Abstract

This study assesses the accuracy of the value-at-risk estimate (VaR). On the basis of posterior distributions of the unknown population parameters, we develop a confidence interval for VaR that reflects the genuine information available about the portfolios for which the VaR is calculated. This approach is more accurate than that in Dowd (2000) as it avoids explaining the behaviour of the population parameters on the basis of distributions of sample parameters. We find that the accuracy of both the confidence interval and the VaR estimate depend more dramatically on the sample size than what Dowd's results suggest. In addition, we not only find that the impact of the confidence level and the holding period at which the VaR is predicated are negligible compared to that of the sample size (as in Dowd), but also that the confidence interval is far from being symmetric.

JEL Classification Number: C15, GOO
Keywords: Bayesian Statistics, Confidence Interval, Monte Carlo Simulations, Value-at-Risk.

A Bayesian Confidence Interval for Value-at-Risk

Patricio Contreras[1] and Stephen Satchell[2]

November 2003

1. Introduction

The disastrous consequences of excessive exposure to market risk experienced in the past years revitalized the search for early-warning and forward-looking indicators of financial vulnerability. The ability of value-at-risk (VaR) models to determine maximum expected losses taking into account a portfolio's overall exposure to different types of risk – e.g. exchange rate, interest rate, maturity – has led to the widespread use of these models for internal monitoring and managing of market risk exposure. Furthermore regulators have advocated the use of VaR models for purposes of financial solvency assessment and, in the case of banks, for capital adequacy determination.[3] At the macroeconomic level, recent literature has proposed the use of a "macroeconomic" value-at-risk approach that focuses on the solvency of a nation's aggregate balance sheet- in order to assess the vulnerability of particular economic regimes.[4]

The widespread and increasing reliance on VaR highlights the critical importance of the accuracy of the value-at-risk estimate. Current implementations of VaR, however, do not recognise the fact that VaR measures are only estimates of risk. Because VaR numbers are estimated from sample data it is likely that sampling error feeds through to the VaR figure, making the estimate risky itself. More critically, most often VaR numbers rely on parameters –means, standard deviations, and quantiles – estimated from historical data, which provide a poor guide to future values.

Consider the following example taken from Britten-Jones (1999). Assume a portfolio is composed of one riskless asset and K risky assets. Minimum variance portfolio weights, $\hat{\omega} = (X'X)^{-1}X'\mathbf{1}$, are obtained by running the "artificial" OLS regression $\underset{Tx1}{\mathbf{1}} = \underset{(Txk)(kx1)}{X\omega} + \underset{Tx1}{u}$, where $\mathbf{1}$ is a vector of ones, X is a matrix of excess return vectors x_t, and u is a

---

[1] The Judge Institute of Management Studies, University of Cambridge, UK. Email: pcontreras@oas.org

[2] Faculty of Economics and Politics, University of Cambridge, UK. Email: stephen.satchell@econ.cam.ac.uk.

[3] The Basle Committee on Banking Supervision permitted at the beginning of 1998 that banks determine their capital adequacy for financial risk exposure using VaR models.

[4] See Blejer and Schumacher (1998).

residual vector. The dependent variable **1** is interpreted as a sample counterpart to arbitrage profits -positive excess return with zero standard deviation; the coefficients ω represent the weights on risky assets in the portfolio; $X\omega$ represents excess returns;[5] and the residual vector u shows deviations in the portfolio's return from **1**. Poor estimates of future returns, such as historical returns, make the weights estimates -and hence the corresponding mean and variance figures - subject to "estimation risk." Note that "estimation risk" arises not only with parametric VaR approaches, but also with non-parametric approaches that rely on simulated (rate of return) data but make use of historical portfolio weights.

To recognise the existence of "estimation error" not only provides a decision criterion for selecting the confidence level and the estimation method with the best sampling characteristics, but it also highlights the relevance of reporting the VaR number with confidence intervals.[6] The aim of the paper is to provide a method for calculating a confidence interval for VaR that minimizes estimation risk and that can accommodate a wide range of probability distributions of portfolio returns. The paper is organized as follows: section two presents the value-at-risk measure, section three summarises the relevant literature and section four proposes an alternative Bayesian Approach to the estimation of a confidence interval for VaR. The estimation procedure and results are presented in sections five to seven.

2. Parametric VaR

For general distributions the VaR of a portfolio can be defined as the absolute pound loss,

$$\mathrm{VaR} = W_0 - W^* = -W_0 r^*, \qquad (1)$$

where W_0 is the initial investment in the portfolio, r its rate of return, and $W^* = (1 + r^*)W_0$ the lowest portfolio value at a given confidence level α.[7]

For location-scale densities[8] a parametric representation of VaR is obtained by resorting to the invariance of such densities to linear

[5] The estimated portfolio weights produce a portfolio return that is closest in terms of least squares distance to the arbitrage return vector **1**. It is straightforward to adapt a mean-standard deviation setting.
[6] See Jorion (1996), page 47.
[7] See Jorion (1997), page 87.

transformations. Let $b = \frac{r^* - \mu}{\sigma}$, so that $r^* = b\sigma + \mu$, where b is the standardised random variable, and σ and μ the (scale) standard deviation and (location) mean of r, respectively. Assume $W_0 = 1$, substitute r^* into (1) and get

$$VaR = -(a\mu + b\sigma), \qquad (2)$$

where a equals one and b is the (1-α)% value of the standardised distribution. Hence, when r_i, i=1,…,N, are identically and independently distributed normal with parameters μ, σ^2, $r_i \sim N(\mu, \sigma^2)$, and the confidence level 1-α is set at 95%, b = 1.645 is readily obtained from the standard normal table.[9] To find the VaR under a Student distribution, equation (2) still applies, but b should be replaced by the appropriate 1-α standard Student deviate. For example, for a standard Student with six degrees of freedom at the 95% confidence level, b equals 1.943.

There are two convenient attributes to this approach. The first is that it can easily accommodate distributions for which the dispersion can be adequately summarized by one parameter, the standard deviation. Even fat-tailed distributions, such as those of stock prices, can be accommodated as long as they are "fairly" symmetric. Strongly asymmetric distributions, such as those of derivatives, invalidate the procedure, as standardisation of the distribution is no longer adequate. For large and diversified portfolios, however, the issue is one of fatness of tails, not asymmetry.[10]

The second compelling attribute of the "sigma-based approach" presented in this section relates to efficiency gains. Jorion (1996) shows that "the sample standard deviation method has uniformly lower standard errors and is, therefore, uniformly superior to the sample quantile method."[11]

---

[8] A location scale density has the form $\sigma^{-1} f\left(\frac{x-\theta}{\sigma}\right)$, where $\theta \in R^1$ and $\sigma > 0$ are the unknown parameters. The class of such densities is invariant under the group of affine transformations $g_{b,c}(x) = cx + b$ (Berger, 1980, pages 88 and 401).

[9] When VaR is measured relative to the mean, $VaR = E(W) - W^* = -W_0(R^* - \mu)$, the parametric representation of VaR for location-scale returns is $VaR = -b\sigma W_0$.

