Journal of Risk **17**(4), 1–18

Research Paper

A simple normal inverse Gaussian-type approach to calculate value-at-risk based on realized moments

Christian Lau

School of Economics and Business, Martin-Luther University Halle-Wittenberg,
Grosse Steinstrasse 73, 06108 Halle (Saale), Germany;
email: christian.lau@wiwi.uni-halle.de

(Received January 10, 2014; revised September 2, 2014; accepted September 3, 2014)

ABSTRACT

Expanding the realized variance concept through realized skewness and kurtosis is a straightforward process. We calculate one-day forecasts for these moments with a simple exponentially weighted moving average approach. Once the forecasts are computed, we apply them in a method of moments to fit a sophisticated distribution. The normal inverse Gaussian distribution is appropriate for this purpose because it exhibits higher moments, and a simple analytical solution for the method of moments exists. We then calculate the value-at-risk for the Deutsche Aktienindex (DAX) using this technique. Although the model is comparatively simple, the empirical analysis shows good results in terms of backtesting.

Keywords: value-at-risk; intraday data; realized moments; normal inverse Gaussian distribution; method of moments; EWMA.

1 INTRODUCTION

Despite well-known drawbacks, such as the lack of subadditivity (Artzner 1999), value-at-risk (VaR) is still the dominant method in risk management for estimating

Print ISSN 1465-1211 | Online ISSN 1755-2842

downside risk. Assuming a certain level of confidence $1 - \theta$, $\mathrm{VaR}(\theta)_{t+1}$ of an asset is defined as the number such that the probability that the asset's next-day return r_{t+1} will be below $-\mathrm{VaR}(\theta)_{t+1}$ is θ: $\Pr(r_{t+1} < -\mathrm{VaR}(\theta)_{t+1} \mid \Omega_t) = \theta$, where Ω_t is the information set available in t. From a statistical point of view, VaR is closely related to the quantile function F^{-1} of a distribution of returns: $\mathrm{VaR}(\theta)_{t+1} = -F^{-1}(\theta)$. Although there are numerous, sophisticated methods for calculating VaR, its computation is often based on simple concepts, such as Gaussian assumptions or historical simulations (Pritsker 2006).

The inherent complexity of the more accurate models is at least one reason for their infrequent use; as a result, an incorrect (but clear) VaR method is often employed. Another, more objective, explanation for avoiding the more complex models is model risk, ie, the fewer parameters employed in an estimate, the less chance there is for something to go wrong. Nevertheless, strong empirical evidence from as early as Mandelbrot (1963) argues that the Gaussian distribution is unable to correctly capture the characteristics of financial returns. Gaussian distributions cannot reproduce stylized facts (see, for example, Cont 2001) of financial time series, such as autocorrelation and leptokurtosis, which is particularly problematic when dealing with the distribution tails that typically occur in a VaR.

The goal of VaR is to determine the regulatory capital required to cover unexpected losses resulting from risks. These regulatory requirements are set out in Basel II/III for banks and in Solvency II for insurance companies. Because equity is expensive, it is desirable to assign only the amount that is necessary (but sufficient) to cover losses. To calculate this amount, we need a model that is accurate; better yet would be one that also keeps the VaR as low as possible. This paper offers an alternative method for calculating VaR that is simple yet comprehensive and keeps the VaR comparably small. Based on intraday data from the Deutsche Aktienindex (DAX), we compute forecasts for the empirical moments and use them to parameterize the normal inverse Gaussian (NIG) distribution within a method of moments. Once this is complete, obtaining the VaR is easy.

Recent efforts to calculate VaR utilize parametric models based on the theory of realized variance (RV). RV (see, for example, Andersen and Bollerslev 1998; Andersen *et al* 2003; Barndorff-Nielsen and Shephard 2002) is defined as the sum of squared intraday returns over a specified time interval, for example, one day.[1] RV is a reliable, model-free proxy for the actual (nonobservable) variance. To calculate VaR within a given financial return time series, a standard procedure is to include RV as an external variable in an autoregressive conditional heteroscedasticity (ARCH) process

[1] The term "realized volatility", which is more common in the literature, is merely the positive square root of the realized variance.

so as to model the variance more accurately. This concept, known as generalized autoregressive conditional heteroscedasticity-X (GARCH-X), can be traced back to Engle (2002). Shephard and Sheppard (2010) state that the squared return term in a GARCH-X model has only marginal explanatory power. They thus introduce a high-frequency-based volatility (HEAVY) model that calculates tomorrow's variance only as a function of today's variance and today's realized variance. Furthermore, Hansen *et al* (2012) formulate the realized GARCH as a generalization of the GARCH-X. Nevertheless, there is some disagreement in the literature as to whether RV actually improves VaR calculations. Giot and Laurent (2004) compare VaR calculations with ARCH models and RV models and find no improvement in the latter. Clements *et al* (2008) argue that models incorporating RV improve VaR calculations in the case of currencies. Beltratti and Morana (2005) investigate long-memory models with high-frequency data and identify superior performance.

A smaller strand of the literature takes a different approach and allows higher moments to have a time-varying property. Time-varying moments are not a new concept (see, for example, Turtle *et al* 1994; Bond and Patel 2003), but Brooks *et al* (2005) were the first to build a structured framework for a time-varying kurtosis and provide applications for different daily financial returns. Incorporating realized skewness and kurtosis in models is a straightforward extension of RV, according to Amaya *et al* (2015), who use both to predict stock returns. As for time-varying moments, Liu (2012) applies the Cornish and Fisher (CF) approximation to standardized realized higher moments when calculating VaR (Cornish and Fisher 1937). The present paper uses the flexible NIG distribution (Barndorff-Nielsen 1997), which provides clearly better results in terms of backtesting.

