Prediction Error of Call Options Pricing in Chinese Market

Yanzhe Wang

November 2023

Abstract

Key Words: Black Scholes Model, CS Model, China A share, Call Options, Prediction Error

This paper focuses on the European option pricing problem on Chinese A-share stock index options. We mainly compare the difference in pricing power between the traditional Black Scholes model and the CS model, which introduces the implied skewness and kurtosis to the Black Scholes model, on two stock index options, CSI300 and SSE50. The time period of the study is mainly from 2018 to 2020, which includes one year of bull market, one year of bear market and one year of market under epidemic in China's stock market. We also compare side-by-side the differences in the predictive ability of the two models in different environments by grouping them according to different dimensions such as time and moneyness.

Analysis of the data suggests that Chinese stock index options exhibit relatively stable and low levels of volatility in the early years relative to more mature Western markets. Volatility within the first year of the epidemic was particularly sharp. There is also a negative skewness and a kurtosis much lower than 3. In particular, we find that Chinese-specific holiday customs allow implied volatility performance to reflect significant variability around the Chinese New Year.

The predicted price scatter plot and the fitted density functional image show us more visually the difference in predictive ability between the two models. It can be seen that the BS model always tends to overestimate OTM options and underestimate ITM options, while the CS model consistently gives relatively more accurate forecasts. Both models are not as good at predicting short-term options as long-term options, due to the fact that short-term options are relatively more volatile and introduce more uncertainty into the predictions. Also, we can get the same conclusion from the fitted density functional image.

Finally, we numerically visualize the difference in the predictive ability of the two models by means of a table of the prediction errors. MAE and MSE errors exhibit similar conclusions, and the conclusions regarding the categorization of moneyness are similar to those previously. At the same time, both models have significantly larger short-term errors compared to long-term errors and show negative means.

1 Introduction

In the European-style option pricing theory, the Black Scholes model assumes that volatility is constant, so that every option's plot of volatility against the strike price is a horizontal line. However, much research in the 21st century has revealed that implied volatility of different strikes or moneyness tended to exhibit like a 'smile', indicating that higher implied volatility is usually associated with low strike prices and lower implied volatility is associated with high strike prices, so that Black Scholes overprices in-the-money calls and out-of-the money puts; it also underprices out-of-the money calls and in-the-money puts.

Moreover, many of these studies also point out that the normality assumptions on stock returns in the Black Scholes model mostly do not hold in real world. In particular, stock returns tend to exhibit negative skewness and leptokurtosis (i.e., "fat tails."). In fact, it has been well documented that the skewness and kurtosis of stock returns largely affect the shape of implied volatility. To investigate the effects of skewness and kurtosis in the pricing of European options, Corrado and Su gave an improved version of the Black Scholes model in 1996 (called the CS model), which incorporates non-zero skewness and non-mesokurtosis into the framework of the Black Scholes model.

The Chinese options market started in 2015 and is still in a phase of rapid development. As an emerging market with a huge volume, its internal characteristics are vastly different from many well-established Western markets, which also makes the findings of this paper highly relevant. The author draws on Yanrui Han's research ideas when studying option pricing errors during the financial crisis in the Australian market in 2018, and make some improvements in the research methodology. This paper aims to investigate the difference in the ability of the Black Scholes model and the modified CS model to predict European option prices in the Chinese options market, and to examine the effect of implied skewness and implied kurtosis on the shape of the implied volatility image of the CS model. In particular, we use data from the Chinese market in 2018 (bear market), 2019 (bull market) and 2020 (Covid) to compare and analyze the impact of different market trends on the model's forecasting ability and implied volatility shape.

In the empirical test, we first use the rolling window of the underlying asset returns over the past 30 days to estimate the historical higher-order moments of the call option, which are then used as the initial values in a gradient descent procedure for minimizing the second-order variance of observed prices and prices estimated by the Black Scholes model and CS model, so that we can get the implied variance, skewness and kurtosis in both models for each day. After that, we use those estimated moments and the linear approximation method to get the call options'prices for the Black Scholes model and CS model. Lastly, we calculate the pricing errors for both models to reach our conclusions.

2 Theory for Models

2.1 Symbols

In this section, all symbols used in the equations and their meanings are shown in the following table:

Table 1: Notations

Symbol	Description
C	Call option price
S_0	Underlying asset price
S_t	Underlying asset price at time t
K	Strike price
r	Continuously compounded risk-free interest rate
t	Time to maturity in years
σ	Volatility of the underlying stock price
μ_n	nth moment of a given random variable or sample set
M	Moneyness of a contract

2.2 Models

2.2.1 Black Scholes Model

When it comes to problems related to the pricing of European-style options, the discovery of the Black Scholes model (hereinafter referred to as BS model) is certainly both groundbreaking and practical. As is known to all, BS model is based on three fundamental hypotheses:

- The volatility of the underlying stock is constant over time
- The instantaneous log return of the stock price is an infinitesimal random walk with a constant drift; more precisely, the stock price follows a geometric Brownian motion
- The underlying stock does not pay a dividend
- The market environment is frictionless with no arbitrage opportunity
- The risk-free interest is constant and borrowing and lending any amount of cash at the risk-free rate is allowed

With those hypotheses and related knowledge of probability theory and stochastic analysis, the price of an European-style call option can be calculated by the following formula with five inputs: underlying stock price, strike price, continuously compounded risk-free interest rate, time to maturity in years and volatility of the underlying stock price.

$$C_{BS} = S_0 N(d_1) - Ke^{-rt} N(d_2) \tag{1}$$

where

$$N(\cdot) = \text{Standard normal cumulative distribution function}$$

$$d_1 = \frac{ln(\frac{S_0}{K}) + (r + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}} \tag{2}$$

$$d_2 = d_1 - \sigma\sqrt{t} \tag{3}$$

Based on the BS formula above, we can quickly obtain the theoretical price of European-style options in the market. Among the five parameters required by the BS formula, the first four can be intuitively calculated from the market, but the volatility of underlying stock prices is difficult to obtain directly. Our approach to estimating implied volatility is described in detail in the subsequent sections of this chapter.

Since the introduction of the BS model, many researchers have done empirical tests of its cumbersome fundamental hypotheses in the market. It turns out that in real life, the first two assumptions are significantly rejected in the vast majority of cases, which means the volatility of stock prices is constantly changing over time, and the assumption of normality in the random walk hypothesis is hardly valid. Specifically, most of the realistic stock returns will show a skewness away from 0 and a kurtosis away from 3, which is against the assumption of normality.

