

VALUE AT RISK WHEN DAILY CHANGES IN MARKET VARIABLES ARE NOT NORMALLY DISTRIBUTED

John Hull

is a professor in the Faculty of Management at the University of Toronto.

Alan White

is a professor in the Faculty of Management at the University of Toronto.

This article proposes a new model for calculating VaR in which the user can choose any probability distributions for daily changes in the market variables, and parameters of the probability distributions are subject to updating schemes similar to GARCH. Transformations of the probability distributions are assumed to be multivariate normal. The model is appealing in that the calculation of VaR is rel-

atively straightforward and can make use of the RiskMetrics™ or a similar data base.

We test a version of the model using nine years of daily data on twelve different exchange rates. When the first half of the data is used to estimate the model's parameters, we find that it provides good predictions for the second half of the data.

In the last few years value at risk (VaR) has become a very popular risk management tool in many different types of organizations. There are a number of reasons for this. One is J.P. Morgan's decision in 1994 to make its RiskMetrics™ data base freely available to all market participants. Another is the climate created by derivatives disasters as represented by Procter & Gamble, Kidder Peabody, Orange County, and Barings. A third reason is the decision by central bank regulators to use VaR in calculating a bank's required capital.¹

A VaR calculation is aimed at making a statement of the form: "We are X% certain that we will

not lose more than V dollars in the next N days." The variable V is the VaR. It is a function of two parameters: N, the time horizon, and X, the confidence level. In defining a bank's required capital, regulators use $N = 10$ and $X = 99$. The required capital is therefore based on the losses over a ten-day period that are expected to happen 1% of the time.

Models for calculating VaR are reviewed by Duffie and Pan [1997]. The model most commonly used assumes that the probability distribution of the daily changes in each market variable is normal, an assumption that is far from perfect. The daily changes in many variables, particularly exchange

rates, exhibit significant amounts of positive excess kurtosis. This means that the probability distributions of daily changes in these variables have "fat tails" so that extreme outcomes happen much more frequently than would be predicted by the normal distribution assumption.

Duffie and Pan [1997] identify jumps and stochastic volatility as possible causes of kurtosis. They point out that under a jump-diffusion model, kurtosis is a declining function of the time horizon, while under a stochastic volatility model, it is an increasing function of the time horizon (at least for the time horizons normally considered in VaR calculations).

In this article we show how the normal distribution assumption can be relaxed. We develop a model that allows the user to choose any probability distribution for the daily changes in each market variable, with parameters of the distribution updated using schemes similar to GARCH. A key aspect of the model is the way that correlations are handled. We transform the daily changes in each market variable into a new variable that is normally distributed. We assume that the new variables are multivariate normal. We illustrate the approach using nine years of daily data on twelve different exchange rates.

I. APPROACHES TO CALCULATING VALUE AT RISK

There are a number of alternative ways to calculate VaR for a portfolio. A popular approach that uses the RiskMetrics™ or similar data base is to assume a model in which the changes in the values of the market variables (such as equity prices, zero-coupon bond prices, exchange rates, and commodity prices) have a multivariate normal distribution. The mean change in the value of each variable is assumed to be zero.

If the instruments in the portfolio are linearly dependent on n market variables, VaR can be calculated analytically when this model is used. The dollar change in the portfolio value in one day, ΔP , has the form:

$$\Delta P = \sum_{i=1}^n a_i \Delta x_i$$

where Δx_i is the proportional change in the value of

the i -th market variable during the day, and the a_i are constants ($1 \leq i \leq n$). The probability distribution of the portfolio value at the end of one day is normal with mean zero and standard deviation σ_p where

$$\sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} a_i a_j \sigma_i \sigma_j$$

The variable σ_i is the daily volatility of the i -th market variable, and ρ_{ij} is the correlation between Δx_i and Δx_j . The VaR for any given confidence level and any time horizon can easily be calculated from σ_p . For example, the VaR with a confidence level of 99% and a horizon of N days is $2.33\sigma_p\sqrt{N}$.

When the portfolio includes instruments that are not linearly dependent on the market variables, there are no exact approaches to calculating VaR. One possibility is to use Monte Carlo simulation. Unfortunately this can be quite time-consuming, since the complete portfolio must be revalued on each simulation trial.

