

Implied Volatility Modeling

Sarves Verma, Gunhan Mehmet Ertosun, Wei Wang, Benjamin Ambruster,

Kay Giesecke

I Introduction

Although Black-Scholes formula is very popular among market practitioners, when applied to call and put options, it often reduces to a means of quoting options in terms of another parameter, *the implied volatility*. Further, the function

$$\sigma_t^{BS} : (K, T) \longrightarrow \sigma_t^{BS} (K, T) \dots\dots\dots(1)$$

is called the *implied volatility surface*. Two significant features of the surface is worth mentioning”: a) the non-flat profile of the surface which is often called the ‘smile’ or the ‘skew’ suggests that the Black-Scholes formula is inefficient to price options b) the level of implied volatilities changes with time thus deforming it continuously. Since, the black-scholes model fails to model volatility, modeling implied volatility has become an active area of research. At present, volatility is modeled in primarily four different ways which are : a) The *stochastic volatility model* which assumes a stochastic nature of volatility [1]. The problem with this approach often lies in finding the market price of volatility risk which can’t be observed in the market. b) The deterministic volatility function (DVF) which assumes that volatility is a function of time alone and is completely deterministic [2,3]. This fails because as mentioned before the implied volatility surface changes with time continuously and is unpredictable at a given point of time. Ergo, the lattice model [2] & the Dupire approach [3] often fail[4] c) a factor based approach which assumes that implied volatility can be constructed by forming basis vectors. Further, one can use implied volatility as a mean reverting Ornstein-Uhlenbeck process for estimating implied volatility[5]. However, estimating parameters for such processes is very difficult and one needs to fit the parameters to such data. Further, one needs to check whether these parameters satisfy the arbitrage bounds as specified by Lee et.al.[6]d) the last but the most commonly used method is an empirical way to fit data (using statistical methods)

involving both parametric [7] & non-parametric regression[8].For most of these models, PCA (principal component analysis) together with GARCH seems an obvious choice. Using these methods mentioned above, researchers have performed an in depth analysis of implied volatility. For example, [9] performed a PCA analysis on different maturity buckets of options to study how different loading factors impact implied volatility for each distinct bucket.

In this work, we extend the idea of [9] and classify options both on the basis of *moneyness and maturity* i.e. we form maturity & moneyness buckets and study the impact of different PCA factors on implied volatility. We believe that this will give a clear idea to a trader; which factor to look for when hedging an option of a specific moneyness and a specific maturity. In this context, we also come across a novel way of looking at gamma and vega (greeks) using principal components. Further, we also develop a comprehensive model to incorporate the effect of maturity on implied volatility. Section II describes the methodology while Section III deals with our results & interpretation of those results. Finally we conclude with Section IV.

II Data Collection & Model

In our work, we consider call option prices on S&P 500 index (ticker SPX) which we re obtained from optionmetrics¹. The data was considered from June 1, 2000 to June 20, 2001. Note that these years were turbulent owing to a bubble burst and resulted in high volatility. Once the data was obtained, following was done to sort data for our use:

- a) All options with less than 15 days of maturity were ignored as they result in high volatility.
- b) Data values with call prices less than 10 cents were also ignored.
- c) Average value of ask & bid price was taken to represent the call price.
- d) All call prices which were less than the theoretical value (calculated using Black-Scholes) were ignored for arbitrage reasons

We then divide the entire data set in moneyness (represented by m) buckets of $m < -1$, $-1 < m < -0.5$, $-0.5 < m < 0$, $0 < m < 0.5$, $0.5 < m < 1$, $1 < m$ where the moneyness is defined as

$$m = \ln\left(\frac{S_t e^{r\tau}}{K}\right) / \sqrt{\tau} \dots\dots\dots(2)$$

Where S_t = Index value at time t

r = the risk free rate of interest (as given by treasury)

T = maturity date

τ = time to maturity of options (T-t)

K = strike price

Taking m as in (2) incorporates the effect of time (τ) & strike price (K) in moneyness.

