Incorporating volatility updating into the historical simulation method for value-at-risk

John Hull and Alan White

This paper proposes a procedure for using a GARCH or exponentially weighted moving average model in conjunction with historical simulation when computing value-at-risk. It involves adjusting historical data on each market variable to reflect the difference between the historical volatility of the market variable and its current volatility. The authors compare the approach using 9 years of daily data on 12 exchange rates and 5 stock indices with more commonly used historical simulation approaches and show that it is a substantial improvement.

1. INTRODUCTION

In recent years value-at-risk (VaR) has become a very popular measure of market risk. It is widely used by financial institutions, fund managers, and nonfinancial corporations to control the market risk in a portfolio of financial instruments. As discussed by Jorion (1997), it has been adopted by central bank regulators as the major determinant of the capital banks are required to keep to cover potential losses arising from the market risks they are bearing.

The VaR of a portfolio is a function of two parameters, a time period and a confidence level. It equals the dollar loss on the portfolio that will not be exceeded by the end of the time period with the specified confidence level. If $X\%$ is the confidence level and N days is the time period, the calculation of VaR is based on the probability distribution of changes in the portfolio value over N days. Specifically VaR is set equal to the loss on the portfolio at the $100 - X$ percentile point of the distribution. Bank regulators have chosen N equal to 10 days and X equal to 99%. They set the capital required for market risk equal to three times the value of VaR calculated using these parameters.

In practice the VaR for N days is almost invariably assumed to be $\sqrt{N}$ times the VaR for one day. A key task for risk managers has therefore been the development of accurate and robust procedures for calculating a one-day VaR.

One common approach to calculating VaR involves assuming that daily percentage changes in the underlying market variables are conditionally multivariate normal with the mean percentage change in each market variable being zero. This is often referred to as the 'model building' approach. If the daily change in the portfolio value is linearly dependent on daily changes in market variables that are normally distributed, its probability distribution is also normal. The variance of the probability distribution, and hence the percentile of the distribution corresponding to VaR, can be calculated in a straightforward way from the variance–covariance matrix for the market variables. In circumstances where the linear assumption is inappropriate, the change in the portfolio value is often approximated as a quadratic function of percentage changes in the market variables. This allows the first few moments of the probability distribution of the change in the portfolio value to be calculated analytically so that the required percentile of the distribution can be estimated.[1] An alternative approach to handling nonlinearity is to use Monte Carlo simulation. On each simulation trial daily changes in the market variables are sampled from their multivariate distribution and the portfolio is revalued. This enables a complete probability distribution for the daily change in the portfolio value to be determined.[2]

Many market variables have distributions with fatter tails than the normal distribution. This has led some risk managers to use 'historical simulation' rather than the model building approach. Historical simulation involves creating a database consisting of the daily movements in all market variables over a period of time. The first simulation trial assumes that the percentage changes in the market variables are the same as on the first day covered by the database; the second simulation trial assumes that they are the same as on the second day; and so on. The change in the portfolio value is calculated for each simulation trial and the required percentile of the probability distribution of this change is estimated.[3] As an example, suppose that 1000 days of data are used and the 1 percentile of the distribution is required. This would be estimated as the tenth worst change in the portfolio value.

The advantage of the model building approach is that the underlying variance–covariance matrix can be updated using an exponentially weighted moving average (EWMA) or GARCH model.[4] The disadvantage is that the market variables are assumed to be conditionally multivariate normal. The model building approach takes no account of skewness or kurtosis in the distributions of market variables and no account of nonlinear correlations between market variables. Historical simulation, by contrast, has the advantage that it accurately reflects the historical multivariate probability distribution for the market variables. Its main disadvantage is that it incorporates no volatility updating.

Hull and White (1998) show how the assumption of multivariate normality in the model building approach can be relaxed. Their approach allows any probability distribution to be assumed for the daily changes in a market variable. A transformation is used to convert the assumed distribution to a

[1] When only two moments are calculated the distribution of the change in the portfolio value is assumed to be normal. When three or more moments are calculated, the Cornish–Fisher expansion can be used to estimate the required percentile.

[2] Revaluing the complete portfolio on each simulation trial is usually not feasible because of the computation time involved. One approach to speeding up calculations is to assume that the change in the portfolio value is a quadratic function of the change in the market variables.

[3] As in the case of Monte Carlo simulation, the quadratic approximation can be used as an alternative to a full portfolio revaluation on each simulation trial.

