Exponential-type GARCH models with linear-in-variance risk premium

Christian M. Hafner* and Dimitra Kyriakopoulou†

Université catholique de Louvain

December 13, 2019

Supplementary Material

6 The likelihood derivatives

Recall that the quasi log-likelihood function is given by $\overline{L}_T(\theta) = -\frac{1}{2}\frac{1}{T}\sum_{t=1}^{T} l_t(\theta)$, where

$$l_t(\theta) = \log(h_t(\theta)) + \frac{[Y_t - \lambda_1 - \lambda_2 h_t(\theta)]^2}{h_t(\theta)}. \tag{6.1}$$

Let $\varepsilon_t(\theta) := Y_t - \lambda_1 - \lambda_2 h_t(\theta)$. The first and second derivatives of $l_t(\theta)$ with respect to θ are given by

$$\frac{\partial l_t(\theta)}{\partial \theta} = \left(1 - \frac{\varepsilon_t^2(\theta)}{h_t(\theta)}\right)\frac{\partial \log(h_t(\theta))}{\partial \theta} + \frac{2\varepsilon_t(\theta)}{h_t(\theta)}\frac{\partial \varepsilon_t(\theta)}{\partial \theta},$$

*Institute of Statistics, Biostatistics and Actuarial Sciences (ISBA), Université catholique de Louvain, Voie du Roman Pays 20, B-1348 Louvain-la-Neuve, Belgium, e-mail: christian.hafner@uclouvain.be.

†Corresponding author. Center for Operations Research and Econometrics (CORE), Université catholique de Louvain, Voie du Roman Pays 34, Bte L1.03.01, B-1348 Louvain-la-Neuve, Belgium, e-mail: dimitra.kyriakopoulou@uclouvain.be.

and

$$
\begin{aligned}
\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta^\top} &= \frac{\varepsilon_t^2(\theta)}{h_t(\theta)}\frac{\partial\log(h_t(\theta))}{\partial\theta}\frac{\partial\log(h_t(\theta))}{\partial\theta^\top} - \frac{2\varepsilon_t(\theta)}{h_t(\theta)}\frac{\partial\log(h_t(\theta))}{\partial\theta}\frac{\partial\varepsilon_t(\theta)}{\partial\theta^\top} \\
&+\frac{2}{h_t(\theta)}\frac{\partial\varepsilon_t(\theta)}{\partial\theta}\frac{\partial\varepsilon_t(\theta)}{\partial\theta^\top} + \frac{2\varepsilon_t(\theta)}{h_t(\theta)}\frac{\partial^2\varepsilon_t(\theta)}{\partial\theta\partial\theta^\top} - \frac{2\varepsilon_t(\theta)}{h_t(\theta)}\frac{\partial\log(h_t(\theta))}{\partial\theta^\top}\frac{\partial\varepsilon_t(\theta)}{\partial\theta} \\
&+\left(1-\frac{\varepsilon_t^2(\theta)}{h_t(\theta)}\right)\frac{\partial^2\log(h_t(\theta))}{\partial\theta\partial\theta^\top},
\end{aligned}
$$

where, denoting e_{in} the i-th unit vector in $\mathbb{R}^n$ with $n = dim(\Theta)$,

$$
\begin{aligned}
\frac{\partial\varepsilon_t(\theta)}{\partial\theta} &= -e_{1n} - e_{2n}h_t(\theta) - \lambda_2\frac{\partial h_t(\theta)}{\partial\theta} \\
\frac{\partial^2\varepsilon_t(\theta)}{\partial\theta\partial\theta^\top} &= -\frac{\partial h_t(\theta)}{\partial\theta}e_{2n}^\top - e_{2n}\frac{\partial h_t(\theta)}{\partial\theta^\top} - \lambda_2\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta^\top}
\end{aligned}
$$

7 Simulation study

In order to investigate the small, moderate and large sample behavior of the QMLE in the EGARCH$(1,1)$-M and Log-GARCH$(1,1)$-M models, we perform the following simulation study. We generate time series of length $T = 500$, 1000, 2500, 5000 and 10000 observations and we tie the data generating process parameterization for both models to empirically relevant values (see Section 4.2), i.e. the true parameter vector for the EGARCH$(1,1)$-M model is given by $\theta_0 = (0, 0.04, -0.10, -0.12, 0.13, 0.98)^\top$, and for the Log-GARCH$(1,1)$-M model by $\theta_0 = (0, 0.07, 0.024, 0.027, 0.016, 0.971)^\top$. For the innovation process, we assume that Z_t has a Generalized Error Distribution (GED) with tail-thickness parameter ν given by either 2, which yields the Gaussian case, or 1.4, which produces fatter tails. The latter choice corresponds to typical estimates in empirical applications, see e.g. Hafner and Linton (2017) for the S&P 500 index.

The QML estimation is applied in a second step to obtain parameter estimates. We simulate 1000 paths and hence get 1000 independent replicates of $\widehat{\theta}_T$. In Figures 1 and 2 we depict the distributions of the vectors $\widehat{\theta}_T$ for the EGARCH$(1,1)$-M model, when the GED parameter ν is 2 and 1.4, respectively, by compiling boxplots for each estimated parameter. These numerical results illustrate the performance of the QMLE in finite samples. For both cases, the dispersion obviously shrinks as T increases, and the median parameter estimates are very close to the true parameter values even for small sample sizes. The corresponding boxplots for the Log-GARCH-M model are presented in Figures 3 and 4, respectively, and show a very similar behavior.

Tables 1 and 2 summarize the bias and RMSE performance of the QMLE of the EGARCH-M and Log-GARCH-M model parameters, respectively. As we can see from the results of Table 1, the bias is small even for small sample sizes and approaches zero as T increases. The RMSE drops quite fast for all estimated parameters as T gets larger, confirming consistency. While the bias is very similar for the Gaussian case and $\nu = 1.4$, the RMSE is slightly larger for most parameters in the case of fat tails. In terms of estimation precision, the parameter that is most difficult to estimate is the risk premium, as the RMSE is about 100% for $T = 500$, and for $T = 10000$ is still 19% and 23% for the Gaussian and GED(1.4) case, respectively.

As for the Log-GARCH-M results of Table 2, the results are similar to the EGARCH-M case, except that the volatility parameter estimates tend to have a higher RMSE, while the risk premium parameter has a lower RMSE across different sample sizes and distributions. This suggests that volatility estimates are more precise for EGARCH-M than log-GARCH-M, but that the reverse is true for the estimation of the risk premium.