[10] Involving the Central Limit Theorem, near-normality can be a reasonable assumption provided the portfolio is well diversified and individual returns are sufficiently independent of each other.

[11] The appropriate quantile can be obtained directly from the historical distribution (Quantile-based VaR) or indirectly by measuring the standard deviation and

3. A Confidence Interval for VaR

Dowd (2000)[12] writes that the "most natural way" to gauge the precision of a VaR estimate is to construct a confidence interval for it. The simplest way to do it is to assume normality and that μ is known (e.g. zero), so that the VaR estimate is given by

$$\hat{VaR} = -bsW_0 . \tag{3}$$

If we draw a random sample of size N from a normal distribution, the variable $\frac{(N-1)s^2}{\sigma^2}$ will be distributed as a chi-squared with $\upsilon = N-1$ degrees of freedom (χ^2_υ), where s^2 is known and σ^2 is unknown. One can say that there is an $\frac{\alpha}{2}\%$ probability that this variable will fall below the $\chi^2_{\upsilon,\frac{\alpha}{2}}$ quantile and an $\frac{\alpha}{2}\%$ probability that it will fall above the corresponding $\chi^2_{\upsilon,1-\frac{\alpha}{2}}$. It follows that the $(1-\alpha)\%$ confidence interval for $\frac{(N-1)s^2}{\sigma^2}$ must be:

$$P_D = \left(\chi^2_{\upsilon,\frac{\alpha}{2}} < \frac{(N-1)s^2}{\sigma^2} < \chi^2_{\upsilon,1-\alpha} \right) = (1-\alpha)\%, \tag{4}$$

where P_D is the $(1-\alpha)\%$ probability value calculated using the method of Dowd.

By transformation, given a sample value of s, the $(1-\alpha)\%$ confidence interval for the sample standard deviation s is given by:

$$P_D = \left(\sqrt{s^2 \frac{N-1}{\chi^2_{\upsilon,1-\frac{\alpha}{2}}}} < \sigma < \sqrt{s^2 \frac{N-1}{\chi^2_{\upsilon,\frac{\alpha}{2}}}} \right) = (1-\alpha)\% \tag{5}$$

Multiplying equation (5) by $-b$ yields a confidence interval for the VaR:

---

multiplying it by an appropriate scaling factor (Sigma-based VaR). Jorion (1997) and (1996).

[12] The confidence interval was originally developed in Chappel and Dowd (1999), however no testing was conducted in this paper.

$$P_D = \left(-b\sqrt{s^2 \frac{N-1}{\chi^2_{\upsilon,1-\frac{\alpha}{2}}}} < VaR < -b\sqrt{s^2 \frac{N-1}{\chi^2_{\upsilon,\frac{\alpha}{2}}}} \right) = (1-\alpha)\%, \qquad (6)$$

where b is a parameter reflecting the confidence level on which the VaR is predicated.

However, if both μ and σ are unknown variables the VaR estimate, $\hat{VaR} = -(a\bar{r} + bs)$, will depend on two stochastic variables, s and $\bar{r}$. The construction of analytic confidence intervals for this more general case is "nearly impossible" (Kendall and Stuart (1973) in Chappel and Dowd (1999)). The alternative is to construct a "confidence distribution for VaR by simulating each of these terms using the information available. This information consists of s and $\bar{r}$ and their distributions.

Statistical theory tells us that $\bar{r} \sim N\left(\mu, \frac{\sigma^2}{N}\right)$ and that $\frac{(N-1)s^2}{\sigma^2} \sim \chi^2_\upsilon$. After a little rearranging, we can therefore treat σ and μ as if they were random in the sense that:

$$\sigma \sim \sqrt{\frac{(N-1)s^2}{\chi^2_\upsilon}} \quad \text{and} \quad \mu \sim N\left(\bar{r}, \frac{(N-1)s^2}{n\chi^2_\upsilon}\right) \qquad (7)$$

Substituting into (2) and making a=1, the "confidence distribution" for VaR is given by:

$$-b\sqrt{\frac{(N-1)s^2}{\chi^2_\upsilon}} - N\left(\bar{r}, \frac{(N-1)s^2}{n\chi^2_\upsilon}\right) \qquad (8)$$

This distribution can be simulated and the confidence intervals read from the quantiles of the simulated distribution.

4. A Bayesian Approach

A weakness of the above procedure is that it relies on distributions of sample parameters, $\bar{r}$ and s, that are conditional on unknown population parameters, μ and σ^2, to explain the behaviour of the unknown parameters. A more realistic approach that reflects the genuine information that is available about the behaviour of the population parameters is presented in this section.

Our procedure follows the above procedure in that a confidence region for VaR is estimated, assuming that μ and σ^2 are distributed according to some posterior probability density functions (*pdfs).* These

functions are estimated through an empirical Bayes approach[13], which enables us to reduce the estimation error arising when the sample parameters are treated as if these were the true (unknown) parameters.

Our approach works as follows: first we let r_i be i.i.d $N(\mu,\sigma^2)$, and μ and σ be independently distributed so that $\Pi(\mu,\sigma^2)=\Pi(\mu)\Pi(\sigma^2)$, $\Pi(.)$ denotes a posterior distribution. This leads to the following confidence distribution for VaR as specified in equation (2) and assuming a=1:

$$-b\Pi(\sigma)-\Pi(\mu) \qquad (9)$$

4.1. Prior and Posterior Distributions

We assume that plausible functional forms for μ and σ^2 are given by a conjugate family of prior distributions and that one can draw on prior information contained in the cross-sectional pattern of the stock returns composing the portfolio to infer the parameters specifying such functional forms.[14] The assumption of a conjugate family of priors is usually robust and simplifies calculations.[15]

We make the assumption that the individual stock returns are statistically independent through time and that the joint distribution of the cross-section of returns is identical across time. Although unrealistic, we make this assumption following the relevant literature and to simplify the calculations.

Once prior distributions have been estimated, posterior distributions are readily calculated by application of Bayes Theorem:[16]

$$\Pi(\theta/x)\,\alpha\,\pi(\theta/x)l(x/\theta), \qquad (10)$$

where $\Pi(\theta/x)$ and $\pi(\theta/x)$ are the posterior and prior *pdfs*, respectively, for the parameter vector θ, given the sample information x, and $l(x/\theta)$ is the likelihood function; α denotes proportionality.

---

[13] See Berger (1980 a and b).

[14] The procedure is a direct application of the approach of Karolyi (1993), who estimates the stock return volatility for a given stock by drawing on prior information from the cross-sectional pattern in the return volatitilities for a whole group of stocks. He draws on the empirical fact that stock return volatilities are generally clustered about some market-wide measure of volatitility and also within subgroups of stocks sorted by a firm's degree of financial leverage, the level of trading volume in a firm's stock, and a firm's capitalisation value.

[15] See Berger (1980b).

[16] See Berger (1980a).

