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

Christian Conrad† and Melanie Schienle‡

†Heidelberg University, Germany

‡Karlsruhe Institute of Technology, Germany

First draft: July 7, 2015
This draft: December 20, 2016

Abstract

We consider the problem of testing for an omitted multiplicative long-term component in a simple GARCH model. 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, Volatility Component Models.

JEL Classification: C53, C58, E32, G12

---

*We would like to thank Torben G. Andersen, Richard Baillie, Ian Dew-Becker, Robert F. Engle, Eric Ghysels, Tilmann Gneiting, Onno Kleen, Enno Mammen, Rasmus S. Pedersen, Peter Schmidt, Karin Stürmer, Robert Taylor, Timo Teräsvirta, Viktor Todorov and Jeffrey Wooldridge for helpful comments and suggestions.

†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.

‡Melanie Schienle, Department of Economics (ECON), Karlsruhe Institute of Technology, Schlossbezirk 12, 76131 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, Christiansen et al. (2012), Paye (2012), Engle et al. (2013) and Conrad and Loch (2015a) provide recent evidence for the counter-cyclical behavior of financial volatility.[1] In particular, 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 empirical analysis on a multiplicative two-component GARCH model. In this model a short-term unit variance GARCH component fluctuates around a smooth long-term component that is driven by macroeconomic conditions.

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.[2]

For this reason, we develop a new misspecification test for the simple GARCH model. While under the null hypothesis the true model is a GARCH(1,1), 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 a potentially omitted long-term component. For our LM test statistic, we provide a detailed derivation of the asymptotic properties. 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). The structure of the proof builds on the arguments used in the proof of Theorem 2 in Halunga and Orme (2009), who consider general misspecifica-

[1]Their findings complement and extend the earlier work of Officer (1973) and Schwert (1989).

[2]For recent results on properties and estimation of GARCH models with explanatory variables that enter in an additive fashion see Han and Kristensen (2014), Han (2015) and Francq and Thieu (2015).

tion tests for GARCH models. The important difference between our test and theirs is that we consider a situation in which the second component is driven by an explanatory variable that may or may not be generated outside the model. In addition, Halunga and Orme (2009) consider additive components only and focus on estimation effects from the correct specification of the conditional mean. In our set-up, the volatility components are multiplicative, causing substantial differences in the likelihood and test statistic. For simplicity, we assume that returns have mean zero, thus abstracting from estimation effects from the mean. In order to derive the asymptotic distribution of the test statistic, we require the standard assumptions on the GARCH parameters and the innovation term for the pure GARCH model. In addition, our test needs assumptions on the moments of the explanatory variable as well as on the observed (return) process. A nice property of our LM test is that it will not depend on the functional form of the long-term component under the alternative. Further, the test statistic is χ^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 which require estimation of a restricted model under the alternative. In a Monte-Carlo simulation, we find good size and power properties in finite samples. Moreover, we illustrate the usefulness of our procedure by two empirical applications to S&P 500 return data.

The model under the alternative hypothesis is closely related to the GARCH-MIDAS of Engle et al. (2013). 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. Therefore, Wald-type tests like simple t- or F-tests are not straightforward to employ in this context. 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 restrictive parameter space which does not admit our null hypothesis. We illustrate how our test can be applied even in settings with mixed-frequency data and, thus, can be used as a preliminary check before estimating a GARCH-MIDAS model.

Our test statistic is also linked to the 'ARCH nested in GARCH' test for evaluating GARCH models as proposed by Lundbergh and Teräsvirta (2002). Although it is important to point out that the test by Lundbergh and Teräsvirta (2002) should be considered as a general misspecification test without a well-specified alternative, it is possible to think of their 'nested ARCH component' as our long-term component with a specific choice for the explanatory variable. Despite this analogy, the specification of their short-term component is fundamentally different from ours. Under the alternative, in their short-term component the squared observations are not divided by the long-term component, which implies that the short-term component is not a GARCH process and, thereby, leads to a different test indicator. In the Monte-Carlo simulation, we show that even if we modify their test in order to allow for a general explanatory variable, the difference in the specification of their short-term component leads to a considerable loss in power in comparison to our test statistic.

Finally, our work complements recent research on misspecification testing in multiplicative component models of the smooth transition type by Amado and Teräsvirta (2015), in the Realized GARCH model by Lee and Halunga (2015) and on the estimation of semiparametric multiplicative component models by Han and Kristensen (2015).

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.

2 Model and Test Statistic

In Section 2.1, we first introduce the multiplicative two-component GARCH specification and then discuss the null hypothesis of our test. The relationship between the two-component model and the GARCH-MIDAS specification is explored in Section 2.2. We derive the likelihood function and the test indicator in Section 2.3 and present our main result on the asymptotic distribution of the test statistic in Section 2.4. Section 2.5 provides a comparison with the 'ARCH nested in GARCH' test and 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.[3] σ_{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 σ_{0t}^2 into a GARCH component ('short-term component') and a component that is driven by an explanatory variable:

$$\sigma_{0t}^2 = \bar{h}_{0t}^{\infty} \tau_{0t} \tag{2}$$

Follow the terminology used in Engle et al. (2013), we refer to the second component as a 'long-term component'. This is, because in our setting the second component is typically much smoother than the GARCH component.

The short-term component is specified as a mean-reverting GARCH(1,1):

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

with $\alpha_0 + \beta_0 < 1$. We denote the vector of true parameters in the GARCH component as $\boldsymbol{\eta}_0 = (\omega_0, \alpha_0, \beta_0)'$.

The $\tau_{0,t}$ component is assumed to depend on the K lagged values of an explanatory variable x_t. It can be thought of as describing smooth movements in the conditional variance as a function of the weighted sum of the lagged values of the explanatory variable:

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

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

Assumption 1. *The parameter space* $\boldsymbol{\Pi}$ *is a compact subset of* $\mathbb{R}^K$ *and* $\boldsymbol{\pi}_0$ *lies in the interior of* $\boldsymbol{\Pi}$.

Assumption 2. *Let* f *be a known positive function, i.e.* $f(\cdot) > 0$, *continuously differentiable, with* $f(0) = 1$, $f_0' = \frac{\partial \tau_t}{\partial \boldsymbol{\pi}' \mathbf{x}_t}\big|_{\boldsymbol{\pi}=\mathbf{0}} \neq 0$.

[3]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 assumption that $f(\cdot) > 0$ allows 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. The main example that we have in mind for $f(\cdot)$ is the exponential specification

$$f(\boldsymbol{\pi}_0'\mathbf{x}_t) = \exp(\boldsymbol{\pi}_0'\mathbf{x}_t). \tag{5}$$

This model has been used, among others, in Engle et al. (2013), Opschoor et al. (2014) and Conrad and Loch (2015a). While Engle et al. (2013) and Conrad and Loch (2015a) used realized volatility as an explanatory variable, Opschoor et al. (2014) opted for the Bloomberg Financial Conditions Index.

Using the above notation, we are interested in testing $H_0 : \boldsymbol{\pi}_0 = \mathbf{0}$ against the two-sided alternative $H_1 : \boldsymbol{\pi}_0 \neq \mathbf{0}$. Under H_0, the long-term component is equal one and the two component model reduces to the nested GARCH(1,1) with unconditional variance $\sigma_0^2 = \omega_0/(1 - \alpha_0 - \beta_0)$.[4]

Note that equation (3) is specified such that it can be rewritten as an ARCH(∞)

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

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. We make the following assumptions about the GARCH parameters and the innovation Z_t.

Assumption 3. $\boldsymbol{\eta}_0 \in \Theta$ *where the parameter space is given by* $\Theta = \{\boldsymbol{\eta} = (\omega, \alpha, \beta)' \in \mathbb{R}^3 | 0 < \omega < \overline{\omega}, 0 < \alpha, 0 < \beta, \alpha + \beta < 1\}$.

Assumption 4. *We denote by* $\mathcal{F}_{t-1}$ *the* σ*-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* κ_Z *is a finite constant. Further,* Z_t^2 *has a nondegenerate distribution.*

Assumptions 3 and 4 imply that $\sqrt{\bar{h}_{0t}^{\infty}} Z_t$ is a covariance-stationary process with unconditional variance σ_0^2. Furthermore, by Jensen's inequality they imply that $\mathbf{E}[\ln(\alpha_0 Z_t^2 + \beta_0)] < 0$ which ensures that under the null ε_t is strictly stationary and ergodic (see, e.g.,

[4]Later on, we also consider the one-sided alternative $H_1 : \boldsymbol{\pi}_0 \neq \mathbf{0},\ \boldsymbol{\pi}_0 \geq \mathbf{0}$ (latter elementwise). See Remark 3 in Section 2.4.

Francq and Zakoïan, 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.

2.2 Relation to GARCH-MIDAS Model

The two-component model presented in the previous section is closely related to the GARCH-MIDAS model suggested in Engle et al. (2013). In their model, the long-term component is typically of the exponential form and the weights in the long-term component are parsimoniously parameterized as $\pi_{0,k} = \tilde{\pi}_0 \psi_{0k}$, where the $\psi_{0k} \geq 0$, $k = 1, \ldots, K$, are typically generated from a Beta weighting scheme. The parameter $\tilde{\pi}_0$ then determines the sign of the effect of x_t on long-term volatility. Alternatively, Engle et al. (2013) consider a linear long-term component. However, the linear specification of $f(\cdot)$ can only be used in combination with non-negative explanatory variables and requires $\tilde{\pi}_0 \geq 0$. For this model, 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 ε_t and establish consistency and asymptotic normality of the QMLE. However, the proof of the asymptotic normality of the QMLE crucially relies on the assumption that $\tilde{\pi}_0 > 0$ and $\psi_{0k} > 0$ for $k = 1, \ldots, K$ and, hence, their framework does not directly allow to test the null that the lagged x_t are jointly insignificant (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 show that it is straightforward to extend our LM test statistic to the mixed-frequency situation.