2.2.2 CS Model

In order to better approximate the distribution of stock returns in real life, many scholars have continuously tried to explore ways to improve the BS model and to quantify the effects of skewness and kurtosis on the pricing model. Jarrow and Rudd's idea of introducing Edgeworth expansions to lognormal probability density functions was a very constructive breakthrough in this area. They obtained a new probability density function for the normal distribution by expanding the third and fourth moments of stock prices as parameters, which is shown as follows:

$$f_{JR}(x) = n(x)[1 + \frac{\mu_3}{3!}(x^3 - 3x) + \frac{\mu_4 - 3}{4!}(x^4 - 6x^2 + 3) + \frac{\mu_3^3}{72}(x^6 - 15x^4 + 45x^2 - 15)] \tag{4}$$

where

$$n(\cdot) = \text{Standard normal density function}$$

Based on their idea, Corrado and Su refined this improved version of the BS model by improving modeling of stock returns, more specifically, using a Gram-Charlier series expansion of a normal probability density function. One more Hermite polynomial is used in the Gram-Charlier series expansion to improve Jarrow and Rudd's model, which results in the following density function:

$$f_{GC}(x) = \sum_{n=0}^{\infty} c_n H_n(x)\Phi(x) \tag{5}$$

Where c_n is determined by the moments of the improved density function, $H_n(\cdot)$ is a Hermite polynomial generated from consecutively higher derivatives of $\Phi(\cdot)$, and $\Phi(\cdot)$ is density function of the standard normal distribution.

Then Corrado and Su intercepted the terms associated with the first four orders of moments in the above equation, and obtained an approximate density function for the non-normal skewness and kurtosis:

$$f_{CS}(x) = n(x)[1 + \frac{\mu_3}{3!}(x^3 - 3x) + \frac{\mu_4 - 3}{4!}(x^4 - 6x^2 + 3)] \tag{6}$$

Corrado and Su used the approximate density function obtained above to calculate the marginal impact of non-normal skewness and kurtosis. Combining the calculation of the pricing process of the BS model and the minor improvements made by Brown and Robinson in the follow-up, we eventually obtained the following European call option pricing formula for the CS model used in the subsequent empirical tests:

$$C_{CS} = C_{BS} + \mu_3 Q_3 + (\mu_4 - 3)Q_4 \tag{7}$$

where

$$C_{BS} = S_0 N(d_1) - Ke^{-rt} N(d_2) \tag{8}$$

$$Q_3 = \frac{1}{3!} S_0 \sigma \sqrt{t}[(2\sigma\sqrt{t} - d_1)n(d_1) + \sigma^2 t N(d_1)] \tag{9}$$

$$Q_4 = \frac{1}{4!} S_0 \sigma \sqrt{t}[(d_1^2 - 3\sigma\sqrt{t} d_1 + 3\sigma^2 t - 1)n(d_1) + \sigma^3 t^{\frac{3}{2}} N(d_1)] \tag{10}$$

$$d_1 = \frac{ln(\frac{S_0}{K}) + (r + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}} \tag{11}$$

$$d_2 = d_1 - \sigma\sqrt{t} \tag{12}$$

When $\mu_3 = 0$ and $\mu_4 = 3$, the CS model is degenerated into the BS model.

For more detail about CS Model, please look up the paper "*Implied volatility skews and stock return skewness and kurtosis implied by stock option prices*" by Corrado and Su.

3 Research Methodology Details

3.1 Parameter Estimation Methods

In the process of estimating prices using the BS model, all inputs including underlying asset price S_0, strike price K, continuously compounded risk-free interest rate r and time to maturity in years t can be found from the market easily, so all we need to estimate is the implied volatility σ. In our study, we assume that option contracts on the same asset have the same implied volatility throughout the day. Therefore, we take as input the above parameters and observed prices for all similar option contracts in a day and minimize the following second-order error Q_{BS} to obtain the implied volatility of this underlying asset for that day:

$$Q_{BS} = \frac{1}{n}\sum_{i=1}^{n}(C_{OBS,i} - C_{BS,i})^2 \tag{13}$$

After that, we use the implied volatility obtained from the previous day's estimation as a parameter when calculating the estimated price of the BS model.

In CS model, similarly, we need to estimate three parameters which we cannot directly observe from the market, the implied volatility and the third-order and fourth-order moments. We minimize the error function of these three variables as independent variables as follows:

$$Q_{CS} = \frac{1}{n}\sum_{i=1}^{n}(C_{OBS,i} - C_{CS,i})^2 \tag{14}$$

We use the implied volatility and moments obtained from the previous day's estimation as parameters when calculating the estimated price of the CS model.

3.2 Pricing Predictions

In order to investigate the difference in the predictive ability of different models and the impact of the introduction of higher-order moments, we do a length one translation of the estimated data between trading days. Specifically, taking the same approach as in Han's work (2018), we use the implied volatility and higher-order moments estimated on the previous trading day as model parameters for each trading day's option contract to obtain the estimated value of the model for the current contract.

After what, we compute for each contract the mean absolute error (MAE) and the mean squared error (MSE) between its theoretical estimated price and its observed price, which are:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|C_{OBS} - C_{Predict}| \tag{15}$$

and

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(C_{OBS} - C_{Predict})^2 \tag{16}$$

Finally, we combine the obtained error results with the previous categorization of the market conditions and the moneyness and duration of the contract to obtain the comparative results.

3.3 Moneyness

Existing studies tell us that the impact of moneyness differences on option pricing problems is usually significant. Therefore, we also divide the moneyness properties into the following categories and focus on the impact of kurtosis and skewness on model pricing errors under different categories in a cross-sectional comparison.

Following Han (2018) and Tripathi and Gupta (2011), we define $\log \frac{S}{K}$ as moneyness M, thus categorizing the different option contracts. Specifically, we first filter out contracts with an absolute value of M greater than 0.15. For the remaining contracts, we categorize those with M between 0.05 and 0.15 as OTM(Out of The Money), those with M between -0.15 and -0.05 as ITM(In The Money), and those with M not exceeding 0.05 in absolute value approximated as ATM(At The Money).

4 A Brief Introduction to Chinese Stock Options

In order to better contextualize the market environment in the subsequent study of the empirical results, we here provide a succinct overview of the market.

Unlike the rapid development and high trading volume of commodity options in China, the development of stock index options and stock options has been relatively slow. It was not until Feburary 2015 that the first stock or stock index option, the SSE 50 Index option, was available. In the first few years, the trading volume of SSE50 was far less than that of commodity options, basically not in the same order of magnitude. The second stock index option, the CSI 300 stock index option, was listed and traded on the SSE in December 2019 after almost five years. At that time, options in China are all European options. They are usually listed in batches and have a maximum maturity of no more than four months. After that, trading of stock index options entered a phase of more rapid development, with the third stock index option, the CSI 1000 option, being launched in July 2022. All three stock index options now rank roughly between 15 and 30 in terms of options turnover in the Chinese market.

