Rev Deriv Res (2009) 12:169–191
DOI 10.1007/s11147-009-9039-0

Microstructural biases in empirical tests of option pricing models

Patrick Dennis · Stewart Mayhew

Published online: 13 June 2009

Abstract This paper examines how noise in observed option prices arising from discrete prices and other microstructural frictions affects empirical tests of option pricing models and risk-neutral density estimation. The discrete tick size alone introduces enough noise to make model comparisons difficult, especially for lower-priced stocks. We demonstrate that microstructural noise can lead to incorrect inferences in the univariate diffusion test of Bakshi et al. (Rev Financ Stud 13:549–584, 2000), the transition density diffusion test of Aït-Sahalia (J Financ 57:2075–2112, 2002), and the speed-of-convergence test of Carr and Wu (J Financ 58:2581–2610, 2003). We also show that microstructural noise induces a bias into the implied risk-neutral moment estimators of Bakshi et al. (Rev Financ Stud 16:101–143, 2003). Even in active, liquid option markets, observation error is likely to reduce significantly the power of tests, and in some cases represents an important source of bias.

Keywords Option pricing · Microstructure · Jump-diffusion · Risk-neutral moments

JEL Classification G13

P. Dennis (✉)
McIntire School of Commerce, University of Virginia, P.O. Box 400173, Charlottesville,
VA 22904-4173, USA
e-mail: pjd9v@virginia.edu

S. Mayhew
Office of Economic Analysis, U.S. Securities and Exchange Commission, 100 F St. NE,
Washington, DC 20549, USA
e-mail: mayhews@sec.gov

1 Introduction

In this paper, we investigate the extent to which microstructural noise is likely to contaminate empirical tests of option pricing models. Specifically we look at recent techniques that have been developed for characterizing the underlying asset's stochastic process and estimating risk-neutral densities. Microstructural noise will be more severe for options on low-priced stocks, where the relative tick size is large and where the interval between adjacent strike prices represents a large percentage of the stock price.

Since most of the approaches developed for empirically testing specific option pricing models or classes of models require as inputs the observed market prices of exchange-traded options, microstructural frictions can make it difficult to observe accurate option prices and therefore affect tests of option pricing models (see, for example, Phillips and Smith 1980; Figlewski 1989). Observed option prices are also used as inputs for estimating implied volatility, calibrating stochastic volatility models, and estimating implied risk-neutral density functions. At a minimum, noise in observed option prices translates into decreased power of empirical tests to distinguish between alternative models, and introduces error in risk-neutral density estimates. In certain circumstances, noise in observed option prices can introduce biases into empirical methodologies.

One way to address this is by focusing on the most actively traded options, the prices of which are presumed to be more representative of the option's "true" value. Many authors have restricted their attention to the S&P 500 or S&P 100 index option market where microstructural problems are believed to be relatively unimportant. Intraday data are often used to ensure synchronous observation of prices across options of different strikes, and spread midpoints are used in place of transaction prices in order to mitigate problems induced by bid-ask bounce.

However, it is not clear whether such safeguards are sufficient to ensure reliable results, or on the contrary whether microstructural errors are severe enough to cast doubt on the conclusions drawn in the empirical literature. The work of Christensen and Prabhala (1998) suggests that the amount of noise in option prices is large enough to introduce a significant attenuation bias to regression-based implied volatility forecast tests, such as those conducted by Canina and Figlewski (1993). Research by Hentschel (2003) demonstrates that noise in option prices can cause large biases in estimates of the implied volatility smile. Moreover, there are reasons to believe that the bid ask quote midpoint is likely to be a biased estimate of the fair option price, with the bias a function of the strike price.[1]

Our study is most closley related to those of Bliss and Panigirtzoglou (2002) and Hentschel (2003). Bliss and Panigirtzoglou (2002) perturb the prices of options that trade on the LIFFE and compare two methods of extracting the implied risk-neutral probability density functions (PDFs) from option prices. They find that extracting the

[1] Chan and Chung (2000) find this result in a sequential order arrival model with an informed trader—the asymmetry between buy and sell orders arises from the convexity of the option's payoff function. Systematic biases in quote midpoints also arise if bid and ask prices are determined by the "good-deal bounds" of Cochran and SaaRequejo (2002).

PDF from the smoothed implied volatility function is a better method than using the parametric double log-normal method. In contrast to examining how microstructural noise affects errors in implied volatility as in Hentschel (2003) or errors in the implied PDF as in Bliss and Panigirtzoglou (2002), we investigate three methodologies that have recently been proposed for testing whether options are priced according to a diffusion model. First, Bakshi et al. (2000) test whether the underlying asset price and the call option price always move in the same direction, as predicted by all univariate diffusion models. Second, Aït-Sahalia (2002) tests whether options are priced according to a diffusion model by testing whether the second partial cross-derivative of the log risk-neutral density with respect to the stock price and strike price is a strictly positive function. Third, Carr and Wu (2003) introduce a methodology for testing for jumps in the underlying process by examining the speed of convergence of at-the-money and out-of-the-money option prices as they approach maturity.

To test the accuracy of these methods, we generate simulated option prices using the Black and Scholes (1973) model and then add microstructural frictions designed to resemble real-world conditions. We then apply the techniques to the noisy data, and test whether the method comes to the correct conclusions regarding the underlying model. We find that microstructural noise in option prices is large enough to raise concerns as to the applicability of these methods.

We also examine the effect of discrete strike prices, which can be considered another type of microstructural market imperfection, on risk-neutral density estimation using the risk-neutral moment estimators of Bakshi et al. (2003). Here, we document that significant biases may exist even if option prices are observed without any error. In a world with a continuum of strike prices, the method converges to obtain the correct risk-neutral moments, but with discrete intervals between strike prices, the bias can be large.

The remainder of this article is organized as follows. In Sect. 2, we summarize several categories of empirical tests of option pricing models that rely on observed option prices. In Sect. 3, we examine the methods of Bakshi et al. (2000), Aït-Sahalia (2002), and Carr and Wu (2003) for testing whether options are priced according to a diffusion model. In Sect. 4, we evaluate the measurement errors and biases inherent in methods for computing implied skewness and kurtosis of the risk-neutral density, focusing on the method proposed by Bakshi et al. (2003). In Sect. 5 we conclude with a summary of our results, and a few further comments on the implications of our findings.

2 Empirical tests of option pricing models

Several categories of empirical tests of option pricing models have been proposed.[2] Early studies, beginning with Black and Scholes (1972), used historical data to estimate a volatility parameter, and then tested whether the resulting theoretical prices from their model differ systematically from market prices. Other authors, surveyed by Figlewski (1997), have used regression analysis to test whether Black–Scholes implied

[2] See Bates (2003) for a recent survey.

volatility is an accurate forecast of subsequent realized volatility. In these tests, the Black–Scholes model is consistently rejected.

There are some inherent difficulties associated with these early approaches. For one, they do not work well for testing models that rely on parameters that cannot easily be estimated from time-series data, such as a market price of risk or the risk aversion of the representative investor. Certainly, one can back out implied risk parameters from option prices and demonstrate that a particular model is sufficiently flexible to eliminate a forecast bias in Black–Scholes, but this does not constitute an empirical test of the model. In such a case, one can only attempt to assess whether the implied parameter values are stable over time, or whether the magnitude of the number is economically plausible (see, for example, Bates 2000).

There is also a joint hypothesis problem in trying to interpret the results of these simple tests. One possible interpretation is that the parameter inputs may be mis-measured. Another is that the Black–Scholes model is wrong because of an omitted risk premium, or misspecified due to omitted features such as jumps or stochastic volatility. Other authors have suggested that a portion of the apparent bias in the Black–Scholes model may be attributed to a Jensen's Inequality bias, measurement errors in realized volatility (Poteshman 2000), an error-in-variables problem that causes attenuation bias in forecasting regressions (Christensen and Prabhala 1998), or a "peso problem" (Penttinen 2001).

