On the Stability of Implied Probability Density Functions: Empirical Evidence Using Alternative Measures

Byung Jin Kang[*]
Graduate School of Management, Korea Advanced Institute of Science and Technology
Phone: 82-2-958-3689
E-mail address: bjk@kaist.ac.kr

Tong Suk Kim
Graduate School of Management, Korea Advanced Institute of Science and Technology
Phone: 82-2-958-3018
E-mail address: tskim@kgsm.kaist.ac.kr

March, 2004

[*] Corresponding author, Byung Jin Kang, Korea Advanced Institute of Science and Technology (KAIST) Graduate School of Management, 207-43 Cheongnyangni 2-dong, Dongdaemun-ku, Seoul, 130-722, Korea.
Tel : 82-2-958-3689, Fax : 82-2-958-3618, E-mail address : bjk@kaist.ac.kr

On the Stability of Implied Probability Density Functions: Empirical Evidence Using Alternative Measures

Abstract

In spite of their rich applicability, little attention has been paid to the stability of probability density functions (PDFs) implied by option prices. This paper examines the stability of the two methods that are most widely used for estimating implied PDFs: the double lognormal approximating function (DLN) method and the smoothed implied volatility smile (SMIV) method. Our study differs from previous ones in three ways. First, to test the stability between the PDFs in each method, we focus not only on their distributional characteristics, but also on the empirical results of their applications, namely, pricing illiquid options and recovering the risk aversion of representative agents. Second, we examine the sensitivity of the implied PDFs to errors that can be embedded in option prices to increase the validity of our results. Finally, our analysis is carried out using the KOSPI 200 index option in the Korean stock market, which is one of the most actively traded markets in the world.

JEL classification: G13; C13; C15

Keywords: Implied probability density functions; Stability of estimates; Measurement errors; KOSPI 200 index options

1. Introduction

Since Breeden and Litzenberger's seminal paper was published in 1978, there have been numerous researches on methods for recovering implied PDFs from the cross-sectional data of option prices. As Jackwerth (1999) mentioned, these methods of estimating implied PDFs can be classified as either parametric or non-parametric. The parametric approach can be further divided into expansion methods, generalized distribution methods, and mixture methods. The non-parametric approach can also be divided into kernel methods, maximum entropy methods, and curve fitting methods. In spite of remarkable developments in estimation methods, we cannot, however, provide statistical properties for estimated parameters. Therefore, even if a new stylized fact was discovered using implied PDFs, we could not help questioning its reliability.

Recently, Söderlind and Svensson (1997), and Melick and Thomas (1998) assumed that the estimated parameters follows multivariate normal and then derived a confidence interval for the implied PDFs. Söderlind (1999), Cooper (1999), Giamouridis and Tamvakis (2001), and Bliss and Panigirtzoglou (2002) also constructed the pseudo prices by randomly perturbing original prices and tested the stability of the implied PDFs. As noted by Bliss and Panigirtzoglou (2002), however, the former approach is dependent on the parametric assumption and does not provide the confidence intervals for possible PDFs. The latter approach also has at least two limitations. The first is that it only provides a lower bound on the confidence intervals of the summary statistics derived from estimated PDFs.[1] The second, and perhaps more critical in our view, is that it only considers the summary statistics of estimated PDFs to test their stability.

In this study, we examine the stability of the two methods that are most widely used to estimate implied PDFs: the double lognormal approximating function (DLN) method and the smoothed implied volatility smile (SMIV) method. Our study differs from previous ones in three main ways.

First, to test the stability between the PDFs derived by the DLN and the SMIV methods, we focus not only on their distributional characteristics, but also on the empirical results of their applications. Even though the summary statistics derived from estimated PDFs may characterize them roughly, the stability of estimated PDFs is not equivalent to the stability of the summary statistics of them. In this sense, to test the stability of implied PDFs we need to examine the more general characteristics, including the summary statistics. In our study, besides the stability of the summary statistics, we also examine the stability of the derivatives of estimated PDFs and of the option pricing

[1] Except for Söderlind (1999), they use the prices randomly perturbed by no more than a half tick. Thus, the error band they derive should be a lower bound on the confidence intervals.

errors between the fitted prices using one-day-ahead implied PDFs and market prices. These measures are related to recovering implied risk aversion and pricing other options, respectively, which are main applications of implied PDFs. As Jackwerth (2000), Aït Sahalia and Lo (2000), and Zigler (2003) noted, the implied risk aversion of representative agents on the asset return states can be derived from the implied PDFs. In this application, the first derivative of the implied PDFs has an effect on the shape of implied risk aversion. Another application of implied PDFs is to price other options. If an implied PDF is recovered from the liquid option prices, we can price illiquid options using this PDF. As there is no illiquid option price data to test in our case, they are applied to forecasting a tomorrow's option price.

Second, we examine the sensitivity of the implied PDFs to the errors that can be embedded in option prices, to increase the validity of our results. To investigate the sensitivity, a price-perturbation method, which was introduced by Cooper (1999), Giamouruids and Tamvakis (2001), and Bliss and Panigirtzoglou (2002), is employed. They slightly (by no more than a half tick) perturb option prices and then re-estimate a PDF for the perturbed data. As pointed out by Bliss and Panigirtzoglou (2002), a confidence region for the estimates can be provided by the distribution of simulated PDFs for perturbed data. The effect of local price changes on the implied PDFs can also be investigated using this method. However, as they only consider cases in which there is a half tick error, the confidence region they derive is just the lower bound on itself. Their tests on the effect of local price changes are also restricted to marginal price changes. In this paper, we extend their scheme to the more generally perturbed data by varying the size of errors in option prices. This extended price-perturbation method enables us to investigate more generally the effect of price changes on the implied PDFs. It also enables us to examine the sensitivity of the implied PDFs to possible measurement errors.

Finally, our analysis differs from others in that it is carried out using the KOSPI 200 index option in the Korean stock market. KOSPI 200 index options market has become one of the most actively traded markets in the world, despite its short history. Its trading volume reached 1.9 billion contracts in 2002 and ranked the 1st in the world from 1999 to 2001. Not like other big markets, near contracts and OTM (ITM) call (put) options are most liquid in this market. Focusing on this market not only enables us to investigate a rapidly-growing market, but also allows us to examine properties specific to the Korean stock market.[2] Furthermore, compared to call options, put options are consistently overvalued in KOSPI 200 index options, so the implied volatility from put

[2] An example of these properties is the excellent liquidity in the near contract. As noted by Bakshi, Cao and Chen (1997), the volatility smiles tend to be the strongest for short-term options, indicating that short-term options can be the most severely mis-priced by the Black and Scholes (1973) model.

options is much higher than that from call options. The implied PDFs recovered from the data, including both of call and put options, are highly unstable and inefficient in terms of goodness-of-fit (GOF). Therefore, to obtain well-behaved implied PDFs, we have to estimate them from only call options or put options.[3]

Our results show that the stability of implied PDFs is assessed differently according to an employed test statistic. In terms of summary statistics, the implied PDF by the SMIV method is found to be more stable than that by the DLN method. The implied PDF by the DLN method, however, is found to be more stable than that by the SMIV method, in terms of its first derivatives. Regarding pricing errors, the difference between the two methods is found to be negligible. The price-perturbation method also assesses the effect of local price changes on the implied PDFs. Unlike the SMIV method, the DLN method is found to alter the shape of the entire distribution with local price changes, rather than only a small part of it.

The remainder of this paper is structured as follows. Section 2 presents an overview for the estimation methods. Section 3 presents various measures to test the stability of implied PDFs. Section 4 describes the data for empirical analyses. Section 5 provides the empirical results. The last Section concludes with a brief summary.

2. Estimation Methods

In this section, we briefly examine two estimation methods that are used in our study. The first is the double log-normal approximating function (DLN) method and the second is the smoothed implied volatility smile (SMIV) method.