5 Data

In this paper, CSI 300 Index and SSE 50 Index (CSI and SSE are used hereafter to refer to these two indexes) information is obtained from Wind database. We are interested in the data for a total of three years between 2018 and 2020, as they can represent three different time periods: bear market, bull market, and the covid epidemic, respectively. Since the second option contract for the Chinese market, the CSI option, was not available until late 2019, only SSE options are available for the study in 2018 and 2019, while both SSE and CSI are used in 2020. The data for the different option contracts include the previous closing price, the closing price, the strike price, the remaining duration and the closing price of the underlying asset. Besides that, the yield on the 10-year Chinese government bond was downloaded as the risk-free rate.

After obtaining the above data, we performed a filtering process. This study covers a total of two options contracts, a total of 1298 contracts and 67,085 sets of daily frequency data. First, we filter out 16,417 sets of data whose moneyness is outside the required range. After that, 6,442 sets of data with a remaining duration shorter than 10 days were also discarded, because they are disproportionately affected by the difference between the strike price and the spot price, far beyond the nature of the option itself. Eventually, a total of 44,226 sets of daily frequency data were used in the study.

For the 44,226 sets of data that met the requirements, we categorized them according to the remaining duration. Due to the limited number of years of options development in China, which is still in its early stages, the vast majority of options contracts between 2018 and 2020 have a maximum remaining duration of less than 100 days. Therefore, in order to ensure the adequacy of the data for the study, we can only divide it into two types of options for long-term and short-term comparative studies. 27,038 sets of data with a remaining duration greater than 60 days are categorized as long-term, while 17,188 groups of contracts with a remaining duration less than or equal to 60 days are categorized as short-term.

6 Empirical Results

6.1 Implied Moments

Using the numerical approach described above, we first obtained the implied volatility of the two stocks over the three years from 2018 to 2020.

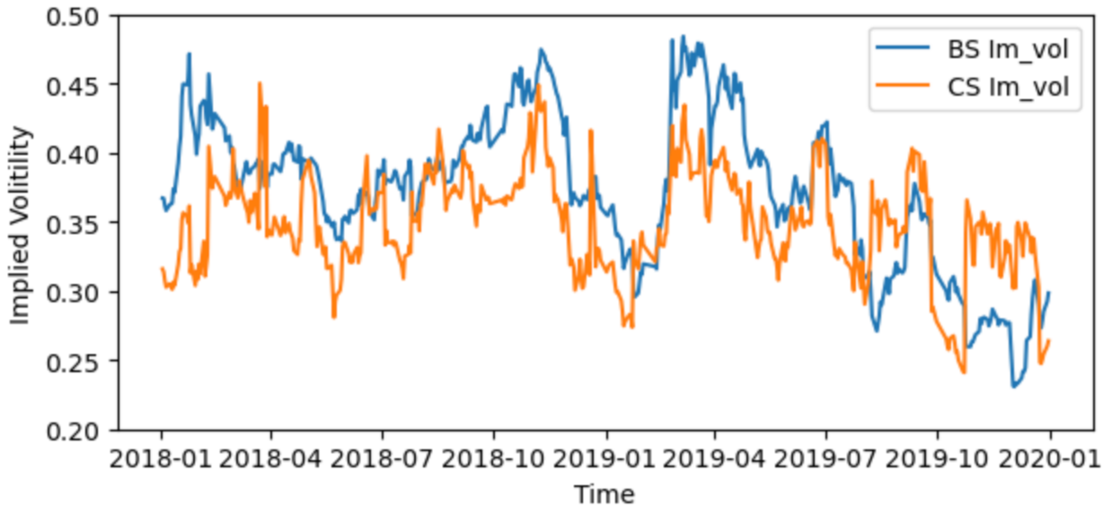

Figure 1: SSE50 Call Option Implied Volatility in 2018 and 2019

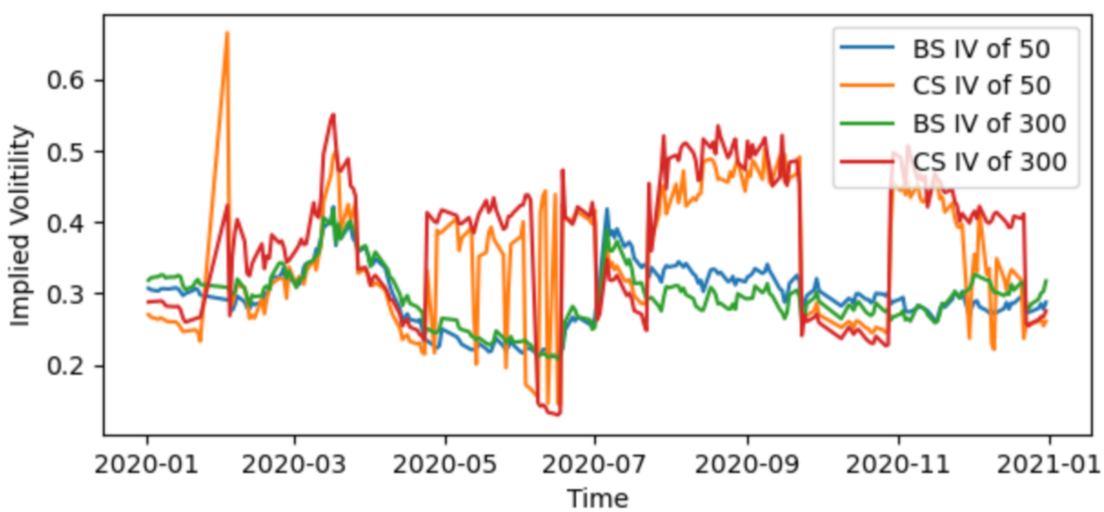

Figure 2: SSE50 and CSI300 Call Option Implied Volatility in 2020

The implied volatility of the SSE50 long-term call option (those with remaining duration of 60 days or more are considered as long-term contracts) over the two-year period from 2018 to 2019 is represented in the chart on the left, where the blue line is obtained from the BS model and the orange line is obtained from the CS model. As described above, there was only one SSE50 in the Chinese stock index options market by the end of 2019, while the line graph on the right representing results in 2020 has the implied volatility obtained from the BS model and CS model calculations for the CSI300 long-term call option, respectively, additionally added by the green and red lines.

Over the course of 2018 and 2019, the implied volatility of SSE50 call options has been relatively stable overall. Implied volatility ranged between 25% and 50% in both years and performed roughly the same across different market environments in both bull and bear markets. This calculation compares favorably with existing findings in the U.S. and Australian markets, where volatility has been more volatile.

