What Will the Likely Range of My Wealth Be?

Raymond Kan and Guofu Zhou*

February, 2009

*Kan is Associate Professor of Finance and National Bank Financial Professor in Capital Market at the University of Toronto, 105 St. George Street Toronto, Ontario Canada M5S 3E6; and Zhou is Professor of Finance at Washington University in St. Louis, St. Louis, MO 63130.

Corresponding Author: Guofu Zhou, Olin School of Business, Washington University, One Brookings Dr., St. Louis, MO 63130. Phone: (314) 935-6384 and e-mail: zhou@wustl.edu.

What Will the Likely Range of My Wealth Be?

The median is often a better measure than the mean in evaluating the long-term value of a portfolio. However, the standard plug-in estimate of the median is too optimistic. It has a substantial upward bias that can easily exceed a factor of two. In this paper, we provide an unbiased forecast of the median of the long-term value of a portfolio. In addition, we also provide an unbiased forecast of an arbitrary percentile of the long-term portfolio value distribution. This allows us to construct the likely range of the long-term portfolio value for any given confidence level. Finally, we provide an unbiased forecast of the probability for the long-term portfolio value falling into a given interval. Our unbiased estimators provide a more accurate assessment of the long-term value of a portfolio than the traditional estimators, and are useful for long-term planning and investment.

Forecasting long-term portfolio return is of great interest to investors and fund managers. However, this is not an easy task. In their thought-provoking paper, Jacquier, Kane, and Marcus (2003) show that when estimating the expected terminal value of a portfolio, both geometric mean and arithmetic mean returns are substantially biased and they provide an approach to correct the bias. Hughson, Stutzer, and Yung (2006), however, point out that the median is of greater interest since the mean is too optimistic compared with the median.[1] Both of these studies provide valuable insights into long-term planning and investment.

In this paper, we address the question of the likely range of the long-term portfolio value. This question seems of even greater interest to investors. For example, in 40 years, what are the 5th and 95th percentiles of the value of my portfolio? The dual question is that, what is the probability that my portfolio value will be, say, between \$1 million and \$5 million when I retire in 40 years? The answers to these questions clearly depend on our estimates of the model parameters. Using the standard lognormal model of long-term wealth that Jacquier, Kane, and Marcus (2003) used, we find that the usual sample estimates of both the range and the probability falling into a given range are significantly biased. Drawing on various contributions in the statistical literature, we provide the minimum-variance unbiased estimators for both of them. These estimators are useful for investors to evaluate their likely ranges of long-term portfolio value as well as to estimate the probabilities for their long-term portfolio value to be in a desired interval.

The Model and Existing Studies

Let $\tilde{R}_t$ be the return on a portfolio between time $t-1$ and time t (a tilde indicates that the return is a random variable). For simplicity, we make the standard lognormal assumption that the return $\tilde{R}_t$ is independently and identically lognormal, which implies that the continuously compounded return $\tilde{r}_t = \ln(1+\tilde{R}_t)$ is normally distributed with mean μ and variance σ^2. Equivalently, we can write

$$\tilde{r}_t = \mu + \tilde{\epsilon}_t, \tag{1}$$

where $\tilde{\epsilon}_t$ is normally distributed, has mean zero and variance σ^2. This is a popular model for asset allocation and is also the model underlying the famous Black-Scholes formula. Jacquier, Kane,

[1] Their paper won both the FAJ Scroll Award and the Readers Choice Award.

and Marcus (2003) also use this lognormal model for asset returns when estimating the expected long-term portfolio value. Bodie, Kane, and Marcus (2005), in Chapter 24 of their popular text, discuss both the model and the estimation issues in detail. Jacquier, Kane, and Marcus (2005) show that different estimation methods can make an important difference in asset allocation decisions. Technically, the assumption that the return is independently and identically distributed is for simplicity. However, Jacquier, Kane, and Marcus (2003, 2005) perform a sensitivity analysis and show that in the context of the estimation of long-term wealth, the results are robust to possible autocorrelations in asset returns.

Assume today is time T. The terminal wealth of investing \$1 for H periods is given by

$$\tilde{W} = \prod_{t=1}^{H}(1+\tilde{R}_{T+t}) = e^{H\mu+\sum_{t=1}^{H}\tilde{\epsilon}_{T+t}}, \tag{2}$$

where $\tilde{\epsilon}_t \sim N(0,\sigma^2)$ is the disturbance of the return at time t. Since $\tilde{W}$ is a random variable, its value is unknown today. The expected value of $\tilde{W}$ is of practical interest and is easily obtained as

$$W_e = E[\tilde{W}] = e^{H\left(\mu+\frac{\sigma^2}{2}\right)} \tag{3}$$

if the true value of μ and σ^2 is known. However, μ and σ^2 are generally unknown and have to be estimated from data.

Let $\hat{\mu}$ and $\hat{\sigma}^2$ be the unbiased estimators of μ and σ^2 based on the historical data with sample size T,

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_t, \qquad \hat{\sigma}^2 = \frac{1}{T-1}\sum_{t=1}^{T}(r_t-\hat{\mu})^2, \tag{4}$$

which are the outputs of almost any statistical software. Although $\hat{\mu}$ and $\hat{\sigma}^2$ are unbiased, a nonlinear function of $\hat{\mu}$ and $\hat{\sigma}^2$ is generally not an unbiased estimator of the corresponding function of μ and σ^2. Indeed, the standard plug-in estimator of W_e based on $\hat{\mu}$ and $\hat{\sigma}^2$ is

$$\hat{W}_e = e^{H\left(\hat{\mu}+\frac{\hat{\sigma}^2}{2}\right)}, \tag{5}$$

and its expected value is given by

$$E[\hat{W}_e] = E\left[e^{H\left(\hat{\mu}+\frac{\hat{\sigma}^2}{2}\right)}\right] \neq e^{H\left(\mu+\frac{\sigma^2}{2}\right)} = W_e. \tag{6}$$

So $\hat{W}_e$ is a biased estimator of W_e. As shown by Jacquier, Kane, and Marcus (2003), the longer the investment horizon H, the greater the bias. Because of the substantial bias, they provide unbiased estimators to resolve the problem. As we will show below, a similar problem also exists for estimating the quantiles and probabilities.

