp. 269–290.

Ling, S. and McAleer, M. (2003). "Asymptotic Theory For A Vector Arma-Garch Model". *Econometric Theory* 19.02, pp. 280–310.

Loudon, G. F., Watt, W. H., and Yadav, P. K. (2000). "An empirical analysis of alternative parametric ARCH models". *Journal of Applied Econometrics* 15.2, pp. 117–136.

Lundbergh, S. and Teräsvirta, T. (2002). "Evaluating GARCH models". *Journal of Econometrics* 110.2, pp. 417–435.

Maheu, J. M. and McCurdy, T. H. (2000). "Identifying Bull and Bear Markets in Stock Returns". *Journal of Business & Economic Statistics* 18.1, pp. 100–112.

Patton, A. J. (2011). "Volatility forecast comparison using imperfect volatility proxies". *Journal of Econometrics* 160.1, pp. 246–256.

Péguin-Feissolle, A. and Sanhaji, B. (2015). *Testing the Constancy of Conditional Correlations in Multivariate GARCH-type Models (Extended Version with Appendix).* AMSE Working Papers 1516. Aix-Marseille School of Economics, Marseille, France.

Péguin-Feissolle, A., Strikholm, B., and Teräsvirta, T. (2013). "Testing the Granger Noncausality Hypothesis in Stationary Nonlinear Models of Unknown Functional Form". *Communications in Statistics - Simulation and Computation* 42.5, pp. 1063–1087.

Teräsvirta, T. (2012). "Nonlinear Models for Autoregressive Conditional Heteroskedasticity". *Handbook of Volatility Models and Their Applications.* John Wiley & Sons, Inc., pp. 47–69. ISBN: 9781118272039.

Wang, J. and Yang, M. (2009). "Asymmetric volatility in the foreign exchange markets". *Journal of International Financial Markets, Institutions and Money* 19.4, pp. 597–615.

Zhang, H. J., Dufour, J.-M., and Galbraith, J. W. (2015). "Exchange rates and commodity prices: measuring causality at multiple horizons". *Journal of Empirical Finance*, forthcoming. ISSN: 0927-5398.

Tables and figures

	Model 1		Model 2		Model 3		Model 4	
$LM_{(1,1,2)}$	0.030	0.044	0.034	0.044	0.069	0.057	0.039	0.032
$LM_{(1,1,3)}$	0.038	0.046	0.041	0.042	0.088	0.065	0.040	0.041
$LM_{P,(2,2,3)}$	0.035	0.047	0.034	0.034	0.090	0.065	0.039	0.041
$L\&T$	0.036	0.051	0.038	0.037	0.034	0.041	0.039	0.042
$H\&0$	0.045	0.050	0.045	0.042	0.044	0.051	0.043	0.047

Table 1 Size: Rejection frequencies of asymptotic and bootstrap (gray) tests, for testing a GARCH(1,1) specification, at nominal level $\alpha = 0.05$, when the null hypothesis is true. The sample T is 1000. Note that we use principal component analysis for the $LM_{P,(2,2,3)}$.

	$LM_{(2,2,3)}$ (No PCA)	$LM_{P,(2,2,3)}$ (PCA)
No bootstrap	0.254	0.034
Bootstrap	0.028	0.034

Table 2 Size: Rejection frequencies of asymptotic and bootstrap $LM_{(2,2,3)}$ test with and without PCA for testing a GARCH(1,1) specification (Model 2), at nominal level $\alpha = 0.05$. $T = 1000$.

	GARCH-t	GJR-N	GJR-t	MS-N	MS-t
$LM_{(1,1,2)}$	0.048	1.000	0.808	0.042	0.016
$LM_{P,(1,1,3)}$	0.040	1.000	0.872	0.214	0.212
$LM_{P,(2,2,3)}$	0.056	1.000	0.774	0.860	0.564
$L\&T$	0.050	0.918	0.498	0.092	0.074
$H\&0$	0.052	1.000	0.844	0.130	0.120

Table 3 Rejection frequencies of bootstrap tests, for testing a GARCH(1,1) specification, at nominal level $\alpha = 0.05$. The disturbance term η_t is Gaussian for GJR-N and MS-N and follows a Student distribution with 5 degrees of freedom for GARCH-t, GJR-t and MS-t. The sample T is 1000.

