Skewness and Kurtosis Trades

Oliver J. Blaskowitz[1]
Wolfgang K. Härdle[1]
Peter Schmidt[2]

[1] Center for Applied Statistics and Economics (CASE), Humboldt–Universität zu Berlin, 10178 Berlin, Germany
[2] Bankgesellschaft Berlin, Quantitative Research, 10178 Berlin, Germany

ABSTRACT In this paper we investigate the profitability of 'skewness trades' and 'kurtosis trades' based on comparisons of implied state price densities versus historical densities. In particular, we examine the ability of SPD comparisons to detect structural breaks in the options market behaviour. While the implied state price density is estimated by means of the Barle and Cakici Implied Binomial Tree algorithm using a cross section of DAX option prices, the historical density is inferred by a combination of a non–parametric estimation from a historical time series of the DAX index and a forward Monte Carlo simulation.

1 Introduction

From a trader's point of view implied state price densities (SPD's) may be used as market indicators and thus constitute a good basis for advanced trading strategies. Deviations of historical SPD's from implied SPD's have led to skewness and kurtosis trading strategies, Ait–Sahalia, Wang and Yared (2001). Blaskowitz and Schmidt (2002) investigated such strategies for the period from 04/97 until 12/99. The trades applied to European options on the German DAX index generated a positive net cash flow.

However, it is market consensus that option markets behavior changed as a consequence of the stock market bubble that burst in March 2000, Figure 1. The purpose of this paper is to examine the trading profitability and the informational content of both the implied and the historical SPD for the extended period from 04/97 to 07/02. Our analysis focuses on the ability of SPD skewness and kurtosis comparisons to detect structural breaks.

For this purpose we use EUREX DAX option settlement prices and DAX closing prices. All data is included in MD*Base (http://www.mdtech.de), a database located at CASE (Center for Applied Statistics and Economics, http://www.case.hu-berlin.de) of Humboldt-Universität zu Berlin.

We start by explaining skewness and kurtosis trades in Section 2. In Section 3 we motivate the transition from Black–Scholes implied and historical volatility comparisons to implied and historical SPD comparisons. The SPD estimation techniques are discussed in Section 4, and Section 5 presents the estimation results. Section 6 investigates the trading performance, Section 7 concludes.

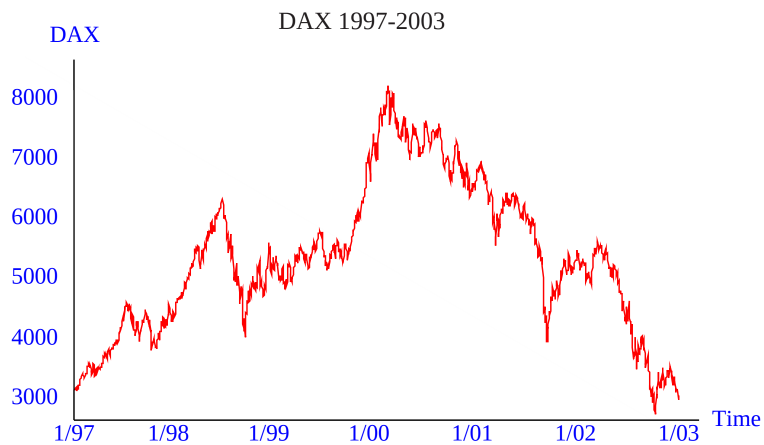

Figure 1. DAX from 01/97 to 01/03.

2 What are Skewness and Kurtosis Trades ?

In derivatives markets option strategies such as risk–reversals and strangles, Willmot (2002), are used to exploit asymmetric and fat–tailed properties of the underlyings' risk–neutral distribution. A risk–reversal is a portfolio of two European options with time to maturity τ. More precisely, it consists of a short position in a put with strike K_1 and a long position in a call with strike K_2, where $K_1 < K_2$. Its payoff profile at maturity as shown in Figure 2 suggests that an investor in this portfolio considers high prices of the underlying to be more likely to occur than low prices. Similarly, an investor believing that large moves of the underlying are likely to occur will buy a long strangle, which consists of a long position in a European put with strike K_1 and time to maturity τ and a long position in a European call with strike K_2 and time to maturity τ.