An alternative is to approximate the relationship between ΔP and the Δx_i variables using the first two terms in a Taylor series expansion so that

$$\Delta P = \sum_{i=1}^n \frac{\partial P}{\partial x_i} \Delta x_i + 0.5 \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 P}{\partial x_i \partial x_j} \Delta x_i \Delta x_j \quad (1)$$

This expression can be used to calculate the moments of ΔP analytically.²

Alternatively a Monte Carlo simulation can be used, with ΔP calculated directly from Equation (1). The latter is referred to as the partial simulation approach. It is much less time-consuming than a full simulation, since it avoids the need for the portfolio to be revalued on each simulation trial.

Recently, Jamshidian and Zhu [1997] have suggested an interesting alternative way of speeding up Monte Carlo simulation: scenario simulation. It involves defining an M -point discrete approximation to the probability distribution of each market variable. On each simulation trial, samples for the changes in each market variable are taken from the full multivariate distribution of the market variables in the usual way. Each sample is then replaced by the closest value in the cor-

responding discrete distribution before valuing the portfolio. The advantage of the approach is that it significantly reduces the number of times individual instruments in a portfolio (particularly those dependent on only one market variable) have to be valued.

Some analysts prefer to use historical data rather than a model to define the statistical behavior of market variables. This involves creating a data base consisting of the daily movements in all market variables over a period of time. The first simulation trial assumes that the percentage changes in each market variable are the same as those on the first day covered by the data base; the second simulation trial assumes that they are the same as those on the second day, and so on.

The change in the portfolio value is calculated for each simulation trial, and the VaR is calculated as the appropriate fractile of the probability distribution of these portfolio changes. The change in the portfolio value can be obtained either by revaluing the portfolio or by using Equation (1).

The historical data approach has the advantage that it accurately reflects the historical multivariate probability distribution of the market variables. Its disadvantage is that the number of simulation trials is limited to the number of days of data that are available. Also, sensitivity analyses are difficult, and variables for which there are no market data cannot easily be included in the analysis.

The purpose of this article is to extend this model-building approach so that the historical behavior of market variables is represented more accurately.

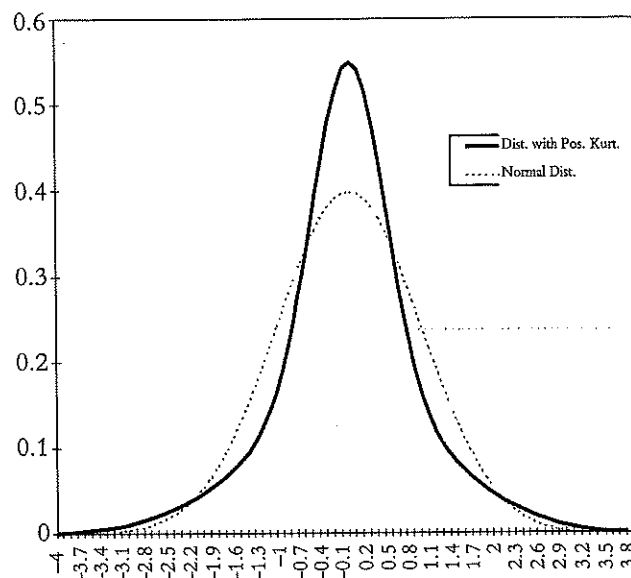
II. NON-NORMALITY IN MARKET VARIABLES

Although the RiskMetrics™ VaR calculation approach assumes multivariate normality, the changes in many market variables, particularly exchange rates, exhibit positive excess kurtosis. This means that extreme movements in the variables are more likely than a normal distribution would predict.

Exhibit 1 compares a normal distribution with a distribution exhibiting positive excess kurtosis.³ Both distributions have the same mean and variance, but the positive excess kurtosis distribution is more peaked and has fatter tails.

It is interesting to note what happens when we

EXHIBIT 1
COMPARISON OF NORMAL DISTRIBUTION AND POSITIVE KURTOSIS DISTRIBUTION — THE DISTRIBUTIONS HAVE THE SAME MEAN AND VARIANCE



move from a normal distribution to a distribution with positive excess kurtosis. Probability mass is added to the central part of the distribution and added to the tails of the distribution. At the same time, probability mass is taken from regions of the probability distribution that are intermediate between the tails and the center. The effect of excess kurtosis is therefore to increase the probability of very large moves and very small moves in the value of the variable, while decreasing the probability of moderate moves.

To illustrate the problem of non-normality in market variables, we use the behavior of exchange rates. The data consist of daily exchange rates for twelve major currencies between January 4, 1988, and August 15, 1997. The total number of trading days covered by the data is 2,425. The currencies are the Australian dollar (AUD), Belgian franc (BEF), Swiss franc (CHF), German deutschemark (DEM), Danish krone (DKK), Spanish peseta (ESP), French franc (FRF), British pound (GBP), Italian lire (ITL), Japanese yen (JPY), Dutch guilder (NGL), and Swedish krone (SEK).