	USD-JPY	USD-RMB	EUR-USD	USD-CLP
Mean	-0.008	-0.014	0.004	-0.002
Median	-0.004	-0.003	0.016	0.000
Stand. Dev.	0.709	0.100	0.661	0.683
Skewness	0.132	0.000	0.057	0.531
Kurtosis	8.822	9.644	4.958	8.063
LM-Autocorr(ε_t)	25.99 (0.166)	22.36 (0.32)	16.31 (0.697)	23.29 (0.275)
LM-Autocorr(ε_t^2)	51.21 (0.000)	68.27 (0.000)	70.46 (0.000)	55.14 (0.00)
LM-ARCH	25.45 (0.00)	24.62 (0.00)	40.49 (0.00)	32.72 (0.00)

Table 4 Summary statistics of the daily exchange rate returns from January 23, 2006 to July 1, 2013. The LM-Autocorr and the LM-ARCH tests are computed with twenty lags on the residuals; we report the p-values in parentheses.

	USD-JPY	USD-RMB	EUR-USD	USD-CLP
$LM_{(1,1,2)}$	0.008	0.194	0.469	0.084
$LM_{P,(1,1,3)}$	0.002	0.295	0.082	0.018
$LM_{P,(2,2,3)}$	0.004	0.016	0.156	0.010
L&T	0.002	0.108	0.450	0.231
H&0	0.001	0.537	0.062	0.793

Table 5 P-values of robust bootstrap tests, for testing a GARCH(1,1) specification in daily exchange rates. In gray, p-values below 5%

	GARCH	GARCH-t	GJR	GJR-t	MS-t
α_{01}	0.012 (0.003)	0.006 (0.003)	0.015 (0.004)	0.008 (0.003)	0.005 (0.003)
α_{11}	0.069 (0.013)	0.066 (0.015)	0.024 (0.009)	0.039 (0.012)	0.098 (0.019)
γ_1	0.909 (0.017)	0.924 (0.017)	0.887 (0.016)	0.915 (0.017)	0.904 (0.018)
ω	-	-	-0.122 (0.025)	-0.065 (0.023)	-
α_{02}	-	-	-	-	0.174 (0.09)
α_{12}	-	-	-	-	0.045 (0.038)
γ_2	-	-	-	-	0.450 (0.273)
dof	-	5.26 (0.615)	-	5.55 (0.470)	5.513 (0.669)
P	-	-	-	-	$\begin{bmatrix} 0.998 & 0.003 \\ 0.002 & 0.997 \end{bmatrix}$
AIC	3.853.8	3686.8	3813.0	3677.0	3682.2
BIC	3870.5	3709.1	3835.3	3704.9	3732.3
LLF	-1923.9	-1839.4	-1902.5	-1833.5	-1832.1

Table 6 Estimation results for the USD-JPY exchange rate returns.
Note: the estimated models are considered with normal and Student distributions: "GARCH" and "GARCH-t" for GARCH(1,1), "GJR" and "GJR-t" for GJR and only "MS-t" for the Markov-Switching model. dof is the degrees of freedom of the Student distribution, P is the transition matrix for the Markov-Switching model, LLF is the value of the log-likelihood function. We report the standard deviation between parentheses.

	GARCH	GARCH-t	GJR	GJR-t	MS-t
α_{01}	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
α_{11}	0.058 (0.010)	0.068 (0.010)	0.039 (0.010)	0.039 (0.015)	0.212 (0.075)
γ_1	0.941 (0.009)	0.932 (0.008)	0.957 (0.004)	0.948 (0.009)	0.812 (0.124)
ω	-	-	-0.007 (0.010)	-0.027 (0.017)	-
α_{02}	-	-	-	-	0.002 (0.001)
α_{12}	-	-	-	-	0.010 (0.003)
γ_2	-	-	-	-	0.877 (0.032)
dof	-	4.67 (0.438)	-	4.56 (0.437)	3.014 (0.202)
P	-	-	-	-	$\begin{bmatrix} 0.972 & 0.031 \\ 0.028 & 0.969 \end{bmatrix}$
AIC	-3999.4	-4331.4	-3997.8	-4332.2	-4363.4
BIC	-3.9827	-4309.1	-3975.5	-4304.3	-4313.3
LLF	2002.7	2169.7	2002.9	2171.1	2136.7

