Tests of Recent Advances in Extracting Information From Option Prices

J. V. Healy

A. Gregoriou

Brighton Business School, University of Brighton.

R. Hudson*

Hull University Business School

September 2017

Abstract

A large literature exists on techniques for extracting probability distributions for future asset prices from option prices. No definitive method has been developed however. The parametric 'mixture of normals', and non-parametric 'smoothed implied volatility' methods remain the most widespread approaches. These though are subject to estimation errors due to discretization, truncation, and noise. Recently, several authors have derived 'model free' formulae for computing the moments of the risk neutral density (RND) directly from option prices, without first estimating the full density. The accuracy of these formulae is studied here for the first time. The Black-Scholes formula is used to generate option prices, and error curves for the first 4 moments of the RND are computed using the 'model-free' formulae. It is found that, in practice, the formulae are prone to large and economically significant errors, because they contain definite integrals that can only be solved numerically. We show that without mathematically equivalent expressions with analytical solutions the formulae are difficult to deploy effectively in practice.

Keywords: Option Pricing, Risk neutral moments, risk neutral density, analytical solutions.

JEL Classifications: C14, G11, G12.

---

* Corresponding author, Robert Hudson. Address: Hull University Business School, University of Hull, Hull HU6 7RX Tel +44 (0) 1482 46 3189 E-Mail: robert.hudson@hull.ac.uk

1.   Introduction

Methods for extracting implied probability distributions for the prices or returns of an asset at a future time, from series of synchronously observed market prices of options on the asset, have been extensively studied since the mid 1990s [see e.g. Bliss and Panigirtzoglou (2002), Jackwerth (2004), and more recently Figlewski (2007) for reviews]. Breeden and Litzenberger (1978) made explicit the exact relationship between option prices and the risk neutral density (RND) [see Appendix A for details and proof]. In the risk-neutral pricing framework the price of an option is equal to its discounted expected payoff under the risk neutral measure. Evaluating the integral of the payoff function over the risk neutral measure and discounting at the risk free rate can thus price an option. Given a continuum of observed option prices, this pricing calculation can be inverted for European exercise options, and the full RND for the price (return) of the underlying asset at maturity extracted. Useful information contained in the shape of the distribution can thus be recovered. RNDs have numerous important applications in finance. These include: Pricing securities [Cox and Ross (1976)]; Estimating value-at-risk (VaR) for risk management purposes [Ait-Sahalia and Lo (2000)]; Studying risk aversion and risk preferences [Bliss and Panigirtzoglou (2004)]; Assessing financial market expectations regarding future asset prices, interest rates, and exchange rates, in connection with setting monetary policy [Lynch and Panigirtzoglou (2008)]. However, existing methods for extracting RNDs are variously, computationally cumbersome, data intensive, and or subject to estimation errors due to discreteization, truncation, and noise issues in the raw options data. No definitive method has been developed, but two approaches are popular with practitioners, namely, the mixture of normals [Ritchey (1990)], and the smoothed implied volatility method [Shimko (1993)]. Tests suggest the latter method produces

better results [see e.g. Bliss and Panigirtzoglou (2002) and Andersson and Lomakka (2003)].

In many applications it is enough to know the first four moments of the RND. Hence a more parsimonious representation will suffice. Based on recent theoretical developments, several authors have derived exact formulae for computing the moments of the RND directly from option prices without first estimating the full density distribution. These formulae have the advantage of being 'model free', in the sense of not being subject to the assumptions of any option pricing model. New approaches for extracting the RND, by using these formulae to compute its moments in a first step, have also been developed. Of course, when these formulae are applied to observational data, they also are subject to estimation errors due to discretization, truncation, and noise issues in the data. Jiang and Tian (2007), and Dennis and Mayhew (2002, 2009), have studied the errors arising from discrete implementation of the 'model free' implied variance, and the implied skewness and kurtosis respectively, and show that they are economically significant. What is perhaps less well appreciated is that the solutions to the formulae themselves exhibit sensitivity to their inputs, even for realistic ranges of values, and are thus capable of being biased estimates.

The contribution of this paper is to demonstrate that applying the 'model free' formulae for the first four moments of the RND produce large and economically significant errors independently of those resulting from the observation issues discussed above. This is shown by solving the formulae as exactly as possible in a continuous strike price framework, for realistic ranges of inputs and constructing error

curves. This is important because it shows that the formulae are of limited applicability in their current forms.

The remainder of the paper is organised as follows. Section 2 reviews the literature on 'model free' implied moments of the RND. Section 3 outlines the methodology used. Section 4 presents the findings. Section 5 contains a summary and conclusions.

2. 'Model free' Implied Moments: Literature Review

The development of "model-free" methods of directly extracting the moments of the RND has emerged from three separate strands of research. First, work on the log contract, volatility, and variance swaps [Neuberger (1994), Carr and Madan (1998), Demeterfi et al (1999)]. Second, extraction of information on the underlying price processes from option prices [Derman and Kani (1994), Britten-Jones and Neuberger (2000)]. Third, studies of the characteristic function of the state price density (discounted RND), as an alternative spanning entity to options for pricing other securities [Bakshi and Madan (2000), Bakshi, Kapadia, and Madan (2003)].

Demeterfi et al (1999) show how hedging an option on the logarithm of the price of an underlying asset (the log contract), provides a payoff equal to the variance of the asset's returns. No such contract is traded in practice; however the log contract can be replicated by a portfolio of European exercise options with a continuous range of strikes and maturities. This portfolio has a value equal to the payoff of the log contract. Dynamically hedging a log contract therefore captures realized variance

(volatility). The value of a variance swap, a forward contract F on future realized variance with strike K, depends on the future payoff $\left(\sigma_R^2 - K_{VAR}\right) \times N$ discounted to its present value under the risk neutral measure, where σ_R^2 is realized variance, N is the notional value and K_{VAR} is the price of variance. K_{VAR} is equal to the value of the portfolio that replicates the log contract. Demeterfi et al (1999) derive formulae for valuing and pricing the variance swap, and directly obtaining the cost of the replicating portfolio. The key result, a formula for the fair value of future variance is given as equation (26) of their paper, and shown here as equation (1).

$$K_{\mathrm{var}} = \frac{2}{T}\left(rT - \left(\frac{S_0}{S_*} e^{rT} - 1 \right) - \log\frac{S_*}{S_0} + e^{rT}\int_0^{S_*}\frac{1}{K^2}P(K)dK + e^{rT}\int_{S_*}^{\infty}\frac{1}{K^2}C(K)dK \right) \quad (1)$$

In equation (1) K_{var} is the fair price of future variance. S_0 is the underlying asset price at time 0. K is the strike price. $P(K)$ and $C(K)$ are the prices of out of the money calls and puts. S_* is the value of the underlying asset at the boundary between the calls and the puts (e.g. at the money). T is the maturity of the option.

Britten-Jones and Neuberger (2000) demonstrate how, given a continuum of European option prices with strikes and maturities ranging from zero to infinity, a condition can be derived which must be satisfied by all price processes consistent with the given set of option prices. Derman and Kani (1994), Dupire (1994, 1997), and Rubinstein (1994) showed that when volatility is deterministic, a unique price process exists that is consistent with option prices. Britten-Jones and Neuberger extended this analysis to a non-deterministic volatility setting, where many consistent price processes are possible. They derived their results in a discrete framework, using a time-price grid, and took limits as the interval sizes approach zero to obtain continuous counterparts.

In an appendix, they also derived their results directly in a diffusion setting. Britten-Jones and Neuberger's simple condition is given as Equation (10) in Proposition 1 of their paper and is shown here as equation 1a[1].

$$E\left[\left(\frac{S_{t+h}-S_t}{S_t}\right)^2 \mid S_t = K\right] = \frac{[C(t+h,K)-C(t,K)](u-1)^2(u+1)/u}{C(t,Ku)-(1+u)C(t,K)+uC(t,K/u)} \qquad (1a)$$

In equation (1a) $C(t, K)$ is the call option price at strike price K and future time t, h is the size of the discrete time intervals used in the setting and u is the geometric factor acting on stock prices that determines the possible stock prices given the discrete time intervals.