2-1. Double Lognormal Approximating Function (DLN) Method

In the double log-normal approximating function (DLN) method, we assume that underlying asset price density, $q(S_T)$, is a weighted sum of 2-component log-normal density functions :

$$q(S_T) \quad = \quad \sum_{i=1}^{2}[\theta_i L(\alpha_i, \beta_i; S_T)], \tag{1}$$

where $L(\alpha_i, \beta_i; S_T)$ is the i^{th} lognormal density function in the double mixture with parameters α_i and β_i;

$$\alpha \quad = \quad \ln S + (\mu_i - \frac{1}{2}\sigma_i^2)\tau \text{ and } \beta_i \quad = \quad \sigma_i\sqrt{\tau}\ , \ i \ = 1, 2 \tag{2}$$

[3] For example, the average of call option-implied volatilities during the year 2001 was lower by 5.39% than that of put option-implied volatilities. Explaining this anomaly is not within the scope of this paper.

$$\sum_{i=1}^{2}\theta_i \; = \; 1 \text{ and } \theta_i > 0 . \tag{3}$$

Bahra (1997) derived closed-form solutions of the European call and put option prices as follows :

$$C(K,\tau)=e^{-r\tau}[\theta\{e^{\alpha_1+\frac{1}{2}\beta_1^2}N(d_1)-KN(d_2)\}+(1-\theta)\{e^{\alpha_2+\frac{1}{2}\beta_2^2}N(d_3)-KN(d_4)\}], \tag{4}$$

$$P(K,\tau)=e^{-r\tau}[\theta\{-e^{\alpha_1+\frac{1}{2}\beta_1^2}N(-d_1)+KN(-d_2)\}+(1-\theta)\{-e^{\alpha_2+\frac{1}{2}\beta_2^2}N(-d_3)+KN(-d_4)\}], \tag{5}$$

where $d \ = \dfrac{-\ln K+\alpha \ +\beta_1^2}{\beta_1}$, $d_2=d_1-\beta_1$, $d_3=\dfrac{-\ln K+\alpha_2+\beta_2^2}{\beta_2}$, $d_4=d_3-\beta_2$.

From equation (4) and (5), we can estimate parameters of the density function by solving following optimization problem:

$$\min_{\alpha_1,\alpha_2,\beta_1,\beta_2,\theta}\sum_{i=1}^{n}[C(K_i,\tau)-\hat{C}_i]^2+\sum_{i=1}^{2}[P(K_i,\tau)-\hat{P}_i]^2+[\theta e^{\alpha_1+\frac{1}{2}\beta_1^2}+(1-\theta)e^{\alpha_2+\frac{1}{2}\beta_2^2}-e^{r\tau}S]^2 . \tag{6}$$

In the above equation (6), the first two terms are just the sum of squared pricing errors between model prices and market prices. In the absence of arbitrage opportunity, the mean of the implied density function should be equal to the forward price of the underlying asset. This property is named a mean-forward price equality condition and is represented by the third term. In this sense, Bahra (1997), and Anagnou, Bedendo, Hodges and Tompkins (2002) include this term using the futures price as a proxy for the forward price in the optimization procedure. On the other hand, Bliss and Panigirtzoglou (2002), and Kim and Kim (2003) do not impose this constraint based on the assertion that it is not required by the mathematics underlying the DLN method. Considering this condition as a product of related, but separate arbitrage arguments, they just use it to see how well the underlying no-arbitrage conditions hold.

As shown by above equations, the DLN method has a computational advantage because of the existence of analytic solutions. Moreover, it is not difficult to estimate the density function with a small data set. DLN method, however, is in lack of flexibility in fitting the implied PDFs when available cross-sectional data of options are enough for estimation.[4] This drawback is due to a pre-defined structure of the DLN method on the density function.

2-2. Smoothed Implied Volatility Smile (SMIV) Method

Smoothed implied volatility smile (SMIV) method, originally introduced by Shimko

---

[4] See Brunner and Hafner (2003)

(1993), fits a function through observed implied volatilities.[5] Fitted implied volatility function is translated into an option price function and then the implied PDFs are recovered using the Breeden and Litzenberger's result:

$$q(S_T) = e^{r\tau} \frac{\partial^2 C(K,\tau)}{\partial K^2}. \qquad (7)$$

As noted by Bahra (1997), Syrdal (2002), and Brunner and Hafner (2003), two related issues have to be considered in the SMIV method. First, we cannot know the implied volatility function beyond the range of traded strike prices. Therefore, we must extrapolate or model the tails of distributions to obtain a well-defined PDF. Second, the estimated implied volatility function should not allow for arbitrage. To guarantee this, implied probabilities should be non-negative, integrate to one and satisfy the martingale restriction, which means that discounted asset prices have to be martingales with respect to these probabilities.

Shimko (1993) assumed that implied volatility is a quadratic function of the strike price, which is equivalent to the assumption of the lognormally distributed tails. His method cannot, however, always satisfy the martingale restriction. Thus, Malz (1997) assumed that implied volatility is a quadratic function with respect to the option's delta. The advantage of the fitting in the delta space is that corresponding implied density integrates to one and satisfies the martingale restriction. Campa et al (1998), and Bliss and Panigirtzoglou (2002), furthermore, used a smoothing cubic spline method to fit the implied volatilities. Campa et al (1998) fitted the implied volatilities as a function of strike prices, while Bliss and Panigirtzoglou (2002) fitted them with respect to option's delta.

In this study, we use the cubic spline method on option's delta spaces. We also extend the spline function outside the range of option data using the first and last polynomial of it.[6] Thus, our method can be constructed as the following optimization problem.[7] :

$$\underset{\Psi}{Min} \quad [\lambda \sum_{i=1}^{n} \varpi_i [y_i - f(x_i;\Psi)]^2 + (1-\lambda)\int f''(x;\Psi)^2 dx], \qquad (8)$$

where x_i and y_i are option deltas and implied volatilities, respectively, $f(x;\Psi)$ is a spline function, Ψ is a parameter matrix of the spline function, ϖ_i is a weighting parameter for observation i and λ is the smoothing parameter.

In equation (8), the first part relates to the GOF of the spline function. The second part

[5] As pointed out by Bates (1991), option prices are not suitable for fitting because of their substantial amount of non-linearity

[6] Bliss and Panigirtzoglou (2002) assumed that the spline function is linear outside the range of observation, while Shimko(1993) fitted lognormal distributions at the tails of the implied density function.

[7] See Syrdal (2002)

relates to the smoothness of the spline function. λ known as the smoothing parameter determines the relative weight between two parts. A large value of λ increases the GOF and a resulting density function may show too much oscillation. With a low value of λ, by contrast, a spline function is too smooth and does not fit the data well. Extremely, the spline function would accurately interpolate the observed data with λ equal to one and minimize the curvature with λ equal to zero.

In conclusion, the SMIV method does not impose any pre-defined structure on a density function. It, especially when there are enough observations available, has more flexibility in fitting the implied PDFs than the DLN method. However, the price to be paid is that the SMIV method may be ineffective if only a few cross-sectional data of options are available. Furthermore, the SMIV method cannot guarantee the non-negativity of estimated density functions.

3. Stability Measures

Besides the summary statistics for estimated PDFs, which were most frequently used in previous researches, other measures could be used to characterize the distributions. In this section, we briefly present measures related to summary statistics first and then discuss those related to other test statistics in detail.