Table 7 Estimation results for the USD-RMB exchange rate returns.
Note: see note of Table 6

	GARCH	GARCH-t	GJR	GJR-t	MS-t
α_{01}	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
α_{11}	0.038 (0.006)	0.039 (0.008)	0.021 (0.008)	0.022 (0.010)	0.041 (0.007)
γ_1	0.960 (0.006)	0.958 (0.008)	0.964 (0.006)	0.962 (0.008)	0.957 (0.007)
ω	-	-	-0.025 (0.009)	-0.025 (0.011)	-
α_{02}	-	-	-	-	0.310 (0.036)
α_{12}	-	-	-	-	0.000 (0.078)
γ_2	-	-	-	-	0.000 (0.067)
dof	-	16.27 (5.51)	-	16.94 (6.04)	17.50 (6.50)
P	-	-	-	-	$\begin{bmatrix} 0.997 & 0.011 \\ 0.003 & 0.989 \end{bmatrix}$
AIC	3531.4	3522.0	3527.6	3519.2	3540.4
BIC	3548.1	3544.3	3549.9	3547.1	3590.5
LLF	-1762.7	-1757.0	-1759.8	-1754.6	-1761.2

Table 8 Estimation results for the EUR-USD exchange rate returns.
Note: see note of Table 6

	GARCH	GARCH-t	GJR	GJR-t	MS-t
α_{01}	0.002 (0.001)	0.003 (0.001)	0.002 (0.001)	0.003 (0.001)	0.002 (0.001)
α_{11}	0.058 (0.008)	0.101 (0.015)	0.062 (0.004)	0.112 (0.019)	0.126 (0.024)
γ_1	0.942 (0.008)	0.899 (0.013)	0.940 (0.004)	0.898 (0.013)	0.883 (0.020)
ω	-	-	-0.004 (0.009)	-0.021 (0.021)	-
α_{02}	-	-	-	-	0.084 (0.043)
α_{12}	-	-	-	-	0.078 (0.041)
γ_2	-	-	-	-	0.730 (0.130)
dof	-	5.51 (0.599)	-	5.50 (0.598)	5.581 (0.956)
P	-	-	-	e-	$\begin{bmatrix} 0.997 & 0.004 \\ 0.003 & 0.996 \end{bmatrix}$
AIC	3478.8	3293.8	3492.4	3294.8	3276.2
BIC	3495.5	3316.1	3514.7	3322.7	3326.3
LLF	-1736.4	-1642.9	-1742.2	-1642.4	-1629.1

Table 9 Estimation results for the USD-CLP exchange rate returns.
Note: see note of Table 6