The authors show that all price processes satisfying their Proposition 1 have the same (risk neutral) expectation for squared price volatility (e.g. price variance) over any given time period, and thus imply the same one-period forecast of volatility. Because this forecast is common to all such processes they refer to it as the "model-free" implied volatility. The analytical formula needed to extract the "model-free" implied volatility for a period between any two arbitrary future dates, from current prices of options expiring on those dates is given as equation (13) of Proposition 2 of their paper shown here as equation 1b[2]. The authors note that this equation was derived

---

[1] PROPOSITION 1: In any continuous risk-neutral process, the expectation of squared return, conditional on the stock price and time, is determined by the initial option prices as

$$E\left[\left(\frac{S_{t+h}-S_t}{S_t}\right)^2 \mid S_t = K\right] = \frac{[C(t+h,K)-C(t,K)](u-1)^2(u+1)/u}{C(t,Ku)-(1+u)C(t,K)+uC(t,K/u)} \qquad (10)$$

The converse is also true; any continuous martingale process for S that satisfies the above condition for all $K \in \mathbf{K}$ and $t \in \mathbf{T}$ will price all European options correctly by their expected payoffs.

[2] PROPOSITION 2: The risk-neutral expected sum of squared returns between two arbitrary dates t_1 and t_2 is given from the set of prices of options expiring on these two dates as

independently by Carr and Madan (1998) in the context of pricing and hedging variance swaps, using results from Neuberger (1994), and the well- known Breeden and Litzenberger (1978) result.

$$E_0\left[\int_0^{t_2}\left(\frac{dS_t}{S_t}\right)^2\right]=2\int_0^\infty\frac{C\left(t_2,K\right)-C\left(t_1-K\right)}{K^2}dK \tag{1b}$$

A simplification of equation 1b for the period between the current time and any arbitrary future date, is given as equation (14) of Britten-Jones and Neuberger (2000), and is reproduced here as equation (2).

$$E_0\left[\int_0^{t_2}\left(\frac{dS_t}{S_t}\right)^2\right]=2\int_0^\infty\frac{C\left(t_2,K\right)-\max\left(S_0-K,0\right)}{K^2}dK \tag{2}$$

In equation (2) max(S_0- K,0) is the intrinsic value of the option at time 0.

Bakshi and Madan (2000) observe that though the payoff functions of other securities are spanned by options, this has not resulted in a simplification of security valuations, because options themselves are complex to value. They propose the use of an alternative spanning entity. Namely; the characteristic function of the state price density (SPD), which they argue, significantly simplifies option pricing. The SPD is the discounted risk-neutral density function, and its characteristic function can be obtained via a Fourier transform[3]. Theorem 1 of Bakshi and Madan (2000) demonstrates that in an arbitrage free setting, the continuum of characteristic functions and the continuum of options are equivalent classes of spanning securities.

$$E_0\left[\int_0^{t_2}\left(\frac{dS_t}{S_t}\right)^2\right]=2\int_0^\infty\frac{C\left(t_2,K\right)-C\left(t_1-K\right)}{K^2}dK \tag{13}$$

[3] Recent research discusses how the use of wavelets has advantages over Fourier transforms in option pricing (Ortiz-Gracia, and Oosterlee, 2013).

It follows as a special case of Theorem 1 that all twice differentiable payoff functions can be algebraically spanned by a continuum of out-of-the-money calls and puts.

Bakshi, Kapadia and Madan (2003), use the above result to derive formulae for the second, third, and fourth moments of the RND. They demonstrate how the risk neutral volatility, skewness, and kurtosis, can be obtained as functions of three payoffs, the quadratic, cubic, and quartic, corresponding to the second, third, and fourth powers of the returns on an asset. By Theorem 1 of Bakshi and Madan (2000), these payoffs can be expressed as linear combinations of prices of out of the money calls and puts on the asset. These formulae are "model free", in that they are independent of any specific option pricing model. They are reproduced here as equations (3), (4), (5), (6), (7), and (8) respectively. The BKM(2003) skewness and kurtosis formulae are[4]:

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

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

In equations (3) and (4);

$$V(t,\tau) = \int_{S_t}^{\infty} \frac{2\left(1 - \ln\left[\frac{K}{S_t}\right]\right)}{K^2} C(t,\tau;K)dK + \int_0^{S_t} \frac{2\left(1 + \ln\left[\frac{S_t}{K}\right]\right)}{K^2} P(t,\tau;K)dK \tag{5}$$

The BKM(2003) cubic and quartic formulae are[5]:

---

[4] These are taken from formulae (5) and (6) of Bakshi, Kapadia and Madan (2003).
[5] These are taken from formulae (8) and (9) of Bakshi, Kapadia and Madan (2003).

$$W(t,\tau) = \int_{S_t}^{\infty} \frac{6\ln\left[\frac{K}{S_t}\right] - 3\left(\ln\left[\frac{K}{S_t}\right]\right)^2}{K^2} C(t,\tau;K)dK$$

$$- \int_0^{S_t} \frac{6\ln\left[\frac{S_t}{K}\right] + 3\left(\ln\left[\frac{S_t}{K}\right]\right)^2}{K^2} P(t,\tau;K)dK \tag{6}$$

$$X(t,\tau) = \int_{S_t}^{\infty} \frac{12\left(\ln\left[\frac{K}{S_t}\right]\right)^2 - 4\left(\ln\left[\frac{K}{S_t}\right]\right)^3}{K^2} C(t,\tau;K)dK$$

$$+ \int_0^{S_t} \frac{12\left(\ln\left[\frac{S_t}{K}\right]\right)^2 + 4\left(\ln\left[\frac{S_t}{K}\right]\right)^3}{K^2} P(t,\tau;K)dK \tag{7}$$

In the above equations S_t and K, have their usual interpretations. C and P are call and put prices respectively. The options are assumed to be trading on date t and maturing at time (t+τ). The first moment $\mu(t,\tau)$ is given as;

$$\mu(t,\tau) = e^{r\tau} - 1 - \frac{e^{r\tau}}{2} V(t,\tau) - \frac{e^{r\tau}}{6} W(t,\tau) - \frac{e^{r\tau}}{24} X(t,\tau) \tag{8}$$

Rompolis and Tzavalis (2004) have generalised the Bakshi, Kapadia and Madan (2003) approach to any moment of the RND. In Corollary 1 of their paper they derive the Breeden and Litzenberger (1978) result from their Proposition 1, which itself is a corollary of Theorem 1 of Bakshi and Madan (2000). Proposition 2 of Rompolis and Tzavalis (2004) and its Corollaries 2 and 3, provide general analytical formulae for conditional asset return moments of order m, where $m = 1$, and $m \geq 2$. These are reproduced here as equations (9) and (10) below. The Rompolis and Tzavalis (2004) formula are[6]:

$$\mu_{\rho,1} = e^{r\tau}\left\{1 - \int_{S_t}^{\infty} \frac{1}{K^2} C_t(\tau,K)dK - \int_0^{S_t} \frac{1}{K^2} P_t(\tau,K)dK\right\} - 1 \tag{9}$$

---

[6] These are taken from formulae (8) and (9) of Rompolis and Tzavalis (2004)

$$\mu_{\rho,m} = me^{r\tau} \left\{ \begin{array}{l} \int_{S_t}^{\infty} \frac{1}{K^2} \left[\ln\left(\frac{K}{S_t}\right) \right]^{m-2} \left[m-1-\ln\left(\frac{K}{S_t}\right) \right] C_t(\tau, K) dK \\ +\int_0^{S_t} \frac{1}{K^2} \left[\ln\left(\frac{K}{S_t}\right) \right]^{m-2} \left[m-1-\ln\left(\frac{K}{S_t}\right) \right] P_t(\tau, K) dK \end{array} \right\} \qquad (10)$$

In equations (9) and (10) ρ is the conditional asset return, where the future log return is defined as $\rho_t(S_T) = \ln(S_T/S_t)$, m is the order of moment, and τ is the (T-t) period option maturity.

Rompolis and Tzavalis (2004) suggest an implementation of their formulae to observational data, by first using a cubic spline to interpolate discrete option prices, and then computing the risk neutral moments using equations (9) and (10) above. They test this approach against the usual procedure of extracting the full RND in a first step and then computing its moments. In a Monte Carlo simulation using data generated by the Heston (1993) stochastic volatility model they find their proposed procedure has a mean percentage error of under 1% for moments of all orders, and outperforms the method based on extracting the full RND in a first step.

Rompolis and Tzavalis (2004), demonstrate that the Britten-Jones and Neuberger (2000) 'model free' variance [equation (2) in this paper] is implied by their Proposition 2, and is a special case of their equation (10). Jiang and Tian (2007) demonstrate that the Britten-Jones and Neuberger (2000) implied variance formula and the model free implied variance formula developed by Demeterfi et al (1999)[7]

---

[7] This is used as the basis for calculating the CBOE's VIX volatility index.

[equation (1) in this paper] are equivalent. Thus, three different expressions for the implied variance which are essentially equivalent, have been separately developed.