An alternative approach that largely avoids the joint hypothesis problem is to test the cross-sectional restrictions imposed by the model on the relative prices of the stock and options with different strikes and maturities. For example, Macbeth and Merville (1979); Rubinstein (1985), and others have tested the Black–Scholes model by testing whether implied volatility varies systematically as a function of moneyness or maturity. Longstaff (1995) tests whether the stock price implied by option prices equals the observed stock price. If the cross-sectional restrictions are systematically violated, there are no possible parameter values that will correctly price all the options. Therefore, the rejection cannot be due to mis-estimating parameters.

In this paper, we are interested in evaluating the extent to which microstructure affects the empirical testing of option pricing models. We focus on evaluating a new approach that has arisen in recent years, one that involves testing restrictions or properties that should hold for entire classes of models. For example, Bakshi et al. (2000) note that all univariate diffusion models predict that the underlying asset price and the call option price should always move in the same direction, and test this on a sample of S&P 500 index options. Aït-Sahalia (2002) demonstrates that if the underlying stock follows a diffusion process, the second partial cross-derivative of the log risk-neutral density with respect to the stock price and strike price must be a strictly positive function, and he tests this property on the implied risk-neutral density from option prices. Carr and Wu (2003) introduce a methodology for testing for jumps in the underlying process by examining the speed of convergence of at-the-money and out-of-the-money option prices as they approach maturity.

We are also interested in examining how microstructure is likely to affect implied risk neutral density estimation. Building on the seminal work of Breeden and Litzenberger (1978), several authors have suggested methods for backing out an implied risk-neutral density from option prices, including Rubinstein (1994) and others surveyed by Jackwerth (1999). We focus on the method introduced by Bakshi et al. (2003). These authors propose a method for estimating the moments of the risk-neutral density from the market price of carefully constructed portfolios of options, designed to approximate the payoffs of options with polynomial payoff functions.

3 Tests of diffusion properties

3.1 BCC approach

Bakshi et al. (2000), hereafter BCC, observe that all univariate diffusion option pricing models have the property that the stock price and the call option price, after correcting for time decay, should always move in the same direction. They proceed to test this property by examining the frequency with which observed call option and underlying prices move in the same direction. They find that this frequency is significantly different from zero, and interpret this as evidence that the prices are not being generated by a univariate diffusion model.

Given that option prices are observed with noise, it is possible that the observed option price may move in the opposite direction as the underlying, even if the true price is moving in the same direction. BCC examine the effect of microstructure on the error rate by stratifying their sample by characteristics such as the bid-ask spread, number of quote revisions, time-of-day and volume. They do not find much cross-sectional variation in the error rate, suggesting that their results are not being driven by microstructure. We take a different approach. Beginning with assumptions about stock-price behavior, noise in the observed price, and tick interval we measure how many violations occur simply by chance. We first develop a simple model of this behavior and then quantify the frequency of the errors under more complicated but realistic assumptions using simulation.

Suppose that the underlying asset price can be observed without noise, and let S_t represent the asset price at time t. And suppose that the true call option price is given by the Black–Scholes formula, or some other univariate diffusion model, in which the call price is monotonically increasing with respect to the underlying price:

$$C_t = bs(S_t, \Phi),$$

where the vector Φ includes the other parameters of the model.

Further, suppose that due to market microstructure effects, option prices are observed with noise:

$$\hat{C}_t = C_t + \epsilon_t.$$

Under the conservative assumption that the observed option price is always equal to the true price rounded off to the nearest tick, we would expect the unconditional distribution of ϵ_t to be uniform over the range $[-D/2, D/2]$, where D is the tick size.

Consider two points in time, designated t_0 and t_1. We are able to observe S_0, S_1, $\hat{C}_0$, and $\hat{C}_1$. Between time zero and one, the stock price and the observed call price can move in opposite directions if the change in the noise term more than offsets the change in the true option price. For example, the probability that the stock price increases and the observed call price decreases is:

$$\begin{aligned} PR(S_1 > S_0; \hat{C}_1 < \hat{C}_0) &= PR(S_1 > S_0; C_1 + \epsilon_1 < C_0 + \epsilon_0) \\ &= PR(S_1 > S_0; bs(S_1; \Phi) - bs(S_0; \Phi) < -\Delta\epsilon) \\ &= \int_{S_0}^{\infty} PR[bs(X; \Phi) - bs(S_0; \Phi) < -\Delta\epsilon] f(X) dX \end{aligned}$$

where $\Delta\epsilon \equiv \epsilon_1 - \epsilon_0$ and $f(\,)$ represents the probability density function of S_1, conditional on S_0. If ϵ_0 and ϵ_1 are independent, a uniform distribution for ϵ_t implies that $\Delta\epsilon$ has a triangular distribution with a range of $[-D, D]$, and the expression simplifies to:

$$PR(S_1 > S_0; \hat{C}_1 < \hat{C}_0) = \int_{S_0}^{S^*} \frac{(D + bs(S_0; \Phi) - bs(X; \Phi))^2}{2D^2} f(X) dX$$

where S^* satisfies the equation

$$bs(S^*; \Phi) = bs(S_0) + D.$$

An analogous expression can be derived for the probability that the stock price declines while the observed option price increases. Under Black–Scholes, the distribution $f(\,)$ is lognormal, but the expression may be applied to univariate diffusion option pricing models in general.

This equation tells us that the probability that the stock price and call price move in opposite directions depends on the tick size and the distribution of the true stock price change over the period. In particular, it depends on the probability that the true option price will change by less than one tick. The magnitude of the typical option price change will be larger if the observation horizon is longer, or if the options are more in the money. Thus, in a world where the Black–Scholes model is true but option prices are noisy, we would expect to see the proportion of spurious violations of the BCC condition to be decreasing with time horizon, and increasing with moneyness. These are the exact findings reported by the authors, suggesting that their results may be influenced by microstructural biases. We would also expect the proportion of spurious violations to be an increasing function of the relative tick size, suggesting that the BCC technique should be more biased for options on lower-priced stocks.

Quantifying the number of violations induced by microstructural noise with realistic assumptions is more difficult than the above analysis suggests. Specifically, the tick interval is discrete and there is noise in the stock price as well as the option price. Furthermore, the time decay in the option's price has to be accounted for. Since the tick size is discrete, there is the possibility that the observed price of either the stock or the option may not change, even if the true price changes. Due to this, BCC distinguish between three different types of errors. Let $\Delta\hat{S}$ be the change in the observed stock price and let $\Delta\hat{C}$ be the change in the observed option price. The three types of errors are defined as follows:

- Type I: The observed stock price and call price move in opposite directions: $\Delta\hat{S}\Delta\hat{C} < 0,\ \Delta\hat{S} \neq 0,\ \Delta\hat{C} \neq 0$.
- Type II: The call price does not move, but the stock price does: $\Delta\hat{S}\Delta\hat{C} = 0$, $\Delta\hat{S} \neq 0,\ \Delta\hat{C} = 0$.
- Type III: The call price moves, but the stock price does not: $\Delta\hat{S}\Delta\hat{C} = 0$, $\Delta\hat{S} = 0,\ \Delta\hat{C} \neq 0$.

We use simulation to estimate the magnitude of these violations as follows. Fix the current stock price at S_0 and, using Black–Scholes, compute the price of a call option on the stock, C_0. Noise, which may originate from random arrival of limit orders, temporary supply/demand imbalances, or other sources, is then added to the option price, but the stock price is observed without noise. Compute the observed option price as the true price plus the noise: $\tilde{C}_0 = C_0 + \epsilon_0^c,\ \epsilon_0^c \sim N(0, \sigma_n^2)$. Round the stock price and $\tilde{C}_0$ to the nearest tick and denote the rounded prices as $\hat{C}_0$ and $\hat{S}_0$. For simplicity a tick interval of \$0.10 is used for both the stock and the option.[3] Generate a new stock price Δt years later based on a geometric Brownian motion: $\Delta S = S\mu\Delta t + S\sigma\Delta z$, where $\Delta z = \tilde{z}\sqrt{\Delta t}$. To be consistent with BCC Δt is set to 1 h (the units are still in years) and $\tilde{z}$ is a draw from a standard normal distribution. Set $\mu = 0.10$ per year and $\sigma = 0.30$ per year which are reasonable market parameters for the S&P 500 index. Update the true stock price, $S_1 = S_0 + \Delta S$ and update the option's time to maturity by subtracting Δt from the previous value. Compute the new true option price, C_1, based on the new true stock price, S_1, and the updated time to maturity. Add noise to the updated option price and round both the stock and option price to the nearest tick to arrive at the observed stock and option price, $\hat{S}_1$ and $\hat{C}_1$. Then define the change in the observed prices as: $\Delta\hat{S} = \hat{S}_1 - \hat{S}_0,\ \Delta\hat{C} = \hat{C}_1 - \hat{C}_0$.

