Testing the Distributional Assumptions Embedded in Risk Models

Jeremy Berkowitz
University of California, Irvine

January 14, 2002

Address correspondence to:

Jeremy Berkowitz
Graduate School of Management
University of California-Irvine
Irvine, CA 92697-3125
jberkowitz@gsm.uci.edu

Acknowledgements: I gratefully acknowledge helpful input from Jim O'Brien and Matt Pritsker. Any remaining errors and inaccuracies are mine.

1. Introduction

The explosion of research into modeling and forecasting portfolio value-at-risk is comprised nearly entirely of highly parameterized factor models (see, for example, Jorion (2001)). In order to use such VaR models for forecasting, practitioners must take a stand on the distribution of the factors. This is because unlike many other areas of Applied Finance, the output of interest is more than just a point forecast. VaR is an interval forecast, typically a one-sided 95 or 99 percent interval. Users generally assume that the factors are Normally or t-distributed or use some variant of historical simulation.[1] It is useful for our purposes to view historical simulation as imposing the empirical distribution -- the histogram of observed historical returns.

The choice of whether to impose a parametric distribution or historical simulation can have important influence on the performance and forecast accuracy of the model. Lopez and Walter (2000) provide evidence that forecasts of portfolio distributions are particularly sensitive to distributional assumptions – indeed, far more so than other design issues that have typically received more attention, such as the method used to calculate covariance matrices. Related Monte Carlo evidence in given in Pritsker (1997) and Berkowitz (2001). Both papers find significant differences between models that assume Normality versus those based on some form of historical simulation.

Further evidence of how contentious the choice of distribution has become can be found in recent reports that practitioners have not converged on a single methodology. A survey by Britain's Financial Services Authority reveals that, in estimating their market risk models, 42 percent of banks use a variance-covariance approach, 31 percent use historical simulation and 23 percent use Monte Carlo methods. Berkowitz and O'Brien (2001) report that 2 of 6 large commercial banks in their study use historical simulation. The remaining banks imposed distributional assumptions on the factor prices.

Since historical simulation permits forecasting of value-at-risk without imposing parametric distributional assumptions, why do anything else? In some cases, historical simulation is a reasonable modeling approach. However, implementation of historical

1 Within this category there are also models comprised of a mixture of Normals such as conditionally Normal GARCH models.

simulation can be subject to pitfalls that do not arise under MC. Historical simulation, like bootstrapping, involves resampling from historically observed changes in factor prices. Such financial time series invariably display autocorrelation – typically in the mean, as well as in the volatility and possibly in higher moments. Dependence makes resampling with replacement invalid because the resampling procedure fails to preserve the underlying dependence structure.

It is undoubtedly the case that nonstationarity (structural breaks of any kind) can be a problem in financial time series. If so, this can make less recent data irrelevant and invalidate historical simulation.

Monte Carlo simulation bypasses both such problems by allowing the researcher to posit a data-generating process. Rather than try to transform the historical factor changes to make them suitable for resampling, under Monte Carlo methods one simply generates data from an assumed process like a GARCH that has built-in dependence of the kind believed important.

Lastly, in recent work, Pritsker (2001) emphasizes that VaR forecasts from HS can be insufficiently responsive to big changes in volatility. This dovetails with Berkowitz and O'Brien (2002) who find that, although historical simulation methods do well, GARCH-based methods do even better.

Despite the significance of these issues, standard statistical methods are not well-suited for testing distributional assumptions. The problem is that no model is literally true and thus *any* VaR models is typically rejected by standard statistical tests. VaR models contain many potentially restrictive modeling assumptions and shortcuts.[2] A simple rejection leaves us with little or no information on the accuracy of the distributional assumption in isolation. Is it the distributional assumption which is inaccurate or have we chosen the wrong set of factors?

In this article, we describe a procedure that has power to isolate rejections arising from distributional assumptions. The key is to generate data that is consistent with all aspects of the model other than the distributional assumption. A rejection of the model using this artificial data, by construction implies a rejection of the distributional assumptions. We go beyond a simple reject/not reject conclusion and give constructive guidance to researchers as to *why* a particular model fails (at least in some dimensions).

