Rev Quant Finan Acc (2010) 35:245–269
DOI 10.1007/s11156-009-0153-8

ORIGINAL RESEARCH

A jump diffusion model for VIX volatility options and futures

Dimitris Psychoyios · George Dotsis · Raphael N. Markellos

Published online: 11 December 2009

Abstract Volatility indices are becoming increasingly popular as a measure of market uncertainty and as a new asset class for developing derivative instruments. Although jumps are widely considered as a salient feature of volatility, their implications for pricing volatility options and futures are not yet fully understood. This paper provides evidence indicating that the time series behaviour of the VIX index is well approximated by a mean reverting logarithmic diffusion with jumps. This process is capable of capturing stylized facts of VIX dynamics such as fast mean-reversion at higher levels, level effects of volatility and large upward movements during times of market stress. Based on the empirical results, we provide closed-form valuation models for European options written on the spot and forward VIX, respectively.

D. Psychoyios
Department of Industrial Management, University of Piraeus, Deligiorgi 107,
185 34 Piraeus, Greece
e-mail: dpsycho@unipi.gr

D. Psychoyios
CAIR, Manchester Business School, University of Manchester, Booth Street East,
Manchester M13 9PL, UK

G. Dotsis (✉)
Essex Business School and Essex Finance Centre,
University of Essex, Wivenhoe Park, Colchester C04 3SQ, UK
e-mail: gdotsis@essex.ac.uk

R. N. Markellos
Department of Management Science and Technology, Athens University of Economics
and Business, Office 915, 47A Evelpidon Str., 113 62 Athens, Greece
e-mail: rmarkel@aueb.gr

R. N. Markellos
Centre for Research in International Economics and Finance (CIFER),
Loughborough University, Loughborough, UK

Keywords Implied volatility · Jump diffusion · Option pricing · Volatility risk

JEL Classification G13 · C51 · C52

1 Introduction

Volatility is undoubtedly one of the most important variables in finance, appearing in a wide spectrum of theories and applications in asset pricing, portfolio theory, risk management, derivatives, corporate finance, investment evaluation and econometrics. A fascinating recent development has been the treatment of volatility as a distinct asset which can be packaged in an index and traded using futures and options (hereinafter collectively referred to as "volatility derivatives").[1] Volatility derivatives provide new ways to trade and hedge volatility risk and are considered by some to "*have the potential to be one of the most important new financial innovations*" (Grünbichler and Longstaff 1996).

The first volatility index, originally named VIX (currently termed VXO), was introduced in 1993 by the Chicago Board Options Exchange (CBOE). The VXO index is constructed according to the methodology proposed by Whaley (1993) and it represents the implied volatility of a synthetic at-the-money option on the S&P 100 with constant 30 calendar days to expiry. The CBOE adopted a new methodology in 2003 to calculate VIX in a model-free manner as a weighted sum of out-of-the-money S&P 500 call and put option prices at two nearby maturities across all available strikes. Carr and Wu (2006) showed that the new VIX squared approximates the 30-day conditional risk-neutral expectation of the return variance. Hence, VIX squared approximates the 30-day variance swap rate. Several other volatility indices similar to VIX have also been developed. These include the VXN and the VXD in the CBOE, which are the equivalent to VIX volatility indices of the NASDAQ and Dow Jones Industrial Average, respectively. The DAX-30 volatility index (VDAX-NEW) in Germany, the CAC-40 volatility index (VCAC) in France and the Dow Jones EURO STOXX 50 volatility index (VSTOXX) in the Eurex. In March 2004, the CBOE introduced volatility futures written on the VIX and while in February 2006 it also launched European volatility options on the same index. Futures on the VXD were introduced in April 2005 and European options are to follow soon. In September 2005, Eurex launched futures on the VDAX-NEW and VSTOXX indices, respectively.

Volatility derivatives were first suggested by Brenner and Galai (1989, 1993) as a response to the growing need for instruments to hedge volatility risk. It has been argued that volatility derivatives make markets more complete since they expand the realm of investment opportunities and allow direct hedging of volatility (vega) risk, without affecting the delta exposure to the underlying asset price risk. Volatility derivatives have a wide range of other important potential applications. Traditionally, volatility could be traded via at-the-money straddles, whose value increases with volatility; but straddles have the disadvantage of creating both market and volatility exposure. The market effect can be removed by rolling forward; however this is done at uncertain future market levels and trading costs. In contrast, volatility derivatives allow pure exposure to volatility changes. Certain classes of investors, such as convertible bond arbitrage funds and structured product issuers, can use these derivatives to insure against their structural exposure to volatility. Volatility

[1] Other types of derivatives used for trading/hedging volatility include variance and volatility swaps, which are traded over-the-counter (see Demeterfi et al. 1999; Chriss and Morokoff 1999, and Carr and Lee 2005, for details on the pricing and hedging aspects of variance/volatility swaps).

derivatives can be used to partially hedge against shifts in transaction costs and tracking error penalties, both of which increase during periods of high uncertainty. Investment managers that are long in equities are usually short in volatility because of the leverage effect. Asset correlations have been found to increase significantly during periods of market stress, making portfolio diversification very difficult. Given that volatility tends to increase when equity markets decline, a long position in a volatility derivative could be used as a hedging vehicle in a high correlation environment. An increase in volatility has a significant impact on the equity risk premium that shareholders require above the risk-free rate. Thus, firms could employ volatility derivatives to protect themselves from unexpected subsequent changes in the marginal cost of new issues. Although not yet available, bond and foreign exchange volatility indices and derivatives, could allow firms to hedge their volatility exposure in these markets. All these applications appear appealing; however, future research must establish empirically the effectiveness and advantages or disadvantages of volatility derivatives in comparison to conventional alternatives.

A number of studies have examined the forecasting power of volatility indices (see, among others, Fleming et al. 1995; Whaley 2000; Blair et al. 2001; Corrado and Miller 2005; Simon 2003; Giot 2005). There has also been a growing interest in modelling the time series dynamics of the autonomous implied volatility process. For example, Bakshi et al. (2006) estimated various diffusion processes with a non-linear drift and a diffusion component on the square of VXO. Wagner and Szimayer (2004) estimated a mean reverting jump diffusion process using VIX and VDAX. They found evidence of significant positive jumps in implied volatilities. However, they adopted the rather restrictive assumption that the jump size is constant. Dotsis et al. (2007) examined the ability of alternative popular continuous-time diffusion and jump diffusion processes to capture the dynamics of eight major European and US volatility indices. They found that the best fit to the data was offered by jump diffusion models with random upward and downward jumps. Finally, Sepp (2008) tried to model the VIX consistently with the dynamics of the variance of the S&P 500 and he found that jumps in variance are important.

Grünbichler and Longstaff (1996) developed the first model for the valuation of futures and European-style options written on volatility, respectively. The authors assumed that the underlying volatility followed a mean reverting square root process, similar to that used earlier by Heston (1993). Detemple and Osakwe (2000) provided analytical formulae for pricing both American and European-style volatility options assuming a mean-reverting process in the logarithms of volatility. On the basis of a discrete-time GARCH process, Heston and Nandi (2000a) derived analytical solutions for pricing European options written on variance. More recently, Daouk and Guo (2004), priced volatility options based on a Switching Regime Asymmetric GARCH process.

Motivated by the growing importance of volatility derivatives, this paper examines two main issues. First, it extends the literature on volatility indices by evaluating the performance of various diffusion and jump diffusion processes in approximating the empirical behaviour of the VIX. Specifically, we show that the best fit to the data is offered by the simple mean reverting logarithmic diffusion process of Detemple and Osakwe (2000) which is augmented by jumps. The diffusion part allows rapid increases followed by fast mean reversion, a salient feature of VIX, while the jump component accounts for the possibility of large upward movements during periods of market stress. Second, on the basis of the empirically favoured jump diffusion process we develop closed form expressions for pricing futures and European options written on VIX. Initially, we provide pricing formulae for futures and options written on spot VIX and then we extend the analysis to the case where the option is written on forward VIX. The option pricing model

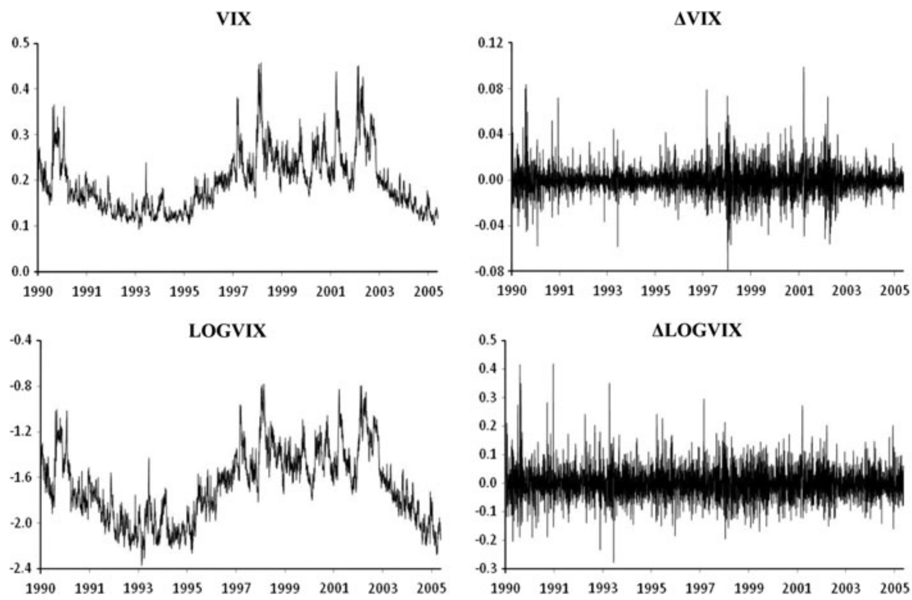

Fig. 1 The VIX in levels, first differences (ΔVIX), logarithmic levels (LOGVIX) and logarithmic differences (ΔLOGVIX), respectively

on spot VIX nests as a special case the model by Detemple and Osakwe (2000). We assess the potential implications of incorrectly omitting jumps from the diffusion process by showing that prices and hedge ratios may differ substantially. When the option is written on the forward VIX, as is the case with the options currently traded in the CBOE, it is shown that the pricing model is similar to that proposed by Kou (2002).

