

# GARCH vs. Stochastic Volatility: Option Pricing and Risk Management\*

Alfred Lehar  
Department of Business Studies  
University of Vienna †

Martin Scheicher  
Central Bank of Austria ‡

Christian Schittenkopf  
Austrian Research Institute  
for Artificial Intelligence §

August 31, 2001

---

\*This work was supported by the Austrian Science Fund (FWF) within the research project “Adaptive Information Systems and Modelling in Economics and Management Science” (SFB 010). The Austrian Research Institute for Artificial Intelligence is supported by the Austrian Federal Ministry of Education, Science and Culture. The opinions in this paper do not reflect those of the Central Bank of Austria. We thank Thomas Kaiser, Peter Schaller, Stephen D. Smith and Peter Vlaar and participants of the Chicago Fed / JBF conference for helpful comments.

†Department of Business Studies, University of Vienna, Brünnerstrasse 72, 1210 Vienna, Austria, Phone: +43-1-4277-38077, Fax: +43-1-4277-38074, Alfred.Lehar@univie.ac.at

‡Corresponding author: Otto-Wagner Platz 3, 1011 Vienna, Austria, Phone: +43-1-40420-7418, Fax: +43-1-40420-7499, Martin.Scheicher@oenb.co.at

§Austrian Research Institute for Artificial Intelligence, Schottengasse 3, 1010 Vienna, Austria, Phone: +43-1-5336112-16, Fax: +43-1-5336112-77, chris@ai.univie.ac.at

# GARCH vs. Stochastic Volatility: Option Pricing and Risk Management

## Abstract

In this paper we compare the out-of-sample performance of two common extensions of the Black-Scholes option pricing model, namely GARCH and stochastic volatility. We calibrate the three models to intraday FTSE 100 option prices and apply two sets of performance criteria, namely out-of-sample valuation errors and Value-at-Risk oriented measures. When we analyze the fit to observed prices, GARCH clearly dominates both stochastic volatility and the benchmark Black-Scholes model. However, the predictions of the market risk from hypothetical derivative positions show sizable errors. The fit to the realized profits and losses is poor and there are no notable differences between the models. Overall, we therefore observe that the more complex option pricing models can improve on the Black-Scholes methodology only for the purpose of pricing, but not for the Value-at-Risk forecasts.

JEL codes: G13, G15, C15

# 1. Introduction

The correct valuation of derivatives is of crucial importance for practitioners in any financial market. Derivatives are now a key component of investors' portfolios. This development is reflected by the fact that turnover and volume in these products experienced phenomenal growth since the 70's. For market participants, the main issue is that the prices obtained from the standard Black-Scholes model differ significantly from observed prices. These systematic valuation errors are documented in a stylized fact called the "smile" effect: When the volatilities backed out from option prices are plotted against moneyness and maturity, the resulting surface deviates significantly from the flat surface, which the theoretical model predicts. These empirical biases reflect the fact that in reality volatility is not constant but time-dependent. This behavior is in contrast to the constant variance framework of geometric Brownian Motion, which underlies the Black-Scholes method. The pricing errors resulting from this unrealistic assumption have serious consequences for market participants measuring the market risk of their portfolios. During the last couple of years regulators have allowed financial institutions to use their internal risk models to measure market risk and to allocate economic capital. For these purposes, Value-at-Risk (VaR)<sup>1</sup> has become the most common concept. It measures the potential portfolio loss caused by adverse price moves during a holding period of one or ten days with a given probability. Based on the Capital Adequacy Directive supervisory authorities require large banks and securities firms to perform daily VaR calculations of their portfolios. Option pricing models are a crucial component of the marking-to-model systems with which banks calculate the regulatory and economic capital for their trading activities. Computing the potential loss for derivative positions is particularly challenging for a number of reasons. First, options are nonlinear instruments because the distribution of option returns is a complex function of the underlying asset's return distribution. The number of risk factors increases as well as the Vega needs

---

<sup>1</sup>Duffie and Pan (1997) or Jorion (1997) survey the VaR concept. See also Vlaar (2000) or Christoffersen, Hahn, and Inoue (2001).

to be captured. Then the distribution of option returns with respect to the underlying is skewed meaning that the normal distribution is an approximation of limited accuracy. For these reasons, practitioners working in risk management have renewed the interest in the pricing methods proposed in the academic literature<sup>2</sup>.

The purpose of this paper is to compare option pricing models based on three commonly used parameterizations of volatility: Constant volatility, GARCH and stochastic volatility (SV)<sup>3</sup>. We perform this evaluation with two sets of criteria: statistical and economic loss functions. First, we measure the errors relative to observed option prices. Then we test the suitability of the three models for forecasting the market risk of hypothetical positions. In this context the constant volatility method is the natural benchmark. By reproducing the smile effect, a more complex specification should lower the pricing errors. Earlier comparative studies<sup>4</sup> document that the pricing errors of the Black-Scholes model can indeed be reduced if better specifications for the stochastic process of the underlying instrument are used. The focus of our comparative study is on the performance of the three valuation methods for risk management. We thus extend the literature by analyzing the VaR for hypothetical options positions. So far, there is no evidence on how the pricing models perform when they are applied to calculate the VaR of derivatives. In this perspective, our study closes a gap which is of particular interest for practitioners.

We proceed in three steps. First, we estimate the three stochastic processes from high-frequency FTSE 100 option prices by means of a Nonlinear Least Squares procedure. Thereby, the one-parameter case of constant volatility is generalized to more complex stochastic processes. The resulting estimates are the risk-neutral parameters. This is a clear advantage over methods where parameters are first estimated from return

---

<sup>2</sup>A detailed overview is offered by Bates (1996).

<sup>3</sup>The GARCH option pricing model was introduced by Duan (1995a). SV models have been proposed by Hull and White (1987) or Heston (1993).

<sup>4</sup>Prominent articles are Engle and Mustafa (1992), Duan and Zhang (1999), Bakshi, Cao, and Chen (1997), Bakshi, Cao, and Chen (2000) and Nandi (1998).

series and then adjusted for risk-neutral valuations. Therefore, our information set is not restricted to historical stock prices but contains the market expectations about the future price process. By means of this unified estimation, our study directly illustrates the advantages and disadvantages of GARCH and SV, which represent the two most popular variance models. In the second step we measure the out-of-sample pricing performance by the median errors relative to observed prices and by the extent to which moneyness and maturity biases are reduced. We find that a more complex volatility model outperforms the simpler implied volatility approach. However, model selection is crucial: Only GARCH offers significant improvements in the pricing performance. In contrast to that, the median pricing errors of SV are close to those of the benchmark model of constant volatility. The third step is a risk management application where we compute the next-day VaR by means of a Monte Carlo simulation with full valuation. We compare the fit to the profits and losses of hypothetical positions. Our tests indicate a weak fit for the modeling of the tails of the profit/loss distribution. In particular, GARCH and SV are unable to improve on the benchmark Black-Scholes model. Thus, we can conclude that the rankings of the models depend on the performance criteria. Although we find remarkable improvements in the out-of-sample pricing performance, these are not reflected in the forecasting of the distribution of the changes in portfolio value. Therefore, despite the decrease in the pricing errors, we can not improve the modeling of actual next-day profits/losses in options trading. These differences in the quality of the methodologies are noteworthy because so far the literature has mostly been oriented towards pricing performance.

The rest of this paper is composed as follows: In section 2 the data sets are described. Section 3 gives a detailed description of the three option pricing models. Section 4 contains the empirical results. In section 5 we report the estimation results and findings from the VaR application. In section 6 we summarize our results and provide suggestions for future research.

## 2. Sample and Estimation Method

Our sample comprises transactions data of European options on the FTSE 100 (index) traded at the London International Financial Futures and Options Exchange (LIFFE). The option prices are recorded synchronously with the FTSE 100 and time-stamped to the nearest second. Our high-frequency observations start on 4 January 1993 and end on 22 October 1997 covering a period of 1210 trading days. The sample comprises 102,211 traded call and put option contracts.