---

2 A well-known problem with CAPM equity pricing, for example, is that the factor loadings must be estimated and are therefore subject to measurement error.

The remainder of the paper is organized as follows. Section 2 formalizes the notions of factors and distributional assumptions within the context of a risk model. Section 3 presents the suggested testing procedures. Section 4 discusses the results of some simulation experiments and section 5 concludes.

2. The Basic Framework

Consider a firm that maintains a risk model which is used to forecast the distribution of possible returns, y_{t+1}, on some portfolio. Denote the distribution of returns, $g(y_{t+1})$. For example, quantile or Value-at-Risk modeling reduces to estimating a single percentile, say the 99^{th}, of $g(y_{t+1})$.

Such risk models are composed of two parts. First, the model contains a set of risk factors, such as interest rates and exchange rates. Let x_t be the kx1 vector of factor returns realized at time t.

The second component of the model is a set of *pricing rules*, $P(\cdot)$, which predict asset returns as a function of the underlying factors. Typically, this will be a simple linear mapping from the space of factors to assets,

$$(1) \qquad \hat{y}_{j,t} = \lambda_j' x_t$$

where $\hat{y}_{j,t}$ is the predicted return of asset j, λ_j is the associated k vector of factor loadings, and x_t is a k-vector of factor changes. We might use more complex functions $\hat{y}_{j,t} = p_j(x_t)$ to accommodate assets with nonlinear dependence such as pricing options via Black-Scholes. The predicted portfolio return is then given by

$$(2) \qquad \hat{y}_t = \sum_j p_j(x_t)\, w_j$$

where w_j is the weight of asset j in the portfolio at time t.

In many cases interest centers on the VaR or distribution of possible portfolio returns. If so, users assume a distribution for the underlying factors f(x). From this distribution artificial factor price changes are generated and then converted into simulated portfolio returns. For example, factor prices changes are commonly assumed to follow a Normal or conditionally Normal distribution such as GARCH. If we denote a draw from $f(\cdot)$ as $\hat{x}_f^i$, then a simulated value can be written $\hat{y}^{i}_{t+1} = \sum p_j(\hat{x}^{i}_{f})\, w_j$. This process is repeated many times to build up a set $\{\hat{y}^{1}_{t+1},...,\hat{y}^{R}_{t+1}\}$ where R is a large number chosen by the user. This set of pseudo values is tabulated into an estimate of the distribution.

3. Testing the Factor Distributions

In this section we suggest a framework for diagnosing the source of model rejections. The trick is to create a test in which the same pricing model appears in both null and alternative. The alternative hypothesis is different from the null only in the assumptions regarding factor distributions.

Under the null $f(\cdot)$ is some parametric distribution such as the Normal. The alternative hypothesis is based on the empirical distribution. If portfolio returns under H_0 are very different than under H_a, we conclude that the parametric assumptions are statistically restrictive.

Write the true but unknown density of portfolio returns as $g[y_t|\, f, p]$ to emphasize that it depends on the factor distributions, $f(\cdot)$, and on the mapping between assets and the factors, $P(\cdot)$. The risk model to be tested is $g[y_t|\, \hat{f}, \hat{p}]$, where $\hat{f}(\cdot)$ is a distributional assumption and $\hat{p}(\cdot)$ is a pricing model. Both the pricing model and the distribution may not be correct.

Suppose the full model is rejected and we would like to see if the distributional assumptions are to blame. If we could study both $g[y_t|\, \hat{f}, \hat{p}]$ and $g[y_t|\, f, \hat{p}]$, we would know that any differences are attributable to $\hat{f}(\cdot)$. In this section we describe a method of doing precisely that.