On the other hand, the results we see during 2020 are quite different compared to previous years. The BS model is still in a relatively stable range and has low overall volatility. The implied volatility calculated by the CS model, however, reflects higher volatility. In particular, 2020 was a year in which China was hit by the full force of the epidemic. During that year, employees from all walks of life in many cities, including Shanghai and Shenzhen, had extended periods of home office or even building closure and quarantine. The liquidity of the already less mature Chinese options market was also hit hard. In this case, since new contracts in the Chinese options market at that time were issued in batches by month, each new contract entering the market would have a certain impact on the volatility of the overall market. As can be seen from the results in the above figure, the BS model is less affected by the introduction of new contracts, whereas the CS model is very sensitive to the entry and exit of new contracts under low liquidity, generating several sizable jumps.

Interestingly, the implied volatility of options is at a higher level around the Chinese New Year each year (around January to February) relative to the rest of the year. This is largely because the Chinese New Year has half a month off, and the traditional thinking of the Chinese people before the holidays is generally controlling the position at a lower level, thus minimizing the risk and getting peace of mind for the New Year. As a result market liquidity will also be generally low during this period, leading to higher implied volatility.

Meanwhile, we also obtained the implied skewness and kurtosis of long-term call options as follow:

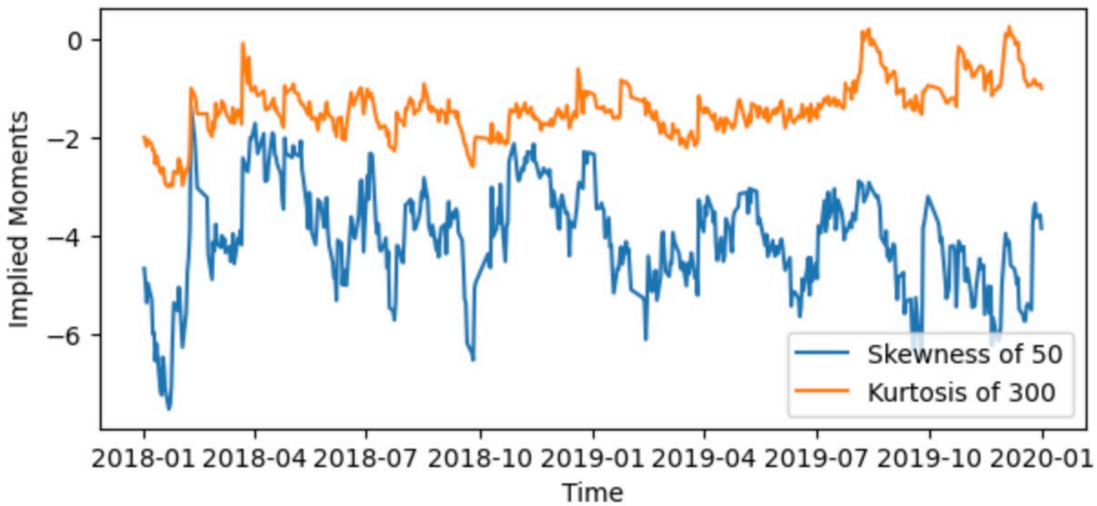

Figure 3: SSE50 Implied Skewness and Kurtosis in 2018 and 2019

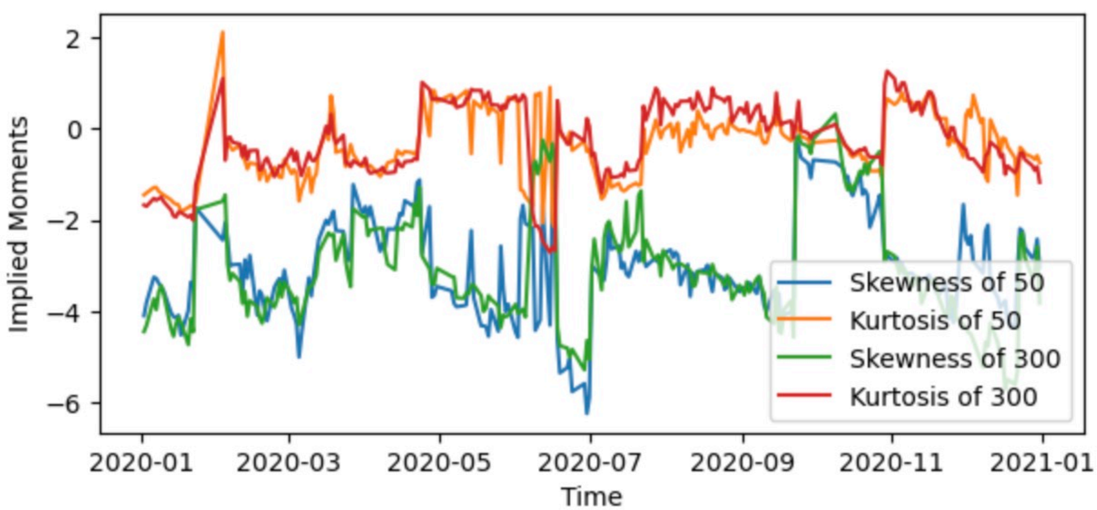

Figure 4: SSE50 and CSI300 Implied Skewness and Kurtosis in 2020

It can be seen that, as the underlying assets of these two call options, the stock indices in the Chinese market generally have a negative skewness and a kurtosis much lower than 3, representing a negative shift in the peak of the stock index returns and a fat tail in its distribution. This conclusion is in line with existing academic research perceptions of the overall environment of the developing stock market.

