Testing Expected Shortfall Models for Derivative Positions

Kerkhof, F.L.J.; Melenberg, B.; Schumacher, J.M.

Publication date:
2003

No. 2003–24

TESTING EXPECTED SHORTFALL MODELS FOR DERIVATIVE POSITIONS

By Jeroen Kerkhof, Bertrand Melenberg and Hans Schumacher

March 2003

ISSN 0924-7815

Testing Expected Shortfall Models for Derivative Positions *

Jeroen Kerkhof[†], Bertrand Melenberg[‡], and Hans Schumacher[§]

Tilburg University

PO Box 90153

5000 LE Tilburg

The Netherlands

ABSTRACT

In this paper we test several risk management models for computing expected shortfall for one-period hedge errors of hedged derivatives positions. Contrary to value-at-risk, expected shortfall cannot be tested using the standard binomial test, since we need information of the distribution in the tail. As derivatives positions change characteristics and thereby the size of risk exposures over time one cannot apply the standard tests based on stationarity. To overcome this problem, we present a transformation procedure. For comparison purposes the tests are also performed for value-at-risk.

Keywords: Risk management, Backtesting, Expected shortfall, Value-at-risk.

JEL codes: C12, G18

*We thank Alessandro Sbuelz for constructive and helpful comments and Mark-Jan Boes for his help with the data. Any remaining errors are ours.

[†]Corresponding author: Department of Econometrics and Operations Research and CentER, Tilburg University, and Product Development Group, ABN-AMRO Bank, Amsterdam. E-mail: F.L.J.Kerkhof@CentER.nl, Phone: +31-13-4662134, Fax: +31-13-4663280.

[‡]Department of Econometrics and Operations Research, Department of Finance, and CentER, Tilburg University. E-mail: B.Melenberg@TilburgUniversity.nl.

[§]Department of Econometrics and Operations Research and CentER, Tilburg University. E-mail: J.M.Schumacher@TilburgUniversity.nl.

Testing Expected Shortfall Models for Derivative Positions

ABSTRACT

In this paper we test several risk management models for computing expected shortfall for one-period hedge errors of hedged derivatives positions. Contrary to value-at-risk, expected shortfall cannot be tested using the standard binomial test, since we need information of the distribution in the tail. As derivatives positions change characteristics and thereby the size of risk exposures over time one cannot apply the standard tests based on stationarity. To overcome this problem, we present a transformation procedure. For comparison purposes the tests are also performed for value-at-risk.

Keywords: Risk management, Backtesting, Expected shortfall, Value-at-risk.

JEL codes: C12, G18

I. Introduction

Managing the risks of derivative assets has always been one of the major challenges in risk management. With the strong increase in derivative positions in the portfolios of financial institutions the task of managing these risks has become more daunting than ever. An equally daunting task is testing the quality of models used to quantify the risk of derivatives positions.

Since the Basle Committee advised the use of value-at-risk (VaR) in the 1996 amendment to the Basle Accord for determination of regulatory capital, many studies have investigated VaR (see, for example, the overviews in Jorion (2000) and Dowd (1998) and the references therein). Recently, a literature emerged advocating alternative risk measures, namely, coherent risk measures and, in particular, expected shortfall (see, for example, Artzner et al. (1999), Delbaen (2000), Acerbi and Tasche (2002), and Tasche (2002)). The advantages of expected shortfall over VaR are that it satisfies the property of subadditivity and the fact that portfolio optimization under expected shortfall constraints yields reasonable portfolios, contrary to VaR (see, for example, Yamai and Yoshiba (2002) for the constrained portfolio optimization). Though most people agree that from a theoretical point of view expected shortfall is to be preferred to VaR, it is still less widely used due to the lack of a solid backtesting procedure. Recently, Kerkhof and Melenberg (2002) introduced a test for expected shortfall and found that for appropriately adjusted levels, expected shortfall has more desirable backtesting properties than VaR.

Though quite a number of studies have tested the performance of several VaR models, derivatives positions were rarely explicitly taken into account (see, for example, McNeil and Frey (2000), Christoffersen et al. (2001), and Berkowitz and O'Brien (2002)). In cases where derivative positions were explicitly taken into account, the literature usually focused on the computation of VaR rather than on the testing of the VaR models, since the standard binomial test can be applied (see, for example, Kupiec (1995) and El-Jahel et al. (1999)). However, the standard binomial test cannot be applied to expected shortfall. In order to test expected shortfall we need information of the distribution of bank's profit and losses (P&L) account, or more specifically its tail behavior.

One of the problems that one faces in determining the P&L distribution of (non-linear) derivatives is that their risk characteristics change over time. For example, an option can change from a 1 year at-the-money option into a 3 month far out-of-the-money option, resulting in completely different risk characteristics. In this paper, we propose a method to take into account the differences in risk exposures between options with different characteristics.

We consider several methods to estimate the risk measure for the one-day hedge error. The first method we consider is a simple Black-Scholes based model which assumes normal asset returns and constant implied volatilities. Method 2 relaxes the assumption of normal asset return and uses a nonparametric asset return distribution based on historical simulation. Method 3 is a full historical simulation method that assumes a nonparametric asset returns distribution and a nonparametric implied volatility distribution. The fourth method is a Vector AutoRegessive (VAR) model for asset returns and implied volatilities returns with Gaussian errors, while method 5 considers nonparametric errors instead.

We test the models on the FX market and, in particular, the mutual exchange rates of the US, the UK, and Japan. Furthermore, we test the models on S&P 500 options. We find that the historical simulation method and the VAR models perform reasonably well.

The remainder of the paper is structured as follows. Section II describes daily market risk for derivative positions. Section III discusses the aging, moneyness, and level effects of derivative positions and a possible transformation to standardize the risk exposures. The models used are described in Section IV. Section V describes the test used and Section VI presents the empirical results. Finally, Section VII concludes.

