Option Pricing under the Mixture of Distributions Hypothesis

Marco Neumann

Diskussionspapier Nr. 208

First Version: December 14, 1997
Current Version: June 27, 1998

Preliminary Version – Comments Welcome

Marco Neumann
Institut für Entscheidungstheorie und Unternehmensforschung
Universität Karlsruhe (TH)
76128 Karlsruhe
Germany
Phone: +49 (721) 608-4767
Fax: +49 (721) 359200
E-mail: marco.neumann@wiwi.uni-karlsruhe.de

Option Pricing under the Mixture of Distributions Hypothesis

Abstract

This paper investigates the pricing of stock index options on the Deutscher Aktienindex (DAX) traded on the Deutsche Terminbörse (DTB) under the mixture of distributions hypothesis. Motivated by the poor empirical behavior of the lognormal distribution assumption under the Black/Scholes model, an option pricing model is constructed by assuming a mixed lognormal distribution for the underlying asset price. This assumption underlies both theoretical and empirical reasoning. From a theoretical point of view, using a mixture of distributions implies an option pricing model with randomly changing volatility. From an empirical point of view, observed phenomena like fat tailed and skewed asset price distributions must be explained. As already shown by Melick/Thomas (97), the broader set of possible shapes of the mixed lognormal distribution explains the empirical asset price distribution better than the simple lognormal distribution.

The option pricing formula under the mixture of lognormal distributions is both theoretically and empirically compared to the Black/Scholes formula. Fortunately, the corresponding option pricing formula is simply a linear combination of Black/Scholes option prices. Nevertheless, additional features like the probability of future extreme negative underlying price movements are incorporated into the pricing formula.

The empirical performance of the mixed lognormal option pricing formula is by construction at least as good as the Black/Scholes formula, since the latter is a special case of the former. Additionaly, the systematic overpricing for out of money calls and the systematic underpricing for in the money calls is avoided by the mixed option pricing model.

1 Introduction

Since the work of Black/Scholes (73) numerous tests of their option pricing model have been carried out. Besides the deviation from observed and theoretical option prices in these tests, the most stringent result is the calculation of an implied volatility smile (eg. Chiras/Manaster (78)). Because Black/Scholes (73) assume a constant volatility for the derivation of their option pricing formula, the existence of a volatility smile makes the empirical application of their formula questionable.

Recent research tried to explain the volatility smile by market imperfections like transaction costs. In the absence of arbitrage opportunities, the price S_t of any security equals its discounted expected terminal value S_T, where expectation is taken with respect to the equivalent martingale measure Q:

$$S_t = E_t^Q\left[S_T e^{-r(T-t)}\right].$$

Longstaff (95) used this relationship to perform what he called the martingale restriction test. From the Black/Scholes formula he calculated the implied index value S_t and the implied volatility σ of the lognormal distribution. Comparison between observed and implied index values shows, that the latter are almost always higher. Neumann/Schlag (96) obtain similar results for the German market. The interpretation is, that due to transaction costs, the duplication of the underlying by a portfolio of options is more expensive than buying the underlying itself. Regressions of the price differences on the bid ask spread show that this interpretation is at least one of the causes.

Since transaction costs are nearly constant over time, the generated Black/Scholes volatility smile should also be constant over time. Jackwerth/Rubinstein (96) show that this is not the case. Indeed the volatility structure changed significantly around the 1987 market crash. Therefore they attribute nearly all of the smile to shifts in probability beliefs which can not be adequately represented by the lognormal distribution. The nonparametric probability distribution calculated using Rubinstein's (94) method indeed has a fat left tail. This probability mass of extreme events, which is interpreted as a crash–o–phobia phenomenon, can not be represented by a lognormal distribution.

Melick/Thomas (97) report similar problems for the lognormal distribution using American oil futures options during the gulf crisis. They show that a mixture of lognormal distributions is able to represent the crash–o–phobia phenomenon. Empirically, this approach leads to better oil future option prices than the Black/Scholes model.

The purpose of this paper is to use a mixture of lognormal distributions for the pricing of European index options. Therefore, similar to Melick/Thomas (97), the parameters of the mixed lognormal distribution are estimated by minimizing the squared difference between observed and theoretical option prices. The resulting distributions and the pricing differences are then compared to the Black/Scholes model. Furthermore, the option pricing formula underlying the mixed lognormal distribution is derived and compared to the Black/Scholes formula. The additional parameters under the mixed lognormal distribution allow for an incorporation of extreme underlying price movements.

The rest of the paper is organized as follows. In Section 2 we describe the mixed lognormal probability distribution and derive the corresponding option pricing formula. For the sake of simplicity, we consider a mixture of two lognormal distributions. The resulting option pricing formula is interpreted and its theoretical features are compared to the Black/Scholes formula. In Section 3 we describe the data and some methodological assumptions used for the empirical study. In Section 4, the empirical results are presented and Section 5 summarizes and concludes.

2 Option pricing under a mixed lognormal distribution

In the absence of arbitrage opportunities, European–style options can be priced without any assumption about the underlying price process by duplicating their state dependent payoffs using the observed prices of the basis assets. This approach results in what is known as risk neutral valuation, where the price $C_t(S_t, X)$ of a European option equals its discounted expected terminal value $\max(S_T - X, 0)$, expectation taken with respect to the equivalent martingale measure Q:

$$C_t(S_t, X) = E_t^Q\left[\max(S_T - X, 0)e^{-r(T-t)}\right].$$

Black/Scholes (73) assume that the asset price S_T is lognormally distributed using

$$\begin{aligned} f_Q^{log}(S_T; \alpha, \beta) &= \frac{1}{\sqrt{2\pi}\beta S_T} \cdot e^{\left(-\frac{(\ln S_T - \alpha)^2}{2\beta^2}\right)}, \\ \text{with} \qquad & \alpha = \ln(S_t) + (r - \frac{\sigma^2}{2})(T-t), \\ & \beta = \sqrt{\sigma^2 \cdot (T-t)}, \end{aligned}$$

which results in the well known Black/Scholes option pricing formula,

$$\begin{aligned} C_t^{log}(S_t, X) &= S_t \cdot N(d_1) - X \cdot e^{-r(T-t)} \cdot N(d_1 - \sigma\sqrt{T-t}), \\ \text{with} \qquad & d_1 = \frac{\ln(\frac{S_t}{X}) + r(T-t) + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}}. \end{aligned}$$

Since the lognormal distribution is not appropriate to represent the real asset price distribution (especially in the lower left tail), the extension to a mixed lognormal distribution is regarded. In its simplest case, a mixed distribution is a linear combination of simple distribution functions. For the sake of simplicity we regard a mixture f_Q^{mix} of two lognormal distribution functions f_1^{log}, f_2^{log} with parameters α_i, β_i $(i = 1, 2)$ and weights $\pi_1 + \pi_2 = 1$ $(\pi_1 \geq 0,\ \pi_2 \geq 0)$:

$$f_Q^{mix}(\pi_1, \alpha_1, \alpha_2, \beta_1, \beta_2) = \pi_1 f_1^{log}(\alpha_1, \beta_1) + (1 - \pi_1) f_2^{log}(\alpha_2, \beta_2).$$

The mixed distribution is achieved by the superposition of the distributional components with regard to the probability π_i of each component f_i^{log}. The mixing weights π_i can be interpreted as the probability distribution of the parameters α_i, β_i for the lognormal distribution. Consequently, the assumption of a mixed distribution leads to randomly changing parameters.

From a theoretical point of view, following Harris (87), in a model where a fixed number of agents trade in response to new information the mixing variable can be interpreted as the number of information events occuring each day. Ignoring the evolution of information and the generated transactions, the approach presented here can be compared to models with stochastically changing volatility. Merton (76) for example assumes that the stock price performs random jumps. Other authors, e.g. Hull/White (87), provide an own diffusion process for the volatility. The approach presented here may be interpreted as consisting of two diffusions for the price process with different volatility parameters. Then one can choose between these two diffusions with probability π_1 and $(1-\pi_1)$, respectively. Consequently volatility changes randomly over time. Since further knowledge about the corresponding price process is not necessary for the valuation of European options, this question is not adressed any further here.

