Bias in the Estimation of Non-Linear Transformations of the Conditional Variance of Returns

Cherif Guermat and Richard D.F. Harris

School of Business and Economics
University of Exeter

January 2003

Abstract

Many applications in finance use a non-linear transformation of the variance of returns. While the sample variance is an unbiased and consistent estimator of the population variance of returns, non-linear transformations of the sample variance will be consistent but biased. For estimates of non-linear transformations of the unconditional variance, this will rarely be a problem in practice, since sample sizes employed in finance are typically large. However, estimators of the conditional variance typically use sample sizes that are effectively much smaller, particularly those that apply an exponential weighting to returns such as GARCH or EMWA. Consequently, the bias is likely to be more important in estimating non-linear transformations of the conditional variance. In this paper, we derive a simple analytical approximation for the unconditional bias of estimators of non-linear transformations of the conditional variance, under the assumption that the true conditional variance is generated by an arbitrary stochastic volatility model. As an illustration, we estimate the bias that is implied by the RiskMetrics approach to the calculation of value at risk.

Keywords: Conditional variance; Conditional standard deviation; Non-linear transformation; Small sample bias; Exponentially weighted moving average; Generalised autoregressive conditional heteroscedasticity; Value at risk.

1. Introduction

Many applications in finance require a measure of the volatility of returns. In many cases this measure is a non-linear function of the variance of returns. For instance, estimation of value at risk uses the square root of the variance (see, for instance, Jorion, 2000), calculation of the Sharpe ratio uses the reciprocal of the square root of the variance (see, for instance, Sharpe, 1994), and the Black-Scholes option pricing model uses the cumulative standard normal function of the variance (see, for instance, Black and Scholes, 1973). The usual approach in practice is to estimate the variance of returns and to transform this estimate using the same non-linear transformation. However, while the sample variance is a consistent and unbiased estimator of the true variance, non-linear transformations of the sample variance will, in general, be biased estimators of the same non-linear transformations of the true variance because of Jensen's inequality. [1]

When estimating non-linear transformations of the unconditional variance, this is not likely to be a problem in practice, since samples sizes used in finance are typically large, and consequently one can rely on the fact that while biased, non-linear transformations of the sample variance are consistent. However, estimators of the conditional variance typically use sample sizes that are effectively much smaller, either explicitly, as in the case of, say, a rolling sample variance estimator, or implicitly, through an exponential weighting of the observations, such as in a generalised autoregressive conditional heteroscedasticity (GARCH) or exponentially weighted moving average (EWMA) model. Consequently, the effect of Jensen's inequality is likely to be more important in estimating non-linear transformations of the conditional variance.

In this paper, we derive a simple analytical approximation for the unconditional bias of estimators of non-linear transformations of the conditional variance, under the assumption that the true conditional variance is generated by an arbitrary stochastic volatility model. The bias is a function of the unconditional variance and kurtosis of

---

[1] Here we mean the sample variance when the mean is assumed to be known, or the sample variance adjusted for the loss in degrees of freedom when the mean is estimated.

returns, the conditional kurtosis of returns, and the variance of the estimated conditional variance, and depends on the non-linear transformation used. The analytical expression can be used to provide upper and lower bounds for the bias, and, when the conditional distribution is assumed to be known, an approximation to the bias under the assumed conditional distribution.

As an illustration, we consider the RiskMetrics approach to the calculation of value at risk (see JP Morgan, 1994). The RiskMetrics approach assumes that returns are conditionally normally distributed and employs a EWMA model to estimate the conditional variance of returns. The value at risk of a portfolio is estimated by multiplying the appropriate quantile of the standard normal distribution by an estimate of the conditional standard deviation from the EWMA model. Using daily returns on several equity, bond and exchange rate portfolios over a fifteen year period, we estimate the bounds for the bias of the EWMA conditional standard deviation. We show that under the maintained assumption of conditional normality, there is a significant upward bias in the estimated conditional standard deviation for almost all the series considered, implying a corresponding upward bias in the estimated value at risk. In some cases the bias is as much as twenty percent.

The outline of this paper is as follows. In Section 2 we derive the unconditional bias of estimators of non-linear transformations of the conditional variance. Section 3 reports the results of the empirical illustration. Section 4 concludes.

2. Theoretical Framework

Suppose that the return on an asset between time t-1 and time t, denoted r_t, is drawn conditionally from a distribution with mean zero and conditional variance σ_t^2.[2] We assume that the data are generated by a stochastic volatility model, and that a conditional variance model is used to generate a forecast of the conditional variance at

---

[2] It is assumed that the mean return is equal to zero, which is a common assumption when modelling the variance of short horizon asset returns, and is justified by the fact that the mean return of an asset is typically several orders of magnitude smaller than its standard deviation (see, for example, Figlewski, 1997, or Jorion, 2000). However, it would be straightforward to accommodate a non-zero mean return.

time $t+1$, $\hat{\sigma}_{t+1}^2$. We assume that while $\hat{\sigma}_{t+1}^2$ may be conditionally biased (because the model may be misspecified with respect to the data generating process), the unconditional bias of $\hat{\sigma}_{t+1}^2$ is zero. Consider, for instance, the following conditional variance models.