In our study we will use a risk–reversal and a modified strangle portfolio to exploit differences in two risk–neutral SPD's. To motivate the SPD comparison we recall the general pricing equations for European put and

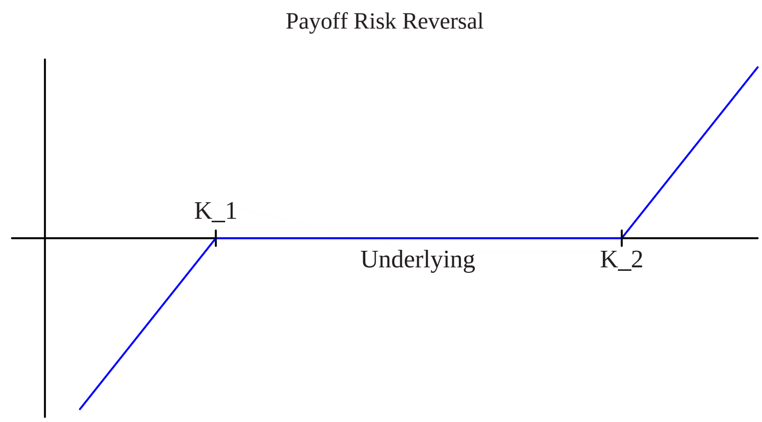

Figure 2. Payoff Risk Reversal

call options. From option pricing theory it follows that:

$$\begin{aligned} P &= e^{-r\tau} \int_0^\infty \max(K_1 - S_T, 0) q(S_T) dS_T \\ C &= e^{-r\tau} \int_0^\infty \max(S_T - K_2, 0) q(S_T) dS_T, \end{aligned} \quad (1.1)$$

where P and C are put respectively call prices, r is the risk–free interest rate, S_T is the price of the underlying at maturity T and q is a risk–neutral density, Franke, Härdle and Hafner (2001). Consider two risk–neutral densities denoted f^* and g^* as in Figure 3 where density f^* is more negatively skewed than g^*. Then equation (1.1) implies that the price of a European call option with strike K_2 computed with density f^* is lower than the price computed with density g^*. The reason for this is that f^* assigns less probability mass to prices $S_T > K_2$ than g^*. If the call is priced using f^* but one regards density g^* as a better approximation of the underlyings' distribution one would buy the option. Along these lines one would sell a put option with strike K_1, what finally results in a risk–reversal portfolio or, as well call it, a skewness 1 trade.

The same probability mass reasoning leads to kurtosis trades. We buy and sell calls and puts of different strikes as shown in Figure 4. The payoff profile at maturity is given in Figure 5. In Section 6 we will specify the regions in which to buy or sell options in terms of the moneyness $K/S_te^{r\tau}$.

Note, in a complete market model admitting no arbitrage opportunities exists exactly *one* risk–neutral density. If markets are not complete, for example when the volatility is stochastic, there are in general many risk–neutral measures. Comparing two risk–neutral densities, as we do, amounts

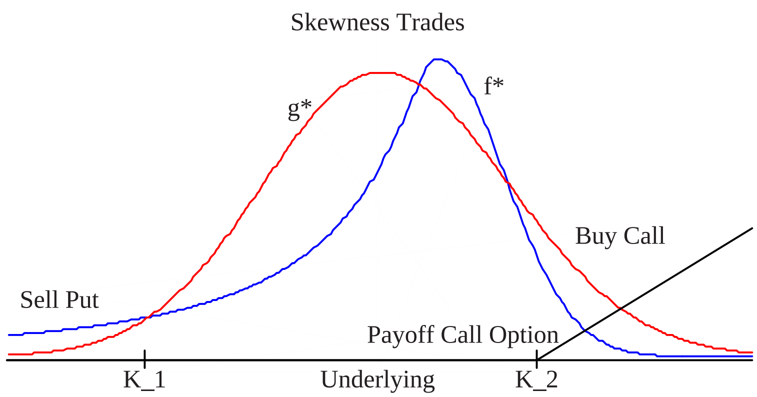

Figure 3. Skewness Trade

rather to comparing two different models, and trades are initiated depending on the model in which one believes more. The next section will discuss briefly how this approach is implemented in practice.

3 Skewness and Kurtosis in a Black–Scholes World

Black–Scholes' assumption that the underlyings' process S_t follows a geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where μ, σ are constants and dW_t is a Wiener Process, implies the underlying to be log–normally distributed with mean $exp\left(\mu dt\right)$ and variance $exp\left(2\mu dt\right)\left\{exp\left(\sigma^2\right)-1\right\}$. Skewness and kurtosis solely depend on the volatility parameter σ. As σ increases, skewness and kurtosis increase, as well. If there is only one implied volatility (IV) for all options, trading differences in Black–Scholes implied and historical volatilities, σ_{imp} respectively σ_{hist}, amounts to a comparison of two log–normal distributions. More precisely, such traders compare two Black–Scholes models with constant parameters σ_{imp} and σ_{hist}. Within this framework of two log–normals and a constant volatility one would buy all options if the historical volatility is higher than the IV and if one believes in such a 'historical volatility model'.