In a first step the data are carefully checked for recording errors and for prices violating arbitrage bounds. Then all contracts with a time to maturity of less than two weeks, with a price below 5 points<sup>5</sup>, with moneyness<sup>6</sup> outside  $[-0.1, 0.1]$ , and with an annualized implied volatility below 5% or above 50% are removed. The final set consists of 65,549 option contracts (33,633 call options and 31,916 put options).

A brief description of the option contracts is given in Table 1. We report the minimum, the 10th, the 25th, the 50th, the 75th and the 90th percentile, and the maximum of several sample characteristics. During the sample period the FTSE 100 rose from about 2700 points to more than 5000 points. Therefore, strike prices range from 2525 to 5875. The average time to maturity of all option contracts is 42 days. The longest contract has a time to maturity of more than one year. In Table 1 we also report the statistics of the risk-free interest rate<sup>7</sup> and the ex post dividend yield of the FTSE 100, which are inputs to the option pricing models. Finally, the percentiles of the implied volatilities of all options are reported. The average implied volatility is about 14%, and 80% of all contracts have an implied volatility between 9% and 20%.

---

<sup>5</sup>The accuracy of the recorded prices is 0.5 points. For options with low prices, pricing errors which are due to rounding effects, may thus be substantial.

<sup>6</sup>The relation of the option's strike price to the observed index value defines the moneyness of the option. This ratio is defined as  $(\text{index value} - \text{strike price})/\text{strike price}$  for calls and  $(\text{strike price} - \text{index value})/\text{strike price}$  for puts, respectively.

<sup>7</sup>Daily GBP-LIBOR quotes for overnight, one week, one month, three months, six months, and one year were linearly interpolated.

[Table 1 about here.]

Since the sample comprises 65,549 option contracts traded on 1210 days, the average number of contracts per day is about 54. However, trading activity is not equally distributed. On the least active day, which is 4 May 1993, only 7 trades were recorded whereas 351 transactions took place on the most active day, which is 16 July 1997. Therefore, it is not meaningful to reestimate the option pricing models every day. On the other hand, we need to reestimate the models within relatively short time intervals to allow for changing parameters. In our experimental setup a *sliding window technique* is applied. The size of the time window is ten trading days meaning that option contracts are collected over a period of ten days. From this set, the parameters are estimated by minimizing the (average) squared relative pricing error

$$\text{SRPE} = \frac{1}{N} \sum_{i=1}^N \left( \frac{\hat{p}_i - p_i}{p_i} \right)^2 \quad (1)$$

where  $N$  is the number of contracts in the time window and  $p_i$  and  $\hat{p}_i$  denote the observed and the theoretical option price, respectively. We choose relative pricing errors because absolute pricing errors put too much weight on options with a high market price in the parameter estimation. Then the models are tested out-of-sample on the next window of ten days. Finally, the time window is shifted by ten days and the models are reestimated. All together, 120 time windows are studied where the number of options ( $N$ ) ranges from 114 to 1357. This procedure seems to be a reasonable trade-off between the size of data sets (the amount of market information available) on the one hand and time dependence of the estimates on the other hand.

### 3. Theoretical Framework

#### 3.1. The Black-Scholes model

In their seminal paper Black and Scholes (1973) derive the result that the dynamics of the underlying stock price  $S$  under the risk-neutral measure is a geometric Brownian motion with the same volatility  $\sigma$  as under the physical measure (under additional assumptions such as a constant risk-free interest rate, no transactions costs etc.). The drift, however, is replaced by the risk-free interest rate  $\bar{r}$  minus the dividend yield  $\delta$

$$\frac{dS}{S} = (\bar{r} - \delta)dt + \sigma dW \quad (2)$$

where  $dW$  is a Wiener Process. In particular, the volatility  $\sigma$  is assumed to be constant. As a consequence of Eq. (2), the price of a European call option with strike price  $K$  is given by

$$C_{BS} = Se^{-\delta(T-t)}N(d_1) - Ke^{-\bar{r}(T-t)}N(d_2) \quad (3)$$

where  $T - t$  is the time to maturity of the option and

$$\begin{aligned} d_1 &= \frac{\log\left(\frac{S}{K}\right) + (\bar{r} - \delta + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned}$$

The price of a European put option can be derived as

$$P_{BS} = C_{BS} + Ke^{-\bar{r}(T-t)} - Se^{-\delta(T-t)}. \quad (4)$$

Given a set of observed option prices, one can estimate the (implied) volatility  $\sigma^*$  yielding the lowest pricing error<sup>8</sup>. As described in section 2, the parameter  $\sigma^*$  is estimated over a period of ten trading days such that the SRPE is minimized.

---

<sup>8</sup>This procedure was introduced by Lamoureux and Lastrapes (1993).



### 3.2. The Hull-White option pricing model

Hull and White (1987) proposed an option pricing model where both the stock price  $S$  and the instantaneous variance  $V$  follow a stochastic process:

$$\frac{dS}{S} = (\bar{r} - \delta)dt + \sqrt{V}dW, \quad (5)$$

$$dV = (a + bV)dt + \xi\sqrt{V}dZ \quad (6)$$

where the parameters  $a$  and  $b$  determine the long-run mean and the speed of mean reversion for the variance and  $\xi$  can be interpreted as the volatility of volatility. The two Wiener processes  $dW$  and  $dZ$  have a correlation of  $\rho$ .

In this model there are two sources of uncertainty, namely the future path of the asset price and the future path of volatility. In the above parameterization the price of volatility risk is set to be zero. This assumption implies that there is no risk premium for the volatility risk. In contrast to that, Heston (1993) has introduced a closed-form model with a non-zero price of volatility risk.

The Hull-White model allows for an impact of correlation and volatility of volatility on option prices. Higher volatility of volatility means that the risk-neutral distribution of returns exhibits higher leptokurtosis. For this model, extreme returns of positive and negative sign are more likely than in the Black-Scholes world where the asset price follows a lognormal distribution. When the correlation between shocks to returns and shocks to volatility is zero, the risk-neutral distribution is symmetric but fat-tailed. The sign of the correlation determines the symmetry of the distribution. Considering the empirically relevant case where the correlation is negative, the left tail of the risk-neutral distribution of returns contains more probability mass than the right tail. As Abken and Nandi (1996) discusses, this negative skewness has consequences for pricing: Out-of-the-money calls are overpriced by the Black-Scholes option pricing model.

To make the parameter estimation computationally feasible we use the Taylor series expansion as in Hull and White (1988). It is assumed that the variance at time  $t$  is at its long-run mean  $V = -a/b$  and that the future path is specified by Eq. (6). In the Hull and White (1988) model prices  $C_{HW}$  of European call options are approximated by

$$C_{HW} = C_{BS} + Q_1\rho\xi + Q_2\xi^2 + Q_3\rho^2\xi^2 \quad (7)$$

where  $C_{BS}$  is the corresponding Black-Scholes price. The factors  $Q_1, Q_2, Q_3$  are derived in Hull and White (1988) and determine the deviation from the Black-Scholes price. For the estimation of the model parameters, we minimize the SRPE (as outlined earlier) using a simulated annealing algorithm. This procedure exhibited better convergence to the optimum than standard algorithms.

