A Jump Diffusion Model for VIX Volatility Options and Futures

Dimitris Psychoyios[a], George Dotsis[b], Raphael N. Markellos[c]

Abstract

Implied volatility indices are becoming increasingly popular as a measure of market uncertainty and as a vehicle for developing derivative instruments to hedge against unexpected changes in volatility. Although jumps are widely considered as a salient feature of volatility, their implications for pricing implied volatility options and futures are not yet fully understood. This paper provides evidence indicating that the time series behavior of the VIX equity implied volatility 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 high levels, level effects of volatility and large upward movements during times of market stress. Based on this process, we develop closed form valuation models for volatility futures and options and show that incorrectly omitting jumps may cause considerable problems to pricing and hedging.

[a] Lecturer (Corresponding author), Manchester Accounting & Finance Group (MAFG), Manchester Business School, University of Manchester, Booth Street East, Manchester, M13 9PL, UK. Tel.+44 161 275 4492. E-mail: Dimitris.Psychoyios@mbs.ac.uk
[b] Lecturer, Dept. of Accounting, Management and Finance and Essex Finance Centre, University of Essex, Wivenhoe Park, Cochester C04 3SQ. Tel.: +44 120 687 2455. Fax.: +44 120 687 4261. E-mail: gdotsis@essex.ac.uk.
[c] Senior Lecturer, Department of Management Science and Technology, Athens University of Economics and Business, Office 915, 47A Evelpidon Str. 113 62, Athens, Greece; Visiting Research Fellow, Centre for Research in International Economics and Finance (CIFER), Loughborough University, UK. Tel. +30 210 8203671; Fax. +30 210 8828078. E-mail: rmarkel@aueb.gr.

1. Introduction

Volatility is undoubtedly the most important variable in finance. It appears consistently across a wide spectrum of theories and applications in asset pricing, portfolio theory, risk management, derivatives, corporate finance, investment evaluation and econometrics. Most of our obsession with the analysis of volatility has to do with the simple fact that it is not directly observable. A myriad of alternative measures and approaches have been developed in academia and industry in order to empirically measure volatility (for a selective review and references see Mills and Markellos, 2007).

A fascinating recent development has been the treatment of volatility as a distinct asset which can be packaged in an index and traded using volatility futures and options (hereafter referred to as "volatility derivatives"). Volatility derivatives are considered by some to "*have the potential to be one of the most important new financial innovations*" (Grünbichler and Longstaff, 1996). Traditionally, derivatives have allowed investors and firms to hedge against factors such as market volatility, interest rate volatility and foreign exchange volatility. Volatility derivatives provide protection against volatility risk, that is, unexpected changes in the volatility level itself. Such changes may arise as a response to changes in macroeconomic or microeconomic conditions (see, for example, Copeland *et al*., 2000).

The first volatility index, named VIX (currently termed VXO), was introduced in 1993 by the Chicago Board Options Exchange (CBOE). This was estimated from implied volatilities from at-the-money options on the SP100 index using a methodology proposed by Whaley (1993). The CBOE adopted a new methodology in 2003 to calculate VIX in a model-free manner as a weighted sum of out-of-money option prices across all available strikes on the S&P 500 index. Carr and Wu (2006) have demonstrated that the new VIX approximates the volatility swap rate, since it can represent the conditional risk-neutral expectation of the return volatility under general market settings. Several other implied volatility indices have been developed ever since, including: the VXN and VXD in the CBOE, the VDAX-NEW in Germany, the VX1 and VX6 in France, the VSTOXX in the Eurex, the VSMI in Switzerland, the MVX in Canada, etc. Volatility derivatives have been traded over the counter for several years, mainly as volatility swaps. However, only recently, in March 2004, the Chicago Board of Exchange (CBOE) introduced volatility futures on the implied volatility measured by the VIX index. The CBOE has announced the imminent introduction of volatility futures on the implied volatility index VXD along with volatility options. Eurex has launched in September 2005 new volatility futures on the VDAX-NEW, VSTOXX and VSMI volatility indices, respectively.

Options and futures written on a volatility index 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 risk, without necessarily resorting to dynamical adjustments. 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 volatility exposure by design. Volatility indices are also particularly useful in monitoring market expectations. The popular financial press, eg., CNBC, Barrons, Wall Street Journal, regularly quotes the VIX volatility index as an “investor fear gauge” (see also Whaley, 2000). Regulatory bodies and central banks, such as the Bank of England, have used the VIX to depict equity uncertainty and relate it to subsequent movements in other variables, such as swap spreads.[1] Volatility derivatives have a wide range of important applications for all market participants. Investment funds employ volatility derivatives for vega hedging their portfolios against movements in volatility. Certain classes of investors, such as convertible bond arbitrage funds and structured product issuers, can use these derivatives to insure against their structural

[1] For example, see Bank of England, *Quarterly Bulletin*, Winter 2003.

exposure to volatility. Investors can employ them to partially insure against shifts in transaction costs and tracking error penalties, both of which increase during periods of high uncertainty. Investment managers may use these derivatives to hedge against the risks of a so-called high-correlation environment. This is because, asset correlations have been found to increase significantly during periods of high volatility, making active asset picking and portfolio diversification very difficult. As volatility is a key input for risk management and capital adequacy models, such as the VaR, volatility derivatives could be used by banks as a shield against changes in volatility and correlation during stress market conditions. Since shifts in volatility have a significant impact on the risk premium that shareholders require above the risk-free rate, firms could employ volatility derivatives to protect themselves from unexpected subsequent changes in cost of capital. Moreover, shifts in market volatility are also likely to influence systematic equity risk and the return that shareholders require from a stock. Although not available yet, bond and foreign exchange volatility indices and derivatives, would allow firms that are exposed to volatility in these markets to hedge against changes in volatility. Finally, ample liquidity in this market is provided by traders and hedge funds since volatility derivatives can provide the most efficient and low-cost way for speculating against changes in volatility.