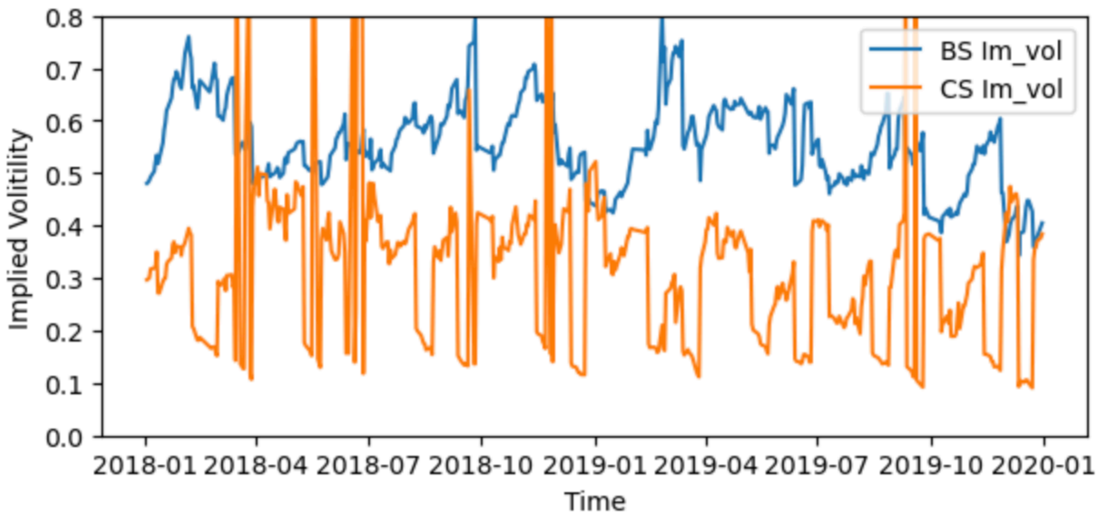

Figure 5: SSE50 short-term IV in 2018 and 2019

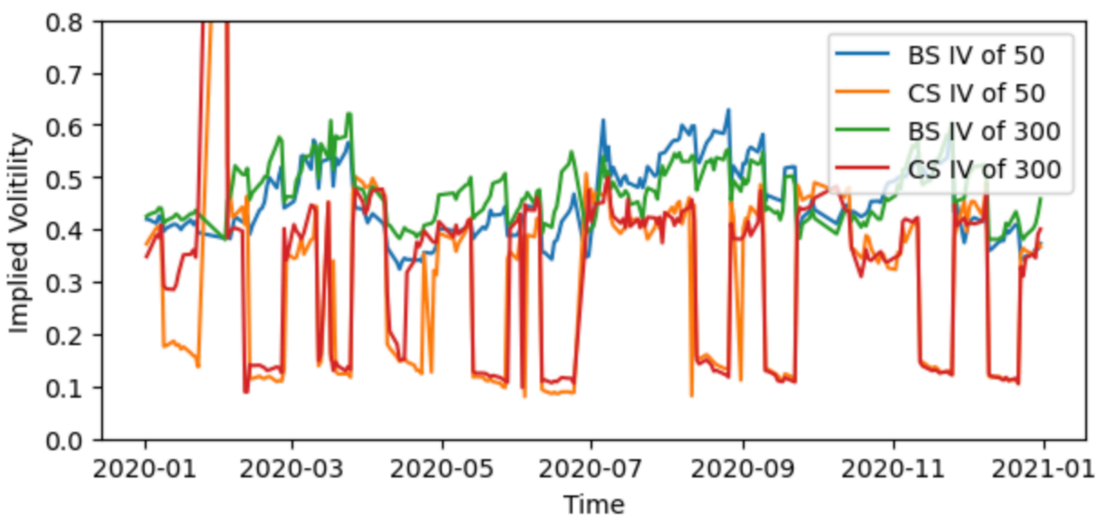

Figure 6: SSE50 and CSI300 short-term IV in 2020

On the other hand, the implied volatility of short-term contracts (with a remaining duration of more than 10 days and less than 60 days) is clearly significantly higher than that of long-term. Moreover, due to the large batch entry and exit of the Chinese options market and the generally shorter remaining duration, the implied volatility of short-term contracts is also subject to more jumps as a result of the entry and exit of the same batch of contracts.

6.2 Model Predicted Prices and Observed Prices

6.2.1 Scatter Plots

In order to more intuitively analyze and compare the pricing errors of different models under the influence of different factors, we pair the market price of each contract on each day with the predicted price of the model, and plot to obtain the scatterplot as follows:

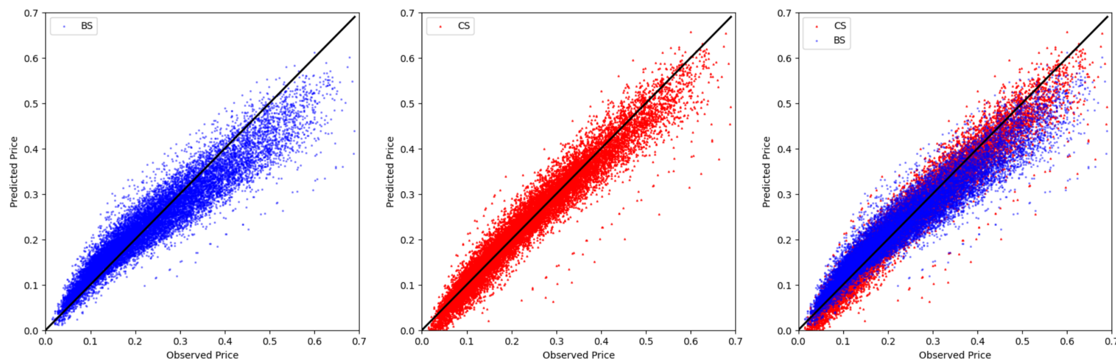

Figure 7: SSE50 long-term Pricing Pairs

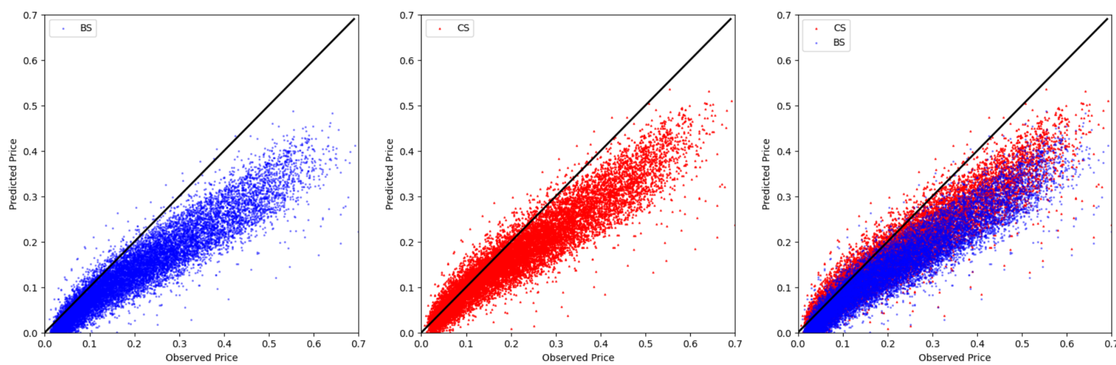

Figure 8: SSE50 short-term Pricing Pairs

In the two graphs above, from left to right, the predicted price vs. observed price for the BS model, the predicted price vs. observed price for the CS model, and the combination of the two are shown. The horizontal axis is the observed price and the vertical axis is the model's predicted price, while the blue dots represent the predicted values of the BS model and the red dots represent the predicted values of the CS model. Ideally a perfect model prediction should fall on the black diagonal line of y=x in the scatter plot.

It can be observed that the predictive power of the CS model is more significantly stronger than that of the BS model for Chinese stock index options. Specifically, for long-term contracts, the BS model tends to overestimate lower-valued call option contracts and underestimate higher-valued call option contracts. In contrast, the CS model gives more stable predicted prices for all values of option contracts. On the other hand, for short-term contracts, although both the BS and CS models tend to underestimate the value of call options, the predicted price given by the CS model is still relatively closer to the true observed value.

