**Non-parametric Tests of Martingale Restriction: A New Approach**

This version: June 25, 2011

**Abstract**

In this study, we use both parametric and non-parametric methods to test the property of martingale restriction in KOSPI 200 index options market. Our results provide strong evidence that the property is violated. Further regression analysis and robustness checks suggest that market friction factors can only marginally explain the relative percentage differences between the option-implied and market observed prices. Our results strongly suggest that there exist arbitrage opportunities in the KOSPI 200 options market.

*JEL classification*: C14; G12; G13; G14

*Keywords*: Martingale restriction; Emerging market; Arbitrage; Option pricing; Non-parametric; Risk-neutral density

**Non-parametric Tests for the Martingale Restriction: A New Approach**

**Biao Guo\*, Qian Han\*\*, and Doojin Ryu\*\*\***

\* Finance & Accounting Division, Business School, Jubilee Campus, University of Nottingham, Nottingham, NG8 1BB, UK, tigerguob@gmail.com

\*\* Corresponding Author, Room A402, Economics Building, Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen, China 361005, Tel: +86 592 2180969, Fax: +86 592 2187708, China, hanqian@gmail.com

\*\*\* Department of International Business, Hankuk University of Foreign Studies, Yongin-si, Gyeonggi-do, Korea 449-791, [doojin.ryu@gmail.com](mailto:doojin.ryu@gmail.com)

**1. Introduction**

This study re-examines the property of martingale restriction proposed by the breakthrough paper of Longstaff (1995). According to the property of the martingale restriction, if a financial market is frictionless and does not allow any arbitrage opportunity, then there exists at least one risk neutral probability measure under which the underlying asset price discounted at the risk-free rate follows a martingale process. Therefore, if we find a violation of this property in the financial market, we can attribute it to either existence of arbitrage or market frictions or to both of them. Either has significant implications for the popular risk neutral valuation approach in derivatives pricing; hence, testing the martingale restriction is of great importance in modern financial economics.

One natural way to test the validity of the martingale restriction is to compare the market observed underlying asset prices with those implied from the prices of its derivatives, for example, options contracts. Depending on the assumptions made on the functional form of the risk neutral density, two approaches have been adopted in the previous literature. The first approach extracts spot prices based on specific option pricing models, mainly the Black-Scholes model (hereafter the “BS model”). This is equivalent to assuming a lognormal risk neutral density function. Manaster and Rendleman (1982) are the first to take this approach though their study does not directly focus on the test of the martingale restriction. Using the S&P 100 index option prices from 1973 to 1976, they find that 63% of the price differences between the options-implied stock prices and the observed stock prices are positive. Longstaff (1995) also extracts the spot prices and implied volatilities simultaneously from the S&P 100 option prices from 1988 to 1989 using the nonlinear programming techniques. He reports that the percentage differences between the implied and observed spot prices are positive for 442 out of 444 observations and claims that this is strong evidence for a violation of the martingale restriction. His regression analysis controlling for the option moneyness, time-to-maturity, and volatility biases of the BS model indicates that the percentage differences are highly related to market frictions such as bid-ask spreads of option prices and other options market liquidity factors. Neumann and Schlag (1996) carry out a similar study for the German DAX index options market and support the rejection of the martingale restriction in the options market for the first six months of 1994. Turvey and Komar (2006) test a variation of the martingale restriction. They show that the martingale restriction under the Black-Scholes setting is equivalent to the condition that the market price of risk equals to the Sharpe ratio, and test this relationship using the implied market price of risk from live cattle options at the Chicago Mercantile Exchange (CME). They find that the implied market prices of risks from option prices vary systematically across strikes and randomly with time and conclude the strong violations of the martingale restriction.

As soon realized by the literature, rejections of the martingale restriction based on the BS model are questionable because the assumptions embedded in the BS model are too restrictive. Therefore the other approach relaxes the log-normality assumption for the risk neutral density function of the underlying spot to allow for more general forms. Longstaff (1995) adopts a four-parameter Hermite polynomial to approximate the risk neutral density and estimates parameters using four call prices closest to the money. He then compares the spot prices computed as a closed form expression of the first moment of the risk neutral density with market observed spot prices. Strong and Xu (1999) repeat similar tests as in Longstaff (1995) but use the S&P 500 index options instead of the S&P 100 options. They claim that the martingale restriction cannot be rejected for the S&P 500 index calls and puts over the period from 1990 to 1994. Although the tests based on the BS model indicate that there are significant price differences across the time-to-maturities and that overall 94% of the differences are positive, the results based on a generalized risk neutral distribution reveal much lower price percentage differences and economically insignificant pricing errors. Using the S&P 500 index options data from 1993 to 1994, Brenner and Eom (1997) compare the price percentage differences under the lognormal, Hermite polynomial, and Laguerre polynomial series risk neutral density functions. They show that the Hermite polynomial expansion creates pricing bias if the true risk neutral density function is not lognormal and that the Laguerre polynomial method is not only unbiased but also has smaller approximation errors for the true distributions. They also find that using the Laguerre density estimator reduces about 80% of the mean percentage differences than using the lognormal specification, while using the Hermite polynomial as in Longstaff (1995) reduces only about 10% of the percentage differences. Based on the t-statistics, they are able to reject the martingale restriction for all cases, but the magnitudes of deviations are much smaller for the Laguerre density estimator than for the Hermite density approximation. Based on the F-statistics, they are able to reject the martingale restriction for the lognormal specification, but not for the general distribution approximated by the Laguerre series. Additionally, they report significant pricing error reductions for the general Laguerre series density compared with the lognormal specification.

As presented so far, the findings of previous studies on the martingale restriction are mixed and the validity of the restriction is still an open empirical questions. In spite of this situation, there has not been much research on testing this important property since the study of Strong and Xu (1999). In addition, the data used in the previous studies are quite outdated as they are all up to the year of 1994 though the global financial markets and the nature of their microstructures have undergone significant changes during the past two decades. Therefore, a martingale restriction test using more recent data is then very necessary.

Another motivation for this study is that so far there have been few studies on the martingale restriction in emerging financial derivatives markets. Since, at least theoretically, the martingale restriction is a fundamental property that should hold in any efficient financial markets, any violation of it, whether due to the market frictions or the existence of arbitrage opportunities, has significant policy implications for the design and regulation of emerging financial derivatives markets. In this paper we mainly focus on KOSPI 200 index options market, which is the most liquid and successful derivative market in the world. If the martingale restriction fails in this market, it can be reasonably expected to fail in other emerging markets.

While there is relatively insufficient research on the martingale restriction considering the importance of its implications, various non-parametric estimation methods to derive the implied risk neutral densities of the underlying assets have been continuously developed. Presumably these non-parametric methods are natural candidates for testing the martingale restriction because they do not depend on any pricing models or on any parametric assumptions on the risk neutral density. However, surprisingly, none of these methods have been applied for testing the martingale restriction.

As Jackwerth (1999) presents a great review on this literature, the non-parametric methods to derive options-implied risk neutral distributions are voluminous. Rubinstein (1994) proposes how to minimize the distance between the terminal risk neutral probability and the lognormal prior probability subject to some pricing constraints. One of the constraints is exactly the martingale restriction: the spot price given by the risk neutral probability has to fall within the observed bid-ask spread of the underlying spot. Jackwerth and Rubinstein (1996) show that different measures of the distance do not severely affect the final risk neutral distribution if there are enough observations. Shimko (1993) provides a parabolic equation to smooth implied volatilities, and gives an analytic expression for the density functions under the parabolic structure framework. Rosenberg and Engle (1997) use a polynomial to fit the log smile curve and obtain a smoothed pricing function using the Black-Scholes formula. Then, according to Breeden and Litzenberger (1978), they take the second derivatives of the pricing function with respect to the strikes to derive the risk neutral density. Aït-Sahalia and Lo (1998) use the kernel regressions across five dimensions of stock price, strike price, time to maturity, dividend yield, and interest rate to fit the observed volatility smile.[[1]](#footnote-1) In this study we adopt the kernel regression methods described in Aït-Sahalia and Lo (1998) and the implied volatility smoothing method in Shimko (1993) as the main non-parametric methods to test the martingale restriction property in the KOSPI 200 index options market.

For comparison, we also investigate the martingale restriction using parametric methods based on the BS model and on the Edgeworth expansion risk neutral density distribution as in Longstaff (1995). Our results strongly reject the martingale restriction property for the sample period from 2006 to 2010. The non-parametric methods generally have much less pricing errors relative to parametric ones. Furthermore, the regressions of the positive price percentage differences on the market frictions factors suggest that unlike developed markets the differences are not so much related to market frictions but more related to speculations in the market. Our overall findings and conclusion are robust when we use the KOSPI 200 index futures instead of the KOSPI 200 index as the underlying.

To the best of our knowledge, this is the first paper to adopt non-parametric methods for testing the martingale restriction. This study is also the first to examine the martingale property in emerging financial markets, especially in the KOSPI 200 options market. Therefore, our findings will have important policy implications to derivatives market designers and regulators.

The rest of the paper goes as follows. Section 2 briefly summarizes the non-parametric methods used in this study. Section 3 describes the property of the KOSPI 200 options market and Section 4 illustrates the sample data. Section 5 presents empirical results on the martingale restriction property. Section 6 is devoted to robustness check. Section 7 concludes this study.

**2. Methodology**

In this section, we introduce the methods to be used in later sections to estimate the risk neutral densities from option prices. We classify these methods into two categories, parametric and nonparametric, according to whether a specific form of risk neutral density is assumed.

Under the frictionless market and no arbitrage assumptions the price (*H*) of a European option is given as in Equation (1).

(1)

Here, *τ*=*T-t* is the time to maturity, *ST* is the price of the underlying asset at the maturity date *T*, *K* is the strike price, *r* is the continuous compounding risk free rate, *P*(*ST*, *K*) is the payoff function of the option, *fT*(*ST*) is the date-*t* risk neutral density for date-*T* payoff. The martingale restriction is the constraint that the above pricing formula applies to the underlying asset. In other words, the martingale restriction implies that the option implied underlying prices should equal the market observed prices. Equation (2) shows this.

(2)

This constraint is also called “internal consistency” in Harrison and Kreps (1979) and it is the definition for the risk neutral measure. Violations of the martingale restriction would imply either the market is not frictionless or there exist arbitrage opportunities or both. Our approach to test the martingale restriction involves calibrating (parametrically) or estimating (non-parametrically) the risk neutral density from option prices using Equation (1), and then use Equation (2) to compare the option-implied underlying prices with the market prices of the underlying assets.

*2.1. Parametric Approaches*

We select the classic BS model and the Longstaff (1995)’s four-factor model to represent parametric methods. European options under the BS model are valued as in Equation (3).

(3)

Here, *C* denotes the call option price and *P* denotes the put option prices; *N*(.) is the cumulative distribution of the standard normal distribution; *σ* is the volatility of underlying asset’s returns, and *St* is the dividend-adjusted underlying asset price at time *t[[2]](#footnote-2)*; *d*1*=*[*ln*(*St*/*K*)+(*r+σ2*/*2*)*τ*]/*σ√τ* and *d2=d*1*- σ√τ*. The BS model-implied spot prices can then be calibrated and compared to the market prices for the underlying. Realizing that the BS model may be subject to the problem of model-misspecification despite of its widespread use and simplicity, Longstaff (1995) extends the log-normal density assumption to a four-factor Edgeworth expansion. It is therefore a general form with the BS model, the Merton (1973)’s stochastic interest rate model, and the Merton (1976)’s jump diffusion model as special cases.

Let *Z*=(*ln*(*ST*)*-μ*)/*σ*, where *μ* is the conditional mean. The risk-neutral density of *Z* is assumed to belong to an Edgeworth expansion family of density functions:

(4)

Here, *β* and *γ* are parameters related to the skewness and kurtosis of the underlying asset price. A call option is priced as follows:

(5)

The price of a put option can be derived based on the following put-call parity.

*P=Ke-rτ-St+C* (6)

Once the parameters are estimated, the first moment of the risk neutral density is given by

Et[ST]=exp(μ+σ2/2)(1+βσ3+γσ4) (7)

and the martingale restriction is expressed as follows.