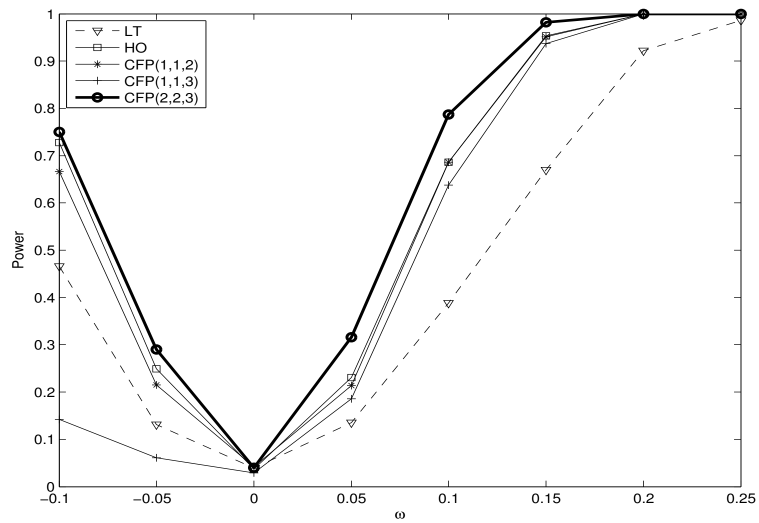


Figure 1: POWER: Rejection frequencies of bootstrap tests, for testing a GARCH(1,1) specification, at nominal level $\alpha = 0.05$, when the null hypothesis is false. Data are generated from a GJR-GARCH model. $T = 1000$. LT, HO, $LM(1,1,2)$, $LM(1,1,3)$ and $LM(2,2,3)$ correspond respectively to statistics $L\&T$, $H\&O$, $LM_{(1,1,2)}$, $LM_{P,(1,1,3)}$ and $LM_{P,(2,2,3)}$.

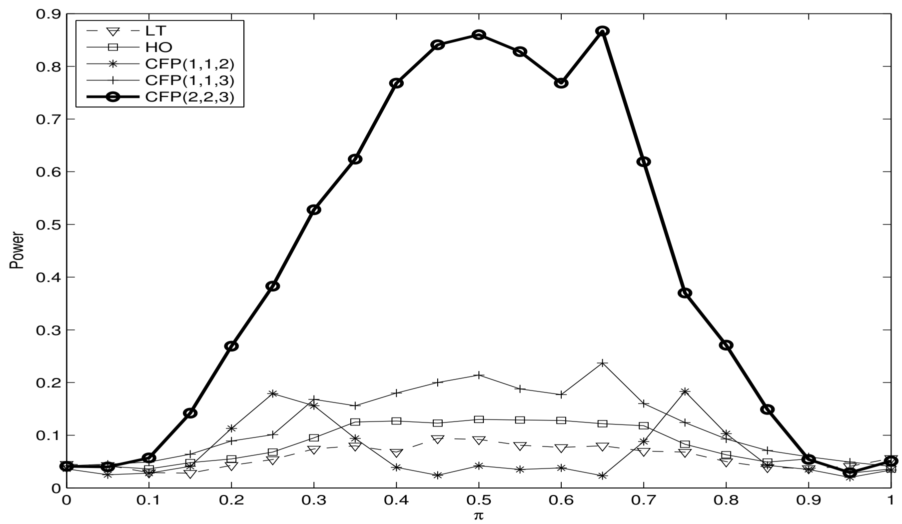
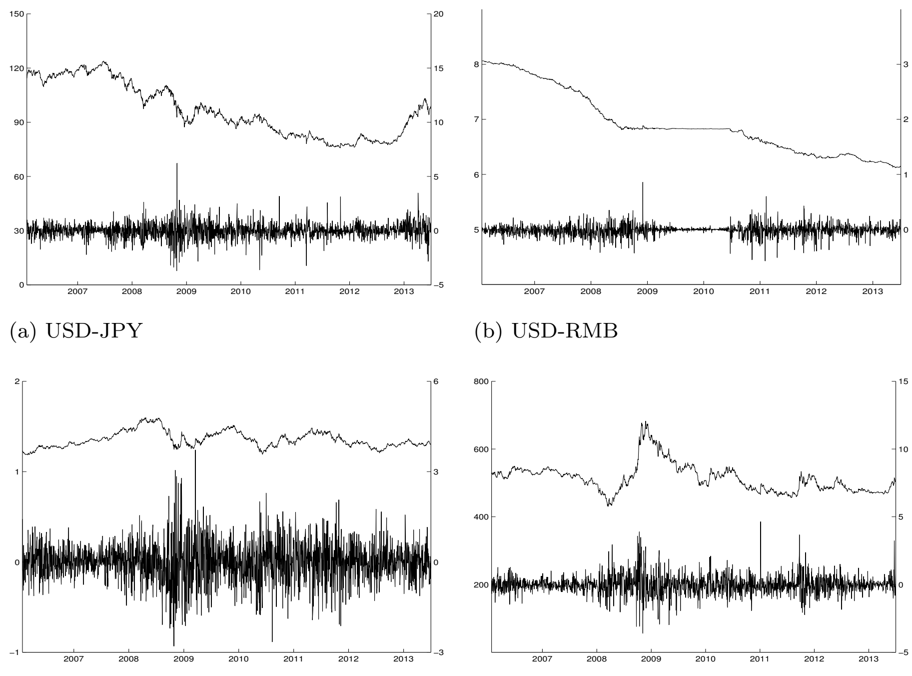


