A benchmark for measuring bias in estimated daily value at risk

Imad A. Moosa\textsuperscript{a,}\textsuperscript{*}, Bernard Bollen\textsuperscript{b}

\textsuperscript{a}Department of Economics and Finance, La Trobe University, Room 424 (level 4) Donald Whitehead Building, Bundoora Campus, Victoria 3086, Australia

\textsuperscript{b}Department of Accounting and Finance, Monash University, Clayton, Australia

Abstract

In this study the notion of “realised volatility,” which has been proposed in the literature recently, is employed to obtain a benchmark to test statistically for bias in estimates of daily value at risk (VAR). A number of estimates based on the parametric and historical approaches to VAR are examined. It is found that the bias in the estimates of VAR are sensitive to the sample period and to the methodology chosen to estimate daily volatility. It appears that the estimates based on the exponentially weighted moving average volatility with a high decay factor are unbiased and relatively efficient. © 2002 Elsevier Science Inc. All rights reserved.

\textit{JEL classification:} G10

\textit{Keywords:} Value at risk; Realised volatility; GARCH; Normal distribution

1. Introduction

Value at risk (VAR) models have become a very popular tool for measuring the market (or price) risk of a portfolio of financial assets. Simply defined, VAR is an estimate of the maximum potential loss to be expected over a given period a certain percentage of the time. Put differently, it is the decline in the market value of an asset or a portfolio that can

\* Corresponding author. Fax: +61-3-9479-1654.
\textit{E-mail address:} i.moosa@latrobe.edu.au (I.A. Moosa).
be expected within a given time interval with a probability not exceeding a given number. This measure of risk has been gaining acceptance by practitioners, regulators, and exchanges, becoming a de facto standard for measuring financial risk. Yet, it has been also recognised that the measurement of VAR is based on arguably unrealistic assumptions, and that the calculation of VAR is sensitive to a large number of factors. Hence, the danger is that overreliance on VAR can give a false sense of security or lead to complacency. Indicative of the interest in VAR by academics and practitioners alike is the mushrooming literature on the subject (see, for example, Beder, 1995; Dowd, 1998; Hendricks, 1996; Jorion, 1996).

A crucial factor for the accuracy of VAR models that are based on the parametric approach is the underlying measure of volatility. Any bias in the measure of volatility would lead to bias in the estimated VAR. Hence, the need arises for recognising and measuring this bias. The problem with testing for bias in VAR, resulting from bias in the underlying volatility measure, is finding a benchmark against which all models can be compared.

In this paper, we suggest a benchmark model based on the concept of “realised volatility” that has been proposed by Andersen, Bollerslev, Diebold, and Labys (1999). With this benchmark in mind, models based on the parametric and historical approaches to the estimation of VAR are tested for bias. For this purpose, a VAR estimate would be biased if the mean value of VAR derived from this model is significantly different from the mean value of VAR derived from the benchmark model. Because the concept of realised volatility used in this paper implies the calculation of daily volatility from intraday data, this exercise is limited to daily VAR models.

2. The basic VAR model

In this section we formally define the concept of VAR for a portfolio of securities. Initially, we assume that the behaviour over time of the daily price of a portfolio of securities in a financial market can be represented by

\[ \ln(P_t) = \mu_t + \ln(P_{t-1}) + \epsilon_t \]  

where\( P_t \) is the closing price on trading day \( t \), \( \mu_t \) is the mean daily return, and \( \text{VAR}(\epsilon_t) = \sigma^2 \). In this study we deal only with daily data, and we shall therefore assume that \( \mu_t = 0. \)\(^1\) We define the daily return, \( r_t \), on a portfolio on trading day \( t \) as

\[ r_t = \ln(P_t) - \ln(P_{t-1}). \]  

By combining Eqs. (1) and (2) we obtain

\[ P_t = P_{t-1}e^{r_t} \]  