[4] For a discussion of these models, see J. P. Morgan (1995) or Engle and Mezrich (1995).

standard normal distribution. This transformation is defined so that the X-percentile point of the assumed distribution is transformed to the X-percentile of a standard normal distribution. The transformed market variables are assumed to be multivariate normal. The approach was tested using 9 years of data on 12 different currencies and found to perform well.

The assumed distribution for each market variable in Hull and White (1998) can be chosen in a variety of ways. One possibility is to select an appropriate standard distribution (for example, a mixture of normals) and use maximum likelihood methods to find the best fit parameters. Another possibility is to use the historical distribution. A third possibility is to smooth the historical distribution; for example, by using a kernel estimator.[5] The Hull and White approach provides one way of bridging the gap between the model building and historical simulation approaches. It shows how the model building approach can be modified to incorporate some of the attractive features of the historical simulation approach. In this paper we propose an alternative approach that allows volatility updating to be incorporated into historical simulation.

2. INCORPORATING VOLATILITY UPDATING SCHEMES INTO A HISTORICAL SIMULATION

The probability distribution of a market variable, when scaled by an estimate of its volatility, is often found to be approximately stationary. This suggests that historical simulation can be improved by taking account of the volatility changes experienced during the period covered by the historical data. If the current volatility of a market variable is 1.5% per day and two months ago the volatility was only 1% per day, the data observed two months ago understates the changes we expect to see now. On the other hand, if the volatility was 2% per day two months ago the reverse is true.

We consider a portfolio dependent on a number of market variables and assume that the variance of each market variable during the period covered by the historical data is monitored using either a GARCH or EWMA model. We are interested in estimating VaR for the portfolio at the end of day $N-1$ (i.e. for day N).

Define:

h_{tj}: the historical percentage change in variable j on day t of the period covered by the historical sample ($t < N$);

σ_{tj}^2: the historical GARCH/EWMA estimate of the daily variance of the percentage change in variable j made for day t at the end of day $t-1$.

The most recent GARCH/EWMA estimate of the daily variance is σ_{Nj}^2. This is the estimate, made at the end of day $N-1$, of the variance of the percentage

[5] An approach involving the use of historical simulation in conjunction with a kernel estimator is suggested by Butler and Schachter (1998).

change in variable j during day N. We assume that the probability distribution of h_{tj}/σ_{tj} is stationary. We therefore replace each h_{tj} by h^*_{tj}, where

$$h^*_{tj} = \sigma_{Nj}\frac{h_{tj}}{\sigma_{tj}}, \tag{1}$$

and set the tth sample percentage change for variable j to h^*_{tj} instead of h_{tj}.

This approach (which will be referred to as HW) is a straightforward extension of traditional historical simulation (which will be referred to as HS). Instead of using the actual historical percentage changes in market variables for the purposes of calculating VaR, we use historical changes that have been adjusted to reflect the ratio of the current daily volatility to the daily volatility at the time of the observation. Suppose that 20 days ago the observed percentage change in a market variable was 1.6% and the daily volatility was estimated to be 1%. If the daily volatility is now estimated to be 1.5%, the sample percentage change corresponding to the observation 20 days ago is 2.4%.

3. THE BRW APPROACH

One of the ways in which risk managers attempt to allow for stochastic volatility is by sampling more frequently from recent observations than from observations generated in the distant past. Boudoukh, Richardson, and Whitelaw (1998) proposed one version of this approach (which will be referred to as BRW). The weight given to the observation $n + 1$ days ago is λ times the weight given to the observation n days ago, where $0 < \lambda < 1$.[6] To determine a particular percentile of the probability distribution in BRW, it is necessary to order the observations over the last N days and then, starting from the lowest one, accumulate weights until the percentile is reached.[7]

We define a '5% tail event' as the occurrence of an observation that lies in the 5% tail of the historical distribution and a '1% tail event' as the occurrence of an observation that lies in the 1% tail of the historical distribution. Boudoukh, Richardson, and Whitelaw (1998) indicate that, when regular historical simulation is used, 5% tail events do happen approximately 5% of the time and 1% tail events do happen approximately 1% of the time. However, there is significant 'bunching'; that is, tail events tend to happen in close succession rather than occurring randomly throughout the days covered by the data. An attractive feature of the BRW approach is that it greatly reduces bunching.

[6] Note that BRW use an EWMA approach to define the weights given to observations, but this is quite different from the EWMA model for updating volatilities.

[7] Note that percentiles can be computed in a variety of ways. Suppose that a data set consists of the numbers 1, 2, 3, and 4. The definition we use throughout this paper implies 1, 2, 3, and 4 are 25, 50, 75, and 100 percentiles, respectively, and the values for intermediate percentiles are calculated using linear interpolation. An alternative definition (used by Microsoft's Excel) implies that 1, 2, 3, and 4 are the 0, 33.33, 66.67, and 100 percentiles, respectively, and the values for intermediate percentiles are calculated using linear interpolation.

