Option Pricing Model with Stochastic Implied Volatility

Tianxiang Liu

School of Economics and Management, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing 100049, China, liugoanywhere@163.com

Desheng Wu

Risklab, University of Toronto, 40 St. George St. BA7262, Toronto ON M5S2E4, Canada, dash@risklab.ca; School of Economics and Management, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing 100049, China, dwu@ucas.ac.cn

Although option pricing theory has made great progress over the past thirty years, there are still large gaps between theory and practice. Practitioners still use the Black-Sholes-Merton model and implied volatility surfaces to manage their positions. We develop a new method for pricing options by describing the dynamics of Black-Scholes-Merton implied volatility rather than instantaneous volatility. Our model nests models from previous studies in this field, and generates volatility surfaces guaranteed to be either “smile” or “smirk”. We evaluate our model using data from S&P 500 index options and discuss parameter selection. Our model outperforms not only previous implied volatility models, but also traditional option pricing theory models (i.e. Variance Gamma model, Hull-White model and Heston model).

Keywords: option; implied volatility surface; stochastic implied volatility model; no-arbitrage condition

JEL Classification: G12

1. Introduction

Research on option pricing has made great strides over the last fifty years. The Black-Scholes-Merton model [1, 2] led to a boom in options trading which is easy to understand, calculate and explain to others. Furthermore, industry financial information providers, such as Bloomberg and Wind, only provide the BSM model in their software. Indeed, most market practitioners still use the traditional BSM model whereas traders apply implied volatilities calculated from the BSM formula to quote options. If the assumptions underlying the BSM model hold for an underlying asset, its volatility surface will be flat and unchanging. However, the volatility surfaces are not only curved but also change continuously and stochastically for most assets. Developing a model integrates how to trade options price and manage their positions would be a great benefit.

Previous studies have directly modelled implied volatility dynamics which have not been satisfied [3, 4, 5]. Unlike stochastic volatility models, implied volatility cannot be specified freely because the specification of implied volatility incorporates into the BSM model. The no-arbitrage condition should be satisfied by the specification of implied volatility. Previous models have been constructed using the no-arbitrage condition; however, these models failed to discuss the shape of generated volatility surfaces [6, 7].

In this paper, we aim to build a framework modelling implied volatility surfaces more practical to generate meaningful shape and broader to accommodate to most situations in industry. We make a more general specification, and our general framework nests these two models. We narrow the specification afterwards to meet the condition that guarantees the implied volatility skew has only one minimum. This condition is logical and certify the U-shape of the volatility skew when the market volatility skew shows a one-way down “smirk” pattern because we can assume that skew is still a U-shape, but the minimum is far right out of the sample. In addition to this, we can only see the left part of the U-shape. To price pathwise options, we propose two ways to use Monte Carlo simulations in our model.

In the empirical analysis of 2003-2016 S&P 500 samples, our model explains over 95% of the variation of implied volatility and the frequency of average pricing error is quite low. We also discuss parameter selection and performance of our model, we have six parameters in U-shape condition. We test some simplified models with fewer parameters by using Akaike Information Criterion to parameter selection, and find that the model with all six parameters performs best without overfitting problem. Even if the number of parameters is reduced from six to three, the results still show a relatively good performance. Moreover, we apply some individual equity to double check that our model can be implemented for different underlying assets. The model still yields robust results when used for individual stocks. We compare our model with other option pricing theories, including Variance Gamma model, Heston model and Hull-White model. As mentioned above, our model can explain over 90% of the variations, whereas the Variance Gamma model only interprets about 70% of the variations and two stochastic volatility models explain less.

1. Literature Review

The BSM model has been improved over the past decades by two types of models: The first kind uses Lévy processes instead of diffusion processes to describe asset movement. There are three approaches to building Lévy processes: specifying the Lévy measure of jumps, time subordination, and indicating the probability distribution of fixed increments directly.

Two popular examples of specifying the Lévy measure of jumps are the Merton [8] jump-diffusion model with Gaussian jumps, and the Kou and Wang [9] model with double exponential jumps. In these models, the dynamic structure of process is easy to understand since the size of the jumps is explicit. These models can easily use Monte Carlo simulations to price options and exotic options but hardly lead to closed-form probability densities. Another example of specifying the Lévy measure of jumps is the stable tempered process, which provides directly jumps structure and shows the probability distribution of the process via the Lévy-Khinchin formula.

Time subordination means to obtain a Lévy process by the subordinating time variable of the diffusion process with an independent increasing Lévy process. Two well-known examples include Madan et al. [10] with Gamma subordinator and Rydberg [11] with the inverse Gaussian subordinator. The primary function of distribution can be easily obtained, but the explicit expression of the Lévy measure of the process is not always available.

One example of indicating the probability distribution of fixed increments is directly the generalized hyperbolic processes. In this approach, it is simple to estimate the parameters and simulate processes with the same time increments. In general, the Lévy measure is unknown, and we do not know the law of the increments at other time scales.

In these models, the process of the specification from the Itô process to the Lévy process is directly expanded, whereas, in the second type of model, the process of specification of the Itô process is unchanged. Instead, these models randomly distribute the volatility of the Itô process. In these models, the volatility follows a random process and is usually named instantaneous volatility. There are many specifications of instantaneous volatility such as the Heston [12] model and the 3/2 model [13, 14].

These models have partly overcome the limitations of Black-Scholes. Researchers have also continued to extend this research by pricing not only options but also other option-related derivatives. Take corridor variance swap as an example. To price a corridor variance swap, under classical exponential Lévy process setting, the most popular way is numerical Fast Fourier Transform (FFT) approach [15]. Theoretically, Liu and Wu [16] have priced corridor variance swaps (CVS) under Black-Scholes assumption. Analytical results of corridor variance swaps have also been discovered under some stochastic volatility settings [17].