Unbiased Estimation of the Median

We are interested in estimating different quantiles of the terminal wealth $\tilde{W}$. We first consider a particularly interesting quantile of $\tilde{W}$ — the median (50th percentile), which is given by

$$W_m = e^{H\mu}. \tag{7}$$

Comparing this with (3), the expected terminal wealth is equal to the median terminal wealth multiplied by a scalar $e^{\frac{H\sigma^2}{2}}$. For reasonable values of σ, this scalar can easily be greater than two with a modest H. For example, with an annual standard deviation of $\sigma = 20\%$ and $H = 40$ years, we have the mean of the terminal wealth to be 2.23 times as large as its median. This says that the terminal wealth is distributed with a large positive skewness, and the mean is located at the extreme right tail of the distribution. As H increases, the probability that the terminal wealth will be greater than its expected value approaches zero. Theoretically, the mean is optimal for quadratic error loss function, and the median is optimal for absolute error loss function. Given the shape of the distribution of terminal loss function, it is difficult to argue for the use of a quadratic error loss function. This leads to Hughson, Stutzer, and Yung's (2006) insightful suggestion that median rather than expected terminal wealth may be more appropriate for assessing long-term wealth outcomes.

Given the median is of interest now, the question is how to estimate it in practice. This is because μ is unknown in the real world, and have to be estimated by the available data. The standard approach of estimating W_m is to use

$$\hat{W}_m = e^{H\hat{\mu}}, \tag{8}$$

where the unknown μ is replaced by its estimate $\hat{\mu}$. This is the usual plug-in estimate of the median. However, from the normality assumption on $\hat{\mu}$, we can show that the expected value of the estimate is

$$E[\hat{W}_m] = e^{H\mu + \frac{H^2\sigma^2}{2T}}, \tag{9}$$

which is different from the true median W_m by having the second term. As a result, the ratio of this expected value to the true median is

$$E[\hat{W}_m]/W_m = e^{\frac{H^2\sigma^2}{2T}}. \tag{10}$$

Hence, the plug-in median estimate is always biased upward, and the bias grows exponentially with a ratio of $H^2\sigma^2/(2T)$. For any given sample size, the greater the investment horizon or the volatility, the greater the bias.

Suppose we use $T = 30$ years of annual data to estimate the median of terminal wealth, with the true annual standard deviation of $\sigma = 20\%$. Assume the investment horizon is $H = 40$ years. Then $e^{H^2\sigma^2/(2T)} = 2.906$. This means that, on average, the estimated median can be almost three times as large as the true value! In comparison with Jacquier, Kane, and Marcus's (2003) finding for estimated expected terminal wealth, the bias in estimating the median is of the same magnitude.

To see how the result varies under alternative parametric assumptions, Table 1 provides the ratio of $E[\hat{W}_m]/W_m$ for varying values of H and σ following Jacquier, Kane, and Marcus (2003). With a sample period of 30 years, the biases are huge, varying from 1.822 to 11.023 for an investment horizon of 40 years as volatility goes from 15% to 30%. Only when the investment horizon is 10 years or less that we can ignore this bias. With a longer sample period of 75 years, the biases are smaller. Nevertheless, at a horizon of 40 years, the biases are still significant, varying from 1.271 to 2.612.

Note that the results in Table 1 only depend on H, T and σ, but it does not depend on how often we sample the data. This is because $\hat{W}_m$ depends only on $\hat{\mu}$ and $\hat{\mu}$ remains the same no matter how often we sample the data.

Since the bias of $\hat{W}_m$ is substantial, it is important to obtain an unbiased estimator. Intuitively, since $\hat{W}_m$ is biased significantly upwards, we need to adjust it down. As a special case of Proposition 1 below, the unbiased estimator of W_m is given by:

$$\hat{W}_m^u = e^{H\hat{\mu}} g_0(\hat{\sigma}), \tag{11}$$

where $g_0(\hat{\sigma})$ is a function of $\hat{\sigma}$ that is used to eliminate the bias, and is defined by

$$g_0(\hat{\sigma}) = \sum_{k=0}^{\infty} \frac{(d\hat{\sigma}^2)^k}{\left(\frac{T-1}{2}\right)_k k!}, \tag{12}$$

Table 1. Ratio of Expected Forecasted Median to True Median

	Horizon (H years)			
σ	10	20	30	40
A. Sample period ($T = 30$ years)				
15%	1.038	1.162	1.401	1.822
20%	1.069	1.306	1.822	2.906
25%	1.110	1.517	2.554	5.294
30%	1.162	1.822	3.857	11.023
B. Sample period ($T = 75$ years)				
15%	1.015	1.062	1.145	1.271
20%	1.027	1.113	1.271	1.532
25%	1.043	1.181	1.455	1.948
30%	1.062	1.271	1.716	2.612

where $d = -H^2(T-1)/4T$ and $(a)_k = a(a+1)\cdots(a+k-1)$ is the rising factorial (or Pochhammer symbol). As shown in the proof of Proposition 1, $\hat{W}_m^u$ is unbiased and also has the smallest variance among all possible unbiased estimators of W_m. Note that unlike the plug-in estimate $\hat{W}_m$ which is independent of the sampling frequency of the returns, $g_0(\hat{\sigma})$ and hence $\hat{W}_m^u$ is not invariant to the sampling frequency of returns. For a given length of sample period, the $\hat{W}_m^u$ that is computed using the annual data is different from the one that is computed using the monthly data.