Define e_i as the proportional change in an

exchange rate between day i and day $i + 1$. Exhibit 2 shows the frequency with which e_i exceeded one, two, three, four, five, and six standard deviations for each of the currencies. (For the purposes of calculating this table, e_i is assumed to have zero mean and constant variance.)

The table illustrates that all currencies exhibit significant excess kurtosis. The hypothesis that the e_i are normal can be rejected with a very high degree of confidence. The probability of a one-standard deviation move is 25.04% on average. This is considerably less than the 31.73% predicted by the normal distribution, and indicates that exchange rates have a more peaked distribution than the normal distribution.

The probability of a three-standard deviation move is 1.34% on average. This compares with 0.27% for the normal distribution and is consistent with exchange rates having fatter tails than the normal distribution.

Exhibit 2 also shows the usual excess kurtosis measure for each currency. For a normal distribution, this measure is zero.⁴

The RiskMetrics™ data base uses an exponentially weighted moving average (EWMA) for the daily

variance so that

$$\sigma_i^2 = \lambda\sigma_{i-1}^2 + (1 - \lambda)e_{i-1}^2 \quad (2)$$

where σ_i^2 is the variance calculated on day i , and λ is set equal to 0.94.⁵

Exhibit 3 shows the frequency with which e_i exceeded one, two, three, four, five, and six standard deviations when the standard deviation is calculated using Equation (2). The table shows that the stochastic volatility model in Equation (2) does lead to a significant reduction in the excess kurtosis measure, but the discrepancies between the observed frequencies and those that would be expected if the e_i were normal are still large. The hypothesis that the e_i are normal can still be rejected with a very high degree of confidence.

III. MODELING NON-NORMALITY AND THE CALCULATION OF VAR

The traditional approach to modeling non-normality in observed outcomes in a time series has

EXHIBIT 2

EXCHANGE RATE MOVES WHEN CONSTANT-VARIANCE MODEL IS USED (%)

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NLG	SEK
> 1 stdev	23.80	24.96	26.24	25.74	25.58	23.72	26.03	25.33	25.08	23.60	25.54	24.83
> 2 stdev	5.49	4.95	5.45	5.65	5.40	4.70	5.53	5.45	4.83	5.40	5.24	5.20
> 3 stdev	1.36	1.32	1.24	1.24	1.32	1.53	1.44	1.44	1.28	1.49	1.20	1.24
> 4 stdev	0.41	0.25	0.17	0.25	0.12	0.37	0.12	0.37	0.29	0.62	0.25	0.25
> 5 stdev	0.21	0.04	0.04	0.00	0.00	0.17	0.00	0.04	0.17	0.12	0.04	0.12
> 6 stdev	0.08	0.04	0.00	0.00	0.00	0.08	0.00	0.00	0.08	0.00	0.00	0.04
Excess Kurtosis	4.40	3.31	1.69	1.91	1.61	7.75	1.65	2.57	5.93	3.34	2.07	4.44

	AVERAGE	NORMAL
> 1 stdev	25.04	31.73
> 2 stdev	5.27	4.55
> 3 stdev	1.34	0.27
> 4 stdev	0.29	0.01
> 5 stdev	0.08	0.00
> 6 stdev	0.03	0.00
Excess Kurtosis	3.39	0.00

EXHIBIT 3

EXCHANGE RATE MOVES WHEN EXPONENTIAL WEIGHTED MOVING AVERAGE MODEL IS USED (%)

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NLG	SEK
> 1 stdev	26.98	28.42	28.80	28.55	29.17	29.00	29.50	27.81	28.96	26.49	29.37	29.29
> 2 stdev	5.73	6.52	6.31	6.68	6.35	6.15	6.52	6.68	6.11	6.52	6.15	6.35
> 3 stdev	1.53	1.40	1.03	1.44	1.28	1.32	1.53	1.44	1.69	1.24	1.11	1.24
> 4 stdev	0.66	0.33	0.17	0.08	0.21	0.29	0.17	0.25	0.33	0.50	0.29	0.08
> 5 stdev	0.21	0.08	0.08	0.04	0.04	0.12	0.04	0.08	0.08	0.17	0.08	0.04
> 6 stdev	0.08	0.08	0.04	0.04	0.04	0.08	0.04	0.04	0.04	0.12	0.04	0.04
Excess Kurtosis	3.59	3.23	2.09	2.03	1.97	2.77	2.12	2.35	2.15	3.47	2.15	1.47

	AVERAGE	NORMAL
> 1 stdev	28.53	31.73
> 2 stdev	6.34	4.55
> 3 stdev	1.36	0.27
> 4 stdev	0.28	0.01
> 5 stdev	0.09	0.00
> 6 stdev	0.06	0.00
Excess Kurtosis	2.45	0.00

been to assume that, although the unconditional returns are not normal, suitably conditioned returns are normal. There have been three popular models of this type: GARCH models, mixed jump-diffusion models, and Markov switching models.