In the existing literature, some researchers have attempted to model the implied volatility surfaces, but they do not achieve satisfactory results. Zhu and Avellaneda [3], Schönbucher [4] and Fengler [5] attempted to describe the volatility surfaces without giving volatility dynamics. Their work is more like data fitting rather than derivative pricing. Daglish et al. [6] and Carr and Wu [7] manage to build their models on implied volatility dynamics and the no-arbitrage condition. Unlike stochastic volatility models, the implied volatility cannot be specified freely because the specification of implied volatility must incorporate with the BSM model. The no-arbitrage condition must be satisfied by the specification of implied volatility. Their studies, however, have some limitations. For example, their models are not broad enough to cover all the possibilities of volatility surface shape, and they do not guarantee that the volatility skew generated by their models is U-shaped to match which is seen in the financial market.

1. Methodology

We consider a market with a world with one risky asset and one riskless bond, and there is no arbitrage between the risky asset and the bond. Thus, risk-neutral probability measure defined in a probability space . We ignore the interest rates and carrying costs for the risky asset. When dealing with a deterministic term structure of financing rates in practice, we can model the forward value of the underlying asset and define moneyness of the option against the forward price.

The BSM model assumes that the underlying asset price follows the equation:

where is a constant that denotes the variance of the risky asset.

The BSM formula shows the price of European option with strike K and time to maturity as

where N() is cumulative distribution function of normal distribution.

Due to the put-call-parity, the price of European put option is

If the assumption of the BSM model holds, the implied volatility

would be flat with respect to strike and time to maturity and invariant concerning time. However, implied volatilities of most assets are not flat but stochastically change. Practitioners use the implied volatilities surface to quote options and manage their positions.

In models with jumps, researchers have attempted to improve the BSM model by adding a Lévy measurement:

 denotes the real line except zero, is the stock price just prior to any jump at time t, and the Poisson random measure on counts the jumps of price by at time . The process is an intensity measure, which reveals the expected jump count for a given x at time t, so that the last term in the equation is the increment of a -pure jump martingale. Researchers can specify directly, or use time subordination, or specify the probability distribution of fixed increments.

Another way to improve the BSM model is a stochastic volatility model. We assume volatility is not a constant but a stochastic process:

 (1)

 is named as instantaneous variance, and stochastic volatility models specify: in Heston [12] or, in Hull and White [18] or, in Scott [19].

In our model, we specify the implied volatility rather than instantaneous volatility in Eq (1). This equation corresponds more strongly with how practitioners manage their options than stochastic volatility models because we model what practitioners quote directly. We offer a more explicit view of option price than models with jumps and stochastic volatility models.

Most traditional models based on options prices will suffer from model calibration problems: option prices across different moneyness and maturity differ by several orders of magnitude; the solutions to inverse problems are difficult to calculate and need not necessarily exist, in which case only local minima can be found. Our model can be easily calibrated by OLS across different maturities and moneyness because implied volatilities have the same order of magnitude. Also, this model does not suffer from the ill-posed inverse problem like in Heston model and many other stochastic volatility models, because this model has simple closed-form solution and gradients are easy to find. Later we show that our models have higher pricing power of volatility surface than traditional models.

We assume the implied volatility for fixed strike K and time to maturity by the following equation.

 (2)

where are Wiener processes driving the volatility surface. Without the loss of generality, we assume that they are uncorrelated. The and are allowed to be correlated:

 (3)

where is the stochastic process taking values in an interval from -1 to 1.

 is the drift of implied volatility and is the volatility of volatility (volvol) process. Both processes can be either deterministic or stochastic, and can depend on implied volatility. Given : the underlying asset price , strike and time to maturity .

**Proposition 1. Under the stock assumption (1), implied volatility assumption (2) and correlation specification (3), the absence of dynamic arbitrage on an option results in the following constraint:**

 **(4)**

Proof: See in the Appendix

Proposition 1 is the general constraint in our model. Eq(4) involves partial derivatives of the BSM formula, including theta for , vega for , dollar gamma for and dollar vanna for , and volga for . This is a generalization of Vega-Gamma-Vanna-Volga Model [7]. We can transform Eq(4) into implied volatility form.

**Proposition 2. The constraint (4) can be transformed into a non-dynamic arbitrage constraint on implied volatility surface . The depends on relative moneyness , time to maturity , instantaneous variance , the drift of volatility , volatility of volatility , and correlation .**

 **(5)**

Proof: See in the Appendix

Here we have our general implied volatility model. Different specifications for the drift ,volvol processes and the correlations result in different shapes of volatility surfaces. For practical purposes, we reduce the number of Wiener processes in (2) to one, and make a particular specification for and in our model:

 (6)

where and are constant, , , and are stochastic processes that do not depend on K, and .

The describes the average drift of volatility and describes the asymmetry of drift to relative moneyness k. The asymmetry structure of drift can be applied in the empirical observation that variance risk premium of options with different strikes are not flat. The is the time-decay coefficient which assumes that long-maturity implied volatility tends to move less.

The specification does not lose generality, and nests models proposed in previous studies. If , and , the model becomes Carr and Wu model [7]. If , and , the model becomes Daglish et al. model [6]. If , , and , this model becomes Heston [12]’s model directly applied to implied volatility.

**Proposition 3. Under the stock assumption (1), implied volatility assumption (6) and correlation specification (3), the absence of dynamic arbitrage on an option derives the following constraint:**

 **(7)**

Proof: See in the Appendix

Eq (7) is an equation of implied volatility , and implied volatility can be decided by Eq(7). In the limit of , we derive:

The implied volatility of at-the-money option can be solved using the following equation:

The necessary condition for the minimum of implied volatility skew at fixed is:

 (8)

Now we try to narrow down the assumption of and to make sure that the model has a clear interpretable solution and the solution is U-shape “volatility smile”. We know from the Abel–Ruffini theorem that there is no algebraic solution to general polynomial equations of degree five or higher with arbitrary coefficients. There are , , , , , and a constant in Eq(7). It is easy for us to pick up and carefully to ensure Eq(7) is quartic equation. Furthermore, to have a clear interpretable solution of Eq(7), we need to make sure that Eq(7) is a biquadratic equation; therefore, we assign the following condition to narrow our model assumption:

We also want to ensure that our parameterization guarantees only one minimum in implied volatility skew with a U-shape. Even if the market volatility skew shows a one-way down “smirk” pattern, it is possible to assume the skew is still U-shaped. We can imagine that only the left part of the U-shape is visible because the minimum is to the far right out of sample. From condition (8), if , there is only one minimum in the volatility skew. However, later we will show condition can also guarantee the U-shape of the volatility skew.