The remainder of the paper is structured as follows: The next section analyses the time series behaviour of the daily VIX over a period of 15 years. Section 3 presents the volatility processes considered in the empirical analysis. Section 4 discusses the empirical results with respect to the competing processes. Section 5, develops valuation formulae for VIX futures and options and discusses some of their key properties. The final section concludes the paper and offers some ideas for future research.

2 Empirical properties of the VIX

We use daily closing values of the VIX from 1/2/1990 to 9/13/2005, a total of 3,957 observations.[2] Figure 1 depicts the evolution of VIX in levels, first differences (ΔVIX), logarithmic levels (LOGVIX) and logarithmic differences (ΔLOGVIX), respectively. The plots suggest a volatile and mean-reverting behaviour for both VIX and LOGVIX with a number of jumps, while ΔVIX and ΔLOGVIX display violent swings.

The summary statistics of the series, shown in Table 1, largely confirm this behaviour. The VIX ranges from 9 to 45%, with an average of 19.6%. The higher moments suggest a leptokurtotic distribution with heavy tails and skeweness to the right for the VIX, ΔVIX

[2] Data were downloaded from the website of the CBOE. For details on the construction methodology of the VIX see Carr and Wu (2006).

Table 1 Descriptive statistics of daily VIX in levels, first differences (ΔVIX), logarithmic levels (LOGVIX) and logarithmic differences (ΔLOGVIX), respectively, from 1/2/1990 to 9/13/2005

	VIX	ΔVIX	LOGVIX	ΔLOGVIX
Mean	0.1957	−0.0000	−1.6808	0.0000
Median	0.1856	−0.0004	−1.6841	−0.0022
Maximum	0.4574	0.0992	−0.7822	0.4169
Minimum	0.0931	−0.0780	−2.3740	−0.2750
Std. Dev.	0.0639	0.0122	0.3123	0.0556
Skewness	0.9382	0.5647	0.2510	0.6006
Kurtosis	3.7411	9.1172	2.3677	6.6136
Jarque–Bera	671.14**	6,378.52**	107.46**	2,390.3**
$\rho(1)$	0.981**	−0.041**	0.983**	0.070**
$\rho(2)$	0.964**	−0.088**	0.969**	0.070**
$\rho(3)$	0.950**	−0.057**	0.957**	0.059**
$\rho^2(1)$	0.975**	0.201**	0.983**	0.125**
$\rho^2(2)$	0.950**	0.189**	0.969**	0.070**
$\rho^2(3)$	0.932**	0.204**	0.958**	0.064**

ρ(q) and ρ^2(q) are autocorrelation and autocorrelation of the squared series coefficients at lag q, respectively

** Statistical significance at the 1% (5%) level

and ΔLOGVIX. The Jarque–Bera test rejects the normality assumption at a high level of confidence for all series. Autocorrelations die out slowly in levels, something consistent with a highly persistent process. ΔVIX (ΔLOGVIX) appears weakly anti-persistent (persistent) with small negative (positive) short-term autocorrelations. The highly significant autocorrelations of the squared series suggest that heteroskedasticity is present in all series.

We proceed in examining the unconditional distribution of the data. As depicted in Fig. 2, the shape of the histogram for the VIX, and to a less extent of the LOGVIX, corresponds to a highly skewed distribution such as the chi-squared or the lognormal. Both differenced series appear to have leptokurtic distributions. A more detailed breakdown of the unconditional distributions is presented in Table 2. Given that the standard deviation of ΔVIX is around 0.0122 with a mean very close to zero, we can observe 22 distinct four-standard deviation events: 9 downward and 13 upward. Under a normal distribution, variations exceeding four-standard deviations should occur with a probability of under 0.005% or once in about every 80 years. In the case of VIX, the probability of these extreme variations occurring is over 0.5%, 100 times higher than that expected under the normal distribution, i.e., such an event occurs, on average, once every 178 days. These findings are expected given the fat-tails in the distribution of ΔVIX and the apparent jumps in the underlying process. It should be noted that almost all of the extreme negative changes in ΔVIX occur after large positive changes. This means that the jumps could be characterised also as positive spikes.[3] The ΔLOGVIX is somewhat closer to the normal distribution since only 14 (12 positive and 2 negative), 4-standard deviation events can be observed.

[3] Jumps are defined here as upward or downward discontinuous shifts in the underlying process which occur infrequently. As positive (negative) spikes we characterize upward (downward) discontinuous variations of the underlying which are immediately followed by a downward (upward) discontinuous reversal.

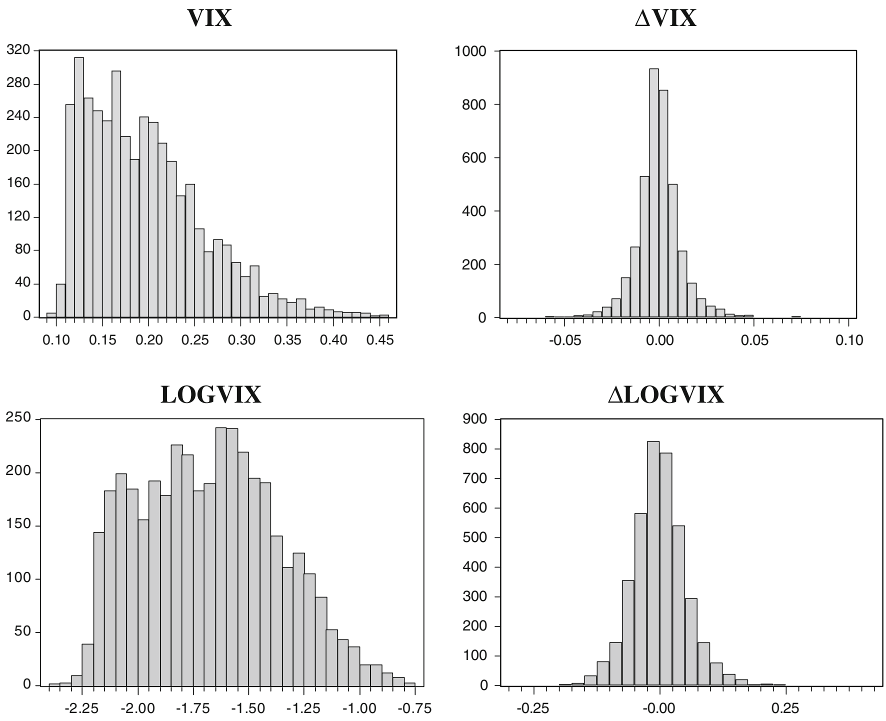

Fig. 2 Histograms of VIX in levels, first differences (ΔVIX), logarithmic levels (LOGVIX) and logarithmic differences (ΔLOGVIX), respectively

Table 2 Conditional tabulation of VIX versus ΔVIX and LOGVIX versus ΔLOGVIX

		VIX					Total
		[0, 0.1)	[0.1, 0.2)	[0.2, 0.3)	[0.3, 0.4)	[0.4, 0.5)	
ΔVIX	[−0.1, −0.05)	0	1	1	6	0	8
	[−0.05, 0)	2	1,244	661	114	8	2,029
	[0, 0.05)	3	1,052	703	130	19	1,908
	[0.05, 0.1)	0	0	4	6	2	12
	Total	5	2,298	1,369	256	29	3,956

		LOGVIX				Total
		[−2.5, −2)	[−2, −1.5)	[−1.5, −1)	[−1, −0.5)	
ΔLOGVIX	[−0.4, −0.2)	0	3	0	0	3
	[−0.2, 0)	411	1,084	516	23	2,034
	[0, 0.2)	354	949	563	37	1,903
	[0.2, 0.4)	0	7	4	3	14
	[0.4, 0.6)	0	1	1	0	2
	Total	765	2,044	1,084	63	3,956

3 Diffusion and jump diffusion processes for the VIX

3.1 Diffusion processes

One of the simplest continuous time processes of volatility is the Mean Reverting Gaussian Process (often called Ornstein–Uhlenbeck process):

$$dV_t = k(\theta - V_t)dt + \sigma dZ_t \tag{1}$$

It was initially proposed in order to capture the evident mean reversion in volatility (e.g., Stein and Stein 1991; Scott 1987; Brenner et al. 2006). Under this process, the volatility changes have a normal distribution, an assumption that is clearly rejected from our empirical analysis of the VIX. Moreover, this process has the significant disadvantage of allowing the possibility of negative volatility values. Two of the most popular alternatives are the Mean Reverting Square Root Process (SR) and the Mean Reverting Logarithmic Process (LR), given by Eqs. (2) and (3), respectively[4]:

$$dV_t = k(\theta - V_t)dt + \sigma\sqrt{V_t}\,dZ_t \tag{2}$$

$$d(\ln V_t) = k(\theta - (\ln V_t))dt + \sigma\, dZ_t \tag{3}$$

where V_t is the value of VIX at time t, Z_t is a standard Wiener process, k is the speed of mean reversion, θ is the long run mean, and σ is the diffusion coefficient. Both equations are defined under the actual probability measure P and can be obtained as a limit of particular ARCH-type processes. In particular, Heston and Nandi (2000b) have shown that SR can be obtained as a limit of a particular GARCH-type process, similar to the NGARCH and VGARCH models of Engle and Ng (1993). Detemple and Osakwe (2000) show that the EGARCH model of Nelson (1990) converges to a Gaussian process that is mean reverting in the log and thus matches the specification of the LR process. These processes should be able to capture two of the basic empirical characteristics of the VIX: mean reversion and heteroskedasticity. Furthermore, volatility follows a non-central Chi-squared distribution under the SR and a log-normal distribution under the LR, respectively (see Cox et al. 1985 and Detemple and Osakwe 2000), which is consistent with our analysis of the VIX unconditional distribution. We do not consider the Constant Elasticity of Variance (CEV) process (see Chan et al. 1992) since option pricing becomes infeasible due to the intractability of the characteristic function (see Duffie et al. 2000).