2.1   Empirical Implementations of the Implied Moments

Option prices are only quoted for finite sets of discrete strike prices and maturities, thus empirical applications of model free implied moments involve approximating integrals, or replacing integration over continuous strike prices and maturities, with summation over their discrete counterparts. This results in the introduction of discretization and truncation errors. Several authors have performed tests to assess the resulting biases.

Dennis and Mayhew (2002) consider the empirical implementation of the Bakshi, Kapadia and Madan (2003) model free skewness formula. They generate option prices using the Black-Scholes model with constant values for maturity, volatility, risk free rate, and the stock price, since this will guarantee a skewness of zero. First, they assess the impact of discrete strike price intervals and find that when integrating over a range of strike prices from $30 to $70, a strike price interval of $5 produces a downward bias in skewness of approximately -0.07, and one of $2.50 produces a downward bias in skewness of approximately -0.05. They observe that since strike price intervals are mostly constant, at least the bias will be the same for all observations. Next, they consider the effects of an asymmetrical domain of observation, where the number of put options and call options are not equal. They find that the bias introduced by asymmetry can be significant, but can be avoided by using

the largest range of strikes consistent with a symmetrical domain of integration. Finally, they consider the truncation bias resulting from a finite set of option prices. In this case, their results indicate that at least 20 strikes are required to avoid a downward bias in the calculated skewness. If there are only ten strikes available the resulting bias in skewness will be approximately -0.1.

Jiang and Tian (2005) empirically tested the Britten-Jones and Neuberger (2000) model free implied volatility using S&P 500 index options data. They concluded that it is an efficient forecast of future volatility, and is also unbiased after a constant adjustment. They performed simulation tests on a discrete version of equation (2) to assess any biases, prior to applying it empirically. Using the structure (e.g. range in moneyness, number, and intervals) of strike prices for S&P 500 options offered on 23 September 1988 as a template, they generated cross sections of option prices from the Heston (1993) stochastic volatility model, parameterised identically to Bakshi, Cao, and Chen (1997). The risk free rate was assumed to be zero, and the unconditional mean of the volatility process was set to 0.2. The following procedure was then used: As proposed by Shimko (1993) they first converted the generated option prices to implied volatilities using the BS formula. Next, they interpolated between available strike prices by fitting a smooth function (cubic splines) to the implied volatilities. The implied volatility function beyond the range of available strikes was assumed to be flat. Finally, the BS formula was used again, to convert the smoothed implied volatilities back to call prices, and the model free implied volatility was extracted. The model free implied volatility was found to be in error by 0.0016, 0.0009, 0.0003, 0.0001, and 0.0001 over time intervals of 30, 45, 60, 75, and 90 days to maturity

respectively. Since the true volatility originally set for the test was 0.2, Jiang and Tian (2005) conclude that the errors are negligible in all cases. They consider that their tests verify the model free implied volatility and their implementation method.

Jiang and Tian (2007) tested the Chicago Board Options Exchange (CBOE)'s implementation of the Demeterfi et al (1999) model free implied variance (volatility) formulae, used to calculate the VIX volatility index. They concluded that the CBOE procedure for constructing the index is flawed, and produces economically significant errors. In a simulation using option prices generated from the BS formula they found that volatility was estimated with a 1.6% error, in the base case considered, which equates to $310 dollars for a single VIX futures contract. They found this error was almost wholly due to discretization, and that the strike price interval of 2.5% of asset price was too coarse for the VIX calculation. When the BS input parameters were varied over typical input ranges, it was found that the errors varied between +3.1% and –13.8% of the true volatility, and that truncation errors could be as large as -15% and discretization errors as large as 6.3% of the true volatility. The approximation errors also tended to increase with maturity. Jiang and Tian (2007) propose an improved procedure for calculating the VIX index based on smoothing as in Jiang and Tian (2005). The procedure in the former is improved by imposing a smooth pasting condition where the linear extrapolation and cubic spline interpolations meet, thereby ensuring a smooth function everywhere. In tests of their proposed procedure using option prices generated by the stochastic volatility with jumps (SVJ) model, Jiang and Tian (2007) find that the maximum error is 0.08% and most errors are within less than 0.05% of the true volatility. In comparison, the VIX procedure

produces errors of +0.79% to –1.98%[8]. For robustness, further tests were performed using prices generated by the BS formula, Heston (1993) model, and Duffie, Pan, and Singleton (2000) stochastic volatility with time varying jumps models. The results suggest the proposed procedure is consistently accurate across all specifications compared to the CBOE procedure, which performs poorly even under the simple Black-Scholes model.

Dennis and Mayhew (2009) presents some additional test results for the Bakshi, Kapadia, and Madan (2003) formulae for (model free) skewness and kurtosis not included in Dennis and Mayhew (2002). Using the same initial parameter values, excepting the risk free rate, which is now set at 5% rather than 7%, the strike price interval is gradually increased from $0.1 to $5 in increments of $0.1, over a strike price range of $30-$70. Skewness diverged from the correct value of zero, in an oscillatory fashion, as the strike price interval increased. This effect was not found by Dennis and Mayhew (2002), who suggest only a smooth downward bias occurs as the strike price interval is increased. Moreover, Dennis and Mayhew (2002) states that the error in skewness is –0.07 for a strike price interval of $5, whereas Dennis and Mayhew (2009) suggests the error in skewness is around ±0.4 for a strike price interval in the range of $5. A similar oscillatory divergence was found for kurtosis, which was 3.0 for strike price intervals close to zero, but ±2.0 for strike price intervals between $4.5 and $5.

---

[8] One VIX index point is equivalent to a 1% change in implied volatility.

In summary, the Dennis and Mayhew (2002, 2009) simulation test results suggest that economically significant errors are introduced by truncation and discretization effects when empirically implementing model free moments for the RND, even without considering the effects of noise and observation errors on option prices. By contrast, the simulation test results of Jiang and Tian (2005, 2007), suggest that these errors can be effectively corrected by employing data smoothing methods. Jiang and Tian (2007) also presents evidence that the implementation of model free variance (volatility) used by the CBOE to compute the widely followed VIX volatility index is seriously flawed.

The results discussed in this section relate to applications of "model free" moment formulae to discrete strike prices, and the biases caused by discretization and truncation errors in such cases. Our methodology, described in the next section, eliminates discretisation and truncation errors by solving the model free moment formulae directly in a continuous strike price framework.

3.   Methodology

We cannot directly observe the empirical pricing functions (or volatility of the underlying assets) at play in options markets. Thus, in order to test the "model free" formulae for the moments of the implied probability distribution, we investigated whether they can accurately recover the moments of the distribution implied by a well-understood closed-form pricing formula namely, the Black-Scholes formula. Implicit in this formula is the assumption that returns on the underlying asset are normally distributed, and that the volatility of these returns is constant. Moreover, the

formula is twice differentiable with respect to the exercise price. Unlike more complicated stochastic volatility models, the BS implied volatility is widely used by market practitioners as a benchmark reference. If the "model free" formulae are unable to recover the moments of the implied (normal) distribution in this simple case, they are unlikely to do so in the more complicated case of stochastic volatility models. Moreover, stochastic volatility models seek to explain the underlying volatility process generating realised (historical) volatilities which are time varying[9]. However, implied volatilities are a constant reflecting the expected volatility of the underlying asset at expiration of the option. Thus, we limited the scope of the enquiry to the BS formula.

As shown by Breeden and Litzenberger (1978) the implied probability distribution for the value of the underlying asset at expiration is simply the second differential (where it exists) of the pricing function with respect to the exercise price of the option, (in the case of discrete prices, the second difference of the option prices). It does not depend on any closed-form pricing model or its assumptions. This distribution and its moments, including the standard deviation (volatility) are expectations over the residual life of the option. Because in this case, we have tested the "model free" formulae for the implied distribution using the BS formula as our pricing function, our results assume the volatilities input to the BS model will be recovered (e.g. the BS implied volatilities).

---

[9] Ole E. Barndorff-Nielsen & Shephard, (2002) have analysed the relationship between realised volatility and stochastic volatility models.

The objective of this paper is to test the accuracy of the proposed model free formulae for the moments of the risk neutral density, as opposed to implementations of them on discrete data, as discussed in the previous section. We used the Rompolis and Tzavalis (2004) derivations for our tests, as these are the most general versions. As discussed above, the Demeterfi et al (1999), Britten-Jones and Neuberger (2000), and Bakshi Kapadia and Madan (2003) derivations are all implied by the Rompolis and Tzavalis (2004) derivations.