II. Quantifying daily market risk

Consider the situation where a financial institution manages a portfolio which is short in options. Due to this position the financial institution is subject to a risk exposure with respect to the value of the options. To decrease this risk exposure the financial institution hedges the derivative using a particular hedge strategy. To illustrate, consider a derivative whose price at day t equals f_t. The financial institution hedges the derivative using some underlying instruments with prices $S_t = \left(S_t^1, ..., S_t^k\right)$. Let the money market account at time t be given by N_t. Let the financial institution hedge the derivative by buying amounts $\gamma_t = \left(\gamma_t^1, ..., \gamma_t^k\right)$ of the underlying instruments. Define

$$\alpha_t = \frac{f_t - \gamma_t \cdot S_t}{N_t}. \tag{1}$$

Then we will have as accounting identity

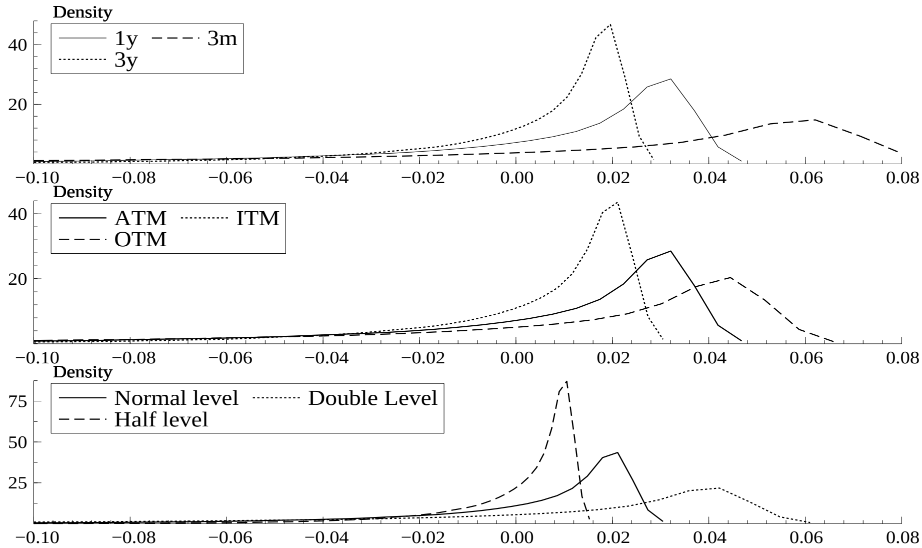
$$f_t = \gamma_t \cdot S_t + \alpha_t N_t. \tag{2}$$

The next day the price of the derivative will be f_{t+1}, while the hedging position (if there are no intermediate adaptations) will be valued $\gamma_t \cdot S_{t+1} + \alpha_t N_{t+1}$. The difference between the next period's derivative's price and the hedge position induces the *daily market risk*. A financial institution can quantify this daily market risk by assuming some method to estimate or calibrate the next day's probability distribution of $(f_{t+1}, S_{t+1}, N_{t+1})$. Taking a numeraire whose future value at $t+1$ is known (for example, a one-period discount bond) reduces the problem to estimating or calibrating the next day's probability distribution of (f_{t+1}, S_{t+1}), but now with respect to the numeraire instead of cash. This allows for estimation or calibration of the daily market risk measures (for instance, value-at-risk or expected shortfall). Our interest in this paper is in risk measures of the daily market risk profile. Specifically, we are interested in the distribution of

$$E_t^1 \equiv \Delta f_t - \gamma_t \Delta S_t, \tag{3}$$

where $\Delta x_t \equiv x_{t+1} - x_t$ for $x = f, S$. E_t^1 denotes the *one-period hedge error* and its distribution which is termed the *daily market risk profile* is denoted by $\mathcal{L}\left(E_t^1\right)$. Examples of the daily market risk profile are given in the upper panel of Figure 1, which

Figure 1. Aging, moneyness, and level effects On the x-axis the return on the hedged portfolio is given in percentages. Total number of simulations = 100,000. The upper panel shows the daily risk profiles of delta hedged ATM call option with a maturities of 3 months, 1 year, and 3 years and level 100. The middle panel shows the daily risk profiles of delta hedged OTM (m = -0.1) , ATM (m = 0), and ITM (m = 0.1) call option with a maturity 1 year and level of 100. The lower panel shows the daily risk profiles of a delta hedged ATM call option with a maturity of 1 year and levels of 50, 100, and 200.



presents daily market risk profiles of a delta hedged 3 month at-the-money, 1 year at-the-money, and 3 year at-the-money (ATM) call option in a Black-Scholes world with annual instantaneous drift $\mu = 0.1$, instantaneous volatility $\sigma = 0.2$, and instantaneous riskless interest rate equal to $r = 0.05$. Time t is measured in days (1 year equals 250 days). In line with Boyle and Emanuel (1980) a shifted non-central χ^2-distribution is found as an approximation for the market risk profile.

III. Aging, moneyness, and level effect

Figure 1 clearly shows that the distribution of hedge errors of options depends on the time to maturity, $\tau = T - t$. We refer to the fact that the daily market risk profile changes with the time to maturity as the *aging effect*. For shorter maturities, the daily risk profile is more spread out. The middle panel of Figure 1 shows the dependence of the daily risk profile on moneyness which is termed the *moneyness effect.*[1] Out-of-the-money options have more variability than in-the-money options. Finally, in the lower panel we see the influence of the level on the risk profiles, the so-called *level effect.* It is easy to show that this effect is linearly dependent on the level.

The three effects shown in Figure 1 indicate the problems one encounters when using time series data of a particular option to extract information of the daily market risk profile of that option. The observations of hedge errors of the option are taken with different times to maturity and potentially different moneyness and levels. Since the distribution differs for these situations, these hedge errors are hard to compare. In order to suppress the level effect we first determine a level-independent distribution of relative hedge errors in the following way:

$$\tilde{E}_t^1 = \frac{E_t^1}{f_t - \gamma_t \cdot S_t}. \tag{4}$$

The dependence on the daily market risk profile on the aging and moneyness effect is more complicated to resolve. To get rid of the aging and moneyness effect, it is natural to use data on derivatives with the same moneyness and time to maturity, if possible. For FX derivatives and interest-rate derivatives these data are available, since these are quoted in the market with a fixed time to maturity. For equity derivatives, however, this is more complicated due to the fact that these derivatives have fixed maturity dates. Therefore, we have to transform our data.