Table 3 presents some parameter evaluations from each model, that is summary statistics of the estimated conditional standard deviations simulated by both the EGARCH-M and Log-GARCH-M models, under GED innovations and for the two cases, i.e. $\nu = 2$ and $\nu = 1.4$. As we can see, all statistics are very close for both models.

8 Proofs

Proof of Lemma 1

Following Bougerol (1993) and Straumann and Mikosch (2006), there is a unique stationary and ergodic solution if the function ϕ_t in (2.6) is contracting on average, i.e. if the Lyapunov exponent of the mapping is negative. The contraction condition of Bougerol (1993) can be used to ensure model invertibility and bounded moments for the filtered sequence, $\log \widehat{h}_t(\theta)$. We consider a new functional SRE that is driven by a function of the observations Y_t

$$s_{t+1} = [\Phi_t(s_t)](\theta), \tag{8.1}$$

where the sequence of random functions Φ_t are given by

$$[\Phi_t(s)](\theta) = \omega + [\gamma(Y_t - \lambda_1 - \lambda_2 \exp(s(\theta))) + \delta(|Y_t - \lambda_1 - \lambda_2 \exp(s(\theta))|)] \exp(-s(\theta)/2) + \beta s(\theta),$$

for each $\theta \in \Theta$. The functions Φ_t map continuous functions $s : \Theta \to \left[\inf_\Theta \omega (1-\beta)^{-1}, +\infty\right)$ onto the class of such functions. Because Φ_0 is Lipschitz continuous, and in general a Lipschitz continuous function is differentiable pointwise almost everywhere due to Rademacher's theorem (Evans and Gariepy, 2015, p. 103), continuous differentiability is not necessary to prove that its first derivative is bounded. The Lipschitz constant Λ, in this case, is given by the essential supremum of its derivative, ignoring any set of elements of Lebesgue measure zero where the derivative of the random functions Φ_t is not defined, which is the following

$$\left|\frac{\partial \Phi_0(s)}{\partial s}\right| = \left| \begin{array}{c} \beta - 2^{-1}[\gamma(Y_0 - \lambda_1 - \lambda_2 \exp(s(\theta))) + \delta(|Y_0 - \lambda_1 - \lambda_2 \exp(s(\theta))|)] \exp(-s(\theta)/2) \\ -\lambda_2 \exp(s(\theta))[\gamma + \delta \operatorname{sgn}(Y_0 - \lambda_1 - \lambda_2 \exp(s(\theta)))] \exp(-s(\theta)/2) \end{array} \right|.$$

Hence,

$$\Lambda = \sup_\Theta \left| \begin{array}{c} \beta - \frac{1}{2}[\gamma(Y_0 - \lambda_1 - \lambda_2 \exp(s(\theta))) + \delta(|Y_0 - \lambda_1 - \lambda_2 \exp(s(\theta))|)] \exp\left(-\frac{1}{2}\inf_\Theta \omega(1-\beta)^{-1}\right) \\ -\lambda_2[\gamma + \delta \operatorname{sgn}(Y_0 - \lambda_1 - \lambda_2 \exp(s(\theta)))] \exp\left(2^{-1}\inf_\Theta \omega(1-\beta)^{-1}\right) \end{array} \right|$$

The condition $E[\log \Lambda] < 0$, which implies continuous invertibility, is pointwise and reads

$$E[\log \max\{\beta, \Psi_0\}] < 0,$$

where

$$\begin{aligned} \Psi_0 &= 2^{-1}[\gamma\varepsilon_0(\theta) + \delta|\varepsilon_0(\theta)|] \exp\left(-2^{-1}\inf_\Theta \omega(1-\beta)^{-1}\right) \\ &\quad + \lambda_2[\gamma + \delta \operatorname{sgn}(\varepsilon_0(\theta))] \exp\left(2^{-1}\inf_\Theta \omega(1-\beta)^{-1}\right) - \beta, \end{aligned}$$

with $\varepsilon_0(\theta) := Y_0 - \lambda_1 - \lambda_2 h_0(\theta)$ and

$$\log h_0(\theta) = \omega + \gamma \frac{\varepsilon_{-1}(\theta)}{\sqrt{h_{-1}(\theta)}} + \delta \left|\frac{\varepsilon_{-1}(\theta)}{\sqrt{h_{-1}(\theta)}}\right| + \beta \log h_{-1}(\theta).$$

Since the uniform log moments exist by continuity of the Lipschitz coefficient in θ, having excluded the zero point of discontinuity, and because $E\left[\log^+(Y_t - \lambda_1 - \lambda_2 h_t(\theta))\right] < \infty$, together with the last result for each $\theta \in \Theta$ the condition (CI) for continuous invertibility of Wintenberger (2013) is satisfied.

The lemma follows from an application of Theorem 3.1 in Bougerol (1993), as its conditions are met since $E[\log \Lambda] < 0$ which implies $E\left[\log^{+} \Lambda\right] < \infty$ and also $E[\log + \|\Phi_t(s) - s\|] < \infty$ for a constant function s by the fact that Y_t has a bounded first moment. The invertibility condition is sufficient for

$$\sup_{\Theta} \left|\log \widehat{h}_t(\theta) - \log h_t(\theta)\right| \overset{e.a.s.}{\to} 0, \quad t \to \infty.$$