To formalize the test, consider the following procedure. Rather than simulated draws from an assumed distribution, use the actual historical factor returns $\{x_1, x_2, ..., x_T\}$ in the portfolio pricing model,

$$\hat{y}_{t-i} = \sum_j p_j(x_{t-i})\, w_j \tag{3}$$

for i=0,..,T-1. Note that the portfolio weights are kept at their time-t values. We can therefore understand the series $\{\hat{y}_1,...,\hat{y}_T\}$ as estimates of what the portfolio returns would have been, had portfolio composition been constant over history.[3] Under stationarity, the distribution of the pseudo-values $\{\hat{y}_1,\ldots,\hat{y}_T\}$ will converge to the distribution of the portfolio $g(y_{t+1})$. It is therefore natural to fit a parametric model to $\{\hat{y}_1,\ldots,\hat{y}_T\}$ as an estimate of $g(y_{t+1})$. For example, we might fit a simple reduced form model such as a conditionally Normal GARCH(1,1).

It is equally possible to fit a "nonparametric" model at this point. That is, we could simply form a histogram of the pseudo-values $\{\hat{y}_1,\ldots,\hat{y}_T\}$ or use a nonparametric kernel density estimate of the density function. In either case, $f(x)$ is set to the empirical distribution of x so that, at least asymptotically, we have estimated the distribution $g[y_t \mid f, \hat{p}]$.

Comparing Distributions

Given the above machinery, how do we compare the two densities of interest $g[y_t \mid f, \hat{p}]$ and $g[y_t \mid \hat{f}, \hat{p}]$? First define the Probability Integral Transform (PIT) of y_t as: $z_t = \int_{-\infty}^{y_t} f(u)\,du$. Equivalently we can write $z_t = F(y_t)$ where $F(\cdot)$ is the cumulative density associated with $f(y_t)$.

The utility of the transform stems from the following result. If y_t is a random variable with probability density $f(\cdot)$, then $F(y_t)$ is a random variable that is distributed iid Uniform (0,1). This gives us a way of testing a density function. To examine whether $f(\cdot)$ is the density of some data y_t, we simply need to test whether the distribution of $F(y_t)$ is iid U(0,1).

These results have proven particularly useful for risk management. Crnkovic and Drachman (1997) formulate a test of forecast densities on the PIT of portfolio returns. Similar approaches are advocated by Diebold, Gunther, and Tay (1998) and Berkowitz (2001). In the present context, however, we would like to test only the distributional assumptions of the underlying risk model. Towards this end, consider the following algorithm:

1. Use the factor model to calculate the forecasted distribution $g[y_{t+1} \mid \hat{f}, \hat{p}]$.

3 The astute reader will recognize that this bears a close relationship with a form of modeling typically referred to as historical simulation. It can also be viewed as a formalization of a procedure used by practitioners whereby portfolios of hypothetical P&L are examined. See, for example, Berkowitz and O'Brien (2001) for further details.

2. Generate a series of pseudo-historical data, $y^*_{t\text{-}i} = \hat{p}(x_{t-i})$, i=0,...,T-1. The actual historical realizations of the factor returns are plugged into the pricing model instead of making any distributional assumptions.

3. Calculate the Probability Integral Transform $z_{t\text{-}i} = \int_{-\infty}^{y^*_{t\text{-}i}} g[u \mid \hat{f}, \hat{p}]du$, for i=0,...,T-1. The transform of the pseudo-historical data is taken with respect to the full model.

Under the null of correct factor distributions, $z_t \sim$ iid U(0,1) *whether or not the pricing model is accurate*. The following proposition formalizes this statement.

Proposition 1. If $\hat{f}(\cdot) = f(\cdot)$, then the transformed data $z_t \sim$ iid U(0,1), regardless of whether the remainder of the risk model is correct.

Proof. See appendix.

In this way, we have reduced the distributional assumptions into the statement that a sequence of data should be iid U(0,1). There are a variety of existing methods for testing this prediction such as the Kuiper statistic of Crnkovic and Drachman (1997) and Berkowitz (2001) likelihood ratio statistic.

Note that our procedure also provides an indirect test of the pricing model. To see this, suppose we have rejected the model on the basis of a backtest but we are unable to reject the distributional assumptions. In this case, we conclude that the pricing model is the problem. We are able to state this because we have split the full model into exactly two distinct components.