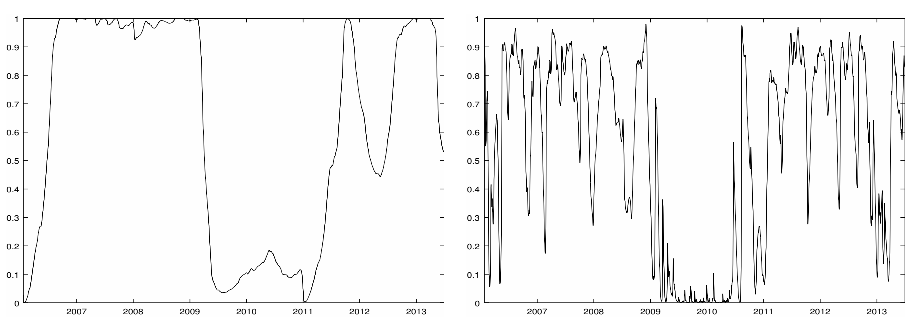
Figure 2: POWER: Rejection frequencies of bootstrap tests for testing a GARCH(1,1) specification, at nominal level $\alpha = 0.05$, when the null hypothesis is false. Data are generated from an MS-GARCH model. $T = 1000$. LT, HO, $LM(1,1,2)$, $LM(1,1,3)$ and $LM(2,2,3)$ correspond respectively to statistics $L\&T$, $H\&O$, $LM_{(1,1,2)}$, $LM_{P,(1,1,3)}$ and $LM_{P,(2,2,3)}$.



(a) USD-JPY (b) USD-RMB

(c) EUR-USD (d) USD-CLP

Figure 3: Daily exchange rates and returns of the fourth series. Left axis labels exchange rates and right axis labels exchange rate returns.



(a) USD-RMB (b) USD-CLP

Figure 4: Smoothed inference probabilities for high volatility regime over time.

Appendix 1: Computation of the LM statistic

$I(\theta)$, given in Equation (10) as $I(\theta) = -E[\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'}]$, can be written:

$$I(\theta) = -E \begin{bmatrix} \frac{\partial^2 l_t(\theta)}{\partial\lambda\partial\lambda'} & \frac{\partial^2 l_t(\theta)}{\partial\lambda\partial\phi'} & \frac{\partial^2 l_t(\theta)}{\partial\lambda\partial B'} \\ (\frac{\partial^2 l_t(\theta)}{\partial\lambda\partial\phi'})' & \frac{\partial^2 l_t(\theta)}{\partial\phi\partial\phi'} & \frac{\partial^2 l_t(\theta)}{\partial\phi\partial B'} \\ (\frac{\partial^2 l_t(\theta)}{\partial\lambda\partial B'})' & (\frac{\partial^2 l_t(\theta)}{\partial\phi\partial B'})' & \frac{\partial^2 l_t(\theta)}{\partial B\partial B'} \end{bmatrix}$$
$$= \begin{bmatrix} I^{11}(\theta) & I^{12}(\theta) \\ I^{12}(\theta)' & I^{22}(\theta) \end{bmatrix}$$

with

$$\begin{cases} I^{11}(\theta) = -E \begin{bmatrix} \frac{\partial^2 l_t(\theta)}{\partial\lambda\partial\lambda'} & \frac{\partial^2 l_t(\theta)}{\partial\lambda\partial\phi'} \\ (\frac{\partial^2 l_t(\theta)}{\partial\lambda\partial\phi'})' & \frac{\partial^2 l_t(\theta)}{\partial\phi\partial\phi'} \end{bmatrix} \\ I^{12}(\theta) = -E \begin{bmatrix} \frac{\partial^2 l_t(\theta)}{\partial\lambda\partial B'} \\ \frac{\partial^2 l_t(\theta)}{\partial\phi\partial B'} \end{bmatrix} \\ I^{22}(\theta) = -E \left[\frac{\partial^2 l_t(\theta)}{\partial B\partial B'}\right] \end{cases} \tag{28}$$