Also, since $\log h_t(\theta)$ is stationary and because

$$E(\log h_0) \leq (\omega + \delta)(1 - \beta)^{-1} < \infty,$$

this implies that

$$\sup_{\Theta} \left|\widehat{h}_t(\theta) - h_t(\theta)\right| \overset{e.a.s.}{\to} 0, \quad t \to \infty, \tag{8.2}$$

by the mean value theorem and Lemma 2.5.4. of Straumann (2005). □

Proof of Lemma 2

The proof follows the lines of Proposition 5.5.1 of Straumann (2005) and Theorem 8 of Wintenberger (2013), which essentially show the existence of a unique stationary solution to the stochastic recurrence equation (SRE), which is ergodic. We need similar contraction techniques as in the proof of Lemma 1. Equation (A.4) is a linear SRE, $\partial \log h_{t+1}(\theta)/\partial\theta = \Phi_t'[\partial \log h_t(\theta)/\partial\theta]$, where $\Phi_t'(s) := \partial\Phi_t(s)/\partial s$ is the stationary approximation of $\partial\widehat{\Phi}_t(s)/\partial s$. By Assumption 2, log-volatility is bounded from below, i.e. there exists a constant s^* such that $\inf_\Theta \log(h_t) \geq s^*$ a.s. Note also that except for a point with Lebesgue measure zero, both A_t and B_t are continuous functions in (s, θ) and that g_θ in (3.3) is differentiable pointwise almost everywhere. There exists a $\delta > 0$ such that

$$\left\|\widehat{\Phi}_t'(0) - \Phi_t'(0)\right\| = \left\|\widehat{B}_t - B_t\right\|_{\|\theta-\theta_0\|\leq\delta} \leq \left\|\sup_{s\geq s^*} \frac{\partial^2 g_\theta}{\partial s \partial\theta}\right\|_{\|\theta-\theta_0\|\leq\delta} \left\|\log \widehat{h}_t(\theta) - \log h_t(\theta)\right\|_{\|\theta-\theta_0\|\leq\delta},$$

where

$$\frac{\partial^2 g_\theta}{\partial\theta\partial s} = \begin{pmatrix} 2^{-1}[\gamma + \delta \operatorname{sgn}(g(s))] \exp(s/2) \\ -2^{-1}[\gamma + \delta \operatorname{sgn}(g(s))] \exp(s/2) \\ 0 \\ [-2^{-1}g(s) - \lambda_2 \exp(s)] \exp(-s/2) \\ \{[-2^{-1}g(s) - \lambda_2 \exp(s)] \exp(-s/2)\} \operatorname{sgn}(g(s)) \\ 1 \end{pmatrix},$$

where $g(s) = y - \lambda_1 - \lambda_2 \exp(s)$. Because

$$E\left[\log^+ \left\| \sup_{s \geq s^*} \frac{\partial^2 g_\theta}{\partial \theta \partial s} \right\| \right] < \infty,$$

as $E\left[\log^+ \left| Y_t - \lambda_1 - \lambda_2 \widehat{h}_t \right| \right] < \infty$ and $E\left[\log^+ \widehat{h}_t\right] < \infty$, then for any $|\rho| < 1$,

$$\sum_{t=0}^{\infty} \left(\rho^t \left\| \sup_{s \geq s^*} \frac{\partial^2 g_\theta}{\partial \theta \partial s} \right\| \right) < \infty,$$

by an application of the Borel-Cantelli lemma (see Lemma 2.2 of Berkes *et al.* 2003). This means that $\sum_{t=0}^{\infty} \left\| \sup_{s \geq s^*} \frac{\partial^2 g_\theta}{\partial \theta \partial s} \right\| \left\| \log \widehat{h}_t(\theta) - \log h_t(\theta) \right\|_{\|\theta - \theta_0\| \leq \delta}$ converges a.s. and hence

$$\left\| \widehat{\Phi}'_t(0) - \Phi'_t(0) \right\| \to 0 \quad e.a.s.$$

because $\left\| \log \widehat{h}_t(\theta) - \log h_t(\theta) \right\|_{\|\theta - \theta_0\| \leq \delta} \to 0$ e.a.s. by Lemma 1.

We also have

$$\Lambda\left(\widehat{\Phi}'_t - \Phi'_t\right) \leq \left\| \widehat{A}_t - A_t \right\|_{\|\theta - \theta_0\| \leq \delta} \leq \left\| \sup_{s \geq s^*} \frac{\partial^2 g_\theta}{\partial s^2} \right\|_{\|\theta - \theta_0\| \leq \delta} \left\| \log \widehat{h}_t(\theta) - \log h_t(\theta) \right\|_{\|\theta - \theta_0\| \leq \delta},$$

where

$$\begin{aligned} \frac{\partial^2 g_\theta}{\partial s^2} &= \left\{4^{-1}\left[\gamma g(s) + 2^{-1}\lambda_2 \exp(s) + \left(\delta g(s) + 2^{-1}\lambda_2 \exp(s)\right) \operatorname{sgn}(g(s))\right]\right\} \exp(-s/2) \\ &\quad + \left\{2^{-1}\lambda_2\left[\gamma + \delta \operatorname{sgn}(g(s))\right]\right\} \exp(s/2). \end{aligned}$$

Again, because

$$E\left[\log^+ \left\| \sup_{s \geq s^*} \frac{\partial^2 g_\theta}{\partial s^2} \right\| \right] < \infty,$$

as $E\left[\log^+ \left| Y_t - \lambda_1 - \lambda_2 \widehat{h}_t \right| \right] < \infty$, then $\Lambda\left(\widehat{\Phi}'_t - \Phi'_t\right) \to 0$ *e.a.s.* by an application of the Borel-Cantelli lemma as before, since we also have that

$$\left\| \log \widehat{h}_t(\theta) - \log h_t(\theta) \right\|_{\|\theta - \theta_0\| \leq \delta} \to 0 \quad e.a.s..$$

Now, because (i) the SRE (A.4) has a unique stationary ergodic solution as it is contractive by virtue of $E(\log \Lambda(\Phi'_t)) = E(\log \|A_t\|) < 0$ from Lemma 1, (ii) the filtered log-volatility process is invertible, and (iii) $\log h_t(\theta)$ is differentiable as a consequence of Proposition 5.5.1 of

Straumann (2005) and Theorem 8 of Wintenberger (2013), it follows that $\partial \log \widehat{h}_t / \partial \theta$ converges to the unique stationary ergodic solution uniformly locally in a neighborhood of θ_0, that is

$$\left\| \frac{\partial \log \widehat{h}_t}{\partial \theta} - \frac{\partial \log h_t}{\partial \theta} \right\|_{\|\theta - \theta_0\| \leq \delta} \overset{e.a.s.}{\rightarrow} 0, \quad t \rightarrow \infty.$$

$\square$

We now consider the second order derivatives and obtain a new linear SRE, that is

$$\frac{\partial^2 \log h_{t+1}(\theta)}{\partial \theta \partial \theta^\top} = A_t \frac{\partial^2 \log h_t(\theta)}{\partial \theta \partial \theta^\top} + C_t, \tag{8.3}$$

where $A_t := \frac{\partial^2 g_\theta}{\partial s \partial s^\top}$ and $C_t := \frac{\partial^2 g_\theta}{\partial \theta \partial \theta^\top} + 2 \frac{\partial^2 g_\theta}{\partial s \partial \theta} \frac{\partial \log h_t(\theta)}{\partial \theta}$. The next lemma implies the existence and uniqueness of a stationary and ergodic solution of the SRE (8.3).