In the GARCH or stochastic volatility models, the returns are normal conditional on knowing the current variance. In the mixed jump-diffusion models, the returns are normal conditional on there being no jump. In the Markov switching models, the returns are normal conditional on knowing the current state. The models have the advantage that they allow one to take advantage of the simple properties of multivariate normal distributions.

We propose a fourth alternative for handling non-normality. Our proposal is that some functional transformation of the observed returns is normal. This is similar to the other three models in that, suitably conditioned, the data are assumed to be normal.

This proposal is not entirely new. For example, it has traditionally been assumed that changes in asset prices are drawn from a lognormal distribution so that changes in the logarithm of the asset price are normally distributed.

Suppose that there are a total of m market variables. Define e_{ij} as the proportional change in variable j on day i ($1 \leq j \leq m$) and G_{ij} as the cumulative probability distribution assumed for e_{ij} . In general, G_{ij} is dependent on a number of parameters, some independent of i and some subject to updating schemes similar to GARCH. We transform e_{ij} into a new variable f_{ij} using the transformation

$$f_{ij} = N^{-1}[G_{ij}(e_{ij})] \quad (3)$$

where N is the cumulative normal distribution function.⁶

The expression in the brackets in Equation (3) is the cumulative probability that the daily change in variable j on day i is less than e_{ij} . Denote this by z so that e_{ij} is the z -th fractile of the distribution being assumed for daily changes. The variable f_{ij} is the same fractile of the standard normal distribution. In moving from e_{ij} to f_{ij} , we are mapping observations from the assumed distribution of daily changes into a standard normal distribution on a "fractile-to-fractile" basis.

This approach can be used in conjunction with either the historical data approach or the model approach to VaR calculations. The model approach to

VaR calculations assumes a particular form for the distribution of outcomes (for example, a displaced chi-square distribution, a t distribution, or a mixture-of-normals distribution), and estimates the parameters of the distribution for each market variable. In this case, the G function is calculated from the assumed distribution.

The historical data approach bases the distribution of future changes on the distribution of past changes. In this case, the G function is the cumulative histogram of historic changes.

We assume that the f_{ij} ($1 \leq j \leq m$) have a multivariate normal distribution. VaR can be calculated using either Monte Carlo simulation or a Taylor series expansion. When Monte Carlo simulation is used, the multivariate normally distributed variables, f_{ij} , are mapped into actual outcomes, e_{ij} , using

$$e_{ij} = G_{ij}^{-1}[N(f_{ij})] \quad (4)$$

On each simulation trial, we sample an f_{ij} and from Equation (4) find the corresponding e_{ij} .⁷ The e_{ij} are then used to calculate the change in the value of the portfolio in the usual way, using either Equation (1) or a full portfolio revaluation.

This model is ideally suited for use in conjunction with the scenario simulation approach of Jamshidian and Zhu [1997]. We first replace the assumed continuous probability distribution for each e-variable by a discrete distribution in a way similar to that suggested by Jamshidian and Zhu. We then carry out a preliminary analysis to determine from Equation (3) the range of values of f_{ij} that correspond to each discrete value of e_{ij} . As the simulation proceeds, we move directly from samples of the f_{ij} that come from a continuous multivariate normal to the corresponding samples of the e_{ij} that come from the discrete approximations to the assumed distributions.

To use the Taylor series expansion, we assume a quadratic relationship between the change in the portfolio value and the e_{ij} . To calculate the first few moments of the change in the portfolio value (so that the Cornish-Fisher expansion can be used), we require terms of the form $E(e_{ij})$, $E(e_{ij}e_{ik})$, $E(e_{ij}e_{ik}^2)$, $E(e_{ij}^2e_{ik}^2)$, and so on. These can be calculated, tabulated, and stored in advance of VaR calculations

by carrying out a single Monte Carlo simulation of the e_{ij} s.