**Proposition 4. Under the stock assumption (1), implied volatility assumption (6), correlation specification (3) and , the absence of dynamic arbitrage on an option derives the following constraint:**

 **(9)**

**At a fixed time to maturity, Eq(9) can be reorganized as the following expression:**

 **(10)**

**with**

 **(11)**

Proof: See in the Appendix

Gatheral [20] proposed an empirical expression of Eq(10) as the stochastic volatility-inspired model but he did not propose any stochastic processes behind this form. Carr and Wu [7] attempted to construct a model to generate this form but the result was failed. Our model in this paper has successfully built the linkage between stochastic processes and the empirical form (10).

Please note that . These conditions are useful in the proof of Proposition 5 below.

**Proposition 5. Eq(10) with Eq(11) has the asymptote.**

 **(12)**

**when k tends to and asymptote**

 **(13)**

**when k tends to .**

**The volatility skew in Eq(10) with Eq(11)has only one minimum when**

 **(14)**

Proof: See in the Appendix

Proposition 5 guarantees the volatility skew generated by the model is U-shaped because the volatility skew has an upslope asymptote to positive infinity and a downslope negative asymptote to negative infinity and the volatility skew has only one minimum.

The minimum also has economic significance because, in practice, delta is usually estimated in the BSM model. But in our frame, the actual delta will be:

 (15)

 is the delta of the BSM model. is the vega of the BSM model, and is always positive whether the option is a call or a put. Because , is always negative. Therefore, if k is left to the minimum point, BSM delta will underestimate the true delta, and if k is right to the minimum point, BSM delta will overestimate the true delta.



**Figure 1 An example of implied volatility surface under the no-arbitrage condition.**

Figure 1 shows a numerical example of Eq(10) and Eq(11) with . The volatility skew, which is the cross-section of the surface, is U-shaped and asymmetric, and the variation of the skew is less volatile as the time to maturity increases.

One advantage of our model over stochastic volatility models is we can easily evaluate the options and variance swaps by putting implied volatility of our model directly into the BSM-based models. When stochastic volatility is set, pricing of options and variance swaps still requires additional work. Heston and Nandi [21] found analytical results of options on the Heston [12] setting. Zhu and Lian [22] presented a highly efficient approach to price variance swaps with discrete sampling times. Zheng and Kwok [17] provided a general analytic approach for pricing discretely sampled generalized variance swaps under stochastic volatility with simultaneous jumps model. Detemple and Kitapbayev [23] calculated the pricing of American options under the generalized 3/2 and 1/2 classes of volatility processes. However, under our model, the calculation is quite simple. To get the pricing, we simply put the implied volatility into the BSM assumption-based formula of options [1, 2], variance swaps [24] and corridor variance swaps [16].

 Our model can also be extended to model-free theories based on stochastic volatility. Here, we incorporate our model into Glasserman and Wu [25] to calculate forward implied volatility. Glasserman and Wu define and calculate forward volatility implied by option prices when the underlying asset is driven by a stochastic volatility process.

In their work, the model-free version of forward rate is a measure of market-implied forward volatility defined as:

 (16)

Where is the implied volatility at t with the maturity of T and strike of K, which can be derived easily from our model. Therefore, we incorporate the concept of model-free forward rate into our model, and other model-free concepts can also be translated into our model.

1. Estimation Method and Parameter Selection

In our model, the shape of implied volatility surface is only impacted by the current state of parameters, or economic states, (, , , , , ), but do not depend on the stochastic dynamics of the parameters. As a result, we can regard parameters as time series without knowing their dynamics specification.

We estimate parameters (, , , , , ) in our specified model (10) and (11) by quasi-maximum likelihood estimation time by time. We assume that the estimation errors of our model at fixed time-t follow an independent normal distribution.

The maximum likelihood estimator of is the residual sum of squares divided by sample size at time-t.

 (17)

The log-likelihood function at time-t:

 (18)

Maximum likelihood estimators can be estimated by minimizing the sum of squared forecasting errors. Generally, we have six parameters in our full model, but we can reduce degrees of freedom down to three states . Moreover, choosing parameters by what parameters mean and how they impact volatility surface, we use Akaike information criterion (AIC) [26] to select parameters.

 (19)

The AIC is a measure of the relative quality of statistical models for a given set of data, founded on information theory, defined as in which is the number of parameters. At time-t estimation, the AIC is after ignoring the constant and it is reasonable to use average AIC () to compare different parameter selections because and are constant across time-t.

We also considered the state-space method mentioned by Carr and Wu [7]. They treat parameters as hidden states and implied volatility as measurements with the error. To keep all covariates positive, parameters are transformed into the values on the whole real line:

Unlike quasi-MLE, the evolvement of the parameters must be specified as state equations. They assume they propagate as random walks:

 (20)

They define the measurement equations on the logarithm of implied volatility:

 (21)

where is a vector of implied volatility for different relative moneyness and time to maturity at time-t. H(.) is a function from parameters to implied volatility by Eq(10) and Eq(11).

 This setup has non-linearity on H(.), so the traditional Gaussian linear Kalman filter [27] is hard to estimate. They use unscented Kalman filter [28] to handle the nonlinearity. Parameter estimators are updated time by time continuously.

The result of the state-space model is quite unsatisfactory because it fails to explain the variation of implied volatility surface. The reason is that update speed cannot catch up with parameters. Because the state-space model provides entirely unsatisfactory performance, we ignore this method in our empirical analysis.

We do not use Carr and Wu’s state-space model because they use the self-defined updating speed and we do not think it is reliable. They use auxiliary parameters to control the updating speed instead of the standard setting of unscented Kalman filter. They choose these auxiliary parameters by minimizing the RSS of their results. We think this set is problematic because it may break the assumptions of unscented Kalman filter and make quasi-MLE results twisted.