Of course, if on the other hand we reject both the full model and the distributional assumptions tells we can say nothing *directly* about the pricing model. The pricing model may or may not be valid. To test it, it is first necessary to have distributional assumptions which match the data. Once that is done and distributional assumptions are no longer rejected, the full model can be re-calculated to see if the pricing model is adequate.[4] The suggested testing sequence is shown schematically in Figure 1.

---

[4] If the model fails to include all of the relevant factors, it will show up as a rejection of the pricing model. It will not show up in a test of the distributional assumptions.

Direct tests of the factor distribution

One could, in principle, directly test the assumed factor distribution. For example, if Normality is assumed, one could use a Jarque-Bera test on the empirical distribution of $\hat{f}(\cdot)$. However, there are significant advantages to the proposed LR test.

Foremost, the LR testing approach is independent of the assumed distribution $\hat{f}(\cdot)$. If, for example, we assume a t-distribution the Jarque-Bera test is not appropriate. Indeed, for most assumptions one would have to resort to moment-based tests which have notoriously bad small sample properties. The proposed LR tests are, by virtue of the inverse-Normal transformation, always conducted within the Normal likelihood framework.

Second, direct tests of the factor distributions do not typically correspond to any economically interpretable criterion. The proposed LR test will, in many cases, deliver not just a rejection but a constructive guide as to the cause of the rejection. A rejection because of correlation indicates persistent forecast errors. A rejection because of the wrong mean implies a forecast that are too high (low) on average (see Berkowitz (2001)).

Lastly, risk managers are often *exclusively* interested in an accurate description of large losses or tail behavior. They do not want to reject a model that forecasts tail events well because of a failure to match the small day-to-day moves that characterize the interior of the forecast distribution. Berkowitz (2001) shows that the LR framework is easily tailored to allow the user to intentionally ignore model failures that are limited to the interior of the distribution by basing the LR tests on a censored likelihood. Loosely speaking, the shape of the forecasted tail of the density is compared to the observed tail.

This stakes a middle-ground between traditional interval VaR evaluation (counting violations) and the full LR approach. Such tests should be more powerful than traditional approaches while still allowing users to ignore model failures which may not be of interest -- failures which take place entirely in the *interior* of the distribution.

4. Monte Carlo Experiments

In this section, we investigate the actual behavior of the suggested procedure for testing distributional assumptions. The risk models are chosen to mimic techniques that are commonly used by risk managers at large financial institutions for constructing interval forecasts.

Data is generated from Heston's (1993) stochastic volatility process,

$$dS(t) = \mu S dt + \sqrt{\sigma_t} S \, dz_1(t)$$

$$d\sigma(t) = (\alpha - \beta \sigma_t) dt + \eta \sqrt{\sigma_t} \, dz_2(t)$$

with the drift, μ, is set to 12% and the volatility of volatility, η, is set to 10%. The long-run average of volatility is also 10% with α set to 20% and β set to 2. The constant risk free rate is set to 7% and the process is simulated over a 6 month horizon. I consider sample sizes 50, 100, 150, 250 and 500 which may be viewed as corresponding to different observation frequencies since the time-to-expiration is kept constant. The initial value of the stock is $40. For each observation, we calculate the true value of a call option, C_t, written on S_t with strike price $44.

Risk forecasts for the call option are estimated with the Black-Scholes model, the delta and delta-gamma approximations. In addition, we consider two ad hoc models that feature modifications designed to capture the stochastic volatility. We take the delta approximation, $\delta\varepsilon_t+\theta_t$, but instead of drawing ε_t from a lognormal, it is drawn from the (correct) stochastic volatility model. This is akin to the widespread practice of plugging in time-varying estimates of volatility into Black-Scholes. The second model is an analogous modification of the delta-gamma approximation.

We subject a portfolio to the statistical test of Crnkovic and Drachman (1997) which is based on the Kuiper statistic, $\max_z |f(z) - g(z)|$, where f(z) is the observed density of the PIT transformed forecast and g(z) the corresponding theoretical density (Uniform(0,1)). If the Kuiper statistic is large, it implies a rejection. We also calculate the Berkowitz (2001) likelihood ratio (LR) test in which the PIT-transformed z_t are further transformed from Uniform into N(0,1) variates. The LR test rejects if these transformed data do not have a mean of 0 or a variance of 1 as they should in theory.