Similarly, the scatterplot made by CSI300 options has the same result:

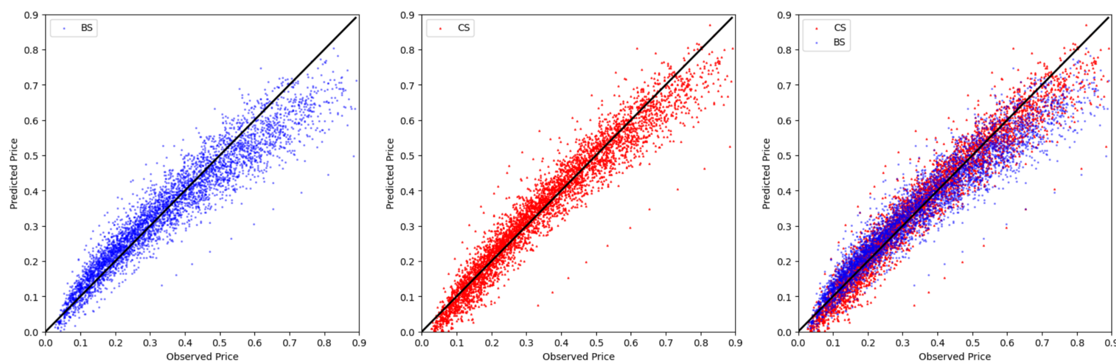

Figure 9: CSI300 long-term Pricing Pairs

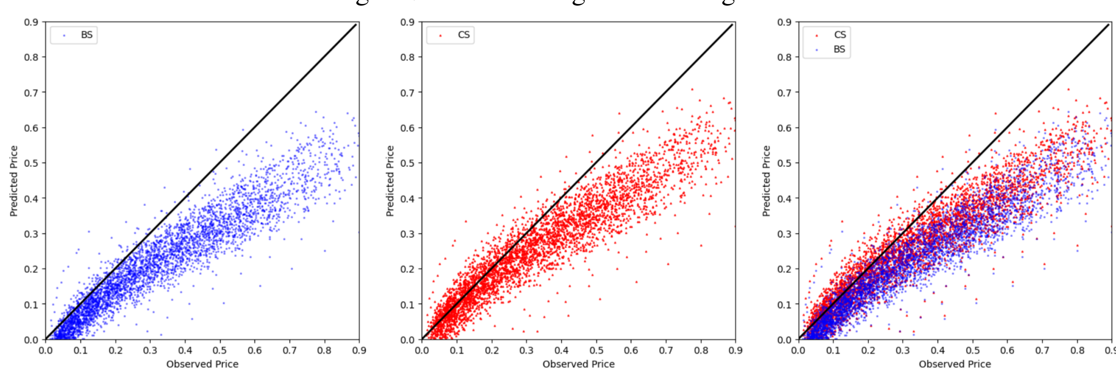

Figure 10: CSI300 short-term Pricing Pairs

This suggests that the CS model has an advantage in its ability to predict prices in different market environments. It is worth noting that this finding is partially contrary to the conclusions of Han's previous study of several stock index options in the Australian market, reflecting the differences in the model's ability to predict prices in different markets and at different times.

Meanwhile, in addition to the categorization in terms of remaining duration, we also plot different scatter plots based on the monetary categorization of the contracts as follows:

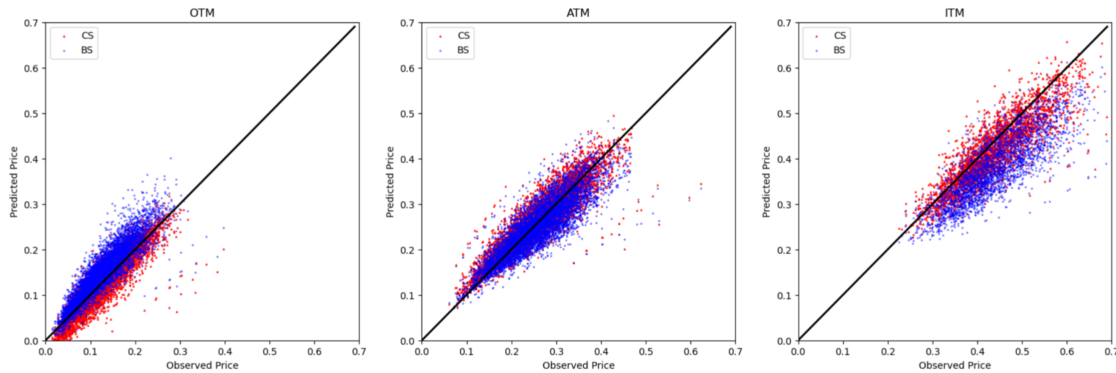

Figure 11: SSE50 Moneyness Pricing Pairs

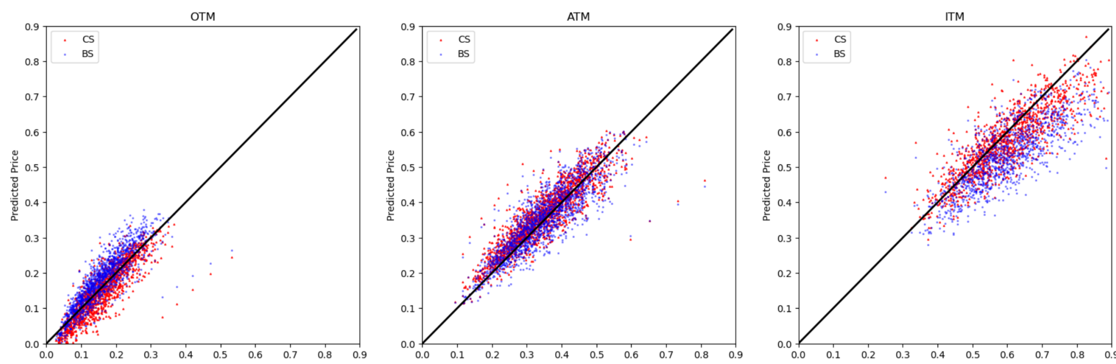

Figure 12: CSI300 Moneyness Pricing Pairs

Referring to what is depicted in Section 3.3, each of the two sets of graphs above shows, from left to right, the modeled predicted price versus the observed price for options in the OTM, ATM, and ITM classes, respectively.

It is not difficult to find that the BS model overestimates those OTM call option contracts, i.e., contracts whose strike prices are smaller than the price of their underlying asset, and underestimates those ITM call option contracts, compared to the CS model. For the ATM call option contracts, the BS model and the CS model perform similarly.