The corresponding option pricing formula is simply a linear combination of Black/Scholes option prices with respect to the distributional components of the mixture:

$$\begin{aligned} C_t^{mix}(S_t, X) &= e^{-r(T-t)} \int_{-\infty}^{\infty} \max(S_T - X, 0) f_Q^{mix}(S_T) dS_T \\ &= e^{-r(T-t)} \Big(\int_{-\infty}^{\infty} \max(S_T - X, 0) \pi_1 f_1^{log}(S_T) dS_T \\ &\quad + \int_{-\infty}^{\infty} \max(S_T - X, 0)(1-\pi_1) f_2^{log}(S_T) dS_T \Big) \\ &= \pi_1 C_t^{log}(S_t^1, X, \sigma_1) + (1-\pi_1) C_t^{log}(S_t^2, X, \sigma_2). \end{aligned}$$

The parameters S_t^1 and S_t^2 being the discounted expected values of the lognormal distribution components[1] are determined by the martingale restriction for the underlying price S_t:

$$e^{-r(T-t)} E_t^Q[S_T] = \pi_1 S_t^1 + (1-\pi_1) S_t^2 = S_t.$$

Consequently the observed underlying price S_t enters the option pricing formula by using

$$S_t^2 = \frac{S_t - \pi_1 S_t^1}{(1-\pi_1)}.$$

Without loss of generality we can assume, that $S_t^1 < S_t^2$ which is equivalent to $S_t^1 < S_t$. Then π_1 can be interpreted as that part of the probability distribution of the underlying asset price, which is mainly attributed to low asset prices. A crash–o–phobia phenomenon should then be represented by the shape of the mixture component f_1^{log} and the weight π_1 given to low asset prices. The degree of crash–o–phobia only depends on the parameters S_t^1 and σ_1 of the mixture component f_1^{log}. A low level of S_t^1 (compared to S_t) implies a

[1]The expected value of a lognormally distributed variable S_T^1 is $E_t^{Q_1}[S_T^1] = e^{\alpha_1 + \frac{\beta_1^2}{2}}$. The discounted expected value is $S_t^1 = e^{-r(T-t)} \cdot E_t^{Q_1}[S_T^1]$.

significant drop of the underlying price in case of a crash. The level of σ_1 is proportional to the possible price range for the crash.

But option prices depend on the risk associated with the whole probability distribution of underlying prices and not only on downside risk corresponding to the first distributional component. Therefore, the volatilities of both distributional components enter the mixed option pricing formula. Unfortunately, there is no single measure of underlying risk entering the option pricing formula. The appropriate risk–measure is the variance of the mixed probability distribution. In Appendix A we show for the variance of the mixed underlying price distribution:

$$\begin{aligned} Var^Q(S_T) &= \pi_1 \cdot Var^{Q_1}(S_T) + (1-\pi_1) \cdot Var^{Q_2}(S_T) \\ &\quad + \pi_1 \cdot (1-\pi_1) \left(E^{Q_1}(S_T) - E^{Q_2}(S_T) \right)^2 . \end{aligned}$$

Obviously, the risk corresponding to the mixed probability distribution of underlying prices in T consists of the weighted dispersions of the component distributions and the distance between the modes of these distributions.

Having this information about the composition of total risk of the mixed distribution, it is possible to calculate the contribution of the first distributional component to the total risk of the mixed distribution. Therefore, we define the share of downside risk SDR as

$$SDR = \frac{Var^Q(S_T) - (1-\pi_1) \cdot Var^{Q_2}(S_T)}{Var^Q(S_T)}.$$

Both the variance of the first distributional component and the distance between the modes of the two distributions contribute to this definition of downside risk. Consequently, SDR may be a good indicator for the influence of the first distributional component to the pricing of options.

So the approach of a mixed probability distribution is quite easy and results in an option pricing formula similar to the Black/Scholes model. Nevertheless, this approach incorporates the possibility of future extreme underlying price movements. The empirical analysis should show whether two mixture components are sufficient for the pricing of DAX index options. If not, an extension of the mixed lognormal option pricing formula for multiple mixture components is straightforward.

3 Data and methodology

The basic sample for this study consists of all best bid and best ask quotes for DAX options traded on the DTB for the first six months of 1994. The quotes are time-stamped to fractions of a second. They were considered good until changed. This yields a time-series of simultaneous best bid and best ask prices. We use the midpoint between bid and ask as an estimate of the true value of the option. The underlying DAX prices from the Kurs–Informations–Service–System (KISS) are time–stamped to the nearest minute.

For the final sample, only option price observations with remaining time to maturity of at least five trading days are considered. We then select one option price observation per trading day from the minute with the highest aggregate quotation activity per day[2]. The option prices are then matched with the DAX prices for the corresponding minute. Finally, for each option series consisting of options with differing strike prices but equal maturity, there remains one series of observations per trading day. The descriptive statistics for our final sample are given in table 1.

There is a total of 7263 observations for calls in 485 series and 6955 observations for puts in 499 series, yielding an average of 14.97 options per series for calls and 13.94 for puts. At least six individual options were available for all the estimations with a maximum of 30 for calls and 29 for puts. Time to maturity ranged from seven calendar days to about nine months for both option types. As expected, the average moneyness of the options defined as the difference between the observed index level and the strike price divided by the strike price is close to zero for both puts (0.0176) and calls (0.0197).

The riskless interest rate used for calculating theoretical option prices is estimated by the put–call parity as suggested by Shimko (93) using the two pairs of put and call options closest at the money with identical maturity for each series.

The parameters for the mixed lognormal distributions are calculated for each series by minimizing the sum of squared differences between observed and theoretical option prices:

$$\min_{\pi_1, S_t^1, S_t^2, \sigma_1, \sigma_2} \sum_{i=1}^{N} \left(C_i - C_i^{mix}(\pi_1, S_t^1, S_t^2, \sigma_1, \sigma_2) \right)^2 .$$

Since market imperfections are not considered, we explicitly impose the martingale restriction by using the observed asset price S_t with $S_t^2 = \frac{S_t - \pi_1 S_t^1}{(1-\pi_1)}$. Parameter estimates are calculated for each option series by performing a gridsearch with non-linear regressions on observed and theoretical option prices. For the special case of only one mixture component we achieve the simple lognormal distribution.

[2]Trading activity is measured by the number of quotes per minute.

Table 1: Descriptive statistics for the final sample

Calls (485 series)				
Variable	Mean	Std Dev	Minimum	Maximum
N[a]	14.97	5.39	6.00	30.00
T[b]	93.26	66.12	7.00	260.00
MONEY[c]	0.0197	0.0819	-0.1649	0.3904
SPREAD[d]	0.0772	0.0898	0.0036	1.7113
Puts (499 series)				
Variable	Mean	Std Dev	Minimum	Maximum
N[a]	13.94	4.75	6.00	29.00
T[b]	97.03	69.32	7.00	262.00
MONEY[c]	0.0176	0.0690	-0.1716	0.3736
SPREAD[d]	0.1083	0.1616	0.0367	1.9683

[a]Number of observations per series. The total number of call (put) price observations is 7263 (6955).

[b]Time to maturity in days.

[c]Relative moneyness calculated for each individual option as $\frac{(S-X)}{X}$ with X as the exercise price and S as the observed DAX price.

[d]Mean relative spread calculated for each individual option as 2 (BAP-BBP)/(BAP+BBP) with BAP (BBP) as the best ask (bid) price.