To illustrate the possible difference between the two estimators $\hat{W}_m$ and $\hat{W}_m^u$, we consider an investment in the U.S. equity market or one of the seven international equity markets. The second and third columns of Table 2 report the sample mean and standard deviations of the returns from different markets, based on available monthly data from 1970/1–2007/12 (456 months). The biased and unbiased estimators $\hat{W}_m$ and $\hat{W}_m^u$, of the median terminal wealth from investing \$1 for a horizon of 30 years on the different markets are given in the fourth and fifth columns. For the U.S. equity market, even with 456 months of data, the plug-in estimate of the median terminal wealth is \$23.388, about 31% higher than the unbiased estimate of \$17.838. The differences are in general greater for other international markets. The greatest difference occurs for the Italian equity market. The biased estimate is now about 101% larger than the unbiased one.[2]

Table 2. Biased and Unbiased Estimates of Median, Quantiles and Probabilities

					$H = 360$ months, $T = 456$ months					
Market	$\hat{\mu}$	$\hat{\sigma}$	$\hat{W}_m$	$\hat{W}_m^u$	$\hat{W}_{0.05}$	$\hat{W}_{0.95}$	$\hat{W}_{0.05}^u$	$\hat{W}_{0.95}^u$	$\hat{p}$	$\hat{p}_u$
U.S.	0.0088	0.0436	23.388	17.838	5.990	91.319	4.557	69.688	0.535	0.660
Australia	0.0094	0.0679	29.413	15.249	3.529	245.132	1.817	127.313	0.446	0.742
Canada	0.0085	0.0539	21.194	14.013	3.937	114.085	2.593	75.497	0.458	0.549
France	0.0097	0.0632	33.020	18.705	4.594	237.352	2.587	134.643	0.484	0.818
Germany	0.0075	0.0553	15.108	9.781	2.690	84.841	1.734	54.976	0.359	0.280
Italy	0.0063	0.0701	9.579	4.760	1.075	85.387	0.530	42.513	0.251	0.114
Japan	0.0085	0.0612	21.081	12.364	3.118	142.545	1.819	83.709	0.428	0.539
U.K.	0.0100	0.0626	36.705	21.000	5.195	259.333	2.956	148.571	0.496	0.866

Estimating Quantiles of the Terminal Wealth

Besides being interested in the median of the terminal wealth, which is a special quantile that permits simple intuition and analytical expression, an investor may also be interested in estimating other quantiles of the terminal wealth. Given a probability level p, the p-th quantile of his terminal wealth, W_p, is defined as

$$\text{Prob}(\tilde{W} < W_p) = p. \tag{13}$$

For example, the median terminal wealth, W_m, is $W_{0.5}$. Investors may be interested in estimating $W_{0.05}$ and $W_{0.95}$ because there is a 90% probability that his terminal wealth will fall in the range $[W_{0.05}, W_{0.95}]$.

Based on Equation (2),

$$\frac{\ln(\tilde{W}) - H\mu}{\sqrt{H}\sigma} \tag{14}$$

is a standard normal random variable with mean zero and variance one. Let C_p be the p-th quantile

[2] The unbiased estimator $\hat{W}_m^u$ also have smaller mean squared errors and mean absolute errors than $\hat{W}_m$. We do not address these statistical issues here for brevity. The technical results are available in a separate paper.

of a standard normal distribution, i.e.,

$$\Phi(C_p) = p, \tag{15}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable.[3] Then,

$$W_p = e^{H\mu + C_p\sqrt{H}\sigma}. \tag{16}$$

Note that W_p is a log-linear function of both μ and σ.

The standard plug-in estimator of W_p is

$$\hat{W}_p = e^{H\hat{\mu} + C_p\sqrt{H}\hat{\sigma}}. \tag{17}$$

As in the case of estimating W_m, this plug-in estimator is substantially biased. Figure 1 provides the bias ratio $E[\hat{W}_p]/W_p$ as a function of investment horizon H for $p = 0.05$, 0.5 and 0.95, respectively. Consider the top left panel with $T = 75$ years and $\sigma = 15\%$ per year. We find that the ratio increases from 1 to 1.3 as the horizon increases from one to 40 years. Although the ratio is significantly different from one, it does not seem too striking. However, when σ is 30% per year, as shown in the top right panel, the ratio is more than two when $H = 40$ years. The worse case occurs when $T = 30$ years and $\sigma = 30\%$ per year. As shown in the bottom right panel, the ratio is now too large to be ignored. It is over three for $H = 30$ years and over 12 for $H = 40$ years. Overall, the plug-in estimator has substantial biases. Another interesting result is that the ratio of $E[\hat{W}_p]/W_p$ for either $p = 0.05$, 0.5 or 0.95 are virtually the same, despite asymmetry of the wealth distribution.

Because of the substantial bias of the standard plug-in estimator, it is important to obtain an unbiased estimator. Shimizu and Iwase (1981) obtained an unbiased estimator of the quantile of a lognormal distribution and their result allows us to obtain an unbiased estimator of W_p for $H = 1$. The following proposition generalizes the result of Shimizu and Iwase (1981) to obtain the unbiased estimator of W_p for general H.

Proposition 1 The unbiased estimator of W_p is

$$\hat{W}_p^u = e^{H\hat{\mu}} g(\hat{\sigma}), \tag{18}$$

[3] C_p can be easily computed using the Excel function for the inverse of the standard normal distribution, `NORMSINV`.

