University of Heidelberg

Department of Economics

Discussion Paper Series | No. 597

Misspecification Testing in GARCH-MIDAS Models

Christian Conrad and Melanie Schienle

July 2015

Misspecification Testing in GARCH-MIDAS Models*

Christian Conrad[†] and Melanie Schienle[‡]

[†]Heidelberg University, Germany

[‡]Karlsruhe Institute of Technology, Germany

July 6, 2015

Abstract

We develop a misspecification test for the multiplicative two-component GARCH-MIDAS model suggested in Engle et al. (2013). In the GARCH-MIDAS model a short-term unit variance GARCH component 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. Hence, under the null hypothesis the long-term component is constant and the GARCH-MIDAS reduces to the simple GARCH model. We derive the asymptotic theory for our test statistic and investigate its finite sample properties by Monte-Carlo simulation. The usefulness of our procedure is illustrated by an empirical application to S&P 500 return data.

Keywords: Volatility Component Models, LM test, Long-term Volatility.
JEL Classification: C53, C58, E32, G12

*We would like to thank Richard Baillie, Tilmann Gneiting, Onno Kleen, Karin Loch, Enno Mammen, Rasmus S. Pedersen, Robert Taylor and Timo Teräsvirta 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) and Conrad and Loch (2014) provide recent evidence for the counter-cyclical behavior of financial volatility.[1] In particular, Conrad and Loch (2014) 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. These findings are clearly relevant from a risk management perspective. Prior to the last financial crisis, risk management exclusively focused on short-run risks – such as Value at Risk at the one or 10-day horizon – and, hence, failed to address the *risk that risk will change* (Engle, 2009). Determining these long-term risks, however, requires statistical models which allow for effects of changes in relevant economic variables on the conditional variance of asset returns.

For this reason, an increasing amount of empirical studies has employed the GARCH-MIDAS framework introduced by Engle et al. (2013) (see, e.g., Asgharian et al., 2013, Conrad and Loch, 2014, 2015, Dorion, 2013, Opschoor et al., 2014). In a GARCH-MIDAS specification, the conditional variance consists of two multiplicative components, whereby economic conditions enter through the smooth long-term component around which a short-term unit variance GARCH component fluctuates. Besides predictive evidence, however, it is still an open question whether and which macroeconomic and financial variables are significant drivers of volatility. This is the case because standard procedures for misspecification testing in GARCH models do not cover the case of (exogenous) explanatory variables. As most of them also require additive separability of the additional component under the alternative, their adaption to a general GARCH-MIDAS structure is not straightforward.[2]

We develop a misspecification test for the multiplicative two-component GARCH-MIDAS model. In particular, we propose a Lagrange Multiplier (LM) statistic for testing the null hypothesis that the long-term component is constant. Thus, under our null hypothesis, the GARCH-MIDAS model reduces to a simple GARCH. Note that Wald-type tests like simple t- or F-tests are not straightforward to employ in this context, as

[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) and Han (2015).

there exists no asymptotic theory yet for the general case of macroeconomic effects in the GARCH-MIDAS model. The most recent theoretical results by Wang and Ghysels (2015) are specific to long-term components that are driven by realized volatility and only hold in a restrictive parameter space which does not admit our null hypothesis. 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 follows similar lines as the arguments in the proof of Theorem 2 in Halunga and Orme (2009), who consider general misspecification tests for GARCH models. However, Halunga and Orme (2009) focus on estimation effects from the correct specification of the conditional mean and consider additive components only. In our set-up, the volatilty 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. In a Monte-Carlo simulation, we find good size and power properties in finite samples. Moreover, we illustrate the usefulness of our procedure by an empirical application to S&P 500 return data.

Our test statistic is also closely related to the 'ARCH nested in GARCH' test for evaluating GARCH models as proposed by Lundbergh and Teräsvirta (2002). While it is possible to consider their 'nested ARCH component' as our long-term component with a specific choice for the explanatory variable, 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. The loss in power is the stronger the larger the ARCH parameter is and the more the long-term component fluctuates.

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 GARCH-MIDAS 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 the Appendix.

2 Model and Test Statistic

In Section 2.1, we first introduce the GARCH-MIDAS specification of Engle et al. (2013) and then discuss the null hypothesis of our test. We derive the likelihood function and the test indicator in Section 2.2 and present our main result on the asymptotic distribution of the test statistic in Section 2.3. Finally, Section 2.4 provides a comparison with the 'ARCH nested in GARCH' test.

2.1 The GARCH-MIDAS 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. 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 short-term and a long-term component:

$$\sigma_{0t}^2 = \tilde{h}_{0t}^{\infty} \tilde{\tau}_{0t} \tag{2}$$

The terminology of decomposing σ_{0t}^2 into a short- and a long-term component follows Engle et al. (2013). In our setting, the long-term component is the one that is driven by (exogeneous) explanatory variables and, typically, is smoother than the short-term component.

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

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

On the other hand, the long-term (MIDAS) component depends on the K lagged values of a (nonnegative) explanatory variable x_t:

$$\tilde{\tau}_{0t} = \sigma_0^2 + \tilde{\pi}_0 \sum_{k=1}^{K} \psi_{0k} x_{t-k} \tag{4}$$

with MIDAS weights $\psi_{0k} \geq 0$ summing to one. A common choice for determining the ψ_{0k} is the Beta weighting scheme. In this case, $\psi_{0k} = \psi_{0k}(\omega_{01}, \omega_{02})$, whereby the parameters ω_{01} and ω_{02} determine the Beta weights. The sign of the effect of x_t on long-term volatility can be inferred from the parameter $\tilde{\pi}_0$. Note that in equation (4) we consider a specification in which the explanatory variable and the returns are observed at the same frequency. Alternatively, one might also assume that the explanatory variable is observed at a lower frequency than the returns (see, e.g., Conrad and Loch, 2014). Nevertheless, our long-term component can be considered as a MIDAS specification in the sense that it parsimoniously models the dependence of $\tilde{\tau}_{0t}$ on (possibly) many lags of x_t in terms of only two parameters ω_{01} and ω_{02} via the flexible weighting scheme $\psi_{0k}(\cdot)$.

Following Conrad and Loch (2014), we denote the model with (exogenous) explanatory variables as GARCH-MIDAS-X. Engle et al. (2013) and Wang and Ghysels (2015) consider a specification with the realized volatility, $RV_t^{(N)} = \sum_{j=0}^{N-1} \varepsilon_{t-j}^2$, of the last N days as the explanatory variable. We refer to this model as GARCH-MIDAS-RV. The long-term component can then be rewritten as

$$\tilde{\tau}_{0t} = \sigma_0^2 + \tilde{\pi}_0 \sum_{l=1}^{N+K-1} c_{0l} \varepsilon_{t-l}^2,$$

where the c_{0l}'s are combinations of the ψ_{0k}'s and $\sum_{l=1}^{N+K-1} c_{0l} = N \sum_{k=1}^{K} \psi_{0k} = N$ (see Wang and Ghysels, 2015).[3] In other words, in this alternative representation the squared returns can be considered as being the explanatory variables.

For the specific case of a GARCH-MIDAS-RV model, Wang and Ghysels (2015) 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 rests 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 us to test the null that the lagged $RV_t^{(N)}$ are jointly insignificant (see Assumption 4.3 in Wang and Ghysels, 2015). In addition, there is no

[3]Let $\nu_t = \sqrt{\tilde{h}_{0t}^{\infty}} Z_t$. Then, $\varepsilon_t = \sqrt{\tilde{\tau}_{0t}} \nu_t$ can be interpreted as a semi-strong ARCH process with multiplicative GARCH error (see Wang and Ghysels, 2015).

asymptotic theory for the general GARCH-MIDAS-X model yet. We circumvent these problems by deriving an LM test for the hypothesis that x_t does not affect the long-term component. The LM test requires estimation of the model under the null only. To derive our test statistic, we re-parameterize equation (2) as follows:

$$\sigma_{0t}^2 = \tilde{h}_{0t}^{\infty}\tilde{\tau}_{0t} = (\sigma_0^2\tilde{h}_{0t}^{\infty})\left(\frac{\tilde{\tau}_{0t}}{\sigma_0^2}\right) = \bar{h}_{0t}^{\infty}\tau_{0t}$$

The short-term component can then be expressed as

$$\bar{h}_{0t}^{\infty} = \omega_0 + \alpha_0\frac{\varepsilon_{t-1}^2}{\tau_{0,t-1}} + \beta_0\bar{h}_{0,t-1}^{\infty} \tag{5}$$

with $\omega_0 = \sigma_0^2(1-\alpha_0-\beta_0)$ (provided that $\alpha_0+\beta_0<1$). We denote the vector of true parameters in the short-term component as $\boldsymbol{\eta}_0 = (\omega_0, \alpha_0, \beta_0)'$. Similarly, the long-term component can be rewritten as

$$\tau_{0t} = 1 + \sum_{k=1}^{K}\pi_{0k}x_{t-k} = 1 + \boldsymbol{\pi}_0'\mathbf{x}_t \tag{6}$$

with $\pi_{0k} = \sigma_0^{-2}\tilde{\pi}_0\psi_{0k}$, $\boldsymbol{\pi}_0 = (\pi_{01},\ldots,\pi_{0K})'$ and $\mathbf{x}_t = (x_{t-1},\ldots,x_{t-K})'$.