4 Empirical results

The empirical part of the paper compares the ability to explain option prices under the mixed lognormal distribution to the Black/Scholes distribution assumption. Therefore, several questions are examined. First, we compare the pricing error of the mixed lognormal formula to the pricing errors of the Black/Scholes formula. We find, that in most times the Black/Scholes formula is adequate for the pricing of DAX index options. But there are situations where the pricing may be significantly improved when using a mixed lognormal distribution. These situations are analyzed in section 4.2. The analysis shows, that the mixed lognormal approach improves pricing by the additional probability mass of the first mixture component in regions of low underlying prices. Consequently, these situations are abnormal in the sense that investors seem to expect a large decline of underlying prices to be possible. Finally, the shape of the corresponding probability distributions are considered in section 4.3.

4.1 Comparison of the pricing errors

In this section, the pricing differences using the mixed lognormal distribution and the single lognormal distribution are compared. Figure 1 shows the relative pricing error $\frac{C^{log}-C}{C}$ for the Black/Scholes model with respect to the moneyness of the option. As was also shown by Neumann/Schlag (96) for the German market, the Black/Scholes model systematically underprices (slightly) in the money calls and overprices (deep) out of money calls. This is not the case for the mixed lognormal model. Figure 2 shows the relative pricing error $\frac{C^{mix}-C}{C}$ for the mixed lognormal model. At first glance, comparison with figure 1 shows, that an increased number of out of money options are underpriced than before.

Since for $\pi_1 = 0$ both formulas are identical, only options with a clear pricing improvement under the mixed lognormal formula should be considered. Therefore, figure 4 shows the relative pricing error for options with improved pricing performance ($\pi_1 > 0$) under the mixed lognormal formula. The corresponding option prices under the Black/Scholes formula are plotted in figure 3. Comparison of both figures shows, that under the mixed lognormal distribution the systematic mispricing of the Black/Scholes formula is weakened for both overpriced out of money options and underpriced in the money options in the direction of a symmetric pricing error for both types of options. Therefore, the source of remaining mispricing under the mixed lognormal distribution may at least partially be explained as being random. The characteristics of the options with improved pricing

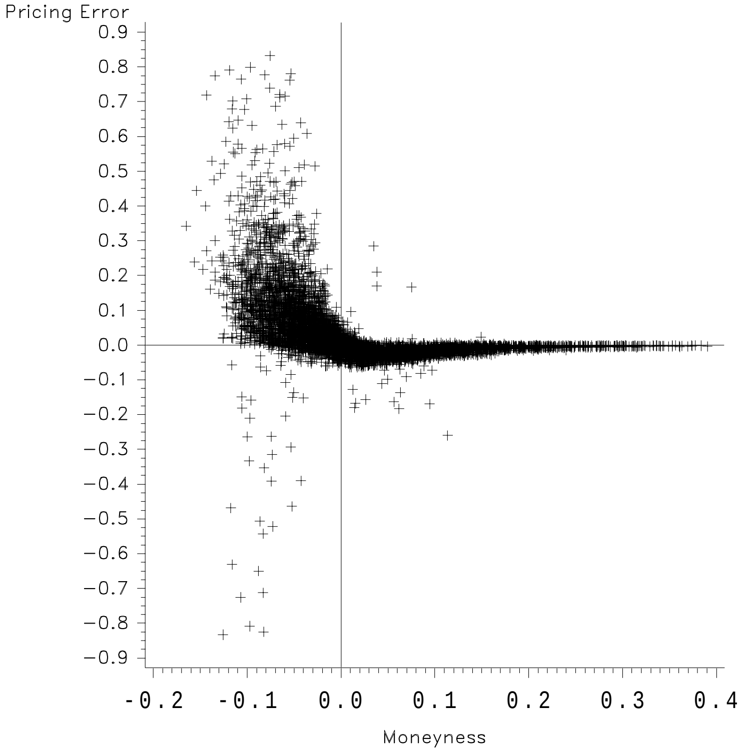

Figure 1: Relative pricing errors with respect to moneyness for the Black/Scholes formula.

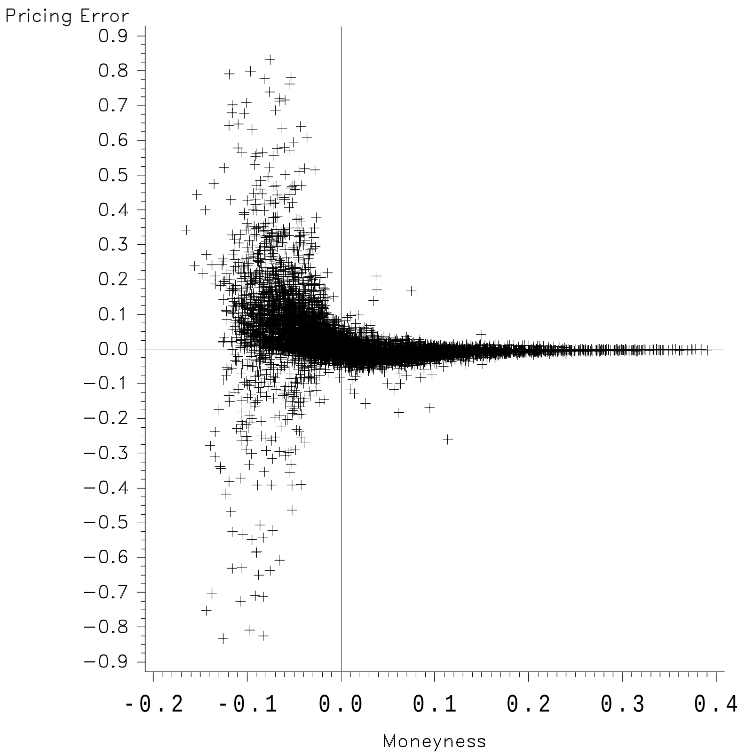

Figure 2: Relative pricing errors with respect to moneyness for the mixed lognormal formula.

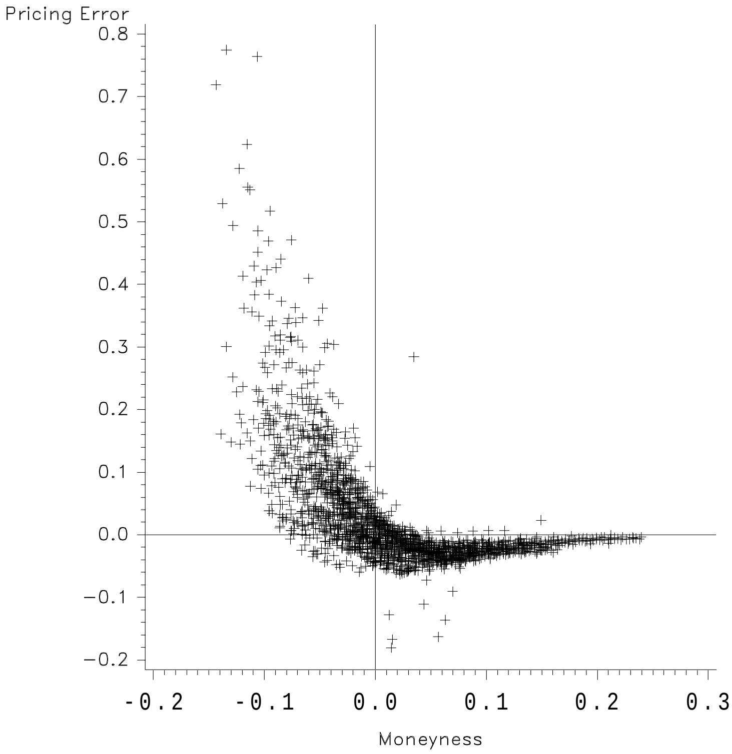

Figure 3: Relative pricing errors with respect to moneyness for the Black/Scholes formula with $\pi_1 > 0$.