1. Monte Carlo Simulation

In order to price pathwise options, i.e., American options, there are many pieces of literature which may incorporate into our model to price. Options in future studies, including empirical research such as Jensen and Pedersen [29], analytical research such as Chockalingam and Muthuraman [30], or numerical research such as like Haugh and Kogan [31]. Therefore, we choose Monte Carlo simulation. We propose two ways to use Monte Carlo simulations in our model. First, we adjust Broadie and Kaya [32]'s method to our model then convert assumptions (1), (2), (3) and into the following form:

where and are two independent Brownian motion processes. Further, we discretize them by Euler Discretization. The stock price at time t, given the values of and for , can be written as:

 (22)

and the implied volatility at time t is given by:

 (23)

Knowing and the initial state of , we can implement Euler discretization of (16) to samples from the distribution of . The conditional Monte Carlo method [33] is an alternative simulation approach that improves the efficiency of the simulation estimators. We use European call options to demonstrate the method.

Assume that is the average instantaneous volatility of the underlying asset over the time horizon:

Willard (2009) observes that the call option can be re-written as:

 where

Thus, the call option can be re-written as conditional on a path of the Brownian motion ,

1. Empirical Results

We use CBOE’s end-of-day listed option market data from 2003 to 2016 to test our model. The end-of-day mid quotes are used in all the related research. Data without valid bid price in a day are removed from the sample.

We construct implied volatility on a grid of five fixed relative strikes from 80 to 120, and three fixed times to maturity, from one month to three months. To avoid weekday effects, we sample the data weekly every Wednesday. During the construction, we used only out-of-the-money options because out-of-the-money options are more actively traded and more sensitive to implied volatility. We use the average of put and call implied volatilities to present the at-the-money implied volatility.

We do not use any smoothing technique on our dataset. There are smoothing methods, such as Nadaraya-Watson smoothing [34, 35] or least squares kernel smoothing[36], applied in previous studies, but we do not apply them to volatility surfaces because we do not want the evaluated pricing performance to be affected by smoothing.

**Table 1 The Average Implied Volatility of Samples across Different Moneyness and Time to Maturity**

|  |  |
| --- | --- |
| Moneyness | Time to Maturity |
|  | 1 month | 2 months | 3 months | 3m-1m |
| 80 | 0.361149 | 0.308269 | 0.293038 | -0.06811 |
| 90 | 0.260827 | 0.243275 | 0.238524 | -0.0223 |
| 100 | 0.173324 | 0.178842 | 0.182146 | 0.008823 |
| 110 | 0.164808 | 0.14774 | 0.147119 | -0.01769 |
| 120 | 0.219449 | 0.17365 | 0.159361 | -0.06009 |
| 80-120 | 0.1417 | 0.134619 | 0.133676 |  |
| 90-110 | 0.096018 | 0.095536 | 0.091405 |  |
| max-min | 0.19634 | 0.16053 | 0.145918 |  |

In all maturities, implied volatility was U-shaped (“smile”) at one month, two months, and three months (Table 1). This pattern in our listed option data is different from the “smirk” pattern of the over-the-counter data reported in Foresi and Wu [37]. We observe 80-120 moneyness implied volatility differences, 90-110 moneyness implied volatility differences; furthermore, the implied volatility difference between maximum and minimum are all reduced from short maturity to long maturity. The above result shows that the implied volatility of options with longer maturity tends to move less than implied volatility of options. We also verify this by the implied volatility difference between three months and one month, which shows a shallow reverse U-shape. Therefore, the reverse U-shape of difference is offset by the one-month U-shape of implied volatility, flattening the U-shape of the three-month maturity options.



**Figure 2 The Plot of At-the-money Volatility, Volatility Term Structure and Volatility Skew.**

Figure 2 shows at-the-moneyness implied volatility, implied volatility term structure and implied volatility skew. At-the-moneyness implied volatility shows spikes corresponding to the 2008 Global Financial Crisis and the 2012 European sovereign debt crisis (Figure 2A). The three months-one month volatility difference is usually positive but can become negative when a crisis occurs and short and short-term volatility climbs much faster than long-term volatility (Figure 2B). The 90-110 volatility skew is constantly positive through our whole sample across all three maturities (Figure 2C). The figure suggests the option implied the distribution of S&P 500 is heavy-tailed to the left compared to normal distribution.

