



## **Impressum**

Karlsruher Institut für Technologie (KIT)  
Fakultät für Wirtschaftswissenschaften  
Institut für Volkswirtschaftslehre (ECON)

Kaiserstraße 12  
76131 Karlsruhe

KIT – Die Forschungsuniversität in der Helmholtz-Gemeinschaft

Working Paper Series in Economics  
**No. 121**, January 2019

ISSN 2190-9806

[econpapers.wiwi.kit.edu](http://econpapers.wiwi.kit.edu)

---

# Testing for an Omitted Multiplicative Long-Term Component in GARCH Models\*

Christian Conrad<sup>†</sup> and Melanie Schienle<sup>‡</sup>

<sup>†</sup>Heidelberg University, Germany

<sup>‡</sup>Karlsruhe Institute of Technology, Germany

This draft: May 20, 2018

## Abstract

We consider the problem of testing for an omitted multiplicative long-term component in GARCH-type models. Under the alternative there is a two-component model with a short-term GARCH component that fluctuates around a smoothly time-varying long-term component which is driven by the dynamics of an explanatory variable. We suggest a Lagrange Multiplier statistic for testing the null hypothesis that the variable has no explanatory power. We derive the asymptotic theory for our test statistic and investigate its finite sample properties by Monte-Carlo simulation. Our test also covers the mixed-frequency case in which the returns are observed at a higher frequency than the explanatory variable. The usefulness of our procedure is illustrated by empirical applications to S&P 500 return data.

**Keywords:** GARCH-MIDAS, LM test, Long-Term Volatility, Mixed-Frequency Data, Volatility Component Models.

**JEL Classification:** C53, C58, E32, G12

---

\*We are most grateful to the Editor, Todd Clark, an Associate Editor and two anonymous referees for comments that greatly improved our paper. We would like to thank Torben G. Andersen, Richard Baillie, Ian Dew-Becker, Robert F. Engle, Christian Francq, Eric Ghysels, Tilmann Gneiting, Onno Kleen, Fabian Krüger, Enno Mammen, Rasmus S. Pedersen, Peter Schmidt, Karin Stürmer, Robert Taylor, Timo Teräsvirta, Viktor Todorov, Jeffrey Wooldridge and Jean-Michel Zakoïan for helpful comments and suggestions.

<sup>†</sup>Christian Conrad, Department of Economics, Heidelberg University, Bergheimer Strasse 58, 69115 Heidelberg, Germany, E-Mail: christian.conrad@awi.uni-heidelberg.de; Phone: +49/6221/54/3173.

<sup>‡</sup>Melanie Schienle, Department of Economics (ECON), Karlsruhe Institute of Technology, Blücherstrasse 17, 76185 Karlsruhe, Germany, E-Mail: melanie.schienle@kit.edu; Phone: +49/721/608/47535.

# 1 Introduction

The financial crisis of 2007/8 has highlighted the need for a better understanding of the interplay between risks in financial markets and economic conditions. Among others, Engle and Rangel (2008), Christiansen et al. (2012), Paye (2012), Engle et al. (2013) and Conrad and Loch (2015a) provide recent evidence for a counter-cyclical behavior of financial volatility. For example, Conrad and Loch (2015a) show that changes in the secular component of stock market volatility can be anticipated from variables such as the term spread, housing starts or survey expectations on future industrial production. While Christiansen et al. (2012) and Paye (2012) employ predictive regressions, Engle et al. (2013) and Conrad and Loch (2015a) base their analyses on a multiplicative two-component GARCH-MIDAS model. In this model, a short-term GARCH component fluctuates around a smooth long-term component that is driven by macroeconomic conditions. For further multiplicative component GARCH models see, for example, Feng (2004), Engle and Rangel (2008), Amado and Teräsvirta (2013, 2017) and Silvennoinen and Teräsvirta (2016). A recent survey on multiplicative component models and their properties is provided by Amado et al. (2018).

The findings in Engle et al. (2013) and Conrad and Loch (2015a) suggest that one-component GARCH models are misspecified in the sense that they omit a multiplicative component that is driven by an explanatory variable. However, standard procedures for misspecification testing in GARCH models do not cover the case of explanatory variables (see, e.g., Bollerslev, 1986, Lundbergh and Teräsvirta, 2002, or Halunga and Orme, 2009). As most of them also require additive separability of the additional component under the alternative, their adaption to a general multiplicative two-component structure is not straightforward (see Han and Kristensen, 2014, and Francq and Thieu, 2018, for properties and estimation of GARCH models with explanatory variables that enter in an additive fashion).

For this reason, we develop a new misspecification test for GARCH-type models such as GARCH( $p, q$ ), GJR-GARCH or fractionally integrated GARCH. While under the null hypothesis the true model is a pure (one-component) GARCH, under the alternative there is a second multiplicative component. We propose a Lagrange Multiplier ( $LM$ ) statistic, which is based on the parameter estimates under the null, and checks for an omitted long-term component. By construction, the implementation of the  $LM$  test statistic does not

depend on the functional form of the long-term component. For the case of a GARCH(1,1) under the null hypothesis, we provide a detailed derivation of the asymptotic properties of the  $LM$  statistic, whereby we allow for a large class of functions under the alternative. The arguments in the derivation rely on the results for the quasi-maximum likelihood estimator (QMLE) for pure GARCH models in Francq and Zakoïan (2004). However, due to the multiplicative structure under the alternative, many technical derivations are non-standard and require additional steps. In order to obtain the asymptotic distribution of the test statistic, we impose the standard assumptions on the GARCH parameters and the innovation term for the pure GARCH model. As is standard in testing, we also require assumptions on the moments of the explanatory variable as well as on the observed return process. Furthermore, the test statistic is  $\chi^2$  distributed independent of whether the alternative hypothesis is two- or one-sided. This feature of the  $LM$  test has been discussed in Francq and Zakoïan (2009) and does not hold for Wald and Likelihood ratio tests.

Moreover, we propose a regression-based version of our test that can be applied in settings with mixed-frequency data and, thus, can be used as a preliminary check before estimating the GARCH-MIDAS model of Engle et al. (2013) with an explanatory variable that is observed at a lower frequency than the returns. Although this model is frequently used in empirical applications (see, e.g., Asgharian et al., 2013, Conrad and Loch, 2015a, 2015b, Dorion, 2016, Opschoor et al., 2014), there exists no asymptotic theory for the QMLE yet. The most recent theoretical results by Wang and Ghysels (2015) are specific to linear long-term components that are driven by realized volatility and only hold in a restricted parameter space which does not admit our null hypothesis. The regression-based version of the test statistic links our paper to the literature on predictive regressions and, more generally, to regression models involving mixed-frequency data (see Andreou et al., 2010 and Andreou, 2016).

We also discuss in detail the relation between our test statistic, the ‘ARCH nested in GARCH’ test proposed by Lundbergh and Teräsvirta (2002) and a specific variant of the misspecification tests for multiplicative time-varying GARCH models introduced in Amado and Teräsvirta (2017). In particular, the specification of the short-term component in Lundbergh and Teräsvirta (2002) is different from ours and, hence, their model does not encompass the GARCH-MIDAS under the alternative. Moreover, Amado and Teräsvirta

(2017) do not provide an asymptotic theory for the test with exogenous explanatory variables.

Thus, to the best of our knowledge, our test is the first with a complete statistical theory that allows to check GARCH specifications against GARCH-MIDAS-type alternatives. Our work also complements recent research on misspecification testing in the HYGARCH model by Li et al. (2011) and in the Realized GARCH model by Lee and Halunga (2015).

In a Monte-Carlo simulation, we cover the case of a GARCH(1,1) as well as a GJR-GARCH(1,1) under the null and find good size and power properties in finite samples. We also show by simulations that the  $LM$  test is relatively insensitive with respect to a misspecification of the short-term component. Finally, we illustrate the usefulness of our procedure by two empirical applications to S&P 500 return data.

The plan of the paper is as follows. In Section 2, the two-component GARCH model is introduced and the  $LM$  test statistic is derived. This section also contains the main asymptotic results. Section 3 provides some finite sample evidence in a Monte-Carlo study. In Section 4, we illustrate how the test can contribute to modeling S&P 500 return data. Section 5 concludes. All proofs are contained in Appendix A. Additional material can be found in Appendices B-D.

## 2 Model and Test Statistic

In Section 2.1, we first introduce the multiplicative two-component GARCH specification, discuss the null hypothesis of our test and explore the relationship between the two-component model and the GARCH-MIDAS specification. We derive the likelihood function and the test indicator in Section 2.2 and present our main result on the asymptotic distribution of the test statistic in Section 2.3. Section 2.4 provides a comparison with the ‘ARCH nested in GARCH’ test and Section 2.5 extends the  $LM$  test to other GARCH-type models under  $H_0$ . Finally, Section 2.6 covers the mixed-frequency case.

### 2.1 The Two-Component GARCH Model

We define the log-returns as given by

$$\varepsilon_t = \sigma_{0t} Z_t, \tag{1}$$

where  $Z_t$  is independent and identically distributed (i.i.d.) with mean zero and variance equal to one.<sup>1</sup>  $\sigma_{0t}^2$  is measurable with respect to the information set  $\mathcal{F}_{t-1}$  and denotes the conditional variance of the returns. We consider the following multiplicative decomposition of  $\sigma_{0t}^2$  into a GARCH component ('short-term component') and a component that is driven by an explanatory variable:

$$\sigma_{0t}^2 = \bar{h}_{0t}^\infty \tau_{0t}. \quad (2)$$

Following the terminology used in Engle et al. (2013), we refer to the second component as a 'long-term component', because the second component is typically much smoother than the GARCH component. As we will discuss below, the long-term component is driven by the lagged values of an explanatory variable  $x_t$ . The short-term component is specified as a mean-reverting GARCH(1,1):

$$\bar{h}_{0t}^\infty = \omega_0 + \alpha_0 \frac{\varepsilon_{t-1}^2}{\tau_{0,t-1}} + \beta_0 \bar{h}_{0,t-1}^\infty \quad (3)$$

with  $\alpha_0 > 0$ ,  $\beta_0 > 0$  and  $\alpha_0 + \beta_0 < 1$ . As will be discussed in Section 2.5, it is straightforward to allow for other types of GARCH processes (e.g., GJR-GARCH, higher-order GARCH or long-memory GARCH) in the short-term component. For simplicity in the notation and derivations, we focus on the simple GARCH(1,1).

We denote the vector of true parameters in the GARCH component as  $\boldsymbol{\eta}_0 = (\omega_0, \alpha_0, \beta_0)'$  and make the following assumptions about  $\boldsymbol{\eta}_0$  and the innovation  $Z_t$ .

**Assumption 1.**  $\boldsymbol{\eta}_0$  lies in the interior of the compact parameter space  $\Theta = \{\boldsymbol{\eta} = (\omega, \alpha, \beta)' \in \mathbb{R}^3 \mid 0 < \underline{\omega} \leq \omega \leq \bar{\omega} < \infty, 0 < \underline{\alpha} \leq \alpha \leq \bar{\alpha}, 0 < \underline{\beta} \leq \beta \leq \bar{\beta}, \alpha + \beta \leq \bar{\phi} < 1\}$ .

**Assumption 2.** We denote by  $\mathcal{F}_{t-1}$  the  $\sigma$ -field generated by  $\{(\varepsilon_s, x_s); s < t\}$ . As defined in equation (1), let  $Z_t$  be i.i.d. with  $\mathbf{E}[Z_t | \mathcal{F}_{t-1}] = 0$ ,  $\mathbf{E}[Z_t^2 | \mathcal{F}_{t-1}] = 1$  and  $\mathbf{E}[Z_t^4 | \mathcal{F}_{t-1}] = \kappa_Z$ , where  $\kappa_Z$  is a finite constant. Further,  $Z_t^2 | \mathcal{F}_{t-1}$  has a non-degenerate distribution.

Assumptions 1 and 2 imply that  $\sqrt{\bar{h}_{0t}^\infty} Z_t$  is a covariance-stationary process with unconditional variance  $\sigma_0^2$ . Furthermore, they imply that  $\mathbf{E}[\ln(\alpha_0 Z_t^2 + \beta_0)] < 0$  which ensures that  $\sqrt{\bar{h}_{0t}^\infty} Z_t$  is strictly stationary and ergodic (see, e.g., Francq and Zakoian, 2004). Finally, the assumption on the existence of a fourth-order moment of  $Z_t$  is necessary to ensure that the variance of the score vector exists.

---

<sup>1</sup>Throughout the paper we assume that the conditional mean of the returns is zero. For GARCH misspecification testing in the presence of a non-zero conditional mean see Halunga and Orme (2009).

The  $\tau_{0,t}$  component describes smooth movements in the conditional variance as a function of the weighted sum of the  $K$  lagged values of an explanatory variable  $x_t$ :

$$\tau_{0,t} = f(\boldsymbol{\pi}'_0 \mathbf{x}_t), \quad (4)$$

where  $\boldsymbol{\pi}_0 = (\pi_{0,1}, \dots, \pi_{0,K})'$  and  $\mathbf{x}_t = (x_{t-1}, \dots, x_{t-K})'$ . We make the following assumptions on the parameter space  $\boldsymbol{\Pi}$  and the function  $f(\cdot)$ .

**Assumption 3.** *The parameter space  $\boldsymbol{\Pi}$  is a compact subset of  $\mathbb{R}^K$  and  $\boldsymbol{\pi}_0 \in \boldsymbol{\Pi}$ .*

**Assumption 4.** *Let  $\tau_t = f(\boldsymbol{\pi}' \mathbf{x}_t)$  be a positive, non-constant, continuously differentiable function with  $f(0) = 1$  and  $f'_0 = \frac{\partial \tau_t}{\partial \boldsymbol{\pi}' \mathbf{x}_t} \Big|_{\boldsymbol{\pi}=\mathbf{0}} \neq 0$ .*

Note that Assumption 4 allows  $\tau_t$  not only to depend on  $\boldsymbol{\pi}$  but also on  $\boldsymbol{\eta}$ , namely if  $x_t$  depends on the GARCH component (see Section 2.4 for an example).

**Assumption 5.**  *$\mathbf{c}' \mathbf{x}_t$  is not degenerate for any non-zero  $\mathbf{c} \in \mathbb{R}^K$ .*

Assumption 5 ensures that the long-term component is minimal in the sense that no equivalent representation of lower order exists. Thus, Assumption 5 is an identification assumption and implies that  $f(\boldsymbol{\pi}'_0 \mathbf{x}_t)$  is non-constant with probability one unless  $\boldsymbol{\pi}_0 = \mathbf{0}$ . The assumption that  $f(\cdot) > 0$  allows us to consider explanatory variables that take positive as well as negative values. Further, we do not have to require that the  $\pi_{0,k}$  are all positive. That is, in our model the explanatory variable  $\mathbf{x}_t$  can have a positive as well as a negative effect on the volatility (see Assumption 3). The main example that we have in mind for  $f(\cdot)$  is the exponential specification

$$\tau_{0,t} = f(\boldsymbol{\pi}'_0 \mathbf{x}_t) = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t). \quad (5)$$

The exponential specification has been used, among others, in Engle et al. (2013), Opschoor et al. (2014) and Conrad and Loch (2015a).

Using the above notation, we are interested in testing  $H_0 : \boldsymbol{\pi}_0 = \mathbf{0}$  against the two-sided alternative  $H_1 : \boldsymbol{\pi}_0 \neq \mathbf{0}$ .<sup>2</sup> Under  $H_0$ , the long-term component is equal to one and the two-component model reduces to the nested GARCH(1,1) with unconditional variance  $\sigma_0^2 = \omega_0 / (1 - \alpha_0 - \beta_0)$ . The test statistic will be based on a fitted version of  $\varepsilon_t^2 / \bar{h}_{0t}^\infty$ , whereby

---

<sup>2</sup>Later on, we also consider the one-sided alternative  $H_1 : \boldsymbol{\pi}_0 \neq \mathbf{0}, \boldsymbol{\pi}_0 \geq \mathbf{0}$  (see Remark 4 in Section 2.3).



the estimate of  $\bar{h}_{0t}^\infty$  is obtained under  $H_0$ . Note that under  $H_0$ ,  $\mathbf{E}[\varepsilon_t^2/\bar{h}_{0t}^\infty|\mathcal{F}_{t-1}] = 1$ , while under the alternative  $\mathbf{E}[\varepsilon_t^2/\bar{h}_{0t}^\infty|\mathcal{F}_{t-1}] = \tau_{0t}$ . Our test uses this insight to check if the fitted version of  $\varepsilon_t^2/\bar{h}_{0t}^\infty$  is predictable using information included in  $\mathcal{F}_{t-1}$ , in particular  $\mathbf{x}_t$ .

Finally, note that equation (3) can be rewritten as an ARCH( $\infty$ )

$$\bar{h}_{0t}^\infty = \omega_0 + (\alpha_0 Z_{t-1}^2 + \beta_0) \bar{h}_{0,t-1}^\infty = \sum_{i=0}^{\infty} \beta_0^i \left( \omega_0 + \alpha_0 \frac{\varepsilon_{t-1-i}^2}{\tau_{0,t-1-i}} \right)$$

which means that  $\varepsilon_t/\sqrt{\tau_{0t}} = \sqrt{\bar{h}_{0t}^\infty} Z_t$  follows a GARCH(1,1) both under the null and under the alternative.

Our two-component model is closely related to the GARCH-MIDAS model suggested in Engle et al. (2013). In their model, the long-term component is typically of linear or exponential form. For the model with a linear long-term component, Wang and Ghysels (2015) use a rolling window realized variance of the last  $N$  days as the explanatory variable, provide conditions for the strict stationarity of  $\varepsilon_t$  and establish consistency and asymptotic normality of the QMLE. However, the proof of the asymptotic normality of the QMLE crucially relies on an assumption which rules out testing  $H_0 : \boldsymbol{\pi}_0 = \mathbf{0}$  (see Assumption 4.3 in Wang and Ghysels, 2015).

Most importantly, the GARCH-MIDAS specification allows for the possibility that the explanatory variable is observed at a lower frequency, say monthly or quarterly, than the daily returns. In this case, the long-term component varies at the lower-frequency only. Although the mixed-frequency version of the GARCH-MIDAS is highly relevant from an empirical perspective, there is no asymptotic theory for the general model yet. However, in Section 2.6, we propose an extension of our *LM* test to the mixed-frequency case.

## 2.2 Likelihood Function and Partial Derivatives

We denote the processes that can be constructed from the parameter vectors  $\boldsymbol{\eta} = (\omega, \alpha, \beta)'$  and  $\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)'$  given initial observations for  $\varepsilon_t$  and  $x_t$  by  $\bar{h}_t$  and  $\tau_t$ . It is important to distinguish between the observed quasi-likelihood which is based on  $\bar{h}_t = \sum_{j=0}^{t-1} \beta^j (\omega + \alpha \varepsilon_{t-1-j}^2 / \tau_{t-1-j}) + \beta^t \bar{h}_0$  and the unobserved quasi-likelihood function based on  $\bar{h}_t^\infty = \sum_{j=0}^{\infty} \beta^j (\omega + \alpha \varepsilon_{t-1-j}^2 / \tau_{t-1-j})$  which depends on the infinite history of all past observations. The unobserved Gaussian quasi-log-likelihood function (omitting the con-

stant) can be written as  $L_T^\infty(\boldsymbol{\eta}, \boldsymbol{\pi} | \varepsilon_T, x_T, \varepsilon_{T-1}, x_{T-1}, \dots) = \sum_{t=1}^T l_t^\infty$  with

$$l_t^\infty = -\frac{1}{2} \left[ \ln(\bar{h}_t^\infty) + \ln(\tau_t) + \frac{\varepsilon_t^2}{\bar{h}_t^\infty \tau_t} \right]. \quad (6)$$

Similarly, conditional on initial values ( $\varepsilon_0, \bar{h}_0 = 0$ ) the observed quasi-log-likelihood can be written as  $L_T(\boldsymbol{\eta}, \boldsymbol{\pi} | \varepsilon_T, x_T, \varepsilon_{T-1}, x_{T-1}, \dots, \varepsilon_1, x_1, x_0, x_{-1}, \dots, x_{K-1}) = \sum_{t=1}^T l_t$  with

$$l_t = -\frac{1}{2} \left[ \ln(\bar{h}_t) + \ln(\tau_t) + \frac{\varepsilon_t^2}{\bar{h}_t \tau_t} \right]. \quad (7)$$

We assume that the sample contains  $T + K$  observations for  $x_t$ . Thus, there is no need to impose initial values for the explanatory variable and, hence, the same  $\tau_t$  appears in the observed and unobserved likelihood.

### 2.2.1 First derivatives

In the following, we consider the unobserved log-likelihood function. We define the average score vector evaluated under the null and at the true GARCH parameters as

$$\mathbf{D}^\infty(\boldsymbol{\eta}_0) = \begin{pmatrix} \mathbf{D}_\eta^\infty(\boldsymbol{\eta}_0) \\ \mathbf{D}_\pi^\infty(\boldsymbol{\eta}_0) \end{pmatrix} = \frac{1}{T} \sum_{t=1}^T \mathbf{d}_t^\infty(\boldsymbol{\eta}_0) = \frac{1}{T} \sum_{t=1}^T \begin{pmatrix} \mathbf{d}_{\eta,t}^\infty(\boldsymbol{\eta}_0) \\ \mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta}_0) \end{pmatrix},$$

where  $\mathbf{d}_{\eta,t}^\infty(\boldsymbol{\eta}_0) = \partial l_t^\infty / \partial \boldsymbol{\eta} |_{\boldsymbol{\eta}_0, \boldsymbol{\pi}=\mathbf{0}}$  and  $\mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta}_0) = \partial l_t^\infty / \partial \boldsymbol{\pi} |_{\boldsymbol{\eta}_0, \boldsymbol{\pi}=\mathbf{0}}$ . Next, we derive explicit expressions for  $\mathbf{d}_{\eta,t}^\infty(\boldsymbol{\eta}_0)$  and  $\mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta}_0)$ . First, consider the partial derivative of the log-likelihood with respect to  $\boldsymbol{\eta}$ :

$$\frac{\partial l_t^\infty}{\partial \boldsymbol{\eta}} = \frac{1}{2} \left[ \frac{\varepsilon_t^2}{\bar{h}_t^\infty \tau_t} - 1 \right] \left( \frac{1}{\bar{h}_t^\infty} \frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\eta}} + \frac{1}{\tau_t} \frac{\partial \tau_t}{\partial \boldsymbol{\eta}} \right) \quad (8)$$

with  $\partial \tau_t / \partial \boldsymbol{\eta} = (\partial f / \partial \boldsymbol{\pi}' \mathbf{x}_t) (\partial \mathbf{x}_t / \partial \boldsymbol{\eta})' \boldsymbol{\pi}$ . Under the null hypothesis, the long-term component reduces to unity and the short-term component simplifies to  $h_t^\infty = \bar{h}_t^\infty |_{\boldsymbol{\pi}=\mathbf{0}} = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^\infty$ . Note that  $h_t^\infty$  corresponds to the standard expression of the conditional variance in a GARCH(1,1). We then distinguish between

$$\mathbf{d}_{\eta,t}^\infty(\boldsymbol{\eta}) = \frac{\partial l_t^\infty}{\partial \boldsymbol{\eta}} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{2} \left[ \frac{\varepsilon_t^2}{h_t^\infty} - 1 \right] \mathbf{y}_t^\infty \quad (9)$$

with

$$\mathbf{y}_t^\infty = \frac{1}{\bar{h}_t^\infty} \frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\eta}} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{h_t^\infty} \sum_{i=0}^{\infty} \beta^i \mathbf{s}_{t-i}^\infty, \quad (10)$$

where  $\mathbf{s}_t^\infty = (1, \varepsilon_{t-1}^2, h_{t-1}^\infty)'$ , and the corresponding quantity evaluated at  $\boldsymbol{\eta}_0$ :

$$\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0) = \frac{1}{2} \left[ \frac{\varepsilon_t^2}{h_{0,t}^\infty} - 1 \right] \mathbf{y}_{0,t}^\infty, \quad (11)$$

with  $h_{0,t}^\infty = \omega_0 + \alpha_0 \varepsilon_{t-1}^2 + \beta_0 h_{0,t-1}^\infty$  and  $\mathbf{y}_{0,t}^\infty = (h_{0,t}^\infty)^{-1} \sum_{i=0}^{\infty} \beta_0^i \mathbf{s}_{0,t-i}^\infty$ .