2.3 Likelihood Function and Partial Derivatives

We denote the processes that can be constructed from the parameter vectors $\boldsymbol{\eta} = (\omega, \alpha, \beta)'$ and $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)'$ given initial observations for ε_t and x_t by $\bar{h}_t$ and τ_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 can be written as

$$L_T^\infty(\boldsymbol{\eta}, \boldsymbol{\pi}|\varepsilon_T, x_T, \varepsilon_{T-1}, x_{T-1}, \ldots) = \sum_{t=1}^{T} l_t^\infty \tag{6}$$

with

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

Similarly, conditional on initial values $(\varepsilon_0, \bar{h}_0 = 0, \mathbf{x}_0)$ the observed quasi-log-likelihood can be written as

$$L_T(\boldsymbol{\eta}, \boldsymbol{\pi}|\varepsilon_T, x_T, \varepsilon_{T-1}, x_{T-1}, \ldots, \varepsilon_1, x_1) = \sum_{t=1}^{T} l_t \tag{8}$$

with

$$l_t = -\frac{1}{2}\left[\ln(\bar{h}_t) + \ln(\tau_t) + \frac{\varepsilon_t^2}{\bar{h}_t \tau_t}\right]. \tag{9}$$

2.3.1 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

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

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

$$\frac{\partial l_t^\infty}{\partial \boldsymbol{\eta}} = \frac{1}{2}\left[\frac{\varepsilon_t^2}{\bar{h}_t^\infty \tau_t} - 1\right]\left(\frac{1}{\bar{h}_t^\infty}\frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\eta}} + \frac{1}{\tau_t}\frac{\partial \tau_t}{\partial \boldsymbol{\eta}}\right) \tag{10}$$

with $\partial\tau_t/\partial\boldsymbol{\eta} = (\partial f_t/\partial\boldsymbol{\pi}'\mathbf{x}_t)(\partial\mathbf{x}_t/\partial\boldsymbol{\eta})'\boldsymbol{\pi}$. 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^∞ corresponds to the standard expression of the conditional variance in a GARCH(1,1). We then distinguish between

$$\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}) = \frac{\partial l_t^\infty}{\partial \boldsymbol{\eta}}\bigg|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{2}\left[\frac{\varepsilon_t^2}{h_t^\infty} - 1\right]\mathbf{y}_t^\infty \tag{11}$$

with

$$\mathbf{y}_t^\infty = \frac{1}{\bar{h}_t^\infty}\frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\eta}}\bigg|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{h_t^\infty}\sum_{i=0}^\infty \beta^i \mathbf{s}_{t-i}^\infty, \tag{12}$$

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

$$\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0) = \frac{1}{2}\left[\frac{\varepsilon_t^2}{h_{0,t}^\infty} - 1\right]\mathbf{y}_{0,t}^\infty, \tag{13}$$

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:

$$\frac{\partial l_t^\infty}{\partial \boldsymbol{\pi}} = \frac{1}{2}\left[\frac{\varepsilon_t^2}{\bar{h}_t^\infty \tau_t} - 1\right]\left(\frac{1}{\bar{h}_t^\infty}\frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\pi}} + \frac{1}{\tau_t}\frac{\partial \tau_t}{\partial \boldsymbol{\pi}}\right), \tag{14}$$

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

$$\frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\pi}} = -\alpha \sum_{j=0}^\infty \beta^j \frac{\varepsilon_{t-1-j}^2}{\tau_{t-1-j}^2}\frac{\partial \tau_{t-1-j}}{\partial \boldsymbol{\pi}}. \tag{15}$$

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 $\partial\tau_t/\partial\boldsymbol{\pi}|_{\boldsymbol{\pi}=\mathbf{0}} = f_0'\mathbf{x}_t$ and, hence,

$$\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}) = \frac{\partial l_t^\infty}{\partial \boldsymbol{\pi}}\bigg|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{2}\left[\frac{\varepsilon_t^2}{h_t^\infty} - 1\right]\mathbf{r}_t^\infty \tag{16}$$

with

$$\mathbf{r}_t^\infty = f_0'\left(\mathbf{x}_t - \alpha\frac{1}{h_t^\infty}\sum_{j=0}^\infty \beta^j \varepsilon_{t-1-j}^2 \mathbf{x}_{t-1-j}\right). \tag{17}$$

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

$$\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0) = \frac{1}{2}\left[\frac{\varepsilon_t^2}{h_{0,t}^\infty} - 1\right]\mathbf{r}_{0,t}^\infty \tag{18}$$

with

$$\mathbf{r}_{0,t}^\infty = f_0'\left(\mathbf{x}_t - \alpha_0\frac{1}{h_{0,t}^\infty}\sum_{j=0}^\infty \beta_0^j \varepsilon_{t-1-j}^2 \mathbf{x}_{t-1-j}\right). \tag{19}$$

In summary, we have

$$\mathbf{D}^\infty(\boldsymbol{\eta}_0) = \frac{1}{T}\sum_{t=1}^T \mathbf{d}_t^\infty(\boldsymbol{\eta}_0) = \frac{1}{2T}\sum_{t=1}^T\left[\frac{\varepsilon_t^2}{h_{0,t}^\infty} - 1\right]\begin{pmatrix}\mathbf{y}_{0,t}^\infty \\ \mathbf{r}_{0,t}^\infty\end{pmatrix}. \tag{20}$$

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

$$\begin{aligned}
\mathbf{Var}[\mathbf{d}_t^\infty(\boldsymbol{\eta}_0)] = \boldsymbol{\Omega} &= \begin{pmatrix} \boldsymbol{\Omega}_{\eta\eta} & \boldsymbol{\Omega}_{\eta\pi} \\ \boldsymbol{\Omega}_{\pi\eta} & \boldsymbol{\Omega}_{\pi\pi} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{E}[\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0)\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0)'] & \mathbf{E}[\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0)\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0)'] \\ \mathbf{E}[\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0)\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0)'] & \mathbf{E}[\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0)\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0)'] \end{pmatrix} \\
&= \frac{1}{4}(\kappa_Z - 1) \begin{pmatrix} \mathbf{E}[\mathbf{y}_{0,t}^\infty(\mathbf{y}_{0,t}^\infty)'] & \mathbf{E}[\mathbf{y}_{0,t}^\infty(\mathbf{r}_{0,t}^\infty)'] \\ \mathbf{E}[\mathbf{r}_{0,t}^\infty(\mathbf{y}_{0,t}^\infty)'] & \mathbf{E}[\mathbf{r}_{0,t}^\infty(\mathbf{r}_{0,t}^\infty)'] \end{pmatrix}. \qquad (21)
\end{aligned}$$

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 $\frac{1}{\sqrt{T}}\sum_{t=1}^T \mathbf{d}_t^\infty(\boldsymbol{\eta}_0)$.

2.3.2 Second derivatives

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

$$\frac{\partial \mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'} = -\frac{1}{2}\frac{\varepsilon_t^2}{h_t^\infty}\mathbf{y}_t^\infty(\mathbf{y}_t^\infty)' + \frac{1}{2}\left[\frac{\varepsilon_t^2}{h_t^\infty} - 1\right]\frac{\partial \mathbf{y}_t^\infty}{\partial \boldsymbol{\eta}'} \qquad (22)$$

and

$$\frac{\partial \mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'} = -\frac{1}{2}\frac{\varepsilon_t^2}{h_t^\infty}\mathbf{r}_t^\infty(\mathbf{y}_t^\infty)' + \frac{1}{2}\left[\frac{\varepsilon_t^2}{h_t^\infty} - 1\right]\frac{\partial \mathbf{r}_t^\infty}{\partial \boldsymbol{\eta}'}. \qquad (23)$$

We then define

$$\mathbf{J}_{\eta\eta} = -\mathbf{E}\left[\frac{\partial \mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0)}{\partial \boldsymbol{\eta}'}\right] = \frac{1}{2}\mathbf{E}[\mathbf{y}_{0,t}^\infty(\mathbf{y}_{0,t}^\infty)'] \qquad (24)$$

and

$$\mathbf{J}_{\pi\eta} = -\mathbf{E}\left[\frac{\partial \mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0)}{\partial \boldsymbol{\eta}'}\right] = \frac{1}{2}\mathbf{E}[\mathbf{r}_{0,t}^\infty(\mathbf{y}_{0,t}^\infty)']. \qquad (25)$$

Note that $\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0)$ corresponds to the score of observation t in a standard GARCH(1,1) model and $\partial \mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0)/\partial \boldsymbol{\eta}'$ to the respective second derivative. Under Assumptions 3 and 4, 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.

2.4 The LM Test Statistic

The LM test statistic will be based on the observed quantity $\mathbf{D}_{\boldsymbol{\pi}}(\hat{\boldsymbol{\eta}}) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{d}_{\boldsymbol{\pi},t}(\hat{\boldsymbol{\eta}})$, 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:

$$\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\hat{\boldsymbol{\eta}}) = [\mathbf{J}_{\boldsymbol{\pi\eta}}\mathbf{J}_{\boldsymbol{\eta\eta}}^{-1} \ : \ \mathbf{I}]\sqrt{T}\mathbf{D}^{\infty}(\boldsymbol{\eta}_0) + o_P(1) \tag{26}$$

In the final step it is necessary to show that the observed quantity $\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}(\hat{\boldsymbol{\eta}})$ has the same asymptotic distribution as $\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\hat{\boldsymbol{\eta}})$. The LM statistic follows the usual χ^2 distribution.

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 3 and 4 hold and the model is estimated under the null, the QMLE of the GARCH(1,1) parameters will be consistent and asymptotically normal:

$$\sqrt{T}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, (\kappa_Z - 1)(\mathbf{E}[\mathbf{y}_{0,t}^{\infty}(\mathbf{y}_{0,t}^{\infty})'])^{-1}) \tag{27}$$

Remark 1. *In principle, we can relax the assumption that Z_t is i.i.d. Following Escanciano (2009) and Francq and Thieu (2015), the asymptotic normality of the QMLE can be also 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 asymtotic distribution of the QMLE is given by*

$$\sqrt{T}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{J}_{\boldsymbol{\eta\eta}}^{-1}\widetilde{\boldsymbol{\Omega}}_{\boldsymbol{\eta\eta}}\mathbf{J}_{\boldsymbol{\eta\eta}}^{-1}), \tag{28}$$

where $\widetilde{\boldsymbol{\Omega}}_{\boldsymbol{\eta\eta}} = \mathbf{E}[(\mathbf{E}[Z_t^4|\mathcal{F}_{t-1}] - 1)\mathbf{y}_{0,t}^{\infty}(\mathbf{y}_{0,t}^{\infty})']$. Clearly, if $\mathbf{E}[Z_t^4|\mathcal{F}_{t-1}]$ is constant, (28) simplifies to (27).

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 assume that x_t has a finite fourth moment. Additionally, we require that the long-term component is minimal in the sense that no equivalent representation which is of lower order exists.

Assumption 5. *x_t is strictly stationary and ergodic with $\mathbf{E}[|x_t|^4] < \infty$. There exist no $a_1, \ldots, a_S$ for the long-term component* (4) *such that $\sum_{k=1}^{K} \pi_{0k} x_{t-k} = \sum_{s=1}^{S} a_s x_{t-s}$ with $S < K$.*

By Assumption 4, 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, the unemployment rate or disagreement among forecasters – has been used (see Engle et al., 2013, or Conrad and Loch, 2015a). Wang and Ghysels (2015) show that the GARCH-MIDAS model with rolling window realized volatility as 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 2.5). For testing the simple GARCH model against the former model, Assumption 5 requires that under the null the observed process has a finite eighth moment: $\mathbf{E}[|\varepsilon_t|^8] < \infty$. The corresponding constraints on the parameters of the GARCH(1,1) are provided in Francq and Zakoïan (2010), equation (2.54).

Theorem 1. *If Assumptions 3-5 hold, then*

$$\sqrt{T}\mathbf{D}^{\infty}(\boldsymbol{\eta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}). \tag{29}$$

In the proof we use the fact that $\boldsymbol{\Omega}_{\boldsymbol{\eta}\boldsymbol{\eta}}$ is finite and positive definite which follows from Theorem 2.2 in Francq and Zakoïan (2004).

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 $\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}$ can be consistently estimated by

$$-\frac{1}{T}\sum_{t=1}^{T}\frac{\partial \mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\tilde{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}'},$$

where $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}_0 + o_P(1)$. The result is presented in Proposition 1 in Appendix A. This requires the following Assumption 6 which ensures that $\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}(\boldsymbol{\eta})$ is finite with a uniform bound for all $\boldsymbol{\eta} \in \Theta$.

Assumption 6. $\mathbf{E}[|\varepsilon_t|^{4(1+s)}] < \infty$ *for some $s \in (0, 1)$.*

Note that in general $\varepsilon_t^2 = \bar{h}_{0t}^{\infty}\tau_{0t}Z_t^2$ depends on $\boldsymbol{\eta}_0$ and $\boldsymbol{\pi}_0$. Under the null, $\varepsilon_t^2 = h_{0,t}^{\infty}Z_t^2$ depends on $\boldsymbol{\eta}_0$ only. In the proof of Proposition 1 we will use this observation to argue that $\mathbf{E}[\sup_{\boldsymbol{\eta}}|\varepsilon_t|^{4(1+s)}] = \mathbf{E}[|\varepsilon_t|^{4(1+s)}]$.