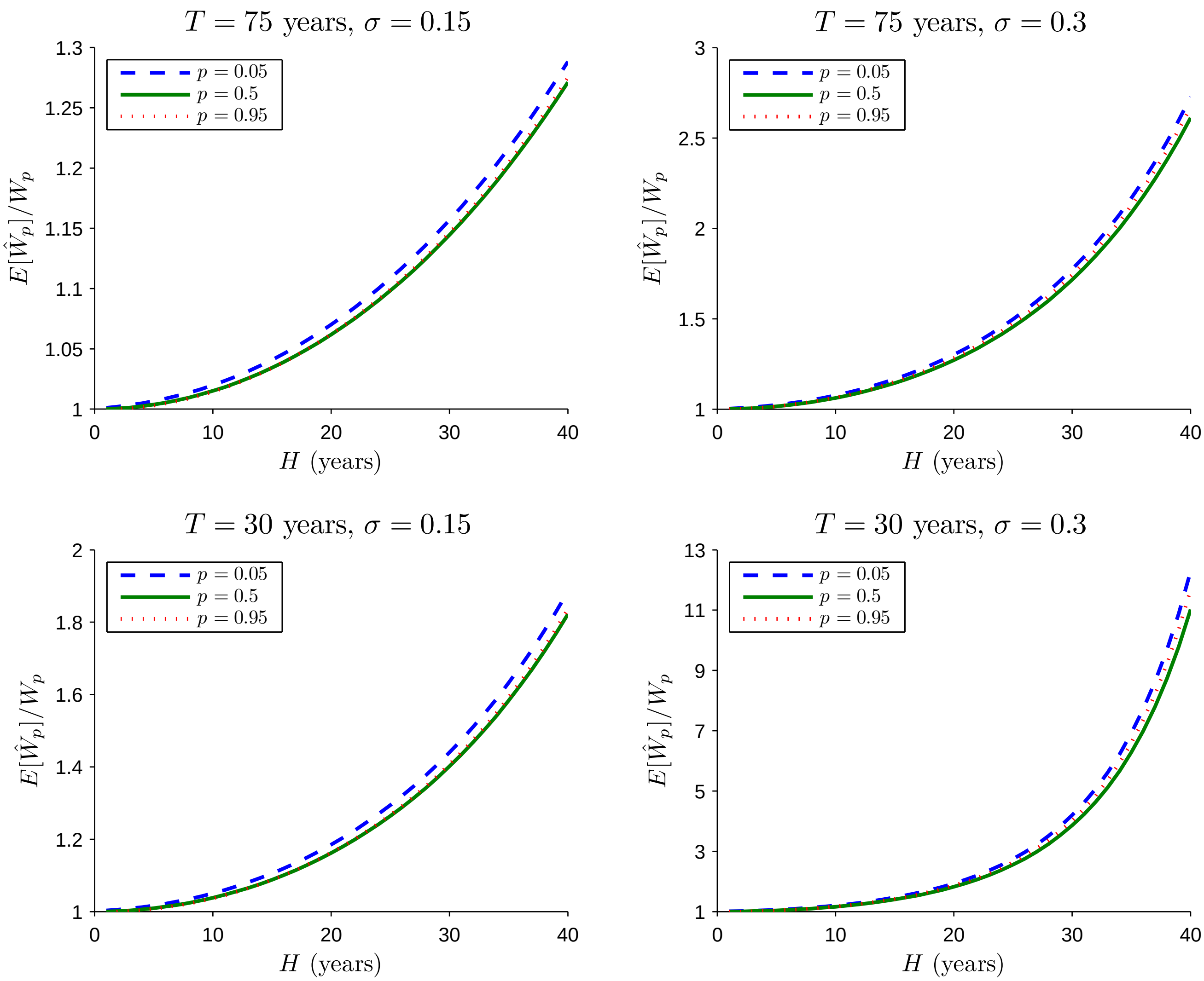

Figure 1
This figure plots the ratio of $E[\hat{W}_p]/W_p$ for $p = 0.05$, $p = 0.5$ and $p = 0.95$ as a function of investment horizon H. The upper panels are for an estimation period of $T = 75$ years. The lower panels are for an estimation period of $T = 30$ years. The left panels assume σ =15% per year and the right panels assume σ =30% per year.

where $g(\hat{\sigma})$ is the adjustment to eliminate the bias, and is defined by

$$g(\hat{\sigma}) = \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{\Gamma\left(\frac{T-1}{2}\right) c^{k-j} d^j \hat{\sigma}^{k+j}}{j!(k-j)!\Gamma\left(\frac{T-1+k+j}{2}\right)}, \tag{19}$$

where $c = C_p\sqrt{H(T-1)/2}$, $d = -H^2(T-1)/4T$, and $\Gamma(\cdot)$ is the gamma function.

The proof of Proposition 1 is given in the Appendix. Like $\hat{W}_m^u$, $\hat{W}_p^u$ also has the minimum variance among all possible unbiased estimators of W_p. Although $\hat{W}_p^u$ looks much more complex than $\hat{W}_m^u$, it can, however, be as easily computed as the latter.[4] In many cases, we can actually correct for most of the bias in $\hat{W}_p$ with a simple adjustment. This adjustment is motivated by the fact that if μ is unknown but σ is known, then the unbiased estimator of W_p is given by

$$\hat{W}_p^u(\sigma) = e^{H\hat{\mu} + C_p\sqrt{H}\sigma - \frac{H^2\sigma^2}{2T}}. \tag{20}$$

Replacing σ by $\hat{\sigma}$, we obtain a much simpler approximate unbiased estimator of W_p as

$$\hat{W}_p^{au} = e^{H\hat{\mu} + C_p\sqrt{H}\hat{\sigma} - \frac{H^2\hat{\sigma}^2}{2T}} = \hat{W}_p e^{-\frac{H^2\hat{\sigma}^2}{2T}}, \tag{21}$$

which is simply equal to the plug-in estimator multiplied by an adjustment factor of $e^{-\frac{H^2\hat{\sigma}^2}{2T}}$. Note that this adjustment factor is independent of p.

To understand the difference between the exact unbiased estimator $\hat{W}_p^u$ and the approximate unbiased estimator $\hat{W}_p^{au}$, we plot in Figure 2 the ratio of $\hat{W}_p^u/\hat{W}_p^{au}$ (which only depends on H, T, and $\hat{\sigma}$) as a function of H for $p = 0.05$, 0.5 and 0.95. In the upper panels, we consider a scenario in which we estimate σ using 75 years of data. The upper left panel assumes we use annual data (i.e., $T = 75$ years) to estimate σ. When $\hat{\sigma} = 0.3$, we can see that for $p = 0.5$ and 0.95, $\hat{W}_p^{au}$ is very close to $\hat{W}_p^u$. However, for $p = 0.05$, $\hat{W}_p^{au}$ can be up to 10% larger than $\hat{W}_p^u$ when the investment horizon is $H = 40$ years. But if we use 75 years of monthly returns (i.e., $T = 900$ months) to estimate σ, the difference between $\hat{W}_p^{au}$ and $\hat{W}_p^u$ becomes negligible. The upper right panel considers this case, and shows that with a monthly sample standard deviation of $\hat{\sigma} = 0.3/\sqrt{12} = 0.0866$, $\hat{W}_p^{au}$ is almost identical to $\hat{W}_p^u$ even for $H = 480$ months and $p = 0.05$.

However, the difference between $\hat{W}_p^{au}$ and $\hat{W}_p^u$ can be significant when the estimation period is short. In the lower panels of Figure 2, we consider a scenario in which we estimate σ using only

[4] For practical situations, the infinite series converges quickly. In all of our calculations, the infinite series converges after summing up less than 15 terms. A set of Matlab programs for computing various estimators used in this paper is available upon request.