Proof of Lemma 3

The proof is analogous to the proof of Lemma 2 for the first derivative, applying Proposition 5.5.1 of Straumann (2005) and Theorem 8 of Wintenberger (2013). Notice that the stationary solution $\partial^2 \log h_t(\theta) / \partial \theta \partial \theta^\top$ is continuous, as it is the locally uniform limit law of $\partial^2 \log \widehat{h}_t(\theta) / \partial \theta \partial \theta^\top$ that is continuous by definition. $\square$

The following two lemmas state that for the limiting processes of the derivatives of $\log \widehat{h}_t(\theta)$, the contraction condition is satisfied for their functional SRE restricted to $\mathcal{V}(\theta_0) := \{\theta : \|\theta - \theta_0\| \leq \delta\}$, having $\theta_T^* \in \mathcal{V}(\theta_0)$ for T sufficiently large.

Proof of Lemma 4

From (A.4) with $h_t := h_t(\theta_0)$ and letting $\theta = \theta_0$, we get

$$\begin{aligned} A_t|_{\theta_0} &= \beta_0 - \left[2^{-1} (\gamma_0 Z_t + \delta_0 |Z_t|) + \lambda_{02} [\gamma_0 + \delta_0 \operatorname{sgn}(Z_t)] \sqrt{h_t} \right], \quad \textit{and} \\ B_t|_{\theta_0} &= \left(-[\gamma_0 + \delta_0 \operatorname{sgn}(Z_t)] h_t^{-1/2}, -[\gamma_0 + \delta_0 \operatorname{sgn}(Z_t)] h_t^{1/2}, 1, Z_t, |Z_t|, \log h_t \right)^\top. \end{aligned}$$

Let $\nabla g_t(\theta) := \partial \log h_t(\theta) / \partial \theta$. The SRE (A.4) can be written as

$$\nabla g_{t+1}(\theta) = \Phi_t'(g_t(\theta), \theta) \nabla g_t(\theta) + \nabla_\theta \Phi_t(g_t(\theta), \theta), \ t \in \mathbb{Z}$$

where $\Phi_t'(g_t(\theta), \theta) =: A_t$ in (A.2) and $\nabla_\theta \Phi_t(g_t(\theta), \theta) =: B_t$ in (A.3).

Notice that we can rewrite (A.4), letting

$$\zeta_t(\theta) := \frac{\varepsilon_t(\theta)}{\sqrt{h_t(\theta)}} = \frac{\sqrt{h_t} Z_t + \lambda_{01} + \lambda_{02} h_t - \lambda_1 - \lambda_2 h_t(\theta)}{\sqrt{h_t(\theta)}}$$

with $\zeta_t(\theta_0) = Z_t$, as

$$\begin{aligned} A_t &= \beta - 2^{-1}\left[\gamma\zeta_t(\theta) + \delta\left|\zeta_t(\theta)\right|\right] + \lambda_2\left[\gamma + \delta \operatorname{sgn}\left(\zeta_t(\theta)\right)\right]\sqrt{h_t(\theta)} \\ B_t &= \left(-\left[\gamma + \delta \operatorname{sgn}\left(\zeta_t(\theta)\right)\right] h_t(\theta)^{-1/2}, -\left[\gamma + \delta \operatorname{sgn}\left(\zeta_t(\theta)\right)\right] h_t(\theta)^{1/2}, 1, \zeta_t(\theta), \left|\zeta_t(\theta)\right|, \log h_t(\theta)\right)^\top \end{aligned}$$

By similar arguments as in the proof of Lemma 3 in Wintenberger (2013) we can argue that there exists a positive random variable a such that

$$\left|A_t - A_t|_{\theta_0}\right| + \left\|B_t - B_t|_{\theta_0}\right\| \leq a\left(\|\theta - \theta_0\| + \left|\log h_t(\theta) - \log h_t(\theta_0)\right|\right),$$

that is, for any θ that belongs to a compact neighborhood of θ_0, $\mathcal{V}(\theta_0)$, we can apply a local continuity argument with respect to the parameters to the Lipschitz coefficients of the SRE (A.4).

Thus, for any sequence $\left(\widehat{\theta}_T\right)$ such that $\left\|\widehat{\theta}_T - \theta_0\right\| \overset{a.s.}{\to} 0$ as $T \to \infty$ and $\widehat{\theta}_T \in \mathcal{V}(\theta_0)$ for T sufficiently large, using the consistency result of Theorem 1, we have

$$\begin{aligned} \left\|\frac{\partial \log h_{t+1}\left(\widehat{\theta}_T\right)}{\partial\theta} - \frac{\partial \log h_{t+1}(\theta_0)}{\partial\theta}\right\| &= \left\|A_t|_{\widehat{\theta}_T}\frac{\partial \log h_t\left(\widehat{\theta}_T\right)}{\partial\theta} + B_t|_{\widehat{\theta}_T} - A_t|_{\theta_0}\frac{\partial \log h_t}{\partial\theta} - B_t|_{\theta_0}\right\| \\ &\leq \left|A_t|_{\theta_0}\right|\left\|\frac{\partial \log h_t\left(\widehat{\theta}_T\right)}{\partial\theta} - \frac{\partial \log h_t}{\partial\theta}\right\| \\ &\quad + \left\|\frac{\partial \log h_t\left(\widehat{\theta}_T\right)}{\partial\theta}\right\|\left|A_t|_{\widehat{\theta}_T} - A_t|_{\theta_0}\right| + \left\|B_t|_{\widehat{\theta}_T} - B_t|_{\theta_0}\right\| \\ &\leq \left|A_t|_{\theta_0}\right|\left\|\frac{\partial \log h_t\left(\widehat{\theta}_T\right)}{\partial\theta} - \frac{\partial \log h_t}{\partial\theta}\right\| + a \times \\ &\quad \left(\left\|\frac{\partial \log h_t}{\partial\theta}\right\|_{\mathcal{V}(\theta_0)} + 1\right)\left(\left\|\widehat{\theta}_T - \theta_0\right\| + \left|\log h_t\left(\widehat{\theta}_T\right) - \log h_t\right|\right) \end{aligned}$$