---

\(^1\) See Figlewski (1997) for a more thorough discussion of this assumption.
where $r_t = \varepsilon_t$. Now, define the critical portfolio return, $r_t^C$, such that the observed return on day $t$ is less than or equal to the critical level with a given probability. Therefore, for a probability of .05, we have

$$\Pr(r_t \leq r_t^C) = .05. \quad (4)$$

The critical portfolio value, $P_t^C$, corresponding to a 5% probability that the observed return on day will be less than or equal to $r_t^C$, is obtained by combining Eqs. (3) and (4). Therefore

$$P_t^C = P_{t-1}e^{r_t^C}. \quad (5)$$

For a portfolio whose market price is $P_t$, the VAR represents the loss in the value of the portfolio with 5% probability. Hence

$$\text{VAR}_t = P_t - P_t^C. \quad (6)$$

By combining Eqs. (5) and (6) we obtain

$$\text{VAR}_t = P_t = P_{t-1}e^{r_t^C}. \quad (7)$$

Since $e^x \approx 1 + x$ when $x$ is small, Eq. (7) can be written as Eq. (8)

$$\text{VAR}_t \approx P_t - P_{t-1}(1 + r_t^C). \quad (8)$$

If it is further assumed that portfolio’s scaled returns are normally distributed (we justify this assumption later), then the critical return at 5% will be given by

$$r_t^C = -1.645\sigma_t. \quad (9)$$

It follows that, under the assumption of normality, VAR can be calculated as

$$\text{VAR}_t = P_t - P_{t-1}(1 - 1.645\sigma_t) \quad (10)$$

where $\sigma_t$ is the volatility of the portfolio on trading day $t$. The expected value of $\text{VAR}_{t+1}$ at time $t$, that is $E_t[\text{VAR}_{t+1}]$, is given by

$$E_t[\text{VAR}_{t+1}] = E_t[P_{t+1}] - E_t[P_t(1 - 1.654\sigma_{t+1})]$$

$$= P_t - P_t + P_t \times 1.654 \times E_t[\sigma_{t+1}] = P_t \times 1.645 \times E_t[\sigma_{t+1}] \quad (11)$$

because $E_t[P_{t+1}] = P_t$ and $E_t[P_t] = P_t$. From Eq. (11), we define an estimator of $\text{VAR}_{t+1}$ as

$$\text{VAR}_{t+1} = P_t \times 1.645 \times \hat{\sigma}_{t+1}. \quad (12)$$

It is clear from Eq. (12) that the estimation of VAR is dependent upon the estimate of the volatility parameter, $\sigma_t$.

---

2 When financial prices are modelled as random walk, as represented by Eq. (1), it can be easily verified that $E_t[P_{t+1}] = P_t$ and $E_t[P_t] = P_t$. 
3. Data

Although this paper deals with daily VAR estimates, intraday data are used to estimate realised volatility, which is the basis of the benchmark model. Thus, this study is based on a sample of intraday data on the S&P 500 futures index for the period January 1993 to December 1995. The data were obtained from the Chicago Mercantile Exchange (CME).

Over the sample period 1,932,585 trades were recorded: these do not necessarily cover each transaction, as transactions with zero price changes are frequently not recorded. The transaction data are then aggregated up into a 5-minute return series for each trading day, where the \( i \)-th 5-minute daily return on trading day \( t \) is defined as \( r_{i,t} = \ln(P_{i,t}) - \ln(P_{i-1,t}) \) where \( P_{i,t} \) is the \( i \)th 5-minute price on day \( t \). The 5-minute data are then employed to estimate daily volatility, as described later in the paper. Furthermore, daily returns, defined as \( r_t = \ln(P_t) - \ln(P_{t-1}) \), are calculated and used to estimate other measures of daily volatility. Finally, daily price differences, defined as \( \Delta P_t = P_t - P_{t-1} \), are used to calculate VAR based on the historical approach.

4. Estimated VAR models

VAR can be estimated by following the parametric approach as represented by Eq. (12). Different results would arise from the use of different volatility measures. For this purpose, we will use five volatility measures: (i) the (sample) standard deviation based on equally weighted moving average; (ii) the standard deviation based on exponentially weighted moving average; (iii) the square root of the conditional variance derived from a GARCH(1,1) model; (iv) the square root of the conditional variance derived from an EGARCH(1,1) model; and (iv) the square root of the conditional variance derived from the GJR model. All of these volatility measures are calculated from historical data. The volatility measures and underlying models are briefly described below.