In Section 2 of this paper, we derive the theoretical background and build the model. Section 3 provides a description and analysis of the data, and we compute, backtest and compare the VaR with the results of other VaR methodologies. Section 4 discusses possible extensions of the presented method.

2 BUILDING THE MODEL

2.1 Standardized realized moments

The availability of intraday data has resulted in this information being integrated in several recent models. Along with meeting the statistical demand of using all available data, RV can also serve as a model-free proxy for the actual (nonobservable) variance. The ith intraday log return on day t is

$$r_{t,i} = \ln p_{t,i} - \ln p_{t,i-1}, \tag{2.1}$$

with $p_{t,i}$ as the ith price on day t. The realized variance is simply calculated as the sum over N squared returns:

$$\mathrm{RV}_t = \sum_{i=1}^{N} r_{t,i}^2. \tag{2.2}$$

With regard to the ordinary moments, generalizing this concept to realized moments yields

$$\mathrm{RM(or)}_t = \sum_{i=1}^{N} r_{t,i}^{\mathrm{or}}, \tag{2.3}$$

with RM(or) the realized moment of order,

$$\mathrm{or} = 1, 2, 3, \dots .$$

Obviously, choosing or $= 2$ results in RV_t.

In Section 2.3, we calculate the parameters of a distribution with a method-of-moments procedure, for which we need standardized rather than unstandardized moments. The term "standardized realized moments" refers to realized variance, realized skewness and realized kurtosis, the latter two of which are standardized by definition. Amaya *et al* (2015) define standardized realized skewness and kurtosis as follows:

$$\mathrm{RS}_t = \frac{\sqrt{N} \sum_{i=1}^{N} r_{t,i}^3}{\mathrm{RV}_t^{3/2}} \tag{2.4}$$

and

$$\mathrm{RK}_t = \frac{N \sum_{i=1}^{N} r_{t,i}^4}{\mathrm{RV}_t^2}. \tag{2.5}$$

The purpose of dividing the realized third (fourth) moment by $\mathrm{RV}_t^{3/2}$ (RV_t^2) is to achieve standardization similar to that of ordinary moments. Scaling by $\sqrt{N}$ (N) ensures that RS_t (RK_t) is on a daily level.

2.2 Forecasting realized moments

In Section 2.1, we calculated only today's realized moments. To compute the VaR for the next day, we need accurate forecasts of tomorrow's realized moments. Ever since the seminal work of Engle (1982) and Engle and Bollerslev (1986), it has become standard procedure in financial time series analysis to use a time-varying variance. The idea of constructing a time-varying framework for the kurtosis to allow for separate behavior dates back to Brooks *et al* (2005). A simple but nonetheless feasible updating schema is the exponentially weighted moving average (EWMA), adapted by JP Morgan and Reuters (1996) to the variance of the daily return r_t:

$$\sigma_{t+1}^2 = \lambda_t \sigma_t^2 + (1 - \lambda_t) r_t^2. \tag{2.6}$$

Given yesterday's variance forecast for today, σ_t^2, and today's squared return, r_t^2, the variance forecast for tomorrow is σ_{t+1}^2. $0 < \lambda_t < 1$ is the decay factor. To begin with, we use RV_t instead of r_t^2 in (2.6) because RV_t is a better proxy than the squared return for the recent variance. Applying and extending the EWMA for the remaining higher realized moments is straightforward. According to Liu (2012),

$$M(\mathrm{or})_{t+1} = \lambda_{\mathrm{or},t} M(\mathrm{or})_t + (1 - \lambda_{\mathrm{or},t}) \mathrm{RM}(\mathrm{or})_t, \tag{2.7}$$

with $M(\mathrm{or})_t$ ($M(\mathrm{or})_{t+1}$) as yesterday's (today's) forecast for the realized moment of order "or" for today (tomorrow) and $\mathrm{RM}(\mathrm{or})_t$ as today's realized moment of order "or", although we are primarily interested in or $= 2, 3, 4$. As for $\lambda_{\mathrm{or},t}$, we assess two different strategies.

(1) $\lambda_{\mathrm{or},t}$ is fixed to $\lambda = 0.94$ for all "or" and t, as suggested by RiskMetrics, for the variance. If this strategy is undertaken, note that λ is not estimated, as in other empirical work, and the amount of data needed for the calculation decreases dramatically.

(2) Find the optimal $\lambda_{\mathrm{or},t}$ for each equation "or" in every t by minimizing average squared errors using the information set available up to and including time t.[2]

After forecasts for the realized moments are obtained, we use them to compute forecasts for the standardized realized moments by simply replacing the sums ($\sum_{i=1}^{N} r_{t,i}^{\mathrm{or}}$) in (2.2), (2.4) and (2.5) with the forecasts for the realized moments ($M(\mathrm{or})_{t+1}$). We denote those forecasts as v_{t+1} (variance), s_{t+1} (skewness) and k_{t+1} (kurtosis). Finally, we apply them to parameterize a distribution that accounts for higher-order moments. The NIG distribution is described below.

2.3 Normal inverse Gaussian distribution

For a distribution to be appropriate for our purposes, it must fulfill the following criteria:

- be broadly accepted in the finance literature;
- exhibit higher moments;
- provide a simple, closed method-of-moments estimation.

[2] To guarantee at least some persistence in the variance process, $\lambda_{\mathrm{or},t}$ can be bounded above 0.5. However, in this study we use $0 < \lambda_{\mathrm{or},t} < 1$.

The NIG distribution meets all three requirements. Because of its flexibility, it is often used in a wide range of financial applications. Forsberg and Bollerslev (2002) formulate a GARCH process with NIG innovations to model daily EUR/USD exchange rates. Venter and de Jongh (2002) compare VaR based on the NIG distribution with extreme value theory VaR and find in favor of the former. Chen *et al* (2005) calculate VaR with the NIG distribution for exchange rate and German bank portfolio data using adaptive volatility estimation and show a perfect fit for their model. Chen and Lu (2012) investigate simulated and real-world data and find that NIG-based VaR estimation is robust and accurate for a forecasting horizon of one day (for other applications, see, for example, Lillestøl (2000), Aas *et al* (2006) and Eriksson *et al* (2009)). Using an NIG distribution instead of a Gaussian distribution results in a more reasonable modeling of financial returns and therefore contributes to a more realistic VaR calculation.