BRW can be criticized on the grounds that it is an indirect and somewhat inefficient way of allowing for stochastic volatility. In BRW (and all schemes that involve sampling more frequently from recent observations) a short run sequence of abnormally large positive (or negative) returns will markedly skew the predicted distribution to the right (or the left). In BRW, when $\lambda = 0.98$, the most recent observation is assigned a probability of about 2% so that a single large outcome is enough to generate this sort of skew. BRW and similar schemes shorten the effective sampling period to capture the behavior of stochastic volatility. Unfortunately, in doing so, they capture the stochastic behavior of all other sample moments of the distribution.

4. COMPARISON OF APPROACHES

We tested the three schemes (HS, BRW, and HW) using daily data on 12 different exchange rates between 4 January 1988 and 15 August 1997 and five different stock indices between 11 July 1988 and 10 February 1998. The currencies were the Australian dollar (AUD), Belgian franc (BEF), Swiss franc (CHF), Deutschmark (DEM), Danish krone (DKK), Spanish peseta (ESP), French franc (FRF), British pound (GBP), Italian lira (ITL), Japanese yen (JPY), Dutch guilder (NLG), and Swedish krona (SEK). The stock indices were the S&P 500, CAC-40, FT-SE 100, Nikkei 225, and Toronto Stock Exchange 300. For each market variable, we had over 2400 daily observations. For BRW we used $\lambda = 0.98$.[8] In HW, the daily variance was updated using the EWMA model

$$\sigma_{tj}^2 = \alpha\sigma_{t-1,j} + (1-\alpha)h_{t-1,j}, \tag{2}$$

with $\alpha = 0.94$.[9]

For all three approaches and all market variables, a probability distribution of the daily percentage change was estimated each day from the most recent 500 days of data. The 5 and 1 percentiles of the distribution were noted. For each market variable, we define indicator functions $I(t)$ and $J(t)$ for day t. We set $I(t) = 1$ if the observation on day t is a 5% tail event, and zero otherwise; and we set $J(t) = 1$ if the observation on day t is a 1% tail event, and zero otherwise.

One issue in historical simulation is whether historical data should be adjusted to bring the mean percentage change to zero. Consider, for example, the S&P 500. During the 500 days ending 10 February 1998 the mean change was 0.09% per day. If we make no adjustment to the historical data, we are implicitly assuming that this is the expected change on 11 February 1998. We tested each of the three approaches with and without a mean adjustment. In the case of regular historical simulation, mean adjustment involved subtracting the mean daily percentage change from each of the 500 observations prior to estimating the 5% and 1% tails of the distribution. In the case of BRW, it

[8] Boudoukh, Richardson, and Whitelaw (1998) tested $\lambda = 0.99$ and $\lambda = 0.97$.

[9] This is the model used by J. P. Morgan in their RiskMetrics database (see J. P. Morgan 1995).

TABLE 1. Percentage of time that change in exchange rate is within 5% and 1% tails of estimated distribution using different approaches (based on 1923 observations). Asterisk indicates that the hypothesis of unbiasedness can be rejected with 95% confidence.

HS: Distribution estimated by giving equal weights to 500 most recent observations.

BRW: Distribution estimated by giving weights that decline exponentially to 500 most recent observations.

HW: Distribution estimated by giving equal weights to 500 most recent observations after they have been adjusted for volatility changes.

5% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
AUD	4.05	4.21	4.94	4.89	4.78	4.68
BEF	4.73	4.57	4.89	4.57	5.09	4.99
CHF	4.68	4.57	5.25	4.89	5.09	4.78
DEM	4.57	4.57	5.04	4.73	4.78	4.73
DKK	4.68	4.37	4.99	4.52	4.89	4.57
ESP	4.78	4.83	4.99	4.78	5.25	5.20
FRF	4.73	4.47	5.20	5.09	4.78	4.68
GBP	4.83	4.68	5.04	4.78	4.68	4.68
ITL	4.78	4.78	5.15	5.15	5.20	4.89
JPY	4.73	4.47	5.20	4.73	5.30	4.83
NLG	4.68	4.68	5.20	4.94	5.09	4.94
SEK	5.41	5.25	5.35	5.15	5.09	4.83
AVE	4.72	4.62*	5.10	4.85	5.00	4.82

1% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
AUD	0.78	0.78	1.40	1.40	1.04	0.99
BEF	0.99	0.94	1.56*	1.35	0.94	0.88
CHF	1.04	0.94	1.35	1.30	0.88	0.94
DEM	0.94	0.83	1.51*	1.25	1.09	1.09
DKK	0.78	0.78	1.40	1.40	1.04	0.99
ESP	0.94	0.94	1.72*	1.72*	0.94	0.94
FRF	0.99	0.94	1.35	1.14	1.09	0.99
GBP	0.94	0.94	1.30	1.14	1.14	1.14
ITL	0.99	0.99	1.40	1.25	0.88	0.88
JPY	0.83	0.83	1.72*	1.46*	0.83	0.78
NLG	0.88	0.88	1.35	1.20	1.04	1.04
SEK	1.14	1.09	1.51*	1.40	1.04	0.94
AVE	0.94	0.91	1.46*	1.33*	1.00	0.97

involved calculating a weighted mean percentage change and subtracting it from each observation. In the case of HW, it involved calculating the mean of the normalized observations, h_{tj}/σ_{tj}, and subtracting this mean from each normalized observation before multiplying by the estimate of the current volatility σ_{Nj}.

Table 1 shows the percentage of days when tail events happen for exchange rates. Table 4 reports similar results for stock indices. If the tails of the estimated

TABLE 2. Mean absolute error in percentage of 5% and 1% tail events for exchange rates in 100 consecutive days. (Results based on a total of 1823 100-day windows).

HS: Distribution estimated by giving equal weights to 500 most recent observations.

BRW: Distribution estimated by giving weights that decline exponentially to 500 most recent observations.

HW: Distribution estimated by giving equal weights to 500 most recent observations after they have been adjusted for volatility changes.

5% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
AUD	2.42	2.56	1.50	1.65	1.53	1.69
BEF	3.15	3.12	1.52	1.76	1.54	1.47
CHF	3.07	3.06	1.75	1.78	1.55	1.42
DEM	3.47	3.47	1.58	1.63	1.50	1.40
DKK	3.48	3.17	1.88	1.90	1.65	1.59
ESP	3.21	3.21	2.10	2.21	1.69	1.62
FRF	3.20	3.14	1.87	1.88	1.54	1.54
GBP	3.02	2.98	1.66	1.45	1.59	1.59
ITL	3.50	3.50	2.02	1.96	1.72	1.82
JPY	2.90	3.02	1.86	1.62	1.54	1.56
NLG	3.56	3.49	1.74	1.69	1.73	1.59
SEK	3.79	3.71	1.69	1.68	1.82	1.71
AVE	3.23	3.20	1.76	1.77	1.62	1.58

1% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
AUD	0.73	0.73	0.77	0.77	0.72	0.73
BEF	1.09	1.09	0.98	0.72	0.48	0.53
CHF	1.36	1.28	0.90	0.99	0.55	0.51
DEM	1.24	1.15	0.91	0.77	0.69	0.69
DKK	0.95	0.95	0.91	0.90	0.79	0.73
ESP	0.91	0.91	1.04	1.05	0.55	0.55
FRF	1.01	0.96	0.78	0.63	0.69	0.69
GBP	0.91	0.91	0.89	0.75	0.70	0.70
ITL	1.22	1.22	0.77	0.71	0.65	0.65
JPY	1.10	1.10	0.98	0.73	0.72	0.72
NLG	1.19	1.19	0.95	0.90	0.68	0.68
SEK	0.99	1.05	0.77	0.76	0.80	0.73
AVE	1.06	1.05	0.89	0.81	0.67	0.66

distributions were unbiased, 5% of tail events would happen on 5% of the days and 1% of tail events would happen on 1% of the days. Each result in Tables 1 and 4 is calculated from 1923 observations.[10] The samples generating the empirical distributions are overlapping, but $I(t)$ and $J(t)$ are each independent

[10] In the case of each currency, the indicator functions can only be calculated from day 500 onward. This explains why, although we started with over 2400 observation per market variable, Tables 1 and 4 are based on 1923 observations per market variable.

TABLE 3. Ljung–Box statistic for autocorrelations between tail events in exchange rates. Results based on autocorrelations with lags of between 1 and 15 days for indicator function which equals 1 if tail event happens and 0 otherwise. Zero autocorrelation cannot be rejected with 95% confidence when statistic is less than 25.

HS: Distribution estimated by giving equal weights to 500 most recent observations.

BRW: Distribution estimated by giving weights that decline exponentially to 500 most recent observations.

HW: Distribution estimated by giving equal weights to 500 most recent observations after they have been adjusted for volatility changes.

5% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
AUD	33.5	32.5	20.5	18.5	27.2	25.3
BEF	85.6	88.2	39.2	41.6	21.8	25.6
CHF	52.5	47.6	13.0	19.6	21.2	24.0
DEM	104.5	104.5	24.9	21.3	13.3	16.2
DKK	112.3	77.9	41.9	36.1	26.6	21.0
ESP	89.5	85.7	22.6	25.9	13.1	17.4
FRF	108.6	88.2	22.5	25.6	20.8	17.1
GBP	86.5	87.7	23.6	22.9	19.5	19.5
ITL	119.6	120.5	20.7	21.0	21.3	21.1
JPY	43.9	46.5	13.1	10.3	14.4	10.4
NLG	87.0	71.3	19.2	22.6	13.2	19.5
SEK	78.8	73.8	26.6	27.2	10.8	7.9
AVE	83.5	77.0	24.0	24.4	18.6	18.7

1% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
AUD	8.5	8.5	7.9	7.9	6.1	6.3
BEF	53.1	60.4	21.7	22.4	6.6	7.1
CHF	153.4	147.8	26.3	37.3	7.4	6.9
DEM	79.2	56.7	18.3	15.6	42.4	42.4
DKK	21.8	21.8	33.5	39.7	14.7	9.7
ESP	67.3	67.3	28.4	27.6	6.6	6.6
FRF	16.5	14.6	13.8	18.0	10.8	9.7
GBP	50.8	50.8	46.6	52.6	7.9	7.9
ITL	134.4	134.4	20.1	17.1	12.6	12.6
JPY	59.1	59.1	10.1	9.4	8.1	9.0
NLG	73.3	73.3	29.1	20.2	11.8	11.8
SEK	83.0	92.5	8.6	10.5	9.0	10.6
AVE	66.7	65.6	22.0	23.2	12.0	11.7

and identically distributed under the null hypothesis that the tail of the distribution is unbiased. The standard deviation of the percentage of days when tail events happen is therefore

$$\sqrt{\frac{p(1-p)}{n}},$$

where p is the probability of a tail event and n is the sample size. Asterisks in

TABLE 4. Percentage of time that change in stock index is within 5% and 1% tails of estimated distribution using different approaches (based on 1923 observations). Asterisk indicates that the hypothesis of unbiasedness can be rejected with 95% confidence.

HS: Distribution estimated by giving equal weights to 500 most recent observations.

BRW: Distribution estimated by giving weights that decline exponentially to 500 most recent observations.

HW: Distribution estimated by giving equal weights to 500 most recent observations after they have been adjusted for volatility changes.

5% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
S&P	5.67	5.09	5.30	4.63	5.04	4.57
CAC	5.61	5.41	5.51	5.25	5.15	5.15
FTSE	5.61	5.25	5.61	5.09	5.09	4.42
NIKKEI	5.51	5.77	5.41	5.93	5.09	5.30
TSE	5.51	4.94	5.77	4.94	4.78	4.31
AVE	5.58*	5.29	5.52*	5.17	5.03	4.75

1% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
S&P	1.40	1.20	1.46*	1.35	0.83	0.83
CAC	1.51*	1.30	1.51*	1.56*	1.09	1.09
FTSE	1.20	1.04	1.40	1.30	1.04	1.04
NIKKEI	1.25	1.30	1.35	1.66*	0.73	0.78
TSE	1.14	1.04	1.20	0.99	0.88	0.78
AVE	1.30*	1.17	1.38*	1.37*	0.91	0.90

Tables 1 and 4 indicate situations where the hypothesis that the tail of the distribution is unbiased can be rejected with 95% confidence. The tables show that the BRW method has a marked tendency to understate 1% tail events. The results reported by Boudoukh, Richardson, and Whitelaw (1998) show a similar phenomenon.

Boudoukh, Richardson, and Whitelaw (1998) propose a mean absolute percentage error measure (MAPE) for measuring bunching. This is calculated as follows. For each period of 100 consecutive days for which estimates are made, the absolute difference between the actual number of tail events and the expected number of tail events is calculated. (For 5% tail events the expected number of tail events is 5; for 1% tail events the expected number of tail events is 1.) The measure is set equal to the mean of these absolute differences.

MAPE is a combined measure of both bias and bunching. The impact of a bias in the measurement of tail events is clear. If the procedure for measuring tail events is biased so that in every 100-day period we observe two 1% tail events then MAPE = 1. To see how the bunching component of the measure works, consider the following example.

TABLE 5. Mean absolute error in percentage of 5% and 1% tail events for stock indices in 100 consecutive days. (Results based on a total of 1823 100-day windows).