As an equally weighted moving average, the sample standard deviation of daily returns over the past \( k \) trading days is calculated as

\[
\hat{\sigma}_t = \sqrt{\frac{1}{(k-1)} \sum_{s=t-k}^{t-1} (r_s - \bar{r})^2}
\]  

(13)

where \( \bar{r} = \frac{1}{k} \sum_{s=t-k}^{t-1} r_s \). In this study we assign the following values to \( k \): 20, 60, 120, and 240, corresponding roughly to 1 month, 3 months, 6 months, and 1 year of daily returns, respectively. Because of the need for a maximum of 240 trading days of historical data to calculate VAR, the estimates begin on trading day 241 (14 December 1993) and end on

---

3 These models are due to Bollerslev (1986), Glosten et al. (1993) and Nelson (1991), respectively.
trading day 753 (29 December 1995), producing 513 daily VAR estimates in each case. Fig. 1 shows the VAR estimates based on equally weighted moving average volatility estimates (Eq. (13)). It is clear from Fig. 1 that the VAR estimates when \( k = 20 \) are prone to sharp swings, but when \( k = 240 \) the estimated VAR has a smoother shape. The VAR estimates when \( k = 60 \) and \( k = 120 \) appear to lie somewhere between these two extremes.

In estimating the standard deviation from Eq. (13), it is assumed that recent and more distant returns are equally weighted for the purpose of measuring current volatility. In contrast to this assumption, the volatility can be estimated as an exponentially weighted moving average, which gives more weight to recent daily returns rather than to more distant daily returns. This measure of volatility is given by Eq. (14)

\[
\hat{\sigma}_t = \sqrt{(1 - \lambda) \sum_{s=1-k}^{t-1} \lambda^{t-s-1} (r_s - \bar{r})^2}
\]  

(14)

where 0 < \( \lambda < 1 \).\(^4\) The lower the decay factor, \( \lambda \), the lower the influence of more distant returns. In this study we let \( \lambda \) assume the values 0.94, 0.97, and 0.99. As before, \( k \) assumes the values 20, 60, 120, and 240. Figs. 2–4 show the VAR estimates based on this measure of volatility with decay factors of 0.94, 0.97, and 0.99, respectively.

It is clear from Fig. 2 that the VAR estimates when the decay factor is 0.94 are virtually identical for \( k = 60, 120, \) and \( 240 \). However, the VAR estimates for \( k = 20 \) appear to be systematically lower relative to the VAR estimates for \( k = 60, 120, \) and \( 240 \). This can be

\(^4\) This model is effectively equivalent to what is obtained by applying the IGARCH model as shown by Hendricks (1996).
explained by the fact that the decay coefficient on day $t - 21$ is equal to $0.94^{21} = 0.27$, which means that a significant amount of information from daily returns will not be incorporated into contemporaneous VAR estimates. It appears that with a decay factor of 0.94 nearly all relevant information is incorporated into contemporaneous VAR estimates using 60 trading days of historical data. Figs. 3 and 4 show the VAR estimates for a decay factor of 0.97 and

![Fig. 2. VAR estimates based on exponentially weighted sample standard deviation (λ = 0.94).](image1)

![Fig. 3. VAR estimates based on exponentially weighted sample standard deviation (λ = 0.97).](image2)
0.99, respectively. It appears that the higher the decay factor, the greater the VAR estimates for a given amount of historical data \((k)\).

The third measure of volatility is derived from a GARCH(1,1) model. The primary motivation behind this model is the belief that volatility clusters in high and low periods. The estimated model is

\[
\begin{align*}
  r_t &= (0.0006016) + \varepsilon_t, \\
  \sigma_t^2 &= 0.000002164 + 0.8897 \sigma_{t-1}^2 + 0.04768 r_{t-1}^2 \\
  & \quad (2.852) \quad (0.3979) \quad (31.94) \quad (1.679)
\end{align*}
\]

Fig. 4. VAR estimates based on exponentially weighted sample standard deviation \((\lambda = 0.99)\).

Fig. 5. VAR estimates based on GARCH (1,1) volatility.
where $\sigma_t^2 = \text{VAR}(r_t)$ and the $t$ statistics are given in parentheses. Daily volatility is then estimated as a one-step ahead forecast from the model defined in Eq. (15). Fig. 5 shows the VAR estimates using the GARCH(1,1) measure of volatility.