A. Transformation of the data

A possible way to correct the daily market risk profile for the aging and moneyness effect is to assume a parametric option pricing model so that one can use the characteristics of such a model to find the appropriate corrections. In this section we correct for the aging and moneyness effect using the Black-Scholes model.[2] Denoting the model price

[1]Moneyness is defined as $m = \log\left(e^{rt} S_t / k\right)$. A call option is called in-the-money (ITM) if $m > 0$, at-the-money if (ATM) $m = 0$, and out-of-the-money (OTM) if $m < 0$.

[2]Other models with sufficiently smooth pricing formulas can also be used.

by $f(\xi_t)$ with $\xi_t = (S_t, t)$ a Taylor series expansion gives

$$f(\xi_{t+\Delta t}) = f(\xi_t) + \Delta(\xi_t)\Delta S_t + \frac{1}{2}\Gamma(\xi_t)\,(\Delta S_t)^2 + \Theta(\xi_t)\Delta t + O(\Delta t^{3/2}), \tag{5}$$

where

$$\Delta S_t = S_{t+\Delta t} - S_t. \tag{6}$$

$\Delta(\xi_t) \equiv \frac{\partial f}{\partial S}(\xi_t)$ denotes the first-order partial derivative of f with respect to the underlying, $\Gamma(\xi_t) \equiv \frac{\partial^2 f}{\partial S^2}(\xi_t)$ denotes the second order partial derivative with respect to the underlying, and $\Theta(\xi_t) \equiv \frac{\partial f}{\partial t}(\xi_t)$ denotes the first partial derivative with respect to the current time.

We take $\Delta t = 1$. Let E_t^1 denote the one-period hedge error from time t to $t+1$ and let $\{\gamma_t\}_{t=1}^T$ denote the hedging strategy. Neglecting the remainder term from now on, we get

$$\begin{aligned} E_t^1 &= \Delta f_t - \gamma_t \Delta S_t \\ &= (\gamma_t - \Delta(\xi_t)\Delta S_t) + \frac{1}{2}\Gamma(\xi_t)\,(\Delta S_t)^2 + \Theta(\xi_t). \end{aligned}$$

In general, the hedge errors $E_1^1, ..., E_T^1$ resulting from the hedge strategy $\{\gamma_t\}_{t=1}^T$ do not have the same distribution. To evaluate the performance of a hedge strategy, we want to "standardize" the hedge errors such that they have the same distribution. As reference distribution, we use the distribution, $\mathcal{L}\left(E_{t^*}^1\right)$, for some t^* such that $0 \le t^* < T$.

We assume strict stationarity of the differenced underlying process, implying,

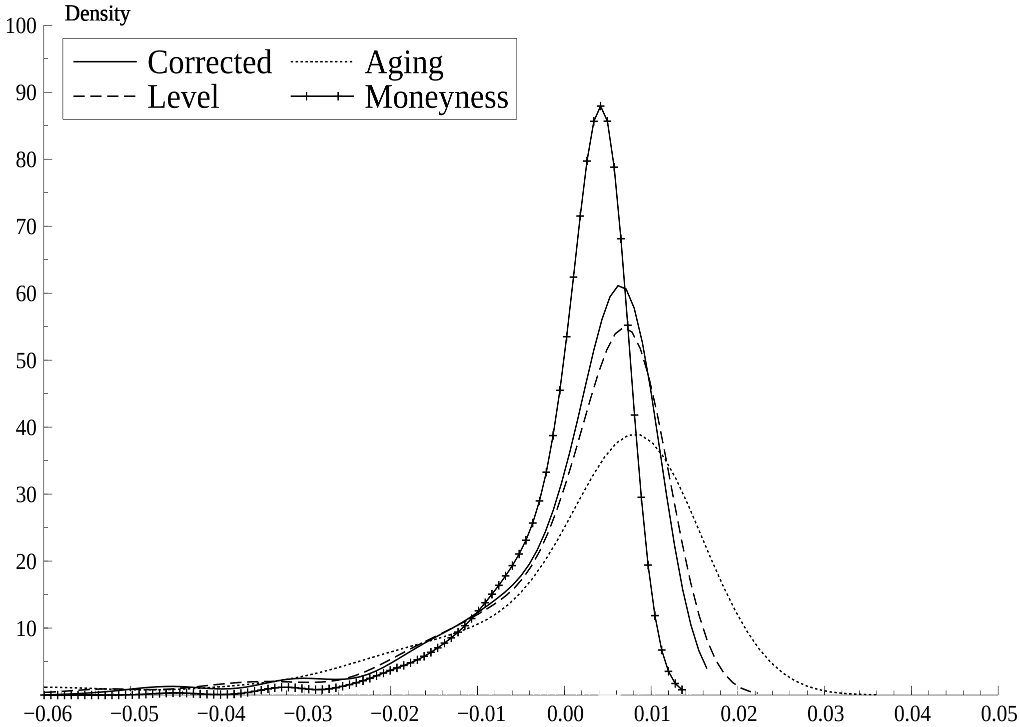
$$\mathcal{L}\,(\Delta S_t) = \mathcal{L}\,(\Delta S_{t^*}) \tag{7}$$

for $t = 1, ..., T$. Using the auxiliary process γ_t^*

$$\gamma_t^* = \Delta(\xi_t) + \frac{\Gamma(\xi_t)}{\Gamma(\xi_{t^*})}\,(\gamma_{t^*} - \Delta(\xi_{t^*}))\,,$$

we find the following relationship between the distributions of the hedge errors at dif-

Figure 2. Transformation This figure shows the effect of the transformation described in III.A. The graph shows the risk profile when the hedge errors of a delta hedged ATM option position are corrected for all effects, and the risk profiles corrected for all effects but the aging effect, the moneyness effect, and the level effect, respectively. The reference distribution is a one year ATM call option.



ferent times.

$$\mathcal{L}\left(E_{t^*}\right) = \mathcal{L}\left(\frac{\Gamma(\xi_{t^*})}{\Gamma(\xi_t)} E_t^1 + \frac{\Gamma(\xi_{t^*})}{\Gamma(\xi_t)}\left(\gamma_t^* - \gamma_t\right)\Delta S_t - \frac{\Gamma(\xi_{t^*})}{\Gamma(\xi_t)}\Theta(\xi_t) + \Theta(\xi_{t^*})\right). \quad (8)$$

In (8) we found a relation between the one-period hedge error from t^* to t^*+1 with characteristics $(S_{t^*}, m_{t^*}, \tau_{t^*})$ and the one-period hedge error from t to $t+1$ with characteristics (S_t, m_t, τ_t). Therefore, we can transform the data set of realizations drawn from not identically distributed distributions to one of realizations drawn from approximately identically distributed distributions. To obtain (8) we neglected the remainder term and used a parametric model in (5), so that this can only be seen as a good practical approximation and not as a strict identity.