HS: Distribution estimated by giving equal weights to 500 most recent observations.

BRW: Distribution estimated by giving weights that decline exponentially to 500 most recent observations.

HW: Distribution estimated by giving equal weights to 500 most recent observations after they have been adjusted for volatility changes.

5% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
S&P	2.68	2.57	1.62	1.48	1.52	1.47
CAC	2.99	2.88	1.61	1.53	1.66	1.69
FTSE	3.23	3.29	1.88	1.67	1.83	1.49
NIKKEI	3.80	3.83	2.19	2.81	2.22	2.29
TSE	2.71	2.33	1.86	1.49	1.58	1.60
AVE	3.08	2.98	1.83	1.79	1.76	1.71

1% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
S&P	1.15	0.98	0.72	0.61	0.63	0.63
CAC	1.14	0.99	0.79	0.77	0.69	0.69
FTSE	0.99	0.83	0.70	0.67	0.58	0.58
NIKKEI	1.21	1.22	0.85	1.20	0.67	0.61
TSE	0.89	0.87	0.72	0.59	0.49	0.42
AVE	1.08	0.98	0.75	0.77	0.61	0.59

Suppose that there are 599 observations numbered 1 to 599. This allows us to compute 500 overlapping 100-day samples. If every 100th observation (observations 100, 200, 300, 400, and 500) are 1% tail events then every 100-day sample will contain exactly one 1% tail event and MAPE will be zero. Now suppose that the 1% tail events are bunched together so that observations 100, 101, 300, 301, and 500 are tail events. Calculations shows that there are 198 samples with no tail events (absolute error = 1), 104 with 1 tail event (absolute error = 0), and 198 with 2 tail events (absolute error = 1). In this case, MAPE = 396/500 or 0.792.

The MAPE measure is similar to a standard deviation measure. If it were based on the difference between the observed number of tail events and the sample mean number of tail events, it would be even closer to a standard deviation measure. The definition used here was chosen to maintain consistency with BRW. The measure is shown in Tables 2 and 5. For each market variable, the results are based on 1823 different 100-day windows. Since the windows are overlapping, standard tests of statistical significance cannot be used.

As an alternative measure of bunching, we calculated the Ljung–Box statistic using the first 15 autocorrelations of the indicator functions $I(t)$ and $J(t)$. In the case of 5% tail events, the autocorrelations were between the $I(t)$; in the case of

TABLE 6. Ljung–Box statistic for autocorrelations between tail events in stock indices. Results based on autocorrelations with lags of between 1 and 15 days for indicator function which equals 1 if tail event happens and 0 otherwise. Zero autocorrelation cannot be rejected with 95% confidence when statistic is less than 25.

HS: Distribution estimated by giving equal weights to 500 most recent observations.

BRW: Distribution estimated by giving weights that decline exponentially to 500 most recent observations.

HW: Distribution estimated by giving equal weights to 500 most recent observations after they have been adjusted for volatility changes.

5% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
S&P	36.5	42.2	22.9	14.0	12.7	10.7
CAC	137.9	132.3	36.3	55.9	48.5	50.0
FTSE	100.1	104.0	15.2	26.0	37.1	17.3
NIKKEI	319.6	293.7	60.2	81.9	57.4	48.2
TSE	104.6	98.0	21.7	17.8	16.6	17.6
AVE	139.7	134.0	31.3	39.1	34.4	28.8

1% Tail	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
S&P	24.6	21.8	30.4	29.2	7.7	7.7
CAC	254.4	206.8	24.2	22.2	29.8	29.8
FTSE	26.8	33.2	7.3	7.4	3.2	3.2
NIKKEI	124.4	112.0	12.4	37.2	17.4	15.1
TSE	52.7	43.0	11.0	13.1	7.1	1.8
AVE	96.6	83.3	17.1	21.8	13.0	11.5

1% tail events, they were between the $J(t)$.[11] The results are reported in Tables 3 and 6. A value of the statistic less than 25 indicates that zero autocorrelation cannot be rejected at the 95% level. The numbers for each market variable are based on 1923 observations of the indicator functions.

Tables 2, 3, 5, and 6 show that BRW and HW both produce big improvements over HS as far as bunching is concerned. On average, HW performs better than BRW. Tables 3 and 6 show that, for 5% tails, the hypothesis of zero autocorrelation cannot be rejected for 12 out of the 17 market variables considered when both HW (no mean adjustment) is used and BRW (no mean adjustment) is used. For 1% tails, the hypothesis of zero autocorrelation cannot be rejected for 15 of the market variables in the case of HW (no mean adjustment) and for 11 of the market variables in the case of BRW (no mean adjustment).