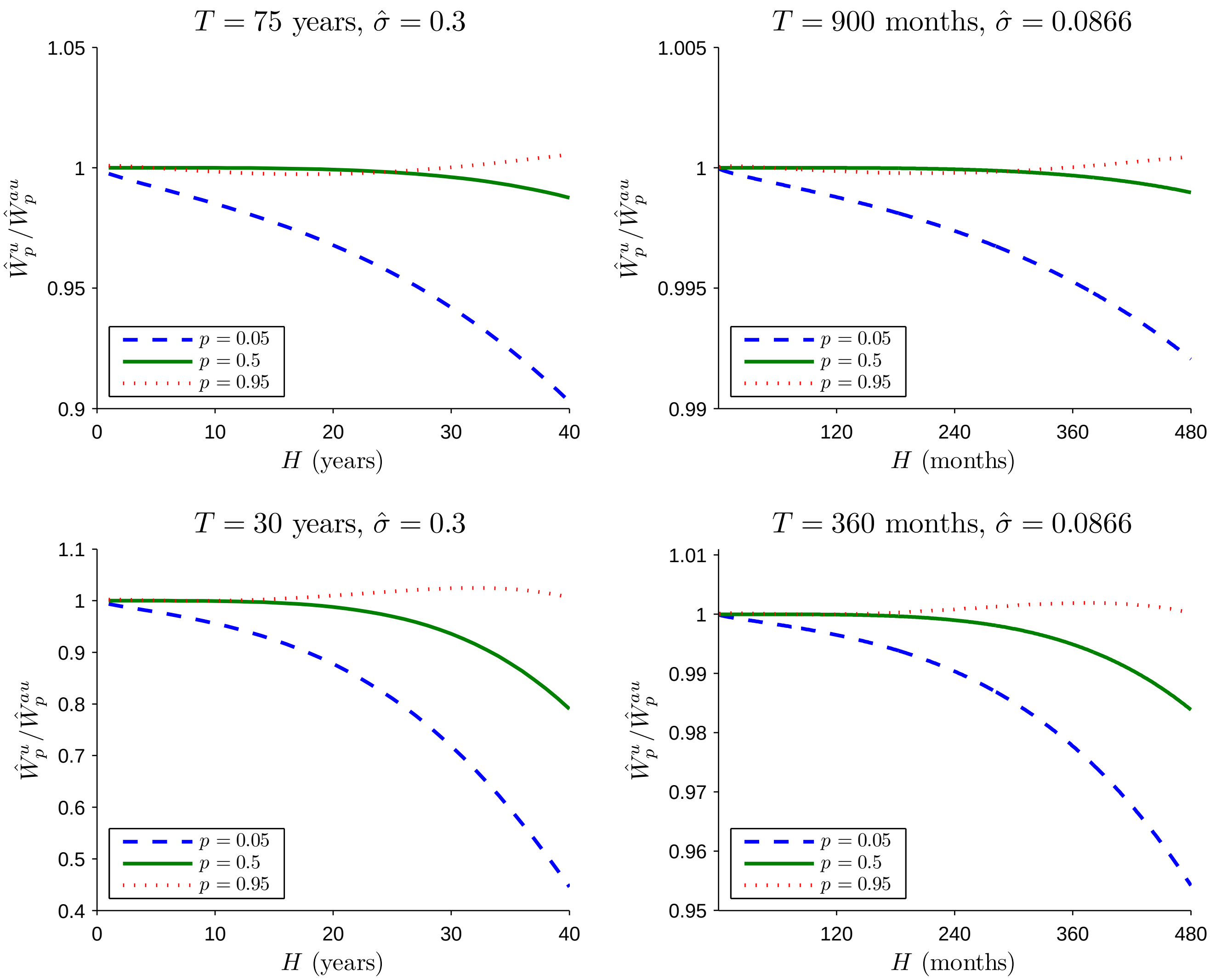

Figure 2
This figure plots the ratio of the exact unbiased estimator ($\hat{W}_p^u$) to the approximate unbiased estimator ($\hat{W}_p^{au}$) of W_p for three different values of p (0.05, 0.5, and 0.95) as a function of investment horizon (H), where W_p is the p-th percentile of the terminal wealth at the end of H periods. The upper panels assume an estimation period of 75 years whereas the lower panels assume an estimation period of 30 years. The left panels assume the use of annual data whereas the right panels assume the use of monthly data. The ratios are obtained by assuming the annual sample standard deviation is 30% or the monthly sample standard deviation is 8.66% (i.e., $0.3/\sqrt{12}$).

30 years of data. When we use annual data to estimate σ, the lower left panel shows that $\hat{W}_p^{au}$ and $\hat{W}_p^u$ can be quite different, especially for long investment horizon and $p = 0.5$ or $p = 0.99$. Even when monthly data are used to estimate σ, the lower right panel shows that $\hat{W}_{0.05}^u$ can still be 5% smaller than $\hat{W}_{0.05}^{au}$ for $H = 480$ months. In summary, Figure 2 suggests that instead of using the more complicated unbiased estimator $\hat{W}_p^u$, we can safely use the simpler approximate unbiased estimator $\hat{W}_p^{au}$ when p is large or when the sample size (T) is relatively large.

The eighth and ninth columns of Table 2 provide $\hat{W}_{0.05}^u$ and $\hat{W}_{0.95}^u$, respectively. These are the unbiased estimates of the levels that the terminal wealth will not exceed with probabilities 5% and 95%, respectively. Hence, $[\hat{W}_{0.05}^u, \hat{W}_{0.95}^u]$ is an unbiased forecast of a range that the terminal wealth will have a 90% probability of falling within. For comparison, Table 2 also reports the plug-in estimates of $W_{0.05}$ and $W_{0.95}$ in the sixth and seventh columns. At both the 5% and 95% quantiles, the usual plug-in estimates are biased upwards. The biases are substantial, and are in general of the same magnitude as the biases observed when estimating the median. For example, for the Italian equity market, $\hat{W}_{0.95}$ is twice as large as $\hat{W}_{0.95}^u$, so relying on the plug-in method can lead to overly optimistic forecast of the terminal wealth.