Suppose we have a time series of hedge errors from a one year ATM call option. In Figure 2 we see the result of correcting the time series hedge errors for aging, moneyness, and level effect. This is in accordance with the "true" distribution determined by cross-sectional simulation. Furthermore, the distribution is given in case one of the corrections is left out. We see that the distribution is more spread out, if we leave out the aging effect correction. This follows from the fact that gamma is higher for short term options. Not correcting for the moneyness results in a less spread out distribution due to the fact that the gamma is lower for ITM and OTM options. Finally, we see that the distribution which is not corrected for level is rather similar to the corrected one. The level effect, however, becomes more important in case the sample is longer and the underlying moves further away from its starting position.

IV. Daily market risk forecasting methods

In this section, we discuss several methods that can be used to compute risk measures, such as value-at-risk and expected shortfall, of the daily market risk. In doing so, the models need to estimate $F \equiv \mathcal{L}\left(\tilde{E}_t^1\right)$ or more specifically the (joint) distribution of $\Delta f_t^{m_i}$ and $\Delta S_t^{m_i}$ for models $i = 1, ..., 5$. After applying the standardizing procedure in (8) we assume a stationary time series of[3]

$$\tilde{E}_t^{1,m_i} = \frac{\Delta f_t^{m_i} - \gamma_t \Delta S_t^{m_i}}{f_t - \gamma_t S_t}, \quad t = 1, ..., T,$$

with $\Delta f_t^{m_i} \equiv f_{t+1}^{m_i} - f_t$ and $\Delta S_t^{m_i} \equiv S_{t+1}^{m_i} - S_t$ where f_t and S_t denote observed prices.

We have returns data available of the underlying (S), the implied volatility (σ), the domestic interest rate (r^d), and the foreign interest rate (r^f), $h_{-N+1} = \left(h_{-N+1}^s, h_{-N+1}^\sigma, h_{-N+1}^{r^d}, h_{-N+1}^{r^f}\right), ..., \left(h_T^s, h_T^\sigma, h_T^{r^d}, h_T^{r^f}\right)$, with $h_t^x = \log\left(x_t / x_{t-1}\right)$. From these data we use $\{h_1, ..., h_T\}$ for testing and denote it as the testing sample. The testing sample is used in the backtest to determine the quality of the method. All the models discussed below are used to estimate the distribution of the relative one-period hedge error $\tilde{E}^1, m_i$, denoted by F^{m_i}.

[3]The time series is also assumed to be ergodic and satisfy the necessary regularity conditions needed for Central limit theorems used later on.

For notational convenience, we neglect the dependence of $\tilde{E}^{1,m_i}$, $\Delta f_t^{m_i}$, and $\Delta S_t^{m_i}$ on m_i in the following enumeration of models.

1. Method 1 is a naive method which more or less follows the Black-Scholes world assumptions, but with potentially changing mean and volatility. It assumes that $\mathcal{L}\left(h_{t+1}^s\right) = \mathcal{N}\left(\mu_t, \sigma_t^2\right)$ and that σ, r^d, and r^f are constant. To estimate μ_t and σ_t^2 we use the returns data of the underlying, $h_t^s, ..., h_{t-N}^s$, to get μ_t and σ_t^2, the so-called rolling window estimators for μ_t and σ_t^2. For $t = 1, ..., T$ we draw h_t^s from $\mathcal{N}\left(\mu_t, \sigma_t^2\right)$ to construct $(\Delta S_t)_{t=1}^T$ and $(\Delta f_t)_{t=1}^T$.[4] Given the hedge strategy γ we construct $\left(\tilde{E}_t^1\right)_{t=1}^T$ from which we produce an estimate $\hat{\varrho}^{m_1} = \varrho\left(\hat{F^{m_1}}\right)$.

2. Method 2 is a historical simulation method for the underlying asset. The implied volatilities, domestic and foreign interest rates are as in method 1. Method 2 allows a distribution for the underlying that differs from the normal distribution. It assumes $(h_t^s)_{t=-N+1}^T$ is an *i.i.d.* sample. We estimate the distribution, $\mathcal{L}\left(h_t^s\right)$, by the empirical distribution of $(h_{t^*}^s)_{t^*=t-N+1}^t$.[5] Drawing (with replacement) from $(h_{t^*}^s)_{t^*=t-N+1}^t$ for $t = 1, ..., T$ allows us to construct $(\Delta S_t)_{t=1}^T$ and $(\Delta f_t)_{t=1}^T$. Given the hedge strategy γ we construct $\left(\tilde{E}_t^1\right)_{t=1}^T$ from which we produce an estimate $\hat{\varrho}^{m_2} = \varrho\left(\hat{F^{m_2}}\right)$.

3. Method 3 is a full historical simulation method. This type of method is often used in practice and assumes that $(h_t)_{t=-N+1}^T$ is an *i.i.d.* sample. We estimate the distribution, $\mathcal{L}\left(h_t\right)$ by the empirical distribution of $(h_{t^*})_{t^*=t-N+1}^t$. Drawing h_t (with replacement) from $(h_{t^*})_{t^*=t-N+1}^t$ for $t = 1, ..., T$ allows us to construct to construct $(\Delta S_t)_{t=1}^T$ and $(\Delta f_t)_{t=1}^T$. Given the hedge strategy γ we construct $\left(\tilde{E}_t^1\right)_{t=1}^T$ from which we produce an estimate $\hat{\varrho}^{m_3} = \varrho\left(\hat{F^{m_3}}\right)$.