The NIG distribution (see, for example, Paolella 2007) has four parameters: steepness (α), asymmetry (β), scale (δ) and location (μ). The density can be written as

$$f_{\mathrm{NIG}}(x;\alpha,\beta,\delta,\mu) = \frac{\alpha\delta\exp(\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu))}{\pi\sqrt{\delta^2+(x-\mu)^2}}K_1(\alpha\sqrt{\delta^2+(x-\mu)^2}), \tag{2.8}$$

with $x \in \mathbb{R}$, $0 \leqslant |\beta| < \alpha$, $\delta > 0$, $\mu \in \mathbb{R}$ and K_1 as the modified Bessel function of the third type with index 1. A smaller value of α implies a fat-tailed density. An increasing value of $|\beta|$ yields skewness. Given the parameters of the NIG distribution, it is possible to calculate the central moments as follows (see, for example, Kalemanova *et al* 2007):

$$m = \mu + \delta\frac{\beta}{\gamma}, \tag{2.9}$$

$$v = \delta\frac{\alpha^2}{\gamma^3}, \tag{2.10}$$

$$s = 3\frac{\beta}{\alpha\sqrt{\delta\gamma}} \tag{2.11}$$

and

$$k = 3 + 3\left(1 + 4\left(\frac{\beta}{\alpha}\right)^2\right)\frac{1}{\delta\gamma}. \tag{2.12}$$

Parameter estimation of the NIG distribution is typically performed with maximum likelihood, but this is not always feasible because of the complexity of the likelihood. By solving (2.9)–(2.12) for the parameters, we obtain a closed analytical method of moments solution under rather fair conditions ($k-\frac{5}{3}s^2-3>0$ and $3k-4s^2-9>0$):[3]

$$\hat{\mu} = m - \frac{3s\sqrt{v}}{3k-4s^2-9}, \tag{2.13}$$

$$\hat{\delta} = \frac{3^{3/2}\sqrt{v(k-5s^2/3-3)}}{3k-4s^2-9}, \tag{2.14}$$

$$\hat{\beta} = \frac{s}{\sqrt{v}(k-5s^2/3-3)} \tag{2.15}$$

and

$$\hat{\alpha} = \frac{\sqrt{3k-4s^2-9}}{\sqrt{v}(k-5s^2/3-3)}. \tag{2.16}$$

When considering short time horizons, such as daily, the mean of returns is dominated by the variance. Therefore, rejection of a zero mean return hypothesis is not possible (see, for example, Christoffersen 2003). Thus, the mean m is set to 0.

2.4 Value-at-risk

Once we have the forecast for the standardized realized moments (v_{t+1}, s_{t+1} and k_{t+1}), we parameterize the NIG distribution with (2.13)–(2.16). In this way, we estimate the future return distribution for $t+1$. On a confidence level $(1-\theta)$, the VaR for the next day is

$$\text{VaR}_{\text{NIG}}(\theta)_{t+1} = -F_{\text{NIG}}^{-1}(\theta; \hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu}), \tag{2.17}$$

where $F_{\text{NIG}}^{-1}(\theta; \hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu})$ is the θ quantile of the NIG distribution. This model is called RM–EWMA–NIG. In general, the approach described in this paper can incorporate any distribution, as long as the method of moments can be feasibly implemented. As in Liu (2012), we can expand the Gaussian distribution to account for higher moments by utilizing the CF approximation (Cornish and Fisher 1937). In this case, the unknown quantile $\tilde{z}_\theta$ can be approximated as

$$\tilde{z}_\theta \approx z_\theta + \tfrac{1}{6}(z_\theta^2-1)s_{t+1} + \tfrac{1}{24}(z_\theta^3-3z_\theta)(k_{t+1}-3) - \tfrac{1}{36}(2z_\theta^3-5z_\theta)s_{t+1}^2, \tag{2.18}$$

with z_θ the θ quantile of the standard normal distribution; consequently,

$$\text{VaR}_{\text{CF}}(\theta)_{t+1} = -\tilde{z}_\theta\sqrt{v_{t+1}}. \tag{2.19}$$

Note that neither procedure needs a complex numerical optimization algorithm.

[3] Ralf Werner offers a free NIG Toolbox that includes this computation. Eriksson *et al* (2009) provide a solution for the method of moments that differs slightly from the one used here. This difference is simply a result of using excess kurtosis instead of kurtosis.

We compare the models with one semiparametric and two parametric models. For this purpose, we choose models that are easy to estimate and allow higher moments and a time-varying variance. Based on the GARCH-X of Engle (2002), Shephard and Sheppard (2010) drop the squared return term because it has only marginal explanatory power. They formulated the HEAVY model as

$$v_{t+1} = \omega + \alpha_H \mathrm{RV}_t + \beta_H v_t, \tag{2.20}$$

with $\omega, \alpha_H \geqslant 0$ and $\beta_H \in [0, 1)$. The second part of their model provides a specification for RV_{t+1} to calculate multistep-ahead forecasts, which is not required in our case.[4] We calculate

$$\mathrm{VaR}_{\mathrm{HEAVY}}(\theta)_{t+1} = -t_{\theta,\nu}\sqrt{v_{t+1}}, \tag{2.21}$$