St=exp(-r*τ*+ μ+σ2/2)( 1+βσ3+γσ4) (8)

*2.2. Non-Parametric Approaches*

Breeden and Litzenberger (1978) show that the risk neutral density *ft*(*ST*) can be obtained from options prices by taking the second order derivative of the option pricing function with respect to the strike price *K*.

(9)

Many non-parametric methods have been proposed to estimate the option pricing function *H*. Unlike a parametric method that assumes a functional form structure, a non-parametric method doesn’t specify a model; instead, its structure is determined purely from data. The literature on non-parametric methods is various. Artificial neural network, radial basis function, and kernel regressions are among the most widely used ones. In this study, we choose the kernel regression as our main non-parametric methods. As shown in Ait-Sahalia and Lo (1998), the kernel regression is better suited for hypothesis testing and requires less restrictive assumptions than other methods.

Given a set of historical European option prices *Hi* and accompanying variables *Vi*=(*Si*, *Ki*, *ri*, *τi*),[[3]](#footnote-3) we look for a non-parametric method that is able to generate option prices *H* with as small mean squared error as possible. Let *h* be a bandwidth used for the kernel regression, an estimator of *H* conditional on *V* is given as the Nadaraya-Watson kernel estimator:

(10)

where *n* is the sample size, *k*((*V-Vi*)/*h*) is a kernel function. It is known that the choice of the kernel function has little effect on the result, while the choice of the bandwidth *h* determines the accuracy of final outcome. In this study, we choose a second order Gaussian kernel as

(11)

and choose *h* by minimizing the least-squares cross-validation.[[4]](#footnote-4)

Because we have four explanatory variables, our multivariate kernel function becomes a product of four univariate kernels, such that the fully non-parametric pricing function is

(12)

differentiating the function twice numerically generates the risk neutral density implicit in the options prices.[[5]](#footnote-5)

A similar but indirect way to estimate the pricing function is to estimate the implied volatility function non-parametrically first. Ait-Sahalia and Lo (1998) express the pricing function *H* as

(13)

Here *BS*(.) denotes the Black-Scholes equation. The implied volatility is estimated non-parametrically in the same manner.

(14)

*σi* is the volatility implied from the observed option price. The risk neutral density is then generated by differentiating twice numerically.[[6]](#footnote-6)

Another well-known non-parametric method is introduced in Shimko (1993), in which the author smoothes the implied volatility structure by a parabolic function of the strike price:

(15)

By fitting the parameters *A*0, *A*1, and *A*2 with least squares, a smooth volatility curve is generated for every *K* and *τ*. These smoothed volatilities are then used to compute a series of smooth option prices via the Black-Scholes equation. Note that just like Ait-Sahalia and Lo (1998), Shimko (1993) doesn’t require the BS model to be correct, he merely uses the equation to interpolate implied volatilities rather than the market prices of options. Given the parabola volatility structure, the risk neutral density function is solved analytically as follows.

(16)

where

And *d*1 and *d*2 are the same as in the Black-Scholes equation, *n*() is the density distribution of the standard normal distribution.

**3. The KOSPI 200 options market**

The KOSPI 200 options market is an ideal place to test the martingale restriction property because of its following characteristics. First, the extreme liquidity and negligible transaction costs of the KOSPI 200 options market reduce the possibility that the market frictions and illiquidity affect the systematic bias of the option-implied underlying prices. Despite its short history, the KOSPI 200 options market has sharply grown to the most actively traded derivative market in the world. For this reason, the Asian Risk magazine nominates the KOSPI 200 options market as the most remarkable derivatives market with the advent of the new millennium. Table 1 presents the trading volume in terms of the number of contracts for the top 10 equity index futures and options products and their exchanges in the world during 2009 and 2010. The table clearly indicates that the trading activity in the KOSPI 200 options dominates that in other markets.

In addition to the ample liquidity, the index options trading in the Korean market only requires negligible transaction costs. For the equity trading, individuals pay substantial trading taxes and commissions and institutions also pay membership and exchange fees. In comparison, the options trading exempts individual investors from trading taxes and often does not require commission fees.[[7]](#footnote-7) The amount of membership and exchange fees is also relatively small in the case of index option trading. When it comes to bid/ask spreads which are often used to measure the transaction cost and market liquidity, the average size of the spreads is very tiny, usually one or two times minimum tick size (See Ahn, Kang, and Ryu (2008, 2010)).

Second, considering that the KOSPI 200 options market is highly speculative and the options traders are known to be easily affected by behavioral and physiological biases, this might generate the systematic pricing bias of the implied spot prices even without market frictions. This second characteristic is possibly related to the high participation rate of domestic individuals, generally considered as noisy and speculative traders. While sophisticated and informed institutional investors are dominant market players in derivatives markets of developed countries, domestic individuals are very active investors in the KOSPI 200 options market. Table 2, which presents the trading activities in terms of the trading volume (in contracts) by investor type, shows that the trading activity of domestic individual investors accounts for more than one third of the total transactions during the sample period. It also shows that these individual investors concentrate on out-of-the-money (OTM) and especially deep-OTM options, which are usually used as high leverage and speculative trading tools (see Ahn et al. (2008, 2010)).[[8]](#footnote-8) Given the coexistence of extreme liquidity, low transaction costs and abundant speculative activities in the KOSPI 200 index options market, it is only natural to ask the question of whether these imply arbitrage opportunities.

The maturity date of the KOSPI 200 options are set to the second Thursday of three consecutive near-term months and one nearest month from the quarterly cycle (March, June, September, and December). For instance, if today is April first, the maturity months of options will be April, May, June, and September. Though the four kinds of option contracts classified by the maturity are always listed, only the nearest maturity contracts are actively traded (see Ryu(2011)). We examine the implication of this phenomena in section 6.3.

The KOSPI 200 options market is classified as a pure order-driven market where the orders are transacted by the central electronic limit order book (CLOB), not by designated market makers. All orders submitted by options traders are either piled up at the CLOB or transacted with other standing orders in the CLOB. Like the underlying index market, the options market begins at 9:00 A.M. per each trading day. However, it closes at 3:15 P.M, 15 minutes later than the closing time of the underlying index market. All orders submitted during the intraday trading secession are transacted under the continuous double auction-trading rule though there exist a one-hour long pre-opening session and a last ten-minute closing session where all orders are transacted under the uniform pricing rule. We take care of the non-synchronous issue in section 6.

**4. Sample Data**

For the period from January 2006 to December 2010, we make daily time series data using only the nearest maturity contracts of the KOSPI 200 options considering that other maturity contracts are barely traded in the Korea Exchange (KRX). Then, we match the options data with the daily time series of the KOSPI 200 index price. On each trading day in the sample, the options contracts are screened with the following filters. First, options with number of transactions less than 30 are excluded. Second, the bid and ask prices should be greater than the minimum tick size (which is described in Section 3). Third, the closing prices of options should be greater than the no-arbitrage lower bound and should lie between 0.03 and 15 points. Equation (17) (Equation (18)) represents the no-arbitrage lower bound for call (put) options.

(17)

(18)

Fourth, Black-Scholes implied volatility (Equation (19)) should be greater than 0.05 and less than 0.95.

*Implied volatility=σ*(*Ot*; *St*, *rf*, *τ*) (19)

Here *Ot* is the observed option price at time *t*; *τ* is the time to maturity as calculated based on the trading days. Fifth, the absolute value of option moneyness defined as |*K*/*St-*1| should be less than 0.1.

*|Option moneyness|=*|*K*/*St-*1|≤0.1 (20)

Lastly, we exclude the trading days with the total numbers of options less than 8. The screening filters listed above leave us 15131 call options in 1206 trading days and 15962 put options in 1202 trading days.

Table 3 summarizes the statistics for the call and put options in our final sample. The option moneyness is newly defined as the difference between index value and strike price for calls and the difference between strike price and index value for puts. There is slightly more puts traded than calls in our sample. The daily average (median) number of puts traded is 13.28 (14) and that of calls is 12.54 (12). The average of moneyness is negative for calls (-0.3296), whereas it is positive for puts (1.6402). Since only nearest maturity contracts are used, both calls and puts have a short time to maturity. The maximum maturity is 0.0976 years, roughly one month, and the median maturity is 0.0405 years, roughly two weeks. Over the sample period, the KOSPI 200 index has fluctuated from a minimum value of 123.54 to a maximum value of 271.19.

**5. Empirical Results**

*5.1. Testing the Martingale Restriction*

This section tests the martingale restriction using the methods described in section 2. We report our results in Table 4. Clearly, the martingale restriction is violated for our sample. At the 95% confidence level, all methods reject the null hypothesis for call options in the sample, and all except the Longstaff model rejects the null hypothesis for the put options. All methods report positive mean and median percentage differences for call options, and 3 out of 5 methods report positive mean differences for put options.

More specifically, the BS model has the largest t-statistic with an average percentage difference of 0.263% for calls and 0.358% for puts. The Longstaff model gives the smallest absolute mean percentage differences of around 0.03% for both calls and puts, corresponding to a difference of only 0.06 point. But it is still a substantial value considering that the spreads in the options market is often just one-tick (i.e. 0.01 point). For nonparametric methods, the Shimko method obtains relatively small mean percentage differences of 0.064% for calls and 0.059% for puts. The PKR and FKR methods produce the largest mean percentage differences for calls (around 0.318% and 0.532%) and the second-largest differences for puts (0.184% and -0.41% respectively).

The result in Table 4 indicates that generally it is more expensive to trade the underlying asset through the options market than from to directly trade it in the spot market. Longstaff (1995) explains that these relatively “expensive” option prices might reflect the higher transaction costs in the options market. However, that explanation does not apply to the KOSPI 200 options market in that the size of bid-ask spread is very tiny and other costs such as commissions, taxes, and transaction fees are negligible in this market.

We may possibly attribute this result to other characteristics of option contracts and the KOSPI 200 options market. First, a trader trying to dynamically replicate the payoffs of options using the underlying index and risk-free assets will incur substantially higher transaction costs in the stock and bond markets than in the index options market. This higher hedging cost will be reflected in option-implied index values. However, since we are looking at the existence of arbitrage in the options market, i.e. we take option prices as given, this argument is of little relevance for our purpose. Second, the excess demand of “crash phobia” institutional investors may force up the option prices. Institutional investors often buy OTM puts to insure against potential future market crashes. In section 5.3 we control the effects of moneyness and number of puts traded on each particular day. Third, the ample liquidity of the KOSPI 200 options market might induce options traders to willingly pay an extra premium. We use the total open interest and total traded value as control variables for this effect in section 5.3. Last but not the least, the expensive option-implied underlying prices may be caused by the behavioral bias and reckless buy trades of domestic individuals. It is well known that many domestic individual investors who have little trading experience and knowledge tend to habitually buy the KOSPI 200 options with simple speculative reasons. They even actively buy the Deep-OTM options though those options have little possibility to result in positive profits. The domestic individuals just buy those “cheap” options as if they buy a lottery or participate in a great casino (see Ahn et al. (2008, 2010)). It is reasonable to expect that under these situations, sophisticated investors (e.g., foreign institutions) can take advantage of the price differences between the option-implied and market-observed spots to make arbitrage profits.

*5.2. Relaxing the Martingale Restriction*

Given the general rejection of the martingale restriction from the data, we follow Longstaff (1995) to see if the model pricing performance can be improved by relaxing the martingale restriction constraint. For the BS Model and Longstaff methods, we calculate both the restricted and unrestricted option prices. Restricted prices are obtained subject to the martingale restriction and unrestricted prices are free from such a constraint. To be specific, only implied volatility is estimated for the restricted BS model, while implied volatility and implied underlying asset value are jointly estimated for the unrestricted BS model.