### 3.3. The GARCH option pricing model

The GARCH option pricing model introduced by Duan (1995b) is based on a discrete-time model of the economy. It starts with a model for one-period returns of the underlying asset which is often a GARCH model (Bollerslev (1986)) and sometimes a more general non-linear asymmetric GARCH model (Engle and Ng (1993)). In the latter case, the continuously compounded returns  $r_t = \ln(S_t/S_{t-1})$ , where  $S_t$  denotes the price of the underlying asset at time  $t$ , are modeled by

$$r_t = (\bar{r} - \delta) + \lambda\sigma_t - \frac{1}{2}\sigma_t^2 + \sigma_t e_t, \quad (8)$$

$$e_t \sim N(0, 1),$$

$$\sigma_t^2 = \alpha_0 + \alpha_1(e_{t-1} - \gamma)^2\sigma_{t-1}^2 + \beta_1\sigma_{t-1}^2 \quad (9)$$

where  $N(0, 1)$  denotes the standard normal distribution.  $\lambda$  can be interpreted as the unit risk premium. This model assumes that returns are drawn from a normal distribution

with time-depending volatility  $\sigma_t$ . Because of this heteroskedasticity the unconditional distribution is fat-tailed. To ensure stationarity of the variance, it is required that the parameters satisfy  $\alpha_1(1 + \gamma^2) + \beta_1 < 1$ . The unconditional variance of the process is then given by  $\alpha_0/(1 - \alpha_1(1 + \gamma^2) - \beta_1)$ . The GARCH process defined by Eqs. (8) – (9) reduces to the standard homoskedastic lognormal process of the Black-Scholes model if  $\alpha_1 = 0$  and  $\beta_1 = 0$ . In other words, the Black-Scholes model is obtained as a special case.

In order to derive the GARCH option pricing model, the conventional risk-neutral valuation relationship has to be generalized to the locally risk-neutral valuation relationship (Duan (1995b)). This concept implies that under the risk-neutral measure the asset return process is given by

$$r_t = (\bar{r} - \delta) - \frac{1}{2}\sigma_t^{*2} + \sigma_t^* e_t^*, \quad (10)$$

$$e_t^* \sim N(0, 1),$$

$$\sigma_t^{*2} = \alpha_0 + \alpha_1(e_{t-1}^* - (\lambda + \gamma))^2 \sigma_{t-1}^{*2} + \beta_1 \sigma_{t-1}^{*2}. \quad (11)$$

The form of the process thus remains essentially the same. Denoting the new noncentrality parameter  $\tilde{\gamma} = \lambda + \gamma$ , the risk-neutral measure is determined by four parameters (if the interest rate  $\bar{r}$  and the dividend yield  $\delta$  are given), namely  $\alpha_0$ ,  $\alpha_1$ ,  $\tilde{\gamma}$ , and  $\beta_1$ .

From Eq. (10), it follows immediately that the price of the underlying asset at maturity  $T$  is

$$S_T = S_t \exp \left( (\bar{r} - \delta)(T - t) - \frac{1}{2} \sum_{i=t+1}^T \sigma_i^{*2} + \sum_{i=t+1}^T \sigma_i^* e_i^* \right). \quad (12)$$

The conditional distribution of  $S_T$  given today's price  $S_t$  can not be determined analytically but has to be approximated by Monte Carlo simulation. More precisely, a set of  $M$  random paths<sup>9</sup> ( $e_{t+1,j}^*, \dots, e_{T,j}^*$ ),  $j = 1, \dots, M$ , is generated and the corresponding

<sup>9</sup>In the simulations  $M$  is set to 10,000. Additionally, a control variate technique is used to reduce the variance of option prices.

prices  $S_{T,j}$  are calculated using Eq. (11) and (12). Then, for example, the price  $C_G$  of a European call option, which is defined by

$$C_G = e^{-\bar{r}(T-t)} E^*[\max(S_T - K, 0)], \quad (13)$$

where  $E^*$  denotes the risk-neutral conditional expectation operator, is approximated by

$$\hat{C}_G = e^{-\bar{r}(T-t)} \frac{1}{M} \sum_{j=1}^M \max(S_{T,j} - K, 0). \quad (14)$$

The price of a European put option can be estimated in a similar way. The conditional distribution of  $S_T$  also depends on the current values of the innovation  $e_t^*$  and the variance  $\sigma_t^{*2}$ . As an approximation, the innovation  $e_t$  and the variance  $\sigma_t^2$  of the asset return series model in Eqs. (8) – (9) are used.

Comparing GARCH and SV, we observe that they nest the geometric Brownian motion of the Black-Scholes option pricing model. Both models are characterized by a time-varying and asymmetric volatility function. Therefore, both models can generate leptokurtosis as well as skewness in the risk-neutral distribution of returns. These deviations from the assumption of conditional normality of the Black-Scholes model are considered to be one explanation for the "smile" effect (cf. Hull (1999), p. 439). GARCH is a deterministic variance model, or in other words, the same source of randomness drives returns and variances. In contrast to that, SV has a second source of randomness in addition to that in returns.

The relation of time-discrete (GARCH) and time-continuous (SV) option pricing models has been studied extensively in Duan (1996). In fact, a unified theory can be formulated where GARCH-type models converge to SV models as the length of the time intervals goes to zero (Duan (1996)). Beyond the theoretical insight, this result is also essential from a practical point of view because it gives rise to the possibility of inter-

changing GARCH and SV models for the purpose of estimation and option valuation. Our study is also intended to provide empirical results in this context.

## 4. Empirical Results

### 4.1. Parameter estimates

As described in section 2, the parameters of each model are estimated on 120 disjoint time windows each of which covers a period of ten trading days. The mean values and standard deviations of the obtained parameters are reported in Table 2. The time variation in the estimated coefficients indicates the usefulness of the applied sliding window technique.

[Table 2 about here.]

The average (annualized) implied volatility obtained from the Black-Scholes model is 13.8%. For the SV models and the GARCH models, the average volatility (calculated from the parameters of the 120 setups) is 14.1% and 16.2%<sup>10</sup>, respectively. While the volatility of the SV model is only slightly larger than the implied volatility of the Black-Scholes model, the increase of volatility for the GARCH model is substantial.

---

<sup>10</sup>For two time windows, the estimated GARCH models are integrated. The average volatility is calculated from the parameters of the remaining 118 models.

## 4.2. Out-of-sample pricing performance

In the first part of our evaluation, two statistical error measures are applied to quantify the deviation of theoretical option prices from prices observed at the market. The RPE and ARPE are defined as

$$\text{RPE} = \frac{\hat{p}_i - p_i}{p_i}, \quad (15)$$

$$\text{ARPE} = \left| \frac{\hat{p}_i - p_i}{p_i} \right| \quad (16)$$

We focus on the overall fit and on the fit in subsets of our sample. The errors are categorized according to moneyness and time to maturity. The RPE is a measure of the bias of the pricing model. A non-zero RPE may therefore indicate the existence of systematic errors. The ARPE measures both the bias and the efficiency of pricing. The advantage of the ARPE relative to the root mean squared error (RMSE) which is a key criterion for model selection in econometrics, is that it gives a percentage value of the pricing error. Therefore, it can be interpreted more easily and it provides more insight into the economic significance of performance differences.

[Table 3 about here.]

Table 3 shows that the median and the mean of the RPE differ. This indicates that the distribution of errors is not symmetric but skewed. Comparing the RPEs for the three different models, we find systematic underpricing across all models. In other words, the volatility of the underlying is systematically underestimated. In terms of the RPE and its standard deviation, the GARCH model performs best. The means and the medians of all models are smaller than 10%. GARCH and SV can value options with an error of about 4%. This shows that the more complex specifications can better reproduce the volatility smile and the term structure of volatility. The Hull-White model achieves a lower RPE but a higher standard deviation than the Black-Scholes model. Regarding the ARPE,

we observe values between 11% and 15% for the median. Again, GARCH dominates the two alternatives. According to this criterion, the performance of the Hull-White model is disappointing. In particular, it does not achieve a significant improvement relative to the Black-Scholes model but requires considerable computational effort. Having examined the overall out-of-sample fit, we now turn to the results on individual subsets.

[Table 4 about here.]