Using this notation, we are interested in testing $H_0 : \boldsymbol{\pi}_0 = \mathbf{0}$ against $H_1 : \boldsymbol{\pi}_0 \neq \mathbf{0}$. Under H_0 the long-term component is equal one and the GARCH-MIDAS-X reduces to the nested GARCH(1,1) with unconditional variance $\sigma_0^2 = \omega_0/(1-\alpha_0-\beta_0)$. Note that equation (5) is specified such that we can write

$$\bar{h}_{0t}^{\infty} = \omega_0 + (\alpha_0 Z_{t-1}^2 + \beta_0)\bar{h}_{0,t-1}^{\infty}$$

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 data generating process under H_0.

Assumption 1. $\boldsymbol{\eta}_0 \in \Theta$ *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 2. *As defined in equation (1), let* Z_t *be i.i.d. with* $\mathbf{E}[Z_t] = 0$ *and* $\mathbf{E}[Z_t^2] = 1$. *Further,* Z_t^2 *has a nondegenerate distribution and* $\kappa_Z = \mathbf{E}[Z_t^4] < \infty$.

Assumptions 1 and 2 imply that $\sqrt{\bar{h}_{0t}^{\infty}}Z_t$ is a covariance-stationary process with unconditional variance σ_0^2. Further, 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., 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 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

$$L_T^\infty(\boldsymbol{\eta}, \boldsymbol{\pi}|\varepsilon_T, x_T, \varepsilon_{T-1}, x_{T-1}, \ldots) = \sum_{t=1}^{T} l_t^\infty \tag{7}$$

with

$$l_t^\infty = \ln(f(\varepsilon_t|\boldsymbol{\eta}, \boldsymbol{\pi})) = -\frac{1}{2}\left[\ln(\bar{h}_t^\infty) + \ln(\tau_t) + \frac{\varepsilon_t^2}{\bar{h}_t^\infty \tau_t}\right]. \tag{8}$$

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

$$L_T(\boldsymbol{\eta}, \boldsymbol{\pi}|\varepsilon_T, x_T, \varepsilon_{T-1}, x_{T-1}, \ldots, \varepsilon_1, x_1) = \sum_{t=1}^{T} l_t \tag{9}$$

with

$$l_t = \ln(f(\varepsilon_t|\boldsymbol{\eta}, \boldsymbol{\pi})) = -\frac{1}{2}\left[\ln(\bar{h}_t) + \ln(\tau_t) + \frac{\varepsilon_t^2}{\bar{h}_t \tau_t}\right]. \tag{10}$$

2.2.1 First derivatives

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

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

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 $\boldsymbol{\eta}$:

$$\frac{\partial l_t^\infty}{\partial \boldsymbol{\eta}} = \frac{1}{2}\left[\frac{\varepsilon_t^2}{\bar{h}_t^\infty \tau_t} - 1\right]\left(\frac{1}{\bar{h}_t^\infty}\frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\eta}} + \frac{1}{\tau_t}\frac{\partial \tau_t}{\partial \boldsymbol{\eta}}\right) \tag{11}$$

with $\partial\tau_t/\partial\boldsymbol{\eta} = (\partial\mathbf{x}_t/\partial\boldsymbol{\eta})'\boldsymbol{\pi}$. 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

$$\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}) = \left.\frac{\partial l_t^\infty}{\partial \boldsymbol{\eta}}\right|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{2}\left[\frac{\varepsilon_t^2}{h_t^\infty} - 1\right]\mathbf{y}_t^\infty \tag{12}$$

with

$$\mathbf{y}_t^\infty = \frac{1}{\bar{h}_t^\infty}\left.\frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\eta}}\right|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{h_t^\infty}\sum_{i=0}^{\infty}\beta^i \mathbf{s}_{t-i}^\infty, \tag{13}$$

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

$$\mathbf{d}_{\boldsymbol{\eta},t}^\infty(\boldsymbol{\eta}_0) = \frac{1}{2}\left[\frac{\varepsilon_t^2}{h_{0,t}^\infty} - 1\right]\mathbf{y}_{0,t}^\infty, \tag{14}$$

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:

$$\frac{\partial l_t^\infty}{\partial \boldsymbol{\pi}} = \frac{1}{2}\left[\frac{\varepsilon_t^2}{\bar{h}_t^\infty \tau_t} - 1\right]\left(\frac{1}{\bar{h}_t^\infty}\frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\pi}} + \frac{1}{\tau_t}\frac{\partial \tau_t}{\partial \boldsymbol{\pi}}\right), \tag{15}$$

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

$$\frac{\partial \bar{h}_t^\infty}{\partial \boldsymbol{\pi}} = -\alpha\sum_{j=0}^{\infty}\beta^j \frac{\varepsilon_{t-1-j}^2}{\tau_{t-1-j}^2}\frac{\partial \tau_{t-1-j}}{\partial \boldsymbol{\pi}}. \tag{16}$$

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

$$\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}) = \left.\frac{\partial l_t^\infty}{\partial \boldsymbol{\pi}}\right|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{2}\left[\frac{\varepsilon_t^2}{h_t^\infty} - 1\right]\mathbf{r}_t^\infty \tag{17}$$

with

$$\mathbf{r}_t^\infty = \mathbf{x}_t - \alpha\frac{1}{h_t^\infty}\sum_{j=0}^{\infty}\beta^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j}. \tag{18}$$

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

$$\mathbf{d}_{\boldsymbol{\pi},t}^\infty(\boldsymbol{\eta}_0) = \frac{1}{2}\left[\frac{\varepsilon_t^2}{h_{0,t}^\infty} - 1\right]\mathbf{r}_{0,t}^\infty \tag{19}$$

with

$$\mathbf{r}_{0,t}^\infty = \mathbf{x}_t - \alpha_0\frac{1}{h_{0,t}^\infty}\sum_{j=0}^{\infty}\beta_0^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j}. \tag{20}$$

In summary, we have

$$\mathbf{D}^\infty(\boldsymbol{\eta}_0) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{d}_t^\infty(\boldsymbol{\eta}_0) = \frac{1}{2T}\sum_{t=1}^{T}\left[\frac{\varepsilon_t^2}{h_{0,t}^\infty} - 1\right]\begin{pmatrix}\mathbf{y}_{0,t}^\infty \\ \mathbf{r}_{0,t}^\infty\end{pmatrix}. \tag{21}$$

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

$$\begin{aligned} \mathbf{Var}[\mathbf{d}_t^\infty(\boldsymbol{\eta}_0)] = \boldsymbol{\Omega} &= \begin{pmatrix} \boldsymbol{\Omega}_{\eta\eta} & \boldsymbol{\Omega}_{\eta\pi} \\ \boldsymbol{\Omega}_{\pi\eta} & \boldsymbol{\Omega}_{\pi\pi} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{E}[\mathbf{d}_{\eta,t}^\infty(\boldsymbol{\eta}_0)\mathbf{d}_{\eta,t}^\infty(\boldsymbol{\eta}_0)'] & \mathbf{E}[\mathbf{d}_{\eta,t}^\infty(\boldsymbol{\eta}_0)\mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta}_0)'] \\ \mathbf{E}[\mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta}_0)\mathbf{d}_{\eta,t}^\infty(\boldsymbol{\eta}_0)'] & \mathbf{E}[\mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta}_0)\mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta}_0)'] \end{pmatrix} \\ &= \frac{1}{4}(\kappa_Z - 1)\begin{pmatrix} \mathbf{E}[\mathbf{y}_{0,t}^\infty(\mathbf{y}_{0,t}^\infty)'] & \mathbf{E}[\mathbf{y}_{0,t}^\infty(\mathbf{r}_{0,t}^\infty)'] \\ \mathbf{E}[\mathbf{r}_{0,t}^\infty(\mathbf{y}_{0,t}^\infty)'] & \mathbf{E}[\mathbf{r}_{0,t}^\infty(\mathbf{r}_{0,t}^\infty)'] \end{pmatrix}. \end{aligned} \quad (22)$$