3.2 Jump-diffusion processes

Since the preliminary analysis in Sect. 2 suggests the possibility of positive spikes in the VIX, we also consider three basic types of mean reverting processes augmented with upward jumps:

Square-Root Process with Jumps (SRJ) $\quad dV_t = k(\theta - V_t)dt + \sigma\sqrt{V_t}\,dZ_t + y\,dq_t \qquad (4)$

[4] See, among others, Heston (1993), Grünbichler and Longstaff (1996), and Jones (2003) for the case of square root (SR) process, and Detemple and Osakwe (2000) for the case of mean reverting logarithmic (LR) process.

$$\text{Square-Root Process with proportional Jumps (SRPJ)} \quad dV_t = k(\theta - V_t)dt + \sigma\sqrt{V_t}dZ_t + y\,dq_t^* \tag{5}$$

$$\text{Logarithmic Process with Jumps (LRJ)} \quad d(\ln V_t) = k(\theta - (\ln V_t))dt + \sigma dZ_t + y\,dq_t \tag{6}$$

where Z_t is a standard Brownian motion, dq_t and dq_t^* are compound Poisson processes, and y is the jump size. In the SRJ and LRJ processes, dq_t has a constant arrival parameter λ, i.e. the probability of a jump is independent of the current level of V_t. In the SRPJ process the arrival parameter of dq_t^* is proportional to V_t, $\Pr\{dq_t{=}1\} = \lambda V_t dt$, i.e., the probability of a jump is proportional to the current level of V_t. dZ is assumed to be independent from both dq and dq_t^*. We further assume that the jump size is drawn from exponential distribution with density:

$$f(y) = \eta e^{-\eta y} 1_{\{y \geq 0\}} \tag{7}$$

where $1/\eta$ is the mean of the upward jump. The exponential distribution allows us to capture upward jumps in VIX and to derive the characteristic function in closed form (see the "Appendix 1" for the derivation of the characteristic functions). The one-sided exponential distribution adopted in (7) is a version of the double exponential distribution used by Kou (2002) in modelling the dynamics of stock and index prices. Equations (4), (5) and (6) are all defined under the actual probability measure P. After an application of Ito's lemma in Eq. (6), the process of V_t can be written as:

$$dV_t = V_t\left[k\left(\tilde{\theta} - \ln(V_t)\right)dt + \sigma\, dZ_t + (e^y - 1)dq\right] \tag{8}$$

where $\tilde{\theta} = (k\theta + 0.5\sigma^2)/k$. Inspection of Eq. (8) shows that LRJ, in contrast to the other two processes, has a proportional structure, i.e., the mean reversion, the diffusion coefficient and the jump size, all depend on the current level of V_t. The proportional structure of this model has three important implications. First, the model can account for a level effect of VIX, i.e., the condition whereby when VIX increases then its diffusion coefficient increases proportionally. Note that the diffusion part of Eq. (8) depends on V_t so that some of the large values of V_t can be captured by the diffusion part instead of the jump part. In order to examine the empirical relevance of the level effect, we depict in Fig. 3 |ΔVIX| against VIX (Panel A), and |ΔLOGVIX| against LOGVIX (Panel B), respectively. To further facilitate interpretation we fitted an OLS regression line to each pair of variables. The graph suggests that a level effect exists mainly in the first case, i.e., there is a positive relationship between the magnitude of the absolute VIX changes and the VIX levels. The graph indicates that in the second case, the relationship between the absolute logarithmic changes and levels is much weaker. The latter implies that the logarithmic transformation corrects some of the heteroskedasticity in the data, which can be expected since the LRJ process expressed in logarithms is homoskedastic. In contrast, as it can be easily shown by Ito's lemma, the square root type processes continue to imply heteroskedasticity under logarithmic transformations (see also Christoffersen et al. 2006). Second, since mean reversion depends on the level of V_t, i.e., the larger the V_t, the larger the effect of the mean reversion is. The LRJ is able to produce "spikes", rather than jumps, which is consistent with our preliminary descriptive analysis of the VIX in Sect. 2. Third, the LRJ process allows for the size of jumps to depend on the level of V_t and is thus capable of generating large upward movements, again a phenomenon that is consistent with the behaviour of the VIX during times of market stress. Finally, since we have only weak indications of abrupt downward movements, we do not include negative jumps. It must be noted that the log

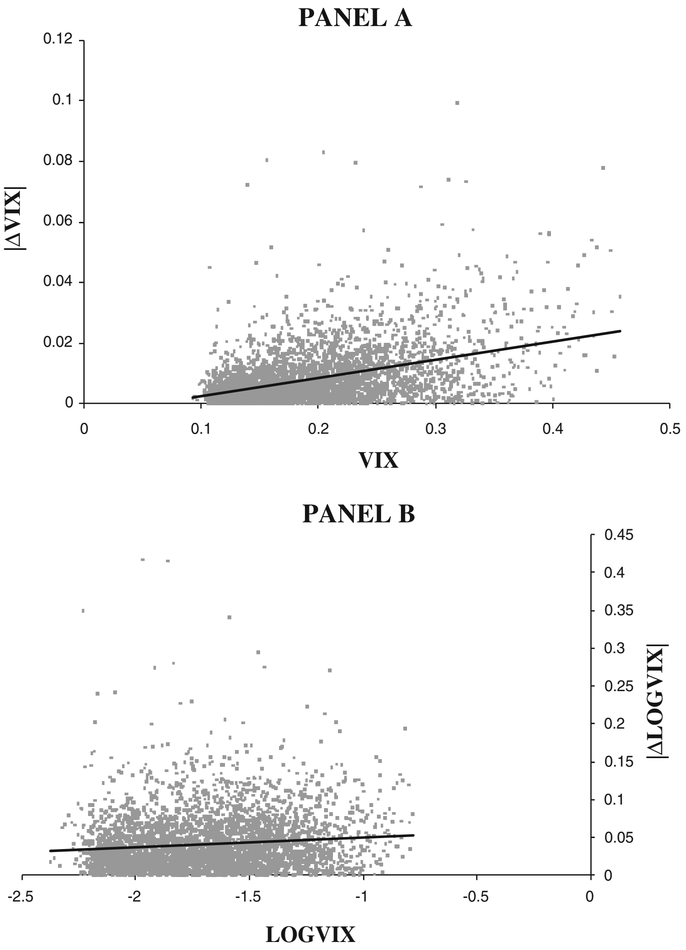

Fig. 3 PANEL A: Absolute daily VIX changes (|ΔVIX|) against the daily VIX level. PANEL B: Absolute daily logarithmic VIX changes (|ΔLOGVIX|) against the daily logarithmic VIX level (LOGVIX)

type processes are able to partially capture this behaviour through their fast mean reversion.

4 Estimation results

Table 3 shows the Maximum Likelihood (ML) estimation results using the VIX sample under study (see "Appendix 2" for details on the estimation methodology). For each process we report: the annualized parameter estimates, the asymptotic *t*-statistics in brackets, the log-likelihood (LL) values, the Akaike Information Criterion (AIC) and the Bayes Information Criterion (BIC). A further comparison of the models under consideration can be made using the Vuong (1989) closeness likelihood-ratio-based test, the results of which are contained in Table 4. It should be noted that the LL can be employed only for

Table 3 Parameter estimates of diffusion and jump diffusion processes over full sample (1/2/1990 to 9/13/2005)

Parameter	SR	SRJ	SRPJ	LR	LRJ
k	4.5496	7.3800	10.5004	3.9598	4.4887
	(5.97)	(9.51)	(11.13)	(5.48)	(6.60)
θ	0.1945	0.1505	0.1379	−1.6853	−2.1326
	(19.95)	(21.75)	(24.04)	(−29.84)	(−19.49)
σ	0.4048	0.3502	0.3294	0.9611	0.7504
	(88.07)	(61.32)	(51.33)	(79.67)	(50.31)
λ	–	19.4080	263.8877	–	41.9585
		(4.50)	(9.13)		(3.10)
$1/\eta$	–	0.0170	0.0125	–	0.068
		(8.22)	(4.56)		(6.74)
LL	12,263.12	12,422.37	12,459.24	12,485	12,627
AIC	−24,520	−24,835	−24,908	−24,964	−25,244
BIC	−24,501	−24,803	−24,877	−24,929	−25,229

Numbers in brackets denote t-statistics

The table also gives the Log-Likelihood value (LL), the Akaike Information Criterion (AIC) and the Bayes Information Criterion (BIC)

Table 4 Vuong likelihood-ratio-based test statistics for model selection

	LR	LRJ
SR	9.67	7.29
SRJ	2.94	7.08
SRPJ	0.90	6.56

The test can be used to evaluate the performance of a model in a given column (LR, LRJ) against that of a model in a given row (SR, SRJ, SRPJ). According to the null hypothesis, the two models under consideration are as close to the true model against the alternative that one model is closer. Under the null, the test statistic is distributed as a standard normal variate. If the test statistic exceeds the critical value at the chosen significance level, then the null hypothesis can be rejected in favour of the model in a given column. For a = 5% the critical value is 1.96

comparisons between nested models, i.e., between LR and LRJ, and, between SR, SRJ and SRPJ, respectively.

According to the two information criteria, the best fit is provided by the LRJ, followed by the LR, SRPJ, SRJ and SR. Moreover, the Vuong test suggests that the performance of the LRJ is significantly better than that of the competing non-nested models SRPJ, SRJ and SR, respectively. The null hypothesis of equal model performance is also rejected with high significance for the LR in all cases except when compared against the SRPJ. In line with the previous analysis, this ranking implies that the VIX is characterised by: (a) fast mean reversion at high levels, (b) level effects, i.e., as VIX increases its volatility increases proportionally, and, (c) jumps that are proportional to the level of VIX.

Amongst the square-root type processes, the SRPJ process displays the highest log-likelihood value. Since SRPJ, SRJ and SR are nested, the likelihood ratio test can be employed to compare the relative goodness-of-fit.[5] We find that the likelihood of the SRJ is significantly higher than that of the SR with the relevant likelihood ratio test statistic being 318.5 and the 5% critical value $\chi^2(df = 2) = 5.99$. Allowing the probability of jumps to be proportional to the level of VIX, produces a further statistically significant improvement (likelihood ratio = 73.74). The information criteria also suggest that the addition of jumps in proportion to the level of VIX improves fitting. Another point worth emphasizing is that the introduction of the jump component raises significantly the speed of the mean reversion parameter for both the SRJ and SRPJ. This is caused by the fact that jumps do not have a persistent effect and hence the speed of mean reversion increases "artificially" so as to pull back the process to its long run mean.