where $t_{\theta,\nu}$ is the θ quantile of the t-distribution with ν degrees of freedom. We are interested in discovering whether using intraday data actually improves the VaR. Thus, the fourth model under consideration is a GARCH(1,1) framework with t-distributed residuals and daily data (T-GARCH). We replace v_{t+1} in (2.21) with the GARCH forecast to calculate the VaR. A filtered historical simulation (Barone-Adesi *et al* 1999; Hull and White 1998) that includes the same T-GARCH process is the final and most complex alternative (T-FHS). The VaR is calculated as in (2.21), but we use the $N(1-\theta)$th value from the list of the descending-ordered standardized returns instead of $t_{\theta,\nu}$. This model produces very good results (Kuester *et al* 2006). However, in their analysis, Kuester *et al* (2006) assume a skewed t instead of a t-distributed innovation process in the GARCH model. Since it is well-established that heavy tails are far more crucial in financial modeling than skewness, we refrain from modeling skewness because, in contrast to the RM–EWMA–NIG model, incorporating skewness disproportionally increases complexity. We use the data from 250 (1000) days to compute the RV (daily) models. When we fix λ at 0.94, the data for one day is sufficient for calculating the VaR for the RM–EWMA models.

Using intraday data, by definition, means that there is a great deal of data to be processed. What would happen if we instead used end-of-day data in the RM–EWMA models? That is, we could simply replace the realized moments in (2.7) by the daily end-of-day return to the power of the order of the corresponding moment. However, this approach returns very poor p-values in the evaluation process. Squared returns, for instance, tend to be a very noisy indicator of variance (Andersen *et al* 2001) and, consequently, computed forecasts will be not reliable. RV, on the other hand, is a reliable proxy for the actual variance and is, therefore, the better choice.

[4] Kevin Sheppard's website offers the MFE Toolbox, which includes the code for the HEAVY model (www.kevinsheppard.com/MFE_Toolbox).

3 EMPIRICAL ANALYSIS

3.1 Data

This analysis uses intraday data for the DAX from January 2, 2006 through December 30, 2011.[5] Note that this time span includes the global financial crisis, which was an extremely volatile period. On a typical working day the first (final) value is fixed at 09:00 (17:45); no trading occurs on weekends or public holidays. The time series provides one unique value for every second.

To simplify the calculation, we forgo the use of an elaborate sampling algorithm (see, for example, Zhang *et al* 2005). As the sampling interval tends toward zero (due to market microstructure noise), the RV is a biased estimator of variance. We sample every 300th value (this corresponds to sampling every five minutes).[6]

The unconditional higher-order moments of the end-of-day data show minimal skewness (0.1006) and considerable kurtosis (8.9781), which is typical for financial time series. This supports the use of sophisticated distributions, such as the NIG distribution, which exhibit higher moments to model the data. To decide whether to choose a conditional or unconditional model, we must determine if there is autocorrelation in the data. Figure 1 on the next page shows the realized moments for the entire period. All three parts of the figure are dominated by fluctuations at the end of 2008. The figure nicely reveals how peaks in the realized variance coincide with peaks in both other realized moments. Autocorrelation is obviously an issue for the realized variance and fourth moment, but no concrete conclusion can be drawn as to whether autocorrelation is present in the realized third moment; thus, we need statistical tests. An analysis of the autocorrelation function (ACF) and partial autocorrelation function (PACF) reveals that the realized variance exhibits strong autocorrelation, while we find weak (considerable) autocorrelation in the realized third (fourth) moment.[7]

Table 1 on the next page shows the p-values of Ljung–Box and Breusch–Godfrey tests with the null hypothesis of no autocorrelation in the realized mean,[8] realized variance and realized third and fourth moment. For the realized mean, the Breusch–Godfrey test clearly suggests acceptance of the null hypothesis for the first three lags. The p-value of the Ljung–Box test is 0.0476, indicating that the null hypothesis is barely rejected, assuming a significance level of 0.05. However, due to the clear result from the Breusch–Godfrey test, and based on the extant literature, we abandon modeling the realized mean and subsequently assume it to be zero (see Section 2.3).

---

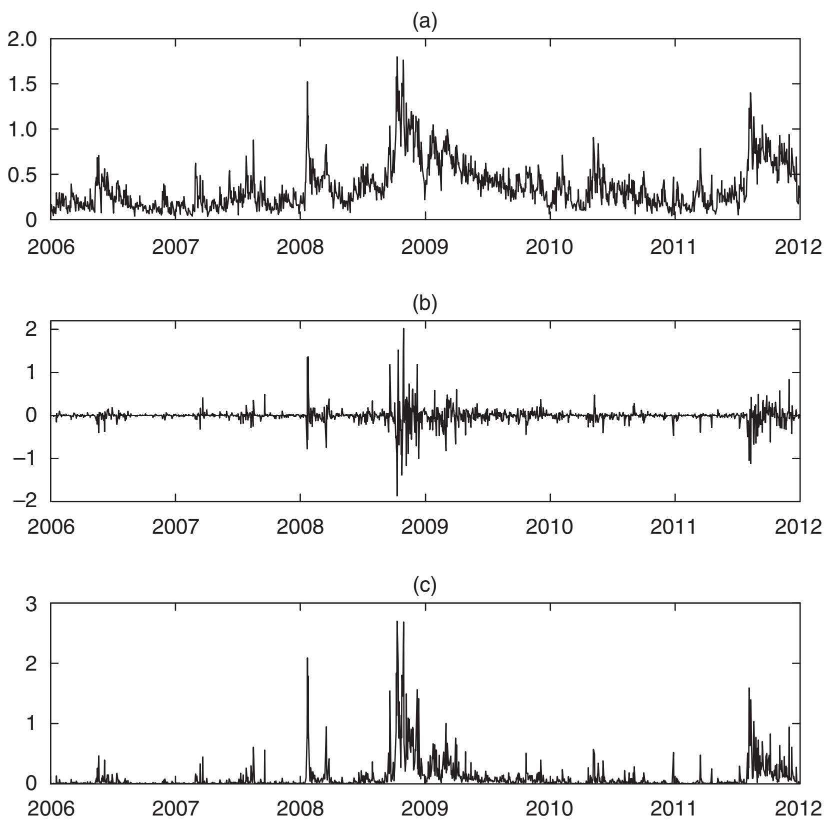
[5] The data is provided by the Karlsruhe Institute of Technology (KIT).

