Supplementary Material:

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

Christian Conrad (Heidelberg University)

Melanie Schienle (Karlsruhe Institute of Technology)

July 18, 2018

The following supplementary materials contain proofs of Theorems 1-3 and additional simulations. All sections, equations, and tables referred to are those of the main paper.

A Proofs

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

Finiteness of $\mathbf{\Omega}$:

From Francq and Zakoïan (2004) it follows that $\mathbf{\Omega}_{\boldsymbol{\eta\eta}}$ 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. Recall from equation (19) 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 $\mathbf{E}[||\mathbf{r}_{0,t}^{\infty}(\mathbf{r}_{0,t}^{\infty})'||] < \infty$ (throughout the paper $||\cdot||$ denotes the Euclidean norm). A typical element of the $K \times 1$ vector $\mathbf{r}_{0,t}^{\infty}$ is given by

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

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

$$\begin{aligned}
\left(\mathbf{E}\left|\frac{\sum_{j=0}^{\infty} \alpha_0 \beta_0^j \varepsilon_{t-1-j}^2 x_{t-1-k-j}}{h_{0,t}^{\infty}}\right|^2\right)^{1/2} &\leq \left(\sum_{j=0}^{\infty} \mathbf{E}\left|\frac{\alpha_0 \beta_0^j \varepsilon_{t-1-j}^2}{h_{0,t}^{\infty}} x_{t-1-k-j}\right|^2\right)^{1/2} && (45)\\
&\leq \sum_{j=0}^{\infty} \left(\mathbf{E}\left|\frac{\alpha_0 \beta_0^j \varepsilon_{t-1-j}^2}{\left(\omega_0 + \alpha_0 \beta_0^j \varepsilon_{t-1-j}^2\right)} x_{t-1-k-j}\right|^2\right)^{1/2} && (46)\\
&\leq \sum_{j=0}^{\infty} \left(\mathbf{E}\left|\left(\frac{\alpha_0 \beta_0^j}{\omega_0} \varepsilon_{t-1-j}^2\right)^s x_{t-1-k-j}\right|^2\right)^{1/2} && (47)\\
&\leq \frac{\alpha_0^s}{\omega_0^s} \left(\mathbf{E}\left[|\varepsilon_{t-1-j}|^{4sp}\right]\right)^{1/(2p)} \left(\mathbf{E}\left[|x_{t-1-k-j}|^{2q}\right]\right)^{1/(2q)} \\
&\qquad \sum_{j=0}^{\infty} \beta_0^{js} < \infty
\end{aligned}$$

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

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

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

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

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

$$0 = \mathbf{c}'\begin{pmatrix} (h^{\infty}_{0,t})^{-1}\mathbf{s}^{\infty}_{0,t} \\ f'_0\mathbf{x}_t \end{pmatrix} + \mathbf{c}'\begin{pmatrix} (h^{\infty}_{0,t})^{-1}\beta_0 \frac{\partial \bar{h}^{\infty}_{t-1}}{\partial \boldsymbol{\eta}}\Big|_{\boldsymbol{\pi}=\mathbf{0}} \\ -f'_0\alpha_0(h^{\infty}_{0,t})^{-1}\varepsilon^2_{t-1}\mathbf{x}_{t-1} + f'_0(h^{\infty}_{0,t})^{-1}\beta_0\frac{\partial \bar{h}^{\infty}_{t-1}}{\partial \boldsymbol{\pi}}\Big|_{\boldsymbol{\pi}=\mathbf{0}} \end{pmatrix}. \quad (48)$$

