Improving financial volatility nowcasts*

Robinson Kruse-Becher[a] and Yuze Liu[b]

[a]University of Hagen and CREATES, Aarhus University†
[b]University of Hagen‡

September 30, 2022

Abstract

Recently, a simple nowcasting model for volatility has been proposed by Breitung and Hafner (2016). They suggest a model in which today's volatility is not only driven by past returns, but also by the current information from the same day. Empirical results demonstrate the relevance of the current squared return for volatility nowcasting. Their model obeys an ARMA representation estimable by maximum likelihood. However, their estimation approach builds on a number of simplifications and we suggest improvements. Rather than assuming normality of the innovations in the ARMA representation for highly skewed and leptokurtic log-squared returns, we take non-normality explicitly into account. Contrary to most situations regarding volatility estimation and forecasting, the distribution actually plays a crucial role in the construction of volatility nowcasts. We devise a one-step exact maximum likelihood estimator which offers significant improvements in estimation efficiency and volatility nowcast accuracy in finite samples. In our empirical application, we investigate five major international stock markets from 2000 to 2019 (including sub-samples relating to the Great Financial Crisis). The results suggest that our estimation approach significantly outperforms the one by Breitung and Hafner (2016) in all cases. Financial volatility can be nowcasted more accurately by applying our suggested estimation approach.

Keywords: Volatility, Nowcasting, Distribution, Estimation, Evaluation

JEL classification: C13, C22, C58

*The authors would like to thank Joerg Breitung, Nalan Bastürk, Christian Conrad, Roxana Halbleib, Roman Liesenfeld, Esther Ruiz, Genaro Sucarrat, Marco Kerkemeier, Eva Amon and the participants of the Computational and Financial Econometrics Conference in Pisa 2018, the Quantitative Finance and Financial Econometrics conference 2018 in Marseille, the Nordic Econometric Meeting in Stockholm 2019 and the Workshop in Time Series Econometrics in Zaragoza 2022 for their helpful comments and inspiring discussions.

†University of Hagen, Faculty of Business Administration and Economics, Universitätsstr. 41, 58097 Hagen, Germany, e-mail address: `robinson.kruse-becher@fernuni-hagen.de` and CREATES, Aarhus University, School of Economics and Management, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark.

‡**Corresponding author:** University of Hagen, Faculty of Business Administration and Economics, Universitätsstr. 41, 58097 Hagen, Germany, e-mail address: `yuze.liu@fernuni-hagen.de`

1 Introduction

Financial volatility is key in a wide range of theoretical and applied fields including risk management, portfolio choice, asset and option pricing. The voluminous literature on volatility modelling and forecasting has experienced many major breakthroughs in the last decades including the (generalized) autoregressive conditional heteroskedasticity model (Engle, 1982 and Bollerslev, 1986), the stochastic volatility model (Taylor, 1986), and the realized and implied volatility measures (Andersen and Bollerslev, 1998 and Latane and Rendleman, 1976).

ARCH-type volatility models based on daily returns are indisputable the leading workhorse in empirical studies and very often used as a benchmark, see e.g. Hansen and Lunde (2005). In these models conditional variances are driven by its lags and past returns. Breitung and Hafner (2016) emphasize that the current (squared) return is not included in the information set when estimating the volatility even though it contains presumably the most important information, see also Politis (2007). Breitung and Hafner (2016) make an important contribution to the literature by suggesting a simple model which includes the currently observed return and thus enables so-called volatility nowcasts.[1] Quick changes in volatility are typically better captured by including the contemporaneous return information. A similar idea was picked up by Smetanina (2017) and introduced as Real-Time GARCH (RT-GARCH). Ding (2021) studies the weak diffusion limits of these two approaches and finds that only the one by Breitung and Hafner (2016) converges to a meaningful process, i.e. the same limit to which the exponential GARCH by Nelson (1991) converges.

Breitung and Hafner (2016) demonstrate the relevance of the current observation for volatility nowcasting in an application to daily S&P 500 returns from 1950 to 2012. Their suggested simple structural nowcasting model obeys an reduced-form ARMA representation (see also Ruiz, 1994, Asai, 1998 and Francq and Sucarrat, 2018) for the log-squared returns. Thereby, the effect of the contemporaneous observation can be estimated from reduced-form ARMA parameters. The existence of the ARMA representation is also beneficial for a comparison to the log-GARCH model independently proposed by Pantula (1986) and Geweke (1986) which also admits an ARMA representation for log-squared returns, but excludes the current return from the information set. For this model, Francq and Sucarrat (2018) suggest a maximum likelihood estimator which takes the

[1]The term "nowcasting" is borrowed from the macroeconomic literature and probably somewhat confusing in the financial context. The basic idea of volatility nowcasting is to have a volatility model in which the current squared return enters. In this sense, the volatility nowcasts enable a contemporaneous news effect.

distributional characteristics of log-squared return innovations explicitly into account. Asymptotic results on strong consistency and asymptotic normality for the estimator are provided in Francq and Sucarrat (2018). Their simulation results clearly indicate efficiency gains which certainly outweigh the slightly increased computational efforts.

In the estimation part, Breitung and Hafner (2016) make some simplifications which motivates our suggested modifications leading to improvements in parameter estimation efficiency and volatility nowcasting accuracy. Contrary to most volatility estimation and forecasting situations (see e.g. Corsi et al., 2008), in the particular case considered here, the distribution actually matters for the construction of volatility nowcasts. In our work, we follow Breitung and Hafner (2016) and combine it with the maximum likelihood estimation approach of Francq and Sucarrat (2018) to explicitly account for the inherent non-normality of log-squared return innovations. Due to a large moving average root, we do use stationary initial recursion values obtained from a nonlinear and non-Gaussian state space filter and thus devise an exact maximum likelihood estimator. It improves estimation efficiency of the parameters involved in the ARMA representation and provides more accurate volatility nowcasts. The following points address differences and similarities to existing approaches.

First, standardized log-squared returns are highly negatively skewed and leptokurtic and thus far from normality. Even under normality of returns, the consequential distribution would be a standardized logarithmic χ^2-distribution which strongly deviates from normality. Under fat-tailedness of returns, as commonly observed as a stylized fact in financial markets (Cont, 2001), say via a $t(\nu)$-distribution, log-squared standardized returns would follow a standardized logarithmic $F(1, \nu)$-distribution. Simulation results in Francq and Sucarrat (2018) are encouraging and motivate us to account for non-normality in the context of Breitung's and Hafner's volatility nowcasting model in a similar fashion. We emphasize that the distribution of the reduced form ARMA innovations is directly and uniquely determined by the distribution of the structural innovations. It is thus only required to specify a distribution for the structural innovations as the reduced form ARMA innovation distribution is then determined by the log-squared transformation automatically. Given that symmetry is a requirement in this model framework[2], we restrict our attention to the class of symmetric location-scale distributions for the structural innovations. We thus show how to obtain the distribution for the log-squared transformation needed for the exact maximum likelihood

[2]The distribution of the structural innovations needs to be symmetric such that the returns follow a martingale difference sequence, see Breitung and Hafner (2016). Such a property implies that the reduced form innovations are serially uncorrelated.

estimation of reduced form ARMA model in general and work out two leading examples in this work. Besides normality, we provide the leading case of a t-distribution (see e.g. Bollerslev, 1987) and thus derive estimators based on the $\log - \chi^2$- and $\log - F$-distributions. We thus compare the results of different distributions and also explicitly state which distribution suits best for the structural innovations. For this issue, we refer to the selection of densities as in Marin and Sucarrat (2015) and use information criteria. The authors find that the Schwarz criterion performs very well, especially for the relatively large sample sizes we consider. This procedure addresses the question how to select the best suited distribution for the structural innovations. Thereby, the distribution of reduced form innovations is automatically given by the log-squared transformation.

Second, we also improve upon estimation of the expected value of standardized log-squared returns which plays a crucial role in the construction of volatility nowcasts via a mean-adjusting localization parameter. Due to an identification problem of the intercept in the ARMA representation for log-squared returns, estimation is carried out for a mean-corrected ARMA representation, see also Breitung and Hafner (2016) and Francq and Sucarrat (2018). The construction of volatility nowcasts require the inverse operation of adding an appropriate localization parameter which hinges also on the expected value of standardized log-squared returns. Breitung and Hafner (2016) use a two-step estimation approach, while we suggest a one-step estimator. Therefore, we also improve estimation and nowcasting accuracy from this angle. As a by-product, we provide an estimator for the standard deviation of the estimator for this expected value. Such an estimator is not available in Breitung and Hafner (2016).

Third, in contrast to Breitung and Hafner (2016) and Francq and Sucarrat (2018), we rely on a non-Gaussian state-space filter (Kitagawa, 1987, see also Calzolari and Halbleib, 2018) to obtain stationary initial recursion values. As the typical reduced-form AR and MA parameters are relatively close to, but still below, the boundary of unity, the initial conditions impact the estimation in a non-negligible way even in larger samples although the effect vanishes asymptotically.

We investigate the newly proposed estimation routine in comparison to existing ones in a Monte Carlo study. The results clearly confirm that noticeable efficiency gains are achievable in practice. In a broad empirical application it turns out that for five major stock markets the approach by Breitung and Hafner (2016) is significantly outperformed over the full sample from 2000 to 2019. This finding is robust to various sub-samples relating to the Great Financial Crisis. By considering possible time-variation explicitly we are also able to show that our newly suggested modifications outperform its simpler version for all markets in almost all time points. Mean squared nowcast errors can be significantly reduced by up to 30%. Generally, we find that the inclusion of the current observation is quite important for volatility nowcasting. While the log-GARCH model (excluding current information) is routinely outperformed, the RT-GARCH model by Smetanina (2017) is good competitor, but outperformed in most situations as well. We would like to stress that we use a 5-minute realized volatility measure as a benchmark for volatility nowcasts rather than daily squared returns even though the framework of Breitung and Hafner (2016) is designed for the latter. Realized volatility has proven to be a robust and reliable benchmark for such comparisons, see Andersen and Bollerslev (1998). As the model by Breitung and Hafner (2016) is about return data on a daily frequency, it is widely applicable, even in markets in which intraday data is either not broadly available at all or only in questionable quality. Moreover, realized volatility is not available for long time spans and costly for a large number of (individual) assets. Besides financial applications, the model by Breitung and Hafner (2016) could also be applied in a macro-financial context, say to output and inflation volatility.[3]

Our paper is organized as follows. Section 2 gives a detailed exposition of the nowcasting model, its ARMA representation and estimation approaches. Section 3 contains our empirical analysis. Conclusions are drawn in Section 4. The appendices contain further empirical and simulation results.

2 Econometric methods

In the following, we present the nowcasting model for volatility and the corresponding ARMA representation. We provide the framework for obtaining volatility nowcasts, discuss distributional properties and devise an exact maximum likelihood estimation approach.