Theorem 2. *If Assumptions 3-6 hold, then*

$$\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\hat{\boldsymbol{\eta}}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \tag{30}$$

with

$$\begin{aligned}\boldsymbol{\Sigma} &= \boldsymbol{\Omega}_{\boldsymbol{\pi}\boldsymbol{\pi}} - \mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\boldsymbol{\Omega}_{\boldsymbol{\pi}\boldsymbol{\eta}}' \\ &= \frac{1}{4}(\kappa_Z - 1)\left(\mathbf{E}[\mathbf{r}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})'] - \mathbf{E}[\mathbf{r}_{0,t}^{\infty}(\mathbf{y}_{0,t}^{\infty})']\left(\mathbf{E}[\mathbf{y}_{0,t}^{\infty}(\mathbf{y}_{0,t}^{\infty})']\right)^{-1}\mathbf{E}[\mathbf{y}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})']\right).\end{aligned} \tag{31}$$

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

Theorem 3. *If Assumptions 3-6 hold, then*

$$\begin{aligned} LM &= T\mathbf{D}_{\boldsymbol{\pi}}(\hat{\boldsymbol{\eta}})'\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{D}_{\boldsymbol{\pi}}(\hat{\boldsymbol{\eta}}) \\ &= \frac{1}{4T}\left(\sum_{t=1}^{T}\left[\frac{\varepsilon_t^2}{\hat{h}_t} - 1\right]\hat{\mathbf{r}}_t\right)'\widehat{\boldsymbol{\Sigma}}^{-1}\left(\sum_{t=1}^{T}\left[\frac{\varepsilon_t^2}{\hat{h}_t} - 1\right]\hat{\mathbf{r}}_t\right) \overset{a}{\sim} \chi^2(K)\end{aligned} \tag{32}$$

where $\hat{\boldsymbol{\eta}} = (\hat{\omega}, \hat{\alpha}, \hat{\beta})'$ is the vector of parameter estimates from the model under the null, $\hat{h}_t = \hat{\omega} + \hat{\alpha}\varepsilon_{t-1}^2 + \hat{\beta}\hat{h}_{t-1}$, $\hat{\mathbf{r}}_t = f_0'(\mathbf{x}_t - \hat{\alpha}/\hat{h}_t\sum_{j=0}^{t-1}\hat{\beta}^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j})$ and

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{4T}\widehat{(\kappa_Z - 1)}\left(\sum_{t=1}^{T}\hat{\mathbf{r}}_t\hat{\mathbf{r}}_t' - \sum_{t=1}^{T}\hat{\mathbf{r}}_t\hat{\mathbf{y}}_t'\left(\sum_{t=1}^{T}\hat{\mathbf{y}}_t\hat{\mathbf{y}}_t'\right)^{-1}\sum_{t=1}^{T}\hat{\mathbf{y}}_t\hat{\mathbf{r}}_t'\right) \tag{33}$$

with $\widehat{(\kappa_Z - 1)} = 1/T\sum_{t=1}^{T}(\varepsilon_t^2/\hat{h}_t - 1)^2$ is a consistent estimator of $\boldsymbol{\Sigma}$.

Note that the LM test statistic does not depend on the constant f_0' because the $(f_0')^2$ in the 'numerator' and the 'denominator' of the test statistic cancels out.

Remark 2. *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 ε_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 monotonically decaying weights, $\pi_{0,k}$, choosing $K = 1$ is sufficient in order to detect whether x_t has an effect on long-term volatility or not. Given that in applications the explanatory variable is likely to be persistent, this result is not surprising because for persistent x_t all entries of $\hat{\mathbf{r}}_t$ will basically carry the same information so that choosing $K = 1$ is sufficient.[5]

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

$$LM = T\,\frac{SSR_0 - SSR_1}{SSR_0}, \tag{34}$$

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 $\hat{\mathbf{y}}_t'$, where $\hat{\mathbf{y}}_t$ is obtained by inserting the respective estimated quantities in equation (12). Hence, LM is simply T times the uncentered R^2 of the regression.

Remark 3. *Wang and Ghysels (2015) consider a specification for $f(\cdot)$ which is linear in the lagged explanatory variable. In this case the long-term component is specified as*

$$f(\boldsymbol{\pi}'\mathbf{x}_t) = 1 + \boldsymbol{\pi}'\mathbf{x}_t \tag{35}$$

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. 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 directly follows from the discussion in Francq and Zakoïan (2009) who consider testing the nullity of coefficients in GARCH processes. While the asymptotic distribution of the Lagrange multiplier test remains the same (because the score vector is asymptotically Gaussian under the null), Francq and Zakoïan (2009) show that the asymptotic distribution of the Wald and Likelihood ratio test would no longer be χ^2 since the asymptotic distribution of the QMLE of

[5] In the extreme case that x_t is constant, $\hat{\mathbf{r}}_t$ collapses to a vector of zeros.

the unrestricted model is non-standard under the null hypothesis. 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 4. *In principle, we could also use our test for testing the joint significance of several explanatory variables. In this case $\mathbf{x}_t$ would not include the lagged values of x_t but different explanatory variables, say, $\mathbf{x}_t = (v_{t-1}, w_{t-1}, \ldots)$ and K would represent the number of different explanatory variables.*

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 alterative, 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 (20) can be rewritten as

$$\mathbf{D}^{\infty}(\boldsymbol{\eta}_0) = \frac{1}{2T}\sum_{t=1}^{T}\left[\frac{\varepsilon_t^2}{\sigma_0^2} - 1\right]\begin{pmatrix} \sigma_0^{-2} \\ f_0'\mathbf{x}_t \end{pmatrix}. \tag{36}$$

Then, the regression-based test simplifies to regressing the squared returns on a constant and $\mathbf{x}_t$ and to computing TR^2 which resembles the Godfrey (1978) test for multiplicative heteroskedasticity. Finally, the Engle (1982) test for ARCH effects is obtained if we choose $x_{t-k} = \varepsilon_{t-k}^2$.

2.5 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}^{∞} is defined as before and $\tau_{0t} = 1 + \boldsymbol{\pi}'\mathbf{x}_t$, i.e. they assume that the long-term component is linear. Lundbergh and Teräsvirta (2002) make the specific choice of $x_t = \xi_{0t}^2 = \varepsilon_t^2/h_{0t}^{\infty}$ for the explanatory variable in the long-term component. 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}^{∞} (instead of $\bar{h}_{0t}^{\infty}$), the squared observation ε_{t-1}^2

is *not* divided by $\tau_{0,t-1}$. Because of this, $\sqrt{h_{0t}^{\infty}}Z_t$ follows a GARCH(1,1) process under the null but *not* under the alternative.[6] Moreover, it follows that $\partial h_t^{\infty}/\partial \boldsymbol{\pi} = \mathbf{0}$ and, hence, in the Lundbergh and Teräsvirta (2002) setting equation (19) reduces to $\mathbf{r}_{0,t}^{\infty} = (\varepsilon_{t-1}^2/h_{0,t-1}^{\infty}, \varepsilon_{t-2}^2/h_{0,t-2}^{\infty}, \ldots, \varepsilon_{t-K}^2/h_{0,t-K}^{\infty})'$. Thus, their LM test statistic is based on

$$\left[\frac{\varepsilon_t^2}{\hat{h}_t} - 1\right] \hat{\mathbf{r}}_t^{LT}, \tag{37}$$

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 $\varepsilon_{t-k}^2/\hat{h}_{t-k}$, $k = 1, \ldots, K$. Intuitively, equation (37) 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 test in situations in which the true data generating process (DGP) has a two-component structure. We implement a regression-based version of the test as in equation (34) but with $\hat{\mathbf{r}}_t^{LT}$ instead of $\hat{\mathbf{r}}_t$. We denote the test statistic 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 (19) 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.[7]

2.6 Mixed-Data Sampling

As discussed in Section 2.2, 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 denote

[6] The observation that h_{0t}^{∞} does not follow a GARCH process under the alternative is closely related to the argument in Halunga and Orme (2009) that the alternative models considered in Lundbergh and Teräsvirta (2002) are not "recursive" in nature.

[7] Amado and Teräsvirta (2015) discuss testing the null of no remaining ARCH effects in multiplicative time-varying GARCH models. Our test is closely related to the model they discuss in Section 4.4 of their paper. However, they provide no asymptotic theory for the test with exogenous explanatory variables.

by x_t 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 $i = 1, \ldots, M$ refers to the M days within each month/quarter.[8] Equation (1) can then be rewritten as

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

whereby Assumption 4 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:

Alternative 1: 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:

$$\tilde{\varepsilon}_t = \sum_{i=1}^{M} \frac{\varepsilon_{i,t}}{\sqrt{\bar{h}^{\infty}_{0,i,t}}} = \sqrt{\tau_{0,t}}Z_t, \tag{39}$$

where $Z_t = \sum_{i=1}^{M} Z_{i,t}$ is i.i.d. with mean zero and variance M by Assumption 4. This leads to the score vector:

$$\mathbf{d}_t(\boldsymbol{\eta}_0) = \sum_{t=1}^{T} \left(\frac{\tilde{\varepsilon}_t^2}{M} - 1 \right) \begin{pmatrix} M^{-1} \\ f_0' \mathbf{x}_t \end{pmatrix} \tag{40}$$

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

$$\hat{\tilde{\varepsilon}}_t = \sum_{i=1}^{M} \frac{\varepsilon_{i,t}}{\sqrt{\hat{h}_{i,t}}}, \tag{41}$$

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

[8]For simplicity in the notation, we assume that each month/quarter has the same number of (trading) days.

pre-estimating $\bar{h}^{\infty}_{0,i,t}$:

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

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

Alternative 2: The second alternative is based on what we call the *volatility-adjusted realized variance* which we define as the monthly/quarterly realized variance of the deGARCHed daily returns. As we will discuss below, this approach is closely related to the empirical literature on predictive regressions for financial volatility. Consider again the deGARCHed daily returns, but now sum the squares:

$$\widetilde{RV}_t = \sum_{i=1}^{M} \frac{\varepsilon^2_{i,t}}{\bar{h}^{\infty}_{0,i,t}} = \tau_{0,t} \sum_{i=1}^{M} Z^2_{i,t} = \tau_{0,t} \tilde{Z}_t, \tag{42}$$

where $\widetilde{RV}_t$ denotes the monthly/quarterly realized variance of the deGARCHed returns. To simplify the analysis further, we assume that the long-term component is given by $\tau_{0,t} = \exp(\boldsymbol{\pi}\mathbf{x}_t)$. This Assumption is plausible 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 (42) as a regression model:

$$\ln(\widetilde{RV}_t) \quad = \quad \boldsymbol{\pi}\mathbf{x}_t + \ln(\tilde{Z}_t) = \tilde{c} + \boldsymbol{\pi}\mathbf{x}_t + \tilde{\zeta}_t, \tag{43}$$

where $\tilde{c} = \mathbf{E}[\ln(\tilde{Z}_t)]$ and $\tilde{\zeta}_t = \ln(\tilde{Z}_t) - \tilde{c}$. Note that by Assumption 4 the innovation $\tilde{\zeta}_t$ is i.i.d. This is, under $H_0 : \boldsymbol{\pi} = \mathbf{0}$ the volatility-adjusted realized variance $\ln(\widetilde{RV}_t)$ should be unpredictable. Of course, we have to replace $\widetilde{RV}_t$ in equation (43) with the estimate $\widehat{\widetilde{RV}}_t = \sum_{i=1}^{M} \varepsilon^2_{i,t}/\hat{h}^{\infty}_t$ in order to obtain an estimable version. Again, this will introduce measurement error.