Now we turn our attention to the logarithmic processes. The estimation results show that the square-root type processes display lower log likelihood values relative to the LR process. The better fit of the LR process is also verified by almost all the information criteria. This result should not come as a surprise since, as mentioned previously, the LR process is capable of generating large increases in V_t at high levels, followed by rapid mean reversion. Essentially, changes that appear as jumps can also be generated by suitable diffusion components.[6]

The inclusion of jumps in the LR process enhances statistical goodness-of-fit even further. Once the drift and the diffusion components are correctly specified, the inclusion of a jump component helps to capture additional skewness. The σ drops from 0.96 to 0.75, which implies that jumps account for a substantial component of volatility, as one would expect intuitively. The estimate of the Poisson arrival rate implies 40 jumps per year with jump amplitude of approximately 7%. In contrast to the SRJ and SRPJ, the speed of mean reversion in the LRJ increases only slightly. This is an advantage, since the drift of the process is capable of generating rapid mean reversion, without inducing unrealistically high levels of k due to the presence of the jump component.

In order to check the stability of the parameters, we divided the sample into two equal parts and we re-estimated the processes. The results for the first and second subsamples are reported in Table 5.[7] For all processes, we can draw the following general conclusions. First, the ranking of the processes remains the same in both subsamples with LRJ continuing to dominate all other models. Second, the diffusion coefficient (σ) displays a stable behaviour in both subsamples when compared to the full sample. Third, the mean reversion parameter is higher in both subsamples. However, it is known in the literature that the mean reversion parameter is biased upwards in finite samples and accurate

[5] The likelihood ratio test statistic for comparing the nested models is $LR = -2 \times (LL_R - LL_U) \sim \chi^2(df)$, where df is the number of parameter restrictions and $LL_R,\ LL_U$ are the log-likelihoods of the restricted and unrestricted model, respectively. The 5% level critical values are: $\chi^2(df) = [3.84\,(df = 1), 5.99(df = 2), 7.82\,(df = 3)]$'. In order to be able to compare the directly the performance of the LRJ and LR processes with that of the SR, SRJ and SRPJ we apply the following change of variable: $LL_R = \sum_{t=1}^{T} \log(V_{t+\tau}) + \max_{\Theta} \sum_{t=1}^{T} \log[f(V_{t+\tau}|V_t, \Theta)]$ where $x_{t+\tau} \equiv \log\left(\frac{V_{t+\tau}}{V_t}\right)$ and $g(x_{t+\tau}|V_t, \Theta),\ f(V_{t+\tau}|V_t, \Theta)$ are the conditional probability density functions of the log-returns and levels of volatility, respectively, and $\Im_R = \max_{\Theta} \sum_{t=1}^{T} \log[g(x_{t+\tau}|V_t, \Theta)]$.

[6] This result implies that the key difference is whether the arithmetic Brownian motion or the Geometric Brownian motion is a better description of the volatility process. To this end, we have also estimated the Ornstein–Uhlenbeck process and it was found to be misspecified. These results are not reported in the paper but are available from the authors upon request. Dotsis et al. (2007) find a similar result. According to the authors, implied volatility follows a Geometric Brownian Process with jumps.

[7] Since the results remain the same, and due to space limitations, we have not include the tables with the Vuong statistic for each subsample. However the tables are available from the authors upon request.

Table 5 Parameter estimates of diffusion and jump diffusion processes over subsample A (1/2/1990 to 1/13/97) and subsample B (1/14/97 to 9/13/2005). Numbers in brackets denote *t*-statistics

Parameter	SR		SRJ		SRPJ		LR		LRJ	
	A	B	A	B	A	B	A	B	A	B
k	6.7809	4.7561	10.0246	7.6445	11.1900	12.6767	6.2675	4.0656	6.5474	5.2050
	(4.86)	(4.26)	(7.91)	(6.37)	(5.82)	(7.17)	(4.72)	(3.89)	(5.68)	(4.97)
θ	0.1682	0.2209	0.1367	0.1780	0.1353	0.1465	−1.8188	−1.5585	−2.0848	−2.2330
	(20.05)	(14.84)	(22.69)	(14.01)	(20.93)	(12.29)	(−34.88)	(−20.58)	(−28.17)	(−7.60)
σ	0.3843	0.4254	0.3167	0.3840	0.3122	0.3446	0.9972	0.8589	0.7402	0.7444
	(61.90)	(62.26)	(46.22)	(38.78)	(35.29)	(26.28)	(71.98)	(66.98)	(43.31)	(25.18)
λ	–	–	18.7032	21.8095	149.9526	478.5668			26.1822	132.5710
			(3.51)	(2.37)	(1.90)	(2.51)			(3.16)	(1.60)
$1/\eta$	–	–	0.0172	0.01569	0.0142	0.0090			0.072	0.027
			(5.76)	(5.09)	(3.52)	(5.74)			(5.71)	(4.25)
LL	6,388	5,888	6,538	5,920	6,547	5,939	6,567	6,010	6,612	6,031
AIC	−12,770	−11,770	−13,066	−11,868	−13,084	−11,868	−13,128	−12,014	−13,214	−12,052
BIC	−12,753	−11,753	−13,038	−11,840	−13,056	−11,840	−13,096	−11,982	−13,201	−12,039

The table also gives the Log-Likelihood value (LL), the Akaike Information Criterion (AIC) and the Bayes Information Criterion (BIC) estimation requires large data sets (e.g., Phillips and Yu 2005). Fourth, the long run mean estimate is lower in the first subsample when compared to the second subsample and the full sample, respectively. By visual inspection of the VIX time series, it appears that indeed the index is characterised by two different regimes: a low volatility regime until the mid 90s followed by a high volatility regime thereafter. Fifth, the estimation in the subsamples reveals some changes in the parameters of the Poisson arrival rates. Yet, this phenomenon may not be due to structural changes because the parameters governing the jump component are known to be rather "noisy" and large samples may be required for disentangling accurately the diffusion from jumps.

5 Pricing of volatility derivatives

Before we proceed to the derivation of the pricing formulae, we have to formally define the underlying asset of these contracts. Recall that VIX is the risk neutral expected volatility from t until the future date $T = t + 30$. However, the underlying asset of the volatility futures and options traded in CBOE is the forward VIX, which is defined as the risk neutral expected volatility between dates T and $T_1 = T + 30$. Forward VIX is similar to the concept of a forward interest rate and can be extracted from the term structure of implied volatilities.[8] In the remainder of this section we derive analytical formulae for pricing option and futures contracts on volatility indices when spot VIX follows the empirically favoured Mean Reverting Logarithmic Process with Jumps (LRJ). First, we derive pricing formulas for futures and options written on spot VIX. Then we extend the analysis to the case where the underlying of the option is forward VIX.

5.1 Futures on spot VIX

Denote F_t the price of a futures contract at time t with maturity T. The VIX futures is given by

$$F_t = E_t^Q(V_T) \tag{9}$$

where Q is the risk neutral probability measure and V_T is the forward VIX. Since VIX is not a tradable asset, the fair value of VIX futures cannot be derived by the cost-of-carry relationship.

We derive the fair value of VIX futures and options on the basis of general equilibrium. Cox et al. (1981) show that the futures price of any asset is the expectation, under the risk-neutral measure, of the asset's value on the expiration date. Hence, before proceeding to futures valuation, we must rewrite Eq. (6) under the risk neutral probability measure Q.

By analogy to Heston's (1993) volatility risk premium specification, we assume that the volatility risk premium is proportional to the logarithm of the current volatility level, i.e., $\zeta_t = \zeta \ln V_t$ (see also Christoffersen et al. 2006). As it is common in the literature (e.g., Pan 2002), we choose this volatility risk premium specification in order to preserve the affine structure of the process under the risk neutral probability measure. We also assume that "volatility of volatility" and "jump" risk, respectively, are not priced and that there is no model risk, i.e., the assumed dynamic is the true one. The volatility process under the risk neutral probability measure Q is then given by:

[8] The CBOE site provides details on the calculation of forward VIX.

$$d(\ln V_t) = (k+\zeta)\left(\frac{k\theta}{k+\zeta} - (\ln V_t)\right)dt + \sigma d\tilde{Z}_t + y\,dq_t \tag{10}$$

or, equivalently,

$$d(\ln V_t) = k^*(\theta^* - (\ln V_t)) + \sigma d\tilde{Z}_t + y\,dq \tag{11}$$

where $k^* = k+\zeta$, $\theta^* = \frac{k\theta}{k+\zeta}$ and $\tilde{Z}_t$ is a standard Brownian motion under the risk neutral probability measure Q.

As the conditional density function is not known in closed form, the characteristic function of V_T can be used to derive the expectation of $E_t^Q(V_T)$. This is done by evaluating the characteristic function at $s = -i$ (see "Appendix 1 for the derivation of the characteristic function of the LRJ.):

$$\begin{aligned} E_t^Q(V_T) = Exp\Bigg[& e^{-k^*(T-t)}\ln(V_t) + \theta^*\left(1 - e^{-k^*(T-t)}\right) + \frac{\left(1 - e^{-2k^*(T-t)}\right)}{4k^*}\sigma^2 \\ & + \frac{\lambda}{k^*}\ln\left(\frac{\eta - e^{-k^*(T-t)}}{\eta - 1}\right)\Bigg] \end{aligned} \tag{12}$$

Equation (12) consists of four terms: the first, the second, and the third term correspond to the diffusion part of the LRJ, respectively, while the fourth term corresponds to the jump part.