2.1 Model specification

We work with the following simple model for financial returns y_t as in Breitung and Hafner (2016):

$$y_t = \exp(h_t/2)\xi_t \,, \qquad t = 1, 2, ..., T \tag{1}$$

with h_t being the latent log-volatility process and ξ_t an i.i.d. process with zero mean and unit

[3] Historically, the ARCH model was applied to inflation volatility, see Engle (1982).

variance. The dynamic volatility process is given as

$$h_t = \alpha + \beta h_{t-1} + \kappa \varepsilon_t \tag{2}$$

where α is the intercept in this autoregressive specification and β is the autoregressive parameter measuring the persistence. The third parameter κ measures the impact of the current information on volatility in the same period. Stationarity is ensured by excluding a unit root via the restriction $|\beta| < 1$.[4] The innovation term ε_t is given by $\varepsilon_t = \log(\xi_t^2) - C$ with zero mean and variance σ_ε^2. The parameter C is defined via $C = E[\log(\xi_t^2)]$ in order to ensure that ε_t is a zero mean process. The expectation depends on the distributional properties of ξ_t. For instance, if ξ_t were Gaussian, then ε_t is distributed as a $\log -\chi^2(1)$ random variable and $C \approx -1.27$. Under fat-tailedness, say ξ_t would be distributed as a $t(5)$ random variable and hence $\varepsilon_t \sim \log -F(1,5)$, C equals approximately -1.57. Similar to Breitung and Hafner (2016), we assume that C is unknown, but contrary to their approach, we estimate all parameters jointly in one-step rather than resorting to a two-step procedure where C is estimated in the second step separately from the other parameters in the volatility nowcasting model.

2.2 ARMA representation

The nowcasting model can be transformed to obtain a simple linear ARMA(1,1) representation. To this end, define log-squared returns $x_t = \log y_t^2$ and one obtains directly

$$x_t = C + h_t + \varepsilon_t \ . \tag{3}$$

Replacing h_{t-1} with $x_{t-1} - C - \varepsilon_{t-1}$ in the volatility process equation $h_t = \alpha + \beta h_{t-1} + \kappa \varepsilon_t$, we get after re-arrangement

$$x_t = \alpha^* + \beta x_{t-1} + (1+\kappa)\varepsilon_t - \beta \varepsilon_{t-1} \ .$$

Here, the intercept α^* is defined as $\alpha^* \equiv \alpha + (1-\beta)C = \mu(1-\beta)$, where μ is the expected value of x_t. The log-squared return series x_t hence obeys an ARMA(1,1) representation where the current information in ε_t impacts x_t contemporaneously and its strength is measured by the parameter κ. It is interesting and insightful to compare this representation to the one for the log-GARCH model

[4] Stationarity is investigated in the empirical application by the unit root test under measurement errors as suggested by Hansen and Lunde (2014). Their test is based on instrumental variable estimation and the results suggest that realized volatility is highly persistent, but stationary, see Section 3.

(Francq and Sucarrat, 2018), i.e.

$$h_t = \alpha + \beta h_{t-1} + \psi \log y_{t-1}^2$$

which we are considering in our simulations and empirical applications as well. Similarly, the ARMA representation is given by

$$x_t = \alpha^* + (\psi + \beta) x_{t-1} + \varepsilon_t - \beta \varepsilon_{t-1} \ .$$

Here, the impact of the lagged return series is measured by the parameter ψ and hence appears in front of x_{t-1}.

The above ARMA(1,1) representation of the nowcasting model can be expressed as the observationally equivalent version

$$x_t = \alpha^* + \beta x_{t-1} + u_t - \theta u_{t-1} \tag{4}$$

where $u_t = (1+\kappa)\varepsilon_t$ has zero mean and $\theta = \beta/(1+\kappa)$. Stationarity and invertibility of this ARMA process is achieved when $1 + \kappa > \beta$ implying $\theta < 1$. The structural parameter κ is related to the reduced-form parameters in the following nonlinear way

$$\kappa = \frac{\beta}{\theta} - 1 \ .$$

The nowcast of volatility is obtained by noting that $\varepsilon_t = u_t/(1+\kappa) = (\theta/\beta)u_t$ and

$$h_t = x_t - (\theta/\beta)u_t - C \ .$$

Noteworthy, the nowcasting model enables a contemporaneous impact of information in t on volatility in the same period. Re-expressing the volatility nowcast as

$$h_t = \frac{\alpha^* \theta}{\beta(1-\theta)} + \left(1 - \frac{\theta}{\beta}\right) \sum_{j=0}^{\infty} \theta^j x_{t-j} \tag{5}$$

shows that it is a linear filter of current and past information with weights that are exponentially decaying, see Breitung and Hafner (2016).

If we replace the dynamic volatility process in Equation (2) with

$$h_t = \alpha + \beta h_{t-1} + \eta_t$$

where η_t and ξ_t are mutually independent. The approach is a classical stochastic volatility (SV) model. Breitung and Hafner (2016) shows this approach still has an ARMA representation with

$$\sigma_\eta^2 = \left[1 - \frac{\theta}{\beta} - \theta(\beta - \theta)\right] \sigma_u^2$$

where the SV parameters are fully transformed into ARMA parameters. Equation (5) shows volatility can be solely defined be the log squared returns and ARMA parameters, therefore the SV filtered volatility using Kalman filter technique is equivalent to the filtered volatility using ARMA representation.

2.3 Volatility nowcasts

Volatility nowcasts $\widehat{h}_t$ are obtained via

$$\widehat{h}_t = x_t - \widehat{\varepsilon}_t - \widehat{C} \tag{6}$$

where x_t is observed and $\widehat{\varepsilon}_t$ is obtained from $\widehat{u}_t/(1+\widehat{\kappa})$ where $\widehat{u}_t$ are the residuals from the estimated ARMA(1,1) model and $\widehat{\kappa}$ is estimated as $\widehat{\kappa} = \widehat{\beta}/\widehat{\theta}-1$. The constant $C = E[\log(\xi_t^2)]$ can be estimated by first assuming a zero mean process for h_t and then shifting the nowcasts by an estimate of C as outlined in Breitung and Hafner (2016). Another option is to estimate it directly with the other ARMA parameters as we suggest. One of the advantages of doing so is that we also build the likelihood function properly on the basis of a non-Gaussian density function and we estimate the parameter C accordingly.

Clearly, the accuracy of the volatility nowcasts $\widehat{h}_t$ hinges directly on the estimation accuracy of the ARMA parameters β and θ and the constant C. The latter directly demonstrates the importance of the distributional characteristics for the nowcasts of volatility in this framework. Our target is to improve estimation accuracy of all parameters, and thereby the volatility nowcasting performance, by using an exact maximum likelihood approach with non-Gaussian log-squared innovation densities. Breitung and Hafner (2016) acknowledge that efficiency gains are possible by imposing appropriate distributional assumptions (e.g. logarithmic χ^2-distribution for ε_t based on normality of ξ_t or logarithmic F-distribution under a fat-tailed t-distribution of ξ_t) when building the likelihood function. As the authors point out, an exact maximum likelihood procedure (with stationary initial values) should be preferred as the moving average parameter θ is typically close to unity in empirical applications. Overall, the increased computational cost is almost zero on modern computers, while the gains in accuracy are remarkable, even in larger samples. Our empirical results clearly show that significant gains are achievable in practice.

2.4 Exact Maximum Likelihood estimation

We start with the reduced-form ARMA(1,1) model

$$x_t = \alpha^* + \beta x_{t-1} + u_t - \theta u_{t-1} \ .$$

Similar to Breitung and Hafner (2016) and Francq and Sucarrat (2018), we estimate the sample mean of x_t ($\widehat{\mu} = T^{-1} \sum_{t=1}^{T} x_t$) and work with the mean-adjusted counterpart (as indicated by $\tilde{x}_t$)

$$\widetilde{x}_t = \beta \widetilde{x}_{t-1} + \widetilde{u}_t - \theta \widetilde{u}_{t-1}.$$

The general log-likelihood function with (non-Gaussian) density f for estimating the parameters $\vartheta = \{\beta, \theta, C\}$ is given as

$$\log L_T(\vartheta) = \log f(x_1; \vartheta) + \sum_{t=2}^{T} \log f(x_t | x_{t-1}, ..., x_1; \vartheta) \ . \qquad (7)$$

A crucial point is the specification of the density f.[5] The innovation term of the reduced form ARMA process for log-squared returns might follow an unfamiliar distribution even if the structural return innovation distribution is well known. In some cases (like for the Gaussian and Student-t distribution), the densities of the squared standardized random variable belong to the set of standard distributions (e.g. χ^2- and F-distribution).[6] The logarithmic form is, however, only available in a few cases, but can be derived from the original distribution by using the exponentiated squared standardized random variable and associated standardization parameters. The general exact maximum likelihood approach we suggest allows for symmetric location-scale distributions of the structural innovation distribution and uses their log-squared transformation.[7] Besides normality of structural return innovations, we provide the leading case for modelling fat tails in finance, namely

[5] Please note that this density is determined via the log-squared transformation of the structural innovation density, see below.

[6] A counterexample could be the generalized error distribution (see e.g. Nelson, 1991) whose squared form needs to be derived as an intermediate step.

[7] Symmetry is required to retain the martingale difference sequence properties of y_t, see the discussion in Breitung and Hafner (2016). The authors also study an asymmetric extension covering the leverage effect. Such a specification might absorb some of potential asymmetry. Hence, an asymmetric version might be interesting in such situations.

the t-distribution, see Bollerslev (1987) for the GARCH context. This leads us to the $\log-\chi^2(1)$- and $\log-F(1,\nu_2)$-distribution for the reduced form ARMA innovations in the representation for log-squared returns. The selection of the density for structural innovations is addressed at the end of this section.

Similar to Francq and Sucarrat (2018), the reduced form innovation density is a $\log-\chi^2(1)$ distribution, given Gaussianity of ξ_t. The standardized density of the $\log-\chi^2(1)$-distribution is given as:

$$\frac{1}{\sqrt{2\pi}}\exp\left(\frac{1}{2}\tilde{u}_t\right)\exp\left(-\frac{1}{2}\exp(\tilde{u}_t)\right)\ , \tag{8}$$

see e.g. Comte (2004) for its probabilistic properties. Alternatively, under fat-tailedness of ξ_t, say via $\xi_t \sim t(\nu_2)$, the resulting density for the reduced form innovations is the standardized density of the $\log-F(1,\nu_2)$-distribution following Jones (2008):

$$\frac{1}{B\left(\frac{1}{2},\frac{\nu_2}{2}\right)}\cdot\exp\left(\frac{1}{2}\tilde{u}_t-\frac{1+\nu_2}{2}\cdot\log\left(1+\exp(\tilde{u}_t)\right)\right) \tag{9}$$

where $B(\cdot,\cdot)$ is the beta function. In this case, the parameter vector ϑ is augmented by the additional parameter ν_2.