3-1. Measures related to Summary Statistics

Since there is no definite way to compare distributions themselves, our empirical results are assessed, in the first place, based on the summary statistics (denoted by Z) of estimated PDFs. As frequently used in most previous researches, we accept the followings:

- $\hat{\mu}$: The sample mean or the first central sample moment of the distribution.
- $\hat{\sigma}$: The sample standard deviation or the second central sample moment of the distribution.
- $\hat{\delta}_1$: The sample skewness coefficients, defined as the third central sample moment of the distribution normalized by the cube of the standard deviation:

$$\hat{\delta}_1 = \frac{\hat{m}^3}{\hat{\sigma}^3}. \tag{9}$$

- $\hat{\delta}_2$: The Pearson mode-based skewness measure, defined as the difference between the sample mean and mode normalized by the standard deviation:

$$\hat{\delta}_2 = \frac{\hat{\mu} - \mathrm{mode}}{\hat{\sigma}}. \tag{10}$$

- $\hat{\delta}_3$: The Pearson median-based skewness measure, defined as the difference between the sample mean and median normalized by the standard deviation:

$$\hat{\delta}_3 = \frac{\hat{\mu} - median}{\hat{\sigma}}. \tag{11}$$

- $\hat{\delta}_4$: A measure of asymmetry, defined as

$$\hat{\delta}_4 = \frac{\hat{X}_{75} - \hat{X}_{50}}{\hat{X}_{50} - \hat{X}_{25}}, \tag{12}$$

where $\hat{X}_n$ is the n^{th} percentile of the estimated PDF.

- $\hat{\xi}$: The sample kurtosis coefficient, defined as the fourth central sample moment normalized by the square of the sample variance:

$$\hat{\xi} = \frac{\hat{m}^4}{\hat{\sigma}^4}. \tag{13}$$

After these summary statistics are derived from the implied PDFs, we test the stability of them from two points of view. First, the stability of the implied PDFs during the entire sample period can be assessed by the sample standard deviation of summary statistics, defined as

$$\sqrt{\frac{1}{T}\sum_{t}^{T}(Z_t - \bar{Z})^2}. \tag{14}$$

Second, the stability of daily changes of the implied PDFs can be assesed by the sample mean and the sample standard deviation of the absolute daily change in summary statistics, defined as

$$\frac{1}{T-1}\sum_{t}^{T-1}\left|Z_{t+1} - Z_t\right| \tag{15}$$

and

$$\sqrt{\frac{1}{T-1}\sum_{t}^{T-1}\left(\left|Z_{t+1} - Z_t\right| - \overline{\left|Z_{t+1} - Z_t\right|}\right)^2}. \tag{16}$$

3-2. Measures related to the Empirical Risk Aversion

As was discussed in the introduction, a risk aversion function of representative agents can be derived from option prices.[8] Under the complete market economy which implies the existence of a representative agent,[9] an expression for the absolute risk aversion can be written in terms of the subjective and implied PDFs:

$$ARA = -\frac{U''(S)}{U'(S)} = \frac{p'(S)}{p(S)} - \frac{q'(S)}{q(S)}, \tag{17}$$

[8] See Leland (1980), Jackwerth (2000), Aït-Sahalia and Lo (2000), Rosenberg and Engle (2002), Bliss and Panigirtzoglou (2002), Kliger and Levy (2002), and Ziegler (2003).
[9] See Constantinides (1982).

where U is a state-independent utility function across states, p is a subjective PDF, q is a risk neutral PDF, i.e. a state price density function which can be implied by option prices and S is the return on the market portfolio across states.

As easily can be seen in equation (17), the first derivatives of the implied PDFs as well as the level of them are important factors that determine the shape of the risk aversion function. If we focus on this application of implied PDFs, the stability of PDFs should be also assessed by their derivatives. Besides the summary statistics of the implied PDFs, in this sense, we suggest following additional measures to test the stability of them for a given S:

$$\sqrt{\frac{1}{T}\sum_{t}^{T}(q^{'}(S,t)-\overline{q^{'}(S,t)})^{2}}\ , \qquad (18)$$

$$\frac{1}{T-1}\sum_{t}^{T-1}\left|q^{'}(S,t+1)-q^{'}(S,t)\right|, \qquad (19)$$

$$\sqrt{\frac{1}{T-1}\sum_{t}^{T-1}(\left|q^{'}(S,t+1)-q^{'}(S,t)\right|-\overline{\left|q^{'}(S,t+1)-q^{'}(S,t)\right|})^{2}}\ . \qquad (20)$$

Each of these measures provides the sample standard deviation of the first derivative of the implied PDFs and the sample mean and standard deviation of their absolute daily changes for a given return state.

In the procedure of deriving the test statistics in (18) - (20), we should pay attention to the smoothing parameter of estimation methods. As the curvature of estimated PDFs can fluctuate greatly according to the choice of the smoothing parameter, an unreasonable input of smoothing parameter may produce misleading results. In this paper, we use the parameter that makes the GOF of both estimation methods equal.

3-3. Measures related to the Option Pricing

Using the implied PDFs which are derived from liquid option prices, we can calculate the prices of illiquid options with the same maturities. Regarding these pricing issues, implied PDFs with stable summary statistics may not always guarantee better performance than those with unstable summary statistics. In this sense, we suggest following additional measures:

$$\frac{1}{T-1}\sum_{t}^{T-1}\left|\hat{C}_{t}-C_{t+1}\right|, \qquad (21)$$

$$\frac{1}{T-1}\sum_{t}^{T-1}\left|\frac{\hat{C}_{t}-C_{t+1}}{\hat{C}_{t}}\right|, \qquad (22)$$

where $\hat{C}_t$ is a fitted price of options calculated by implied PDFs at time t and C_{t+1} is a market price of options at time $t+1$.

The first test statistic measures the absolute pricing error between the fitted price and the market price and the second test statistic measures the absolute normalized one. With the same changes in underlying asset price, we expect that the estimation method guaranteeing more stable implied PDFs would provide smaller pricing errors than another.

3-4. Measures related to the Perturbed Price

As discussed in the introduction, we employ the price-perturbation method to increase the reliability of our tests on the stability of implied PDFs. The perturbed data are usually constructed from possible measurement errors in option prices. The measurement errors originate mostly from non-synchronicity, illiquidity, price discreteness (due to tick size), and operational problems. Among these possible causes of measurement errors, only price discreteness has been considered in most previous researches.[10]

The changes of the stability measures resulting from randomly perturbing the option prices represent the effect of a presumed amount of errors on the implied PDFs. By varying the amount of errors, we can examine the sensitivity of implied PDFs to the errors and can investigate the effect of local price changes on the implied PDFs.

To begin with, we construct pseudo prices by randomly perturbing the market prices with the following types of errors:

$$\varepsilon(K_i, t) \sim \kappa \times uniform(-\eta_i, \eta_i), \quad (23)$$

where $\varepsilon(K_i, t)$ is a measurement error embedded in the price of option with strike price K_i at time t, η_i is a half tick size and κ is the scale parameter which controls the size of measurement errors.

Random numbers are sampled from each given uniform distribution and 200 implied PDFs for a given κ per day are recovered, using the DLN and the SMIV methods. We test the effects of the measurement errors on implied PDFs by following measures:

$$AD(\mathrm{M}, \kappa, t) \equiv \left| \overline{\mathrm{M}}_{perturbed}(\kappa, t) - \mathrm{M}_{unperturbed}(t) \right|, \quad (24)$$

$$SD(\mathrm{M}, \kappa, t) \equiv \sigma(\mathrm{M}_{perturbed}(\kappa, t)), \quad (25)$$

$$PR_{0.05}(\mathrm{M}, \kappa, t) \equiv \mathrm{M}_{(0.95)perturbed}(\kappa, t) - \mathrm{M}_{(0.05)perturbed}(\kappa, t), \quad (26)$$

where, M is a stability measure discussed in previous section, such as summary statistics and first derivatives, $\overline{\mathrm{M}}$ is the sample mean and $\mathrm{M}_{(i)}$ is the i^{th} % order

[10] See Cooper (1999), Bliss and Panigirtzoglou (2002), and Giamourids and Tamvakis (2001).

statistic of the stability measure.

The AD (absolute deviation) measures the convergence of an implied PDF for perturbed data to that for unperturbed data. The high value of AD for a given κ means that it tends to fail to converge to the implied PDF for unperturbed data. The SD (standard deviation) and PR (percentile range) measure the dispersion of implied PDFs for perturbed data around their mean. The higher the SD and PR for a given κ, the more dispersed implied PDFs are.