Rolling sample variance estimator

$$\hat{\sigma}_{t+1}^2 = \frac{1}{T} \sum_{s=0}^{T-1} r_{t-s}^2 \tag{1}$$

GARCH (1,1) *estimator*

$$\begin{aligned}\hat{\sigma}_{t+1}^2 &= \alpha_0 + \alpha_1 \hat{\sigma}_t^2 + \alpha_2 r_t^2 \\ &= \frac{\alpha_0}{1-\alpha_1} + \alpha_2 \sum_{i=0}^{\infty} \alpha_1^i r_{t-i}^2\end{aligned} \tag{2}$$

with $\alpha_0 = (1 - \alpha_1 - \alpha_2)\sigma^2$

EWMA or IGARCH (1,1) *estimator*

$$\begin{aligned}\hat{\sigma}_{t+1}^2 &= \lambda \hat{\sigma}_t^2 + (1-\lambda) r_t^2 \\ &= (1-\lambda) \sum_{i=0}^{\infty} \lambda^i r_{t-i}^2\end{aligned} \tag{3}$$

(For a fuller description of these models, see, for instance, Alexander and Leigh, 1997). The conditional bias of $\hat{\sigma}_{t+1}^2$ from each of these models is given by

$$E_t(\hat{\sigma}_{t+1}^2 - \sigma_{t+1}^2) = A_{t+1} \tag{4}$$

where $E_t(.)$ is the expectation conditional on the time-t information set. In general, the conditional bias, A_{t+1}, will be non-zero. However, the unconditional bias of $\hat{\sigma}_{t+1}^2$, $A = E(A_{t+1})$, is given by

$$A = E(\hat{\sigma}^2_{t+1}) - E(\sigma^2_{t+1})$$
$$= \sigma^2 - \sigma^2 \qquad\qquad (5)$$
$$= 0$$

where $E(\hat{\sigma}^2_{t+1}) = \sigma^2$ from the definition of the estimator given by (1), (2) or (3) and

$E(\sigma^2_{t+1}) = \sigma^2$ by the law of iterated expectations.[3]

In order to estimate a non-linear transformation of the true conditional variance, $g(\sigma^2_{t+1})$, we assume that the same non-linear transformation is applied to the forecast conditional variance, $g(\hat{\sigma}^2_{t+1})$. The conditional bias of $g(\hat{\sigma}^2_{t+1})$ is given by

$$E_t(g(\hat{\sigma}^2_{t+1}) - g(\sigma^2_{t+1})) = B_{t+1} \qquad\qquad (6)$$

Again, the conditional bias, B_{t+1}, will in general be non-zero. This may be true even if the conditional variance model is correctly specified because of Jensen's inequality.[4] The unconditional bias of $g(\hat{\sigma}^2_{t+1})$ can be approximated using second order Taylor series expansions for both $g(\hat{\sigma}^2_{t+1})$ and $g(\sigma^2_{t+1})$, around σ^2, which yield

$$g(\hat{\sigma}^2_{t+1}) \approx g(\sigma^2) + g'(\sigma^2)(\hat{\sigma}^2_{t+1} - \sigma^2) + \frac{1}{2}g''(\sigma^2)(\hat{\sigma}^2_{t+1} - \sigma^2)^2 \qquad (7)$$

and

$$g(\sigma^2_{t+1}) \approx g(\sigma^2) + g'(\sigma^2)(\sigma^2_{t+1} - \sigma^2) + \frac{1}{2}g''(\sigma^2)(\sigma^2_{t+1} - \sigma^2)^2 \qquad (8)$$

---

[3] The rolling sample variance estimator and the EWMA estimator satisfy this property by construction. However, the GARCH estimator will only satisfy this property when $\alpha_0 = (1 - \alpha_1 - \alpha_2)\sigma^2$. While in general the true unconditional variance of returns will not be known, this condition can be imposed using the sample variance in place of the true variance.

[4] The only case in which the conditional bias would be zero is when the data are generated by a deterministic process, and that this process is used to forecast the conditional variance.

where $g'(.)$ and $g''(.)$ are the first and second derivatives of $g(\sigma^2)$ with respect to σ^2. Taking the unconditional expectation of (7) and (8), we can therefore approximate the unconditional bias of $g(\hat{\sigma}^2_{t+1})$, $B = E(B_{t+1})$, as

$$B \approx \frac{1}{2} g''(\sigma^2) \left\{ \mathrm{var}(\hat{\sigma}^2_{t+1}) - \mathrm{var}(\sigma^2_{t+1}) \right\} \tag{9}$$

Therefore, the unconditional bias of $g(\hat{\sigma}^2_{t+1})$ is proportional to the difference between the unconditional variance of the estimated conditional variance, and the unconditional variance of the actual conditional variance.