Equation (43) 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, Conrad and Loch, 2015a, 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:

$$\ln(RV_t) = c + \boldsymbol{\pi}\mathbf{x}_t + \zeta_t \tag{44}$$

where $RV_t = \sum_{i=1}^{M} \varepsilon_{i,t}^2$. From equation (38) it follows that the innovation in equation (44) 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 ζ_t is a low-frequency process that corresponds to the sum of a squared high-frequency GARCH process. That is, while $\tilde{\zeta}_t$ is i.i.d., we can expect that ζ_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 (44) rather than equation (43) as a regression model.

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

$$\ln(RV_t) = c + \boldsymbol{\pi}\mathbf{x}_t + \rho \ln(RV_{t-1}) + \zeta_t. \tag{45}$$

Since the additional regressor $\ln(RV_{t-1})$ provides a parsimonious (but noisy) summary of the information included in $\mathbf{x}_{t-1}$ and requires to estimate only one additional parameter, it will reduce the relevance of including the whole $\mathbf{x}_t$ vector in the regression. This might explain, why in the predictive regressions literature $\mathbf{x}_t$ is often found to be insignificant once the lagged realized volatility is controlled for. The point that predictive regressions like in equation (45) might be problematic has been made already by Engle et al. (2013) who argue that $\ln(RV_{t-1})$ is a noisy measure of the true unobservable long-term component which creates problems due to measurement error on both the left as well as the right hand sight of the equation.

3 Simulation

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 $M =$

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.[9] The sample is chosen such that $T = 1000$. 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.

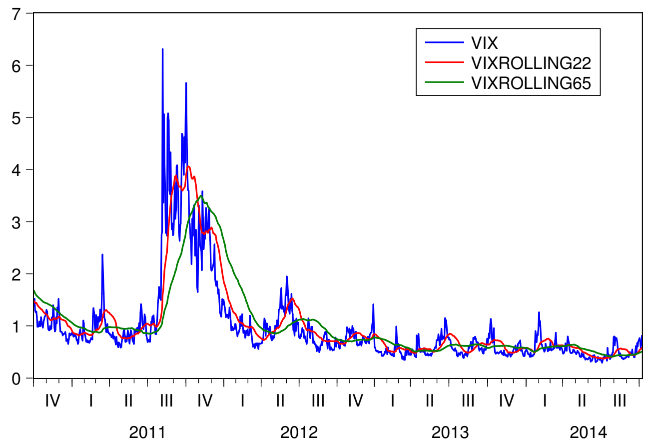

Figure 1: The figure shows the evolution of VIX_t (blue), $VIX_t^{(22)}$ (red) and $VIX_t^{(65)}$ (green) for the period October 2010 to October 2014. The three variables are presented in daily units.

3.1 Size properties

We first discuss the size properties of the different versions of the test statistic. Three alternative GARCH(1,1) specifications are considered. These three specifications reflect different degrees of persistence (Low: L, Moderate: M, High: H) in the conditional variance, whereby we measure persistence by $\alpha_0 + \beta_0$. We keep β_0 fixed at 0.9 and

[9] 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.

increase α_0 from 0.05 to 0.09. ω_0 is always chosen such that under the null $\sigma_0^2 = 1$.

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

Table 1: Empirical size.

		$Z_t \sim \mathcal{N}(0,1)$			$Z_t \sim t(7)$		
		L	M	H	L	M	H
		Panel A: $x_t = \varepsilon_t^2/\hat{h}_t$					
	1%	0.9	1.2	1.3	0.7	0.9	0.7
LM	5%	4.6	5.0	5.2	3.1	3.7	3.9
	10%	9.0	9.7	10.1	7.2	7.4	7.3
	1%	0.9	1.2	1.3	0.9	1.1	1.1
LM_{LT}	5%	5.2	5.2	5.1	3.4	3.8	3.9
	10%	10.2	10.0	10.6	6.7	7.1	7.4
		Panel B: $x_t = VIX_t$					
	1%	1.9	2.2	1.7	2.3	2.6	2.4
LM	5%	6.4	6.0	5.7	6.0	6.1	5.5
	10%	11.2	11.4	11.6	11.4	11.0	9.5
	1%	2.0	2.1	2.3	1.1	1.2	1.9
$LM_{LT,mod}$	5%	7.0	7.1	7.9	5.5	6.4	7.1
	10%	12.2	12.8	13.8	11.3	12.0	12.6

Notes: Entries are rejection rates in percent over the 1000 replications at the 1%, 5% and 10% nominal level. The model for the conditional variance is a GARCH(1,1) with $\beta_0 = 0.90$. L, I and H refer to GARCH models with low ($\alpha = 0.05$), moderate ($\alpha = 0.07$) and high ($\alpha = 0.09$) persistence. ω is chosen such that $\sigma_0^2 = 1$. All test statistics are based on $K = 1$.

To implement the test, we have to specify the explanatory variable. In Panel A of Table 1 we test for remaining ARCH effects by choosing $x_t = \varepsilon_t^2/\hat{h}_t$ and in Panel B we choose x_t to be equal to the VIX index. We report the empirical size for the LM test given in equation (32) as well as for the original, LM_{LT}, and modified, $LM_{LT,mod}$, test statistics of Lundbergh and Teräsvirta (2002). Also, we have to choose the dimension of $\hat{\mathbf{r}}_t$ and $\hat{\mathbf{r}}_t^{LT}$ ($\hat{\mathbf{r}}_t^{LT,mod}$), respectively. We opt for a dimension of $K = 1$.[10] We first discuss the results when testing for remaining ARCH effects. As Panel A of Table 1 shows, 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 errors, 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). As panel B shows, both tests are modestly oversized when the VIX is used as an explanatory variable.

3.2 Power properties

We simulate the model under the alternative using the exponential specification given by

$$\tau_{0,t} = \exp(\boldsymbol{\pi}_0' \mathbf{x}_t). \tag{46}$$

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 scheme are shown in Figure 2. The red and blue lines represent weights that either have a fast (F) or a slow (S) decay. The second and third weighting scheme are scaled such that their weights add up to 0.3.[11]

Table 2 presents size-adjusted rejection rates that were obtained from the Monte-Carlo simulations. The LM test statistics are based on $\hat{\mathbf{r}}_t$ with $x_t \in \{VIX_t, VIX_t^{(22)}, VIX_t^{(65)}\}$. We present two versions of the Lundbergh and Teräsvirta (2002) test. The *modified* Lundbergh and Teräsvirta (2002) test, $LM_{LT,mod}$, is based on $\hat{\mathbf{r}}_t^{LT,mod}$ but uses the same x_t as in LM. 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'. In order to analyze to what extent the rate of decay affects the power, we simulate processes under the three alternative weighting schemes. For all three test statistics we choose $K = 1$, i.e. the tests are based on the first lag of x_t only.

[10] The results presented below are robust with respect to increasing the dimension of $\hat{\mathbf{r}}_t$ and $\hat{\mathbf{r}}_t^{LT}$ ($\hat{\mathbf{r}}_t^{LT,mod}$). The corresponding tables are available upon request.

[11] The weights were generated using a Beta weight scheme. For a detailed discussion of the Beta weighting scheme see Ghysels et al. (2006).

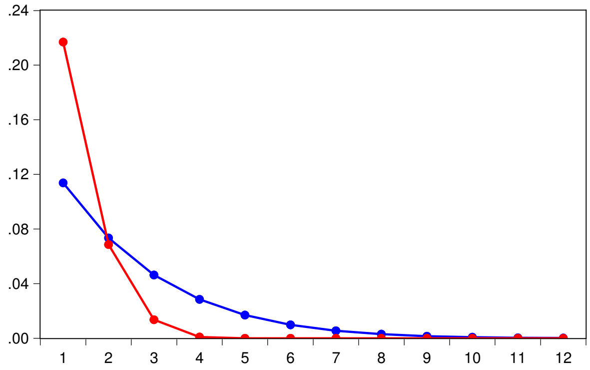

Figure 2: Alternative weighting schemes $\pi_{0,k}$ with fast (F; red) and slow (S; blue) decay.

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.

We first consider the squared VIX as the explanatory variable, i.e. we choose $x_t = VIX_t$. In the GARCH equation we employ models with $\alpha_0 = 0.09$ (high persistence) and $\alpha_0 = 0.07$ (moderate persistence). Besides the size-adjusted power for the different weighting schemes, we also report the variance ratio: $VR = \mathbf{Var}(\ln(\tau_{0t}))/\mathbf{Var}(\ln(\tau_{0t}\bar{h}_{0t}))$, which reflects the fraction of the variance of the log conditional variance that is due to the variance of the log long-term component.[12] For example, for $\alpha_0 = 0.09$ in combination with an immediately decaying weighting scheme, 15.6% of the total conditional variance is due to the long-term component. When α_0 is decreased to 0.07, the VR increases to 35.5%. Intuitively, decreasing α_0, means reducing the variability with which the short-term component fluctuates around τ_{0t}.

First, consider the case where $\alpha_0 = 0.09$. 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 all three tests deteriorates. Nevertheless, the LM still has considerably higher power than $LM_{LT,mod}$. When α_0 is decreased to

[12]This variance ratio has been employed in Conrad and Loch (2015a) as a measure for the relevance of the long-term component. For example, using the realized volatility as an explanatory variable, they find a VR of roughly 13% for data on the S&P 500 for the 1973 to 2010 period.

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

x_t		VIX_t						$VIX_t^{(22)}$	$VIX_t^{(65)}$
		$\omega_{01} = 1$, $\omega_{02} = 10$							
		$\alpha_0 = 0.09$			$\alpha_0 = 0.07$			$\alpha_0 = 0.09$	
weighting scheme		I	F	S	I	F	S	I	I
	1%	55.4	51.1	27.9	60.3	57.1	34.1	12.6	9.2
LM	5%	74.4	71.0	49.9	79.9	77.3	59.4	36.8	19.9
	10%	81.7	79.3	62.0	88.7	87.3	73.0	49.8	32.2
	1%	20.8	18.8	12.9	44.3	41.5	30.2	5.9	3.4
$LM_{LT,mod}$	5%	42.2	39.9	31.2	70.9	69.5	59.0	15.6	11.1
	10%	55.7	52.8	42.5	82.2	80.7	73.0	22.1	18.6
	1%	0.7	0.9	0.8	0.7	0.8	0.8	0.9	0.8
LM_{LT}	5%	5.6	5.6	5.9	5.9	5.6	5.6	5.0	4.7
	10%	9.5	10.1	10.8	11.2	11.2	11.2	9.4	9.3
VR		15.6	15.5	14.8	35.5	35.2	34.1	14.7	12.0

Notes: The table reports the size-adjusted power. 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 $\beta_0 = 0.9$ and $\omega_0 = 1-\alpha_0-\beta_0$. Innovations Z_t are standard normal distributed. The variance ratio, $VR = \mathbf{Var}(\ln(\tau_{0t}))/\mathbf{Var}(\ln(\tau_{0t}\bar{h}_{0t}))$, is the fraction of the variance of the log conditional variance that is due to the variance of the log long-term component. All test statistics are based on $K = 1$.

0.07, 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 α_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 2.5, the difference in the power of the LM and $LM_{LT,mod}$ statistics is less strong when α_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 no power to detect deviation from the null. This is not surprising since LM_{LT} is searching for an omitted ARCH component and, hence, is simply 'searching in the wrong place'.

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 α_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 B.1). 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.

Additionally, we performed several robustness checks.

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 of increasing the sample size on the power, 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.