4. Data

The KOSPI 200 option contracts used in this study has been traded since July 7, 1997. They are European-style options on the KOSPI 200 stock index traded on the Korean Stock Exchange. In spite of their short history, the KOSPI 200 options market has became one of the biggest option markets in the world. The KOSPI 200 options expire on the second Thursday of the expiry month. In terms of time to maturity, they are listed for the three consecutive near-term delivery months and one additional month from the quarterly cycle (March, June, September and December). They also have at least five strike prices, in terms of moneyness. The number of strike prices may, however, change according to an underlying asset price movement.

The sample period extends from December 11, 1998 through September 27, 2002. We obtain minute–by–minute transaction prices for the KOSPI 200 options from the Korean Stock Exchange. The 91-day certificate deposit (CD) rate is used as a risk-free interest rate.[11] The following criteria are applied to filter data needed for the empirical test.

To obtain well-behaved implied PDFs, only call options are included. For each day in the sample, we select the minute–by–minute transaction data at time 14: 50.[12] To remove the illiquid options, we select only the data which are traded actually between 14: 30 and 14: 50. As the liquidity of KOSPI 200 option contracts is concentrated in the nearest expiration contract, the maturity of most options which are used in our study is not more than one month. Because options with less than 7 days to expiration may produce biases due to low prices and high bid-ask spreads, they are excluded from the sample. As a minimum of five strikes is required to estimate the five-parameter double-lognormal function, option cross-sections with less than six 'good'[13] option strikes are excluded. To mitigate the impact of price discreteness on option value,

[11] As Korea does not have a liquid Treasury bill market, 91-day CD rate have been used as a proxy of the risk-free interest rate in most of empirical researches.
[12] In the Korea stock market, simultaneous bids and offers begin at 14:50
[13] Contracts that fulfill no arbitrage conditions are regarded as 'good' options.

prices lower than 0.3 are excluded from the sample. Finally, Data not satisfying the following arbitrage restrictions are not included.

$$C(K_i,\tau_i) \geq S_t - \sum_{s=1}^{\tau_i} e^{-r_{t,s}s} D_{t+s} - K_i B_{t,\tau_i} \ , \tag{27}$$

where B_{t,τ_i} is a zero-coupon bond that pays 1 in τ_i periods from time t and D_t is daily dividends at time t.

Table I shows the summary statistics of our sample filtered by above rules. As expected, the number of OTM options is twice as much as that of ITM or ATM options and the maturity of most options is not more than one month.

5. Empirical Results

5-1. Estimated Summary Statistics of Implied PDFs for Unperturbed Data

The means and standard deviations of GOF measures and estimated summary statistics are presented in panel A of Table II.[14] By construction, the GOF of the DLN method and the SMIV method must be almost the same.[15] In our sample, the SMIV method does not, however, provide well-behaved PDFs with a smoothing parameter that produces the same GOF as that provided by the DLN method. The GOF in the SMIV method is lower than in the DLN method in panel A, and this is the price paid for deriving a PDF without a spike. Most of the summary statistics except for means are slightly different in their means and their standard deviations. The DLN method produces a slightly greater variation in estimated standard deviation and kurtosis than the SMIV method. However, the SMIV method produces a slightly greater variation in estimated skewness measures than the DLN method, except for the $\hat{\delta}_4$. From these results, we generally infer that two methods are similar in performance.

The means and dispersions of the absolute day-to-day changes in estimated summary statistics are presented in panel B of Table II. The difference in performance across PDF estimation methods is more clearly discovered in panel B. The average values of day-to-day changes in most of estimated summary statistics except for skewness measures are greater in the DLN method than in the SMIV method. The DLN method also produces much greater variation except for skewness measures than the SMIV method, by a factor of about two or more. Furthermore, in the performance of four skewness measures, the DLN method does not clearly outperform the SMIV method.

[14] They are just a simple average and a standard deviation for the entire sample period. Since the number of cross-sectional data is different according to option maturity, the simple average and the standard deviation may be inappropriate measures. However, the values derived by maturity provide similar results to ours.
[15] At first, a smoothing parameter that produces the same GOFs is selected in our optimization algorithm.

5-2. Estimated First Derivatives of Implied PDFs for Unperturbed Data

Table III presents the means and the standard deviations of implied PDFs' first derivatives across stock index return states. Our sample return states are divided into twenty groups, excluding extremely positive and negative return states.[16] The value in each return state is the average of the first derivatives included in it.[17] Generally, the implied probability densities increase on negative return states and diminish on positive return states. They are concave as their first derivatives diminish across return states. Table III, thus, indicates that estimated PDFs are inverse U-shaped functions across sample return states from 0.90 to 1.10. The standard deviations of implied PDFs' first derivatives exhibit different patterns according to estimation methods. The DLN method produces a similar variation in first derivatives across entire sample return states, while the SMIV method produces a much greater variation on negative return states than on positive return states. That is, the variation in first derivatives by the DLN method is almost constant regardless of any return state, while the variation in first derivatives by the SMIV method diminishes as the return increases. In comparing these two estimation methods, we find that the SMIV method produces a much greater variation on negative return states than the DLN method, by a factor of two or more. On positive return states, the two methods produce a similar variation in implied PDFs' first derivatives.

The means and dispersions of absolute day-to-day changes in first derivatives are presented in Table IV. As with the results in Table III, absolute day-to-day changes in first derivatives by the DLN method do not differ across entire return states in their means and dispersions. In the case of the SMIV method, absolute day-to-day changes in first derivatives on negative return states are much greater than those on positive return states. On negative return states, the SMIV method produces great and unstable day-to-day changes compared to the DLN method. However, differences between the DLN and the SMIV methods are either negligible or non-existent on positive return states.

Based on the results gathered from Tables II to IV, we find that the performance of the two methods could be assessed differently according to the test statistics employed. In terms of the summary statistics, the SMIV method produces more stable PDFs than the DLN method. We find, however, that the DLN method produces much more stable PDFs in terms of their first derivatives than are produced by the SMIV method. The former finding could be interpreted as resulting form the superior performance of the SMIV

[16] In our paper, a positive (negative) return state refers to the state in which future asset prices are higher (lower) than those of today.

[17] This is straightforward under the assumption that the implied PDFs are piecewise linear within each return state.

method at macro levels, and the latter finding could be interpreted as resulting from the superior performance of the DLN method at micro levels. This interpretation of our results has something in common with arguments presented in previous researches. As pointed out by various authors,[18] it is likely that the instability of the DLN method in estimated summary statistics is due to its parametric nature, which may change the shape of the entire distribution with local price changes. To verify our interpretation of the results, the sensitivity of the implied PDFs to local price changes is examined in the next two subsections.

5-3. Test Statistics of the Implied PDFs for Perturbed Data: the Case of a Half Tick Errors

In this subsection, we perturb the original option prices by no more than a half tick and derive the implied PDFs from the perturbed data. The effect of marginal price changes on the implied PDFs is examined. The lower bound on the estimated test statistics is also obtained. Table V presents the simulation results using the perturbed price data. In this procedure, the sample periods are shortened from September 6, 2002 through September 27, 2002 because of long computer processing time. Despite only 15 days are included in sample periods, the results do not generally depend on the length of sample periods.[19]

The results reported in the first column of Table V imply the converging ability of implied PDFs for perturbed data to the implied PDFs for unperturbed data. The absolute deviations of estimated summary statistics between the two PDFs are generally greater in the DLN method than in the SMIV method. The only exceptions are the Pearson-mode based skewness measure ($\hat{\delta}_2$) and the $\hat{\delta}_4$, which are quite similar. It may be inferred that, with minimum measurement errors in option prices, summary statistics estimated by the SMIV method generally converge to the true values more easily than those derived by the DLN method. This inference, however, is not valid in the case of first derivatives of implied PDFs. The absolute deviations between the first derivatives of implied PDFs for perturbed data and those for unperturbed data are much greater in the SMIV method than in the DLN method, by at least a factor of two, and even as high as five to twenty. Their differences are intensified on negative return states and are nearly negligible on deeply positive return states.