Looking at different maturity and moneyness classes in Table 4, we see that prices for short-term out-of-the-money options are strongly biased for the Black-Scholes and the Hull-White model. For these two models, prices of long-term options are generally more biased over all moneyness groups than prices of short-term options. The Black-Scholes dominates the other models for one category ( $T - t \geq 0.8$ ,  $0.05 \leq \text{moneyness} \leq 0.08$ ), and the Hull-White model is superior in four categories. In the remaining 30 categories the Hull-White model and the BS model are clearly dominated by the GARCH model.

[Table 5 about here.]

When we examine the absolute relative pricing errors, there is again a difference between mean values and medians. Once more, the GARCH model outperforms the Black-Scholes model and the Hull-White model in terms of median, mean and standard deviation. Table 5 indicates that especially the Black-Scholes model and the Hull-White model can not explain the prices of short-term out-of-the-money options. The Hull-White model has the best performance in a single category ( $T - t \geq 0.8$ ,  $0.08 \leq \text{moneyness} \leq 0.1$ ) and the GARCH model is superior in the remaining cases. Comparing the ARPE across different moneyness groups, we see a clear pattern of falling ARPEs with rising moneyness. The ARPE also tends to decrease with increasing time to maturity for all models.

The economic significance of pricing errors is strongly related to trading costs, in particular to bid-ask spreads. We want to shed some light on this issue by reporting spreads for two options of the first test set. On 18 January 1993 at 9 am, a short-term out-of-the-money put option (maturing in February) had a price of 7 points. The bid price was 5 points and the ask price was 9 points. The bid-ask spread of 4 points thus represents more than 57% of the actual price of the option. On the other hand, a long-term out-of-the-money call option (maturing in December) was traded at a price of 148 points where the bid price and the ask price were 145 points and 150 points, respectively. The bid-ask spread of 5 points is roughly 3% in this case.

Comparing RPE and ARPE, we find that the latter is much larger than the former. Since the RPE measures the bias from the theoretical price and the ARPE measures the bias as well as the variance, we see that the variance around the theoretical price dominates the systematic errors.

In order to examine the pricing errors of the option pricing models in more detail, the ARPEs are regressed on the time to maturity (in years), the moneyness of the option, and a binary variable that is set to unity, if the option is a call and to zero in the case of a put:

$$\text{ARPE} = a_0 + a_1 \text{Time to maturity} + a_2 \text{Moneyness} + a_3 \text{Call} + \epsilon, \quad (17)$$

$$\epsilon \sim N(0, \sigma^2).$$

The estimated coefficients  $a_0$ ,  $a_1$ ,  $a_2$ , and  $a_3$  are reported in Table 6. We observe that for all models the absolute relative pricing errors are significantly lower for longer maturities and increased moneyness. Therefore, in-the-money options are priced more precisely than out-of-the-money options and longer maturities better than shorter ones. It is also interesting to see that the GARCH model prices calls and puts with good accuracy in contrast to the Hull-White and the Black-Scholes models. Hence, these two models do



not reproduce the different characteristics of call and put options. Finally, we observe that the  $R^2$  statistics, i.e. the measure of determination, is the lowest for the GARCH model. This indicates a lower explanatory value of moneyness, time to maturity and the put-call dummy. There is thus less systematic structure in the pricing errors of the GARCH model than in the ones of the Hull-White and the Black-Scholes model. We conclude once again that the theoretical prices of the GARCH model are closer to the values observed in the market than the prices of the SV model and the Black-Scholes model. After the evaluation of the statistical performance measures, we now investigate the consequences for risk management.

[Table 6 about here.]

## 5. Value-at-Risk Performance

The Value-at-Risk (VaR) is defined as the maximum loss that will not be exceeded with a certain probability. Let  $\alpha$  be this probability - often referred to as the confidence level - and let  $\Delta v$  be the change in the portfolio value until the next trading day. Then the following equation holds:

$$\mathbb{P}(\Delta v \leq -\text{VaR}) = \alpha. \quad (18)$$

While this measure is very intuitive in its interpretation, its computation in practice is not uniquely determined. We implement a very general approach, which consists of the following steps visualized in Figure 1:

- Every ten days, we reestimate the parameters of the three models from the option prices.
- Within each window we calculate the VaR for all options which are traded on two consecutive days<sup>11</sup>. The reason is that we require realized profit and loss

---

<sup>11</sup>In contrast to a ten day horizon, for instance, this choice results in the largest subsample.

observations in order to evaluate the performance of the VaR models. The sample contains 11,904 pairs for which we compute profits and losses from holding a short position. We focus on short positions because of their unlimited potential loss and their exposure to Gamma risk.

- For each pair, we compare the realized profit or loss to the VaR estimate, which is computed by Monte Carlo (MC) simulation as follows:
  - To avoid in-sample effects, the parameter set of the previous ten day rolling window is used.
  - To get the VaR estimate, the distribution of prices for a particular option for the next trading day is needed. We apply a three step procedure:
    - First, we compute 10,000 different scenarios for the underlying index value<sup>12</sup>. This is done consistently in the sense that the same stochastic process is assumed for both the simulation of the underlying and the option pricing model. For instance, when computing the VaR for the GARCH option pricing model, a GARCH price process is assumed for this simulation.
    - Second, for each scenario, we value the options under the three alternative pricing models. In the case of the Black-Scholes model, this is straightforward because of the closed form solution. For the GARCH model, this step requires a second, nested MC simulation<sup>13</sup>.
    - Third, for each of the option prices calculated for time  $t + 1$ , the price change relative to time  $t$  is computed. Then the VaR is obtained as the  $\alpha$  quantile of the distribution of price changes.

---

<sup>12</sup>An Euler scheme (see, e.g., Kloeden and Platen (1995)) is used to simulate the time-continuous Hull-White process.

<sup>13</sup>This increases the computational effort considerably. Therefore, the number of scenarios was chosen as 2,500 for the GARCH models. For single option contracts, the difference to the VaR estimates obtained from 10,000 simulations was checked and found to be negligible.

- All VaR figures are computed for a one-day holding period using two popular confidence levels, namely  $\alpha = 1\%$  and  $\alpha = 5\%$ . To quantify the performance of the VaR models, we use the proportion of failures test and a distribution test.

[Figure 1 about here.]

Our procedure as outlined above allows us to compute the VaR without simplifying the underlying model. The VaR is obtained by MC simulation accepting high computational effort. This simulation approach is particularly appropriate for options in contrast to the Riskmetrics system, which is a variance-covariance based approach. This alternative method, which is also termed the Delta-Gamma model, provides only an approximation to the distribution of option returns by means of a Taylor series expansion. It is based on the assumption of an unconditional normal distribution of portfolio returns. This distributional assumption is needed for computational tractability but it is frequently rejected in empirical studies. For instance, Pritsker (1997) and Pichler and Selitsch (1999) have documented the weaknesses of the variance-covariance approach. In contrast to that, the MC method is based on the full valuation of derivatives. Therefore, it does not hinge on the assumption of unconditional normality and eliminates one possible source of model error.

## 5.1. Proportion of failures test

The most important criterion of a risk management system is to fulfill the regulatory requirements. Under the current regulation, banks calculate their VaR figures and evaluate whether the realized trading losses exceed the VaR. Since the VaR is the expected loss that will only be exceeded with probability  $\alpha$ , such observations can be modeled as draws from a binomial random variable (under the assumption of independence across time), where the probability of realizing a loss greater than the VaR is equal to  $\alpha$ . The Basle Committee on Banking Supervision proposed a binomial test to verify the accuracy

of internal models for capital requirements. Following Kupiec (1995), we implemented the more powerful likelihood ratio test. The test statistic is given by:

$$LR = -2 \log \left( (1 - \alpha^*)^{(n-x)} (\alpha^*)^x \right) + 2 \log \left( \left( 1 - \frac{x}{n} \right)^{(n-x)} \left( \frac{x}{n} \right)^x \right) \quad (19)$$

where  $\alpha^*$  is the probability of failure under the null hypothesis,  $n$  is the sample size and  $x$  is the number of failures in the sample. A failure is defined as an observation where the realized loss exceeds the VaR. Under the null hypothesis the test statistic has a  $\chi^2$  distribution with one degree of freedom.