If this Monte Carlo simulation has to be carried out every day, the Taylor series expansion approach is clearly inferior to the use of Monte Carlo simulation approach in conjunction with a quadratic approximation. If only one parameter in G_{ij} is updated from day to day, however, and this is a variance parameter that effectively "scales" the e_{ij} , then the Monte Carlo simulation has to be carried out only once (or at least only when the G_{ij} functions are changed). This is the situation in the example given below and is likely to be the case in most applications of our approach.

An alternative approach is to assume that the correlation between e_{ij}^m and e_{ik}^n is the same as it would be if the e variables were normally distributed, and calculate the required expectations from these correlations and the moments of the assumed probability distributions of the e variables.

IV. A PARTICULAR EXAMPLE OF THE MODEL

Here we illustrate how the G functions can be chosen.⁸ The model parameters are estimated from actual foreign exchange data, and the quality of approximation is then tested by comparing the model fit for a holdout sample of data. We use a mixture-of-normals distribution to represent the G functions.⁹

We consider two models. In the first model, the distribution of daily changes is stationary. In the second model, the variance of the distribution of daily changes can change, but parameters describing the kurtosis of the distribution are constant. The parameter estimation is done by fitting the fractiles of the distribution.

We emphasize that the functional form for the G functions and the estimation procedure are illustrations of the general approach that can be followed. Many other different assumptions and estimation procedures can be used. It is encouraging, however, that the results we present for the second model are quite good.

Define σ_i^2 as the variance of the proportional daily change e_i in a variable on day i . We assume that the probability density of daily changes is¹⁰

$$g_i(e_i) = \frac{p}{\sqrt{2\pi u}\sigma_i} \exp\left(-\frac{e_i^2}{2u^2\sigma_i^2}\right) + \frac{1-p}{\sqrt{2\pi v}\sigma_i} \exp\left(-\frac{e_i^2}{2v^2\sigma_i^2}\right) \quad (5)$$

so that

$$G_i(e_i) = pN\left(\frac{e_i}{u\sigma_i}\right) + (1-p)N\left(\frac{e_i}{v\sigma_i}\right)$$

The probability density function is a weighted average of two zero-mean normal distributions. The first distribution has weight p and standard deviation $u\sigma_i$. The second distribution has weight $1-p$ and standard deviation $v\sigma_i$.

Note that we are not assuming a Markov switching model, where on any given day either one normal distribution or the other applies. Equation (5) defines a family of distributions in terms of four parameters: p , u , v , and σ_i .

The variance of the distribution in (5) is

$$pu^2\sigma_i^2 + (1-p)v^2\sigma_i^2$$

Since this must equal σ_i^2 , the parameters underlying the distribution must satisfy

$$pu^2 + (1-p)v^2 = 1 \quad (6)$$

In fitting the model to the data we consider two cases. In the first case, the variance of daily changes in a market variable is assumed to be constant and equal to σ^2 so that $\sigma_i = \sigma$. (This is the assumption in Exhibit 2.) In the second case, the variance is assumed to be given by the EWMA model in Equation (2) with $\lambda = 0.94$. Other GARCH updating schemes for σ_i could also be considered.

To test the models we choose values of p , u , and v that are consistent with Equation (6) and provide a best fit to the distribution of e_i over the period covered by the first half of the data (January 4, 1988, to October 19, 1992). A natural approach here would

appear to be to maximize the log-likelihood function:

$$\sum_i \log \left[\frac{p}{u\sigma_i} \exp\left(-\frac{e_i^2}{2u^2\sigma_i^2}\right) + \frac{1-p}{v\sigma_i} \exp\left(-\frac{e_i^2}{2v^2\sigma_i^2}\right) \right]$$

One problem with this, pointed out by Hamilton [1991], is that attempts to maximize the function can lead to instability, local solutions, and non-convergence problems.¹¹ Another problem is that the best-fit values of p , u , and v are greatly influenced by extreme data items.¹²

To overcome these problems, we decided to estimate the parameters by fitting the fractiles of the distribution. The data on exchange rate changes are divided into four categories: less than one standard deviation ($|e_i| \leq \sigma_i$); one to two standard deviations ($\sigma_i < |e_i| \leq 2\sigma_i$); two to three standard deviations ($2\sigma_i < |e_i| \leq 3\sigma_i$); and more than three standard deviations ($3\sigma_i < |e_i|$). We compare the number of data items observed in each category with the number that would be predicted for particular values of p , u , and v , and select the values of these parameters that maximize the log-likelihood function¹³

$$\sum_{k=1}^4 \alpha_k \log(\beta_k)$$

where α_k is the actual proportion of observations in the k -th category, and β_k is the predicted proportion. This procedure is carried out for each of the twelve currencies as well as for the pooled results from all exchange rates. Parameter estimates are made for the constant-variance case and the EWMA case. The best-fit parameters for the EWMA case values are shown in Exhibit 4.