Equation (9) was used directly to obtain the "model free" implied mean. Equation (10) was simplified to give the following formulae for $\mu_{\rho,2}$ (variance), $\mu_{\rho,3}$ (skew), and $\mu_{\rho,4}$ (kurtosis) respectively;

$$\mu_{\rho,2} = 2e^{r(t)}\left\{\int_S^\infty \frac{1}{K^2}\left(1-\ln\left(\frac{K}{S_t}\right)\right)C_t(\tau,K)dK + \int_0^S \frac{1}{K^2}\left(1-\ln\left(\frac{K}{S_t}\right)\right)P_t(\tau,K)dK\right\} \quad (11)$$

$$\mu_{\rho,3} = 3e^{r(t)}\left\{\begin{array}{l}\int_{S_t}^\infty \frac{1}{K^2}\ln\left(\frac{K}{S_t}\right)\left(2-\ln\left(\frac{K}{S_t}\right)\right)C_t(\tau,K)dK \\ +\int_0^{S_t} \frac{1}{K^2}\ln\left(\frac{K}{S_t}\right)\left(2-\ln\left(\frac{K}{S_t}\right)\right)P_t(\tau,K)dK\end{array}\right\} \quad (12)$$

$$\mu_{\rho,4} = 4e^{r(t)}\left\{\begin{array}{l}\int_{S_t}^\infty \frac{1}{K^2}\left(\ln\left(\frac{K}{S_t}\right)\right)^2\left(3-\ln\left(\frac{K}{S_t}\right)\right)C_t(\tau,K)dK \\ +\int_0^{S_t} \frac{1}{K^2}\left(\ln\left(\frac{K}{S_t}\right)\right)^2\left(3-\ln\left(\frac{K}{S_t}\right)\right)P_t(\tau,K)dK\end{array}\right\} \quad (13)$$

The methodology used was straightforward. The Black-Scholes formula was used to represent the option pricing function. Equations (9), (11), (12), and (13), were solved as exactly as possible at multiple node points on a grid of representative input values.

16

The nodes were sufficiently densely spaced to allow plotting of error curves across the input domain. Sets of input values of maturity, volatility, and risk-free rate, spanning the ranges found in actual options markets were used. The strike price input domain was continuous between zero and infinity. Table 1 shows the known input values used. Table 2 shows the number of nodes in each grid.

Table 1. Domain of Input Values

	Min. Value	Max. Value	Interval	Number	Constant
Asset Price	50	5050	500	10	2500
Strike price	0	∞	0	∞	n/a
Maturity	1 month	24 months	1 month	24	12months
Risk-free Rate	0.01	0.16	0.02	16	0.05
Volatility	0.1	1	0.1	10	0.4
Skewness	0	0	0	1	0
Kurtosis	3	3	0	1	3

Table 1: Shows the known input values used. Volatility (the square root of variance) was used in the tests rather than variance itself. Skewness and kurtosis only have constant single values of 0 and 3 respectively. This is because the BS formula assumes a normal distribution of asset returns. Note that the mean is equal to the forward price of the underlying asset. The column Constant shows the values used for each input, when set as a constant. BS Call and Put pricing functions were integrated over the continuum of strike prices from 0 to infinity to compute the required moment at each node of the grid of input values.

Three plots of errors were produced for each moment considered: Volatility v. Risk-free rate; Volatility v. Time to expiration; Risk free rate v. time to expiration.

Table 2. Numbers of Node Points in Each Grid

	Mean	Volatility	Skewness	Kurtosis
Risk-free Rate	160	160	16	16
Maturity	240	240	24	24
Asset Price	100	100	10	10
Volatility	100	100	10	10

Table 2. Shows the number of node points in each grid of input variables used. The model-free moment formulae were solved exactly at each node of every grid. There are fewer node points in the grids for Skewness and Kurtosis, which always takes the same value of 0 and 3 respectively.

Asset prices ranging from 50 to 5050 at intervals of 500 were used. However, performance was found to be unaffected by asset price level. The recovered values of skewness and kurtosis, were easily compared with the underlying values of zero and

three, implicit in the BS model assumption of normality. The recovered volatility was compared with the volatility input to the BS model, which assumes constant volatility. The BS Call and Put pricing functions were integrated over the continuum of strike prices from zero to infinity, to recover computed values of the moments at each node point. This required the evaluation of more than 5,280 integrals in total. For each of the first four moments, the computed values from the "model-free" formulae, were subtracted from the known input values, at each node point. Since both of these should be identical, any resulting errors are expressed as deviations from zero. To visualize the error curves across the input domain 2D and 3D plots were then constructed. The plots, presented in the following section, show the percentage error curves for the recovered moments of the RND, e.g [(actual value – computed value)/actual value].

4. Error Curves: Risk Neutral Mean

If the (T-t) period return $\rho_t(S_T)$ is defined as the log price difference $Ln\left(\dfrac{S_T}{S_t}\right)$, then from the lognormal property of stock prices assumed in the Black-Scholes formula it follows that the distribution of returns is normal[10];

$$\rho_t\left(S_T\right) \square \ \phi\left(\left(\mu - \frac{\sigma^2}{2}\right)(T-t), \frac{\sigma}{\sqrt{(T-t)}}\right) \tag{14}$$

---

[10] For a discussion see e.g. Hull (2015).

In a risk neutral world the expected return μ is the risk free rate thus $\mu_{\rho,1}$ in equation

(9) should equal $\left(r - \dfrac{\sigma^2}{2} \right)(T - t)$ and we should expect to find that;

$$\left(r - \frac{\sigma^2}{2} \right)(T - t) - \mu_{\rho,1} = 0 \tag{15}$$

for all r, σ, and (T-t). To test if this is so, the percentage error in the expected mean recovered using equation (9) was computed as follows;

$$E_{r_{\rho,1}} = \frac{\left(r - \dfrac{\sigma^2}{2} \right)(T - t) - \mu_{\rho,1}}{\left(r - \dfrac{\sigma^2}{2} \right)(T - t)} \tag{16}$$

Fig. 1 is a 3D plot of the resulting errors across the Maturity / Volatility input space. In Fig. 1 the mean is affected by different combinations of both maturity and volatility.

Fig. 1 Model-free risk Neutral Mean: % error v. Maturity & Volatility

Error=2.78 (278%))

Error=-10.1 (-1010%)

$$E_{r_{\rho,1}} = \frac{\left(r - \dfrac{\sigma^2}{2}\right)(T-t) - \mu_{\rho,1}}{\left(r - \dfrac{\sigma^2}{2}\right)(T-t)}$$
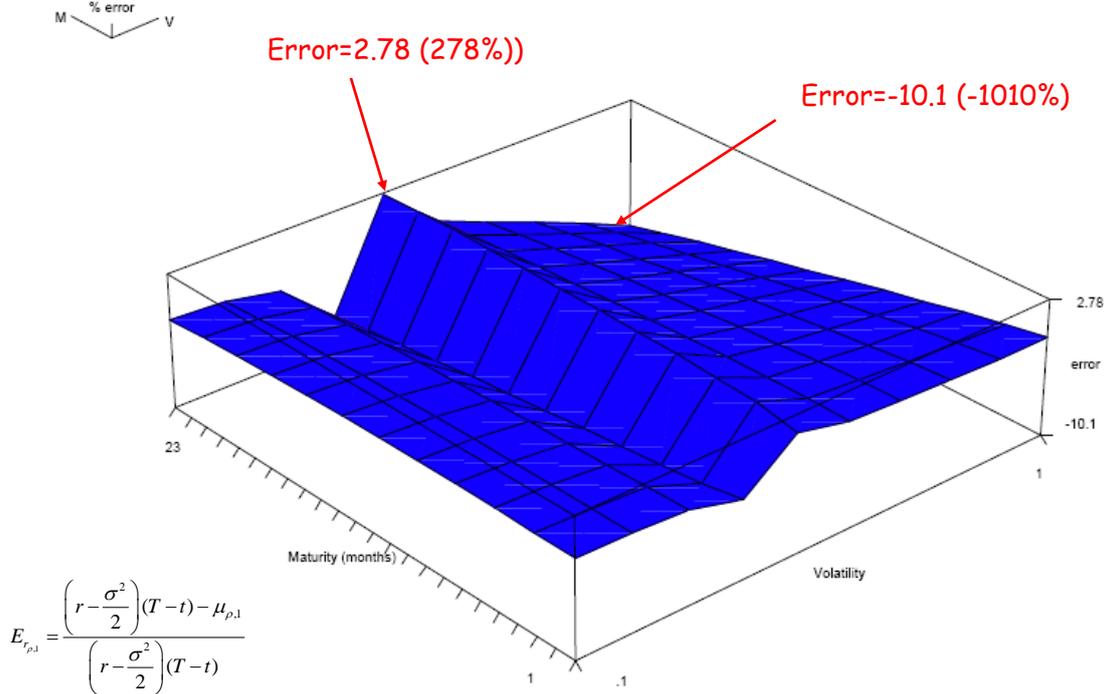
Fig. 2 shows the error curve for volatilities ranging from 0.1 to 1, for three months maturity options with underlying asset price of 2500, and risk-free rate ranging from 1% to 16% per annum. In Fig. 2 the mean is unaffected by the level of the risk-free rate but is sensitive to the level of volatility.