A number of recent empirical studies have examined the properties of implied volatility indices (e.g., Fleming et al., 1995; Moraux et al., 1999; Whaley, 2000; Blair, *et al.*, 2001; Corrado and Miller, 2003; Simon, 2003, and, Giot, 2005). This research has demonstrated the practical importance of at-the-money implied volatility as an efficient, yet biased, forecast of future realized volatility. There has been also been a growing interest in modeling the time series dynamics of the autonomous implied volatility process. Bakshi *et al.* (2006) estimated various general specifications of diffusion processes with a non-linear drift and diffusion component. The authors considered the squared implied volatility index VIX as a proxy to the unobserved instantaneous variance. Wagner and Szimayer (2004) investigated the presence of jumps in implied volatility by estimating an autonomous mean reverting jump diffusion process using data on the implied volatility indices VIX and VDAX. They found evidence of significant positive jumps in implied volatilities. However, they adopted the rather restrictive assumption that the volatility jump size is constant rather than being random. Finally, 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 U.S. volatility indices. They found that the best models in terms of fitting were those with random upward and downward jumps.

In response to the developments in the industry and academia,, Grünbichler and Longstaff (1996) developed the first models for the valuation of futures and European-style options written on instantaneous volatility. 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 formulas to price both American and European-style volatility options assuming a mean-reverting in log volatility model. The discrete time analogs in the limit of the volatility process used by these two studies are the GARCH and EGARCH processes, respectively. Heston and Nandi (2000a) derived analytical solutions in both discrete and continuous time for pricing European options written on variance. These were based on a discrete-time GARCH volatility process and its continuous time counterpart developed by Heston and Nandi (2000b). Recently, Daouk and Guo (2004), studied the valuation of volatility options based on a Switching Regime Asymmetric GARCH process for the underlying.

Motivated by the growing importance of volatility derivatives, this paper examines two main issues. First, it extends the empirical literature on implied volatility indices and evaluates the empirical performance of various diffusion and jump diffusion processes. This analysis is particularly useful in understanding the dynamics of the VIX and building an appropriate model for pricing options and futures. More specifically, we estimate the square root mean reverting process proposed by Grünbichler and Longstaff (1996) along with various jump diffusion variations. Overall, in line with previous research, we find that the addition of jumps improves fitting. Surprisingly, we find that the simple mean reverting log diffusion processes proposed by

Detemple and Osakwe (2000) outperforms all the square root type diffusion and jump diffusion processes, respectively. This result may be caused by misspecifications of the drift and the diffusion components of the square root mean reverting type processes. One must also consider that the log process is able to generate fast mean reversion at high levels. Moreover, its diffusion component takes into account the level effect of volatility as the level of implied volatility increases the volatility of implied volatility increases proportionally. In this manner, the simple mean reverting log diffusion allows rapid increases followed by fast mean reversion, which constitutes a salient feature of the VIX dynamics. In order to account for the possibility of large upward movements in the VIX during periods of market stress, we add a jump component to the log processes and we show that this provides the best fit in terms of various statistical metrics. Jumps are particularly important for accurately pricing options, especially in the short-term, since pure diffusions are not capable of producing realistic levels of higher moments at short horizons. Second, on the basis of the preliminary analysis and estimation results the paper develops closed form expressions for pricing futures and European options on implied volatility assuming that the logarithm of volatility follows a mean reverting process with jumps. The option pricing model proposed nests as a special case the model by Detemple and Osakwe (2000). We also assess the potential implications of incorrectly omitting jumps from the diffusion process by showing that prices and hedge ratios may differ substantially. In particular, the model without jumps in volatility (i.e., the Detemple and Osakwe, 2000, model) undervalues (overvalues) short (long) maturity options, on average, by 10% (6%), respectively. Moreover, it is far more sensitive to changes in the underlying with the delta hedging parameter being about 40% larger.

The remainder of the paper is structured as following. The next section analyses the empirical behavior of the daily VIX over a period of 10 years. Section 3, describes the mean-reverting volatility process considered along with two jump diffusion extensions. It also discusses estimation issues and empirical results. Section 4, develops valuation formulae for volatility futures and European options when the underlying volatility follows a mean-reverting jump-diffusion process. It also discusses the properties of these models and explores the potential importance of jumps from the perspective of pricing and risk management, respectively. The final section concludes the paper.

2. Empirical Properties of the VIX

We use daily closing index values of the VIX from 1/2/1990 to 9/13/2005, a total of 3,957 observations.[2] This implied volatility index is traded in the CBOE and is calculated from the weighted average of out-of-the money put and call S&P 500 Index option prices at two nearby maturities using a wide range of strikes. It should be noted that the construction of the VIX is independent of the model used to price the options and that the squared values of the index approximate the 30-day variance swap rate (Carr and Wu, 2006). Figure 1 depicts the evolution of the VIX and of its first differences for the period under study. The plots suggest a volatile mean-reverting behavior for the levels with violent swings while the first differences appear heteroskedastic with a number of spikes.