**Table 2 The Explanation of Variation and Average Pricing Error of Models with Different Parameter Selection.**

|  |  |  |
| --- | --- | --- |
| Explained Variation | Average Pricing Error | AIC |
| Moneyness | Time to Maturity | Moneyness | Time to Maturity |  |
| Model 1: 　 |
| 　 | 1m | 2m | 3m | AVG　　 | 　 | 1m | 2m | 3m | AVG　　 | RSS |
| 80 | 0.986 | 0.988 | 0.982 | 0.985 | 80 | 0.824 | 1.083 | -0.432 | 0.492 | 0.003 |
| 90 | 0.994 | 0.993 | 0.988 | 0.991 | 90 | -0.256 | -0.719 | -1.557 | -0.844 | k |
| 100 | 0.991 | 0.994 | 0.995 | 0.994 | 100 | 0.387 | -0.330 | -0.786 | -0.243 | 6.000 |
| 110 | 0.952 | 0.984 | 0.985 | 0.974 | 110 | -0.084 | 1.090 | 1.021 | 0.676 | AIC |
| 120 | 0.956 | 0.974 | 0.968 | 0.966 | 120 | -2.009 | 0.427 | 0.857 | -0.242 | -77.6 |
| AVG　 | 0.976 | 0.987 | 0.984 | 0.982 | AVG　　 | -0.227 | 0.311 | -0.179 | -0.032 | 　 |
| Model 2:  |
| 　 | 1m | 2m | 3m | AVG　　 | 　 | 1m | 2m | 3m | AVG　　 | RSS |
| 80 | 0.984 | 0.981 | 0.958 | 0.974 | 80 | 0.801 | 1.256 | -0.207 | 0.617 | 0.003 |
| 90 | 0.988 | 0.994 | 0.988 | 0.990 | 90 | -0.531 | -0.892 | -1.678 | -1.034 | k |
| 100 | 0.991 | 0.994 | 0.994 | 0.993 | 100 | 0.374 | -0.342 | -0.790 | -0.253 | 5.000 |
| 110 | 0.962 | 0.977 | 0.978 | 0.972 | 110 | 0.251 | 1.301 | 1.169 | 0.907 | AIC |
| 120 | 0.936 | 0.904 | 0.939 | 0.926 | 120 | -2.086 | 0.244 | 0.727 | -0.372 | -75.4 |
| AVG　　 | 0.972 | 0.970 | 0.971 | 0.971 | AVG　　 | -0.238 | 0.313 | -0.156 | -0.027 | 　 |
| Model 3:  |
| 　 | 1m | 2m | 3m | AVG　 | 　 | 1m | 2m | 3m | AVG　　 | RSS |
| 80 | 0.980 | 0.985 | 0.968 | 0.978 | 80 | 0.281 | 1.095 | 0.079 | 0.485 | 0.004 |
| 90 | 0.988 | 0.992 | 0.985 | 0.988 | 90 | 0.034 | -0.746 | -1.640 | -0.784 | k |
| 100 | 0.987 | 0.992 | 0.986 | 0.988 | 100 | 1.166 | -0.491 | -1.371 | -0.232 | 5.000 |
| 110 | 0.956 | 0.983 | 0.982 | 0.974 | 110 | 0.255 | 1.012 | 0.645 | 0.637 | AIC |
| 120 | 0.925 | 0.884 | 0.943 | 0.918 | 120 | -2.280 | 0.439 | 1.133 | -0.236 | -74.3 |
| AVG　　 | 0.967 | 0.967 | 0.973 | 0.969 | AVG　　 | -0.109 | 0.262 | -0.231 | -0.026 | 　 |
| Model 4:  |
| 　 | 1m | 2m | 3m | AVG　　 | 　 | 1m | 2m | 3m | AVG　　 | RSS |
| 80 | 0.982 | 0.984 | 0.952 | 0.973 | 80 | 0.830 | 1.238 | -0.135 | 0.644 | 0.004 |
| 90 | 0.986 | 0.994 | 0.983 | 0.988 | 90 | -0.577 | -0.924 | -1.633 | -1.045 | k |
| 100 | 0.979 | 0.994 | 0.989 | 0.987 | 100 | 0.291 | -0.380 | -0.761 | -0.284 | 4.000 |
| 110 | 0.948 | 0.983 | 0.977 | 0.969 | 110 | 0.164 | 1.281 | 1.234 | 0.893 | AIC |
| 120 | 0.931 | 0.855 | 0.920 | 0.902 | 120 | -2.144 | 0.263 | 0.868 | -0.338 | -75.4 |
| AVG　　 | 0.965 | 0.962 | 0.964 | 0.964 | AVG　　 | -0.287 | 0.295 | -0.085 | -0.026 | 　 |
| Model 5:  |
| 　 | 1m | 2m | 3m | AVG　　 | 　 | 1m | 2m | 3m | AVG　　 | RSS |
| 80 | 0.942 | 0.973 | 0.915 | 0.943 | 80 | -1.506 | 1.912 | 2.017 | 0.807 | 0.007 |
| 90 | 0.981 | 0.993 | 0.981 | 0.985 | 90 | -1.579 | -0.680 | -0.798 | -1.019 | k |
| 100 | 0.981 | 0.994 | 0.984 | 0.987 | 100 | 0.254 | -0.428 | -0.813 | -0.329 | 3.000 |
| 110 | 0.937 | 0.981 | 0.969 | 0.963 | 110 | -0.191 | 1.353 | 1.368 | 0.844 | AIC |
| 120 | 0.877 | 0.842 | 0.901 | 0.874 | 120 | -3.239 | 0.524 | 1.637 | -0.360 | -69.5 |
| AVG　　 | 0.944 | 0.957 | 0.950 | 0.950 | AVG　　 | -1.252 | 0.536 | 0.682 | -0.011 | 　 |

Note: In Table 2, the left panel reports the explained variation, defined as defined as one minus the variance ratio of the model’s pricing error to the original volatility series. This measure is analogous to the R-squared measure for regression and reveals the proportion of variation explained by the model. The right panel reports the average pricing error, defined as the average difference between the observed volatility and the model-generated volatility. The first row shows the performance of our full model with six parameters.

The average pricing error of our full model is -0.032. On average, the model explains 98.2% of volatility variations and is stable across all moneyness and maturities. The model has relatively weak explanatory power in high moneyness. This deficiency may be because high moneyness volatility is more volatile than other parts of volatility skew, i.e., sometimes upslope, sometimes downslope, because the skew in the listed options market is usually U-shaped, but skew could be a smirk in a crisis. The changeable shape of volatility skew in the listed options market makes the model hard to fit the high moneyness volatility.

The second model ignores the asymmetry of volatility drift. The third model ignores time decay. The fourth model ignores both the volatility drift and asymmetry of volatility drift. The fifth model ignores volatility drift, the symmetry of volatility drift and time decay, and it contains only the instantaneous volatility, volatility of volatility and the correlation, and represents the minimal set of parameter selection. We find the explained variation is reduced from 98.2% to 95.0% as the model is simplified to 3-economic-state. The absolute value of average pricing error is even reduced, but the variation of average pricing error across different moneyness and maturity is increased as the number of parameters is reduced to 3. According to AIC, the best model is the full model, which has the least AIC (-77.6). Compared with simplified models, the optional parameters increase the accuracy of the model without overfitting problems. In practice, we always recommend the full model, as the AIC suggests, unless the simplicity of the model is emphasized more than the accuracy.