Fig. 2 Model-free risk Neutral Mean: % error v. risk-free rate & Volatility

$$E_{r_{\rho,1}} = \frac{\left(r - \dfrac{\sigma^2}{2}\right)(T-t) - \mu_{\rho,1}}{\left(r - \dfrac{\sigma^2}{2}\right)(T-t)}$$
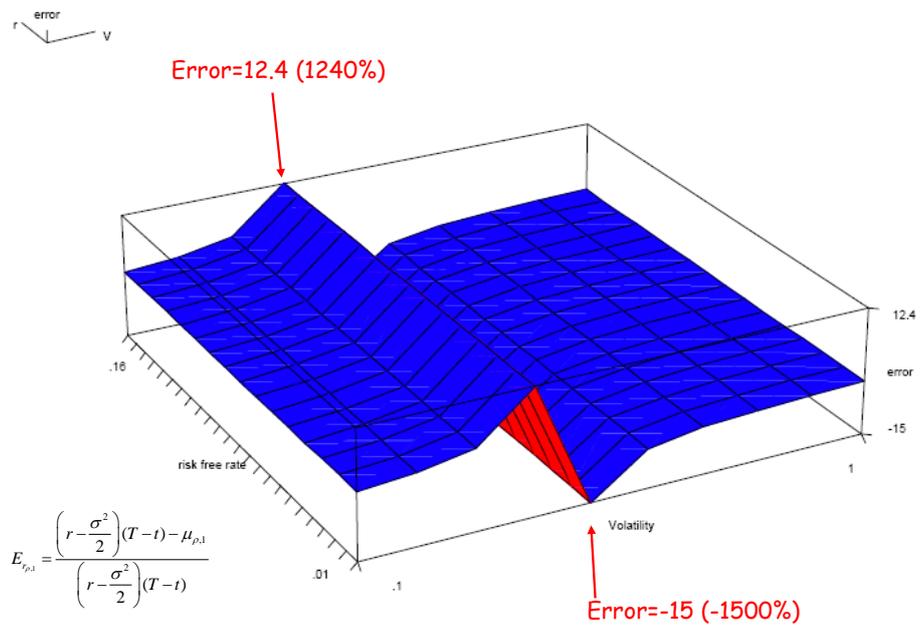
Fig. 3 plots the error curves for the risk free rate and maturity space. Here the mean is again unaffected by the level of the risk free rate but is sensitive to maturity.

Fig. 3 Model-free risk Neutral Mean: % error v. Maturity & Risk-free Rate

$$E_{r_{\rho,1}} = \frac{\left(r - \dfrac{\sigma^2}{2}\right)(T-t) - \mu_{\rho,1}}{\left(r - \dfrac{\sigma^2}{2}\right)(T-t)}$$
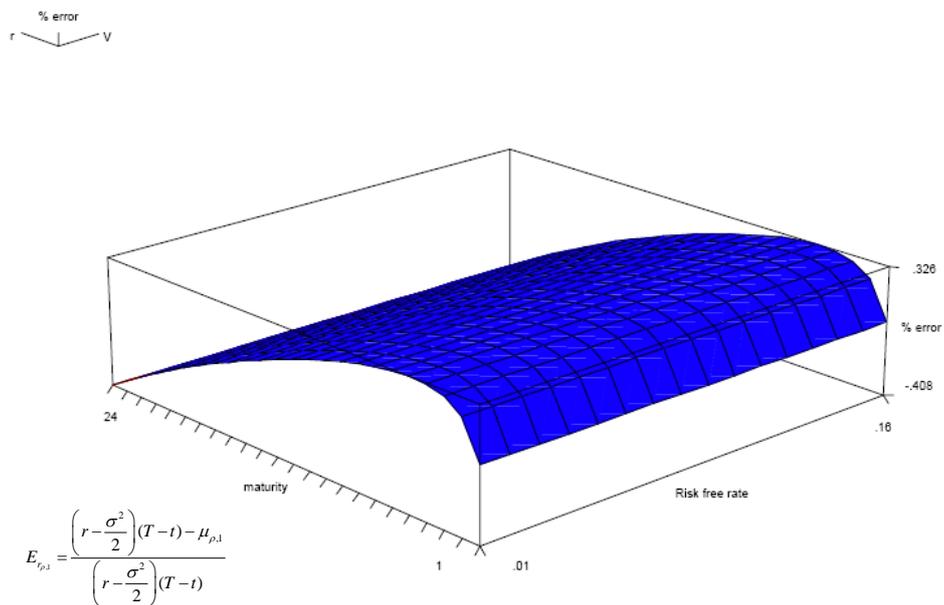
4.1 Error Curves: Risk Neutral Variance

The model-free second moment $\mu_{\rho,2}$ is given by equation (11). This is the variance of the risk neutral distribution of returns. However, the standard deviation, which is the square root of this quantity, is used instead. This is because the standard deviation or volatility of returns is an input to the BS formula, so a direct comparison of the computed values and the volatilities used is easily made. The error in volatility from the model-free formula was calculated as

$$E_{r_{\rho,2}} = \frac{\sigma - \sqrt{\mu_{\rho,2}}}{\sigma} \tag{17}$$

Figures 4, 5, and 6 plot the errors across the risk-free rate and maturity, risk-free rate and volatility, and maturity and volatility input spaces respectively.

Fig. 4 Model-free risk Neutral Variance: % error v. Risk-free Rate % Maturity
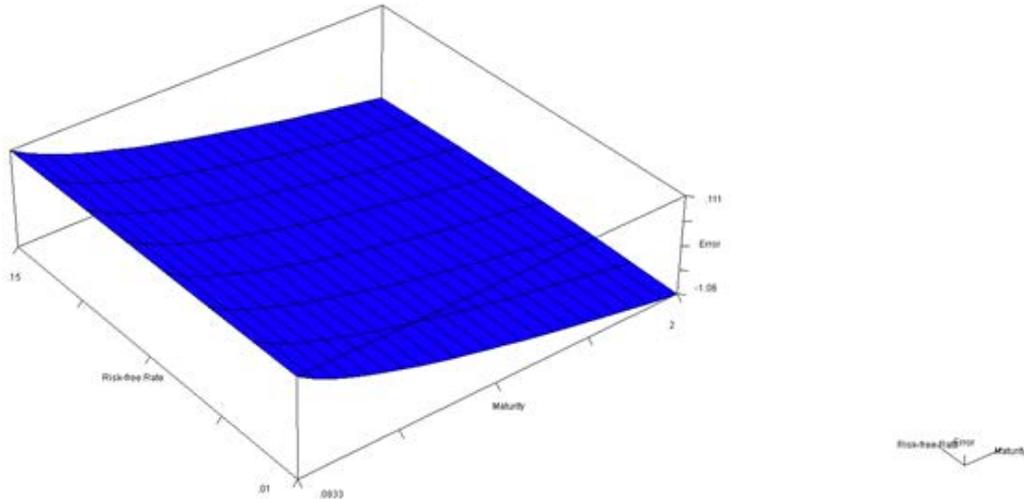
Fig. 4 suggests the error in the risk neutral variance is insensitive to the level of the risk-free rate but sensitive to time to maturity. Fig. 5 indicates some sensitivity to risk-free rate at low levels of volatility of the underlying asset. Fig. 6 shows sensitivity to maturity and volatility together.