4.1.1 Prior and Posterior *Pdfs* of μ

In order to estimate the prior distribution of μ we apply a type II maximum likelihood (ML-II) prior (Berger 1980a). This empirical Bayes approach assumes that a suitable prior is obtained as the result of maximising the likelihood function of the marginal distribution of the independent components of the data given the unknown true prior distribution.

Consider a portfolio composed of P independent stock returns $(r_1,...,r_P)$, each with density $f(r_i \mid \mu_i) = N(\mu_i, \sigma_f^2)$, and let the conjugate prior $\pi(\mu)$ be $N(\mu_\pi, \sigma_\pi^2)$.[17] Then the marginal density of each r_i, $m_0(r_i \mid \pi_0)$ is $N(\mu_\pi, \sigma_\pi^2 + \sigma_f^2)$. The corresponding likelihood function is

$$m(r \mid \pi) = \prod_{i=1}^{P} m_0(r_i \mid \pi_0) = [2\pi(\sigma_\pi^2 + \sigma_f^2)]^{-\frac{P}{2}} \exp\left\{\frac{-Ps^2}{2(\sigma_\pi^2 + \sigma_f^2)}\right\} * \exp\left\{\frac{-P(\bar{r} - \mu_\pi)^2}{2(\sigma_\pi^2 + \sigma_f^2)}\right\} \tag{11}$$

where $\bar{r} = \sum_{i=1}^{P} \frac{r_i}{P}$ and $s^2 = \sum_{i=1}^{P} \frac{(r_i - \bar{r})^2}{P}$.

Berger shows that maximisation of (11) leads to the ML-II prior $\hat{\pi}_0$ to be $N(\hat{\mu}_\pi, \hat{\sigma}_\pi^2)$, where $\hat{\mu}_\pi = \bar{r}$ and $\hat{\sigma}_\pi^2 = \max\{0, s^2 - \sigma_f^2\}$. A suitable estimator of σ_f^2 is given by $\hat{\sigma}_f^2 = \frac{1}{N(N-1)P} \sum_{i=1}^{P} \sum_{j=1}^{N} (r_i^j - \bar{r}_i)^2$ (Berger (1980a) page 172).

The posterior *pdf* of μ is readily obtained after substituting the prior *pdf* of μ and the likelihood of $f(r \mid \mu) = N(\mu, \sigma_f^2)$ into the proportional relation (10). This leads to a Normal *pdf* with mean $\frac{\bar{r}\hat{\sigma}_f^2 + \hat{\mu}_\pi \frac{\hat{\sigma}_f^2}{N}}{\hat{\sigma}_\pi^2 + \frac{\hat{\sigma}_f^2}{N}}$ and variance $\frac{\hat{\sigma}_\pi^2 \frac{\hat{\sigma}_f^2}{N}}{\hat{\sigma}_\pi^2 + \frac{\hat{\sigma}_f^2}{N}}$ (Zellner, 1971, page 15).

[17] In dealing with a normal mean, the class of $N(\mu_\pi, \sigma_\pi^2)$ priors is rich enough to include approximations to most reasonable priors. See Berger (1980a) section 4.7 for cases when this leads to unappealing conclusions.

4.1.2. Prior and Posterior *Pdfs* of σ^2

In order to estimate a suitable prior for σ^2 we follow Karolyi (1993). Assuming independent, linear stochastic processes for the stock returns of each stock i, i=1,…,P; we know that the statistic $\frac{\upsilon_i s_i^2}{\sigma_i^2}$ is distributed chi-squared with $\upsilon_i = N_i - 1$ degrees of freedom ($\chi^2_{\upsilon_i}$). By transformation, s_i^2 has the following density

$$f_i(s_i^2 \mid \sigma_i^2, \upsilon_i) = \frac{1}{\Gamma\left(\frac{\upsilon_i}{2}\right)} \left(\frac{\upsilon_i s_i^2}{2}\right)^{\frac{\upsilon_i}{2}} \left(\frac{1}{\sigma_i^2}\right)^{\frac{\upsilon_i}{2}} \left(\frac{1}{s_i^2}\right) \exp\left(-\frac{\upsilon_i s_i^2}{2\sigma_i^2}\right) \tag{12}$$

where $\Gamma(y) = \int_o^\infty x^{y-1} \exp(-x)dx$, the Gamma function.

Assume now that the σ_i^2 are distributed in the population according to an inverse gamma distribution (the conjugate prior) about an unknown location parameter, τ, and with some unknown parameter of dispersion, ν:[18]

$$\pi(\sigma_i^2) = \frac{\exp\left(-\frac{\nu\tau}{2\sigma_i^2}\right)\left(\frac{\nu\tau}{2}\right)^{\frac{\nu}{2}}}{\sigma_i^{2\left(\frac{\nu}{2}+1\right)}\Gamma\left(\frac{\nu}{2}\right)}, \tag{13}$$

where $\nu, \tau > 0$ and $0 < \sigma_i^2 < \infty$. The expected value, variance and skewness of σ_i^2 are given by

$$E(\sigma_i^2) = \frac{\nu\tau}{2\left(\frac{\nu}{2}-1\right)}, \ \nu > 2 \tag{14}$$

$$V(\sigma_i^2) = \frac{(\nu\tau)^2}{4\left(\frac{\nu}{2}-1\right)^2\left(\frac{\nu}{2}-2\right)}, \ \nu > 4. \tag{15}$$

[18] Empirical Bayes analysis is usually fairly robust with respect to the functional form chosen for π. See Berger (1980) and Zellner (1971).

$$\text{Pearson measure of skewness} = \frac{2\left(\frac{\nu}{2}-2\right)^{\frac{1}{2}}}{\left(\frac{\nu}{2}+1\right)} \tag{16}$$

Note that as ν gets large the mean $E(\sigma_i^2)$ converges to τ and the variance $V(\sigma_i^2)$ and skewness both converge to zero, so that σ_i^2 converges in mean square to τ. For ν greater than four the prior for σ_i^2 has a rather long tail to the right.

The squared coefficient of variation, $\theta^2 = \frac{V(\sigma_i^2)}{E^2(\sigma_i^2)} = \frac{2}{\nu - 4}$, enables us to infer a suitable value for the dispersion parameter, ν, from prior information. For instance, assume that the average volatility of the portfolio stock returns is approximately 25% and that its variance is approximately 3%, then ν is approximately eight.

The marginal densities

$$m_i(s_i^2 \mid \nu, \tau) = \int f_i(s_i^2 \mid \upsilon_i, \sigma_i^2)\pi(\sigma_i^2 \mid \nu, \tau)d\sigma_i^2 , \tag{17}$$

provide the vehicle for the estimation of ν and τ.