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

2.2.2 Second derivatives

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

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

and

$$\frac{\partial \mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'} = -\frac{1}{2}\frac{\varepsilon_t^2}{h_t^\infty}\mathbf{r}_t^\infty(\mathbf{y}_t^\infty)' + \frac{1}{2}\left[\frac{\varepsilon_t^2}{h_t^\infty} - 1\right]\frac{\partial \mathbf{r}_t^\infty}{\partial \boldsymbol{\eta}'}. \quad (24)$$

We then define

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

and

$$\mathbf{J}_{\pi\eta} = -\mathbf{E}\left[\frac{\partial \mathbf{d}_{\pi,t}^\infty(\boldsymbol{\eta}_0)}{\partial \boldsymbol{\eta}'}\right] = \frac{1}{2}\mathbf{E}[\mathbf{r}_{0,t}^\infty(\mathbf{y}_{0,t}^\infty)']. \quad (26)$$

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

2.3 The *LM* Test Statistic

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

$$\sqrt{T}\mathbf{D}^{\infty}_{\boldsymbol{\pi}}(\hat{\boldsymbol{\eta}}) = [\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}^{-1}_{\boldsymbol{\eta}\boldsymbol{\eta}} \ : \ \mathbf{I}]\sqrt{T}\mathbf{D}^{\infty}(\boldsymbol{\eta}_0) + o_P(1) \tag{27}$$

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 $\sqrt{T}\mathbf{D}^{\infty}_{\boldsymbol{\pi}}(\hat{\boldsymbol{\eta}})$. The LM statistic follows the usual χ^2 distribution.

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

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

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 assume that the long-term component is minimal in the sense that no equivalent representation which is of lower order exists.

Assumption 3. *$x_t \geq 0$ 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* (6) *such that $\sum_{k=1}^{K} \pi_{0k} x_{t-k} = \sum_{s=1}^{S} a_s x_{t-s}$ with $S < K$.*

For simplicity, we also assume that the explanatory variable takes nonnegative values only. This assumption is in line with the GARCH-MIDAS-RV model or the specification of Lundbergh and Teräsvirta (2002) with $x_t = \varepsilon_t^2/h_{0t}$ (see Section 2.4). For testing the GARCH-MIDAS-RV against the simple GARCH model, Assumption 3 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). Further, Conrad and Loch (2014) have shown that nonnegative explanatory variables such as the unemployment rate or disagreement among forecasters are important predictors of financial volatility. Other potential variables could be interest rates, the VIX or measures of political uncertainty. However, in order to allow for variables that take negative values, it is straightforward to extend our results to the case that the long-term component is given by some general nonnegative function $\tau_{0t} = f(\boldsymbol{\pi}_0'\mathbf{x}_t)$. For example, Opschoor et al. (2014) employ the specification $\tau_{0t} = \exp(\boldsymbol{\pi}_0'\mathbf{x}_t)$

with the Bloomberg Financial Conditions Index as the explanatory variable. In addition, see Remark 2 below.

Theorem 1. *If Assumptions 1-3 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

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

where $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}_0 + o_P(1)$. The result is presented in Proposition 1 in the Appendix. For doing so, the following Assumption 4 is required which ensures that $\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}(\boldsymbol{\eta})$ is finite with a uniform bound for all $\boldsymbol{\eta} \in \Theta$.

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

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 1-4 hold, then*

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

with

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

Note that the covariance matrix $\boldsymbol{\Sigma}$ in equation (31) takes the same form as in Lundbergh and Teräsvirta (2002). 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 1-4 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,* $\hat{h}_t = \hat{\omega} + \hat{\alpha}\varepsilon_{t-1}^2 + \hat{\beta}\hat{h}_{t-1}$, $\hat{\mathbf{r}}_t = \mathbf{x}_t - \hat{\alpha}/\hat{h}_t \sum_{j=0}^{t-1}\hat{\beta}^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j}$ *and*

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

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

Remark 1. *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 MIDAS lag length is unknown. Although the LM statistic can be easily computed for a variety of K's, our simulation experiments have shown that choosing* $K = 1$ *is typically 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.*[4] *These considerations also suggest that our test is not suited for selecting the true lag order of the GARCH-MIDAS-X.*

Remark 2. *Our result can be directly extended to the case that the long-term component is given by some general nonnegative function* $\tau_{0t} = f(\boldsymbol{\pi}_0'\mathbf{x}_t)$ *with* $f(0) = 1$. *In this situation, it directly follows that equation (20) can be rewritten as*

$$\mathbf{r}_{0,t}^{\infty} = f_0' \cdot (\mathbf{x}_t - \alpha_0\frac{1}{h_{0,t}^{\infty}}\sum_{j=0}^{\infty}\beta_0^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j}), \tag{34}$$

where $f_0' = \frac{\partial\tau_t}{\partial\boldsymbol{\pi}'\mathbf{x}_t}|_{\boldsymbol{\pi}=\mathbf{0}}$. *The expression for* $\mathbf{y}_{0,t}^{\infty}$ *is not affected. Since the factor* f_0' *cancels out in equation (32), the LM statistic remains unchanged. Note that in this case we can drop the assumption that* x_t *is a nonnegative explanatory variable.*

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

$$\widetilde{LM} = T\,\frac{SSR_0 - SSR_1}{SSR_0}, \tag{35}$$

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

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, $\widetilde{LM}$ is simply T times the uncentered R^2 of the regression.

Remark 3. *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 (21) can be rewritten as*

$$\mathbf{D}^{\infty}(\boldsymbol{\eta}_0) = \frac{1}{2T}\sum_{t=1}^{T}\left[\frac{\varepsilon_t^2}{\sigma_0^2} - 1\right]\begin{pmatrix} \sigma_0^{-2} \\ \mathbf{x}_t \end{pmatrix}. \tag{36}$$

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 $x_{t-k} = \varepsilon_{t-k}^2$.

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

Next, we compare our test statistic to the Lundbergh and Teräsvirta (2002) test for misspecification in GARCH models. Their test is based on the following specification $\varepsilon_t = \sqrt{h_{0t}^{\infty}}\xi_{0t} = \sqrt{h_{0t}^{\infty}\tau_{0t}}Z_t$, where h_{0t}^{∞} and τ_{0t} are defined as before. However, 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 τ_{0t} in equation (6). Because under this assumption $\xi_{0t}^2 = \sqrt{\tau_{0t}}Z_t$ 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 the GARCH-MIDAS-X, 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 τ_{0t}. Because of this, $\sqrt{h_{0t}^{\infty}}Z_t$ follows a GARCH(1,1) process under the null but *not* under the alternative.[5] Moreover,

[5]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.

it follows that $\partial h_t^\infty / \partial \boldsymbol{\pi} = \mathbf{0}$ and, hence, in the Lundbergh and Teräsvirta (2002) setting equation (20) 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, i.e. follow an ARCH process.

In the following section, 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) is a GARCH-MIDAS-X. We implement a regression-based version of the test as in equation (35) but with $\hat{\mathbf{r}}_t^{LT}$ instead of $\hat{\mathbf{r}}_t$. We denote the test statistic by $\widetilde{LM}_{LT}$. 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 (20) is simply given by $\mathbf{r}_{0,t}^\infty = \mathbf{x}_t = \hat{\mathbf{r}}_t^{LT,mod}$. We denote the corresponding test statistic $\widetilde{LM}_{LT,mod}$. 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, $\widetilde{LM}$, and $\widetilde{LM}_{LT,mod}$ can be expected to perform similarly if, for example, $\hat{\alpha} \approx 0$. 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. Moreover, the expression for $\hat{\mathbf{r}}_t - \hat{\mathbf{r}}_t^{LT,mod}$ suggests that the difference in power between our new test and the modified Lundbergh and Teräsvirta (2002) test should be stronger in situations in which x_t (and, hence, the long-term component) is more volatile.[6]

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.

[6]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.

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, Intermediate: I, High: H) in the conditional variance, whereby we measure persistence by $\alpha_0 + \beta_0$. We keep β_0 fixed at 0.9 and increase α_0 from 0.05 to 0.09. ω_0 is always chosen such that under the null $\sigma_0^2 = 1$.

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

Table 1: Empirical size.