Using $l_t(\theta)$ given by (5), we have

$$\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'} = \frac{1}{h_t}\frac{\partial m_t}{\partial\theta}\frac{\partial m_t}{\partial\theta'} + \frac{1}{2h_t^2}\frac{\partial h_t}{\partial\theta}\frac{\partial h_t}{\partial\theta'}.$$

Thus, the elements in (28) can be derived as

$$I^{11}(\theta) = -E \begin{bmatrix} \frac{1}{h_t}\frac{\partial m_t}{\partial\lambda}\frac{\partial m_t}{\partial\lambda'} + \frac{1}{2h_t^2}\frac{\partial h_t}{\partial\lambda}\frac{\partial h_t}{\partial\lambda'} & \frac{1}{2h_t^2}\frac{\partial h_t}{\partial\lambda}\frac{\partial h_t}{\partial\phi'} \\ \frac{1}{2h_t^2}\frac{\partial h_t}{\partial\phi}\frac{\partial h_t}{\partial\lambda'} & \frac{1}{2h_t^2}\frac{\partial h_t}{\partial\phi}\frac{\partial h_t}{\partial\phi'} \end{bmatrix}$$

$$I^{12}(\theta) = I^{21}(\theta)' = -E \begin{bmatrix} \frac{\partial^2 h_t}{\partial\lambda\partial B'} \\ \frac{\partial^2 h_t}{\partial\phi\partial B'} \end{bmatrix} = -E \begin{bmatrix} \frac{1}{h_t^2}\frac{\partial h_t}{\partial\lambda}\frac{\partial h_t}{\partial B'} \\ \frac{1}{h_t^2}\frac{\partial h_t}{\partial\phi}\frac{\partial h_t}{\partial B'} \end{bmatrix},$$

and

$$I^{22}(\theta) = -E \left[\frac{\partial^2 l_t(\theta)}{\partial B\partial B'}\right] = -E \left[\frac{1}{2h_t^2}\frac{\partial h_t}{\partial B}\frac{\partial h_t}{\partial B'}\right].$$

Under conditional symmetry of the error term, I^{11} becomes diagonal. The first derivative of the log-likelihood with respect to the parameters θ, $\frac{\partial l_t(\theta)}{\partial \theta}$, is defined as follows:

$$\frac{\partial l_t(\theta)}{\partial \theta} = \begin{bmatrix} \frac{\partial l_t(\theta)}{\partial \lambda} \\ \frac{\partial l_t(\theta)}{\partial \phi} \\ \frac{\partial l_t(\theta)}{\partial B} \end{bmatrix}.$$

Under the null hypothesis, $\frac{\partial l_t(\theta)}{\partial \lambda} = 0$ and $\frac{\partial l_t(\theta)}{\partial \phi} = 0$, thus we have:

$$\frac{\partial l_t(\theta)}{\partial \theta}|_{\theta=\widehat{\theta}_R} = \begin{bmatrix} 0 \\ 0 \\ \frac{\partial l_t(\theta)}{\partial B} \end{bmatrix}.$$

We can derive the LM statistic:

$$\begin{aligned} LM &= \frac{1}{T} \begin{bmatrix} 0 \\ 0 \\ \frac{\partial l_t(\theta)}{\partial B} \end{bmatrix}' I(\widehat{\theta}_R)^{-1} \begin{bmatrix} 0 \\ 0 \\ \frac{\partial l_t(\theta)}{\partial B} \end{bmatrix} \\ &= -\frac{1}{T} \begin{bmatrix} 0 \\ 0 \\ \frac{\partial l_t(\theta)}{\partial B} \end{bmatrix}' \left[E \begin{pmatrix} \frac{1}{h_t} \frac{\partial m_t}{\partial \lambda} \frac{\partial m_t}{\partial \lambda'} + \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \lambda} \frac{\partial h_t}{\partial \lambda'} & 0 & \frac{1}{h_t^2} \frac{\partial h_t}{\partial \lambda} \frac{\partial h_t}{\partial B'} \\ 0 & \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \phi} \frac{\partial h_t}{\partial \phi} & \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \phi} \frac{\partial h_t}{\partial B'} \\ \frac{1}{h_t^2} \frac{\partial h_t}{\partial \lambda} \frac{\partial h_t}{\partial B'} & (\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \phi} \frac{\partial h_t}{\partial B'})' & \frac{1}{2h_t^2} \frac{\partial h_t}{\partial B} \frac{\partial h_t}{\partial B'} \end{pmatrix} \right]^{-1} \begin{bmatrix} 0 \\ 0 \\ \frac{\partial l_t(\theta)}{\partial B} \end{bmatrix} \end{aligned}$$