A mixture of method of moments and maximum likelihood methods are used to estimate ν and τ. Using the method of moments approach, an estimator of τ is given by

$$\hat{\tau} = \frac{\sum_{i=1}^{P} \frac{\upsilon_i s_i^2}{\upsilon_i + \nu + 2}}{\sum_{i=1}^{P} \frac{\upsilon_i}{\upsilon_i + \nu + 2}}, \tag{18}$$

which simplifies to an arithmetic average of the individual stock's sample return variances, $\sum_{i=1}^{P} \frac{s_i^2}{P}$, when υ_i is the same for each stock.[19] The joint log-likelihood of ν, τ from the product of the marginal conditional densities of (17) is

[19] This method is however limited to stocks with common characteristics. See footnote 14.

$$\text{Log L}(\nu,\tau \mid \upsilon_i, s_i^2) = \text{constant} + P\frac{\nu}{2}\text{Log}\left(\frac{\nu\tau}{2}\right) + \sum_{i=1}^{P}\text{Log}\left(\frac{\Gamma\left(\frac{\upsilon_i+\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\right) \qquad (19)$$

$$-\sum_{i=1}^{P}\frac{\nu+\upsilon_i}{2}\text{Log}\left(\frac{\upsilon_i s_i^2+\nu\tau}{2}\right)$$

Replacing τ with $\hat{\tau}$ from (18) and maximising (19) over ν yields $\hat{\nu}$. It is possible though that maximisation of (19) leads to $\nu = \infty$ or ν near zero (see Hui and Berger, 1983).

As before, the posterior *pdf* of σ_i^2 is readily obtained after substituting the likelihood equation obtained by transformation of the $\chi^2_{\upsilon_i}$ *pdf* of the statistic $\frac{\upsilon_i s_i^2}{\sigma_i^2}$, and the prior *pdf* of σ_i^2 into the proportional relation (10). This leads to an inverse gamma *pdf* with parameters $\nu_i'' = \upsilon_i + \hat{\nu} + 2$ and $s_i^{2''} = \frac{\upsilon_i s^2 + \hat{\nu}\hat{\tau}}{\nu''}$.

5. Data

The data employed in this study consists of daily prices for all stocks listed in the London Stock Exchange for the period January 1993 to December 2002. The prices for each day are taken at the close of market and are adjusted for subsequent capital actions. Stock returns are then calculated as the difference in logarithms.

In conducting the Bayesian analysis stocks are grouped on the basis of cross-sectional prior information. Karolyi (1993) summarizes the evidence supporting the idea that each of leverage, volume, and size represents instrumental variables related to the true return variances so that a reasonable alternative choice of prior density should include subgroups of return variances sorted by the stock's measures of leverage, volume and size. This study employs these three grouping criteria, in addition to a separate group containing all stocks in the LSE. The latter takes account of the case where no prior information about the stock is available, other than the fact that it comes from the population of all stocks in the LSE. Each one of these groups is later divided into three subgroups (high, medium, low), which provide the basis for the estimation of ten equally weighted portfolios.

The balance sheet and trading information used to group the stocks is selected in order to ensure that such information is known at the moment of estimating the results. Although different cutting points are tested, the results presented below are calculated on the basis of balance

sheet and trading information published for the third quarter of 2002. Prior hyperparameters are estimated for the last day of the time series. The fact that the results are virtually unchanged when different cutting points are considered, suggests a reasonable level of stationarity.

All data is taken from Datastream and simulations are conducted using Gauss Aptech Systems.

6. Simulations

Confidence regions (8) and (9) above are estimated for each of the ten portfolios described in the previous section by running 10,000 Monte Carlo simulations in each case.[20] This produces a histogram for each portfolio, from which the mean and upper and lower boundaries are read. As in Dowd, the confidence intervals are estimated at the 95% confidence level. The mean of each distribution is regarded as an estimate of the unknown "true" VaR and the 2.5% and the 97.5% quantiles of the distributions are regarded as estimates of the lower and upper bounds. VaRs are predicated on the 95% and 99% confidence levels and on one-day and 30-day holding periods. The simulations are carried out for a sample size (N) ranging from 100 to 2580 (the complete sample).

The accuracy of the VaR estimate and the lower and upper bounds is assessed by measuring their distance (i.e. biasedness) from the "true" VaR and their sensitivity to the sample size, holding period and confidence level.

7. Results and Conclusions

The results of the simulations are shown in tables 1 to 3 below for VaRs predicated on a 30-day holding period. Three additional models were also estimated, however these were excluded from the body of the study (and presented in Annex I) due to the weak results obtained from the simulations. All simulated equations are shown in Annex II in functional form.

The tables show simulation results for sample sizes ranging from 100 to 2580. For N below 600, results are missing for some portfolios as the maximisation of equation (19) fails to converge in these cases. This is consistent with Hui and Berger in the sense that it is possible that maximisation of this equation leads to $\nu = \infty$ or ν near zero.

Our results basically validate the tendencies found in Dowd, however we find that the precision of both the VaR estimate and the confidence interval are more dramatically linked to the sample size than

[20] Dowd uses the Latin Hypercube routine in @Risk.

what Dowd suggests. For instance, Dowd finds that the conventional VaR estimate (equation 2) suffers from a downward bias, but this bias gets smaller as the sample size increases. Dowd reports that the bias of the conventional VaR for a 95% confidence interval decreases from 1% when N=100 to 0.1% when N=1000. When we run Dowd's model with our sample data the bias across groups ranges between 0.35% and 1.6% when N equals 100 and rapidly approaches zero (for all groups) as N increases (Table 1). Only in one case is the bias higher that 0.1% when N is 600 or higher. Our model, on the other hand, suggests that the bias is much larger than what Dowd reports and that it decreases much more dramatically as the sample size increases. For instance, the bias of the conventional VaR estimate for the portfolio "Lev high" decreases from 89% when N=100 to 1% when the complete sample is used. Likewise, when N is maximum the bias is less than 1% in all cases, except for portfolios "Lev high" and "Lev med." This evidence indicates that a "reasonably" precise VaR estimator requires a sample size considerably larger than what Dowd suggests.

The accuracy of the VaR estimate also differs considerably across portfolios in our model. We assess the variability of the bias across portfolios by calculation of the standard deviation, although we acknowledge that this measure of variability provides limited information when calculated for such a small number of data points. We find that the standard deviation of the bias across portfolios is 13% when N equals 600, and decreases to a mere 0.3% when all the sample information is considered. This suggests that the impact of the different groupings becomes considerably more relevant as the sample size decreases. Dowd's model does not support this conclusion as the standard deviation of the bias across portfolios in his model is always less than 0.1%, except for N=100, when the standard deviation is 0.35%. These results remain virtually unchanged when the VaR is predicated at the 99% confidence level.

Similar results can be reported for the lower and upper bounds of the confidence interval. Both the lower and upper bound move further away from the mean of the distribution as each of the sample size and the confidence level decreases. This tendency is less clear in both Dowd's and our model when the holding period increases. For instance, for "Lev high" when N=2580, the lower bound under both models moves closer to the mean of the distribution when the holding period increases. The same happens with "Size high" and "Size med" when N is 1500 or higher. In two additional cases, "Vol high" and "Lev med", the lower bound moves closer to the mean of the distribution when the holding period increases in our model, but not in Dowd's.

Similarly, the upper bound moves closer to the mean of the distribution when N =2580 for "Vol high." This also happens for "Size high" when N is 1500 or higher and again the results go in opposite directions for "Lev med" when N=2580, although this time it is Dowd's model which shows the "right" direction.