To construct the volatility nowcasts given in (6), we need to transform the estimated ARMA error term $\widehat{\tilde{u}}_t$ into $\widehat{u}_t=\widehat{\tilde{u}}_t-\widehat{\alpha}^*$, and $\widehat{\varepsilon}_t=(\widehat{\theta}/\widehat{\beta})\widehat{u}_t$. The intercept α^* is estimated as $\widehat{\alpha}^*=\widehat{\mu}(1-\widehat{\beta})$. Henceforth, volatility nowcasts are given as $\widehat{h}_t=x_t-\widehat{\varepsilon}_t-\widehat{C}$, as described above. Finally, κ is estimated via $\widehat{\kappa}=\widehat{\beta}\widehat{\theta}^{-1}-1$.

We resort to an exact maximum likelihood [EML] estimator with stationary initial recursion values. Several choices can be made to estimate the initial recursion value x_1 and to get a non-zero innovation u_1, i.e. via backcasting, an EM-algorithm or a state-space approach. After having experimented, we opt for the latter approach and rely on the non-Gaussian Kitagawa filter.

A Monte Carlo simulation with 5000 replications shows the exact maximum likelihood estimator advantage over the conditional one. The data generating process follows the proposed structural model with standard normally distributed ξ_t innovations. Two settings of parameters are employed here: β equals 0.95 (or 0.99) and κ is set to 0.056 (or 0.042). The two settings imply a moving average root θ of 0.9 (or 0.95) near unity. These settings resemble typical empirical situations. The intercept α is set to zero.

First, the MSE of estimated reduced form ARMA parameters are reported in Table B.1. In the first two columns, the parameters are estimated using the Gaussian approximation using CML. In columns three to eight, the parameters are estimated by using the $\log -\chi^2(1)$-distribution with either zero initial recursion value (CML) or "estimated" initial recursion values (EML and backcasting). The advantage of the EML estimator can be clearly seen as it offers a gain in MSE of about 30% compared to the CML estimator by using the $\log -\chi^2(1)$-distribution. Moreover, a gain of over 40% is possible for the case of $T = 2000$. The advantage of accurately estimated parameters by using an EML estimator with the $\log -\chi^2(1)$-distribution is also reflected in the evaluation of the nowcast performance. In Table B.2, the volatility nowcast performance by using the Gaussian approximation and the $\log -\chi^2(1)$-distribution with EML is compared by means of R^2, MSE and MAE:

$$R^2 = 1 - \frac{\sum_{t=1}^{T} \left(h_t - \widehat{h}_t\right)^2}{\sum_{t=1}^{T} \left(h_t - \overline{h}\right)^2} , \tag{10}$$

$$MSE = \frac{1}{T} \sum_{t=1}^{T} \left(h_t - \widehat{h}_t\right)^2 , \tag{11}$$

$$MAE = \frac{1}{T} \sum_{t=1}^{T} \left|h_t - \widehat{h}_t\right| , \tag{12}$$

where h_t is the generated log volatility and $\bar{h}_t$ denoting the sample average of h_t. The EML estimator using $\log -\chi^2(1)$-distribution gives a higher R^2 and lower MSE/MAE over all sample sizes. Even with a sample size of $T = 2000$, the R^2 is more than 10% higher by using the $\log -\chi^2(1)$ distribution together with the EML estimator in comparison to the Gaussian approximation. Further simulations are conducted in which both estimation approaches are misspecified. We use Student-t distribution with ten degrees of freedom for the structural innovations. Results are reported in Table B.3. In all cases, the $\log -\chi^2(1)$ approach outperforms the Gaussian approximation. Finally, the MSE of the estimation of the constant parameter C is reported in Table B.4. In comparison to the Gaussian approximation and log-GARCH model, the proposed ARMA nowcasting model offers the smallest MSE overall, especially for the smaller sample sizes. Overall, we observe clear evidence of the advantage of i) using the non-Gaussian distribution, ii) the estimation of the initial recursion value and iii) the joint estimation of parameters $\vartheta = \{\beta, \theta, C\}$. All these techniques of the non-Gaussian distribution, the jointly estimated parameter C and the estimation of initial recursion value (EML) are applied in our empirical analysis in the next Section.

We employ the density selection procedure proposed by Marin and Sucarrat (2015) to select the best suited distribution for the structural innovations ξ_t. The reduced form innovation density is then obtained via the log-squared transformation. The method uses information criteria. According to the simulation study in Marin and Sucarrat (2015), the Schwarz criterion provides the most accurate density selection. The procedure consists of three steps: for each candidate density, its parameters (collected in the vector λ, including mean and variance parameters) are estimated via maximum likelihood in their standardized form. Next, the Schwarz criterion, $SC = -2 \log L_T(\lambda) + \dim(\lambda) \log T$, is computed. The density yielding the smallest value of SC is selected for ξ_t. The density for the reduced form innovations is then obtained via the log-squared transformation.

We illustrate the accuracy of the density selection procedure in a small Monte Carlo simulation study. We draw the innovations ξ_t from the standard normal distribution or the t-distribution with ten degrees of freedom. To be noticed, we use the standardized form of the t-distribution as given in the R (R Core Team, 2020) package "fGarch" Wuertz et al. (2013) with three parameters $\lambda = \{\mu, \sigma^2, \nu\}$ in both, the simulation study and the empirical analysis. The results are reported in Table B.5. The procedure performs quite well. With a sample size of $T = 2000$, the procedure selects the correct distribution in more than 99% of the cases. The estimated parameters lie close to the true parameters.

3 Application to five stock indices

We apply the proposed volatility nowcasting methods to the daily returns of five major stock indices: Dow Jones Industrial Average, Nasdaq, S&P 500, FTSE and Nikkei 225. Thereby, we cover three different US stock markets, a European and an Asian one. These markets have been widely analyzed in the related literature and we study the merits and limitations of several volatility nowcasting approaches. In particular, we compare our approach using exact maximum likelihood for the $\log -\chi^2$-distribution (and $\log -F(1, \nu_2)$-distribution) to the simplified version by Breitung and Hafner (2016) using normality assumptions and QML estimation.[8] We provide an empirical study of our volatility nowcasting approach building upon non-normality of log-squared returns to a version assuming normality (Breitung and Hafner, 2016) and a specification which accounts for non-normality but excludes the contemporaneous effect (Francq and Sucarrat, 2018). In detail, we include a log-GARCH specification in our comparison whose maximum likelihood estimation builds upon the $\log -\chi^2$-distribution. Thereby, we are able to isolate the effect of including the contemporaneous effect from simultaneously accounting for non-normality of log-squared returns.

[8]The authors exploit the information that the theoretical variance equals $\frac{\pi^2}{2}$ in their estimation procedure.

Table 1: Descriptive statistics of returns and log-squared returns

		r_t				$\log r_t^2$			
	# obs.	Mean	Var	Skewness	Kurtosis	Mean	Var	Skewness	Kurtosis
Dow Jones	5010	0.000	1.174	-0.141	11.250	-1.889	6.999	-1.187	5.990
Nasdaq	5011	-0.020	1.721	-0.049	10.224	-1.385	6.295	-1.084	5.466
S&P 500	5013	0.009	1.224	-0.333	11.168	-1.825	6.702	-1.158	6.398
FTSE	5041	-0.005	1.251	-0.388	9.109	-1.571	5.802	-1.148	6.106
Nikkei 225	4866	0.000	1.235	-0.747	14.182	-1.635	5.829	-1.117	5.993

All specifications are estimated via their ARMA representation and are thus directly comparable.

The daily data is sampled from 01/03/2000 to 12/30/2019 and retrieved from the Oxford-Man Realized Library.[9] We use open-to-close returns.[10] We conduct a full sample analysis and further studies on sub-samples (see below). The full sample data is visualized in Figure 1. Here, we plot the returns and the log-squared returns for the five markets under consideration. Descriptive statistics of (log-squared) returns are provided in Table 1. It can be seen that the returns share the common characteristics of an almost zero mean, slightly negative skewness and considerable excess kurtosis. Typical features like volatility clusters are present in the data. Regarding the log-squared returns we also find remarkable skewness and kurtosis which further motivates the use of an exact maximum likelihood estimation approach.

Now, we proceed with the estimation of volatility models and nowcast evaluation.[11] Estimation results are reported in Table 2. In the first column, the result for the Breitung and Hafner (2016) [BH] approach is given, followed by our proposed versions with a $\log -\chi^2$-distribution and a $\log -F(1,\nu_2)$-distribution and finally, the log-GARCH model as implemented in Francq and Sucarrat (2018) by means of the R (R Core Team, 2020) package "lgarch" Sucarrat (2015). R^2 of each estimation is computed. Moreover, we present estimation results for the reduced-form ARMA(1,1) model with autoregressive parameter β, moving average parameter θ and the implied measure κ for the importance of the contemporaneous information. For the $\log -F(1,\nu_2)$-distribution, an estimate of ν_2 is reported as well. Standard errors computed via the Hessian matrix are reported in parentheses.

The estimates of the ARMA parameters β and θ are plausible and highly significant. The implied estimates of $\kappa = \beta/\theta - 1$ have also plausible sign and magnitude in all markets. We find the contemporaneous effect to be important and somewhat larger in $\log -\chi^2$-specifications (up to 0.093

[9] Data is retrieved from: https://realized.oxford-man.ox.ac.uk/

[10] Returns are calculated as $r_t = 100 \cdot (\ln P_t - \ln P_{t-1})$ with P_t denoting the level of the stock market index in time point t.

[11] Estimation of the covariance matrix of the parameters is done via the Hessian matrix. As κ is defined as the ratio between the autoregressive and the moving average parameter (minus one), the variance of $\widehat{\kappa}$ is obtained by the delta method.