The partial derivative with respect to  $\boldsymbol{\pi}$  leads to:

$$\frac{\partial l_t^\infty}{\partial \boldsymbol{\pi}} = \frac{1}{2} \left[ \frac{\varepsilon_t^2}{\bar{h}_t^\infty \tau_t} - 1 \right] \left( \frac{1}{\bar{h}_t^\infty} \frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\pi}} + \frac{1}{\tau_t} \frac{\partial \tau_t}{\partial \boldsymbol{\pi}} \right), \quad (12)$$

whereby the partial derivative of  $\bar{h}_t^\infty$  is given by

$$\frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\pi}} = -\alpha \sum_{j=0}^{\infty} \beta^j \frac{\varepsilon_{t-1-j}^2}{\tau_{t-1-j}^2} \frac{\partial \tau_{t-1-j}}{\partial \boldsymbol{\pi}}. \quad (13)$$

Since  $\partial \tau_t / \partial \boldsymbol{\pi} = (\partial f / \partial \boldsymbol{\pi}' \mathbf{x}_t)(\mathbf{x}_t + (\partial \mathbf{x}_t / \partial \boldsymbol{\pi})' \boldsymbol{\pi})$ , we have  $\partial \tau_t / \partial \boldsymbol{\pi}|_{\boldsymbol{\pi}=\mathbf{0}} = \mathbf{x}_t f'_0$  and, hence,

$$\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}) = \frac{\partial l_t^\infty}{\partial \boldsymbol{\pi}} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{2} \left[ \frac{\varepsilon_t^2}{h_t^\infty} - 1 \right] \mathbf{r}_t^\infty \quad (14)$$

with

$$\mathbf{r}_t^\infty = \left( \mathbf{x}_t - \alpha \frac{1}{h_t^\infty} \sum_{j=0}^{\infty} \beta^j \varepsilon_{t-1-j}^2 \mathbf{x}_{t-1-j} \right) f'_0. \quad (15)$$

Similarly as before, the corresponding expression evaluated at  $\boldsymbol{\eta}_0$  is given by:

$$\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0) = \frac{1}{2} \left[ \frac{\varepsilon_t^2}{h_{0,t}^\infty} - 1 \right] \mathbf{r}_{0,t}^\infty \quad (16)$$

with

$$\mathbf{r}_{0,t}^\infty = \left( \mathbf{x}_t - \alpha_0 \frac{1}{h_{0,t}^\infty} \sum_{j=0}^{\infty} \beta_0^j \varepsilon_{t-1-j}^2 \mathbf{x}_{t-1-j} \right) f'_0. \quad (17)$$

In summary, we have

$$\mathbf{D}^\infty(\boldsymbol{\eta}_0) = \frac{1}{T} \sum_{t=1}^T \mathbf{d}_t^\infty(\boldsymbol{\eta}_0) = \frac{1}{2T} \sum_{t=1}^T \left[ \frac{\varepsilon_t^2}{h_{0,t}^\infty} - 1 \right] \begin{pmatrix} \mathbf{y}_{0,t}^\infty \\ \mathbf{r}_{0,t}^\infty \end{pmatrix}. \quad (18)$$

Using that under  $H_0$ :  $\mathbf{E}[\varepsilon_t^2 / h_{0,t}^\infty] = \mathbf{E}[Z_t^2] = 1$ , it follows that  $\mathbf{E}[\mathbf{d}_t^\infty(\boldsymbol{\eta}_0) | \mathcal{F}_{t-1}] = \mathbf{0}$  and

$$\begin{aligned} \mathbf{Var}[\mathbf{d}_t^\infty(\boldsymbol{\eta}_0)] = \boldsymbol{\Omega} &= \begin{pmatrix} \boldsymbol{\Omega}_{\boldsymbol{\eta}\boldsymbol{\eta}} & \boldsymbol{\Omega}_{\boldsymbol{\eta}\boldsymbol{\pi}} \\ \boldsymbol{\Omega}_{\boldsymbol{\pi}\boldsymbol{\eta}} & \boldsymbol{\Omega}_{\boldsymbol{\pi}\boldsymbol{\pi}} \end{pmatrix} = \begin{pmatrix} \mathbf{E}[\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0) \mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0)'] & \mathbf{E}[\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0) \mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0)'] \\ \mathbf{E}[\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0) \mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0)'] & \mathbf{E}[\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0) \mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0)'] \end{pmatrix} \\ &= \frac{1}{4} (\kappa_Z - 1) \begin{pmatrix} \mathbf{E}[\mathbf{y}_{0,t}^\infty (\mathbf{y}_{0,t}^\infty)'] & \mathbf{E}[\mathbf{y}_{0,t}^\infty (\mathbf{r}_{0,t}^\infty)'] \\ \mathbf{E}[\mathbf{r}_{0,t}^\infty (\mathbf{y}_{0,t}^\infty)'] & \mathbf{E}[\mathbf{r}_{0,t}^\infty (\mathbf{r}_{0,t}^\infty)'] \end{pmatrix}. \end{aligned} \quad (19)$$

In the proof of Theorem 1 we will show that  $\boldsymbol{\Omega}$  is finite and positive definite. This will allow us to apply a central limit theorem for martingale difference sequences to  $\frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbf{d}_t^\infty(\boldsymbol{\eta}_0)$ .

### 2.2.2 Second derivatives

In the subsequent analysis we also make use of the following second derivatives:

$$\frac{\partial \mathbf{d}_{\eta,t}^{\infty}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'} = -\frac{1}{2} \frac{\varepsilon_t^2}{h_t^{\infty}} \mathbf{y}_t^{\infty} (\mathbf{y}_t^{\infty})' + \frac{1}{2} \left[ \frac{\varepsilon_t^2}{h_t^{\infty}} - 1 \right] \frac{\partial \mathbf{y}_t^{\infty}}{\partial \boldsymbol{\eta}'} \quad (20)$$

and

$$\frac{\partial \mathbf{d}_{\pi,t}^{\infty}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'} = -\frac{1}{2} \frac{\varepsilon_t^2}{h_t^{\infty}} \mathbf{r}_t^{\infty} (\mathbf{y}_t^{\infty})' + \frac{1}{2} \left[ \frac{\varepsilon_t^2}{h_t^{\infty}} - 1 \right] \frac{\partial \mathbf{r}_t^{\infty}}{\partial \boldsymbol{\eta}'} \quad (21)$$

We then define

$$\mathbf{J}_{\eta\eta} = -\mathbf{E} \left[ \frac{\partial \mathbf{d}_{\eta,t}^{\infty}(\boldsymbol{\eta}_0)}{\partial \boldsymbol{\eta}'} \right] = \frac{1}{2} \mathbf{E}[\mathbf{y}_{0,t}^{\infty} (\mathbf{y}_{0,t}^{\infty})'] \quad (22)$$

and

$$\mathbf{J}_{\pi\eta} = -\mathbf{E} \left[ \frac{\partial \mathbf{d}_{\pi,t}^{\infty}(\boldsymbol{\eta}_0)}{\partial \boldsymbol{\eta}'} \right] = \frac{1}{2} \mathbf{E}[\mathbf{r}_{0,t}^{\infty} (\mathbf{y}_{0,t}^{\infty})'] \quad (23)$$

Note that  $\mathbf{d}_{\eta,t}^{\infty}(\boldsymbol{\eta}_0)$  corresponds to the score of observation  $t$  in a standard GARCH(1,1) model and  $\partial \mathbf{d}_{\eta,t}^{\infty}(\boldsymbol{\eta}_0)/\partial \boldsymbol{\eta}'$  to the respective second derivative. Under Assumptions 1 and 2, it then directly follows from the results for the pure GARCH model in Francq and Zakoïan (2004) that  $\mathbf{J}_{\eta\eta}$  is finite and positive definite. Finally, note that  $\boldsymbol{\Omega}_{\eta\eta} = \frac{1}{2}(\kappa_Z - 1)\mathbf{J}_{\eta\eta}$  and  $\boldsymbol{\Omega}_{\pi\eta} = \frac{1}{2}(\kappa_Z - 1)\mathbf{J}_{\pi\eta}$ . If  $Z_t$  is normally distributed (i.e. the quasi-log-likelihood is correctly specified), then  $\kappa_Z = 3$  and  $\boldsymbol{\Omega}_{\eta\eta} = \mathbf{J}_{\eta\eta}$  and  $\boldsymbol{\Omega}_{\pi\eta} = \mathbf{J}_{\pi\eta}$ , respectively.

### 2.3 The $LM$ Test Statistic

The  $LM$  test statistic will be based on the observed quantity  $\mathbf{D}_{\pi}(\hat{\boldsymbol{\eta}}) = \frac{1}{T} \sum_{t=1}^T \mathbf{d}_{\pi,t}(\hat{\boldsymbol{\eta}})$ , where  $\hat{\boldsymbol{\eta}}$  is the QMLE of  $\boldsymbol{\eta}_0$  estimated under the null. We derive the asymptotic distribution of the test statistic in three steps. In the first step, we derive the asymptotic normality of the average score evaluated at  $\boldsymbol{\eta}_0$ . We then show that the lower part of the score evaluated at the QMLE can be related to the average score evaluated at  $\boldsymbol{\eta}_0$  in the following way:

$$\sqrt{T} \mathbf{D}_{\pi}^{\infty}(\hat{\boldsymbol{\eta}}) = [\mathbf{J}_{\pi\eta} \mathbf{J}_{\eta\eta}^{-1} \quad \mathbf{I}] \sqrt{T} \mathbf{D}^{\infty}(\boldsymbol{\eta}_0) + o_P(1). \quad (24)$$

In the final step it is necessary to show that the observed quantity  $\sqrt{T} \mathbf{D}_{\pi}(\hat{\boldsymbol{\eta}})$  has the same asymptotic distribution as  $\sqrt{T} \mathbf{D}_{\pi}^{\infty}(\hat{\boldsymbol{\eta}})$ . The  $LM$  statistic follows the usual  $\chi^2$  distribution. The structure of the proof builds on the arguments used in the proof of Theorem 2 in Halunga and Orme (2009). However, our set-up differs substantially from theirs because

we focus on an omitted multiplicative component that is driven by an explanatory variable, which may or may not be generated ‘outside the model’. Both features lead to substantial differences in the likelihood and test statistic.

Since the test statistic is based on the QMLE of  $\boldsymbol{\eta}_0$ , we can rely on the following result from Francq and Zakoïan (2004). If Assumptions 1 and 2 hold and the model is estimated under the null, the QMLE of the GARCH(1,1) parameters will be consistent and asymptotically normal:

$$\sqrt{T}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, (\kappa_Z - 1)(\mathbf{E}[\mathbf{y}_{0,t}^\infty(\mathbf{y}_{0,t}^\infty)'])^{-1}). \quad (25)$$

**Remark 1.** *In principle, we can relax the assumption that  $Z_t$  is i.i.d. Following Escanciano (2009) and Francq and Thieu (2018), the asymptotic normality of the QMLE can also be obtained under the weaker assumption that  $Z_t$  is strictly stationary and ergodic with  $\mathbf{E}[Z_t|\mathcal{F}_{t-1}] = 0$  and  $\mathbf{E}[Z_t^2|\mathcal{F}_{t-1}] = 1$ . This allows for a time-varying conditional kurtosis of  $Z_t$ . Under this weaker assumption the asymptotic distribution of the QMLE is given by*

$$\sqrt{T}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1} \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\eta}\boldsymbol{\eta}} \mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}), \quad (26)$$

where  $\tilde{\boldsymbol{\Omega}}_{\boldsymbol{\eta}\boldsymbol{\eta}} = \mathbf{E}[(\mathbf{E}[Z_t^4|\mathcal{F}_{t-1}] - 1)\mathbf{y}_{0,t}^\infty(\mathbf{y}_{0,t}^\infty)']$ . Clearly, if  $\mathbf{E}[Z_t^4|\mathcal{F}_{t-1}]$  is constant, (26) simplifies to (25).

In the following theorem, we derive the asymptotic distribution of the average score evaluated at  $\boldsymbol{\eta}_0$ . In order to ensure the finiteness of the covariance matrix of the average score, we impose the following assumption on  $x_t$ .

**Assumption 6.**  *$x_t$  is strictly stationary and ergodic with  $\mathbf{E}[|x_t|^{2q}] < \infty$  for some  $q > 1$ .*

By Assumption 2, the explanatory variable  $x_t$  is assumed to be weakly exogenous, i.e.  $\mathbf{E}[Z_t|\mathbf{x}_t] = 0$ . This allows for explanatory variables from ‘outside the model’, but also covers the case that  $x_t$  is ‘generated within the model’. In the empirical literature a variety of explanatory variables from outside the model – such as GDP growth, the term spread or the unemployment rate – has been used. Wang and Ghysels (2015) show that the GARCH-MIDAS model with a rolling window realized volatility as the explanatory variable can be rewritten such that  $x_t = \varepsilon_t^2$ , while the specification of Lundbergh and Teräsvirta (2002) selects  $x_t = \varepsilon_t^2/h_{0t}$  which is generated inside the model (see Section 2.4). For testing the simple GARCH model against the GARCH-MIDAS with  $x_t = \varepsilon_t^2$ , Assumption 6 requires that under the null  $\mathbf{E}[|\varepsilon_t|^{4q}] < \infty$  for some  $q > 1$ .

**Theorem 1.** *If Assumptions 1-6 hold, then  $\sqrt{T}\mathbf{D}^\infty(\boldsymbol{\eta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$ .*

Next, we consider the asymptotic distribution of the relevant lower part of the score vector evaluated at  $\hat{\boldsymbol{\eta}}$ . As an intermediate step, we show that  $\mathbf{J}_{\pi\boldsymbol{\eta}}$  can be consistently estimated by  $-\frac{1}{T} \sum_{t=1}^T \frac{\partial \mathbf{d}_{\pi,t}^\infty(\tilde{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}'}$ , where  $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}_0 + o_P(1)$ . The result is presented in Proposition 1 in Appendix A. The proof requires the following Assumption 7 which ensures that  $\mathbf{J}_{\pi\boldsymbol{\eta}}(\boldsymbol{\eta})$  is finite with a uniform bound for all  $\boldsymbol{\eta} \in \Theta$ .

**Assumption 7.**  $\mathbf{E}[|\varepsilon_t|^{4(1+s)}] < \infty$  for some  $0 < s \ll 1$ .

**Remark 2.** *Recall that no moment conditions on the observed process,  $\varepsilon_t$ , are needed for the asymptotic normality of the QMLE of the GARCH(1,1) when the GARCH parameters lie in the interior of the parameter space. However, as shown in Francq and Zakoïan (2007), if one is interested in testing whether some of the GARCH parameters are zero this requires the existence of  $\mathbf{E}[|\varepsilon_t|^6]$ . Our setting requires a weaker moment condition on  $\varepsilon_t$  (see also Halunga and Orme, 2009).*

**Theorem 2.** *If Assumptions 1-7 hold, then  $\sqrt{T}\mathbf{D}_\pi^\infty(\hat{\boldsymbol{\eta}}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$  with*

$$\begin{aligned} \boldsymbol{\Sigma} &= \boldsymbol{\Omega}_{\pi\pi} - \mathbf{J}_{\pi\boldsymbol{\eta}} \mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1} \boldsymbol{\Omega}'_{\boldsymbol{\eta}\pi} \\ &= \frac{1}{4} (\kappa_Z - 1) \left( \mathbf{E}[\mathbf{r}_{0,t}^\infty (\mathbf{r}_{0,t}^\infty)'] - \mathbf{E}[\mathbf{r}_{0,t}^\infty (\mathbf{y}_{0,t}^\infty)'] \left( \mathbf{E}[\mathbf{y}_{0,t}^\infty (\mathbf{y}_{0,t}^\infty)'] \right)^{-1} \mathbf{E}[\mathbf{y}_{0,t}^\infty (\mathbf{r}_{0,t}^\infty)'] \right). \end{aligned} \quad (27)$$

To prove Theorem 2, we use the fact that

$$\sqrt{T}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) = \mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1} \sqrt{T}\mathbf{D}_\boldsymbol{\eta}^\infty(\boldsymbol{\eta}_0) + o_P(1), \quad (28)$$

which again follows from Assumptions 1 and 2. As discussed in Francq and Zakoïan (2007, p.1280) the above equation holds only if the true parameters are not on the boundary of the parameter space. Assumption 1 thus excludes the ARCH(1) case where  $\beta_0 = 0$ .

The actual test statistic will be based on the observed quantity  $\mathbf{D}_\pi(\hat{\boldsymbol{\eta}})$ . The following theorem states the test statistic and its asymptotic distribution.

**Theorem 3.** *Consider the model given by equations (1-3), where  $\tau_{0,t} = f(\boldsymbol{\pi}'_0 \mathbf{x}_t)$ , and let Assumptions 1-7 be satisfied. Then, under  $H_0 : \boldsymbol{\pi}_0 = \mathbf{0}$  the test statistic*

$$\begin{aligned} LM &= T\mathbf{D}_\pi(\hat{\boldsymbol{\eta}})' \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{D}_\pi(\hat{\boldsymbol{\eta}}) \\ &= \frac{1}{4T} \left( \sum_{t=1}^T \left[ \frac{\varepsilon_t^2}{\hat{h}_t} - 1 \right] \hat{\mathbf{r}}_t \right)' \hat{\boldsymbol{\Sigma}}^{-1} \left( \sum_{t=1}^T \left[ \frac{\varepsilon_t^2}{\hat{h}_t} - 1 \right] \hat{\mathbf{r}}_t \right), \end{aligned} \quad (29)$$

where  $\hat{\boldsymbol{\eta}} = (\hat{\omega}, \hat{\alpha}, \hat{\beta})'$  is the vector of parameter estimates from the model under the null,  $\hat{\mathbf{r}}_t = (\mathbf{x}_t - (\hat{\alpha}/\hat{h}_t) \sum_{j=0}^{t-1} \hat{\beta}^j \varepsilon_{t-1-j}^2 \mathbf{x}_{t-1-j}) f'_0$ ,  $\hat{h}_t = \hat{\omega} + \hat{\alpha} \varepsilon_{t-1}^2 + \hat{\beta} \hat{h}_{t-1}$  and

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{4T} (\widehat{\kappa_Z - 1}) \left( \sum_{t=1}^T \hat{\mathbf{r}}_t \hat{\mathbf{r}}_t' - \sum_{t=1}^T \hat{\mathbf{r}}_t \hat{\mathbf{y}}_t' \left( \sum_{t=1}^T \hat{\mathbf{y}}_t \hat{\mathbf{y}}_t' \right)^{-1} \sum_{t=1}^T \hat{\mathbf{y}}_t \hat{\mathbf{r}}_t' \right) \quad (30)$$

with  $(\widehat{\kappa_Z - 1}) = 1/T \sum_{t=1}^T (\varepsilon_t^2/\hat{h}_t - 1)^2$  and  $\hat{\mathbf{y}}_t = (\hat{h}_t)^{-1} \sum_{j=0}^{t-1} \hat{\beta}^j (1, \varepsilon_{t-1-j}^2, \hat{h}_{t-1-j})'$  is a consistent estimator of  $\boldsymbol{\Sigma}$ , has an asymptotic  $\chi^2$  distribution with  $K$  degrees of freedom.

Theorem 3 illustrates that the *LM* test is a general misspecification test which is robust with respect to the functional form of  $f(\cdot)$ . That is, the test statistic in equation (29) does not depend on the constant  $f'_0$  because the  $(f'_0)^2$  in the ‘numerator’ and the ‘denominator’ of the test statistic cancel out. Moreover, our simulations in Section 3.2 suggest that the power of the test is only mildly affected by the specific functional form of  $f(\cdot)$  under the alternative.

**Remark 3.** The covariance matrix  $\widehat{\boldsymbol{\Sigma}}$  in Theorem 3 takes the same form as in Lundbergh and Teräsvirta (2002). The fact that we can factor out the term  $(\widehat{\kappa_Z - 1})$  follows from the assumption that  $Z_t$  is *i.i.d.* A modified version of the test statistic can be obtained under the weaker assumption discussed in Remark 1. However, this would require to further strengthen the assumptions on  $x_t$  and  $\varepsilon_t$ .

Essentially, the test statistic checks for a correlation between the squared standardized residuals from the model estimated under the null and the elements of the  $K$ -dimensional vector  $\hat{\mathbf{r}}_t$ . In empirical applications, the true lag length is unknown. Although the *LM* statistic can be easily computed for a variety of  $K$ 's, our simulation experiments have shown that for reasonable weighting schemes, choosing  $K = 1$  is sufficient in order to detect whether  $\mathbf{x}_t$  has an effect on long-term volatility or not (for details see Appendix C.3).

Moreover, it is straightforward to construct a regression version of our test (see also Lundbergh and Teräsvirta, 2002). The corresponding test statistic is given by  $LM = T(SSR_0 - SSR_1)/SSR_0$ , where  $SSR_0 = \sum_{t=1}^T (\varepsilon_t^2/\hat{h}_t - 1)^2$  and  $SSR_1$  is the sum of squared residuals from a regression of  $(\varepsilon_t^2/\hat{h}_t - 1)$  on  $\hat{\mathbf{r}}_t'$  and  $\hat{\mathbf{y}}_t'$ . Hence, *LM* is simply  $T$  times the uncentered  $R^2$  of the regression.

**Remark 4.** Wang and Ghysels (2015) consider the linear long-term component

$$f(\boldsymbol{\pi}' \mathbf{x}_t) = 1 + \boldsymbol{\pi}' \mathbf{x}_t, \quad (31)$$

which again ensures that  $f(0) = 1$ . However, this specification requires  $\boldsymbol{\pi} \geq \mathbf{0}$  as well as non-negative explanatory variables, i.e.  $x_t \geq 0$ , almost surely, in order to ensure the positivity of the conditional variance. This is in analogy to the ARCH( $K$ ) model in which  $\varepsilon_{t-j}^2 \geq 0$  corresponds to our  $x_{t-j}$  and where the assumption that all ARCH parameters are non-negative is needed to ensure that the conditional variance is positive. Although the alternative hypothesis becomes one-sided in this case, i.e. is given by  $H_1 : \boldsymbol{\pi}_0 \neq \mathbf{0}, \boldsymbol{\pi}_0 \geq \mathbf{0}$ , this does not affect the asymptotic distribution of our test statistic which is still  $\chi^2(K)$ . This result follows from the fact that the score vector is still asymptotically Gaussian under the null (see Francq and Zakoïan, 2009). However, as suggested by Demos and Sentana (1998), it may be possible to construct a one-sided version of our LM test that would be more powerful.