The best-fit estimates of p , u , and v are then assumed to apply to the second half of the data sample, and a chi-square statistic is calculated to determine the goodness of fit. For the purposes of calculating the chi-square statistic the data are categorized as described above.¹⁴

We carry out a separate analysis for each of the twelve currencies. For the parameters estimated from the pooled results from all exchange rates, the

EXHIBIT 4

BEST-FIT VALUES OF p, u, AND v WHEN THE EXPONENTIALLY WEIGHTED MOVING AVERAGE MIXTURE OF NORMALS MODEL IS FITTED TO FIRST HALF OF DATA

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NLG	SEK	ALL
p	0.85	0.57	0.51	0.57	0.30	0.62	0.49	0.35	0.68	0.75	0.79	0.48	0.62
u	0.78	0.69	0.69	0.66	0.14	0.68	0.61	0.14	0.73	0.70	0.81	0.61	0.70
1 - p	0.15	0.43	0.49	0.43	0.70	0.38	0.51	0.65	0.32	0.25	0.21	0.52	0.38
v	1.77	1.30	1.24	1.32	1.19	1.37	1.27	1.24	1.41	1.58	1.52	1.26	1.36
χ^2	5.81	6.14	1.90	1.86	9.24	9.92	8.77	9.34	4.27	5.34	5.64	4.50	48.24

Note: Chi-square statistic measures the goodness of fit to the second half of the data.

chi-square statistic is calculated by applying the pooled parameter estimates to each currency separately. These results for the EWMA case also appear in Exhibit 4. The results for the constant-variance case are poor because of substantial changes in the volatilities of many of the currencies between the first period and the second period.¹⁵

In testing the validity of the model and the estimated parameters for a single currency, the model can be rejected with 95% confidence when the chi-square statistic is greater than 7.8. When we attempt to fit the distributions of all currencies simultaneously,

the model can be rejected with 95% confidence when the chi-square statistic is greater than 51.

Exhibit 4 shows that for the exponentially weighted moving average model the single-currency model can be rejected for only four of the currencies. When the same p, u, and v are used for all currencies, the value of the chi-square statistic is 48.24, indicating that the model cannot be rejected with 95% confidence.

The sum of the chi-squares for the individual currency models is 72.73, indicating that we can reject at the 95% confidence level the hypothesis that

EXHIBIT 5

PROBABILITY DISTRIBUTIONS FOR SECOND HALF OF DATA COMPARED WITH PROBABILITY DISTRIBUTION FOR THE MIXTURE OF NORMALS EWMA MODEL THAT PROVIDES BEST OVERALL FIT TO FIRST HALF OF DATA (%)

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NLG	SEK
0 to 1 stdev	72.44	72.19	72.28	72.11	70.96	70.30	69.97	72.77	70.71	73.18	71.45	70.79
1 to 2 stdev	21.37	21.29	21.53	21.78	23.10	24.26	24.09	21.04	23.51	20.05	22.36	23.68
2 to 3 stdev	4.70	4.95	5.20	4.79	4.54	4.46	4.70	4.54	4.37	5.28	5.28	4.62
> 3 stdev	1.49	1.57	0.99	1.32	1.40	0.99	1.24	1.65	1.40	1.49	0.91	0.91

	BEST-FIT MODEL	NORMAL MODEL
0 to 1 stdev	73.11	68.27
1 to 2 stdev	21.31	27.18
2 to 3 stdev	4.55	4.28
> 3 stdev	1.03	0.27

the EWMA model is correct with a different p , u , and v for each currency.

Our analysis therefore provides support for using the EWMA model with the same p , u , and v parameters for all currencies. The fit to the first half of the data is of course slightly worse when the p , u , and v parameters are constrained to be the same for all currencies.¹⁶

The fit to the second half of the data is much better, however. This is because the best-fit parameter values for an individual currency are not particularly

stable, while the best-fit parameters for the all-currency model are quite stable. For the exponentially weighted moving average model, the latter change from $p = 0.62$, $u = 0.70$, and $v = 1.36$ for the first half of the data to $p = 0.66$, $u = 0.73$, and $v = 1.38$ for the second half.