		$Z_t \sim \mathcal{N}(0,1)$			$Z_t \sim t(7)$		
		L	I	H	L	I	H
LM	1%	0.9	1.2	1.3	0.7	0.9	0.7
	5%	4.6	5.0	5.2	3.1	3.8	3.9
	10%	9.0	9.7	10.2	7.2	7.4	7.3
$\widetilde{LM}$	1%	0.9	1.2	1.3	0.7	0.9	0.7
	5%	4.6	5.0	5.2	3.1	3.7	3.9
	10%	9.0	9.7	10.1	7.2	7.4	7.2
$\widetilde{LM}_{LT}$	1%	0.9	1.2	1.3	0.9	1.1	1.1
	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

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 ($\alpha = 0.05$), intermediate ($\alpha = 0.07$) and high ($\alpha = 0.09$) persistence. ω is chosen such that $\sigma_0^2 = 1$.

Under the null, we test for remaining ARCH effects by choosing $x_t = \varepsilon_t^2/\hat{h}_t$. In Table 1, we report the empirical size of the LM test given in equation (32), the regression version of the test, $\widetilde{LM}$, and the Lundbergh and Teräsvirta (2002) test statistic $\widetilde{LM}_{LT}$. For the three test statistics we have to choose the dimension of $\hat{\mathbf{r}}_t$ and $\hat{\mathbf{r}}_t^{LT}$, respectively. We opt for a dimension of one, which implies that the model under the alternative has an 'ARCH(1)' long-term component.[7] As Table 1 shows, the empirical size of all three 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 three test statistics are slightly undersized. For the $\widetilde{LM}_{LT}$ test statistic, this is an observation also made in Lundbergh and Teräsvirta (2002) and Halunga and Orme (2009).

3.2 Power properties

In order to consider a realistic example under the alternative, we 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.[8] 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.

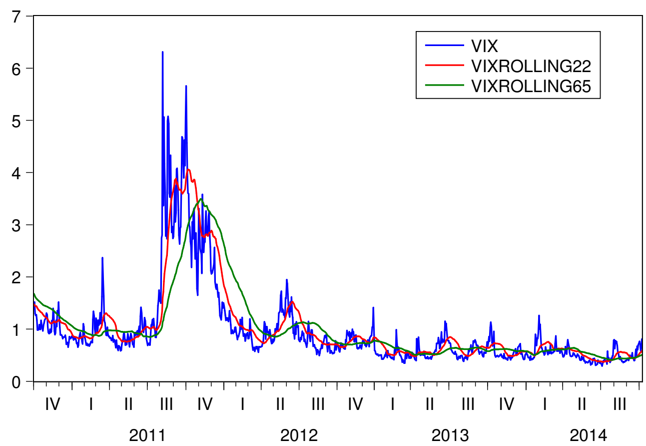


Figure 1: The figure shows the evolution of VIX_t (blue), $VIX_t^{(22)}$ (red) and $VIX_t^{(65)}$ (green) for the period October 2010 to October 2014. The three variables are presented in daily units.

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

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

The long-term component is given by $\tau_{0t} = 1 + 0.5 \sum_{k=1}^{K} \psi_{0k}(\omega_{01}, \omega_{02}) x_{t-k}$, whereby we specify the MIDAS weights via the following Beta weighting scheme:[9]

$$\psi_{0k}(\omega_{01}, \omega_{02}) = \frac{(k/(K+1))^{\omega_{01}-1} \cdot (1 - k/(K+1))^{\omega_{02}-1}}{\sum_{j=1}^{K} (j/(K+1))^{\omega_{01}-1} \cdot (1 - j/(K+1))^{\omega_{02}-1}}. \tag{38}$$

Table 2 presents the results of the Monte-Carlo simulations. The LM and the $\widetilde{LM}$ test statistics are based on $\hat{\mathbf{r}}_t$ with $x_t \in \{VIX_t, VIX_t^{(22)}, VIX_t^{(65)}\}$. We present size-adjusted rejection rates for two versions of the Lundbergh and Teräsvirta (2002) test. The *modified* Lundbergh and Teräsvirta (2002) test, $\widetilde{LM}_{LT,mod}$, is based on $\hat{\mathbf{r}}_t^{LT,mod}$ but uses the same x_t as in LM and $\widetilde{LM}$. As before, $\widetilde{LM}_{LT}$ 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 all four test statistics we choose $K = 1$. We denote the true MIDAS lag length in the data generating processes under the alternative by $K^\star$. We simulate processes with $K^\star \in \{1, 5, 22\}$. Thus, the results for size-adjusted power in Table 2 illustrate the performance of the test statistics when K is correctly chosen but also when $K < K^\star$.

We first consider the squared VIX as the explanatory variable, i.e. we choose $x_t = VIX_t$. The parameters in the Beta weighting scheme are given by $\omega_{01} = 1$ and $\omega_{02} = 10$. These values ensure that the ψ_{0k} decay monotonically from the first lag. In the GARCH equation we employ models with $\alpha_0 = 0.09$ (high persistence) and $\alpha_0 = 0.07$ (intermediate persistence). Besides the size-adjusted power for different values of $K^\star$, we also report the following 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.[10] For example, for $K^\star = 1$ and $\alpha_0 = 0.09$, 12.4% of the total conditional variance is due to the long-term component. When α_0 is decreased to 0.07, the VR increases to 29.5%. 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 $K^\star = 1$, the LM and the $\widetilde{LM}$ tests reject the null hypothesis in 57.2% of the simulations at the nominal 5% level. In contrast, the rejection rate of the modified Lundbergh and Teräsvirta (2002) test, $\widetilde{LM}_{LT,mod}$, is 34.8% only. When $K^\star$ is increased to 5 and 22, the long-term component is based on a weighted average of the values of the VIX during the last week or month. That is, the long-term

[9]For a detailed discussion of the Beta weighting scheme see Ghysels et al. (2006).

[10]This variance ratio has been employed in Conrad and Loch (2014) 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 period 1973 to 2010.

Table 2: Empirical size-adjusted power for long-term components based on the VIX.

x_t		VIX_t						$VIX_t^{(22)}$	$VIX_t^{(65)}$
		$\omega_{01}=1$, $\omega_{02}=10$							
		$\alpha_0=0.09$			$\alpha_0=0.07$			$\alpha_0=0.09$	
$K^\star$		1	5	22	1	5	22	1	1
	1%	34.8	33.1	21.1	44.7	42.7	29.9	14.3	9.9
LM	5%	57.2	54.8	39.2	66.5	64.6	51.5	36.1	26.5
	10%	66.4	64.8	50.7	75.6	74.2	63.0	47.2	35.4
	1%	34.8	33.0	21.3	44.5	42.7	29.9	14.4	10.1
$\widetilde{LM}$	5%	57.2	54.7	39.0	66.6	64.6	51.7	36.1	26.6
	10%	66.4	64.8	50.8	75.6	74.2	62.9	47.1	35.6
	1%	16.0	15.9	13.0	32.6	32.0	27.1	6.6	4.7
$\widetilde{LM}_{LT,mod}$	5%	34.8	34.1	30.3	59.2	58.1	51.1	18.9	15.4
	10%	44.2	43.3	38.6	71.1	70.0	65.6	27.1	23.8
	1%	0.9	0.9	0.9	1.0	1.0	0.9	0.9	0.8
$\widetilde{LM}_{LT}$	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: The table reports the size-adjusted power. The specification of the long term component is given by $\tau_{0,t} = 1 + 0.5\sum_{k=1}^{K}\psi_{0k}(\omega_{01},\omega_{02})x_{t-k}$ with Beta weighting scheme (see equation (38)). The GARCH parameters are $\beta_0 = 0.9$ and $\omega_0 = 1-\alpha_0-\beta_0$. Innovations Z_t are standard normal distributed. $K^\star$ denotes the true MIDAS lag length in the DGP. All test statistics are based on $K = 1$.

component becomes less variable and, hence, more difficult to detect. Consequently, the power of all three tests deteriorates. Although, the LM and the $\widetilde{LM}$ test still have considerably higher power than $\widetilde{LM}_{LT,mod}$, the difference in power is decreasing when the long-term component gets smoother. Next, when α_0 is decreased to 0.07, this increases the power of the three tests. For example, for $K^\star = 1$ the size-adjusted power at the nominal 5% level is 66.5% 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, increasing $K^\star$, i.e. increasing the smoothness of the long-term component, reduces the power of the tests. In line with the arguments at the end of Section 2.4, the difference in the power of the LM and $\widetilde{LM}_{LT,mod}$ statistics is less strong when α_0 is decreased to

0.07. Again, the difference in power is the larger the smoother the long-term component (i.e. the smaller $K^\star$) is. 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 $K^\star = 1$ the long-term components are very smooth and the lowest VR's are observed.[11] 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, $\widetilde{LM}_{LT}$, has no power to detect deviation from the null.