Overall, the results with mean adjustment are similar to those without mean adjustment. Most of the rest of the paper will focus on the results obtained without mean adjustment.

[11] The Ljung–Box statistic is $m\sum_{k=1}^{15} w_k\eta_k^2$, where m is the number of observations, η_k is the autocorrelation with a time lag of k days, and $w_k = (m-2)/(m-k)$.

5. CAPITAL UTILIZATION

Define P as the 1 percentile of daily changes in a market variable. The regulatory risk capital for an investment of \$1 in a long position in the market variable is three times the 10-day 99% VaR, or $-3\sqrt{10}\,P$.

Figure 1 illustrates the three approaches for calculating P by showing the regulatory capital that would be required for an investment of \$1 in DEM. The DEM exchange rate is also shown. As illustrated by the figure, the capital is significantly more variable under BRW and HW than under HS.

In the HS method, the risk capital is unchanged for long periods of time due to the length of the window. The risk capital is affected by large observations that appear in, and drop off from, the window. An interesting aspect of BRW, illustrated in Figure 1, is that the capital tends to increase to a new level for a period of time and then drop off sharply. To understand the reason for this, consider a day (day d) where there is a sharp decrease in the value of a portfolio, worse than anything seen on the previous 500 days. Assuming no worse observation occurs subsequently, when $\lambda = 0.98$ this increases the capital for exactly 35 subsequent days. The reason is as follows. On day $d+1$ the observation for day d is given a weight of about 0.02, so that the 1 percentile of the distribution of daily returns equals the day d observation. Between day 1 and day 35, the weight given to day d remains above 1%, so that the 1 percentile

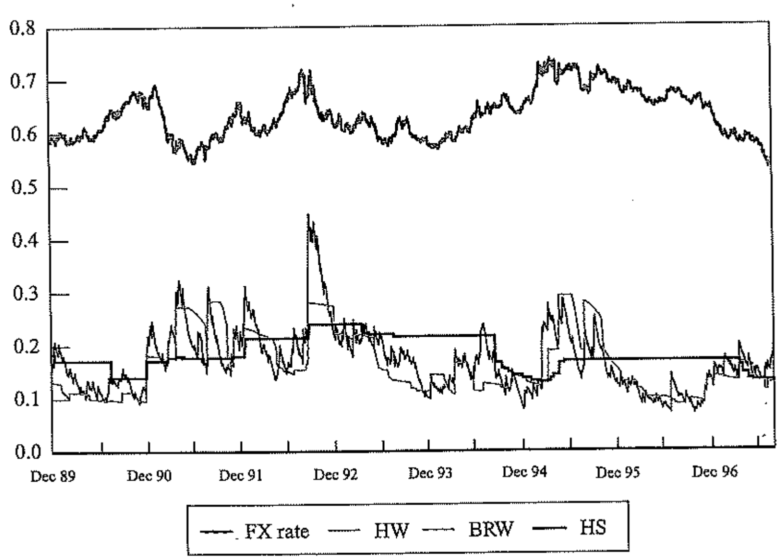

FIGURE 1. Capital required under three methods for an investment of \$1 in DEM.

HS: Distribution estimated by giving equal weights to 500 most recent observations.

BRW: Distribution estimated by giving weights that decline exponentially to 500 most recent observations.

HW: Distribution estimated by giving equal weights to 500 most recent observations after they have been adjusted for volatility changes.

continues to equal the day d observation. On day 36, the weight given to the observation is just below 1% and the next worst observation starts to influence the 1 percentile point. BRW can be contrasted with HW where capital requirements are determined by the most recent estimate of volatility and are therefore much more responsive to recent observations.

For long positions in a single foreign currency, we found that the capital required under BRW is on average 11.0% less than under HS and that the capital required under HW is on average 7.8% less than under HS. For long positions in a stock index, we found that the capital required under BRW is on average 0.2% higher than under HS and the capital required under HW is on average 6.7% higher than under HS. Clearly BRW scores high marks if the objective is to minimize capital. However, this is hardly surprising. Tables 1 and 4 show that BRW's 1% tail are not extreme enough. A consistent result from our data is that there is a greater than 1% chance of an observation being in BRW's 1% tail.

Of course, the objective should not be to choose the method that minimizes capital requirements. We contend that the best method is the one that, for a given average investment of capital, maximizes the protection against losses. Define:

A: average capital required under regular historical simulation;

C_{ave}: average capital required under an alternative scheme;

C_t: capital required on day t under the alternative scheme;

L_{tn}: losses incurred on the n days following day t.