The unconditional variance of σ^2_{t+1} can be written as

$$\begin{aligned}
\mathrm{var}(\sigma^2_{t+1}) &= E(\sigma^4_{t+1}) - E(\sigma^2_{t+1})^2 \\
&= \frac{E(E_t(r^4_{t+1}))}{k_C} - \sigma^4 \\
&= \frac{E(r^4_{t+1})}{k_C} - \sigma^4 \\
&= \frac{k_U \sigma^4}{k_C} - \sigma^4 \\
&= \sigma^4 \left\{ \frac{k_U - k_C}{k_C} \right\}
\end{aligned} \tag{10}$$

where $k_U = E(r^4_{t+1})/\sigma^4$ is the unconditional kurtosis coefficient of returns and $k_C = E_t(r^4_{t+1})/\sigma^4_{t+1}$ is the conditional kurtosis coefficient of returns. The unconditional bias of $g(\hat{\sigma}^2_{t+1})$ can therefore be written as

$$B \approx \frac{1}{2} g''(\sigma^2) \left\{ \mathrm{var}(\hat{\sigma}^2_{t+1}) - \frac{k_U - k_C}{k_C} \sigma^4 \right\} \tag{11}$$

The conditional kurtosis, k_C, is unobserved, but is bounded from below by unity and from above by the unconditional kurtosis, k_U. This therefore yields the following bounds for the unconditional bias of $g(\hat{\sigma}_{t+1}^2)$.

$$\frac{1}{2}g''(\sigma^2)\left\{\text{var}(\hat{\sigma}_{t+1}^2) - (k_U - 1)\sigma^4\right\} \le B \le \frac{1}{2}g''(\sigma^2)\,\text{var}(\hat{\sigma}_{t+1}^2) \qquad (12)$$

When the conditional distribution is assumed to be known, as it is in many applications in finance such as those that assume conditional normality, equation (11) can be used to evaluate the bias that is implied by the assumed conditional distribution.

3. Empirical Illustration

In this section, we illustrate the above results by considering the bias that is implied by the RiskMetrics approach to the calculation of value at risk (see JP Morgan, 1996). The RiskMetrics approach assumes that returns are conditionally normal, and uses a EWMA model with a decay factor of 0.94 to forecast the conditional variance. The value at risk of a portfolio is then estimated by multiplying the appropriate quantile of the standard normal distribution by the estimated conditional standard deviation from the EWMA model. As discussed above, the EWMA model generates forecasts of the conditional variance that are unconditionally unbiased. However, estimates of the conditional standard deviation from the EWMA model will in general be biased. Consequently, estimates of value at risk based on the estimated conditional standard deviation will also be biased leading, under the Basle Committee rules, to either too little or too much capital being held on average (see Basle Committee, 1996).

In order to illustrate the effects of this bias, we use daily returns for several equity, bond and exchange rate portfolios. Specifically, we consider the SP500, FTSE100 and Nikkei225 equity indices, the Datastream long bond return indices for the US, UK and Japan, and the USD/GBP, USD/JPY and GBP/JPY exchange rates. Data were obtained from Datastream for the period 30/12/87 to 4/12/02, which is the longest common sample available (a total of 3895 observations), and used to calculate

continuously compounded daily returns. The EWMA model (3) is used to estimate the conditional variance with $\lambda = 0.94$. The EWMA model is initialised using the first 500 observations, which are then discarded, yielding a sample of 3394 observations.

In order to estimate the unconditional bias of the EWMA conditional standard deviation, we require sample estimates of $g''(\sigma^2)$, $\text{var}(\hat{\sigma}_{t+1}^2)$ and k_U. For the standard deviation, $g(\sigma^2) = (\sigma^2)^{1/2}$ and $g''(\sigma^2) = -(1/4)\sigma^{-3}$. The estimation results are reported in Table 1. Columns 1-3 report the sample estimates of σ, k_U, and $\text{var}(\hat{\sigma}_{t+1}^2)$. Columns 4 and 5 report the bounds for the unconditional bias, evaluated using equation (12). Column 6 reports the unconditional bias under the maintained assumption that returns are drawn from a conditional normal distribution, evaluated using equation (10) with $k_C = 3$.

[Table 1]

The bounds for the bias in the estimated conditional standard deviation show that the bias is potentially large. For the Japanese long bond index, for instance, the range for the bias is about –20% (when $k_C = k_U$) to +133% (when $k_C = 1$). In all cases, the lower bound is much smaller in absolute value than the upper bound, implying that the potential for overestimation of the bias is much greater than the potential for underestimation. Under the maintained assumption of conditional normality, there is, in all but one case, a substantial upward bias in the estimated conditional standard deviation. For the Japanese long bond index, the bias is more than 20%. For the FTSE100 the bias is negative, but close to zero.