[INSERT FIGURE 1 HERE]

The summary statistics of the series, shown in Table 1, largely confirm this behavior. The VIX ranges between about 9% to 45%, with an average of 19.6%. The higher moments suggest a leptokurtotic distribution skewed to the right for both levels and differences. The Jarque-Bera test rejects the normality assumption at a high level of confidence. Autocorrelations die out slowly in levels, something consistent with a smooth, possibly mean reverting process. Differences appear anti-persistent with small negative short-term autocorrelations. The highly significant squared autocorrelations strongly suggest heteroskedasticity.

[2] Data are drawn from the website of the CBOE.

[INSERT TABLE 1 HERE]

[INSERT TABLE 2 HERE]

Given that simple Brownian motion processes have also been employed in the literature to model volatility indices, we examine the stationarity of the VIX levels. The Augmented Dickey-Fuller (Dickey and Fuller, 1979) and Phillips-Perron (1988) tests both reject the null hypothesis of a unit root with a high level of confidence. However, the null hypothesis of stationarity cannot be accepted on the basis of the Kwiatkowski-Phillips-Schmidt-Shin (KPSS, 1992). Since the KPSS test is known to be sensitive to long-memory (see, for example, Lee and Schmidt, 1996) and motivated by relevant empirical findings in the literature with respect to long-memory in historical volatility (e.g., Ding, *et al*, 1993; Baillie *et al*. 1996, by Breidt, *et al*., 1998), we examine further this possibility. Lo's (1991) modified R/S test statistic for long range dependence is significant at the 5% level with a value of 9.4875. The Geweke and Porter-Hudak (1983) log-periodogram method implemented with the trimming and smoothing options proposed by Robinson (1995), produced an estimate of fractional unit root d equal to 0.7236 ($p = 0.0736$). One must view this evidence with caution since long-memory tests are sensitive to a variety of factors such as structural breaks, outliers, regime switching and nonlinear transformations (see, for example, Diebold and Inoue, 2001; Engle and Smith, 1999; Dittmann and Granger, 2002). Moreover, it is possible that long-memory behavior is the result of aggregation in constructing the VIX. Granger (1980) pointed out that the summation of low-order ARMA processes will yield ARMA processes of increasing, and eventually infinite order which can be well approximated using an ARFIMA model. Notwithstanding, on the basis of the results presented, although the possibility of long-memory characteristics in the VIX cannot be excluded, it will not be further entertained

[INSERT FIGURE 2 HERE]

We proceed in examining the unconditional distribution of the VIX levels and differences. As shown by the results contained in Figure 2, the unconditional distribution of the VIX closely resembles the shape of a highly skewed distribution, such as the chi-squared. The distribution of differences is clearly leptokurtotic. Fitting a variety of distributions via maximum likelihood (ML) is consistent with these suggestions, the results given in Table 3. The distribution of VIX levels appears to be well approximated by a skewed *t*-student and a non-central chi-squared distribution. The log-normal and the extreme (max) distributions also appear to fit relatively well the VIX levels. The unconditional distribution of VIX differences is well approximated by a *t*-student.[3] The normal distribution offers a relatively poor fit for both levels and differences.

[INSERT TABLE 3 HERE]

A more detailed breakdown of the unconditional distributions is presented in Table 4. Given that the standard deviation of differences is around 0.0122 with a mean very close to zero, we can observe 20 distinct four-standard deviation events, 8 downward and 12 upward. Under a normal distribution, which is consistent with some diffusion models of volatility, the variance implies that these events should occur with probability under 0.005% or once in about every 80 years. Here, we observe a much higher probability of occurrence, 100 times higher, of over 0.5%, or, once in every 164 days. These findings are expected, given the fat-tails in the ΔVIX distribution and could be due also to jumps in the underlying process. One must be careful in interpreting large negative changes as downward jumps since they are not that unexpected: in all but two cases they are preceded by large increases in the VIX. Hence, they could also be the result of

[3] Although results are not shown here, differences remain highly non-normal even if estimated as logarithmic ratios.

heteroskedasticity and mean reversion. Finally, we can also see that the likelihood of large upward movements in volatility seems to increase with the volatility level. For example, large volatility changes over 5% appear with probability 0.29% (4/1,369), 2.34% (6/256) and 6.9% (2/29) for volatility levels in the [0.2, 0.3), [0.3, 0.4) and [0.4, 0.5) range, respectively.