Beginning with a stock price of S_0, 100 stock price increments are simulated and the number of Type I, II and III errors are counted. Then the stock price and time to maturity are reset to their original values and another sample path of 100 increments is generated. The limit of 100 stock price increments per sample path was chosen so that time decay does not significantly reduce the option's price for the duration of the sample path. Hundred hours is roughly 4 days, which is small relative to the option's life of 6 months.

A total of 100 sample paths are generated to obtain 100 independent estimates of the number of each type of violation. Then the mean of the 100 estimates is computed, and an estimate of the standard deviation of the mean is then 1/10 the standard deviation of the 100 estimates. Table 1 contains the frequency of observing each type of error

[3] While the tick size for option priced under \$3.00 is \$0.05 in U.S. markets, using this tick interval makes it difficult to determine what is happening when we examine the sensitivity of the error rate to various parameters. For example, as the moneyness decreases, the option price gets lower and a discontinuity in the error rate is introduced at a call price of \$3.00.

Table 1 Simulation results of Type I, II and III errors

	Type I	Type II	Type III
$S_0 = 50,\ \sigma_n = \$0.05$	4.3% (0.2)	21.1% (0.4)	12.6% (0.4)
$S_0 = 500,\ \sigma_n = \$0.05$	0.7% (0.1)	3.2% (0.3)	1.4% (0.2)

This table contains the percentage of Type I, II and III errors, as defined in Bakshi et al. (2000), for a hypothetical stock option and index option. A Type I error results when the observed stock and option price move in opposite directions. A Type II (III) error results when the option price does not (does) move and the stock price does (does not) move. Both options are at-the-money and have a maturity of 6 months. The volatility of stock returns is 30% per year. The risk-free rate is set to 5% and the sampling interval is one hour. For each case the stock follows the process $\Delta S = S\mu\Delta t + S\sigma\Delta z$ where $\mu = 0.10,\ \Delta z = \tilde{z}\sqrt{\Delta t}$ and Δt equals one hour. The amount of noise added to the option price is denoted by σ_n. No noise is added to the stock price. The table contains the frequency of observing each type of error using 100 stock-price increments per sample path for 100 sample paths. The standard deviation of the mean is in parentheses below each estimate

using 100 stock-price increments per sample path for 100 sample paths. The standard deviations of the mean frequencies are in parentheses.

We run the simulation twice, once for S_0 equal to \$50 (an individual stock) and once for S_0 equal to \$500 (the S&P 500 index). We chose 500 for the index level since that was approximately the level of the S&P 500 during the time period covered by the BCC data. The standard deviation of the noise for both the stock and index options was set at a conservative level of 1/2 of a tick or \$0.05. Since the bid-ask spread for index options is often one to two dollars, setting σ_n to \$0.05 is very conservative. It is important to note that if the noise increases in proportion to the stock price, a different error rate results since the tick size is fixed. For example, relative to $S_0 = 50$ and $\sigma_n = 0.05$, $S_0 = 500$ and $\sigma_n = 0.50$ will have more Type I and fewer Type II and III errors since the tick size is smaller as a percentage of the asset level.

Comparing the error rate in the table to the results of BCC, our simulated Type I (II) error rate is 0.7% (3.2%), whereas BCC found an error rate of 13.9% (23%). Hence, while a number of these violations occur simply by chance, there is still a component that is unexplained. The unexplained violations could be economically significant, or they could simply be due to the assumptions about model parameters that we made. For example, if σ_n is increased to \$0.30, the Type I error rate increases to 11%, which is the same order of magnitude that BCC find. Hence, it is important to understand how sensitive the error rates are to the choice of different model parameters. We turn to this next.

Table 2 contains the sensitivity of the error rate to changes in five parameters: the noise level, option maturity, stock return volatility, sampling interval between observations, and the option's moneyness. The base option for panel A (B) is an at-the-money stock (index) call option with 6 months to maturity on an asset whose price is \$50 (\$500). The risk-free rate is set to 5%, the stock return volatility is set to 30%, and the sampling interval is one hour and the volatility of the noise (σ_n) is set to \$0.05 for both the stock and index option. The stock follows the process $\Delta S = S\mu\Delta t + S\sigma\Delta z$ where $\mu = 0.10,\ \Delta z = \tilde{z}\sqrt{\Delta t}$ where Δt equals one hour.

The range of each parameter is given in the table. The parameter range is divided up into 10 equally spaced intervals and 10,000 simulations are run at each value of

Table 2 Sensitivity analysis for microstructural impact on diffusion properties

Parameter name	Parameter range	Type I error slope estimate	Type II error slope estimate	Type III error slope estimate
Panel A: $S = 50$, $\sigma_n = \$0.05$				
Noise volatility (σ_n)	\$0.05–0.50	56 (0.01)	−34 (0.00)	18 (0.00)
		[3.8–31.3]	[20.9–4.3]	[12.6–23.2]
Maturity (T)	1/12–12/12 yrs	−0.90 (0.02)	−3.4 (0.00)	0.4 (0.53)
		[5.0–3.7]	[23.1–20.3]	[12.1–13.5]
Annual return volatility (σ)	0.10–0.50	−7.4 (0.00)	−11.2 (0.07)	−50.3 (0.00)
		[5.0–2.8]	[17.5–15.5]	[30.2–8.5]
Interval between observations	1–12 h	−0.24 (0.00)	−1.18 (0.00)	−0.73 (0.00)
		[4.3–0.6]	[19.7–2.9]	[12.7–1.1]
Moneyness (K/S)	0.7–1.3	18 (0.00)	42 (0.00)	−2.5 (0.00)
		[1.9–12.9]	[12.4–35.0]	[14.2–12.6]
Panel B: $S = 500$, $\sigma_n = \$0.20$				
Noise volatility (σ_n)	\$0.05–0.50	40 (0.00)	−0.68 (0.48)	2.2 (0.00)
		[0.6–17.7]	[2.8–3.3]	[1.0–2.3]
Maturity (T)	1/12–12/12 yrs	−0.16 (0.03)	−0.88 (0.00)	0.12 (0.32)
		[0.6–0.4]	[3.6–3.0]	[1.3–1.5]
Annual return volatility (σ)	0.10–0.50	−2.4 (0.00)	−13.2 (0.00)	−7.0 (0.00)
		[1.5–0.2]	[7.6–2.0]	[4.2–0.7]
Interval between observations	1–12 h	−0.025 (0.01)	−0.16 (0.00)	−0.072 (0.00)
		[0.5–0.1]	[2.9–0.3]	[1.3–0.1]
Moneyness (K/S)	0.7–1.3	4.2 (0.00)	19 (0.00)	−0.065 (0.67)
		[0.3–3.1]	[1.6–14.6]	[1.2–1.4]

This table contains slope estimates from the regression of the percentage of Type I, II and III errors as defined in Bakshi et al. (2000), on various parameters. The base case for Panel A (B) is a 6-month at-the-money option on a stock whose price is \$50 (\$500) with a return volatility of 30% per year. The risk-free rate is set to 5% and the sampling interval is one hour. For each case the stock follows the process $\Delta S = S\mu\Delta t + S\sigma\Delta z$ where $\mu = 0.10$, $\Delta z = \tilde{z}\sqrt{\Delta t}$ and Δt equals one hour. The amount of noise added to the stock (index) option prices in Panel A (B) is $\sigma_n = \$0.05$ ($\sigma_n = \$0.20$). The range of each parameter is given in the table. The range is divided up into 10 equally spaced intervals and 50,000 simulations are run at each value of the parameter. Then the percentage of Type I, II and III errors are regressed on the parameter value and a constant. The slope estimate from the regression is given in the table. The p-value for the slope is in parentheses next to the estimate. The percentage of each type of error at the low and high end of the parameter interval is given in brackets below each slope estimate

the parameter (100 increments per sample path, 100 sample paths). The resulting estimates for Type I, II and III errors are graphed vs. the parameter values to ensure that the relationship is approximately linear. The percentage of Type I, II and III errors are then regressed on the corresponding parameter value and a constant, and the slope estimate from the regression is given in the table. The p-value corresponding to the slope estimate is in parentheses. The percentage of each type of error at the low and high end of the parameter interval is given in brackets below each slope estimate.