[6] This sampling frequency is in accordance with the literature (see, for example, Andersen and Bollerslev 1998). However, the results are robust to variations in sampling frequency.

[7] For brevity, the ACF and PACF are omitted.

[8] This is the sum over all sampled returns of one day.

FIGURE 1 Realized moments.



(a) Realized variance. (b) Realized third moment. (c) Realized fourth moment. All on a log scale y-axis for clarity. Corresponding formulas are given in (2.3).

TABLE 1 Tests on autocorrelation in realized moments.

		Breusch–Godfrey		
	LB	**Lag 1**	**Lag 2**	**Lag 3**
Realized mean	0.0476	0.2955	0.4882	0.6997
Realized variance	0.0000	0.0000	0.0000	0.0000
RM(3)	0.0000	0.0050	0.0157	0.0000
RM(4)	0.0000	0.0000	0.0000	0.0000

The table shows p-values of the Ljung–Box (LB) test and the Breusch–Godfrey test with lags from 1 to 3 under the null hypothesis of no autocorrelation in the realized moments.

For all other realized moments, the null hypothesis is rejected at a significance level of 0.05 for all tests. This result differs from Liu (2012), who found no evidence of autocorrelation in the realized third moment when analyzing the intraday data of IBM. In any event, modeling conditional skewness in this instance comes at no additional cost, in contrast to ARCH processes, where additional parameters always increase the difficulty of estimation. In summary, autocorrelation is a crucial feature of our data. We consider this property by including the EWMA approach (see Section 2.2) in our model to compute conditional forecasts for the realized moments. In the event the null hypothesis is accepted, we would need to use an unconditional rather than a conditional model.

3.2 Backtesting

We calculate the VaR for the next day according to (2.17), (2.19) and (2.21) for all models. If the next day's return violates this VaR, 1 (otherwise 0) is noted. This procedure is carried out for the full time series so as to generate a hitting sequence.

This paper employs a state-of-the-art technique for backtesting VaR violations. For a detailed description of the following three tests, the reader should refer to Christoffersen (2003). T_0 (T_1) is the number of 0s (1s) and $T = T_0 + T_1$. An unconditional coverage (uc) test is initially performed to determine whether the expected fraction of VaR violations θ differs from the realized fraction $\theta_{\text{real}} = T_1/T$. Under the null hypothesis $\theta = \theta_{\text{real}}$, the likelihood ratio test-statistic,

$$\text{LR}_{\text{uc}} = -2\ln\left[\frac{(1-\theta)^{T_0}\theta^{T_1}}{(T_0/T)^{T_0}(T_1/T)^{T_1}}\right], \tag{3.1}$$

is asymptotically χ^2 distributed with one degree of freedom and compared with a critical value of a given significance level. However, the "uc" test contains no information on a clustering of VaR violations. To check for such dependence in the violations, we conduct the independence (ind) test. Let T_{ij}, $i, j = 0, 1$, be the number of observations where a j follows an i. Under the null hypothesis of independence, the likelihood ratio test statistic is

$$\begin{aligned}\text{LR}_{\text{ind}} = -2\ln\Biggl[\Biggl(\Bigl(\frac{T_0}{T}\Bigr)^{T_0}\Bigl(\frac{T_1}{T}\Bigr)^{T_1}\Biggr) \\ \times\Biggl(\Biggl(\frac{T_{00}}{T_{00}+T_{01}}\Biggr)^{T_{00}}\Biggl(\frac{T_{01}}{T_{00}+T_{01}}\Biggr)^{T_{01}} \\ \times\Biggl(\frac{T_{10}}{T_{10}+T_{11}}\Biggr)^{T_{10}}\Biggl(\frac{T_{11}}{T_{10}+T_{11}}\Biggr)^{T_{11}}\Biggr)^{-1}\Biggr].\end{aligned} \tag{3.2}$$

Again, the statistic is asymptotically χ^2 distributed with one degree of freedom. To account for both demands (correct fraction and independence) simultaneously, the conditional coverage (cc) test combines both of the above tests:

$$\mathrm{LR}_{\mathrm{cc}} = \mathrm{LR}_{\mathrm{uc}} + \mathrm{LR}_{\mathrm{ind}}. \tag{3.3}$$

Consequently, its test statistic is χ^2-distributed with two degrees of freedom.

The fourth backtest is the dynamic quantile (dq) test of Engle and Manganelli (2004). Given the VaR confidence level $1 - \theta$, we redefine the hitting sequence (see above) by subtracting θ from each item of the sequence. The expected value of this modified hitting sequence, as well as its conditional expectation given any information known at time t, must be zero. Moreover, the modified hitting sequence must be uncorrelated with its lagged values and lagged VaRs. If these conditions are satisfied, the VaR valuations are uncorrelated and will have the correct fraction of violations. To verify this, we set up a regression framework for the modified hitting sequence and include the last four values of both explanatory variables. Under the null hypothesis that all coefficients of this regression are zero, we can then construct a likelihood ratio test that is χ^2 distributed with six degrees of freedom.

3.3 Results

Even a simple Gaussian assumption will suffice to calculate VaR for lower confidence levels (see, for example, Jorion 2007). This is because, even though financial returns are leptokurtic, cumulative distribution functions of empirical and Gaussian distributions tend to intersect around their 0.05 quantile. Therefore, this paper focuses on higher confidence levels. Table 2 on the facing page shows the results of backtests at confidence levels of $(1 - \theta) = 0.99$, 0.995 and 0.999. p-values are given for "uc", "ind", "cc" and "dq" tests with the null hypothesis that the VaR model cannot be rejected.