**Remark 5.** In principle, we could also use our approach to test for the joint significance of several explanatory variables, say,  $v_t, w_t, \dots$ . In this case  $\mathbf{x}_t$  would include the lagged values of the different explanatory variables, i.e.  $\mathbf{x}_t = (v_{t-1}, w_{t-1}, \dots)$ , and  $K$  would represent the number of different explanatory variables times the lags to be included. In this case, the exponential long-term component from equation (5) could be further multiplicatively decomposed as  $\tau_{0,t} = \exp(\pi_{0,1}v_{t-1}) \cdot \exp(\pi_{0,2}w_{t-1}) \cdot \dots$

**Remark 6.** Although we are primarily interested in testing for an omitted multiplicative component, for certain specifications of the long-term component the test can also be used to check for an additive misspecification of the short-term component. For example, if the long-term component is given by equation (31) and we choose  $K = 1$  and  $\mathbf{x}_t = (\mathbb{1}_{\{\varepsilon_{t-1} < 0\}}\varepsilon_{t-1}^2)/\bar{h}_{0,t}^\infty$ , where  $\mathbb{1}$  is an indicator function, we obtain

$$\sigma_{0t}^2 = \bar{h}_{0t}^\infty + \pi_{0,1}\mathbb{1}_{\{\varepsilon_{t-1} < 0\}}\varepsilon_{t-1}^2. \quad (32)$$

Under  $H_0 : \pi_{0,1} = 0$ ,  $\sigma_{0t}^2$  follows a GARCH(1,1), while under the alternative there is an additive asymmetry term. Indeed, equation (32) is similar to the models considered in Engle and Ng (1993) for testing for asymmetry in the conditional variance.

Finally, it is interesting to consider two special cases that are nested within our framework when there are no GARCH effects, i.e. when  $\alpha_0 = \beta_0 = 0$  and  $\bar{h}_{0t}^\infty = \omega_0$ . In this case, the model under  $H_0$  has constant conditional and unconditional variance equal to  $\sigma_0^2 = \omega_0$ . Under the alternative, the conditional variance is given by  $\mathbf{Var}[\varepsilon_t | \mathcal{F}_{t-1}] = \sigma_0^2 \tau_t$ . Without



GARCH effects and under  $H_0$ , the average score in equation (18) can be rewritten as

$$\mathbf{D}^\infty(\boldsymbol{\eta}_0) = \frac{1}{2T} \sum_{t=1}^T \left[ \frac{\varepsilon_t^2}{\sigma_0^2} - 1 \right] \begin{pmatrix} \sigma_0^{-2} \\ f_0' \mathbf{x}_t \end{pmatrix}. \quad (33)$$

As a result, the regression-based test simplifies to regressing the squared returns on a constant and  $\mathbf{x}_t$  and computing  $TR^2$  which resembles the Godfrey (1978) test for multiplicative heteroskedasticity. Finally, the Engle (1982) test for ARCH effects is obtained when choosing  $\mathbf{x}_t = (\varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \dots, \varepsilon_{t-K}^2)'$ .

## 2.4 Relation to *LM* test of Lundbergh and Teräsvirta (2002)

Next, we compare our test statistic to the Lundbergh and Teräsvirta (2002) test for misspecification in GARCH models. Their test is based on the following specification  $\varepsilon_t = \sqrt{h_{0t}^\infty} \xi_{0t} = \sqrt{h_{0t}^\infty \tau_{0t}} Z_t$ , where  $h_{0t}^\infty$  is defined as before and  $\tau_{0t} = 1 + \boldsymbol{\pi}_0' \mathbf{x}_t$ , i.e. they assume a linear long-term component. Lundbergh and Teräsvirta (2002) make the specific choice of  $x_t = \xi_{0t}^2 = \varepsilon_t^2 / h_{0t}^\infty$  for the explanatory variable. Because under this assumption  $\xi_{0t} = \sqrt{\tau_{0t}} Z_t$  follows an ARCH( $K$ ), Lundbergh and Teräsvirta (2002) refer to this specification as ‘ARCH nested in GARCH’ and test the null hypothesis  $H_0 : \boldsymbol{\pi}_0 = \mathbf{0}$ . Although the ‘ARCH nested in GARCH’ is remarkably similar to our model, there is an important conceptual difference. Since the short-term component is based on  $h_{0t}^\infty$  (instead of  $\bar{h}_{0t}^\infty$ ), it depends on the squared observation  $\varepsilon_{t-1}^2$  (instead of  $\varepsilon_{t-1}^2 / \tau_{0,t-1}$ ). Because of this,  $\sqrt{h_{0t}^\infty} Z_t$  follows a GARCH(1,1) process under the null but *not* under the alternative. Due to this difference in the specification, their model does not encompass the GARCH-MIDAS under the alternative. Moreover, it follows that  $\partial h_t^\infty / \partial \boldsymbol{\pi} = \mathbf{0}$  and, hence, in the Lundbergh and Teräsvirta (2002) setting equation (17) reduces to  $\mathbf{r}_{0,t}^\infty = (\varepsilon_{t-1}^2 / h_{0,t-1}^\infty, \varepsilon_{t-2}^2 / h_{0,t-2}^\infty, \dots, \varepsilon_{t-K}^2 / h_{0,t-K}^\infty)'$ . Thus, their *LM* test statistic is based on

$$\begin{bmatrix} \varepsilon_t^2 \\ \hat{h}_t \end{bmatrix} \hat{\mathbf{r}}_t^{LT}, \quad (34)$$

where  $\hat{h}_t = \hat{\omega} + \hat{\alpha} \varepsilon_{t-1}^2 + \hat{\beta} \hat{h}_{t-1}$  and  $\hat{\mathbf{r}}_t^{LT}$  has entries  $\varepsilon_{t-k}^2 / \hat{h}_{t-k}$ ,  $k = 1, \dots, K$ . Intuitively, equation (34) is used to test whether the squared standardized returns are still correlated. In this sense, the test is intended to be a very general misspecification test with omitted ARCH under the alternative (instead of a well-specified alternative).

In the Section 3, we will compare the ‘ARCH nested in GARCH’ test of Lundbergh and Teräsvirta (2002) to our new  $LM$  test. We denote the test statistic based on  $\hat{\mathbf{r}}_t^{LT}$  instead of  $\hat{\mathbf{r}}_t$  by  $LM_{LT}$ . In addition, we consider a modified version of the Lundbergh and Teräsvirta (2002) test, in which we allow for a general regressor  $x_t$ . In this case, equation (17) is simply given by  $\mathbf{r}_{0,t}^\infty = \mathbf{x}_t = \hat{\mathbf{r}}_t^{LT,mod}$ . We denote the corresponding test statistic  $LM_{LT,mod}$ . Since  $\hat{\mathbf{r}}_t - \hat{\mathbf{r}}_t^{LT,mod} = (\hat{\alpha}/\hat{h}_t) \sum_{j=0}^{t-1} \hat{\beta}^j \varepsilon_{t-1-j}^2 \mathbf{x}_{t-1-j}$ , our new test,  $LM$ , and  $LM_{LT,mod}$  can be expected to perform similarly if, for example,  $\hat{\alpha}$  is small. On the other hand, we expect that our test will have better power properties than the modified Lundbergh and Teräsvirta (2002) test when the ARCH effect is strong. The modified test,  $LM_{LT,mod}$ , coincides with a test proposed in Section 4.4 of Amado and Teräsvirta (2017), if their deterministic component is assumed to be one (in their notation:  $g_t \equiv 1$ ) and  $x_t$  is used as the explanatory variable (see their equation (6)).

## 2.5 Generalization of $LM$ Test: Alternative $\bar{h}_{0t}^\infty$ Components

Next, we discuss how the  $LM$  test statistic can be constructed when assuming an alternative short-term component. We focus our attention on the GJR-GARCH which allows positive and negative innovations to have an asymmetric effect on volatility. When assuming a GJR-GARCH(1,1) short-term component, we have to replace equation (3) by

$$\bar{h}_{0t}^\infty = \omega_0 + (\alpha_0 + \gamma_0 \mathbf{1}_{\{\varepsilon_{t-1} < 0\}}) \frac{\varepsilon_{t-1}^2}{\tau_{0,t-1}} + \beta_0 \bar{h}_{0,t-1}^\infty, \quad (35)$$

where  $\omega_0 > 0$ ,  $\alpha_0 > 0$ ,  $\alpha_0 + \gamma_0 > 0$ ,  $\beta_0 > 0$  and  $\alpha_0 + \gamma_0/2 + \beta_0 < 1$ . The long-term component remains unchanged. For the GJR-GARCH under  $H_0$ , equation (11) has the same structure as before but with

$$h_{0,t}^\infty = \omega_0 + \alpha_0 \varepsilon_{t-1}^2 + \gamma_0 \mathbf{1}_{\{\varepsilon_{t-1} < 0\}} \varepsilon_{t-1}^2 + \beta_0 h_{0,t-1}^\infty \quad (36)$$

and  $\mathbf{y}_{0,t}^\infty = (h_{0,t}^\infty)^{-1} \sum_{i=0}^\infty \beta_0^i \mathbf{s}_{0,t-i}^\infty$ , where now  $\mathbf{s}_{0,t}^\infty = (1, \varepsilon_{t-1}^2, \mathbf{1}_{\{\varepsilon_{t-1} < 0\}} \varepsilon_{t-1}^2, h_{0,t-1}^\infty)'$ . That is, by taking the asymmetry into account,  $\mathbf{y}_{0,t}^\infty$  and  $\mathbf{s}_{0,t}^\infty$  become  $(4 \times 1)$  vectors. Similarly, equation (17) becomes

$$\mathbf{r}_{0,t}^\infty = \left( \mathbf{x}_t - \frac{1}{h_{0,t}^\infty} \sum_{j=0}^\infty \beta_0^j (\alpha_0 + \gamma_0 \mathbf{1}_{\{\varepsilon_{t-1-j} < 0\}}) \varepsilon_{t-1-j}^2 \mathbf{x}_{t-1-j} \right) f'_0, \quad (37)$$

but remains a  $(K \times 1)$  vector. In order to implement the  $LM$  test, one now estimates the GJR-GARCH under the null hypothesis. As before, the actual  $LM$  test statistic can be

computed based on the respective estimates  $\hat{h}_t$ ,  $\hat{\mathbf{y}}_t$  and  $\hat{\mathbf{r}}_t$ . The asymptotic distribution of the test statistic will again be  $\chi^2$  with  $K$  degrees of freedom. To formally show this, our assumptions on the model parameters and our proofs could be adapted to the case of a GJR-GARCH under  $H_0$  based on the asymptotic theory for the corresponding QMLE provided in Hamadeh and Zakoïan (2011).

In a similar fashion, the  $LM$  test can be constructed for other GARCH-type specifications of the short-term component such as higher-order GARCH models or long-memory specifications (e.g. the FIGARCH model).

## 2.6 Mixed-Data Sampling

As mentioned in Section 2.1, the two-component model is often applied in settings where the explanatory variable is observed at a lower frequency than the daily returns. In order to capture such a setting we have to slightly adapt our notation. As before, we use  $x_t$  to denote the explanatory variable, but now  $t$  refers to, for example, a monthly or quarterly frequency. We denote the daily returns by  $\varepsilon_{i,t}$ , where  $i = 1, \dots, M$  refers to the  $M$  days within each month/quarter. Equation (1) can then be rewritten as

$$\varepsilon_{i,t} = \sqrt{\bar{h}_{0,i,t}^\infty \tau_{0,t}} Z_{i,t}, \quad (38)$$

whereby Assumption 2 now holds for  $Z_{i,t}$  with  $\mathcal{F}_{i,t}$  defined accordingly. Note that the long-term component has an index  $t$  only, since it is constant within each month/quarter. On the other hand, the GARCH component varies at the daily frequency.

We propose two versions of the test for the mixed-frequency case. The first version again leads to an  $LM$  test and, for space considerations, is presented in Appendix B. The second version is regression-based and links our test to the empirical literature on predictive regressions for financial volatility. This version is presented in the following.

The regression-based test makes use of what we call the *volatility-adjusted realized variance*,  $\widetilde{RV}_t$ . We define this quantity as the sum of the squared daily deGARCHed returns,  $\varepsilon_{i,t}^2 / \bar{h}_{0,i,t}^\infty$ , within each month/quarter:

$$\widetilde{RV}_t = \sum_{i=1}^M \frac{\varepsilon_{i,t}^2}{\bar{h}_{0,i,t}^\infty} = \tau_{0,t} \sum_{i=1}^M Z_{i,t}^2 = \tau_{0,t} \tilde{Z}_t. \quad (39)$$

To further simplify the analysis, we assume that the long-term component is given by  $\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$ . This assumption is reasonable since it is the most common specification

of  $\tau_{0,t}$  in the empirical literature. It is then natural to consider the log of equation (39) as a regression model:

$$\ln(\widetilde{RV}_t) = \boldsymbol{\pi}'_0 \mathbf{x}_t + \ln(\widetilde{Z}_t) = \tilde{c} + \boldsymbol{\pi}'_0 \mathbf{x}_t + \tilde{\zeta}_t, \quad (40)$$

where  $\tilde{c} = \mathbf{E}[\ln(\widetilde{Z}_t)]$  and  $\tilde{\zeta}_t = \ln(\widetilde{Z}_t) - \tilde{c}$ . Note that by Assumption 2 the innovation  $\tilde{\zeta}_t$  is i.i.d.. That is, under  $H_0 : \boldsymbol{\pi}_0 = \mathbf{0}$  the volatility-adjusted realized variance  $\ln(\widetilde{RV}_t)$  should be unpredictable. In order to obtain an estimable version, we have to replace  $\widetilde{RV}_t$  in equation (40) with the estimate  $\widehat{RV}_t = \sum_{i=1}^M \varepsilon_{i,t}^2 / \hat{h}_{i,t}$ , where  $\hat{h}_{i,t}$  is the estimate from the model under the null.

Equation (40) is very much in analogy to the predictive regression model often used when analyzing the link between financial volatility and macro conditions (see Paye, 2012, Christiansen et al, 2012, and others). The important difference is that predictive regressions directly try to explain the realized variance, i.e. are based on regressions of the following type:

$$\ln(RV_t) = c + \boldsymbol{\pi}'_0 \mathbf{x}_t + \zeta_t, \quad (41)$$

where  $RV_t = \sum_{i=1}^M \varepsilon_{i,t}^2$ . From equation (38) it follows that the innovation in equation (41) is given by  $\zeta_t = \ln(\sum_{i=1}^M \bar{h}_{0,it}^\infty Z_{i,t}^2) - \mathbf{E}[\ln(\sum_{i=1}^M \bar{h}_{0,it}^\infty Z_{i,t}^2)]$ . Note that  $\zeta_t$  is a low-frequency process that depends on the sum of a squared high-frequency GARCH process. That is, while  $\tilde{\zeta}_t$  is i.i.d., we can expect that  $\zeta_t$  has a higher variance and is strongly autocorrelated. This intuition is in line with the fact that  $\ln(RV_t)$  is typically found to be highly persistent. These considerations suggest that the relationship between  $\mathbf{x}_t$  and financial volatility is more difficult to detect when using equation (41) rather than equation (40) as a regression model.

Finally, note that in predictive regressions typically the lagged realized variance is also included as an additional explanatory variable. This leads to the regression

$$\ln(RV_t) = c + \boldsymbol{\pi}'_0 \mathbf{x}_t + \rho \ln(RV_{t-1}) + \zeta_t. \quad (42)$$

The point that predictive regressions like in equation (42) might be problematic has already been made by Engle et al. (2013) who argue that  $\ln(RV_t)$  is a noisy measure of the true unobservable long-term component. This creates problems due to measurement error on both the left as well as the right hand side of the equation. In the empirical application in Section 4.2, we estimate a variant of equation (42) in which we replace

$\ln(RV_t)$  with  $\ln(\widehat{RV}_t)$ . Providing a complete treatment of the asymptotic theory of the predictive regression version of the test is beyond the scope of the current paper. However, in Appendix C.7 we provide detailed simulation results suggesting that the test works well in practice.

### 3 Simulations

In this section, we examine the finite sample behavior of the proposed test in a Monte-Carlo experiment. We simulate return series with  $T = 1000$  observations and use  $R = 1000$  Monte-Carlo replications. The innovation  $Z_t$  is assumed to be either standard normally distributed or (standardized)  $t$ -distributed with seven degrees of freedom. In order to consider a realistic example under the alternative, we will base the long-term component on actual data. As an explanatory variable, we use the squared daily VIX index,  $VIX_t$ , for the period October 2010 to October 2014. More specifically, we define  $VIX_t$  as  $1/365$  times the squared VIX index so that the squared annualized observations are transformed to daily units. In addition, we construct monthly and quarterly rolling window versions of the squared VIX as  $VIX_t^{(N)} = \frac{1}{N} \sum_{j=0}^{N-1} VIX_{t-j}$ , with  $N = 22$  and  $N = 65$ . Figure 1 shows the evolution of the VIX and its rolling window versions over the sample period. The spikes in the third quarter of 2011 correspond to the financial turmoil during the European sovereign debt crisis.

#### 3.1 Size Properties

We first discuss the size properties of the test statistic. Three alternative GARCH(1,1) specifications are considered. The three GARCH specifications reflect different degrees of persistence (Low: G-L, Moderate: G-M, High: G-H) in the conditional variance, whereby we measure persistence by  $\alpha_0 + \beta_0$ . We keep  $\beta_0$  fixed at 0.9 and choose  $\alpha_0 \in \{0.05, 0.07, 0.09\}$ .  $\omega_0$  is always chosen such that under the null  $\sigma_0^2 = 1$ .

$$\begin{aligned} \text{G-L: } \quad \bar{h}_{0t}^\infty &= 0.05 + 0.05 \frac{\varepsilon_{t-1}^2}{\tau_{0,t-1}} + 0.90 \bar{h}_{0,t-1}^\infty \\ \text{G-M: } \quad \bar{h}_{0t}^\infty &= 0.03 + 0.07 \frac{\varepsilon_{t-1}^2}{\tau_{0,t-1}} + 0.90 \bar{h}_{0,t-1}^\infty \\ \text{G-H: } \quad \bar{h}_{0t}^\infty &= 0.01 + 0.09 \frac{\varepsilon_{t-1}^2}{\tau_{0,t-1}} + 0.90 \bar{h}_{0,t-1}^\infty \end{aligned}$$

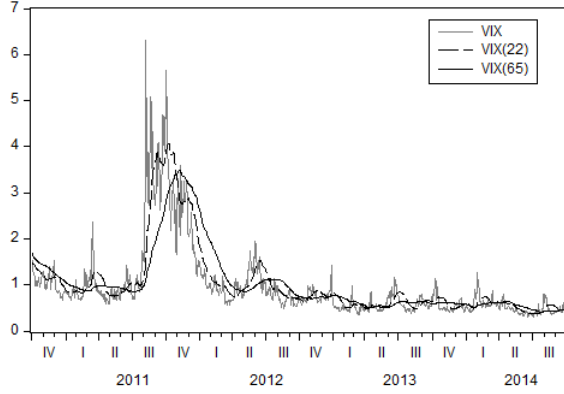


Figure 1: The figure shows the evolution of  $VIX_t$  (grey line),  $VIX_t^{(22)}$  (dashed black line) and  $VIX_t^{(65)}$  (solid black line) for the period October 2010 to October 2014. The three variables are presented in daily units.

In addition, we consider GJR-GARCH(1,1) specifications for the short-term component. These models are given by

$$\begin{aligned}
 \text{GJR-L: } \quad \bar{h}_{0t}^\infty &= 0.05 + (0.02 + 0.06 \cdot \mathbb{1}_{\{\varepsilon_{t-1} < 0\}}) \frac{\varepsilon_{t-1}^2}{\tau_{0,t-1}} + 0.90\bar{h}_{0,t-1}^\infty \\
 \text{GJR-M: } \quad \bar{h}_{0t}^\infty &= 0.03 + (0.02 + 0.10 \cdot \mathbb{1}_{\{\varepsilon_{t-1} < 0\}}) \frac{\varepsilon_{t-1}^2}{\tau_{0,t-1}} + 0.90\bar{h}_{0,t-1}^\infty \\
 \text{GJR-H: } \quad \bar{h}_{0t}^\infty &= 0.01 + (0.02 + 0.14 \cdot \mathbb{1}_{\{\varepsilon_{t-1} < 0\}}) \frac{\varepsilon_{t-1}^2}{\tau_{0,t-1}} + 0.90\bar{h}_{0,t-1}^\infty.
 \end{aligned}$$

For the GJR-GARCH models persistence is measured by  $\alpha_0 + \gamma_0/2 + \beta_0$ . Note that we have chosen  $\alpha_0$  and  $\gamma_0$  such that the persistence in GJR-L, GJR-M and GJR-H coincides with the persistence in models G-L, G-M and G-H above.

To implement the test, we have to specify the explanatory variable. In Panels A and C of Table 1 we test for remaining ARCH effects by choosing  $x_t = \varepsilon_t^2 / \hat{h}_t$  and in Panels B and D we choose  $x_t = VIX_t$ . In Panels A/B the model under  $H_0$  is the GARCH(1,1), while in Panels C/D the model under  $H_0$  is the GJR-GARCH(1,1). We report the empirical size for the  $LM$  test given in equation (29) as well as for the original,  $LM_{LT}$ , and modified,  $LM_{LT,mod}$ , test statistics of Lundbergh and Teräsvirta (2002). For all three tests we choose  $K = 1$ . We first discuss the results when testing for remaining ARCH effects. As Panels A/C of Table 1 show, the empirical size of both versions of the test statistic is very close to the nominal size when  $Z_t$  is normally distributed. In case of Student- $t$  distributed

Table 1: Empirical size.