[Table 7 about here.]

The results of the proportion of failures test are presented in Table 7. In all cases except one, the number of outliers is larger than the prespecified value. All models are rejected at the 99% level. A similar result is obtained for the 95% level where failure rates range from 4.8% to 10.8%. Only the Hull-White model has a relatively good fit. However, this result is limited to the 95% level as the model is rejected at the 99% level. It is also interesting to see that while the GARCH model clearly outperformed the Hull-White model in terms of pricing, it is - apart from the larger numerical effort - not better suited for risk management.

## 5.2. Distribution test

The disadvantage of the proportion of failures test is that the information provided by the predicted distribution of portfolio losses is reduced to a binary variable. In the above test the size of outliers has no impact. Only the number of outliers is taken into account. The choice of the VaR confidence level, i.e., 99% or 95%, is thus crucial as a model may reproduce the 95% but not the 99% values. Indeed, this is exactly the case for the SV model. In this situation, where the tails of the densities intersect at certain points, there

is no clear decision rule. To overcome these problems Crnkovic and Drachman (1996) proposed a test procedure which is based on the entire distribution as predicted by the risk management model. This procedure complements the Kupiec test for the quality of market risk measurement. Eq. (18) can be rewritten as

$$\alpha = \Phi_{\Delta v}(-\text{VaR}) \quad (20)$$

where  $\Phi_{\Delta v}(z)$  is the cumulative distribution function of option returns. Since realized returns  $\varphi$  should be random draws from this distribution, the transformed returns

$$\pi = \Phi(\varphi) \quad (21)$$

should be uniformly distributed over the unit interval<sup>14</sup>. Therefore, this test uses the portfolio changes as they are predicted by the models to measure the goodness-of-fit to the observed density.

[Figure 2 about here.]

Figure 2 shows the histogram of the transformed returns  $\pi$  plotted against the expected uniform distribution for the three option pricing models. All distributions differ from the uniform distribution with respect to the tails. The higher probability mass in the tails shows that more large losses and gains are observed than predicted by the risk management models. Additionally, there is a higher number of observations around 0.7 for the Hull-White model which deviates substantially from the hypothesized uniform distribution. The GARCH and the Black-Scholes model show the biggest deviations not in the centre of the distribution but in the tails. Overall these graphs indicate that all models are unable to capture the leptokurtosis of option returns and that their fit to the tails of the distribution is quite weak. To quantify the deviation from the uniform

---

<sup>14</sup>This transformation was initially proposed by Rosenblatt (1952).

distribution, which is visualized in Figure 2, Crnkovic and Drachman (1996) proposed a test based on the Kuiper statistic which measures the deviation between the empirical and the theoretical cumulative distribution function<sup>15</sup>. The smaller this distance, the better the fit to the theoretical distribution. A key advantage of the Kuiper method is that all returns are equally weighted. Therefore, also the tails receive the necessary attention. Let  $D(x)$  be the cumulative distribution function of the transformed returns. Then the Kuiper statistic is given by

$$K = \max_{0 \leq x \leq 1} (D(x) - x) + \max_{0 \leq x \leq 1} (x - D(x)) \quad (22)$$

and the distribution of  $K$  for  $n$  observations is given by<sup>16</sup>

$$\mathbb{P}(k > K) = G \left( \left[ \sqrt{n} + 0.155 + \frac{0.24}{\sqrt{n}} \right] K \right) \quad (23)$$

where

$$G(\lambda) = 2 \sum_{j=1}^{\infty} (4j^2 \lambda^2 - 1) e^{-2j^2 \lambda^2}. \quad (24)$$

[Table 8 about here.]

The test results in Table 8 show that the null hypothesis of a uniform distribution is rejected at a level of more than 99.9% for all three models. In other words, all models fail to reproduce the distribution of option returns. The relative differences between the GARCH, the SV and the constant volatility framework are minor. Concerning the risk management performance, we therefore conclude that there is no superior model among the three approaches<sup>17</sup>.

Both the failure rates and the distribution test indicate that our models are not well suited to predict the one-day VaR. Although the GARCH model prices the FTSE

---

<sup>15</sup>See also Lopez (1999).

<sup>16</sup>See, e.g., Press, Teukolsky, Vetterling, and Flannery (1992), page 627.

<sup>17</sup>Similar results were obtained with the test proposed by Berkowitz (1998).

options with some accuracy, its forecasts match the distribution of option returns rather poorly. This low VaR prediction performance is expressed by the fact that there is no measurable improvement of GARCH on the benchmark Black-Scholes method. This observation shows a pronounced divergence between suitability for pricing and for risk management which has not been noted in the literature so far.

To interpret these findings we briefly analyze in which aspects the pricing methodology differs from the VaR methodology. As indicated in Figure 1, the key divergence concerns the stock price. The theoretical option price depends on two inputs, namely the current stock price and the set of parameters for the stochastic process. The pricing error is not affected by how we obtain the stock price because the true price (at time  $t$ ) is available. However, in order to forecast the expected profit/loss distribution for tomorrow, we require a prediction of the stock price distribution at  $t + 1$ . We obtain this forecast by MC simulation using the three asset price processes. The discrepancy between the pricing and the VaR results indicates that the forecast of tomorrow's stock price distribution has a negative effect on the VaR performance. Therefore, the problem is not how the derivatives are priced but how the process of the underlying is modeled in the VaR methodology. All three specifications of the stock price at  $t + 1$  assume conditional normality. Our VaR results indicate that this assumption is not adequate. A similar observation has been made by Engel and Gizycki (1999). Their comprehensive study on the VaR of FX portfolios documents that the distribution obtained from "a Gaussian MC simulation does not match the realized profit/loss distribution very well and that "a fat-tailed MC simulation, which is based on a mixture of normals, clearly dominates the VaR model with a single Gaussian. Therefore, the findings of Engel and Gizycki (1999) seem to coincide with our tentative conclusions.

## 6. Conclusion

In this paper the performance of three option pricing models is evaluated. We compare the benchmark Black-Scholes model to valuation methods using GARCH and SV. The three processes are calibrated to intraday market prices of FTSE 100 contracts. The performance is evaluated by means of pricing errors and by the fit for Value-at-Risk applications. Our first conclusion concerns the behavior of the pricing errors. We document that GARCH dominates SV and the benchmark Black-Scholes model. It achieves significant overall improvements in pricing performance. The pricing differences between SV and constant volatility are smaller. We thus observe that the choice of the volatility model is crucial for the pricing performance. However, these differences in performance are not apparent in the context of risk management. To examine the competing methodologies in this context we use the VaR framework. A Monte Carlo simulation with full valuation provides the forecasts of the expected profit/loss over a given horizon. As our second main conclusion, we find no measurable difference between the VaR forecasts of the three models. In particular, we document that all three approaches exhibit a weak fit to the realized profits and losses. Overall we therefore conclude that the performance of a model strongly depends on the loss function applied. Despite the reasonable fit to observed prices, the GARCH model fails to forecast the tails of the distribution of option returns with satisfying accuracy. We discuss this divergence of results and we conclude that the assumption of conditional normality of the underlying is problematic for VaR predictions.

For future research, it seems promising to use a fat-tailed distribution (such as a Student's  $t$ -distribution) to specify the price process of the underlying. The option pricing model is then extended from a GARCH model with a conditional normal distribution to a model with a conditional  $t$ -distribution.