We performed the same analysis as in Table 2 for the case of Student-t distributed innovations Z_t (see Table 5 in the Appendix). 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. We also performed simulations in which we increased K such that it approaches the true lag length, say $K^\star = 5$ or $K^\star = 22$. As discussed in Remark 1, given the smoothness of our explanatory variable, this did not lead to gains in power relative to simply choosing $K = 1$.[12]

In Table 3, we investigate the effects of changing the weighting scheme, ψ_{0k}, on the power of the test statistics. In the first two columns the weighting scheme is given by $\omega_{01} = 1$ and $\omega_{02} = 5$, i.e. the weights are still monotonically decreasing, but more slowly than before. This implies that the long-term component becomes less volatile and, hence, the power of the tests decreases (compare columns one and two of Table 3 with columns two and three of Table 2). In columns three and four of Table 3 the parameters ω_{01} and ω_{02} are chosen such that the weighting schemes are hump-shaped. For example, comparing column one with column three of Table 3 suggests that for the same lag length $K^\star = 5$, replacing a monotonically decaying weighting scheme with a hump-shaped one decreases the power of all versions of the test. Again, for all specifications the size-adjusted power of the $\widetilde{LM}$ test is higher than the one of $\widetilde{LM}_{LT,mod}$.

[11] Note that the model based on the $VIX_t^{(N)}$ with lag length $K^\star = 1$ can be thought of as representing a long-term component based on the VIX_t as the explanatory variable but with $K^\star = N$ and weights equal to $1/N$. Thus, by 'selecting' an appropriate explanatory variable one can always ensure that $K = 1$ is sufficient in the test.

[12] The first order autocorrelation of VIX_t is 0.95.

Table 3: Empirical size-adjusted power for different weighting schemes.

$x_t = VIX_t$		$\alpha_0 = 0.09$			
		$\omega_{01} = 1$, $\omega_{02} = 5$		$\omega_{01} = 3, \omega_{02} = 5$	$\omega_{01} = 3, \omega_{02} = 20$
$K^\star$		5	22	5	22
	1%	28.2	13.8	19.8	16.7
LM	5%	49.2	29.9	38.6	33.8
	10%	60.6	39.8	50.2	45.2
	1%	28.2	13.7	19.7	16.7
$\widetilde{LM}$	5%	49.2	29.9	38.7	34.0
	10%	60.6	39.9	50.2	45.2
	1%	15.0	10.8	12.7	11.7
$\widetilde{LM}_{LT,mod}$	5%	32.8	27.4	30.0	28.6
	10%	41.7	35.8	38.5	36.8
	1%	0.9	0.9	0.9	0.9
$\widetilde{LM}_{LT}$	5%	5.8	5.3	5.2	5.3
	10%	10.5	10.0	10.4	10.4
VR		12.2	11.9	12.2	12.1

Notes: See Table 2.

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).

4 Empirical Application

Finally, we apply our test and estimate a GARCH-MIDAS-X model for daily log-returns on the S&P 500 for the period January 2000 to October 2014. As explanatory variables, we employ the squared VIX, realized volatility as well as the ADS Business Conditions Index (see Aruoba et al., 2009).[13] As before, we construct monthly rolling window versions of all three variables denoted by $VIX_t^{(22)}$, $RV_t^{(22)}$ and $ADS_t^{(22)}$. Since $ADS_t^{(22)}$ can take positive as well as negative values, we specify a 'log-version' of equation (4) for the

[13]Dorion (2013) shows that a GARCH-MIDAS model based on the ADS Business Conditions Index is informative for the valuation of options.

long-term component with $\ln(\tilde{\tau}_{0t}) = \sigma_0^2 + \tilde{\pi}_0 \sum_{k=1}^{K} \psi_{0k} x_{t-k}$.[14] We estimate the GARCH-MIDAS-X models using a restricted Beta weighting scheme (i.e. we impose $\omega_{01} = 1$ in equation (38)) and select a MIDAS lag length of 252 (i.e. one year of lagged observations). Table 4, Panel A, shows the estimates of the parameters of interest for the three GARCH-MIDAS-X models. First, note that for all three cases the ARCH/GARCH parameter estimates are basically the same. The estimates of $\tilde{\pi}$ are positive for the $VIX_t^{(22)}$ and $RV_t^{(22)}$, but negative for the $ADS_t^{(22)}$. That is, higher expected/realized volatility leads to an increase in long-term volatility, while an improvement in business conditions reduces long-term volatility. These findings are perfectly in line with the counter-cyclical behavior of long-term volatility as observed in Conrad and Loch (2014). The estimated weighting parameters ω_2 imply slowly decreasing weights for the $VIX_t^{(22)}$ and $RV_t^{(22)}$, but rapidly decaying weights for the $ADS_t^{(22)}$.[15] However, strictly speaking the parameter estimates and standard errors reported in Panel A do not allow us to formally test the null that the explanatory variables have no significant effect on long-term volatility, since the asymptotic theory is either not available or does not permit this null hypothesis.

Figure 2, left, shows the estimated long-term component ($\sqrt{\tilde{\tau}_t}$) as well as the conditional volatility ($\sqrt{\tilde{h}_t \tilde{\tau}_t}$) based on the $ADS_t^{(22)}$ at an annualized scale. The figure clearly reveals that the long-term component based on $ADS_t^{(22)}$ captures the increase in financial market volatility during the Great Recession, but not during the European sovereign debt crisis. Figure 2, right, allows for a comparison of all three long-term components. Note that the long-term components based on the $VIX_t^{(22)}$ and $RV_t^{(22)}$ – which are based on an estimated weighting parameter around two – are much smoother than the long-term component based on the $ADS_t^{(22)}$. This is also reflected in the corresponding variance ratios which are below 20% for the former variables but above 40% for the latter. This suggests that it should be more easy to detect the effect of the $ADS_t^{(22)}$ on long-term volatility, than the effect of the $VIX_t^{(22)}$ and $RV_t^{(22)}$. This intuition is confirmed in Table 4, Panel B, which presents the test results. Both versions of our test reject the null that the three variables have no significant effect on long-term volatility at the 1% or 2% level. In stark contrast, the modified Lundbergh and Teräsvirta (2002) test rejects only

[14] Recall from Remark 2 that our test statistic applies to this situation as well.

[15] Originally, we also estimated models with the daily VIX_t, RV_t and ADS_t as explanatory variables. However, for the VIX_t and RV_t the estimated MIDAS weights were declining so quickly, that the $\tilde{\tau}_t$ components were effectively given by the VIX or the realized volatility of the last day and, hence, highly volatile. In this sense, they no longer represented a smooth 'long-term' component.

Table 4: GARCH-MIDAS-X for S&P 500 returns.

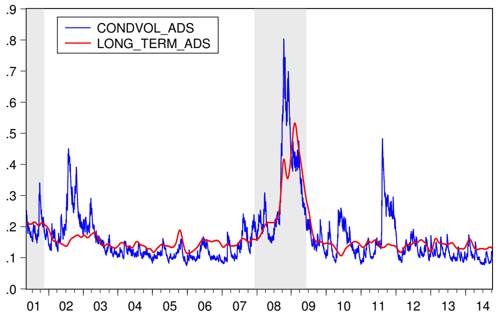
x_t	$VIX_t^{(22)}$	$RV_t^{(22)}$	$ADS_t^{(22)}$
	Panel A: Parameter Estimates		
α	0.091 (0.012)	0.091 (0.012)	0.088 (0.011)
β	0.885 (0.018)	0.885 (0.017)	0.885 (0.013)
$\tilde{\pi}$	0.281 (0.111))	0.261 (0.087)	-0.708 (0.100)
ω_2	2.453 (1.931)	1.900 (1.026)	15.552 (10.883)
VR	17.2	19.3	42.7
	Panel B: Misspecification Tests		
LM	6.40 [0.01]	6.50 [0.01]	5.42 [0.02]
$\widetilde{LM}$	6.40 [0.01]	6.50 [0.01]	5.42 [0.02]
$\widetilde{LM}_{LT,mod}$	1.78 [0.18]	1.98 [0.16]	9.75 [0.00]

Notes: The table presents estimation results for the GARCH-MIDAS-X model with $\ln(\tilde{\tau}_{0t}) = \sigma_0^2 + \tilde{\pi}_0 \sum_{k=1}^{252} \psi_{0k} x_{t-k}$. All estimations are based on daily data from January 2000 to October 2014. We include a restricted Beta weighting scheme ($\omega_{01} = 1$). The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The reported LM tests for misspecification are based on $K = 1$. Numbers in brackets are p-values.

in the case of the $ADS_t^{(22)}$. This result confirms our findings from Section 3 that our test is more sensitive to smooth movements in the long-term component. In summary, in all three cases the newly proposed LM test clearly rejects the null of a constant long-term component and, thereby, confirms the GARCH-MIDAS-X specifications.