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

$$\mathbf{c}'_1\mathbf{s}^{\infty}_{0,t} + f'_0 h^{\infty}_{0,t}\mathbf{c}'_2\mathbf{x}_t - f'_0\alpha_0\varepsilon^2_{t-1}(\mathbf{c}'_2\mathbf{x}_{t-1}) = -\left(\beta_0\mathbf{c}'_1\frac{\partial \bar{h}^{\infty}_{t-1}}{\partial \boldsymbol{\eta}}\Big|_{\boldsymbol{\pi}=\mathbf{0}} + f'_0\beta_0\mathbf{c}'_2\frac{\partial \bar{h}^{\infty}_{t-1}}{\partial \boldsymbol{\pi}}\Big|_{\boldsymbol{\pi}=\mathbf{0}}\right) \quad (49)$$

or

$$\begin{aligned} c_{11} + c_{12}Z^2_{t-1}h^{\infty}_{0,t-1} + f'_0(\omega_0 + \alpha_0 Z^2_{t-1}h^{\infty}_{0,t-1} + \beta_0 h^{\infty}_{0,t-2})(\mathbf{c}'_2\mathbf{x}_t) - f'_0\alpha_0 Z^2_{t-1}h^{\infty}_{0,t-1}(\mathbf{c}'_2\mathbf{x}_{t-1}) \\ = -c_{13}h^{\infty}_{0,t-1} - \left(\beta_0\mathbf{c}'_1\frac{\partial \bar{h}^{\infty}_{t-1}}{\partial \boldsymbol{\eta}}\Big|_{\boldsymbol{\pi}=\mathbf{0}} + f'_0\beta_0\mathbf{c}'_2\frac{\partial \bar{h}^{\infty}_{t-1}}{\partial \boldsymbol{\pi}}\Big|_{\boldsymbol{\pi}=\mathbf{0}}\right) = F_{t-2}, \end{aligned}$$

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

$$Z^2_{t-1} = \frac{-c_{11} + f'_0(\omega_0 + \beta_0 h^{\infty}_{0,t-2})(\mathbf{c}'_2\mathbf{x}_t)}{h^{\infty}_{0,t-1}(c_{12} - f'_0\alpha_0(\mathbf{c}'_2\mathbf{x}_{t-1}) + f'_0\alpha_0)} = A_{t-2} + B_{t-2}(\mathbf{c}'_2\mathbf{x}_t)$$

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

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

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

∎

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

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

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

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

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

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

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

Using equation (21) we obtain

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

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

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

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

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

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

Hence, using the Cauchy-Schwarz inequality, the first summand in equation (51), 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

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

and

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

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

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

and Assumption 7 applies again.

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

$$\begin{aligned}(f_0')^{-1}\frac{\partial\mathbf{r}_t^\infty}{\partial\boldsymbol{\eta}'} &= \frac{\partial}{\partial\boldsymbol{\eta}'}\mathbf{x}_t-\frac{\partial}{\partial\boldsymbol{\eta}'}\left(\frac{1}{h_t^\infty}\sum_{j=0}^{\infty}\alpha\beta^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j}\right) \\ &= \frac{\partial}{\partial\boldsymbol{\eta}'}\mathbf{x}_t-\frac{1}{h_t^\infty}\left(\sum_{j=0}^{\infty}\alpha\beta^j\varepsilon_{t-1-j}^2\frac{\partial}{\partial\boldsymbol{\eta}'}\mathbf{x}_{t-1-j}\right) \\ &\quad+\left(\frac{1}{h_t^\infty}\sum_{j=0}^{\infty}\alpha\beta^j\varepsilon_{t-1-j}^2\mathbf{x}_{t-1-j}\right)(\mathbf{y}_t^\infty)' \\ &\quad-\frac{1}{h_t^\infty}\sum_{j=0}^{\infty}\mathbf{x}_{t-1-j}\left(\frac{\partial}{\partial\boldsymbol{\eta}'}\alpha\beta^j\varepsilon_{t-1-j}^2\right). \end{aligned} \tag{56}$$

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

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

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

The fourth term can be written as:

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

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

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

and

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

by similar arguments as used before.

