NYU STERN
Salomon Center for the Study of Financial Institutions

Working Paper Series DERIVATIVES Research Project

ESTIMATION ERROR IN THE ASSESSMENT OF FINANCIAL RISK EXPOSURE

Stephen Figlewski

S-DRP-03-10

Version of June 29, 2003

"Estimation Error in the Assessment of Financial Risk Exposure"

by

Stephen Figlewski*

* Professor of Finance
NYU Stern School of Business
44 West 4th Street, Suite 9-160
New York, NY 10012-1126

212-998-0712
sfiglews@stern.nyu.edu

I wish to thank James Park and Antonios Sangvinatsos for research assistance, and Eric Jacquier, Lasse Pedersen, Rob Engle, Jim Ramsey and participants at the NYU "Half-Baked" seminar for comments that have proved very helpful in the baking process.

"Estimation Error in the Assessment of Financial Risk Exposure"

Abstract

Value at Risk and similar measures of financial risk exposure require predicting the tail of an asset returns distribution. Assuming a specific form, such as the normal, for the distribution, the standard deviation (and possibly other parameters) are estimated from recent historical data and the tail cutoff value is computed. But this standard procedure ignores estimation error, which we find to be substantial even under the best of conditions. In practice, a "tail event" may represent a truly rare occurrence, or it may simply be a not-so-rare occurrence at a time when the predicted volatility underestimates the true volatility, due to sampling error. This problem gets worse the further in the tail one is trying to predict.

Using a simulation of 10,000 years of daily returns, we first examine estimation risk when volatility is an unknown constant parameter. We then consider the more realistic, but more problematical, case of volatility that drifts stochastically over time. This substantially increases estimation error, although strong mean reversion in the variance tends to dampen the effect. Non-normal fat-tailed return shocks makes overall risk assessment much worse, especially in the extreme tails, but estimation error per se does not add much beyond the effect of tail fatness. Using an exponentially weighted moving average to downweight older data hurts accuracy if volatility is constant or only slowly changing. But with more volatile variance, an optimal decay rate emerges, with better performance for the most extreme tails being achieved using a relatively greater rate of downweighting.

We first simulate non-overlapping independent samples, but in practical risk management, risk exposure is estimated day by day on a rolling basis. This produces strong autocorrelation in the estimation errors, and bunching of apparently extreme events. We find that with stochastic volatility, estimation error can increase the probabilities of multi-day events, like three 1% tail events in a row, by several orders of magnitude. Finally, we report empirical results using 40 years of daily S&P 500 returns which confirm that the issues we have examined in simulations are also present in the real world.

1. Introduction

Formal procedures for assessing exposure to financial risk, such as value at Risk (VaR), have become standard in recent years.[1] Regulatory authorities are using them for setting capital requirements for banks and much research has been done to examine the performance of different methods.[2,3]

In all cases, the effort involves predicting the size and/or frequency of low probability events--the tails of the probability distributions governing asset returns. Statistical techniques are employed to estimate the relevant parameters of the underlying distribution. A standard VaR calculation, for example, begins with the assumption that over a short horizon this distribution is normal (lognormal for end of period asset values). Recent historical data is used to estimate the volatility, and the mean is either estimated or, more typically, constrained to 0.[4] The estimated parameter values are then plugged into a standard calculation, $\mathrm{VaR} = \alpha_c\hat{\sigma} + \hat{\mu}$, where α_c denotes the 5% or 1% critical value of the normal distribution (-1.645 or -2.326) and $\hat{\sigma}$ and $\hat{\mu}$ are the sample volatility and mean ($\hat{\mu} = 0$, if the mean is not calculated).[5]

But the sample parameter values are only statistical estimates of the true parameters. The standard procedure takes no account of the estimation error in these figures. The 5% cutoff for the tail of a normal distribution is at -1.645 standard deviations, but when the sample volatility is an underestimate of the true volatility, there is more than 5% probability that the next observation will fall in the predicted 5% tail. Moreover, the effect of sampling error increases the further into the tail one is trying to forecast. In fact, under the assumption of normally distributed returns with constant mean and variance, the distribution of the next period return using estimated parameters is not normal at all,

---

[1] Among many references, see: Hull [2002] ch. 16, for a textbook discussion; Jorion [1997]; or Risk Publications [1996]. Schachter's website gives an extensive bibliography on the subject.

[2] See Basle Committee on Bank Supervision [1996], for example.

[3] See, for example, Duffie and Pan [1997] or Kupiec [1995]. The Journal of Banking and Finance devoted its entire July 2002 issue to VaR-related research.

[4] In classical statistics, the best estimator for the mean return is the sample average, but the sampling error on this estimate is surprisingly large. If an asset price follows a lognormal diffusion with mean μ and volatility σ, the sample average from a sample spanning T years has expected value μ and standard deviation $\sigma T^{-1/2}$. With a sample size like those typically used in VaR-type calculations, it is common for the sample average to make no sense economically: It may be negative, which is inconsistent with market equilibrium for normal assets, or outlandishly large. For example, with annual volatility of 20% and a 3-month returns sample, the standard deviation of the sample average as an estimate of the true mean is $0.20 / (1/4)^{1/2} = 40\%$. A confidence interval of 2 standard deviations on either side of the true mean, would cover $\mu \pm 80\%$. Constraining the mean to 0 amounts to a kind of Bayesian procedure, with a strong prior that the true mean daily return for financial assets is much closer to 0 than to a value randomly drawn from the sampling distribution for the mean.

[5] Kupiec [1995] discusses the pros and cons of measuring VaR relative to zero, or relative to the expected value.

but rather, a Student-t distribution, which has fatter tails than the normal.[6] Even the fact that the volatility parameter is obtained by taking the square root of the sample variance introduces a bias. The sample variance is an unbiased estimate of the true variance, but due to Jensen's Inequality, the sample standard deviation is biased low as an estimate of the true volatility, because the square root is a concave function:

$$E[\hat{\sigma}] = E[\sqrt{\hat{\sigma}^2}] < \sqrt{E[\hat{\sigma}^2]} = \sqrt{\sigma^2} = \sigma$$

Estimation risk is an important factor in evaluating statistical measures of risk exposure, that seems to be largely ignored in practical risk management.[7]

We begin in the next section with an examination of this basic sampling problem. But, while we will see that estimation risk is more serious than might have been recognized previously, the "Baseline case" with constant parameters and normally distributed return shocks actually represents the best situation one could reasonably hope for. Since with constant parameters, the classical estimators for mean and volatility are consistent, the solution to estimation risk in the Baseline case is simply to use more data in the calculations. Estimation error would be negligible from a sample of, say, 5 years of returns. Although research has shown that more accurate volatility forecasts can often be obtained by using considerably longer historical returns samples than what is typical (see Figlewski [1997] or Green and Figlewski [1999]), this is not done in practice out of concern that volatility is not constant over time.

A great deal of empirical evidence shows that volatility is not time-invariant (see, for example Schwert [1989]. Figure 1 plots the sample volatility for the S&P 500 stock index in a moving 63-day window from 1992 - 2002, during which time the sample volatility ranged from around 6.0 percent to over 33.0 percent. If the true volatility drifts, extending a sample backward to bring in more, but older, data may simply contaminate the calculation. Common practice is to use relatively short sample periods (generally less than a year), and often to downweight older observations relative to recent ones with a weighting technique such as an exponentially weighted moving average.

The problem I focus on in this paper is how estimation risk affects our ability to predict the probabilities of rare events accurately when the parameters of the underlying returns process are nonstationary. If we know the form of the returns distribution a priori, normal perhaps, then we can accumulate information quickly, since every observation drawn from that distribution will yield information about its tails. On the other hand, if we do not know the distribution and simply want to tabulate an empirical distribution, we have to wait for many observations before we can learn much about the tails, because the world does not generate information about rare events very rapidly. Now if the data

---

[6] A proof is provided in the Appendix.

[7] The impact of estimation risk on optimal investment and portfolio choice was explored many years ago by Bawa, Brown and Klein [1979], but they did not address risk management per se. Jorion [1996] raises the issue in a VaR context, but his focus is primarily on offering an alternative estimation technique that can improve accuracy in the case in which the true form of the returns distribution is known (e.g., normal) and the unknown parameters are assumed to be constant.

generating process itself is changing over time, estimating the empirical tails becomes much harder by either technique. It also becomes much less tenable to assume that we know the form of the distribution a priori.

I use simulation to explore the problem of evaluating risk exposure when the parameters of the underlying returns distribution vary stochastically over time. In the next section, we begin by considering the problem of estimating risk exposure on a single data sample using standard techniques. This gives a framework for thinking about the problem and introduces the measures of prediction accuracy that we will use throughout the paper. We then present simulation results to show the overall performance of the standard approach in a repeated sample, but with nonstochastic volatility. Section 3 then extends the simulation to the case in which volatility evolves stochastically over time, following a mean-reverting square root process. In Section 4, we examine the impact of non-Gaussian return shocks, drawn from a fat-tailed Student-t distribution. We also consider whether weighting past observations unequally using an exponentially weighted moving average can reduce the effect of estimation error. Section 5 looks at the effect of autocorrelation in the errors when risk exposure is estimated from a rolling sample of returns, as is the common practice in real-world risk management. We find that this exacerbates the problem of estimation risk, and can greatly increase the probabilities of multi-day events, such as three tail occurrences in a row. In this section, we also compare our simulation results to estimation risk in the real world, using a rolling volatility estimate on 40 years of Standard and Poor's 500 Index returns. Section 6 concludes.

2. Estimating Risk Exposure when Volatility is Constant

To illustrate the nature of the forecasting problem and to introduce the type of analysis we will employ in the paper, this section looks at estimation error in using the standard VaR approach to predict the tails of a returns distribution when the underlying returns process is well-behaved.

We assume the asset value follows a standard lognormal diffusion:

$$\frac{dS}{S} = \mu\, dt + \sigma\, dz \tag{1}$$

S is the value of the security (or asset, liability, portfolio, etc.) that we are interested in, constants μ and σ are the instantaneous mean and volatility, at annualized rates, and dz represents standard Brownian motion.

For the simulation, we discretize (1) as

$$r_{t+1} = \ln(S_{t+1}/S_t) = \mu\,\Delta t + \sigma\,\tilde{z}_t\sqrt{\Delta t}\,; \qquad \tilde{z} \sim N(0,1) \tag{2}$$

The time interval Δt represents one trading day, and is set to the value $\Delta t = 1/250$. We will refer to a period of 21 days as a "month," 63 days as "3 months" and 250 days as a "year."

Value at Risk addresses the following question: For a specified holding period (typically 1 day) and probability level α, what is the return such that the probability of experiencing a worse return over the holding period is no more than α? The process in (1) produces a normal returns distribution, for which the one-day 5% VaR is given by $-1.645\,\sigma/\sqrt{250}$ and 1% VaR = $-2.326\,\sigma/\sqrt{250}$. Typically the minus sign is suppressed and VaR is stated as a positive number.

Use of VaR as a practical tool for risk measurement has grown rapidly. It is intuitive and relatively easy to calculate. It suffers from a variety of shortcomings, however, including the problem that it is not a "coherent" risk measure, in the terminology of Artzner, et al [1999]. A practical difficulty is that while VaR specifies where the α-tail of the returns distribution begins, it says nothing about the distribution of outcomes that fall in the tail. A better (and coherent) measure of tail risk is the expected value of the return conditional on being in the α-tail. This is called by various names, including "Conditional Value at Risk" (C-VaR), or "expected shortfall." Figure 2 illustrates the two concepts for the standard normal distribution. Our investigation of estimation error is not specifically tied to Value at Risk, but VaR is a useful concept in discussing the general problem of assessing the risk exposure associated with the occurrence of large, but low probability, adverse events.

Let us first consider calculating the 1-day 5% and 1% VaR using a single 3-month sample of returns. We simulated 63 consecutive returns from equation (2), with the true parameter values set to $\mu = 0$, $\sigma = 0.20$. The sample mean (which for this particular run turned out to be -30.75%, see footnote 4) was suppressed and the volatility was estimated from equation (3).

$$(3) \qquad \hat{\sigma} = \sqrt{\frac{250}{63}\sum_{t=1}^{63} r_t^2}$$

This sample produced a volatility estimate $\hat{\sigma} = 16.62\%$. Multiplying by $-1.645/\sqrt{250}$ and $-2.326/\sqrt{250}$, respectively, gave estimated 1-day values of 5% VaR = -1.73%, 1% VaR = -2.45%. These are shown in Panel A of Table 1.

Standard VaR focuses on the location of the extreme tail of the returns distribution. But, because the sample volatility is an underestimate of the true volatility in this case, the predicted VaR values also underestimate the true values, which are -2.08% and -2.94%, respectively. Figure 3, showing the left tails of the true and the estimated distributions, illustrates the sampling error problem graphically. In this case, the 5% tail of the true distribution starts at a loss of -2.08%, 0.35% larger than is predicted by the standard VaR calculation, and the average loss in the 5% tail is -2.61%, 0.44% worse than expected.

Another way to think about the estimation error in this problem is in terms of the underprediction of the probability that the next period return will be worse than α, the target VaR probability. Using the sample volatility, one would predict that there is only a 5% probability of a loss worse than -1.73%. But under the true distribution, the probability of a loss that large is actually 8.58%. The last column in Table 1 gives the ratio of the true tail probability to the predicted probability. For the 5% tail, the true probability is 1.72 times greater than the predicted 5%, and for the 1% tail, the ratio is 2.66. Figure 4 illustrates this way of evaluating estimation error.

The 5% and 1% cutoff values, that are by far the most frequent choices in practical VaR applications, actually represent relatively common outcomes. An event with a 5% probability will occur on average one time out of 20. For daily returns, therefore, 5%-tail events should average more than one per month. A 1% event should occur every few months, about 2 1/2 times a year. While it is obviously important to be prepared for these events, a cautious risk manager needs to be concerned about losses that are less frequent but much more serious.