For the 0.99 confidence level, the T-FHS and the RM–EWMA–NIG prove to be the best models. θ_{real} is close to θ, resulting in high p-values for the "uc" test ($p_{\mathrm{uc}} = 0.6101$ and $p_{\mathrm{uc}} = 0.5417$). Both models pass the "ind", "cc" and "dq" tests with p-values not less than 0.5. The third best is the HEAVY model, which shows high p-values except for the "dq" test ($p_{\mathrm{dq}} = 0.2347$). Nevertheless, p_{dq} is still well above a significance level of 0.1. The HEAVY model has the highest p-value for the "uc" test ($p_{\mathrm{uc}} = 0.7334$) because it meets the expected fraction of violations very well ($\theta_{\mathrm{real}} = 0.011$). At a significance level of 0.1, we must reject all other models for at least one test. The RM–EWMA–NIG model, where we fix λ, does surprisingly well. Given a 0.05 significance level, we cannot reject the model. The RM–EWMA–CF model fails the "dq" test ($p_{\mathrm{dq}} = 0.0007$). When we fix λ, the model passes the "dq" test, but it still has to be rejected due to weak results in the "uc" and "cc" tests ($p_{\mathrm{uc}} = 0.0009$ and $p_{\mathrm{cc}} = 0.0042$). The T-GARCH is only average because it misses the expected fraction of violations ($\theta_{\mathrm{real}} = 0.0055$).

TABLE 2 Backtests on VaR calculations with confidence levels of 0.99, 0.995 and 0.999.

(a) 0.99

	θ_{real}	θ	p_{uc}	p_{ind}	p_{cc}	p_{dq}
T-GARCH	0.0055	0.010	*0.0759*	0.7812	0.1992	0.7932
T-FHS	0.0086	0.010	0.6101	0.6620	0.7981	0.9695
HEAVY	0.0110	0.010	0.7334	0.5774	0.8080	0.2347
RM–EWMA–CF$_{\lambda=0.94}$	0.0023	0.010	**0.0009**	0.9054	**0.0042**	0.1555
RM–EWMA–NIG$_{\lambda=0.94}$	0.0078	0.010	0.4180	0.6911	0.6658	*0.0572*
RM–EWMA–CF	0.0078	0.010	0.4180	0.6911	0.6658	**0.0007**
RM–EWMA–NIG	0.0117	0.010	0.5417	0.5504	0.6945	0.9448

(b) 0.995

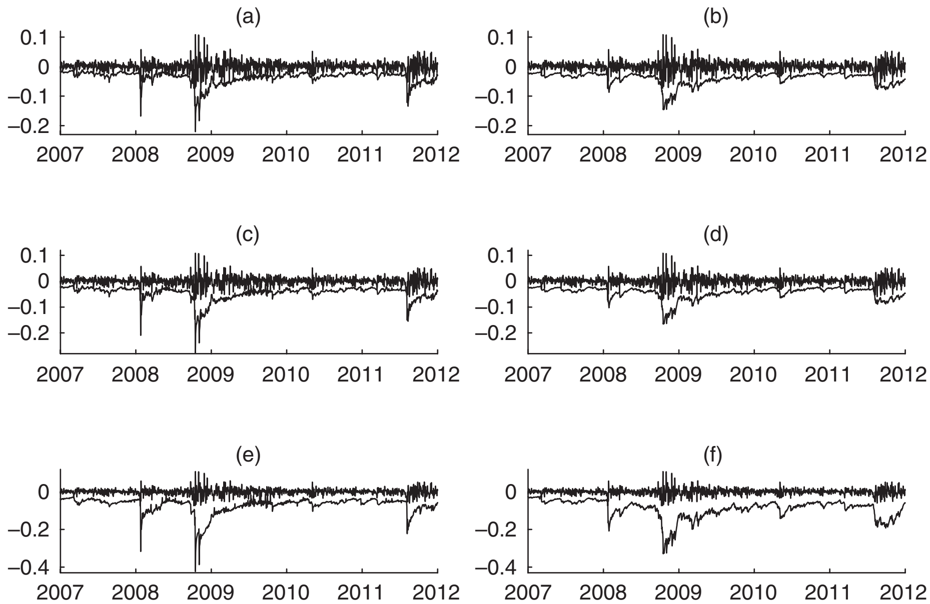
	θ_{real}	θ	p_{uc}	p_{ind}	p_{cc}	p_{dq}
T-GARCH	0.0023	0.005	0.1339	0.9054	0.3229	0.9233
T-FHS	0.0039	0.005	0.5678	0.8428	0.8329	0.9942
HEAVY	0.0070	0.005	0.3286	0.7208	0.5821	0.8787
RM–EWMA–CF$_{\lambda=0.94}$	0.0016	0.005	<u>0.0418</u>	0.9369	0.1257	0.5936
RM–EWMA–NIG$_{\lambda=0.94}$	0.0039	0.005	0.5678	0.8428	0.8329	0.9531
RM–EWMA–CF	0.0047	0.005	0.8774	0.8119	0.9606	**0.0000**
RM–EWMA–NIG	0.0039	0.005	0.5678	0.8428	0.8329	0.9977

(c) 0.999

	θ_{real}	θ	p_{uc}	p_{ind}	p_{cc}	p_{dq}
T-GARCH	0.0016	0.001	0.5547	0.9369	0.8373	0.8131
T-FHS	0.0016	0.001	0.5547	0.9369	0.8373	0.6617
HEAVY	0.0016	0.001	0.5547	0.9369	0.8373	0.9186
RM-EWMA-CF$_{\lambda=0.94}$	0.0008	0.001	0.7987	0.9684	0.9672	0.9989
RM-EWMA-NIG$_{\lambda=0.94}$	0.0016	0.001	0.5547	0.9369	0.8373	0.8051
RM-EWMA-CF	0.0063	0.001	**0.0001**	0.7508	**0.0003**	**0.0000**
RM-EWMA-NIG	0.0008	0.001	0.7987	0.9684	0.9672	0.9985