**Figure 3 The Plot of Parameters of the Six-Parameter Model.**

Figure 3 shows the plots of each parameter in our full model. The instantaneous volatility of the full model peaks during in the 2007-2008 global financial crisis and the 2010-2011 European debt crisis, as shown by the at-the-money implied volatility (Figure 3A). The peak of instantaneous volatility during the global financial crisis is higher than that of the implied volatility, whereas the two peaks of instantaneous volatility during European financial crisis are much smoother than implied volatility. The volatility of volatility shows high variation, from 0.5 to 2.5, during our sample period (Figure 3B). The volvol swings quickly before 2013, but its variation slows after 2013. The correlation of stock movement and instantaneous variance stays strongly negative during the whole sample period, which is consistent with previous studies, such as GARCH(Figure 3C). The drift of instantaneous volatility shows mean-reverting features (Figure 3D). The drift is around zero during normal times but becomes very negative during the crisis, when instantaneous volatility reaches its peaks and starts to fall back. Panel E shows the asymmetry of drift. Before 2007 and after 2013, the drift of higher moneyness drifts more positively than lower moneyness. Between 2007 and 2013, the drift sometimes falls to a very negative position and shows a different picture compared with its performance in normal times. The time decay factor fails to reveal much information about the model, which shows a random walk, likely because our sample contains only up to three-month maturities (Figure 3F). But we confirm that the time decay factor is constantly positive, which means long-maturity skew changes less than the short-maturity skew.



**Figure 4 The Illustration of Volatility Smile and Smirk under our Model.**

We assess the robustness of our model by evaluating the performance of our model (1) during the crisis and (2) for individual stock options.

We discuss the minimum point of one-month volatility skew and find the moneyness when the minimum is reached. Due to the discretization of moneyness in our sample, our sample minimum moneyness is also discrete. All one-month volatility minima are at money=100 or 110 or 120. When the minimum is at money=120, the skew will be a “smirk"; otherwise, the skew is a “smile”. Normally, the minimal theoretical moneyness follows the minimal factual moneyness well. During financial crises, however, the skew turns into the “smirk” pattern, and the minimal theoretical moneyness goes beyond the sample moneyness ceiling 120. That shows that our model fits well when skew becomes the “smirk” pattern because the model treats “smile” pattern skew as the left part of the U-shape, and the minimal moneyness is right out of sample, as is shown in Figure 4.



**Figure 5 The k of Minimum Point in Theory and in Fact.**

Normally, the k of minimum point follows the simple k of the minimum point well (Figure 5). During the global financial crisis and the European debt crisis, sample k reaches the coiling 120, and the model treats the k higher than sample coiling. In 2008, our model evaluates the k of the minimum point above 200 moneyness, which successfully deals with the steep downslope of volatility skew in 2008. No matter which shape the skew is, smile or smirk, our model works well with the data.

**Table 3 The Explanation of Variation and Average Pricing Error of IBM Options.**

|  |  |  |
| --- | --- | --- |
| Explained Variation | Average Pricing Error | AIC |
| Moneyness | Time to Maturity | Moneyness | Time to Maturity |  |
| Full Model: 　 |
| 　 | 1m | 2m | 3m | AVG　　 | 　 | 1m | 2m | 3m | AVG　　 | RSS |
| 80 | 0.9957 | 0.9888 | 0.9907 | 0.992 | 80 | 0.5539 | 0.5396 | -0.401 | 0.2308 | 0.001 |
| 90 | 0.9932 | 0.9933 | 0.9898 | 0.992 | 90 | -0.202 | -0.347 | -0.745 | -0.432 | k |
| 100 | 0.9919 | 0.9955 | 0.9947 | 0.994 | 100 | 0.0338 | -0.225 | -0.436 | -0.209 | 6.000 |
| 110 | 0.9858 | 0.9884 | 0.9936 | 0.989 | 110 | 0.6512 | 0.7256 | 0.342 | 0.5729 | AIC |
| 120 | 0.9606 | 0.9874 | 0.9909 | 0.980 | 120 | -1.595 | 0.594 | 0.4595 | -0.18 | -85.6 |
| AVG　 | 0.985 | 0.991 | 0.992 | 0.989 | AVG　　 | -0.112 | 0.2574 | -0.156 | -0.003 | 　 |

Our study also evaluates the performance of our model in individual stock options. We select IBM as our target for our robustness check. IBM is a well-known, widely-traded listed corporation; also, IBM options have relatively large trading volumes and long data history. Individual stock options at the CBOE are American options. But in this study, we ignore American option premiums because the premium is usually not large in short-term maturity.

Table 3 shows the pricing performance of the full model with six parameters of our model in IBM options. We see explained the variation of 0.989 and average pricing error of -0.0032 percentage points, even better than the result of the index options. We can also see stable performance across moneyness and time to maturity. Here the robustness of our model across different types of options is shown.

We attempt to compare our model with other mainstream option pricing theories. We choose Variance Gamma model as the benchmark of jump-diffusion models, Heston model and Hull-White model as the benchmark of stochastic volatility model. In Variance Gamma model and Hull-White model, we use the methods provided by the original paper to estimate vanilla options. In Heston model, we apply a constraint on the parameters to prevent the instantaneous volatility reaching zero [38].

We use ordinary least squares to estimate implied parameters, and no model calibration techniques, such as regularization or nonlinear least squares, are applied due to fair competition. We use the same data sample and estimate the implied parameters in three benchmark models every day. If the benchmark model can successfully describe the skewness and kurtosis of the return distribution, it can also describe the implied volatility surface. For convenience, we compare the explained variations of out-of-the-money option price between our model and benchmark models. The result is shown in Table 4.