The fourth measure of volatility is based on the EGARCH(1,1) model of Nelson (1991). The key motivation for the development of the EGARCH model was the insight that bad news on day $t - 1$ (as reflected in a negative return on day $t - 1$) has a greater impact upon volatility on day $t$ than good news. The estimated model is

$$
\ln(\sigma_t^2) = -2.98 + 0.722 \ln(\sigma_{t-1}^2) - 0.259 w_{t-1} + 0.052 [\mid w_{t-1} \mid - E(w_{t-1})] \\
(-5.88) (15.27) (-7.53) (0.96)
$$

(16)

where the $t$ statistics are given in parentheses and $w_t = r_t/\sigma_t$. It is the significance of the coefficient on $w_{t-1}$ that reflects the existence of asymmetric information effects on contemporaneous volatility.
The fifth measure of volatility is based on the GJR model of Glosten, Jagannathan, and Runkle (1993), which also attempts to capture the effect of asymmetric information on volatility. The estimated model is

\[
\sigma_t^2 = 0.0000845 + 0.681\sigma_{t-1}^2 - 0.0378 r_{t-1}^2 + 0.260 S_{t-1}^- r_{t-1}^2
\]

(17)

where \( t \) statistics are given in parentheses and \( S_{t-1}^- = 1 \) if \( r_{t-1} < 0 \) and \( S_{t-1}^- = 0 \) otherwise. Engle and Ng (1993) find that the GJR model is well specified with respect to capturing asymmetric information effects. VAR estimates based upon EGARCH(1,1) and GJR volatilities are shown in Figs. 6 and 7, respectively.

In addition to using the parametric approach and various volatility measures, we also use the historical approach to estimate VAR.\(^5\) This approach involves taking historical data over a sample period and ordering daily returns in a descending order. If the underlying probability is .05, then the daily return at the 95th percentile of the ordered sample of daily returns should be an estimate of VAR. Note that the historical approach makes no distributional assumptions about daily returns, and hence it is effectively a nonparametric

---

\(^5\) One possibility that we did not try is using the volatility implied by the prices of option on futures. There are two reasons for our decision not to use this measure of volatility. First, the implied volatility is a biased estimator because it is based on the highly restrictive assumptions used to derive the Black–Scholes formula (such as the normality of returns, constant volatility, and nonstochastic interest rate). Second, although we accept the proposition that implied volatility contains information about future volatility, we feel that it has to be corrected for bias as suggested by Gwilyn and Buckle (1999).
method. For sample sizes of 20, 60, 120, and 240 the 2nd, 4th, 7th, and 15th lowest ordered daily return observation should be the VAR estimate, respectively. Fig. 8 shows VAR estimates using the historical approach, allowing for the sample size, \( k \), to assume the values 20, 60, 120, and 240. It is clear from the graph that the VAR estimates based on the historical approach with large samples has smoother time series properties than estimates based on small samples.

5. The benchmark VAR model

The benchmark model is based on the concept of realised volatility, recently proposed by Andersen et al. (1999). This measure of volatility is calculated as the sum of squares of 5-minute returns. Hence,

\[
\hat{\sigma}^2_t = \sum_{i=1}^{n_t} r_{i,t}^2
\]

where \( r_{i,t} \) is the \( i \)th intraday return on day \( t \) and \( n_t \) is the number of intraday trades on day \( t \) (Eq. (18)). Andersen et al. show that this is an efficient estimator of daily volatility and that it is justified as an estimator of daily volatility because

\[
\text{plim}_{n_t \to \infty} \sum_{i=1}^{n_t} r_{i,t}^2 = \sigma^2_t.
\]

Furthermore, this unconditional, ex post estimator of daily volatility does not require the intraday return series to be homoscedastic, although it does require intraday returns to be uncorrelated (see Eq. (19)). Fig. 9 shows the VAR estimates based on the concept of realised volatility.

Why would a VAR model that is based on realised volatility be considered a benchmark model? This choice can be justified on the basis of the argument put forward
by Andersen et al. (1999) that realised volatility is the actual and observed volatility measure. With respect to summing sufficiently many high frequency discrete time intraday returns, they argue that “the volatility estimation error is in principle under our control and can be made arbitrarily small by summing sufficiently finely sampled intraday returns.” They further argue that “the resulting small amount of measurement error in our daily volatility estimates means that for practical purposes we can treat the daily volatility as observed.” If this view is accepted, then we have a natural benchmark to test VAR models since a VAR estimate based on realised volatility can, in effect, be treated as an actual observation.