This way traders can trade volatility differences, but the assumption of log–normality does not allow the implementation of skewness and kurtosis trades. Comparing skewness and kurtosis requires the information contained in the Black–Scholes IV smile that is observed on option markets. Traders often use this smile to asses the markets' view on the underlyings' risk–neutral probabilistic behavior. Applying the inverted Black–Scholes

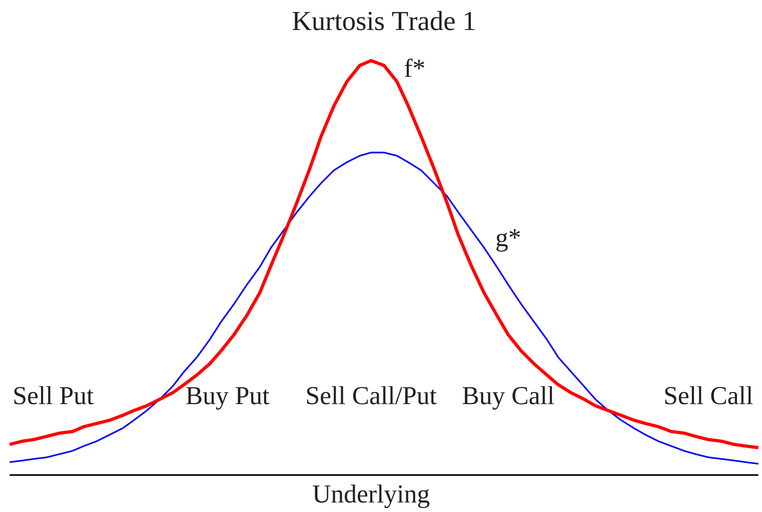

Figure 4. Kurtosis Trade 1

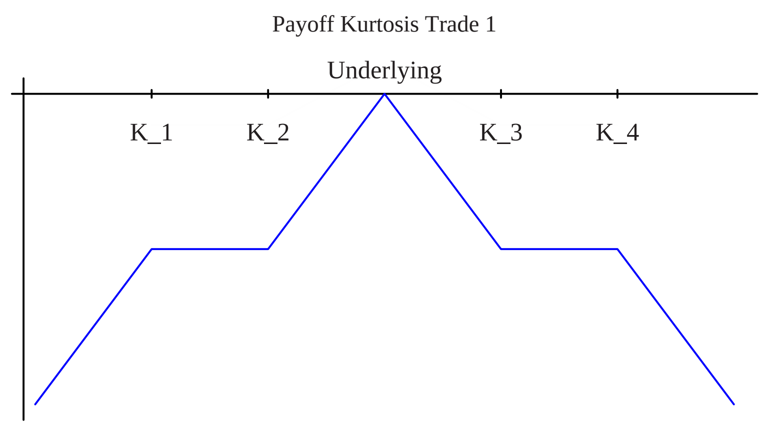

Figure 5. Payoff Kurtosis Trade 1

option pricing formula to prices obtained from a model with a left–skewed risk–neutral distribution, for example, entails out–of–the–money (OTM) calls to have a higher IV than at–the–money (ATM) calls or puts and the latter to have higher IVs than OTM puts. If the unknown risk–neutral distribution has more kurtosis than the log–normal distribution then OTM calls and OTM puts have higher Black–Scholes IVs than ATM options. Depending on the smiles' location, slope and curvature a skewness or a kurtosis trade is set up.

Relaxing the assumption that neither the implied nor the historical risk–neutral distribution is log–normal it is possible to extend trading rules based on Black–Scholes IVs to a framework where comparisons within a more general class of risk–neutral distributions are possible. In light of this, the approach we follow in this paper amounts to a generalization of a comparison of implied and historical volatility. In the following section we will briefly motivate the notion of implied and historical SPD's and describe the methods we used to extract them.

4 Implied and Historical DAX SPD's

Modern options markets have a high degree of market liquidity, i. e. prices on these markets are determined by supply and demand. This is particularly true for DAX options traded on the EUREX, the world's leading market for the trading of futures and options on stocks and stock indices. Some 425 participants in 17 countries traded more than 106 million options and futures contracts in March 2003. DAX options belong to the most frequently traded contracts (www.eurexchange.com).

Given a set of market option prices an implied distribution q is the distribution that simultaneously satisfies the pricing equation for all observed options in the set. As we work with European options, prices are given by equation (1.1). The implied state price density of an asset should be viewed as a way of characterizing the prices of derivatives contingent upon this asset. It is the density used to price options and has therefore a 'forward looking character', Cont (1998).

We will later see that the historical SPD is inferred from a time series of past underlyings' prices without involving option prices at all. Since we will use this distribution to compar it to the implied SPD, we call it a SPD too, a 'historical SPD'.