∎

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

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

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

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

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

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

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

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

Combining equation (60) and Proposition 1 leads to

$$\begin{aligned}
\sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\hat{\boldsymbol{\eta}}) &= \sqrt{T}\mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\boldsymbol{\eta}_0) - \mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\sqrt{T}\mathbf{D}_{\boldsymbol{\eta}}^{\infty}(\boldsymbol{\eta}_0) + o_P(1) && (62)\\
&= [-\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1} \ : \ \mathbf{I}]\sqrt{T}\begin{pmatrix}\mathbf{D}_{\boldsymbol{\eta}}^{\infty}(\boldsymbol{\eta}_0)\\ \mathbf{D}_{\boldsymbol{\pi}}^{\infty}(\boldsymbol{\eta}_0)\end{pmatrix} + o_P(1) && (63)\\
&= [-\mathbf{J}_{\boldsymbol{\pi}\boldsymbol{\eta}}\mathbf{J}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1} \ : \ \mathbf{I}]\sqrt{T}\mathbf{D}^{\infty}(\boldsymbol{\eta}_0) + o_P(1). && (64)
\end{aligned}$$

Applying Theorem 1 gives the asymptotic distribution

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

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

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

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

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

∎

Proof of Theorem 3. We show that

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

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

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

Since

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

we establish equation (67) by showing that

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

Consider the following decomposition:

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

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

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

First, note that

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

Next, consider a typical element:

$$
\begin{aligned}
(f_0')^{-1}\left(\mathbf{E}\sup_{\boldsymbol{\eta}}|r_{k,t}^{\infty}-r_{k,t}|^2\right)^{1/2} &= \left(\mathbf{E}\sup_{\boldsymbol{\eta}}\left|\alpha\frac{1}{h_t^{\infty}}\sum_{j=t}^{\infty}\beta^j\varepsilon_{t-1-j}^2 x_{t-1-k-j}\right|^2\right)^{1/2} \\
&\leq \sum_{j=t}^{\infty}\left(\mathbf{E}\sup_{\boldsymbol{\eta}}\left|\frac{\alpha\beta^j\varepsilon_{t-1-j}^2}{\omega+\alpha\beta^j\varepsilon_{t-1-k-j}^2}x_{t-1-k-j}\right|^2\right)^{1/2} \\
&\leq \sum_{j=t}^{\infty}\left(\mathbf{E}\sup_{\boldsymbol{\eta}}\left|\left(\frac{\alpha\beta^j}{\omega}\varepsilon_{t-1-j}^2\right)^s x_{t-1-k-j}\right|^2\right)^{1/2} \\
&\leq \left(\mathbf{E}[|\varepsilon_{t-1-j}|^{4sp}]\right)^{1/(2p)}\left(\mathbf{E}[|x_{t-1-k-j}|^{2q}]\right)^{1/(2q)} \\
&\qquad \sup_{\boldsymbol{\eta}}\left(\frac{\alpha}{\omega}\right)^s\sum_{j=t}^{\infty}\beta^{js} \qquad (71) \\
&= \left(\mathbf{E}[|\varepsilon_{t-1-j}|^{4sp}]\right)^{1/(2p)}\left(\mathbf{E}[|x_{t-1-k-j}|^{2q}]\right)^{1/(2q)} \\
&\qquad \sup_{\boldsymbol{\eta}}\left(\frac{\alpha}{\omega}\right)^s\frac{\beta^{st}}{1-\beta^s}, \qquad (72)
\end{aligned}
$$

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

Hence,

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

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

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

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

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

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

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

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

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

The third term can be treated as follows:

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

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

∎

B Mixed-Frequency LM Test

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

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

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

$$\mathbf{d}_t(\boldsymbol{\eta}_0) = \sum_{t=1}^{T} \left(\frac{\tilde{\varepsilon}_t^2}{M} - 1 \right) \begin{pmatrix} M^{-1} \\ f_0' \mathbf{x}_t \end{pmatrix}. \tag{77}$$

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

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

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

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

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

C Simulations

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

C.4 Persistence in the GARCH component.

In this section, we investigate the power properties of the LM test when $\hat{\alpha} + \hat{\beta}$ is close to or even above one. As in Appendix C.3, we first simulate an AR(1) process with autoregressive coefficient δ and i.i.d. normal innovations with mean zero and variance 0.025 and use the generated time series as the explanatory variable. We choose $\delta \in \{0.9, 0.98\}$ and impose an immediately decaying weighting scheme with $\boldsymbol{\pi}_0 = 0.3$ or $\boldsymbol{\pi}_0 = 0.4$. The GARCH component has either moderate (G-M), high (G-H) or extreme (G-E, $\alpha_0 = 0.095$, $\beta_0 = 0.90$) persistence. As before, we choose T and R as 1000.