The sensitivity of the lower and upper bounds to the different groupings of stocks is extremely low (0.2%) and virtually identical in both models for large N. As N gets smaller the standard deviation across portfolios of both the distance between the lower bound and the mean of the distribution and the upper bound and the mean gets slightly bigger, although it remains below the 2% level under both models. Only when N equals 100 does the standard deviation rise to levels of 12%.

Dowd concludes that the confidence interval depends mostly on the sample size and gets smaller as N gets larger. He reports that the confidence interval decreases from roughly the VaR estimate plus or minus 20% when N=100 to the VaR estimate plus or minus 9% when N=1000. We verify these results when we run Dowd's model with our data although with some variability across portfolios (Table 3). When N is 1500 or higher, our model behaves similarly to that of Dowd, although we find a considerable degree of asymmetry between lower and upper bounds. When N is less than 1500 our model becomes increasingly less precise as the VaR estimate falls outside the interval in an amount that increases as N gets smaller. This also happens in one case, "Lev med", when N equals 1500. All of these cases are characterized by a very high v value, whereas when v is small (say below 30) the model provides precise results even for small N.

We also verify Dowd's conclusion that the impact on the confidence interval of the confidence level and holding period are negligible compared to that of the sample size.

Confidence level = 95%

VaR (distance from mean)

	ALL LSE	Vol high	Vol med	Vol low	Lev high	Lev med	Lev low	Size high	Size med	Size low
					N=2580					
Dowd	-0.01%	-0.01%	-0.03%	-0.02%	-0.01%	-0.04%	-0.03%	0.00%	0.02%	0.01%
C&S	-0.34%	-0.56%	-0.75%	-0.71%	-1.00%	-1.40%	-0.97%	-0.58%	-0.48%	-0.67%
					N=1500					
Dowd	0.04%	-0.04%	-0.04%	-0.02%	-0.03%	-0.04%	-0.06%	-0.01%	-0.05%	-0.02%
C&S	-2.12%	-1.73%	-2.23%	-2.19%	-3.45%	-6.01%	-3.00%	-2.41%	-1.66%	-1.74%
					N=1000					
Dowd	-0.07%	0.02%	-0.02%	-0.02%	-0.09%	-0.07%	-0.08%	0.01%	-0.03%	-0.02%
C&S	-6.72%	-6.11%	-7.63%	-6.49%	-13.99%	-26.47%	-10.96%	-9.18%	-5.81%	-5.46%
					N=600					
Dowd	-0.06%	0.02%	-0.04%	-0.11%	-0.08%	-0.02%	-0.09%	-0.02%	-0.04%	-0.21%
C&S	-31.45%	-26.00%	-36.00%	-36.80%	-52.32%		-56.10%		-17.38%	-33.72%
					N=100					
Dowd	-0.54%	-0.78%	-0.72%	-0.70%	-1.62%	-0.84%	-0.62%	-0.46%	-0.35%	-0.63%
C&S					-88.09%			-30.29%		

Lower boundary (distance from mean)

					N=2580					
Dowd	2.98%	2.80%	2.92%	3.12%	2.76%	2.73%	3.02%	2.53%	2.62%	3.15%
C&S	3.06%	2.89%	2.89%	3.15%	2.73%	2.84%	2.98%	2.59%	2.65%	3.12%
					N=1500					
Dowd	4.22%	4.01%	4.04%	4.68%	4.12%	3.98%	4.19%	3.52%	3.52%	4.52%
C&S	4.21%	3.95%	4.13%	4.51%	4.01%	3.97%	4.23%	3.51%	3.55%	4.37%
					N=1000					
Dowd	5.50%	5.69%	5.16%	5.95%	5.21%	5.08%	5.47%	4.96%	4.73%	5.71%
C&S	5.46%	5.30%	5.10%	5.64%	5.00%	4.79%	5.26%	4.52%	4.67%	5.56%
					N=600					
Dowd	9.95%	9.43%	8.91%	11.88%	8.84%	8.32%	9.86%	6.73%	6.74%	11.89%
C&S	8.31%	8.06%	7.34%	9.55%	6.14%		6.26%		6.23%	9.41%
					N=100					
Dowd	31.25%	26.75%	28.51%	39.24%	66.38%	31.98%	27.79%	19.71%	20.20%	36.87%
C&S					17.07%			14.86%		

Upper boundary (distance from mean)

					N=2580					
Dowd	-2.83%	-2.70%	-2.87%	-3.02%	-2.70%	-2.73%	-2.92%	-2.43%	-2.50%	-3.03%
C&S	-2.86%	-2.68%	-2.81%	-3.01%	-2.72%	-2.63%	-2.89%	-2.42%	-2.49%	-3.00%
					N=1500					
Dowd	-3.99%	-3.81%	-3.87%	-4.29%	-3.69%	-3.76%	-3.95%	-3.32%	-3.48%	-4.26%
C&S	-3.90%	-3.81%	-3.90%	-4.42%	-3.71%	-3.75%	-3.93%	-3.24%	-3.36%	-4.28%
					N=1000					
Dowd	-5.12%	-5.16%	-4.71%	-5.46%	-5.11%	-5.00%	-5.08%	-4.40%	-4.24%	-5.32%
C&S	-5.05%	-4.88%	-4.80%	-5.25%	-4.81%	-4.48%	-4.92%	-4.27%	-4.37%	-5.31%
					N=600					
Dowd	-9.11%	-8.25%	-8.32%	-10.94%	-8.01%	-7.32%	-8.90%	-6.18%	-6.24%	-10.76%
C&S	-7.30%	-7.13%	-6.82%	-8.78%	-5.66%		-5.69%		-5.58%	-8.38%
					N=100					
Dowd	-23.97%	-21.25%	-22.95%	-31.02%	-52.27%	-25.12%	-21.82%	-15.40%	-15.72%	-29.00%
C&S					-14.23%			-12.23%		

Table 1

Confidence level = 99%

VaR (distance from mean)

	ALL LSE	Vol high	Vol med	Vol low	Lev high	Lev med	Lev low	Size high	Size med	Size low
					N=2580					
Dowd	-0.02%	-0.02%	-0.02%	0.00%	-0.02%	-0.02%	-0.02%	-0.02%	-0.01%	-0.01%
C&S	-0.33%	-0.56%	-0.74%	-0.69%	-1.05%	-1.39%	-1.01%	-0.59%	-0.50%	-0.66%
					N=1500					
Dowd	-0.02%	-0.03%	-0.02%	-0.04%	-0.03%	-0.07%	-0.05%	-0.01%	-0.02%	-0.04%
C&S	-2.03%	-1.60%	-2.16%	-2.00%	-3.38%	-5.82%	-2.89%	-2.42%	-1.62%	-1.63%
					N=1000					
Dowd	-0.07%	-0.09%	-0.04%	0.01%	-0.03%	-0.06%	-0.05%	-0.05%	-0.03%	-0.04%
C&S	-6.36%	-5.75%	-7.35%	-6.02%	-13.40%	-25.61%	-10.46%	-9.08%	-5.78%	-5.09%
					N=600					
Dowd	-0.12%	-0.07%	-0.08%	-0.07%	-0.09%	-0.14%	-0.05%	-0.04%	-0.05%	-0.18%
C&S	-27.73%	-23.22%	-32.76%	-31.03%	-49.02%		-51.74%		-16.91%	-28.50%
					N=100					
Dowd	-0.55%	-0.36%	-0.36%	-0.76%	-0.89%	-0.42%	-0.43%	-0.35%	-0.22%	-0.50%
C&S					-79.07%			-28.85%		