Now $\|\partial \log h_t(\widehat{\theta}_T)/\partial\theta - \partial \log h_t(\theta_0)/\partial\theta\|$ is Césaro summable, i.e. its arithmetic mean of the first T partial sums tends to a limit as $T \to \infty$, and there exist a positive random variable a^* and a random continuous function b satisfying $b(\theta_0) = 0$ a.s.,

$$T^{-1}\sum_{t=1}^{T}\left\|\frac{\partial \log h_t\left(\widehat{\theta}_T\right)}{\partial\theta} - \frac{\partial \log h_t(\theta_0)}{\partial\theta}\right\| \leq a^*\left\|\widehat{\theta}_T - \theta_0\right\| + b\left(\widehat{\theta}_T\right).$$

Hence, for any sequence $\left(\widehat{\theta}_T\right)$ converging a.s. to θ_0,

$$P\left[\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T}\left\|\frac{\partial \log h_t\left(\widehat{\theta}_T\right)}{\partial\theta}-\frac{\partial \log h_t\left(\theta_0\right)}{\partial\theta}\right\|=0\right]=1$$

which implies that, as stated, $\left\|\frac{\partial \log h_t\left(\widehat{\theta}_T\right)}{\partial\theta}-\frac{\partial \log h_t(\theta_0)}{\partial\theta}\right\|=o_p\left(1\right).$ □

Proof of Lemma 5 By Lemma 4, taking the derivative in $\left\|\frac{\partial \log h_{t+1}\left(\widehat{\theta}_T\right)}{\partial\theta}-\frac{\partial \log h_{t+1}(\theta_0)}{\partial\theta}\right\|$, the second derivative of the log volatility is also Césaro summable, such that

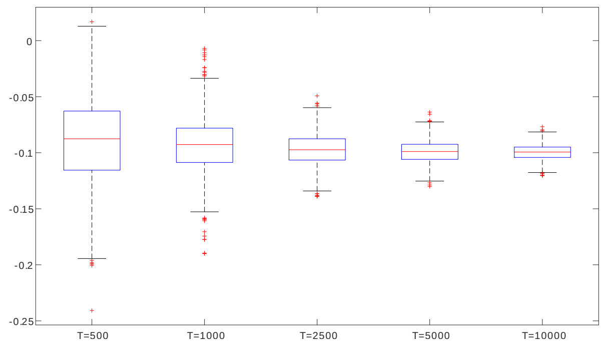
$$T^{-1}\sum_{t=1}^{T}\left\|\frac{\partial^2 \log h_t\left(\widehat{\theta}_T\right)}{\partial\theta\partial\theta^\top}-\frac{\partial^2 \log h_t\left(\theta_0\right)}{\partial\theta\partial\theta^\top}\right\|\leq a'\left\|\widehat{\theta}_T-\theta_0\right\|+b'\left(\widehat{\theta}_T\right),$$

with a' a positive random variable and b' a random continuous function which satisfies $b'\left(\theta_0\right)=0$ a.s. Then

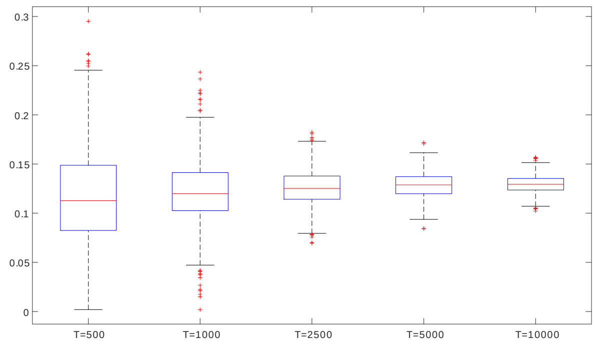
$$\lim_{n\to\infty}\sum_{t=I}^{n}\left\|\frac{\partial^2 \log h_t\left(\widehat{\theta}_T\right)}{\partial\theta\partial\theta^\top}-\frac{\partial^2 \log h_t\left(\theta_0\right)}{\partial\theta\partial\theta^\top}\right\|=0$$

a.s. for a random integer $I>1$. Due to the fact that $\partial^2 \log h_t/\partial\theta\partial\theta^\top$ is continuous in its arguments, for any sequence $\left(\widehat{\theta}_T\right)$ converging a.s. to θ_0, $\frac{\partial^2 \log h_t\left(\widehat{\theta}_T\right)}{\partial\theta\partial\theta^\top}-\frac{\partial^2 \log h_t(\theta_0)}{\partial\theta\partial\theta^\top}\to 0$ a.s. □

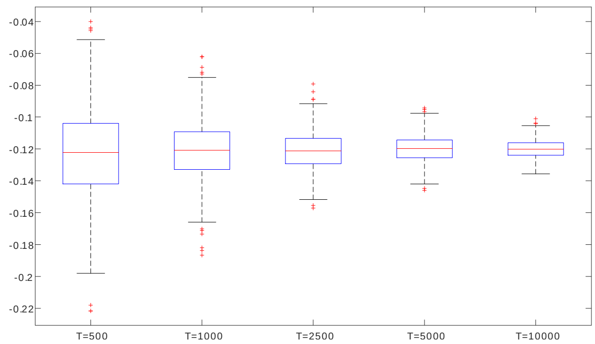
Figure 1: Boxplots of independent realizations of the QML estimators of the EGARCH(1,1)-M parameters with GED innovations.