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

Further, the table shows that for all specifications in which the GARCH component has moderate persistence, we never observe that the sum of the estimated GARCH parameters is greater than or equal to one. For example, in Panel B when $\alpha_0 = 0.07$, $\beta_0 = 0.90$ (G-M) and $\boldsymbol{\pi}_0 = 0.4$, the GARCH component is severely misspecified and the LM test rejects in 92.9% of the simulations, there is not a single simulation in which $\hat{\alpha} + \hat{\beta} \geq 1$.

The picture changes slightly when the persistence in the GARCH component is high (G-H). For this specification $\hat{\alpha} + \hat{\beta}$ is greater than or equal to one in 6 out of the 1000 simulations when $\delta = 0.9$ and in 11 ($\boldsymbol{\pi}_0 = 0.3$) or 19 ($\boldsymbol{\pi}_0 = 0.4$) cases when $\delta = 0.98$. However, for these cases the LM test has rejection rates which are (in all but one case) even higher than the average rejections rates over all 1000 simulations. For example, when $\delta = 0.98$ and $\boldsymbol{\pi}_0 = 0.4$, the LM test rejects in all 19 cases in which $\hat{\alpha} + \hat{\beta} \geq 1$.

Finally, when the persistence in the GARCH component is extreme (G-E), $\hat{\alpha} + \hat{\beta}$ is greater than or equal to one in 54 simulations when $\delta = 0.9$ and in 80 ($\boldsymbol{\pi}_0 = 0.3$) or 92 ($\boldsymbol{\pi}_0 = 0.4$) cases when $\delta = 0.98$. Nevertheless, we find that the LM test has very reasonable power and that the rejection frequency among those cases in which $\hat{\alpha} + \hat{\beta} \geq 1$ is typically higher than the average power.

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

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

Table 8: Size-adjusted power: persistent GARCH component

	Panel A: x_t is AR(1) with $\delta = 0.90$					
	G-M		G-H		G-E	
	$\alpha_0 = 0.07$, $\beta_0 = 0.90$		$\alpha_0 = 0.09$, $\beta_0 = 0.90$		$\alpha_0 = 0.095$, $\beta_0 = 0.90$	
	$\boldsymbol{\pi}_0 = 0.3$	$\boldsymbol{\pi}_0 = 0.4$	$\boldsymbol{\pi}_0 = 0.3$	$\boldsymbol{\pi}_0 = 0.4$	$\boldsymbol{\pi}_0 = 0.3$	$\boldsymbol{\pi}_0 = 0.4$
median($\hat{\alpha}$)	0.072	0.073	0.091	0.092	0.096	0.097
median($\hat{\beta}$)	0.890	0.888	0.894	0.893	0.894	0.893
median($\hat{\alpha} + \hat{\beta}$)	0.965	0.964	0.986	0.986	0.991	0.991
LM	38.7	62.1	37.0	60.8	36.2	59.7
VR	7.68	12.9	2.78	4.82	1.97	3.44
$\#(\hat{\alpha} + \hat{\beta} \geq 1)$	0	0	6	6	54	54
$\%reject \vert (\hat{\alpha} + \hat{\beta} \geq 1)$	-	-	50.0%	50.0%	33.3%	64.8%
	Panel B: x_t is AR(1) with $\delta = 0.98$					
	G-M		G-H		G-E	
	$\alpha_0 = 0.07$, $\beta_0 = 0.90$		$\alpha_0 = 0.09$, $\beta_0 = 0.90$		$\alpha_0 = 0.095$, $\beta_0 = 0.90$	
	$\boldsymbol{\pi}_0 = 0.3$	$\boldsymbol{\pi}_0 = 0.4$	$\boldsymbol{\pi}_0 = 0.3$	$\boldsymbol{\pi}_0 = 0.4$	$\boldsymbol{\pi}_0 = 0.3$	$\boldsymbol{\pi}_0 = 0.4$
median($\hat{\alpha}$)	0.076	0.081	0.095	0.099	0.100	0.103
median($\hat{\beta}$)	0.888	0.885	0.889	0.885	0.891	0.887
median($\hat{\alpha} + \hat{\beta}$)	0.968	0.970	0.986	0.986	0.991	0.992
LM	73.4	92.9	63.3	86.9	58.2	83.9
VR	26.7	39.4	11.3	18.6	8.32	14.1
$\#(\hat{\alpha} + \hat{\beta} \geq 1)$	0	0	11	19	80	92
$\%reject \vert (\hat{\alpha} + \hat{\beta} \geq 1)$	-	-	72.7%	100%	70.0%	91.3%