22

Fig. 5 Model-free risk Neutral Variance: % error v. Risk-free Rate & Volatility
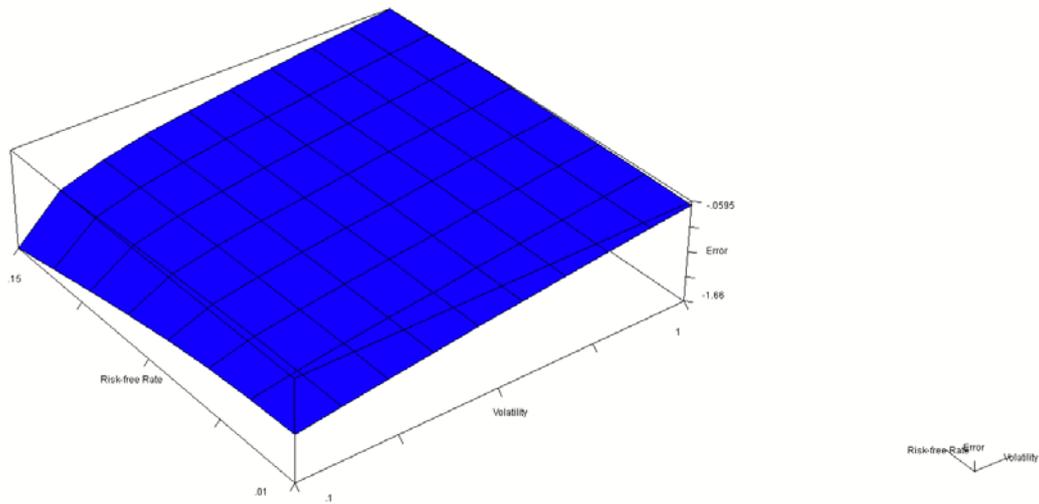

Fig. 6 Model-free risk Neutral Variance: % error v. Maturity & Volatility
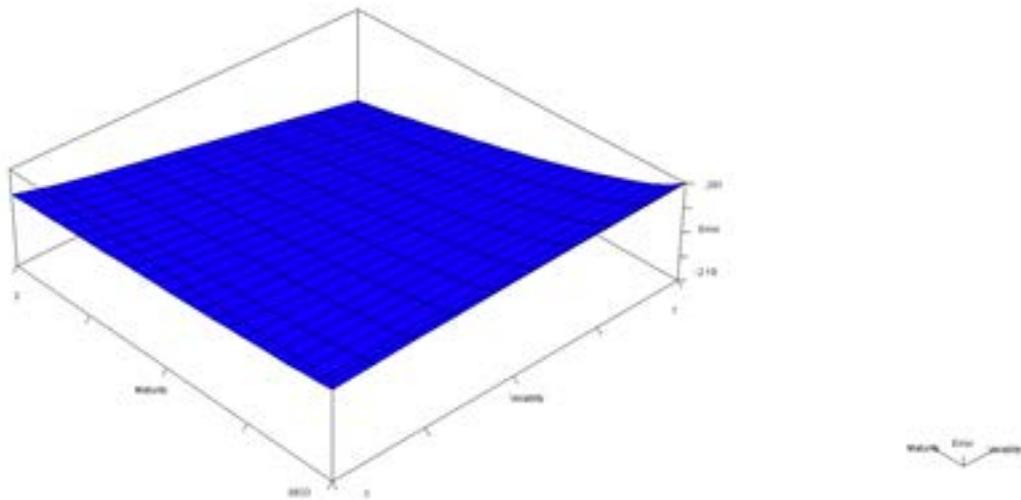
4.2 Error Curves: Risk Neutral Skewness

The model-free second and third moment formula, given by equations (11) and (12) were used to compute the skewness of the RND. Skewness, is the third moment about the mean, standardised by the cube of the standard deviation;

$$\gamma_1 = \left(\frac{\mu_{\rho,3}}{\left(\mu_{\rho,2} \right)^{3/2}} \right) \tag{18}$$

Since the BS formula assumes a normal distribution of returns, the skewness should equal zero, and any deviation from zero is a direct error measure for the formula. Figures 7 to 9 plot the error curves over the input spaces for the risk neutral skewness.

**Fig. 7 Model-free Risk Neutral Skewness: % error v. Risk-free Rate & Volatility**
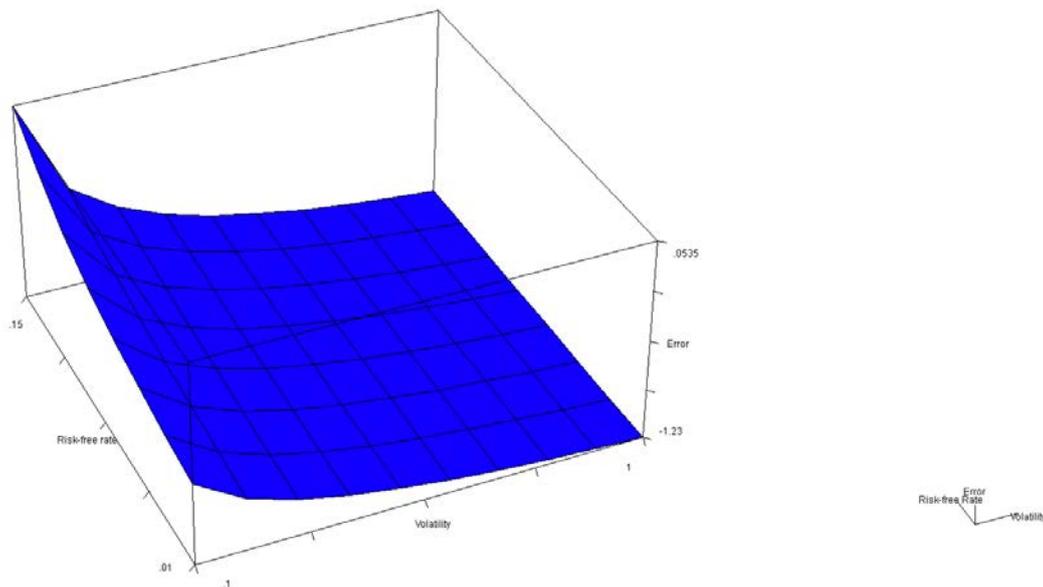
Fig 7 Suggests the risk neutral skewness is relatively insensitive to changes in the risk free rate, when combined with volatility as an input. However, in Fig. 8 the risk neutral skewness is observed to be sensitive to the risk free rate when combined with time to maturity. Fig. 9 reveals the sensitivity of skewness to the combination of Maturity and volatility.

**Fig. 8 Model-free Risk Neutral Skewness: % error v. Risk-free Rate & Maturity**
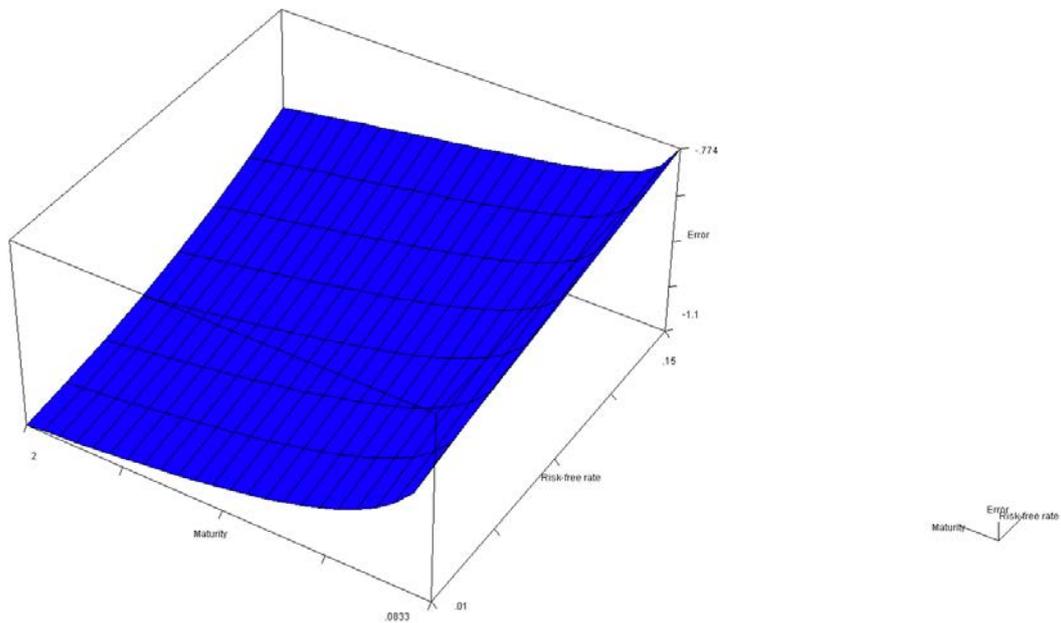
Fig. 9 Model-free Risk Neutral Skewness: % error v. Maturity & Volatility
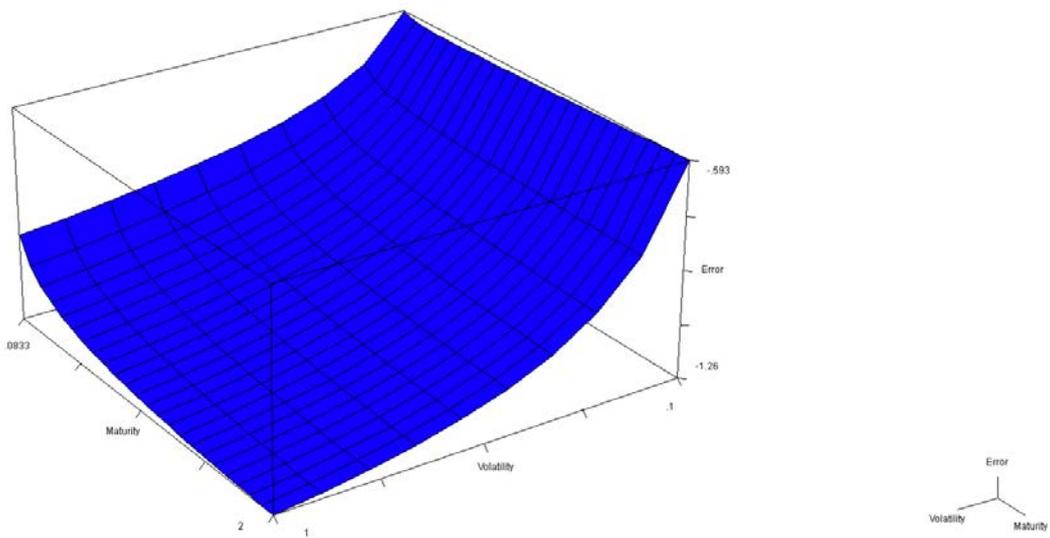
4.3 Error Curves: Risk Neutral Kurtosis