4. In method 4 a first-order Vector AutoRegressive (VAR) model for estimation of

[4]Note that the sequence $(\Delta S_t)_{t=1}^T$ is not used to produce a price path $(S_t)_{t=1}^T$ of the underlying. It only serves to compute a series of hedge errors. The price path of the underlying is given by the data.

[5]Considering the stationarity assumption, it would be more efficient to use all available data, but we use this nonparametric rolling window estimator because it is often used in practice.

the distribution of $(S_{t+1}, \sigma_{t+1})_{t=1}^{T}$ is estimated using $(h_{t^*})_{t^*=t-N+1}^{t}$ for $t = 1, ..., T$

$$\begin{aligned} \begin{bmatrix} h_{t+1}^{s} \\ h_{t+1}^{\sigma} \end{bmatrix} &= \Phi_0 + \Phi_1 \begin{bmatrix} h_t^{s} \\ h_t^{\sigma} \end{bmatrix} + \begin{bmatrix} u_{t+1}^{s} \\ u_{t+1}^{\sigma} \end{bmatrix} & (9) \\ h_{t+1} &= \Phi_0 + \Phi_1 h_t + u_{t+1}, \quad t = 1, ..., T & (10) \end{aligned}$$

with

$$\mathcal{L}\left(u_{t+1} | \mathcal{F}_t\right) = \mathcal{N}\left(0, \Sigma_t\right),$$

where $\mathcal{F}_t$ denotes the information set at point t. This gives $(\Phi_{0,t}, \Phi_{1,t})_{t=1}^{T}$, and $(\Sigma_t)_{t=1}^{T}$ to generate $(\Delta S_{t+1}, \Delta\sigma_{t+1})$ for $t = 1, ..., T$ and $(\Delta f_t)_{t=1}^{T}$. Given the hedge strategy γ we construct $\left(\tilde{E}_t^1\right)_{t=1}^{T}$ from which we produce an estimate $\hat{\varrho}^{m_4} = \varrho\left(F^{\hat{m}_4}\right)$.

5. In method 5 a first-order Vector AutoRegressive (VAR) model for estimation of the distribution of $(S_{t+1}, \sigma_{t+1})_{t=1}^{T}$ is estimated using $(h_{t^*}^{s}, h_{t^*}^{\sigma})_{t^*=t-N+1}^{t}$ for $t = 1, ..., T$

$$\begin{aligned} \begin{bmatrix} h_{t+1}^{s} \\ h_{t+1}^{\sigma} \end{bmatrix} &= \Phi_0 + \Phi_1 \begin{bmatrix} h_t^{s} \\ h_t^{\sigma} \end{bmatrix} + \begin{bmatrix} u_{t+1}^{s} \\ u_{t+1}^{\sigma} \end{bmatrix} & (11) \\ h_{t+1} &= \Phi_0 + \Phi_1 h_t + u_{t+1}, \quad t = 1, ..., T & (12) \end{aligned}$$

with

$$\mathcal{L}\left(u_{t+1} | \mathcal{F}_t\right) = F_t^N,$$

where F_t^N denotes the empirical distribution function of u at time t estimated from $u_{t-N+1}, ..., u_{t-1}$. This gives $(\Phi_{0,t}, \Phi_{1,t})_{t=1}^{T}$, and $(\Sigma_t)_{t=1}^{T}$ to generate $(\Delta S_{t+1}, \Delta\sigma_{t+1})$ for $t = 1, ..., T$ and $(\Delta f_t)_{t=1}^{T}$. Given the hedge strategy γ we construct $\left(\tilde{E}_t^1\right)_{t=1}^{T}$ from which we produce an estimate $\hat{\varrho}^{m_5} = \varrho\left(F^{\hat{m}_5}\right)$.

V. Test procedure

In this section, we present a test to evaluate daily market risk evaluation models described in Section IV. Time t runs from $-N+1$ to T. The last T observations are used for testing. At each point in time the method is estimated from the previous N observations, that is, we use the so-called rolling window estimator.

The predicted daily market risk will be obtained from the distribution F^{m_i} of[6]

$$\tilde{E}_t^{1,m_i} = \frac{\Delta f_t^{m_i} - \gamma_t \Delta S_t^{m_i}}{f_t - \gamma_t S_t}, \tag{13}$$

while the actual daily market risk is induced by the distribution F of

$$\tilde{E}_t^{1} = \frac{\Delta f_t - \gamma_t \Delta S_t}{f_t - \gamma_t S_t} \tag{14}$$

We would like to test whether the predicted risk measures are the same for the method hedge errors as for the empirical hedge errors. Let $\varrho\left(F^{m_i}\right)$ represent the characteristic of interest of F^{m_i} and let $\varrho\left(F\right)$ represent the corresponding characteristic of interest of F.

Denote by $\varrho\left(\hat{F}^{m_i}\right)$ an appropriate estimator for $\varrho\left(F^{m_i}\right)$ such that

$$\sqrt{T}\left(\varrho\left(\hat{F}^{m_i}\right) - \varrho\left(F^{m_i}\right)\right) = \sqrt{T}\frac{1}{T}\sum_{t=1}^{T}\Psi_t^{m_i} + o_p\left(1\right), \qquad \mathbb{E}\Psi_t^{m_i} = 0, \mathbb{E}\left(\Psi_t^{m_i}\right)^2 < \infty, \tag{15}$$

and, similarly, let $\varrho\left(\hat{F}\right)$ be an appropriate corresponding estimator for $\varrho\left(F\right)$ such that

$$\sqrt{T}\left(\varrho\left(\hat{F}\right) - \varrho\left(F\right)\right) = \sqrt{T}\frac{1}{T}\sum_{t=1}^{T}\Psi_t + o_p\left(1\right), \qquad \mathbb{E}\Psi_t = 0, \mathbb{E}\left(\Psi_t\right)^2 < \infty, \tag{16}$$

where $\Psi_t^{m_i}$ and Ψ_t are called the influence functions. In Appendix A the influence functions for VaR and expected shortfall are given. Then, under the null hypothesis

[6] Notice the difference in notation: $\Delta x_t^{m_i} \equiv x_{t+1}^{m_i} - x_t$ and $\Delta x_t \equiv x_{t+1} - x_t$ for $x = S, f$. In both cases we use the observed prices as starting point.