Panel 1B of Table 1 extends the VaR calculations from Panel A into the more extreme tails of the returns distribution. It shows that the farther one looks into the tail, the greater is the effect of the estimation error on the volatility parameter. For example, while the actual 5% tail is -0.35% worse than predicted, the actual 0.5% tail is -0.55% worse and the 0.05% tail is -0.70% worse. For these very low probability events, the probability ratio gives a clearer picture of the impact of estimation error than the location of the tail does. The true probabilities are several times the predicted values, and the ratio increases for rarer events, such that a return that is predicted to occur only 1 time in 10,000 is actually almost 10 times more likely than that.

Another way to understand the probability results we are developing is in terms of how frequent an event with a given probability is. Table 2 shows the frequency of occurrence for the tail values examined in Table 1, both the predicted frequencies using the sample volatility of 16.02%, and the true frequencies based on the true $\sigma = 20\%$. For example, an event with a probability of only 0.002 can be expected to occur on average every 2 years (where, as mentioned above, we take a "year" to be 250 days). Even a 1 in 10,000 event falling in the 0.0001 tail happens on average every 40 years: not a common occurrence by any means, but not out of the range of concern for a prudent risk manager with a long term view.

One can express the frequency of occurrence in a different way that takes into account the fact that the probability of an event on any given day is independent of the outcomes on other days. Given the 1-day probability, we can calculate the likelihood that there will be no event over a period of K days. If P is the probability of an event on a given day, the probability of no event in K days is $(1 - P)^K$, making the probability that there will be at least one event in K days

(4) $q = 1 - (1 - P)^K$

For a specified probability of occurrence q, equation (4) can be solved for K_q, the shortest time period such that the probability an event will be observed during that interval is greater than q.

(5) $K_q = \log(1 - q) / \log(1 - P)$

Table 2 presents predicted and true values for $K_{50\%}$, the shortest period going forward such that there is more than 50% probability that an event will be observed. For example, a 1 in 1000 event will happen on average every 4 years, but there is more than 50% chance of one within the next 2.8 years. An event in the 0.0002 tail of the distribution might not seem so unusual after all, if the odds are greater than 50/50 of seeing one in less than 14 years.

What is more striking in Table 2 is how much the estimation error in using the sample volatility for this problem changes the calculation. For example, given our parameters, an event whose probability is estimated to be 0.001 is actually more than 5 times more likely than predicted. So, while an event is expected on average only every four years, the true frequency is one every 195 days, and there is more than 50% chance of an occurrence within less than 6 months. A once-in-40-years 0.0001 event is actually more likely than not within less than 3 years.

These VaR estimates and probabilities have all been computed from a single 63-day sample of returns, in which the sample volatility was substantially lower than the true value. They do not represent the average forecasting performance for a standard VaR calculation. In order to examine that much more relevant issue, we use equation (2) to simulate a long sequence of 2,500,000 consecutive returns and consider repeating the process of estimating volatility every 63 days from equation (3) (i.e., setting the estimate of the mean to zero). The 2 1/2 million simulated days yields 39,682 non-overlapping 63-day periods. To maximize comparability across simulations based on different assumptions and a variety of estimation strategies, we use the same random seed in every simulation run throughout the paper.

The estimated volatility is used to compute 1-day VaR and C-VaR values for the first day following the sample period. The realized return for that day (simulated using the true volatility) is then converted into a number of standard deviations by dividing it by the predicted volatility. That is, the return shocks are standardized by expressing them in terms of the forecasted standard deviations. When the full set of returns data has been processed, for each value of α, we determine the location of the cutoff for the true α-tail (i.e., the number of predicted standard deviations that would have captured exactly α percent of the returns), the true mean return in that α-tail (true C-VaR) in units of $\hat{\sigma}$, and the actual percentage of returns that fell within the predicted α-tails (true probability).

Table 3 presents the results. Overall, the root mean squared error (RMSE) in the volatility forecast is 0.0178. This is approximately 9 percent of the true volatility value of 0.20. How large this volatility forecast error is judged to be depends on what use is to be made of the estimate: for pricing options it may be considered quite large, but for assessing overall risk exposure, it seems at first glance to be relatively small. Table 3 shows that for the 10% and 5% tails, estimation error does not seriously affect the standard VaR calculation. For example, the 5% VaR level is estimated to be -1.645 times the sample volatility, while in the simulation, the true 5% VaR was at -1.649 times the sample volatility and the C-VaR estimate was also very close to the true value. The true probability of a return in the predicted 5% tail was 5.05%, only 1.01 times the predicted probability.

Estimation error is a little greater for the 1% VaR calculation. The actual 0.01-tail was at -2.384 standard deviations, and the true probability that the next period return would fall in the predicted 1% tail was actually 1.17%. However, as we saw above, the effect of estimation error is greater for more extreme values of α. For $\alpha = .0005$ (a 1-in-2000, or once-in-8 years, event) the true probability is more than twice as large as predicted.

The simulation results reported in Table 3, using a 63-day sample period and a constant value of 0.20 for the true volatility, will be our standard of comparison in later experiments, so we refer to this as the Baseline simulation.

We have chosen a period of 3 months (63 trading days), which is a common sample size for estimating volatility. RiskMetrics, for example, uses an exponentially weighted moving average for volatility on a historical sample of 74 days, with a decay rate of 0.94. That methodology puts about 99% of the weight on observations within the first 63 days. (We will examine the effect of downweighting old data in this calculation below.) The reason to limit the data sample is to reduce the effect of time variation in the volatility. However, in the Baseline case, volatility is a constant parameter, so the way to reduce sampling error is simply to increase sample size.

Table 4 shows the effect of varying the sample size from 21 days to 250 days, with Panel A showing the effect on the true cutoff values for the α-tail. Panel B shows the ratios of true to predicted probabilities, which are plotted in Figure 5. With a fixed 2 1/2 million returns, the number of non-overlapping samples decreases with the sample length, increasing the sampling error in our simulated "true" tail values. A 250-day sample would produce only 10,000 observations, making locating the 0.0001 tail problematical, so we increased the simulation sample size to 10 million for that run. Even so, the expected number of 0.0001 tail events was only 4, leading one to suspect that Table 4 may understate the true probability ratio for the $\alpha = 0.0001$, $K = 250$ combination (there were no events in the predicted 0.0001 tail in the first 2 1/2 million returns runs, for example). Sampling error probably also plays a significant role in our 125-day simulation, for which only 20,000 runs were used. Even so, the results indicate that using more data in the sample substantially reduces the problem of estimation error, making it relatively unimportant when volatility is estimated over the past year, until one gets out beyond the 0.001 tail.

Unfortunately, the Baseline case probably represents the best possible conditions for predicting the extreme tails of a returns distribution using the standard methodology. The true distribution is normal, the true volatility is constant, and the true mean is 0, so computing sample volatility as if the sample mean were zero actually saves a degree of freedom by imposing a true constraint. In the real world, it is safe to say that none of these conditions holds. Predicting the standard deviation of next period's return requires aiming at a moving target, and it is no longer automatically true that a longer sample period produces a more accurate estimate. Procedures that limit the amount of past data used, either simply cutting off the sample by using a fixed window, or downweighting older observations, may improve accuracy. We now turn to the estimation problem when volatility is allowed to change over time.

3. Estimating Risk Exposure when Volatility is Time-Varying

The returns model in equation (1) has become the workhorse of continuous-time asset pricing. The empirical evidence suggests, however, that real-world asset price processes are more complex than this. In this section we will consider the estimation problem when volatility varies stochastically over time.

We will continue to leave aside the behavior of the drift term μ. It is well-known that μ is hard to predict. Moreover, the drift does not have a very large effect on volatility estimation when returns are sampled over short intervals, because it is of smaller order than the volatility (Δt versus $\sqrt{\Delta t}$, for Δt close to zero). Only when the sample mean is very different from the true mean does it have much effect on the sample volatility. This is why substituting 0 for the sample mean is an adequate "fix" for the sampling error problem (even though it contradicts the principle that the true mean return for a normal asset in equilibrium should be positive, not 0).

A variety of alternatives to (1) that allow volatility to change over time have been explored in the literature. We will model time-varying volatility as a mean-reverting square root process, as in Heston [1993]. We regard this as a reasonable assumption, that will allow us to explore the estimation risk problem when the parameter of interest drifts over time, but it is only one of many alternatives.

The model of the real world that I have in mind is one in which virtually nothing is ever truly stationary. Even if a mean-reverting square root process were to be the best description for volatility movements, we should expect that the model parameters would drift over time. And if we tried deal with the problem by building a model of the parameter drift, the parameters of the parameter-drift process would themselves be nonconstant. In short, I believe that parameter drift is endemic to the financial system, and that one is always trying to predict the future value of a variable of interest using data drawn from a previous time, when its value was different. Since we can not expect that important parameters will stand still while we measure them, our best hope is that information will arrive rapidly relative to the rate at which the parameters are changing, so that forecast error can be limited. It is this need for rapid information arrival that fails when we are trying to assess the probabilities of rare events.

Equation (1) now becomes

(6) $$\frac{dS}{S} = \mu\, dt + \sigma_t\, dz$$

(7) $$\sigma_t = \sqrt{V_t}$$

(8) $$dV = \kappa(\bar{V} - V_t)\, dt + \theta \sqrt{V_t}\, dw$$

$\bar{V}$ is the long run variance, κ is the rate of reversion of the current variance V_t toward that long run value, $\theta\sqrt{V_t}$ is the volatility of the variance process and dw is a second Brownian motion, independent of dz.[8]

Equations (6) and (8) are discretized for the simulation study as follows.

(9) $$r_{t+1} = \ln S_{t+1} / S_t = \mu\, \Delta t + \sigma_t\, \tilde{z}_t \sqrt{\Delta t}\, ; \qquad \tilde{z} \sim N(0,1)$$

(10) $$V_{t+1} = \kappa(\bar{V} - V_t)\, \Delta t + \theta\, \tilde{w}_t \sqrt{\Delta t}\, ; \qquad \tilde{w} \sim N(0,1)$$

In choosing values for mean reversion, κ, and the volatility of variance parameter, θ, in the simulation we must be aware of the Feller condition for overall stability of the variance process.[9] For the variance process to remain positive over the long run, we must have

(11) $$\kappa \bar{V} > \frac{\theta^2}{2}$$

Otherwise, in finite time, variance converges to zero. Although we choose parameter values that satisfy (11), in a discrete simulation of equation (10), we still get occasional random draws for w_t that would produce negative variances. When that happens, we set V_{t+1} to 10^{-8}, essentially imposing the constraint that annual volatility can not be less than 0.01%.

---

[8] For equity returns, it is common to allow negative correlation between dz and dw, see for example Bakshi, et al [1997]. We do not do that here, in the interest of limiting the amount of results to be presented. Explorations allowing correlation between return and variance shocks in this model did not indicate striking differences from the results presented here.

[9] See Feller [1951].

Table 5 presents simulation results for the standard estimation technique examined above, but with volatility that evolves over time according to equation (10). That is, the user is assumed to calculate sample volatility from the last 63 days of simulated returns as if it were a constant parameter. He then estimates the location of the α-tails of the period t+1 returns distribution for a range of αs, treating the sample volatility as if it were the true volatility. As in the Baseline simulation, Table 5 Panel A shows the true α-tails, expressed in terms of the estimated standard deviations; and Panel B gives the ratios of true to predicted probabilities.

The first column in Panel A shows the theoretical tails for a normal distribution, which is what the user incorrectly assumes will apply to this problem (and which would be correct, if the true volatility could be used in the VaR calculation instead of the estimated volatility). The second column reproduces the Baseline results for a 63-day sample. Runs 1-9 assume 4 different values for the variance mean-reversion parameter κ: 0.20, 0.40, 1.0, and 2.5. For each κ, we show two or three values for θ, with the largest one in each case being within 0.05 of a value that would violate the Feller condition.

Not surprisingly, allowing variance to change over time increases the RMSE of the forecasted volatility, to about 2.3% when $\theta = 0.10$ and to more than 3% with $\theta = 0.20$. The cutoff values for the true α-tails are distinctly more negative than predicted. Panel B shows the substantial impact on predicted probabilities of tail events. As we suggested above, the Baseline simulation represents the best case for the standard approach to risk assessment.

We do not know, however, what values of κ and θ would be typical for real world asset returns. Bakshi, Cao, and Chen [1997] present estimation results for a variety of stochastic volatility models for the S&P 500 stock index, including one that is close to (10). Their values are $\kappa = 1.15$ and $\theta = 0.39$. However, there are several serious qualifications, which make it not completely appropriate to take these values as good real world estimates for the parameters in our problem. First, the parameter values are obtained by implying them out from S&P 500 index option prices, not by fitting them to actual returns data. Second, the Bakshi et al specification of the returns process includes a strong negative correlation between dz and dw, while we have modeled them as independent. Finally, with our value for long term variance, the combination of parameters in Bakshi, et al, would violate the Feller condition and cause the variance process to be unstable.

Figure 6 plots the effect of volatility of variance on the true / predicted probability ratios. We set $\kappa = 1.0$ and plot the results for a range of θ values, from 0 (the Baseline case) to 0.25. (It may be useful to refer back to Table 2 to get a feel for what the errors in the estimated tail probabilities mean in practical terms.)

In Figure 7, we examine the effect of changing variance mean reversion while holding θ fixed at 0.10. It is not surprising that larger volatility of variance makes forecasting harder, as we see in Figure 6. It is less clear what one should expect for the rate of reversion toward long run variance. On the one hand, a rapid rate of reversion tends to keep the process closer to its long term (constant) value, which should make forecasting easier. On the other hand, a larger κ also means that when instantaneous variance differs from $\overline{V}$, it will drift more rapidly over the sample period under the force of mean reversion, which could make post-sample forecasting harder. As Figure 7 shows, the former effect appears to win out, at least with these parameter values: higher κ reduces the impact of estimation error in calculating the location of the α-tails. Even so, in Table 5 it is clear that, overall, time varying volatility makes the problem of assessing risk exposure worse.