Table 5 lists the summary statistics for the differences between the actual and fitted call and put option prices. For call options, while the unrestricted BS model greatly reduces the mean pricing error to -0.0135 from 0.0443 in the restricted BS model, the unrestricted Longstaff model only marginally reduces the errors from 0.0049 to 0.0032. A similar pattern is found for put options. The restricted Longstaff model significantly outperforms the restricted BS model with a 0.0049 mean pricing error, compared with the 0.0443 pricing error in the latter. We also include non-parametric methods for compassion[[9]](#footnote-9). Nonparametric models generally outperform parametric models with both smaller mean pricing error and lower standard deviation. The PKR method has a mean error 0.0033 and standard deviation 0.1295 for calls, 0.0013 and 0.1258 for puts. Compared with the BS model, the mean pricing error of the PKR method is only 7.45% as big as the error of the BS model, and shrinks the standard deviation shrinks by 279%. Overall, the PKR method has the smallest absolute error for call options, and the FKR method has the smallest absolute error for put options. Within non-parametric methods, the PKR method performs better than the Shimko method possibly due to the parabola volatility structure constraint of the Shimko method. The mean pricing error reduces from 0.0063 for the Shimko method to 0.0033 for the PKR in case of call options, and from 0.0030 to 0.0013 for put options.

Figure 1 and Figure 2 visually illustrate the pricing performances of our methods against moneyness. Notice that we observe upward moneyness bias for calls and downward moneyness bias for puts in the BS model, consistent with Longstaff (1995) and Brenner and Eom (1997). Similar patterns, although not as obvious as for the BS model, are found for other methods.

*5.3. Properties of the Pricing Differences*

Option implied underlying asset values can be greater than the observed spot prices due to market frictions. Longstaff (1995) regresses the differences between the implied and the actual index values on several selected market friction variables such as average bid-ask spread, open interest, and total trading volume of the options while controlling for time to maturity, moneyness, current and first two lagged absolute daily returns on the S&P 100 index. He finds all market friction variables are significant and the R2 is as high as 64.4%, a demonstration that market frictions are highly related to the rejection of martingale restriction. However, using S&P 500 options data, Brenner and Eom (1997) find no relation between estimated pricing differences and market frictions for their proposed general distribution of the risk neutral density.

In this section, we regress the index differences on several measures of market frictions[[10]](#footnote-10), including the total open interest of all KOSPI 200 options, the total trading value, and the total number of options used to compute the pricing differences for each day in our sample period. As discussed in section 5.1, we also include moneyness to control for the fact that there are strong interests in trading OTM options from both institutional and individual investors. Table 7 reports the regression results. The coefficient for the total open interest is negative and significant for all except for the FKR method. The negative sign makes sense because as the total open interest increases, the option market becomes more liquid. The coefficient for the total trading value for call options is insignificant except for the FKR method. However, it remains positively significant for put options for four of our methods. The coefficient for the number of options isn’t as significant as for the total open interest, and the sign is rather mixed for both call and put options, unlike the strong negative relation found in Longstaff (1995). The unrestricted BS model has the highest R2 18.95% and 15% for calls and puts respectively. This shows that market frictions has the largest influence for the BS model. The smallest R2 occurs for the Longstaff model with only 1.03% explaining power for calls and 3.7% for puts. The significance of the intercept C for all methods indicates that there are other important factors that are attributable to the violation of the martingale restriction property but are left out of our regression model.

In summary, our regression analysis shows that market frictions have rather small explanatory power for the failure of martingale restriction. This implies that the KOSPI 200 options market may not be arbitrage-free.

**6. Robustness Tests**

*6.1. Using KOSPI 200 Index Futures as the underlying asset*

There may be some possible criticisms of our previous test using the KOSPI 200 index prices. First, we assume that investors can perfectly forecast the future dividends until the maturity of each option contract. This assumption does not seem to reflect the reality. However, the effect of this assumption might be negligible considering that the time-to-maturity of options is usually less than a month and the average present dividend value is very small compared to the level of KOSPI 200 index price.[[11]](#footnote-11) Second, the KOSPI 200 index prices might not reflect the fundamental value of the underlying asset due to the problems related to the non-synchronous trading and stale price, which are often brought up when index prices are used. These problems may also be negligible because the KOSPI 200 index consists of the most actively traded 200 underlying stocks in the Korean stock market. Third, there is 15-minute difference between the closing times of the stock market and of the options market. Nevertheless, this 15-minute difference may not be serious in that many of previous research still use the closing prices of underlying index and index options.

Though these minor problems embedded in the KOSPI 200 index prices do not seem to affect the empirical results, for the robustness check, we carry out the same empirical tests using the KOSPI 200 futures contracts to avoid these potential issues. The KOSPI 200 futures contracts are also very actively traded and its daily closing time is the same as that of the options market. Also in practice, traders often use cheaper futures contracts instead of the index itself to hedge an index option. Therefore we compare implied futures values with observed futures values for the robustness test of the martingale restriction. This methodology has been adopted in many studies including Longstaff (1995), Brenner and Eom (1997) and Strong and Xu (1999).

Despite of the many similar features between the KOSPI 200 options and futures, they have different time to maturities. The maturity months of the KOSPI 200 futures are March, June, September, and December. Therefore, we only select the trading days where nearest maturity dates of the options and futures contracts are the same to each other. We match them by expiration dates and end up with 403 daily call options with 5259 observations, and 403 daily put options with 5381 observations. Table 8 reports the results for the martingale restriction test using KOSPI index futures. In line with the test using index values, all our methods reject the null hypothesis for call options, and all except Longstaff models reject the null hypothesis for put options at the 95% confidence level. The Longstaff model now rejects the null hypothesis at the 90% confidence level and has positive price percentage differences compared with the results in Table 4.

Table 9, together with Figure 3 and 4, report the option pricing errors for call and put options with index futures as the underlying. Overall, using index futures greatly reduces the pricing errors. Consistent with the previous results, non-parametric methods outperform parametric methods with smaller mean pricing errors. The mean error of call options for restricted BSM is 0.0261 and is -0.0004, 0.0018 and 0.0007 for the Shimko, PKR and FKR, respectively. Although restricted Longstaff has -0.0003 mean error, its pricing error is more volatile than non-parametric methods. For both call and put options, FKR has the smallest absolute errors, demonstrating the robustness of kernel regression methods and its ability to deal with outliers. However, the superiority of kernel regression methods over others diminishes. One possible explanation is that the intensive data requirement nature of kernel regression methods produce worse performance when the number of sample size decreases from over 15000 to about 5000 due to maturity dates matching.

Similar results are obtained when we regress the index futures differences on those variables for market frictions. Table 10 lists the coefficient estimates and t-statistics. Again, the coefficient for the total open interest is negatively significant for all methods except FKR. The total trading value becomes less significant, especially for put options. The model explanatory power, as captured by the R square, increases for all methods, but still not very high. The largest R2 value of 25% is achieved for the BS model, compared with the smallest R2 of 4.14% for the Longstaff model.

*6.2. Time to maturity, moneyness and volatility biases*

To further check the validity of our results, in analogous with Longstaff (1995), we add a few control variables to capture three possible model bias: the average time to maturity for the time to expiration bias, the average moneyness for the moneyness bias[[12]](#footnote-12), and the current, first and second lagged absolute daily returns of the KOSPI 200 index for the volatility bias. Table 11 reports the regression coefficients and p-values for call and put options using index values[[13]](#footnote-13). As expected, adding extra variables increases the value of R square, but the adjusted R2 for all methods are still much smaller than the 64.4% value in Longstaff (1995). Both Longstaff and Shimko models have adjusted R2 less than 10%. The results also indicate that once these biases are controlled, the market frictions factors such as the total traded value (TTV) and open interest (OI) become less significant. The coefficient for TTV is significant only in two models (the Longstaff and Shimko methods for calls and the BS model and the FKR method for puts). The sign of coefficient for the number of options remains mixed.

*6.3. Nearest Maturity Bias*

In all our previous tests, we have ignored one unique feature of the KOSPI 200 index options market. Unlike its US peer, the trading volume of the nearest maturity contracts is much higher than that of the second nearest maturity contracts. It is now well known that option prices for these nearest maturity contracts are usually far from their fair values. We redo the martingale test and re-run the regression analysis by excluding those contracts that mature within three days to the immediate next maturity date. The martingale restriction test is very close to the previous results, so we do not present them here. Table 12 shows the results for regression analysis using the index futures as the underlying.

The first thing to notice is that the intercept C now becomes very significant for all cases at a 95% level, meaning that those omitted nearest maturity contracts indeed represent part of the arbitrage opportunities in the KOSPI 200 index options market. Based on the BS model, we compute the R square differences in Table 10 and Table 12 for calls and puts. These differences indicate that nearest maturity contracts contribute roughly 6% to the whole arbitrage opportunities for calls and 11% for puts. Second, the R square is higher for all cases than without excluding those nearest maturity contracts, with the highest value being 31% for the BS model. Third, market friction variables become more significant, especially for the TTV. However, the relatively low R square combined with a strong rejection of the martingale restriction property suggest that even after we exclude those obvious speculative behaviors in the market, the market is still far from arbitrage free[[14]](#footnote-14).

**7. Conclusions**

We examine the property of martingale restriction for the Korean KOSPI 200 index options market. To avoid model misspecifications, we use both parametric and non-parametric methods to estimate the risk neutral density function, and then use it to compute the option implied index values. Positive differences between these option-implied values and market observed values and the t-tests suggest that martingale restriction is violated in this market for almost all methods that we use. Regression analysis shows that market frictions factors such as total trading value, open interest and the number of options have rather small explanatory power for these percentage differences. We conclude that the fundamental assumption of no arbitrage is not valid for the Korean options market. This finding is of critical importance to the practitioners in the Korean options market as well as the regulators.

Our conclusion is affirmative because: (1) we use both parametric and non-parametric methods to test for martingale restriction. Previous studies only consider parametric ones. Since non-parametric methods do not rely on any specific models, our findings cannot be easily challenged; (2) we show that non-parametric methods have much better pricing performance relative to the parametric ones. Hence, non-parametric methods are more reliable for the martingale restriction test. All tests based on non-parametric methods reject the null hypothesis that option implied index values are equal to the actual index values; (3) we do a battery of robustness checks to ensure our results are not subject to certain factors. We replace index values by index futures values to get rid of market non-synchronous and stale price issues and find even stronger rejections of the property. We also control for maturity, moneyness and volatility biases when examining the relation between the price percentage differences and market friction factors, but our major results stay the same. It is possible that we have missed some critical factors in our regression analysis, but given an average of 30% of explanatory power with eight independent variables already, significant improvements seem very unlikely.

**Bibliography**

Ahn, H.-J., J. Kang, and D. Ryu (2008). "Informed trading in the index option market: The case of KOSPI 200 options." Journal of Futures Markets 28(12): 1118-1146.

Ahn, H.-J., J. Kang, and D. Ryu (2010). "Information Effects of Trade Size and Trade Direction: Evidence from the KOSPI 200 Index Options Market." Asia-Pacific Journal of Financial Studies 39(3): 301-339.

Ait-Sahalia, Y. and A. W. Lo (1998). "Nonparametric estimation of state-price densities implicit in financial asset prices." Journal of Finance 53(2): 499-547.

Bahra, B. (1997). "Implied Risk-neutral Probability Density Functions From Option Prices: Theory and Application." Bank of England Working Paper No 66.

Breeden, D. T. and R. H. Litzenberger (1978). "Prices of State-Contingent Claims Implicit in Option Prices." Journal of Business 51(4): 621-651.

Brenner, M. and Y. H. Eom (1997). "No-Arbitrage Option Pricing: New Evidence on the Validity of the Martingale Property." NYU Working Paper No. FIN-98-009.

Choe, H. and Y. Eom (2009). "The disposition effect and investment performance in the futures market." Journal of Futures Markets 29(6): 496-522.

Figlewski, S. (2008). "Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio." Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle.

Harrison, J. M. and D. M. Kreps (1979). "Martingales and Arbitrage in Multiperiod Securities Markets." Journal of Economic Theory 20(3): 381-408.

Jackwerth, J. C. (1999). "Option Implied Risk-Neutral Distributions and Implied Binomial Trees: A Literature Review." Journal of Derivatives 7(2): 66-82.

Jackwerth, J. C. and M. Rubinstein (1996). "Recovering Probability Distributions from Option Prices." The Journal of Finance 51(5): 1611-1631.

Longstaff, F. A. (1995). "Option pricing and the martingale restriction." Review of Financial Studies 8(4): 1091-1124.

Manaster, S. and J. R. J. Rendleman (1982). "Option Prices as Predictors of Equilibrium Stock Prices." Journal of Finance 37(4): 1043-1057.