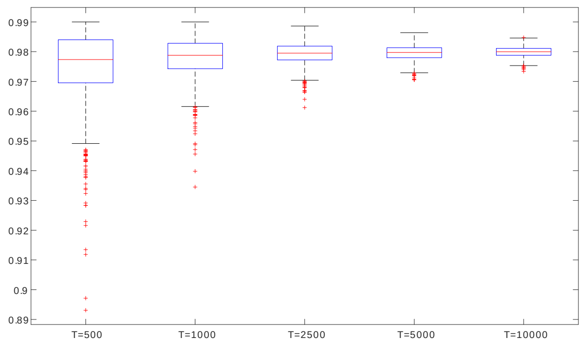
(a) constant parameter ω



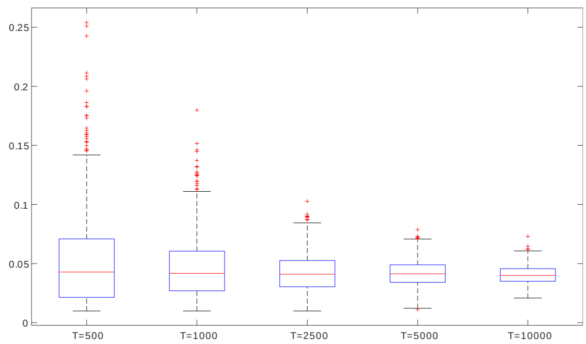
(b) size parameter δ



(c) leverage parameter γ



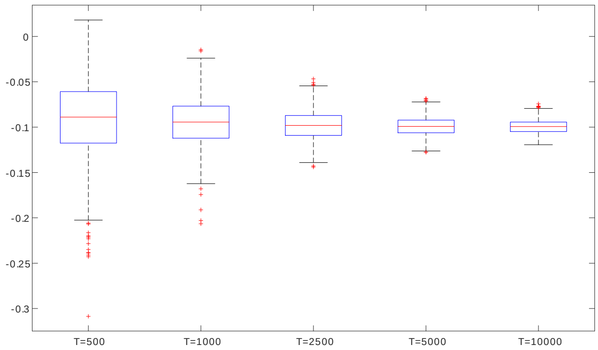
(d) GARCH parameter β



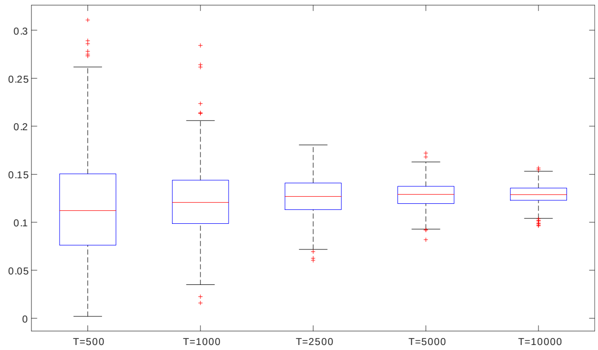
(e) risk premium parameter λ_2

Notes: GED errors with $\nu = 2$. Various sample sizes are compared. The red line represents the median.

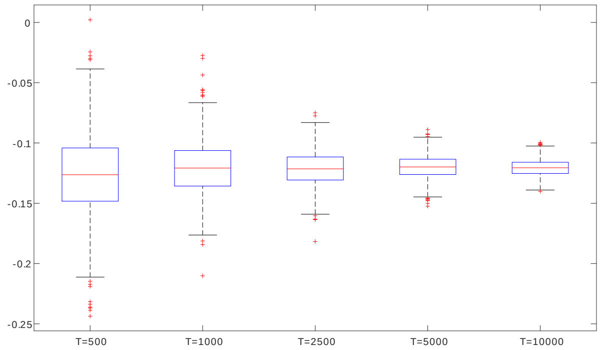
Figure 2: Boxplots of independent realizations of the QML estimators of the EGARCH(1,1)-M parameters with GED innovations.



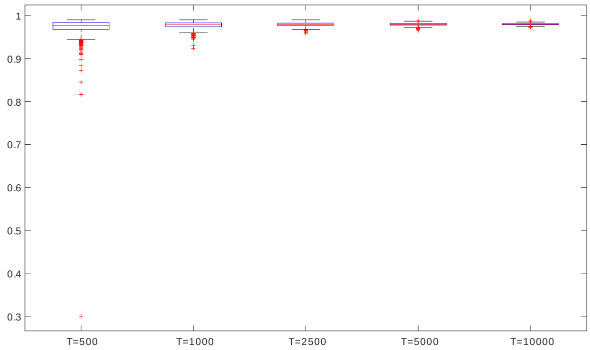
(a) constant parameter ω



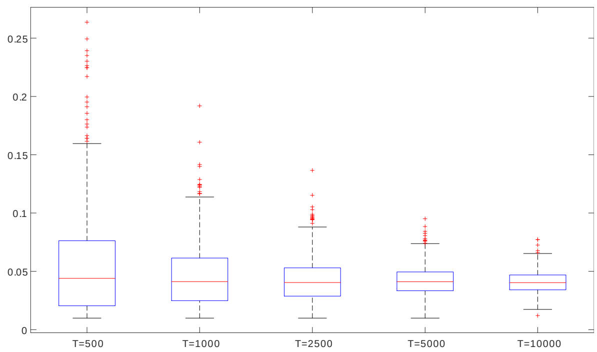
(b) size parameter δ



(c) leverage parameter γ



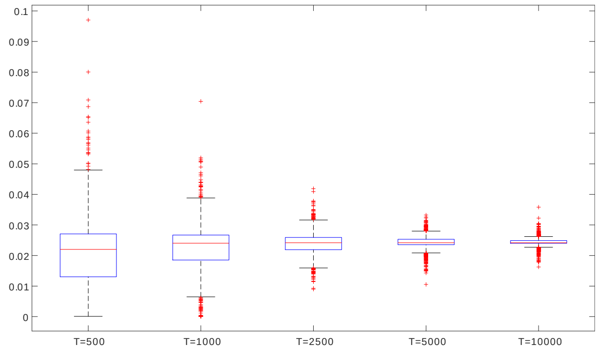
(d) GARCH parameter β



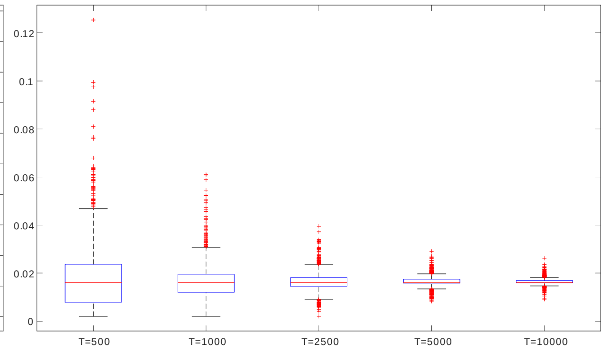
(e) risk premium parameter λ_2

Notes: GED errors with $\nu = 1.4$. Various sample sizes are compared. The red line represents the median.