Estimating the Probability of the Terminal Wealth Falling within a Given Range

Instead of estimating the quantiles of the terminal wealth, an investor may instead want to estimate the probability that his terminal wealth will be below a certain level. That is, for a given value of W, he is interested in estimating

$$p = \text{Prob}(\tilde{W} < W). \tag{22}$$

If μ and σ are known, it is an easy matter to compute p as

$$p = \Phi\left(\frac{w - H\mu}{\sqrt{H}\sigma}\right), \tag{23}$$

where $w = \ln(W)$.

Similarly, if an investor would like to estimate the probability that the terminal wealth will be within a given interval $[W_1, W_2]$ at time H, he then just needs to compute

$$\text{Prob}(W_1 < \tilde{W} < W_2) = \Phi\left(\frac{w_2 - H\mu}{\sqrt{H}\sigma}\right) - \Phi\left(\frac{w_1 - H\mu}{\sqrt{H}\sigma}\right), \tag{24}$$

where $w_1 = \ln(W_1)$, $w_2 = \ln(W_2)$.

When μ and σ^2 are unknown, we need to estimate p. The standard approach plugs in the estimates of μ and σ^2 to obtain

$$\hat{p} = \Phi\left(\frac{w - H\hat{\mu}}{\sqrt{H}\hat{\sigma}}\right) \tag{25}$$

To examine the bias of $\hat{p}$, Figure 3 plots the difference of $E[\hat{p}] - p$ as a function of investment horizon H.[5] When $p = 0.05$, the two graphs on the left show that $\hat{p}$ is biased upward. The bias can be as large as 200% when the sample size (T) is 30. When $p = 0.95$, while the biases are of the same magnitude, they are biased downward, as shown by the two graphs on the right. This is different from estimating the quantiles where the bias is fairly symmetric around the median. It should be pointed out the graphs depend only on p and T, and not on the choices of μ and σ.

The unbiased estimator of p for $H = 1$ is solved in the statistics literature by Kolmogorov (1950) (see also Barton (1961), Ellison (1964), and Johnson, Kotz, and Balakrishnan (1994, p.141)). Here we solve the problem for $H < T$.[6]

Proposition 2 When $H < T$, the unbiased estimator of p is

$$\hat{p}_u = I_a\left(\frac{T-2}{2}, \frac{T-2}{2}\right), \tag{28}$$

where

$$a = \frac{1}{2} + \frac{\sqrt{T}(w - H\hat{\mu})}{2\sqrt{H(T-H)(T-1)}\hat{\sigma}}, \tag{29}$$

and $I_x(m, n)$ is the incomplete beta function ratio, and we use the convention that $I_x(m, n) = 0$ for $x < 0$ and $I_x(m, n) = 1$ for $x > 1$.

The last column of Table 2 reports $\hat{p}_u$, the unbiased estimate of the probability that the terminal wealth in 30 years will be between \$20 and \$100 for an initial investment of \$1. For comparison, we

[5] It can be shown that

$$E[\hat{p}] = \int_0^\infty \Phi\left(\frac{C_p}{\left(\frac{u}{T-1} + \frac{H}{T}\right)^{\frac{1}{2}}}\right) f(u)\mathrm{d}u, \tag{26}$$

where $f(u)$ is the density function of a chi-squared random variable with $T-1$ degrees of freedom.

[6] When $H = T$, the unbiased estimator of p is given by

$$\hat{p}_u = \begin{cases} 1 & \text{if } H\hat{\mu} \le w, \\ 0 & \text{if } H\hat{\mu} > w. \end{cases} \tag{27}$$

For $H > T$, we are unable to obtain an unbiased estimator for p, so we leave this problem for future research.