	$Z_t \sim \mathcal{N}(0, 1)$			$Z_t \sim t(7)$		
$H_0$ : GARCH	G-L	G-M	G-H	G-L	G-M	G-H
	Panel A: $x_t = \varepsilon_t^2 / \hat{h}_t$					
$LM$	4.6	5.0	5.2	3.1	3.7	3.9
$LM_{LT}$	5.2	5.2	5.1	3.4	3.8	3.9
	Panel B: $x_t = VIX_t$					
$LM$	6.4	6.0	5.7	6.0	6.1	5.5
$LM_{LT,mod}$	7.0	7.1	7.9	5.5	6.4	7.1
$H_0$ : GJR-GARCH	GJR-L	GJR-M	GJR-H	GJR-L	GJR-M	GJR-H
	Panel C: $x_t = \varepsilon_t^2 / \hat{h}_t$					
$LM$	5.3	5.4	6.0	3.7	4.0	4.0
$LM_{LT}$	5.7	6.0	5.9	4.1	4.4	4.0
	Panel D: $x_t = VIX_t$					
$LM$	7.2	6.2	6.1	6.5	6.8	6.3
$LM_{LT,mod}$	7.5	8.1	7.7	5.9	6.2	6.1

**Notes:** The number of observations is  $T = 1000$ . Entries are rejection rates in percent over the  $R = 1000$  replications at the 5% nominal level. In Panels A and B the model for the conditional variance is a GARCH(1,1) with  $\beta_0 = 0.90$ . G-L, G-M and G-H refer to GARCH models with low ( $\alpha_0 = 0.05$ ), moderate ( $\alpha_0 = 0.07$ ) and high ( $\alpha_0 = 0.09$ ) persistence. In Panels C and D the model for the conditional variance is a GJR-GARCH(1,1) with  $\alpha_0 = 0.02$  and  $\beta_0 = 0.90$ . GJR-L, GJR-M and GJR-H refer to models with low ( $\gamma_0 = 0.06$ ), moderate ( $\gamma_0 = 0.10$ ) and high ( $\gamma_0 = 0.14$ ) persistence, respectively.  $\omega_0$  is chosen such that  $\sigma_0^2 = 1$ . All test statistics are based on  $K = 1$ .

innovations, the two test statistics are slightly undersized. For the  $LM_{LT}$  test statistic, this is an observation also made in Lundbergh and Teräsvirta (2002) and Halunga and Orme (2009). However, in Appendix C.1, we show that the size distortion fades away with increasing sample size. As Panels B/D show, both tests are modestly oversized when the squared VIX is used as an explanatory variable.

### 3.2 Power Properties

We simulate the model under the alternative using the exponential specification from equation (5). Three alternative weighting schemes  $\boldsymbol{\pi}_0$  are considered. The first one includes only the first lag of  $x_t$  with a weight of  $\pi_{0,1} = 0.3$ . We refer to this weighting scheme as one with immediate (I) decay. The second and third weighting schemes are

shown in Figure 2. The solid and dashed lines represent weights that either have a fast (F) or a slow (S) decay. The second and third weighting schemes are scaled such that their weights add up to 0.3. The weights were generated using a Beta polynomial (see Ghysels et al., 2007).

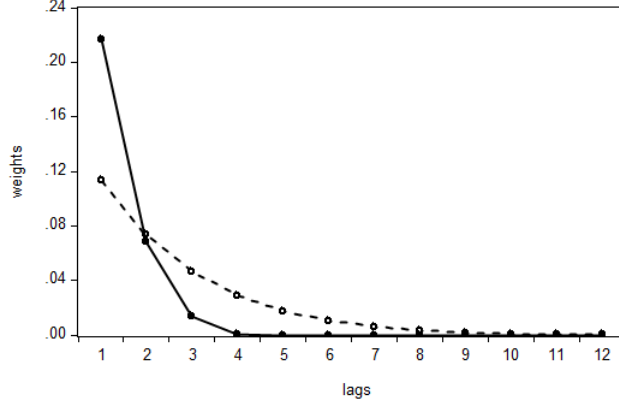


Figure 2: Alternative weighting schemes  $\pi_{0,k}$ , for lags  $k = 1, \dots, 12$ , with fast (F; solid line) and slow (S; dashed line) decay.

Table 2 presents size-adjusted rejection rates that were obtained from the Monte-Carlo simulations. The  $LM$  ( $LM_{LT,mod}$ ) test statistics are based on  $\hat{\mathbf{r}}_t$  ( $\hat{\mathbf{r}}_t^{LT,mod}$ ) with  $x_t \in \{VIX_t, VIX_t^{(22)}, VIX_t^{(65)}\}$ . As before,  $LM_{LT}$  is based on  $\hat{\mathbf{r}}_t^{LT}$  with  $x_t = \varepsilon_t^2 / \hat{h}_t$  and, hence, tests for ‘ARCH nested in GARCH’. For all three test statistics we choose  $K = 1$ , i.e. the tests are based on the first lag of  $x_t$ . Thus, the results in Table 2 illustrate the performance of the test statistics when  $K$  is correctly chosen but also when  $K$  is smaller than the true lag length. The upper/lower part of Table 2 shows the size-adjusted power for the test with GARCH/GJR-GARCH under  $H_0$ .

For the time being, consider the upper part of Table 2. At first, we choose  $x_t = VIX_t$  and report the size-adjusted power for the two GARCH models with high (G-H) and moderate (G-M) persistence in combination with the different weighting schemes. In addition, we report the variance ratio:  $VR = \widehat{\mathbf{Var}}(\ln(\tau_{0t})) / \widehat{\mathbf{Var}}(\ln(\tau_{0t} \bar{h}_{0t}^\infty))$ , which reflects the fraction of the sample variance of the log conditional variance that is due to the sample variance of the log long-term component (see Conrad and Loch, 2015a). For example, for  $\alpha_0 = 0.09$  (G-H) in combination with an immediately decaying weighting scheme, 15.6% of the total conditional variance is due to the long-term component. When  $\alpha_0$  is decreased



Table 2: Empirical size-adjusted power for an exponential long-term component.

$x_t$ weighting scheme	$VIX_t$						$VIX_t^{(22)}$	$VIX_t^{(65)}$
	I	F	S	I	F	S	I	I
$H_0$ : GARCH	G-H			G-M			G-H	
$LM$	74.4	71.0	49.9	79.9	77.3	59.4	36.8	19.9
$LM_{LT,mod}$	42.2	39.9	31.2	70.9	69.5	59.0	15.6	11.1
$LM_{LT}$	5.6	5.6	5.9	5.9	5.6	5.6	5.0	4.7
$VR$	15.6	15.5	14.8	35.5	35.2	34.1	14.7	12.0
$H_0$ : GJR-GARCH	GJR-H			GJR-M			GJR-H	
$LM$	77.1	74.9	54.2	82.8	80.6	62.1	41.5	24.8
$LM_{LT,mod}$	51.6	49.6	39.2	71.6	70.3	59.6	21.9	14.4
$LM_{LT}$	5.7	5.7	5.8	5.8	5.8	5.5	5.3	5.5
$VR$	11.7	11.6	11.1	26.9	26.7	25.7	11.0	8.9

**Notes:** The table reports the empirical size-adjusted power at the 5% nominal level. The specification of the long term component is given by  $\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$  with weighting schemes with immediate (I), fast (F) and slow (S) decay. The GARCH parameters are  $\alpha_0 = 0.09$  (G-H) or  $\alpha_0 = 0.07$  (G-M) in combination with  $\beta_0 = 0.90$  and  $\omega_0 = 1 - \alpha_0 - \beta_0$ . The GJR-GARCH parameters are  $\alpha_0 = 0.02$ ,  $\gamma_0 = 0.14$  (GJR-H) or  $\gamma_0 = 0.10$  (GJR-M) in combination with  $\beta_0 = 0.90$  and  $\omega_0 = 1 - \alpha_0 - \gamma_0/2 - \beta_0$ . Innovations  $Z_t$  are standard normal distributed. The variance ratio,  $VR = \widehat{\mathbf{Var}}(\ln(\tau_{0t})) / \widehat{\mathbf{Var}}(\ln(\tau_{0t} \bar{h}_{0t}^\infty))$ , is the fraction of the sample variance of the log conditional variance that is due to the sample variance of the log long-term component. All test statistics are based on  $K = 1$ . The number of observations is  $T = 1000$ . Results are based on  $R = 1000$  replications.

to 0.07 (G-M), the  $VR$  increases to 35.5%. Clearly, decreasing  $\alpha_0$  means to weaken the GARCH effect and, hence, to reduce the variability with which the short-term component fluctuates around  $\tau_{0t}$ .<sup>3</sup>

First, consider the case where  $\alpha_0 = 0.09$  (G-H). For the immediately decaying weighting scheme, the  $LM$  test rejects the null hypothesis in 74.4% of the simulations at the nominal 5% level. In contrast, the rejection rate of the modified Lundbergh and Teräsvirta (2002) test,  $LM_{LT,mod}$ , is 42.2% only. Next, we consider the weighting schemes with fast and slow decay. In these cases, the long-term component becomes less variable and, hence, more difficult to detect. Consequently, the power of both tests deteriorates. Nevertheless,

<sup>3</sup>For  $\alpha_0 \rightarrow 0$  the GARCH component approaches a constant and all the variation in  $\sigma_{0,t}^2$  is due to movements in the long-term component.

the  $LM$  test still has considerably higher power than  $LM_{LT,mod}$ . When  $\alpha_0$  is decreased to 0.07 (G-M), this increases the power of both tests. For example, for the immediately decaying weighting scheme the size-adjusted power at the nominal 5% level is now 79.9% for the  $LM$  test. Clearly, with lower  $\alpha_0$  and thus less volatile GARCH component, the long-term component can be detected more easily. As before, having more slowly decaying weights, i.e. increasing the smoothness of the long-term component, reduces the power of the tests. In line with the arguments at the end of Section 2.4, the difference in the power of the  $LM$  and  $LM_{LT,mod}$  statistics is less strong when  $\alpha_0$  is decreased to 0.07. Finally, the last two columns of Table 2 show the rejection rates for the case that the long-term component is based on the monthly and quarterly rolling window versions of the squared VIX. Then, even for the immediately decaying weighting scheme, the long-term components are very smooth and the lowest  $VR$ 's are observed. As expected, the size-adjusted powers are the lowest for these two cases. Note that in all eight scenarios the original version of the Lundbergh and Teräsvirta (2002) test,  $LM_{LT}$ , has only trivial power against deviations from the null hypothesis. This is not surprising since  $LM_{LT}$  is searching for an omitted ARCH component and, hence, is ‘searching in the wrong place’.

The lower part of Table 2 confirms all findings for GJR-GARCH models with high (GJR-H) and moderate (GJR-M) persistence.

In summary, the size-adjusted power of the newly proposed test,  $LM$ , is higher the more volatile the long-term component is and the less volatile the short-term component fluctuates around the long-term component (i.e. the lower  $\alpha_0$  is).

We performed the same analysis as in Table 2 for the case of Student- $t$  distributed innovations  $Z_t$  (see Table 6 in Appendix C.2). As the table shows, for each specification the  $t$  distributed innovations decrease the  $VR$  in comparison to the one that we obtained for normally distributed innovations. The lower  $VR$ 's then lead to a loss of power, i.e. under  $t$  distributed innovations the long-term component is more difficult to detect. However, all qualitative results regarding the different versions of the test statistics remain unchanged.

### 3.3 Practical Implementation and Robustness Checks

First, we consider two issues related to the practical implementation of the  $LM$  test.

**Choice of  $K$ :** We performed simulations to investigate how important it is to correctly choose  $K$ . The simulations suggest that for all practically relevant weighting schemes the

choice of  $K = 1$  is very reasonable. Even in extreme cases where all weight is put on lags greater than one, choosing  $K = 1$  only leads to a modest loss in power relative to the correct choice of  $K$  (see Appendix C.3). Nevertheless, if the  $LM$  test does not reject for  $K = 1$  and  $x_t$  has low persistence, it may be advisable to successively increase  $K$  to check whether there is some dependence on higher lags.

**Persistence in the GARCH component:** In empirical studies, the sum of the estimated GARCH parameters,  $\hat{\alpha} + \hat{\beta}$ , is often found to be close to one or, occasionally, even greater than one. The simulations in Appendix C.4 investigate the power of the  $LM$  test in situations in which  $\alpha_0 + \beta_0$  is close to (but below) one. Using various specifications of the long-term component, we show that – despite an omitted (stationary) long-term component – the median of the parameter estimates of  $\alpha$  and  $\beta$  for the GARCH model under the null hypothesis is close to the true parameter values. Only if the sum of the true GARCH parameters is very close to one, some of the simulations lead to parameter estimates such that  $\hat{\alpha} + \hat{\beta} \geq 1$ . This effect is the stronger the more relevant the long-term component is. Although the variance ratio decreases considerably in cases in which  $\alpha_0 + \beta_0$  is close to (but below) one (because the GARCH component becomes more volatile), the  $LM$  test has still very reasonable power to reject the null hypothesis. Thus, the simulation results suggest that the  $LM$  test can be applied even if  $\hat{\alpha} + \hat{\beta} \geq 1$  (as long as  $\alpha_0 + \beta_0 < 1$ ).

However, whenever  $\hat{\alpha} + \hat{\beta}$  is close to one, we suggest to investigate more than one source of potential misspecification. For example, it is well known that neglected parameter changes (see Hillebrand, 2005) or neglected long-memory (Baillie et al., 1996) can lead to the so-called ‘IGARCH effect’. In particular, if a researcher estimates  $\hat{\alpha} + \hat{\beta} \geq 1$ , he would not want to rely on the GARCH(1,1) specification but rather check for potential sources of misspecification. If  $\hat{\alpha} + \hat{\beta} \geq 1$  and the  $LM$  test rejects the null hypothesis, then the misspecification may indeed be due to an omitted long-term component. If  $\hat{\alpha} + \hat{\beta} \geq 1$  and the test does not reject, neglected parameter changes or neglected long-memory in the GARCH component are likely sources of the misspecification and should be tested for.

Finally, we performed several robustness checks.

**Linear long-term component:** In order to evaluate the effect of different choices for  $f(\cdot)$  on the power of the test statistic, we replaced the exponential specification of  $\tau_{0,t}$  with the linear specification in equation (31). The corresponding results for normally and

Student- $t$  distributed innovations can be found in Tables 9/10 in Appendix C.5 and, again, qualitatively confirm our previous findings. Note that (for the same weighting schemes as in the exponential specification) the linear specification leads to lower variance ratios which explains the difference in power.

**Sample size:** Given that a sample size of  $T = 1000$  is relatively modest for applications in financial econometrics, the power of the  $LM$  test is very satisfactory. However, in order to evaluate the effect on the power of increasing the sample size, we performed the same simulations as before but with  $T = 2000$ . As expected, in the larger sample the power of  $LM$  and  $LM_{LT,mod}$  increased substantially under all scenarios.

**Misspecification of the Short-Term Component:**

We also investigated how sensitive the  $LM$  test is with respect to a misspecification of the short-term component. A potential objection against our  $LM$  test might be that the test could ‘confuse’ a misspecified short-term component with an omitted long-term component. In Appendix C.6, we investigate the consequences of implementing the  $LM$  test under the null of a GARCH(1,1) although the true short-term component is a GJR-GARCH(1,1), a higher-order GARCH or a fractionally integrated GARCH (FIGARCH). In summary, the simulation results show that in the absence of a long-term component the  $LM$  test based on an exogenous explanatory variable,  $x_t$ , is insensitive with respect to a misspecified short-term component. When testing for remaining ARCH effects, both tests,  $LM$  and  $LM_{LT}$ , correctly detect a deviation from the null if the misspecification is sufficiently severe. That is, by choosing  $x_t = \varepsilon_t^2 / \hat{h}_t$ , our test is able to detect other forms of misspecification than omitted multiplicative components only (see also Remark 6). Finally, in the presence of a long-term component the power of the  $LM$  test decreases only slightly when using a misspecified GARCH model under  $H_0$ .

**Mixed-frequency test:** In Appendix C.7, we provide simulations for the mixed-frequency version of the test. In line with our arguments in Section 2.6, the simulations clearly show that predictive regressions based on  $\widehat{RV}_t$  are more powerful than predictive regressions based on  $RV_t$ .

## 4 Empirical Application

We consider two empirical applications. The first one deals with daily, weekly and monthly return data that are combined with explanatory variables which are available at the *same* frequency. The second one applies the test in a *mixed-frequency* setting. For both applications we use log-returns on the S&P 500.

### 4.1 Daily, Weekly and Monthly Data

First, we apply our test to four variables that are observed at a daily frequency and check whether these variables might be useful in a two-component GARCH specification. The data sample starts in December 1987 (i.e. after the crash in October 1987) and ends in June 2016. The first explanatory variable is the squared VXO,  $VXO_t$ . The VXO volatility index is disseminated by the Chicago Board Options Exchange (Cboe) and is based on S&P 100 index options. In contrast to the VIX, the VXO is already available from 1986 onwards. As a crude measure of daily realized variance (RV), we rely on the squared daily return,  $RV_t = \varepsilon_t^2$ . While the VXO and RV measure stock market volatility, we employ the index suggested by Baker et al. (2016) as a measure of general economic policy uncertainty (EPU,  $EPU_t$ ). Last, we use the ADS Business Conditions Index,  $ADS_t$ , suggested by Aruoba et al. (2009). This variable is meant to proxy for macroeconomic conditions. Data on the ADS were obtained from the website of the Federal Reserve Bank of Philadelphia. All other data were obtained from the FRED database maintained by the Federal Reserve Bank of St. Louis. In addition to the daily variables, we also consider the 22-days rolling window versions, defined as  $x_t^{(22)} = 1/22 \sum_{j=0}^{21} x_{t-j}$ . The contemporaneous correlations between the four variables are in line with intuition: financial volatility (VXO, RV) and political uncertainty (EPU) are positively correlated, while the three variables are negatively correlated with business conditions (ADS) (see Tables 14/15 in Appendix D).

For the daily log-returns on the S&P 500, we estimate a GJR-GARCH(1,1) and a GARCH(1,1). The parameter estimates for the GJR-GARCH(1,1) are

$$h_t = 0.019 + \underbrace{(0.002)}_{(0.004)} + \underbrace{0.136}_{(0.020)} \cdot \mathbf{1}_{\{\varepsilon_{t-1} < 0\}} \varepsilon_{t-1}^2 + \underbrace{0.913}_{(0.013)} h_{t-1}$$

with robust standard errors in parentheses. Thus, the estimates suggest that there is significant asymmetry in the volatility of S&P 500 returns. The parameter estimates for

the simple GARCH are

$$h_t = \underset{(0.004)}{0.014} + \underset{(0.011)}{0.075}\varepsilon_{t-1}^2 + \underset{(0.012)}{0.913}h_{t-1}.$$

We then apply our test with the GJR-GARCH under  $H_0$ ,  $LM^{GJR}$ , and alternatively with the GARCH under  $H_0$ ,  $LM^{GA}$ . In addition, we show the outcome of the  $LM_{LT,mod}$  test.

As Panel A of Table 3 shows, all three  $LM$  tests reject the null hypothesis for  $VXO_t$ ,  $EPU_t$  and  $ADS_t$  at the 1% level. Thus, these variables might be useful predictors of stock market volatility and could be drivers of an omitted second component. While  $LM^{GJR}$  rejects the null for  $RV_t$ ,  $LM^{GA}$  and  $LM_{LT,mod}$  do not. Overall, the test outcome is in line with the previous literature: long-term volatility is well known to be closely related to expected future volatility/uncertainty and behaves in a counter-cyclical fashion (see, e.g., Engle et al., 2013, Conrad and Loch, 2015a, and Dorion, 2016).

All three  $LM$  tests lead to the same conclusions when applied to the 22-day rolling window versions of the explanatory variables. The tests reject the null in case of  $VXO_t^{(22)}$  and  $ADS_t^{(22)}$ , but do not reject for  $RV_t^{(22)}$  and  $EPU_t^{(22)}$ . For all variables but  $ADS_t^{(22)}$ , the test statistics now take smaller values than when using the non-smoothed data. This is in line with the simulations in Section 3.2, which have shown that the power of the tests decreases when the explanatory variables are smoother.

It is important to note that (with only one exception), the  $LM$  tests based in the GJR-GARCH and the simple GARCH lead to the same conclusion. That is, although  $LM^{GA}$  is based on the misspecified GARCH model, this does not distort the test outcome.

Finally, we apply the  $LM$  test to weekly and monthly data. For this, we calculate weekly/monthly returns as the sum of the daily log-returns within each week/month. We construct weekly/monthly explanatory variables as the average of the explanatory variables within each week,  $x_t^{(W)}$ , or month  $x_t^{(M)}$ . For the weekly and monthly return data, the preferred specification is a simple GARCH with the following parameter estimates:

$$\begin{aligned} \text{weekly:} \quad h_t &= \underset{(0.061)}{0.117} + \underset{(0.025)}{0.084}\varepsilon_{t-1}^2 + \underset{(0.029)}{0.894}h_{t-1} \\ \text{monthly:} \quad h_t &= \underset{(0.315)}{0.592} + \underset{(0.045)}{0.143}\varepsilon_{t-1}^2 + \underset{(0.040)}{0.831}h_{t-1}. \end{aligned}$$

The time index  $t$  now refers to a weekly or monthly frequency. Panels B and C of Table 3 show that the  $LM^{GA}$  test now rejects the null hypothesis for all four variables. An explanation could be that the fraction of the total conditional variance that is due to

Table 3: *LM* test for S&P 500 returns.

$x_t$	Panel A: daily returns							
	$VXO_t$	$RV_t$	$EPU_t$	$ADS_t$	$VXO_t^{(22)}$	$RV_t^{(22)}$	$EPU_t^{(22)}$	$ADS_t^{(22)}$
$LM^{GJR}$	28.82 [<0.01]	5.39 [0.02]	6.97 [<0.01]	5.99 [0.01]	4.55 [0.03]	1.26 [0.25]	0.00 [0.96]	6.60 [0.01]
$LM^{GA}$	87.84 [<0.01]	0.83 [0.36]	17.22 [<0.01]	5.05 [0.02]	12.71 [<0.01]	0.13 [0.72]	1.56 [0.21]	6.30 [0.01]
$LM_{LT,mod}$	40.93 [<0.01]	0.93 [0.33]	9.76 [<0.01]	5.61 [0.02]	4.44 [0.04]	0.00 [0.97]	0.25 [0.62]	4.72 [0.03]
$x_t$	Panel B: weekly returns				Panel C: monthly returns			
	$VXO_t^{(W)}$	$RV_t^{(W)}$	$EPU_t^{(W)}$	$ADS_t^{(W)}$	$VXO_t^{(M)}$	$RV_t^{(M)}$	$EPU_t^{(M)}$	$ADS_t^{(M)}$
$LM^{GA}$	38.16 [<0.01]	34.21 [<0.01]	17.91 [<0.01]	7.97 [<0.01]	7.04 [0.01]	15.78 [<0.01]	9.56 [<0.01]	19.85 [<0.01]
$LM_{LT,mod}$	20.63 [<0.01]	29.12 [<0.01]	7.48 [0.01]	7.93 [<0.01]	2.98 [0.08]	8.76 [<0.01]	1.28 [0.25]	11.88 [<0.01]

**Notes:** The table reports test statistics for GJR-GARCH under  $H_0$ ,  $LM^{GJR}$ , for GARCH under  $H_0$ ,  $LM^{GA}$ , and for  $LM_{LT,mod}$ . Numbers in brackets are  $p$ -values. All tests are based on  $K = 1$  and the 1987M12 - 2016M06 period.

the long-term component and, hence, due to  $x_t$ , is larger for low-frequency (weekly or monthly) than for high-frequency (daily) returns.

In summary, our test results provide convincing evidence that a simple one-component GARCH is misspecified for the given sample. However, which variable and frequency should be selected for modelling the second component will ultimately depend on the specific application. For example, one variable might dominate when one is interested in forecasting tomorrow's conditional variance, but another one when the interest lies in forecasting next month's variance.