Choice of K: As a further robustness check, we also performed simulations in which we increased K such that it approaches the true lag length of the fast and slow decaying weighting schemes. Given the smoothness of our explanatory variable (the first order autocorrelation of VIX_t is 0.95), this did not lead to significant gains in power relative to simply choosing $K = 1$ (see also the discussion below Theorem 3).

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 following linear specification: $\tau_{0t} = 1 + \sum_{k=1}^{K} \pi_{0k} x_{t-k}$. Besides the exponential one, the linear specification is most often used in empirical applications. The corresponding results for normally and student-t distributed innovations can be found in Tables 7/8 in Appendix B.2 and, again, qualitatively confirm our previous findings. Note that the linear specification leads to lower variance ratios which explains the difference in power.

Alternative specification under H_1: Under the alternative, we also simulated the *additive* two-component GARCH(1,1) model of Engle and Lee (1999) and applied all three tests. However, neither the original LM_{LT} test nor LM and $LM_{LT,mod}$ detected a deviation from the null in this case. Since the additive two-component GARCH(1,1) has a GARCH(2,2) representation, this result is not surprising. Even thought the GARCH(1,1) under the null is misspecified it might adequately capture the volatility persistence of the GARCH(2,2) by choosing α and β such that the sum is close to one. Hence, the tests that check for multiplicative misspecification search in the wrong place and, consequently, do not reject.

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* daily 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 seven variables that are observed at a daily frequency and check whether these variables might be useful in a two-component GARCH specification. The first explanatory variable is the squared VIX, $VIX_t^{(1)}$. We construct two measures of realized variance. The first one is simply the daily squared return, $RV_t^{(1)} = \varepsilon_t^2$. The second one is the daily realized variance, $\overline{RV}_t^{(1)}$, defined as the sum of the squared five-minute returns within each day. This measure is obtained from the Oxford-Man Institute's "realised library". While the VIX and realized volatility measure stock market uncertainty, we use the Baker et al. (2016) daily index, $EPU_t^{(1)}$, as a measure of general economic policy uncertainty. The last three variables are meant to proxy for macroeconomic conditions. Here, we use the ADS Business Conditions Index, $ADS_t^{(1)}$, suggested by Aruoba et al. (2009) as well as the surprise, $Surp_t^{(1)}$, and uncertainty, $Unc_t^{(1)}$ indices of Scotti

(2016).[13] All seven variables might potentially be useful for predicting future stock market volatility. With the exception of $\overline{RV}_t^{(1)}$, our sample starts in January 1991 and ends in June 2016, i.e. covers 25 years. Unfortunately, $\overline{RV}_t^{(1)}$ is available for the period January 2000 to June 2016 only. In addition to the daily variables, we also consider the 22-days rolling window versions, defined as $x_t^{(22)} = 1/22 \sum_{j=0}^{21} x_{t-j}$.

Table 3 shows the contemporaneous correlations between the seven variables. *Below* the diagonal the correlations for $N = 1$ and *above* the diagonal the correlations for $N = 22$ are provided. For $N = 1$, $VIX_t^{(1)}$ and $\overline{RV}_t^{(1)}$ have the highest correlation (0.76) among all variables. Interestingly, the correlation between $\overline{RV}_t^{(1)}$ and $RV_t^{(1)}$ is only 0.54, presumably due to the fact that $RV_t^{(1)}$ is a noisy measure of daily variance. The other correlations have the expected signs: $VIX_t^{(1)}$ is positively correlated with economic policy uncertainty, $EPU_t^{(1)}$, and uncertainty related to the state of the economy, $Unc_t^{(1)}$, but negatively correlated with the business conditions index, $ADS_t^{(1)}$, and economic data surprises, $Surp_t^{(1)}$. For $N = 22$ the correlations between all variables increase in absolute value. Interestingly, the correlation between $\overline{RV}_t^{(22)}$ and $RV_t^{(22)}$ is now 0.98, showing that the measurement error arising from using daily returns instead of high-frequency returns is much less pronounced when estimating monthly realized variances.

Next, we estimate a GARCH(1,1) for the daily log-returns on the S&P 500 and then apply our *LM* test to each of the variables. As Panel A of Table 4 shows, the *LM* test rejects the null for $VIX_t^{(1)}$, $\overline{RV}_t^{(1)}$, $EPU_t^{(1)}$ and $ADS_t^{(1)}$ at the 1% level. Thus, these variables might be useful predictors of stock market volatility and could be drivers of an omitted second component. The test outcome is line with the previous literature: realized (and expected) variances are found to be useful in GARCH-MIDAS models in Engle et al. (2013) and Conrad and Loch (2015a). Similarly, Dorion (2016) shows that a GARCH-MIDAS model based on the ADS Business Conditions Index is informative for the valuation of options. The finding that the test does not reject the null for $RV_t^{(1)}$ is likely to be due to the fact that $RV_t^{(1)}$ is a noisy measure of the daily variance. At first sight, it might appear counterintuitive that the test rejects for $VIX_t^{(1)}$ and $EPU_t^{(1)}$ but not for $Unc_t^{(1)}$. However, although the three series are positively correlated, $Unc_t^{(1)}$

[13]Data on the ADS can be obtained from the website of the Federal Reserve Bank of Philadelphia. The surprise and uncertainty indices can be downloaded from `https://sites.google.com/site/chiarascottifrb/`.

Table 3: Correlations between explanatory variables, $x_t^{(N)}$, for $N = 1$ and $N = 22$.

	$VIX_t^{(N)}$	$RV_t^{(N)}$	$\overline{RV}_t^{(N)}$	$EPU_t^{(N)}$	$ADS_t^{(N)}$	$Unc_t^{(N)}$	$Surp_t^{(N)}$
$VIX_t^{(N)}$	1.00	0.90	0.91	0.60	-0.68	0.52	-0.36
$RV_t^{(N)}$	0.54	1.00	0.98	0.51	-0.59	0.44	-0.36
$\overline{RV}_t^{(N)}$	0.76	0.53	1.00	0.52	-0.57	0.44	-0.38
$EPU_t^{(N)}$	0.43	0.23	0.31	1.00	-0.43	0.36	-0.23
$ADS_t^{(N)}$	-0.62	-0.31	-0.41	-0.30	1.00	-0.51	0.26
$Unc_t^{(N)}$	0.47	0.22	0.30	0.24	-0.46	1.00	-0.24
$Surp_t^{(N)}$	-0.30	-0.16	-0.24	-0.16	0.20	-0.23	1.00

Notes: The table presents the correlations between the different explanatory variables. Correlations *below* the diagonal correspond to $N = 1$ and correlations *above* the diagonal to $N = 22$. Correlations involving $\overline{RV}_t^{(N)}$ are for the 2000-2016 period. All other correlation figures are for the 1991-2016 period.

often spikes (e.g. in the years 2004, 2005 and 2012) when the other two series do not increase.[14] Potentially this difference is due to the fact that $VIX_t^{(1)}$ (and partly $EPU_t^{(1)}$) are forward-looking, while $Unc_t^{(N)}$ is based on current surprises in macroeconomic releases. Similarly, the news revealed by $Surp_t^{(N)}$ might be instantaneously incorporated in stock markets and, therefore, $Surp_t^{(1)}$ may not be useful for predicting future long-term volatility. Also, all variables for which the test rejects reveal a pronounced cyclical ($ADS_t^{(1)}$) or counter-cyclical ($VIX_t^{(1)}$, $\overline{RV}_t^{(1)}$, $EPU_t^{(1)}$) pattern, while $Unc_t^{(1)}$ and $Surp_t^{(1)}$ are less (counter-)cyclical. Given the empirical observation that long-term stock market volatility is counter-cyclical, this provides a further rationalization for the test outcomes.

The LM tests for the 22-day rolling window versions of the explanatory variables point into the same direction as before. While the test rejects for $VIX_t^{(22)}$ and $\overline{RV}_t^{(22)}$ at the 1% level, it rejects for $EPU_t^{(22)}$ and $ADS_t^{(22)}$ at the 8% level only. As discussed in the simulation section, the test appears to loose power when the explanatory variables become smoother. Interestingly, the test does still not reject for $\overline{RV}_t^{(22)}$. Although the correlation

[14]Scotti (2016, p.16) compares $Unc_t^{(1)}$ with the VIX and economic policy uncertainty and notes that $Unc_t^{(1)}$ "exceeds 1.65 standard deviations above its mean only few times but the peaks do not always correspond with the peaks of the other series suggesting that these uncertainty measures might indeed carry slightly different information."

between $RV_t^{(22)}$ and $\overline{RV}_t^{(22)}$ is 0.98, the two measures differ substantially during crises periods such as October 2008. This is because large daily movements in the S&P 500 typically lead to much stronger increases in $RV_t^{(22)}$ than in $\overline{RV}_t^{(22)}$. Since the fraction of the total conditional variance of daily returns that is potentially due to variation in the 22-days realized variance is relatively small, the test does not detect an effect of $RV_t^{(22)}$ which systematically 'overshoots' during turbulent times.

Also, it is important to note that the LM and $LM_{LT,mod}$ tests lead to the same decision for $N = 1$ but $LM_{LT,mod}$ never rejects for $N = 22$.

Table 4: LM test for S&P 500 returns for the 1991-2016 period.

	Panel A: daily returns						
x_t	$VIX_t^{(1)}$	$RV_t^{(1)}$	$\overline{RV}_t^{(1)}$	$EPU_t^{(1)}$	$ADS_t^{(1)}$	$Unc_t^{(1)}$	$Surp_t^{(1)}$
LM	92.46 [<0.01]	1.63 [0.20]	21.19 [<0.01]	21.63 [<0.01]	6.21 [0.01]	0.02 [0.90]	1.09 [0.30]
$LM_{LT,mod}$	32.62 [<0.01]	1.69 [0.19]	9.97 [<0.01]	21.63 [<0.01]	4.29 [0.04]	0.25 [0.62]	0.09 [0.76]
x_t	$VIX_t^{(22)}$	$RV_t^{(22)}$	$\overline{RV}_t^{(22)}$	$EPU_t^{(22)}$	$ADS_t^{(22)}$	$Unc_t^{(22)}$	$Surp_t^{(22)}$
LM	9.51 [<0.01]	0.08 [0.77]	6.53 [0.01]	3.14 [0.08]	3.08 [0.08]	0.62 [0.43]	0.17 [0.68]
$LM_{LT,mod}$	2.38 [0.12]	0.01 [0.91]	1.66 [0.20]	0.16 [0.69]	0.42 [0.52]	1.68 [0.20]	0.70 [0.41]
	Panel B: weekly returns						
x_t	$VIX_t^{(W)}$	$RV_t^{(W)}$	$\overline{RV}_t^{(W)}$	$EPU_t^{(W)}$	$ADS_t^{(W)}$	$Unc_t^{(W)}$	$Surp_t^{(W)}$
LM	28.77 [<0.01]	33.52 [<0.01]	11.76 [<0.01]	16.59 [<0.01]	5.69 [0.02]	0.00 [0.97]	4.63 [0.03]
$LM_{LT,mod}$	14.90 [<0.01]	27.89 [<0.01]	11.49 [<0.01]	6.65 [<0.01]	6.3 [0.01]	0.11 [0.75]	1.93 [0.16]
	Panel C: monthly returns						
x_t	$VIX_t^{(M)}$	$RV_t^{(M)}$	$\overline{RV}_t^{(M)}$	$EPU_t^{(M)}$	$ADS_t^{(M)}$	$Unc_t^{(M)}$	$Surp_t^{(M)}$
LM	6.44 [0.01]	14.77 [<0.01]	7.86 [<0.01]	18.70 [<0.01]	9.71 [<0.01]	1.87 [0.17]	0.20 [0.66]
$LM_{LT,mod}$	2.46 [0.11]	8.17 [<0.01]	4.48 [0.03]	10.56 [<0.01]	1.54 [0.21]	2.44 [0.11]	0.33 [0.56]

Notes: The table reports LM and $LM_{LT,mod}$ test statistics for seven explanatory variables based on $K = 1$. Numbers in brackets are p-values. For all variables but $\overline{RV}_t^{(\cdot)}$ the sample covers the 1991-2016 period. For $\overline{RV}_t^{(\cdot)}$ the sample is based on the 2000-2016 period.