[INSERT TABLE 4 HERE]

3. Diffusion and Jump Diffusion Processes for the VIX

3.1 Diffusion Processes

One of the simplest processes to model volatility is the Mean Reverting Gaussian Process (also called Ornstein – Uhlenbeck). It was initially proposed in order to capture the mean reverting empirical property of volatility (e.g., Hull and White, 1987; Stein and Stein, 1991; Scott, 1987; Brenner *et al*, 2006). Under this process, the implied volatility changes are normal, something that is clearly rejected from our empirical analysis of the VIX. Moreover, this process has the significant disadvantage of allowing negative values. Two of the most popular alternative processes that have been developed in the literature are the Mean Reverting Square Root Process (SR) and the Mean Reverting Logarithmic Process (LR), given by equations (1), (2), respectively:[4]

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

$$d\left(lnV_t\right) = k\left(\theta - \left(lnV_t\right)\right)dt + \sigma dZ_t \tag{2}$$

[4] See Hull and White (1988), Heston (1993), Ball and Roma (1994), Heynen et al. (1994), Grünbichler and Longstaff (1996), Bates (2000), and, Jones (2003) for the case of SR, and Wiggins (1987) and Detemple and Osakwe (2000) for he case of LR.

where V_t is the value of VIX at time t, dZ_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. Equations (1) and (2) are defined under the actual probability measure P.

Both processes can be obtained as a limit of ARCH-type processes. In particular, Heston and Nandi (2000b) have shown that a degenerate case of the 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 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 the two 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 diffusion processes belonging to the CEV class (see Chan *et al.*, 1992) since option pricing becomes infeasible due to the intractability of the characteristic function (see Duffie *et al.*, 2000). The density function can be approximated via Taylor expansion in the time domain (see Ait-Sahalia, 1999; Bakshi *et al.*, 2006) and econometric estimation is possible using discretely sampled data. However, using this approach, only very short term options can be accurately priced.

3.2 Jump-Diffusion Processes

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

Square-Root Process with Jumps (SRJ) $$dV_t = k\left(\theta - V_t\right)dt + \sigma\sqrt{V_t}dZ_t + ydq_t \quad (3)$$

Square-Root Process with proportional Jumps (SRPJ) $$dV_t = k\left(\theta - V_t\right)dt + \sigma\sqrt{V_t}dZ_t + ydq_t \quad (4)$$

Logarithmic Process with Jumps (LRJ) $$d\left(lnV_t\right) = k\left(\theta - \left(lnV_t\right)\right)dt + \sigma dZ_t + ydq_t \quad (5)$$

where dZ_t is a standard Brownian motion, dq_t is a compound Poisson process and y is the jump amplitude. dZ and dq are assumed to be independent processes. Again, equations (3), (4), and (5) are all defined under the actual probability measure P. We further assume that the jump size is drawn from exponential distribution:

$$f(y) = p\eta e^{-\eta y} 1_{\{y \geq 0\}} \quad (6)$$

where $1/\eta$, is the mean of the upward jump. The exponential distribution allows us to capture upward jumps in implied volatility and to derive the characteristic function in closed form (see the Appendix A1 for the derivation of the characteristic functions). The one sided exponential distribution adopted is a version of the double exponential distribution used by Kou (2004) in modeling the dynamics of stock and index prices.

In the SRJ and LRJ processes, dq_t has a constant arrival parameter λ, whereas in the SRPJ process the arrival parameter is proportional to V_t, that is, $\Pr\{dq_t=1\} = \lambda V_t dt$. The latter means that in cases of the SRJ and LRJ, the probability of a jump is independent of the current level of implied volatility, while in the case of SRPJ the probability of jump is proportional to the current level of implied volatility. It can easily be verified that in term of volatility levels, the process in (5) can be written as:

$$dV_t = V_t \left[k\left(\theta - \ln(V_t)\right) dt + \sigma dW_t + (e^y - 1) dq \right] \quad (7)$$

Inspection of equation (7) 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 depend on the current level of implied volatility, respectively. The proportional structure of this model has three important implications. First, the model accounts for the level effect of the volatility, i.e., when implied volatility increases then the volatility increases proportionally. This allows capturing relative large changes of V_t which are likely to be characterized as jumps under the SRJ or the SRPJ. Second, since mean reversion depends on the level of V_t, i.e., the larger the V_t , the larger the mean reversion, the LRJ is able to produce “spikes”, rather than jumps, which is consistent with our preliminary descriptive analysis of the VIX. Third, the LRJ process allows for size jumps to depend on the level of implied volatility and is thus capable of generating large upward movements which are consistent with the behavior 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 type processes are able to partially capture this behavior through their fast mean reversion. Also, we do not attempt to account for long-memory or more complicated nonlinear dynamics in the data since these have been examined in detail by other studies and are outside the scope of this paper (see, for example, Bakshi *et al.*, 2006; Daouk and Guo, 2004; Bollerslev and Mikkelsen, 1996).

4. Estimation Results

Table 5 shows the ML estimation results using the VIX sample (see Appendix A2 for details on ML Estimation). For each process we report: the estimated parameters (annualized), the asymptotic *t*-statistics (within brackets), the log-likelihood (LL) values, the Akaike Information

Criterion (AIC) and the Bayes Information Criterion (BIC). The two information criteria are employed for comparing non-nested models. According to all statistical criteria, the best fit is provided by the LRJ, followed by the LR, SRPJ, SRJ and SR, respectively. This ranking implies that the VIX is characterised by: a) fast mean reversion at high levels, b) increase in volatility when implied volatility increases, and, c) jumps proportional to the current level of implied volatility.