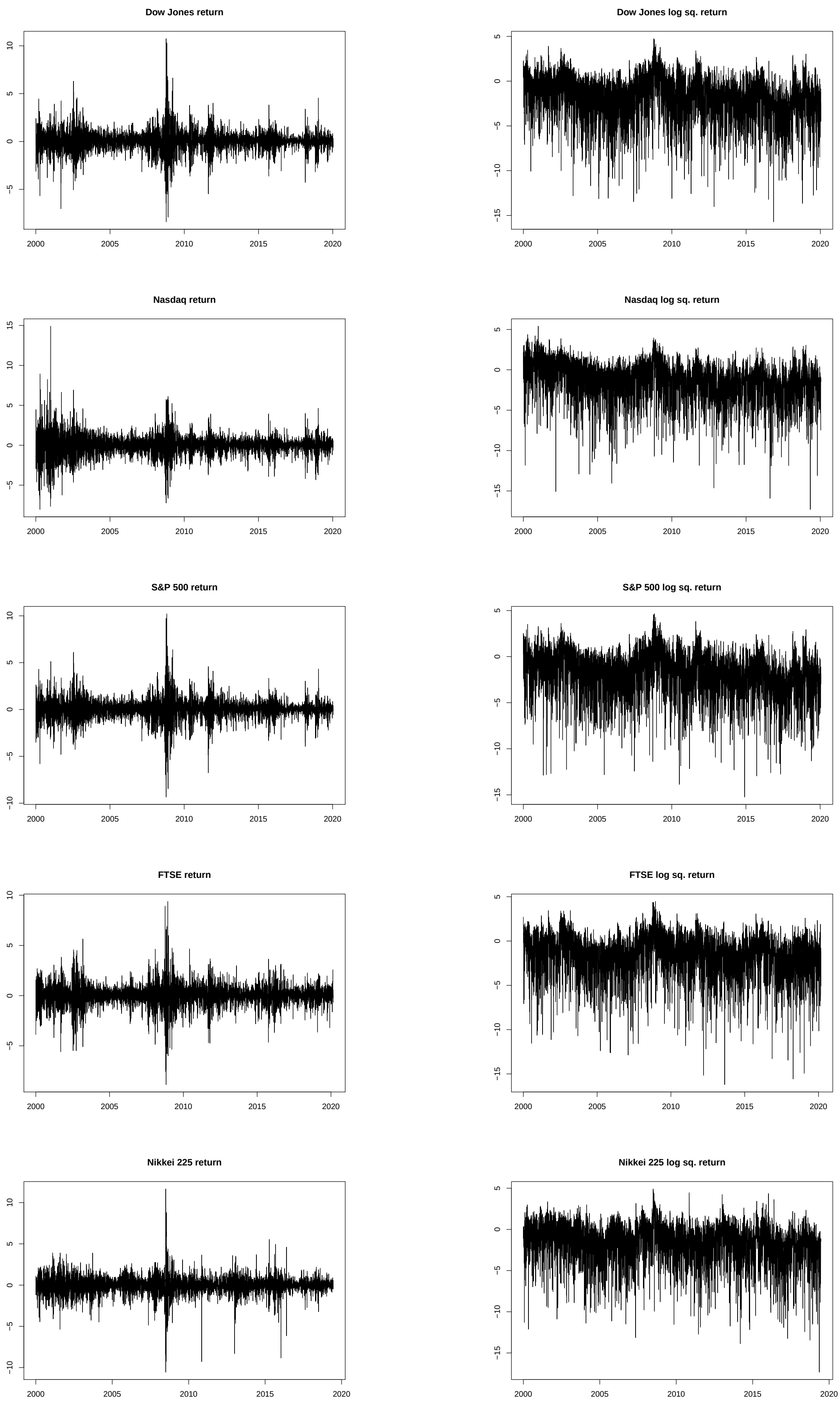


Figure 1: Returns and log-squared returns

for Dow Jones). Furthermore, the estimated degree of freedom ν_2 in the $\log -F(1, \nu_2)$-specifications resemble the previously detected non-normality in returns. The estimates vary between 6.2 and 9.0. This result is consistent with the result obtained from the density selection procedure which prefers the fat-tailed t-distribution over the Gaussian distribution in all markets, see Table 3. The estimated degrees of freedom stay in line with the ones in $\log -F$-distribution. Finally, the estimated values for the constant C controlling the volatility nowcasting mean adjustment also suggest deviations from normality as their estimates differ (significantly) from the theoretical value of -1.27. By construction, the Breitung and Hafner (2016) approach does not offer a standard error for the estimate of C, while it is available for our exact maximum likelihood approach and also the log-GARCH model by Francq and Sucarrat (2018).

In Table 4, as evaluation measures for volatility nowcasts $\widehat{h}_t$ we compute the MSE and MAE by using log-realized volatility h_t (based on 5-minutes sampling) as a benchmark. Contrary to Breitung and Hafner (2016), we evaluate the nowcasts against realized volatility (5-min).[12] For both loss functions we also provide relative measures in which the Breitung and Hafner (2016) approach is taken as the benchmark (with value 1). Values below unity indicate that the benchmark is outperformed and vice versa. Additionally, we compare the volatility nowcast obtained from the RT-GARCH model following Smetanina (2017). In the RT-GARCH model the volatility is also driven by the information set including the current period and QML is employed for parameter estimation.[13] The nowcasting performance is compared by means of the model confidence set (MCS) approach introduced by Hansen et al. (2011). The p-values for models included in the final model confidence set are given in the table below MSEs and MAEs, respectively. Missing entries indicate that the respective model is eliminated from the model confidence set.

[12] An application of the Hansen and Lunde (2014) instrumental variable-based unit root test confirms the absence of a unit root at all conventional significance levels, see Appendix C.

[13] Estimation results for the RT-GARCH model are available upon request.

Table 2: Volatility nowcasting estimation of full sample data

	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
				Dow Jones				
R^2	0.484		**0.591**		0.588		0.393	
κ	0.060	*(0.006)*	0.093	*(0.005)*	0.060	*(0.005)*		
β	0.990	*(0.003)*	0.967	*(0.004)*	0.985	*(0.003)*	0.973	*(0.003)*
θ	0.934	*(0.007)*	0.885	*(0.003)*	0.929	*(0.003)*	0.907	*(0.006)*
C	-1.472	*(—)*	-1.395	*(0.023)*	-1.474	*(0.032)*	-1.632	*(0.020)*
ν_2					6.213	*(0.523)*		
				Nasdaq				
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.434		0.520		**0.624**		0.333	
κ	0.047	*(0.005)*	0.060	*(0.004)*	0.048	*(0.004)*		
β	0.994	*(0.002)*	0.984	*(0.002)*	0.991	*(0.002)*	0.982	*(0.002)*
θ	0.949	*(0.006)*	0.928	*(0.003)*	0.945	*(0.003)*	0.925	*(0.005)*
C	-1.356	*(—)*	-1.315	*(0.021)*	-1.335	*(0.026)*	-1.606	*(0.020)*
ν_2					9.002	*(1.005)*		
				S&P 500				
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.503		0.580		**0.613**		0.394	
κ	0.062	*(0.006)*	0.083	*(0.005)*	0.061	*(0.004)*		
β	0.989	*(0.003)*	0.977	*(0.003)*	0.988	*(0.002)*	0.980	*(0.002)*
θ	0.932	*(0.007)*	0.902	*(0.003)*	0.931	*(0.003)*	0.915	*(0.005)*
C	-1.438	*(—)*	-1.387	*(0.022)*	-1.441	*(0.030)*	-1.476	*(0.020)*
ν_2					6.650	*(0.582)*		
				FTSE				
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.506		0.570		**0.589**		0.404	
κ	0.057	*(0.006)*	0.081	*(0.005)*	0.066	*(0.004)*		
β	0.985	*(0.004)*	0.976	*(0.003)*	0.982	*(0.003)*	0.981	*(0.003)*
θ	0.932	*(0.007)*	0.903	*(0.003)*	0.921	*(0.003)*	0.915	*(0.006)*
C	-1.337	*(—)*	-1.280	*(0.022)*	-1.328	*(0.026)*	-1.482	*(0.020)*
ν_2					9.003	*(0.908)*		
				Nikkei 225				
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.447		**0.558**		0.549		0.331	
κ	0.057	*(0.007)*	0.090	*(0.005)*	0.046	*(0.004)*		
β	0.987	*(0.004)*	0.967	*(0.005)*	0.988	*(0.003)*	0.975	*(0.003)*
θ	0.934	*(0.009)*	0.887	*(0.004)*	0.944	*(0.003)*	0.911	*(0.006)*
C	-1.459	(—)	-1.387	*(0.023)*	-1.471	*(0.031)*	-1.604	*(0.020)*
ν_2					6.193	*(0.507)*		

Note: Bold-faced values indicate the best performing specification. Numbers in parentheses are standard deviations.

Table 3: Density selection of full sample data

	Gaussian distribution			*t*-distribution			
	SC	μ	σ^2	SC	μ	σ^2	ν
Dow Jones	14988.823	0.026	1.085	**14575.010**	0.065	1.096	4.789
Nasdaq	14703.280	-0.006	1.054	**14476.281**	0.034	1.055	6.508
S&P 500	14994.256	0.011	1.085	**14644.012**	0.053	1.092	5.182
FTSE	14961.910	-0.002	1.072	**14829.378**	0.020	1.076	7.602
Nikkei 225	14564.930	0.005	1.086	**14111.200**	0.029	1.076	5.438

Note: Bold-faced value indicates the selected density.

Table 4: Volatility nowcasting performance of full sample data

	Dow Jones				
	BH	$\log-\chi^2(1)$	$\log-F(1,\nu_2)$	log-GARCH	RT-GARCH
MSE	1	*0.794*	*0.799*	*1.177*	*0.793*
MCS p-value		**1.000**	0.956		**1.000**
MAE	1	*0.883*	*0.799*	*1.103*	*0.882*
MCS p-value		0.100	**1.000**		0.690

	Nasdaq				
	BH	$\log-\chi^2(1)$	$\log-F(1,\nu_2)$	log-GARCH	RT-GARCH
MSE	1	*0.849*	*0.666*	*1.177*	*0.834*
MCS p-value			**1.000**		
MAE	1	*0.911*	*0.787*	*1.097*	*0.903*
MCS p-value			**1.000**		

	S&P 500				
	BH	$\log-\chi^2(1)$	$\log-F(1,\nu_2)$	log-GARCH	RT-GARCH
MSE	1	*0.846*	*0.780*	*1.221*	*0.804*
MCS p-value			**1.000**		0.199
MAE	1	*0.916*	*0.858*	*1.122*	*0.886*
MCS p-value			**1.000**		0.059

	FTSE				
	BH	$\log-\chi^2(1)$	$\log-F(1,\nu_2)$	log-GARCH	RT-GARCH
MSE	1	*0.870*	*0.833*	*1.206*	*0.878*
MCS p-value		0.242	**1.000**		0.323
MAE	1	*0.933*	*0.895*	*1.123*	*0.933*
MCS p-value			**1.000**		

	Nikkei 225				
	BH	$\log-\chi^2(1)$	$\log-F(1,\nu_2)$	log-GARCH	RT-GARCH
MSE	1	*0.798*	*0.815*	*1.210*	*0.774*
MCS p-value		0.964	0.464		**1.000**
MAE	1	*0.884*	*0.861*	*1.127*	*0.864*
MCS p-value			**1.000**		0.846

Note: Bold-faced value(s) indicate the model(s) with the largest MCS p-value. Numbers in italics are relative to the benchmark (BH). Missing entries indicate that the respective model is eliminated from the model confidence set.