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 $x_t^{(M)}$. The time index t now refers to a weekly or monthly frequency. Panels B and C of Table 4 show that the test results remain qualitatively unchanged. Interestingly, now the test rejects the null for $RV_t^{(W)}$ as well as $RV_t^{(M)}$. An explanation could be as follows: as noted before, $RV_t^{(W)}$ and $RV_t^{(M)}$ are more accurate estimates of the weekly and monthly variance than $RV_t^{(1)}$ is for the daily variance. In addition, the fraction of the total conditional variance that is due to the long-term component and, hence, due to x_t is larger for low-frequency (weekly or monthly) than for high-frequency (daily) returns. This intuition is confirmed when estimating GARCH-MIDAS models for daily or weekly/monthly data (results not reported).

In summary, our test results provide convincing evidence that a simple GARCH(1,1) 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 the continuously compounded 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 (see Alternative 2 in

Section 2.6). Based on the following predictive regression

$$\ln(\sqrt{RV_t}) = c + \pi_1 x_{t-1} + \rho \ln(\sqrt{RV_{t-1}}) + \zeta_t, \tag{47}$$

Conrad and Loch (2015a) find that the π_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 (43) for the same data (again with $K = 1$) and found a significant effect for six out the eleven variables (results not reported). Table 5 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:[15]

$$\ln(\widehat{\widetilde{RV}}_t) = \tilde{c} + \pi_1 x_{t-1} + \rho \ln(\widehat{\widetilde{RV}}_{t-1}) + \tilde{\zeta}_t \tag{48}$$

Although the estimate of the volatility-adjusted realized variance, $\widehat{\widetilde{RV}}_t$, has measurement error, for simplicity we rely on the usual critical values when testing for the significance of π_1. Since the test is then undersized, still finding a significant effect is a strong result.

More specifically, 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 (48) instead of equation (47) suggests that the volatility-adjusted realized variance is indeed the appropriate dependent variable. Although $\ln(\widehat{\widetilde{RV}}_t)$ as well as RV_t suffer from measurement error, the effect of the measurement error appears to be much stronger for RV_t. When reestimating equation (48) by including more lags of the macro variables the picture remains the same.[16] In conclusion, we provide strong evidence that the apparent inability of macro conditions to forecast financial volatility which is document using predictive regressions as in equation (47) seems to be driven by the strong measurement error in RV_t which masks the existing relationship.

[15] We also estimated the same equation with $\sqrt{\widehat{\widetilde{RV}}_t}$ replacing $\widehat{\widetilde{RV}}_t$. The results remain qualitatively unchanged.

[16] For some variables, the results further improve. For example, when including lags of housing starts the third lag is highly significant. This is in line with the finding in Conrad and Loch (2015a) that housing starts is a leading variable and, hence, affects financial volatility with some lag.

Table 5: Predictive Regressions

Variable	$\tilde{c}$	π_1	ρ	adj. R^2
Δ real GDP	6.4007^{***} (0.6885)	-0.0107^{***} (0.0040)	0.1829^{**} (0.0866)	4.81
Δ Ind. prod.	6.4174^{***} (0.6891)	-0.0052^{***} (0.0018)	0.1789^{**} (0.0866)	4.85
Δ Unemp.	6.4461^{***} (0.6644)	0.0799^{**} (0.0349)	0.1735^{**} (0.0834)	4.39
Δ Housing	6.2327^{***} (0.7114)	-0.0004 (0.0004)	0.2014^{**} (0.0894)	3.29
Δ Corp. prof.	6.3006^{***} (0.7037)	-0.0009^{*} (0.0005)	0.1939^{**} (0.0883)	4.14
Δ 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
Δ Cons. sent.	6.2262^{***} (0.7171)	0.0010 (0.0034)	0.2020^{**} (0.0900)	2.89
Δ 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 (48). 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 simple GARCH model against a multiplicative two-component GARCH specification. The test provides a first solution to statistically evaluate if there is a separate long-term time-varying volatility component driven by a macroeconomic explanatory variable, besides the standard short-term GARCH part. We derive the asymptotic properties of our test and study its finite sample performance. 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 for why standard predictive regressions might fail to find a relationship between macro conditions and financial volatility.

References

Amado, C., Teräsvirta, T., 2015. Specification and testing of multiplicative time-varying GARCH models with applications. *Econometric Reviews*, forthcoming.

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 and Economic Statistics* 27, 417-427.

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., 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., Ghysels, E., Sohn, B., 2013. Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics* 95, 776-797.

Engle, R., and Lee, G., 1999. A permanent and transitory component model of stock return volatility. In: Engle, R., White, H. (Ed.), Cointegration, Causality and Forecasting: A Festschrift in Honor of Clive W.J. Granger. Oxford University Press, 475-497.

Escanciano, J. C., 2009. Quasi-maximum likelihood estimation of semi-strong GARCH models. *Econometric Theory* 25, 561-570.

Francq, C., and Thieu, L.Q., 2015. QML inference for volatility models with covariates. MPRA Paper No. 63198.

Francq, C., Zakoïan, J.-M., 2010. *GARCH models: structure, statistical inference and financial applications.* Wiley.

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., 2004. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* 10, 605-637.

Ghysels, E., Sinko, A., Valkanov, R., 2006. 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.

Han, H., 2015. Asymptotic properties of GARCH-X processes. *Journal of Financial Econometrics*, 13, 188-221.

Han, H., Kristensen, D., 2015. Semiparametric multiplicative GARCH-X model: adopting economic variables to explain volatility. Working Paper. Sungkyunkwan University.

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.

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.

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.

Officer, R. R., 1973. The variability of the market factor of the New York Stock Exchange. *Journal of Business* 46, 434-453.

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.

Schwert, W., 1989. Why does stock market volatility change over time? *Journal of Finance* 44, 1115-1153.

Scotti, C., 2016. Surprise and uncertainty indexes: real-time aggregation of real-activity macro surprises. *Journal of Monetary Economics* 82, 1-19.

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. From Francq and Zakïoan (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.

Finiteness of $\mathbf{\Omega}_{\pi\pi}$:

Recall from equation (21) 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 4 that $0 < \kappa_Z - 1 < \infty$. Moreover, $||\mathbf{E}[\mathbf{r}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})']||$ is finite if $\mathbf{E}[||\mathbf{r}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})'||] < \infty$.[17] A typical element of the $K \times 1$ vector $\mathbf{r}_{0,t}^{\infty}$ is given by

$$r_{0,kt}^{\infty} = f_0'(x_{t-k} - \alpha_0 \frac{1}{h_{0,t}^{\infty}} \sum_{j=0}^{\infty} \beta_0^j \varepsilon_{t-1-j}^2 x_{t-1-k-j}). \tag{49}$$

First, f_0' is bounded by Assumption 4 and $\mathbf{E}[|x_{t-k}|^2] < \infty$ by Assumption 5. Second,

$$\begin{aligned}
\left(\mathbf{E}\left|\frac{\sum_{j=0}^{\infty}\alpha_0\beta_0^j\varepsilon_{t-1-j}^2 x_{t-1-k-j}}{h_{0,t}^{\infty}}\right|^2\right)^{1/2} &\leq \left(\mathbf{E}\left|\sum_{j=0}^{\infty}\frac{\alpha_0\beta_0^j\varepsilon_{t-1-j}^2}{(\omega_0+\alpha_0\beta_0^j\varepsilon_{t-1-j}^2)} x_{t-1-k-j}\right|^2\right)^{1/2} && (50)\\
&\leq \sum_{j=0}^{\infty}\left(\mathbf{E}\left|\frac{\alpha_0\beta_0^j\varepsilon_{t-1-j}^2}{(\omega_0+\alpha_0\beta_0^j\varepsilon_{t-1-j}^2)} x_{t-1-k-j}\right|^2\right)^{1/2} && (51)\\
&\leq \sum_{j=0}^{\infty}\left(\mathbf{E}\left|\left(\frac{\alpha_0\beta_0^j}{\omega_0}\varepsilon_{t-1-j}^2\right)^{s/4} x_{t-1-k-j}\right|^2\right)^{1/2} && (52)\\
&\leq \frac{\alpha_0^{s/4}}{\omega_0^{s/4}}\left(\mathbf{E}\left[\varepsilon_{t-1-j}^{2s}\right]\right)^{1/4}\left(\mathbf{E}\left[|x_{t-1-k-j}|^4\right]\right)^{1/4} \\
&\qquad \sum_{j=0}^{\infty}\beta_0^{js/4} < \infty.
\end{aligned}$$

The arguments used above are similar to the ones in Francq and Zakïoan (2004, Eq. (4.19), p.619). In particular, in equation (50) we use that $h_{0,t}^{\infty} \geq \omega_0 + \alpha_0\beta_0^j\varepsilon_{t-1-j}^2$. In equation (51) we use Minkowski's inequality. Next, in equation (52) we use the fact that $w/(1+w) \leq w^s$ for all $w > 0$ and any $s \in (0,1)$. Finally, Assumption 3 implies that there exists some $s > 0$ such that $\mathbf{E}\left[\varepsilon_{t-1-j}^{2s}\right] < \infty$ (see Proposition 1 in Francq and Zakïoan, 2004, p.607). By Assumption 5, $\mathbf{E}\left[|x_{t-1-k-j}|^4\right] < \infty$.

[17]Throughout the paper $||\cdot||$ denotes the euclidean norm.

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 means that $\boldsymbol{\Omega}_{\boldsymbol{\pi}\boldsymbol{\pi}}$ is finite.

Positive definiteness of $\boldsymbol{\Omega}_{\boldsymbol{\pi}\boldsymbol{\pi}}$:

As $\kappa_Z - 1 > 0$, it remains to show that $\mathbf{c}'\mathbf{E}[\mathbf{r}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})']\mathbf{c} > 0$ for any non-zero $\mathbf{c} \in \mathbb{R}^{K\times 1}$. Assume the contrary, i.e., there exists a $\mathbf{c} \neq \mathbf{0}$ such that $\mathbf{c}'\mathbf{E}[\mathbf{r}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})']\mathbf{c} = 0$. This implies $\mathbf{E}[(\mathbf{c}'\mathbf{r}_{0,t}^{\infty})^2] = 0$ and, thus, $\mathbf{c}'\mathbf{r}_{0,t}^{\infty} = 0$ a.s.. Hence, there exists a linear combination of $r_{0,1t}^{\infty}, \ldots, r_{0,Kt}^{\infty}$ which equals zero a.s., i.e.,

$$0 = \sum_{k=1}^{K} c_k \left(x_{t-k} - \frac{\alpha_0}{h_{0,t}^{\infty}} \sum_{j=0}^{\infty} \beta_0^j \varepsilon_{t-1-j}^2 x_{t-1-k-j} \right) \quad \text{a.s.} \tag{53}$$

Using that $0 < \beta_0 < 1$ by Assumption 3 and rearranging, this requires

$$\mathbf{c}'\mathbf{x}_t = \left[\frac{\alpha_0}{h_{0,t}^{\infty}} (1 - \beta_0 L)^{-1} L \right] (\varepsilon_t^2 \mathbf{c}'\mathbf{x}_t) \quad \text{a.s.,} \tag{54}$$

where L denotes the lag operator. Clearly, the operator in square brackets cannot have an eigenvalue of 1. Moreover, Assumption 4 imposes Z_t^2 and, therefore, also ε_t^2 to be non-degenerate. Hence, the only way to fulfill the above equation is by $\mathbf{c}'\mathbf{x}_t = 0$ a.s.. This would imply that we can write $c_K = -\sum_{k=1}^{K-1} c_k / c_K x_{t-k}$ and, hence, τ_{0t} would have a representation which is of the order $K-1$. However, this contradicts Assumption 5. Thus, $\boldsymbol{\Omega}_{\boldsymbol{\pi}\boldsymbol{\pi}}$ must be invertible and hence positive definite.