Table 1 shows rejection rates of the various models. In all cases, the confidence level of the test is fixed at .05. The top panel of the Table reports the Monte Carlo rejection rates when the model is Black-Scholes and thus *incorrectly* assumes lognormality. The first column presents rejection rates for the Kuiper statistic of Crnkovic and Drachman (1997) and column 2 the LR statistic of Berkowitz (2001). Although there is some variation across columns, the distribution test rejects in the range of about 30 to 40 percent in samples of 500.

The next two panel show rejection rates if the model is a delta approximation or a delta-gamma approximation *again incorrectly* assuming lognormality. There is some variation across tests, but rejection rates are roughly similar in reaching about 30 percent. This is the case even though the largest sample size we consider is 500.

What about incorrect models with the correct (stochastic volatility) assumption on the factor distribution? The lower two columns in Table 1 report rejection rates for two such models. The rates should be near .05 because despite the model being wrong, the factor distribution is correct. Indeed, we see from the Table that rates are quite close in samples of about 500 observations. Results are roughly comparable whether we look at delta or delta-gamma approximation.

To a certain extent, Table 1 understates the usefulness of the testing procedure. If we mimic the sequential procedure (Figure 1) and test the distributional assumption *conditional on the full model being rejected*, power rises substantially. For Table 2 we again generate a total of 2000 iterations. The distributional test is run only if the full model is rejected. Column 1 shows the LR rejection rates for the various models. The results indicate rejection of the lognormal distribution at a rate as high as 88%, given that we have already rejected the overall Black-Scholes forecast.

In the bottom two panels of Table 2 we examine two models that are *false* but contain the *correct factor* distribution. In this case we expect the test to reject only about 5 percent of the time. Suppose, for example, the delta model with stochastic volatility is rejected – presumably because the delta part is an approximation. Given this rejection, we then test whether the factor distribution is to blame. The Table indicates that in samples of 500, we would only reject at a rate 5.2%. Thus, the size properties of the test appear quite good.

5. Discussion

Given that no risk model is literally correct, it is not surprising that risk managers and other practitioners have shied away from using formal statistical tests. Such tests, when enough historical data is available, will inevitably deliver a rejection of the full risk model. A more constructive approach is to ask whether a model is acceptably accurate for whatever purpose it is designed. If not, the next step should be to isolate those approximations, shortcuts and assumptions that are most detrimental.

In this paper, we have taken an initial step in this direction. The aspect of risk modeling that we are interested in isolating in the present paper is the distributional assumption made on factors. Risk models by their very nature require taking a stand on the factor distributions and yet, as we have emphasized above, no single approach appears to have gained a consensus. It may well be that the best procedure is to test a variety of distributional assumptions against the data and then impose the assumption that fits best. We hope that the tools of the present paper encourage users to explore this as an avenue to improving forecasts of portfolio risk.

References

Berkowitz, J. (2000), "A Coherent Framework for Stress-Testing," *Journal of Risk*, 2, 5-15.

Berkowitz, J. (2001), "Testing the Accuracy of Density Forecasts in Risk Management," *Journal of Business and Economic Statistics*, 19, 465-474.

Berkowitz, J. and J. O'Brien (2001), "How Accurate are the Value-at-Risk Models at Commercial Banks?" forthcoming, *Journal of Finance*.

Christoffersen, P. F. (1998), "Evaluating Interval Forecasts," *International Economic Review*, 39, 841-862.

Courant, R. (1961). *Differential and Integral Calculus, Volume I*. New York: Interscience Publishers.

Crnkovic, C. and Drachman, J. (1997), "Quality Control," in *VAR: Understanding and Applying Value-at-Risk*. London: Risk Publications.

Diebold, F. X., Gunther, T. A. and Tay, A. S. (1998),"Evaluating Density Forecasts," *International Economic Review*, 39, 863-883.

Jorion, P. (2001). *Value-at-Risk: the New Benchmark for Controlling Market Risk*. New York: McGraw Hill.