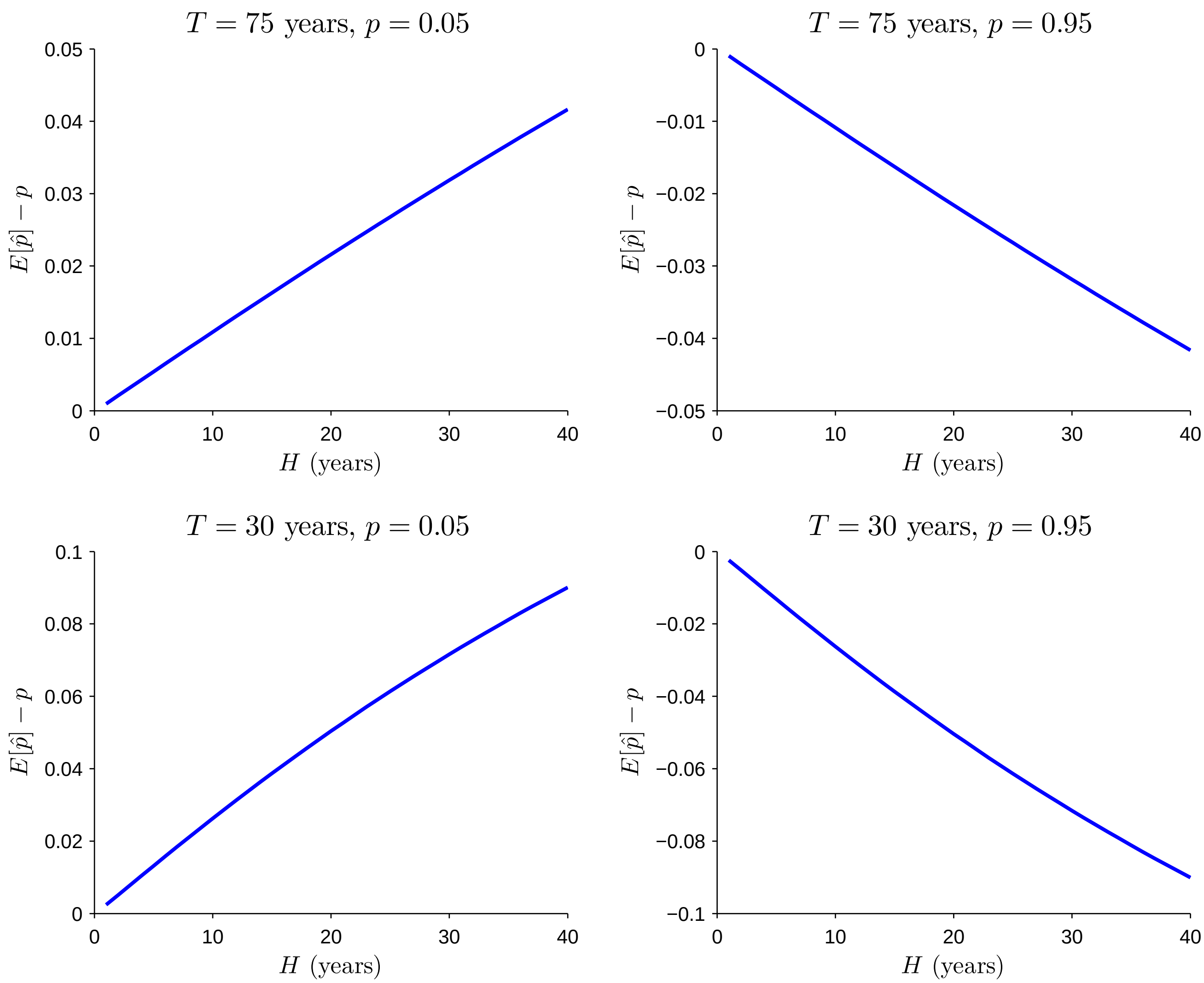

Figure 3
This figure plots the bias of $\hat{p}$, i.e., $E[\hat{p}] - p$, as a function of investment horizon H, where $p = P[\tilde{W} < W_p]$ with W_p as the p-th percentile of the terminal wealth at the end of H years, and $\hat{p}$ is the sample estimator of p. The upper panels are for an estimation period of $T = 75$ years. The lower panels are for an estimation period of $T = 30$ years. The left panels are for $p = 0.05$ and the right panels are for $p = 0.95$.

also report $\hat{p}$, the standard plug-in estimate of probability in the second last column. The difference between $\hat{p}$ and $\hat{p}_u$ can be substantial. For example, if the U.K. equity market is the investment asset, the unbiased estimate of the probability, $\hat{p}_u$, is 86.6% but the plug-in estimate, $\hat{p}$, suggests that the probability is only 49.6%, much less than the the unbiased estimate.

Conclusion

Forecasting long-term portfolio value is of great interest to investors and fund managers. Jacquier, Kane, and Marcus (2003) show that both geometric or arithmetic estimates are substantially biased for estimating the expected terminal wealth, and they provide an approach to correct the bias. Hughson, Stutzer, and Yung (2006) point out that the median terminal wealth, rather than the expected terminal wealth, is often more useful in evaluating the long-term value of a portfolio.

However, the usual plug-in median forecast is overly optimistic. It has a substantial upward bias that can easily exceed a factor of two. In this paper, we provide unbiased forecasts of the median as well as the percentiles of the terminal wealth, and the probability of the terminal wealth falling into any given interval.

Using the U.S. equity market or one of the seven international equity markets as the investment asset, we show that, even with available 456 months of data (from 1970/1 to 2007/12), the usual estimates of the median, range and probabilities are substantially biased. In contrast, the unbiased forecasts provide more accurate assessments of the terminal wealth, and are much more useful for long-term planning and investment.

• • •

The authors wish to thank Lukasz Pomorski, Michael Stutzer, Alan White, Chu Zhang, and especially an anonymous referee for many helpful comments and discussions. Kan also gratefully acknowledges financial support from the National Bank Financial of Canada.

• • •

A Appendix A. The Proofs and Unbiased Estimator of Expected Terminal Wealth

Proof of Proposition 1: Proposition 1 is a special case of obtaining the minimum-variance unbiased estimator for

$$\theta = \exp(a\mu + b\sigma), \tag{A1}$$

where a and b are constants. Since $a\hat{\mu} \sim N(a\mu, a^2\sigma^2/T)$, we have

$$E[\exp(a\hat{\mu})] = \exp\left(a\mu + \frac{a^2\sigma^2}{2T}\right). \tag{A2}$$

To eliminate the bias caused by the second term, we need to find an explicit adjustment, which must be a function of $\hat{\sigma}$ alone. Under the normality assumption, $\hat{\sigma}^2 \sim \sigma^2\chi^2_{T-1}/(T-1)$. For $r > -n$, we have (see Johnson, Kotz, Balakrishnan (1994, Eq.18.13))

$$E[(\chi^2_n)^{\frac{r}{2}}] = \frac{2^{\frac{r}{2}}\Gamma\left(\frac{n+r}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}. \tag{A3}$$

Hence, the r-th moments of $\hat{\sigma}$ is given by

$$E[\hat{\sigma}^r] = \left(\frac{2}{T-1}\right)^{\frac{r}{2}} \frac{\Gamma\left(\frac{T-1+r}{2}\right)}{\Gamma\left(\frac{T-1}{2}\right)}\sigma^r. \tag{A4}$$

Now we define a function

$$g(x) = \sum_{r=0}^{\infty}\sum_{j=0}^{r} \frac{\Gamma\left(\frac{T-1}{2}\right)c^{r-j}d^j x^{r+j}}{j!(r-j)!\Gamma\left(\frac{T-1+r+j}{2}\right)}, \tag{A5}$$

where c and d are constant scalars. The expectation of this function at $\hat{\sigma}$ is given by

$$\begin{aligned} E[g(\hat{\sigma})] &= \sum_{r=0}^{\infty}\sum_{j=0}^{r} \frac{\Gamma\left(\frac{T-1}{2}\right)c^{r-j}d^j E[\hat{\sigma}^{r+j}]}{j!(r-j)!\Gamma\left(\frac{T-1+r+j}{2}\right)} \\ &= \sum_{r=0}^{\infty}\sum_{j=0}^{r} \left(\frac{2}{T-1}\right)^{\frac{r+j}{2}} \frac{c^{r-j}d^j\sigma^{r+j}}{j!(r-j)!} \\ &= \exp\left(\sqrt{\frac{2}{T-1}}c\sigma + \left(\frac{2}{T-1}\right)d\sigma^2\right). \end{aligned} \tag{A6}$$

Therefore, if we set

$$c = \sqrt{\frac{T-1}{2}}b, \qquad d = -\frac{a^2(T-1)}{4T}, \tag{A7}$$

we have

$$E[g(\hat{\sigma})] = \exp\left(b\sigma - \frac{a^2\sigma^2}{2T}\right). \tag{A8}$$

Then using the fact that $\hat{\mu}$ and $\hat{\sigma}$ are independent of each other, we obtain

$$E[\exp(a\hat{\mu})g(\hat{\sigma})] = E[\exp(a\hat{\mu})]E[g(\hat{\sigma})] = \exp(a\mu + b\sigma). \tag{A9}$$

This completes the proof of Proposition 1.