Merton, R. C. (1973). "Theory of Rational Option Pricing." Bell Journal of Economics 4(1): 141-183.

Merton, R. C. (1976). "Option Pricing When Underlying Stock Returns Are Discontinuous." Journal of Financial Economics 3(1-2): 125-144.

Neumann, M. and C. Schlag (1996). "Martingale restrictions and implied distributions for German stock index option prices." Working Paper, Universit at Karlsruhe.

Rosenberg, J. V. and R. F. Engle (1997). "Option Hedging Using Empirical Pricing Kernels." NBER Working Papers 6222, National Bureau of Economic Research, Inc.

Rubinstein, M. (1994). "Implied Binomial Trees." Journal of Finance 49(3): 771-818.

Ryu, D. (2011). "Intraday Price Formation and Bid/Ask Spreads: A Cross Market Approach." Journal of Futures Markets, forthcoming.

Shimko, D. (1993). "Bounds of Probability." Risk Magazine 6(4).

Stone, C. J. (1984). "An Asymptotically Optimal Window Selection Rule for Kernel Density Estimates." Annals of Statistics 12(4): 1285-1297.

Strong, N. and X. Xu (1999). "Do S&P 500 index options violate the martingale restriction?" Journal of Futures Markets 19(5): 499-521.

Turvey, C. G. and S. Komar (2006). "Martingale Restrictions and the Implied Market Price of Risk." Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie 54(3): 379-399.

**Table 1: Top 10 Equity Index Futures and Options Worldwide**

Notes: This table shows the ten most active index derivatives contracts, which are ranked by the number of contracts traded and/or cleared in 2009 and 2010. The SPDR S&P 500 ETF Options are traded on multiple U.S. options exchanges. Source: *Futures Industry Association* ([www.futuresindustry.org](http://www.futuresindustry.org))

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Rank | Contract | Index Multiplier | 2010 | 2009 |
| 1 | KOSPI 200 Options, KRX | 100,000 Korean won | 3,525,898,562 | 2,920,990,655 |
| 2 | E-mini S&P 500 Index Futures, CME | 50 U.S. dollars | 555,328,670 | 556,314,143 |
| 3 | SPDR S&P 500 ETF Options | NA | 456,863,881 | 347,697,659 |
| 4 | S&P CNX Nifty Index Options, NSE India | 100 Indian rupees | 529,773,463 | 321,265,217 |
| 5 | Euro Stoxx 50 Futures, Eurex | 10 Euros | 372,229,766 | 333,407,299 |
| 6 | Euro Stoxx 50 Index Options, Eurex | 10 Euros | 284,707,318 | 300,208,574 |
| 7 | RTS Index Futures, RTS | 2 U.S. dollars | 224,696,733 | 150,019,917 |
| 8 | S&P 500 Index Options, CBOE | 100 U.S. dollars | 175,291,508 | 154,869,646 |
| 9 | S&P CNX Nifty Index Futures, NSE India | 100 India rupees | 156,351,505 | 195,759,414 |
| 10 | Nikkei 225 Mini Futures, OSE | 100 Yen | 125,113,769 | 104,738,309 |

**Table 2: Trading Activities by Investor Types**

Notes: The table presents summary statistics of the trading activities in the KOSPI 200 options market by investor type among domestic individuals, domestic institutions, and foreigners from January of 2006 to December of 2010. The domestic institutional investors are further classified into seven sub-categories, which are securities and futures, insurance companies, investment trusts, banks, merchant banks, pension funds, government, and other corporations. The trading activity is gauged in terms of the number contracts.

|  |  |  |
| --- | --- | --- |
| Investor Group | Total (Contract) | Percentage (%) |
| Securities & Futures | 10,391,992,867 | 36.24 |
| Insurance | 177,698,763 | 0.62 |
| Invest Trusts | 209,119,324 | 0.73 |
| Banks | 28,216,583 | 0.1 |
| Merchant | 4,639,461 | 0.02 |
| Pension Funds | 74,002,132 | 0.26 |
| Government | 1,256,383 | 0 |
| Other | 136,873,761 | 0.48 |
| Domestic Individuals | 10,112,553,157 | 35.27 |
| Foreigners | 7,538,908,875 | 26.29 |
| Total | 28,675,261,306 | 100 |

**Table 3: Summary statistics for the options sample**

Notes: This table shows summary statistics for the KOSPI 200 call and put options included in the sample from January 2006 to December 2010. “Number” denotes the daily number of call options; “Index value” denotes the KOSPI 200 index price; “Moneyness” for call options denotes the average difference between the index value and the strike price of the options for a given day; “Moneyness” for put options denotes the average difference between the strike price of the option and the index value for a given day; “Time to maturity” denotes the (annualized) average time to expiration for a given day.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Panel A. Call options** | | | | | |
|  | Mean | Std. Dev. | Min | Median | Max |
| Number | 12.55 | 3.49 | 1 | 12 | 20 |
| Index value | 201.33 | 31.07 | 123.54 | 204.65 | 271.19 |
| Moneyness | -0.330 | 2.970 | -11.160 | -0.275 | 13.280 |
| Time to maturity | 0.0416 | 0.0246 | 0.0000 | 0.0405 | 0.0976 |
|  |  |  |  |  |  |
| **Panel B. Put options** | | | | | |
|  | Mean | Std. Dev. | Min | Median | Max |
| Number | 13.28 | 3.33 | 1 | 14 | 20 |
| Index value | 201.44 | 30.95 | 123.54 | 204.76 | 271.19 |
| Moneyness | 1.640 | 2.928 | -12.967 | 1.176 | 11.690 |
| Time to maturity | 0.0417 | 0.0245 | 0.0000 | 0.0405 | 0.0976 |

**Table 4: Percentage price differences between the implied and observed underlying index values**

Notes: This table presents the summary statistics for the relative percentage differences between the implied and actual index value for call and put options. t-statistic values are used for the null hypothesis *H0* that the difference equals to zero, and the critical values are 1.646, 1.962, and 2.580 for 10%, 5%, and 1% confidence level for a two-side t-test.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Panel A. Call options** | | | | | | |
|  | Mean | Std.dev | Min | Median | Max | t-stat. |
| **Parametric** |  |  |  |  |  |  |
| BSM | 0.263 | 0.451 | -1.520 | 0.170 | 3.332 | 20.242 |
| Longstaff | 0.035 | 0.522 | -4.441 | 0.049 | 4.627 | 2.284 |
|  |  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |  |
| Shimko | 0.064 | 0.193 | -0.686 | 0.059 | 2.362 | 10.427 |
| Partial Kernel Regression | 0.318 | 0.639 | -3.055 | 0.222 | 5.260 | 17.151 |
| Full Kernel Regression | 0.532 | 1.306 | -4.921 | 0.689 | 5.803 | 14.148 |
|  |  |  |  |  |  |  |
| **Panel B. Put options** | | | | | | |
|  | Mean | Std.dev | Min | Median | Max | t-stat. |
| **Parametric** |  |  |  |  |  |  |
| BSM | 0.358 | 0.857 | -0.919 | 0.126 | 7.510 | 14.467 |
| Longstaff | -0.031 | 0.825 | -2.325 | -0.097 | 10.107 | -1.270 |
|  |  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |  |
| Shimko | 0.059 | 0.247 | -0.667 | 0.048 | 5.139 | 6.611 |
| Partial Kernel Regression | 0.184 | 0.748 | -2.637 | 0.030 | 7.032 | 8.477 |
| Full Kernel Regression | -0.405 | 1.231 | -4.097 | -0.575 | 8.696 | -11.416 |

**Table 5: Price differences between the actual and fitted options using index values**

Notes: This table shows summary statistics for the differences between the fitted and actual call and put option prices using index as underlying, where restricted models require the implied index value be equal to the actual index price, while unrestricted models do not have such a constraint.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Panel A. Call options** |  |  |  |  |  |
|  | Mean | Std.dev | Min | Median | Max |
| **Parametric** |  |  |  |  |  |
| Restricted BSM | 0.0443 | 0.3608 | -1.4738 | 0.0040 | 6.5900 |
| Unrestricted BSM | -0.0135 | 0.1843 | -2.4221 | -0.0165 | 3.5631 |
| Restricted Longstaff | 0.0049 | 0.1619 | -1.9153 | -0.0003 | 3.1441 |
| Unrestricted Longstaff | 0.0032 | 0.1494 | -1.9351 | 0.0028 | 3.5493 |
|  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |
| Shimko | 0.0063 | 0.1608 | -1.5746 | -0.0005 | 3.6107 |
| Partial Kernel Regression | 0.0033 | 0.1295 | -1.2900 | -0.0040 | 2.3214 |
| Full Kernel Regression | 0.0050 | 0.2206 | -1.5620 | -0.0111 | 2.9765 |
|  |  |  |  |  |  |
| **Panel B. Put options** |  |  |  |  |  |
|  | Mean | Std.dev | Min | Median | Max |
| **Parametric** |  |  |  |  |  |
| Restricted BSM | 0.0610 | 0.3261 | -2.6654 | 0.0700 | 5.5768 |
| Unrestricted BSM | 0.0438 | 0.2174 | -1.6974 | 0.0574 | 5.1668 |
| Restricted Longstaff | 0.0037 | 0.1652 | -1.7725 | -0.0008 | 5.7371 |
| Unrestricted Longstaff | 0.0046 | 0.1489 | -1.7752 | 0.0036 | 5.7010 |
|  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |
| Shimko | 0.0030 | 0.1648 | -1.7983 | 0.0018 | 5.1181 |
| Partial Kernel Regression | 0.0013 | 0.1258 | -1.4973 | 0.0000 | 3.4227 |
| Full Kernel Regression | 0.0058 | 0.2031 | -1.1708 | -0.0096 | 2.4132 |

**Table 6: Estimates for bandwidth *h* in Kernel Regressions**

Notes: This table shows estimates for bandwidth *h* by minimizing least-squares cross-validation. The Partial Kernel Regression (PKR) is three-dimensional, and Full Kernel Regression (FKR) is four dimensional. The dependent variable in PKR is implied volatility, and is option price in FKR.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Panel A. Call options** |  |  |  |  |  |
|  | Index value | Strike price | Time to maturity | Risk free rate | R2 |
| Partial Kernel Regression | 0.0867 | 2.7968 | 0.0001 | - | 0.9842 |
| Full Kernel Regression | 0.6477 | 1.239 | 0.0073 | 0.0005 | 0.9987 |
|  |  |  |  |  |  |
| **Panel B. Put options** |  |  |  |  |  |
|  | Index value | Strike price | Time to maturity | Risk free rate | R2 |
| Partial Kernel Regression | 0.1037 | 2.3675 | 0.0001 | - | 0.9861 |
| Full Kernel Regression | 0.6805 | 1.26 | 0.0069 | 0.0001 | 0.9987 |

**Table 7: Regression analysis using KOSPI 200 index values**

Notes: This table shows coefficients estimates and t-statistics for regressing the percentage differences between the implied index value and the actual index value on market frictions related independent variables. C is the regression intercept, OI is open interest, TTV is the total trading value, and N is the number of call/put options for that day. The coefficients for OI and CV are multiplied by 107 and 109, respectively.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Panel A. Call options** | | | | | |
|  | C | OI | TTV | N | R2 |
| **Parametric** |  |  |  |  |  |
| BSM | 0.2698\*\*\* | -4.0833\*\*\* | 0.0011 | 0.0280\*\*\* | 0.1895 |
| P-value | 0 | 0 | 0.1215 | 0 |  |
| Longstaff | 0.1093 | -1.3286\*\*\* | 0.0004 | 0.0029 | 0.0103 |
| P-value | 0.1253 | 0.0002 | 0.6208 | 0.5412 |  |
|  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |
| Shimko | 0.1103\*\*\* | -0.8251\*\*\* | -0.0004 | 0.0042\*\* | 0.0471 |
| P-value | 0 | 0 | 0.234 | 0.015 |  |
| Partial Kernel Regression | 0.6731\*\*\* | -2.7231\*\*\* | 0.0015 | -0.0124\*\* | 0.0334 |
| P-value | 0 | 0 | 0.1699 | 0.0294 |  |
| Full Kernel Regression | 1.8070\*\*\* | -0.6435 | -0.0056\*\* | -0.0800\*\*\* | 0.0512 |
| P-value | 0 | 0.4575 | 0.0119 | 0 |  |
|  |  |  |  |  |  |
| **Panel B. Put options** | | | | | |
|  | C | OI | TTV | N | R2 |
| **Parametric** |  |  |  |  |  |
| BSM | 1.1206\*\*\* | -8.2292\*\*\* | 0.0032\*\*\* | -0.0127\* | 0.1494 |
| P-value | 0 | 0 | 0.0046 | 0.0841 |  |
| Longstaff | 0.7580\*\*\* | -1.9581\*\*\* | 0.0004 | -0.0456\*\*\* | 0.0377 |
| P-value | 0 | 0.0013 | 0.7562 | 0 |  |
|  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |
| Shimko | 0.2154\*\*\* | -1.5113\*\*\* | 0.0007\*\* | -0.0041\* | 0.0555 |
| P-value | 0 | 0 | 0.0257 | 0.0777 |  |
| Partial Kernel Regression | 0.2663\*\*\* | -8.2164\*\*\* | 0.0057\*\*\* | 0.0316\*\*\* | 0.1892 |
| P-value | 0.0048 | 0 | 0 | 0 |  |
| Full Kernel Regression | -1.4750\*\*\* | 0.0043 | 0.0069\*\*\* | 0.0599\*\*\* | 0.0455 |
| P-value | 0 | 0.9962 | 0.0001 | 0 |  |