## 4.2 Mixed-Frequency Data

For the mixed-frequency application we use the same data as in Conrad and Loch (2015a). We construct quarterly realized variances  $RV_t$  from daily S&P 500 stock returns for the 1973Q1 to 2010Q4 period. Eleven macroeconomic variables are then used to test whether macroeconomic conditions can predict financial volatility. The macro variables are: real GDP, industrial production, the unemployment rate, housing starts, corporate profits, the GDP deflator, the Chicago Fed national activity index (NAI), the new orders index of the Institute for Supply Management, the University of Michigan consumer sentiment index, real personal consumption and the term spread. All variables are considered at

the quarterly frequency. We include the NAI and the new orders index in levels and take the first difference of the respective level for the unemployment rate and the consumer sentiment index. For all other variables, we calculate annualized quarter-over-quarter percentage changes. For a more detailed description of the macro variables see Section 3 in Conrad and Loch (2015a).

We focus on the predictive regression version of our test statistic. Conrad and Loch (2015a) estimate a predictive regression as in equation (42) and find that the  $\pi_1$  parameter estimate is insignificant for each macro variable (see their Section 4.4). This result is in line with the common notion that macro conditions do not help to forecast quarterly stock market volatility once one controls for lagged stock market volatility. We now show that this conclusion is premature. Following the discussion in Section 2.6, we first estimated equation (40) for the same data (again with  $K = 1$ ) and found a significant effect for six out of the eleven variables (results not reported). Table 4 shows that these results are robust to including the first lag of the volatility-adjusted realized variance as an additional regressor, i.e. we consider the regression:

$$\ln(\widehat{RV}_t) = \tilde{c} + \pi_1 x_{t-1} + \rho \ln(\widehat{RV}_{t-1}) + \tilde{\zeta}_t. \quad (43)$$

For simplicity, we rely on the critical values from the  $t$  distribution when testing for the significance of  $\pi_1$  although  $\widehat{RV}_t$  is an estimated quantity.

As Table 4 shows, real GDP, industrial production, the NAI and new orders are significant at the 1% level. The unemployment rate and corporate profits are significant at the 5% and 10% level. The fact that we do find a significant relationship between macro conditions and financial volatility when estimating equation (43) instead of equation (42) suggests that (the log of) the volatility-adjusted realized variance is indeed the appropriate dependent variable. When reestimating equation (43) by including more lags of the macro variables the picture remains the same. In conclusion, we provide strong evidence that the apparent inability of macro conditions to forecast financial volatility which is documented when using predictive regressions as in equation (42) seems to be driven by the fact that  $RV_t$  is a very noisy measure of the underlying long-term component which then masks the existing relationship.



Table 4: Predictive regressions.

$x_t$	$\tilde{c}$	$\pi_1$	$\rho$	adj. $R^2$
$\Delta$ real GDP	6.4007*** (0.6885)	-0.0107*** (0.0040)	0.1829** (0.0866)	4.81
$\Delta$ Ind. prod.	6.4174*** (0.6891)	-0.0052*** (0.0018)	0.1789** (0.0866)	4.85
$\Delta$ Unemp.	6.4461*** (0.6644)	0.0799** (0.0349)	0.1735** (0.0834)	4.39
$\Delta$ Housing	6.2327*** (0.7114)	-0.0004 (0.0004)	0.2014** (0.0894)	3.29
$\Delta$ Corp. prof.	6.3006*** (0.7037)	-0.0009* (0.0005)	0.1939** (0.0883)	4.14
$\Delta$ GDP deflator	6.2222*** (0.7196)	-0.0029 (0.0075)	0.2039** (0.0899)	2.94
NAI	6.5619*** (0.6545)	-0.0521*** (0.0173)	0.1586** (0.0823)	6.35
New orders	6.8507*** (0.7127)	-0.0058*** (0.0021)	0.1622*** (0.0866)	6.07
$\Delta$ Cons. sent.	6.2262*** (0.7171)	0.0010 (0.0034)	0.2020** (0.0900)	2.89
$\Delta$ real cons.	6.3365*** (0.7083)	-0.0072 (0.0060)	0.1905** (0.0895)	3.64
Term spread	6.3032*** (0.6857)	-0.0186 (0.0149)	0.1961** (0.0867)	3.84

**Notes:** The table reports parameter estimates for the predictive regression given by equation (43). Robust standard errors are presented in parentheses and \*\*\*, \*\*, \* indicate significance at the 1%, 5%, and 10% level. The adjusted  $R^2$  is reported in percentages. The sample covers the 1973Q1 - 2010Q4 period.

## 5 Conclusions

We develop a Lagrange Multiplier test for the null hypothesis of a one-component GARCH model against a multiplicative two-component specification. The test provides a first solution to statistically evaluate if there is a separate long-term time-varying volatility component driven by an explanatory variable, besides the standard short-term GARCH part. For the case of a GARCH(1,1) under the null, we derive the asymptotic properties of our test and study its finite sample performance. We also illustrate how the test can be implemented for other GARCH-type models under the null. In particular, we consider the GJR-GARCH which allows for asymmetry in the conditional variance. Our simulation results show that the  $LM$  test is robust with respect to a misspecification of the GARCH component and that it is powerful even if the GARCH component is highly persistent. The test covers the case that the returns as well as the explanatory variable are observed

at the same frequency but also the empirically relevant mixed-frequency setting. In an application to S&P 500 returns, we find that the test provides useful guidance in model specification. We also provide an explanation as to why standard predictive regressions might fail to find a relationship between macro conditions and financial volatility.

## References

- Amado, C., Silvennoinen, A., Teräsvirta, T., 2018. Models with multiplicative decomposition of conditional variances and correlations. CREATES Research Paper 2018-14.
- Amado, C., Teräsvirta, T., 2013. Modelling volatility by variance decomposition. *Journal of Econometrics* 175, 153-165.
- Amado, C., Teräsvirta, T., 2017. Specification and testing of multiplicative time-varying GARCH models with applications. *Econometric Reviews* 36, 421-446.
- Andreou, E., 2016. On the use of high frequency measures of volatility in MIDAS regressions. *Journal of Econometrics* 193, 367-389.
- Andreou E., Ghysels, E., Kourtellos, A., 2010. Regression models with mixed sampling frequencies. *Journal of Econometrics* 158, 246-261.
- Asgharian, H., Hou, A. J., Javed, F., 2013. The importance of the macroeconomic variables in forecasting stock return variance: a GARCH-MIDAS approach. *Journal of Forecasting* 32, 600-612.
- Aruoba, S., Diebold, F., Scotti, C., 2009. Real-time measurement of business conditions. *Journal of Business & Economic Statistics* 27, 417-427.
- Baker, S.R., Bloom, N., Davis, S.J., 2016. Measuring economic policy uncertainty. *The Quarterly Journal of Economics* 131, 1593-1636.
- Baillie, R.T., Bollerslev, T., Mikkelsen, H.O., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74, 3-30.
- Billingsley, P., 1961. The Lindeberg-Levy theorem for martingales. *Proceedings of the American Mathematical Society* 12, 788-792.

- Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307-327.
- Christiansen, C., Schmeling, M., Schrimpf, A., 2012. A comprehensive look at financial volatility prediction by economic variables. *Journal of Applied Econometrics* 27, 956-977.
- Conrad, C., Haag, B., 2006. Inequality constraints in the fractionally integrated GARCH model. *Journal of Financial Econometrics* 3, 413-449.
- Conrad, C., Loch, K., 2015a. Anticipating long-term stock market volatility. *Journal of Applied Econometrics* 30, 1090-1114.
- Conrad, C., Loch, K., 2015b. The variance risk premium and fundamental uncertainty. *Economics Letters* 132, 56-60.
- Demos, A., Sentana, E., 1998. Testing for GARCH effects: a one-sided approach. *Journal of Econometrics* 86, 97-127.
- Dorion, C., 2016. Option valuation with macro-finance variables. *Journal of Financial and Quantitative Analysis* 51, 1359-1389.
- Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987-1007.
- Engle, R.F., Rangel, J. G. 2008. The Spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies* 21, 1187-1222.
- Engle, R.F., Ghysels, E., Sohn, B., 2013. Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics* 95, 776-797.
- Engle, R.F., Ng, V.K., 1993. Measuring and testing the impact of news on volatility. *The Journal of Finance* 48, 1749-1778.
- Escanciano, J. C., 2009. Quasi-maximum likelihood estimation of semi-strong GARCH models. *Econometric Theory* 25, 561-570.
- Feng, Y., 2004. Simultaneously modeling conditional heteroskedasticity and scale change. *Econometric Theory* 20, 563-596.

- Francq, C., Thieu, L.Q., 2018. QML inference for volatility models with covariates. *Econometric Theory*, forthcoming.
- Francq, C., Zakoïan, J.-M., 2009. Testing the nullity of GARCH coefficients: correction of the standard tests and relative efficiency comparisons. *Journal of the American Statistical Association* 104, 313-324.
- Francq, C., Zakoïan, J.-M., 2007. Quasi-maximum likelihood in GARCH processes when some coefficients are equal to zero. *Stochastic Processes and their Applications* 117, 1265-1284.
- Francq, C., Zakoïan, J.-M., 2004. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* 10, 605-637.
- Ghysels, E., Sinko, A., Valkanov, R., 2007. MIDAS regressions: further results and new directions. *Econometric Reviews* 26, 53-90.
- Godfrey, L. G., 1978. Testing for multiplicative heteroskedasticity. *Journal of Econometrics* 8, 227-236.
- Halunga, A. G., Orme, C. D., 2009. First-order asymptotic theory for parametric misspecification tests of GARCH models. *Econometric Theory* 25, 364-410.
- Hamadeh, T., Zakoïan, J.-M., 2011. Asymptotic properties of LS and QML estimators for a class of nonlinear GARCH processes. *Journal of Statistical Planning and Inference* 141, 488-507.
- Han, H., Kristensen, D., 2014. Asymptotic theory for the QMLE in GARCH-X models with stationary and non-stationary covariates. *Journal of Business & Economic Statistics* 32, 416-429.
- Hillebrand, E., 2005. Neglecting parameter changes in GARCH models. *Journal of Econometrics* 129, 121-138.
- Lee, S., Halunga, A. G., 2015. Misspecification tests for Realised GARCH models. Working Paper. University of Exeter.

- 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.
- Li, M., Li, G., Li, W., 2011. Score tests for hyperbolic GARCH models. *Journal of Business & Economic Statistics* 29, 579-586.
- Ling, S., McAleer, M., 2003. Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory* 19, 280-310.
- Lundbergh, S., Teräsvirta, T., 2002. Evaluating GARCH models. *Journal of Econometrics* 110, 417-435.
- Nelson, D.B., Cao., C.Q., 1992. Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics* 10, 229-235.
- Opschoor, A., van Dijk, D., van der Wel, M., 2014. Predicting volatility and correlations with financial conditions indexes. *Journal of Empirical Finance* 29, 435-447.
- Paye, B. S., 2012. ‘Déja Vol’: Predictive regressions for aggregate stock market volatility using macroeconomic variables. *Journal of Financial Economics* 106, 527-546.
- Silvennoinen, A., Teräsvirta, T., 2016. Testing constancy of unconditional variance in volatility models by misspecification and specification tests. *Studies in Nonlinear Dynamics & Econometrics* 20, 347-364.
- Wang, F., Ghysels, E., 2015. Econometric analysis of volatility component models. *Econometric Theory* 31, 362-393.

# A Proofs

*Proof of Theorem 1.* First, we show that  $\mathbf{\Omega}$  is finite and positive definite.

## Finiteness of $\mathbf{\Omega}$ :

From Francq and Zakoïan (2004) it follows that  $\mathbf{\Omega}_{\eta\eta}$  is finite and positive definite. What remains to be shown is that  $\mathbf{\Omega}_{\pi\pi}$  is finite and positive definite. If this is true, then by the Cauchy-Schwarz inequality the “off-diagonal matrices” will also be finite and positive definite. Recall from equation (19) that  $\mathbf{\Omega}_{\pi\pi} = \frac{1}{4}(\kappa_Z - 1)\mathbf{E}[\mathbf{r}_{0,t}^\infty(\mathbf{r}_{0,t}^\infty)']$ . It follows from Assumption 2 that  $0 < \kappa_Z - 1 < \infty$ . Moreover,  $\|\mathbf{E}[\mathbf{r}_{0,t}^\infty(\mathbf{r}_{0,t}^\infty)']\|$  is finite if  $\mathbf{E}[\|\mathbf{r}_{0,t}^\infty(\mathbf{r}_{0,t}^\infty)'\|] < \infty$  (throughout the paper  $\|\cdot\|$  denotes the Euclidean norm). A typical element of the  $K \times 1$  vector  $\mathbf{r}_{0,t}^\infty$  is given by

$$r_{0,kt}^\infty = (x_{t-k} - \alpha_0 \frac{1}{h_{0,t}^\infty} \sum_{j=0}^{\infty} \beta_0^j \varepsilon_{t-1-j}^2 x_{t-1-k-j}) f_0'. \quad (44)$$

First,  $f_0'$  is bounded by Assumption 4 and  $\mathbf{E}[|x_{t-k}|^2] < \infty$  by Assumption 6. Second,

$$\left( \mathbf{E} \left| \frac{\sum_{j=0}^{\infty} \alpha_0 \beta_0^j \varepsilon_{t-1-j}^2 x_{t-1-k-j}}{h_{0,t}^\infty} \right|^2 \right)^{1/2} \leq \left( \sum_{j=0}^{\infty} \mathbf{E} \left| \frac{\alpha_0 \beta_0^j \varepsilon_{t-1-j}^2 x_{t-1-k-j}}{h_{0,t}^\infty} \right|^2 \right)^{1/2} \quad (45)$$

$$\leq \sum_{j=0}^{\infty} \left( \mathbf{E} \left| \frac{\alpha_0 \beta_0^j \varepsilon_{t-1-j}^2}{(\omega_0 + \alpha_0 \beta_0^j \varepsilon_{t-1-j}^2)} x_{t-1-k-j} \right|^2 \right)^{1/2} \quad (46)$$

$$\leq \sum_{j=0}^{\infty} \left( \mathbf{E} \left| \left( \frac{\alpha_0 \beta_0^j}{\omega_0} \varepsilon_{t-1-j}^2 \right)^s x_{t-1-k-j} \right|^2 \right)^{1/2} \quad (47)$$

$$\leq \frac{\alpha_0^s}{\omega_0^s} (\mathbf{E} [|\varepsilon_{t-1-j}|^{4sp}])^{1/(2p)} (\mathbf{E} [|x_{t-1-k-j}|^{2q}])^{1/(2q)}$$

$$\sum_{j=0}^{\infty} \beta_0^{js} < \infty$$

for any  $p > 1$  and  $q > 1$  such that  $p^{-1} + q^{-1} = 1$ . The arguments used above are similar to the ones in Francq and Zakoïan (2004, Eq. (4.19), p.619). In particular, in equation (45) we employ Minkowski's inequality. In equation (46) we use that  $h_{0,t}^\infty \geq \omega_0 + \alpha_0 \beta_0^j \varepsilon_{t-1-j}^2$ . Next, in equation (47) we use the fact that  $w/(1+w) \leq w^s$  for all  $w > 0$  and any  $s \in (0, 1)$ . In the next equation, we apply the Hölder inequality. Finally, Assumptions 1 and 2 imply that under the null there exists some  $u > 0$  such that  $\mathbf{E} [|\varepsilon_{t-1-j}|^{2u}] < \infty$  (see Proposition

1 in Francq and Zakoian, 2004, p.607). Thus, for any  $p > 1$ , we can always choose an  $s$  small enough such that  $2sp = u$ . By Assumption 6,  $\mathbf{E} [|x_{t-1-k-j}|^{2q}] < \infty$ .

This implies  $\mathbf{E}[|r_{0,kt}^\infty|^2] < \infty$  and  $\mathbf{E}[|r_{0,kt}^\infty r_{0,jt}^\infty|] < \infty$  by Cauchy-Schwarz inequality which yields that  $\boldsymbol{\Omega}_{\pi\pi}$  is finite.

**Positive definiteness of  $\boldsymbol{\Omega}$ :**

As  $\kappa_Z - 1 > 0$ , it remains to be shown that  $\mathbf{c}'\mathbf{E} \left[ \begin{pmatrix} \mathbf{y}_{0,t}^\infty \\ \mathbf{r}_{0,t}^\infty \end{pmatrix} \begin{pmatrix} (\mathbf{y}_{0,t}^\infty)' & (\mathbf{r}_{0,t}^\infty)' \end{pmatrix} \right] \mathbf{c} > 0$  for any non-zero  $\mathbf{c} \in \mathbb{R}^{(3+K) \times 1}$ . Assume the contrary, i.e., there exists a  $\mathbf{c} \neq \mathbf{0}$  such that  $\mathbf{c}'\mathbf{E} \left[ \begin{pmatrix} \mathbf{y}_{0,t}^\infty \\ \mathbf{r}_{0,t}^\infty \end{pmatrix} \begin{pmatrix} (\mathbf{y}_{0,t}^\infty)' & (\mathbf{r}_{0,t}^\infty)' \end{pmatrix} \right] \mathbf{c} = 0$ . This implies  $\mathbf{E} \left[ \left( \mathbf{c}' \begin{pmatrix} \mathbf{y}_{0,t}^\infty \\ \mathbf{r}_{0,t}^\infty \end{pmatrix} \right)^2 \right] = 0$  and, thus,  $\mathbf{c}' \begin{pmatrix} \mathbf{y}_{0,t}^\infty \\ \mathbf{r}_{0,t}^\infty \end{pmatrix} = 0$  a.s.. The last expression can be written as

$$0 = \mathbf{c}' \begin{pmatrix} (h_{0,t}^\infty)^{-1} \mathbf{s}_{0,t}^\infty \\ f_0' \mathbf{x}_t \end{pmatrix} + \mathbf{c}' \begin{pmatrix} (h_{0,t}^\infty)^{-1} \beta_0 \frac{\partial \bar{h}_{t-1}^\infty}{\partial \boldsymbol{\eta}} \Big|_{\boldsymbol{\pi}=\mathbf{0}} \\ -f_0' \alpha_0 (h_{0,t}^\infty)^{-1} \varepsilon_{t-1}^2 \mathbf{x}_{t-1} + f_0' (h_{0,t}^\infty)^{-1} \beta_0 \frac{\partial \bar{h}_{t-1}^\infty}{\partial \boldsymbol{\pi}} \Big|_{\boldsymbol{\pi}=\mathbf{0}} \end{pmatrix}. \quad (48)$$

Using the notation  $\mathbf{c} = (\mathbf{c}'_1 \ \mathbf{c}'_2)'$  where  $\mathbf{c}_1 = (c_{11} \ c_{12} \ c_{13})'$  and  $\mathbf{c}_2 = (c_{21} \dots \ c_{2K})'$  this can be expressed as

$$\mathbf{c}'_1 \mathbf{s}_{0,t}^\infty + f_0' h_{0,t}^\infty \mathbf{c}'_2 \mathbf{x}_t - f_0' \alpha_0 \varepsilon_{t-1}^2 (\mathbf{c}'_2 \mathbf{x}_{t-1}) = - \left( \beta_0 \mathbf{c}'_1 \frac{\partial \bar{h}_{t-1}^\infty}{\partial \boldsymbol{\eta}} \Big|_{\boldsymbol{\pi}=\mathbf{0}} + f_0' \beta_0 \mathbf{c}'_2 \frac{\partial \bar{h}_{t-1}^\infty}{\partial \boldsymbol{\pi}} \Big|_{\boldsymbol{\pi}=\mathbf{0}} \right) \quad (49)$$

or

$$\begin{aligned} c_{11} + c_{12} Z_{t-1}^2 h_{0,t-1}^\infty + f_0' (\omega_0 + \alpha_0 Z_{t-1}^2 h_{0,t-1}^\infty + \beta_0 h_{0,t-2}^\infty) (\mathbf{c}'_2 \mathbf{x}_t) - f_0' \alpha_0 Z_{t-1}^2 h_{0,t-1}^\infty (\mathbf{c}'_2 \mathbf{x}_{t-1}) \\ = -c_{13} h_{0,t-1}^\infty - \left( \beta_0 \mathbf{c}'_1 \frac{\partial \bar{h}_{t-1}^\infty}{\partial \boldsymbol{\eta}} \Big|_{\boldsymbol{\pi}=\mathbf{0}} + f_0' \beta_0 \mathbf{c}'_2 \frac{\partial \bar{h}_{t-1}^\infty}{\partial \boldsymbol{\pi}} \Big|_{\boldsymbol{\pi}=\mathbf{0}} \right) = F_{t-2}, \end{aligned}$$

where  $F_{t-2}$  is a measurable function of  $\{Z_{t-1-j}, \mathbf{x}_{t-1-j}, j \geq 1\}$ . This implies that the expression in the upper line must be degenerate. Hence,

$$Z_{t-1}^2 = \frac{-c_{11} + f_0' (\omega_0 + \beta_0 h_{0,t-2}^\infty) (\mathbf{c}'_2 \mathbf{x}_t)}{h_{0,t-1}^\infty (c_{12} - f_0' \alpha_0 (\mathbf{c}'_2 \mathbf{x}_{t-1}) + f_0' \alpha_0)} = A_{t-2} + B_{t-2} (\mathbf{c}'_2 \mathbf{x}_t)$$

with  $A_{t-2}$  and  $B_{t-2}$  measurable functions of  $\{Z_{t-1-j}, \mathbf{x}_{t-1-j}, j \geq 1\}$  is degenerate. This equation could only be fulfilled either is left and right hand side are both degenerate, or  $\mathbf{c}'_2 \mathbf{x}_t$  is a linear function of  $Z_{t-1}^2$ . The latter case, however, implies that  $Z_{t-1}^2$  is measurable

with respect to  $\{Z_{t-1-j}, \mathbf{x}_{t-1-j}, j \geq 1\}$  which contradicts Assumption 2. The former case is ruled out since  $\mathbf{c}'_2 \mathbf{x}_t$  is non-degenerate by Assumption 5 and  $Z_t^2$  is non-degenerate by Assumption 2. Thus,  $\mathbf{\Omega}$  must be positive definite.

Next,  $\mathbf{E}[\mathbf{d}_t^\infty(\boldsymbol{\eta}_0)|\mathcal{F}_{t-1}] = \mathbf{0}$ . From Francq and Zakoian (2004) and Assumptions 1-6 it then follows that  $\mathbf{d}_t^\infty(\boldsymbol{\eta}_0)$  is a stationary and ergodic martingale difference sequence with finite second moment. Applying Billingsley's (1961) central limit theorem for martingale differences gives the result.

■

The following proposition will be used in the proof of Theorem 2.

**Proposition 1.** *Under Assumptions 1-7, we have that*

$$-\frac{1}{T} \sum_{t=1}^T \frac{\partial \mathbf{d}_{\pi,t}^\infty(\tilde{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}'} \xrightarrow{P} \mathbf{J}_{\pi\boldsymbol{\eta}} = -\mathbf{E} \left[ \frac{\partial \mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta}_0)}{\partial \boldsymbol{\eta}'} \right], \quad (50)$$

where  $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}_0 + o_P(1)$ .