5 Conclusions

We develop a Lagrange Multiplier test for the null hypothesis of a GARCH volatility against the alternative of a GARCH-MIDAS 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



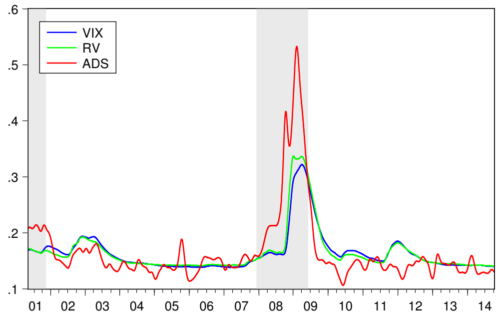


Figure 2: The left figure shows the conditional volatility $\sqrt{\tilde{h}_t \tilde{\tau}_t}$ (blue line) and the long-term volatility component $\sqrt{\tilde{\tau}_t}$ (red line) based on the $ADS_t^{(22)}$. The right figure provides a comparison of the long-term volatility components based on the $VIX_t^{(22)}$ (blue), $RV_t^{(22)}$ (green) and $ADS_t^{(22)}$ (red). Both figures are for the January 2000 to October 2014 period and graph the volatilities at an annualized scale. Shaded areas represent NBER recession periods.

performance. In an application to S&P 500 returns, we find that the test provides useful guidance in model specification.

There are several interesting extensions that we would like to address in future work. Clearly, we could allow for asymmetries in the short-term component. More importantly, it would be interesting to extend our test to the case that the explanatory variable varies at a different frequency than the returns. Also, the case of more than one explanatory variable at a time could be explored.

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.

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., 2015. The variance risk premium and fundamental uncertainty. *Economics Letters* 132, 56-60.

Conrad, C., Loch, K., 2014. Anticipating long-term stock market volatility. *Journal of Applied Econometrics*, forthcoming.

Dorion, C., 2013. Option valuation with macro-finance variables. *Journal of Financial and Quantitative Analysis*, forthcoming.

Engle, R. F., 2009. The risk that risk will change. *Journal of Investment Management* 7, 24-28.

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.

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

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.

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.

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 $\mathbf{\Omega}_{\boldsymbol{\eta\eta}}$ is finite and positive definite. What remains to be shown is that $\mathbf{\Omega}_{\boldsymbol{\pi\pi}}$ 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 $\mathbf{\Omega}_{\boldsymbol{\pi\pi}}$:

Recall from equation (22) that $\mathbf{\Omega}_{\boldsymbol{\pi\pi}} = \frac{1}{4}(\kappa_Z - 1)\mathbf{E}[\mathbf{r}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})']$. It follows from Assumption 2 that $0 < \kappa_Z - 1 < \infty$. Moreover, $||\mathbf{E}[\mathbf{r}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})']||$ is finite if $\mathbf{E}[||\mathbf{r}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})'||] < \infty$.[16] A typical element of the $K \times 1$ vector $\mathbf{r}_{0,t}^{\infty}$ is given by

$$r_{0,kt}^{\infty} = x_{t-k} - \alpha_0 \frac{1}{h_{0,t}^{\infty}} \sum_{j=0}^{\infty} \beta_0^j \varepsilon_{t-1-j}^2 x_{t-1-k-j}. \tag{39}$$

First, $\mathbf{E}[|x_{t-k}|^2] < \infty$ by Assumption 3. Second,

$$\begin{aligned}
\left(\mathbf{E}\left|\frac{\sum_{j=0}^{\infty}\alpha_0\beta_0^j\varepsilon_{t-1-j}^2 x_{t-1-k-j}}{h_{0,t}^{\infty}}\right|^2\right)^{1/2} &\leq \left(\mathbf{E}\left|\sum_{j=0}^{\infty}\frac{\alpha_0\beta_0^j\varepsilon_{t-1-j}^2}{(\omega_0+\alpha_0\beta_0^j\varepsilon_{t-1-j}^2)}x_{t-1-k-j}\right|^2\right)^{1/2} &(40)\\
&\leq \sum_{j=0}^{\infty}\left(\mathbf{E}\left|\frac{\alpha_0\beta_0^j\varepsilon_{t-1-j}^2}{(\omega_0+\alpha_0\beta_0^j\varepsilon_{t-1-j}^2)}x_{t-1-k-j}\right|^2\right)^{1/2} &(41)\\
&\leq \sum_{j=0}^{\infty}\left(\mathbf{E}\left|\left(\frac{\alpha_0\beta_0^j}{\omega_0}\varepsilon_{t-1-j}^2\right)^{s/4} x_{t-1-k-j}\right|^2\right)^{1/2} &(42)\\
&\leq \frac{\alpha_0^{s/4}}{\omega_0^{s/4}}\left(\mathbf{E}\left[\varepsilon_{t-1-j}^{2s}\right]\right)^{1/4}\left(\mathbf{E}\left[|x_{t-1-k-j}|^4\right]\right)^{1/4} \\
&\qquad \sum_{j=0}^{\infty}\beta_0^{js/4} < \infty.
\end{aligned}$$

The arguments used above are similar to the ones in Francq and Zakïoan (2004, Eq. (4.19), p.619). In particular, in equation (40) we use that $h_{0,t}^{\infty} \geq \omega_0 + \alpha_0\beta_0^j\varepsilon_{t-1-j}^2$. In equation (41) we use Minkowski's inequality. Next, in equation (42) we use the fact that $w/(1+w) \leq w^s$ for all $w > 0$ and any $s \in (0,1)$. Finally, Assumption 1 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 3, $\mathbf{E}\left[|x_{t-1-k-j}|^4\right] < \infty$.

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

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

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

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.,

$$0 = \sum_{k=1}^{K} c_k \left(x_{t-k} - \frac{\alpha_0}{h_{0,t}^{\infty}} \sum_{j=0}^{\infty} \beta_0^j \varepsilon_{t-1-j}^2 x_{t-1-k-j} \right) \quad \text{a.s.} \tag{43}$$

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

$$\mathbf{c}'\mathbf{x}_t = \left[\frac{\alpha_0}{h_{0,t}^{\infty}} (1 - \beta_0 L)^{-1} L \right] (\varepsilon_t^2 \mathbf{c}'\mathbf{x}_t) \quad \text{a.s.,} \tag{44}$$

where L denotes the lag operator. Clearly, the operator in square brackets cannot have an eigenvalue of 1. Moreover, Assumption 2 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 3. Thus, $\mathbf{\Omega}_{\pi\pi}$ 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 1-3 it then follows that $\mathbf{d}_t^{\infty}(\boldsymbol{\eta}_0)$ is a stationary and ergodic martingale difference sequence with finite second moment. Applying Billingsley's (1961) central limit theorem for martingale differences gives the result.

∎

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

Proposition 1. *Under Assumptions 1-4, we have that*

$$-\frac{1}{T} \sum_{t=1}^{T} \frac{\partial \mathbf{d}_{\pi,t}^{\infty}(\tilde{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}'} \xrightarrow{P} \mathbf{J}_{\pi\eta} = -\mathbf{E}\left[\frac{\partial \mathbf{d}_{\pi,t}^{\infty}(\boldsymbol{\eta}_0)}{\partial \boldsymbol{\eta}'} \right], \tag{45}$$

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

Proof of Proposition 1. We obtain (45) by showing that $\mathbf{J}_{\pi\eta}(\boldsymbol{\eta}) = -\mathbf{E}\left[\frac{\partial \mathbf{d}_{\pi,t}^{\infty}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'}\right]$ is finite with a uniform bound for all $\boldsymbol{\eta} \in \Theta$. Then a uniform weak law of large numbers (see, e.g., Theorem 3.1. in Ling and McAleer, 2003) implies

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

Equation (45) follows from the triangle inequality and the fact that $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}_0 + o_P(1)$.