**Table 8: Percentage price differences between the implied and observed index futures values**

Notes: This table shows summary statistics for the relative percentage differences between the implied and actual futures value for call and put options. t-statistics is used for the null hypothesis H0 that the difference equals to zero, and the critical values are 1.6487, 1.9659 and 2.5882 for 10%, 5% and 1% confidence level for a two-side t-test.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Panel A. Call options** | | | | | |  |
|  | Mean | Std.dev | Min | Median | Max | t-stat. |
| **Parametric** |  |  |  |  |  |  |
| BSM | 0.3562 | 0.4542 | -0.2572 | 0.2433 | 3.161 | 10.7562 |
| Longstaff | 0.1266 | 0.4577 | -1.52 | 0.0682 | 3.7918 | 5.464 |
|  |  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |  |
| Shimko | 0.0621 | 0.1754 | -0.6074 | 0.0727 | 0.8632 | 8.2337 |
| Partial Kernel Regression | 0.5022 | 0.6223 | -1.4244 | 0.3587 | 3.3486 | 16.0179 |
| Full Kernel Regression | 0.2417 | 1.4665 | -4.8382 | 0.3998 | 4.7186 | 3.3091 |
|  |  |  |  |  |  |  |
| **Panel B. Put options** | | | | | |  |
|  | Mean | Std.dev | Min | Median | Max | t-stat. |
| **Parametric** |  |  |  |  |  |  |
| BSM | 0.4823 | 0.8793 | -0.6556 | 0.2634 | 6.7052 | 5.9911 |
| Longstaff | 0.0802 | 0.8991 | -2.6361 | -0.0122 | 7.3514 | 1.7614 |
|  |  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |  |
| Shimko | 0.0704 | 0.3342 | -0.7194 | 0.0721 | 5.5918 | 4.2687 |
| Partial Kernel Regression | 0.1755 | 0.8168 | -3.0236 | 0.0771 | 6.7747 | 4.2656 |
| Full Kernel Regression | -0.2319 | 1.4537 | -4.1153 | -0.4218 | 4.656 | -3.202 |

**Table 9: Price differences between the actual and fitted options using index futures values**

Notes: This table shows summary statistics for the differences between the actual and fitted call and put option prices using index futures as the underlying, where restricted models require the implied index futures value to be equal to the actual index futures price, while unrestricted models do not have such a constraint.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Panel A. Call options** | | | | | |
|  | Mean | Std.dev | Min | Median | Max |
| **Parametric** |  |  |  |  |  |
| Restricted BS | 0.0261 | 0.3538 | -1.3249 | -0.0129 | 2.841 |
| Unrestricted BS | -0.0163 | 0.1892 | -1.3342 | -0.021 | 1.8406 |
| Restricted Longstaff | -0.0003 | 0.1642 | -1.3234 | -0.0017 | 2.0589 |
| Unrestricted Longstaff | 0.0037 | 0.1511 | -1.1712 | 0.0021 | 3.1218 |
|  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |
| Shimko | -0.0004 | 0.1788 | -1.2742 | -0.0014 | 1.819 |
| Partial Kernel Regression | 0.0018 | 0.1421 | -0.9631 | -0.0081 | 2.0406 |
| Full Kernel Regression | 0.0007 | 0.1377 | -0.6892 | -0.008 | 1.1399 |
|  |  |  |  |  |  |
| **Panel B. Put options** | | | | | |
|  | Mean | Std.dev | Min | Median | Max |
| **Parametric** |  |  |  |  |  |
| Restricted BS | 0.0384 | 0.3778 | -3.2712 | 0.0665 | 5.5709 |
| Unrestricted BS | 0.0473 | 0.2444 | -1.6974 | 0.0585 | 5.1668 |
| Restricted Longstaff | -0.0031 | 0.2169 | -2.9613 | 0.0009 | 5.7385 |
| Unrestricted Longstaff | 0.0069 | 0.1785 | -1.7916 | 0.0042 | 5.701 |
|  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |
| Shimko | -0.0027 | 0.2217 | -3.2707 | 0 | 5.1196 |
| Partial Kernel Regression | -0.0054 | 0.1849 | -3.2707 | 0.002 | 4.148 |
| Full Kernel Regression | 0.0041 | 0.1684 | -0.8808 | -0.0077 | 1.651 |

**Table 10: Regression analysis using KOSPI 200 index futures values**

Notes: This table shows coefficients estimates and t-statistics for regressing the percentage differences between the implied index futures value and the actual index futures value on market frictions related independent variables. C is the regression intercept, OI is open interest, TTV is the total trading value, and N is the number of options for all call/put options for that day. The coefficients for OI and TTV are multiplied by 107 and 109, respectively.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Panel A. Call options** | | | | | |
|  | Constant | OI | TTV | N | R2 |
| **Parametric** |  |  |  |  |  |
| BSM | 0.5812\*\*\* | -4.8857\*\*\* | -0.0005 | 0.0220\*\*\* | 0.2499 |
| P-value | 0 | 0 | 0.6134 | 0.0003 |  |
| Longstaff | 0.415\*\*\* | -2.0194\*\*\* | -0.0006 | -0.0049 | 0.0417 |
| P-value | 0.0001 | 0.0002 | 0.602 | 0.4941 |  |
|  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |
| Shimko | 0.0743\* | -1.2476\*\*\* | -0.0005 | 0.0106\*\*\* | 0.1415 |
| P-value | 0.0528 | 0 | 0.2488 | 0.0001 |  |
| Partial Kernel Regression | 0.8074\*\*\* | -4.6857\*\*\* | 0.0042\*\* | -0.0012 | 0.078 |
| P-value | 0 | 0 | 0.0159 | 0.895 |  |
| Full Kernel Regression | 1.9947\*\*\* | 0.3737 | -0.0017 | -0.1365\*\*\* | 0.0889 |
| P-value | 0 | 0.8253 | 0.65 | 0 |  |
|  |  |  |  |  |  |
| **Panel B. Put options** | | | | | |
|  | Constant | OI | TTV | N | R2 |
| **Parametric** |  |  |  |  |  |
| BSM | 1.2618\*\*\* | -9.0635\*\*\* | 0.0005 | 0.0028 | 0.1754 |
| P-value | 0 | 0 | 0.7773 | 0.8091 |  |
| Longstaff | 0.9605\*\*\* | -3.6961\*\*\* | 0.0027 | -0.0494\*\*\* | 0.0414 |
| P-value | 0 | 0.0027 | 0.213 | 0.0004 |  |
|  |  |  |  |  |  |
| **Non-Parametric** |  |  |  |  |  |
| Shimko | 0.2597\*\*\* | -2.0817\*\*\* | 0.0001 | 0.0003 | 0.0568 |
| P-value | 0.0012 | 0 | 0.9391 | 0.9501 |  |
| Partial Kernel Regression | 0.003 | -8.5907\*\*\* | 0.0038\*\* | 0.0613\*\*\* | 0.2248 |
| P-value | 0.986 | 0 | 0.0356 | 0 |  |
| Full Kernel Regression | -2.3234\*\*\* | -0.1109 | 0.0080\*\* | 0.1342\*\*\* | 0.1313 |
| P-value | 0 | 0.9505 | 0.0159 | 0 |  |

**Table 11: Full regression analysis using KOSPI 200 index values**

Notes: This table shows coefficients estimates and p-values for regressing the percentage differences between the option-implied index value and the actual index value on selected independent variables. Panel A shows the results for the call options and Panel B shows the results for the put options. C is the regression intercept, T is the average time to maturity, M is the average moneyness of all call/put options for that day, OI is open interest, TTV is the total trading value for all call/put options for that day, N is the number of calls/puts for that day, ARet, ARet-1, ARet-2 are the current, first lagged, and the second lagged absolute daily returns on the KOSPI 200 index. The coefficients for OI and TTV are multiplied by 107 and 109, respectively.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Panel A. Call options** | C | T | M | OI | TTV | N | ARet | ARet-1 | ARet-2 | R2 |
| **Parametric** |  |  |  |  |  |  |  |  |  |  |
| BSM | -0.0557 | 1.412\*\* | 0.0235\*\*\* | -1.6668\*\*\* | -0.001 | 0.0188\*\*\* | 7.6462\*\*\* | 7.5615\*\*\* | 6.7698\*\*\* | 0.3142 |
| P-value | 0.3004 | 0.0205 | 0 | 0 | 0.1518 | 0 | 0 | 0 | 0 |  |
| Longstaff | -0.0306 | -3.2157\*\*\* | 0.0169\*\*\* | -0.152 | -0.0028\*\*\* | 0.0142\*\* | 4.598\*\*\* | 4.6058\*\*\* | 5.8247\*\*\* | 0.0699 |
| P-value | 0.6938 | 0.0001 | 0.0021 | 0.7172 | 0.0037 | 0.01 | 0.0006 | 0.0005 | 0 |  |
| **Non-Parametric** |  |  |  |  |  |  |  |  |  |  |
| Shimko | 0.0161 | 0.3517 | -0.0027 | -0.2302 | -0.0009\*\*\* | 0.0025 | 3.0687\*\*\* | 1.5779\*\*\* | 0.6664 | 0.0979 |
| P-value | 0.5702 | 0.244 | 0.1721 | 0.1533 | 0.0093 | 0.1972 | 0 | 0.0011 | 0.1649 |  |
| Partial Kernel Regression | 0.372\*\*\* | -0.7494 | -0.0738\*\*\* | -1.4709\*\*\* | -0.0003 | -0.0097 | 8.5099\*\*\* | 5.0884\*\*\* | 3.0906\*\* | 0.1932 |
| P-value | 0 | 0.4268 | 0 | 0.0035 | 0.7936 | 0.1165 | 0 | 0.0007 | 0.0393 |  |
| Full Kernel Regression | 1.0948\*\*\* | 14.8843\*\*\* | -0.1569\*\*\* | 2.2058\*\* | -0.0003 | -0.1515\*\*\* | 11.2229\*\*\* | 13.4349\*\*\* | 6.0493\*\* | 0.3092 |
| P-value | 0 | 0 | 0 | 0.0132 | 0.8813 | 0 | 0 | 0 | 0.0311 |  |
| **Panel B. Put options** | C | T | M | OI | TTV | N | ARet | ARet-1 | ARet-2 | R2 |
| **Parametric** |  |  |  |  |  |  |  |  |  |  |
| BSM | 0.2081 | 9.1112\*\*\* | -0.0842\*\*\* | -5.7593\*\*\* | 0.0034\*\*\* | -0.0219\*\*\* | 13.3387\*\*\* | 17.7085\*\*\* | 15.2428\*\*\* | 0.3545 |
| P-value | 0.0521 | 0 | 0 | 0 | 0.0019 | 0 | 0 | 0 | 0 |  |
| Longstaff | 0.4463\*\*\* | 2.9349\*\* | -0.0597\*\*\* | -2.6956\*\*\* | 0.0019 | -0.0271\*\*\* | 1.0891 | 5.6603\*\* | 8.0123\*\*\* | 0.0596 |
| P-value | 0.0005 | 0.0205 | 0 | 0.0003 | 0.1479 | 0 | 0.6303 | 0.011 | 0.0003 |  |
| **Non-Parametric** |  |  |  |  |  |  |  |  |  |  |
| Shimko | 0.0215 | 1.4362\*\*\* | 0.0062\*\* | -0.1843 | -0.0005 | -0.0031\*\* | 4.1405\*\*\* | 3.5259\*\*\* | 2.012\*\*\* | 0.1574 |
| P-value | 0.5763 | 0 | 0.0233 | 0.3955 | 0.1357 | 0.0274 | 0 | 0 | 0.0009 |  |
| Partial Kernel Regression | -0.2801\*\*\* | 0.8276 | 0.1023\*\*\* | -2.2348\*\*\* | -0.0013 | 0.0244\*\*\* | 8.4811\*\*\* | 9.5508\*\*\* | 6.5553\*\*\* | 0.3979 |
| P-value | 0.0046 | 0.3696 | 0 | 0 | 0.1428 | 0 | 0 | 0 | 0 |  |
| Full Kernel Regression | -1.0861\*\*\* | -14.8144\*\*\* | 0.2305\*\*\* | 4.4079\*\*\* | -0.0034\*\* | 0.0668\*\*\* | -1.0583 | -7.0152\*\* | -7.8367\*\*\* | 0.2744 |
| P-value | 0 | 0 | 0 | 0 | 0.04 | 0 | 0.7124 | 0.0119 | 0.0048 |  |