One further issue that needs to be addressed is the use of the normal critical values in association with Eqs. (9)–(12). The literature overwhelmingly supports the notion that asset returns have a leptokurtic or fat-tailed distribution.⁶ On the surface, therefore, the use of the normal critical values would appear to be extremely inappropriate. However, the use of these critical values for the purpose of this paper is not justified by assuming that asset returns are normally distributed, but rather by assuming that scaled returns are normally distributed. Specifically, if \( \sigma_t \) is an adequate estimator of daily volatility, we would expect \( \hat{w}_t = \frac{r_t}{\sigma_t} \) to be normally distributed.

The hypothesis that scaled returns are normally distributed, provided that they are conditional on an adequate measure of volatility is accepted in the finance literature. Supporting arguments and evidence can be found in the RiskMetrics document of J.P. Morgan (1996, p 73), Andersen, Bollerslev, Diebold, and Labys (2000) and in Bollen and Inder (1998).⁷ In this case we argue that scaled returns conditional on realised volatility are normally distributed, whereas scaled returns conditional on volatility estimates from the ARCH family of models are nonnormal. This is because the ARCH-based volatility estimates, unlike realised volatility, fail to incorporate contemporaneous information as it is clear from Eqs. (15)–(17). To confirm the validity of our assumption we applied the Jarque and Bera (1980) test to the time series on scaled returns, \( \hat{w}_t \). The value of the test statistic turned out to be 4.0316, implying that the hypothesis of the normality of scaled returns cannot be rejected at the 5% significance level.⁸

6. Testing for bias in estimated VAR

The test can be conducted as follows. If the mean of the VAR estimated from a particular model is not significantly different from the mean of the VAR estimated from the benchmark

---

⁶ On this issue, see Campbell, Lo, and MacKinlay (1997, Chapter 2).
⁷ The title of the paper by Andersen et al. (2000) sums it all up: “Exchange Rate Returns Standardised by Realized Volatility are (Nearly) Gaussian.”
⁸ If the condition of normality is not satisfied, then other testing procedures must be employed. For example, Jorion (1996) uses the \( t \) distribution, whereas Glasserman, Heidelberger, and Shahabuddin (2000) employ a multivariate \( t \) distribution.
model, then the first model produces results that are unbiased. Let $\mu_B$ and $\mu_M$ be the (population) mean values of VAR calculated from the benchmark model and another model, respectively. Hence, the null and alternative hypotheses are given by Eqs. (20) and (21)

$$H_0 : \mu_B - \mu_R = 0$$

and

$$H_1 : \mu_B - \mu_R \neq 0.$$  

A natural test statistic is thus defined for the difference of two population means. This is given by Eq. (22)

$$z = \frac{\bar{X}_B - \bar{X}_R}{\sqrt{\hat{V}(\text{VAR}_B - \text{VAR}_R)}} = \frac{\bar{X}_B - \bar{X}_R}{\sqrt{\hat{V}(\text{VAR}_B) + \hat{V}(\text{VAR}_R) - 2\hat{C}(	ext{VAR}_B, \text{VAR}_R)}}$$

where $\bar{X}$, $\hat{V}(\cdot)$, and $\hat{C}(\cdot)$ are the sample mean, variance, and covariance, respectively.