4.1 Extracting the Options Implied SPD

In recent years a number of methods have been developed to infer implied SPD's from cross–sectional option prices, see Cont (1998) and Jackwerth (1999) for an overview. As Binomial Trees are widely used and accepted by practitioners, we use Implied Binomial Trees (IBT) in order to obtain a proxy for the option implied SPD, which is denoted by f^* from now on. The IBT algorithm is a modification of the Cox–Ross–Rubinstein (CRR) algorithm. The numerous IBT techniques proposed by Rubinstein (1994), Derman and Kani (1994), Dupire (1994) and Barle and Cakici (1998) represent discrete versions of a continuous time and space diffusion model

$$\frac{dS_t}{S_t} = \mu\,(S_t, t)\,dt + \sigma(S_t, t)dW_t.$$

Whereas the classical CRR binomial tree assumes the instantaneous local volatility function to be constant, i. e. $\sigma(S_t, t) = \sigma$, the IBT allows $\sigma(S_t, t)$ to dependent on time and space.

Relying on the work of Härdle and Zheng (2002) we decided to work with Barle & Cakici's method for two reasons. First, the authors provide interactive XploRe quantlets to compute the IBT's proposed by Derman & Kani and Barle & Cakici. Second, according to the authors the latter method proved to be more robust.

The procedure works as follows: From a cross–section of two weeks of options data the XploRe quantlet **volsurf.xpl** estimates the IV surface over 'forward' moneyness and time to maturity, which we measure assuming 250 trading days per year. The quantlet **IBTbc.xpl** computes the IBT assuming a flat yield curve, a constant time to maturity of three months and taking the IV surface as input.

Furthermore, the IBT consists of three trees, the tree of stock prices, the tree of transition probabilities and finally the tree of Arrow–Debreu prices. If the tree is discretised by N time steps of length $\Delta t = \tau/N$, the tree consists of $N+1$ final stock prices $S_{N+1,i}, i \in \{1, 2, \ldots, N+1\}$, and $N+1$ final Arrow–Debreu prices $\lambda_{N+1,i}, i \in \{1, 2, \ldots, N+1\}$. Compounding the Arrow–Debreu prices to maturity $e^{r\tau}\lambda_{N+1,i}, i \in \{1, 2, \ldots, N+1\}$, and associating them to annualized stock returns

$$u_{N+1,i} = \{\log(S_{N+1,i}) - \log(S_t)\}\,\tau^{-1}\,,\; i \in \{1, 2, \ldots, N+1\}.$$

we obtain the option implied risk–neutral SPD f^* over log–returns. A more detailed description of the procedure is given in Blaskowitz and Schmidt (2002).

Figure 6 displays the implied SPD on Monday, June 23, 1997, and $N = 10$ time steps. This is the fourth Monday in June 1997. On that day, the DAX index S_t was at 3748.79 and the risk–free three month rate r was at 3.12. The plot shows the three months ahead risk–neutral SPD for Friday, September 19, 1997. This is the third Friday of September 1997, the expiration day of September 97 options. The SPD's standard deviation is 0.5, its skewness is -0.45 and its kurtosis is 4.17.

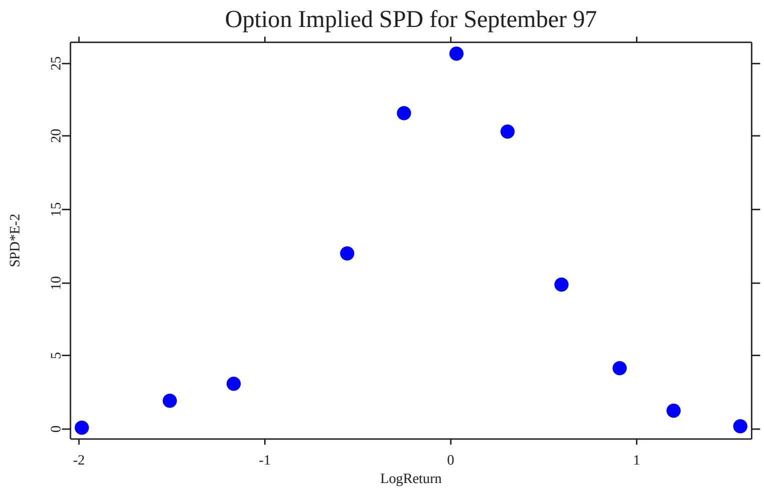

Figure 6. Option Implied SPD on Monday, June 23, 1997, for Friday, September 19, 1997, with $S_t = 3748.79, r = 3.12, N = 10$

We are interested in the SPD on the third Friday of the expiry month since later we will design the trading strategies such that we set up skewness and kurtosis portfolios on the 4th Monday of each month. These portfolios will consist of long and short positions in call and put options expiring on the 3rd Friday three months later.