Lopez, J. A. and C. Walter (2000), "Evaluating Covariance Matrix Forecasts in a Value-at-Risk Framework," Federal Reserve Bank of San Francisco, paper 2000-21.

Pritsker, Matthew, 1997, "Evaluating Value at Risk Methodologies: Accuracy versus Computational Time," *Journal of Financial Services Research*, 12, 201-242.

Pritsker, Matthew, 2001, "The Hidden Dangers of Historical Simulation," Working Paper, Federal Reserve Board of Governors.

R. Velu and G. Zhou (1999), "Testing Multi-Beta Asset Pricing Models," *Journal of Empirical Finance*, 219-241.

Appendix

Proof of Proposition 1. Using the Jacobian of transformation, the distribution of z_t is given by

$$(4) \qquad g_0(y_t^*)\left|\frac{\partial}{\partial z_t} G^{-1}[y_t^* | \hat{f}, \hat{p}]\right|$$

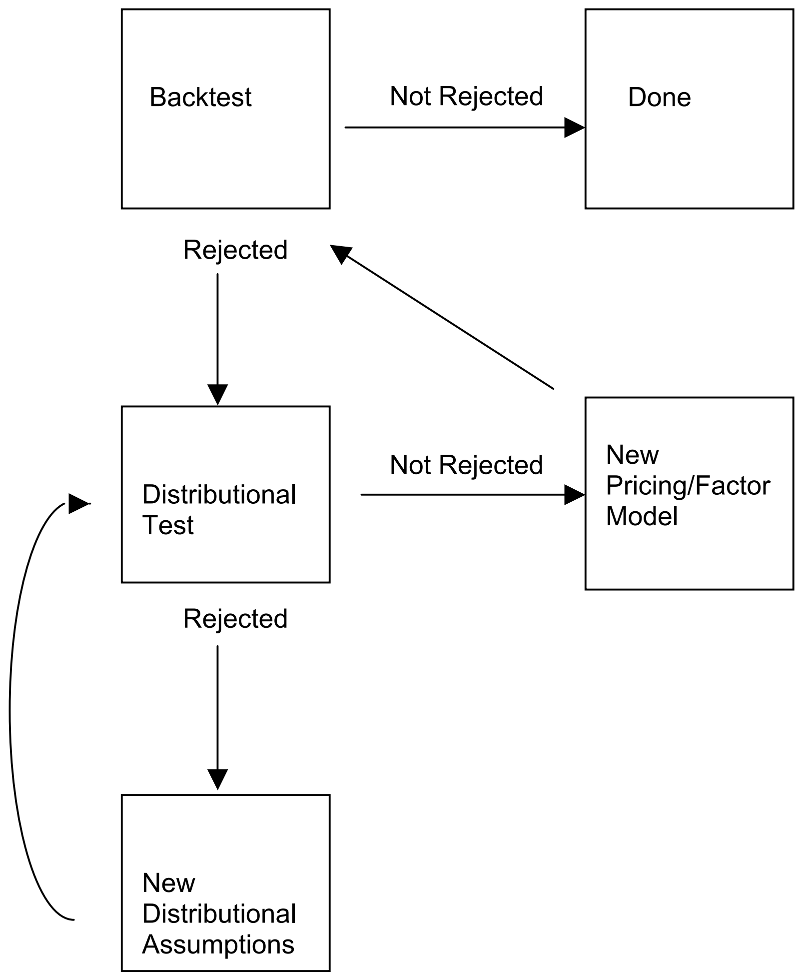
where g_0 is the distribution of y_t^* given $\hat{p}$ and the true factor distribution, f. $G[y_t^* | \hat{f}, \hat{p}] = \int_{-\infty}^{y_t^*} g[u \,|\, \hat{f}, \hat{p}]du$ and G^{-1} is its inverse. Because G is monotone we can simplify this to $g_0(y_t^*) / \left|\frac{\partial}{\partial z_t} G[y_t^* | \hat{f}, \hat{p}]\right|$. Application of the fundamental theorem of calculus (e.g., Courant (1961), p.111) yields $\frac{g_0(y_t^*)}{g[y_t^* | \hat{f}, \hat{p}]}$ where we have dropped the absolute value sign because $g(\cdot)$ is a density. Now by definition $y_t^* = \hat{p}(x_t)$ so that its distribution $g_0(y_t^*) = g_0[y_t^* | \hat{f}, \hat{p}]$. If $\hat{f}(\cdot) = f(\cdot)$ then $g(\cdot) \equiv g_0(\cdot)$, so that the quantity (4) simplifies to 1. The result follows because the density of a U(0,1) is identically equal to 1.