We will focus on the results in panel A. The sensitivity results for the index option in panel B are qualitatively the same as for the individual stock, but the rate of violations tends to be lower. The first parameter of interest is the noise level. At a zero noise level we observe no Type I errors, but as the noise level increases the probability of a Type I error increases. The slope of 56 means that for every one-half tick of additional noise, we observe $(56)(0.05) = 2.8\%$ more Type I errors. Not surprisingly, the Type I error rate is quite sensitive to the amount of noise in option prices. As the noise level in the observed option price increases, the rate of Type II errors decreases (since this counts the number of times that the option price *does not* move when the stock price does), while the rate of Type III errors increases (since this counts the number of times that the option price *does move* when the stock price does not).

The second parameter of interest is the option's maturity. As the maturity gets longer, the price of the option increases and the noise (which has a fixed dollar standard deviation) becomes relatively less important. Hence the Type I error rate decreases. The Type II error rate also decreases. The delta of an ATM call increases slightly as maturity increases, decreasing the chance that the stock price moves and the call price does not, which slightly lowers the Type II error rate. The sensitivity of the Type III error rate to option maturity is not significantly different from zero.

The third parameter of interest is the stock's return volatility, σ. The Type I error rate decreases as volatility increases. As the level of stock volatility increases relative to the noise component, the large price changes swamp the error in observing the price, decreasing the chances of observing the stock price and option price moving in opposite directions. For the same reason, the chance of observing a Type II (III) error, where the stock (option) price moves but the option (stock) price does not move, decreases as volatility increases.

The fourth parameter of interest is how often the prices are sampled. Since we vary the interval of observation up to one-half of a day for these simulations, we use only 20 independent stock-price increments per sample path and increase the option's maturity to one year. This is to ensure that the option's price does not decrease too much due to time decay. Time decay is relevant since the decrease in price due to the time to maturity getting shorter tends to increase the relative effect of a fixed amount of noise added to the option's true price. We find that Type I, II and III errors all decrease as the interval of observation increases from one to 12 h. The intuition here is very similar to that for the stock return volatility. As the time interval increases the variation in stock and option prices between successive observations increases relative to the noise in observing the price, which decreases the rate of all three types of errors.

The last parameter of interest is the option's moneyness, measured by K/S. We vary the moneyness from 0.7 (in-the-money) to 1.3 (out-of-the-money). As the option moves out of the money, both the price of the option and its delta decrease, so a given change in the stock's price results in a very small change in the option's price. This means that a larger proportion of the variation in the observed price of the option is coming from noise rather than true price movements, and the Type I error rate increases. For the same reason, the chances that the observed option price does not move at all increases as delta decreases, increasing the number of Type II violations and decreasing the number of Type III violations.

The results for the index option in panel B are qualitatively the same as in panel A, but the overall error rate for Type I, II and III violations is lower. Compared with the case of an individual stock, the rate of Type I errors for index options appears to be relatively low. However, since index options have larger bid-ask spreads than stock options, it is reasonable to assume that the amount of microstructural noise is higher for index options. A modest amount of noise can result in error rates comparable to what BCC find. For example, while the base case for the index option uses $\sigma_n = 0.05$, If we increase the standard deviation of the noise by only \$0.30, then we would observe a Type I violation rate of $0.6 + (0.30)(40) = 12.6\%$ which is the same order of magnitude as what BCC find for index options.

In sum, microstructural effects can account for a significant portion of the observed cases where the stock price and call price move in opposite directions and the observed phenomena can still be consistent with a univariate diffusion model.

3.2 Transition density approach

The primary consideration when constructing an option pricing model is selecting a stochastic process for the underlying asset's price. There are two main classes of price processes: those that are pure diffusions (with either constant or stochastic volatility) and those that include jumps. In an ideal world, the nature of the asset's price process would be determined by a complete specification of all agent's endowments, utilities and the information filtration process. As a practical matter, however, we must look to historical price data to estimate the asset's stochastic process. The problem that we have is that we only observe the price at discrete intervals, so how can we hope to distinguish whether we should select from the class of processes that are pure diffusions, or those that have jumps?

Aït-Sahalia (2002) presents an elegant solution to this problem which is briefly reviewed here. Let there be two sample paths for the asset's price starting at time t and ending at time $t+\Delta$. The first (second) path starts at a price of x ($\tilde{x}$), with $x < \tilde{x}$. Divide the set of all possible terminal asset prices at time $t+\Delta$ into two regions, Y and $\tilde{Y}$, with all prices in region Y less than those in region $\tilde{Y}$. Aït-Sahalia relies on total positivity of order two as a necessary and sufficient condition for a diffusion. In essence, if the asset's price process starts at a value of $x < \tilde{x}$ at time t, then to be a diffusion the probability that x ends up in region Y and $\tilde{x}$ ends up in region $\tilde{Y}$ must be greater than the probability that x ends up in region $\tilde{Y}$ and $\tilde{x}$ ends up in region Y at time $t+\Delta$ (see Fig. 4, p. 2085 in Aït-Sahalia 2002). This result makes it possible to determine if the process is a pure diffusion by simply examining the transition density function.

If options trade on the underlying asset, the test is very convenient since we can obtain the transition density function directly from the option prices. If $C(K, S, \Delta)$ is the price of a call option with strike K maturing in Δ years on stock with price S, then from Breeden and Litzenberger (1978) we know that the transition density is:

$$p(\Delta, K|S) = \exp(r\Delta)\frac{\partial^2}{\partial K^2}C(S, K, \Delta).$$

From Aït-Sahalia (2002) we know that for the process to be a diffusion, the transition density must satisfy the total positivity condition:

$$\frac{\partial^2}{\partial S \partial K}\ln(p(\Delta, K|S)) > 0. \tag{1}$$

As a practical matter, we need a continuum of option prices at each strike price, but, of course, market data have discrete strike price intervals. Aït-Sahalia overcomes this limit by using the Black–Scholes model to map prices into implied volatilities, using a cubic polynomial in moneyness to capture the variation of volatility with strike price. With volatility as a continuous function of moneyness, the Black–Scholes formula is then used to map volatility back into an option price for any given strike price. We will follow this method here.

One reason that observed option prices can differ from "true" prices is that the prices are rounded to the nearest tick. In U.S. markets the tick interval is currently \$0.05 for options priced lower than \$3.00 and \$0.10 otherwise. To examine the effect of this type of noise, imagine that stock prices follow a diffusion and assumptions of the Black–Scholes are true. We do not observe the true Black–Scholes price, but rather the price that is rounded to the nearest tick. The question that we wish to address is whether this type of noise prevents us from distinguishing whether the true process is a diffusion or contains jumps.