*Proof of Proposition 1.* We obtain (50) by showing that  $\mathbf{J}_{\pi\boldsymbol{\eta}}(\boldsymbol{\eta}) = -\mathbf{E} \left[ \frac{\partial \mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'} \right]$  is finite with a uniform bound for all  $\boldsymbol{\eta} \in \Theta$ . Then a uniform weak law of large numbers (see, e.g., Theorem 3.1. in Ling and McAleer, 2003) implies

$$\sup_{\boldsymbol{\eta}} \left\| -\frac{1}{T} \sum_{t=1}^T \frac{\partial \mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'} - \mathbf{J}_{\pi\boldsymbol{\eta}}(\boldsymbol{\eta}) \right\| = o_P(1).$$

Equation (50) follows from the triangle inequality and the fact that  $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}_0 + o_P(1)$ .

Using equation (21) we obtain

$$\begin{aligned} \left\| \frac{\partial \mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'} \right\| &\leq \frac{1}{2} \left( \left| \frac{\varepsilon_t^2}{h_t^\infty} \right| \cdot \|\mathbf{r}_t^\infty\| \cdot \|(\mathbf{y}_t^\infty)'\| + \left| \frac{\varepsilon_t^2}{h_t^\infty} - 1 \right| \cdot \left\| \frac{\partial \mathbf{r}_t^\infty}{\partial \boldsymbol{\eta}'} \right\| \right) \\ &\leq C|\varepsilon_t^2 + \omega| \left( \|\mathbf{r}_t^\infty\| \cdot \|(\mathbf{y}_t^\infty)'\| + \left\| \frac{\partial \mathbf{r}_t^\infty}{\partial \boldsymbol{\eta}'} \right\| \right). \end{aligned} \quad (51)$$

The last inequality follows with a generic constant  $0 < C < \infty$  and  $h_t^\infty \geq \omega > 0$ .

First, consider the three elements of  $\|(\mathbf{y}_t^\infty)'\|$ . To simplify the notation note that  $\frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\eta}}|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{\partial h_t^\infty}{\partial \boldsymbol{\eta}}$ . Since  $\frac{\partial h_t^\infty}{\partial \omega} = 1/(1-\beta)$ , we have  $|\frac{1}{h_t^\infty} \frac{\partial h_t^\infty}{\partial \omega}| \leq 1/(\omega(1-\beta)) < \infty$ . Then  $\alpha \frac{\partial h_t^\infty}{\partial \alpha} = \sum_{j=0}^{\infty} \alpha \beta^j \varepsilon_{t-1-j}^2 \leq h_t^\infty$  and, therefore,  $|\frac{1}{h_t^\infty} \frac{\partial h_t^\infty}{\partial \alpha}| \leq 1/\alpha < \infty$ . Finally,  $\frac{\partial h_t^\infty}{\partial \beta} =$



$\sum_{j=0}^{\infty} j\beta^{j-1}(\omega + \alpha\varepsilon_{t-1-j}^2)$ . We then obtain

$$\begin{aligned} \left| \frac{1}{h_t^\infty} \frac{\partial h_t^\infty}{\partial \beta} \right| &\leq \left| \frac{1}{\beta} \sum_{j=0}^{\infty} \frac{j\beta^j(\omega + \alpha\varepsilon_{t-1-j}^2)}{\omega + \beta^j(\omega + \alpha\varepsilon_{t-1-j}^2)} \right| \\ &\leq \frac{1}{\beta\omega^s} \sum_{j=0}^{\infty} j |\beta^{js}(\omega + \alpha\varepsilon_{t-1-j}^2)^s|, \end{aligned} \quad (52)$$

where we again use the fact that  $w/(1+w) \leq w^s$  for all  $w > 0$  and any  $s \in (0, 1)$ . It follows that  $\|(\mathbf{y}_t^\infty)'\| \leq C'(1 + \sum_{j=0}^{\infty} j |\beta^{js}(\omega + \alpha\varepsilon_{t-1-j}^2)^s|)$  for some constant  $C' > 0$ .

Hence, using the Cauchy-Schwarz inequality, the first summand in equation (51), i.e.  $\mathbf{E} [\sup_{\boldsymbol{\eta}} |\varepsilon_t^2 + \omega| \cdot \|\mathbf{r}_t^\infty\| \cdot \|(\mathbf{y}_t^\infty)'\|]$ , can be bounded from above by the terms

$$\sqrt{\mathbf{E}[\sup_{\boldsymbol{\eta}} |\varepsilon_t^2 + \omega|^2] \mathbf{E}[\sup_{\boldsymbol{\eta}} \|\mathbf{r}_t^\infty\|^2]} \quad (53)$$

and

$$\begin{aligned} &\sup_{\boldsymbol{\eta}} \sum_{j=0}^{\infty} j\beta^{js} \mathbf{E}[\sup_{\boldsymbol{\eta}} (\omega + \alpha\varepsilon_{t-1-j}^2)^s |\varepsilon_t^2 + \omega| \|\mathbf{r}_t^\infty\|] \leq \\ &\sup_{\boldsymbol{\eta}} \sum_{j=0}^{\infty} j\beta^{js} \sqrt{\mathbf{E}[\sup_{\boldsymbol{\eta}} (\omega + \alpha\varepsilon_{t-1-j}^2)^{2s} |\varepsilon_t^2 + \omega|^2] \mathbf{E}[\sup_{\boldsymbol{\eta}} \|\mathbf{r}_t^\infty\|^2]}. \end{aligned} \quad (54)$$

The finiteness of (53) follows from Assumption 7 and similar arguments as in the proof of Theorem 1. The finiteness of (54) follows by applying Hölder's inequality, since for the elements in the sum which involve expectations of the squared observations we have

$$\begin{aligned} &\mathbf{E}[\sup_{\boldsymbol{\eta}} (\omega + \alpha\varepsilon_{t-1-j}^2)^{2s} |\varepsilon_t^2 + \omega|^2] \leq \\ &(\mathbf{E}[\sup_{\boldsymbol{\eta}} (\omega + \alpha\varepsilon_{t-1-j}^2)^{2(1+s)}])^{s/(1+s)} (\mathbf{E}[\sup_{\boldsymbol{\eta}} |\varepsilon_t^2 + \omega|^{2(1+s)}])^{1/(1+s)} \end{aligned} \quad (55)$$

and Assumption 7 applies again.

Using the Cauchy-Schwarz inequality for the two factors in the second term in (51), we are left with the need to show that  $\mathbf{E} \left[ \sup_{\boldsymbol{\eta}} \left\| \frac{\partial \mathbf{r}_t^\infty}{\partial \boldsymbol{\eta}'} \right\|^2 \right]$  is finite. This follows from

$$\begin{aligned} (f'_0)^{-1} \frac{\partial \mathbf{r}_t^\infty}{\partial \boldsymbol{\eta}'} &= \frac{\partial}{\partial \boldsymbol{\eta}'} \mathbf{x}_t - \frac{\partial}{\partial \boldsymbol{\eta}'} \left( \frac{1}{h_t^\infty} \sum_{j=0}^{\infty} \alpha\beta^j \varepsilon_{t-1-j}^2 \mathbf{x}_{t-1-j} \right) \\ &= \frac{\partial}{\partial \boldsymbol{\eta}'} \mathbf{x}_t - \frac{1}{h_t^\infty} \left( \sum_{j=0}^{\infty} \alpha\beta^j \varepsilon_{t-1-j}^2 \frac{\partial}{\partial \boldsymbol{\eta}'} \mathbf{x}_{t-1-j} \right) \\ &\quad + \left( \frac{1}{h_t^\infty} \sum_{j=0}^{\infty} \alpha\beta^j \varepsilon_{t-1-j}^2 \mathbf{x}_{t-1-j} \right) (\mathbf{y}_t^\infty)' \\ &\quad - \frac{1}{h_t^\infty} \sum_{j=0}^{\infty} \mathbf{x}_{t-1-j} \left( \frac{\partial}{\partial \boldsymbol{\eta}'} \alpha\beta^j \varepsilon_{t-1-j}^2 \right). \end{aligned} \quad (56)$$

The first two terms vanish in the model with an explanatory variable  $\mathbf{x}_t$  from outside the model as  $\frac{\partial \mathbf{x}_t}{\partial \boldsymbol{\eta}'} = \mathbf{0}$  or in a model with  $x_{t-k} = \varepsilon_{t-k}^2$ .

**Remark 7.** *There also exists a bound for  $\mathbf{E} \left[ \sup_{\boldsymbol{\eta}} \left\| \frac{\partial \mathbf{r}_t^\infty}{\partial \boldsymbol{\eta}'} \right\|^2 \right]$  in the case of  $\mathbf{x}_t$  with elements  $x_{t-k} = \frac{\varepsilon_{t-k}^2}{h_{t-k}^\infty}$  (the ‘ARCH nested in GARCH’ case). Here, in the first two terms in equation (56) we have  $\frac{\partial x_{t-k}}{\partial \boldsymbol{\eta}'} = -\frac{\varepsilon_{t-k}}{(h_{t-k}^\infty)^2} \frac{\partial h_{t-k}^\infty}{\partial \boldsymbol{\eta}'}$  and, hence, explicit bounds for terms of this type can be obtained as before.*

Boundedness of the norm of the third term follows for all  $\boldsymbol{\eta}$  in expectation with a combination of the argument directly above and the considerations in the proof of Theorem 1.

The fourth term can be written as:

$$\frac{1}{h_t^\infty} \begin{pmatrix} 0 & \sum_{j=0}^{\infty} \beta^j \varepsilon_{t-1-j}^2 x_{t-2-j} & \alpha \sum_{j=0}^{\infty} j \beta^{j-1} \varepsilon_{t-1-j}^2 x_{t-2-j} \\ 0 & \sum_{j=0}^{\infty} \beta^j \varepsilon_{t-1-j}^2 x_{t-3-j} & \alpha \sum_{j=0}^{\infty} j \beta^{j-1} \varepsilon_{t-1-j}^2 x_{t-3-j} \\ \vdots & \vdots & \vdots \\ 0 & \sum_{j=0}^{\infty} \beta^j \varepsilon_{t-1-j}^2 x_{t-1-K-j} & \alpha \sum_{j=0}^{\infty} j \beta^{j-1} \varepsilon_{t-1-j}^2 x_{t-1-K-j} \end{pmatrix}. \quad (57)$$

Hence, for typical elements of the second and third column it follows that

$$\mathbf{E} \sup_{\boldsymbol{\eta}} \left| \frac{1}{h_t^\infty} \sum_{j=0}^{\infty} \beta^j \varepsilon_{t-1-j}^2 x_{t-1-k-j} \right|^2 < \infty$$

and

$$\mathbf{E} \sup_{\boldsymbol{\eta}} \left| \frac{1}{h_t^\infty} \alpha \sum_{j=0}^{\infty} j \beta^{j-1} \varepsilon_{t-1-j}^2 x_{t-1-k-j} \right|^2 < \infty$$

by similar arguments as used before.

■

*Proof of Theorem 2.* First, consider a mean value expansion of  $\sqrt{T} \mathbf{D}_{\boldsymbol{\eta}}^\infty(\hat{\boldsymbol{\eta}})$  around the true value  $\boldsymbol{\eta}_0$

$$\mathbf{0} = \sqrt{T} \mathbf{D}_{\boldsymbol{\eta}}^\infty(\hat{\boldsymbol{\eta}}) = \sqrt{T} \mathbf{D}_{\boldsymbol{\eta}}^\infty(\boldsymbol{\eta}_0) + \frac{1}{T} \sum_{t=1}^T \frac{\partial \mathbf{d}_{\boldsymbol{\eta},t}^\infty(\tilde{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}'} \sqrt{T}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) \quad (58)$$

with  $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}_0 + o_P(1)$ . Under Assumptions 1 and 2, Francq and Zakoian (2004) have shown that

$$-\frac{1}{T} \sum_{t=1}^T \frac{\partial \mathbf{d}_{\boldsymbol{\eta},t}^\infty(\tilde{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}'} \xrightarrow{P} \mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}} = -\mathbf{E} \left[ \frac{\partial \mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0)}{\partial \boldsymbol{\eta}'} \right] \quad (59)$$

and, hence, equation (58) can be written as

$$\sqrt{T}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) = \mathbf{J}_{\eta\eta}^{-1} \sqrt{T} \mathbf{D}_{\eta}^{\infty}(\boldsymbol{\eta}_0) + o_P(1). \quad (60)$$

Similarly, a mean value expansion of  $\sqrt{T} \mathbf{D}_{\pi}^{\infty}(\hat{\boldsymbol{\eta}})$  around the true value  $\boldsymbol{\eta}_0$  leads to

$$\sqrt{T} \mathbf{D}_{\pi}^{\infty}(\hat{\boldsymbol{\eta}}) = \sqrt{T} \mathbf{D}_{\pi}^{\infty}(\boldsymbol{\eta}_0) + \frac{1}{T} \sum_{t=1}^T \frac{\partial \mathbf{d}_{\pi,t}^{\infty}(\tilde{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}'} \sqrt{T}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0). \quad (61)$$

Combining equation (60) and Proposition 1 leads to

$$\sqrt{T} \mathbf{D}_{\pi}^{\infty}(\hat{\boldsymbol{\eta}}) = \sqrt{T} \mathbf{D}_{\pi}^{\infty}(\boldsymbol{\eta}_0) - \mathbf{J}_{\pi\eta} \mathbf{J}_{\eta\eta}^{-1} \sqrt{T} \mathbf{D}_{\eta}^{\infty}(\boldsymbol{\eta}_0) + o_P(1) \quad (62)$$

$$= [-\mathbf{J}_{\pi\eta} \mathbf{J}_{\eta\eta}^{-1} : \mathbf{I}] \sqrt{T} \begin{pmatrix} \mathbf{D}_{\eta}^{\infty}(\boldsymbol{\eta}_0) \\ \mathbf{D}_{\pi}^{\infty}(\boldsymbol{\eta}_0) \end{pmatrix} + o_P(1) \quad (63)$$

$$= [-\mathbf{J}_{\pi\eta} \mathbf{J}_{\eta\eta}^{-1} : \mathbf{I}] \sqrt{T} \mathbf{D}^{\infty}(\boldsymbol{\eta}_0) + o_P(1). \quad (64)$$

Applying Theorem 1 gives the asymptotic distribution

$$\sqrt{T} \mathbf{D}_{\pi}^{\infty}(\hat{\boldsymbol{\eta}}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, [\mathbf{J}_{\pi\eta} \mathbf{J}_{\eta\eta}^{-1} : \mathbf{I}] \boldsymbol{\Omega} [\mathbf{J}_{\pi\eta} \mathbf{J}_{\eta\eta}^{-1} : \mathbf{I}]') \quad (65)$$

which has the form of  $\mathbf{A} \boldsymbol{\Omega} \mathbf{A}'$  in Halunga and Orme (2009, p.372/373). The covariance matrix can be written as

$$\begin{aligned} \boldsymbol{\Sigma} &= [-\mathbf{J}_{\pi\eta} \mathbf{J}_{\eta\eta}^{-1} : \mathbf{I}] \boldsymbol{\Omega} [-\mathbf{J}_{\pi\eta} \mathbf{J}_{\eta\eta}^{-1} : \mathbf{I}]' \\ &= \boldsymbol{\Omega}_{\pi\pi} + \mathbf{J}_{\pi\eta} \mathbf{J}_{\eta\eta}^{-1} \boldsymbol{\Omega}_{\eta\eta} \mathbf{J}_{\eta\eta}^{-1} \mathbf{J}'_{\pi\eta} - \mathbf{J}_{\pi\eta} \mathbf{J}_{\eta\eta}^{-1} \boldsymbol{\Omega}_{\eta\pi} - \boldsymbol{\Omega}_{\pi\eta} \mathbf{J}_{\eta\eta}^{-1} \mathbf{J}'_{\pi\eta}. \end{aligned}$$

Finally, using equations (19), (22) and (23) the expression for  $\boldsymbol{\Sigma}$  simplifies to:

$$\boldsymbol{\Sigma} = \frac{1}{4} (\kappa_Z - 1) \left( \mathbf{E}[\mathbf{r}_{0,t}^{\infty} (\mathbf{r}_{0,t}^{\infty})'] - \mathbf{E}[\mathbf{r}_{0,t}^{\infty} (\mathbf{y}_{0,t}^{\infty})'] (\mathbf{E}[\mathbf{y}_{0,t}^{\infty} (\mathbf{y}_{0,t}^{\infty})'])^{-1} \mathbf{E}[\mathbf{y}_{0,t}^{\infty} (\mathbf{r}_{0,t}^{\infty})'] \right). \quad (66)$$

■

*Proof of Theorem 3.* We show that

$$\sqrt{T} \mathbf{D}_{\pi}(\hat{\boldsymbol{\eta}}) = \sqrt{T} \mathbf{D}_{\pi}^{\infty}(\hat{\boldsymbol{\eta}}) + o_P(1). \quad (67)$$

Hence, the observed quantity  $\sqrt{T} \mathbf{D}_{\pi}(\hat{\boldsymbol{\eta}})$  will have the same asymptotic distribution as the unobserved  $\sqrt{T} \mathbf{D}_{\pi}^{\infty}(\hat{\boldsymbol{\eta}})$ . The asymptotic distribution of the test statistic then follows directly from Theorem 2. Standardization with the consistent estimator  $\hat{\boldsymbol{\Sigma}}$  instead of the

theoretical  $\Sigma$ , has no effect on the final  $\chi^2$ -distribution of the  $LM$  test statistic. This can be easily seen from similar considerations as the ones outlined above and below in detail.

Since

$$\sup_{\boldsymbol{\eta}} \|\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\boldsymbol{\eta}) - \sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}(\boldsymbol{\eta})\| \leq \frac{1}{\sqrt{T}} \sum_{t=1}^T \sup_{\boldsymbol{\eta}} \|\mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\boldsymbol{\eta}) - \mathbf{d}_{\boldsymbol{\pi},t}(\boldsymbol{\eta})\|, \quad (68)$$

we establish equation (67) by showing that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \sup_{\boldsymbol{\eta}} \|\mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\boldsymbol{\eta}) - \mathbf{d}_{\boldsymbol{\pi},t}(\boldsymbol{\eta})\| = o_P(1). \quad (69)$$

Consider the following decomposition:

$$\begin{aligned} 2(\mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\boldsymbol{\eta}) - \mathbf{d}_{\boldsymbol{\pi},t}(\boldsymbol{\eta})) &= \left(\frac{\varepsilon_t^2}{h_t^{\infty}} - 1\right) \mathbf{r}_t^{\infty} - \left(\frac{\varepsilon_t^2}{h_t} - 1\right) \mathbf{r}_t \\ &= \left(\frac{\varepsilon_t^2}{h_t^{\infty}} - 1\right) \mathbf{r}_t^{\infty} - \left(\frac{\varepsilon_t^2}{h_t} - 1\right) \mathbf{r}_t + \left[\left(\frac{\varepsilon_t^2}{h_t} - 1\right) \mathbf{r}_t^{\infty} - \left(\frac{\varepsilon_t^2}{h_t} - 1\right) \mathbf{r}_t^{\infty}\right] \\ &= \left(\frac{\varepsilon_t^2}{h_t^{\infty}} - \frac{\varepsilon_t^2}{h_t}\right) \mathbf{r}_t^{\infty} + \left(\frac{\varepsilon_t^2}{h_t} - 1\right) (\mathbf{r}_t^{\infty} - \mathbf{r}_t) \\ &= \varepsilon_t^2 \left(\frac{h_t - h_t^{\infty}}{h_t^{\infty} h_t}\right) \mathbf{r}_t^{\infty} + \left(\frac{\varepsilon_t^2}{h_t} - 1\right) (\mathbf{r}_t^{\infty} - \mathbf{r}_t) + \\ &\quad \left[\left(\frac{\varepsilon_t^2}{h_t^{\infty}} - 1\right) (\mathbf{r}_t^{\infty} - \mathbf{r}_t) - \left(\frac{\varepsilon_t^2}{h_t^{\infty}} - 1\right) (\mathbf{r}_t^{\infty} - \mathbf{r}_t)\right] \\ &= \varepsilon_t^2 \left(\frac{h_t - h_t^{\infty}}{h_t^{\infty} h_t}\right) \mathbf{r}_t^{\infty} - \varepsilon_t^2 \left(\frac{h_t - h_t^{\infty}}{h_t^{\infty} h_t}\right) (\mathbf{r}_t^{\infty} - \mathbf{r}_t) + \left(\frac{\varepsilon_t^2}{h_t^{\infty}} - 1\right) (\mathbf{r}_t^{\infty} - \mathbf{r}_t). \end{aligned}$$

Since  $h_t \geq \omega > 0$  and  $h_t^{\infty} \geq \omega > 0$  we have

$$\|\mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\boldsymbol{\theta}) - \mathbf{d}_{\boldsymbol{\pi},t}(\boldsymbol{\theta})\| \leq \frac{1}{\omega} \left\{ |\varepsilon_t^2 + \omega| \|\mathbf{r}_t^{\infty} - \mathbf{r}_t\| + \varepsilon_t^2 \|\mathbf{r}_t^{\infty}\| \left| \frac{h_t^{\infty} - h_t}{h_t^{\infty}} \right| + \varepsilon_t^2 \|\mathbf{r}_t^{\infty} - \mathbf{r}_t\| \left| \frac{h_t^{\infty} - h_t}{h_t^{\infty}} \right| \right\}.$$

First, note that

$$(f_0')^{-1}(\mathbf{r}_t^{\infty} - \mathbf{r}_t) = -\alpha \frac{1}{h_t^{\infty}} \sum_{j=t}^{\infty} \beta^j \varepsilon_{t-1-j}^2 \mathbf{x}_{t-1-j}. \quad (70)$$

Next, consider a typical element:

$$\begin{aligned}
(f'_0)^{-1} (\mathbf{E} \sup_{\boldsymbol{\eta}} |r_{k,t}^\infty - r_{k,t}|^2)^{1/2} &= \left( \mathbf{E} \sup_{\boldsymbol{\eta}} \left| \alpha \frac{1}{h_t^\infty} \sum_{j=t}^{\infty} \beta^j \varepsilon_{t-1-j}^2 x_{t-1-k-j} \right|^2 \right)^{1/2} \\
&\leq \sum_{j=t}^{\infty} \left( \mathbf{E} \sup_{\boldsymbol{\eta}} \left| \frac{\alpha \beta^j \varepsilon_{t-1-j}^2}{\omega + \alpha \beta^j \varepsilon_{t-1-k-j}^2} x_{t-1-k-j} \right|^2 \right)^{1/2} \\
&\leq \sum_{j=t}^{\infty} \left( \mathbf{E} \sup_{\boldsymbol{\eta}} \left| \left( \frac{\alpha \beta^j}{\omega} \varepsilon_{t-1-j}^2 \right)^s x_{t-1-k-j} \right|^2 \right)^{1/2} \\
&\leq (\mathbf{E}[|\varepsilon_{t-1-j}|^{4sp}])^{1/(2p)} (\mathbf{E}[|x_{t-1-k-j}|^{2q}])^{1/(2q)} \\
&\quad \sup_{\boldsymbol{\eta}} \left( \frac{\alpha}{\omega} \right)^s \sum_{j=t}^{\infty} \beta^{js} \tag{71}
\end{aligned}$$

$$\begin{aligned}
&= (\mathbf{E}[|\varepsilon_{t-1-j}|^{4sp}])^{1/(2p)} (\mathbf{E}[|x_{t-1-k-j}|^{2q}])^{1/(2q)} \\
&\quad \sup_{\boldsymbol{\eta}} \left( \frac{\alpha}{\omega} \right)^s \frac{\beta^{st}}{1 - \beta^s}, \tag{72}
\end{aligned}$$

where in equation (71) we have used the Hölder inequality with the same  $p$  and  $q$  as in the proof of Theorem 1. This shows that  $\mathbf{E} \sup_{\boldsymbol{\eta}} \|\mathbf{r}_{k,t}^\infty - \mathbf{r}_{k,t}\|^2 = O(\beta^{ts/2})$ .