6.2.2 Probability Distribution Functions

In order to visualize the distributional characteristics of the predicted prices, we fit the predicted prices of the BS model and the CS model to the probability density functions as follow:

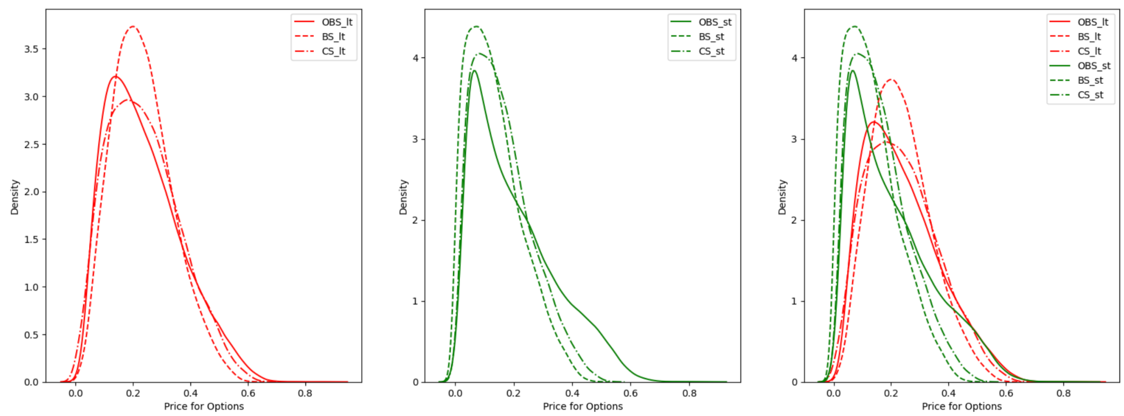

Figure 13: SSE50 Pricing PDFs

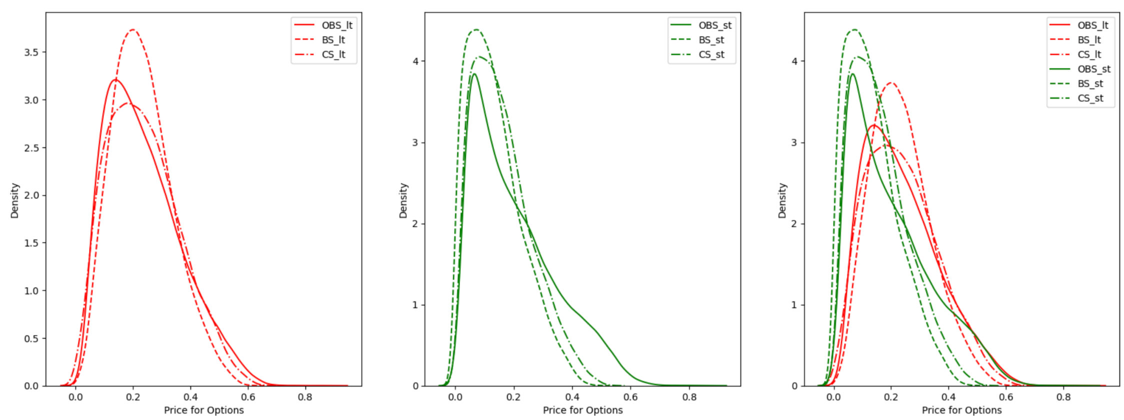

Figure 14: CSI300 Pricing PDFs

In each of the two sets of graphs above, from left to right, the long-term contract price pdfs, the short-term contract price pdfs, and the combination of the two are shown. Solid lines are used in each panel to indicate observed prices, dashed lines to indicate BS model predicted prices, and dotted dashed lines to indicate CS model predicted prices.

It can be seen that, intuitively, short-term contracts have smaller average prices than long-term contracts. For all contracts, the CS model fits the price distribution more closely to the observed proces compared to the BS model, reflecting the advantage of the CS model in terms of predictive ability. Regardless of the type of model, the distribution of prices given reflects a thin tail compared to the observed prices. This phenomenon is particularly evident in short-term contracts. This may be due to the fact that financial products within the Chinese market generally exhibit high kurtosis, i.e., thick tails, making it difficult for model predictions to simulate the true thickness of their tail.

6.3 Pricing Errors

Regarding the pricing error analysis of the models, we categorize the option contracts according to their moneyness and remaining duration, respectively, and obtain the pricing error conclusions under two different error measures, ME and MSE, as mentioned above. The results are shown in the tables as follow:

	Pricing Error of BS and OBS		Pricing Error of CS and OBS		Pricing Error of BS and CS	
	MAE	MSE	MAE	MSE	MAE	MSE
OTM	0.0272	0.00118	0.0170	0.00055	0.0278	0.00142
ATM	0.0271	0.00130	0.0219	0.00089	0.0187	0.00049
ITM	0.0619	0.00570	0.0342	0.00217	0.0387	0.00175
Average	0.0335	0.00205	0.0222	0.00099	0.0259	0.00109

Table 1: SSE50 Pricing Error Based on Moneyness

	Pricing Error of BS and OBS		Pricing Error of CS and OBS		Pricing Error of BS and CS	
	MAE	MSE	MAE	MSE	MAE	MSE
LT	0.0335	0.00206	0.0222	0.00099	0.0259	0.00109
ST	0.0702	0.00861	0.0487	0.00498	0.0271	0.00091
Average	0.0514	0.00525	0.0351	0.00293	0.0265	0.00101

Table 2: SSE50 Pricing Error Based on Expiration Time

What can be seen is that the CS model gives more accurate pricing than the BS model, regardless of the measure of error. The same pricing advantage applies under all different moneyness categories.

First of all, as far as MAE is concerned, it can be seen that BS performs comparably on OTM options and ATM options, and significantly outperforms ITM options. CS Model is likewise performing the worst on ITM options, but it has significantly better performance on OTM options than ATM options (in the sense of 1% of observations). This conclusion is not consistent with Black's (1975) conclusion that the BS model is more accurate in pricing ATM options than both OTM and ITM options. A comparison of the two shows that the gap between the BS model and the CS model on ATM options is the smallest. While on OTM and ITM options, the introduction of implied third order moments and implied fourth order moments helps the CS model to give significantly more accurate pricing.

On the comparison between the long and short term, the prediction errors of both models are larger on the short term contract, which is consistent with the findings of the existing studies. Interestingly, the relative prediction errors of the BS and CS models are essentially the same in the short and long run, reflecting the fact that the gap between the two is not significantly affected by the length of the remaining duration of the contract.

Secondly, since MSE will be more affected by extreme values, we can also get more information from MSE results. Most of the above MAE conclusions hold under MSE, with the obvious difference being the only point below. Namely, it is that the BS model outperforms ATM options for pricing OTM options. This illustrates that compared to OTM options and ITM options,, the BS model produces more extreme values away from the observed value for the pricing of ATM options, i.e., its relative variance is larger.