4.2 Extracting the Historical SPD

The risk–neutral historical SPD g^* is estimated by assuming that the underlying S_t follows a continuous diffusion process:

$$dS_t = \mu(S_t)dt + \sigma(S_t)dW_t.$$

If we assume, as above, a flat yield curve and the existence of a bank account, which evolves according to $B_t = B_0 e^{rt}$, then from Itô's formula and Girsanov's theorem we obtain the risk–neutral dynamics:

$$dS_t^* = rS_t^* dt + \sigma(S_t^*)dW_t^*. \tag{1.2}$$

Note, since here the underlying is the DAX performance index, we do not take dividend yields into account.

The instantaneous local volatility function is identical under both the actual and the risk–neutral dynamics. It is estimated by means of Härdle and Tsybakov's (1997) non–parametric version of the minimum contrast estimator:

$$\hat{\sigma}^2(S) \quad = \quad \frac{\sum_{i=1}^{N^*-1} K_1(\frac{S_i - S}{h_1}) N^* \{S_{(i+1)/N^*} - S_{i/N^*}\}^2}{\sum_{i=1}^{N^*} K_1(\frac{S_i - S}{h_1})},$$

where K_1 is a kernel function, h_1 is a bandwidth parameter, S_i are discretely observed daily DAX closing prices and N^* is the total number of observed daily DAX closing prices. In the model specified in equation (1.2) $\hat{\sigma}^2(S)$ is an unbiased estimator of $\sigma^2(S)$.

Using three months of past daily DAX closing prices we estimate $\sigma^2(S)$ and then simulate $M = 10000$ paths of the diffusion process for a time period of 3 months:

$$dS_t^* \quad = \quad rS_t^* dt + \hat{\sigma}_a(S_t^*) dW_t^*,$$

with $\hat{\sigma}_a(S) = \hat{\sigma}(S)\,\tau^{-1}$ being the estimated annualized diffusion coefficient. As the DAX is a performance index, the continuous dividend yield is 0.

Collecting the endpoints of the simulated paths, we compute annualized log–returns:

$$u_{m,t} \quad = \quad \{\log(S_{m,T}) - \log(S_t)\}\,\tau^{-1}, m = 1, \ldots, M.$$

Using the notation $u = \log(S_T/S_t)$ and knowing that

$$P(S_T \le S) \quad = \quad P(u \le \log(S/S_t)) \quad = \quad \int_{-\infty}^{\log(S/S_t)} p_t^*(u) du$$

g^* is obtained by

$$g^*(S) \quad = \quad \frac{\partial}{\partial S} P(S_T < S) \quad = \quad \frac{\hat{p}^*\{\log(S/S_t)\}}{S},$$

where $\hat{p}^*$ is a non–parametric kernel density estimation of the continuously compounded log–returns. $\hat{p}^*$ is given by

$$\hat{p}^*(u) \quad = \quad \frac{1}{Mh_2} \sum_{m=1}^{M} K_2\Big(\frac{u_{m,t} - u}{h_2}\Big),$$

with K_2 being a kernel function and h_2 a bandwidth parameter. g^* is $\sqrt{N^*}$–consistent for $M \to \infty$ even though $\hat{\sigma}^2$ converges at a slower rate,

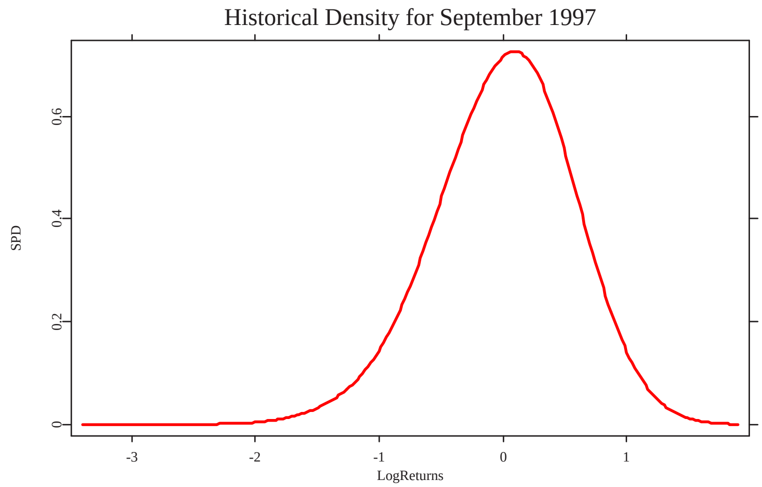

Figure 7. Historical SPD on Monday, June 23, 1997, for Friday, September 19, 1997, with $S_t = 3748.79, r = 3.12$

Ait–Sahalia, Wang and Yared (2001).