Proof of Proposition 2: To prove Proposition 2, we make use of Theorem 1 of Ellison (1964), which shows that if $Y \sim \sqrt{\chi^2_{T-1}}$, $U \sim \beta\left(\frac{T-2}{2}, \frac{T-2}{2}\right)$, and they are independent of each other, then $Y(2U-1) \sim N(0,1)$. Based on this theorem, we can establish that, when U is independent of $\hat{\mu}$ and $\hat{\sigma}^2$,

$$\tilde{Z} = H\hat{\mu} + \sqrt{\frac{H(T-H)(T-1)}{T}}\hat{\sigma}(2U-1) \sim N(H\mu, H\sigma^2). \tag{A10}$$

This is because the first term is $H\hat{\mu} \sim N(H\mu, H^2\sigma^2/T)$ and the second term is

$$\sqrt{\frac{H(T-H)}{T}}\sigma Y(2U-1) \sim N\left(0, \frac{H(T-H)\sigma^2}{T}\right), \tag{A11}$$

with $Y = \sqrt{T-1}\hat{\sigma}/\sigma \sim \sqrt{\chi^2_{T-1}}$. Then, using the independence between the first and the second term, we have $\tilde{Z} \sim N(H\mu, H\sigma^2)$.

Note that $p = P[\tilde{Z} < w]$, so we have

$$p = E[P[\tilde{Z} < w|\hat{\mu}, \hat{\sigma}^2]] = E[P[U < a|\hat{\mu}, \hat{\sigma}^2]] = E[\hat{p}_u], \tag{A12}$$

where the second equality follows because

$$\begin{aligned} \tilde{Z} < w \quad &\Rightarrow \quad H\hat{\mu} + \sqrt{\frac{H(T-H)(T-1)}{T}}\hat{\sigma}(2U-1) < w \\ &\Rightarrow \quad 2U - 1 < \frac{\sqrt{T}(w - H\hat{\mu})}{\sqrt{H(T-H)(T-1)}\hat{\sigma}} \\ &\Rightarrow \quad U < \frac{1}{2} + \frac{\sqrt{T}(w - H\hat{\mu})}{2\sqrt{H(T-H)(T-1)}\hat{\sigma}}. \end{aligned} \tag{A13}$$

This completes the proof of Proposition 2.

Finally, since the estimation of the expected terminal wealth is of great interest, we provide the explicit formula for computing the minimum-variance unbiased estimator of $E[\tilde{W}]$, to complement the earlier studies by Jacquier, Kane, and Marcus (2003, 2006). Jacquier (2006) further analyzes the impact of using estimated σ. While they use the large sample distribution of $\hat{\sigma}$, we use its exact distribution. Bradu and Mundlak (1970) seems the first to provide such an estimator. Our presentation below follows from Shimizu and Iwase (1981).

The estimation of $E[\tilde{W}]$ is a special case of estimating a general function,

$$\theta = \exp(a\mu + b\sigma^2) \tag{A14}$$

with $a = H$ and $b = H/2$. For these choices of a and b, the unbiased estimator of θ is

$$\hat{\theta}_u = e^{H\hat{\mu}}{}_0F_1\left(\frac{T-1}{2}; \frac{(T-1)(T-H)H}{4T}\hat{\sigma}^2\right), \tag{A15}$$

where

$${}_0F_1(m, z) = \sum_{i=0}^{\infty} \frac{1}{(m)_i} \frac{z^i}{i!} \tag{A16}$$

is a generalized hypergeometric function. The proof follows from Corollary 3.1 of Shimizu and Iwase (1980).

References

Barton. D. E. 1961. "Unbiased Estimation of a Set of Probabilities." *Biometrika*, vol. 48, no. 1/2 (June):227–229.

Bodie, Zvi, Alex Kane, and Alan J. Marcus. 2005. *Investments*, New York: McGraw Hill.

Bradu, Dan, and Yair Mundlak. 1970. "Estimation in Lognormal Linear Models." *Journal of the American Statistical Association*, vol. 65, no. 329 (March):198–211.

Ellison, Bob E. 1964. "Two Theorems for Inferences about the Normal Distribution with Applications in Acceptance Sampling." *Journal of the American Statistical Association*, vol. 59, no. 305 (March):89–95.

Hughson, Eric, Michael Stutzer, and Chris Yung. 2006. "The Misuse of Expected Returns." *Financial Analysts Journal*, vol. 62, no. 6 (November/December):89–96.

Jacquier, Eric. 2006. "Estimates of Long Term Expected Returns: Economic vs. Statistical Rationales." Working paper, HEC Montreal.

Jacquier, Eric, Alex Kane, and Alan Marcus. 2003. "Geometric Mean or Arithmetic Mean: A Reconsideration." *Financial Analysts Journal*, vol. 59, no. 6 (November/December):46–53.

Jacquier, Eric, Alex Kane, and Alan Marcus. 2005. "Optimal Estimation of the Risk Premium for the Long Run and Asset Allocation: A Case of Compounded Estimation Risk." *Journal of Financial Econometrics*, vol. 3, no. 1 (winter):37–55.

Kolmogorov, Andrey Nikolaevich. 1950. "Unbiased Estimates." *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya*, vol. 14, no. 4 (August):303–326.

Johnson, Norman L., Samuel Kotz, and N. Balakrishnan. 1994. *Continuous Univariate Distributions*, vol. 1, New York: Wiley.

Shimizu, K., and T. Iwase. 1981. "Uniformly Minimum Variance Unbiased Estimation in Lognormal and Related Distributions." *Communications in Statistics–Theory and Methods*, vol. 10, no. 11 (May):1127–1147.