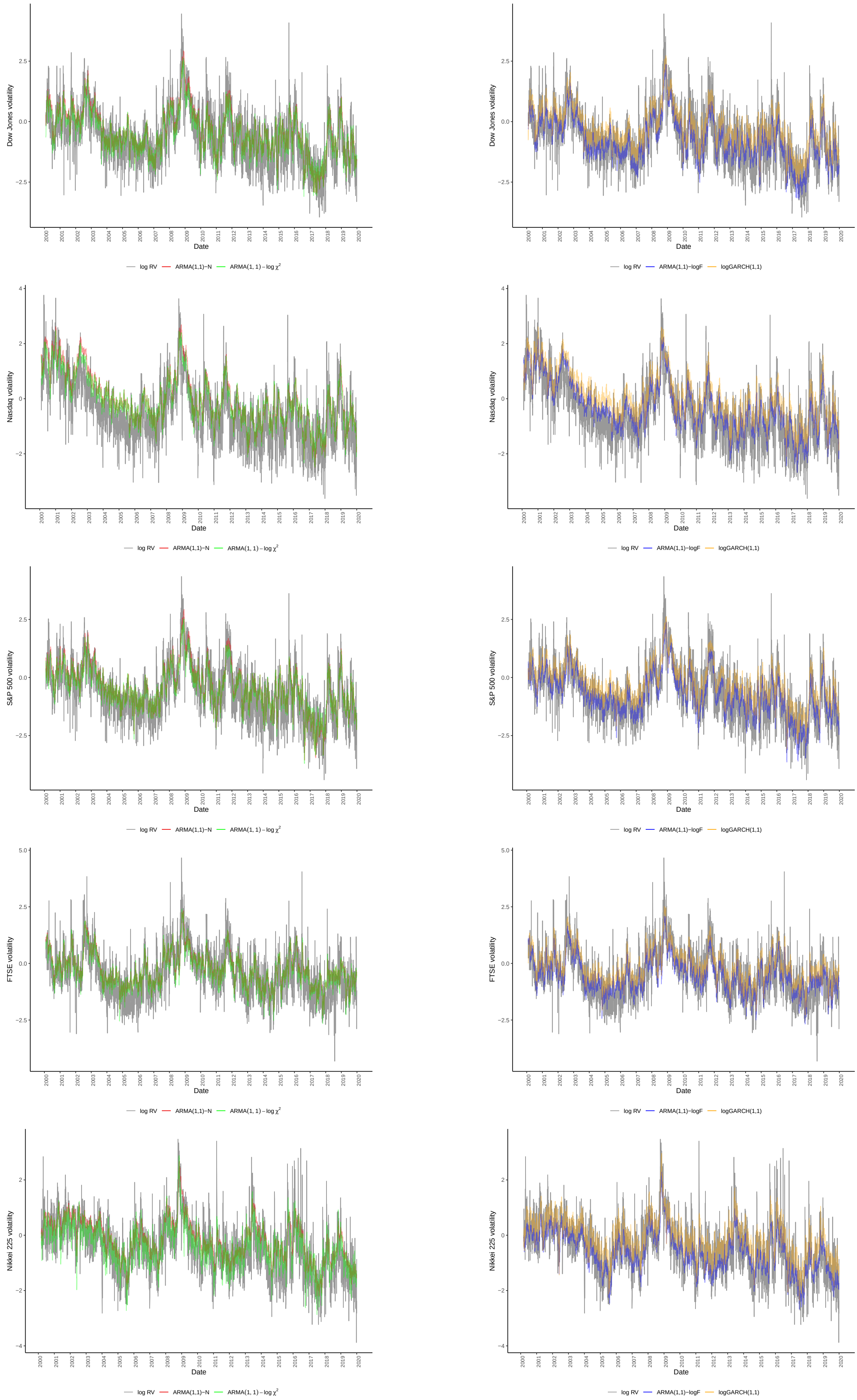
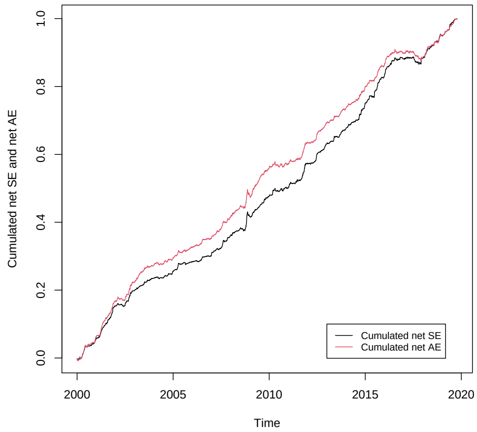
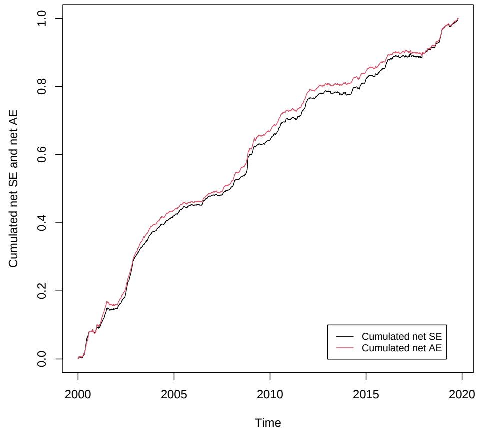


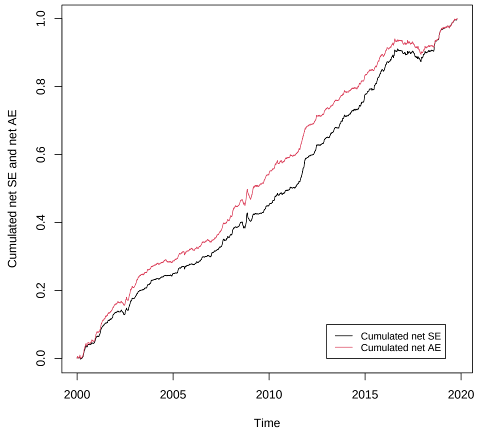
Figure 2: Volatility nowcasting (full sample)



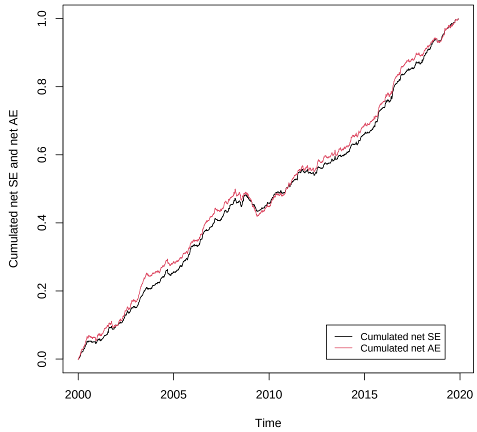
(a) Dow Jones



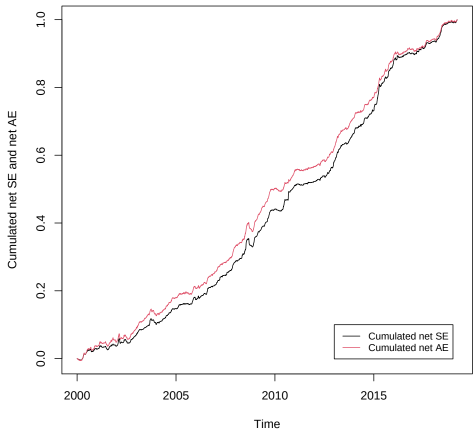
(b) Nasdaq



(c) S&P 500



(d) FTSE



(e) Nikkei 225

Figure 3: Cumulative net squared and absolute error plots (full sample)

Strikingly, the Breitung and Hafner (2016) approach is outperformed in each case by both suggested variants ($\log -\chi^2(1)$ and $\log -F$) using exact maximum likelihood estimation. This result holds irrespectively of the particular evaluation measure (R^2, MSE or MAE). The coefficient of determination is increased by 8.2 (FTSE) to 18.9 (Nasdaq) percentage points relative to the benchmark. Overall, the best performing specifications (estimated via exact maximum likelihood and non-Gaussian distributions) offer R^2 measure in the range from 55.8% (Nikkei) to 62.4% (Nasdaq). The log-GARCH specification performs somewhat worse, suggesting that the inclusion of the contemporaneous information is indeed quite important. However, additionally accounting for non-normality of log-squared returns via an exact maximum likelihood approach further enhances the nowcast performance via the $\log -\chi^2$ and $\log -F$ specifications. The MSE and MAE measures can be reduced by around twenty and fifteen percent (in comparison to the Breitung and Hafner (2016) approach) across the different markets, respectively. Volatility nowcasts are displayed in Figure 2. It can be seen that realized volatility is tracked quite well by the ARMA-based models using only daily data rather than intra-day information. This is also supported by the overall R^2-values of around 60%. Based on the MCS results, the best nowcasting model with $\log -F$-distribution is included for all five markets (being consistent with the density selection procedure), while the second best is the RT-GARCH model which is included less often. Neither the Gaussian nowcasting model nor the log-GARCH model are included in the model confidence set for any market.

Before we turn to our sub-sample analyses, we first investigate potential time-variation in the relative nowcasting performance: It is possible that even though the newly proposed exact maximum likelihood approach dominates the extant one by Breitung and Hafner (2016) over the full sample, the relative nowcasting performance might be time-varying within the sample. This is studied by considering simple cumulative net squared (or absolute) errors. The squared and absolute errors $(h_t - \widehat{h}_t)^2$ and $|h_t - \widehat{h}_t|$ enter the three evaluation measures R^2, MSE and MAE. Therefore, it is informative to consider the cumulative net difference (V_t) of these errors between the BH and $x = \{\log -\chi^2\}$ approach over time

$$V_t = \sum_{j=1}^{t} \left| h_j - \widehat{h}_j^{BH} \right|^p - \left| h_j - \widehat{h}_j^{x} \right|^p ,$$

with $p = \{1, 2\}$. These are provided in Figure 3. For convenience, we rescale the values to a terminal value of ± 1 for ease of comparison.[14] Values above unity indicate that the Breitung and

[14] As the BH approach is outperformed in all cases, terminal values V_T are always rescaled to +1.

Hafner (2016) approach is outperformed and vice versa. While the sign of the cumulative net errors might vary over time, the terminal value is positive as the Breitung and Hafner (2016) approach is outperformed in all markets. Remarkably, the displayed cumulative net errors clearly demonstrate a dominance by the newly proposed methods over nearly each single time period.

Overall, the results indicate the advantages of the newly proposed methods over existing ones, also dynamically over time. We observe a mostly monotonic and sometimes even linear shape of the cumulative net errors. There is very little difference in the results for squared and absolute errors.

We now turn to our sub-sample analyses relating to the Great Financial Crisis, we choose for simplicity a similar sample split point as in the related work of Smetanina (2017), i.e. 07/31/2008. The pre-crisis sample contains about 2100 observations, while the crisis (and post-crisis) sample ends in 12/30/2019 covering about 2860 data points.

Descriptive statistics of the data on sub-samples are given in Table A.1. The overall descriptive findings in the full sample hold approximately for all considered sub-samples with some interesting variations. The stock returns during the (post-)crisis sub-sample exhibit the largest kurtosis, while the pre-crisis has smaller kurtosis in contrast. The difference in kurtosis between the samples is clearly smaller, when we consider the log-squared returns. In contrast, the skewness of the series is increased through the log-squared transformation nearly in all markets and across all sub-samples.