In order to satisfy the condition that under the absence of arbitrage the mean of the underlyings' risk–neutral density is equal to the futures price, we translate the Monte Carlo simulated historical density:

$$\bar{S} \quad = \quad S - E(S) + S_t e^{r_{t,\tau}}.$$

As for the SPD comparison later on we are only interested in the standard deviation, skewness and kurtosis measures of g^*. Because the annualized log–returns contain already all the necessary information, we finally compute only these statistics for the simulated log–returns.

Consider the period in the example above. On Monday, June 23, 1997, we used past daily DAX closing prices of the period of time between Monday, March 23, 1997, and Friday, June 20, 1997, to estimate σ^2. Following, on Monday, June 23, 1997, we simulate $M = 10000$ paths to obtain the three months ahead SPD, shown in Figure 7, whose standard deviation is 0.52, skewness is -0.39 and kurtosis is 3.23. Figure 8 illustrates both procedures.

5 SPD Comparison

In order to compare both SPD's we computed the three month implied and historical densities every three months. More precisely, in March we compute implied and historical densities for June. Following, we compute

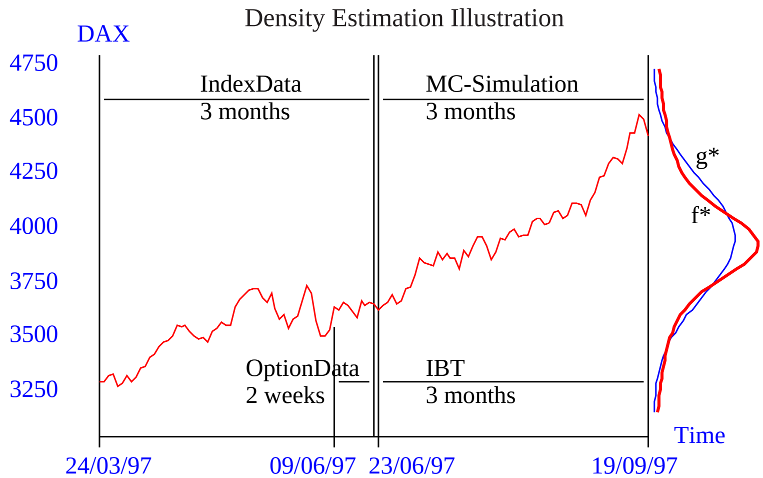

Figure 8. Comparison of procedures to estimate historical and implied SPD of Friday, 19/09/97. SPD's estimated on Monday, 23/06/97, by means of 3 months of index data respectively 2 weeks of option data.

in June the densities for September etc. The reason is that DAX options maturing in March, June, September and December are most liquid, thus containing most information. Starting in June 97 we estimate the first SPD's for September 97. We compare both SPD's by looking at the standard deviation, skewness and kurtosis.

Figure 9 shows the standard deviations of the implied (blue line with triangles) and the historical SPD (red line with circles). Although difficult to interpret, it appears that differences in standard deviations are less significant at the end of 1997 and in 1998. It seems that deviations in the dispersion become more pronounced from 1999 on.

In contrast, skewness and kurtosis measures of implied and historical SPD as shown in Figures 10 and 11 give a less unambiguous picture. Whereas the skewness signal changes in the beginning of 2001, the kurtosis comparison yields in almost all periods a one-sided signal, ignoring the outlier in September 2001.

Given that market participants agree upon a structural break occurring in March 2000, the methodology applied above does not seem to provide useful information about such a break in the options markets behavior.

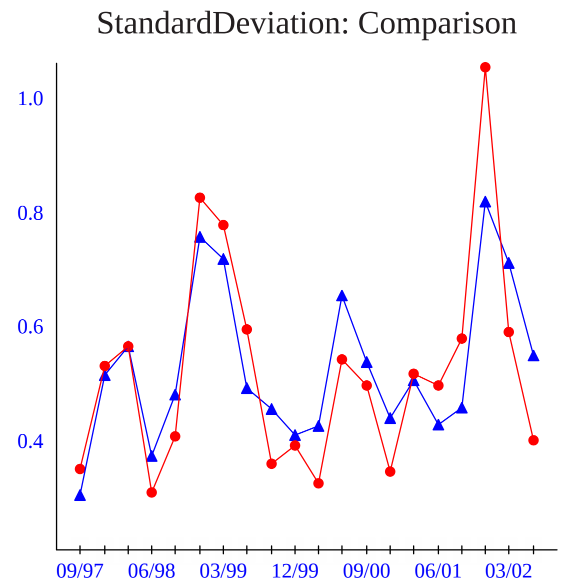

Figure 9. Comparison of Standard Deviations of Implied and Historical Densities. Historical and implied SPD are denoted by a circle respectively a triangle.