We also divide the entire data into different maturity buckets of 15-30, 30-60, 60-90, 90-150, 150-250 days. We then build a model to incorporate the effect of maturity and moneyness in implied volatility. The following four different models were constructed, simulated and compared:

$$\text{Log}(I) = \eta_0 + \varepsilon \dots\dots\dots(3)$$

$$\text{Log}(I) = \eta_0 + \eta_1 m + \eta_2 m^2 + \varepsilon \dots\dots\dots(4)$$

$$\text{Log}(I) = \eta_0 + \eta_1 m + \eta_2 m^2 + \eta_3 \tau + \eta_4 \tau m + \varepsilon \dots\dots\dots(5)$$

$$\text{Log}(I) = \eta_0 + \eta_1 m + \eta_2 m^2 + \eta_3 \tau + \eta_4 \tau m + \eta_5 \tau^2 + \varepsilon \dots\dots\dots(6)$$

We considered these models and then compared the results to see whether the later models are accurate. Further, we performed PCA on the sorted data mentioned above, both in terms of moneyness bucket as well as maturity bucket.

III PCA (Principal Component Analysis)

A) PCA based on moneyness buckets

The PCA analysis was done both on moneyness buckets and maturity buckets, we first consider the moneyness bucket. Fig. 1 shows the implied volatility Vs strike price for a fixed maturity of 60 days. One can see the ‘skew’ in the figure. Table I summarizes the percentage contribution of first three principal components to the total variance.

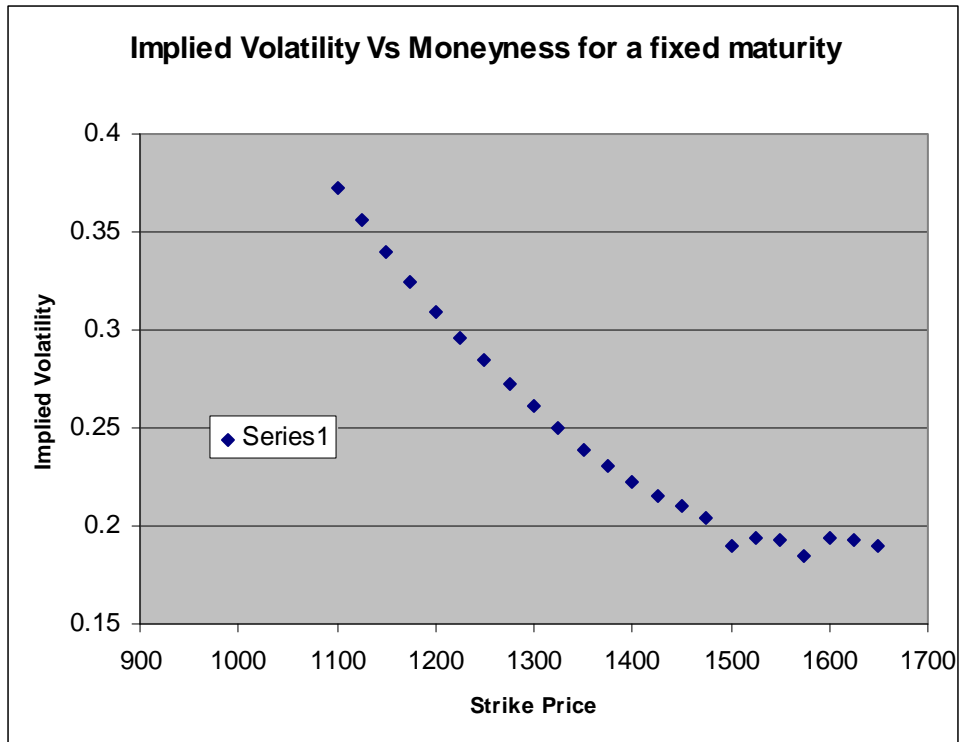


Fig. 1 Implied Volatility Vs Strike Price

Moneyness of Call Option	1st PC (in %)	2nd PC (in %)	3rd PC (in %)	Total explained Variance By 1st three PCs (in %)
m<-1	51.561	39.379	9.0596	100
-1<m<-0.5	50.548	26.729	11.646	88.923
-0.5<m<0	45.248	23.932	18.656	87.836
0<m<0.5	50.017	19.536	16.346	85.899
0.5<m<1	37.732	24.999	22.1	84.831
m>1	62.871	23.417	10.996	97.284

Table 1. Contribution of different principal components towards the total variance of implied volatility for call options

The following are the worth noting observations:

- a) As we traverse from 'out of moneyness' ($m < -1$) to 'at the moneyness' and then to 'in the moneyness' ($m > 1$) for call options, we find (fig.2) that the total variance in the implied volatility explained by the first three components first decreases (from 100% for $m < -1$ to 84.831% for $0.5 < m < 1$) and then increases (to 97% for $m > 1$). We believe that this happens because 'at the moneyness' is highly 'unstable' or most sensitive to hedging and hence can be hardly explained by just the first three components. In contrast, 'out of money' and 'in the money' options are relatively illiquid and stable and just need the first three factors to fully explain it.
- b) The first principal component which represents the mean level; decreases initially till around the 'at the money' level and then increases sharply. As we traverse from out of money to in the money, we find (fig. 3) the implied volatility flattening out & hence, the contribution of first principal component towards total variance increases sharply. Again as described before, the 'at the money' regime is highly sensitive (has higher variance) and the first principal component is not sufficient to explain it alone. Therefore, we expect a dip in the contribution of the first principal component in that regime.
- c) The second principal component (contribution towards total variance) which represents slope or tilt is expected to flatten out as we move from at the money to in the money. This is because; the skew flattens out itself, resulting in a lower contribution from the slope (representing the second principal component) towards the total variance. We find what we expect in our results (fig. 4).
- d) The third principal component which represents the curvature of the implied volatility is crucial for hedging. We find (fig. 5) that the contribution of the third component towards total variance peaks near the 'at the money' regime. To understand this, we can look at 'gamma', or the curvature of the call price Vs index price curve which also peaks around the 'at the money'. Implied volatility Vs moneyness can be understood as call price (which is directly proportional to volatility) Vs Index price (moneyness has a log dependence on the index price and is very sensitive to it).

Hence, we believe that the third principal component is a novel way of thinking about the gamma hedging.

The figures summarize most of the details mentioned above.

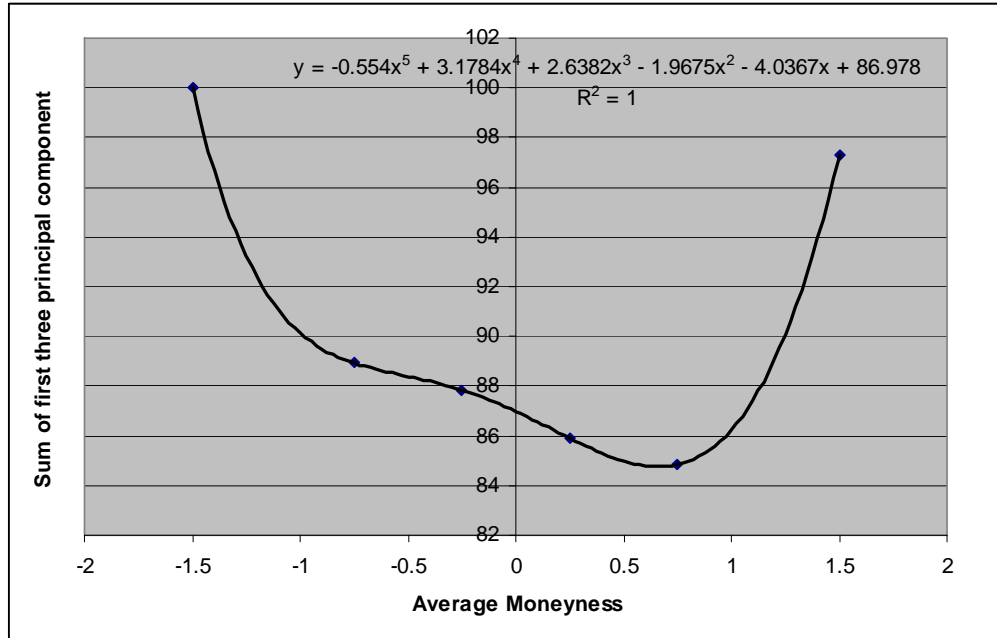


Fig. 3 Percentage contribution towards total variance by the first three components