What does the sub-sample analysis of pre- and post-crisis reveal? Tables A.2 to A.5 report the results. First of all, most of the previous findings can also be established for two sub-samples relating to the Great Financial Crisis. The evaluation of performance measures leads to similar conclusions as for the full sample. Interestingly, the importance of the contemporaneous information is much more important in the sample ranging from 08/01/2008 to 12/30/2019. Here, we find that in most markets the estimate of κ is doubled in comparison to the pre-crisis sub-sample. In a similar vein, the estimated degree of freedom ν_2 is reduced by around two units indicating fatter tails as expected and confirmed by descriptive statistics. This is also reflected and further substantiated by the estimates of C which deviate stronger from -1.27 (under normality). Again, the density selection procedure follows Marin and Sucarrat (2015) is in favour of t-distribution in all cases, see Table A.6. The cumulative net error plots in Figures A.1 and A.2 reveal that the BH approach performs best in the beginning of the early 2000s for the Nasdaq and during the Great Financial Crisis for the FTSE, while it is outperformed in all other cases and time points. The findings within sub-samples are consistent with the results obtained for the full sample. The MCS results also show a similar pattern to the previous analysis. In the pre-crisis sub-sample, the nowcasting model with $\log -F$-distribution performs best for all markets (one exception is the Nikkei with MAE loss). The benchmark model is always excluded, while th RT-GARCH model is excluded for the Nasdaq and S&P 500. In the post-crisis sub-sample, we find a more diverse picture. The RT-GARCH model is best performing in three markets (Dow Jones, S&P 500 and Nikkei), while the log-F nowcasting model performs best for Nasdaq and FTSE. Still, the benchmark model (BH) is significantly outperformed and thus excluded for each market.

4 Conclusions

In this work, we consider a simple nowcasting model for daily financial volatility. The model allows the current return observation to impact volatility of the same day. Contrary to many situations involving volatility estimation and forecasting, the distributional assumptions in the estimation of the model plays a decisive role. We suggest a one-step exact maximum likelihood estimator which takes non-normal distributional characteristics of log-squared return innovations explicitly into account. Overall, estimation efficiency can be improved from several angles. In simulations we show that noticeable improvements are achievable. Our empirical application to five major stock markets underline the merits of the suggested approach. We find the benchmark to be significantly outperformed in all cases. The current information matters and so does the estimation routine for the construction of volatility nowcasts. These claims are clearly confirmed by our empirical results.

As the proposed volatility nowcasting method does not require intraday data, it can be of additional interest in markets where such information is difficult to obtain in a sufficiently high quality. Furthermore, the use of daily data is (almost) free of charge and does not require tedious intra-day data cleaning procedures. Given that the volatility nowcasts extracted from the exact maximum likelihood approach perform quite well and may explain around 60% percent of 5-min realized volatility, it offers a good benefit/cost-ratio for practical purposes.

An interesting extension for practical situations with many assets in a high-dimensional setting would be a multivariate equation-by-equation framework similar to Francq and Sucarrat (2017). This avenue is left for future research.

References

Andersen, T. G. and Bollerslev, T. (1998), 'Answering the skeptics: Yes, standard volatility models do provide accurate forecasts', *International Economic Review* **39**(4), 885–905.

Asai, M. (1998), 'A new method to estimate stochastic volatility models: A log-garch approach', *Journal of the Japan Statistical Society* **28**(1), 101–114.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics* **31**(3), 307–327.

Bollerslev, T. (1987), 'A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return', *The Review of Economics and Statistics* **69**(3), 542–547.

Breitung, J. and Hafner, C. M. (2016), 'A simple model for now-casting volatility series', *International Journal of Forecasting* **32**(4), 1247–1255.

Calzolari, G. and Halbleib, R. (2018), 'Modelling realized covariance matrices with stochastic volatility latent factors: Filter, likelihood, forecast'. Working Paper.

Comte, F. (2004), 'Kernel deconvolution of stochastic volatility models', *Journal of Time Series Analysis* **25**(4), 563–582.

Cont, R. (2001), 'Empirical properties of asset returns: stylized facts and statistical issues', *Quantitative Finance* **1**(2), 223–236.

Corsi, F., Mittnik, S., Pigorsch, C. and Pigorsch, U. (2008), 'The Volatility of Realized Volatility', *Econometric Reviews* **27**(1-3), 46–78.

Ding, Y. (2021), 'Weak Diffusion limit of real-time GARCH models'. Working Paper, University of Cambridge.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation', *Econometrica* **50**(4), 987–1007.

Francq, C. and Sucarrat, G. (2017), 'An equation-by-equation estimator of a multivariate log-GARCH-X model of financial returns', *Journal of Multivariate Analysis* **153**, 16–32.

Francq, C. and Sucarrat, G. (2018), 'An exponential chi-squared qmle for log-GARCH models via the ARMA representation', *Journal of Financial Econometrics* **16**(1), 129–154.

Geweke, J. (1986), 'Modeling the persistence of conditional variances: a comment', *Econometric Reviews* **5**(1), 57–61.

Hansen, P. R. and Lunde, A. (2005), 'A forecast comparison of volatility models: does anything beat a GARCH(1, 1)?', *Journal of Applied Econometrics* **20**(7), 873–889.

Hansen, P. R. and Lunde, A. (2014), 'Estimating the persistence and the autocorrelation function of a time series that is measured with error', *Econometric Theory* **30**(1), 60–93.

Hansen, P. R., Lunde, A. and Nason, J. M. (2011), 'The model confidence set', *Econometrica* **79**(2), 453–497.

Jones, M. (2008), 'On a class of distributions with simple exponential tails', *Statistica Sinica* **18**(3), 1101–1110.

Kitagawa, G. (1987), 'Non-gaussian state-space modeling of nonstationary time series', *Journal of the American Statistical Association* **82**(400), 1032–1041.

Latane, H. A. and Rendleman, R. J. (1976), 'Standard deviations of stock price ratios implied in option prices', *The Journal of Finance* **31**(2), 369–381.

Marin, J. M. and Sucarrat, G. (2015), 'Financial density selection', *The European Journal of Finance* **21**(13-14), 1195–1213.

Nelson, D. B. (1991), 'Conditional heteroskedasticity in asset returns: A new approach', *Econometrica* **59**(2), 347–370.

Pantula, S. G. (1986), 'Modeling the persistence of conditional variances: a comment', *Econometric Reviews* **5**(1), 79–97.

Politis, D. N. (2007), 'Model-free versus model-based volatility prediction', *Journal of Financial Econometrics* **5**(3), 358–359.

R Core Team (2020), 'R: A language and environment for statistical computing'. Vienna, Austria: R Foundation for Statistical Computing.

Ruiz, E. (1994), 'Quasi-maximum likelihood estimation of stochastic volatility models', *Journal of Econometrics* **63**(1), 289–306.

Smetanina, E. (2017), 'Real-time GARCH', *Journal of Financial Econometrics* **15**(4), 561–601.

Sucarrat, G. (2015), *lgarch: Simulation and Estimation of Log-GARCH Models.* R package version 0.6-2. URL: https://CRAN.R-project.org/package=lgarch.

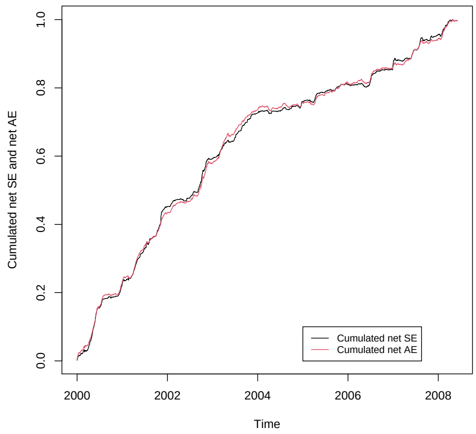
Taylor, S. (1986), 'Modeling financial time series John Wiley & Sons'. Great Britain.

Wuertz, D., Setz, T., Chalabi, Y., Boudt, C., Chausse, P., Miklovac, M., Setz, M. T. and RUnit, S. (2013), Package 'fGarch', Technical report, Technical report, working paper/manual, 09.11. 2009. URL: https://cran.r-project.org/web/packages/fGarch/index.html.