Using equation (24) we obtain

$$\begin{aligned} \left|\left|\frac{\partial \mathbf{d}^{\infty}_{\boldsymbol{\pi},t}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'}\right|\right| &\leq \frac{1}{2}\left(\left|\frac{\varepsilon_t^2}{h_t^{\infty}}\right| \cdot ||\mathbf{r}_t^{\infty}|| \cdot ||(\mathbf{y}_t^{\infty})'|| + \left|\frac{\varepsilon_t^2}{h_t^{\infty}} - 1\right| \cdot \left|\left|\frac{\partial \mathbf{r}_t^{\infty}}{\partial \boldsymbol{\eta}'}\right|\right|\right) \\ &\leq C|\varepsilon_t^2 + \omega| \left(||\mathbf{r}_t^{\infty}|| \cdot ||(\mathbf{y}_t^{\infty})'|| + \left|\left|\frac{\partial \mathbf{r}_t^{\infty}}{\partial \boldsymbol{\eta}'}\right|\right|\right). \end{aligned} \tag{46}$$

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

$$\begin{aligned} \left|\frac{1}{h_t^{\infty}}\frac{\partial h_t^{\infty}}{\partial \beta}\right| &\leq \left|\frac{1}{\beta}\sum_{j=0}^{\infty} \frac{j\beta^j(\omega + \alpha\varepsilon_{t-1-j}^2)}{\omega + \beta^j(\omega + \alpha\varepsilon_{t-1-j}^2)}\right| \\ &\leq \frac{1}{\beta\omega^s}\sum_{j=0}^{\infty} j\left|\beta^{js}(\omega + \alpha\varepsilon_{t-1-j}^2)^s\right|, \end{aligned} \tag{47}$$

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 (46), 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

$$\sqrt{\mathbf{E}[\sup_{\boldsymbol{\eta}}|\varepsilon_t^2 + \omega|^2]\mathbf{E}[\sup_{\boldsymbol{\eta}}||\mathbf{r}_t^{\infty}||^2]} \tag{48}$$

and

$$\begin{aligned} &\sup_{\boldsymbol{\eta}} \sum_{j=0}^{\infty} j\beta^{js}\mathbf{E}[\sup_{\boldsymbol{\eta}}(\omega + \alpha\varepsilon_{t-1-j}^2)^s|\varepsilon_t^2 + \omega|\ ||\mathbf{r}_t^{\infty}||] \leq \\ &\sup_{\boldsymbol{\eta}} \sum_{j=0}^{\infty} j\beta^{js}\sqrt{\mathbf{E}[\sup_{\boldsymbol{\eta}}(\omega + \alpha\varepsilon_{t-1-j}^2)^{2s}|\varepsilon_t^2 + \omega|^2]\mathbf{E}[\sup_{\boldsymbol{\eta}}||\mathbf{r}_t^{\infty}||^2]}. \end{aligned} \tag{49}$$

The finiteness of (48) follows from Assumption 4 and similar arguments as in the proof of Theorem 1. The finiteness of (49) follows by applying Hölder's inequality, since for the elements in the sum which involve expectations of the squared observations we have

$$\begin{aligned} &\mathbf{E}[\sup_{\boldsymbol{\eta}}(\omega + \alpha\varepsilon_{t-1-j}^2)^{2s}|\varepsilon_t^2 + \omega|^2] \leq \\ &\left(\mathbf{E}[\sup_{\boldsymbol{\eta}}(\omega + \alpha\varepsilon_{t-1-j}^2)^{2(1+s)}]\right)^{s/(1+s)} \left(\mathbf{E}[\sup_{\boldsymbol{\eta}}|\varepsilon_t^2 + \omega|^{2(1+s)}]\right)^{1/(1+s)} \end{aligned} \tag{50}$$

and Assumption 4 applies again.

Using the Cauchy-Schwarz-Inequality for the two factors in the second term in (46), 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

$$\begin{aligned}
\frac{\partial \mathbf{r}_t^\infty}{\partial \boldsymbol{\eta}'} &= \frac{\partial}{\partial \boldsymbol{\eta}'}\mathbf{x}_t - \frac{\partial}{\partial \boldsymbol{\eta}'}\left(\frac{1}{h_t^\infty}\sum_{j=0}^{\infty}\alpha\beta^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j}\right) \\
&= \frac{\partial}{\partial \boldsymbol{\eta}'}\mathbf{x}_t - \frac{1}{h_t^\infty}\left(\sum_{j=0}^{\infty}\alpha\beta^j\varepsilon_{t-1-j}^2\frac{\partial}{\partial \boldsymbol{\eta}'}\mathbf{x}_{t-1-j}\right) \\
&\quad + \left(\frac{1}{h_t^\infty}\sum_{j=0}^{\infty}\alpha\beta^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j}\right)(\mathbf{y}_t^\infty)' - \frac{1}{h_t^\infty}\sum_{j=0}^{\infty}\mathbf{x}_{t-1-j}\left(\frac{\partial}{\partial \boldsymbol{\eta}'}\alpha\beta^j\varepsilon_{t-1-j}^2\right)
\end{aligned} \tag{51}$$

The first two terms vanish in the GARCH-MIDAS-X with exogenous explanatory variable $\mathbf{x}_t$ as $\frac{\partial \mathbf{x}_t}{\partial \boldsymbol{\eta}'} = \mathbf{0}$ or in the GARCH-MIDAS-RV with $x_{t-t} = \varepsilon_{t-k}^2$.

Remark 4. *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* (51) *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:

$$\frac{1}{h_t^\infty}\begin{pmatrix}
0 & \sum_{j=0}^{\infty}\beta^j\varepsilon_{t-1-j}^2 x_{t-2-j} & \alpha\sum_{j=0}^{\infty}j\beta^{j-1}\varepsilon_{t-1-j}^2 x_{t-2-j} \\
0 & \sum_{j=0}^{\infty}\beta^j\varepsilon_{t-1-j}^2 x_{t-3-j} & \alpha\sum_{j=0}^{\infty}j\beta^{j-1}\varepsilon_{t-1-j}^2 x_{t-3-j} \\
\vdots & & \\
0 & \sum_{j=0}^{\infty}\beta^j\varepsilon_{t-1-j}^2 x_{t-1-K-j} & \alpha\sum_{j=0}^{\infty}j\beta^{j-1}\varepsilon_{t-1-j}^2 x_{t-1-K-j}
\end{pmatrix} \tag{52}$$

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 $\boldsymbol{\eta}_0$

$$\mathbf{0} = \sqrt{T}\mathbf{D}_{\boldsymbol{\eta}}^{\infty}(\hat{\boldsymbol{\eta}}) = \sqrt{T}\mathbf{D}_{\boldsymbol{\eta}}^{\infty}(\boldsymbol{\eta}_0) + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial \mathbf{d}_{\boldsymbol{\eta},t}^{\infty}(\tilde{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}'}\sqrt{T}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) \tag{53}$$

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

$$-\frac{1}{T}\sum_{t=1}^{T}\frac{\partial \mathbf{d}_{\boldsymbol{\eta},t}^{\infty}(\tilde{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}'} \xrightarrow{P} \mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}} = -\mathbf{E}\left[\frac{\partial \mathbf{d}_{\boldsymbol{\eta},t}^{\infty}(\boldsymbol{\eta}_0)}{\partial \boldsymbol{\eta}'}\right] \tag{54}$$

and, hence, equation (53) can be written as

$$\sqrt{T}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) = \mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\mathbf{D}_{\boldsymbol{\eta}}^{\infty}(\boldsymbol{\eta}_0) + o_P(1). \tag{55}$$

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

$$\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\hat{\boldsymbol{\eta}}) = \sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\boldsymbol{\eta}_0) + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial \mathbf{d}_{\boldsymbol{\pi},t}^{\infty}(\tilde{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}'}\sqrt{T}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0). \tag{56}$$

Combining equation (55) and Proposition 1 leads to

$$\begin{aligned}
\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\hat{\boldsymbol{\eta}}) &= \sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\boldsymbol{\eta}_0) - \mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\sqrt{T}\mathbf{D}_{\boldsymbol{\eta}}^{\infty}(\boldsymbol{\eta}_0) + o_P(1) & (57)\\
&= [-\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1} \; : \; \mathbf{I}]\sqrt{T}\begin{pmatrix}\mathbf{D}_{\boldsymbol{\eta}}^{\infty}(\boldsymbol{\eta}_0)\\ \mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\boldsymbol{\eta}_0)\end{pmatrix} + o_P(1) & (58)\\
&= [-\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1} \; : \; \mathbf{I}]\sqrt{T}\mathbf{D}^{\infty}(\boldsymbol{\eta}_0) + o_P(1). & (59)
\end{aligned}$$

Applying Theorem 1 gives the asymptotic distribution

$$\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\hat{\boldsymbol{\eta}}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, [\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1} \; : \; \mathbf{I}]\boldsymbol{\Omega}[\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1} \; : \; \mathbf{I}]') \tag{60}$$