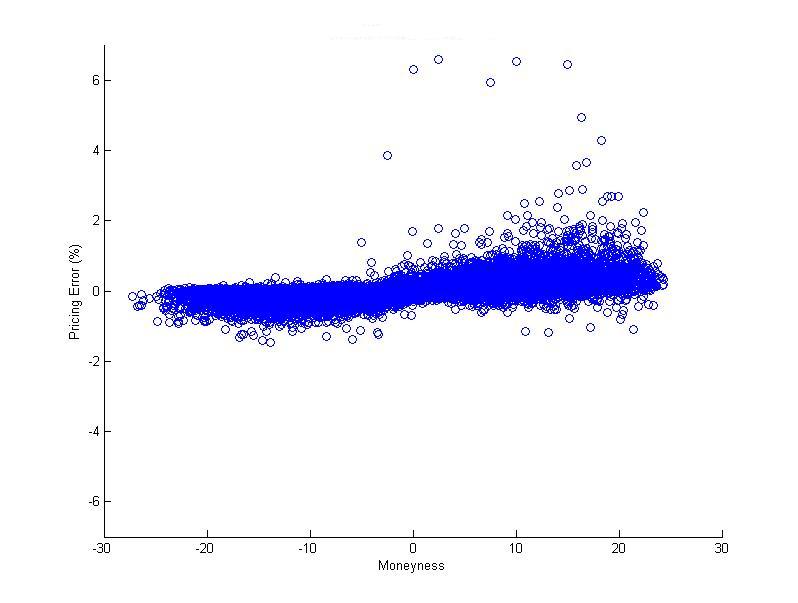
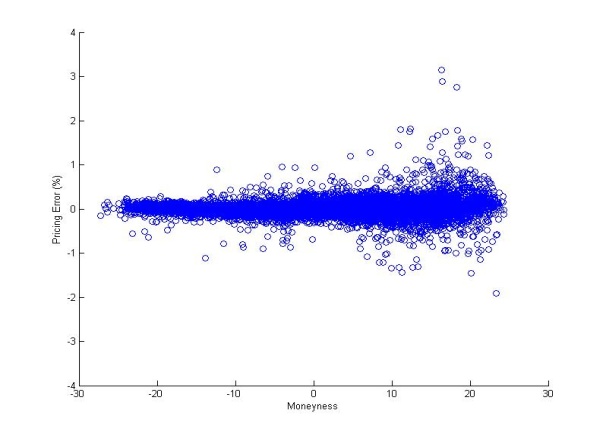
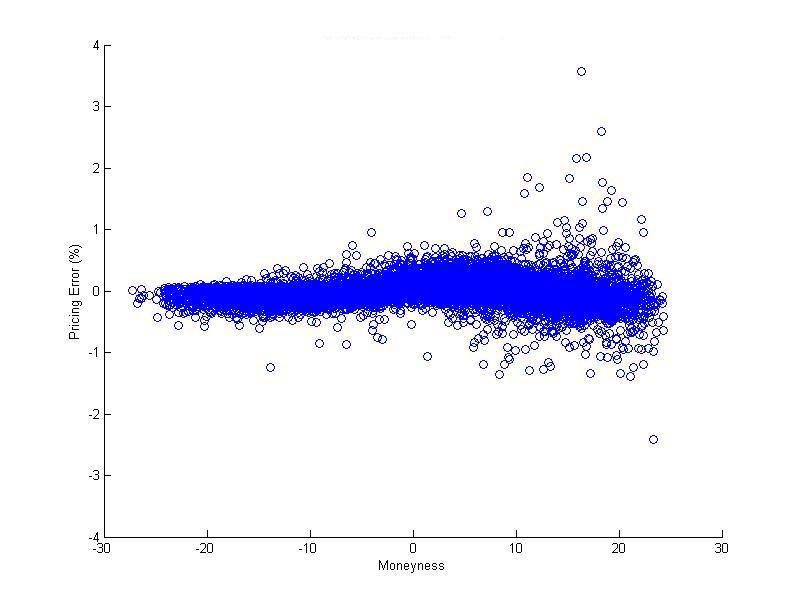
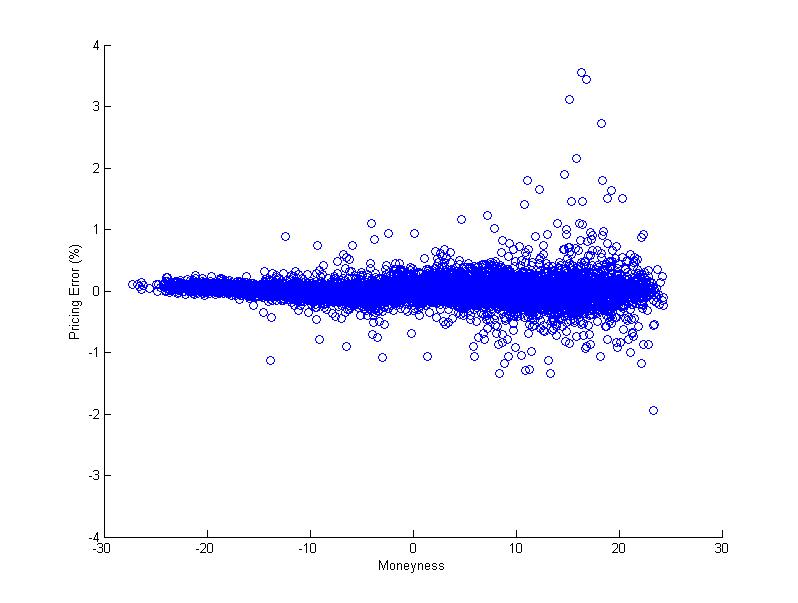
**Table 12: Regression analysis using KOSPI 200 index futures values excluding maturity less than three days**

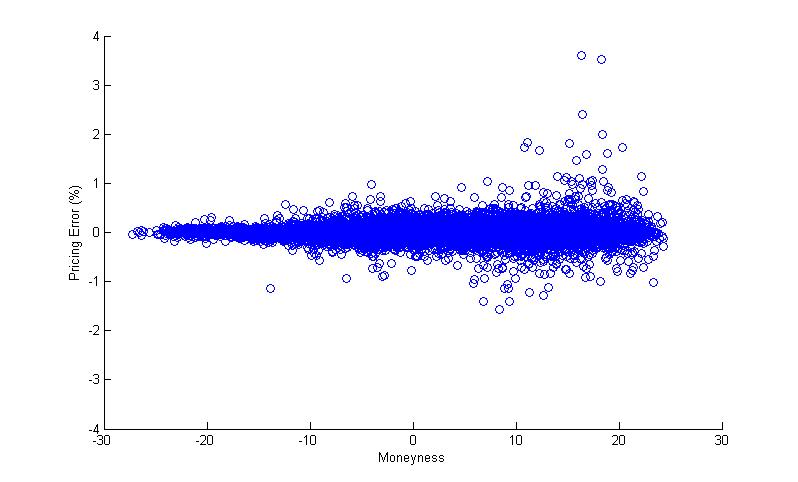
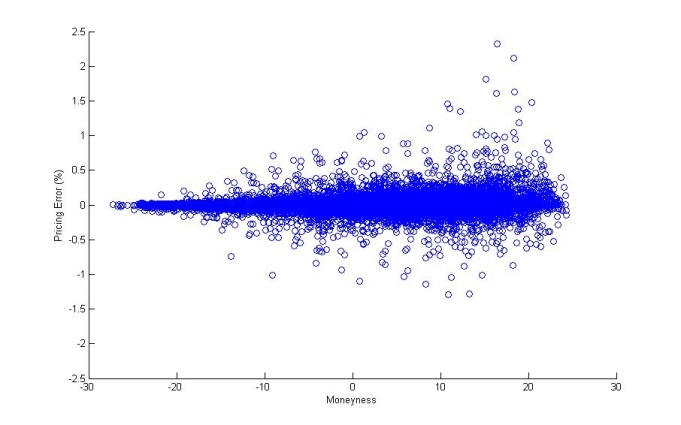
Notes: This table shows coefficients estimates and t-statistics for regressing the percentage differences between the implied index futures value and the actual index futures value on market frictions related independent variables, all options with maturity less than three day are excluded. C is the regression intercept, OI is open interest, TTV is the total trading value, and N is the number of options for all call/put options for that day. The coefficients for OI and TTV are multiplied by 107 and 109, respectively.

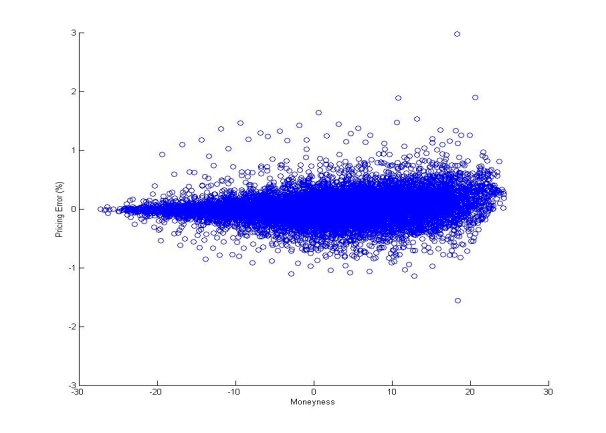
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Panel A. Call options** | | | | | | | |
|  | Constant | OI | | TTV | | NumO | R2 |
| **Parametric** |  |  | |  | |  |  |
| BSM | 0.9618\*\*\* | -6.4978\*\*\* | | 0.0035\*\*\* | | -0.0047 | 0.3121 |
| P-value | 0.0000 | 0.0000 | | 0.0081 | | 0.5424 |  |
| Longstaff | 0.5698\*\*\* | -2.7481\*\*\* | | 0.0010 | | -0.0154\* | 0.0571 |
| P-value | 0.0000 | 0.0000 | | 0.5158 | | 0.0939 |  |
| **Non-Parametric** |  |  | |  | |  |  |
| Shimko | 0.1998\*\*\* | -1.5259\*\*\* | | 0.0006\* | | 0.0007 | 0.2201 |
| P-value | 0.0000 | 0.0000 | | 0.0816 | | 0.7490 |  |
| Partial Kernel Regression | 1.2411\*\*\* | -5.8653\*\*\* | | 0.0082\*\*\* | | -0.0337\*\*\* | 0.1287 |
| P-value | 0.0000 | 0.0000 | | 0.0000 | | 0.0042 |  |
| Full Kernel Regression | 3.5326\*\*\* | -6.0033\*\*\* | | 0.0157\*\*\* | | -0.2471\*\*\* | 0.2152 |
| P-value | 0.0000 | 0.0012 | | 0.0004 | | 0.0000 |  |
|  |  |  | |  | |  |  |
|  |  |  | |  | |  |  |
| **Panel B. Put options** | | | | | | | |
|  | Constant | | OI | | TTV | NumO | R2 |
| **Parametric** |  | |  | |  |  |  |
| BSM | 2.5766\*\*\* | | -13.6914\*\*\* | | 0.0076\*\*\* | -0.0779\*\*\* | 0.2940 |
| P-value | 0.0000 | | 0.0000 | | 0.0006 | 0.0000 |  |
| Longstaff | 1.7564\*\*\* | | -6.1635\*\*\* | | 0.0069\*\*\* | -0.1001\*\*\* | 0.0904 |
| P-value | 0.0000 | | 0.0000 | | 0.0074 | 0.0000 |  |
| **Non-Parametric** |  | |  | |  |  |  |
| Shimko | 0.5778\*\*\* | | -2.8356\*\*\* | | 0.0011 | -0.0194\*\*\* | 0.1035 |
| P-value | 0.0000 | | 0.0000 | | 0.1815 | 0.0022 |  |
| Partial Kernel Regression | 0.4497\*\* | | -10.2531\*\*\* | | 0.0061\*\*\* | 0.0349\*\* | 0.2245 |
| P-value | 0.0487 | | 0.0000 | | 0.0036 | 0.0194 |  |
| Full Kernel Regression | -3.9973\*\*\* | | 3.2923\* | | -0.0013 | 0.2517\*\*\* | 0.2493 |
| P-value | 0.0000 | | 0.0845 | | 0.6907 | 0.0000 |  |