Hence,

$$\mathbf{E} \sup_{\boldsymbol{\eta}} |\varepsilon_t^2| \|\mathbf{r}_t^\infty - \mathbf{r}_t\| \leq \sqrt{\mathbf{E} \sup_{\boldsymbol{\eta}} |\varepsilon_t^4| \mathbf{E} \sup_{\boldsymbol{\eta}} \|\mathbf{r}_t^\infty - \mathbf{r}_t\|^2} = O(\beta^{ts/4})$$

by Assumption 1 and equation (72). Therefore,  $\frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbf{E} \sup_{\boldsymbol{\eta}} |\varepsilon_t^2| \|\mathbf{r}_t^\infty - \mathbf{r}_t\| = o(1)$  and, hence, by Markov's inequality  $\frac{1}{\sqrt{T}} \sum_{t=1}^T \sup_{\boldsymbol{\eta}} |\varepsilon_t^2| \|\mathbf{r}_t^\infty - \mathbf{r}_t\| = o_P(1)$ .

For the treatment of the second term we use the fact that

$$\left| \frac{h_t^\infty - h_t}{h_t^\infty} \right| \leq \frac{\alpha^s}{\omega^s} \sum_{j=t}^{\infty} (\beta^s)^j \varepsilon_{t-j}^{2s}, \tag{73}$$

where again we use that  $w/(1+w) \leq w^s$  for all  $w > 0$  and any  $s \in (0, 1)$ . Then,

$$\begin{aligned}
\mathbf{E} \sup_{\boldsymbol{\eta}} \varepsilon_t^2 \|\mathbf{r}_t^\infty\| \left| \frac{h_t^\infty - h_t}{h_t^\infty} \right| &\leq \mathbf{E} \sup_{\boldsymbol{\eta}} \|\varepsilon_t^2 \mathbf{r}_t^\infty \varepsilon_{t-j}^{2s}\| \sup_{\boldsymbol{\eta}} \frac{\alpha^s}{\omega^s} \sum_{j=t}^{\infty} (\beta^s)^j \\
&\leq \sqrt{\mathbf{E} \sup_{\boldsymbol{\eta}} \|\mathbf{r}_t^\infty\|^2 \mathbf{E} |\varepsilon_t^4 \varepsilon_{t-j}^{4s}|} \sup_{\boldsymbol{\eta}} \frac{\alpha^s}{\omega^s} (\beta^s)^t \sum_{j=0}^{\infty} (\beta^s)^j \\
&= \sqrt{\mathbf{E} \sup_{\boldsymbol{\eta}} \|\mathbf{r}_t^\infty\|^2 \mathbf{E} |\varepsilon_t^4 \varepsilon_{t-j}^{4s}|} \sup_{\boldsymbol{\eta}} \frac{\alpha^s}{\omega^s (1 - \beta^s)} (\beta^s)^t \\
&= O((\beta^s)^t). \tag{74}
\end{aligned}$$

The last line follows because it can be shown by similar arguments as in the proof of Theorem 1 that  $\mathbf{E}\sup_{\boldsymbol{\eta}}\|\mathbf{r}_t^\infty\|^2 < \infty$  and because Hölder's inequality and Assumption 7 imply that  $\mathbf{E}|\varepsilon_t^4\varepsilon_{t-j}^{4s}| \leq \left(\mathbf{E}|\varepsilon_t^{4(1+s)}|\right)^{1/(1+s)} \left(\mathbf{E}|\varepsilon_{t-j}^{4(1+s)}|\right)^{s/(1+s)} < \infty$ . Equation (74) implies that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbf{E}\sup_{\boldsymbol{\eta}} \varepsilon_t^2 \|\mathbf{r}_t^\infty\| \left| \frac{h_t^\infty - h_t}{h_t^\infty} \right| = o(1), \quad (75)$$

and, again, by Markov's inequality  $\frac{1}{\sqrt{T}} \sum_{t=1}^T \sup_{\boldsymbol{\eta}} \varepsilon_t^2 \|\mathbf{r}_t^\infty\| |(h_t^\infty - h_t)/h_t^\infty| = o_P(1)$ .

The third term can be treated as follows:

$$\begin{aligned} \frac{1}{\sqrt{T}} \sum_{t=1}^T \sup_{\boldsymbol{\eta}} \varepsilon_t^2 \|\mathbf{r}_t^\infty - \mathbf{r}_t\| \left| \frac{h_t^\infty - h_t}{h_t^\infty} \right| &\leq \sqrt{\frac{1}{T} \sum_{t=1}^T \sup_{\boldsymbol{\eta}} \varepsilon_t^4 \|\mathbf{r}_t^\infty - \mathbf{r}_t\|^2 \sum_{t=1}^T \sup_{\boldsymbol{\eta}} \left| \frac{h_t^\infty - h_t}{h_t^\infty} \right|^2} \\ &\leq \left\{ \frac{1}{\sqrt{T}} \sum_{t=1}^T \sup_{\boldsymbol{\eta}} \varepsilon_t^2 \|\mathbf{r}_t^\infty - \mathbf{r}_t\| \right\} \left\{ \sum_{t=1}^T \sup_{\boldsymbol{\eta}} \left| \frac{h_t^\infty - h_t}{h_t^\infty} \right| \right\} \end{aligned}$$

because  $\sum_{t=1}^T w_t^2 \leq \left\{ \sum_{t=1}^T w_t \right\}^2$  when  $w_t \geq 0$  for all  $t$ . Above, we have already shown that  $\sum_{t=1}^T \mathbf{E}\sup_{\boldsymbol{\eta}} \varepsilon_t^2 \|\mathbf{r}_t^\infty - \mathbf{r}_t\| = O(1)$  and  $\mathbf{E}\sup_{\boldsymbol{\eta}} \left| \frac{h_t^\infty - h_t}{h_t^\infty} \right| = O(\beta^{ts})$ .

■

## B Mixed-Frequency $LM$ Test

Here, we present the first variant of the  $LM$  test for the mixed-frequency setting from Section 2.6. Since  $\tau_{0,t}$  varies at the lower frequency only, we calculate the volatility adjusted low-frequency returns  $\tilde{\varepsilon}_t$  from the ‘deGARCHed’ high-frequency returns as follows:

$$\tilde{\varepsilon}_t = \sum_{i=1}^M \frac{\varepsilon_{i,t}}{\sqrt{\bar{h}_{0,i,t}^\infty}} = \sqrt{\tau_{0,t}} Z_t, \quad (76)$$

where  $Z_t = \sum_{i=1}^M Z_{i,t}$  is i.i.d. with mean zero and variance  $M$  by Assumption 2. This leads to the score vector:

$$\mathbf{d}_t(\boldsymbol{\eta}_0) = \sum_{t=1}^T \left( \frac{\tilde{\varepsilon}_t^2}{M} - 1 \right) \begin{pmatrix} M^{-1} \\ f'_0 \mathbf{x}_t \end{pmatrix}. \quad (77)$$

Thus, if  $\tilde{\varepsilon}_t$  were observable, we could implement the test by simply regressing  $\tilde{\varepsilon}_t^2$  on a constant and  $\mathbf{x}_t$ . Again, this would be a test for heteroscedasticity in the spirit of Godfrey (1978). To actually implement the test, we need to replace the unobservable  $\tilde{\varepsilon}_t$  by

$$\hat{\tilde{\varepsilon}}_t = \sum_{i=1}^M \frac{\varepsilon_{i,t}}{\sqrt{\hat{h}_{i,t}}}, \quad (78)$$

where the  $\hat{h}_{i,t}$  are obtained by estimating the GARCH model under the null for the daily data. However, a simple Taylor expansion shows that  $\hat{\tilde{\varepsilon}}_t$  has measurement error due to pre-estimating  $\bar{h}_{0,i,t}^\infty$ :

$$\hat{\tilde{\varepsilon}}_t = \sum_{i=1}^M \frac{\varepsilon_{i,t}}{\sqrt{\bar{h}_{0,i,t}^\infty}} \left( 1 - \left( \sqrt{\hat{h}_{i,t}} - \sqrt{\bar{h}_{0,i,t}^\infty} \right) / \sqrt{\bar{h}_{0,i,t}^\infty} + o_P(\sqrt{T}) \right) \approx \tilde{\varepsilon}_t + W_t,$$

where  $W_t$  has mean zero but non-zero variance. Higher-order terms are negligible for the test performance. Thus, tests based on the critical values from the  $\chi^2$ -distribution (derived in Theorem 3) will be size distorted (see also Li and Mak, 1994). However, the correct distribution of the test statistic based on  $\hat{\tilde{\varepsilon}}_t$  can be obtained by simulation.

## C Simulations

### C.1 Empirical Size as a Function of Sample Size $T$ .

The following table shows the empirical size for model G-L with  $x_t = \varepsilon_t^2/\hat{h}_t$  and  $Z_t \sim t(7)$ . For this specification, we observed the strongest size distortion in Table 1. The column labelled  $T = 1000$  contains the same figures as in the respective column of Table 1. The other columns show that the size distortion diminishes with increasing sample size.

Table 5: Empirical size for model G-L,  $Z_t \sim t(7)$  depending on sample size  $T$ .

$x_t = \varepsilon_t^2/\hat{h}_t$		$T = 1000$	$T = 2500$	$T = 5000$	$T = 7500$	$T = 10000$
$LM$	1%	0.7	0.9	0.9	0.9	1.3
	5%	3.1	4.2	4.6	4.4	5.2
	10%	7.2	7.7	8.5	9.4	10.0
$LM_{LT}$	1%	0.9	1.1	0.9	1.1	1.3
	5%	3.4	3.5	4.4	4.6	5.1
	10%	6.7	8.0	8.6	9.3	9.5

**Notes:** The number of observations is  $T \in \{1000, 2500, 5000, 7500, 10000\}$ . Entries are rejection rates in percent over  $R = 1000$  replications at the 1%, 5% and 10% nominal level. The model for the conditional variance is a GARCH(1,1) with  $\omega_0 = 0.05$ ,  $\alpha_0 = 0.05$  and  $\beta_0 = 0.90$  (i.e. model G-L). The  $LM$  tests are performed for a GARCH(1,1) under  $H_0$ . Otherwise see Table 1.

### C.2 Size-Adjusted Power: Exponential Long-Term Component and $t$ Distributed Innovations.

The following table provides simulation results on the size-adjusted power for the case that the innovation  $Z_t$  is Student- $t$  distributed with 7 degrees of freedom.



Table 6: Size-adjusted power: exponential  $\tau_{0,t}$  component,  $t$  distributed innovations.

$x_t$	$VIX_t$						$VIX_t^{(22)}$	$VIX_t^{(65)}$
	$\alpha_0 = 0.09$			$\alpha_0 = 0.07$			$\alpha_0 = 0.09$	
	I	F	S	I	F	S	I	I
$LM$	54.9	44.1	32.1	59.3	57.1	39.3	20.1	13.9
$LM_{LT,mod}$	30.5	25.8	22.7	50.7	48.7	39.4	12.2	9.5
$LM_{LT}$	5.8	5.7	5.7	5.0	5.2	5.1	5.6	5.3
$VR$	12.8	12.4	12.1	28.2	27.9	27.0	12.0	9.7

**Notes:** Innovations  $Z_t$  are Student- $t$  distributed with 7 degrees of freedom. The specification of the long-term component is given by  $\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$ . The number of observations is  $T = 1000$ . Results are based on  $R = 1000$  replications. The  $LM$  tests are performed for a GARCH(1,1) under  $H_0$ . Otherwise see Table 2.

### C.3 Size-Adjusted Power for Different Values of $K$ .

Table 7 illustrates how a misspecification of  $K$  affects the power of the  $LM$  test. We simulate return data with the short-term component G-H (high persistence) and the long-term component as in equation (5). We either choose  $\boldsymbol{\pi}_0 = 0.3$ ,  $\boldsymbol{\pi}_0 = (0, 0.3)'$  or  $\boldsymbol{\pi}_0 = (0, 0, 0.3)'$ . The first option corresponds to the immediately decaying weighting scheme from Table 2. The second and third weighting schemes are extreme in the sense that all weight is put on lag 2 or 3, respectively. Clearly, the correct choice of  $K$  in the  $LM$  test is either  $K = 1$ ,  $K = 2$  or  $K = 3$ . In Panel A, we use the VIX as the explanatory variable. In Panels B-D, we first simulate an AR(1) process with autoregressive coefficient  $\delta$  and i.i.d. normal innovations with mean zero and variance 0.025 and use the generated time series as the explanatory variable. We vary  $\delta$  between 0.98, 0.9 and 0.8 to check whether the persistence of the AR(1) process affects our findings.

As Table 7 shows, for all specifications we observe the highest size-adjusted power when  $K$  is chosen correctly. This finding is also independent of the persistence of the AR(1) process. Clearly, when the persistence of the AR(1) process decreases, the long-term component becomes less variable relative to the short-term component and, hence, the variance ratio (VR) decreases. For example, for  $\delta = 0.8$  the variance ratio is less than 2%. The low variance ratio then leads to a decline of the power of the test. Nevertheless, the simulations show that choosing  $K = 1$  always delivers a reasonable power even in the extreme case when all weight is put on the second or third lag. At first, it might

be surprising that the power of the test is reasonably high for  $K = 1$ , even though zero weight is attached to the first lag in the weighting scheme. However, for persistent  $x_t$ , the information in  $x_{t-1}$  is very similar to that in  $x_{t-2}$  and so the test works reasonably well despite the misspecification of  $K$ . Given that in most real applications we can expect that the true weighting scheme is declining from the first lag, we recommend always starting with  $K = 1$ . If the test does not reject for  $K = 1$  and  $x_t$  has low persistence, it may be advisable to redo the test for  $K = 2, K = 3, \dots$

Table 7: Size-adjusted power: exponential long-term component, variation in  $K$ .

$K$	<b>1</b>	2	1	<b>2</b>	3	1	2	<b>3</b>	4
weighting scheme	$\boldsymbol{\pi}_0 = 0.3$		$\boldsymbol{\pi}_0 = (0, 0.3)'$			$\boldsymbol{\pi}_0 = (0, 0, 0.3)'$			
	Panel A: $x_t = VIX_t$								
<i>LM</i>	74.4	68.3	51.4	68.6	67.5	42.7	46.7	67.0	61.1
<i>VR</i>	15.6	15.6	15.6	15.6	15.6	15.6	15.6	15.6	15.6
	Panel B: $x_t$ is AR(1) with $\delta = 0.98$								
<i>LM</i>	63.3	54.8	52.6	53.4	46.6	44.1	43.4	48.0	45.2
<i>VR</i>	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3
	Panel C: $x_t$ is AR(1) with $\delta = 0.90$								
<i>LM</i>	37.0	30.3	24.0	29.0	23.6	13.8	18.2	24.8	21.5
<i>VR</i>	2.78	2.78	2.77	2.77	2.77	2.78	2.78	2.78	2.78
	Panel D: $x_t$ is AR(1) with $\delta = 0.80$								
<i>LM</i>	36.0	29.0	21.3	29.1	23.8	13.2	16.2	25.0	21.2
<i>VR</i>	1.76	1.76	1.75	1.75	1.75	1.74	1.74	1.74	1.74

**Notes:** The number of observations is  $T = 1000$ . The table reports the size-adjusted power in percent over the  $R = 1000$  replications at the 5% nominal level. The model for the conditional variance is a GARCH(1,1) with  $\omega_0 = 0.05$ ,  $\alpha_0 = 0.09$  and  $\beta_0 = 0.90$  (i.e. model G-H). The specification of the long-term component is given by  $\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$  with parameter  $\boldsymbol{\pi}_0$  as specified in the table. We consider the *LM* test with the GARCH(1,1) under the null hypothesis.  $K$  denotes the number of lags that are used in the test. The bold number indicates the correct lag length. Otherwise see Table 2.

## C.4 Persistence in the GARCH component.

In this section, we investigate the power properties of the  $LM$  test when  $\hat{\alpha} + \hat{\beta}$  is close to or even above one. As in Appendix C.3, we first simulate an AR(1) process with autoregressive coefficient  $\delta$  and i.i.d. normal innovations with mean zero and variance 0.025 and use the generated time series as the explanatory variable. We choose  $\delta \in \{0.9, 0.98\}$  and impose an immediately decaying weighting scheme with  $\pi_0 = 0.3$  or  $\pi_0 = 0.4$ . The GARCH component has either moderate (G-M), high (G-H) or extreme (G-E,  $\alpha_0 = 0.095$ ,  $\beta_0 = 0.90$ ) persistence. As before, we choose  $T$  and  $R$  as 1000.

Table 8 shows that – despite the fact that the simulation is under the alternative – the median of the estimates of  $\alpha$  and  $\beta$  over the  $M = 1000$  simulations is close to the true parameter values. In particular, in all scenarios the median of  $\hat{\alpha} + \hat{\beta}$  is below (or equal to)  $\alpha_0 + \beta_0$ . This suggests that the misspecified GARCH model does not suffer from the so-called IGARCH effect. This is true even for cases in which the variance ratio is as high as  $VR = 39.4$ .

Further, the table shows that for all specifications in which the GARCH component has moderate persistence, we never observe that the sum of the estimated GARCH parameters is greater than or equal to one. For example, in Panel B when  $\alpha_0 = 0.07$ ,  $\beta_0 = 0.90$  (G-M) and  $\pi_0 = 0.4$ , the GARCH component is severely misspecified and the  $LM$  test rejects in 92.9% of the simulations, there is not a single simulation in which  $\hat{\alpha} + \hat{\beta} \geq 1$ .

The picture changes slightly when the persistence in the GARCH component is high (G-H). For this specification  $\hat{\alpha} + \hat{\beta}$  is greater than or equal to one in 6 out of the 1000 simulations when  $\delta = 0.9$  and in 11 ( $\pi_0 = 0.3$ ) or 19 ( $\pi_0 = 0.4$ ) cases when  $\delta = 0.98$ . However, for these cases the  $LM$  test has rejection rates which are (in all but one case) even higher than the average rejection rates over all 1000 simulations. For example, when  $\delta = 0.98$  and  $\pi_0 = 0.4$ , the  $LM$  test rejects in all 19 cases in which  $\hat{\alpha} + \hat{\beta} \geq 1$ .

Finally, when the persistence in the GARCH component is extreme (G-E),  $\hat{\alpha} + \hat{\beta}$  is greater than or equal to one in 54 simulations when  $\delta = 0.9$  and in 80 ( $\pi_0 = 0.3$ ) or 92 ( $\pi_0 = 0.4$ ) cases when  $\delta = 0.98$ . Nevertheless, we find that the  $LM$  test has very reasonable power and that the rejection frequency among those cases in which  $\hat{\alpha} + \hat{\beta} \geq 1$  is typically higher than the average power.

Our results suggest that the main reason for obtaining estimates  $\hat{\alpha} + \hat{\beta} \geq 1$  is unlikely to be an omitted long-term component, but rather an extreme persistence in the true

GARCH component. Nevertheless, the effect is strengthened if the omitted long-term component is more relevant. However, the simulations also clearly suggest that – for a given specification – the power of the  $LM$  test does not decrease in the persistence of the estimated parameters (as measured by  $\hat{\alpha} + \hat{\beta}$ ).

Table 8: Size-adjusted power: persistent GARCH component

Panel A: $x_t$ is AR(1) with $\delta = 0.90$						
	G-M		G-H		G-E	
	$\alpha_0 = 0.07, \beta_0 = 0.90$		$\alpha_0 = 0.09, \beta_0 = 0.90$		$\alpha_0 = 0.095, \beta_0 = 0.90$	
	$\pi_0 = 0.3$	$\pi_0 = 0.4$	$\pi_0 = 0.3$	$\pi_0 = 0.4$	$\pi_0 = 0.3$	$\pi_0 = 0.4$
median( $\hat{\alpha}$ )	0.072	0.073	0.091	0.092	0.096	0.097
median( $\hat{\beta}$ )	0.890	0.888	0.894	0.893	0.894	0.893
median( $\hat{\alpha} + \hat{\beta}$ )	0.965	0.964	0.986	0.986	0.991	0.991
$LM$	38.7	62.1	37.0	60.8	36.2	59.7
$VR$	7.68	12.9	2.78	4.82	1.97	3.44
$\#(\hat{\alpha} + \hat{\beta} \geq 1)$	0	0	6	6	54	54
$\%reject (\hat{\alpha} + \hat{\beta} \geq 1)$	-	-	50.0%	50.0%	33.3%	64.8%
Panel B: $x_t$ is AR(1) with $\delta = 0.98$						
	G-M		G-H		G-E	
	$\alpha_0 = 0.07, \beta_0 = 0.90$		$\alpha_0 = 0.09, \beta_0 = 0.90$		$\alpha_0 = 0.095, \beta_0 = 0.90$	
	$\pi_0 = 0.3$	$\pi_0 = 0.4$	$\pi_0 = 0.3$	$\pi_0 = 0.4$	$\pi_0 = 0.3$	$\pi_0 = 0.4$
median( $\hat{\alpha}$ )	0.076	0.081	0.095	0.099	0.100	0.103
median( $\hat{\beta}$ )	0.888	0.885	0.889	0.885	0.891	0.887
median( $\hat{\alpha} + \hat{\beta}$ )	0.968	0.970	0.986	0.986	0.991	0.992
$LM$	73.4	92.9	63.3	86.9	58.2	83.9
$VR$	26.7	39.4	11.3	18.6	8.32	14.1
$\#(\hat{\alpha} + \hat{\beta} \geq 1)$	0	0	11	19	80	92
$\%reject (\hat{\alpha} + \hat{\beta} \geq 1)$	-	-	72.7%	100%	70.0%	91.3%

**Notes:** The number of observations is  $T = 1000$ . median( $\hat{\alpha}$ ), median( $\hat{\beta}$ ) and median( $\hat{\alpha} + \hat{\beta}$ ) present the median of the parameter estimates over the  $R = 1000$  replications.  $LM$  is the size-adjusted power in percent at the 5% nominal level.  $VR$  is the variance ratio.  $\#(\hat{\alpha} + \hat{\beta} \geq 1)$  gives the number of simulations in which the condition  $\hat{\alpha} + \hat{\beta} < 1$  is violated.  $\%reject|(\hat{\alpha} + \hat{\beta} \geq 1)$  presents the percentage of cases in which the  $LM$  test rejects given that  $\hat{\alpha} + \hat{\beta} \geq 1$ . The model for the conditional variance is a GARCH(1,1) with moderate (G-M), high (G-H) or extreme (G-E) persistence. The specification of the long-term component is given by  $\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$  with parameter  $\boldsymbol{\pi}_0$  as specified in the table.  $x_t$  is an AR(1) with autoregressive parameter  $\delta$ . Otherwise see Table 2.