The futures pricing formula (12) has the following limiting properties:

$$\lim_{(T-t)\to 0} E_t^Q(V_T) = V_t, \tag{13}$$

$$\lim_{(T-t)\to +\infty} E_t^Q(V_T) = Exp\left(\theta^* + \frac{\sigma^2}{4k^*} + \frac{\lambda}{k^*}\ln\left(\frac{\eta}{\eta-1}\right)\right), \tag{14}$$

$$\lim_{V_t\to 0} E_t^Q(V_T) = 0. \tag{15}$$

Equation (13) shows the standard convergence property of the futures price to the spot price at maturity. Equation (14) shows that as the time-to-maturity increases, the futures price tends to the constant long-run volatility mean $Exp\left(\theta^* + \frac{\sigma^2}{4k^*} + \frac{\lambda}{k^*}\ln\left(\frac{\eta}{\eta-1}\right)\right)$. The latter means that as time-to-maturity increases, futures prices become less sensitive to current volatility changes and fail to capture the stochastic evolution of the VIX. Finally, Eq. (15) shows that as volatility tends to zero, futures prices also converge to zero.

5.2 Volatility options on spot VIX

In this section we develop a pricing formula when the volatility option is written on spot VIX. The discussion focuses on the additional impact that is due to the jump component since the properties of volatility options under diffusion processes are already well understood (see Grünbichler and Longstaff 1996; Detemple and Osakwe 2000). In order to obtain the valuation formula for a European volatility call, we follow the approach of Bakshi and Madan (2000) (page 219, Eqs. 22, 23 and 24). The price $C(V_t, T-t; K)$of the call option with strike price K and time to maturity $T-t$ is given by:

$$C(V_t, T-t; K) = e^{-r(T-t)}\left[V_t^{e^{-k^*(T-t)}} W(t, T-t)\Pi_1(t, T-t) - K\Pi_2(t, T-t)\right] \tag{16}$$

where r is the risk-free interest rate. The probabilities Π_1 and Π_2 are determined by

$$\Pi_j(t, T-t) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left[\frac{e^{-is(\ln K)} \times \psi_j(t, T-t; s)}{is}\right] ds, \quad j = 1, 2 \tag{17}$$

where $W(t, T-t) = Exp\left(\theta^*\left(1 - e^{-k^*(T-t)}\right) + \frac{\left(1-e^{-2k^*(T-t)}\right)}{4k^*} \times \sigma^2 + \frac{\lambda}{k^*} \times \ln\left(\frac{\eta - e^{-k^*(T-t)}}{\eta - 1}\right)\right)$, $\psi_2(t, T-t; s) = \frac{\psi(t,T-t;s)}{\psi(t,T-t;0)}$, $\psi_2(t, T-t; s) = \frac{\psi(t,T-t;s)}{\psi(t,T-t;0)}$,

$$\psi(t, T-t; s) = e^{-r(T-t)}\phi(\ln(V_t), T-t; s)$$

where $\phi(lnV_t, T-t; s)$ is the characteristic function of $ln(V_t)$ given in "Appendix 1". The call pricing formula has the following limiting properties:

$$\lim_{T-t \to 0} C(V_t, T-t; K) = \max(V_t - K, 0), \tag{18}$$

$$\lim_{T-t \to +\infty} C(V_t, T-t; K) = 0, \tag{19}$$

$$\lim_{V_t \to 0} C(V_t, T-t; K) = 0. \tag{20}$$

Equation (18) shows the standard convergence property of the option price to the option's payoff at maturity. Equation (19) implies that for very long maturities, the volatility call option is going to be worthless, as in the models of Grünbichler and Longstaff (1996) and Detemple and Osakwe (2000). Finally, Eq. (20) suggests (We see no economic reason to investigate the case $\lim_{V_t \to \infty} C(V_t, T-t; K)$. The assumption that volatility tends to infinity makes no economic sense, as it implies that volatility can drift to arbitrarily high levels in finite time. This is the same as assuming a priori that the stock market breaks down in some catastrophic fashion within a short time span.) that as V_t tends to zero, the volatility option price converges also to zero. Although under the model of Detemple and Osakwe (2000) a similar result is obtained, the model of Grünbichler and Longstaff (1996) predicts a non-zero value since the later assumes that V_t follows a SR. Similar to Detemple and Osakwe (2000), our model has an absorbing barrier at zero due to the multiplicative structure of the logarithmic process.

Using the estimated parameters from the previous section, Fig. 4 shows the value of a volatility call option as a function of V_t for three different levels of moneyness. This is a purely theoretical exercise since the parameters correspond to the real distribution and not to the risk-neutral one. We consider the diffusion model of Detemple and Osakwe (2000) along with our jump-diffusion specification. We can see that for short (long) maturities the logarithmic diffusion model underprices (overprices) the volatility call in comparison to the logarithmic jump-diffusion model. This is because the jump part affects mainly the value of short-term volatility calls, whilst the diffusion part affects mainly the value of long-term volatility calls.[10]

Figures 5 and 6 depict the *delta* (Δ) of the diffusion and jump-diffusion models as a function of maturity and V_t, respectively. The formula for the *delta* is the following[11]:

[10] Das and Sundaram (1999) and Pan (2002) provide similar results in the case of index options, where jumps improve the pricing mainly of short-term options. The pricing of intermediate and long maturity options is mainly improved by the assumption that the volatility of returns is stochastic.

[11] The derivation of Δ is straightforward and requires taking the partial derivative of $C(V_t, T-t, K)$, with respect to V_t (see also Proposition 2 in Detemple and Osakwe 2000). The diffusion model's Δ can be derived by setting $\lambda = 0$.

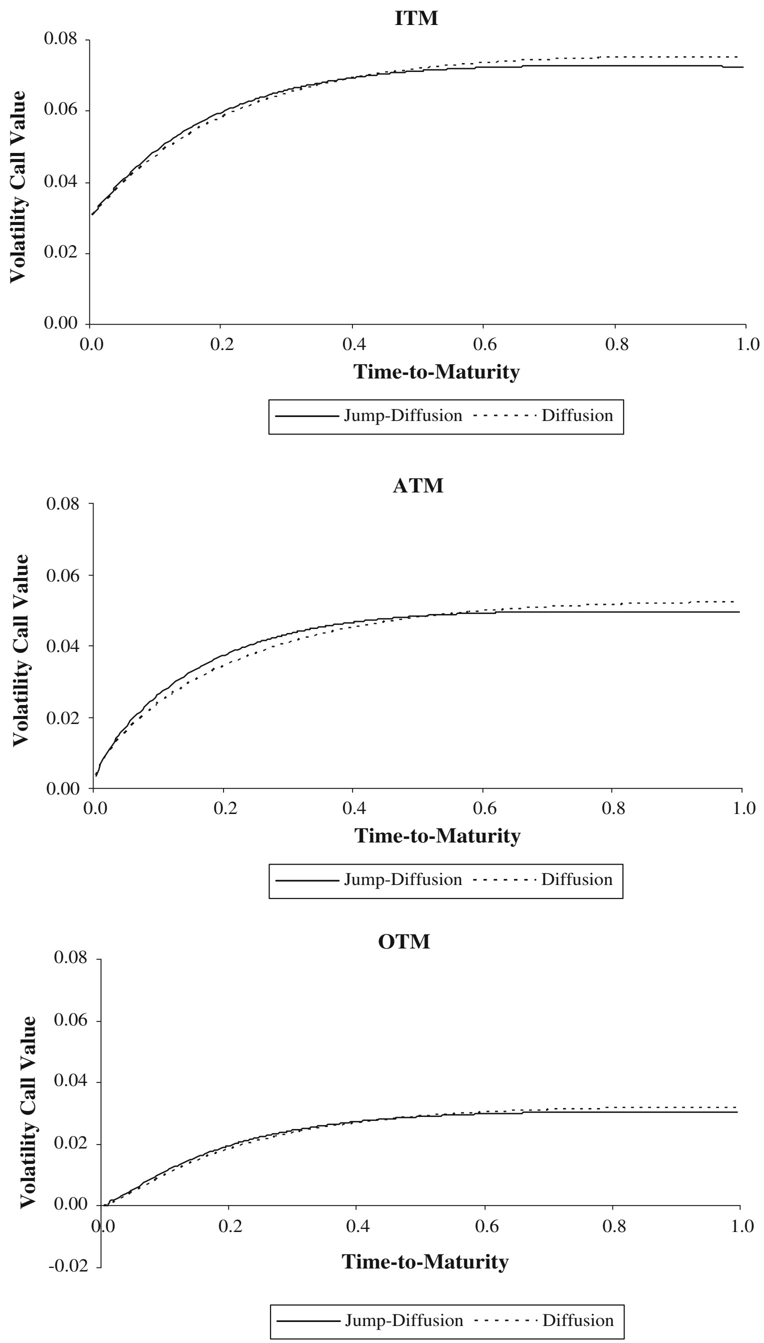

Fig. 4 Value of the volatility call option as a function of time-to-maturity estimated for three different moneyness levels: 20% in-the-money, at-the-money and 20% out-of-the-money. The dotted line corresponds to the case where there are no jumps in the volatility process (i.e., model of Detemple and Osakwe, 2000) using the estimated parameters k, θ, and σ from Table 3, fifth column. The solid line corresponds to the case where there are upwards jumps in the volatility process using the estimated parameters k, θ, σ, η, λ from Table 3, sixth column. We assume that r = 5% and V_t =15%