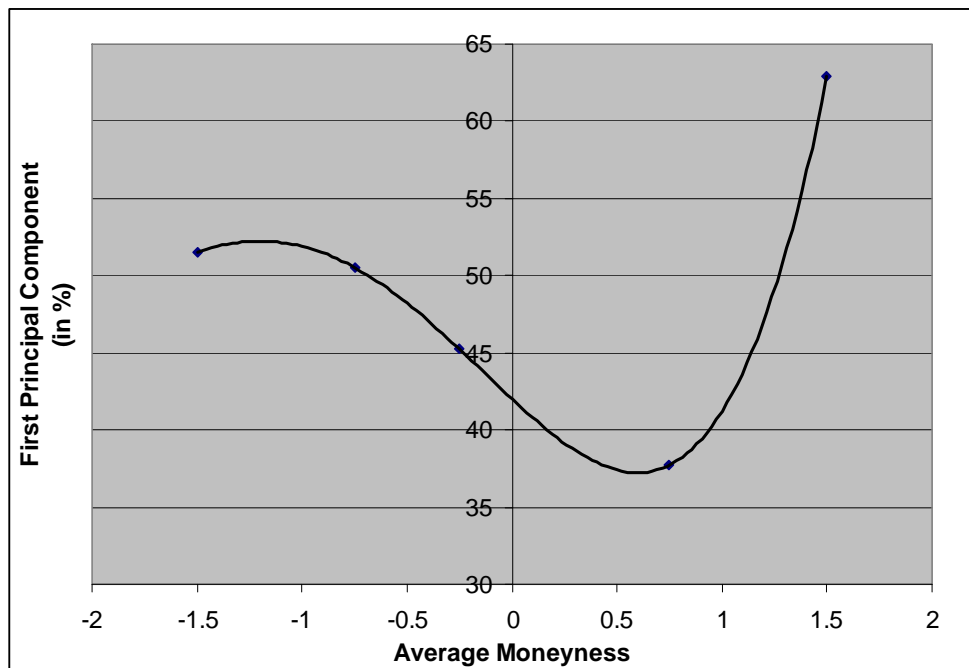


Fig. 4 Percentage contribution towards total variance by the first component

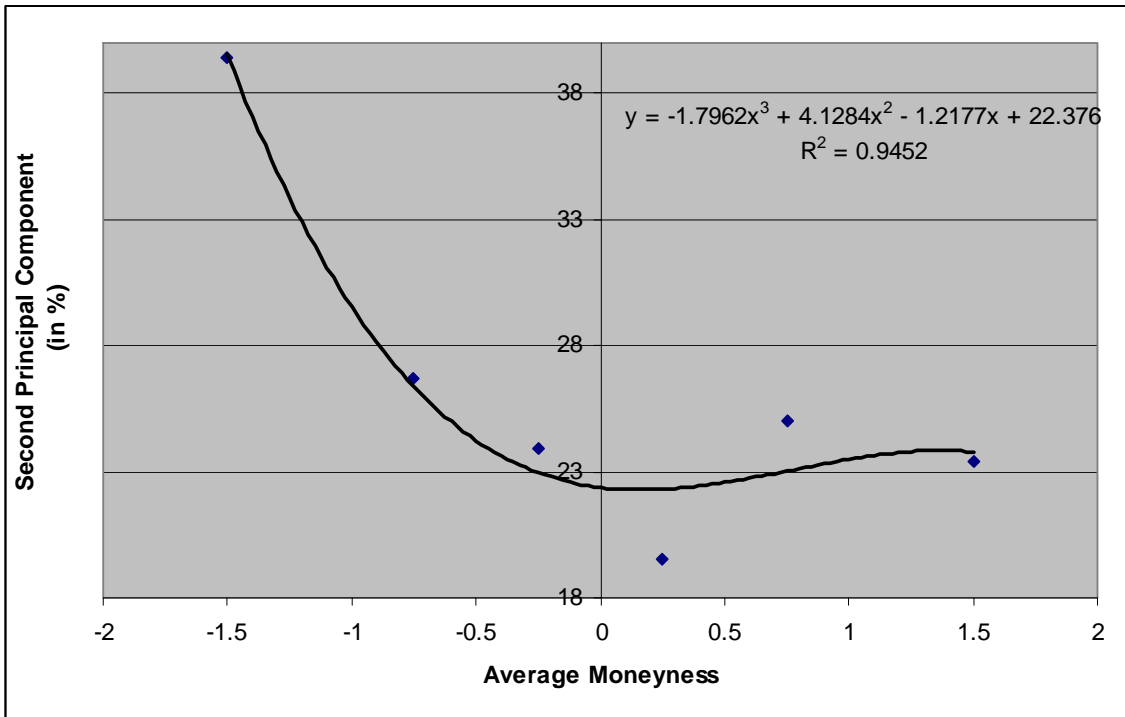


Fig. 5 Percentage contribution of Second principal component towards total variance