4. Further Departures from the Standard Model

The assumption that returns come from a normal distribution is widely made for mathematical convenience, but a large amount of statistical evidence indicates that the true distribution is more fat-tailed than the normal. Time-variation in the variance would be one reason for apparent non-normality, of course. But even when models with explicitly stochastic volatility are estimated, it often seems that the returns shocks are fat-tailed. A convenient alternative to the normal is the Student-t distribution. It has one additional parameter, the degrees of freedom(d.f.), that governs the tail behavior. The distribution converges to the standard normal as d.f. goes to infinity, but for small values, the t-distribution is distinctly fatter-tailed than the normal. In fact, in a t-distribution all moments greater than d.f. are infinite. For example, a t(3) has finite mean, variance and skewness, but infinite kurtosis and higher moments.

Table 6 examines the effect of drawing the disturbances in the returns equation from a t-distribution with either 7 or 4 degrees of freedom, compared with normal (0,1) shocks. The shocks to the variance equation are still drawn from a normal distribution. We consider three cases: constant parameters, stochastic volatility with a relatively low volatility of variance of 0.05 and slow mean reversion of 0.40, and more strongly stochastic volatility with $\theta = 0.20$ and $\kappa = 2.5$. The results show that fat-tailed disturbances from a t-distribution significantly worsen the underestimation of exposure to extreme returns. Even with the low volatility of variance process, the true probability of experiencing a 0.0001-tail event is in excess of 20 times greater than is predicted under the assumption that returns are normal. However, comparing across the different runs it becomes clear that this result is due much more to the tail-fatness of the t-distribution than to the problem of sampling error that we have been examining above. Given the degrees of freedom in the distribution, allowing time variation in the volatility makes little difference to the results. For example, with t(7) shocks, when variance is constant the true probability of a return in the 0.0001 tail is about 23 times the probability predicted from a normal distribution. When variance has a volatility parameter $\theta = 0.05$, the same multiple of about 23 applies, and increasing the volatility of variance to 0.20 only moves the ratio to 25.5. Assuming the returns distribution is normal when return shocks actually come from a t-distribution leads to huge underestimates of the exposure to large low probability events. But the additional risk exposure that can be attributed to estimation risk is not very important.

The fact that variance is known to be time-varying has led to the use of alternative estimation techniques to reduce the problem. Restricting the returns sample to a fixed window, such as 63 days, only makes sense because it is felt that more recent data give a better estimate of a moving parameter. A fixed window imposes a rather artificial weighting of past data in the estimation, either 1 for an observation in the window, or 0 for one outside the window. A common alternative that downweights data more smoothly as it ages is to use an exponentially weighted moving average (EWMA).

Under EWMA, each observation is downweighted at a fixed rate of decay as it ages. The volatility estimate is given by

$$(12) \qquad \hat{\sigma} = \sqrt{\sum_{t=1}^{K\max} \lambda^{t-1} (r_t)^2 / \sum_{t=1}^{K\max} \lambda^{t-1}} \ ,$$

where λ is the decay parameter and Kmax is the maximum lag included in the calculation. RiskMetrics, which has made a profitable and influential business out of estimating volatilities and correlations for use in Value at Risk calculations, uses $\lambda = 0.94$ and Kmax = 74 for all of its daily volatility estimates, the latter chosen because if an infinite number of past observations were available, the total weight applied to those more than 74 days old would be less than 1%. We use Kmax = 63 here, which captures most of the weight of an infinite sample, with the fraction ranging from about 72% for $\lambda = 0.98$ to more than 99.8% for $\lambda = 0.90$. According to RiskMetrics, tests on volatilities from a large number of return series indicate that $\lambda = 0.94$ seems to give the best average forecast performance (see RiskMetrics [1996]).

An EWMA offers the possibility of extracting some volatility information from comparatively old data while recognizing that more recent data probably contains more information that is relevant for predicting next period's volatility. The optimal decay factor should be a function of the rate of change of volatility and the size of the stochastic component. A relatively large value for λ, close to 1.0, would be appropriate for a stable and slow moving variance process, while if volatility changes rapidly over time, one would like to reduce the weighting of older data by using a smaller λ.

In Table 7, we present simulation results for three volatility regimes, comparing four decay factors. In results not shown here, we found that varying the rate of mean reversion κ had very little effect on the estimation error in the tail estimates for this case. For example, with $\lambda = 0.94$ and $\theta = 0.10$, κ values of 0.20, 1.0, and 2.5 produced forecast RMSEs of 0.0270, 0.0269, and 0.0267, respectively, and the 0.0001-tails fell at 4.269, 4.277, and 4.273. Given the minuscule effect of changing the rate of variance mean reversion over a broad range, we report only results with $\kappa = 1.0$. θ values were set to 0 (constant volatility), 0.05 (relatively stable variance) or 0.25 (volatile variance), and we considered λs of 1.0 (no downweighting in a fixed 63-day window), 0.97, 0.94 and 0.90.

Not surprisingly, if volatility is constant, downweighting older data points simply throws away useful information, since every observation contains an equal amount of information. The RMSE results show that with $\theta = 0$ forecast accuracy diminishes monotonically as λ is reduced from 1.0 to 0.90. The tail statistics and probability ratios are consistent with this. The same result holds for the low volatility of variance $\theta = 0.05$ regime. However, in the high θ regime, the pattern is different. Forecast accuracy is better for $\lambda = 0.97$ and $\lambda = 0.94$ than it is for either $\lambda = 1.0$ or $\lambda = 0.90$. Evidently, in this case no downweighting ($\lambda = 1.0$) allows too much noise from obsolete data points into the calculation, while too much downweighting ($\lambda = 0.90$) excludes too much useful information that could have be extracted from observations that are not old enough to have lost their value. This suggests that we would have found a similar result for the $\theta = 0.05$ case, if we had checked λ values between 0.97 and 1.0.

Figures 8 and 9 do just that, reporting the true / predicted probability ratios for the low and high θ regimes, respectively. Figure 8 shows that while for the nearby 0.05 to 0.0005 tails, the best performance is achieved with no downweighting (or at least, with a decay factor over 0.99), for the further 0.0002 and 0.0001 tails, it is better to use a decay factor of 0.99 than to weight each observation equally. Figure 9 shows that this general pattern is similar and more pronounced when θ is relatively high. The most extreme tails are estimated more accurately using EWMA, with a decay parameter of 0.92 or 0.94. The nearer tails also are more accurate, but with less downweighting. The best λ values are $\lambda = 0.96$ for the 0.01 tail and $\lambda = 0.97$ for the 0.05 tail. Thus, EWMA appears to give some improvement in tail estimation under conditions of time-varying variance. The overall impact of estimation error on predictions of risk exposure, however, is still very large.

5. Autocorrelation in the Volatility Forecast Errors from a Rolling Sample

One feature of our research design that affects the results substantially is the fact that so far we used only non-overlapping samples. This allowed us to compute the effect of estimation error on the predicted probabilities without the problem of serial dependence that overlapping samples would produce. This is both good and bad. We get better statistical behavior of our estimation with non-overlapping samples, but the forecasting problem that we are modeling is different from what risk managers face in the real world. A firm that uses VaR as a risk management tool will recompute volatility each day to estimate the exposure for the immediate future. Each time, the most recent day's observation is added to the sample and the oldest day is dropped. This means that the prediction error on date t will be highly correlated with the error on date t-1, perhaps generating a string of volatility underestimates, and multiple tail events. This section explores that issue.

We simulate 250,000 daily returns (1000 years) using the same procedures as above, but then consider estimating volatility from a rolling 63-day sample, updating each day. This produces (250,000 - 63) = 249,937 1-day VaR forecasts, whose prediction errors will be serially correlated. Table 8 presents results for the same set of parameter values as in Table 5. Panel A shows the actual tail cutoff values and Panel B gives the probability ratios.

Average prediction accuracy, as measured by the root mean squared error of the volatility forecast, is very similar for the overlapping and non-overlapping samples. Although there are periods when volatility is underestimated, which increases the likelihood of observing what appear to be multiple tail events within a short time interval, these times are balanced by periods of overestimated volatility, with a lower than expected chance of an event. Overall, the RMSE of the volatility estimate is not affected very much. In other words, using a rolling sample does not increase the bias of the volatility estimate. Nor does it seem to make much difference in the Baseline constant volatility case. But once volatility is allowed to vary over time, the offsetting of under- and overestimates in the rolling sample does not produce offsetting errors in estimating the tails of the distribution. Under stochastic volatility, a rolling sample produces substantially more tail events than were shown in Table 5.

Panel B of Table 8 is set up to illustrate clearly the difference a rolling sample makes to risk estimation. For each probability cutoff, we show the probability ratio for both the rolling sample and the corresponding non-overlapping sample result from Table 5. It is evident that even a low value of θ leads to a substantial increase in the probability of a tail event, with the difference increasing as one looks further into the tail. For example, with $\theta = 0.10$ and $\kappa = 0.4$ (Run 4), a rolling 63-day sample would experience 58% more 5% tails events than expected (versus only 3% more with no overlap). But at the 1-in-a-thousand 0.1% level, the rolling sample would produce more than 11 times as many events as expected, while the non-overlapping sample only experiences twice as many. A faster rate of volatility mean reversion κ mitigates the effect considerably, but even with $\kappa = 2.5$, a rolling sample still produces considerably larger tail probabilities than expected and than the non-overlapping sample does.

So far we have been examining results only from simulations. This raises the question of whether these experiments really reflect what happens in the real world. To provide some evidence on this issue, I fitted rolling volatility forecasts on about 40 years of S&P 500 stock index returns, and examined the tail predictions using the same kind of analysis we have been considering. The sample period is July 2, 1962 through August 30, 2002, which yields 10113 observations. Rolling estimations were done using returns from the previous 21, 63 and 250 days. Table 9 presents the results.

Table 9 shows clearly that the standard procedure of estimating volatility over a relatively short historical period and rolling the sample forward each day leads to serious underestimates of the tail probabilities in the real world, just as it does in our simulations. The more remote tails are underestimated to a larger degree, but even the 1% tail had more than 70% more events than were predicted. One noteworthy feature here is that adding more data by going from 21 to 63 to 250 day estimation periods only improves the tail predictions very slightly. This suggests that the problem is not just sampling error, which can be made to go away by using more data points, as was shown in Table 4. Time variation in the volatility, which does not go to zero with a longer estimation sample, is likely to be playing an important role, as well (and probably fat tails in the returns distribution, too).

Risk exposure is not limited to getting one really bad day. When statistically independent volatility forecasts are produced from non-overlapping data, the probability of getting two tail events in a row is just the square of the probability of one event. But with a rolling sample, the volatility forecast errors will be highly positively autocorrelated. This will produce a much greater chance of getting multiple events over a short period than with independent forecasts. Table 10 examines this phenomenon.

Panel A presents results on the occurrence of two tail events in a row and Panel B does the same for three events in a row. The Theoretical Probability is just the tail probability raised to the power 2 or 3. The remaining columns show the probability ratios for other asset price processes. The first is the constant volatility Baseline run. These results show that sampling error alone leads to a substantially higher multi-day risk than expected. For example, three 1% tail events in 3 days should be a one in a million event, but because of sampling error it is 12 times more probable than that, even when volatility is constant and returns are normal.

The next four columns give the probability ratios for different values of θ and κ, ranging from a relatively low θ of 0.10 with κ of 0.2 or 1.0, to high values of θ = 0.40 and κ = 2.50. The effect is striking. If either θ is high, or θ is moderate but mean reversion is slow, the probability of two or three events in a row grows sharply. For both Run 1 and Run 4, the "one in a million" occurrence of three 1% tail events in a row is actually much closer to 1 in 1000--something that might happen about every 4 years on average.

Finally, the last column gives statistics on actual multi-day tail events observed for the S&P 500 stock index over the 40 year sample period. The results are not as extreme as some of the simulations, but are more extreme than others. Three 5% tail events in a row, for example, should happen only once in 8000 days, or about 32 years. But three in a row was actually almost 10 times more frequent than that, averaging 1 in about 3 1/2 years.

Panels C and D show the same kind of results for two relatively more likely multi-day tail events: 3 events in 5 days and 3 events in 10 days. The theoretical probability of K events in N days can be computed directly from the binomial distribution, with the probability of a single event set equal to the tail cutoff.

If α is the tail cutoff probability, the probability of observing K (or more) events in N days $P_{K,N}$ is given by

$$(13) \qquad P_{K,N} = \sum_{k=K}^{N} \binom{k}{N} \alpha^{k} (1-\alpha)^{N-k}$$

One thing that complicates the interpretation of these results a little is the fact that with a rolling estimation, the same events can be counted more than once. For example, if there are three events in a row, there will be 7 days in the sample in which those 3 events will fall within a 10 day window. However, these "multiple counting" cases will tend to be offset by multiple-<u>under</u> counting periods that will occur, as well. Asymptotically, this should not bias the probability ratios as reported in Table 10. Again, these results indicate that the use of a rolling sample to estimate sequential volatilities and risk exposures increases the problem of estimation risk.

6. Conclusion

Use of statistical procedures to quantify exposure to financial risk has been spreading rapidly among real-world risk managers, with Value at Risk probably representing the single most common technique at present. Many of the alternatives to VaR also involve trying to estimate the tails of a probability distribution of asset values or returns. This effort inevitably entails estimation error, but the effect of that error is seldom considered explicitly. We have seen that even with constant volatility and normal distributions, the best situation for the standard technique based on samples of a few months of historical data leads to substantial misestimation of tail probabilities. The problem grows worse the farther into the extreme tails one looks.