$H_0 : \varrho\left(F^{m_i}\right) = \varrho\left(F\right)$, we have

$$\sqrt{T}\left(\varrho\left(\hat{F}^{m_i}\right) - \varrho\left(F^{m_i}\right)\right) - \sqrt{T}\left(\varrho\left(\hat{F}\right) - \varrho\left(F\right)\right) = \\ \begin{bmatrix} 1 & -1 \end{bmatrix} \sqrt{T}\frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} \Psi_t^{m_i} \\ \Psi_t \end{bmatrix} + o_p\left(1\right) \xrightarrow[H_0]{d} N\left(0, V\right) \quad (17)$$

with

$$V = \begin{bmatrix} 1 & -1 \end{bmatrix}\left[\lim_{T\to\infty}\mathbb{E}\left[T^{-1}\left(\sum_{t=1}^{T}\begin{bmatrix} \Psi_t^{m_i} \\ \Psi_t \end{bmatrix}\begin{bmatrix} \Psi_t^{m_i} \\ \Psi_t \end{bmatrix}'\right)\right]\right]\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (18)$$

So, with $\hat{V}$ (using, for example, the estimator of Newey and West (1987)) satisfying $\hat{V} \xrightarrow{p} V$, we can take as a test statistic

$$T\frac{\left(\varrho\left(\hat{F}^{m_i}\right) - \varrho\left(\hat{F}\right)\right)^2}{\hat{V}} \xrightarrow[H_0]{d} \chi_1^2 \quad (19)$$

Since we can simulate from F^{m_i} as often as we would like, we can strengthen the test above by using $\bar{\varrho} \equiv \frac{1}{K}\sum_{k=1}^{K}\varrho_k(\hat{F}_t^{m_i})$, with K equal to the number of trials, instead of $\hat{\varrho}(F_t^{m_i})$. This gives for fixed K

$$\sqrt{T}\left(\bar{\varrho} - \varrho(F^{m_i})\right) = \sqrt{T}\frac{1}{TK}\sum_{k=1}^{K}\sum_{t=1}^{T}\Psi_{t,k}^{m_i} + o_p\left(1\right), \\ \mathbb{E}\Psi_{t,k}^{m_i} = 0, \mathbb{E}\left(\Psi_{t,k}^{m_i}\right)^2 < \infty. \quad (20)$$

The expression in (20) converges in probability to zero as $K \to \infty$ and so we can take as a test statistic

$$T\frac{\left(\bar{\varrho} - \varrho\left(\hat{F}\right)\right)^2}{\hat{v}} \xrightarrow[H_0]{d} \chi_1^2, \quad (21)$$

where $\hat{v}$ denotes a consistent estimator for

$$v = \lim_{T\to\infty}\mathbb{E}\left[T^{-1}\sum_{t=1}^{T}\Psi_t^2\right]. \quad (22)$$

VI. Empirical Results

A. FX market

The FX market is by far the most liquid market in the world with a daily turnover of about 1.5 trillion US dollars (for comparison, the NYSE has a daily turnover of about 30 billion US dollar). In this section, we apply the test outlined above to call options on the dollar-yen, dollar-pound, and pound-dollar exchange rates. Quotes are in implied volatilities in the FX market and prices can be computed using the Garman-Kohlhagen model (see Garman and Kohlhagen (1983)). This is a version of the Black-Scholes model applicable to currency options. Call option prices are given by

$$c^{GK}\left(S,k,r^d,r^f,\sigma,\tau\right) = S_t e^{-r^f\tau}\Phi\left(d_+\right) - Xe^{-r^d\tau}\Phi\left(d_-\right), \tag{23}$$

where

$$d_\pm = \frac{\log\left(S_t/k\right)+\left(r^d-r^f\right)\tau}{\sigma\sqrt{\tau}} \pm \frac{1}{2}\sigma\sqrt{\tau}. \tag{24}$$

r^d is the domestic instantaneous riskless interest rate, r^f is the foreign instantaneous riskless interest rate, σ denotes the instantaneous volatility of the exchange rate and $\Phi\left(\cdot\right)$ denotes the Gaussian cumulative distribution function.

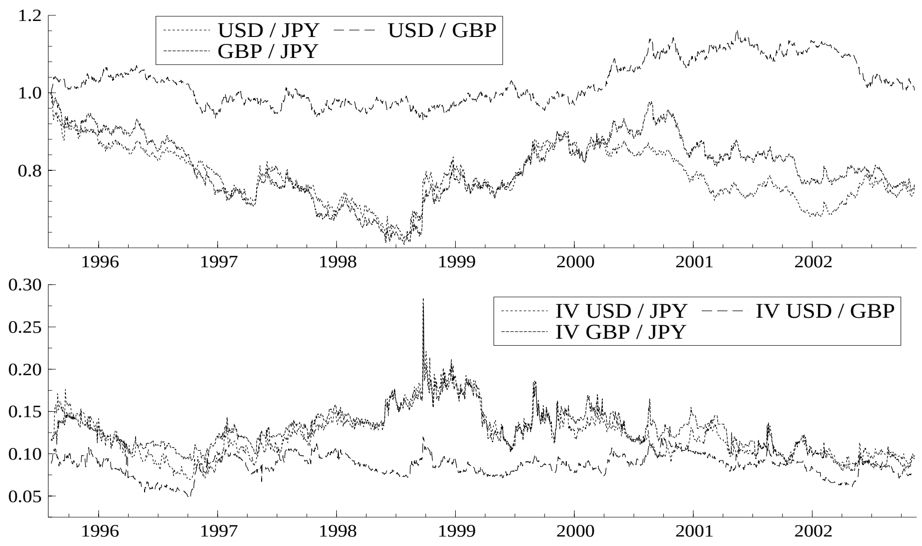
The daily data available consist of implied volatilities of 3 month ATM call options on dollar-yen, dollar-pound, and pound-dollar exchange rates, the corresponding exchange rates, and the US, UK, and Japanese interest rates.[7] The data run from August 9, 1995 until December 13, 2002 and are shown in Figure 3.