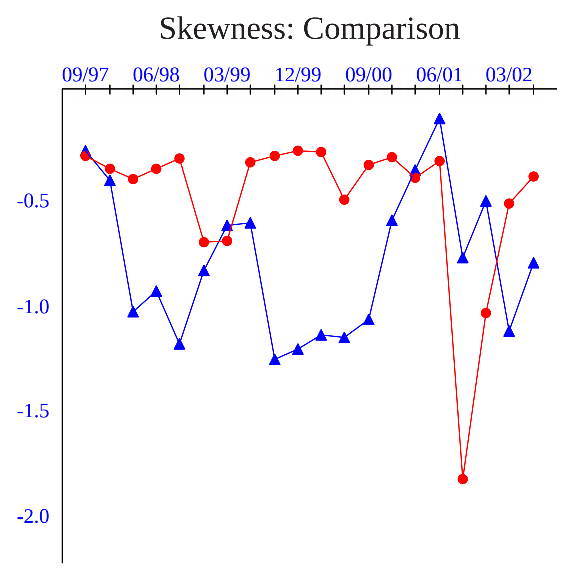

Figure 10. Comparison of Skewness of Implied and Historical Densities. Historical and implied SPD are denoted by a circle respectively a triangle.

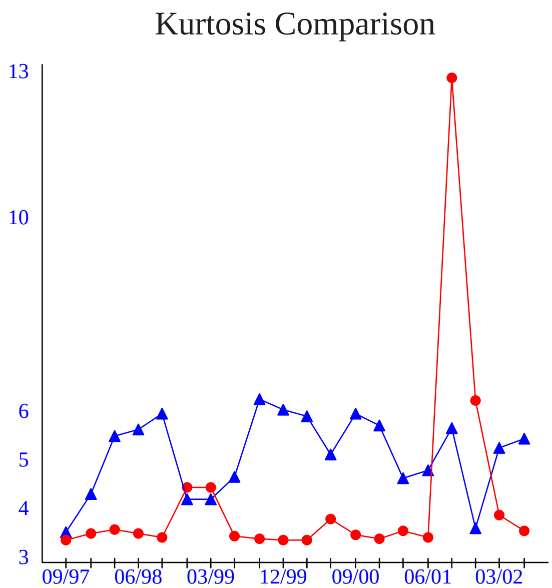

Figure 11. Comparison of Kurtosis of Implied and Historical Densities. Historical and implied SPD are denoted by a circle respectively a triangle.

6 Skewness and Kurtosis Trades

The results from the previous section indicate that the implied SPD is more negatively skewed and has a higher kurtosis than the historical SPD. In this section, we investigate the skewness 1 and kurtosis 1 trade performance for the periods from June 1997 to March 2000, from June 2000 to March 2002, and for the overall period from June 1997 to March 2002. Each period consists of non–overlapping three month subperiods, in which we set up portfolios of calls and puts with a time to maturity of three months. For each subperiod we measure the portfolio return by:

$$\text{portfolio return} \quad = \quad \frac{\text{net cash flow at maturity}}{\text{net cash flow at initiation}} - 1.$$

The investment to set up the portfolio comprises the net cash flow from buying and selling calls and puts. Whenever the options sold are worth more than the options bought entailing a positive cash inflow, it is not possible to compute a return measure. To ensure that the net cash flow at initiation is negative, we buy one share of the underlying for each call option sold and deposit on a bank account the cash value of each sold put options strike. Such an approach amounts to a very careful performance measurement. Applying margin deposits required by EUREX, for example, would lower the cash outflow at initiation and thus increase the profitability. Since a DAX option contract on EUREX consists of five options and one index point has a value of 5 EUR, we 'charge' 1 EUR for an index point. At maturity, we sum up all option payoffs, the bank account balance and the DAX value. For simplicity, we assume that the bank accounts interest rate is zero.

As for the skewness trade, we consider put options with a moneyness, $K/S_te^{r\tau}$, of less than 0.95 as OTM. We sell all OTM put options available in the market. We buy all available ITM call options, i. e. call options with a moneyness of more than 1.05, see Table 1.1. In our trading simulation one call or put is traded on each moneyness. As Table 1.2 shows, the performance for the two subperiods reversed. The annualized total returns as well as Sharpe ratios turned from positive to negative.

A kurtosis 1 portfolio is set up by selling and buying puts and calls as given in Table 1.1. The kurtosis trade performed similarly to the skewness trade. In the first period it was highly profitable. In the second period it turned out to be a bad strategy compared to a risk–free investment.