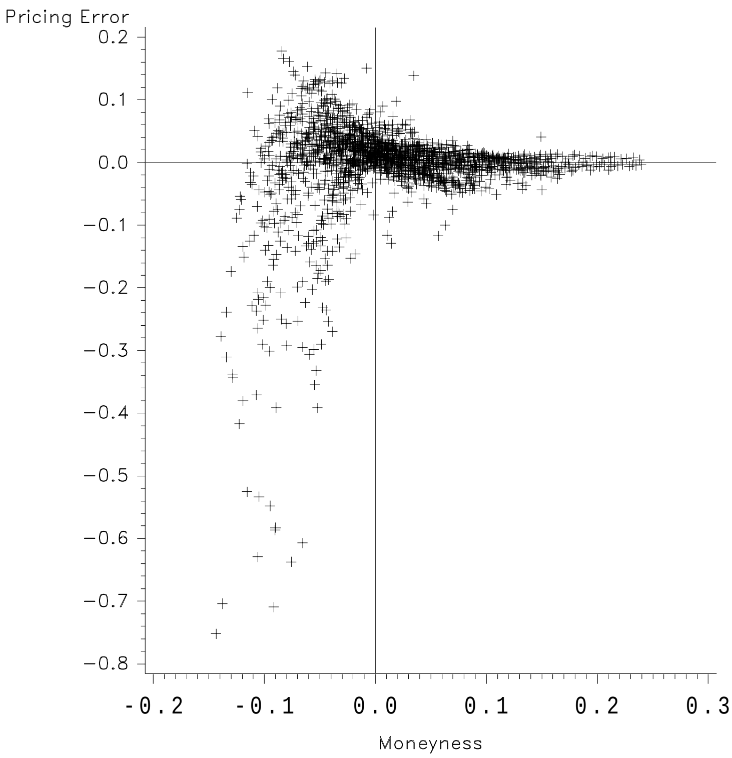

Figure 4: Relative pricing errors with respect to moneyness for the mixed lognormal formula with $\pi_1 > 0$.

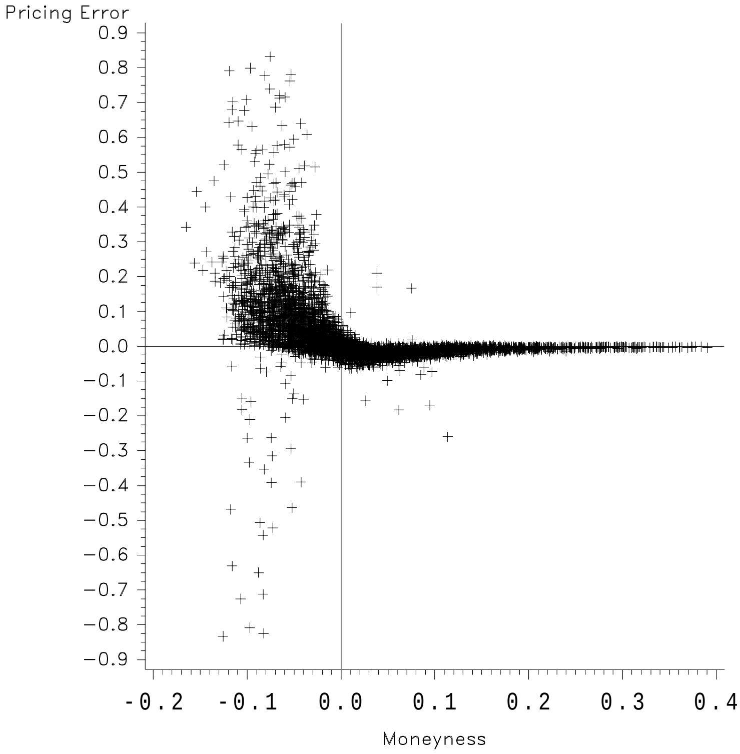

Figure 5: Relative pricing errors with respect to moneyness with $\pi_1 = 0$.

performance are shown in table 2. Comparison with table 1 shows, that on average the pricing of options with longer maturity and lower moneyness is improved. For 1794 out of 7263 options (in 132 out of 485 series), the mixed lognormal formula performs better than the Black/Scholes formula. Figure 5 shows those options where no pricing improvement could be achieved. These options are still priced according to the Black/Scholes formula. Especially deep in the money options have a good pricing performance under the Black/Scholes model.

So the first result is, that the mixed lognormal formula never worsens the pricing performance when compared to the Black/Scholes model. This is an implication of the construction of the mixed lognormal model, which includes for $\pi_1 = 0$ the Black/Scholes model as a special case. For options where a pricing improvement is achieved, the pattern of systematic mispricing by the Black/Scholes formula is avoided. But out of money calls still have a larger pricing error than in the money calls. The reason for this pricing improvement comes from the flexible shape of the mixed probability distribution especially in the range of low underlying prices, as will be shown later.

Additionally to the graphical analysis, the pricing errors are considered in greater detail for the whole sample in table 3 and for a reduced sample containing only observations with improved pricing performance ($\pi_1 > 0$) in table 4.

For the whole sample, the mean relative pricing error between mixed lognormal and observed option prices, $\frac{C^{mix}-C}{C}$, is 2.1%. The mixed lognormal option prices are on average 2.1% higher than the observed prices. Comparison between mixed and Black/Scholes relative option pricing errors shows, that the former perform on average better. Since the extreme underpricing of 77.62 points is identical under both models, the conclusion is that for the corresponding option no pricing improvement could be achieved. But the maximal overpricing is reduced from 37.24 to 31.89 under the mixed lognormal model. The maximum difference of 0 for the sum of squared pricing errors for the mixed lognormal and the Black/Scholes model, $\sum(C^{mix} - C)^2 - \sum(C^{log} - C)^2$, in table 3 also confirms the expected improvement under the mixed lognormal model. Consequently the mixed lognormal option pricing formula performs slightly better than the Black/Scholes formula for the considered period.

Since both models are identical for $\pi_1 = 0$, the outperformance of the mixed lognormal formula is observed in table 4 for the reduced sample with $\pi_1 > 0$. Clearly, the improvement is much better. Whereas the Black/Scholes formula overprices those options on average by 3.4%, the mixed lognormal formula underprices those options on average by only 0.82%.

Obviously, the pricing improvement of the mixed lognormal formula for the whole sample is rather poor. But seperating observations with improved pricing performance from those still priced according to the Black/Scholes formula yields a much better pricing improvement. One possible conclusion from this observation is, that in most times the Black/Scholes formula is adequate for the pricing of DAX index options. But there are situations where the lognormal distribution is not able to represent the investors probability assessments correctly. This leads to better theoretical option prices under the more flexible mixed lognormal distribution assumption.

4.2 Situations with improved pricing performance under the mixed lognormal distribution

The following analysis of the parameters S^1, S^2, σ_1, σ_2, π_1 of the mixed lognormal formula shows, that pricing performance is improved almost in situations with additional probability mass in regions with low asset prices. Table 5 shows some descriptive statistics for those parameters. Since for $\pi_1 = 0$ the parameters S^1 and σ_1 have no meaning, the corresponding values are only shown for $\pi_1 > 0$ in table 6.