In addition, we also plot the fitted probability density function images of the pricing error separately as follows:

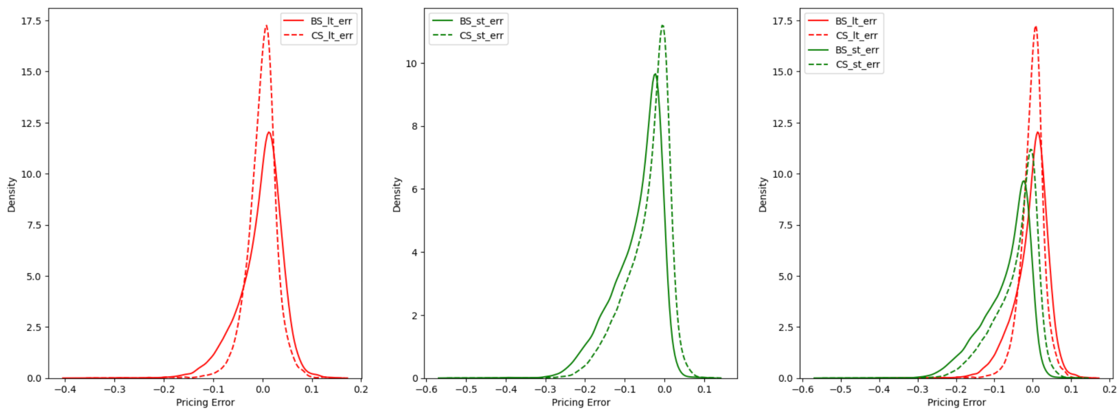

Figure 15: SSE50 Pricing Error PDFs

It is easy to see from the images that the BS model has a larger pricing error and the density function shows thick tails compared to the CS model. At the same time, both models have significantly larger short-term errors compared to long-term errors and show negative means (i.e., tend to undervalue).

As a supplement, we calculate pricing error data and the fitted probability density function images for CSI300 as below:

	Pricing Error of BS and OBS		Pricing Error of CS and OBS		Pricing Error of BS and CS	
	MAE	MSE	MAE	MSE	MAE	MSE
OTM	0.0324	0.00177	0.0274	0.00145	0.0367	0.00226
ATM	0.0373	0.00248	0.0347	0.00215	0.0227	0.00075
ITM	0.0756	0.00908	0.0502	0.00462	0.0391	0.00191
Average	0.0462	0.00407	0.0366	0.00260	0.0318	0.00156

Table 3: CSI300 Pricing Error Based on Moneyness

	Pricing Error of BS and OBS		Pricing Error of CS and OBS		Pricing Error of BS and CS	
	MAE	MSE	MAE	MSE	MAE	MSE
LT	0.0462	0.00406	0.0366	0.00260	0.0318	0.00156
ST	0.1136	0.02173	0.0876	0.01481	0.0354	0.00174
Average	0.0780	0.01241	0.0607	0.00837	0.0335	0.00165

Table 4: CSI300 Pricing Error Based on Expiration Time

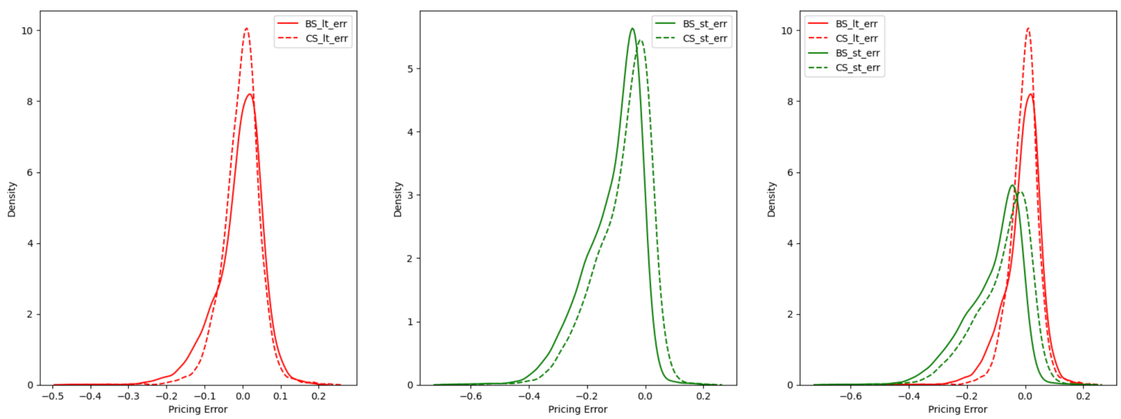

Figure 16: CSI300 Pricing Error PDFs

7 Conclusions

This paper focuses on the European option pricing problem on Chinese A-share stock index options. We mainly compare the difference in pricing power between the traditional Black Scholes model and the CS model, which introduces the implied skewness and kurtosis to the Black Scholes model, on two stock index options, CSI300 and SSE50. The time period of the study is mainly from 2018 to 2020, which includes one year of bull market, one year of bear market and one year of market under epidemic in China's stock market. We also compare side-by-side the differences in the predictive ability of the two models in different environments by grouping them according to different dimensions such as time and moneyness.

Analysis of the data suggests that Chinese stock index options exhibit relatively stable and low levels of volatility in the early years relative to more mature Western markets. Volatility within the first year of the epidemic was particularly sharp. There is also a negative skewness and a kurtosis much lower than 3, representing a negative shift in the peak of the stock index returns and a fat tail in its distribution. In particular, we find that Chinese-specific holiday customs allow implied volatility performance to reflect significant variability around the Chinese New Year.

The predicted price scatter plot shows us more visually the difference in predictive ability between the two models. It can be seen that the BS model always tends to overestimate OTM options and underestimate ITM options, while the CS model consistently gives relatively more accurate forecasts. Both models are not as good at predicting short-term options as long-term options, due to the fact that short-term options are relatively more volatile and introduce more uncertainty into the predictions. Also, we can get the same conclusion from the fitted density functional image.

Finally, we numerically visualize the difference in the predictive ability of the two models by means of a table of the different means of the prediction errors. MAE and MSE errors exhibit similar conclusions, and the conclusions regarding the categorization of monetariness are similar to those in the previous section. BS model has a larger pricing error and its distribution shows thick tails compared to the CS model. At the same time, both models have significantly larger short-term errors compared to long-term errors and show negative means.

Regarding the future direction of research, the authors believe that the main points in the following: First, on the establishment of the results of the study in the markets of other countries. Second, on the phenomenon and the deeper reasons for the role of the holiday for China's financial market. Third, along the development history of China's stock index options market.

8 Reference

1. Implied volatility skews and stock return skewness and kurtosis implied by stock option prices - C. J. Corrado & Tie Su, 1997

2. Effectiveness of the Skewness and Kurtosis Adjusted Black-Scholes Model in Pricing Australian Options - Yanrui Han, 2018

3. Wind Database: www.wind.com