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

This observation is also confirmed by the plots in Figure 3. The figure shows a scatterplot of the estimate of the persistence $(\hat{\alpha} + \hat{\beta})$ in the GARCH component (x-axis) and the corresponding LM statistics (y-axis). The horizontal green line indicates the 5% critical value of the LM test and the vertical red line a persistence of one. In both plots the true GARCH component has extreme persistence (G-E), the AR(1) parameter of the explanatory variable is either $\delta = 0.90$ (left plot) or $\delta = 0.98$ (right plot) and $\boldsymbol{\pi}_0 = 0.4$. Again, the figure shows that there is no indication that the power of the test decreases when the estimated persistence is increasing.

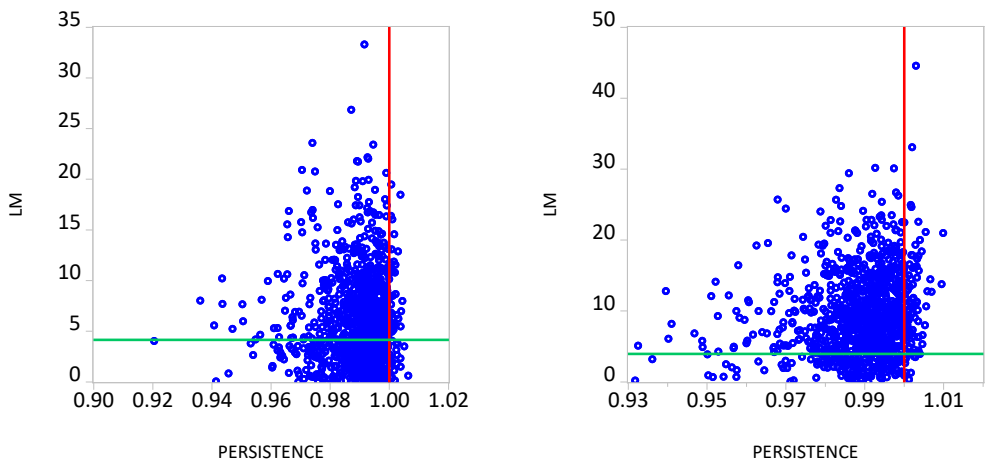

Figure 3: Scatterplot of estimated persistence $(\hat{\alpha} + \hat{\beta})$ and LM test statistics. Green horizontal line: critical value 5% level. Vertical red line: persistence of one. The true GARCH component has extreme persistence (G-E specification). We choose an immediately decaying weighting scheme with $\boldsymbol{\pi}_0 = 0.4$. The explanatory variable is an AR(1) process with autoregressive parameter $\delta = 0.90$ (left) and $\delta = 0.98$ (right).

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

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

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

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

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

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

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

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

C.6 Misspecification of the Short-Term Component.

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

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

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

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

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

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

Table 11: Misspecified short-term component.

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

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

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

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

C.7 Simulation Mixed-Frequency Test.

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

$$x_t = \delta x_{t-1} + \nu_t, \tag{81}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

D Correlations of Explanatory Variables

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

Table 14: Correlations between daily explanatory variables.

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

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

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

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

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

References (cited in the Appendix but not in the main paper)

- Billingsley, P., 1961. The Lindeberg-Levy theorem for martingales. *Proceedings of the American Mathematical Society* 12, 788-792.

- Conrad, C., Haag, B., 2006. Inequality constraints in the fractionally integrated GARCH model. *Journal of Financial Econometrics* 3, 413-449.
- Ling, S., McAleer, M., 2003. Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory* 19, 280-310.