## References

- Abken, P.A., Nandi, S., 1996. Options and Volatility, *Economic Review - Federal Reserve Bank of Atlanta* December, 21–35.
- Bakshi, G., Cao, C., Chen, Z., 1997. Empirical Performance of Alternative Option Pricing Models, *Journal of Finance* 52, 2003–2049.
- Bakshi, G., Cao, C., Chen, Z., 2000. Pricing and hedging long-term options, *Journal of Econometrics* 94, 277–318.
- Bates, D., 1996. Testing Option Pricing Models, in: C.S. Maddala, and C.R. Rao, eds., *Handbook of Statistics*, North Holland, New York.
- Berkowitz, J., 1998. Evaluating the Forecasts of Risk Models, mimeo.
- Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities, *Journal of Political Economy* 8, 1637–1659.
- Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics* 31, 307–327.
- Christoffersen, P., Hahn, J., Inoue, A., 2001. Testing and comparing Value-at-Risk measures, *Journal of Empirical Finance* 8, 325–342.
- Crnkovic, C., Drachman, J., 1996. A Universal Tool to Discriminate Among Risk Measurement Techniques, *Risk* 9, 138–143.
- Duan, J., 1995a. The GARCH Option Pricing Model, *Mathematical Finance* 5, 13–32.
- Duan, J.C., 1995b. The GARCH Option Pricing Model, *Mathematical Finance* 5, 13–32.
- Duan, J.C., 1996. A Unified Theory of Option Pricing under Stochastic Volatility - from GARCH to Diffusion, Working paper, Hong Kong University of Science and Technology.
- Duan, J.C., Zhang, H., 1999. Pricing Hang Seng Index Options - A GARCH Approach, mimeo.

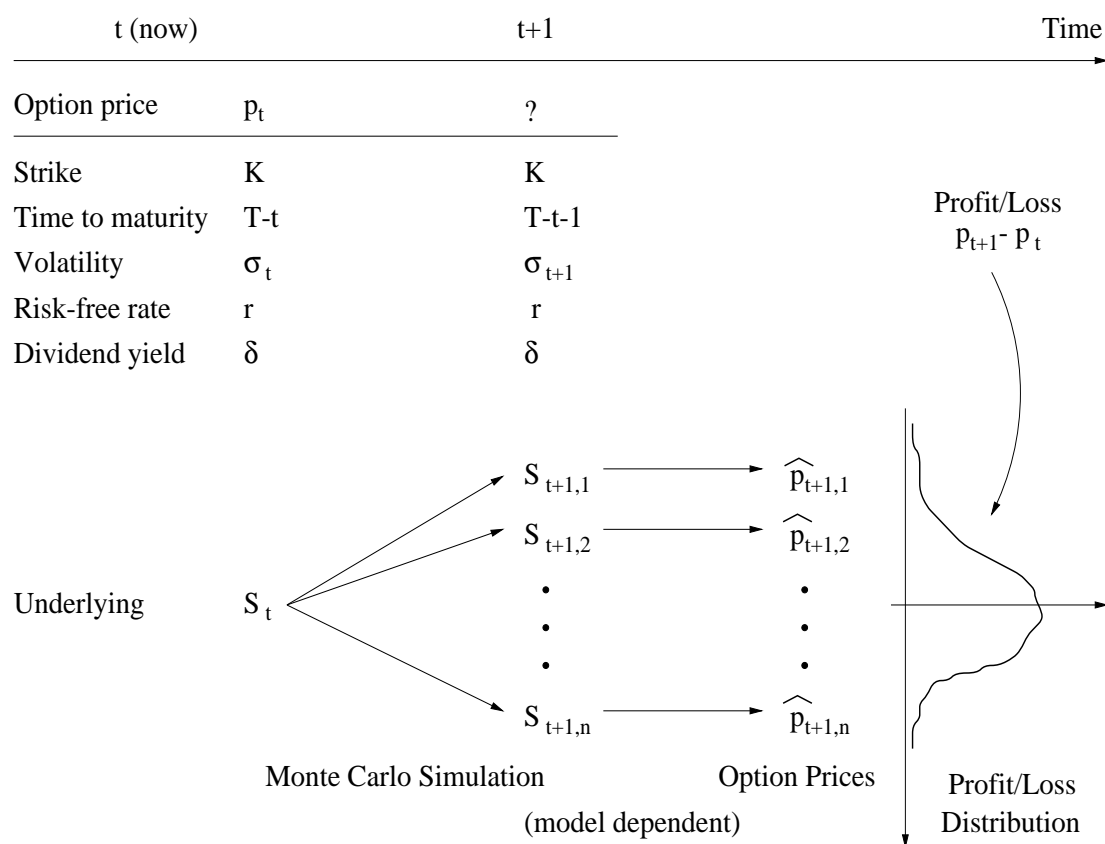
- Duffie, D., Pan, J., 1997. An Overview of Value at Risk, *Journal of Derivatives* 4 pp. 7–49.
- Engel, J., Gizycki, M., 1999. Conservatism, Accuracy and efficiency: comparing value-at-risk models, Working Paper, Australian Prudential Regulation Authority.
- Engle, R.F., Mustafa, C., 1992. Implied ARCH models from Option prices, *Journal of Econometrics* 52, 5–59.
- Engle, R.F., Ng, V.K., 1993. Measuring and Testing the Impact of News on Volatility, *Journal of Finance* 48, 1749–1778.
- Heston, S., 1993. A closed form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, 27–44.
- Hull, J., 1999. *Options, Futures and Other Derivatives*. Prentice Hall, London.
- Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatility, *Journal of Finance* 42, 281–300.
- Hull, J., White, A., 1988. An Analysis of the Bias in Option Pricing Caused by a Stochastic Volatility, in: *Advances in Futures and Options Research*, Greenwich CT: JAI Press, .
- Jorion, P., 1997. *Value at Risk : the new benchmark for controlling market risk*. Irwin.
- Kloeden, P.E., Platen, E., 1995. *Numerical Solution of Stochastic Differential Equations*. Springer.
- Kupiec, P.H., 1995. Techniques for Verifying the Accuracy of Risk Measurement Models, *Journal of Derivatives* pp. 73–84.
- Lamoureux, C., Lastrapes, W., 1993. Forecasting Stock Return Variances: Towards an Understanding of stochastic Implicit Volatilities, *Review of Financial Studies* 6, 293–326.
- Lopez, J.A., 1999. Regulatory Evaluation of Value-at-Risk Models, *Journal of Risk* 1, 37–64.

- Nandi, S., 1998. How important is the correlation between returns and volatility in a stochastic volatility model? Empirical evidence from pricing and hedging in the S&P 500 index options market., *Journal of Banking & Finance* 22, 589–610.
- Pichler, S., Selitsch, K., 1999. A Comparison of Analytical VaR Methodologies for Portfolios that Include Options, Working Paper, Technische Universität Wien.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. *Numerical Recipes in C: the art of scientific computing*. Cambridge University Press.
- Pritsker, M., 1997. Evaluating Value at Risk Methodologies: Accuracy versus Computational Time, *Journal of Financial Services Research* 12, 201–242.
- Rosenblatt, M., 1952. Remarks on a Multivariate Transformation, *Annals of Mathematical Statistics* 23, 470–472.
- Vlaar, P., 2000. Value-atRisk Models for Dutch Bond Portfolios, *Journal of Banking & Finance* 24, 1131–1154.