Amongst the square-root type processes, the SRPJ process displays the highest log-likelihood value. Since the models are nested, the likelihood ratio (LR) test can be employed to compare the relative goodness-of-fit. We find that the likelihood of the SRJ is significantly higher than that of the SR, the relevant LR test statistic being 318.5 (the critical value at the 1% level from a Chi-squared with two degrees of freedom is 9.21). Allowing the probability of jumps to be proportional to volatility, produces a further statistically significant improvement in likelihood (LR= 73.74). The information criteria also suggest that the addition of jumps in proportion to the volatility level improves fitting. Another point to be emphasized 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. The better fit of the LR is also verified by the information criteria. This result should not come as a surprise since, as mentioned previously, the LR is capable of generating large increases in implied volatility at high levels, followed by rapid mean reversion. Essentially, changes that appear as jumps can also be generated by suitable diffusion components. This result points to the conclusion that before adding a jump factor, it is crucial to specify correctly the drift and diffusion.

The inclusion of jumps in the LR enhances statistical fitting even further. Once the drift and the diffusion components are correctly specified, the inclusion of jumps allows capturing additional skewness. This is also evident from the fact that σ drops from 0.88 to 0.75, which implies that jumps account for a substantial component of volatility, as expected intuitively. The estimate of the Poisson arrival rate implies 40 jumps per year with a 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.

[INSERT TABLE 5 HERE]

In order to check the stability of the parameters we divide the sample into two equal parts and we re-estimate the processes. The results for the first and second subsample are reported in Table 6. For all processes we can draw the following general conclusions. First, the diffusion coefficient (σ) displays a stable behavior in both subsamples when compared to the complete sample. Second, 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 estimation requires large data sets (e.g., Phillips and Yu, 2005). Third, the long run mean is higher (lower) in the first (second) subsample when compared to the complete sample. 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. An interesting extension for future research would be to build a two factor model and allow the long run mean to follow another stochastic process. Under this set up, one should take into account the fact that the number of parameters to be estimated increases substantially. Since the long run mean is unobserved, the two factor model can be estimated by means of Kalman filter.

Fourth, 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.

[INSERT TABLE 6 HERE]

5. Pricing of Volatility Derivatives

In this section we derive analytical formulae for pricing option and futures contracts on volatility indices when the underlying follows a Mean Reverting Logarithmic Process with Jumps (LRJ). We also examine the properties of the proposed models and investigate the potential implications of incorrectly omitting jumps from the volatility process.

5.1 Volatility Futures

Before proceeding to futures valuation, we must rewrite equation (5) under the risk neutral probability measure *Q*. Since VIX is not a tradable asset, implying that the market is not complete, the equivalent martingale measure *Q* is not unique and the actual measure that should be used to price derivatives is determined by preferences. 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., $\zeta_t=\zeta \; lnV$ (see also Christoffersen, *et al*, 2006). So, the volatility process under the risk neutral probability measure *Q* is given by:

$$d\left(lnV_t\right)=\left(k+\zeta\right)\left(\frac{k\theta}{k+\zeta}-\left(lnV_t\right)\right)dt+\sigma d\tilde{Z}_t+ydq_t \quad (8)$$

or, equivalently,

$$d\left(lnV_t\right) = k^*(\theta^* - \left(lnV_t\right)) + \sigma d\tilde{Z}_t + Jdq \tag{9}$$

where $k^* = k + \zeta$ and $\theta^* = \dfrac{k\theta}{k+\zeta}$.

Now denote *$F_t(V,T)$* the price of a futures contract on V_t at time t with maturity T. Under the risk neutral probability measure Q, *$F_t(V,T)$* is determined by the conditional expectation of V_T at time t. This expectation is conditional on the information up to time t:

$$F_t(V,T) = E_t^Q(V_T), \qquad t < T \tag{10}$$

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

$$E_t\left(V_T\right) = Exp\left[e^{-k^*(T-t)}ln(V_t) + \theta^*\left(1-e^{-k^*(T-t)}\right) + \frac{\left(1-e^{-2k^*(T-t)}\right)}{4k^*}\times\sigma^2 + \frac{\lambda}{k^*}\times ln\left(\frac{\eta - e^{-k^*(T-t)}}{\eta-1}\right)\right] \tag{11}$$

Equation (11) consists of four terms: the first, the second, and the third term correspond to the diffusion part of the LRJ, while the third term corresponds to the jump part.

The futures pricing formula (11) has the following limiting properties:

i. $$\lim_{(T-t)\to 0} E_t\left(V_T\right) = V_t, \tag{12}$$

ii. $$\lim_{(T-t)\to+\infty} E_t\left(V_T\right) = Exp\left(\theta^* + \frac{\sigma^2}{4k^*} + \frac{\lambda}{k^*} ln\left(\frac{\eta}{\eta-1}\right)\right), \tag{13}$$

iii. $$\lim_{V\to 0} E_t\left(V_T\right) = Exp\left[e^{-k^*(T-t)} + \theta^*\left(1-e^{-k^*(T-t)}\right) + \frac{\left(1-e^{-2k^*(T-t)}\right)}{4k^*}\times\sigma^2 + \frac{\lambda}{k^*}\times ln\left(\frac{\eta - e^{-k^*(T-t)}}{\eta-1}\right)\right]. \tag{14}$$

Equation (12) shows the standard convergence property of the futures price to the spot price at maturity. Equation (13) 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 are becoming less sensitive to current volatility changes and fail to capture the stochastic evolution of the VIX. Finally, equation (14) shows that as volatility tends to zero, futures prices does not converge to zero, as in the case of futures on stocks or stock indices. The intuition of this property of futures prices can be partially related to the mean reverting nature of volatility (see Grünbichler and Longstaff, 1996, for a discussion on the effect of mean reversion on volatility futures prices) and to the existence of jumps. Even if V equals zero, there is always a probability that V is going to jump to a positive number prior to the expiration of the contract.