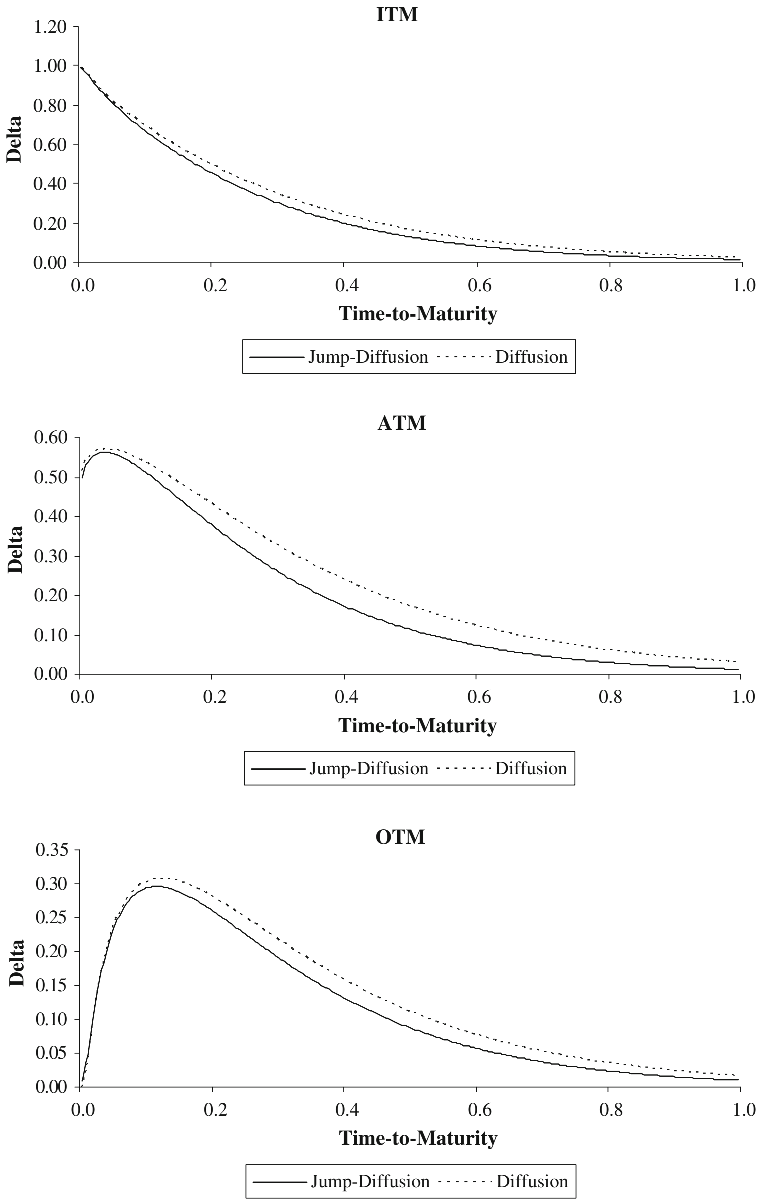

Fig. 5 Delta of the volatility call option as a function of time-to-maturity estimated for three different moneyness levels: 20% in-the-money, at-the-money and 20% out-of–the-money. The dotted line corresponds to the case where there are no jumps in the volatility process (i.e., model of Detemple and Osakwe, 2000) using the estimated parameters k, θ, and σ from Table 3, fifth column. The solid line corresponds to the case where there are upwards jumps in the volatility process using the estimated parameters k, θ, σ, η, λ from Table 3, sixth column. We assume that $r = 5\%$ and $V_t = 15\%$

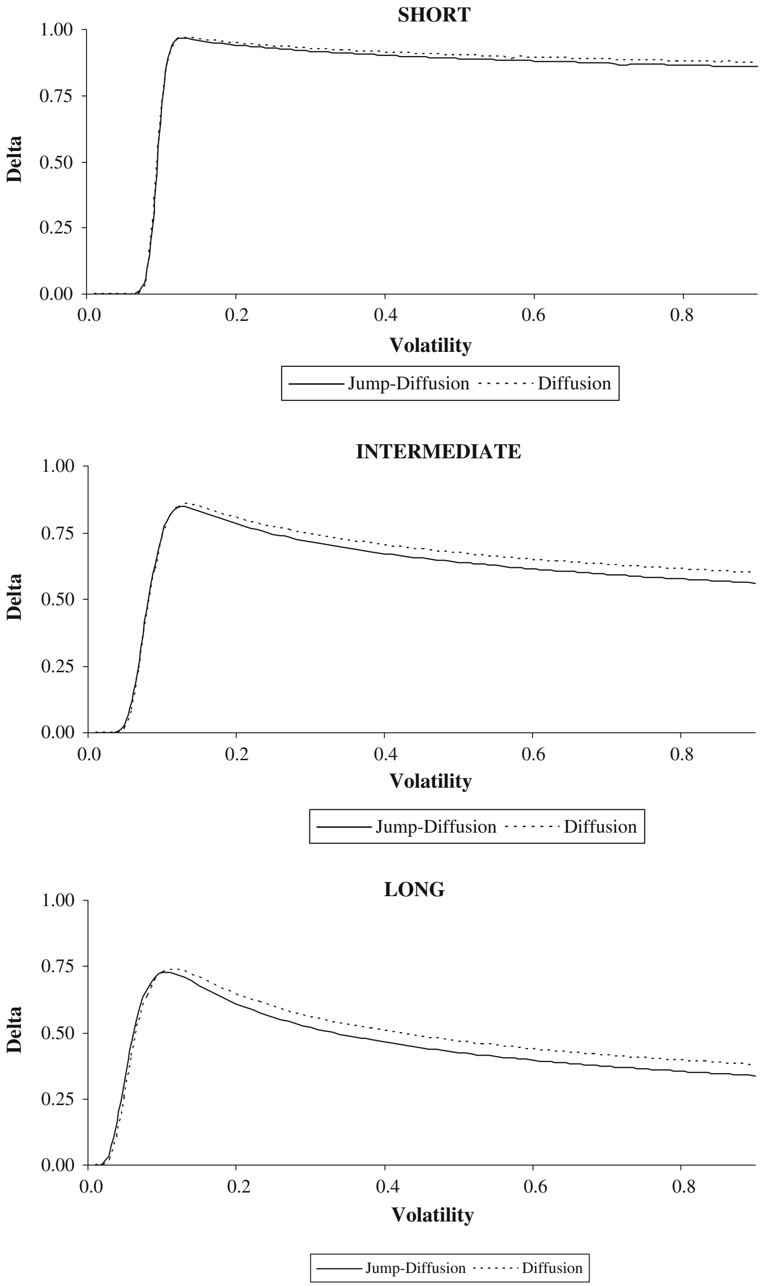

Fig. 6 Delta of the volatility call option as a function of volatility, estimated for three different maturities: short (5 days), intermediate (20 days) and long (40 days). The dotted line corresponds to the case where there are no jumps in the volatility process (i.e., model of Detemple and Osakwe, 2000) using the estimated parameters k, θ, and σ from Table 3, fifth column. The solid line corresponds to the case where there are upwards jumps in the volatility process using the estimated parameters k, θ, σ, η, λ from Table 3, sixth column. We assume that r = 5% and X =15%

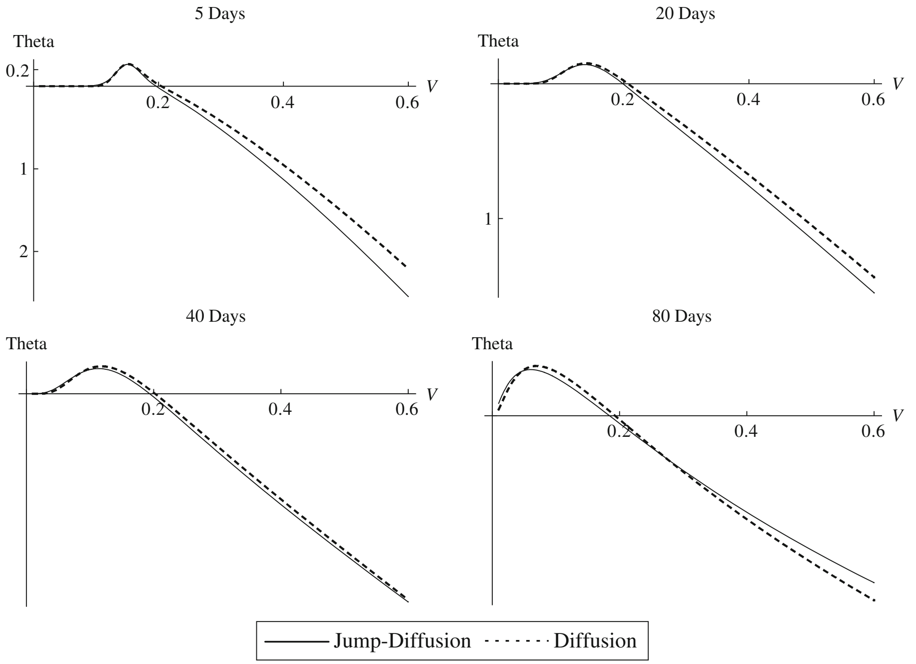

Fig. 7 Theta of the volatility call option as a function of volatility, estimated for four different maturities: 5, 20, 40 and 80 days. The dotted line corresponds to the case where there are no jumps in the volatility process (i.e., model of Detemple and Osakwe, 2000) using the estimated parameters k, θ, and σ from Table 3, fifth column. The solid line corresponds to the case where there are upwards jumps in the volatility process using the estimated parameters k, θ, σ, η, λ from Table 3, sixth column. We assume that r = 5% and X =15%

$$\Delta\left(V_t, T-t; K\right) = \frac{\partial C(V_t, T-t; K)}{\partial V_t} = e^{-(r+k^*)(T-t)} V_t^{e^{-k^*(T-t)}-1} W(t, T-t)\Pi_1(t, T-t) \quad (21)$$

We can easily observe that *delta* is always positive. Interestingly, the *delta* of the diffusion model is significantly higher in all cases, indicating that the diffusion model is more sensitive to VIX changes than the jump-diffusion model. This result may be attributed to the fact that the conditional variance of the diffusion model is larger than that of the jump-diffusion model.[12]

As Figs. 5 and 6 show, short maturity deep in-the-money calls have the highest *delta* (almost equal to 1), while long maturity deep out-of-the-money calls have the lowest *delta* (almost equal to 0). It can also be observed that as time-to-maturity increases, or, as calls move from in-the-money to out-of the-money, the value of the *delta* decreases and flattens out. This means that as time to maturity increases, or, as moneyness decreases, the volatility call option loses its hedging effectiveness. Another interesting implication of the mean-reverting nature of VIX is that the *delta* of a call option written on spot VIX, in contrast to ordinary options, is not a monotonic increasing function of the underlying. As

[12] This can be verified by replacing the estimated parameters to the conditional variance of the LRJ process, which is given by: $\frac{1-e^{-2k(T-t)}}{k}\left(\frac{\lambda}{\eta^2}+\frac{\sigma^2}{2}\right)$. However, we should note that this is not a general result and holds only for the estimated set of parameters.

the underlying increases, *delta* initially increases as well and then, for high levels of the underlying, it decreases. Decreasing *delta* means that the incremental potential of the call option becomes smaller and smaller as the underlying increases, or, in other words, for high levels of VIX, volatility options become more insensitive to the changes of VIX. Given the leverage effect, the hedging effectiveness of volatility options is actually diminishing at the time it is most needed.