This results in 1918 data points. We use a two year rolling window estimation period for all the models. Taking the number of trading days per year equal to 250 gives us estimation periods of 500 observations and 1418 observations for testing. In Kerkhof and Melenberg (2002) it is argued that for fair comparison with a 1% value-at-risk the level of expected shortfall should be about 2.5%.[8] The quality of the models is

[7]The data have been kindly shared by ABN-AMRO Bank.

[8]This argument is based on the normal distribution, but seems to be approximately correct in our sample.

Figure 3. FX data In the upper panel the normalized price paths of the USD/JPY, USD/GBP, and GBP/JPY are given. In the lower panel the implied volatilities for the 3m ATM call options are given.



tested by tests whether the variances, the 1%-value-at-risk, and 2.5% expected shortfall of the hedge error as predicted by the models and empirical hedge errors are equal.[9] The level for value-at-risk is chosen at 1% such that it equals the current level in the 1996 amendment to the Basle Accord (see Basle Committee on Banking Supervision (1996)). Table I reports the variance, 1% value-at-risk, and 2.5% expected shortfall for an investment of \$100 in a portfolio of ATM call options and the underlying exchange rate with as ratio the hedge strategy.

For all exchange rates we find that the methods 1 and 2 is rejected for all risk measures. The full historical simulation method (method 3) performs well for all exchange rates and all risk measures. The parametric VAR method, method 4, is rejected for the USD/JPY exchange rate for being too conservative, while it is rejected in the GBP/JPY exchange rate for underestimating the risk. The nonparametric VAR method, method 5, is conservative in all markets and is rejected for the USD/JPY and USD/GBP exchange

[9]In the absence of data on ITM and OTM options, we have assumed a flat volatility smile for the FX options. Since we are looking at one-day hedge errors and the FX volatility smile is rather flat near the money, this should not lead to severe biases.

Table I
Tests of Risk measures for the 3 month exchange rates

This table shows the empirical standard deviations, $\mathrm{VaR}_{0.01}$, and $\mathrm{ES}_{0.025}$ and those obtained from methods 1,...,5 for the USD/JPY, GBP/JPY, and USD/GBP exchange rate. We test whether the method predictions correspond to the empirical quantities. The p-values of these tests are given in parentheses. In order to reduce sampling error we used $K = 10,000$.

	emp	method 1	method 2	method 3	method 4	method 5
USD/JPY						
std. dev.	0.23	0.05 (0.00)	0.09 (0.00)	0.20 (0.36)	0.22 (0.85)	0.21 (0.63)
$\mathrm{VaR}_{0.01}$	-0.58	−0.18 (0.00)	−0.25 (0.00)	−0.62 (0.39)	−0.80 (0.00)	−0.72 (0.00)
$\mathrm{ES}_{0.025}$	-0.74	−0.19 (0.00)	−0.33 (0.00)	−0.70 (0.74)	−0.90 (0.43)	−0.85 (0.81)
USD/GBP						
std. dev.	0.11	0.04 (0.00)	0.04 (0.00)	0.11 (0.19)	0.12 (0.09)	0.13 (0.00)
$\mathrm{VaR}_{0.01}$	-0.30	−0.13 (0.00)	−0.17 (0.00)	−0.35 (0.11)	−0.35 (0.12)	−0.44 (0.00)
$\mathrm{ES}_{0.025}$	-0.38	−0.14 (0.00)	−0.19 (0.00)	−0.40 (0.19)	−0.36 (0.75)	−0.47 (0.02)
GBP/JPY						
std. dev.	0.19	0.06 (0.00)	0.08 (0.00)	0.17 (0.27)	0.16 (0.07)	0.17 (0.30)
$\mathrm{VaR}_{0.01}$	-0.52	−0.19 (0.00)	−0.30 (0.00)	−0.54 (0.78)	−0.40 (0.03)	−0.52 (0.91)
$\mathrm{ES}_{0.025}$	-0.61	−0.20 (0.00)	−0.34 (0.00)	−0.60 (0.83)	−0.41 (0.02)	−0.61 (0.92)

rates.

B. S&P 500 options

We have available option data on the S&P 500 ranging from January 2, 1992 till August 29, 1997. Quotes on the options are the end-of-day quotes with synchronous observations of the underlying index. For the FX options we have data on fixed time to maturity and moneyness options available. For the S&P 500 we have fixed maturity and varying moneyness option data. Therefore, we apply the transformation method of Section III.A. We analyze the models for calculating the risk measures for 3 month ATM options. For this we use the options with time to maturity closest to 3 months and closest to the ATM level. Again we investigate a portfolio of $100 invested in options and the underlying asset. As hedge ratio we apply the standard Black-Scholes delta with continuous dividend yield.

We find that the empirical risks for S&P500 options are higher than for the FX options. We find that the positions in the 1 year options are more risky than the positions in the 3 months options. For the tests of the S&P500 options we get more or less the same results as for the FX options. Only models 3, 4, and 5 have a acceptable prediction behavior.

Overall, we see that models 1 and 2 do not perform well and underestimate the risk of delta hedged derivatives positions in almost all cases. This can be explained by the fact that they do not take fluctuations in the levels of implied volatilities into account. The historical simulation method and both VAR models perform about the same, although the VAR models for changes in the underlying and implied volatilities are sometimes a bit too conservative. Since the historical simulation method and the VAR model with historical simulation take more time to compute than the Gaussian VAR model where VaR and ES can be computed analytically, it seems easiest to compute both VaR and ES based on the Gaussian VAR model.