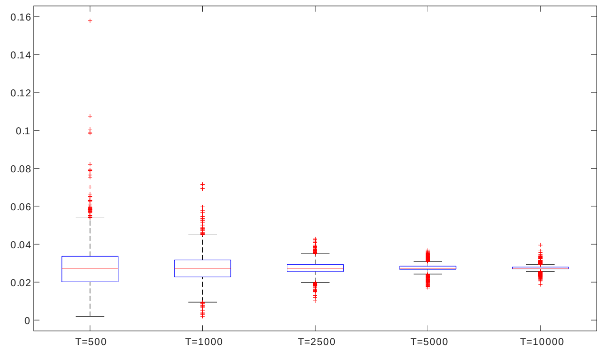
Figure 3: Boxplots of independent realizations of the QML estimators of the Log-GARCH(1,1)-M parameters with GED innovations.



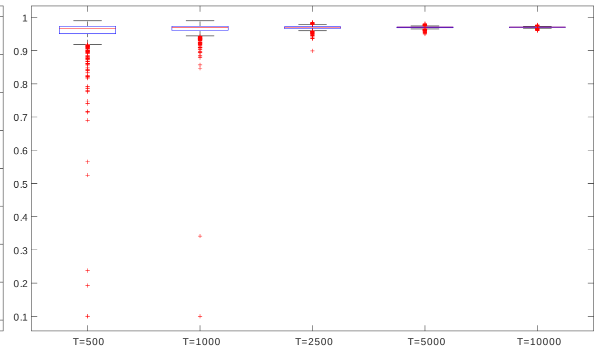
(a) constant parameter ω



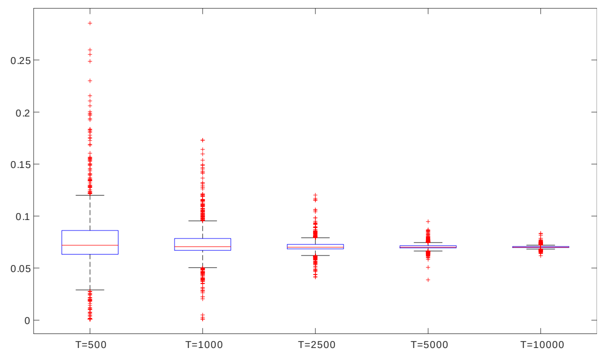
(b) leverage parameter α_{-}



(c) leverage parameter α_{+}



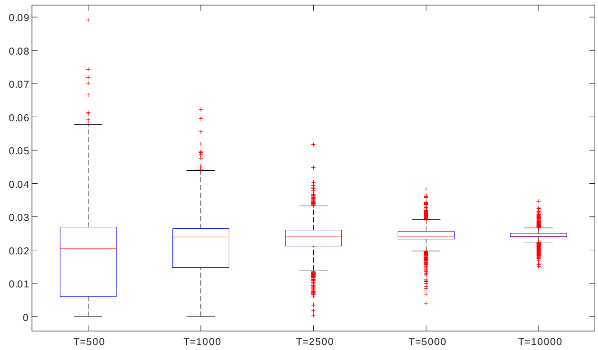
(d) GARCH parameter β



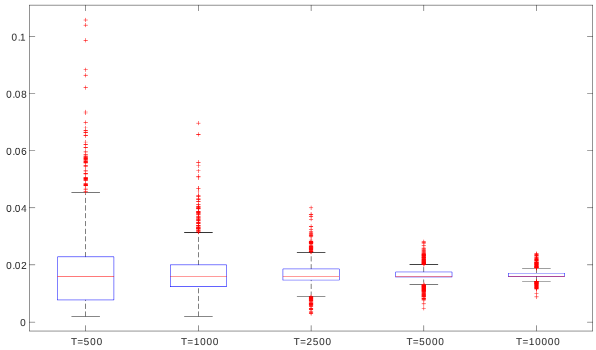
(e) risk premium parameter λ_2

Notes: GED errors with $\nu = 2$. Various sample sizes are compared. The red line represents the median.

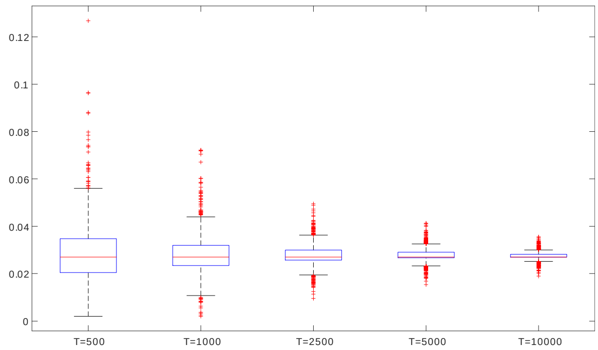
Figure 4: Boxplots of independent realizations of the QML estimators of the Log-GARCH(1,1)-M parameters with GED innovations.



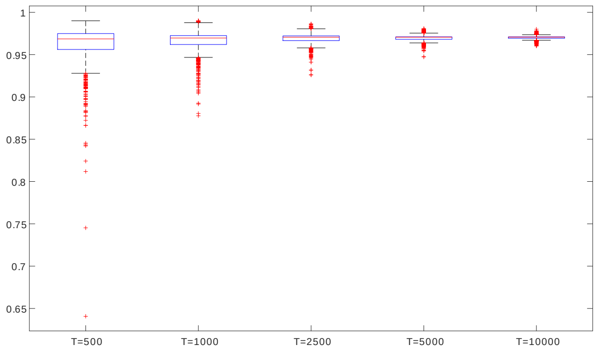
(a) constant parameter ω



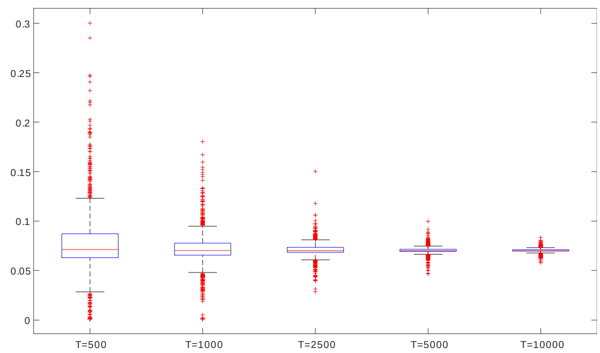
(b) leverage parameter α_-



(c) leverage parameter α_+



(d) GARCH parameter β



(e) risk premium parameter λ_2

Notes: GED errors with $\nu = 1.4$. Various sample sizes are compared. The red line represents the median.