The ratio L_{tn}/C_t is the proportion of the capital required to cover the losses (if any) during the n days following day t under the alternative scheme. The extreme values of this expression measure the chances of financial difficulties being experienced when the alternative scheme is used. However, it does not take account of the average amount of capital used by the alternative scheme. We propose that the measure

$$\frac{L_{tn}}{C_t} \cdot \frac{C_{\mathrm{ave}}}{A} \tag{3}$$

be used for choosing between schemes. This is the proportion of capital that would be required on day t to cover losses in the subsequent n days under the alternative scheme if a constant multiplier is applied to the capital each day so that it is the same on average as the capital used under regular historical simulation. We will refer to this as the 'capital utilization ratio'. The extreme values of the capital utilization ratio provide a measure of capital adequacy.

We calculated the 99.5 percentile and 99 percentile of the probability distribution of the capital utilization ratio for investments in each of the currencies and each of the stock indices for $n = 1$ and $n = 10$. The average of the results for each currency are shown in Table 7 and the average of the results for each stock index are shown in Table 8. The results indicate that for currencies HW and BRW provide better capital utilization than HS. HW performs better than BRW and the mean adjustment appears to improve the performance of

TABLE 7. Percentiles of the capital utilization ratio for investments in currencies. The capital utilization ratio measures the proportion of the capital used to cover losses during the specified time period. Results are the averages of those obtained from investments in each of 12 different currencies.

Time period (days)	Percentile	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
1	99.5	12.54	12.44	12.08	11.93	11.71	11.66
1	99.0	10.35	10.33	10.50	10.29	9.87	9.79
10	99.5	38.49	38.25	38.12	37.50	38.05	37.61
10	99.0	32.38	32.16	32.33	32.48	32.12	31.88

TABLE 8. Percentiles of the capital utilization ratio for investments in stock indices. The capital utilization ratio measures the proportion of the capital used to cover losses during the specified time period. Results are the averages of those obtained from investments in each of 5 different stock indices.

Time period (days)	Percentile	HS No Mean Adj	HS Mean Adj	BRW No Mean Adj	BRW Mean Adj	HW No Mean Adj	HW Mean Adj
1	99.5	13.68	13.52	14.36	14.56	13.24	13.30
1	99.0	11.26	11.31	11.92	11.97	11.17	11.06
10	99.5	41.14	40.56	43.28	43.18	43.40	43.00
10	99.0	33.61	33.26	36.17	35.95	36.16	35.93

HW slightly. For stock indices the results are less clear. HW performs slightly better than HS for a 1-day time horizon, but worse for a 10-day time horizon. BRW performs worse than either HS or HW for a 1-day time horizon, and about the same as HW for a 10-day time horizon.

6. SUMMARY

This paper shows how a volatility updating scheme such as GARCH can be used in conjunction with historical simulation for calculating value-at-risk. Risk managers sometimes attempt to allow for stochastic volatility by sampling from recent observations more frequently than from those generated in the distant past. One such approach, BRW, involves applying weights that decline exponentially to the observations. The approach we propose is more direct. It involves adjusting observations to reflect the difference between the volatility at the time of the observation and the current volatility.

We compare our approach with BRW. We use approximately 9 years of daily data on 12 different exchange rates and 5 different stock indices. Our approach provides better 1-percentile estimates of daily returns and is as good, if not better, at eliminating the bunching of tail events.

We have proposed a new way of assessing the effectiveness of a scheme for calculating VaR. It is designed to test the proportion of the capital likely to be used in extreme situations. We find that our method performs better than BRW. It is superior to regular historical simulation for investments in currencies. For investments in stock indices the results are mixed.

Acknowledgement
We are grateful to the editor, Philippe Jorion, for many suggestions that improved this paper.

REFERENCES

Boudoukh, J., Richardson, M., and Whitelaw, R. (1998). The best of both worlds. *Risk*, May, pp. 64–67.

Butler, J. S., and Schachter, B. (1998). Estimating value-at-risk with a precision measure by combining kernel estimation with historical simulation. *Review of Derivatives Research*, **1**, 371–390.

Engle, R. F., and Mezrich, J. (1995). Grappling with GARCH. *Risk*, September, pp. 112–117.

Hull, J., and White, A. (1998). Value-at-risk when daily changes in market variables are not normally distributed. *Journal of Derivatives*, 5(3), pp. 9–19.

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

J. P. Morgan (1995). *RiskMetrics Technical Manual*. J. P. Morgan.

J. Hull and A. White
University of Toronto