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 (22), (25) and (26) the expression for $\boldsymbol{\Sigma}$ simplifies to:

$$\boldsymbol{\Sigma} = \frac{1}{4}(\kappa_Z - 1)\left(\mathbf{E}[\mathbf{r}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})'] - \mathbf{E}[\mathbf{r}_{0,t}^{\infty}(\mathbf{y}_{0,t}^{\infty})']\left(\mathbf{E}[\mathbf{y}_{0,t}^{\infty}(\mathbf{y}_{0,t}^{\infty})']\right)^{-1}\mathbf{E}[\mathbf{y}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})']\right). \tag{61}$$

∎

Proof of Theorem 3. We show that

$$\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}(\hat{\boldsymbol{\eta}}) = \sqrt{T}\mathbf{D}^{\infty}_{\boldsymbol{\pi}}(\hat{\boldsymbol{\eta}}) + o_P(1). \tag{62}$$

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

$$\sup_{\boldsymbol{\eta}}||\sqrt{T}\mathbf{D}^{\infty}_{\boldsymbol{\pi}}(\boldsymbol{\eta}) - \sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}(\boldsymbol{\eta})|| \leq \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\sup_{\boldsymbol{\eta}}||\mathbf{d}^{\infty}_{\boldsymbol{\pi},t}(\boldsymbol{\eta}) - \mathbf{d}_{\boldsymbol{\pi},t}(\boldsymbol{\eta})||, \tag{63}$$

we establish equation (62) by showing that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\sup_{\boldsymbol{\eta}}||\mathbf{d}^{\infty}_{\boldsymbol{\pi},t}(\boldsymbol{\eta}) - \mathbf{d}_{\boldsymbol{\pi},t}(\boldsymbol{\eta})|| = o_P(1). \tag{64}$$

Consider the following decomposition:

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

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

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

First, note that

$$\mathbf{r}_t^{\infty} - \mathbf{r}_t = -\alpha\frac{1}{h_t^{\infty}}\sum_{j=t}^{\infty}\beta^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j}. \tag{65}$$

Next, consider a typical element:

$$
\begin{aligned}
\left(\mathbf{E}\sup_{\boldsymbol{\eta}}|r^{\infty}_{k,t}-r_{k,t}|^2\right)^{1/2} &= \left(\mathbf{E}\sup_{\boldsymbol{\eta}}\left|\alpha\frac{1}{h^{\infty}_t}\sum_{j=t}^{\infty}\beta^j\varepsilon^2_{t-1-j}x_{t-1-k-j}\right|^2\right)^{1/2} \\
&\leq \sum_{j=t}^{\infty}\left(\mathbf{E}\sup_{\boldsymbol{\eta}}\left|\frac{\alpha\beta^j\varepsilon^2_{t-1-j}}{\omega+\alpha\beta^j\varepsilon^2_{t-1-k-j}}x_{t-1-k-j}\right|^2\right)^{1/2} \\
&\leq \sum_{j=t}^{\infty}\left(\mathbf{E}\sup_{\boldsymbol{\eta}}\left|\left(\frac{\alpha\beta^j}{\omega}\varepsilon^2_{t-1-j}\right)^{s/4}x_{t-1-k-j}\right|^2\right)^{1/2} \\
&\leq \left(\mathbf{E}[|\varepsilon_{t-1-j}|^{2s}]\right)^{1/4}\left(\mathbf{E}[|x_{t-1-k-j}|^4]\right)^{1/4} \\
&\qquad \sup_{\boldsymbol{\eta}}\left(\frac{\alpha}{\omega}\right)^{s/4}\sum_{j=t}^{\infty}\beta^{js/4} \\
&= \left(\mathbf{E}[|\varepsilon_{t-1-j}|^{2s}]\right)^{1/4}\left(\mathbf{E}[|x_{t-1-k-j}|^4]\right)^{1/4} \\
&\qquad \sup_{\boldsymbol{\eta}}\left(\frac{\alpha}{\omega}\right)^{s/4}\frac{(\beta^{s/4})^t}{1-\beta^{s/4}} \qquad (66)
\end{aligned}
$$

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

Hence,

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

by Assumption 1 and equation (66). Therefore, $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{E}\sup_{\boldsymbol{\eta}}|\varepsilon^2_t|\ ||\mathbf{r}^{\infty}_t-\mathbf{r}_t|| = o(1)$ and, hence, by Markov's inequality $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\sup_{\boldsymbol{\eta}}|\varepsilon^2_t|\ ||\mathbf{r}^{\infty}_t-\mathbf{r}_t|| = o_P(1)$.

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

$$\left|\frac{h^{\infty}_t-h_t}{h^{\infty}_t}\right| \leq \frac{\alpha^s}{\omega^s}\sum_{j=t}^{\infty}(\beta^s)^j\varepsilon^{2s}_{t-j}, \qquad (67)$$

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

$$
\begin{aligned}
\mathbf{E}\sup_{\boldsymbol{\eta}}\varepsilon^2_t||\mathbf{r}^{\infty}_t||\left|\frac{h^{\infty}_t-h_t}{h^{\infty}_t}\right| &\leq \mathbf{E}\sup_{\boldsymbol{\eta}}||\varepsilon^2_t\mathbf{r}^{\infty}_t\varepsilon^{2s}_{t-j}||\ \sup_{\boldsymbol{\eta}}\frac{\alpha^s}{\omega^s}\sum_{j=t}^{\infty}(\beta^s)^j \\
&\leq \sqrt{\mathbf{E}\sup_{\boldsymbol{\eta}}||\mathbf{r}^{\infty}_t||^2\mathbf{E}|\varepsilon^4_t\varepsilon^{4s}_{t-j}|}\ \sup_{\boldsymbol{\eta}}\frac{\alpha^s}{\omega^s}(\beta^s)^t\sum_{j=0}^{\infty}(\beta^s)^j \\
&= \sqrt{\mathbf{E}\sup_{\boldsymbol{\eta}}||\mathbf{r}^{\infty}_t||^2\mathbf{E}|\varepsilon^4_t\varepsilon^{4s}_{t-j}|}\ \sup_{\boldsymbol{\eta}}\frac{\alpha^s}{\omega^s(1-\beta^s)}(\beta^s)^t \\
&= O((\beta^s)^t). \qquad (68)
\end{aligned}
$$

The last line follows because it can be shown by similar arguments as in the proof of Theorem 1 that $\mathbf{E}\sup_{\boldsymbol{\eta}}||\mathbf{r}^{\infty}_t||^2 < \infty$ and because Hölder's inequality and Assumption 4 imply that $\mathbf{E}|\varepsilon_t^4 \varepsilon_{t-j}^{4s}| \le \left(\mathbf{E}|\varepsilon_t^{4(1+s)}|\right)^{1/(1+s)} \left(\mathbf{E}|\varepsilon_{t-j}^{4(1+s)}|\right)^{s/(1+s)} < \infty$. Equation (68) implies that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathbf{E}\mathrm{sup}_{\boldsymbol{\eta}} \varepsilon_t^2 ||\mathbf{r}_t^\infty|| \left| \frac{h_t^\infty - h_t}{h_t^\infty} \right| = o(1), \tag{69}$$

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

The third term can be treated as follows:

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

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

∎

B Simulation: size-adjusted power for 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 5: Empirical size-adjusted power for t distributed innovations.

x_t		VIX_t						$VIX_t^{(22)}$	$VIX_t^{(65)}$
		$\omega_{01} = 1$, $\omega_{02} = 10$							
		$\alpha_0 = 0.09$			$\alpha_0 = 0.07$			$\alpha_0 = 0.09$	
$K^\star$		1	5	22	1	5	22	1	1
LM	1%	24.3	20.0	14.6	30.0	28.6	18.7	14.9	10.5
	5%	39.7	34.5	28.4	48.3	46.7	37.9	26.2	22.1
	10%	52.7	46.5	39.2	60.8	59.3	47.7	39.4	32.4
$\widetilde{LM}$	1%	24.3	19.9	14.5	30.0	28.5	18.7	14.9	10.5
	5%	39.8	34.5	28.4	48.3	46.7	37.9	26.3	22.4
	10%	52.7	46.5	39.2	60.7	59.4	47.8	39.4	32.4
$\widetilde{LM}_{LT,mod}$	1%	10.3	9.2	8.3	20.1	19.6	16.4	5.4	3.9
	5%	27.4	25.9	24.6	43.0	42.8	38.5	18.5	16.8
	10%	37.3	35.0	32.9	53.8	52.9	49.4	26.3	24.2
$\widetilde{LM}_{LT}$	1%	1.0	1.1	1.1	1.0	1.0	1.2	1.0	1.0
	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. Otherwise see Table 2.