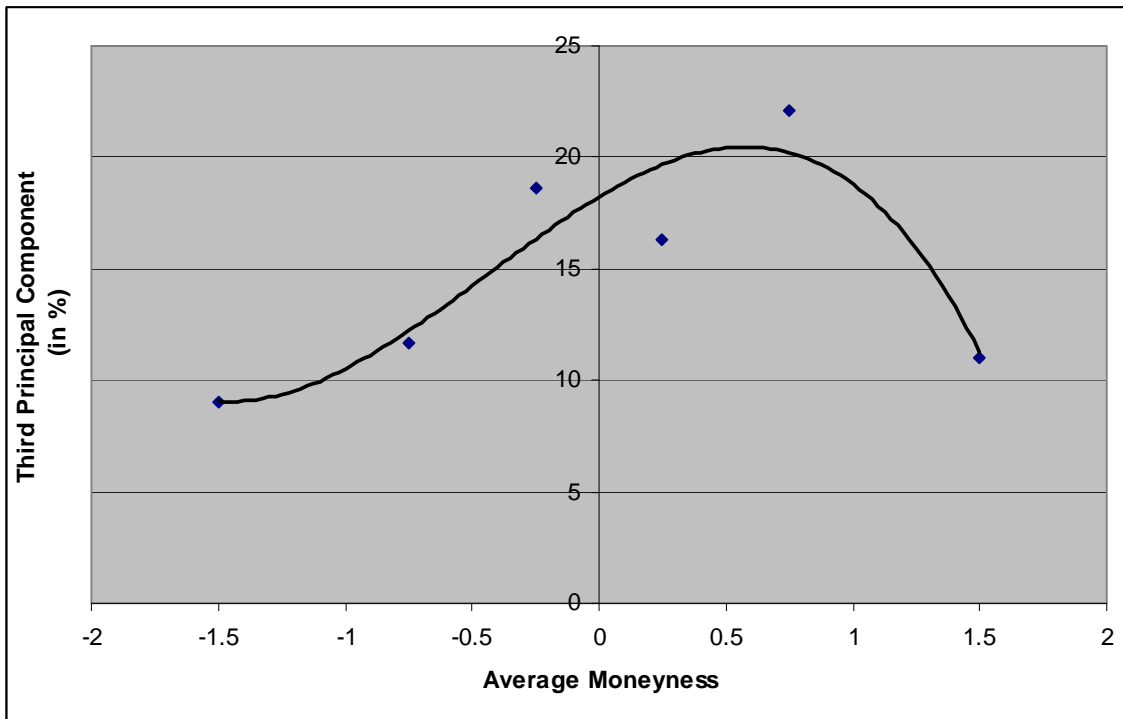
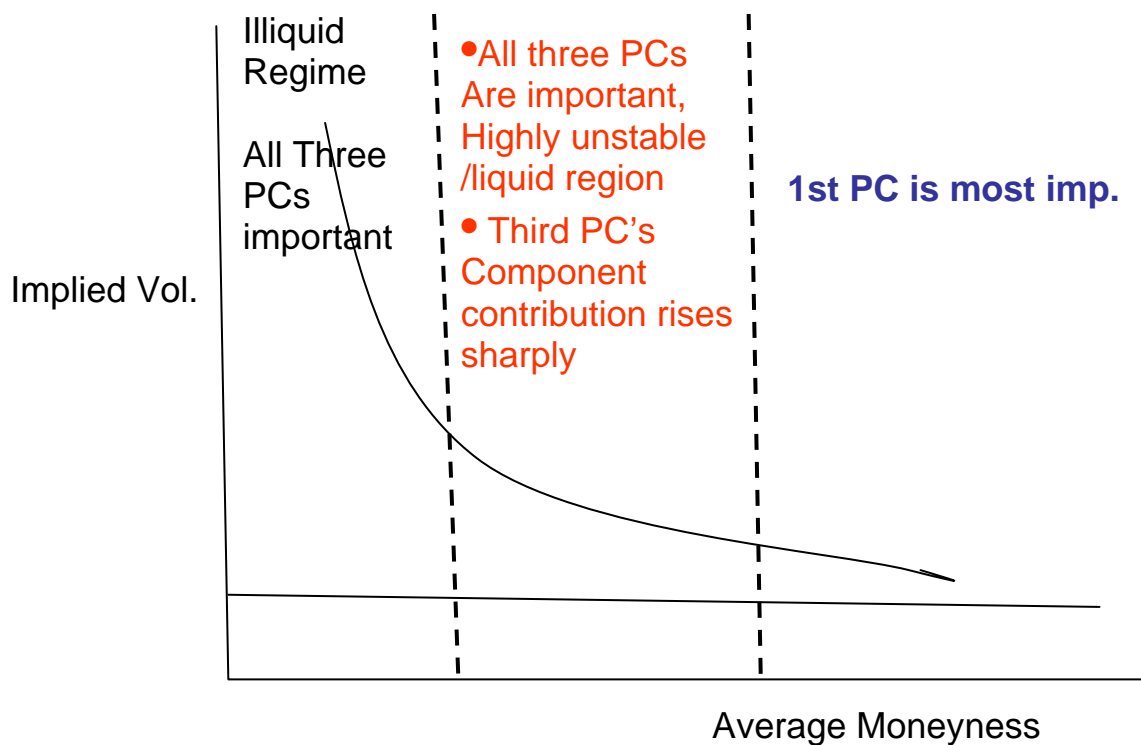