For a normal distribution Kurtosis is 3. Excess kurtosis is used here, since it is zero

for a normal distribution and errors produced by the model free formula are thus

easily observed as deviations from zero. The following formula is used to obtain the

excess kurtosis;

$$\gamma_2 = \left(\frac{\mu_{\rho,4}}{\left(\mu_{\rho,2} \right)^2} \right) - 3 \tag{19}$$

Figures 10, 11, and 12 plot the error curves for the excess kurtosis for each pair of inputs

Fig. 10 Model-free Risk Neutral (excess) Kurtosis: % error v. Risk-free Rate & Volatility
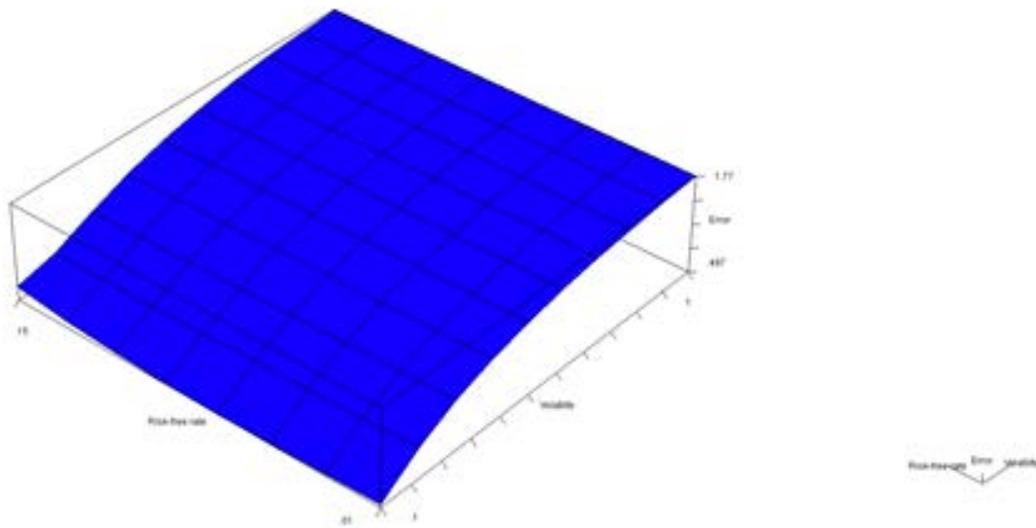

Fig. 11 Model-free Risk Neutral (excess) Kurtosis: % error v. Risk-free Rate & Maturity
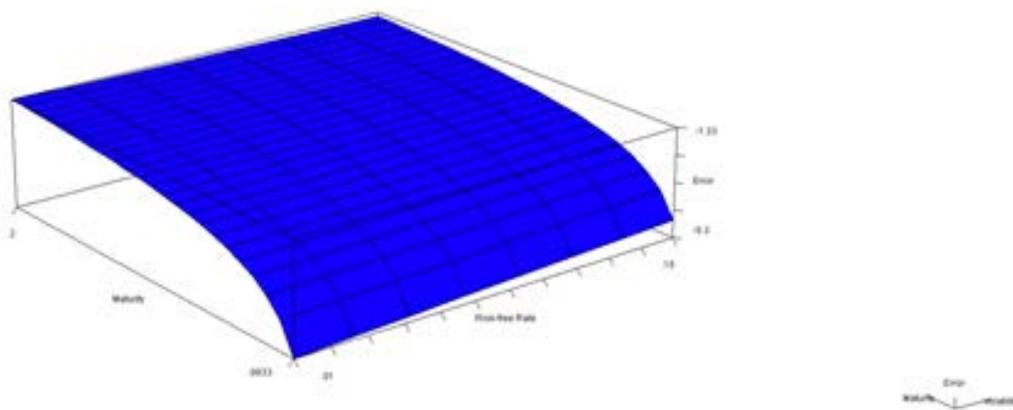
Fig. 12 Model-free Risk Neutral (excess) Kurtosis: % error v. Maturity & Volatility
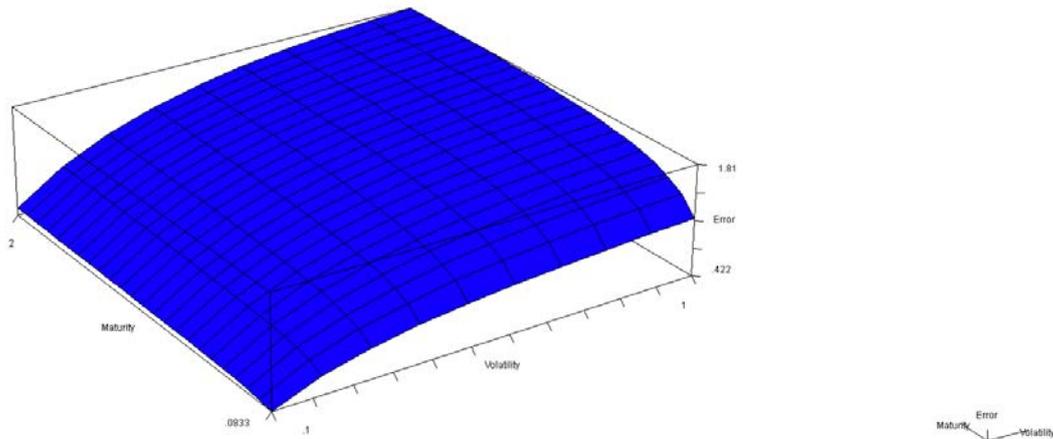

Fig 10 and 11 indicate that kurtosis is relatively insensitive to the risk free rate when combined with volatility and maturity respectively. Fig 12 indicates sensitivity to the combined inputs of maturity and volatility.

5.   Summary and Conclusions

Several authors have derived exact formulae for computing the moments of the RND directly from option prices without first estimating the full density, based on recent theoretical developments reviewed in this paper. These formulae have the advantage of being 'model free', in the sense of not being subject to the assumptions of any option pricing model, and offer the possibility of significantly simplifying option pricing and the extraction of market expectations from option market data. Dennis and Mayhew (2002, 2009) applied these formulae to discrete data and found that economically significant errors are introduced by truncation and discretization effects, even without considering the effects of noise and observation errors on option prices. Jiang and Tian (2005, 2007) however, suggest that these errors can be corrected by employing data smoothing methods. Jiang and Tian (2007) presents additional

evidence that the implementation of model free variance (volatility) used by the CBOE to compute the widely followed VIX volatility index is seriously flawed.

The above authors tested discrete data versions of the 'model free' formulae for the moments of the RND, where integrals are replaced by summation. In this paper by contrast, we tested the most general versions of the exact formulae; Namely the Rompolis and Tzavalis (2004) derivations. To perform our tests we used the Black Scholes formula to generate option prices, and solved the formulae as exactly as possible for realistic ranges of known inputs, assuming a continuum of values from zero to infinity for the strike price. We then plotted error curves for the results over the input domains.

Our results demonstrate that that the "exact" model-free formulae for the first four moments of the RND, like their discrete counterparts, produce major and economically significant errors. This is due to the presence of definite integrals in the formulae that can only be solved numerically. To realise the benefits of significant simplifications in option pricing, recover risk neutral distributions from option prices, and effectively operationalise the promised advances, we thus conclude that either mathematically equivalent expressions with analytical solutions are required or extensive work on the biases and errors introduced by the relevant numerical methods.