We conduct the following experiment. We generate a series of option prices using the Black–Scholes model and strike-price intervals typical in U.S. markets, for a number of strike prices, $K \in [K_L, K_U]$, then round the prices to the nearest tick using the standard tick interval. The lower bound on rounded prices is set at 1/2 of a tick or \$0.025 to avoid arbitrage possibilities. Using the rounded prices, we follow the procedure in Aït-Sahalia (2002). Specifically, first we use the rounded prices to recover Black–Scholes implied volatilities (IVs). Then, using these IVs, we estimate the volatility as a function of moneyness (K/S) using the following cubic polynomial:

$$\sigma(K/S) = \beta_0 + \beta_1(K/S) + \beta_2(K/S)^2 + \beta_3(K/S)^3. \tag{2}$$

Keeping in mind that volatility is a function of moneyness ($\sigma \equiv \sigma(K/S)$), we recover the risk-neutral transition density by starting with the Black–Scholes equation and taking the second derivative of the call price w.r.t. the strike price:

$$\begin{aligned} p(\Delta, K|S) &= \exp(r\Delta)\frac{\partial^2}{\partial K^2}C(S, K, \Delta) \\ &= \frac{S\sqrt{\Delta}}{\sqrt{2\pi}}\exp\left(-\frac{d_1^2}{2} + r\Delta\right)\left[\frac{d_1}{\sigma}\frac{\partial \sigma}{\partial K}\left(\frac{1}{K\sqrt{\Delta}} + d_2\frac{\partial \sigma}{\partial K}\right) + \frac{\partial^2 \sigma}{\partial K^2}\right] \\ &\quad + \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{d_2^2}{2}\right)\left(\frac{1}{\sigma}\right)\left(\frac{1}{K\sqrt{\Delta}} + d_1\frac{\partial \sigma}{\partial K}\right). \end{aligned}$$

We then apply Aït-Sahalia's test in (1) to determine whether of not the option prices are consistent with a model where the underlying asset follows a diffusion. Given S, r, Δ and $\sigma(K/S)$, if (1) is negative for any strike price $K \in [K_L, K_U]$, we reject the hypothesis that the asset's price process is a diffusion. This, of course, is a Type I error, as by design the option prices were generated from a diffusion model. Intuition tells us that this may be a problem for low-priced stocks that have a short time to maturity since the error induced by rounding the option's price to the nearest tick is largest for these stocks. In fact, the noise can cause an incorrect conclusion even for relatively high priced stocks that have a long time to maturity. For example, consider

Table 3 Total positivity diffusion criteria example

Strike (K)	Black–Scholes		
	Price	Rounded price	Implied volatility
50	15.129	15.10	0.1961
55	11.259	11.30	0.2029
60	7.962	8.00	0.2019
65	5.348	5.30	0.1979
70	3.421	3.40	0.1991

This table contains data used in the example that illustrates the total positivity diffusion criteria from Aït-Sahalia (2002). The Black–Scholes prices are for an option series on a stock whose price is \$60 per share with $\sigma = 0.20$/yr, $T = 1$ year, and $r = 0.10$/yr. The price is rounded to the nearest tick. The Black–Scholes implied volatilities are then recovered using the rounded prices and are used as inputs for the total positivity diffusion criteria example

one-year options on a \$60 stock where $\sigma = 0.20$ and $r = 0.10$. Table 3 contains the Black–Scholes price, price rounded to the nearest tick, and implied volatility for the 5 nearest strikes.

These implied volatilities give us the following implied volatility function:

$$\sigma(K/S) = -1.7538 + 5.8046(K/S) - 5.7006(K/S)^2 + 1.8514(K/S)^3,$$

which makes criterion (1) negative at several strikes. For example $\frac{\partial^2}{\partial S \partial K} \ln(p(\Delta, K|S))|_{K=56} = -0.0076$.

To gain a sense of how serious this problem is, we repeat the above experiment for stock prices ranging from \$20 to \$100 and maturities ranging from two weeks to one year. The results are shown in Fig. 1. If the second partial derivative in (1) is positive for all strike prices $K \in [K_L, K_U]$, then the underlying process is a diffusion and this is indicated by a solid circle at that stock-price-maturity pair. If the partial derivative is non-positive for at least one strike price, then we incorrectly conclude that jumps are present and plot an empty box at that stock-price-maturity pair. Often for low-priced options when the price is rounded to the nearest tick an arbitrage violation occurs and the implied volatility does not exist. If only three or fewer observations are available, then the moneyness-implied volatility relation (2) cannot be estimated and no symbol is plotted at the respective stock-price-maturity pair.

As can be seen in Fig. 1 for short-maturity options on low-priced stocks, the no-arbitrage boundary is often violated and we cannot recover enough implied volatilities to estimate (2). For those stock-price-maturity pairs that have at least four implied volatilities, we often arrive at an incorrect conclusion for options with maturities up to four months, with the problem being worse for low-priced stocks. As the maturity gets longer and the stock price gets higher, committing Type-I errors becomes less frequent, but it still happens fairly often. For example, we arrive at incorrect conclusions for 3 out of 17 stock-price levels for the three longest maturities of 48, 50 and 52 weeks.

Adding more strike prices does not help the problem. For example, if we increase the number of available strikes from 5 to 7, we are adding deep out-of-the-money and in-the-money options where the vega is very small. In these cases a small change in

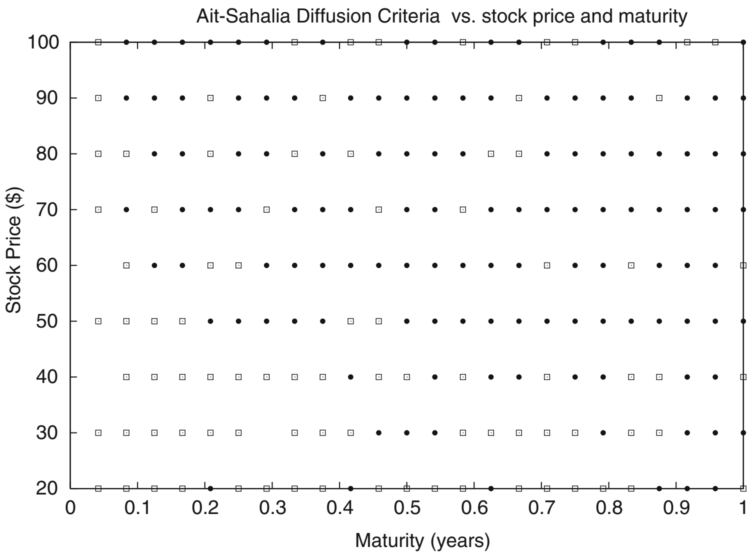

Fig. 1 Tick size calibrated to US option markets. Key: Hollow square = diffusion criteria violated at least one strike price. Solid dot = diffusion criteria not violated. No symbol = less than four options were available due to arbitrage violations. True volatility = 0.20 per year, risk-free rate = 0.10. Strike spacing conforms to the standard used in US markets

price due to rounding to the nearest tick causes a large change in implied volatility, resulting in the implied volatility-moneyness relation (2) deviating even more from the constant-volatility case.

As the tick size is reduced, the microstructural noise induced by rounding is also reduced. In the limit if we reduce the tick size to 0 we find criteria (1) is positive in all cases. To study the case of a very small, but finite, tick size we reduce the tick size to one cent and repeat the above experiment. We find that while the problem caused by the noise is much less than in the previous example, errors are still made for low-priced stocks (less than 80) and very short time to maturities (less than 3 months). Specifically, violations only occur for stock price (S) and maturity (t) combinations where $S < -200t + 80$, which are those combinations below the line connecting (S, t) points (0, 80) and (0.3, 20) in Fig. 1. However this type of noise may not be a problem for S&P 500 index options. An asset price of 1000 with an option-price tick size of 0.1 (the index) is a scaled transformation of an asset price of 100 with a option-price tick size of 0.01, and we find no that no incorrect conclusions are drawn at this price-level and tick-size combination.

3.3 Speed of convergence approach

An alternative approach to testing the type of stochastic process that the underlying asset follows is developed by Carr and Wu (2003). Whereas Aït-Sahalia's approach requires a cross-section of option prices to estimate the risk-neutral PDF, Carr and Wu propose an approach based on the speed of convergence to zero of at-the-money and out-of-the-money option prices as they approach maturity. If the stochastic process is a pure diffusion, then the probability of the option finishing in-the-money as the time to maturity approaches zero is very small and the price decays to zero relatively quickly. However, if the process includes jumps, then the probability that the option finishes in-the-money is higher and the price does not decay to zero as fast. By observing the rate at which the prices of OTM and ATM options decay to zero, Carr and Wu develop a test that allows one to distinguish between alternative specifications for the asset's stochastic process. Their main result is summarized in Table 1 of their paper, which contains the order of the speed of convergence to zero for each type of option and process type. For OTM options, the speed of convergence for diffusion, pure jump, and jump-diffusions processes is $O(e^{-c/T}), c > 0$; $O(T)$; and $O(T)$, respectively. For ATM options, the speed of convergence for diffusion, pure jump, and jump–diffusions processes is $O(\sqrt{T})$; $O(T^p) p \in (0, 1]$; and $O(T^p)$, $p \in (0, 1/2]$, respectively.