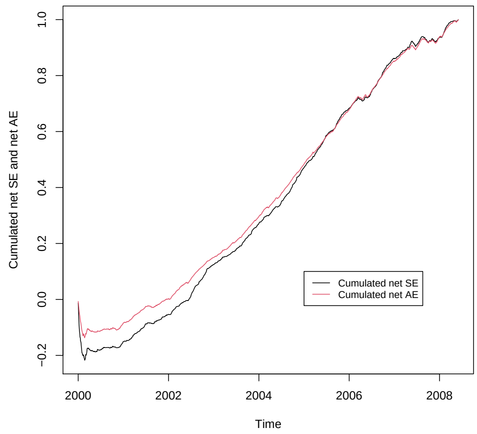
Appendix

A Further empirical results

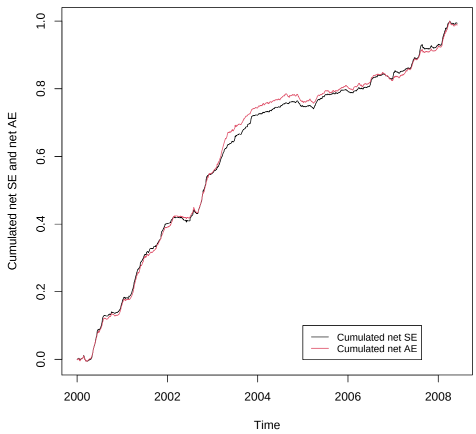
Figures



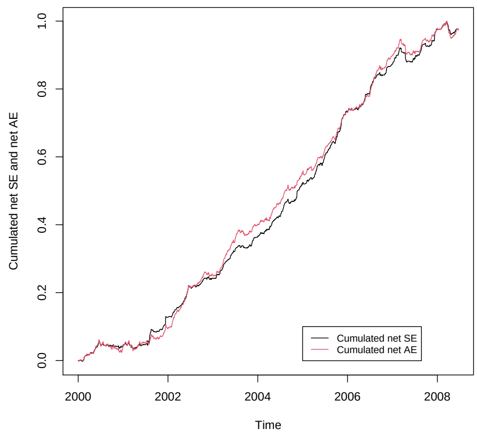
(a) Dow Jones



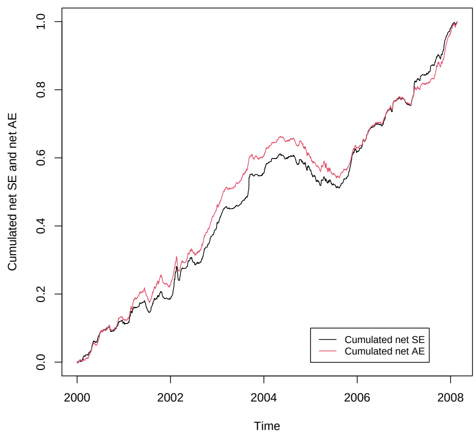
(b) Nasdaq



(c) S&P 500

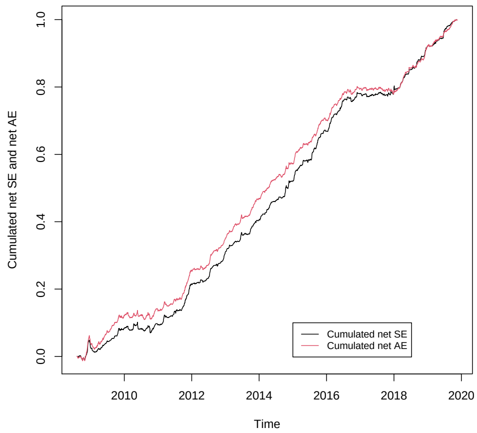


(d) FTSE

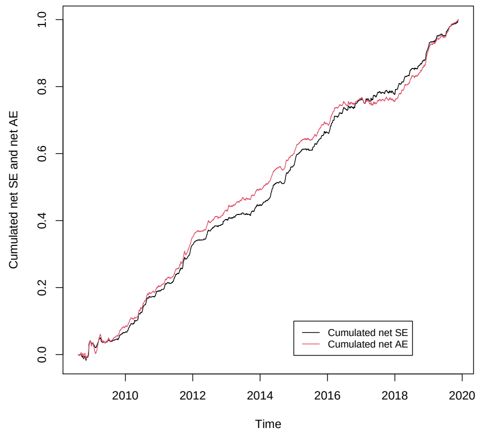


(e) Nikkei 225

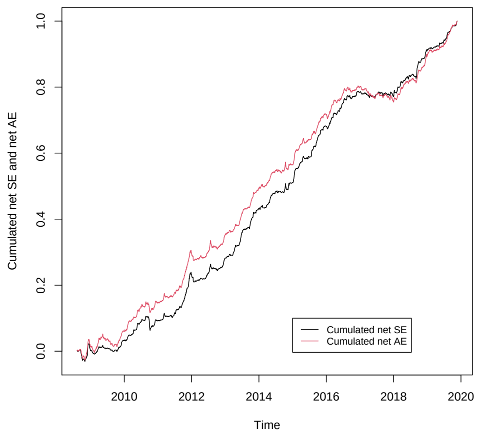
Figure A.1: Cumulative net squared and absolute error plots (pre-crisis sub-sample)



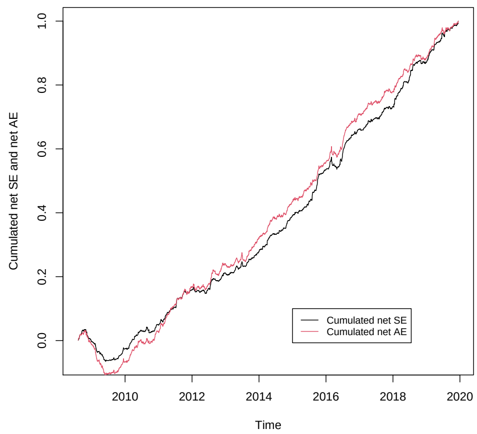
(a) Dow Jones



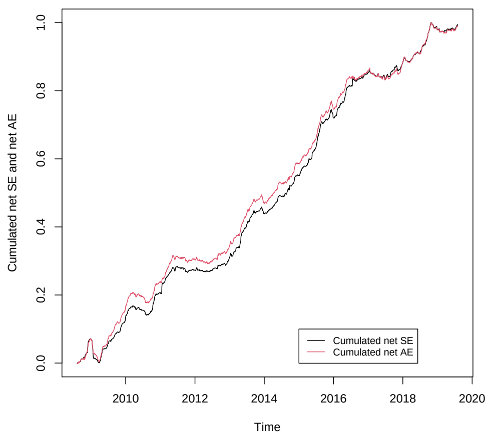
(b) Nasdaq



(c) S&P 500



(d) FTSE



(e) Nikkei 225

Figure A.2: Cumulative net squared and absolute error plots (post-crisis sub-sample)

Tables

Table A.1: Descriptive statistics of returns and log-squared returns (pre/post-crisis sub-sample)

		pre-crisis sub-sample 01/03/2000 to 07/31/2008							
		r_t				$\log r_t^2$			
	# obs.	Mean	Var	Skewness	Kurtosis	Mean	Var	Skewness	Kurtosis
Dow Jones	2145	0.011	1.146	-0.162	6.505	-1.583	5.906	-1.088	4.877
Nasdaq	2144	0.000	2.497	0.193	8.705	-0.813	5.646	-1.024	4.787
S&P 500	2145	-0.012	1.159	-0.065	5.460	-1.534	5.721	-1.050	4.888
FTSE 100	2157	-0.031	1.256	-0.339	5.902	-1.460	5.672	-1.170	5.581
Nikkei 225	2075	0.000	1.218	-0.237	4.306	-1.345	5.714	-1.356	6.416
		post-crisis sub-sample 08/01/2008 to 12/30/2019							
		r_t				$\log r_t^2$			
	# obs.	Mean	Var	Skewness	Kurtosis	Mean	Var	Skewness	Kurtosis
Dow Jones	2865	0.000	1.193	-0.146	14.171	-2.069	7.016	-0.875	4.413
Nasdaq	2867	0.020	1.134	-0.470	9.431	-1.800	6.328	-1.199	6.132
S&P 500	2868	0.024	1.269	-0.478	14.583	-2.028	7.062	-0.967	4.989
FTSE	2884	0.014	1.242	-0.385	11.609	-1.693	6.535	-1.383	6.874
Nikkei 225	2791	-0.028	1.250	-1.120	21.104	-1.895	6.440	-1.211	6.408

Table A.2: Volatility nowcasting estimation of pre-crisis sub-sample data

				Dow Jones				
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.388		0.515		**0.542**		0.330	
κ	0.033	*(0.007)*	0.054	*(0.006)*	0.038	*(0.005)*		
β	0.995	*(0.003)*	0.981	*(0.004)*	0.988	*(0.003)*	0.981	*(0.004)*
θ	0.963	*(0.007)*	0.930	*(0.004)*	0.951	*(0.004)*	0.931	*(0.008)*
C	-1.431	*(—)*	-1.362	*(0.032)*	-1.414	*(0.042)*	-1.536	*(0.031)*
ν_2					7.356	*(0.935)*		
				Nasdaq				
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.308		0.423		**0.561**		0.296	
κ	0.021	*(0.005)*	0.034	*(0.004)*	0.025	*(0.004)*		
β	0.998	*(0.001)*	0.996	*(0.001)*	0.997	*(0.001)*	0.995	*(0.002)*
θ	0.978	*(0.005)*	0.963	*(0.003)*	0.973	*(0.004)*	0.951	*(0.005)*
C	-1.298	*(—)*	-1.223	*(0.030)*	-1.274	*(0.038)*	-1.513	*(0.031)*
ν_2					9.003	*(1.166)*		
				S&P 500				
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.369		0.465		**0.546**		0.274	
κ	0.032	*(0.007)*	0.047	*(0.005)*	0.037	*(0.005)*		
β	0.994	*(0.003)*	0.986	*(0.004)*	0.991	*(0.003)*	0.983	*(0.004)*
θ	0.963	*(0.008)*	0.943	*(0.004)*	0.955	*(0.004)*	0.935	*(0.007)*
C	-1.403	*(—)*	-1.346	*(0.032)*	-1.391	*(0.041)*	-1.346	*(0.031)*
ν_2					7.611	*(0.997)*		
				FTSE				
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.515		0.582		**0.603**		0.412	
κ	0.051	*(0.008)*	0.076	*(0.006)*	0.067	*(0.006)*		
β	0.987	*(0.005)*	0.982	*(0.004)*	0.983	*(0.004)*	0.986	*(0.004)*
θ	0.939	*(0.009)*	0.912	*(0.005)*	0.922	*(0.005)*	0.920	*(0.008)*
C	-1.273	*(—)*	-1.199	*(0.032)*	-1.259	*(0.039)*	-1.415	*(0.030)*
ν_2					9.003	*(1.155)*		
				Nikkei 225				
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.387		0.468		**0.498**		0.306	
κ	0.030	*(0.007)*	0.043	*(0.006)*	0.030	*(0.005)*		
β	0.992	*(0.005)*	0.981	*(0.006)*	0.990	*(0.004)*	0.979	*(0.005)*
θ	0.962	*(0.010)*	0.940	*(0.005)*	0.961	*(0.004)*	0.942	*(0.008)*
C	-1.361	*(—)*	-1.323	*(0.033)*	-1.363	*(0.040)*	-1.437	*(0.031)*
ν_2					9.003	*(1.348)*		

Note: Bold-faced values indicate the best performing specification. Numbers in parentheses are standard deviations.

Table A.3: Volatility nowcasting performance of pre-crisis sub-sample data

			Dow Jones		
	BH	$\log - \chi^2(1)$	$\log - F(1, \nu_2)$	log-GARCH	RT-GARCH
MSE	1	*0.792*	*0.748*	*1.093*	*0.814*
MCS p-value		0.603	**1.000**		0.176
MAE	1	*0.792*	*0.832*	*1.058*	*0.889*
MCS p-value			**1.000**		
			Nasdaq		
	BH	$\log - \chi^2(1)$	$\log - F(1, \nu_2)$	log-GARCH	RT-GARCH
MSE	1	*0.792*	*0.635*	*1.018*	*0.829*
MCS p-value			**1.000**		
MAE	1	*0.792*	*0.760*	*1.014*	*0.894*
MCS p-value			**1.000**		
			S&P 500		
	BH	$\log - \chi^2(1)$	$\log - F(1, \nu_2)$	log-GARCH	RT-GARCH
MSE	1	*0.792*	*0.719*	*1.150*	*0.826*
MCS p-value			**1.000**		
MAE	1	*0.792*	*0.812*	*1.087*	*0.899*
MCS p-value			**1.000**		
			FTSE		
	BH	$\log - \chi^2(1)$	$\log - F(1, \nu_2)$	log-GARCH	RT-GARCH
MSE	1	*0.792*	*0.818*	*1.211*	*0.865*
MCS p-value		0.311	**1.000**		0.680
MAE	1	*0.792*	*0.873*	*1.131*	*0.923*
MCS p-value		0.084	**1.000**		0.313
			Nikkei 225		
	BH	$\log - \chi^2(1)$	$\log - F(1, \nu_2)$	log-GARCH	RT-GARCH
MSE	1	*0.792*	*0.818*	*1.132*	*0.813*
MCS p-value			0.914		**1.000**
MAE	1	*0.792*	*0.878*	*1.075*	*0.896*
MCS p-value			**1.000**		0.379

Note: Bold-faced value(s) indicate the model(s) with the largest MCS p-value. Numbers in italics are relative to the benchmark (BH). Missing entries indicate that the respective model is eliminated from the model confidence set.