Lower boundary (distance from mean)

					N=2580					
Dowd	3.00%	2.82%	2.86%	3.02%	2.69%	2.80%	2.95%	2.58%	2.67%	2.99%
C&S	3.04%	2.84%	2.95%	3.08%	2.71%	2.78%	2.89%	2.71%	2.66%	2.99%
					N=1500					
Dowd	4.10%	3.93%	3.91%	4.28%	3.86%	3.90%	4.06%	3.53%	3.63%	4.13%
C&S	3.94%	3.87%	4.01%	4.15%	3.77%	3.83%	3.94%	3.55%	3.66%	4.14%
					N=1000					
Dowd	5.12%	5.12%	4.98%	5.35%	5.20%	4.97%	5.12%	4.65%	4.68%	5.46%
C&S	5.22%	5.24%	5.07%	5.41%	4.90%	4.58%	5.10%	4.48%	4.61%	5.29%
					N=600					
Dowd	8.15%	7.97%	7.87%	9.22%	7.82%	7.36%	8.33%	6.53%	6.49%	9.18%
C&S	7.33%	7.24%	6.64%	7.79%	5.71%		5.79%		6.02%	7.74%
					N=100					
Dowd	24.80%	22.21%	22.91%	27.51%	34.53%	24.64%	23.13%	17.85%	18.63%	25.98%
C&S					15.60%			13.89%		

Upper boundary (distance from mean)

					N=2580					
Dowd	-2.9%	-2.7%	-2.7%	-2.9%	-2.6%	-2.7%	-2.8%	-2.6%	-2.6%	-2.9%
C&S	-2.8%	-2.7%	-2.8%	-2.9%	-2.7%	-2.6%	-2.8%	-2.5%	-2.5%	-2.9%
					N=1500					
Dowd	-3.9%	-3.7%	-3.7%	-4.1%	-3.6%	-3.7%	-3.8%	-3.2%	-3.4%	-3.9%
C&S	-3.7%	-3.6%	-3.7%	-4.1%	-3.6%	-3.7%	-3.8%	-3.3%	-3.4%	-4.0%
					N=1000					
Dowd	-5.0%	-4.9%	-4.6%	-4.9%	-4.7%	-4.7%	-4.8%	-4.4%	-4.2%	-5.0%
C&S	-4.7%	-4.8%	-4.6%	-5.0%	-4.6%	-4.3%	-4.7%	-4.3%	-4.2%	-4.9%
					N=600					
Dowd	-7.5%	-7.3%	-7.2%	-8.4%	-7.2%	-6.7%	-7.5%	-6.0%	-5.9%	-8.2%
C&S	-6.6%	-6.5%	-6.1%	-7.2%	-5.4%		-5.4%		-5.4%	-7.0%
					N=100					
Dowd	-19.2%	-17.2%	-18.0%	-21.8%	-26.6%	-19.2%	-17.9%	-14.2%	-14.3%	-20.7%
C&S					-12.9%			-11.3%		

Table 2

	Distance from VaR estimate									
	ALL LSE	Vol high	Vol med	Vol low	Lev high	Lev med	Lev low	Size high	Size med	Size low
	N=2580									
Dowd										
Lower bound	2.99%	2.81%	2.95%	3.14%	2.77%	2.77%	3.05%	2.52%	2.59%	3.14%
Upper bound	-2.82%	-2.69%	-2.84%	-3.00%	-2.69%	-2.68%	-2.89%	-2.44%	-2.53%	-3.04%
C&S										
Lower bound	3.42%	3.47%	3.68%	3.89%	3.77%	4.30%	4.00%	3.20%	3.15%	3.81%
Upper bound	-2.53%	-2.14%	-2.07%	-2.32%	-1.73%	-1.25%	-1.94%	-1.85%	-2.01%	-2.35%
	N=1500									
Dowd										
Lower bound	4.18%	4.05%	4.08%	4.70%	4.15%	4.02%	4.25%	3.53%	3.57%	4.54%
Upper bound	-4.02%	-3.78%	-3.83%	-4.26%	-3.66%	-3.73%	-3.90%	-3.31%	-3.43%	-4.24%
C&S										
Lower bound	6.47%	5.78%	6.51%	6.85%	7.73%	10.62%	7.46%	6.07%	5.30%	6.23%
Upper bound	-1.82%	-2.12%	-1.71%	-2.28%	-0.27%	2.40%	-0.96%	-0.85%	-1.73%	-2.58%

Table 3

Bibliography

Berger, J.O.[a] (1980) *Statistical Decision Theory and Bayesian Analysis*, Second Ed., Springer-Verlag Publishers.

---------, [b] (1980) *Statistical Decision Theory: Foundations, Concepts and Methods*, Springer-Verlag Publishers.

Blejer, M.I. and Schumacher, L. (May 1998), 'Central Bank Vulnerability and the Credibility of Commitments: A Value-at-Risk Approach to Currency Crises', WP/98/65, *IMF*.

Britten-Jones, M. (April 1999) 'The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights', *Journal of Finance*, 44, 2, 655-671.

Chappel, D. and Dowd, K. (March 1999) 'Confidence Intervals for VaR', *Financial Engineering News*, 1-2.

Cornelius, P.K. (May 2000) 'Trade in Financial Services, Capital Flows, and the Value-at-Risk of Countries', *World Economy* (U.K.); 23, No. 5:649-72.

Dowd, K., (Spring 2000) 'Assessing VaR Accuracy', *Derivatives Quarterly*, 61-63.

Efron, B. and Morris, C. (March 1973) 'Stein's Estimation Rule and Its Competitors-An Empirical Bayes Approach', *Journal of the American Statistical Association*, Vol. 68, No. 341, 117-130.

Frost. P.A. and Savarino, J.E. (September 1986) 'An Empirical Bayes Approach to Efficient Portfolio Selection', *Journal of Financial and Quantitative Analysis*, Vol. 21, No. 3, 293-305.

Hui, S.L. and Berger, J.O. (December 1983) 'Empirical Bayes Estimation of Rates in Longitudinal Studies', *Journal of the American Statistical Association*, Vo. 78, No. 384, 753-760.

Jorion, P. (1997) *Value at Risk, The New Benchmark for Controlling Derivatives Risk*, Chicago, Irwin.

--------- (1996) 'Risk^2:Measuring the Risk in Value at Risk', *Financial Analyst Journal*, 47-56.