The second column of the Table V presents the standard deviations of estimated summary statistics and first derivatives of implied PDFs for perturbed data. The 5th to

[18] For example, see Bliss and Panigirtzoglou (2002)
[19] In the case of half-tick error, we conduct the same analysis in the longer sample periods; from March 11, 2002 through June 28, 2002. The results do not systematically differ from those in shorter sample periods.

95^{th} and the 25^{th} to 75^{th} percentile ranges of estimated results are presented in the third and fourth columns of Table V. All these three columns show the dispersions of implied PDFs for perturbed data. In the case of estimated summary statistics, the DLN method produces generally greater variations than the SMIV method. On the contrary, the SMIV method produces much greater variation in the first derivatives than the DLN method, by at least a factor of two, and even as high as three to ten.

From the above results, we find that the effect of small price changes on the implied PDFs is different according to the estimation method used. From the low converging abilities and high dispersions around the mean of estimated summary statistics by the DLN method, we can infer that it alters the shape of an entire distribution with small price changes. From the highly stable first derivatives of implied PDFs, we can also infer that the DLN method does not sharply alter a specific part of the distribution with small price changes. Because of this, the parametric DLN method produces a PDF with a more stable slope, although it does produce a more unstable PDF in terms of summary statistics. The nature of the non-parametric SMIV method is found to be directly opposite to the parametric DLN method. The SMIV method produces a more stable PDF in terms of summary statistics, although it does produce a PDF with a more unstable slope.

Next, the lower bound on the confidence intervals of the estimated results can be also derived from Table V. For example, for the PDFs estimated by the DLN method, the 90% confidence intervals for $\hat{\mu}$ and $\hat{\sigma}$ are 0.1805 and 0.2559 and are 6.5% and 12.9% of their respective mean absolute daily changes (2.7918 and 1.9804). For the PDFs by the SMIV method, they are 0.0089 and 0.0580 and are only 0.4% and 5.9%, respectively. The 90% confidence intervals for first derivatives on the return states of 0.92~0.93 and 1.07~1.08 are 0.0251 and 0.0255 and are 27.3% and 32.4% of their respective mean absolute daily changes (0.0918 and 0.0788) in the case of the DLN method. For the PDFs by the SMIV method, they are 0.2067 and 0.0198 and are 87.1% and 24.8%, respectively. These results shown in Table V indicate that measurement errors of option prices can create a significant degree of uncertainty to the accuracy of implied PDFs, although a half tick error is only considered.

5-4. Test Statistics of the Implied PDFs for Perturbed Data: A General Case

In this subsection, we examine the sensitivity of the implied PDFs to possible measurement errors by varying the size of κ. Figure I plots the convergence (AD) and dispersion (SD) measures of estimated summary statistics of implied PDFs for perturbed data by varying the value of κ from 0.5 to 5.0.[20] As expected, the degree of uncertainty to the estimated summary statistics is generally increased with increasing value of the κ. Except for the skewness, the converging ability of implied PDFs for perturbed data to PDFs for unperturbed data is found to be consistently better in the SMIV method than in the DLN method in panel A. The dispersion around the mean of estimated summary statistics is also found to be consistently much greater in the DLN method in panel B. To summarize, the relative instability of the summary statistics derived by the DLN method is valid regardless of the value of κ and probably tends to be more intensified with increasing value of the κ, especially for the dispersion measure.

The stability measures of first derivatives are plotted in Figure II. Panel A plots the convergence measure (AD) and panel B plots the dispersion measure (SD) of them by varying the value of κ from 0.5 to 5.0. Because of the limited space, we plot only the results on four return states: deeply positive, slightly positive, deeply negative, and slightly negative ones. Irrespective of the value of κ, the slope of PDFs estimated by the DLN method is much more stable than the slope estimated by the SMIV method. The only exception is the slope at deeply negative return states. Therefore, the relative instability of the implied PDFs' first derivatives by the SMIV method is valid regardless of the value of κ. It is also likely that it tends to be more intensified with increasing value of the κ, especially for the dispersion measure on the deeply negative return states.[21]

Finally, from the results shown in Figures I and II, we see that the findings reported in subsection 5-3 are valid, regardless of the amount of price changes. The differences in performance between the DLN and the SMIV methods reported in Tables II to V are found to be consistent and systematic.

5.5. Pricing Errors between the Fitted Prices and the Market Prices

Another stability measure, pricing errors between market prices and fitted prices of options using the one-day-ahead implied PDF as a today's PDF, is presented in Table VI. Sample data are divided into eight groups according to their moneyness. The means of absolute pricing errors tend to be high in ITM call options and tend to be low in OTM call options, regardless of the estimation method. The difference between two methods in performance is almost negligible. By contrast, the means of absolute pricing errors

---

[20] The skewness measure plotted in Figure I is $\hat{\delta}_1$. The other skewness measures exhibit similar patterns to $\hat{\delta}_1$.

[21] Although the other dispersion measures (PR) are not presented in this paper, they, like absolute deviations and standard deviations, give similar results to above.

normalized by their market prices tend to be low in ITM and high in OTM call options. The difference between the DLN and the SMIV methods in performance is also negligible. Therefore, only focusing on the application of implied PDFs to an option pricing, estimated PDFs by the DLN method can be considered as stable as those by the SMIV method.

6. Conclusions

This paper has developed and applied various measures for assessing the stability of implied PDFs recovered from option prices. We present three different measures: the estimated summary statistics; the first derivatives of implied PDFs; and the pricing errors between fitted prices and market prices of options. The latter two measures are related to the applications of option-implied PDFs, namely, recovering the risk aversion of investors and pricing other options.

The estimation methodologies, employed in this paper for assessing the stability of implied PDFs, are the parametric DLN method and the non-parametric SMIV method. These estimation methods are investigated on a set of option data: KOSPI 200 index options in the Korean market, which is one of the most actively traded options in the world.

The results of our study are different from previous ones in the following ways. Even if the SMIV method, viewed in summary statistics, provides more stable implied PDFs than the DLN method, the results of the other two measures are different. First, when we focus on the stability of the implied PDFs' slope, it is found that the DLN method is much superior to the SMIV method in overall performances. It is likely that this finding results from the parametric nature of the DLN method, which may result in local price changes affecting the entire distribution rather than sharply altering a small part of the distribution. Second, the stability of implied PDFs in the two methods does not systematically differ in terms of pricing errors. Their differences are not greater than a half tick and thus can be considered negligible.

To summarize, the stability between the parametric DLN and the non-parametric SMIV methods cannot be assessed by an absolute criterion. The selection of an estimation method to derive the implied PDFs should be based on the purpose for the estimate. To investigate the changes of investors' beliefs on a financial market using the implied PDFs, it is advisable to use a non-parametric method, such as the SMIV. This suggestion is based on the fact that the SMIV method produces implied PDFs with more stable summary statistics than are produced in the DLN method. However, focusing on the shape of implied PDFs on the specific return state, we find that a parametric method is much more desirable than a non-parametric method. This finding is based on the fact that the DLN method produces implied PDFs with a more stable slope than are produced in the SMIV method. One of drawbacks to using this method is that a theoretical constraint, such as on a process for pricing kernels, may be imposed immediately after the adoption of a parametric method to recover a PDF. In addition, for the purpose of pricing other options using the estimated PDFs, it is also supposed that the same results would be induced regardless of the estimation method used.

Finally, the sensitivity analysis of the implied PDFs to possible measurement errors by varying the size of errors strengthens the reliability of out results: the stability of estimated PDFs by the DLN and the SMIV methods is assessed differently according to the test statistics employed.