In the real world, it is not plausible to expect volatility to be time-invariant, and much empirical evidence indicates that it is not. But when volatility changes stochastically over time, the estimation error in predicting the probabilities of rare events can get much larger, and the possibility of substantially increasing accuracy simply by using longer data samples disappears. That is the general problem we have examined here. Using an extensive simulation (10,000 years of daily returns), we found that events with probabilities on the order of 1 in 1000 or less can easily be twice as likely as predicted. This tends to happen when volatility is underpredicted because it has increased during the sample period, so the true volatility of tomorrow's return exceeds that of the data sample. The more volatile variance is, the worse the problem becomes.

We extended the analysis to consider how things would change when the true returns distribution was more fat-tailed than the normal, as much evidence suggests it is. Drawing returns shocks from a t-distribution with either moderate or large tail fatness, we found that the estimation errors for the extreme tails grew enormously. Events that would be predicted to occur on the order of once in a decade under a normal distribution can be more than 15 times more likely when shocks come from one of the t-distributions we looked at. We also examined the alternative exponentially weighted moving average estimation strategy, which is used in practice to mitigate the expected instability of returns volatilities. EWMA did improve predictive accuracy when volatility was strongly stochastic, especially for the remote tails which have focused on.

The results reported above were obtained using only non-overlapping samples. By contrast, real world risk managers who use VaR as a risk management tool will recompute volatility each day, adding in the most recent observation and dropping the oldest one from the sample. The volatility prediction error on date t will be highly correlated with the error on date t-1, which can generate a string of volatility underestimates, and multiple tail events. Using a rolling sample for volatility estimation substantially increases the frequency of apparent tail events. It also makes multiple events within a few days much more likely.

One alternative approach to estimating volatility that is commonly used is historical simulation. This involves analyzing past returns from a much longer sample period, typically several years, and simply tabulating the empirical returns distribution. This may ameliorate some of the estimation risk problems that we have seen with parametric estimation on a short data sample. For example, if return shocks are non-normal, using a long sample period will allow a better empirical fit to the actual tail behavior. Also, if volatility varies stochastically but the rate of mean reversion is fairly rapid, a sample of several years may produce a reasonably good estimate of the ergodic distribution. On the other hand, it is not possible to say anything about the remote tails from a limited sample. For example, in 2 years of data, one would not expect to see even one 0.1% event. And the problem of serial correlation of the errors should be considerably worse. Estimation risk in historical simulation as a strategy for assessing risk exposure will be explored in subsequent research.

References

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath. "Coherent Measures of Risk." Mathematical Finance 9, 1999, 203-28.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen. "Empirical Performance of Alternative Option Pricing Models." Journal of Finance 52, 1997, 2003-2049.

Basle Committee on Bank Supervision. "Amendment to the Capital Accord to Incorporate Market Risks." Basle: Bank for International Settlements, January 1996. (http://www.bis.org/publ/bcbs24a.htm)

Bawa, Vijay, Stephen Brown and Roger Klein. Estimation Risk and Optimal Portfolio Choice. Amsterdam: North Holland Publishing Co., 1979.

Bollerslev, Tim, R. Chou and Ken Kroner. "ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence." Journal of Econometrics 52, 1992, 5-59.

Boudoukh, Kobi, Matthew Richardson, and Robert Whitelaw. "The Best of Both Worlds." RISK, May 1998, 64-67.

Duffie, Darrell and Jun Pan. "An Overview of Value at Risk." Journal of Derivatives 4, Spring 1997, 7-49.

Feller, W. "Two Singular Diffusion Problems." Annals of Mathematics 54, 1951, 173-82.

Heston, Steven. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." Review of Financial Studies 6, 1993, 327-343.

Jorion, Philippe. "Risk^2: Measuring the Risk in Value at Risk." Financial Analysts Journal, Nov/Dec 1996, 47-56.

Jorion, Philippe. Value at Risk: The New Benchmark for Controlling Market Risk. Chicago: Irwin, 1997.

Karatzas, Ioannis and Steven Shreve. Brownian Motion and Stochastic Calculus 2nd ed. New York: Springer-Verlag, 1991.

Kupiec, Paul. "Techniques for Verifying the Accuracy of Risk Measurement Models." Journal of Derivatives 2, Winter 1995, 73-84.

RISK Publications, "Value at Risk," RISK Supplement, June 1996.

RiskMetrics. RiskMetrics Technical Document, Fourth Edition. New York. 1996. (http://www.riskmetrics.com/techdoc.html).

Schachter, Barry. Website: <http://www.gloriamundi.org/>

Schwert, G. William. "Why Does Stock Market Volatility Change over Time?" Journal of Finance 44, 1989, 1115-1154.

Theil, Henri. Principles of Econometrics. New York: John Wiley and Sons, 1971.

Appendix

Proof that when r is drawn from a normal distribution with constant mean and variance, the conditional distribution of r_{t+1}, given $\{r_{t-K+1}, \ldots, r_t\}$ is Student-t.

The returns $\{r_\tau\}$ are independent draws from a normal distribution with mean μ and standard deviation σ.

We will use the following result from Theil [1971, p. 82]. Let X be a standardized normal variate, let Y^2 have a χ^2 distribution with K degrees of freedom, and let X and $Y = +\sqrt{Y^2}$ be independent. Then $X\sqrt{K} / Y$ is distributed as Student-t with K degrees of freedom. Here, X will be the standardized forecast error for r_{t+1} and Y will be the estimated standard deviation computed from the most recent K returns $\{r_{t-K+1}, \ldots, r_t\}$.

<u>Case 1:</u> Both mean and standard deviation are sample estimates.

$$\hat{\mu} = \frac{1}{K}\sum_{\tau=1}^{K} r_{t-\tau+1} \qquad\qquad \hat{\sigma} = \sqrt{\frac{1}{K-1}\sum_{\tau=1}^{K}(r_{t-\tau+1}-\hat{\mu})^2}$$

The first post-sample return, r_{t+1}, is independent of the returns used in computing the sample mean and variance. The expected value of $\hat{\mu}$ is μ and variance of $\hat{\mu}$ is σ^2/K. This means

$$(r_{t+1} - \hat{\mu}) \sim N(0, \sigma^2 + \sigma^2/K) \quad \text{and} \quad \sqrt{\frac{K}{K+1}}\,\frac{r_{t+1}-\hat{\mu}}{\sigma} \sim N(0,1).$$

From basic statistics, (e.g., Theil [1971], p. 91), we have $\dfrac{\hat{\sigma}^2}{\sigma^2(K-1)} \sim \chi^2(K-1)$.

Applying the result stated above and simplifying gives $\sqrt{\dfrac{K}{K+1}}\,\dfrac{r_{t+1}-\hat{\mu}}{\hat{\sigma}} \sim t(K-1)$. The return r_{t+1} is distributed like a Student-t with K-1 degrees of freedom, but scaled up by the factor $\sqrt{\dfrac{K+1}{K}}$. It has the same zero mean as a t(K-1) variate but its standard deviation is larger, making the distribution a mean-preserving spread on a standard Student-t(K-1). The distribution of r_{t+1} has fatter tails than the normal, and because of the scaling factor each quantile in the tail (5%, 1%, etc.) is more negative than the corresponding quantile for a standard Student-t(K-1).

Case 2: The sample mean is set to zero; only the standard deviation is estimated:

$\hat{\sigma} \;=\; \sqrt{\frac{1}{K} \sum_{t-K+1}^{t} r_{\tau}^{2}}$. If the true mean $\mu = 0$, the constraint is true and a similar computation as in Case 1 yields $\frac{r_{t+1}}{\hat{\sigma}} \sim t(K)$

Case 3: If the true mean is nonzero, suppressing calculation of the sample mean in the estimation procedure introduces a specification error and the proof does not go through. The sample variance is a biased estimate of the true variance and $\frac{r_{t+1}}{\hat{\sigma}}$ will not satisfy the conditions of the theorem.

Table 1: Estimating Value at Risk and Conditional Value at Risk on a Simulated 63-Day Sample of Returns

One run of 63 consecutive returns was simulated from the discretized model

$$(2) \qquad r_{t+1} = \ln S_{t+1} / S_t = \mu \, \Delta t + \sigma \, \tilde{z}_t \sqrt{\Delta t} \; ; \qquad \tilde{z} \sim N(0,1)$$

with parameter values: $\mu = 0.0\%$, $\sigma = 20.0\%$, $\Delta t = 1 / 250$

Estimated values are based on the sample (zero mean) volatility estimate: $\hat{\sigma}_{mean0} = 16.62\%$;
True values are based on the true $\sigma = 20.0\%$.

Panel A: Estimated and True Values for Standard 5% and 1% VaR and C-VaR

Probability	Estimated VaR	TrueVaR	Estimated C-VaR	True C-VaR	True / Predicted Probability
.05	-1.73%	-2.08%	-2.17%	-2.61%	1.72
.01	-2.45%	-2.94%	-2.81%	-3.38%	2.66

Panel B: Estimated and True Risk Exposures for Extreme Tails

Probability	Estimated VaR	True VaR	Estimated C-VaR	True C-VaR	True / Predicted Probability
.005	-2.71%	-3.26%	-3.05%	-3.67%	3.23
.002	-3.03%	-3.64%	-3.36%	-4.05%	4.19
.001	-3.25%	-3.91%	-3.49%	-4.19%	5.12
.0005	-3.46%	-4.16%	-3.68%	-4.43%	6.25
.0002	-3.72%	-4.48%	-3.91%	-4.70%	8.16
.0001	-3.91%	-4.70%	-4.21%	-5.06%	9.98

Table 2: Event Frequency as a Function of Event Probability

The table shows the predicted and true frequency of events with a given predicted probability, based on the sample volatility. One "year" is 250 "days."

For a given probability, $K_{50\%}$ is the time interval such that there is more than 50% probability of experiencing an event within that period.

$K_{50\%}$ is the solution to: $0.50 = 1 - (1 - P)^{K_{50\%}}$, where P is the probability of an event.

Probability	Estimated Frequency 1 / P	Estimated $K_{50\%}$	True P / Predicted P	True Frequency 1 / P	True $K_{50\%}$
.05	20 days	14 days	1.72	12 days	8 days
.01	100 days	69 days	2.66	38 days	26 days
.005	200 days	139 days	3.23	62 days	43 days
.002	2 years	1.4 years	4.19	119 days	83 days
.001	4 years	2.8 years	5.12	195 days	136 days
.0005	8 years	5.5 years	6.25	1.28 years	222 days
.0002	20 years	13.9 years	8.16	2.45 years	1.7 years
.0001	40 years	27.7 years	9.98	4.01 years	2.78 years

Table 3: Constant Volatility Baseline Simulation

Simulation: Sequential returns are simulated for a period of 2,500,000 days (10,000 years).
Estimation sample: K = 63 days; 39,682 non-overlapping intervals.
True Volatility: $\sigma = 0.20$; True mean: $\mu = 0$; Sample mean is not estimated

All tail statistics are reported in standard deviations.

RMSE of volatility estimate = 0.0178

Probability α	α-Tail cutoff for Normal	C-VaR for Normal	Actual α-tail cutoff	True C-VaR	True Prob / Predicted Prob
0.10	-1.282	-1.747	-1.285	-1.780	1.01
0.05	-1.645	-2.062	-1.649	-2.107	1.01
0.01	-2.326	-2.673	-2.384	-2.770	1.17
0.005	-2.576	-2.898	-2.660	-3.041	1.23
0.002	-2.878	-3.198	-3.015	-3.373	1.46
0.001	-3.090	-3.316	-3.297	-3.599	1.76
0.0005	-3.290	-3.499	-3.514	-3.802	2.03
0.0002	-3.540	-3.716	-3.777	-4.120	2.21
0.0001	-3.719	-4.003	-4.144	-4.317	2.37

Table 4: Tail Cutoffs and Probability Ratios for Different Estimation Samples Sizes

Simulation: Sequential returns are simulated for a period of 2,500,000 days. (10,000,000 days for K = 250)
Estimation sample: Non-overlapping samples of K days, K = 21, 42, 63, 125, 250
True Volatility: $\sigma = 0.20$; True mean: $\mu = 0$; Sample mean is not estimated

All tail statistics are reported in standard deviations.

Panel A: Comparison of Tail Estimates

	α-Tail cutoff for Normal	Actual α-tail cutoff K = 21	Actual α-tail cutoff K = 42	Actual α-tail cutoff K = 63	Actual α-tail cutoff K = 125	Actual α-tail cutoff K = 250
Runs		119047	59523	39682	20000	40000
Forecast RMSE		0.0307	0.0219	0.0178	0.0127	0.0089
Prob= 0.05	-1.645	-1.718	-1.668	-1.649	-1.666	-1.658
0.01	-2.326	-2.534	-2.422	-2.384	-2.333	-2.349
0.001	-3.090	-3.582	-3.251	-3.297	-3.043	-3.129
0.0002	-3.540	-4.257	-3.758	-3.777	-3.706	-3.659
0.0001	-3.719	-4.550	-3.982	-4.144	-3.879	-3.888

Table 4: Tail Cutoffs and Probability Ratios for Different Estimation Samples Sizes, p.2

Panel B: Ratio of True Probability / Predicted Probability

Probability α	Probability ratio K = 21	Probability ratio K = 42	Probability ratio K = 63	Probability ratio K = 125	Probability ratio K = 250
Prob= 0.05	1.15	1.05	1.01	1.04	1.03
0.01	1.51	1.24	1.17	1.03	1.07
0.001	3.03	1.71	1.76	0.85	1.14
0.0002	5.43	1.99	2.21	1.55	1.66
0.0001	6.94	2.54	2.37	1.93	1.64

Table 5: Tail Cutoffs and Probability Ratios for Different θ and κ Values

Simulation: Sequential returns for 2,500,000 days; Estimation sample: 63 days; 39,682 non-overlapping periods
True Volatility: $\sigma = 0.20$; True mean: $\mu = 0$; Sample mean is not estimated
All tail statistics are reported in standard deviations.