We are concerned about the effect of microstructural noise for two reasons. First, the noise may be so large that it is not possible to reject the null hypothesis. Second, it may bias the results. Since short maturity ATM and OTM options have low prices, noise such as discrete tick size and bid-ask bounce may make a pure diffusion process appear to have jumps, or a mixed jump–diffusion process appear as a pure jump process.

To investigate this, we conduct the following experiment. First, we generate short-maturity option prices using one of three models: Black–Scholes, jump-diffusion (Merton 1976), and a pure jump model which is the limiting case of Merton's model as the variance of the diffusion component goes to zero. Prices were generated on the domain $\ln(T) \in [-4, -2]$ at an interval of $\Delta \ln(T) = 0.2$, corresponding from 7 to 50 days to maturity. In each case, prices are generated for at-the-money calls and for out-of-the-money calls with a strike price of 1.1 times the current stock price. We compute both the exact theoretical prices and the theoretical price rounded to the nearest tick, calibrating the tick size to that currently used on the six option exchanges. If a price of zero was generated due to rounding down, that observation was dropped.

For each model, prices were generated assuming a risk-free rate of 5% per year. For the Black–Scholes model, we use a volatility of 20% per year. For the pure jump model, we use a total volatility of 30% per year, $\lambda = 1$ and $\gamma = 1$ (all the volatility due to jumps). We use the same parameters for Merton's jump–diffusion model, except that we set $\gamma = 0.5$ (half the total volatility is due to jumps). As in Carr and Wu (2003), we plot the log of the ratio of the option price to maturity vs. the log of the time to maturity in a term-decay plot. The rate of decay at a particular T is the slope of the term decay plot at that time. To estimate the slope, we regress $\ln(P/T)$ on $\ln(T)$ using the following form:

$$\ln(P/T) = a_0 + a_1 \ln(T) + a_2 \ln(T)^2$$

and estimate the limiting rate by evaluating the slope of the term decay plot at $\ln(T) = -4$ (7 days to maturity). The results are shown in Table 4.

Using actual prices, the estimated slope for the Black–Scholes, pure jump and jump–diffusion models are essentially identical to Carr and Wu (2003) (see Table 2 in their paper). When microstructural noise is present, however, the results are mixed. In the case of ATM options when the underlying process is a pure diffusion, we would

Table 4 Asymptotic slope estimates for the prices of short-maturity options

	ATM		OTM	
	Slope	R^2	Slope	R^2
Black–Scholes—Actual prices	−0.488 (0.002)	1.000	2.554(0.515)	0.999
Black–Scholes—Discrete ticks	−0.509 (0.255)	0.997	1.163(3.714)	0.923
Pure jump—Actual prices	−0.003 (0.007)	0.999	0.001(0.001)	1.000
Pure jump—Discrete ticks	0.013 (0.677)	0.388	0.442(2.042)	0.321
Jump diffusion—Actual prices	−0.438 (0.002)	1.000	0.001(0.001)	1.000
Jump diffusion—Discrete ticks	−0.413 (0.257)	0.997	−0.020(2.286)	0.683

This table contains slope estimates from the term decay plot of $\ln(P/T)$ vs. $\ln(T)$ using the Black–Scholes, Merton's pure jump and Merton's jump–diffusion models. Prices were generated on the domain $\ln(T) \in [-4, -2]$ at an interval of $\Delta \ln(T) = 0.2$, corresponding from 7 to 50 days to maturity. OTM calls have a strike price of 1.1 times the current stock price. Both actual prices and prices rounded to the nearest tick were used. If a price of zero was generated due to rounding down, that observation was dropped. Prices generated from the option pricing models were fitted using the following specification: $\ln(P/T) = a_0 + a_1 \ln(T) + a_2 \ln(T)^2$ and the slope of the term decay plot is determined by evaluating $a_1 + 2a_2 ln(T)$ at $\ln(T) = -4$. Standard errors are in parentheses

expect a slope of −0.5, which is what we find. For OTM options we would expect a statistically significant positive slope since, for small T, terms of order $e^{-c/T}$ grow quite rapidly as T increases. While the slope is 1.163, we cannot reject the null that it is different from zero at any reasonable level of significance. The R^2 is quite good for both the pure diffusion cases.

In the case of ATM options where the underlying stochastic process is a pure jump process, the rate of decay is order T^p and we expect p to be one since the jump component has finite variation. This means that the decay is of order T and the limiting slope of the term decay plot should be zero. We find that the slope is 0.013 and is not significantly different than 0. However, there is significant noise introduced by the price discretization. The R^2 is only 0.39 and the standard error is 0.677 so though the slope is not statistically significantly different from 0, it is not that precise. For example, we cannot reject the null that the slope is equal to −0.5, making it difficult to distinguish from the pure diffusion case. In the case of out-of-the-money options under the pure jump model, we expect a decay of order T and a limiting slope of zero, but the same criticism applies as in the case of the at-the-money option; the noise is so great that we cannot reject the null that the slope is positive.

The last row in Table 4 contains the results from Merton's jump–diffusion model. With both a jump and a diffusion component, the convergence speed is dominated by the component with the slowest convergence to zero (i.e. the least positive limiting slope from the term decay plot). Thus, for the ATM case we predict a coefficient of −0.5. We find a coefficient of −0.413, which is not statistically different from −0.5 at the 5% level, but is statistically different from 0 (the pure jump case). Hence, we can detect the diffusion component in the presence of microstructural noise. In the case of OTM options, the decay is of order T so we would predict a limiting slope of the term decay plot of zero. We find a coefficient of −0.020 and cannot reject the null that it is zero. However, due to the large standard error that the noise introduces, we cannot reject the null that the slope is positive (pure diffusion case), so the noise make it difficult to draw the correct inference.

In sum, the noise introduced by price discretization does not affect making the correct inference in the pure diffusion Black–Scholes case, but it does make it difficult to make the correct inference in the pure jump case. The presence of a diffusion component can be detected in the presence of the noise, but the noise makes it difficult to distinguish the jump–diffusion from the pure diffusion models.

4 Implied risk-neutral moments

A substantial body of literature has addressed the question of how to extract from observed option prices an estimate of the risk-neutral density of the underlying asset returns (see Jackwerth (1999) for a survey). Other authors, such as Bakshi et al. (2003) (BKM) have suggested methods for estimating the higher moments of the risk-neutral density, without having to estimate the entire density. Dennis and Mayhew (2002) have used the BKM technique to examine the cross sectional determinants of risk-neutral skewness, and Bakshi and Cao (2004) have used the approach to further explore risk-neutral kurtosis, for individual stock options.

In their paper, BKM derive a relation between the moments of the risk-neutral density and the price of securities with quadratic, cubic, and quartic payoff functions, which can be replicated using a portfolio of out-of-the-money calls and puts with a continuum of strike prices. Specifically, they demonstrate that the risk-neutral skewness and kurtosis of the continuously-compounded return distribution at horizon τ are given by

$$SKEW(t,\tau) = \frac{e^{r\tau}W(t,\tau) - 3\mu(t,\tau)e^{r\tau}V(t,\tau) + 2\mu(t,\tau)^3}{\left[e^{r\tau}V(t,\tau) - \mu(t,\tau)^2\right]^{3/2}}$$

$$KURT(t,\tau) = \frac{e^{r\tau}X(t,\tau) - 4\mu(t,\tau)e^{r\tau}W(t,\tau) + 6\mu(t,\tau)^2V(t,\tau) - 3\mu(t,\tau)^4}{\left[e^{r\tau}V(t,\tau) - \mu(t,\tau)^2\right]^2}$$

where the prices of the expected return, volatility, cubic, and quartic payoffs are given by:

$$\mu(t,\tau) \approx e^{r\tau} - 1 - \frac{e^{r\tau}}{2}V(t,\tau) - \frac{e^{r\tau}}{6}W(t,\tau) - \frac{e^{r\tau}}{24}X(t,\tau),$$

$$V(t,\tau) = \int_{S(t)}^{\infty} \frac{2\left(1 - \ln\left[\frac{K}{S(t)}\right]\right)}{K^2} C(t,\tau;K)dK + \int_{0}^{S(t)} \frac{2\left(1 + \ln\left[\frac{S(t)}{K}\right]\right)}{K^2} P(t,\tau;K)dK,$$

$$W(t,\tau)=\int_{S(t)}^{\infty}\frac{6\ln\left[\frac{K}{S(t)}\right]-3\left(\ln\left[\frac{K}{S(t)}\right]\right)^2}{K^2}C(t,\tau;K)dK$$
$$-\int_{0}^{S(t)}\frac{6\ln\left[\frac{S(t)}{K}\right]+3\left(\ln\left[\frac{S(t)}{K}\right]\right)^2}{K^2}P(t,\tau;K),dK$$
$$X(t,\tau)=\int_{S(t)}^{\infty}\frac{12\left(\ln\left[\frac{K}{S(t)}\right]\right)^2-4\left(\ln\left[\frac{K}{S(t)}\right]\right)^3}{K^2}C(t,\tau;K)dK$$
$$+\int_{0}^{S(t)}\frac{12\left(\ln\left[\frac{S(t)}{K}\right]\right)^2+4\left(\ln\left[\frac{S(t)}{K}\right]\right)^3}{K^2}P(t,\tau;K)dK.$$