The tables document that for $\pi_1 > 0$ the expected underlying price of the second lognormal distribution is always higher than the observed underlying price. Only for the

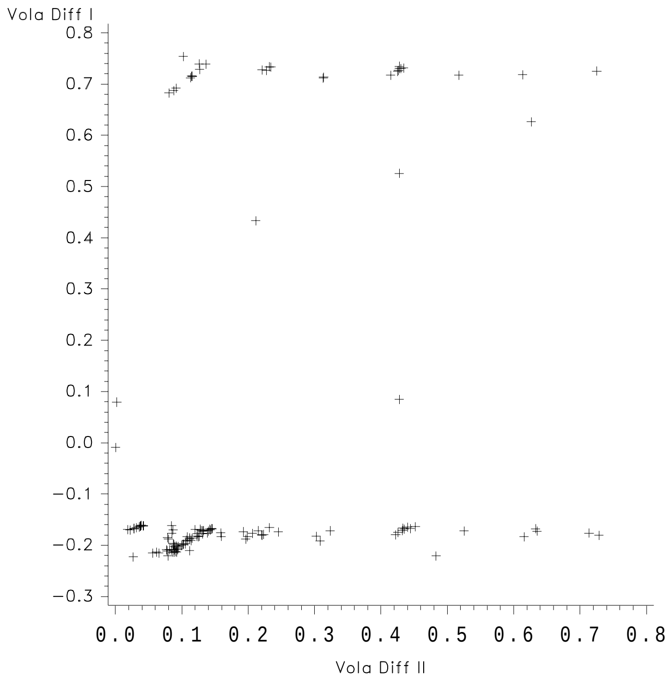

Figure 6: Volatility differences $\sigma_1 - \sigma$, $\sigma_2 - \sigma$ for both distributional components with respect to implied Black/Scholes volatilities for series with $\pi_1 > 0$.

case $\pi_1 = 0$ they are identical. The expected underlying price S^1 of the first lognormal distribution clearly (again by construction) corresponds to low underlying prices since it always lies below S_t. Therefore, a crash–o–phobia situation should be represented by the first lognormal distribution.

The volatility σ_1 of the first component ranges from very low (0.01) to very high (0.96) and is on average 0.22. The behaviour of implied volatility is analysed in greater detail in Figure 6, which shows the volatility changes for both distributional components with respect to the implied Black/Scholes volatility. The graph shows that σ_2 is always higher than σ, leading to a wider range for the distributional component corresponding to high underlying prices when compared to the single lognormal distribution. For σ_1 two cases are identified. σ_1 can either be by approximately 0.7 higher than σ or by approximately 0.2 lower than σ. The interpretation follows from the properties of the Black/Scholes formula which implies, that a portfolio of options is more expensive than a single option written on the portfolio of the corresponding underlyings[3] ($S_t = \pi_1 S^1 + (1-\pi_1)S^2$):

$$C^{log}(\pi_1 S^1 + (1-\pi_1)S^2, X, \sigma) \le \pi_1 C^{log}(S^1, X, \sigma) + (1-\pi_1)C^{log}(S^2, X, \sigma).$$

[3]This follows from the convexity of the max operator. See Merton (90), S. 265. Following Merton, the economic interpretation is that for options diversification "hurts". Since diversification reduces risk, only lower option prices are obtained. I am grateful to my colleague Nicole Branger for giving me this hint.

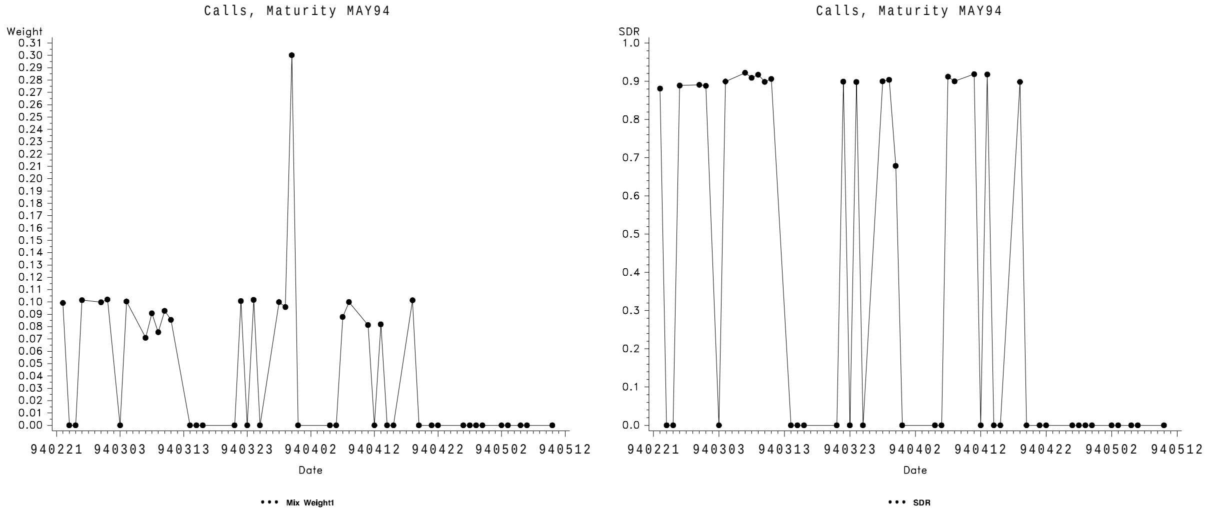

Figure 7: Time series for mixing weight π_1 (left) and time series for share of downside risk SDR (right).

Therefore, using the mixed lognormal formula with identical volatility as in the Black/Scholes case leads to systematically higher option prices compared to Black/Scholes prices. Since the second component always has a higher volatility, lower option prices may only be achieved by reducing the volatility σ_1 of the first component. But the aggregated effect on option prices also depend on the level of the weighting factor π_1.

The mixing weight π_1 equals the amount of probability mass attributed to low underlying prices, which is on average 19% (table 6). Figure 7 (left) shows the time series of π_1 for options maturing in May 94. Until April 18, 1994 probability mass is allocated to low underlying prices, leading to an implied probability distribution with two modes (figure 10). The remaining parameters S^1, S^2, σ_1 and σ_2 for the May 94 contract are drawn in figure 8. As also shown earlier, S^2 always lies above S by approximately 200 points when $\pi_1 > 0$. For the Black/Scholes case, both S and S^2 are identical. Considering only values where $\pi_1 > 0$, S^1 lies in the range of approximately 5% of the current underlying price. Together with very low σ_1 values of approximately 0.5% a crash–o–phobia situation is indicated.

The aggregated effect of all these parameter variations for the pricing of options can be expressed by SDR, the share of downside risk. SDR indicates the contribution of the first mixture component to total risk and consequently the importance of the first component for the pricing of the options. Figure 7 (right) shows the time series of SDR. For the estimation of SDR see appendix B.

Table 7 shows that SDR is on average 0.756. This means, that 75% of the risk associated

Table 2: Descriptive statistics for the final sample with $\pi_1 > 0$

Calls (132 series)				
Variable	Mean	Std Dev	Minimum	Maximum
N[a]	13.59	4.68	6.00	23.00
T[b]	121.29	63.36	7.00	260.00
MONEY[c]	0.0115	0.0696	-0.1561	0.2397
SPREAD[d]	0.0715	0.0534	0.0082	0.7023

[a]Number of observations per series. The total number of call price observations is 1794 .

[b]Time to maturity in days.

[c]Relative moneyness calculated for each individual option as $\frac{(S-X)}{X}$ with X as the exercise price and S as the observed DAX price.

[d]Mean relative spread calculated for each individual option as 2 (BAP-BBP)/(BAP+BBP) with BAP (BBP) as the best ask (bid) price.