Reference

Aït-Sahalia, Y., Lo, A., 2000, "Nonparametric Risk Management and Implied Risk Aversion," Journal of Econometrics, 94, 9-51.
Anagnou, I., Bendendo, M., Hodges, S., Tompkins, R., 2002, "The Relation between Implied and Realized Probability Density Function," working paper, University of Warwick.
Bahra, B., 1997, "Implied Risk-Neutral Probability Density Functions from Option Prices: Theory and Application," working paper, Bank of England.
Bakshi, G. S., Cao, C., Chen, Z. W., 1997, "Empirical Performance of Alternative Option Pricing Models," Journal of Finance, 52, 2003-2049.
Bates, D. S., 1991, "The Crash of '87: Was it Expected? The Evidence from Option Markets," Journal of Finance, 46, 1009-1044.
Black, F., Scholes, M., 1973, "The Pricing of Options and Corporate Liabilities," Journal of Political Economy, 81, 634-659.
Bliss, R. R., Panigirtzoglou, N., 2002, "Testing the Stability of Implied Probability Density Functions," Journal of Banking and Finance, 26, 381-422.
Bliss, R. R., Panigirtzoglou, N., 2004, "Option Implied Risk Aversion Estimate," Journal of Finance, 59.
Breeden, D. T., Litzenberger, R. H., 1978, "Prices of State-Contingent Claims Implicit in Option Prices," Journal of Business, 51, 621-651.
Brunner, B., Hafner, R., 2003, "Arbitrage-Free Estimation of the Risk-Neutral Density from the Implied Volatility Smile," Journal of Computational Finance, 7
Campa, J. M., Chang, P. H. K., Reider, R. L., 1997, "Implied Exchange Rate Distributions: Evidence from OTC Option Markets," NBER working paper.
Constantinides, G. M., 1982, "Intertemporal Asset Pricing with Heterogeneous Consumers and without Demand Aggregation," Journal of Business, 55, 253-267.
Cooper, N., 1999, "Testing Techniques for Estimating Implied RNDs from the Prices of European-style Options," working paper, Bank of England.
Giamouridis, D. G., Tamvakis, M. N., 2001, "Can We Rely on Implied Distributions? Evidence from the Interest Rates Market," working paper, City University Business School, London.
Heston, S., 1993, "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," Review of Financial Studies, 6, 327-343.
Jackwerth, J. C., 1999, "Option-Implied Risk-Neutral Distributions and Implied Binomial Trees: A Literature Review," Journal of Derivatives, Winter, 66-82.
Jackwerth, J. C., 2000, "Recovering Risk Aversion from Option Prices and Realized Returns," Review of Financial Studies, 13, 433-451.
Kim, I. J., Kim, S., 2003, "On the Usefulness of Implied Risk Neutral Distributions – Evidence from Korean KOSPI 200 Index Options Market," Journal of Risk, 6.
Kliger, D., Levy, O., 2002, "Risk Preferences Heterogeneity: Evidence from Asset Markets," European Finance Review, 6, 277-290.

Leland, H., 1980, “Who should Buy Portfolio Insurance,” Journal of Finance, 35, 581-594.
Malz, A. M., 1997, “Estimating the Probability Distributions of the Future Exchange Rate from Options Prices,” Journal of Derivatives, Winter, 18-36.
Melick, W. R., Thomas, C. P., 1998, “Confidence Intervals and Constant Maturity Series for Probability Measures Extracted from Option Prices,” working paper, Proceedings of the Bank of Canada Conference on ‘Information Contained in Prices of Financial Assets’.
Melick, W. R., 1999, “Results of the Estimation of Implied PDFs from a Common Data Set,” Proceedings of the BIS Workshop on ‘Estimating and Interpreting Probability Density Functions’.
Rosenberg, J. V., Engle, R. F., 2002, “Empirical Pricing Kernels,” Journal of Financial Economics, 64, 341-372.
Salmon, M., 2001, “Estimating Implied Risk Neutral Densities from Options Prices,” Lecture note, City University of Business School.
Shimko, D., 1993, “Bounds of Probability,” Risk, 6, 34-37.
Soderlind, P., Svensson, L. E. O., 1997, “New Techniques to Extract Market Expectations from Financial Instruments,” Journal of Monetary Economics, 40, 383-429.
Soderlind, P., 1999, “Market Expectations in the United Kingdom before and after the ERM Crisis,” working paper, Stockholm School of Economics.
Syrdal, S. A., 2002, “A Study of Implied Risk-Neutral Density Functions in the Norwegian Option Market,” working paper, Norges Bank.
Ziegler, A., 2003, “Why does Implied Risk Aversion Smile?,” working paper

Table I. Summary statistics of option cross-sectional data samples

The sample period extends from December 11, 1998 through September 27, 2002. Only call options are included and their maturities are not more than 41 days.

Final sample		1998-99	2000	2001	2002	All
Number of cross-sections	ITM	394	343	330	484	1551
	ATM	415	430	223	390	1458
	OTM	673	1003	395	641	2712
Strikes per cross-sections						
Range		6~14	6~13	6~12	6~15	6~15
Average		8.142	8.663	7.128	9.469	8.413
Time to expiry(years)						
Range		0.019~0.096	0.019~0.112	0.019~0.112	0.019~0.112	0.019~0.112
Average		0.052	0.059	0.062	0.060	0.058

Table II. Means and standard deviations of estimated PDF summary statistics for unperturbed data

Panel A presents means and standard deviations of the estimated PDF summary statistics. Panel B presents means and three different dispersion measures of the estimated PDF summary statistics absolute day-to-day changes.

<Panel A>

Summary statistic	Mean		Standard deviation	
	DLN	SMIV	DLN	SMIV
GOF	0.0142	0.0222	0.0552	0.0631
$\hat{\mu}$	92.0741	92.3208	17.9094	17.8227
$\hat{\sigma}$	10.0802	9.2727	3.8363	2.7527
$\hat{\delta}_1$	0.3351	0.4185	0.3590	0.5088
$\hat{\delta}_2$	0.0969	0.2134	0.2261	0.3299
$\hat{\delta}_3$	0.0386	0.0439	0.1036	0.1154
$\hat{\delta}_4$	1.0958	1.1907	0.6926	0.3147
$\hat{\xi}$	3.9368	3.4202	1.0141	0.6198

<Panel B>

Summary statistic	Mean		Standard deviation		$\hat{X}_{05}$ to $\hat{X}_{95}$ range		$\hat{X}_{25}$ to $\hat{X}_{75}$ range	
	DLN	SMIV	DLN	SMIV	DLN	SMIV	DLN	SMIV
$\hat{\mu}$	2.7918	2.0187	3.4943	1.9635	8.4770	5.1976	2.7185	2.1476
$\hat{\sigma}$	1.9804	0.9809	3.2329	1.1750	7.1972	3.6404	1.5994	0.8884
$\hat{\delta}_1$	0.3378	0.3966	0.2969	0.3937	0.8254	1.1606	0.3399	0.4080
$\hat{\delta}_2$	0.1973	0.2914	0.2362	0.2557	0.7258	0.7962	0.2161	0.3445
$\hat{\delta}_3$	0.0833	0.0926	0.1060	0.0826	0.2832	0.2552	0.0849	0.0951
$\hat{\delta}_4$	0.2323	0.2285	0.9425	0.2973	0.7547	0.7129	0.1542	0.2215
$\hat{\xi}$	0.8748	0.4309	1.0376	0.5881	2.4392	1.5193	0.9578	0.4116

Table III. Means and standard deviations of estimated PDF first derivatives for unperturbed data

Extremely positive and negative return states are excluded and sample return states are divided into twenty groups.