Exhibit 5 compares the actual distributions of e_i for the second half of the data assuming the EWMA model with the distributions predicted by the normal distribution and by a model based on a single set of p , u , and v parameters. The table illus-

EXHIBIT 6

CORRELATIONS BETWEEN EXCHANGE RATE MOVEMENTS AND BETWEEN TRANSFORMED EXCHANGE RATE MOVEMENTS (SECOND HALF OF DATA EWMA MODEL)

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NLG	SEK
EXCHANGE RATE MOVEMENTS												
AUD	1.000											
BEF	-0.046	1.000										
CHF	-0.054	0.862	1.000									
DEM	-0.057	0.926	0.916	1.000								
DKK	-0.057	0.886	0.839	0.900	1.000							
ESP	-0.049	0.821	0.762	0.831	0.816	1.000						
FRF	-0.055	0.924	0.887	0.945	0.902	0.846	1.000					
GBP	0.087	0.566	0.576	0.605	0.586	0.526	0.599	1.000				
ITL	0.049	0.601	0.594	0.627	0.621	0.651	0.668	0.466	1.000			
JPY	-0.116	0.513	0.531	0.536	0.495	0.448	0.520	0.294	0.333	1.000		
NLG	-0.059	0.934	0.914	0.984	0.901	0.836	0.946	0.595	0.628	0.534	1.000	
SEK	-0.006	0.573	0.571	0.600	0.598	0.578	0.625	0.449	0.588	0.324	0.594	1.000
TRANSFORMED EXCHANGE RATE MOVEMENTS												
AUD	1.000											
BEF	-0.053	1.000										
CHF	-0.061	0.866	1.000									
DEM	-0.068	0.932	0.915	1.000								
DKK	-0.071	0.897	0.844	0.908	1.000							
ESP	-0.055	0.823	0.765	0.833	0.820	1.000						
FRF	-0.065	0.925	0.884	0.943	0.907	0.847	1.000					
GBP	0.084	0.580	0.587	0.614	0.595	0.533	0.608	1.000				
ITL	0.037	0.606	0.600	0.627	0.629	0.657	0.672	0.478	1.000			
JPY	-0.123	0.513	0.528	0.532	0.498	0.451	0.516	0.299	0.339	1.000		
NLG	-0.071	0.938	0.913	0.983	0.908	0.838	0.944	0.605	0.628	0.530	1.000	
SEK	-0.017	0.581	0.575	0.601	0.607	0.582	0.630	0.450	0.596	0.327	0.596	1.000

trates that the mixture of normals is a big improvement over the pure normal as a representation of daily exchange rate changes.

V. USING THE RISKMETRICS™ OR SIMILAR DATA BASE

Exhibit 6 shows estimates for the correlations between the *e* variables and the correlations between the *f* variables for the second half of the data in our example when the EWMA model is used to update variance estimates. The correlations are very similar.¹⁷

This suggests that, rather than estimating the correlation between the *f* variables directly, we can reasonably assume that the correlation matrix for the *f* variables is the same as that for the *e* variables.

With this small approximation, all calculations can be carried out using the RiskMetrics™ or similar data base once the *G* functions have been determined. In general, G_{ij} depends on the estimate of the standard deviation of e_{ij} made on day *i*. Our example uses an EWMA model with $\lambda = 0.94$ for this, but other models can be used. The covariance matrix for the *f* variables is known because by construction the variance of each *f* is 1.0, and by assumption the correlations between the *f* variables are the same as those between the *e* variables.

VI. CONCLUSIONS

We have shown how the usual multivariate-normal assumption in the calculation of VaR can be replaced by a "transform-to-multivariate-normal" assumption. The new model has the advantage that it enables the third, fourth, and higher moments of the returns on market variables to be reflected in VaR calculations.

We illustrate the model using nine years of daily exchange rate data for twelve currencies, and find that it is capable of accurately reflecting the probabilities of one-, two-, and three-standard deviation moves in the exchange rates. The model can be applied to calculating VaR relatively easily using the RiskMetrics™ or a similar data base. Either a Monte Carlo simulation or a Taylor series expansion approach can be used.

ENDNOTES

The authors are grateful to Tom McCurdy for comments and helpful suggestions. An earlier version of this article was titled "Taking Account of the Kurtosis in Market Variables When Calculating Value at Risk."

¹See Jackson et al. [1997] for a discussion of this.

²The Cornish-Fisher expansion provides an approximate way of converting the moments of ΔP to the appropriate fractile of the distribution of ΔP . For a description, see Johnson and Kotz [1972].

³The positive kurtosis distribution in Exhibit 1 is constructed as an equally weighted mixture of two normal distributions.

⁴The excess kurtosis of a distribution is defined as $(m_4/v^2) - 3$

where m_4 is its fourth central moment, and v is its variance.