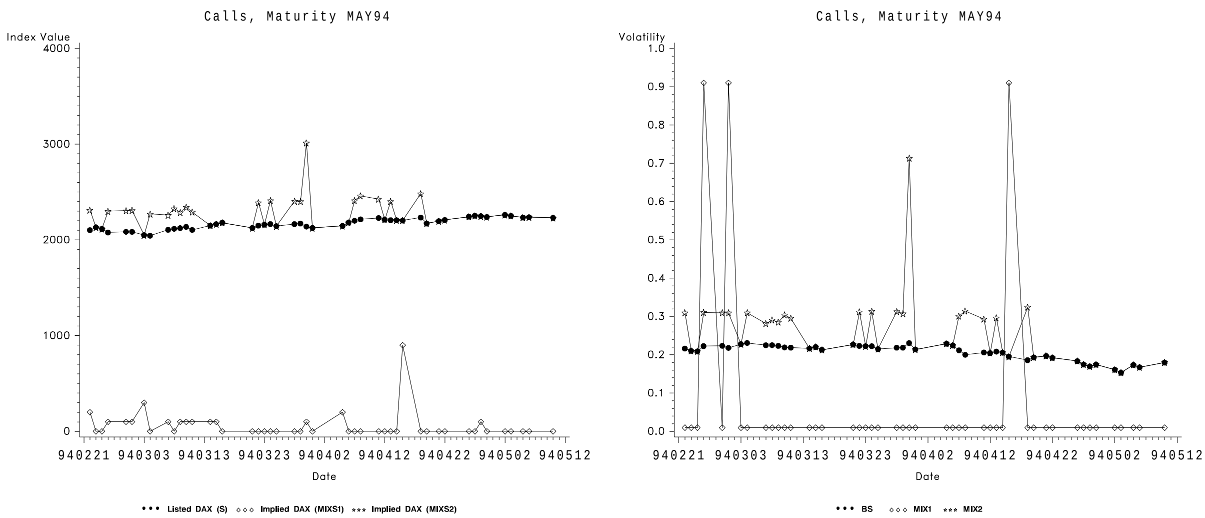

Figure 8: Time series for implied index values (left) and for implied volatilities (right).

Table 3: Call option pricing errors

Variable	Mean	Std Dev	Minimum	Maximum
Option prices (7263 observations)				
$C^{mix} - C^a$	-0.3042	3.5941	-77.6279	31.8922
$(C^{mix} - C)/C$	0.0214	0.1090	-1.8329	1.2786
$C^{log} - C^b$	-0.4702	3.9566	-77.6279	37.2392
$(C^{log} - C)/C$	0.0317	0.1061	-0.8326	1.2786
$C^{mix} - C^{log\,c}$	0.1659	2.9541	-26.1129	14.0496
$(C^{mix} - C^{log})/C^{log}$	-0.0071	0.0665	-1.5560	0.1166
$\sum(C^{mix} - C)^2 - \sum(C^{log} - C)^{2d}$	-42.90	117.84	-1082.41	0

[a]Difference between theoretical mixed lognormal and observed option prices (midpoint between bid and ask).

[b]Difference between theoretical Black/Scholes and observed option prices (midpoint between bid and ask).

[c]Difference between theoretical mixed lognormal and Black/Scholes option prices.

[d]Difference of Sum of squared pricing errors between mixed lognormal and Black/Scholes option prices. First the sum of squared pricing errors $\sum(C^{mix} - C)^2$, $\sum(C^{log} - C)^2$ is calculated for each series. Then the differences of these values are calculated for all series in the sample.

with the mixed distribution emerge from the additional distributional component corresponding to low underlying prices. Comparison to the minimum (≈ 0) and maximum ($= 0.916$) value of SDR shows, that the pricing of options is mainly improved for high levels of SDR. Consequently, the mixed lognormal distribution improves the pricing of DAX index options with additional probability mass in regions of low underlying prices. Therefore it seems, that investors expect a large decline of underlying prices to be possible with strictly positive probability.

The shape of the corresponding implied probability distribution will be analysed later. First, considering the valuation of the May 94 contract, figure 9 (left) shows that under the mixed lognormal distribution, valuation is only improved until approximately one month prior to maturity. The same result can be observed for other contracts. Therefore, the additional probability mass attributed to low underlying prices not only shows reactions to extreme events (as shown by Melick/Thomas (97) during the gulf crisis) but also expresses the uncertainty about future underlying price evolutions which grow with maturity. Options with lower maturity can be priced using the Black/Scholes formula, because there is not much volatility uncertainty nor is there much probability for a crash in

Table 4: Call option pricing errors for $\pi_1 > 0$

Variable	Mean	Std Dev	Minimum	Maximum
Option prices (1794 observations)				
$C^{mix} - C^{a}$	0.1032	3.9786	-26.2820	18.1786
$(C^{mix} - C)/C$	-0.0082	0.1089	-1.8329	0.1777
$C^{log} - C^{b}$	-0.5686	5.2078	-36.8405	37.2392
$(C^{log} - C)/C$	0.0336	0.1025	-0.1803	1.0964
$C^{mix} - C^{log c}$	0.6718	5.9165	-26.1129	14.0496
$(C^{mix} - C^{log})/C^{log}$	-0.0291	0.1315	-1.5560	0.1166
$\sum(C^{mix} - C)^2 - \sum(C^{log} - C)^{2d}$	-157.6397	181.8703	-1082.41	-0.0079

[a]Difference between theoretical mixed lognormal and observed option prices (midpoint between bid and ask).

[b]Difference between theoretical Black/Scholes and observed option prices (midpoint between bid and ask).

[c]Difference between theoretical mixed lognormal and Black/Scholes option prices.

[d]Difference of Sum of squared pricing errors between mixed lognormal and Black/Scholes option prices. First the sum of squared pricing errors $\sum(C^{mix} - C)^2$, $\sum(C^{log} - C)^2$ is calculated for each series. Then the differences of these values are calculated for all series in the sample.

the remaining time. Figure 9 (right) shows the sum of squared pricing errors between the mixed lognormal and Black/Scholes formula with respect to maturity. For the last thirty trading days, there is no pricing difference between mixed lognormal and Black/Scholes option prices.

So up to now, two shortcomings of the Black/Scholes model are avoided when using the mixed lognormal model. First, the constant Black/Scholes volatility is replaced by a randomly changing volatility (σ_1 and σ_2 with probability π_1 and $(1-\pi_1)$) under the mixed lognormal model, leading to improved prices for options with later maturities. Second, the strange Black/Scholes pricing pattern with respect to the moneyness of the options (which is also related to the volatility smile) disappears when using the more flexible mixed lognormal distribution. The analysis shows, that this shortcoming of the Black/Scholes model is eliminated under the mixed lognormal distribution by the additional probability mass of the first mixture component in regions of low underlying prices. The shape of these distributions and its pricing implications for options are considered in the next section.

Table 5: Estimated parameters for the mixed lognormal model

Variable[a]	Mean	Std Dev	Minimum	Maximum
Index values (485 observations)				
S^b	2152.56	67.5592	1967.36	2274.00
S^{1c}	.	.	.	.
S^{2d}	2303.75	378.46	1967.36	5343.71
${\pi_1}^e$	0.0528	0.1093	0	0.5999
Implied volatilities (485 observations)				
$\sigma^{log\,f}$	0.1983	0.0182	0.1532	0.2478
$\sigma_1^{mix\,g}$	.	.	.	.
$\sigma_2^{mix\,h}$	0.2535	0.1260	0.1532	0.9189

[a] The estimated values for S^1 and σ_1^{mix} are omitted because for cases with $\pi_1 = 0$ they distort the results.

[b] Observed DAX price.

[c] Implied DAX price for the first component of the mixed lognormal distribution.

[d] Implied DAX price for the second component of the mixed lognormal distribution.

[e] Mixing weight for the mixed lognormal distribution.

[f] Implied volatility using the Black and Scholes model.

[g] Implied volatility for the first component of the mixed lognormal distribution.

[h] Implied volatility for the second component of the mixed lognormal distribution.