Panel A: Comparison of Tail Estimates

	Normal	Baseline	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9
θ (Vol'y of variance)	-	0	0.05	0.10	0.05	0.10	0.10	0.20	0.10	0.20	0.40
κ (mean reversion)	-	0	0.2	0.2	0.4	0.4	1.0	1.0	2.5	2.5	2.5
RMSE	-	0.0178	0.0193	0.0230	0.0192	0.0229	0.0227	0.0331	0.0222	0.0318	0.0545
Prob= 0.05	-1.645	-1.649	-1.654	-1.663	-1.653	-1.659	-1.660	-1.684	-1.658	-1.674	-1.728
0.01	-2.326	-2.384	-2.388	-2.414	-2.385	-2.398	-2.392	-2.504	-2.392	-2.455	-2.702
0.002	-2.878	-3.015	-3.036	-3.107	-3.036	-3.091	-3.065	-3.239	-3.057	-3.183	-3.693
0.001	-3.090	-3.297	-3.295	-3.401	-3.294	-3.383	-3.356	-3.580	-3.324	-3.448	-4.000
0.0005	-3.290	-3.514	-3.538	-3.632	-3.534	-3.596	-3.582	-3.810	-3.587	-3.715	-4.397
0.0002	-3.540	-3.777	-3.789	-3.912	-3.780	-3.770	-3.750	-4.225	-3.776	-3.944	-4.846
0.0001	-3.719	-4.144	-3.928	-4.103	-3.942	-3.933	-3.915	-4.472	-3.957	-4.520	-5.261

Table 5: Tail Cutoffs and Probability Ratios for Different θ and κ Values, p.2

Panel B: Ratio of True Probability / Predicted Probability

	Baseline	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9
θ (Vol'y of variance)	0	0.05	0.10	0.05	0.10	0.10	0.20	0.10	0.20	0.40
κ (mean reversion)	0	0.2	0.2	0.4	0.4	1.0	1.0	2.5	2.5	2.5
Prob= 0.05	1.01	1.02	1.03	1.02	1.03	1.03	1.07	1.02	1.05	1.15
0.01	1.17	1.14	1.26	1.15	1.21	1.19	1.43	1.19	1.32	1.86
0.002	1.46	1.55	1.82	1.56	1.72	1.69	2.43	1.66	2.07	3.76
0.001	1.76	1.70	2.12	1.70	2.03	1.88	2.95	1.84	2.34	5.56
0.0005	2.03	2.02	2.69	2.05	2.36	2.18	3.59	2.17	3.12	7.57
0.0002	2.21	2.39	3.50	2.39	3.27	3.06	5.36	2.70	4.05	12.65
0.0001	2.37	2.45	3.35	2.40	2.59	2.95	7.07	2.70	4.94	19.06

Table 6: Tail Cutoffs and Probability Ratios under Student-t Return Shocks

Simulation: Sequential returns for 2,500,000 days; Estimation sample: 63 days; 39,682 observations
Return shocks are drawn from a Normal (0,1) and Student-t distributions with 7 and 4 degrees of freedom.
True Volatility: σ = 0.20; True mean: μ = 0; Sample mean is not estimated
All tail statistics are reported in standard deviations.

Panel A: Comparison of Tail Estimates

	Normal	Baseline	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8
θ	-	0	0	0	0.05	0.05	0.05	0.20	0.20	0.20
κ	-	0	0	0	0.40	0.40	0.40	2.5	2.5	2.5
Shocks	-	N(0,1)	t(7)	t(4)	N(0,1)	t(7)	t(4)	N(0,1)	t(7)	t(4)
RMSE	-	0.0178	0.0243	0.0389	0.0192	0.0254	0.0396	0.0318	0.0360	0.0473
Prob= 0.05	-1.645	-1.649	-1.627	-1.580	-1.653	-1.633	-1.584	-1.674	-1.641	-1.595
0.01	-2.326	-2.384	-2.615	-2.794	-2.385	-2.615	-2.802	-2.455	-2.710	-2.915
0.002	-2.878	-3.015	-3.812	-4.712	-3.036	-3.855	-4.758	-3.183	-3.882	-4.778
0.001	-3.090	-3.297	-4.362	-5.621	-3.294	-4.275	-5.539	-3.448	-4.332	-5.438
0.0005	-3.290	-3.514	-4.683	-6.515	-3.534	-4.747	-6.484	-3.715	-4.949	-6.499
0.0002	-3.540	-3.777	-5.499	-8.401	-3.780	-5.704	-8.144	-3.944	-5.884	-8.407
0.0001	-3.719	-4.144	-6.774	-11.917	-3.942	-7.133	-12.548	-4.520	-7.324	-12.680

Table 6: Tail Cutoffs and Probability Ratios under Student-t Return Shocks, p.2

Panel B: Ratio of True Probability / Predicted Probability

	Baseline	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8
θ	0	0	0	0.05	0.05	0.05	0.20	0.20	0.20
κ	0	0	0	0.40	0.40	0.40	2.5	2.5	2.5
Shocks	N(0,1)	t(7)	t(4)	N(0,1)	t(7)	t(4)	N(0,1)	t(7)	t(4)
Prob= 0.05	1.01	0.98	0.91	1.02	0.98	0.92	1.05	0.99	0.93
0.01	1.17	1.61	1.83	1.15	1.61	1.80	1.32	1.71	1.88
0.002	1.46	3.43	4.64	1.56	3.54	4.74	2.07	4.08	5.13
0.001	1.76	5.39	7.68	1.70	5.56	7.63	2.34	5.93	8.55
0.0005	2.03	8.49	12.75	2.05	8.61	12.94	3.12	9.07	14.02
0.0002	2.21	14.72	26.84	2.39	14.52	26.43	4.05	15.78	27.20
0.0001	2.37	23.21	45.79	2.40	22.67	46.21	4.94	25.52	48.02

Table 7: EWMA Volatility Tail Cutoffs and Probability Ratios

Simulation: Sequential returns for 2,500,000 days; Estimation sample: 63 days; 39,682 observations
Volatility is calculated using an exponentially weighted moving average with decay factors D = 0.97, 0.94, 0.90
True Volatility: $\sigma = 0.20$; True mean: $\mu = 0$; Sample mean is not estimated
All tail statistics are reported in standard deviations.

Panel A: Comparison of Tail Estimates

	Normal	Run 1 Baseline	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10	Run 11	Run 12
θ	-	0	0	0	0	0.05	0.05	0.05	0.05	0.25	0.25	0.25	0.25
κ	-	0	0	0	0	1	1	1	1	1	1	1	1
Decay	-	1.0	0.97	0.94	0.90	1.0	0.97	0.94	0.90	1.0	0.97	0.94	0.90
RMSE	-	0.0178	0.0202	0.0251	0.0319	0.0191	0.0210	0.0256	0.0322	0.0390	0.0345	0.0340	0.0372
Prob= 0.05	-1.645	-1.649	-1.659	-1.676	-1.706	-1.651	-1.658	-1.676	-1.708	-1.705	-1.677	-1.685	-1.714
0.01	-2.054	-2.097	-2.113	-2.146	-2.198	-2.093	-2.115	-2.136	-2.194	-2.195	-2.164	-2.176	-2.220
0.002	-2.326	-2.384	-2.402	-2.444	-2.519	-2.379	-2.400	-2.442	-2.526	-2.602	-2.533	-2.513	-2.586
0.001	-2.576	-2.660	-2.686	-2.746	-2.817	-2.683	-2.706	-2.733	-2.823	-2.944	-2.880	-2.860	-2.889
0.0005	-2.878	-3.015	-3.024	-3.098	-3.164	-3.029	-3.043	-3.101	-3.163	-3.463	-3.303	-3.245	-3.320
0.0002	-3.090	-3.297	-3.317	-3.348	-3.437	-3.298	-3.306	-3.319	-3.444	-3.748	-3.589	-3.581	-3.625
0.0001	-3.290	-3.514	-3.529	-3.632	-3.809	-3.553	-3.538	-3.688	-3.817	-4.023	-3.799	-3.939	-4.214

Table 7: EWMA Volatility Tail Cutoffs and Probability Ratios, p.2

Panel B: Ratio of True Probability / Predicted Probability

	Run 1 Baseline	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10	Run 11	Run 12
θ	0	0	0	0	0.05	0.05	0.05	0.05	0.25	0.25	0.25	0.25
κ	0	0	0	0	1	1	1	1	1	1	1	1
Decay	1.0	0.97	0.94	0.90	1.0	0.97	0.94	0.90	1.0	0.97	0.94	0.90
Prob= 0.05	1.01	1.03	1.06	1.14	1.01	1.02	1.06	1.14	1.11	1.06	1.09	1.14
0.01	1.11	1.14	1.22	1.32	1.10	1.13	1.21	1.33	1.31	1.23	1.25	1.37
0.002	1.17	1.20	1.33	1.52	1.15	1.21	1.32	1.54	1.56	1.45	1.51	1.64
0.001	1.23	1.25	1.46	1.82	1.26	1.27	1.45	1.81	2.11	1.83	1.76	2.05
0.0005	1.46	1.64	1.75	2.18	1.57	1.57	1.76	2.12	3.03	2.52	2.42	2.56
0.0002	1.76	1.81	2.04	2.65	1.73	1.81	2.04	2.38	3.92	3.19	2.74	2.97
0.0001	2.03	2.16	2.18	2.81	2.02	2.13	2.13	2.75	5.45	4.08	3.83	4.25

Table 8: Tail Cutoffs and Probability Ratios for a 63-Day Rolling Sample with Different θ and κ Values

Simulation: Sequential simulated returns for 250,000 days; Estimation sample: 63-day rolling sample
True Volatility: σ = 0.20; True mean: μ = 0; Sample mean is not estimated
All tail statistics are reported in standard deviations.

Panel A: Comparison of Tail Estimates

	Normal	Baseline	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9
θ (Vol'y of variance)	-	0	0.05	0.10	0.05	0.10	0.10	0.20	0.10	0.20	0.40
κ (mean reversion)	-	0	0.2	0.2	0.4	0.4	1.0	1.0	2.5	2.5	2.5
RMSE	-	0.0177	0.0188	0.0233	0.0192	0.0223	0.0224	0.0328	0.0219	0.0313	0.0540
Prob= 0.05	-1.645	-1.660	-1.802	-2.169	-1.704	-1.987	-1.721	-2.039	-1.699	-1.797	-2.232
0.01	-2.326	-2.384	-2.707	-3.889	-2.513	-3.199	-2.560	-3.390	-2.248	-2.720	-3.890
0.002	-2.878	-2.998	-3.523	-6.482	-0.201	-4.504	-3.315	-4.967	-3.142	-3.621	-5.737
0.001	-3.090	-3.200	-3.907	-7.924	-3.483	-5.133	-3.909	-5.710	-3.388	-4.054	-6.686
0.0005	-3.290	-3.458	-4.306	-9.454	-3.749	-3.770	-4.282	-6.501	-3.622	-4.385	-7.499
0.0002	-3.540	-3.719	-4.786	-12.070	-4.108	-6.687	-4.282	-7.884	-3.969	-4.871	-8.884
0.0001	-3.719	-3.949	-5.102	-14.445	-4.925	-7.125	-4.662	-8.760	-4.223	-5.310	-10.283

Table 8: Tail Cutoffs and Probability Ratios for a 63-Day Rolling Sample, p.2

Panel B: Ratio of True Probability / Predicted Probability

		Baseline	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9
θ (Vol'y of variance)		0	0.05	0.10	0.05	0.10	0.10	0.20	0.10	0.20	0.40
κ (mean reversion)		0	0.2	0.2	0.4	0.4	1.0	1.0	2.5	2.5	2.5
Prob= 0.05	No overlap	1.01	1.02	1.03	1.02	1.03	1.03	1.07	1.02	1.05	1.15
	overlap	1.03	1.30	1.80	1.12	1.58	1.15	1.64	1.09	1.30	1.88
0.01	No overlap	1.17	1.14	1.26	1.15	1.21	1.19	1.43	1.19	1.32	1.86
	overlap	1.16	2.03	4.25	1.49	3.17	1.58	3.52	1.40	2.00	4.52
0.002	No overlap	1.46	1.55	1.82	1.56	1.72	1.69	2.43	1.66	2.07	3.76
	overlap	1.41	3.62	12.09	2.28	7.56	2.56	8.98	1.84	3.86	12.95
0.001	No overlap	1.76	1.70	2.12	1.70	2.03	1.88	2.95	1.84	2.34	5.56
	overlap	1.50	4.74	19.99	2.66	11.41	3.20	14.01	2.26	5.08	20.96
0.0005	No overlap	2.03	2.02	2.69	2.05	2.36	2.18	3.59	2.17	3.12	7.57
	overlap	1.54	6.57	33.17	3.22	17.84	4.12	22.51	2.59	7.13	34.46
0.0002	No overlap	2.21	2.39	3.50	2.39	3.27	3.06	5.36	2.70	4.05	12.65
	overlap	1.90	9.74	66.48	4.48	32.19	6.06	42.61	3.04	11.54	68.40
0.0001	No overlap	2.37	2.45	3.35	2.40	2.59	2.95	7.07	2.70	4.94	19.06
	overlap	1.32	10.64	101.87	3.68	44.01	6.04	60.02	2.60	13.44	101.95

Notes: "No overlap" lines duplicate results from Table 5, for comparison.