The prices of these polynomial contracts represent weighted sums of prices of out-of-the-money calls and puts, across a continuum of strike prices. Thus, the risk-neutral moments can be estimated directly from the prices of traded options. The accuracy of the resulting estimates depends on the accuracy with which the traded option prices are observed. Even if we can observe option prices without error, however, there can still be error in the risk-neutral moment estimation from two sources—the fact that the interval between adjacent strike prices is not infinitely small and the fact that we do not observe all strike prices from 0 to ∞.

Since options do not trade with a continuum of strike prices, the interval between strike prices matters since we have to approximate the integral with a sum. Namely, due to contract specifications in the index and stock options market, ΔK is usually \$2.50 or \$5. To examine the impact of a discrete strike price interval, assume that we are in a Black–Scholes world and can observe option prices with no noise. Furthermore, assume that there is a stock that has a price of 50, a return volatility of 20% per year, the riskless rate is 5% per year, and we can observe option prices that have strikes from 30 to 70. If we set ΔK to some small number, say \$0.10, and compute the risk-neutral skewness, we recover the correct value of 0. If we increase ΔK from this value and repeat the experiment, the risk-neutral skewness gradually diverges from 0 in an oscillatory fashion shown in Fig. 2.

Furthermore, the economic effect is large. Since we are starting with the Black–Scholes model, the skewness should be zero, yet for strike price intervals in the range of \$5 the error in skewness can be plus or minus 0.4, causing the researcher to reject the null of zero skewness even though it is true.

Examining Fig. 2, we see that the skewness looks like it is zero at a strike price interval of 5, in which case the BKM metric is unbiased at the strike price interval that we observe most frequently in the market. However, the value of the skewness at this strike price interval is actually −0.026; that it appears to be zero is simply due to the scale of the graph. More importantly, the skewness and kurtosis metrics are homogeneous of degree 0 with respect to K, ΔK and S. If we compute the skewness metric at αK, $\alpha\Delta K$, and αS for some positive constant α, the values of the integrand

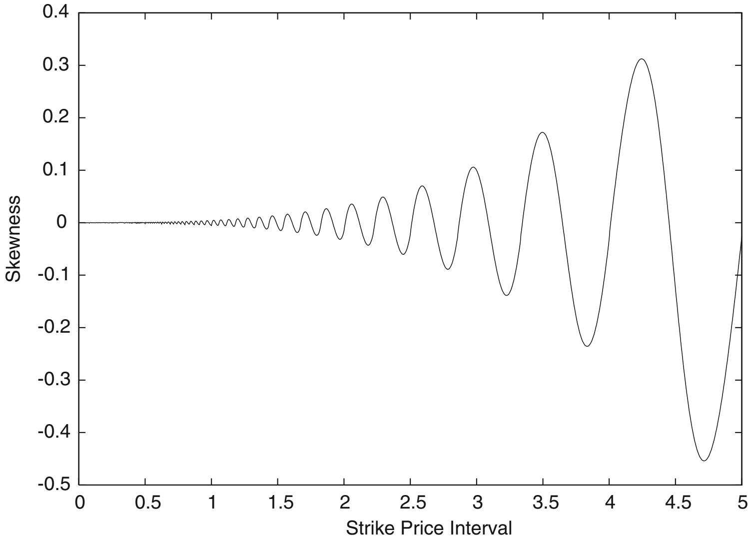

Fig. 2 Skewness vs. strike price interval. This figure contains a plot of the risk-neutral skewness computed by the method of Bakshi, Kapadia and Madan vs. the strike price interval. Strike prices are generated at intervals of ΔK from a low strike of 30 to a high strike of 70. The strike price interval is initially set to $\Delta K = 0.1$ and then increased to $\Delta K = 5$ in increments of 0.1. For each set of strike prices, corresponding option prices are generated from the Black–Scholes model with a time to maturity of one month, a risk-free rate of 5%, a stock price of \$50, and a volatility of 20% per year. The BKM method is then used to compute the risk-neutral skewness from these prices

of the volatility, cubic, and quartic payoffs will be the same as if it were evaluated at K, ΔK and S, and, as a result, we will get the same skewness and kurtosis. This means that evaluating the skewness at a stock price of \$50 and a strike price interval of \$4.75 will be the same as evaluating the skewness at a stock price of $(\$50)\left(\frac{5}{4.75}\right) = \$52.63$ and a strike price interval of \$5. In either case we would (incorrectly) measure the skewness as -0.45.

Above we assumed that we could measure the option price with no error and had a large set of strike prices (from 30 to 70) available to us. The only error was introduced by the discrete strike price interval. In option markets, however, often we do not have a full set of strike prices available, either because a contract does not exists at a certain strike price, or a contract exists but is relatively illiquid. Next we examine the impact of limiting the number of observed strike prices on the risk-neutral skewness. Our method is essentially the same as before, except we fix the strike price interval at a very small value of \$0.10 and gradually decrease the number of strike prices available. The domain of available strikes is from $50 - w$ to $50 + w$. Initially we set the domain half-width, w at 20 and then gradually decrease the domain half-width from 20 down to 1. The results for skewness are shown in Fig. 3. At a large domain half-width, the skewness is 0, however as the number of available strikes drops, the skewness is biased downward.

We repeated the above experiments to examine the effect on kurtosis and found similar results. At small strike price intervals, the kurtosis is 3.0, but as the strike price

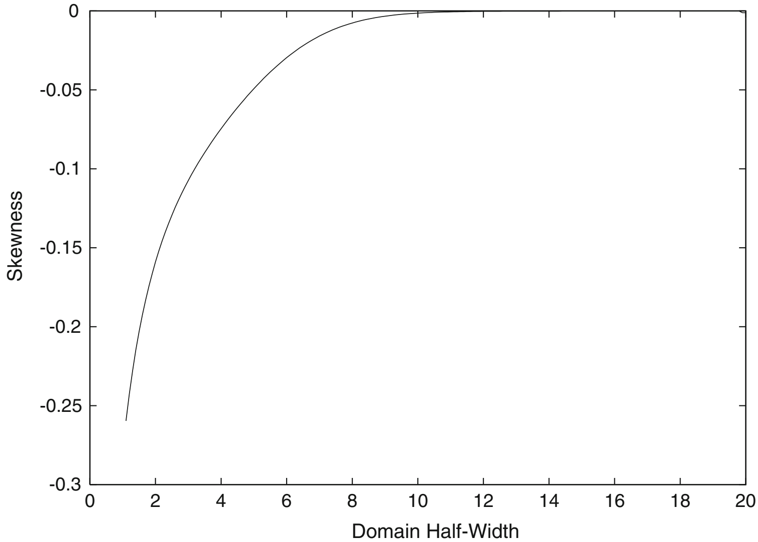

Fig. 3 Skewness vs. domain half-width. This figure contains a plot of the risk-neutral skewness computed by the method of Bakshi, Kapadia and Madan vs. the half-domain width (w). Strike prices are generated at intervals of 0.1 from a low strike of $50 - w$ to a high strike of $50 + w$. The domain half-width is initially set to 20, and then decreased in steps of 0.1 to a value of 1.0. For each set of strike prices, from $50 - w$ to $50 + w$ corresponding option prices are generated from the Black–Scholes model with a time to maturity of one month, a risk-free rate of 5%, a stock price of \$50, and a volatility of 20% per year. The BKM method is then used to compute the risk-neutral skewness from these prices

interval is increased from \$0 to \$5, the kurtosis oscillates around 3.0 in a manner similar to Fig. 2, with the amplitude off the oscillation increasing with strike price interval. At a strike price interval of \$5, the kurtosis swings between 1.0 and 5.5. Also, when the strike-price domain half with is large ($w = 20$), kurtosis is 3.0, but as the domain half-width decreases, kurtosis montonically decreases in a manner similar to Fig. 3, with kurtosis equal to 2.9 at $w = 8$ and 1.0 at $w = 1$.