Next, $\mathbf{E}[\mathbf{d}_t^{\infty}(\boldsymbol{\eta}_0)|\mathcal{F}_{t-1}] = \mathbf{0}$. From Francq and Zakoïan (2004) and Assumptions 3-5 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 3-6, we have that*

$$-\frac{1}{T} \sum_{t=1}^{T} \frac{\partial \mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\tilde{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}'} \xrightarrow{P} \mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}} = -\mathbf{E}\left[\frac{\partial \mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\boldsymbol{\eta}_0)}{\partial \boldsymbol{\eta}'} \right], \tag{55}$$

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

Proof of Proposition 1. We obtain (55) by showing that $\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}(\boldsymbol{\eta}) = -\mathbf{E}\left[\frac{\partial \mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'}\right]$ 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

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

Equation (55) follows from the triangle inequality and the fact that $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}_0 + o_P(1)$.

Using equation (23) we obtain

$$\begin{aligned}\left|\left|\frac{\partial \mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'}\right|\right| &\leq \frac{1}{2}\left(\left|\frac{\varepsilon_t^2}{h_t^{\infty}}\right|\cdot ||\mathbf{r}_t^{\infty}||\cdot||(\mathbf{y}_t^{\infty})'|| + \left|\frac{\varepsilon_t^2}{h_t^{\infty}} - 1\right|\cdot\left|\left|\frac{\partial \mathbf{r}_t^{\infty}}{\partial \boldsymbol{\eta}'}\right|\right|\right)\\ &\leq C|\varepsilon_t^2+\omega|\left(||\mathbf{r}_t^{\infty}||\cdot||(\mathbf{y}_t^{\infty})'|| + \left|\left|\frac{\partial \mathbf{r}_t^{\infty}}{\partial \boldsymbol{\eta}'}\right|\right|\right). \qquad (56)\end{aligned}$$

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

$$\begin{aligned}\left|\frac{1}{h_t^{\infty}}\frac{\partial h_t^{\infty}}{\partial \beta}\right| &\leq \left|\frac{1}{\beta}\sum_{j=0}^{\infty}\frac{j\beta^j(\omega+\alpha\varepsilon_{t-1-j}^2)}{\omega+\beta^j(\omega+\alpha\varepsilon_{t-1-j}^2)}\right|\\ &\leq \frac{1}{\beta\omega^s}\sum_{j=0}^{\infty} j\left|\beta^{js}(\omega+\alpha\varepsilon_{t-1-j}^2)^s\right|, \qquad (57)\end{aligned}$$

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 $||(\mathbf{y}_t^{\infty})'|| \leq C'(1+\sum_{j=0}^{\infty} j\left|\beta^{js}(\omega+\alpha\varepsilon_{t-1-j}^2)^s\right|)$ for some constant $C' > 0$.

Hence, using Cauchy-Schwarz inequality, the first summand in equation (56), i.e. $\mathbf{E}\left[\sup_{\boldsymbol{\eta}}|\varepsilon_t^2+\omega|\cdot||\mathbf{r}_t^{\infty}||\cdot||(\mathbf{y}_t^{\infty})'||\right]$, can be bounded from above by the terms

$$\sqrt{\mathbf{E}[\sup_{\boldsymbol{\eta}}|\varepsilon_t^2+\omega|^2]\mathbf{E}[\sup_{\boldsymbol{\eta}}||\mathbf{r}_t^{\infty}||^2]} \qquad (58)$$

and

$$\begin{aligned}&\sup_{\boldsymbol{\eta}}\sum_{j=0}^{\infty} j\beta^{js}\mathbf{E}[\sup_{\boldsymbol{\eta}}(\omega+\alpha\varepsilon_{t-1-j}^2)^s|\varepsilon_t^2+\omega|\ ||\mathbf{r}_t^{\infty}||] \leq\\ &\sup_{\boldsymbol{\eta}}\sum_{j=0}^{\infty} j\beta^{js}\sqrt{\mathbf{E}[\sup_{\boldsymbol{\eta}}(\omega+\alpha\varepsilon_{t-1-j}^2)^{2s}|\varepsilon_t^2+\omega|^2]\mathbf{E}[\sup_{\boldsymbol{\eta}}||\mathbf{r}_t^{\infty}||^2]}. \qquad (59)\end{aligned}$$

The finiteness of (58) follows from Assumption 6 and similar arguments as in the proof of Theorem 1. The finiteness of (59) follows by applying Hölder's inequality, since for the elements in the sum which involve expectations of the squared observations we have

$$\mathbf{E}[\sup_{\boldsymbol{\eta}}(\omega+\alpha\varepsilon_{t-1-j}^2)^{2s}|\varepsilon_t^2+\omega|^2] \leq$$
$$\left(\mathbf{E}[\sup_{\boldsymbol{\eta}}(\omega+\alpha\varepsilon_{t-1-j}^2)^{2(1+s)}]\right)^{s/(1+s)}\left(\mathbf{E}[\sup_{\boldsymbol{\eta}}|\varepsilon_t^2+\omega|^{2(1+s)}]\right)^{1/(1+s)} \tag{60}$$

and Assumption 6 applies again.

Using the Cauchy-Schwarz-Inequality for the two factors in the second term in (56), we are left with the need to show that $\mathbf{E}\left[\sup_{\boldsymbol{\eta}}||\frac{\partial \mathbf{r}_t^\infty}{\partial \boldsymbol{\eta}'}||^2\right]$ is finite. This follows from

$$\begin{aligned}
(f_0')^{-1}\frac{\partial \mathbf{r}_t^\infty}{\partial \boldsymbol{\eta}'} &= \frac{\partial}{\partial \boldsymbol{\eta}'}\mathbf{x}_t - \frac{\partial}{\partial \boldsymbol{\eta}'}\left(\frac{1}{h_t^\infty}\sum_{j=0}^{\infty}\alpha\beta^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j}\right)\\
&= \frac{\partial}{\partial \boldsymbol{\eta}'}\mathbf{x}_t - \frac{1}{h_t^\infty}\left(\sum_{j=0}^{\infty}\alpha\beta^j\varepsilon_{t-1-j}^2\frac{\partial}{\partial \boldsymbol{\eta}'}\mathbf{x}_{t-1-j}\right)\\
&\quad + \left(\frac{1}{h_t^\infty}\sum_{j=0}^{\infty}\alpha\beta^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j}\right)(\mathbf{y}_t^\infty)'\\
&\quad - \frac{1}{h_t^\infty}\sum_{j=0}^{\infty}\mathbf{x}_{t-1-j}\left(\frac{\partial}{\partial \boldsymbol{\eta}'}\alpha\beta^j\varepsilon_{t-1-j}^2\right)
\end{aligned} \tag{61}$$

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 5. *There also exists a bound for* $\mathbf{E}\left[\sup_{\boldsymbol{\eta}}||\frac{\partial \mathbf{r}_t^\infty}{\partial \boldsymbol{\eta}'}||^2\right]$ *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* (61) *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:

$$\frac{1}{h_t^\infty}\begin{pmatrix} 0 & \sum_{j=0}^\infty \beta^j\varepsilon_{t-1-j}^2 x_{t-2-j} & \alpha\sum_{j=0}^\infty j\beta^{j-1}\varepsilon_{t-1-j}^2 x_{t-2-j}\\ 0 & \sum_{j=0}^\infty \beta^j\varepsilon_{t-1-j}^2 x_{t-3-j} & \alpha\sum_{j=0}^\infty j\beta^{j-1}\varepsilon_{t-1-j}^2 x_{t-3-j}\\ \vdots & & \\ 0 & \sum_{j=0}^\infty \beta^j\varepsilon_{t-1-j}^2 x_{t-1-K-j} & \alpha\sum_{j=0}^\infty j\beta^{j-1}\varepsilon_{t-1-j}^2 x_{t-1-K-j}\end{pmatrix} \tag{62}$$

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 $\boldsymbol{\eta}_0$

$$\mathbf{0}=\sqrt{T}\mathbf{D}_{\boldsymbol{\eta}}^\infty(\hat{\boldsymbol{\eta}})=\sqrt{T}\mathbf{D}_{\boldsymbol{\eta}}^\infty(\boldsymbol{\eta}_0)+\frac{1}{T}\sum_{t=1}^{T}\frac{\partial\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\tilde{\boldsymbol{\eta}})}{\partial\boldsymbol{\eta}'}\sqrt{T}(\hat{\boldsymbol{\eta}}-\boldsymbol{\eta}_0) \tag{63}$$

with $\tilde{\boldsymbol{\eta}}=\boldsymbol{\eta}_0+o_P(1)$. Under Assumptions 3 and 4, Francq and Zakoïan (2004) have shown that

$$-\frac{1}{T}\sum_{t=1}^{T}\frac{\partial\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\tilde{\boldsymbol{\eta}})}{\partial\boldsymbol{\eta}'}\xrightarrow{P}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}=-\mathbf{E}\left[\frac{\partial\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0)}{\partial\boldsymbol{\eta}'}\right] \tag{64}$$

and, hence, equation (63) can be written as

$$\sqrt{T}(\hat{\boldsymbol{\eta}}-\boldsymbol{\eta}_0)=\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\sqrt{T}\mathbf{D}_{\boldsymbol{\eta}}^\infty(\boldsymbol{\eta}_0)+o_P(1). \tag{65}$$

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

$$\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^\infty(\hat{\boldsymbol{\eta}})=\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^\infty(\boldsymbol{\eta}_0)+\frac{1}{T}\sum_{t=1}^{T}\frac{\partial\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\tilde{\boldsymbol{\eta}})}{\partial\boldsymbol{\eta}'}\sqrt{T}(\hat{\boldsymbol{\eta}}-\boldsymbol{\eta}_0). \tag{66}$$

Combining equation (65) and Proposition 1 leads to

$$\begin{aligned}\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^\infty(\hat{\boldsymbol{\eta}}) &= \sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^\infty(\boldsymbol{\eta}_0)-\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\sqrt{T}\mathbf{D}_{\boldsymbol{\eta}}^\infty(\boldsymbol{\eta}_0)+o_P(1) &(67)\\ &= [-\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\ :\ \mathbf{I}]\sqrt{T}\begin{pmatrix}\mathbf{D}_{\boldsymbol{\eta}}^\infty(\boldsymbol{\eta}_0)\\ \mathbf{D}_{\boldsymbol{\pi}}^\infty(\boldsymbol{\eta}_0)\end{pmatrix}+o_P(1) &(68)\\ &= [-\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\ :\ \mathbf{I}]\sqrt{T}\mathbf{D}^\infty(\boldsymbol{\eta}_0)+o_P(1). &(69)\end{aligned}$$

Applying Theorem 1 gives the asymptotic distribution

$$\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^\infty(\hat{\boldsymbol{\eta}})\xrightarrow{d}\mathcal{N}(\mathbf{0},[\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\ :\ \mathbf{I}]\boldsymbol{\Omega}[\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\ :\ \mathbf{I}]') \tag{70}$$