To establish independence, consider the sequence $\{y_t\}_{t=1}^{m}$. We can write the joint density $g(y_m, \ldots, y_1) = g_m(y_m | y_{m-1}, \ldots, y_1)\, g_{m-1}(y_{m-1} | y_{m-2}, \ldots, y_1)\, g_1(y_1)$. The joint density of the transformed variables is now given by the multivariate Jacobian of transformation

$$g_m(y_m)\, g_{m-1}(y_{m-1}) \,.\,.\, g_1(y_1)\left|\frac{1}{\frac{\partial G_1}{\partial z_1}} \cdots \frac{1}{\frac{\partial G_m}{\partial z_m}}\right|.$$

Each of the ratios, $g_m(y_m) / \frac{\partial G_m}{\partial z_m}$, is the density of a U(0,1), the product of which yields m-variate U(0,1). The joint density equals the product of the marginal densities■

Figure 1. Sequential Testing Procedure



Notes to Figure: The flowchart indicates the sequence of tests suggested in the text for evaluating both a pricing model and a set of distributional assumptions. Arrows labeled "reject" indicate that the test statistic led to a rejection of the null hypothesis.

Table 1
Backtesting Distributional Assumptions:
Stochastic Volatility

		Kuiper Statistic	LR Statistic
Black-Scholes			
	T = 50	0.063	0.181
	T=100	0.092	0.197
	T=150	0.121	0.226
	T=250	0.184	0.262
	T=500	0.276	0.358
Delta Model (lognormal)			
	T = 50	0.048	0.052
	T=100	0.092	0.063
	T=150	0.132	0.095
	T=250	0.190	0.156
	T=500	0.303	0.265
Delta-Gamma (lognormal)			
	T = 50	0.044	0.050
	T=100	0.089	0.063
	T=150	0.122	0.083
	T=250	0.184	0.137
	T=500	0.294	0.239
Delta Model - Stochastic Volatility			
	T = 50	0.030	0.025
	T=100	0.048	0.025
	T=150	0.049	0.023
	T=250	0.057	0.024
	T=500	0.057	0.051
Delta-Gamma - Stochastic volatility			
	T = 50	0.027	0.025
	T=100	0.049	0.025
	T=150	0.045	0.021
	T=250	0.062	0.022
	T=500	0.077	0.024

Notes: The Table compares the Monte Carlo performance of alternative techniques for validating forecast models over 2000 simulations. In each simulation, the portfolio of interest is comprised of a call options on an underlying diffusion process. Kuiper denotes the forecast test suggested by Crnkovic and Drachman (1997). LR is the likelihood ratio tests of Berkowitz (2001). The desired size is 5% in all cases.

Table 2: Sequential Testing of Distributional Assumptions
when the Factor Displays Stochastic Volatility

	Sample Size	Rejection Rate
Reject lognormal given Black-Scholes Rejected		
	T = 50	0.709
	T=100	0.815
	T=150	0.791
	T=250	0.870
	T=500	0.881
Reject stoch. vol. given Lognormal Delta rejected		
	T = 50	0.067
	T=100	0.062
	T=150	0.053
	T=250	0.048
	T=500	0.052
Reject stoch. vol. given Lognormal Delta-Gamma rejected		
	T = 50	0.085
	T=100	0.064
	T=150	0.053
	T=250	0.060
	T=500	0.052

Notes: The Table compares the Monte Carlo rejection rates of the LR test described in the text and displayed schematically in Figure 1. A total of 2000 simulations were generated with the distributional test run only if the full model is rejected. The portfolio of interest is comprised of a call option on an underlying stochastic volatility process for which the Heston (1993) model provides an analytic option price.