**Figure 1: Pricing errors across moneyness for call options using KOSPI 200 index values**

Notes: This figure presents differences between the fitted and actual call option prices using index as underlying, where x-axis is moneyness, defined as index value minus strike price, y-axis is pricing error in percentage. The plots from top left to bottom right are for restricted Black Scholes, unrestricted Black Scholes, restricted Longstaff, unrestricted Longstaff, Shimko, Partial Kernel Regression and Full Kernel Regression, respectively.

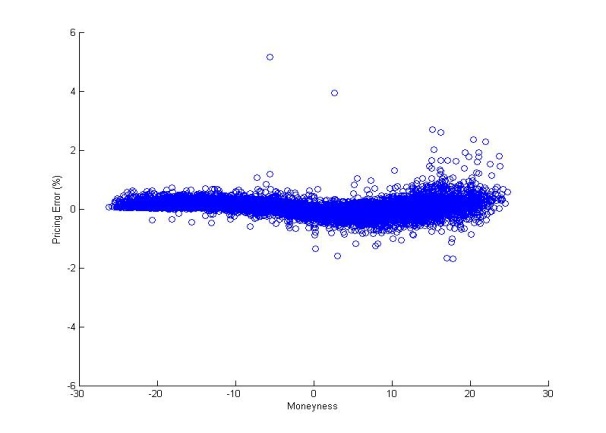
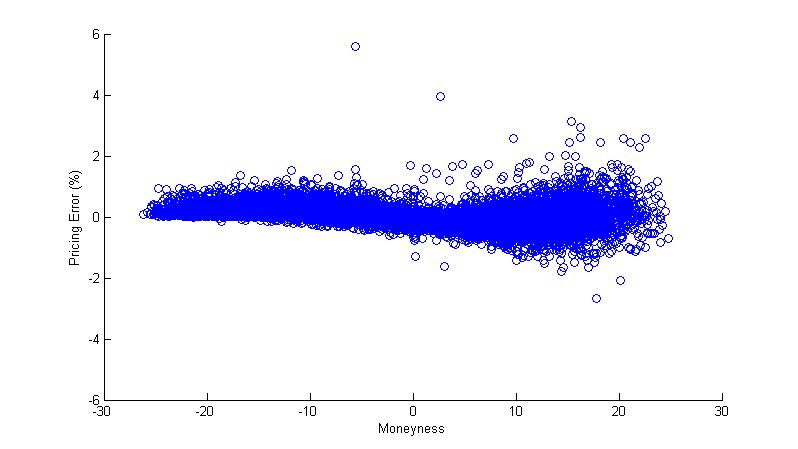
  

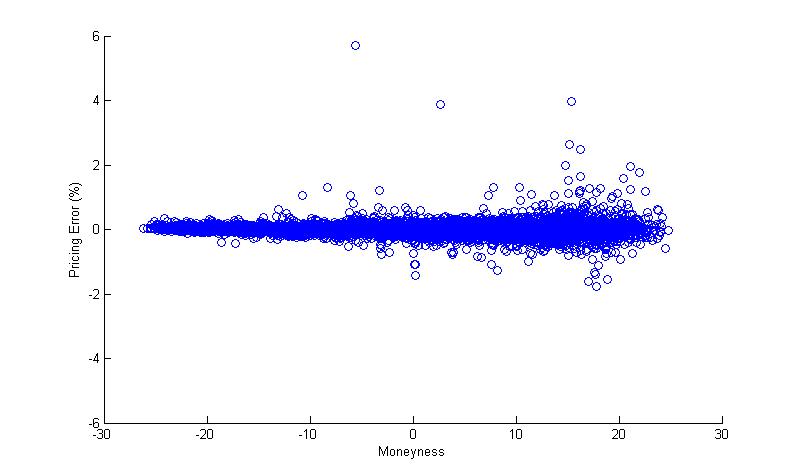
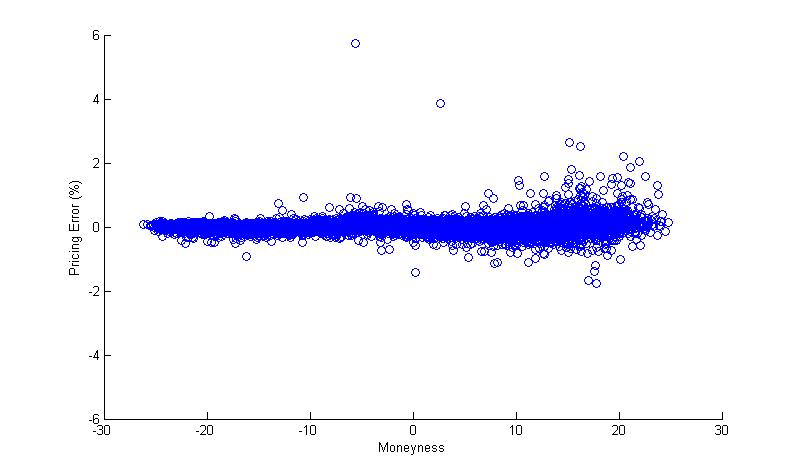
 

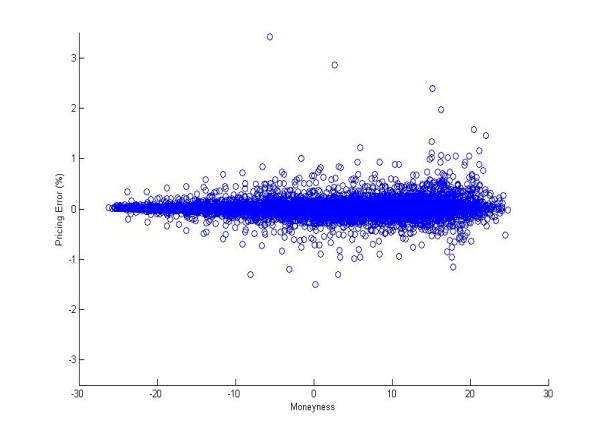
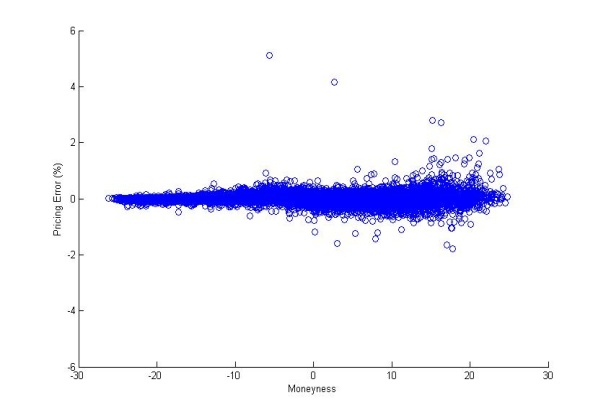


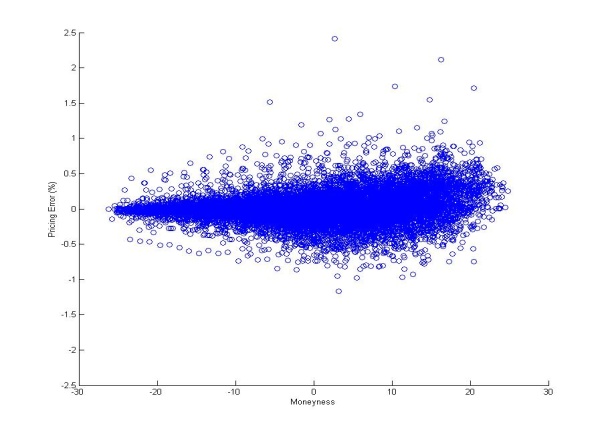
**Figure 2: Pricing errors across moneyness for put options using KOSPI 200 index values**

Notes: This figure presents differences between the fitted and actual put option prices using index as underlying, where x-axis is moneyness, defined as strike price minus index value, y-axis is pricing error in percentage. The plots from top left to bottom right are for restricted Black Scholes, unrestricted Black Scholes, restricted Longstaff, unrestricted Longstaff, Shimko, Partial Kernel Regression and Full Kernel Regression, respectively.



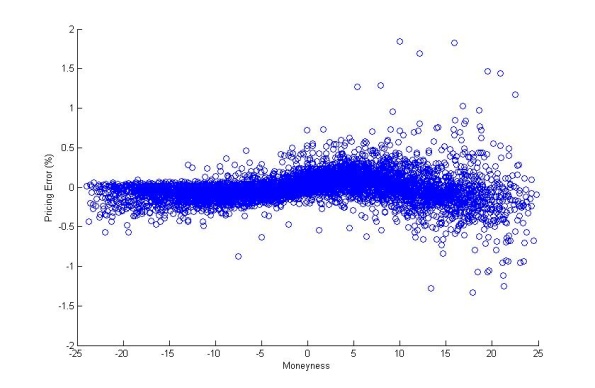
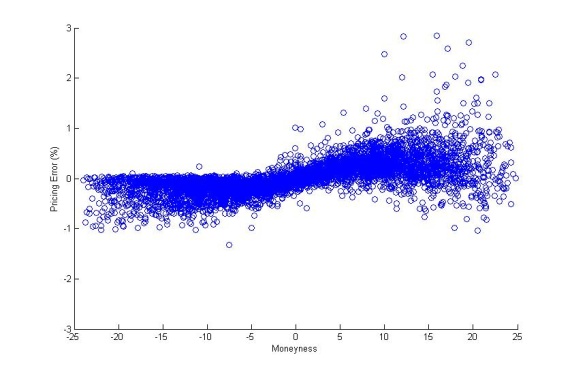


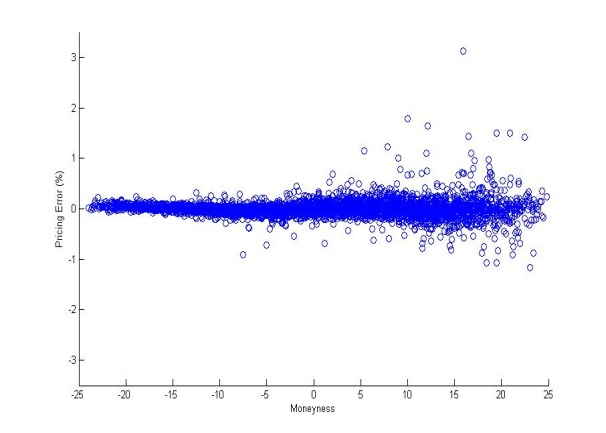
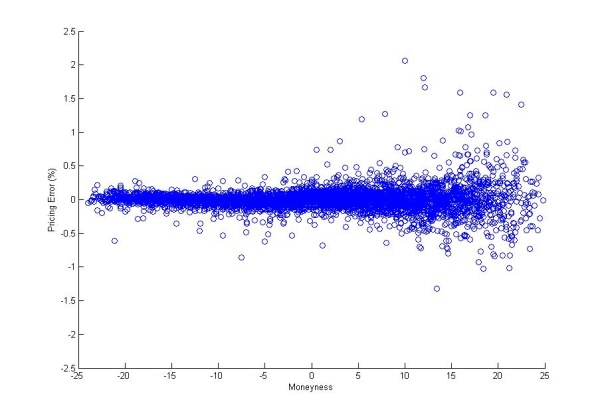