Table A.4: Volatility nowcasting estimation of post-crisis sub-sample data

	Dow Jones							
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.500		**0.604**		0.574		0.375	
κ	0.082	*(0.009)*	0.122	*(0.008)*	0.078	*(0.007)*		
β	0.985	*(0.005)*	0.959	*(0.006)*	0.982	*(0.004)*	0.963	*(0.004)*
θ	0.910	*(0.010)*	0.855	*(0.005)*	0.911	*(0.005)*	0.875	*(0.008)*
C	-1.485	*(—)*	-1.406	*(0.030)*	-1.512	*(0.047)*	-1.724	*(0.026)*
ν_2					5.199	*(0.023)*		
	Nasdaq							
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.392		0.492		**0.574**		0.232	
κ	0.061	*(0.007)*	0.082	*(0.007)*	0.057	*(0.005)*		
β	0.985	*(0.005)*	0.967	*(0.005)*	0.985	*(0.003)*	0.968	*(0.004)*
θ	0.929	*(0.009)*	0.894	*(0.004)*	0.932	*(0.004)*	0.900	*(0.008)*
C	-1.387	*(—)*	-1.347	*(0.030)*	-1.412	*(0.040)*	-1.684	*(0.026)*
ν_2					6.233	*(0.657)*		
	S&P 500							
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.536		**0.611**		0.609		0.415	
κ	0.080	*(0.009)*	0.112	*(0.007)*	0.078	*(0.006)*		
β	0.985	*(0.004)*	0.971	*(0.004)*	0.987	*(0.003)*	0.975	*(0.003)*
θ	0.913	*(0.010)*	0.874	*(0.005)*	0.915	*(0.005)*	0.893	*(0.007)*
C	-1.435	*(—)*	-1.347	*(0.031)*	-1.437	*(0.043)*	-1.539	*(0.026)*
ν_2					5.959	*(0.482)*		
	FTSE							
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.498		0.565		**0.577**		0.374	
κ	0.063	*(0.009)*	0.085	*(0.007)*	0.057	*(0.007)*		
β	0.982	*(0.006)*	0.973	*(0.005)*	0.984	*(0.004)*	0.971	*(0.005)*
θ	0.924	*(0.012)*	0.896	*(0.005)*	0.931	*(0.005)*	0.895	*(0.009)*
C	-1.388	*(—)*	-1.339	*(0.031)*	-1.427	*(0.039)*	-1.544	*(0.026)*
ν_2					7.175	*(0.852)*		
	Nikkei 225							
	BH		$\log-\chi^2(1)$		$\log-F(1,\nu_2)$		log-GARCH	
R^2	0.419		**0.569**		0.525		0.273	
κ	0.059	*(0.009)*	0.128	*(0.008)*	0.056	*(0.007)*		
β	0.987	*(0.005)*	0.957	*(0.007)*	0.984	*(0.004)*	0.972	*(0.005)*
θ	0.932	*(0.011)*	0.848	*(0.005)*	0.931	*(0.005)*	0.887	*(0.009)*
C	-1.526	*(—)*	-1.346	*(0.032)*	-1.506	*(0.049)*	-1.675	*(0.027)*
ν_2					5.080	*(0.484)*		

Note: Bold-faced values indicate the best performing specification. Numbers in parentheses are standard deviations.

Table A.5: Volatility nowcasting performance of post-crisis sub-sample data

	BH	$\log-\chi^2(1)$	$\log-F(1,\nu_2)$	log-GARCH	RT-GARCH
			Dow Jones		
MSE	1	*0.791*	*0.852*	*1.249*	*0.780*
MCS p-value		**1.000**	0.059		**1.000**
MAE	1	*0.889*	*0.902*	*1.143*	*0.882*
MCS p-value		**1.000**	0.494		**1.000**
			Nasdaq		
MSE	1	*0.836*	*0.701*	*1.263*	*0.769*
MCS p-value			**1.000**		
MAE	1	*0.905*	*0.821*	*1.138*	*0.871*
MCS p-value			**1.000**		
			S&P 500		
MSE	1	*0.837*	*0.843*	*1.260*	*0.791*
MCS p-value		0.368	0.550		**1.000**
MAE	1	*0.916*	*0.902*	*1.135*	*0.887*
MCS p-value			0.301		**1.000**
			FTSE		
MSE	1	*0.866*	*0.842*	*1.247*	*0.901*
MCS p-value		**1.000**	**1.000**		0.116
MAE	1	*0.935*	*0.907*	*1.140*	*0.953*
MCS p-value			**1.000**		
			Nikkei 225		
MSE	1	*0.741*	*0.819*	*1.251*	*0.702*
MCS p-value		0.139			**1.000**
MAE	1	*0.851*	*0.856*	*1.150*	*0.821*
MCS p-value		0.235	0.168		**1.000**

Note: Bold-faced value(s) indicate the model(s) with the largest MCS p-value. Numbers in italics are relative to the benchmark (BH). Missing entries indicate that the respective model is eliminated from the model confidence set.

Table A.6: Density selection of pre- and post-crisis sub-sample data

	Gaussian distribution			t-distribution			
	pre-crisis sub-sample						
	SC	μ	σ^2	SC	μ	σ^2	ν
Dow Jones	6182.727	0.017	1.034	**6108.159**	0.036	1.032	7.620
Nasdaq	6047.199	0.019	1.002	**6033.608**	0.026	1.002	13.796
S&P 500	6168.085	-0.010	1.031	**6117.306**	0.007	1.030	8.430
FTSE	6299.512	-0.020	1.054	**6280.211**	-0.002	1.055	11.507
Nikkei 225	6009.261	0.002	1.042	**5967.624**	0.017	1.040	9.614

	post-crisis sub-sample						
	SC	μ	σ^2	SC	μ	σ^2	ν
Dow Jones	8753.005	0.031	1.125	**8417.432**	0.088	1.163	3.834
Nasdaq	8558.424	0.021	1.086	**8351.026**	0.086	1.100	4.831
S&P 500	8732.930	0.029	1.119	**8440.743**	0.087	1.146	4.137
FTSE	8604.832	0.011	1.085	**8499.733**	0.034	1.094	6.202
Nikkei 225	8364.680	-0.025	1.093	**7985.776**	0.005	1.086	4.428

Note: Bold-faced value indicates the selected density.

B Simulation results

Estimation of ARMA parameters

Table B.1: MSE of parameter estimation

	Gaussian CML		log $-\chi^2(1)$ CML		EML		Backcasting	
	$\beta = 0.95$, $\kappa = 0.056$ ($\theta = 0.9$)							
T	β	θ	β	θ	β	θ	β	θ
500	0.757	0.903	0.293	0.328	0.386	0.435	0.907	1.137
1000	0.362	0.478	0.171	0.204	0.173	0.207	0.217	0.314
1500	0.245	0.333	0.125	0.154	0.102	0.124	0.121	0.186
2000	0.164	0.233	0.093	0.116	0.067	0.082	0.072	0.114
	$\beta = 0.99$, $\kappa = 0.042$ ($\theta = 0.95$)							
T	β	θ	β	θ	β	θ	β	θ
500	0.418	0.592	0.109	0.190	0.123	0.211	0.217	0.401
1000	0.083	0.156	0.041	0.085	0.029	0.063	0.033	0.097
1500	0.030	0.069	0.023	0.054	0.013	0.031	0.014	0.047
2000	0.017	0.045	0.014	0.038	0.007	0.019	0.008	0.030

Note: All values are given in $\times 100$.

Evaluation criteria

Table B.2: Nowcasting performance of ARMA(1, 1) model with $\xi_t \overset{i.i.d.}{\sim} N(0,1)$.

		Gaussian			$\log -\chi^2(1)$	
T	R^2	MSE	MAE	R^2	MSE	MAE
			$\beta = 0.95$, $\kappa = 0.056$ ($\theta = 0.9$)			
500	0.729	0.046	0.165	0.877	0.021	0.113
1000	0.744	0.039	0.157	0.880	0.019	0.108
1500	0.766	0.036	0.151	0.886	0.017	0.106
2000	0.788	0.033	0.145	0.907	0.014	0.098
			$\beta = 0.99$, $\kappa = 0.042$ ($\theta = 0.95$)			
500	0.814	0.053	0.171	0.886	0.034	0.138
1000	0.882	0.034	0.138	0.938	0.018	0.098
1500	0.923	0.025	0.120	0.959	0.014	0.083
2000	0.937	0.022	0.112	0.964	0.014	0.074

Table B.3: Nowcasting performance of ARMA(1, 1) model with $\xi_t \overset{i.i.d.}{\sim} t(10)$.

		Gaussian			$\log -\chi^2(1)$	
T	R^2	MSE	MAE	R^2	MSE	MAE
			$\beta = 0.95$, $\kappa = 0.056$ ($\theta = 0.9$)			
500	0.715	0.058	0.185	0.806	0.041	0.152
1000	0.744	0.046	0.168	0.807	0.035	0.147
1500	0.765	0.040	0.159	0.810	0.032	0.145
2000	0.772	0.038	0.157	0.812	0.032	0.145
			$\beta = 0.99$, $\kappa = 0.042$ ($\theta = 0.95$)			
500	0.831	0.056	0.174	0.847	0.054	0.165
1000	0.882	0.039	0.145	0.909	0.031	0.123
1500	0.909	0.030	0.129	0.934	0.023	0.107
2000	0.933	0.026	0.120	0.955	0.017	0.093

Estimation of the constant C

Table B.4: MSE of $\widehat{C}$.

	$\beta = 0.95$, $\kappa = 0.056$ ($\theta = 0.9$)			$\beta = 0.99$, $\kappa = 0.042$ ($\theta = 0.95$)		
T	Gaussian	$\log -\chi^2(1)$	log-GARCH	Gaussian	$\log -\chi^2(1)$	log-GARCH
500	6.054	0.007	0.043	1.648	0.010	0.213
1000	1.360	0.003	0.022	0.279	0.005	0.122
1500	0.137	0.002	0.014	0.071	0.003	0.085
2000	0.088	0.002	0.011	0.002	0.002	0.065

Density selection

Table B.5: Density selection performance

	Gaussian distribution			t-distribution			
T	Frequency	μ	σ^2	Frequency	μ	σ^2	ν
500	0.995	0.000	0.999	0.583	-0.001	1.248	12.232
1000	0.997	0.001	0.999	0.866	0.000	1.248	11.485
1500	0.997	0.000	0.999	0.967	0.001	1.248	10.930
2000	0.998	0.000	1.000	0.991	0.000	1.250	10.631

C Unit root test

Table C.1: Hansen and Lunde (2014) test for RV

	full sample		pre-crisis		post-crisis	
Index	$\hat{\pi}$	HL stat	$\hat{\pi}$	HL stat	$\hat{\pi}$	HL stat
Dow Jones	0.963	-177.563	0.955	-94.911	0.964	-94.509
Nasdaq	0.947	-251.238	0.955	-95.284	0.928	-188.433
S&P 500	0.956	-207.697	0.952	-100.856	0.956	-115.028
FTSE	0.960	-188.828	0.950	-107.323	0.968	-83.888
Nikkei 225	0.930	-321.978	0.917	-170.125	0.928	-182.289

Notes: $\hat{\pi}$ stands for the largest AR root and HL is the Hansen and Lunde (2014) unit root statistic under measurement errors. The 1% and 5% critical values are -20.7 and -14.1, respectively.