Using the properties of the partitioned matrix, the statistic is equal to:

$$\begin{aligned} LM = \frac{1}{T} \frac{\partial l_t(\theta)}{\partial B}' \Bigg\{ & E\left(\frac{1}{2h_t^2} \frac{\partial h_t}{\partial B} \frac{\partial h_t}{\partial B'}\right) \\ & - \left[E\left(\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \phi} \frac{\partial h_t}{\partial B'}\right)' E\left(\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \phi} \frac{\partial h_t}{\partial \phi'}\right)^{-1} E\left(\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \phi} \frac{\partial h_t}{\partial B'}\right) \right. \\ & \left. + E\left(\frac{1}{h_t^2} \frac{\partial h_t}{\partial \lambda} \frac{\partial h_t}{\partial B'}\right)' E\left(\frac{1}{h_t} \frac{\partial m_t}{\partial \lambda} \frac{\partial m_t}{\partial \lambda'} + \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \lambda} \frac{\partial h_t}{\partial \lambda'}\right)^{-1} E\left(\frac{1}{h_t^2} \frac{\partial h_t}{\partial \lambda} \frac{\partial h_t}{\partial B'}\right) \right] \Bigg\}^{-1} \frac{\partial l_t(\theta)}{\partial B} \end{aligned}$$

with

$$\begin{aligned}
\frac{\partial l_t(\theta)}{\partial B} &= -\frac{1}{2}\sum_{t=1}^{T}\frac{\partial}{\partial B}\left[\ln(h_t)+\frac{\varepsilon_t^2}{h_t}\right] \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left[\frac{\partial}{\partial B}\ln(h_t)+\varepsilon_t^2\frac{\partial}{\partial B}\frac{1}{h_t}\right] \\
&= \frac{1}{2}\sum_{t=1}^{T}\frac{1}{h_t}\left[\frac{\varepsilon_t^2}{h_t}-1\right]\frac{\partial h_t}{\partial B}.
\end{aligned}$$

$\frac{\partial h_t}{\partial \phi}$ and $\frac{\partial h_t}{\partial B}$ are respectively given by equations (14) and (15). If we note $f_t = \frac{\partial m(x_t,\lambda)}{\partial \lambda}$, $c_t = \frac{1}{h_t}\frac{\partial h_t}{\partial \lambda}$, $v_t = \frac{1}{h_t}\frac{\partial h_t}{\partial B}$ and $z_t = \frac{1}{h_t}\frac{\partial h_t}{\partial \phi}$, the statistic will be computed as

$$LM = \frac{1}{4}\left(\sum_{t=1}^{T}\left[\frac{\varepsilon_t^2}{h_t}-1\right]v_t\right)' V^{-1}\left(\sum_{t=1}^{T}\left[\frac{\varepsilon_t^2}{h_t}-1\right]v_t\right)$$

with

$$\begin{aligned}
V =& \frac{1}{2}\sum_{t=1}^{T}v_t v_t' - \frac{1}{2}\sum_{t=1}^{T}(z_t v_t')'\sum_{t=1}^{T}(z_t z_t')^{-1}\sum_{t=1}^{T}(z_t v_t') + \\
&\sum_{t=1}^{T}(c_t v_t')'\sum_{t=1}^{T}\left(\frac{1}{h_t}f_t f_t' + \frac{1}{2}c_t c_t'\right)^{-1}\sum_{t=1}^{T}(c_t v_t').
\end{aligned}$$

The LM statistic has an asymptotic χ^2 distribution with N^* degrees of freedom (because the dimension of the vector B is N^* given by (9).)