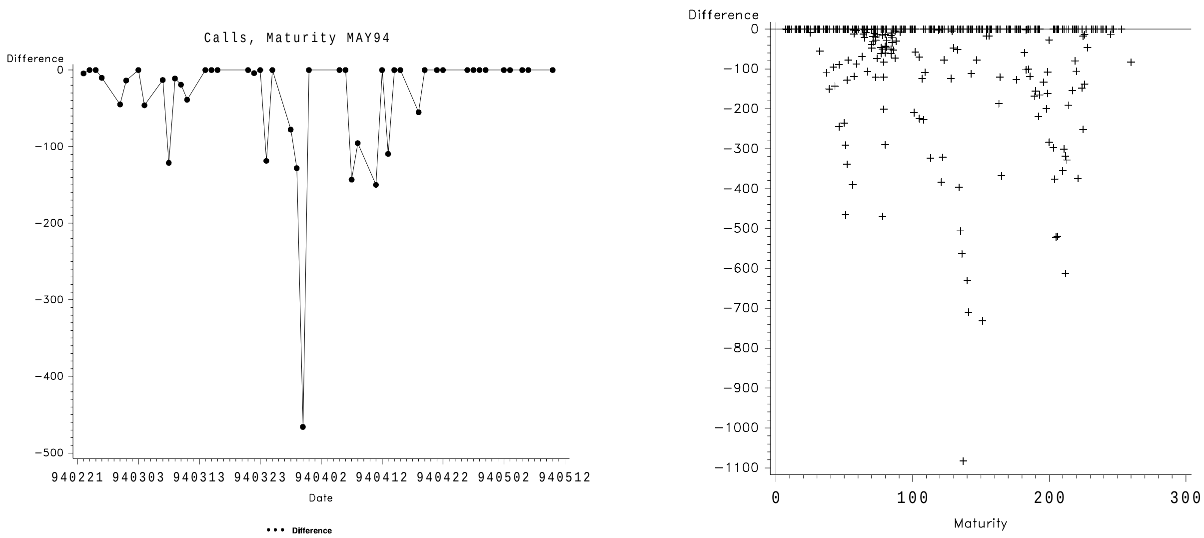

Figure 9: Time series of difference between mixed lognormal and Black/Scholes sum of squared residuals (left) and difference between mixed lognormal and Black/Scholes sum of squared residuals with respect to maturity (right).

Table 6: Estimated parameters for the mixed lognormal model with $\pi_1 > 0$

Variable	Mean	Std Dev	Minimum	Maximum
Index values (132 observations)				
S^a	2164.96	73.1503	1983.27	2274.00
S^{1b}	154.26	242.55	1.00	1801.00
S^{2c}	2720.44	526.80	2118.49	5343.71
${\pi_1}^d$	0.1939	0.1286	≈ 0	0.5999
Implied volatilities (132 observations)				
$\sigma^{log\,e}$	0.1937	0.0170	0.1710	0.2326
$\sigma_1^{mix\,f}$	0.2164	0.3714	0.0100	0.9610
$\sigma_2^{mix\,g}$	0.3964	0.1717	0.1979	0.9189

[a]Observed DAX price.

[b]Implied DAX price for the first component of the mixed lognormal distribution.

[c]Implied DAX price for the second component of the mixed lognormal distribution.

[d]Mixing weight for the mixed lognormal distribution.

[e]Implied volatility using the Black and Scholes model.

[f]Implied volatility for the first component of the mixed lognormal distribution.

[g]Implied volatility for the second component of the mixed lognormal distribution.

Table 7: Estimated crash–o–phobia parameters for the mixed lognormal model with $\pi_1 > 0$

Variable	Mean	Std Dev	Minimum	Maximum
Index values (132 observations)				
$S - S_1^{mix\,a}$	2010.70	271.85	347.09	2273.00
$(S - S_1^{mix})/S$	0.9276	0.1139	0.1616	0.9996
SDR^b	0.7565	0.1376	≈ 0	0.9161

[a]Difference between observed and implied DAX price of the first component of the mixed lognormal distribution.

[b]Share of downside risk.

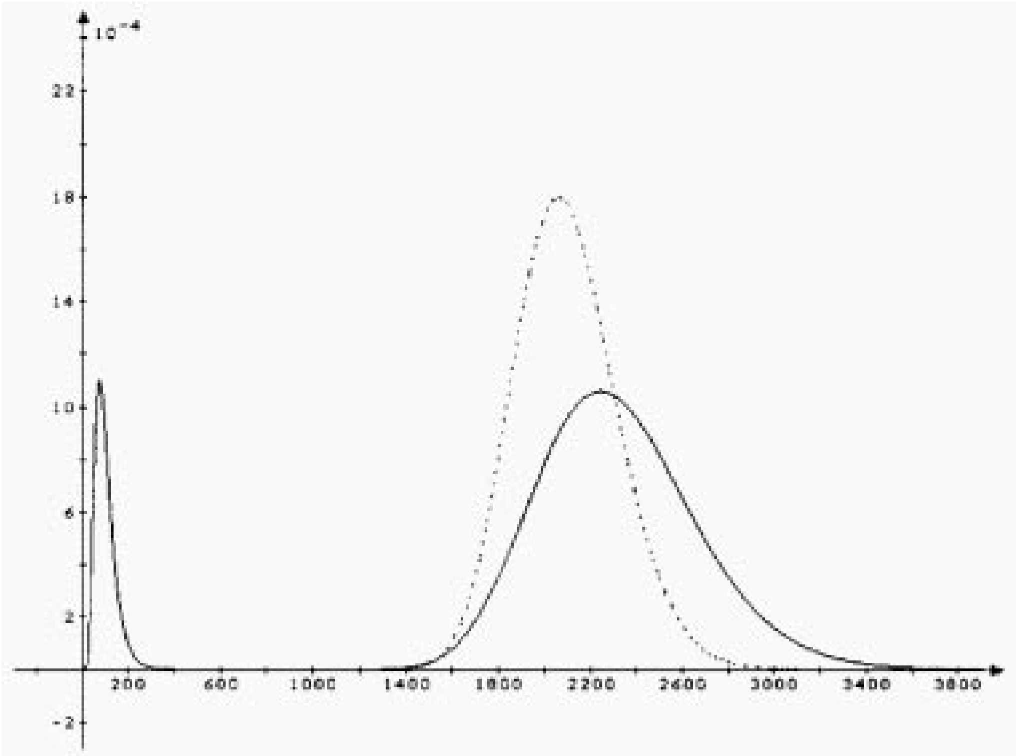

Figure 10: Implied probability distribution for the Black/Scholes (dotted line) and mixed lognormal model corresponding to the call prices at February 25, 1994 maturing in May 94.

4.3 The mixed lognormal probability distribution

Only for $\pi_1 > 0$ the distributions underlying the Black/Scholes model and the mixed lognormal model differ. Figure 10 shows a typical implied probability distribution for both the Black/Scholes and the mixed lognormal model. The shape of the probability distribution shows, that compared to the Black/Scholes formula, probability mass is attributed to low underlying prices and the main part of the probability mass is shifted to the right.

Before interpreting the shape of the probability distribution, the implications for the valuation of call options are considered. For calls, only the probability mass to the right of the exercise price is relevant for pricing. Since the empirically observed exercise prices are concentrated around the observed underlying price, the probability mass π_1 reduces the prices for all call options in the sample, because all options are out of money there. This effect can be offset by the shift of the remaining probability mass to higher underlying prices. Consequently, one has to be careful with the interpretation of the first mixture component, because this part of the distribution does not contribute directly to call option prices. Therefore, the second crash–o–phobia mode is not an indication of underlying prices falling to a level of 200 in case of a crash. The second mode may rather be interpreted as a risk premium for the general possibility of falling prices expressed in probabilistic terms. This interpretation is also confirmed by the comparison of the mixed distribution to the nonparametric underlying probability distribution calculated using Rubinstein's (94) method. First tests show, that crash–o–phobia situations are almost always exaggerated by the mixed distribution, because two mixture components are not sufficient to represent the markets probability assessments correctly. But further empirical research is needed to verify this assertion.

5 Summary and conclusion

The poor empirical result of the Black/Scholes model motivated the construction of an option pricing model with a mixed probability distribution. For the mixture of two lognormal probability distributions, the corresponding option pricing formula is simply a linear combination of Black/Scholes option prices. Nevertheless, the restrictive assumption of a constant volatility under the Black/Scholes model is avoided when using the mixed lognormal model. The mixture of lognormal distributions implies a randomly changing volatility and a flexible shape for the underlying price distribution. The additional parameters π_1, S^1 and σ_1 of the distribution represent the uncertainty of future underlying price decreases. To calculate the contribution of the first mixture component to total risk, the share of downside risk SDR is defined. Since the Black/Scholes formula is a special case of the mixed lognormal formula, the latter explains observed option prices at least as good as the former.