5.2 Volatility Options

In this section 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). The price $C(V_t, \tau; K)$ of the call option with strike price *K* and time to maturity *τ* is given by:

$$C(t, T-t; K) = e^{-r(T-t)}\left[V_t^{e^{-k(T-t)}} G(t, T-t)\Pi_1(t, T-t) - K\Pi_2(t, T-t)\right] \tag{15}$$

The probabilities Π_1 and Π_2 are determined by

$$\Pi_j(t,\tau) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-is(lnK)} \times f_j(t, T-t; s)}{is}\right] ds\ ,\ j=1,2 \tag{16}$$

where

$$G(t, T-t) = Exp\left(\theta^*\left(1 - e^{-k^*(T-t)}\right) + \frac{\left(1 - e^{-2k^*(T-t)}\right)}{4k^*} \times \sigma^2 + \frac{\lambda}{k^*} \times \ln\left(\frac{\eta - e^{-k^*(T-t)}}{\eta - 1}\right)\right),$$

$$f_1(t, T-t; s) = \frac{f(t, T-t; s-i)}{f(t, T-t; -i)},$$

$$f_2(t, T-t; s) = \frac{f(t, T-t; s)}{f(t, T-t; 0)},$$

$$f(t, T-t; s) = e^{-r(T-t)} F(ln(V_t), T-t; s)$$

The call pricing formula has the following limiting properties:

i. $$\lim_{T-t \to 0} C(V_t, T-t; K) = max(V_t - K, 0), \tag{17}$$

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

iii. $\lim_{V \to 0} C(V_t, T-t; K) = 0.$[5] (19)

Equation (17) shows the standard convergence property of the option price to the option's payoff at maturity. Equation (18) 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, equation (19) suggests 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 that a similar option preserves a non-zero value since the later assumes that V follows a SR. 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, Figure 3 shows the value of a volatility call option as a function of volatility for three different levels of moneyness. We consider the diffusion model of Detemple and Osakwe (2000) along with jump-diffusion specifications. We can see that for short (long) maturities the diffusion model underprices (overprices) the volatility call in comparison to the jump-diffusion. The overpricing occurs because in the jump-diffusion model the volatility of the process consists of two parts: the diffusion and the jump part. The jump part affects the value of the volatility call mainly in the short-run, whilst the diffusion part affects the value of the volatility call mainly in the long-run.[6] On the other hand, the volatility of the diffusion model is driven only by the diffusion part. Note that although the total volatility is almost the same for both the diffusion and jump diffusion model, σ is significantly larger in the case of the diffusion model. In this manner, the diffusion

[5] We see no economic reason to investigate the case $\lim_{V \to \infty} C(V_t, T-t; K)$. 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.

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

model underprices the volatility call for short maturities where jumps in volatility still affect the call value.

[INSERT FIGURE 3 HERE]

Figures 4 and 5 depict the *delta* of the diffusion and jump-diffusion models as a function of τ and V_t, respectively. Interestingly, the *delta* of the diffusion model is significantly higher in all cases. The latter indicates that the diffusion model is more sensitive to volatility changes than the jump-diffusion model. The explanation follows again from the fact that σ is significantly larger in the case of the diffusion model. Differentiation shows that the delta of the volatility calls depends mainly on σ rather than λ or η. This finding has important implications in terms of hedging. For example, suppose that you have a long position in a call option and you use volatility options in order to hedge the vega risk of your position. Recall that the diffusion model overestimates the *delta* of the volatility option. So, if you incorrectly use the diffusion model to calculate the *delta*, then you will use less volatility options for hedging than those that are actually required.

[INSERT FIGURE 4 HERE]

[INSERT FIGURE 5 HERE]

5.3 Basis Risk

As we have already mentioned, futures and options written on the VIX were introduced as more effective instruments than traditional approaches (e.g., straddles, butterfly spreads) for hedging volatility risk. However, strictly speaking, VIX derivatives can be used directly only for hedging volatility of positions on the underlying index, i.e., the S&P500. To the extent that the VIX is a good proxy for market risk, VIX derivatives can also be used also to hedge against shifts in market volatility for positions on other broad stock market indices or on widely diversified portfolios. However, in the case of volatility hedging for individual stocks, a basis risk problem arises. This risk is due to the fact that the VIX is fundamentally linked to the volatility of options written on S&P500 index which may be different that the volatility of an individual stock. So, using futures and options written on the VIX to hedge the volatility risk of individual stocks or of undiversified portfolios may expose the hedger to basis risk. Assume an investor is holding a stock or an undiversified portfolio of stocks denoted by *i*. Following Bakshi and Kappadia (2003), we assume that the relationship between the implied volatility of the portfolio and the index is described by:[7]

$$V_i^t = \beta_i V_m^t + Z_i^t \tag{20}$$

Where V_i^t is the implied volatility of the portfolio, V_m^t is the implied volatility index VIX, and Z_i^t is the idiosyncratic volatility component. The dynamics of the portfolios' implied volatility are given by:

$$dV_i^t = \beta_i dV_m^t + dZ_i^t \tag{21}$$

According to equation (21), we can consider two scenarios for the dynamics of the portfolio's implied volatility. If idiosyncratic volatility is constant for all *t* then dV_i^t can be described by a one dimensional process. Assuming a perfect correlation between volatility futures and the underlying, the portfolio can be hedged by buying β futures contracts. This assumes also that the portfolio's dollar value equals the dollar value of the index. Apart from the basis risk arising from