Karolyi, G.A., (December 1993) 'A Bayesian Approach to Modeling Stock Return Volatility for Option Valuation', *Journal of Financial and Quantitative Analysis*, Vol. 28, No. 4, 579-594.

Morris, C.N., (March 1983) *'Parametric Empirical Bayes Inference: Theory and Applications'*, *Journal of the American Statistical Association*, Vol. 78, No. 381, 47-55.

Raiffa, H., and Schlaifer, R. (1961) *Applied Statistical Decision Theory*, Cambridge, Massachusetts, Harvard University Press.

Zellner, A. (1971) *An Introduction to Bayesian Inference in Econometrics*, New York, NY: John Wiley and Sons Publishers.

Appendix I: Additional Models

Three additional models were tested, however the results obtained from their simulations are either weaker than those obtained for equation (9) or they fail to achieve "normal" convergence. The models tested are the following:

- Confidence region for VaR on the basis of prior distributions:

$$-b\pi(\sigma)-\pi(\mu), \tag{20}$$

where $\pi(.)$ denotes some prior *pdf* of the unknown parameters μ and σ^2.

- Confidence region for $\mathrm{V\hat{a}R}$

Given the conditional joint distribution $f(\bar{r},s|\mu,\sigma^2)$, the "unconditional" joint distribution $g(\bar{r},s)\equiv \underset{\mu,\sigma^2}{E}[f(\bar{r},s|\mu,\sigma^2)$ may be estimated by taking expectations over the unknown parameters, μ and σ^2. We apply this procedure to *Vâ**R*** (equation 3) and obtain two alternative "unconditional" *pdfs*[21] for $\mathrm{V\hat{a}R}$:

$$-b\underset{\pi(\sigma)}{E}\left[f\left(s/\sigma\right)\right]-\underset{\pi(\mu,\sigma^2)}{E}\left[f\left(\bar{r}/\mu,\sigma^2\right)\right], \tag{21}$$

where $E_{\pi(.)}[f(.)]$ denotes the expected value of the *pdf* $f(.)$ over the unknown parameters, which are distributed according to some prior distributions $\pi(.)$.

$$-b\underset{\Pi(\sigma)}{E}\left[f\left(s/\sigma\right)\right]-\underset{\Pi(\mu,\sigma^2)}{E}\left[f\left(\bar{r}/\mu,\sigma^2\right)\right], \tag{22}$$

where $E_{\Pi(.)}[f(.)]$ denotes the expected value of the pdf $f(.)$ over the unknown parameters, which are distributed according to some posterior distributions $\Pi(.)$.

Equations (21) and (22) require estimation of the joint distribution of μ and σ^2. When both parameters of an Independent Normal process are unknown, the most convenient joint distribution of the two variables –

[21] Although the sample parameters are no longer conditional on the population parameters, they are conditional on the hyperparameters of the prior or posterior distributions.

the natural conjugate of the Normal pdf – is the Normal-gamma distribution[22]:

$$\pi_{NG}\left(\mu,\sigma^2/\hat{\mu}_\pi,\hat{h},\hat{\nu},\hat{\tau}\right)=\pi_{IG}\left(\sigma^2/\hat{\nu},\hat{\tau}\right)\pi_N\left(\mu/\hat{\mu}_\pi,\sigma^2,\hat{h}\right), \qquad (23)$$

where π_{IG} and π_N are an inverse gamma and a normal prior *pdfs*, respectively, and h is an unknown precision parameter, $h>0$. The maximum likelihood estimator of $\hat{h}=\frac{\sigma^2}{s^2}$ can be obtained from the maximisation of the likelihood of π_N evaluated at $r_i=\bar{r}$. As before, a suitable estimator for σ^2 is $\hat{\sigma}_f^2$.

The joint posterior is obtained by combining the likelihood function of the joint Normal and the joint Normal-gamma prior. This results in a joint Normal-gamma *pdf* with parameters $\mu_\pi^{"}, h^{"}, s^{2"}$ and $\nu^{"}$, where

$$\mu_\pi^{"}=\frac{\hat{h}\hat{\mu}_\pi+N\bar{r}}{h^{"}},$$

$$h^{"}=\hat{h}+N,$$

$$s^{2"}=\frac{\left(\hat{\nu}\hat{\tau}+\hat{\mu}_\pi^2 h\right)+\left((N-1)s^2+N\bar{r}^2\right)-h^{"}\hat{\mu}_\pi^{2"}}{\nu^{"}}, \text{ and}$$

$$\nu^{"}=N+\hat{\nu}+1.$$

[22] Raffa and Schleifer work with the Gamma *pdf* of $h=\frac{1}{\sigma^2}$. For consistency purposes, here we work with the inverse-gamma *pdf* of σ^2. See Raffa and Schleifer, page 300.

Appendix II: Main Equations in Extended Form

- *Pdfs* of $\bar{r}$ and s:

$$f(\bar{r} \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi \frac{\sigma^2}{N}}} \text{Exp}\left(- \frac{(\bar{r} - \mu)^2}{2\frac{\sigma^2}{N}} \right) \tag{24}$$

$$f(s \mid \sigma^2, \upsilon) = 2 \frac{1}{\Gamma\left(\frac{\upsilon}{2}\right)} \left(\frac{\upsilon}{2}\right)^{\frac{\upsilon}{2}} \left(\frac{1}{\sigma^2}\right)^{\frac{\upsilon}{2}} s^{\upsilon - 1} \exp\left(- \frac{\upsilon s^2}{2\sigma^2} \right) \tag{25}$$

Prior and posterior distributions of μ and σ^2:

$$\pi_N(\mu \mid \mu_\pi, \sigma_\pi^2) = \frac{1}{\sqrt{2\pi\hat{\sigma}_\pi^2}} \text{Exp}\left(- \frac{(\mu - \hat{\mu}_\pi)^2}{2\hat{\sigma}_\pi^2} \right), \tag{26}$$

where $\pi_N(.)$ denotes a prior Normal distribution.

$$\pi_{IG}(\sigma^2 \mid \tau, \nu) = \frac{\left(\frac{\nu\tau}{2}\right)^{\frac{\nu}{2}} \text{Exp}\left(- \frac{\nu\tau}{2\sigma^2} \right)}{\Gamma\left(\frac{\nu}{2}\right) \sigma^{2(\frac{\nu}{2}+1)}}, \tag{27}$$

where $\pi_{IG}(.)$ denotes an inverted gamma prior distribution.

- Confidence regions

The confidence regions (9) and (20)-(22) above take the functional forms presented below once the estimated prior and posterior *pdfs* are substituted into each equation. Closed form solutions are obtained for equations (9) and (20), both of which are readily simulated using the appropriate software. Simulations of equations (21) and (22) involve greater complication, as it is not possible to arrive to a close form solution due to the presence of μ. These equations have first to be integrated numerically, which compromises the precision of the results.