### Table 1
Results of testing for bias in VAR estimates

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean</th>
<th>S.D.</th>
<th>$z$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Equally weighted S.D.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k = 20$</td>
<td>4.71</td>
<td>1.24</td>
<td>-0.13</td>
</tr>
<tr>
<td>$k = 60$</td>
<td>4.77</td>
<td>0.76</td>
<td>-0.95</td>
</tr>
<tr>
<td>$k = 120$</td>
<td>4.80</td>
<td>0.56</td>
<td>-1.26</td>
</tr>
<tr>
<td>$k = 240$</td>
<td>4.86</td>
<td>0.49</td>
<td>-2.05</td>
</tr>
<tr>
<td><strong>Exponentially weighted S.D.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k = 20, \lambda = 0.94$</td>
<td>3.86</td>
<td>1.02</td>
<td>11.81</td>
</tr>
<tr>
<td>$k = 60, \lambda = 0.94$</td>
<td>4.58</td>
<td>0.93</td>
<td>1.76</td>
</tr>
<tr>
<td>$k = 120, \lambda = 0.94$</td>
<td>4.64</td>
<td>0.93</td>
<td>0.91</td>
</tr>
<tr>
<td>$k = 240, \lambda = 0.94$</td>
<td>4.63</td>
<td>0.91</td>
<td>0.89</td>
</tr>
<tr>
<td>$k = 20, \lambda = 0.97$</td>
<td>3.15</td>
<td>0.81</td>
<td>22.04</td>
</tr>
<tr>
<td>$k = 60, \lambda = 0.97$</td>
<td>4.33</td>
<td>0.73</td>
<td>5.27</td>
</tr>
<tr>
<td>$k = 120, \lambda = 0.97$</td>
<td>4.68</td>
<td>0.67</td>
<td>0.30</td>
</tr>
<tr>
<td>$k = 240, \lambda = 0.97$</td>
<td>4.74</td>
<td>0.66</td>
<td>0.58</td>
</tr>
<tr>
<td>$k = 20, \lambda = 0.99$</td>
<td>2.01</td>
<td>0.51</td>
<td>37.95</td>
</tr>
<tr>
<td>$k = 60, \lambda = 0.99$</td>
<td>3.21</td>
<td>0.49</td>
<td>20.54</td>
</tr>
<tr>
<td>$k = 120, \lambda = 0.99$</td>
<td>4.02</td>
<td>0.46</td>
<td>9.198</td>
</tr>
<tr>
<td>$k = 240, \lambda = 0.99$</td>
<td>4.60</td>
<td>0.45</td>
<td>1.35</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>4.94</td>
<td>0.62</td>
<td>4.34</td>
</tr>
<tr>
<td>EGARCH(1,1)</td>
<td>4.42</td>
<td>0.95</td>
<td>-3.96</td>
</tr>
<tr>
<td>GJR</td>
<td>4.83</td>
<td>1.04</td>
<td>1.92</td>
</tr>
<tr>
<td><strong>Historical</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k = 20$</td>
<td>3.68</td>
<td>1.71</td>
<td>11.21</td>
</tr>
<tr>
<td>$k = 60$</td>
<td>4.11</td>
<td>1.37</td>
<td>6.60</td>
</tr>
<tr>
<td>$k = 120$</td>
<td>4.19</td>
<td>0.88</td>
<td>5.91</td>
</tr>
<tr>
<td>$k = 240$</td>
<td>4.06</td>
<td>0.49</td>
<td>8.04</td>
</tr>
</tbody>
</table>
We can appeal to the central limit theorem to use the standard normal critical values for the test.9 At 95% confidence, the appropriate critical values are ±1.96. The results reported in Table 1 clearly show that the VAR estimates based on the equally weighted moving average volatility are unbiased relative to the benchmark model except when \( k = 240 \), that is when a sample period of about 1 year is taken. VAR estimates based on exponentially weighted moving average volatility seem to be systematically biased when short sample periods are taken together with a large decay factor. With a decay factor of \( \lambda = 0.99 \), it takes 240 trading days of daily data to obtain an unbiased estimate of VAR. The VAR estimates based on GARCH(1,1) and EGARCH(1,1) volatility as well as those derived from the historical approach appear to be systematically biased. VAR estimates based upon GJR volatility

---

9 Given the large sample size utilised for this test (513), this asymptotic result would appear to be a reasonable assumption.
estimates perform well relative to the benchmark model, a result consistent with the finding of Engle and Ng (1993) that the GJR volatility model is well specified. Given a standard deviation of 0.45, the VAR estimates based on the exponentially weighted moving average volatility (with a high decay factor, $\lambda = 0.99$, and long sample period, $k = 240$) appear to be the most efficient of the unbiased VAR estimates.