⁵This is a particular case of a GARCH (1, 1) process.

⁶A similar idea is in Hull [1977] and has been used by Duan [1997].

⁷This may involve an iterative search.

⁸Our objective is to show that it is not difficult to find a reasonable representation of the actual distribution of returns for foreign exchange data. We are not attempting to find the "true" form of *G*.

⁹Other researchers who have used mixtures of normals in VaR calculations are Zangari [1996] and Venkataraman [1997]. Zangari assumes probability distributions for each of the parameters describing the mixture of normals and uses a Bayesian updating scheme. Venkataraman uses a quasi-Bayesian maximum likelihood estimation procedure. Our model is non-Bayesian and conceptually simpler than that of either Zangari or Venkataraman. Also it has a quite different approach to handling correlations.

¹⁰For ease of notation, we omit the *j* subscript in this part of the article.

¹¹Consider, for example, when one of the e_i is zero. As u tends to zero, the two exponential functions are both 1, while $p/(u\sigma_i)$ tends to infinity, so that the log-likelihood function also tends to infinity.

¹²It might seem strange that we are concerned about the adverse effect of extreme data items on our parameter estimates when the purpose of the model is to capture the probability of extreme outcomes more accurately. Our argument is as follows. When calculating VaR, we are not interested in modeling the really extreme (e.g., 0.1% tails) of the distribution of a market variable accurately at the expense of decreased accuracy in modeling the 1% and 5% tails. This is because the 0.1% tail is not known accurately from historical data and does not influence a VaR calculation for the confidence levels usually chosen.

¹³The probability that, out of a total of N observations, n_i are in category i ($1 \leq i \leq 4$) is proportional to

$$\beta_1^{n_1} \beta_2^{n_2} \beta_3^{n_3} \beta_4^{n_4}$$

Taking the logarithm of this function and dividing by N, we obtain the log-likelihood function.

¹⁴The chi-square statistic is

$$\sum_{i=1}^4 \frac{(A_i - E_i)^2}{E_i}$$

where A_i is the actual number of observations in category i in the second half of the data, and E_i is the expected number of observations based on the distribution estimated from the first half of the data.

¹⁵It is interesting to note that, if we base the calculation of the chi-square statistic on the volatility estimate of each currency during the second period instead of the first, the model works very well. Among the twelve currencies, the constant-variance version of the model can be rejected only for the Swedish krone. When a single model is used for all currencies, the chi-square statistic is 29.55, well under the 95% confidence level of 51. This suggests that, although the variance of a currency's return is non-stationary, the p , u , and v parameters describing the kurtosis are fairly stationary.

¹⁶The maximum value of the total log-likelihood function (all currencies) decreases from -9.46 to -9.47 when we constrain the parameters to be the same for all currencies in the case of the exponentially weighted moving average model.

¹⁷Note that the instantaneous correlations between the e variables are identical to those between the f variables.

REFERENCES

- Duan, J.-C. "Conditionally Fat-Tailed Distributions and the Volatility Smile in Options." Working paper, Hong Kong University of Science and Technology, 1997.
- Duffie, D., and J. Pan. "An Overview of Value at Risk." *Journal of Derivatives*, Vol. 4, No. 3 (Spring 1997), pp. 7-49.
- Hamilton, J. "A Quasi-Bayesian Approach to Estimating Parameters for Mixtures of Normal Distributions." *Journal of Business and Economic Statistics*, Vol. 9, No. 1 (1991), pp. 27-39.
- Hull, J.C. "Dealing with Dependence in Risk Simulations." *Operational Research Quarterly*, Vol. 28, No. 1 (1977), pp. 201-203.
- Jackson, P., D.J. Maude, and W. Perraudin. "Bank Capital and Value at Risk." *Journal of Derivatives*, Vol. 4, No. 3 (Spring 1997), pp. 73-90.
- Jamshidian, F., and Y. Zhu. "Scenario Simulation Model: Theory and Methodology." *Finance and Stochastics*, Vol. 1 (1997), pp. 43-67.
- Johnson, S.L., and S. Kotz. *Distributions in Statistics: Continuous Univariate Distributions I*. New York: John Wiley & Sons, 1972.
- Venkataraman, S. "Value at Risk for a Mixture of Normal Distributions: The Use of Quasi-Bayesian Estimation Techniques." *Economic Perspectives*, Federal Reserve Bank of Chicago, March/April 1997, pp. 2-13.
- Zangari, P. "An Improved Methodology for Measuring VaR." *RiskMetrics Monitor*, Reuters/J.P. Morgan, 1996.