Equation (20): $-b\pi_{IG}(\sigma/\hat{\nu},\hat{\tau})-\pi_{N}(\mu/\hat{\mu}_{\pi},\hat{\sigma}_{\pi}^{2})$

$$= -b\left[2\frac{\left(\frac{\hat{\nu}\hat{\tau}}{2}\right)^{\frac{\hat{\nu}}{2}}e^{-\frac{\hat{\nu}\hat{\tau}}{2\sigma^{2}}}}{\sigma^{(\frac{\hat{\nu}+1}{2})}\Gamma\left(\frac{\hat{\nu}}{2}\right)}\right]-\left(\frac{1}{2\Pi\sigma_{\pi}^{2}}\right)^{\frac{1}{2}}e^{-\frac{1}{2}\frac{(\mu-\mu_{\pi})^{2}}{\sigma_{\pi}^{2}}}$$

Equation (9): $-b\Pi_{IG}(\sigma/\nu'',s^{2''})-\Pi_{N}\left(\mu/\frac{\bar{r}\hat{\sigma}_{f}^{2}+\hat{\mu}_{\pi}\frac{\hat{\sigma}_{f}^{2}}{N}}{\hat{\sigma}_{\pi}^{2}+\frac{\hat{\sigma}_{f}^{2}}{N}},\frac{\hat{\sigma}_{\pi}^{2}\frac{\hat{\sigma}_{f}^{2}}{N}}{\hat{\sigma}_{\pi}^{2}+\frac{\hat{\sigma}_{f}^{2}}{N}}\right)$

$$= -b\left[2\frac{\left(\frac{\nu''s^{2''}}{2}\right)^{\frac{\nu''}{2}}e^{-\frac{\nu''s^{2''}}{2\sigma^{2}}}}{\sigma^{(\frac{\nu''+1}{2})}\Gamma\left(\frac{\nu''}{2}\right)}\right]-\left(\frac{1}{2\Pi\frac{\hat{\sigma}_{\pi}^{2}+\frac{\hat{\sigma}_{f}^{2}}{N}}{\hat{\sigma}_{\pi}^{2}\frac{\hat{\sigma}_{f}^{2}}{N}}}\right)^{\frac{1}{2}}e^{-\frac{1}{2}\frac{\left(\mu-\frac{\bar{r}\hat{\sigma}_{f}^{2}+\hat{\mu}_{\pi}\frac{\hat{\sigma}_{f}^{2}}{N}}{\hat{\sigma}_{\pi}^{2}+\frac{\hat{\sigma}_{f}^{2}}{N}}\right)^{2}}{\frac{\hat{\sigma}_{\pi}^{2}\frac{\hat{\sigma}_{f}^{2}}{N}}{\hat{\sigma}_{\pi}^{2}+\frac{\hat{\sigma}_{f}^{2}}{N}}}}$$

Equation (21): $-b\underset{\pi(\sigma)}{E}\left[f(s/\sigma)\right]-\underset{\pi(\mu,\sigma^{2})}{E}\left[f\left(\bar{r}/\mu,\sigma^{2}\right)\right]$

$$=-b\left[\int_{\sigma}f(s/\sigma)\pi_{IG}(\sigma/\hat{\nu},\hat{\tau})d\sigma\right]-\iint_{\mu\sigma}f\left(\bar{r}/\mu,\sigma^{2}\right)\pi_{NG}\left(\mu,\sigma^{2}/\hat{\mu}_{\pi},\hat{h},\hat{\nu},\hat{\tau}\right)d\sigma d\mu$$

$$= -b\left[2\left(\frac{u}{2}\right)^{\frac{u}{2}}\left(\frac{\hat{\nu}\hat{\tau}}{2}\right)^{\frac{\hat{\nu}}{2}}\frac{\Gamma\left(\frac{u+\hat{\nu}}{2}\right)}{\Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{\hat{\nu}}{2}\right)}s^{u-1}\left(\frac{us^{2}+\hat{\nu}\hat{\tau}}{2}\right)^{-\left(\frac{u+\hat{\nu}}{2}\right)}\right]-$$

$$\left[\left(\frac{1}{2\Pi}\right)\left(\frac{\hat{\nu}\hat{\tau}}{2}\right)^{\frac{\hat{\nu}}{2}}\frac{\Gamma\left(\frac{\hat{\nu}+2}{2}\right)}{\Gamma\left(\frac{\hat{\nu}}{2}\right)}\left(N\hat{h}\right)^{\frac{1}{2}}\int_{\mu}\left[\frac{\left(N\bar{r}+\hat{h}\hat{\mu}_{\pi}^{2}+\hat{\nu}\hat{\tau}\right)}{2}-\frac{\left(N\bar{r}+\hat{h}\hat{\mu}_{\pi}\right)^{2}}{2(N+\hat{h})}+\frac{(N+\hat{h})}{2}\left(\mu-\frac{N\bar{r}+\hat{h}\hat{\mu}_{\pi}}{N+\hat{h}}\right)^{2}\right]^{-\frac{\hat{\nu}+2}{2}}d\mu\right]$$

Equation (22): $-b\underset{\Pi(\sigma)}{E}\left[f(s/\sigma)\right]-\underset{\Pi(\mu,\sigma^2)}{E}\left[f(\bar{r}/\mu,\sigma^2)\right]$

$$=-b\left[\int_{\sigma} f(s/\sigma)\Pi_{IG}(\sigma/\nu'',s^{2''})d\sigma\right]-\iint_{\mu\sigma} f(\bar{r}/\mu,\sigma^2)\Pi_{NG}(\mu,\sigma^2/\mu''_{\pi},h'',\nu'',s^{2''})d\sigma d\mu$$

$$=-b\left[2\left(\frac{u}{2}\right)^{\frac{u}{2}}\left(\frac{\hat{\nu}''s^{2''}}{2}\right)^{\frac{\hat{\nu}''}{2}}\frac{\Gamma\left(\frac{u+\hat{\nu}''}{2}\right)}{\Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{\hat{\nu}''}{2}\right)}s^{u-1}\left(\frac{us^2+\hat{\nu}''s^{2''}}{2}\right)^{-\left(\frac{u+\hat{\nu}''}{2}\right)}\right]-$$

$$\left[\left(\frac{1}{2\Pi}\right)\left(\frac{\hat{\nu}''s^{2''}}{2}\right)^{\frac{\hat{\nu}''}{2}}\frac{\Gamma\left(\frac{\hat{\nu}''+2}{2}\right)}{\Gamma\left(\frac{\hat{\nu}''}{2}\right)}\left(N\hat{h}''\right)^{\frac{1}{2}}\int_{\mu}\left[\frac{\left(N\bar{r}+\hat{h}''\mu''^{2}_{\pi}+\hat{\nu}''s^{2''}\right)}{2}-\frac{\left(N\bar{r}+\hat{h}''\hat{\mu}''_{\pi}\right)^2}{2(N+\hat{h}'')}+\frac{(N+\hat{h}'')}{2}\left(\mu-\frac{N\bar{r}+\hat{h}''\hat{\mu}''_{\pi}}{N+\hat{h}''}\right)^2\right]^{-\frac{\hat{\nu}''+2}{2}}d\mu\right]$$