Fig. 6 Percentage contribution of Third principal component towards total variance

Note that the figure above is skewed towards left

To summarize, we find that PCA analysis of the moneyness buckets give a good insight of the underlying call option itself. It also provides a new way to understand option hedging. As we observe from the figures above, the first principal component is most important for the ‘in the money’ case, while in the ‘at the money’ regime, the third component becomes increasingly important. Still in the ‘at the money’ regime, none of the three components are sufficient to fully describe the variance in the implied volatility. Further, the out of the money regime is relatively illiquid and has all aspects of levelness, steepness & curvature to it (i.e. all three components are important).



The summary has been put in the form of a cartoon above.

B) PCA based on Maturity Buckets

Performing PCA on maturity buckets have already been done by [9]. The aim of our work is to verify the result. We find that for short term maturities, all three principal components are equally important while for long term maturities only the first principal component matters. The result obtained by us has been summarized in Table 2.

Maturity of Call Option	1st PC (in %)	2nd PC (in %)	3rd PC (in %)	Total explained Variance By 1st three PCs (in %)
15-30	56.929	21.359	12.072	90.41
30-60	69.426	15.266	10.496	95.18
60-90	88.71	5.41	2.79	96.92
90-150	81.419	10.712	7.2489	98.83
150-250	77.38	15.55	4.58	97.5

Table 2 Percentage contribution of principal components towards total variance

We also performed PCA on both moneyness and maturity bucket, as shown in table 3. From the table we can see that for long term maturity and out of money bucket the first principal component dominant. However, as the options step out of money the second and third component become more and more important.

	-1<m<-0.5	-0.5<m<0	0<m<0.5	0.5<m<1
Short Maturity (T<90)	38.1616 32.2998 29.5445	50.1882 39.8367 15.9751	42.7924 36.4123 20.7953	42.3005 31.1038 26.5957
Long Maturity (90<T<250)	71.6205 15.9579 12.4215	56.222 27.8128 15.9652	38.4003 36.4862 25.1135	41.1661 31.7541 27.0798

Table 3 Percentage contribution of principal components towards total variance with respect to both moneyness and maturity

IV Results on Model & Comparison

As discussed earlier, we developed a model incorporating both the effect of maturity and moneyness (described by equations 3-6). The results for each model are shown below. Note that in each figure the red colored points represent the real data while the blue points represent the fit. On the left one can see the resulting plot during the fitting process to extract the parameters and on the right one can see the result of the out of sample prediction using the next years' (June 2001-2002) data.