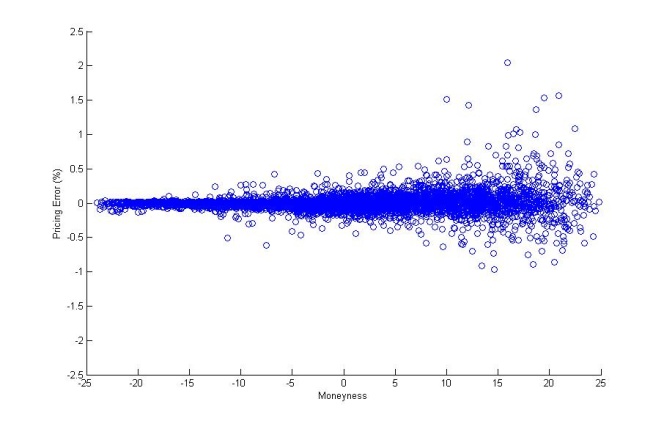
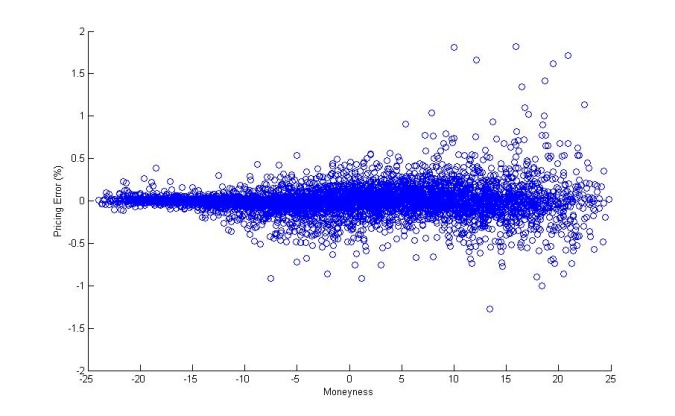


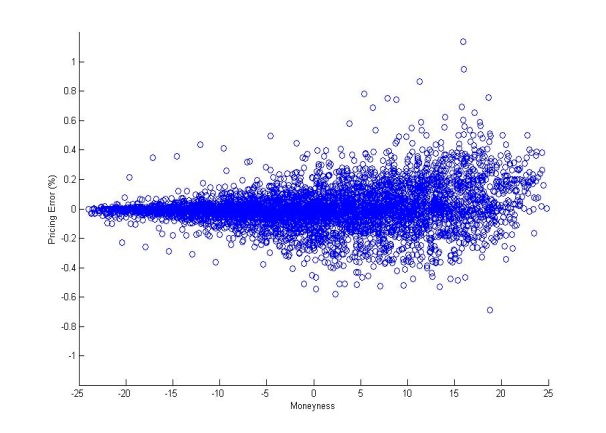
**Figure 3: Pricing errors across moneyness for call options using KOSPI 200 index futures value**

This Figure shows the differences between the fitted and actual call option prices using index futures as underlying, where x-axis is moneyness, defined as index futures value minus strike price, y-axis is pricing error in percentage. The plots from top left to bottom right are for restricted Black Scholes, unrestricted Black Scholes, restricted Longstaff, unrestricted Longstaff, Shimko, Partial Kernel Regression and Full Kernel Regression, respectively.



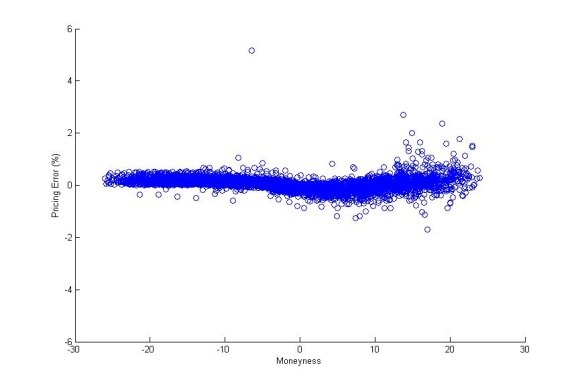
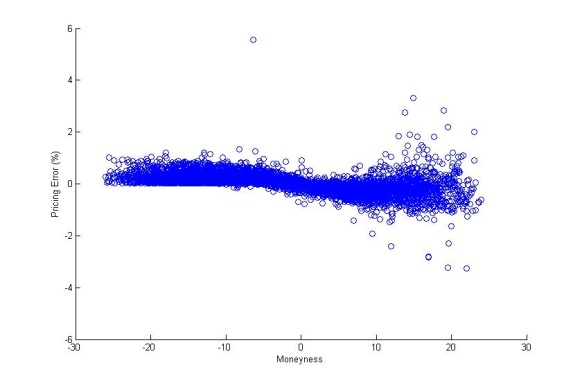


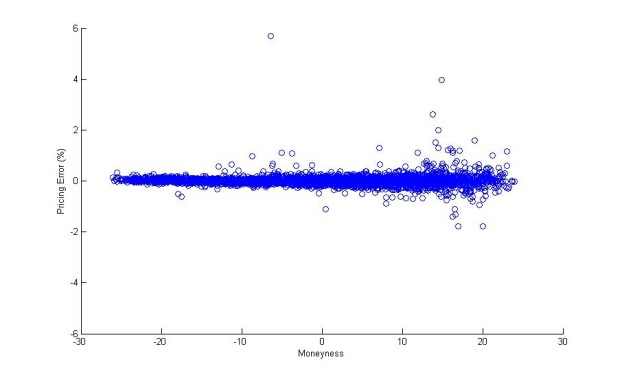
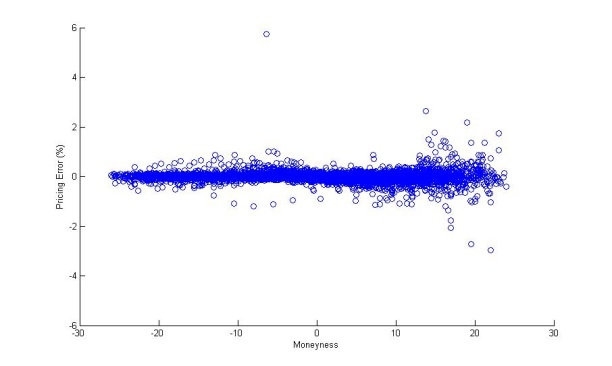


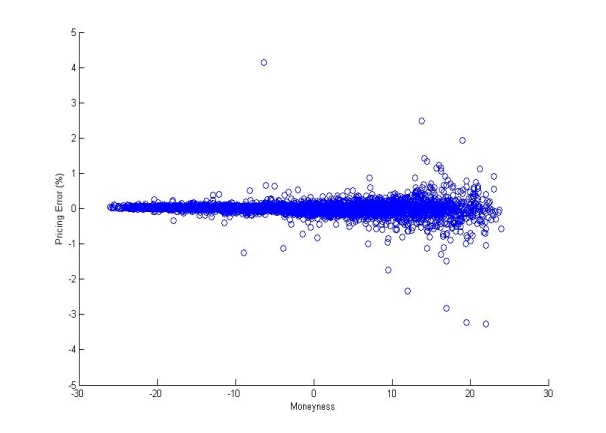
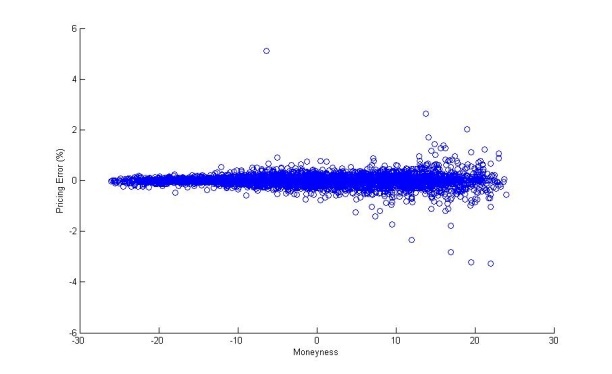


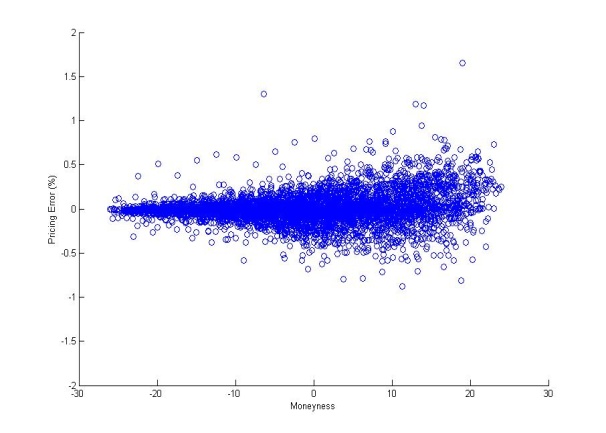
**Figure 4: Pricing errors across moneyness for put options using KOSPI 200 index futures values**

Notes: This Figure shows differences between the fitted and actual put option prices using index futures as underlying, where x-axis is moneyness, defined as strike price minus index futures value, y-axis is pricing error in percentage. The plots from top left to bottom right are for restricted Black Scholes, unrestricted Black Scholes, restricted Longstaff, unrestricted Longstaff, Shimko, Partial Kernel Regression and Full Kernel Regression, respectively.









1. For more detailed reviews, refer to Bahra (1997), and Figlewski (2008). [↑](#footnote-ref-1)
2. In this paper *St* always denotes dividend adjusted price. [↑](#footnote-ref-2)
3. We suppress the time variable *t* for easier notation. [↑](#footnote-ref-3)
4. See, for example, Stone (1984). [↑](#footnote-ref-4)
5. This method is noted as Full Kernel Regression (FKR) in this study to differentiate it from a later Partial Kernel Regression (PKR) method. [↑](#footnote-ref-5)
6. This method is noted as Partial Kernel Regression (PKR) in this study. [↑](#footnote-ref-6)
7. Investment companies who provide the Home Trading System (HTS) to the individual investors usually require very small commission fees. Recently, many companies start to waive such commission fees which were charged to the individual investors. [↑](#footnote-ref-7)
8. The Korean options market adopts two different tick sizes based on the options prices. If the option price is more than 3 points, which equal to 300,000 Korean Won (KRW), the minimum tick size is 0.05 points (i.e. KRW 5,000) while the tick size is 0.01 points (i.e. KRW 1,000) otherwise. [↑](#footnote-ref-8)
9. Since the choice of the kernel function has little effect on the result, and the choice of the bandwidth *h* is crucial, we estimate *h* by minimizing the least-squares cross-validation. Table 6 lists those parameters for both regressions on calls and puts. The high R2 value demonstrates the accuracy of our Kernel Regression methods. [↑](#footnote-ref-9)
10. We don't add the bid-ask spreads as a proxy for option transaction cost because the spreads in our sample are too small to carry useful information. The inclusion of the bid-ask spreads generally performs worse due to the singular problem in regression. [↑](#footnote-ref-10)
11. The average present dividend value is only 0.130 point in our whole sample, while the average value of KOSPI 200 index price is 201 points. [↑](#footnote-ref-11)
12. We compare the pricing performances across moneyness for all methods by dividing the calls (puts) into ATM, ITM and OTM groups. An upward (downward) bias is shown for calls (puts). [↑](#footnote-ref-12)
13. We also did a regression using index futures values. The results are similar so omitted here to save space, but they are available upon requests. [↑](#footnote-ref-13)
14. We run a separate regression analysis including additional control variables as in section 6.2 and get very similar conclusions. [↑](#footnote-ref-14)