Backtesting results for the T-GARCH, T-FHS, HEAVY, RM–EWMA–CF and RM–EWMA–NIG models assuming confidence levels of 0.99, 0.995 and 0.999. The subscript $\lambda = 0.94$ indicates that the decay factor is fixed. $\theta_{\text{real}}(\theta)$ represents the realized (expected) fraction of VaR violations to the total number of observations. p-values are given for unconditional coverage (uc), independence (ind), conditional coverage (cc) and dynamic quantile (dq) tests with the null hypothesis that the VaR model cannot be rejected: $p \leqslant 0.01$ (boldface), $0.01 < p \leqslant 0.05$ (underlined) and $0.05 < p \leqslant 0.1$ (italic).

FIGURE 2 Comparison of VaR and returns.



Calculated VaR based on RM–EWMA–NIG for confidence levels of (a) 0.99, (c) 0.995 and (e) 0.999, together with the corresponding returns, and calculated VaR based on T-FHS for confidence levels of (b) 0.99, (d) 0.995 and (f) 0.999, together with the corresponding returns; average VaR is the average over all VaRs. (a) Average VaR = 0.03663. (b) Average VaR = 0.04008. (c) Average VaR = 0.04603. (d) Average VaR = 0.04509. (e) Average VaR = 0.07107. (f) Average VaR = 0.09038.

The results are not significantly different at the 0.995 confidence level. The T-FHS and RM–EWMA–NIG models are still the best, but the RM–EWMA–NIG model with fixed λ is not too far behind. All three models pass all tests with high p-values not less than 0.5. The HEAVY and T-GARCH models also perform better at this level, though the former is still superior. The performance of the RM–EWMA–CF models improves, especially the version with fixed λ, but it is nonetheless still weak. Not surprisingly, because fewer violations are expected and occur, across all models, the p-values of the "ind" test (and, in most instances, for the "dq" test) are high.

At the 0.999 confidence level, all analyzed models do well. Because θ is hit properly and violations occur even less frequently, p-values are higher than 0.5 across all models, with one exception. The RM–EWMA–CF model fails three of four tests at a 0.01 significance level because it clearly misses the expected fraction of violations.

In summary, three models cannot be rejected at the 0.1 significance level for all tests: the T-FHS, the RM–EWMA–NIG and the HEAVY models. The results of the

RM–EWMA–NIG model with fixed λ are remarkable, considering that we use no optimization algorithm for the variance.

Figure 2 on the facing page shows the returns and the corresponding VaRs for confidence levels of 0.99, 0.995 and 0.999 for RM–EWMA–NIG and T-FHS, which are the superior models. In general, both models show a similar pattern and adjust rapidly to changes in volatility. The VaR of RM–EWMA–NIG tends to respond more sensitively to massive price fluctuations and appears slightly more "jumpy". This is due to the non-mean-reverting nature of the EWMA, in contrast to the GARCH models. From the perspective of the requirement that a VaR should be as high as necessary but as low as possible, the VaR of RM–EWMA–NIG recovers much faster from shocks, while the VaR of T-FHS tends to remain at a high level. Particularly at the 0.999 confidence levels, both models show a conspicuously different pattern in the last half of 2011, during which there was a long period of high price fluctuation. The VaR of T-FHS during this period looks more jumpy and stays at a high level, whereas the VaR of RM–EWMA–NIG has one high amplitude and decreases more or less constantly to a normal level, which appears to be the more appropriate behavior. In general, the VaR of T-FHS seems to show higher values than the VaR of RM–EWMA–NIG. To quantify this visual impression, we calculate the average over all VaRs. For the important confidence levels of 0.99 and 0.999, the average VaR is much lower for the RM–EWMA–NIG (0.03663 and 0.07107) compared with the T-FHS (0.04008 and 0.09038); that is, utilizing the T-FHS would require too much equity to back up risk compared with the RM–EWMA–NIG.

4 CONCLUSION

Gaussian distributions do not adequately capture the behavior of financial returns, especially when it comes to the tails of a distribution, which is the important region in risk management and for VaR. The fit is poor because financial returns exhibit kurtosis and occur in volatility clusters. The literature contains many methods designed to take these stylized facts into account, but the models are often complex and sometimes barely feasible. This paper presents an alternative approach for VaR calculation based on realized moments. We computed forecasts for the realized moments with an EWMA and used them to parameterize the NIG distribution in a method of moments.

Using this technique, we calculated the VaR for the DAX at confidence levels of 0.99, 0.995 and 0.999. Although our alternative VaR methodology is comparatively simple, our results, in terms of backtesting, stack up well against the long-established T-FHS. Even if we fix the decay factor of the EWMA, and thus simplify the presented method by considerably reducing the amount of data required, the results are still convincing. The VaR is closely linked to the amount of (equity) capital a financial company must provide. In light of the maxim that a VaR should be as high as necessary but as low as possible, the RM–EWMA–NIG model is the best choice for its calculation, because its average VaR is lower than the one from the T-FHS.

Although the presented approach already shows very interesting results, there are a few extensions that might further improve it. Replacing the simple EWMA forecast for the moments with a more sophisticated realized GARCH (Hansen *et al* 2012), HEAVY or Corsi-type (Corsi 2009) forecast will allow mean reversion for the moments and enable the derivation of multistep-ahead forecasts. But doing so will, of course, require more complex estimation techniques. Although the NIG distribution is widely used for financial data, we could use other distributions that meet the criteria outlined in Section 2.3. For example, one possible second candidate is the tempered stable distribution. A convenient feature of the method presented in this paper is that it can encompass a portfolio point of view with realized covariance, coskewness and cokurtosis.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

An earlier version of this paper was titled “Keep value-at-risk calculation simple: an alternative approach based on realized variance”. The author thanks the participants of the 7th International Conference on Computational and Financial Econometrics in London and the 20th International Conference on Forecasting Financial Markets 2013 in Hanover, as well as Claudia Becker, Steffen Günther, Jörg Laitenberger and three anonymous referees for their helpful comments and discussion.