Figure 7 shows the *theta* of the diffusion and jump-diffusion model as a function of V_t, for four different maturities. The sign of the *theta* of the volatility calls is inconclusive. In particular, for out-of-the-money options *theta* is close to zero, for at-the-money options *theta* is positive and as V_t grows larger *theta* becomes negative. These results are similar to the results of both Grunbichler and Longstaff (1996), and Detemple and Osakwe (2000). The sign of *theta* relies on the time-to-maturity, as well as on the moneyness level of the volatility call. As Grunbichler and Longstaff (1996) state "… the notion of moneyness is subtly different in the case of volatility options. The moneyness of the volatility call is not only the difference between V_t and K, but also of the difference between the long run mean of V_t and K…"

5.3 Volatility options on forward VIX

As mentioned, the volatility options traded in CBOE are written on forward VIX. Carr and Wu (2006) use no-arbitrage arguments to derive upper and lower bounds on VIX futures. They also showed that Eq. (9) can be written as $F_t = E_t^Q(V_T) = E_t^Q\sqrt{E_T^Q(RV_{T,T_1})}$, where $E_t^Q(RV_{T,T_1})$ is the risk neutral expectation of realized variance between T and T_1. From the concavity of the square root and Jensen's inequality, the upper and lower bounds of VIX futures are given by:

$$E_t^Q\sqrt{(RV_{T,T_1})} \le F_t \le \sqrt{E_t^Q(RV_{T,T_1})} \tag{22}$$

Hence, the lower bound is a forward volatility swap, which can be approximated by a forward starting option. These bounds are model-free and do not depend on any particular specification of spot or forward VIX.

Zhu and Zhang (2007) and Lin (2007) start from the specification of instantaneous variance and then derive pricing formulae for VIX futures written on forward VIX. However, in our case, because we start from a specification of spot VIX, it is not feasible to derive an exact pricing formula for VIX futures. In order to be able to use the information from spot VIX dynamics for option pricing purposes, we make the assumption that an unbiased expectation hypothesis holds, i.e. that the forward VIX is the expectation of future spot VIX, $E_t(V_T) = V_t^{forward} = F_t$. Though this assumption may be questionable or may not be perfectly accurate, it greatly facilitates the analysis because it allows us to derive simple closed form solutions for VIX options written on forward VIX. Specifically, if we apply Ito's lemma in Eq. (12), and set $\lambda = 0$, the stochastic process of forward VIX under the risk neutral measure is given by:

$$d(\ln F_t) = -\frac{1}{2}\sigma_F^2 dt + \sigma_F dZ_t, \tag{23}$$

where $\sigma_F = \sigma e^{-k^*(T-t)}$. Equation (23) shows that, in the absence of jumps, forward VIX is lognormally distributed. Note that the volatility risk premium still appears in the process. The volatility of the futures is scaled down by a factor $e^{-k^*(T-t)}$ relative to the volatility of spot VIX. Hence, forward VIX is less volatile compared to spot VIX and this may well be one of the reasons that CBOE chose to write the options on forward VIX. Since forward VIX is lognormally distributed, we can use Black's (1976) formula to price a volatility call option with maturity T and strike price K:

$$C_t = e^{-r(T-t)}[F_t N(d_1) - KN(d_2)] \tag{24}$$

where $d_1 = \frac{\ln(F_t)/K + 1/2\sigma_F^2(T-t)}{\sigma_F\sqrt{T-t}}$ and $d_2 = d_1 - \sigma_F(T-t)$. A similar formula has been obtained by Carr and Wu (2006). However, in their analysis they made the somewhat arbitrary assumption that forward VIX is log-normally distributed. In this paper, the log-normality of the VIX futures is derived from the dynamics of spot VIX. Note that under the square-root type process the derivation of the option pricing formula in a closed or semi-closed form is not feasible (see, Grünbichler and Longstaff 1996). When the spot VIX follows the LRJ, then by applying Ito's lemma for jump diffusions in Eq. (12) we can derive the dynamics of forward VIX:

$$d(\ln F_t) = \left(a_\lambda - \frac{1}{2}\sigma_F^2\right)dt + \sigma_F dZ_t + y_F y dq_t \tag{25}$$

where $a_\lambda = -\frac{\lambda}{\eta e^{k^*(T-t)} - 1}$, $\sigma_F = \sigma e^{-k^*(T-t)}$, $y_F = e^{-k^*(T-t)}y$. Note that now both volatility and the jump size are scaled down by $e^{-k^*(T-t)}$. The option can still be easily priced, using the semi-closed formulae of Kou (2002).[13]

5.4 Basis risk

As mentioned previously, futures and options written on the VIX were introduced as more effective volatility hedging instruments than stock option positions, e.g., straddles, butterfly spreads. However, strictly speaking, VIX derivatives can be used only for hedging the volatility of positions with respect to the underlying index, i.e., the S&P500. To the extent that the VIX is a good proxy for overall market risk, VIX derivatives can also be used to hedge against shifts in market volatility for positions on other broad based equity portfolios. In the case of volatility hedging for individual stocks, a basis risk problem could arise if the individual stock volatility does not move in tandem with the S&P 500. Apart from the basis risk arising from the cross-hedge, since we have assumed that VIX is best described by a logarithmic jump diffusion process, the hedge may also be exposed to basis risk from the jump component of the market if S&P 500 and the individual stock do not "jump" together.

Another cause for basis risks arises from the fact that the VIX is not a traded asset. Hence, in the absence of arbitrage, as a self-correcting mechanism, the futures price may not be tied to the movements of VIX, possibly resulting in a substantial basis risk for the hedger. The arbitrage bounds of Carr and Wu (2006) hold under the assumption that one can actively trade a basket of options on SPX and exotic OTC derivatives, such as forward-start at-the-money forward call options. In conclusion, a comprehensive treatment of the issues involved in hedging volatility risk of individual stocks is both interesting and important. However, this is beyond the scope of this paper and is left for future research.

[13] The actual importance of jumps in VIX option pricing can only be verified by studying the distributional shape of forward VIX implied by the VIX options. This is a good strand for future research, when enough VIX options data will be available.

6 Conclusions

Motivated by the growing literature on volatility derivatives and their recent introduction in major exchanges, this paper examines the empirical relevance and potential impact of volatility jumps in volatility derivative pricing and hedging.

We estimate, via maximum likelihood, various continuous time processes using daily closing prices on the VIX volatility index over a period of 15 years. The results suggest that a logarithmic mean reverting diffusion process provides the best fit compared to square root diffusion and the jump diffusion processes, respectively. Moreover, performance is further enhanced when a jump component is added to the logarithmic mean reverting process. On the basis of the estimation results, we develop closed form models for pricing futures and options on spot VIX. The proposed volatility derivative pricing models nest, as a special case, those proposed by Detemple and Osakwe (2000) and appear to have comparable properties. The pricing model without jumps in volatility undervalues (overvalues) short (long) maturity options, on average, by 10% (6%). Moreover, it is more sensitive to changes in the underlying. Indicatively, the delta hedging parameter for an at-the-money volatility call with intermediate time-to-maturity (15–25 days), which is most likely to be used as hedging instrument, is 8% larger. Finally, we show that when the option is written on the forward VIX, the pricing model with jumps is similar to that proposed by Kou (2002).

The findings in this paper do not necessarily support criticism against the specific structural form assumed by existing volatility future and option pricing models. Rather, they attempt to demonstrate that pricing derivatives on a volatility index should carefully account for salient features of the data since the results obtained are particularly sensitive to the specification used to approximate the underlying dynamics. Testing against actual market prices will provide more definitive evidence on the merit of alternative pricing models. In the case of futures, this is possible since some data do exist for futures on volatility indices (e.g., see Dotsis et al. 2007). However, since market traded volatility options prices are not in abundance, we cannot fully test the empirical relevance of alternative option pricing models.

We believe that much more research is needed on the practical usefulness of volatility derivatives, especially for corporate finance. Although some ideas have been proposed in the literature and discussed in this paper, it is not yet clear how financial managers can use these instruments and what the actual benefits they may expect are. This is not a trivial problem, since the implications of volatility for a firm are widespread and complex. For example, a short futures position on the VIX index buys insurance against changes in the volatility of the US equity market. A US firm assuming this position, would be affected directly and indirectly in a number of ways with respect to factors including: firm value, cost of equity, cost of debt, optimal finance mix, employee stock option value, value and effectiveness of existing hedges, value of investments, and investment hurdle rates.

Appendix 1: Derivation of the characteristic functions for the jump-diffusion processes

Duffie et al. (2000) prove that, under technical regularity conditions, the characteristic function for affine diffusion/jump diffusion processes, such as the SRJ and SRPJ, has the following exponential affine form[14]:

[14] See Duffie et al. page 1,351, Eqs. 2.4, 2.5, 2.6.

$$\phi(V_t, T-t; s) = \exp(A(T-t;s) + B(T-t;s)V_t) \tag{26}$$

Thus, for the case of the SRJ, $A(T-t;s)$ and $B(T-t;s)$ are given by[15]:

$$A(T-t;s) = a(T-t;s) + z(T-t;s) \tag{27}$$

$$a(T-t;s) = -\frac{2k\theta}{\sigma^2} \times ln\left(\left(k - \frac{1}{2}i\sigma^2 s\left(1 - e^{-k(T-t)}\right)\right)\Big/ k\right) \tag{28}$$

$$z(T-t;s) = \frac{2\lambda\rho}{2k - \eta\sigma^2} \times ln\left(\left(k - \frac{1}{2}i\sigma^2 s + is\left(\frac{\sigma^2}{2} - \frac{k}{\eta}\right)e^{-k(T-t)}\right)\Big/\left(k - \frac{isk}{\eta}\right)\right) \tag{29}$$

and,

$$B(T-t;s) = \frac{ksie^{-k(T-t)}}{k - \frac{1}{2}i\sigma^2 s(1 - e^{-k(T-t)})} \tag{30}$$