Table 9: Realized Tail Events for the Standard and Poor's 500 Index

Simulation: Historical sample of returns on the S&P 500 Index July 2, 1992 - August 30, 2002 (10,113 days)
Estimation sample: 63 day rolling sample; Sample mean is not estimated

Sample		Tail Probability								
		.05	.02	.01	.005	.002	.001	.0005	.0002	.0001
21-day	Events predicted	505	202	101	50	20	10	5	2	1
	Actual events	587	304	195	128	85	65	50	42	30
	Probability ratio	1.16	1.51	1.93	2.54	4.21	6.44	9.91	20.81	29.73
63-day	Events predicted	502	201	100	50	20	10	5	2	1
	Actual events	548	283	184	130	75	59	43	30	22
	Probability ratio	1.09	1.41	1.83	2.59	3.73	5.87	8.56	14.93	21.89
250-day	Events predicted	493	197	99	49	20	10	5	2	1
	Actual events	507	256	171	114	77	56	43	29	22
	Probability ratio	1.03	1.30	1.73	2.31	3.90	5.68	8.72	14.70	22.31

Table 10: Probability Ratios for a 63-Day Rolling Sample with Different θ and κ Values

Simulation: Sequential simulated returns for 250,000 days; Historical sample 1992 - 2002 for S&P 500 Index (10,113 days);
Estimation sample: 63 day rolling sample
True Volatility: σ = 0.20; True mean: μ = 0; Sample mean is not estimated

Panel A: Two Events in Two Days

	Theoretical Probability	Ratio of Realized Probability to Theoretical Probability					
	Normal	Baseline	Run 1	Run 2	Run 3	Run 4	S&P500
θ (Vol'y of variance)	-	0	0.10	0.10	0.20	0.40	-
κ (mean reversion)	-	0	0.2	1.0	1.0	2.50	-
Prob= 0.10	0.01	1.04	2.61	1.32	2.28	2.79	1.65
0.05	0.0025	1.12	5.66	1.82	4.68	6.03	3.03
0.02	0.00040	1.39	19.54	3.10	14.10	20.11	6.72
0.01	0.00010	1.68	54.65	4.56	35.21	55.13	16.92
0.005	0.000025	2.56	160.36	8.48	90.90	152.20	47.77
0.002	0.000004	2.00	690.18	22.01	360.09	631.16	74.64

Panel B: Three Events in Three Days

	Theoretical Probability	Ratio of Realized Probability to Theoretical Probability					
	Normal	Baseline	Run 1	Run 2	Run 3	Run 4	S&P500
θ (Vol'y of variance)	-	0	0.10	0.10	0.20	0.40	-
κ (mean reversion)	-	0	0.2	1.0	1.0	2.50	-
Prob= 0.10	0.001	1.08	5.75	1.81	4.67	6.44	3.19
0.05	0.0001	1.25	22.89	3.52	17.09	24.39	9.56
0.02	0.000008	3.00	184.55	11.00	120.03	173.55	24.88
0.01	0.000001	12.00	980.25	36.01	576.15	908.24	=
0.005	1.25E-7	32.01	5665.5	160.0	2944.8	4641.2	=
0.002	8.0E-9	=	59515.5	1000.3	26506.9	42010.9	=

Table 10: Probability Ratios for a 63-Day Rolling Sample, continued

Panel C: Three Events in Five Days

	Theoretical Probability	Ratio of Realized Probability to Theoretical Probability					
	Normal	Baseline	Run 1	Run 2	Run 3	Run 4	S&P500
θ (Vol'y of variance)	-	0	0.10	0.10	0.20	0.40	-
κ (mean reversion)	-	0	0.2	1.0	1.0	2.50	-
Prob= 0.10	0.009	1.06	4.61	1.68	3.79	4.93	2.34
0.05	0.0012	1.22	17.24	3.26	12.93	17.83	7.13
0.02	0.00008	1.75	132.53	9.23	86.86	130.93	20.52
0.01	0.000010	6.90	716.1	27.21	388.7	655.2	80.85
0.005	1.24E-06	16.12	3873.2	77.4	1947.9	3354.0	481.5
0.002	7.98E-08	=	41886.8	351.1	18159.3	34111.4	2496.3

Panel D: Three Events in Ten Days

	Theoretical Probability	Ratio of Realized Probability to Theoretical Probability					
	Normal	Baseline	Run 1	Run 2	Run 3	Run 4	S&P500
θ (Vol'y of variance)	-	0	0.10	0.10	0.20	0.40	-
κ (mean reversion)	-	0	0.2	1.0	1.0	2.50	-
Prob= 0.10	0.070	1.03	2.70	1.39	2.39	2.88	1.39
0.05	0.0115	1.10	8.97	2.50	7.33	9.67	3.58
0.02	0.00086	1.37	64.72	7.34	46.02	69.30	14.07
0.01	0.000114	2.60	341.4	19.40	214.7	339.6	52.49
0.005	1.46E-05	5.20	1895.0	62.7	1080.9	1840.2	231.8
0.002	9.50E-07	=	20082.2	290.6	10171.7	18384.9	838.8

Notes: The symbol = in a cell indicates that there were no events observed in the sample and that the theoretically expected number of events in a sample of this size was also 0.

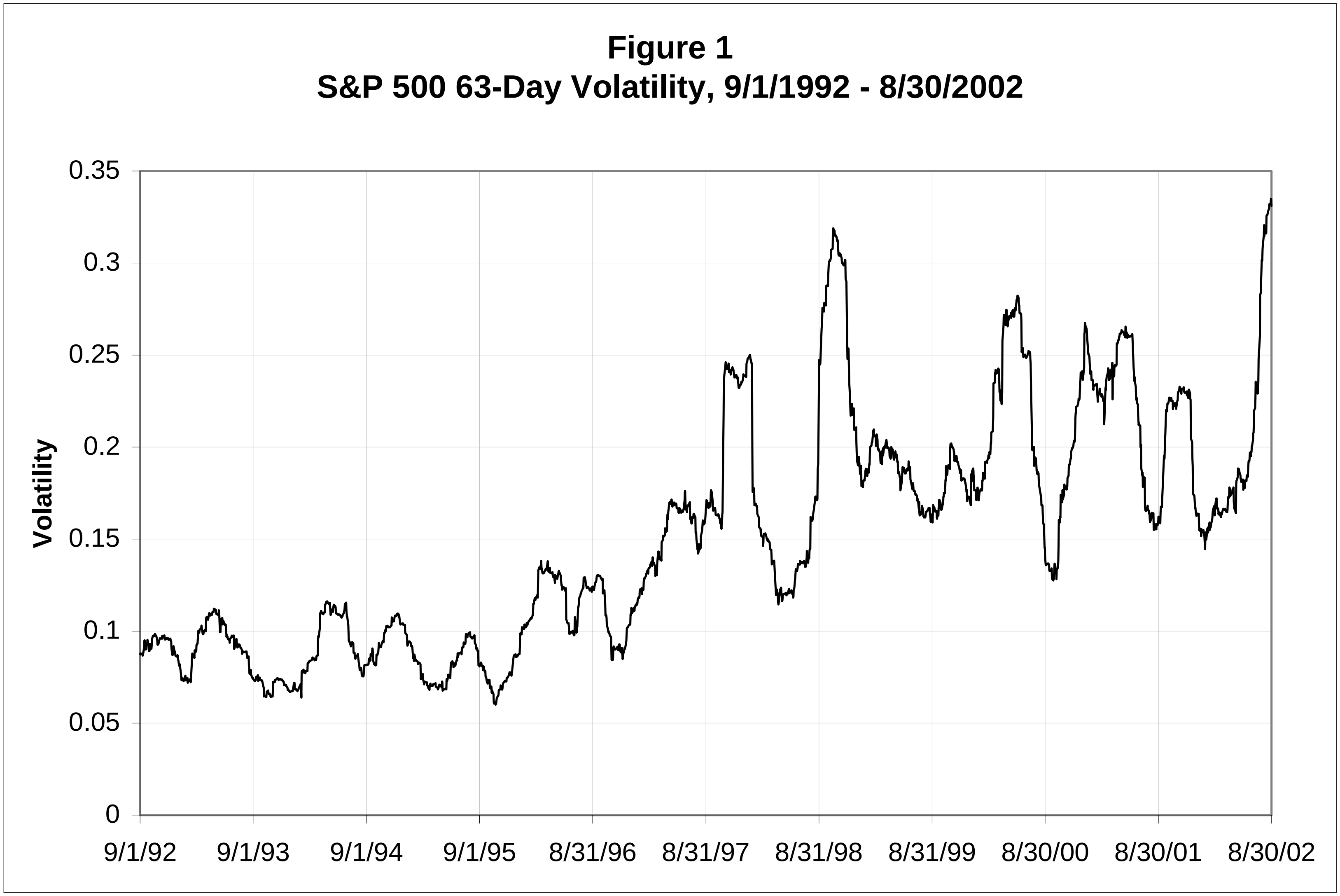


Figure 1
S&P 500 63-Day Volatility, 9/1/1992 - 8/30/2002
Volatility
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
9/1/92
9/1/93
9/1/94
9/1/95
8/31/96
8/31/97
8/31/98
8/31/99
8/30/00
8/30/01
8/30/02