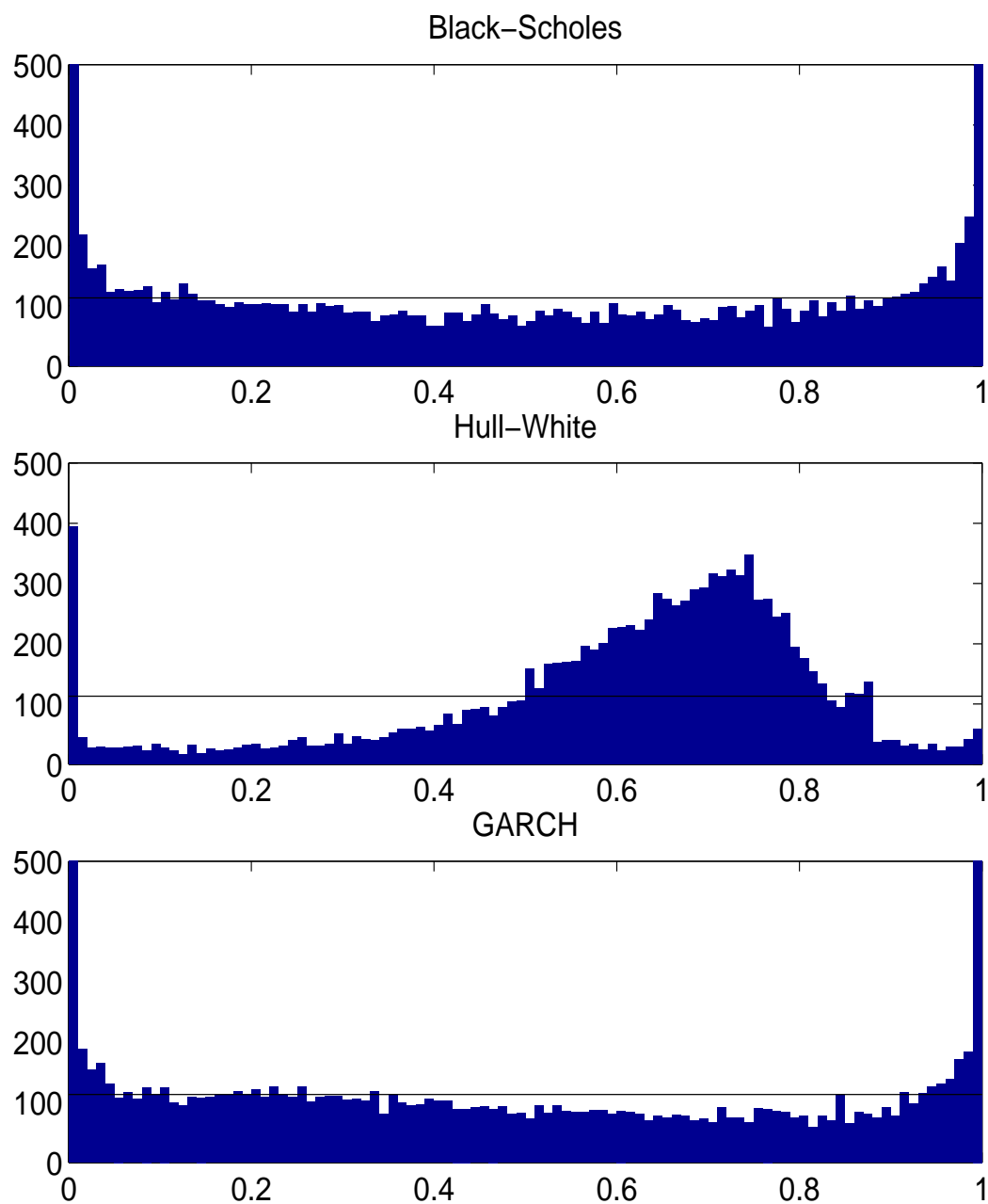
## List of Figures

- 1 Illustration of the procedure to obtain model-dependent Value-at-Risk estimates by Monte Carlo simulation. . . . . 28
- 2 Histograms of the transformed returns under the distributions predicted by the three option pricing models. The classes in the histograms have a width of 0.01 and are plotted against the density of the uniform distribution. 29

**Figure 1.** Illustration of the procedure to obtain model-dependent Value-at-Risk estimates by Monte Carlo simulation.



**Figure 2.** Histograms of the transformed returns under the distributions predicted by the three option pricing models. The classes in the histograms have a width of 0.01 and are plotted against the density of the uniform distribution.



## List of Tables

1	Sample descriptive statistics . . . . .	31
2	Mean values and standard deviations of the parameter estimates of the risk-neutral processes . . . . .	32
3	Out-of-sample pricing performance . . . . .	33
4	Relative pricing errors for different moneyness and maturity groups . . . . .	34
5	Absolute relative pricing errors for moneyness and maturity groups . . . . .	35
6	Regression Results . . . . .	36
7	Results of the proportion of failures test . . . . .	37
8	Results of the distribution test . . . . .	38

Table 1  
Sample descriptive statistics

	Min.	10	25	50	75	90	Max.
FTSE 100	2727.7	3015.4	3204.0	3740.3	4274.1	4836.5	5366.5
Strike Price	2525	2975	3225	3725	4225	4775	5875
Option Price	5	10	22	48	92	153	843
Time to Maturity	14	17	28	42	78	170	368
Interest Rate	4.28	5.24	5.64	5.89	6.35	6.77	7.58
Dividend Yield	3.01	3.32	3.64	3.88	4.02	4.14	4.36
Implied Volatility	5.02	9.10	10.80	14.07	17.28	19.97	48.52

The minimum, the 10th, the 25th, the 50th, the 75th and the 90th percentile, and the maximum of some characteristics of the cleaned FTSE data set used in the empirical analysis, namely the FTSE 100, the strike price, the option price (in points), the time to maturity (in days), the (annualized) interest rate (in %), the (ex post annualized) dividend yield (in %), the (annualized) implied volatility (in %). The data set consists of 65,549 option contracts (33,633 call options and 31,916 put options).



Table 2

Mean values and standard deviations of the parameter estimates of the risk-neutral processes

	BS	Hull-White				GARCH			
	$\sigma$	$a$	$b$	$\xi$	$\rho$	$\alpha_0$	$\alpha_1$	$\beta_1$	$\tilde{\gamma}$
Mean	0.138	0.498	-24.951	4.506	-0.716	1.561E-05	0.317	0.187	1.227
Std.	0.026	0.221	6.771	2.794	0.375	1.199E-05	0.196	0.282	1.116

Parameter estimates for the Black-Scholes model:  $dS/S = (\bar{r} - \delta)dt + \sigma dW$ , the Hull-White model:  $dS/S = (\bar{r} - \delta)dt + \sqrt{V}dW$ ,  $dV = (a + bV)dt + \xi\sqrt{V}dZ$ , and the GARCH model:  $r_t = (\bar{r} - \delta) - \frac{1}{2}\sigma_t^2 + \sigma_t e_t$ ,  $\sigma_t^2 = \alpha_0 + \alpha_1(e_{t-1} - \tilde{\gamma})^2\sigma_{t-1}^2 + \beta_1\sigma_{t-1}^2$  where  $dW$  and  $dZ$  denote standard Wiener processes with correlation  $\rho$  and where the  $e_t$  are normally distributed random variables with zero mean and unit variance.

Table 3  
Out-of-sample pricing performance

<b>RPE</b>	Median	Mean	Std. Dev.
Black-Scholes	-6.67%	-7.71%	32.09%
Hull-White	-4.24%	-4.61%	32.52%
GARCH	-3.09%	-2.63%	23.01%
<b>ARPE</b>	Median	Mean	Std. Dev.
Black-Scholes	14.69%	22.72%	23.93%
Hull-White	14.58%	22.42%	24.00%
GARCH	11.77%	16.12%	16.63%

Out-of-sample pricing performance of the Black-Scholes model, the Hull-White model and the GARCH model for FTSE 100 call and put option contracts: The median, the mean and the standard deviation of the relative pricing errors  $RPE = (\hat{p}_i - p_i)/p_i$  and the absolute relative pricing errors  $ARPE = |\hat{p}_i - p_i|/p_i$  are given.  $p_i$  and  $\hat{p}_i$  denote the observed price and the model price, respectively.