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

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

Finally, using equations (21), (24) and (25) the expression for $\mathbf{\Sigma}$ simplifies to:

$$\mathbf{\Sigma} = \frac{1}{4}(\kappa_Z - 1)\left(\mathbf{E}[\mathbf{r}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})'] - \mathbf{E}[\mathbf{r}_{0,t}^{\infty}(\mathbf{y}_{0,t}^{\infty})']\left(\mathbf{E}[\mathbf{y}_{0,t}^{\infty}(\mathbf{y}_{0,t}^{\infty})']\right)^{-1}\mathbf{E}[\mathbf{y}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})']\right). \tag{71}$$

∎

Proof of Theorem 3. We show that

$$\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}(\hat{\boldsymbol{\eta}}) = \sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\hat{\boldsymbol{\eta}}) + o_P(1). \tag{72}$$

Hence, the observed quantity $\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}(\hat{\boldsymbol{\eta}})$ will have the same asymptotic distribution as the unobserved $\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\hat{\boldsymbol{\eta}})$. The asymptotic distribution of the test statistic then follows directly from Theorem 2. Standardization with the consistent estimator $\widehat{\mathbf{\Sigma}}$ instead of the theoretical $\mathbf{\Sigma}$, has no effect on the final χ^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

$$\sup_{\boldsymbol{\eta}}||\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\boldsymbol{\eta}) - \sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}(\boldsymbol{\eta})|| \leq \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\sup_{\boldsymbol{\eta}}||\mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\boldsymbol{\eta}) - \mathbf{d}_{\boldsymbol{\pi},t}(\boldsymbol{\eta})||, \tag{73}$$

we establish equation (72) by showing that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\sup_{\boldsymbol{\eta}}||\mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\boldsymbol{\eta}) - \mathbf{d}_{\boldsymbol{\pi},t}(\boldsymbol{\eta})|| = o_P(1). \tag{74}$$

Consider the following decomposition:

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

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

$$(f_0')^{-1}(\mathbf{r}_t^\infty - \mathbf{r}_t) = -\alpha \frac{1}{h_t^\infty} \sum_{j=t}^{\infty} \beta^j \varepsilon_{t-1-j}^2 \mathbf{x}_{t-1-j}. \tag{75}$$

Next, consider a typical element:

$$\begin{aligned}
(f_0')^{-1}\left(\mathbf{E}\sup_{\boldsymbol{\eta}}|r_{k,t}^\infty - r_{k,t}|^2\right)^{1/2} &= \left(\mathbf{E}\sup_{\boldsymbol{\eta}}\left|\alpha\frac{1}{h_t^\infty}\sum_{j=t}^{\infty}\beta^j\varepsilon_{t-1-j}^2 x_{t-1-k-j}\right|^2\right)^{1/2} \\
&\le \sum_{j=t}^{\infty}\left(\mathbf{E}\sup_{\boldsymbol{\eta}}\left|\frac{\alpha\beta^j\varepsilon_{t-1-j}^2}{\omega+\alpha\beta^j\varepsilon_{t-1-k-j}^2}x_{t-1-k-j}\right|^2\right)^{1/2} \\
&\le \sum_{j=t}^{\infty}\left(\mathbf{E}\sup_{\boldsymbol{\eta}}\left|\left(\frac{\alpha\beta^j}{\omega}\varepsilon_{t-1-j}^2\right)^{s/4} x_{t-1-k-j}\right|^2\right)^{1/2} \\
&\le \left(\mathbf{E}[|\varepsilon_{t-1-j}|^{2s}]\right)^{1/4}\left(\mathbf{E}[|x_{t-1-k-j}|^4]\right)^{1/4} \\
&\qquad \sup_{\boldsymbol{\eta}}\left(\frac{\alpha}{\omega}\right)^{s/4}\sum_{j=t}^{\infty}\beta^{js/4} \\
&= \left(\mathbf{E}[|\varepsilon_{t-1-j}|^{2s}]\right)^{1/4}\left(\mathbf{E}[|x_{t-1-k-j}|^4]\right)^{1/4} \\
&\qquad \sup_{\boldsymbol{\eta}}\left(\frac{\alpha}{\omega}\right)^{s/4}\frac{(\beta^{s/4})^t}{1-\beta^{s/4}}
\end{aligned} \tag{76}$$

which shows that $\mathbf{E}\sup_{\boldsymbol{\eta}}||\mathbf{r}_{k,t}^\infty - \mathbf{r}_{k,t}||^2 = O(\beta^{ts/2})$.

Hence,

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

by Assumption 3 and equation (76). 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

$$\left|\frac{h_t^\infty - h_t}{h_t^\infty}\right| \le \frac{\alpha^s}{\omega^s}\sum_{j=t}^{\infty}(\beta^s)^j\varepsilon_{t-j}^{2s}, \tag{77}$$

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

$$\begin{aligned}
\mathbf{E}\text{sup}_{\boldsymbol{\eta}}\varepsilon_t^2||\mathbf{r}_t^\infty|| \left|\frac{h_t^\infty - h_t}{h_t^\infty}\right| &\leq \mathbf{E}\text{sup}_{\boldsymbol{\eta}}||\varepsilon_t^2\mathbf{r}_t^\infty\varepsilon_{t-j}^{2s}|| \ \text{sup}_{\boldsymbol{\eta}}\frac{\alpha^s}{\omega^s}\sum_{j=t}^{\infty}(\beta^s)^j \\
&\leq \sqrt{\mathbf{E}\text{sup}_{\boldsymbol{\eta}}||\mathbf{r}_t^\infty||^2\mathbf{E}|\varepsilon_t^4\varepsilon_{t-j}^{4s}|} \ \text{sup}_{\boldsymbol{\eta}}\frac{\alpha^s}{\omega^s}(\beta^s)^t\sum_{j=0}^{\infty}(\beta^s)^j \\
&= \sqrt{\mathbf{E}\text{sup}_{\boldsymbol{\eta}}||\mathbf{r}_t^\infty||^2\mathbf{E}|\varepsilon_t^4\varepsilon_{t-j}^{4s}|} \ \text{sup}_{\boldsymbol{\eta}}\frac{\alpha^s}{\omega^s(1-\beta^s)}(\beta^s)^t \\
&= O((\beta^s)^t). \qquad (78)
\end{aligned}$$

The last line follows because it can be shown by similar arguments as in the proof of Theorem 1 that $\mathbf{E}\text{sup}_{\boldsymbol{\eta}}||\mathbf{r}_t^\infty||^2 < \infty$ and because Hölder's inequality and Assumption 6 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 (78) implies that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{E}\text{sup}_{\boldsymbol{\eta}}\varepsilon_t^2||\mathbf{r}_t^\infty|| \left|\frac{h_t^\infty - h_t}{h_t^\infty}\right| = o(1), \qquad (79)$$

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

The third term can be treated as follows:

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

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 $\sum_{t=1}^T \mathbf{E}\text{sup}_{\boldsymbol{\eta}}\varepsilon_t^2||\mathbf{r}_t^\infty - \mathbf{r}_t|| = O(1)$ and $\mathbf{E}\text{sup}_{\boldsymbol{\eta}}\left|\frac{h_t^\infty - h_t}{h_t^\infty}\right| = O(\beta^{ts})$.

∎

B Simulation:

B.1 Size-adjusted power for 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 t distributed with 7 degrees of freedom.

Table 6: Empirical size-adjusted power for exponential long-term component and t distributed innovations.

x_t		VIX_t						$VIX_t^{(22)}$	$VIX_t^{(65)}$
		$\omega_{01} = 1$, $\omega_{02} = 10$							
		$\alpha_0 = 0.09$			$\alpha_0 = 0.07$			$\alpha_0 = 0.09$	
weighting scheme		I	F	S	I	F	S	I	I
	1%	24.3	16.5	9.4	34.6	32.4	16.5	5.3	3.5
LM	5%	54.9	44.1	32.1	59.3	57.1	39.3	20.1	13.9
	10%	66.8	58.1	45.7	71.3	68.7	53.1	36.0	22.6
	1%	15.6	12.6	10.1	29.9	27.7	21.6	5.2	3.8
$LM_{LT,mod}$	5%	30.5	25.8	22.7	50.7	48.7	39.4	12.2	9.5
	10%	42.0	36.3	32.0	60.2	58.9	51.2	19.2	15.8
	1%	0.9	1.0	1.1	0.8	1.0	1.1	1.1	1.0
LM_{LT}	5%	5.8	5.7	5.7	5.0	5.2	5.1	5.6	5.3
	10%	9.7	9.8	9.9	9.5	9.7	10.0	9.8	9.8
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. Otherwise see Table 2.

B.2 Size-adjusted power for linear long-term component.

Table 7: Empirical size-adjusted power for linear long-term component and normally distributed innovations.

x_t		VIX_t						$VIX_t^{(22)}$	$VIX_t^{(65)}$
		$\omega_{01} = 1$, $\omega_{02} = 10$							
		$\alpha_0 = 0.09$			$\alpha_0 = 0.07$			$\alpha_0 = 0.09$	
weighting scheme		I	F	S	I	F	S	I	I
LM	1%	34.8	33.1	21.1	44.7	42.7	29.9	10.1	8.4
	5%	57.2	54.8	39.2	66.5	64.6	51.5	28.9	18.0
	10%	66.4	64.8	50.7	75.6	74.2	63.0	42.2	30.4
$LM_{LT,mod}$	1%	16.0	15.9	13.0	32.6	32.0	27.1	4.8	3.0
	5%	34.8	34.1	30.3	59.2	58.1	51.1	14.3	107
	10%	44.2	43.3	38.6	71.1	70.0	65.6	18.9	17.2
LM_{LT}	1%	0.9	0.9	0.9	1.0	1.0	0.9	0.9	0.8
	5%	5.9	5.9	5.4	5.6	5.6	5.3	4.8	4.6
	10%	10.3	10.5	10.3	10.5	10.7	10.8	9.5	9.4
VR		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 $\tau_{0,t} = 1 + \sum_{k=1}^{K} \pi_{0k} x_{t-k}$. Otherwise see Table 2.

Table 8: Empirical size-adjusted power for linear long-term component and t distributed innovations.

x_t		VIX_t						$VIX_t^{(22)}$	$VIX_t^{(65)}$
		$\omega_{01} = 1$, $\omega_{02} = 10$							
		$\alpha_0 = 0.09$			$\alpha_0 = 0.07$			$\alpha_0 = 0.09$	
weighting scheme		I	F	S	I	F	S	I	I
	1%	24.3	20.0	14.6	30.0	28.6	18.7	4.4	3.2
LM	5%	39.7	34.5	28.4	48.3	46.7	37.9	17.4	12.3
	10%	52.7	46.5	39.2	60.8	59.3	47.7	28.3	20.5
	1%	10.3	9.2	8.3	20.1	19.6	16.4	4.5	3.5
$LM_{LT,mod}$	5%	27.4	25.9	24.6	43.0	42.8	38.5	11.1	9.1
	10%	37.3	35.0	32.9	53.8	52.9	49.4	16.6	15.0
	1%	1.0	1.1	1.1	1.0	1.0	1.2	1.0	1.0
LM_{LT}	5%	5.5	5.5	5.6	5.3	5.3	5.3	5.7	5.4
	10%	9.9	9.8	9.7	9.9	10.0	10.0	9.6	9.7
VR		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 $\tau_{0,t} = 1 + \sum_{k=1}^{K} \pi_{0k} x_{t-k}$. Otherwise see Table 2.