Figure 2
Value at Risk and Conditional Value at Risk

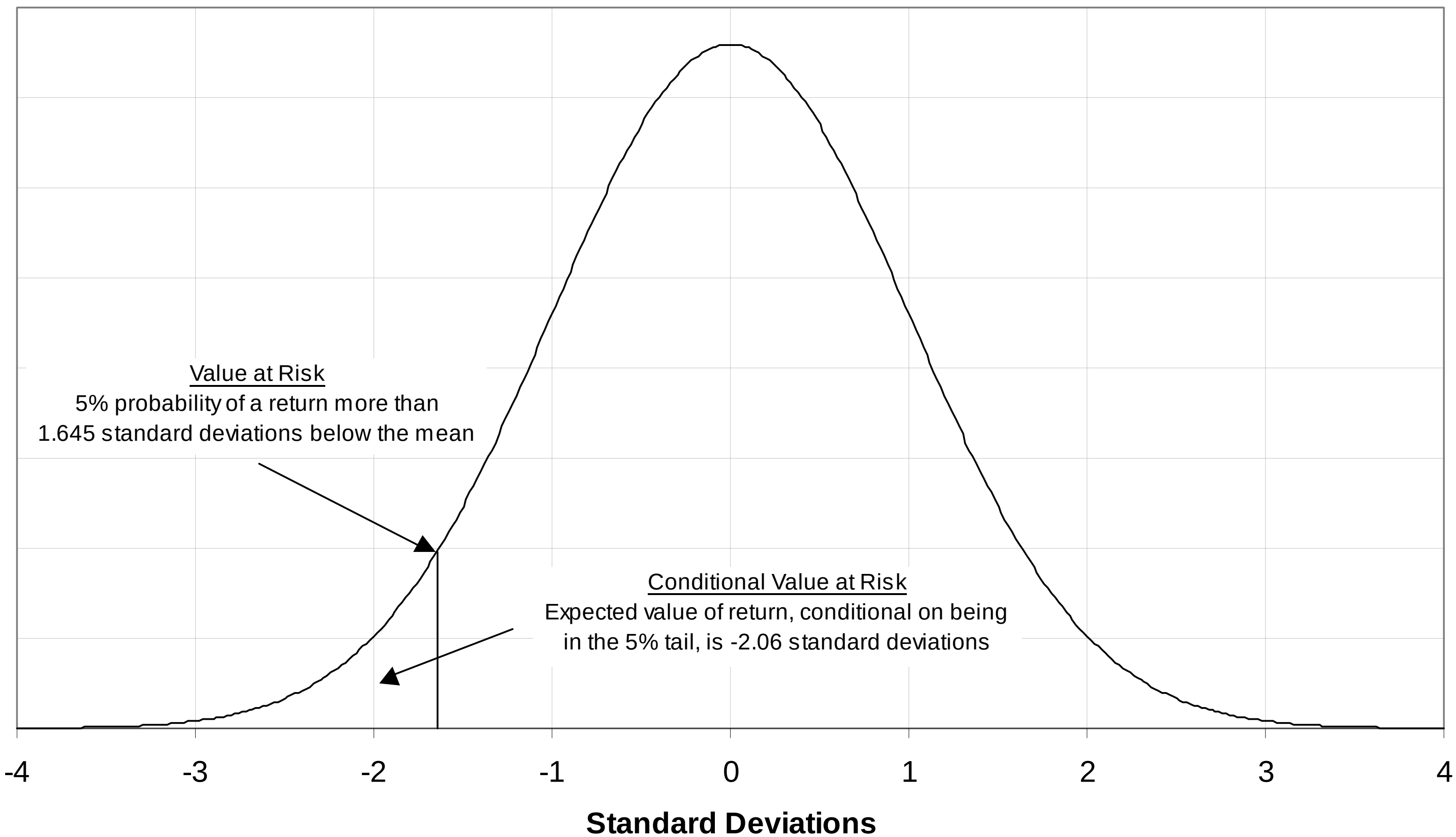


Figure 3
Left Tails of Predicted and True Returns Distributions

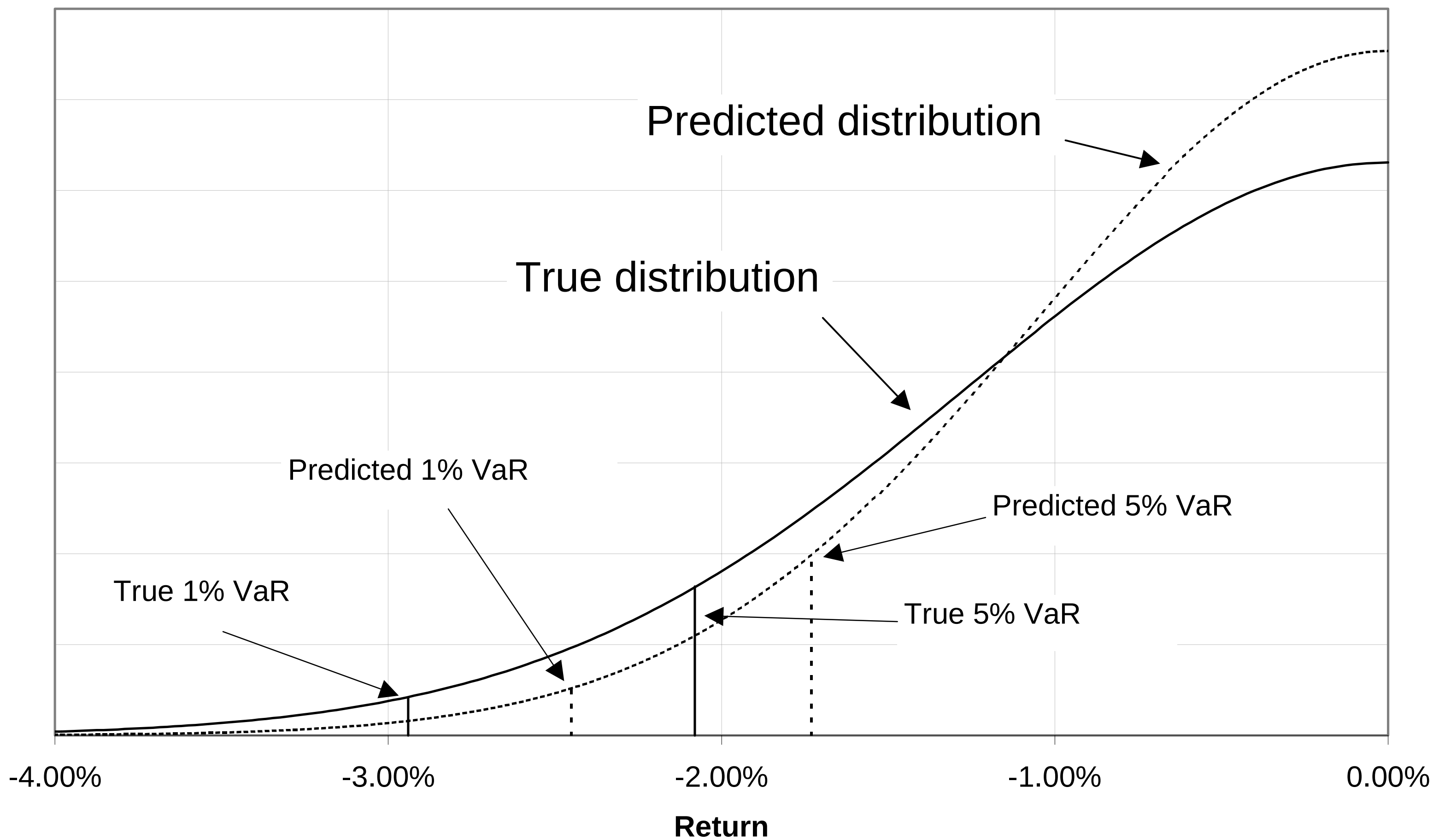
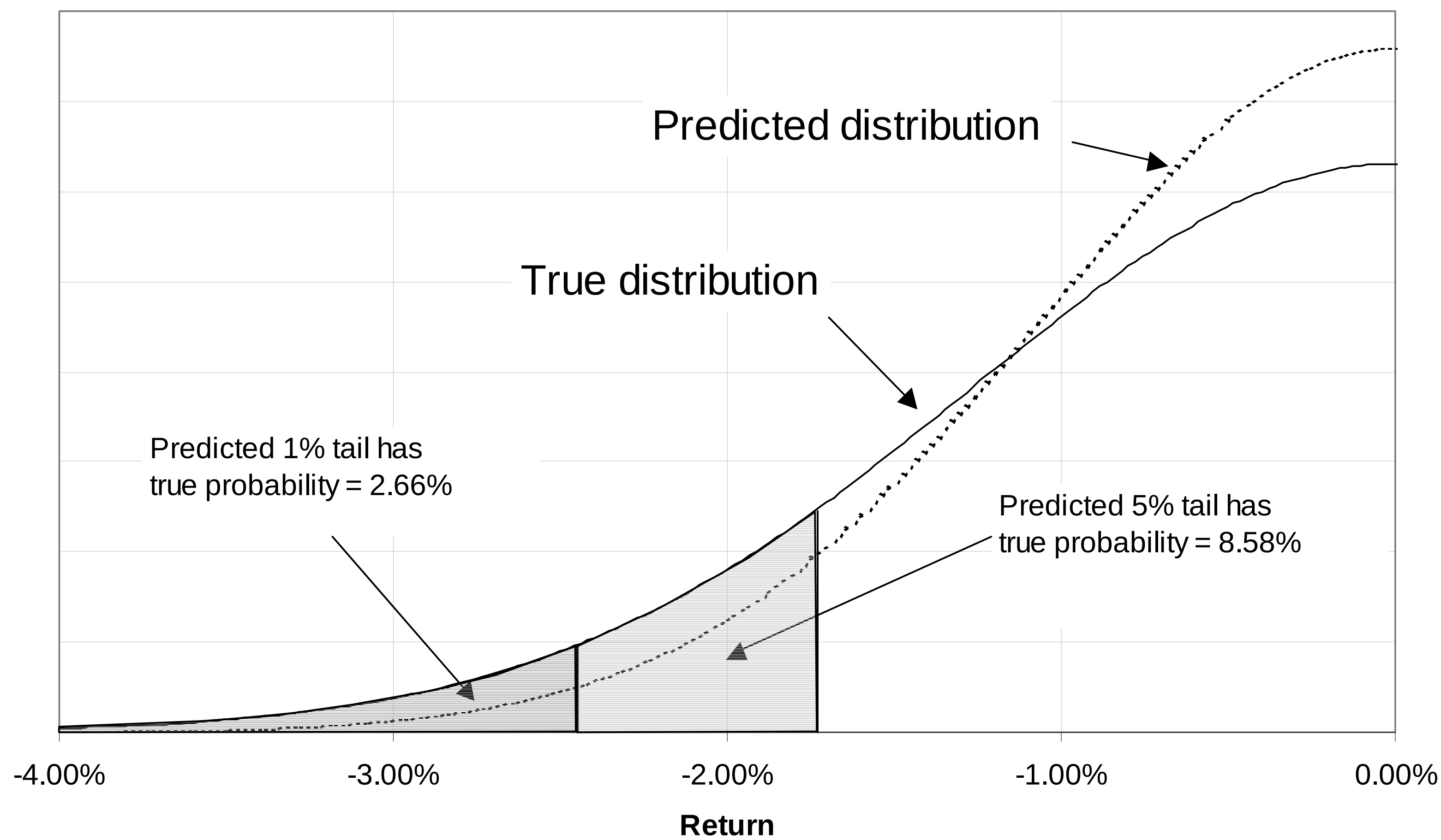


Figure 4
Predicted and True Tail Probabilities



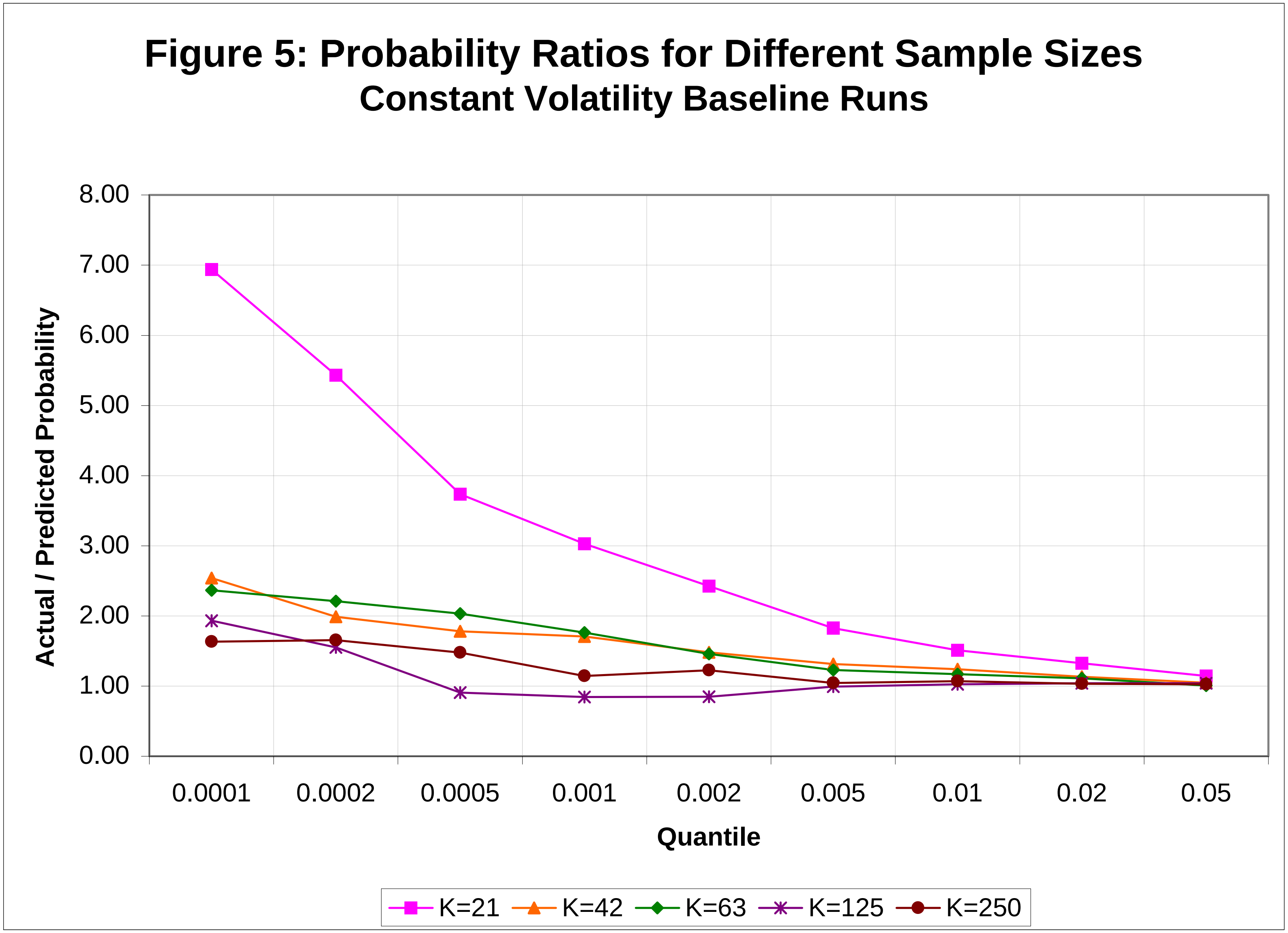


Figure 5: Probability Ratios for Different Sample Sizes
Constant Volatility Baseline Runs
Actual / Predicted Probability
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
0.0001
0.0002
0.0005
0.001
0.002
0.005
0.01
0.02
0.05
Quantile
K=21
K=42
K=63
K=125
K=250

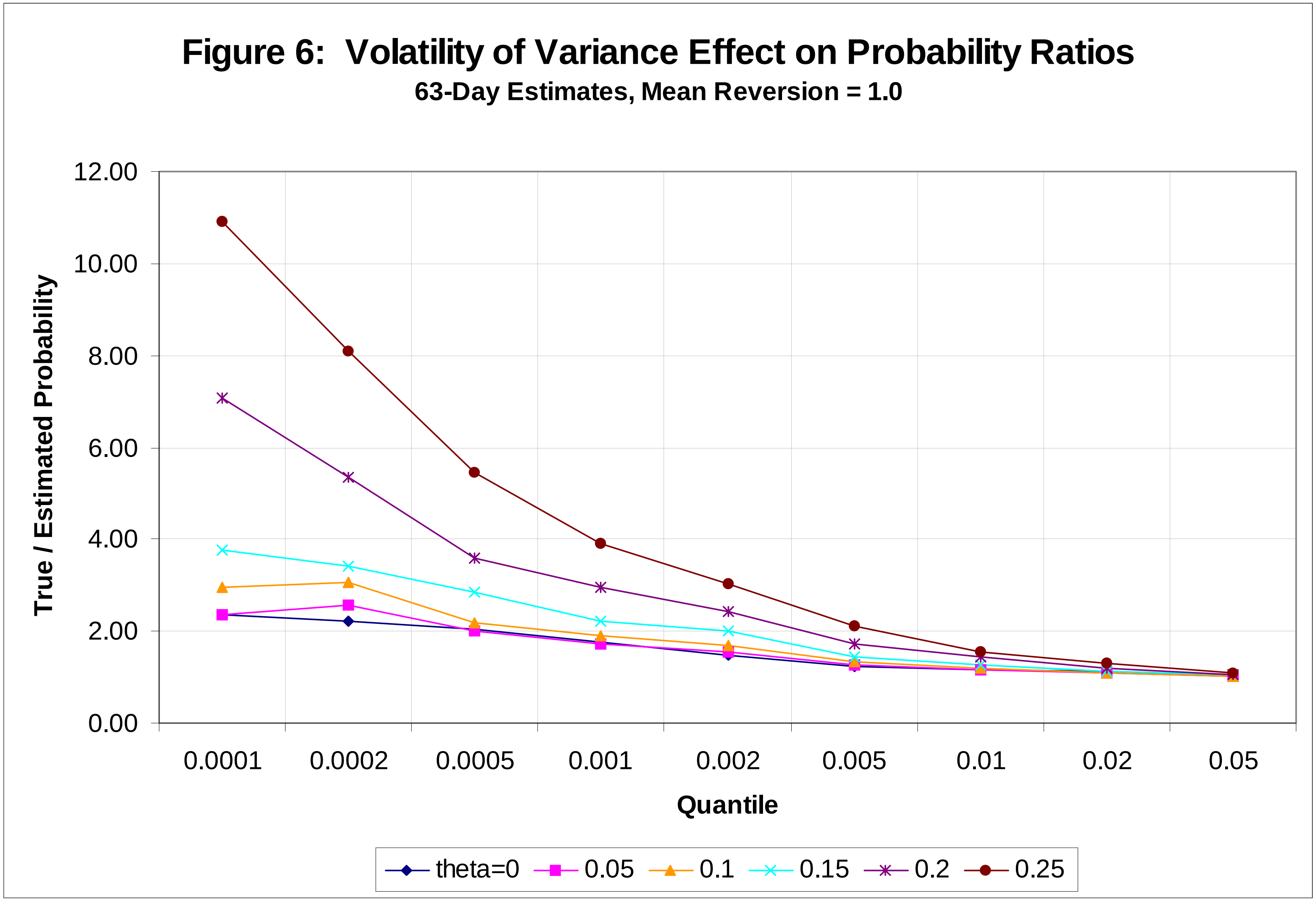


Figure 6: Volatility of Variance Effect on Probability Ratios
63-Day Estimates, Mean Reversion = 1.0
True / Estimated Probability
12.00
10.00
8.00
6.00
4.00
2.00
0.00
0.0001
0.0002
0.0005
0.001
0.002
0.005
0.01
0.02
0.05
Quantile
theta=0
0.05
0.1
0.15
0.2
0.25