References

Ait-Sahalia, Y., and Lo, A.W. 2000. "*Nonparametric Risk Management and Implied Risk Aversion*", Journal of Econometrics 94, p 9-51.

Andersson, M., and Lomakka, M. 2003. "*Evaluating Implied RND's by Some New Confidence Interval Estimation Techniques*", Sveriges Riksbank Working Paper Series, No. 146.

Bakshi, G., Kapadia, N., and Madan, D. 2003. "*Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options*", Review of Financial Studies 16, 101-143.

Bakshi, G. and Madan, D. 2000. "*Spanning and derivative-security valuation*", Journal of Financial Economics 55, p205-238.

Bliss, R. and Panigirtzoglou, N. 2002. "*Testing the Stability of Implied Probability Density Functions*", Journal of Banking and Finance 26, p 381-422.

Bliss, R., and Panigirtzoglou, N. 2004. "*Option-Implied Risk Aversion Estimates*", The Journal of Finance, Vol, LIX, No.1 p407-446.

Breeden, D. T., and Litzenberger, R. H. 1978. *"Prices of State-contingent Claims Implicit in option Prices"*, Journal of Business 4, p 621-651.

Britten-Jones, M., and Neuberger, A. 2000. "*Option Prices, Implied Price Processes, and Stochastic Volatility*", The Journal of Finance, Vol. 55, No. 2, p839-866.

Carr, P. and Madan, D., 1998. Towards a theory of volatility trading. *Volatility: New estimation techniques for pricing derivatives*, 29, pp.417-427.

Cox, J.C., and Ross, S.A. 1976. "*The Valuation of Options for Alternative Stochastic Processes*", Journal of Financial Economics 3, p 145-166.

Dennis, P., and Mayhew, S. 2002. "*Risk-Neutral Skewness: Evidence From Stock Options*", Journal of Financial and Quantitative Analysis", 37(3), p471-494.

Dennis, P., and Mayhew, S. 2009. "*Microstructural Biases in Empirical Tests of Option Pricing Models*", Review of Derivatives Research, DOI 10.1007/s11147-009-9039-0.

Demeterfi, K., Derman, E., Kamal, M. and Zou, J. 1999. "*A guide to Volatility and Variance Swaps*", Journal of Derivatives, 6, p9–32.

Derman, E. and Kani, I., 1994. "Riding on a smile", Risk Magazine, 7, p32-39.

Dupire B., 1994, "Skewness and kurtosis in S&P500 index returns implied by option prices", The Journal of Financial Research, 19, p175-92.

Dupire, B., 1997. *Pricing and hedging with smiles* (pp. 103-112). Mathematics of derivative securities. Dempster and Pliska eds., Cambridge Uni. Press.

Figlewski, S. 2008. "*Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio*", in Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle (eds. Tim Bollerslev, Jeffrey R. Russell and Mark Watson). Oxford, UK: Oxford University Press.

Hull, J.C. 2015. Options, Futures and Other Derivatives. Pearson.

Jackwerth, J.C. 2004."*Option-Implied Risk-Neutral Distributions and Risk Aversion*". Charlotteville: Research Foundation of AIMR.

Jiang, G., and Tian, Y. 2005. "*Model-free Implied Volatility and Its Information Content*", Review of Financial Studies, 18, p1305–1342.

Jiang, G.J., and Tian, Y. S. 2007. "*Extracting Model-Free Volatility from Option Prices An Examination of the VIX Index*", Journal of Derivatives, Vol. 14, No. 3: p35-60.

Lynch, D., and Panigirtzoglou, N. 2008. "*Summary Statistics of Option implied Probability Density Functions and Their Properties*", Bank of England, Working Paper No. 345.

Ole E. Barndorff-Nielsen & Shephard, 2002. "*Econometric analysis of realized volatility and its use in estimating stochastic volatility models*" Journal of the Royal Statistical Society Series B, Royal Statistical Society, vol. 64(2), pages 253-280.

Ortiz-Gracia, L., and Oosterlee C. W. 2013 "Robust pricing of European options with wavelets and the characteristic function." *SIAM Journal on Scientific Computing* 35.5 B1055-B1084.

Ritchey, R. J. 1990. "*Call Option Valuation for Discrete Normal Mixtures*", Journal of Financial Research 13, p 285-296.

Rompolis, L. S. and Tzavalis, E. 2004. "*Retrieving risk neutral moments from option prices*", MIMEO, Department of Accounting and Finance, Athens University of Economics and Business.

Rubinstein, M., 1994. "Implied binomial trees", The Journal of Finance, *49*(3), pp.771-818.

Shimko, D. 1993. "*Bounds of Probability*", Risk, Vol.6, No.4, pp 33-37.

APPENDIX A

Cox and Ross (1976)[11] show that the prices of European exercise options can be expressed as the expected value of their payoffs, discounted at the risk-free interest rate.

$$C(X,t,T) = e^{-r(T-t)} \int_X^{\infty} P(S_T)(S_T - X)dS_T \tag{A1}$$

$$P(X,t,T) = e^{-r(T-t)} \int_{-\infty}^{X} P(S_T)(X - S_T)dS_T \tag{A2}$$

In equations (A1) and (A2) $C(X,t,T)$ and $P(X,t,T)$ are the prices of calls and puts trading at time t for expiration at some later time T. X is the exercise price, and r is the risk-free interest rate. $P(S_T)$ is the RND for the value of the underlying asset S at time t. Given a known functional form for $P(S_T)$, options can be priced for any value of exercise price X. Conversely, given a series of synchronous market prices observed at some time t, for options expiring at some later time T, this calculation can be inverted and an estimate of $P(S_T)$ extracted. Breeden and Litzenberger (1978)[12] showed that the cumulative distribution function (CDF) (negatively signed) for the value of the underlying asset S at time t is given by the first partial differential with respect to X of $f(X,t,T)$, where $f(X,t,T)$ represents the call (put) option pricing function, as shown in equation (A3),

$$\frac{\partial f(X,t,T)}{\partial X} = -e^{-r(T-t)} \int_X^{\infty} P(S_T)dS_T \tag{A3}$$

and the risk neutral distribution (RND) is obtained by differentiating $f(X,t,T)$ twice with respect to X as shown in equation (A4).

---

[11] Cox, J.C., and Ross, S.A. 1976. "*The Valuation of Options for Alternative Stochastic Processes*", Journal of Financial Economics 3, p 145-166

[12] Breeden, D. T., and Litzenberger, R. H. 1978*. "Prices of State-contingent Claims Implicit in option Prices",* Journal of Business 4, p 621-651.

$$\frac{\partial^2 f(X,t,T)}{\partial X^2} = e^{-r(T-t)} P(X). \tag{A4}$$

To understand why equation (A4) gives the RND consider the portfolio known as a 'butterfly spread'. This is given by

$$C(X+\varepsilon,t,T) - 2C(X,t,T) + C(X-\varepsilon,t,T) \tag{A5}$$

Where ε in (A5) is a small increment. The portfolio is created by selling two call options at exercise price X, and by purchasing a single call option at exercise price $(X + \varepsilon)$ and another at $(X - \varepsilon)$. The portfolio makes no payout except in the interval $[X - \varepsilon, X + \varepsilon]$. Consider $1/\varepsilon^2$ shares of this portfolio; in the limit as ε tends to zero, the payoff function tends to a Dirac delta function with mass at X, thus the portfolio will pay £1/\$1 if $S_T = X$ and nothing otherwise. The price of the portfolio (A5) must be

$$\frac{1}{\varepsilon^2}[C(X+\varepsilon,t,T) - 2C(X,t,T) + C(X-\varepsilon,t,T)] \tag{A6}$$

and

$$\underset{\varepsilon\to 0}{Lim}\left(\frac{1}{\varepsilon^2}[C(X+\varepsilon,t,T) - 2C(X,t,T) + C(X-\varepsilon,t,T)]\right) = \frac{\partial^2 C(X,t,T)}{\partial X^2} \tag{A7}$$

Thus the risk-neutral probability that $S_T = X$ is the price of a butterfly spread centered at X, in the limit as $\varepsilon\to 0$, and this is equal to $e^{-r(T-t)}p(X)$. In reality, X is not continuous and options are only available for a limited number of exercise prices at discrete intervals. However, Breeden and Litzenberger (1978) have shown that for discrete data, finite difference methods can be used to obtain a numerical solution to equation (A4. In addition, Neuhaus (1995)[13] has shown how the RND can be obtained via equation (A4) using finite differences.

---

[13] Neuhaus, H., 1995. "*The Information Content of Derivatives for Monetary Policy*", Discussion Paper 3/95, Economic Research Group of the Deutsche Bundesbank.