Return states	Mean		Standard deviation	
	DLN	SMIV	DLN	SMIV
0.90~0.91	0.1976	0.2229	0.1585	0.3043
0.91~0.92	0.1895	0.1918	0.1606	0.3378
0.92~0.93	0.1759	0.1557	0.1603	0.3701
0.93~0.94	0.1566	0.1190	0.1580	0.3910
0.94~0.95	0.1321	0.0802	0.1535	0.3939
0.95~0.96	0.1030	0.0381	0.1477	0.3161
0.96~0.97	0.0706	0.0055	0.1431	0.3342
0.97~0.98	0.0365	-0.0094	0.1413	0.2889
0.98~0.99	0.0021	-0.0278	0.1422	0.2316
0.99~1.00	-0.0313	-0.0510	0.1437	0.1822
1.00~1.01	-0.0625	-0.0787	0.1445	0.1563
1.01~1.02	-0.0910	-0.1070	0.1450	0.1571
1.02~1.03	-0.1162	-0.1323	0.1455	0.1650
1.03~1.04	-0.1380	-0.1518	0.1462	0.1698
1.04~1.05	-0.1561	-0.1667	0.1463	0.1687
1.05~1.06	-0.1708	-0.1764	0.1451	0.1603
1.06~1.07	-0.1820	-0.1821	0.1421	0.1485
1.07~1.08	-0.1901	-0.1839	0.1370	0.1364
1.08~1.09	-0.1952	-0.1837	0.1302	0.1239
1.09~1.10	-0.1978	-0.1826	0.1222	0.1128

Table IV. Means and three different dispersion measures of estimated PDF first derivatives absolute day-to-day changes

Return states	Mean		Standard deviation		$\hat{X}_{05}$ to $\hat{X}_{95}$ range		$\hat{X}_{25}$ to $\hat{X}_{75}$ range	
	DLN	SMIV	DLN	SMIV	DLN	SMIV	DLN	SMIV
0.90~0.91	0.0968	0.1904	0.1323	0.2534	0.3584	0.7283	0.0927	0.1959
0.91~0.92	0.0963	0.2141	0.1312	0.3078	0.3555	0.8748	0.0961	0.2293
0.92~0.93	0.0918	0.2373	0.1323	0.3321	0.2895	0.9271	0.0878	0.2554
0.93~0.94	0.0882	0.2466	0.1336	0.3267	0.2945	0.8534	0.0829	0.2725
0.94~0.95	0.0870	0.2486	0.1348	0.3357	0.2832	0.8181	0.0754	0.2645
0.95~0.96	0.0839	0.2475	0.1401	0.2873	0.2959	0.8732	0.0714	0.2765
0.96~0.97	0.0863	0.2444	0.1473	0.3278	0.3363	0.8590	0.0727	0.2329
0.97~0.98	0.0890	0.2116	0.1578	0.2782	0.3344	0.7467	0.0753	0.1964
0.98~0.99	0.0918	0.1728	0.1675	0.2191	0.3529	0.5815	0.0775	0.1548
0.99~1.00	0.0962	0.1369	0.1696	0.1662	0.3779	0.4320	0.0761	0.1328
1.00~1.01	0.0989	0.1172	0.1637	0.1290	0.3814	0.3442	0.0832	0.1104
1.01~1.02	0.0990	0.1114	0.1519	0.1203	0.3814	0.3454	0.0890	0.1108
1.02~1.03	0.0977	0.1076	0.1363	0.1273	0.3570	0.3529	0.0921	0.1023
1.03~1.04	0.0946	0.1053	0.1207	0.1321	0.3223	0.3757	0.0860	0.1012
1.04~1.05	0.0910	0.1030	0.1066	0.1285	0.3066	0.3763	0.0849	0.0977
1.05~1.06	0.0870	0.0963	0.0956	0.1176	0.2835	0.3432	0.0841	0.0892
1.06~1.07	0.0832	0.0881	0.0870	0.1045	0.2599	0.2955	0.0780	0.0849
1.07~1.08	0.0788	0.0797	0.0805	0.0940	0.2383	0.2601	0.0786	0.0797
1.08~1.09	0.0742	0.0705	0.0754	0.0819	0.2204	0.2188	0.0736	0.0747
1.09~1.10	0.0696	0.0624	0.0714	0.0720	0.1999	0.1893	0.0709	0.0645

Table V. Test statistics of estimated PDFs for perturbed data

All the test statistics reported in this table are derived with the κ equal to 0.5.

	AD		SD		$PR_{0.05}$		$PR_{0.25}$	
	DLN	SMIV	DLN	SMIV	DLN	SMIV	DLN	SMIV
Summary statistics								
$\hat{\mu}$	0.0234	0.0008	0.0578	0.0037	0.1805	0.0089	0.0761	0.0070
$\hat{\sigma}$	0.0409	0.0021	0.0844	0.0190	0.2559	0.0580	0.1037	0.0297
$\hat{\delta}_1$	0.0177	0.0038	0.0405	0.0262	0.1289	0.0762	0.0495	0.0448
$\hat{\delta}_2$	0.0113	0.0154	0.0209	0.0419	0.0616	0.1199	0.0223	0.0716
$\hat{\delta}_3$	0.0036	0.0027	0.0080	0.0091	0.0244	0.0265	0.0098	0.0155
$\hat{\delta}_4$	0.0073	0.0107	0.0142	0.0240	0.0435	0.0743	0.0175	0.0362
$\hat{\xi}$	0.0579	0.0029	0.0922	0.0190	0.2786	0.0604	0.1045	0.0264
First derivatives on return states								
0.90~0.91	0.0034	0.0727	0.0082	0.0540	0.0251	0.1674	0.0098	0.0871
0.91~0.92	0.0029	0.0689	0.0079	0.0507	0.0247	0.1600	0.0093	0.0738
0.92~0.93	0.0033	0.0663	0.0081	0.0662	0.0251	0.2067	0.0094	0.0999
0.93~0.94	0.0039	0.0449	0.0087	0.0737	0.0257	0.2281	0.0097	0.1192
0.94~0.95	0.0043	0.0341	0.0092	0.0561	0.0262	0.1768	0.0101	0.0871
0.95~0.96	0.0047	0.0220	0.0093	0.0467	0.0262	0.1471	0.0097	0.0726
0.96~0.97	0.0050	0.0228	0.0093	0.0337	0.0255	0.1073	0.0093	0.0499
0.97~0.98	0.0051	0.0168	0.0093	0.0273	0.0258	0.0871	0.0095	0.0406
0.98~0.99	0.0048	0.0158	0.0092	0.0229	0.0266	0.0726	0.0095	0.0339
0.99~1.00	0.0044	0.0093	0.0091	0.0189	0.0268	0.0595	0.0101	0.0279
1.00~1.01	0.0039	0.0058	0.0089	0.0156	0.0275	0.0502	0.0104	0.0215
1.01~1.02	0.0037	0.0062	0.0087	0.0147	0.0277	0.0475	0.0110	0.0204
1.02~1.03	0.0036	0.0073	0.0086	0.0146	0.0276	0.0475	0.0111	0.0196
1.03~1.04	0.0034	0.0074	0.0088	0.0115	0.0278	0.0372	0.0111	0.0155
1.04~1.05	0.0031	0.0058	0.0091	0.0099	0.0278	0.0318	0.0110	0.0136
1.05~1.06	0.0031	0.0036	0.0092	0.0079	0.0274	0.0256	0.0108	0.0105
1.06~1.07	0.0035	0.0040	0.0091	0.0072	0.0267	0.0228	0.0106	0.0098
1.07~1.08	0.0037	0.0038	0.0088	0.0064	0.0255	0.0198	0.0101	0.0088
1.08~1.09	0.0038	0.0034	0.0082	0.0055	0.0237	0.0171	0.0095	0.0079
1.09~1.10	0.0037	0.0029	0.0074	0.0047	0.0215	0.0146	0.0087	0.0068