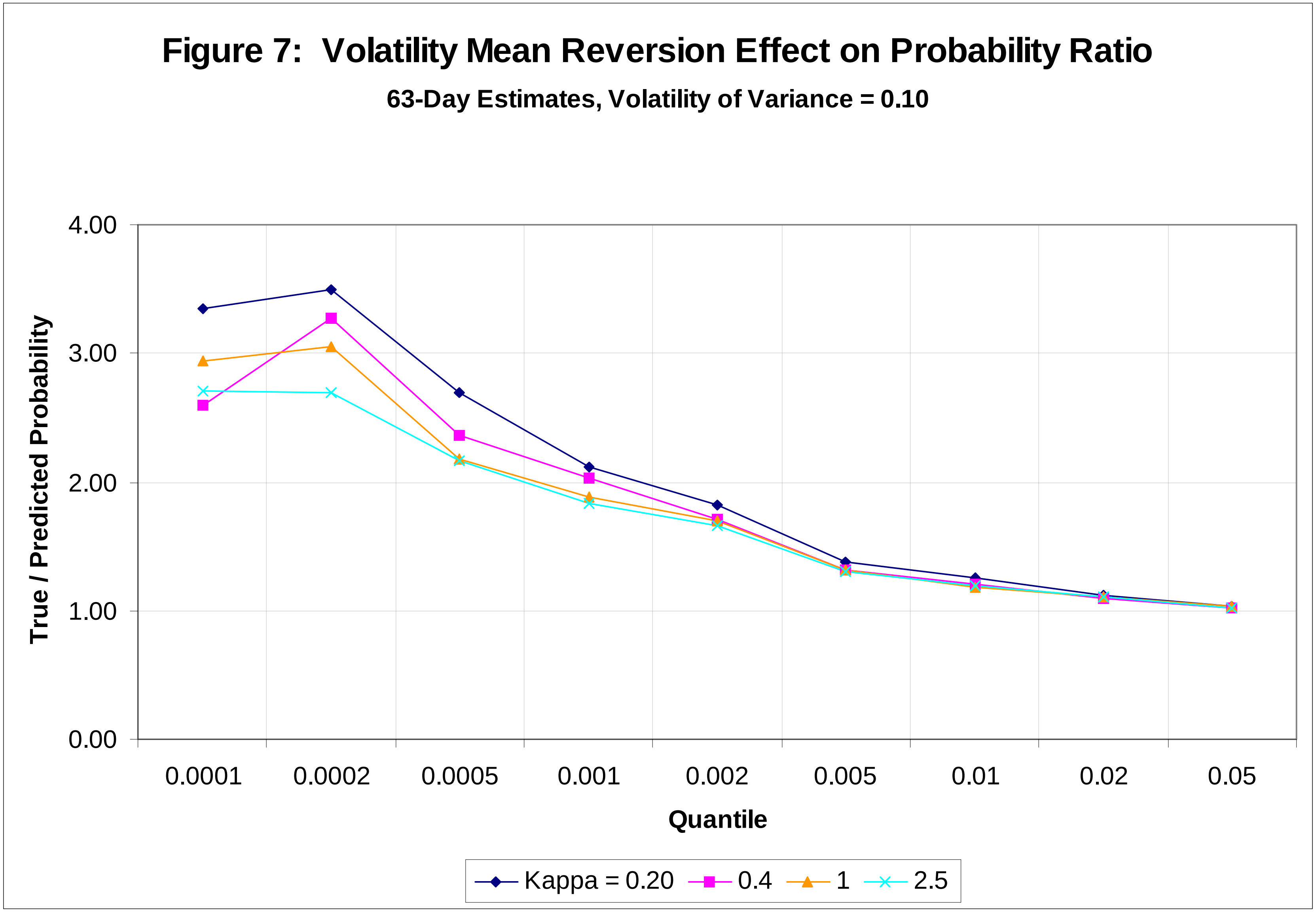


Figure 7: Volatility Mean Reversion Effect on Probability Ratio
63-Day Estimates, Volatility of Variance = 0.10
True / Predicted Probability
4.00
3.00
2.00
1.00
0.00
0.0001
0.0002
0.0005
0.001
0.002
0.005
0.01
0.02
0.05
Quantile
Kappa = 0.20
0.4
1
2.5

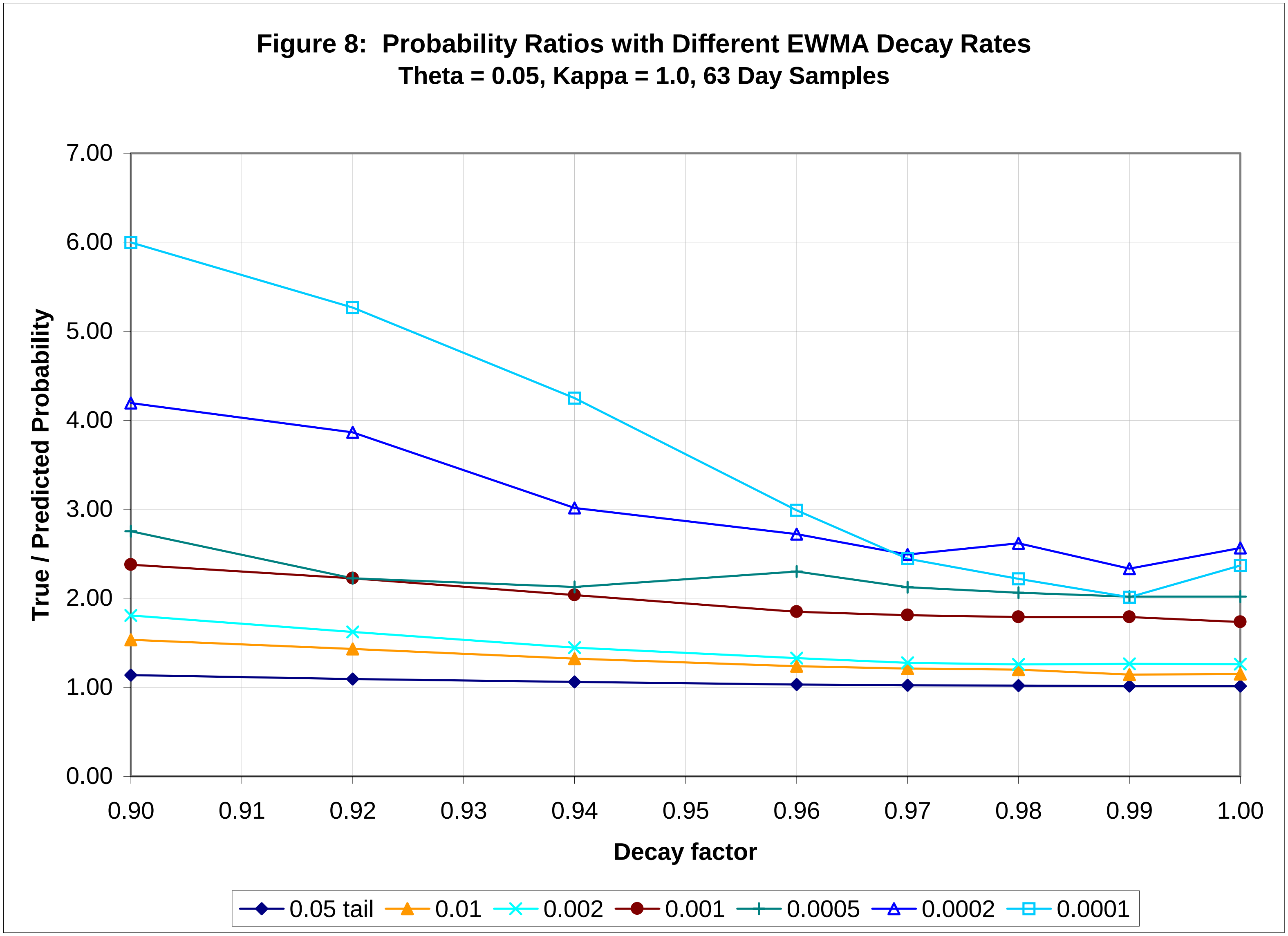


Figure 8: Probability Ratios with Different EWMA Decay Rates
Theta = 0.05, Kappa = 1.0, 63 Day Samples
True / Predicted Probability
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
0.90
0.91
0.92
0.93
0.94
0.95
0.96
0.97
0.98
0.99
1.00
Decay factor
0.05 tail
0.01
0.002
0.001
0.0005
0.0002
0.0001

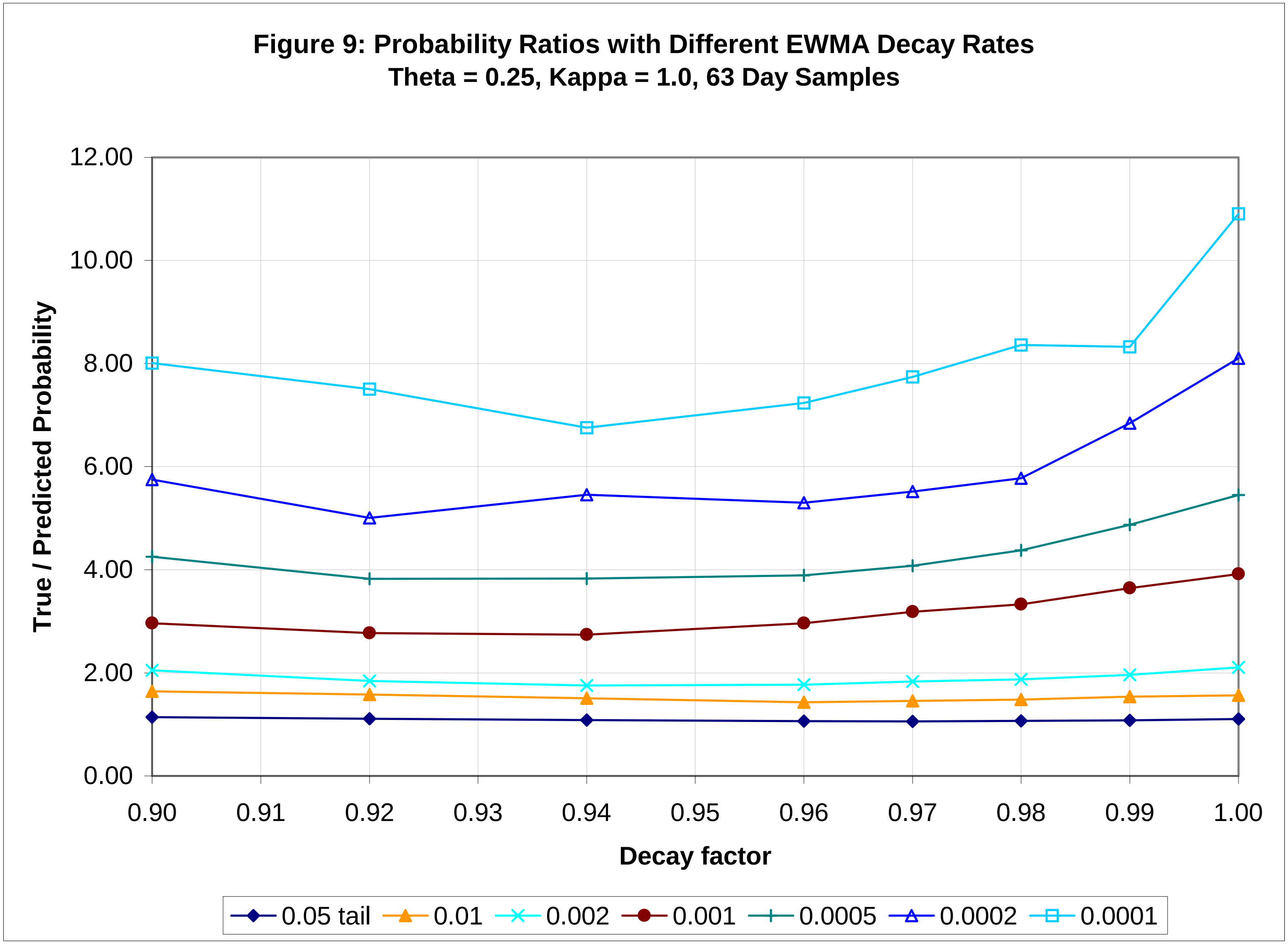


Figure 9: Probability Ratios with Different EWMA Decay Rates
Theta = 0.25, Kappa = 1.0, 63 Day Samples
True / Predicted Probability
12.00
10.00
8.00
6.00
4.00
2.00
0.00
0.90
0.91
0.92
0.93
0.94
0.95
0.96
0.97
0.98
0.99
1.00
Decay factor
0.05 tail
0.01
0.002
0.001
0.0005
0.0002
0.0001