**Table 4 The Explanation of Price Variations Comparison between our Model and other Models**

|  |  |  |  |
| --- | --- | --- | --- |
| Explained Variation |   |   |   |
| Our Model |   |   |   |   |
|   | 1m | 2m | 3m | AVG　　 |
| 80Put | 0.992 | 0.992 | 0.992 | 0.992 |
| 90Put | 0.994 | 0.991 | 0.975 | 0.987 |
| 100Put | 0.963 | 0.957 | 0.943 | 0.955 |
| 100Call | 0.961 | 0.940 | 0.940 | 0.947 |
| 110Call | 0.997 | 0.994 | 0.994 | 0.995 |
| 120Call | 1.000 | 0.998 | 0.998 | 0.999 |
| AVG　　 | 0.985 | 0.979 | 0.974 | 0.979 |
| Variance Gamma |   |   |   |
|   | 1m | 2m | 3m | AVG　　 |
| 80Put | 0.480 | 0.604 | 0.649 | 0.578 |
| 90Put | 0.644 | 0.662 | 0.670 | 0.659 |
| 100Put | 0.915 | 0.902 | 0.888 | 0.902 |
| 100Call | 0.834 | 0.867 | 0.857 | 0.853 |
| 110Call | 0.922 | 0.743 | 0.676 | 0.780 |
| 120Call | 0.971 | 0.683 | 0.765 | 0.806 |
| AVG　　 | 0.794 | 0.743 | 0.751 | 0.763 |
| Hull-White |   |   |   |   |
|   | 1m | 2m | 3m | AVG　　 |
| 80Put | 0.258 | 0.473 | 0.549 | 0.427 |
| 90Put | 0.598 | 0.757 | 0.776 | 0.710 |
| 100Put | 0.892 | 0.940 | 0.947 | 0.926 |
| 100Call | 0.706 | 0.717 | 0.773 | 0.732 |
| 110Call | 0.511 | 0.750 | 0.907 | 0.723 |
| 120Call | 0.000 | 0.000 | 0.956 | 0.319 |
| AVG　　 | 0.494 | 0.606 | 0.818 | 0.639 |
| Heston |   |   |   |   |
|   | 1m | 2m | 3m | AVG　　 |
| 80Put | 0.313 | 0.687 | 0.614 | 0.538 |
| 90Put | 0.196 | 0.651 | 0.622 | 0.490 |
| 100Put | 0.469 | 0.626 | 0.599 | 0.565 |
| 100Call | 0.277 | 0.398 | 0.506 | 0.394 |
| 110Call | 0.000 | 0.424 | 0.443 | 0.289 |
| 120Call | 0.000 | 0.822 | 0.756 | 0.526 |
| AVG　　 | 0.209 | 0.601 | 0.590 | 0.467 |

We can see our model provides higher and stable explained variation of price across all three maturities and all strikes (Table 4). Our model can explain over 90% of the variations, whereas the Variance Gamma model only explains about 70% of the variations and two stochastic volatility models explains even less. The Variance Gamma model and Hull-White model tend to have less explanatory power of short-maturity options and out-of-the-money options because these options have tiny values and put less weights in least squares. Our model is much more stable than these two models because we model implied volatility directly. The inverse problem of Heston model is ill-posed, and computation of a global minimum of OLS is difficult. So the Heston model is very unstable across different maturities and moneyness. Therefore, our model shows a better empirical option pricing performance than the benchmark models.

1. Conclusions

We develop a new model to price options by stochastic implied volatility that integrates how traders quote options in the industry. The general setting of our model is broad, and nests models from previous studies. We derive the no-arbitrage condition for the model and show the properties of volatility surfaces generated by the model. To keep only one minimum in the volatility, we narrow the model to a six-parameter-model and discuss the shape of volatility surface in this setting. We examine the model by using 2003-2016 S&P 500 index options data, and our model shows great performance with an average pricing error of -0.032 and explained volatility variation of 98.2%. Our model is stable across all moneyness and maturities. We discuss parameter selection and show that the model with all six parameters is the best model in AIC. This study shows that our model can accommodate both “smile” skew and “smirk” skew. Furthermore, our model outperforms the traditional Variance Gamma model, Hull-White model, and Heston model by providing higher explained variations of out-of-the-money option prices empirically. To price pathwise options, we propose two ways to use Monte Carlo simulation in our model.

Our model gives a more intuitive approach for practitioners than either Lévy process models or stochastic volatility models, given the heavy reliance on BMS implied volatility surfaces in industry. In our model, the options and option-based derivatives are easier to price than other affine models without the BMS-based models being replaced. It can calculate the implied volatility from our model, and get the result by substituting the implied volatility with the BSM-based models.

To conclude, our model outperforms previous models in three ways. Firstly, we specify the stochastic dynamics of implied volatility which makes our work different from previous data-fitting models. Secondly, our model is more general and nests some past attempts. These make it possible for the model to be widely applied in reality. Finally, unlike previous studies, we guarantee that the surface is U-shaped which is reasonable to describe the volatility surfaces in practice.

1. Appendix

*Proof of Proposition 1.* We treat the option price as a composite function: . By Ito’s Lemma, the risk-neutral drift of a call option is:

Because we ignore the interest rate, to avoid dynamic arbitrage opportunity, the drift of the call option must be zero.

*Proof of Proposition 2.* From BSM formula, we can get the following relationship between partial derivatives:

We substitute these derivatives at into propostion1, and divide both sides of the equation by the common denominator .

*Proof of Proposition 3.* From assumption Eq(6), we know:

Substitute these into Proposition 2, and we get Proposition 3.

*Proof of Proposition 4.* Substitute into Proposition 3, we get Eq(9). By quadratic formula, we get the form of Eq(10) and Eq(11).

*Proof of Proposition 5.*We get the limits easily:

By oblique asymptote formula, we get the asymptotes. Because , the volatility skew has an upslope asymptote to positive infinity and a downslope negative asymptote to negative infinity.

Also, we get the partial derivative:

To get the extreme value, let derivative be zero and rearrange:

Please note, in order to get real solution, need to be positive. Because , need to be less than . Square both side of equation and use quadratic formula, we get two solutions:

But because only one solution is less than , which is:

Due to the slope of two asymptotes, we know this extreme value is a minimum.

1. Reference

[1] F. Black and M. Scholes, "The pricing of options and corporate liabilities," The Journal of Political Economy, 1973, pp. 637-654.

[2] R. C. Merton, "Theory of Rational Option Pricing," Bell Journal of Economics & Management Science, vol. 4, no. 1, 1973, pp. 141-183.