The empirical investigation shows that for most times the Black/Scholes model is adequate for the pricing of DAX index options. But in 25% of the cases the mixed lognormal formula improves the explanation of observed option prices by shifting probability mass to regions of low underlying prices. In these cases the systematic overpricing for out of money calls and the systematic underpricing for in the money calls under the Black/Scholes model is avoided when using the mixed lognormal distribution. The better pricing performance of the mixed lognormal model is an implication of the more flexible shape of the underlying mixed probability distribution.

Several questions are adressed to further research:

- First, put options should be included for the empirical investigation. Since for puts the probability mass attributed to low index levels is much higher than for calls (see Neumann/Schlag (96)), the valuation of puts should improve significantly under a mixed probability distribution.

- Second, the implied probability distributions should be analyzed in greater detail. Especially the fit of the mixed lognormal distribution to the data should be examined. Therefore, the mixed distribution can be compared to the nonparametric underlying probability distribution calculated, for example, using Rubinstein's (94) method. The latter comparison shows, whether the mixed lognormal distributions shape is flexible enough to represent the markets probability assessments correctly. In the case of bad representation more mixture components with multiple modes must be considered.

- Third, the price process(es) which are compatible to the mixed lognormal distribution should be considered. The question is, whether jump diffusion models, stochastic volatility models or models with deterministically changing volatility are compatible to a mixed lognormal distribution. The identification of the corresponding price process is important especially for the valuation of American options.

A The volatility of the mixed probability distribution

In Section 2 we consider the mixed probability distribution f_Q^{mix} for the underlying price S_T which consists of the two distributional components f_1^{log} and f_2^{log}:

$$f_Q^{mix}(S_T) = \pi_1 f_1^{log}(S_T) + (1-\pi_1) f_2^{log}(S_T). \tag{1}$$

For the expected value of the mixed distribution we calculate

$$\begin{aligned} E^Q(S_T) &= \int_{-\infty}^{\infty} S_T f_Q^{mix}(S_T) dS_T \\ &= \pi_1 E^{Q_1}(S_T) + (1-\pi_1) E^{Q_2}(S_T). \end{aligned} \tag{2}$$

Therefore, the expected value of the mixed distribution is a linear combination of the expected values of its distributional components.

The variance of the mixed distribution can be calculated using the relationship

$$Var^Q(S_T) = \int_{-\infty}^{\infty} (S_T - E^Q(S_T))^2 f_Q^{mix}(S_T) dS_T.$$

Inserting equation (1) results in

$$Var^Q(S_T) = \pi_1 \int_{-\infty}^{\infty} (S_T - E^Q(S_T))^2 f_1^{log}(S_T) dS_T + (1-\pi_1) \int_{-\infty}^{\infty} (S_T - E^Q(S_T))^2 f_2^{log}(S_T) dS_T. \tag{3}$$

Now inserting $E^Q(S_T)$ from equation (2) and performing simple calculations we achieve

$$\begin{aligned} Var^Q(S_T) &= \pi_1 \int_{-\infty}^{\infty} (S_T - \pi_1 E^{Q_1}(S_T))^2 f_1^{log}(S_T) dS_T \\ &\quad + (1-\pi_1) \int_{-\infty}^{\infty} (S_T - (1-\pi_1) E^{Q_2}(S_T))^2 f_2^{log}(S_T) dS_T \\ &\quad + \left(\pi_1 E^{Q_1}(S_T) - (1-\pi_1) E^{Q_2}(S_T)\right)^2 \\ &\quad - \left(\pi_1^3 E^{Q_1}(S_T)^2 + (1-\pi_1)^3 E^{Q_2}(S_T)^2\right). \end{aligned}$$

Further calculations lead to

$$\begin{aligned} Var^Q(S_T) &= \pi_1 \cdot E^{Q_1}(S_T^2) + (1-\pi_1) \cdot E^{Q_1}(S_T^2) \\ &\quad - \pi_1^2 \cdot E^{Q_1}(S_T)^2 - (1-\pi_1)^2 \cdot E^{Q_2}(S_T)^2 \\ &\quad - 2\pi_1(1-\pi_1) E^{Q_1}(S_T) E^{Q_2}(S_T) \end{aligned}$$

Now using $E(x^2) = Var(x) + E(x)^2$ finally leads to

$$\begin{aligned} Var^Q(S_T) &= \pi_1 \cdot Var^{Q_1}(S_T) + (1-\pi_1) \cdot Var^{Q_2}(S_T) \\ &\quad + \pi_1 \cdot (1-\pi_1) \left(E^{Q_1}(S_T) - E^{Q_2}(S_T)\right)^2. \end{aligned}$$

For an interpretation of this formula see section 2.

B Empirical estimation of the parameters of the mixed probability distribution

In our empirical study we estimate the parameters S_t^1, S_t^2, σ_1, σ_2 and π_1 implicitly from the mixed option pricing formula. These values can be used to calculate the parameters of the mixed probability distribution. Therefore, we use the following relationship between the variance of the lognormally distributed underlying asset price and the volatility σ under the risk neutral measure for the component distributions:

$$Var^{Q_i}(S_T) = (S_t^i)^2 \cdot e^{2r(T-t)} \left(e^{\sigma_i^2(T-t)} - 1 \right), \quad i = 1, 2.$$

Consequently, the variance with respect to the parameters of the pricing formula is:

$$\begin{aligned} Var^Q(S_T) &= \Big(\pi_1 \cdot (S_t^1)^2 (e^{\sigma_1^2(T-t)} - 1) + (1 - \pi_1) \cdot (S_t^2)^2 (e^{\sigma_2^2(T-t)} - 1) \\ &\quad + \pi_1 \cdot (1 - \pi_1) \left(S_t^1 - S_t^2 \right)^2 \Big) e^{2r(T-t)}. \end{aligned}$$

References

[1] **Black, F.; M. Scholes:** *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy 81 (1973), 637-659.

[2] **Chiras, D.P.; S. Manaster:** *The Information Content of Option Prices and a Test of Market Efficiency*, Journal of Financial Economics 6 (1978), 213-234.

[3] **Harris, L.:** *Transaction Data Tests of the Mixture of Distributions Hypothesis,* Journal of Financial and Quantitative Analysis 22 (1987), 127-141.

[4] **Hull, J.; A. White:** *The Pricing of Options on Assets with Stochastic Volatilities,* Journal of Finance 42 (1987), 281-300.

[5] **Jackwerth, C.; M. Rubinstein:** *Recovering Probability Distributions from Option Prices*, Journal of Finance 5 (1996), 1611-1631.

[6] **Longstaff, F.:** *Option Pricing and the Martingale Restriction*, Review of Financial Studies 8 (1995), 1091-1124.

[7] **Melick, W.R.; C.P. Thomas:** *Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis*, Journal of Financial and Quantitative Analysis 32 (1997), 91-115.

[8] **Merton, R.C.:** *Option Pricing when Underlying Stock Returns are Discontinous*, Journal of Financial Economics 3 (1976), 125-143.

[9] **Merton, R.C.:** *Continuous–Time Finance*, Blackwell Publishers, 1990.

[10] **Neumann, M.; C. Schlag:** *Martingale Restrictions and Implied Distributions for German Stock Index Option Prices,* Universität Karlsruhe, Diskussionspapier Nr. 196, 1996.

[11] **Rubinstein, M.:** *Implied Binomial Trees*, Journal of Finance 49 (1994), 771-818.

[12] **Rubinstein, M.:** *As Simple as One, Two, Three*, RISK 8 (1995), No. 1, 44-48.

[13] **Shimko, D.:** *Bounds of Probability,* RISK 6 (1993), No. 4, 33-38.