Table 4  
Relative pricing errors for different moneyness and maturity groups

Moneyness	Model	Time to Maturity					all maturities
		0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	$\geq 0.8$	
-0.1 to -0.08	BS	-66.38%	-33.57%	-24.63%	-31.51%	-16.85%	-46.06%
	HW	-63.83%	-28.37%	-19.31%	-26.26%	-11.67%	-42.77%
	GARCH	-16.40%	-4.70%	-1.40%	-5.40%	-2.98%	-7.48%
-0.08 to -0.05	BS	-37.51%	-13.83%	-17.01%	-14.63%	-18.01%	-25.98%
	HW	-33.82%	-9.13%	-13.87%	-12.14%	-15.27%	-22.62%
	GARCH	-9.18%	-5.10%	0.22%	0.02%	-2.64%	-6.02%
-0.05 to -0.02	BS	-6.42%	-11.16%	-11.60%	-10.79%	-11.51%	-8.08%
	HW	-3.99%	-8.17%	-9.15%	-8.35%	-9.18%	-5.58%
	GARCH	-4.98%	-5.39%	-0.09%	-0.49%	-2.88%	-4.32%
-0.02 to 0.02	BS	-1.58%	-6.03%	-10.71%	-9.02%	-8.77%	-3.48%
	HW	0.86%	-3.74%	-8.75%	-7.06%	-7.62%	-1.16%
	GARCH	-2.82%	-1.20%	-1.59%	-0.45%	-4.06%	-2.41%
0.02 to 0.05	BS	-3.91%	-4.86%	-3.74%	-2.88%	-1.29%	-3.96%
	HW	0.11%	-2.89%	-2.40%	-1.94%	-0.34%	-0.71%
	GARCH	-1.79%	-1.03%	0.42%	1.79%	0.15%	-1.34%
0.05 to 0.08	BS	-3.76%	-2.28%	0.22%	-2.49%	-0.09%	-3.46%
	HW	-1.07%	-2.01%	0.05%	-2.51%	0.41%	-1.23%
	GARCH	-2.02%	-1.81%	2.19%	0.02%	0.88%	-1.65%
0.08 to 0.1	BS	-3.46%	0.07%	1.69%	1.70%	-0.23%	-2.61%
	HW	-2.18%	0.05%	1.93%	3.46%	-0.09%	-1.70%
	GARCH	-2.13%	-1.08%	1.05%	-1.25%	-1.79%	-1.89%
all moneyness groups	BS	-5.38%	-8.55%	-10.65%	-11.27%	-10.15%	-6.67%
	HW	-2.82%	-5.90%	-8.48%	-8.70%	-7.94%	-4.24%
	GARCH	-3.60%	-2.87%	-0.60%	-3.67%	-2.92%	-3.09%

Out-of-sample pricing performance of the Black-Scholes model, the Hull-White model and the GARCH model for FTSE 100 call and put option contracts: The median of the relative pricing errors  $RPE = (\hat{p}_i - p_i)/p_i$  is broken down into different moneyness and maturity groups.  $p_i$  and  $\hat{p}_i$  denote the observed price and the model price, respectively.

Table 5  
Absolute relative pricing errors for moneyness and maturity groups

Moneyness	Model	Time to Maturity					all maturities
		0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	$\geq 0.8$	
-0.1 to -0.08	BS	67.59%	37.57%	30.54%	32.73%	22.28%	49.81%
	HW	64.89%	37.39%	29.61%	30.23%	20.51%	46.94%
	GARCH	22.88%	16.96%	15.20%	15.93%	12.56%	18.31%
-0.08 to -0.05	BS	44.74%	24.74%	20.77%	18.71%	18.87%	34.59%
	HW	41.82%	23.55%	19.59%	17.94%	15.63%	32.63%
	GARCH	22.61%	15.46%	11.32%	12.28%	11.18%	18.15%
-0.05 to -0.02	BS	22.91%	14.82%	13.03%	12.36%	12.58%	20.21%
	HW	21.36%	13.45%	12.61%	10.82%	11.38%	18.89%
	GARCH	18.35%	12.27%	8.21%	7.97%	11.36%	16.28%
-0.02 to 0.02	BS	10.89%	9.29%	11.37%	10.12%	10.12%	10.66%
	HW	11.58%	8.36%	9.77%	9.22%	9.14%	10.91%
	GARCH	10.62%	8.34%	6.29%	8.41%	7.19%	9.88%
0.02 to 0.05	BS	6.46%	7.18%	7.32%	5.93%	6.94%	6.54%
	HW	7.12%	6.68%	6.25%	5.53%	7.48%	6.85%
	GARCH	4.85%	5.96%	4.94%	4.95%	6.53%	5.03%
0.05 to 0.08	BS	5.21%	4.97%	5.62%	6.56%	5.12%	5.25%
	HW	4.50%	4.28%	5.64%	6.83%	5.80%	4.65%
	GARCH	3.05%	3.62%	3.83%	7.04%	4.26%	3.33%
0.08 to 0.1	BS	3.86%	3.95%	4.60%	4.06%	2.35%	3.89%
	HW	2.64%	3.23%	4.61%	5.40%	2.08%	2.88%
	GARCH	2.37%	2.79%	3.16%	3.16%	2.77%	2.48%
all moneyness groups	BS	15.46%	13.29%	12.90%	13.42%	12.56%	14.69%
	HW	15.76%	12.39%	11.98%	12.04%	11.04%	14.58%
	GARCH	12.81%	10.47%	7.92%	9.29%	8.88%	11.77%

Out-of-sample pricing performance of the Black-Scholes model, the Hull-White model and the GARCH model for FTSE 100 call and put option contracts: The median of the absolute relative pricing errors  $ARPE = |\hat{p}_i - p_i|/p_i$  is broken down into different moneyness and maturity groups.  $p_i$  and  $\hat{p}_i$  denote the observed price and the model price, respectively.

Table 6  
Regression Results

	Black-Scholes	Hull-White	GARCH
$a_0$	0.2290 (162.91)	0.2262 (155.544)	0.1602 (144.40)
$a_1$	-0.2039 (-50.56)	-0.2291 (-54.913)	-0.1568 (-49.13)
$a_2$	-3.7060 (-151.922)	-3.4185 (-135.491)	-1.7942 (-91.03)
$a_3$	-0.0659 (-42.602)	-0.0460 (-28.737)	-0.0055 (-4.47)
$R^2$	28.28%	23.72%	13.12%
F-test	9114.84	7188.53	3287.18

Results from regressing out-of-sample absolute relative pricing errors (ARPEs) on the options' time to maturity, their moneyness and a binary variable set to unity for calls and to zero for puts:  $ARPE = a_0 + a_1 \text{ Time to maturity} + a_2 \text{ Moneyness} + a_3 \text{ Call} + \epsilon$  where  $\epsilon$  denotes the residual error. The  $t$ -statistics of the parameter estimates are shown in parentheses. Additionally, the  $R^2$ -statistic and the results of the corresponding F-test are reported.

Table 7  
 Results of the proportion of failures test

Model	Failures $\alpha = 1\%$			Failures $\alpha = 5\%$		
	total	percent	$p$ -value	total	percent	$p$ -value
Black-Scholes	570	4.79%	1	1281	10.76%	1
Hull-White	436	3.66%	1	576	4.83%	0.5842
GARCH	535	4.49%	1	1207	10.14%	1

The table shows the total number and the percentage of failures for Value-at-Risk calculations using a 99% and 95% confidence level. The associated  $p$ -values show the probability at which the model can be rejected.

Table 8  
Results of the distribution test

Model	Kupier Statistic
Black-Scholes	0.15036
Hull-White	0.34495
GARCH	0.15527

Kuiper statistic for the models under the null hypothesis that the distribution is uniform (the 99.9% critical value is 0.0201195).