Skewness 1 Trade		Kurtosis 1 Trade	
Position	Moneyness	Position	Moneyness
short puts	< 0.95	short puts	< 0.90
		long puts	$0.90 - 0.95$
		short puts	$0.95 - 1.00$
		long calls	$1.00 - 1.05$
		short calls	$1.05 - 1.10$
long calls	> 1.05	long calls	> 1.10

Table 1.1. Skewness 1 Trade: Definitions of moneyness regions.

Skewness 1 Trade			
Period	06/97–03/00	06/00–03/02	Overall
Number of Subperiods	12	8	20
Total Return	4.85	-8.53	-2.05
Return Volatility	3.00	9.79	6.78
Minimum Return	-3.66	-25.78	-25.78
Maximum Return	7.65	7.36	7.65
Sharpe Ratio (Strategy)	0.10	-0.46	-0.24
Sharpe Ratio (DAX)	0.38	-0.35	0.02

Table 1.2. Skewness 1 Trade Performance. Only Total Return is annualized. Returns are given in percentages.

Kurtosis 1 Trade			
Period	06/97–03/00	06/00–03/02	Overall
Number of Subperiods	12	8	20
Total Return	14.49	-7.48	2.01
Return Volatility	3.87	13.63	9.33
Minimum Return	-4.54	-28.65	-28.65
Maximum Return	8.79	18.14	18.14
Sharpe Ratio (Strategy)	0.55	-0.32	-0.05
Sharpe Ratio (DAX)	0.38	-0.35	0.02

Table 1.3. Kurtosis 1 Trade Performance. Only Total Return is annualized. Returns are given in percentages.

7 Conclusion

Given that the trading performance of the skewness as well as the kurtosis trade differ significantly in both subperiods, it is disappointing that the SPD comparison does not reveal a similar pattern. One could argue that what we see within the two subperiods is just a feature of the risk premium as pointed out by Ait–Sahalia, Wang and Yared (2001). However, as market participants agree that options markets behave differently since March 2000, we believe that there is more to exploit from a SPD comparison. In light of this, a topic for future research will be to investigate different methodologies with respect to their potential to improve and fine tune such a SPD comparison.

8 References

Ait–Sahalia, Y., Wang, Y. and Yared, F. (2001). Do Option Markets correctly Price the Probabilities of Movement of the Underlying Asset?, *Journal of Econometrics* **102**: 67–110.

Barle, S. and Cakici, N., (1998). How to Grow a Smiling Tree, *The Journal of Financial Engineering* **7**: 127–146.

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

Blaskowitz, O. and Schmidt, P. (2002). *Trading on Deviations of Implied and Historical Density*, in W. Härdle, T. Kleinow, G. Stahl: Applied Quantitative Finance, Springer Verlag, Heidelberg.

Cont, R. (1998). *Beyond Implied Volatility: Extracting Information from Options Prices*, in Econophyiscs, Dodrecht: Kluwer.

Derman, E. and Kani, I. (1994). The Volatility Smile and Its Implied Tree, *http://www.gs.com/qs/*

Dupire, B. (1994). Pricing with a Smile, *Risk* **7**: 18–20.

Florens–Zmirou, D. (1993). On Estimating the Diffusion Coefficient from Discrete Observations, *Journal of Applied Probability* **30**: 790–804.

Franke, J., Härdle, W. and Hafner, C. (2001). *Einführung in die Statistik der Finanzmärkte*, Springer Verlag, Heidelberg.

Härdle, W. and Simar, L. (2003). *Applied Multivariate Statistical Analysis*, Springer Verlag, Heidelberg.

Härdle, W. and Tsybakov, A., (1997). Local Polynomial Estimators of the Volatility Function in Nonparametric Autoregression, *Journal of Econometrics*, **81**: 223–242.

Härdle, W. and Zheng, J. (2002). *How Precise Are Price Distributions Predicted by Implied Binomial Trees?*, in W. Härdle, T. Kleinow, G. Stahl: Applied Quantitative Finance, Springer Verlag, Heidelberg.

Jackwerth, J.C. (1999). Option Implied Risk Neutral Distributions and Implied Binomial Trees: A Literature Review, *The Journal of Derivatives* **Winter**: 66–82.

Kloeden, P., Platen, E. and Schurz, H. (1994). *Numerical Solution of SDE Through Computer Experiments*, Springer Verlag, Heidelberg.

Rubinstein, M. (1994). Implied Binomial Trees, *Journal of Finance* **49**: 771–818.

Willmot, P. (2002). *Paul Willmot Introduces Quantitative Finance*, Wiley.