Table 1: Bias and RMSE in % of the true parameter values: key characteristics of the finite sample distribution of the QMLE for the EGARCH(1,1)-M model, for various sample sizes and two types of GED errors.

EGARCH(1,1)-M						
$\nu = 2$ (Gaussian)						
	T	$\widehat{\omega}$	$\widehat{\gamma}$	$\widehat{\delta}$	$\widehat{\beta}$	$\widehat{\lambda_2}$
Bias	500	-0.129	0.0267	-0.1308	-0.0050	0.2875
	1000	-0.062	0.0100	-0.0646	-0.0021	0.1375
	2500	-0.028	0.0108	-0.0292	-0.0007	0.0550
	5000	-0.011	-0.0025	-0.0123	-0.0004	0.0425
	10000	-0.005	0.0008	-0.0054	-0.0001	0.0125
RMSE	500	0.437	0.2317	0.4254	0.0139	1.0025
	1000	0.253	0.1542	0.2485	0.0076	0.6550
	2500	0.147	0.0983	0.1423	0.0039	0.4100
	5000	0.01	0.0675	0.0977	0.0027	0.2825
	10000	0.071	0.0483	0.0692	0.0018	0.1900
$\nu = 1.4$						
	T	$\widehat{\omega}$	$\widehat{\gamma}$	$\widehat{\delta}$	$\widehat{\beta}$	$\widehat{\lambda_2}$
Bias	500	-0.099	0.0542	-0.1338	-0.0074	0.3450
	1000	-0.053	0.0083	-0.0677	-0.0024	0.1350
	2500	-0.02	0.0117	-0.0238	-0.0008	0.0450
	5000	-0.008	-0.0008	-0.0115	-0.0005	0.0450
	10000	-0.006	0.0050	-0.0069	-0.0002	0.0175
RMSE	500	0.473	0.2875	0.4700	0.0287	1.1050
	1000	0.273	0.1867	0.2738	0.0083	0.6900
	2500	0.159	0.1175	0.1600	0.0046	0.4625
	5000	0.102	0.0800	0.1023	0.0031	0.3150
	10000	0.076	0.0583	0.0762	0.0021	0.2325

Notes: The number of simulations is 1000. The true values of the parameters are: $\omega = -0.10, \gamma = -0.12, \delta = 0.13, \beta = 0.98, \lambda_2 = 0.04$. We also assume that $\lambda_1 = 0$.

Table 2: Bias and RMSE in % of the true parameter values: key characteristics of the finite sample distribution of the QMLE for the Log-GARCH-M model, for various sample sizes and two types of GED errors.

Log-GARCH(1, 1)-M						
$\nu = 2$ (Gaussian)						
	T	$\widehat{\omega}$	$\widehat{\alpha_-}$	$\widehat{\alpha_+}$	$\widehat{\beta}$	$\widehat{\lambda_2}$
Bias	500	−0.1250	0.1438	0.0333	−0.0194	0.0986
	1000	−0.0542	0.0250	0.0148	−0.0075	0.0500
	2500	−0.0042	0.0313	0.0074	−0.0020	0.0114
	5000	0.0083	0.0313	0.0111	−0.0012	0.0071
	10000	0.0125	0.0250	0.0111	−0.0008	0.0029
RMSE	500	0.5333	0.9313	0.5296	0.0683	0.4843
	1000	0.3417	0.5250	0.3037	0.0388	0.2614
	2500	0.1708	0.2813	0.1556	0.0067	0.0986
	5000	0.1042	0.1563	0.1037	0.0038	0.0500
	10000	0.0667	0.1000	0.0667	0.0023	0.0243
$\nu = 1.4$						
	T	$\widehat{\omega}$	$\widehat{\alpha_-}$	$\widehat{\alpha_+}$	$\widehat{\beta}$	$\widehat{\lambda_2}$
Bias	500	−0.2167	0.1438	0.0519	−0.0101	0.0986
	1000	−0.1125	0.0625	0.0407	−0.0053	0.0229
	2500	−0.0125	0.0438	0.0259	−0.0021	0.0114
	5000	0.0083	0.0313	0.0259	−0.0013	0.0057
	10000	0.0125	0.0313	0.0148	−0.0008	0.0029
RMSE	500	0.6167	0.9375	0.5074	0.0234	0.5200
	1000	0.4333	0.5500	0.3296	0.0153	0.2729
	2500	0.2333	0.2938	0.1815	0.0072	0.1157
	5000	0.1500	0.1813	0.1222	0.0041	0.0586
	10000	0.0875	0.1188	0.0741	0.0026	0.0329

Notes: The number of simulations is 1000. The true values of the parameters are: $\omega = 0.024, \alpha_- = 0.016, \alpha_+ = 0.027, \beta = 0.971, \lambda_2 = 0.07$. We also assume that $\lambda_1 = 0$.

References

[1] Berkes, I., Horváth, L., Kokoszka, P. (2003). GARCH processes: structure and estimation. *Bernoulli* 9: 201–227.

[2] Bougerol, P. (1993). Kalman filtering with random coefficients and contractions. *SIAM Journal on Control and Optimization* 31: 942-959.

[3] Evans, L. C., Gariepy, R.F. (2015). *Measure theory and fine properties of functions.* Studies in Advanced Mathematics, CRC Press, Revised Edition.

[4] Hafner, C. M., Linton, O. (2017). An almost closed form estimator for the EGARCH model. *Econometric Theory* 33: 1013-1038.

[5] Straumann, D. (2005). *Estimation in Conditionally Heteroskedastic Time Series Models.* Springer.

[6] Straumann, D., Mikosch, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroskedastic time series: a stochastic recurrence equations approach. *Annals of Statistics* 34: 2449-2495.

[7] Wintenberger, O. (2013). Continuous invertibility and stable QML estimation of the EGARCH(1,1) model. *Scandinavian Journal of Statistics* 40: 846-867.

Table 3: Summary statistics of the estimated conditional standard deviation for simulated EGARCH-M and Log-GARCH-M models.

		Mean	Min.	Max.	Median	St.Dev.
EGARCH-M	$\nu = 2$	0.148	0.122	1.007	0.126	0.091
	$\nu = 1.4$	0.152	0.109	0.841	0.144	0.053
Log-GARCH-M	$\nu = 2$	0.149	0.124	1.007	0.128	0.090
	$\nu = 1.4$	0.069	0.048	0.715	0.062	0.052