[7] Bakshi and Kappadia (2003) use a similar relationship for the variances and show that this relationship implicitly assumes a single market model for the log returns. In our cases, the correspondence is not exact because we are dealing with volatilities.

the cross-hedge, since we have assumed that VIX is well described by a jump diffusion process, the hedge may also be exposed to basis risk from the jump component of the market. However, because of the mean reverting and "spiky" behavior of the VIX, jumps will expose the hedge to significant basis risk only if they occur just before the expiration of the contract. Moreover, if idiosyncratic volatility is not constant over time, the hedge will be exposed to substantial basis risk. The magnitude of basis risk will depend on the exact dynamics of the idiosyncratic volatility.

Another cause for basis risks arises from the fact that the VIX is not a traded asset. Hence, in the absence of arbitrage the futures price is not directly tied to the VIX, possibly resulting to substantial basis risk for the hedger. Towards this direction, Carr and Wu 2006 show that futures prices are bounded between the forward volatility swap rate and the forward variance swap rate. Even so, the capitalization of the arbitrage profits is rather difficult since we have to actively trade either a basket of options on SPX or exotic OTC derivatives, such as forward-start ATM forward call options (see Carr and Wu 2006). 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.

6. Conclusions

Motivated by the growing literature on volatility derivatives and their imminent introduction in major exchanges, this paper examined the empirical relevance and potential impact of volatility jumps in autonomous volatility option pricing and risk management.

Empirical analysis of the VIX over a period of 10 years provided a wealth of evidence supporting the existence of some stationary, mean-reverting process with jumps. An ML estimation scheme was applied to VIX data for a variety of processes. The results suggested that a simple log mean reverting diffusion performs better compared to square root diffusion or jump diffusion processes, respectively. This surprising result was attributed to the misspecification of the drift and diffusion component of square root processes and to the fact that the log processes are capable of generating fast mean reversion and level effects. Moreover, we showed that when a jump component is added to the log processes performance is further enhanced. On the basis of the estimation results, we developed closed form models for pricing futures and options on the VIX. The proposed volatility option pricing models nests as a special case the Detemple and Osakwe (2000) and appears to have comparable properties. The model without jumps in volatility undervalues (overvalues) short (long) maturity options, on average, by 10% (6%), respectively. Moreover, it is far more sensitive to changes in the underlying with the delta hedging parameter being 40% larger.

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 model 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 no volatility options market data are still inadequate, we cannot fully test the empirical relevance of alternative option pricing models. However, it is crucial to fully understand the dynamics of the underlying and the implications of competing option pricing models in order to understand the peculiarities of this asset class and facilitate a smooth functioning of the market when it operates.

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 are the actual benefits they may expect. This is not a trivial problem since the implications of volatility for a firm are so widespread, complicated 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. This complicates also the accounting treatment of the hedge relationship and effectiveness offered by volatility derivatives. For example, according to FAS 133, the statement issued by F.A.S.B. (Financial Accounting Standards Board) regarding accounting for derivative instruments and hedging accounting, three hedge relationships are recognized: fair value hedge, cash flow hedge and foreign currency hedge. The accounting treatment of derivatives depends on the hedge relationship they participate and the effectiveness of the hedge offered. In the case of volatility derivatives, the determination of the hedge relationship and the effectiveness is a very difficult task.

In closing, we would like to emphasize the growing need for introducing volatility indices and derivatives in more markets. Brenner and Galai (1989) first argued that volatility indices should be developed for equity, bond and foreign exchange markets. However, the recent history has shown that significant volatility risk exists also in other important markets, such as, for example, the market for petrol and for electricity.

Appendix A1: 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, like the SRJ, LRJ and SRPJ, has the following exponential affine form:

$$F(V_t, T-t; s) = \exp\left(A(T-t; s) + B(T-t; s)V_t \right) \tag{22}$$

Thus, for the case of the SRJ, $A(T-t;s)$ and $B(T-t;s)$ are given by[8]:

$$A(T-t;s) = a\left(T-t,s\right) + z\left(T-t,s\right) \tag{23}$$

$$a(T-t;s) = -\frac{2k\theta}{\sigma^2} \times ln\left(\frac{k - \frac{1}{2} i\sigma^2 s\left(1-e^{-k(T-t)}\right)}{k} \right) \tag{24}$$

$$z(T-t;s) = \frac{2\lambda p}{2k - \eta\sigma^2} \times ln\left(\frac{k - \frac{1}{2} i\sigma^2 s + is\left(\frac{\sigma^2}{2} - \frac{k}{\eta}\right)e^{-k(T-t)}}{k - \frac{isk}{\eta}} \right) \tag{25}$$

and,

$$B(T-t;s) = \frac{ksie^{-k(T-t)}}{k - \frac{1}{2} i\sigma^2 s\left(1-e^{-k(T-t)}\right)} \tag{26}$$