4. Conclusion

Applications in finance commonly use non-linear transformations of the variance of returns. While the sample variance is an unbiased and consistent estimator of the population variance of returns, non-linear transformations of the sample variance will be consistent but biased, owing to Jensen's inequality. For estimates of non-linear transformations of the unconditional variance, this will rarely be a problem in

practice, since sample sizes employed in finance are typically large. However, estimators of the conditional variance typically use sample sizes that are effectively much smaller, particularly those that apply an exponential weighting to returns such as GARCH or EMWA. Consequently, the effect of Jensen's inequality is likely to be more important in estimating non-linear transformations of the conditional variance.

In this paper, we derive a simple analytical approximation for the unconditional bias of estimators of non-linear transformations of the conditional variance, under the assumption that the true conditional variance is generated by an arbitrary stochastic volatility model. The bias is a function of the unconditional variance and kurtosis of returns, the conditional kurtosis of returns and the variance of the estimated conditional variance, and depends on the non-linear transformation used. Many applications in finance make the assumption that returns are drawn from a conditional normal distribution. The results of this paper can be used to estimate the bias that is implied under the maintained assumption of conditional normality.

As an illustration, we estimate the bias implied by the RiskMetrics approach to the calculation of value at risk, using daily returns on several equity, bond and exchange rate portfolios over a fifteen year period. We find that in most cases, there is a significant upward bias in estimates of the conditional standard deviation, leading to a corresponding upward bias in the estimated value at risk.

References

Alexander, C. and C. Leigh, 1997, "On the Covariance Matrices Used in Value at Risk Models", *Journal of Derivatives* 4, 50-62.

Black, F., and M. Scholes, 1973, "The Pricing of Options and Corporate liabilities", *Journal of Political Economy* 81, 637-654.

Basle Committee on Banking Supervision, 1996, Overview of the Amendment to the Capital Accord to Incorporate Market Risks, January.

Figlewski, S., 1997, "Forecasting Volatility", *Financial Markets, Institutions and Instruments* Vol. 6, No. 1.

Jorion, P., 2000, *Value at Risk*, McGraw Hill.

JP Morgan, 1996, *RiskMetrics*^{TM} Technical Document, fourth edition, New York.

Sharpe, W., 1994, "The Sharpe Ratio", *Journal of Portfolio Management*, Fall, 49-58.

Table 1 Estimated Unconditional Bias of RiskMetrics Estimates of the Conditional Standard Deviation

	$\hat{\sigma}$ ($\times 10^{-3}$)	$\hat{k}_U$	$\text{vâr}(\hat{\sigma}_t^2)$ ($\times 10^{-8}$)	Bias (Lower) (%)	Bias (Upper) (%)	Bias ($k_C = 3$) (%)
US Bond	5.36	4.77	0.022	-3.35	45.04	4.22
UK Bond	4.95	7.43	0.027	-5.73	64.85	9.43
Japan Bond	3.88	11.71	0.032	-20.15	133.11	21.40
SP500	10.32	7.09	1.040	-12.70	71.42	6.13
FTSE100	10.41	5.98	1.240	-14.68	54.77	-0.82
Nikkei225	15.05	6.51	2.820	-7.87	70.98	8.87
USD/GBP	5.58	6.05	0.064	-8.47	56.29	4.57
USD/JPY	7.01	8.11	0.172	-9.42	84.93	13.18
GBP/JPY	7.47	6.16	0.150	-6.67	64.64	7.88

Notes: The sample period used is 30/12/87 to 4/12/02. The EWMA model is used to estimate the conditional variance of returns using a decay factor of 0.94. The first 500 observations are discarded in order to make the EWMA model insensitive to the initial conditional variance. $\hat{\sigma}$ is the sample standard deviation of returns. $\hat{k}_U$ is the sample kurtosis coefficient of returns, computed as the sample average of $(r_t - \bar{r})^4$ divided by $\hat{\sigma}^4$, where $\bar{r}$ is the sample mean. $\text{vâr}(\hat{\sigma}_t^2)$ is the sample variance of the estimated conditional variance. 'Bias (Lower)' and 'Bias (Upper)' are the lower and upper bound for the unconditional bias in the estimated conditional standard deviation, estimated using equation (12) with $\hat{\sigma}$, $\hat{k}_U$ and $\text{vâr}(\hat{\sigma}_t^2)$, reported as a percentage of the sample average of the estimated conditional standard deviation. 'Bias ($k_C = 3$)' is the unconditional bias in the estimated conditional standard deviation, estimated using equation (11) with $k_C = 3$.