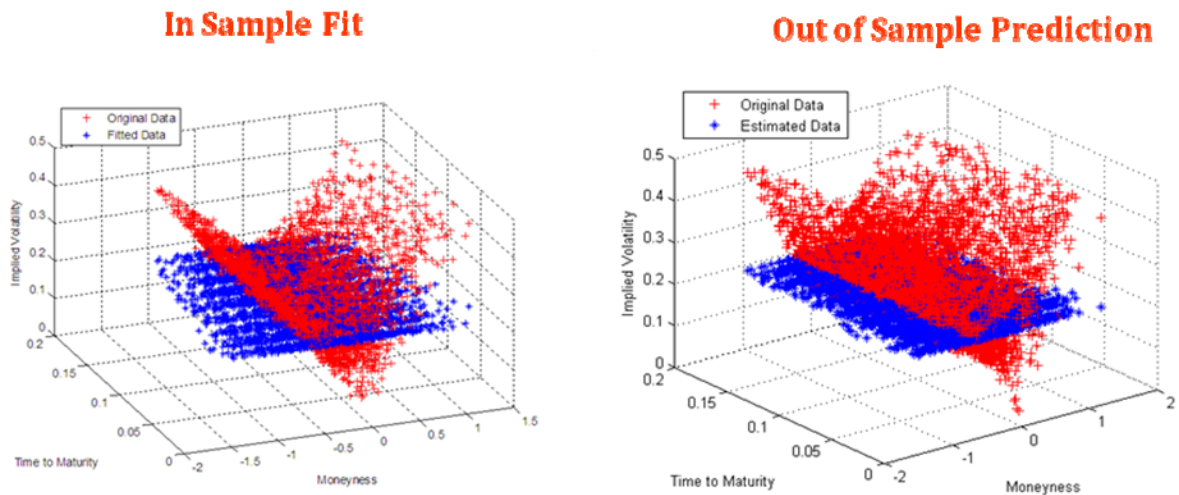
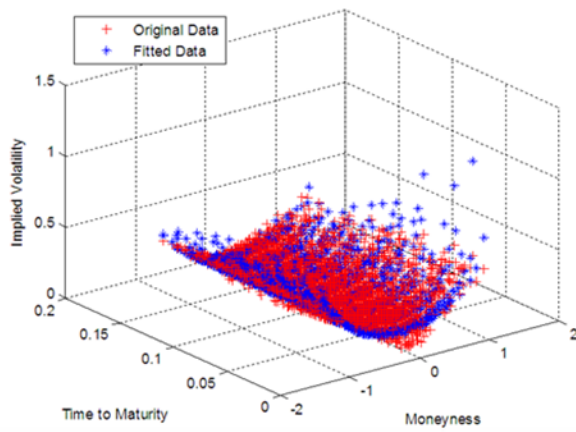


Figure 7 Model representing equation 3 (the Black-Scholes model which assumes constant volatility)

In Sample Fit



Out of Sample Prediction

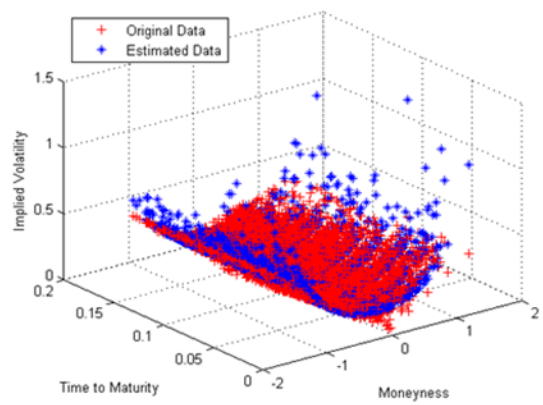
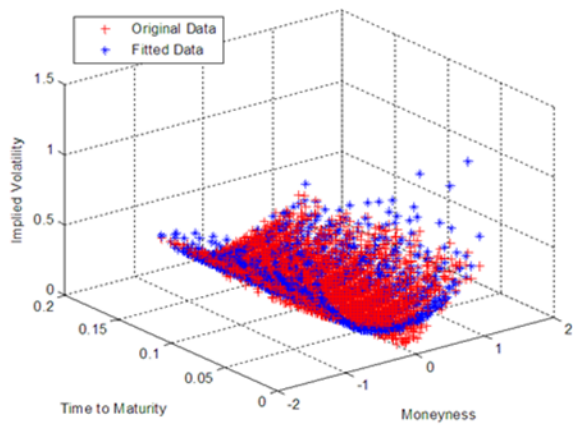


Figure 8 Showing Model represented by equation (4) which takes in account both steepness & curvature.