It is important to note that the only error that we have introduced has been in the form of a coarse strike price interval and a limited number of strike prices. We have assumed that prices come from a Black–Scholes world and are observed without error. In reality, of course, error in option prices is present and can come from many sources. This additional noise makes it more difficult to distinguish between option pricing models on the basis of the moments of the risk-neutral density.

5 Comments and conclusions

Bid-ask spreads in option markets tend to be large, compared to the differences between theoretical predictions of alternative models. For various microstructural reasons, spread midpoints are not only noisy, but are likely to be biased estimates of an option's value. Microstructural features such as discrete prices, the random arrival of limit orders, and a wide interval between adjacent strike prices make it difficult to perform reliable empirical tests of option pricing models that rely on observed option prices, especially options on low-priced stocks.

The theoretical literature provides us a wide variety of potential option pricing models. Many of these models are characterized by a large number of parameters, and the models differ from each other in fairly subtle ways. Rather than focusing our attention on the latest, most sophisticated models, we have tried to address the question of testability on a more fundamental level. We have shown that it is often very difficult to distinguish between simple models that differ in basic ways. Disentangling models with more subtle differences must present an even bigger challenge. If we cannot differentiate between a constant volatility model with and without jumps, how can we hope to distinguish a stochastic volatility jump model with stochastic jump intensity from a stochastic volatility jump model with constant jump intensity?

In the body of this paper, we have pointed out a number of potential problems that may arise in empirical tests of option pricing models. Here, we will offer some additional insights, and summarize some of our main points, trying along the way to offer constructive suggestions as to how future research may be improved in light of the potential problems.

A recurring theme in this paper is that the noise in observed option prices can severely hamper the power of empirical techniques. This problem has been recognized in the past—the response has been correct but insufficient. It makes sense to focus our attention on only the most actively-traded options, use midpoints instead of trade prices, and avoid using closing prices. Beyond this, we can attempt to reduce the noise in observed prices through careful use of intraday data. For example, we can attempt to come up with a more accurate estimate of current option prices by applying time-series techniques to high-frequency intraday data.

Researchers studying individual stock options should be especially careful to recognize the potential biases associated with the tick size and wide interval between adjacent strikes, and should consider using only higher-priced stocks. Bid-ask spreads on individual stock options have became considerably tighter since 1999, as a result of competitive forces in the industry. Also, quotes are updated more frequently than they were in the mid 1990s and before. This suggests it may be a good idea to use current data rather than the Berkeley Options Data Base. On the other hand, some of the most actively traded options in recent years have been on stocks with prices below $20.00, where microstructural problems can be severe.

Let us now comment on the idea of testing an option model by testing the performance of a trading strategy based on the model. Even in the most liquid exchange-traded option markets, bid-ask spreads tend to be sufficiently wide that the round-trip trading cost of a strategy implemented through market orders would obliterate the trading profits of all but the most extremely profitable strategies. This is particularly true for strategies involving out-of-the-money options. Thus, a key question is the degree to which such trading strategies could be implemented at lower cost by option trading firms with a seat on an exchange, or by investors through the use of limit order trading. We believe that a careful study of limit order execution quality in the option market would be particularly valuable.

Finally, we would like to emphasize that all of the potential problems we discuss in this paper result from difficulty in observing option prices, not underlying stock prices.

While there may also be errors associated with observing underlying stock prices, we submit that these are likely to be much less important, perhaps an order of magnitude smaller. Bid-ask spreads for actively-traded stocks have become very tight in recent years, often they are down to one penny. As transaction costs have plummeted, it has become much cheaper to implement dynamic trading strategies. Therefore, we suggest that a productive direction for future research would be to focus on directly testing option models by testing the accuracy of the assumed equation describing the dynamics of the underlying asset, and, in models where it would be appropriate to do so, testing the performance of dynamic replicating strategies.

References

Aït-Sahalia, Y. (2002). Telling from discrete data whether the underlying continuous-time model is a diffusion. *Journal of Finance, 57*, 2075–2112.

Bakshi, G., & Cao, C. (2004). *Risk neutral kurtosis, jumps, and option pricing: Evidence from most actively traded firms*. Working paper, University of Maryland.

Bakshi, G., Cao, C., & Chen, Z. (2000). Do call prices and the underlying stock always move in the same direction? *Review of Financial Studies, 13*, 549–584.

Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies, 16*, 101–143.

Bates, D. S. (2000). Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics, 94*, 181–238.

Bates, D. S. (2003). Empirical option pricing: A retrospection. *Journal of Econometrics, 116*, 387–404.

Black, F., & Scholes, M. (1972). The valuation of option contracts and a test of market efficiency. *Journal of Finance, 27*, 399–417.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*, 637–654.

Bliss, R., & Panigirtzoglou, N. (2002). Testing the stability of implied probability density functions. *Journal of Banking & Finance, 26*, 381–422.

Breeden, D., & Litzenberger, R. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business, 51*, 621–651.

Canina, L., & Figlewski, S. (1993). The informational content of implied volatility. *Review of Financial Studies, 6*, 659–681.

Carr, P., & Wu, L. (2003). What type of process underlies options? A simple robust test. *Journal of Finance, 58*, 2581–2610.

Chan, K., & Chung, Y. P. (2000). *Asymmetric price distribution and bid-ask quotes in the stock options market*. Working paper, Hong Kong University of Science and Technology.

Christensen, B., & Prabhala, N. (1998). The relation between implied and realized volatility. *Journal of Financial Economics, 50*, 125–150.

Cochrane, J. H., & Saa-Requejo, J. (2002). Beyond arbitrage: Good deal asset price bounds in incomplete markets. *Journal of Political Economy, 108*, 79–119.

Dennis, P., & Mayhew, S. (2002). Risk-neutral skewness: Evidence from stock options. *Journal of Financial and Quantitative Analysis, 37*, 471–493.

Figlewski, S. (1989). Options arbitrage in imperfect markets. *Journal of Finance, 44*, 1289–1311.

Figlewski, S. (1997). Forecasting volatility. *Financial Markets, Institutions, and Instruments, 6*, 1–88.

Hentschel, L. (2003). Errors in implied volatility estimation. *Journal of Financial and Quantitative Analysis, 38*, 779–810.

Jackwerth, J. C. (1999). Option-implied risk-neutral distributions and implied binomial trees: A literature review. *Journal of Derivatives, 7*, 66–82.

Longstaff, F. A. (1995). Option pricing and the martingale restriction. *Review of Financial Studies, 8*, 1091–1124.

Macbeth, J. D., & Merville, L. J. (1979). An empirical examination of the Black–Scholes call option formula. *Journal of Finance, 34*, 1173–1186.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics, 3*, 125–144.

Penttinen, A. (2001). *The sensitivity of implied volatility to expectations of jumps in volatility: An explanation to the illusory bias in implied volatility as a forecast*. Working paper, Swedish School of Economics.

Phillips, S., & Smith, C. (1980). Trading costs for listed options: The implications for market efficiency. *Journal of Financial Economics, 8*, 179–189.

Poteshman, A. M. (2000). *Forecasting future volatility from option prices*. Working paper, University of Illinois at Urbana-Champaign.

Rubinstein, M. (1985). Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978. *Journal of Finance, 40*, 455–480.

Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance, 49*, 771–818.