[3] Y. Zhu and M. Avellaneda, "A risk-neutral stochastic volatility model," International Journal of Theoretical and Applied Finance, vol. 01, no. 02, 1998, pp. 289-310.

[4] P. J. Schönbucher, "A market model for stochastic implied volatility," Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, vol. 357, no. 1758, 1999, pp. 2071-2092.

[5] M. R. Fengler, "Semiparametric modeling of implied volatility," Springer Science & Business Media, 2006.

[6] T. Daglish, J. Hull and W. Suo, "Volatility surfaces: Theory, rules of thumb, and empirical evidence," Quantitative Finance, vol. 7, no. 5, 2007, pp. 507-524.

[7] P. Carr and L. Wu, "Analyzing volatility risk and risk premium in option contracts: A new theory," Journal of Financial Economics, vol. 120, no. 1, 2016, pp. 1-20.

[8] R. C. Merton, "Option pricing when underlying stock returns are discontinuous," Journal of Financial Economics, vol. 3, no. 1-2, 1976, pp. 125-144.

[9] S. G. Kou and H. Wang, "Option pricing under a double exponential jump diffusion model," Management Science, vol. 50, no. 9, 2004, pp. 1178-1192.

[10] D. B. Madan, P. P. Carr and E. C. Chang, "The variance gamma process and option pricing," European finance review, vol. 2, no. 1, 1998, pp. 79-105.

[11] T. H. Rydberg, "The normal inverse Gaussian Lévy process: simulation and approximation," Communications in Statistics Part C Stochastic Models, vol. 13, no. 4, 1997, pp. 887-910.

[12] S. L. Heston, "A closed-form solution for options with stochastic volatility with applications to bond and currency options," Review of Financial Studies, vol. 6, no. 2, 1993, pp. 327-343.

[13] S. L. Heston, "A simple new formula for options with stochastic volatility," SSRN Electronic Journal, 1997.

[14] E. Platen, "A non-linear stochastic volatility model," Research report: financial mathematics research report/Centre for mathematics and its applications (Canberra), 1997.

[15] P. Carr and D. Madan, "Option valuation using the fast Fourier transform," Journal of computational finance, vol. 2, no. 4, 1999, pp. 61-73.

[16] T. Liu and D. Wu, "New Approach to Estimating VIX Truncation Error using Corridor Variance Swaps," Journal of Derivatives, 2018(Forthcoming).

[17] W. Zheng and Y. K. Kwok, "Closed form pricing formulas for discretely sampled generalized variance swaps," Mathematical Finance, vol. 24, no. 4, 2014, pp. 855-881.

[18] J. Hull and A. White, "The pricing of options on assets with stochastic volatilities," The Journal of Finance, vol. 42, no. 2, 1987, pp. 281-300.

[19] L. O. Scott, "Option pricing when the variance changes randomly: Theory, estimation, and an application," Journal of Financial and Quantitative analysis, vol. 22, no. 4, 1987, pp. 419-438.

[20] J. Gatheral, "The volatility surface: a practitioner's guide," John Wiley & Sons, 2011.

[21] S. L. Heston and S. Nandi, "Derivatives on volatility: Some simple solutions based on observables," SSRN Electronic Journal, 2000.

[22] S. P. Zhu and G. H. Lian, "A closed‐form exact solution for pricing variance swaps with stochastic volatility," Mathematical Finance, vol. 21, no. 2, 2011, pp. 233-256.

[23] J. Detemple and Y. Kitapbayev, "On American VIX options under the generalized 3/2 and 1/2 models," Mathematical Finance, 2017.

[24] P. Carr and D. Madan, "Optimal positioning in derivative securities," Quantitative Finance, vol. 1, no. 1, 2001, pp. 19-37.

[25] P. Glasserman and Q. Wu, "Forward and future implied volatility," International Journal of Theoretical & Applied Finance, vol. 14, no. 03, 2011, pp. 407-432.

[26] H. Akaike, "A new look at the statistical model identification," Automatic Control IEEE Transactions on, vol. 19, no. 6, 1974, pp. 716-723.

[27] R. E. Kalman, "A new approach to linear filtering and prediction problems," Journal of basic Engineering, vol. 82, no. 1, 1960, pp. 35-45.

[28] D. Simon, "Optimal state estimation: Kalman, H infinity, and nonlinear approaches," John Wiley & Sons, 2006.

[29] M. V. Jensen and L. H. Pedersen, "Early option exercise: Never say never," Journal of Financial Economics, 2016.

[30] A. Chockalingam and K. Muthuraman, "American options under stochastic volatility," Operations Research, vol. 59, no. 4, 2011, pp. 793-809.

[31] M. B. Haugh and L. Kogan, "Pricing American options: A duality approach," Operations Research, vol. 52, no. 2, 2001, pp. 258-270.

[32] M. Broadie and Ö. Kaya, "Exact simulation of stochastic volatility and other affine jump diffusion processes," Operations Research, vol. 54, no. 2, 2006, pp. 217-231.

[33] G. A. Willard, "Calculating prices and sensitivities for path-independent derivative securities in multifactor models," Journal of Derivatives, vol. 5, no. 1, 2009, pp. 45-61.

[34] E. A. Nadaraya, "On estimating regression," Theory of Probability and Its Applications, vol. 9, no. 1, 1964, pp. 157-159.

[35] G. S. Watson, "Smooth regression analysis," Sankhyā: The Indian Journal of Statistics, Series A (1961-2002), vol. 26, no. 4, 1964, pp. 359-372.

[36] M. R. Fengler and Q. Wang, "Fitting the smile revisited: A least squares kernel estimator for the implied volatility surface," Social Science Electronic Publishing, 2003.

[37] S. Foresi and L. Wu, "Crash–O–Phobia: A domestic fear or a worldwide concern?," The Journal of Derivatives, vol. 13, no. 2, 2005, pp. 8-21.

[38] S. Mikhailov and U. Nögel, "Heston’s stochastic volatility model: Implementation, calibration and some extensions," John Wiley and Sons, 2004.