In Sample Fit



Out of Sample Prediction

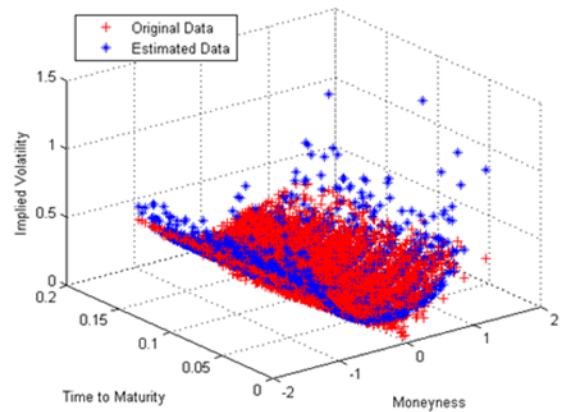
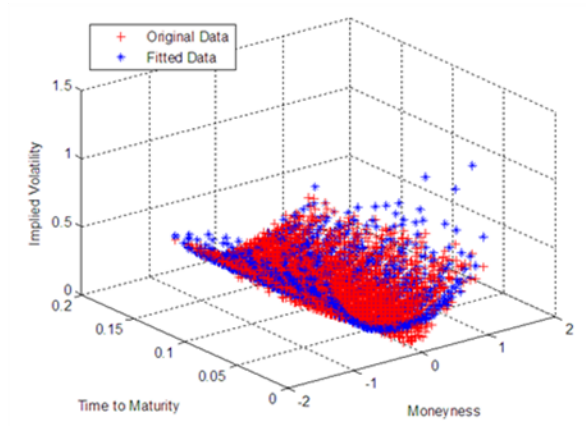


Figure 9 Showing Model represented by equation (5) which takes in account steepness of moneyness and maturity.

In Sample Fit



Out of Sample Prediction

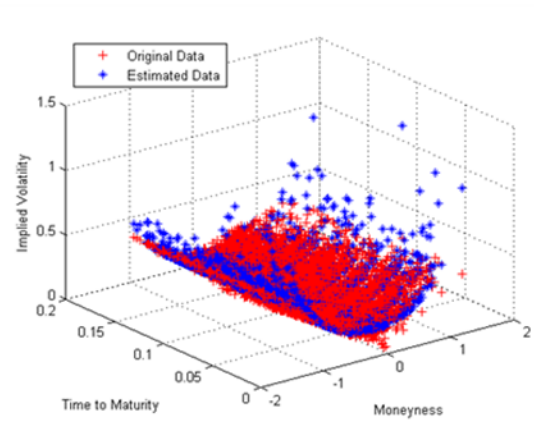


Figure 10 Showing model represented by equation (6) incorporating curvature of maturity

	β_0	β_1	β_2	β_3	β_4	β_5	RMSE (In Sample) (Fitting)	RMSE (Out of Sample) Prediction
Model I	-1.4876						0.3033	0.3362
Model II	-1.6352	0.2702	0.8836				0.1805	0.2001
Model III	-1.6244	0.2504	0.8779	-0.1208	0.2565		0.1802	0.1999
Model IV	-1.6108	0.2538	0.8783	-0.5613	0.2202	2.5269	0.1801	0.1998

From the root mean square (RMS) analysis we find that model 4 & 3 are better than model 2, however, the difference is not appreciable.

References

- [1] J Hull, A White, "The pricing of options on assets with stochastic volatilities", *Journal of Finance*, 1987
- [2] E. Derman, I.Kani "Stochastic Implied Trees: Arbitrage Pricing with Stochastic Term and Strike Structure of Volatility", *International Journal of Theoretical and Applied Finance*, 1998
- [3] B. Dupire, " Pricing with a smile", *RISK*, 1994
- [4] B. Dumas, J. Fleming, R.E.Whaley, "Implied Volatility Functions: Empirical Tests", *Journal of Finance*, 1998
- [5] R Cont, J. Fonseca, V Durrleman "Stochastic models of implied volatility surfaces, *Economic Notes*, 2002.
- [6] R. Lee, "Implied Volatility: Statics, Dynamics, and Probabilistic Interpretation" , *Recent Advances in Applied Probability*, Springer (2004)
- [7] M.L.Roux, "A long term model of the dynamics of the S&P 500 Implied Volatility Suface", working paper ING institutional markets.
- [8]R Cont, J da Fonseca, "Dynamics of implied volatility surfaces", *Quantitative Finance*, 2002
- [9] G. Skiadopoulos, S.Hodges, L.Clelow, " The Dynamics of the S&P 500 Implied Volatility Surface", *Review of Derivative Research*, 1999