This observation is also confirmed by the plots in Figure 3. The figure shows a scatterplot of the estimate of the persistence ( $\hat{\alpha} + \hat{\beta}$ ) in the GARCH component ( $x$ -axis) and the corresponding  $LM$  statistics ( $y$ -axis). The horizontal green line indicates the 5% critical value of the  $LM$  test and the vertical red line a persistence of one. In both plots the true GARCH component has extreme persistence (G-E), the AR(1) parameter of the explanatory variable is either  $\delta = 0.90$  (left plot) or  $\delta = 0.98$  (right plot) and  $\pi_0 = 0.4$ . Again, the figure shows that there is no indication that the power of the test decreases when the estimated persistence is increasing.

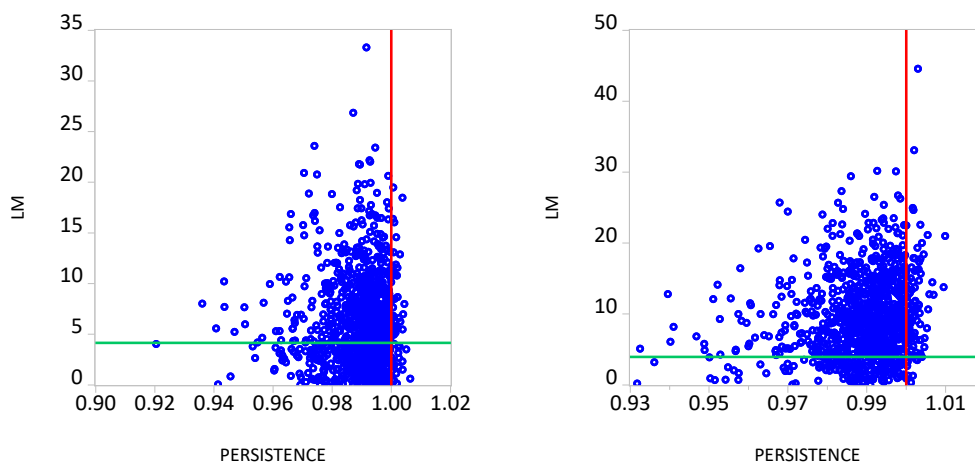


Figure 3: Scatterplot of estimated persistence ( $\hat{\alpha} + \hat{\beta}$ ) and  $LM$  test statistics. Green horizontal line: critical value 5% level. Vertical red line: persistence of one. The true GARCH component has extreme persistence (G-E specification). We choose an immediately decaying weighting scheme with  $\pi_0 = 0.4$ . The explanatory variable is an AR(1) process with autoregressive parameter  $\delta = 0.90$  (left) and  $\delta = 0.98$  (right).

It is important to highlight that empirically  $\hat{\alpha} + \hat{\beta}$  might be close to one for other reasons than an omitted multiplicative component. For example, as shown in Hillebrand (2005) the ‘IGARCH effect’ can be due to neglected parameter changes or, as discussed in Baillie et al. (1996), due to neglected long-memory. Finally, even if the true model is a stationary but very persistent GARCH model, it may happen that the sum of the estimated GARCH parameters is above one.

## C.5 Size-Adjusted Power for Linear Long-Term Component.

Table 9: Size-adjusted power: linear long-term component,  $Z_t$  normally distributed.

$x_t$	$VIX_t$						$VIX_t^{(22)}$	$VIX_t^{(65)}$
	$\alpha_0 = 0.09$			$\alpha_0 = 0.07$			$\alpha_0 = 0.09$	
	I	F	S	I	F	S	I	I
<i>LM</i>	57.2	54.8	39.2	66.5	64.6	51.5	28.9	18.0
<i>LM<sub>LT,mod</sub></i>	34.8	34.1	30.3	59.2	58.1	51.1	14.3	10.7
<i>LM<sub>LT</sub></i>	5.9	5.9	5.4	5.6	5.6	5.3	4.8	4.6
<i>VR</i>	12.4	12.3	12.1	29.5	29.4	29.0	12.0	10.5

**Notes:** Innovations  $Z_t$  are standard normally distributed. The specification of the long term component is given by  $\tau_{0,t} = 1 + \sum_{k=1}^K \pi_{0k} x_{t-k}$ . The number of observations is  $T = 1000$ . Results are based on  $R = 1000$  replications. The *LM* tests are performed for a GARCH(1,1) under  $H_0$ . Otherwise see Table 2.

Table 10: Size-adjusted power: linear long-term component,  $t$  distributed innovations.

$x_t$	$VIX_t$						$VIX_t^{(22)}$	$VIX_t^{(65)}$
	$\alpha_0 = 0.09$			$\alpha_0 = 0.07$			$\alpha_0 = 0.09$	
	I	F	S	I	F	S	I	I
<i>LM</i>	39.7	34.5	28.4	48.3	46.7	37.9	17.4	12.3
<i>LM<sub>LT,mod</sub></i>	27.4	25.9	24.6	43.0	42.8	38.5	11.1	9.1
<i>LM<sub>LT</sub></i>	5.5	5.5	5.6	5.3	5.3	5.3	5.7	5.4
<i>VR</i>	10.0	9.9	9.8	23.0	22.9	22.5	9.7	8.5

**Notes:** Innovations  $Z_t$  are Student- $t$  distributed with 7 degrees of freedom. The specification of the long term component is given by  $\tau_{0,t} = 1 + \sum_{k=1}^K \pi_{0k} x_{t-k}$ . The number of observations is  $T = 1000$ . Results are based on  $R = 1000$  replications. The *LM* tests are performed for a GARCH(1,1) under  $H_0$ . Otherwise see Table 2.

## C.6 Misspecification of the Short-Term Component.

In the following, we investigate the consequences of implementing the  $LM$  test under the null of a GARCH(1,1) although the true short-term component is a GJR-GARCH(1,1), a higher-order GARCH or a fractionally integrated GARCH (FIGARCH). We first consider a situation in which the short-term component is given by a GJR-GARCH(1,1). We simulate data from a model with a short-term component given by equation (35) with parameters as specified in GJR-M and GJR-H and either  $\tau_{0,t} = 1$  or  $\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$ .

Table 11, Panel A, presents the empirical size-adjusted rejection rates. First, consider the case that  $\tau_{0,t} = 1$ . When using  $VIX_t$  as the explanatory variable, we find that the empirical rejection rates are close to the 5% nominal level. That is, using a truly exogenous explanatory variable the  $LM$  test does not detect a deviation from the null hypothesis. Even when testing for ‘ARCH nested in GARCH’, i.e. when using the ‘endogenous’  $x_t = \varepsilon_t^2 / \hat{h}_t$  as the explanatory variable, we obtain the same result. Second, we consider the case that  $\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$ . Although the short-term component is misspecified, the empirical power is only slightly lower than when the short-term component is correctly specified. For example, for the GJR-M model with an immediately decaying weighting scheme the  $LM$  test rejects in 80.4% of cases at the 5% nominal level. The corresponding figure for the correctly specified GJR-GARCH model from Table 2 is 82.8%. On the other hand, when testing for remaining ARCH effects both tests,  $LM$  and  $LM_{LT}$ , do not detect a deviation.

Next, we investigate the performance of the  $LM$  test when the true short-term component is higher-order GARCH or FIGARCH while the long-term component is constant ( $\tau_{0,t} = 1$ ). We consider a GARCH(1,2) and denote the second order ARCH parameter by  $\tilde{\alpha}_0$ . As in model G-L in Section 3.1, we choose  $\omega_0 = 0.01$ ,  $\alpha_0 = 0.05$ ,  $\beta_0 = 0.9$  in combination with  $\tilde{\alpha}_0 \in \{0.02, 0.04\}$ . For the GARCH(2,2) model, we choose the parameter estimates from Nelson and Cao (1992) for the Deutschmark/Dollar exchange rate (see their Table 1):

$$\bar{h}_{0t}^\infty = 0.186 + 0.0573\varepsilon_{t-1}^2 + 0.2262\varepsilon_{t-2}^2 + 0.3833\bar{h}_{0,t-1}^\infty + 0.3100\bar{h}_{0,t-2}^\infty. \quad (79)$$

Finally, we consider a FIGARCH(1,  $d$ , 1) model which features long-memory in the conditional variance. For this model, the conditional variance is given by

$$(1 - \beta_0 L)\bar{h}_{0t}^\infty = \omega_0 + [1 - \beta_0 L - (1 - \phi_0 L)(1 - L)^{d_0}]\varepsilon_{t-1}^2 \quad (80)$$

Table 11: Misspecified short-term component.

Panel A	$\tau_{0,t} = 1$				$\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$			
	GJR-M		GJR-H		GJR-M		GJR-H	
	$VIX_t$	$\varepsilon_t^2/\hat{h}_t$	$VIX_t$	$\varepsilon_t^2/\hat{h}_t$	$VIX_t$	$\varepsilon_t^2/\hat{h}_t$	$VIX_t$	$\varepsilon_t^2/\hat{h}_t$
$x_t$								
$LM$	5.4	5.8	5.0	5.9	80.4	6.4	70.6	6.2
$LM_{LT}/LM_{LT,mod}$		6.2		5.7	69.4	6.2	43.8	5.9
Panel B: $\tau_{0,t} = 1$	GARCH(1,2)				GARCH(2,2)		FIGARCH(1, $d$ , 1)	
	$\alpha_0 = 0.05, \beta_0 = 0.9$				parameters as in		$\phi_0 = 0.95, \beta_0 = 0.9$	
	$\tilde{\alpha}_0 = 0.02$		$\tilde{\alpha}_0 = 0.04$		equation (79)		$d_0 = 0.3$	
$x_t$	$VIX_t$	$\varepsilon_t^2/\hat{h}_t$	$VIX_t$	$\varepsilon_t^2/\hat{h}_t$	$VIX_t$	$\varepsilon_t^2/\hat{h}_t$	$VIX_t$	$\varepsilon_t^2/\hat{h}_t$
$LM$	4.2	9.2	3.2	19.4	5.8	88.0	6.1	40.4
$LM_{LT}$		9.0		20.2		89.3		37.9

**Notes:** The table reports the empirical size-adjusted rejection rates over  $R = 1000$  replications at the 5% nominal level. In Panel A, the data generating process is a GJR-GARCH(1,1) with parameters as given by GJR-M and GJR-H. In Panel B, the model for the conditional variance is a GARCH(1, 2) with  $\omega_0 = 0.01$ , the GARCH(2,2) given in equation (79) and a FIGARCH(1,  $d$ , 1) with  $\omega_0 = 0.05$ . The long-term component  $\tau_{0,t}$  is specified in the table. The  $LM$  tests are performed for a GARCH(1,1) under  $H_0$ . Innovations  $Z_t$  are standard normal distributed. All test statistics are based on  $K = 1$ . The number of observations is  $T = 1000$ .

under  $H_0$ , where  $L$  denotes the lag operator and  $d_0$  the fractional differencing parameter. We set  $\omega_0 = 0.05$ ,  $\phi_0 = 0.95$  and  $\beta_0 = 0.9$ . For  $d_0 = 0$  the FIGARCH reduces to a GARCH(1,1) model with  $\alpha_0 = \phi_0 - \beta_0 = 0.05$  and, hence, to model G-L. Also, note that the parameters satisfy the conditions that ensure the positivity of the conditional variance (see Conrad and Haag, 2006).

Again, Table 11, Panel B, shows that for all short-term specifications the rejection rate of the  $LM$  test is quite close to the 5% nominal level for  $x_t = VIX_t$ . When searching for remaining ARCH effects ( $x_t = \varepsilon_t^2/\hat{h}_t$ ), the  $LM$  test tends to reject the null hypothesis with higher rejection rates for models that are further away from the null hypothesis (GARCH(2,2) and FIGARCH(1,  $d$ , 1)). The table also shows that for this choice of  $x_t$  the  $LM$  test has a similar power as the Lundbergh and Teräsvirta (2002) test.



## C.7 Simulation Mixed-Frequency Test.

In this section we provide simulation evidence for the mixed-frequency version of the test. All results are based on  $R = 1000$  replications. We model  $x_t$  as either evolving at a quarterly or monthly frequency and assume that  $x_t$  follows an AR(1), i.e.

$$x_t = \delta x_{t-1} + \nu_t, \quad (81)$$

with  $\delta = 0.98$  and  $\nu_t \sim \mathcal{N}(0, \sigma_\nu^2)$ . As before, the specification of the long-term component is given by  $\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$ . The model for the short-term component is the GJR-GARCH with high (GJR-H) or moderate (GJR-M) persistence. In the simulations, we employ the immediate (I) and slow (S) decaying weighting schemes presented in Section 3.2. We consider the regression of either  $\widehat{RV}_t$  or  $RV_t$  on a constant,  $x_{t-1}$  and its own first lag:

$$\ln(DV_t) = \tilde{c} + \pi_1 x_{t-1} + \rho \ln(DV_{t-1}) + \tilde{\zeta}_t \quad (82)$$

with  $DV_t \in \{\widehat{RV}_t, RV_t\}$ . Table 12 reports the number of instances in which the null hypothesis that the coefficient on  $x_{t-1}$  is zero is rejected (by comparing the squared  $t$ -statistic with the critical value from the asymptotic  $\chi^2(1)$  distribution).  $\widehat{RV}$  is based on the estimated conditional variances of the correctly specified GJR model.

Recall that in the mixed-frequency setting returns are denoted by  $\varepsilon_{i,t}$ , where  $i = 1, \dots, M$  refers to the trading days within period  $t = 1, \dots, T$ . In the simulations, we first fixed  $T = 172$  and  $M = 66$ , which corresponds to 172 quarters of 66 days each. We then choose  $M = 22$  days which corresponds to monthly data. For  $M = 22$ , we either keep the number of low frequency observations fixed at  $T = 172$  (which reduces the number of daily observations) or keep the number of daily observations fixed and, thereby, extend the low frequency observations to  $T = 516$ .

Table 12 shows that under  $H_0$  ( $\tau_{0,t} = 1$ ) the empirical size is close to the nominal 5% level for all scenarios. Next, for  $M = 66$  and under the alternative, we observe that the test based on  $\widehat{RV}$  does indeed have a higher power than the test based on  $RV$ . For example, for model GJR-H, an immediately decaying weighting scheme and  $\sigma_\nu^2 = 0.025$  the test rejects in 74.8% of cases for  $\widehat{RV}$  but only in 40.6% for  $RV$ . Interestingly, the power decreases only modestly when the true weighting scheme has a slow decay but the regression is still based on  $x_{t-1}$  only. As expected, increasing the variability of the long-term component ( $\sigma_\nu^2 = 0.030$ ) increases the power of both tests. Similarly, reducing

the persistence of the short-term component (i.e. considering GJR-M), strongly increases the power of both tests. The same effect was already observed in Table 2. Nevertheless, the power for the test based on  $\widehat{RV}$  still remains higher than the power of the test based on  $RV$ .

Table 12: Empirical size and power of low-frequency regression-based test.

$H_0$ : GJR-GARCH	GJR-H					GJR-M				
	$\tau_{0,t} = 1$	$\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$				$\tau_{0,t} = 1$	$\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$			
		$\sigma_\nu^2 = 0.025$		$\sigma_\nu^2 = 0.030$			$\sigma_\nu^2 = 0.025$		$\sigma_\nu^2 = 0.030$	
weighting scheme	-	I	S	I	S	-	I	S	I	S
	$M = 66, T = 172$ “quarterly” observations of $x_t$									
$\widehat{RV}$	6.5	74.8	69.4	80.3	77.2	5.7	97.5	94.4	97.0	96.4
$RV$	4.4	40.6	38.1	48.6	44.1	5.8	89.4	86.5	91.7	89.0
$VR$	-	5.81	5.39	7.15	6.48	-	19.2	18.2	22.2	21.1
	$M = 22, T = 172$ “monthly” observations of $x_t$									
$\widehat{RV}$	6.4	38.7	38.0	47.4	43.4	6.5	74.6	70.3	79.9	77.8
$RV$	6.3	26.2	26.4	32.7	30.5	7.1	62.3	59.2	70.1	67.5
$VR$	-	5.91	5.49	6.97	6.33	-	16.6	15.7	18.6	17.9
	$M = 22, T = 516$ “monthly” observations of $x_t$									
$\widehat{RV}$	6.0	89.9	88.9	93.8	92.2	6.1	99.8	99.7	100	99.7
$RV$	6.8	61.4	63.9	67.7	69.6	5.2	99.2	97.9	99.5	98.8
$VR$	-	7.54	7.28	8.76	8.67	-	21.5	20.6	24.4	23.5

**Notes:** The table reports size and power in percent over the  $R = 1000$  replications at the 5% nominal level.  $T$  denotes the number of low-frequency observations and  $M$  the number of days within each period  $t$ .  $\sigma_\nu^2$  is the variance of the innovation of the AR(1) process for  $x_t$ . The low-frequency regression version of the test is based on equation (82) with either  $\widehat{RV}$  or  $RV$  as dependent variable.  $\widehat{RV}$  is based on the estimated conditional variance from the correctly specified GJR-GARCH model. I and S indicate the immediate and slow decaying weighting schemes and  $VR$  denotes the variance ratio. In all tests, we choose  $K = 1$ .

Switching to monthly observations, i.e. choosing  $M = 22$ , reduces the power of both tests when keeping the number of low frequency observations constant ( $T = 172$ ). Intuitively, this is reasonable since the predictive regressions are now based on the same number of low-frequency observations as before but the quality of the estimated conditional variances deteriorates (because the number of high-frequency observations decreases) which means that the precision of  $\widehat{RV}_t$  as an estimator of  $RV_t$  decreases. On the

other hand, when decreasing  $M$  from 66 to 22 while keeping the number of *high-frequency* observations fixed, the power of the test increases. In this scenario, the number of observations in the predictive regression increases ( $T = 516$ ) which makes it easier to detect the omitted component.

Table 13 shows the empirical size and power for the same GJR models as before, but with an  $\widehat{RV}$  that is based on the estimated conditional variances from a misspecified GARCH(1,1). First, note that both tests appear to be slightly oversized in this situation. Second, as a result of the misspecification of the short-term component the power of the test based on  $\widehat{RV}$  is lower than in Table 12, but still higher than the power of the test based on  $RV$ .

Table 13: Size and power of low-frequency regression test based on misspecified GARCH.

$H_0$ : GARCH	GJR-H					GJR-M				
	$\tau_{0,t} = 1$	$\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$				$\tau_{0,t} = 1$	$\tau_{0,t} = \exp(\boldsymbol{\pi}'_0 \mathbf{x}_t)$			
		$\sigma_\nu^2 = 0.025$	$\sigma_\nu^2 = 0.030$			$\sigma_\nu^2 = 0.025$	$\sigma_\nu^2 = 0.030$			
weighting scheme	-	I	S	I	S	-	I	S	I	S
	$M = 66, T = 172$ “quarterly” observations of $x_t$									
$\widehat{RV}$	6.3	62.8	59.2	69.6	64.9	6.9	95.2	91.6	96.6	93.0
$RV$	4.1	43.7	38.1	49.5	44.3	6.8	88.8	87.8	92.6	88.3
$VR$	-	5.87	5.54	7.11	6.50	-	18.4	18.4	22.4	20.9
	$M = 22, T = 172$ “monthly” observations of $x_t$									
$\widehat{RV}$	6.7	32.3	29.7	37.0	34.0	6.5	69.2	63.9	76.2	71.0
$RV$	8.0	28.1	26.2	33.1	28.7	6.4	65.9	61.2	72.5	67.5
$VR$	-	5.86	5.61	7.00	6.40	-	16.6	15.0	19.2	18.2
	$M = 22, T = 516$ “monthly” observations of $x_t$									
$\widehat{RV}$	6.0	80.8	77.9	86.2	84.8	5.3	99.6	98.5	99.8	99.3
$RV$	7.2	64.8	63.3	72.9	70.0	6.5	98.8	98.0	99.3	98.2
$VR$	-	7.52	7.18	9.12	8.49	-	21.5	20.9	24.5	23.7

**Notes:**  $\widehat{RV}$  is based on the estimate of the conditional variance from the misspecified GARCH model. Otherwise see Table 12.

## D Correlations of Explanatory Variables

Tables 14 and 15 show the contemporaneous correlations between  $VXO_t$ ,  $RV_t$ ,  $EPU_t$  and  $ADS_t$ . Among the daily variables  $VXO_t$  and  $RV_t$  have the highest correlation (0.52). The other correlations also have the expected signs:  $VXO_t$  is positively correlated with economic policy uncertainty,  $EPU_t$ , but negatively correlated with the business conditions index,  $ADS_t$ . The correlations of the rolling window versions of the four variables with  $N = 22$  are higher in absolute value but reveal the same relationships.

Table 14: Correlations between daily explanatory variables.

	$VXO_t$	$RV_t$	$EPU_t$	$ADS_t$
$VXO_t$	1.00			
$RV_t$	0.52	1.00		
$EPU_t$	0.31	0.19	1.00	
$ADS_t$	-0.48	-0.26	-0.28	1.00

**Notes:** The table presents the correlations between the daily explanatory variables. All correlation figures are for the 1987M12-2016M06 period.

Table 15: Correlations between explanatory variables,  $x_t^{(N)}$ , for  $N = 22$ .

	$VXO_t^{(22)}$	$RV_t^{(22)}$	$EPU_t^{(22)}$	$ADS_t^{(22)}$
$VXO_t^{(22)}$	1.00	0.92	0.59	-0.68
$RV_t^{(22)}$		1.00	0.52	-0.57
$EPU_t^{(22)}$			1.00	-0.42
$ADS_t^{(22)}$				1.00

**Notes:** The table presents the correlations between the rolling window explanatory variables. All correlation figures are for the 1987M12-2016M06 period.

# Working Paper Series in Economics

---

## recent issues

- No. 121** *Christian Conrad and Melanie Schienle*: Testing for an omitted multiplicative long-term component in GARCH models, January 2019
- No. 120** *Marta Serra-Garcia and Nora Szech*: The (in)elasticity of moral ignorance, December 2018
- No. 119** *Thomas Mariotti, Nikolaus Schweizer, Nora Szech and Jonas von Wangenheim*: Information nudges and self-control, November 2018
- No. 118** *Andranik S. Tangian*: Methodological notes on composite indicators for monitoring working conditions, October 2018
- No. 117** *Andranik S. Tangian*: Testing the improved third vote during the 2018 election of the Karlsruhe Institute of Technology student parliament, September 2018
- No. 116** *Yuri Golubev and Mher Safarian*: On robust stopping times for detecting changes in distribution, May 2018
- No. 115** *Daniel Hoang, Sebastian Gatzer and Martin Ruckes*: The economics of capital allocation in firms: Evidence from internal capital markets, January 2018
- No. 114** *Francesco D'Acunto, Daniel Hoang and Michael Weber*: Unconventional fiscal policy, January 2018
- No. 113** *Alberto Bucci, Levent Eraydin, Moritz Müller*: Dilution effects, population growth and economic growth under human capital accumulation and endogenous technological change, January 2018
- No. 112** *Jochen Schweikert and Markus Höchstötter*: Epidemiological spreading of mortgage default, January 2018
- No. 111** *Armin Falk and Nora Szech*: Diffusion of being pivotal and immoral outcomes, December 2017