REFERENCES

Aas, K., Hobæk Haff, I., and Dimakos, X. K. (2006). Risk estimation using the multivariate normal inverse Gaussian distribution. *The Journal of Risk* **8**(2), 39–60.

Amaya, D., Christoffersen, P., Jacobs, K., and Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, forthcoming.

Andersen, T. G., and Bollerslev, T. (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review* **39**, 885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics* **61**, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica* **71**, 579–626.

Artzner, P. (1999). Coherent measures of risk. *Mathematical Finance* **9**, 203–228.

Barndorff-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics* **24**(1), 1–13.

Barndorff-Nielsen, O. E., and Shephard, N. G. (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society* B **64**(2), 253–280.

Barone-Adesi, G., Giannopoulos, K., and Vosper, L. (1999). VaR without correlations for nonlinear portfolios. *Journal of Futures Markets* **19**, 583–602.

Beltratti, A., and Morana, C. (2005). Statistical benefits of value-at-risk with long memory. *The Journal of Risk* **7**(4), 21–45.

Berkowitz, J., and O'Brien, J. (2002). How accurate are value-at-risk models at commercial banks? *Journal of Finance* **57**, 1093–1111.

Bond, S. A., and Patel, K. (2003). The conditional distribution of real estate returns: are higher moments time varying? *Journal of Real Estate Finance and Economics* **26**, 319–339.

Brooks, C., Burke, S. P., Heravi, S., and Persand, G. (2005). Autoregressive conditional kurtosis. *Journal of Financial Econometrics* **3**(3), 399–421.

Chen, Y. and Lu, J. (2012). Value at risk estimation. In *Handbook of Computational Finance*, Duan, J.-C., Härdle, W., and Gentle, J. E. (eds), pp. 307–333. Springer Handbooks of Computational Statistics. Springer.

Chen, Y., Härdle, W., and Jeong, S.-O. (2005). Nonparametric risk management with generalized hyperbolic distributions. *Journal of the American Statistical Association* **103**, 910–923.

Christoffersen, P. F. (2003). *Elements of Financial Risk Management*. Academic Press.

Clements, M. P., Galvão, A. B., and Kim, J. H. (2008). Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. *Journal of Empirical Finance* **15**(4), 729–750.

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

Cornish, E. A. and Fisher, R. A. (1937). Moments and cumulants in the specification of distributions. *Extrait de la Revue de l'Institute International* **4**, 1–14.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* **7**(2), 174–196.

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

Engle, R. F. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics* **17**(5), 425–446.

Engle, R. F., and Bollerslev, T. P. (1986). Modelling the persistence of conditional variances. *Econometric Reviews* **5**(1), 1–50.

Engle, R. F., and Manganelli, S. (2004). CAViaR: conditional autoregressive value-at-risk by regression quantiles. *Journal of Business and Economic Statistics* **22**, 367–381.

Eriksson, A., Ghysels, E., and Wang, F. (2009). The normal inverse Gaussian distribution and the pricing of derivatives. *Journal of Derivatives* **16**(3), 23–37.

Forsberg, L., and Bollerslev, T. (2002). Bridging the gap between the distribution of realized (ECU) volatility and ARCH modelling (of the euro): the GARCH–NIG model. *Journal of Applied Econometrics* **17**, 535–548.

Giot, P., and Laurent, S. (2004). Modelling daily value-at-risk using realized volatility and ARCH type models. *Journal of Empirical Finance* **11**(3), 379–398.

Hansen, P. R., Huang, Z., and Shek, H. H. (2012). Realized GARCH: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics* **27**(6), 877–906.

Hull, J., and White, A. (1998). Incorporating volatility updating into the historical simulation method for value-at-risk. *The Journal of Risk* **1**(1), 5–19.

Jorion, P. (2007). *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd edn. McGraw-Hill, New York.

JP Morgan and Reuters (1996). RiskMetrics. Technical Document, 4th edn. JP Morgan, New York.

Kalemanova, A., Schmid, B., and Werner, R. (2007). The normal inverse Gaussian distribution for synthetic CDO pricing. *Journal of Derivatives* **14**(3), 80–94.

Kuester, K., Mittnik, S., and Paolella, M. S. (2006). Value-at-risk prediction: a comparison of alternative strategies. *Journal of Financial Econometrics* **4**(1), 53–89.

Lillestøl, J. (2000). Risk analysis and the NIG distribution. *The Journal of Risk* **2**(4), 41–56.

Liu, P. (2012). Volatility, duration and value-at-risk. PhD Thesis, Paper 933, Electronic Thesis and Dissertation Repository, University of Western Ontario.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business* **36**, 394–419.

Paolella, M. S. (2007). *Intermediate Probability: A Computational Approach*. Wiley.

Pritsker, M. (2006). The hidden dangers of historical simulation. *Journal of Banking and Finance* **30**(2), 561–582.

Shephard, N., and Sheppard, K. (2010). Realising the future: forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics* **25**, 197–231.

Turtle, H. J., Buse, A., and Korkie, R. M. (1994). Tests of conditional asset pricing with time-varying moments and risk prices. *Journal of Financial and Quantitative Analysis* **29**, 15–29.

Venter, J. H., and de Jongh, P. J. (2002). Risk estimation using the normal inverse Gaussian distribution. *The Journal of Risk* **4**(2), 1–23.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* **100**, 1394–1411.