To take the matter further, we employ the testing procedure proposed by Hendricks (1996). He suggests that the correlation between the various VAR measures and the absolute value of price changes is indicative of how closely changes in the VAR estimates correspond to actual changes in the risk of the portfolio. In Table 2 we report the estimated correlation coefficient ($\hat{\rho}$) and its $z$ statistic, which is based on the approximation 
\[ \sqrt{n\hat{\rho}} \sim N(0,1), \]
where $n$ is the number of observations in the sample. It is immediately obvious from an inspection of this table that all, except two, estimated correlation coefficients are insignificant at the 5% level. Although VAR based upon the equally weighted moving average volatility estimate is significantly correlated with absolute price changes, it is a biased measure of VAR and hence the reported negative correlation has little meaning. Of considerably more interest, however, is the finding that the correlation coefficient between the VAR estimates based upon realised volatility and absolute price changes is significant. Furthermore, the numerical value of this coefficient is higher than any of the other correlations, offering further evidence in support of the VAR estimates based upon realised volatility as being an appropriate benchmark.

7. Concluding remarks

Utilising the concept of realised daily volatility, as proposed by Andersen et al. (1999), it is possible to obtain a benchmark to test for bias in VAR models. It has been found that bias in the estimates of VAR is sensitive to the methodology employed to estimate daily volatility. Specifically, it was found that:

1. Estimates of VAR based on equally weighted moving average volatility when shorter time periods ($k = 20, 60, 120$) are used are unbiased. When longer time periods are used ($k = 240$), the estimates become biased.
2. VAR estimates based on exponentially weighted moving average volatility are unbiased when longer time periods are used ($k = 240$). Shorter time periods can be used when estimating VAR with exponentially weighted moving average volatility as long as the decay factor is relatively small.
3. When volatility is based on a GARCH(1,1) model, the VAR estimates turn out to be biased.
4. Using GJR volatility estimates produces reasonable VAR estimates.
5. The historical approach produces bias in VAR estimates.
6. The most efficient of the VAR estimates is produced when exponentially weighted moving average volatility estimates are used with a large decay factor.
A cautionary note is warranted here. One must not consider these results to be universal, as they may be specific to the particular time series used in this study. The results, however, demonstrate a weakness in the VAR methodology: the sensitivity of the VAR estimates to the method of calculation and the assumption concerning some underlying parameters. This is indeed one of the drawbacks of the VAR methodology that have led to thinking about alternative measures of risk.

One of the most interesting attempts to remedy VAR has been made by Artzner, Delbaen, Eber, and Heath (1999) who proposed conditional VAR to circumvent the problems associated with (conventional) VAR. These authors recognise the following pitfalls and shortcomings:

- VAR fails to satisfy the subadditivity criterion. In other words, VAR does not behave nicely with respect to the addition of risks, thereby creating severe aggregation problems. This is why it is not a coherent measure of risk.
- VAR fails to recognise the concentration of risk, and so it fails to encourage a reasonable allocation of risk among agents. It does not encourage diversification because it fails to take into account the economic consequences of the events, the probabilities of which it controls.
- Work on VAR, including Dowd (1998) and RiskMetrics (1996) is concerned mainly with the computational and statistical problems, without first considering the implications of this method for measuring risk. This is exactly what this paper is concerned with.

As an alternative, Artzner et al. (1999) propose conditional VAR (also called conditional excess), which is a measure of risk based on the “worst conditional expectation.” In this sense this measure of risk tackles the question of “how bad is bad,” which is not addressed by VAR. This measure of risk depends on \( \min\{0, -\text{VAR}(X) - X\} \), where \( X \) is a random variable representing the final net worth of a position for each element of the set of states of nature. It is also supported by Glasserman et al. (2000) who criticise VAR as being insensitive to the magnitude of losses beyond a certain percentile. The argument put forward by Glasserman et al. is that, unlike VAR, the conditional VAR weights losses by their magnitudes. Similar arguments can be found in Uryasev and Rockafellar (1999).

The fact remains, however, that conventional VAR is a popular and heavily used measure of risk. Despite its drawbacks, this measure is appealing because of its simplicity, which makes it easily comprehensible by the nonspecialist. This paper presented results confirming that VAR estimates can be biased and suggested a method to eliminate the bias. If VAR must be used, then calculating this measure of risk on the basis of the concept of realised volatility will be beneficial in the sense that it produces a more accurate measure of the VAR.

**Acknowledgments**

We are grateful to an anonymous referee for some useful comments. An earlier version of this paper was presented at the Quantitative Methods in Finance Conference, which was held
in Sydney in December 2000. We are grateful to the participants in this conference for their comments and suggestions.

References