The characteristic function of the LRJ expressed in logarithms is given by:

$$\phi(lnV_t, T-t; s) = \exp(A(T-t;s) + B(T-t;s)(lnV_t)) \tag{31}$$

where

$$A(T-t;s) = is\theta(1 - e^{-k(T-t)}) - s^2\sigma^2\left(\frac{1 - e^{-2k(T-t)}}{4\kappa}\right) + \frac{\lambda}{k} \times ln\left(\frac{\eta - ise^{-k(T-t)}}{\eta - is}\right) \tag{32}$$

$$B(T-t;s) = ise^{-k(T-t)} \tag{33}$$

Finally, in the case of the SRPJ the coefficients $A(T-t;s)$ and $B(T-t;s)$ cannot be solved in closed form and are estimated numerically. So, the conditional characteristic function $\phi(V_t, T-t; s) = E(e^{isV_T}|V_t)$ of the SRPJ must satisfy the following Kolmogorov backward differential equation:

$$\frac{\partial\phi}{\partial V_t} + k(\theta - V_t) + \frac{1\partial^2\phi}{2\partial V_t^2}V_t\sigma^2 - \frac{\partial\phi}{\partial\tau} + \lambda V_t \mathrm{E}[F(V_t + y) - F(V_t)] = 0 \tag{34}$$

subject to the boundary condition

$$F(V_t, T-t=0; s) = e^{isV_t} \tag{35}$$

where $i = \sqrt{-1}$. Differentiating the characteristic function given by Eq. (26) yields:

$$\begin{aligned} \phi_V &= BF \\ \phi_{VV} &= B^2F \\ \phi_{T-t} &= F(A_{T-t} + VB_{T-t}) \end{aligned} \tag{36}$$

where the subscripts denote the corresponding partial derivatives. Substituting Eq. (36) into Eq. (34) and rearranging yields:

$$V_t\left(-kB - B_{T-t} + \frac{1}{2}\sigma^2B^2 + \lambda\mathrm{E}\left[e^{yB} - 1\right]\right) + (k\theta B - A_{T-t}) = 0 \tag{37}$$

Also, $\mathrm{E}[e^{yB} - 1] = \int_0^{+\infty} \eta e^{-\eta y} e^{yB} dy - 1 \ = \frac{\eta}{\eta - \mathrm{B}} - 1$, and since $V_t \neq 0$, the expressions in

[15] This characteristic function has also been used for estimating purposes by Bakshi and Cao (2006).

the parentheses in Eq. (37) must equal zero. Therefore, we obtain the following ordinary differential equations (ODEs)

$$-kB - B_{T-t} + \frac{1}{2}\sigma^2 B^2 + \lambda\left(\frac{\eta}{\eta - \mathrm{B}} - 1\right) = 0 \tag{38}$$

$$k\theta B - A_{T-t} = 0 \tag{39}$$

Although the ODEs cannot be solved in a closed form, numerical solutions are possible subject to the boundary conditions $A(T-t=0;s)=0$, and $B(T-t=0;s)=is$.

Appendix 2: Maximum-likelihood estimation

Maximum Likelihood estimation requires the conditional (transition) density function $f[V(t+\tau)|V(t),\Theta]$ $(\tau > 0)$ of the process V_t, where τ denotes the sampling frequency of observations and Θ is the set of parameters to be estimated. For a sample $\{V(t)\}_{t=1}^{T}$, the log-likelihood function that is maximized is given by: $LL = \max\limits_{\Theta} \sum_{t=1}^{T-\tau} \log(f(V(t+\tau))|V(t),\Theta)$.

In the case of the SR and LR processes, the conditional density is known in closed form (see Dotsis et al. 2007 and Detemple and Osakwe 2000, respectively). The conditional density of the jump diffusion processes is derived from the characteristic function as described below (see also Singleton 2001). Assume that we stand at time t, and τ denotes the sampling frequency of observations. Then, the Fourier inversion of the characteristic function $\phi(V(t),T-t;s)$ provides the required conditional density function $f[V(T)|V(t)]$:

$$f[V(T)|V(t)] = \frac{1}{\pi}\int_0^\infty \mathrm{Re}[e^{-isV(T)}\phi(V(t),T-t;s)]ds \tag{40}$$

where Re denotes the real part of complex numbers. For a sample $\{V(t)\}_{t=1}^{T}$, the conditional log-likelihood function to be maximized is given by:

$$LL = \max_{\{\Theta\}} \sum_{t=1}^{T} \log\left(\frac{1}{\pi}\int_0^\infty \mathrm{Re}[e^{-isV(t+\tau)}\phi(V(t),T-t;s)]ds\right) \tag{41}$$

where $\Theta = \{\kappa, \theta, \sigma, \lambda, \eta\}$ is the set of parameters to be estimated.[16] The standard errors of the ML estimators are retrieved from the inverse Hessian evaluated at the obtained estimates.

References

Bakshi G, Cao C (2006) Risk-neutral kurtosis, jumps, and option pricing: evidence from most actively traded firms on the CBOE. Working Paper, University of Maryland

Bakshi G, Madan D (2000) Spanning and derivative-security valuation. J Financ Econ 55:205–238

Bakshi G, Ju N, Ou-Yang H (2006) Estimation of continuous-time models with an application to equity volatility dynamics. J Financ Econ 82:227–249

Black F (1976) The pricing of commodity contracts. J Financ Econ 3:167–179

Blair BJ, Poon S, Taylor SJ (2001) Forecasting S&P 100 volatility: the incremental information content of implied volatilities and high-frequency index returns. J Econom 105:5–26

[16] In the case of the LRJ process the conditional density is given by: $f[\ln V(T)|\ln V(t)] = \frac{1}{\pi}\int_0^\infty \mathrm{Re}[e^{-is\ln V(T)}\phi(\ln V(t),T-t;s)]ds$.

Brenner M, Galai D (1989) New financial instruments for hedging changes in volatility. Financ Anal J 4:61–65
Brenner M, Galai D (1993) Hedging volatility in foreign currencies. J Deriv 1:53–59
Brenner M, Ou E, Zhang J (2006) Hedging volatility risk. J Bank Financ 30:811–821
Carr P, Lee R (2005) Robust replication of volatility derivatives working paper, Courant Institute, NYU and University of Chicago
Carr P, Wu L (2006) A tale of two indices. J Deriv 13:13–29
Chan KC, Karolyi GA, Longstaff FA, Sanders AB (1992) An empirical comparison of alternative models of the short-term interest rate. J Financ 47:1209–1227
Chriss N, Morokoff W (1999) Market risk of variance swaps. RISK 12:55–59
Christoffersen P, Jacobs K, Mimouni K (2006) Models for S&P500 dynamics: evidence from realized volatility, daily returns, and option prices. Working Paper, McGill University
Corrado CJ, Miller T (2005) The forecast quality of CBOE implied volatility indexes. J Futur Mark 25:339–373
Cox JC, Ingersoll JE, Ross SA (1981) The relationship between forward prices and futures prices. J Financ Econ 9:321–346
Cox JC, Ingersoll JE, Ross SA (1985) A theory of the term structure of interest rates. Econometrica 53:385–408
Daouk H, Guo JQ (2004) Switching asymmetric GARCH and options on a volatility index. J Futur Mark 24:251–282
Das RS, Sundaram R (1999) Of smiles and smirks: a term structure perspective. J Financ Quant Anal 34: 211–240
Demeterfi K, Derman E, Kamal M, Zou J (1999) More than you ever wanted to know about volatility swaps. J Deriv 6:9–32
Detemple J, Osakwe C (2000) The valuation of volatility options. Eur Financ Rev 4:21–50
Dotsis G, Psychoyios D, Skiadopoulos G (2007) An empirical comparison of continuous-time models of implied volatility indices. J Bank Financ 31:3584–3603
Duffie D, Pan J, Singleton K (2000) Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68:1343–1376
Engle RF, Ng VK (1993) Measuring and testing the impact of news and volatility. J Financ 48:1749–1778
Fleming J, Ostdiek B, Whaley RE (1995) Predicting stock market volatility: a new measure. J Futur Mark 15:265–302
Giot P (2005) Relationships between implied volatility indices and stock index returns. J Portf Manage 31:92–100
Grünbichler A, Longstaff F (1996) Valuing futures and options on volatility. J Bank Financ 20:985–1001
Heston S (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev Financ Stud 6:327–343
Heston S, Nandi S (2000a) Derivatives on volatility: some simple solutions based on observables. Working Paper, Federal Reserve Bank of Atlanta
Heston S, Nandi S (2000b) A closed-form GARCH option pricing model. Rev Financ Stud 13:585–626
Jones C (2003) The dynamics of stochastic volatility: evidence from underlying and options markets. J Econom 116:181–224
Kou SG (2002) A jump-diffusion model for option pricing. Manage Sci 48:1086–1101
Lin Y (2007) Pricing VIX futures: evidence from integrated physical and risk-neutral probability measures. J Futur Mark 27:1175–1217
Nelson DB (1990) ARCH models as diffusion approximations. J Econom 45:7–39
Pan J (2002) The jump-risk premia implicit in options: evidence from an integrated time-series study. J Financ Econ 63:3–50
Phillips PCB, Yu J (2005) Jackknifing bond option prices. Rev Financ Stud 18:707–742
Scott LO (1987) Option pricing when the variance changes randomly: theory, estimation, and an application. J Financ Quant Anal 22:419–438
Sepp A (2008) VIX option pricing in a jump-diffusion model. RISK 21:84–89
Simon DP (2003) The Nasdaq volatility index during and after the Bubble. J Deriv 11:9–24
Singleton KJ (2001) Estimation of affine asset pricing models using the empirical characteristic function. J Econom 102:111–141
Stein E, Stein J (1991) Stock price distributions with stochastic volatility: an analytic approach. Rev Financ Stud 4:727–752
Vuong Q (1989) Likelihood ratio tests for model selection and nonnested hypotheses. Econometrica 57:307–333
Wagner N, Szimayer A (2004) Local and Spillover shocks in implied market volatility: evidence for the US and Germany. Res Int Bus Financ 18:237–251
Whaley RE (1993) Derivatives on market volatility: hedging tools long overdue. J Deriv 1:71–84
Whaley RE (2000) The investor fear gauge. J Portf Manage 26:12–17
Zhu Y, Zhang JE (2007) Variance term structure and VIX futures pricing. Int J Theor Appl Financ 10:111–127