Table II
Tests of risk measures for delta hedged 3 month S&P 500 options

This table shows the empirical standard deviations, $\mathrm{VaR}_{0.01}$, and $\mathrm{ES}_{0.025}$ and those obtained from methods 1,...,5 for a delta-hedged positions in 3 month ATM S&P 500 options. We test whether the method predictions correspond to the empirical quantities. The p-values of these tests are given in parentheses. In order to reduce sampling error we used $K = 10,000$.

	emp	method 1	method 2	method 3	method 4	method 5
ATM 3 months						
std. dev.	0.24	0.04	0.05	0.23	0.24	0.23
		(0.00)	(0.00)	(0.34)	(0.83)	(0.55)
$\mathrm{VaR}_{0.01}$	-0.74	-0.13	-0.19	-0.67	-0.80	-0.71
		(0.00)	(0.00)	(0.27)	(0.42)	(0.55)
$\mathrm{ES}_{0.025}$	-0.85	-0.13	-0.19	-0.76	-0.83	-0.79
		(0.00)	(0.00)	(0.35)	(0.73)	(0.48)
ATM 1 year						
std. dev.	0.41	0.06	0.06	0.43	0.43	0.43
		(0.00)	(0.00)	(0.50)	(0.49)	(0.28)
$\mathrm{VaR}_{0.01}$	-1.26	-0.13	-0.17	-1.39	-1.12	-1.42
		(0.00)	(0.00)	(0.18)	(0.15)	(0.10)
$\mathrm{ES}_{0.025}$	-1.27	-0.17	-0.19	-1.41	-1.16	-1.44
		(0.00)	(0.00)	(0.22)	(0.12)	(0.12)

VII. Conclusions

In this paper we tested several risk management models for computing expected shortfall and value-at-risk for one-period hedge errors of hedged derivatives positions. Though value-at-risk can be tested using a binomial test, this is not the case for expected shortfall and we need information of the distribution in the tail. By nature, the characteristics of derivatives positions are changeable and as a consequence the size of risk exposures varies over time. To overcome this problem, we present a transformation procedure.

We empirically test the performance of several models, based on tests for standard deviation, value-at-risk, and expected shortfall. We find that in order to get good indication of the risk of a hedged derivative in both the FX and the equity market it is of crucial importance to take the variation in the implied volatilities into account. We find that a historical simulation method, which is commonly used in practice, produces the best results. A parametric and non-parametric VAR model perform reasonably well, but their performance trails that of the historical simulation method.

A. Influence functions for value-at-risk and expected shortfall

Let F_t denote the distribution of the one-day hedge error E_t^1. The influence functions of value-at-risk and expected shortfall are then given by:

1. Value-at-risk: In the case of VaR_p the influence function $\Psi\left(F_t\right)$ is given by

$$\Psi_{\mathrm{VaR}}\left(F_t\right) = \frac{p - \mathbf{I}_{\left[-\infty, F_t^{-1}(p)\right]}\left(x\right)}{q\left(F_t^{-1}\left(p\right)\right)}, \tag{25}$$

   and

$$\mathbb{E}\Psi_{\mathrm{VaR}}^2\left(F_t\right) = \frac{p\left(1-p\right)}{q^2\left(F_t^{-1}\left(p\right)\right)}. \tag{26}$$

2. Expected shortfall: In the case of ES_p the influence function $\Psi\left(F_t\right)$ is given by

$$\begin{aligned}\Psi_{\mathrm{ES}}\left(F_t\right) &= -\frac{1}{p}\left[\left(x - F_t^{-1}\left(p\right)\right)\mathbf{I}_{\left[-\infty, F_t^{-1}(p)\right]}\left(x\right)\right. \\ &\quad \left. + \Psi_{\mathrm{VaR}}\left(F_t\right)\left(p - \int_{-\infty}^{F_t^{-1}(p)} dF\left(x\right)\right)\right] - \mathrm{ES}\left(F_t\right) + \mathrm{VaR}\left(F_t\right)\end{aligned} \tag{27}$$

   and

$$\begin{aligned}\mathbb{E}\Psi_{\mathrm{ES}}^2\left(F_t\right) &= \frac{1}{p}\mathbb{E}\left[X^2 | X \leq F_t^{-1}\left(p\right)\right] - \mathrm{ES}\left(F_t\right)^2 \\ &\quad + 2\left(1 - \frac{1}{p}\right)\mathrm{ES}\left(F_t\right)\mathrm{VaR}\left(F_t\right) - \left(1 - \frac{1}{p}\right)\mathrm{VaR}\left(F_t\right)^2.\end{aligned} \tag{28}$$

See Kerkhof and Melenberg (2002) for derivations.

References

Acerbi, C. and Tasche, D.: 2002, On the coherence of expected shortfall, *Journal of Banking and Finance* **26**, 1487–1503.

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D.: 1999, Coherent measures of risk, *Mathematical Finance* **9**, 203–228.

Basle Committee on Banking Supervision: 1996, *Supervisory Framework for the Use of "Backtesting" in Conjunction with the Internal Models Approach to Market Risk Capital Requirements,* Bank for International Settlements, Basle.

Berkowitz, J. and O'Brien, J.: 2002, How accurate are value-at-risk models at commercial banks?, *Journal of Finance* **57**, 1093–1111.

Boyle, P. and Emanuel, D.: 1980, Discretely adjusted option hedges, *Journal of Financial Economics* **8**, 259–282.

Christoffersen, P., Hahn, J. and Inoue, A.: 2001, Testing and comparing value-at-risk measures, *Journal of Empirical Finance* **8**, 325–342.

Delbaen, F.: 2000, Coherent risk measures on general probability spaces, *Working paper ETH* pp. 1–35.

Dowd, K.: 1998, *Beyond Value at Risk: The New Science of Risk Management*, John Wiley & Sons, England.

El-Jahel, L., Perraudin, W. and Sellin, P.: 1999, Value at risk for derivatives, *Journal of Derivatives* **6**, 7–26.

Garman, M. B. and Kohlhagen, S. W.: 1983, Foreign currency option values, *Journal of International Money and Finance* **2**, 231–237.

Jorion, P.: 2000, *Value at Risk: The New Benchmark for Managing Financial Risk*, 2 edn, McGraw-Hill, New York.

Kerkhof, J. and Melenberg, B.: 2002, Backtesting for risk-based regulatory capital, *mimeo Tilburg University* pp. 1–30.

Kupiec, P. H.: 1995, Techniques for verifying the accuracy of risk management models, *Journal of Derivatives* **3**, 73–84.

McNeil, A. and Frey, R.: 2000, Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach, *Journal of Empirical Finance* **7**, 271–300.

Newey, W. and West, K.: 1987, A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**, 703–708.

Tasche, D.: 2002, Expected shortfall and beyond, *Journal of Banking and Finance* **26**, 1519–1533.

Yamai, Y. and Yoshiba, T.: 2002, Comparative analyses of expected shortfall and value-at-risk: Their estimation error, decomposition, and optimization, *Monetary and Economic Studies* **20**, 87–122.