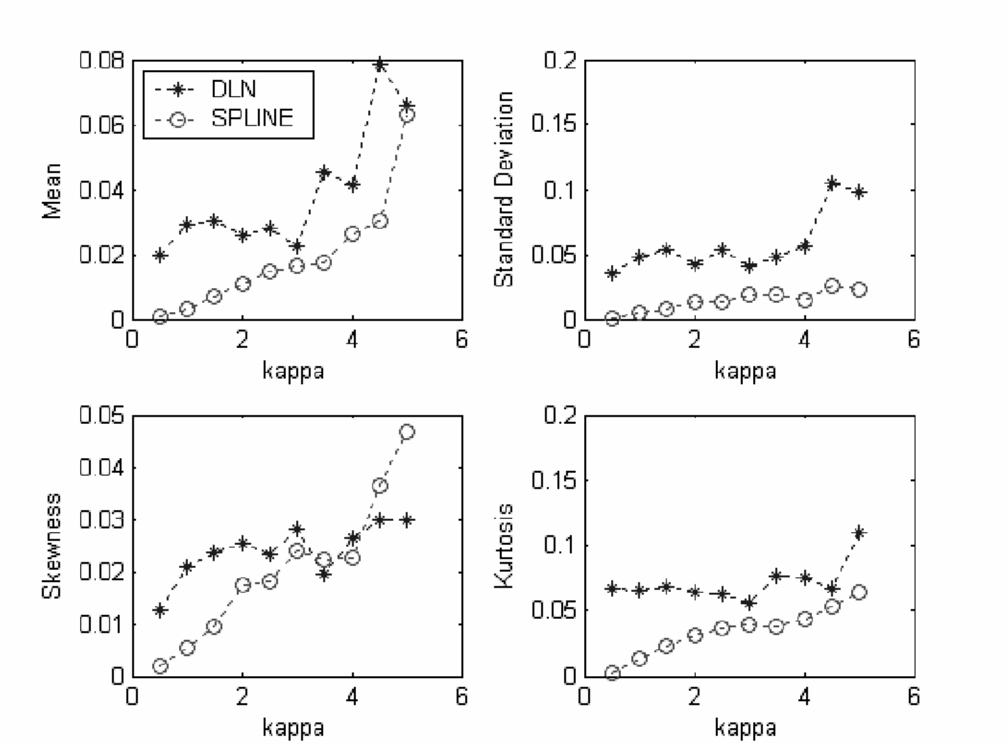
Table VI. Means of pricing errors using the estimated PDFs for unperturbed data

Means of absolute pricing errors between the fitted price and the market price of options are reported in this table. The last two columns report the means of absolute pricing errors normalized by market prices of options.

Moneyness	Number of data	Means of pricing errors		Means of normalized pricing errors	
		DLN	SMIV	DLN	SMIV
K/S < 0.97	1481	1.5181	1.5140	0.1868	0.1869
$0.97 \le K/S < 0.98$	207	1.1750	1.1719	0.2334	0.2325
$0.98 \le K/S < 0.99$	242	1.2259	1.2195	0.2705	0.2686
$0.99 \le K/S < 1.00$	255	0.9851	0.9807	0.2622	0.2604
$1.00 \le K/S < 1.01$	249	0.9843	0.9853	0.2736	0.2731
$1.01 \le K/S < 1.02$	228	0.9761	0.9653	0.3222	0.3181
$1.02 \le K/S < 1.03$	260	0.7755	0.7781	0.2971	0.2972
$1.03 \le K/S$	2641	0.4941	0.4841	0.4101	0.3969
Total	5563	0.9012	0.8945	0.3162	0.3096

Figure I. Comparison of the DLN and the SMIV summary statistics (perturbed data)

〈panel A : absolute deviations between the summary statistics for perturbed and unperturbed data〉



〈panel B : standard deviations of the summary statistics for perturbed data〉

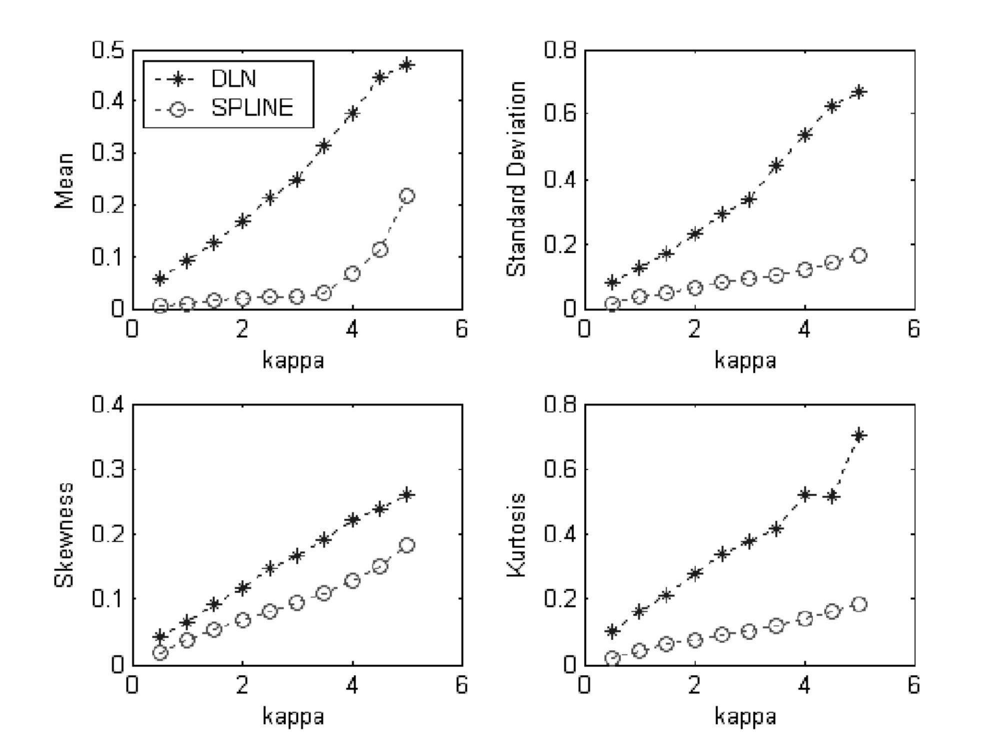
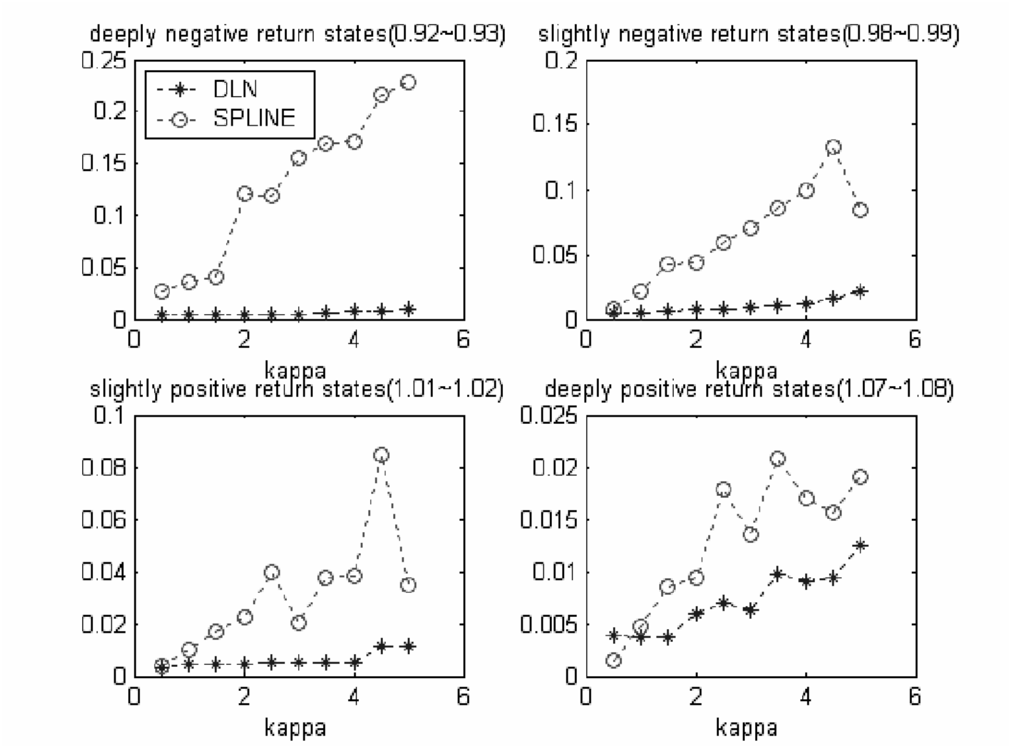


Figure II. Comparison of the first derivatives by the DLN and the SMIV method (perturbed data)

〈panel A : absolute deviations between the first derivatives for perturbed and unperturbed data〉



<panel B : standard deviations of the first derivatives for perturbed data>