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

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

$$F((lnV_t),T-t;s)=\exp\left(A(T-t;s)+B(T-t;s)(lnV_t)\right) \tag{27}$$

where

$$A(T-t;s)=is\theta(1-e^{-k(T-t)})-s^2\sigma^2\left(\frac{1-e^{-2k(T-t)}}{4\kappa}\right)+\frac{\lambda}{k}\times ln\left(\frac{\eta-ise^{-k(T-t)}}{\eta-is}\right) \tag{28}$$

$$B(T-t;s)=ise^{-k(T-t)} \tag{29}$$

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 found numerically. So, the conditional characteristic function $F(V_t,T-t;s)=E(e^{isV_T}\left|V_t\right.)$ of the SRPJ must satisfy the following Kolmogorov backward differential equation:

$$\frac{\partial F}{\partial V_t}+k(\theta-V_t)+\frac{1}{2}\frac{\partial^2 F}{\partial V_t^2}V_t\sigma^2-\frac{\partial F}{\partial \tau}+\lambda V_t\mathrm{E}\left[F(V_t+y)-F(V_t)\right]=0 \tag{30}$$

subject to the boundary condition

$$F(V_t,T-t=0;s)=e^{isV_t} \tag{31}$$

where $i=\sqrt{-1}$. Differentiating the characteristic function given by equation (22) yields

$$\begin{aligned}F_V&=BF\\F_{VV}&=B^2F\\F_{T-t}&=F\left(A_{T-t}+VB_{T-t}\right)\end{aligned} \tag{32}$$

where the subscripts denote the corresponding partial derivatives.

Replacing equations (32) in equation (30) and rearranging yields

$$V_t\left(-kB - B_{T-t} + \frac{1}{2}\sigma^2 B^2 + \lambda \mathrm{E}\left[e^{yB} - 1\right]\right) + \left(k\theta B - A_{T-t}\right) = 0 \tag{33}$$

Also,

$$\mathrm{E}\left[e^{yB} - 1\right] = \int_0^{+\infty} \eta e^{-\eta y} e^{yB} dy - 1 = \frac{\eta}{\eta - \mathrm{B}} - 1$$

Since $V_t \neq 0$, the expressions in the parentheses in equation (33) must equal zero. Therefore, we obtain the following ordinary differential equations (ODEs)

$$-kB - B_{T-t} + \frac{1}{2}\sigma^2 B^2 + \lambda\left(\frac{\eta}{\eta - \mathrm{B}} - 1\right) = \tag{34}$$

$$k\theta B - A_{T-t} = 0 \tag{35}$$

The ODEs cannot be solved in closed form. They are solved numerically subject to the boundary conditions $A(T - t = 0; s) = 0$, and $B(T - t = 0; s) = is$.

Appendix A2: Maximum-Likelihood Estimation

Suppose that $\{V_t\}_{t=1}^{T}$ is a discretely sampled time series of implied volatilities. Assume that we stand at time *t*, and *τ* denotes the sampling frequency of observations. Then, the Fourier inversion of the characteristic function $F(V(t),\tau;s)$ provides the required conditional density function $f[V(t+\tau)|V(t)]$:

$$f[V(t+\tau)|V(t)] = \frac{1}{\pi}\int_0^\infty \mathrm{Re}[e^{-isV(t+\tau)}F(V(t),\tau;s)]ds \tag{36}$$

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:

$$\Im = \max_{\{\Theta\}} \sum_{t=1}^{T} \log\left(\frac{1}{\pi}\int_0^\infty \mathrm{Re}[e^{-isV(t+\tau)}F(V(t),\tau;s)]ds\right) \tag{37}$$

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

References

Ait-Sahalia, Y., (1999), 'Transition Densities for Interest Rate and other Nonlinear Diffusions', *Journal of Finance*, Vol. 54, pp. 1361-1395.

Baillie, R.T, T. Bollerslev and H.O. Mikkelsen (1996), 'Fractionally integrated generalized autoregressive conditional heteroscedasticity', *Journal of Econometrics*, Vol. 74, pp. 3-30.

Bakshi, G., and C. Cao (2006), 'Risk-Neutral Kurtosis, Jumps, and Option Pricing: Evidence from Most Actively Traded Firms on the CBOE', *Working Paper*, University of Maryland.

Bakshi, G. and N. Kappadia (2003), 'Volatility Risk Premium Embedded in Individual Equity Options: Some New Insights', *Journal of Derivatives*, Vol. 3, pp. 45-54.

Bakshi, G., and D. Madan (2000), 'Spanning and derivative-security valuation', *Journal of Financial Economics*, Vol. 55, pp. 205-238

Bakshi, G., N. Ju and H. Ou-Yang (2006), 'Estimation of Continuous-Time Models with an Application to Equity Volatility Dynamics' *Journal of Financial Economics*, Vol. 82, pp. 227-249.

Ball, C., and A. Roma (1994), 'Stochastic Volatility Option Pricing', *Journal of Financial and Quantitative Analysis*, Vol. 29, pp. 589-607.

Bates, D. (2000), 'Post-87 crash fears in S&P 500 futures options', *Journal of Econometrics*, Vol. 94, pp. 181-238.

Blair, B.J., S. Poon and S.J. Taylor (2001), 'Forecasting S&P 100 Volatility: the Incremental Information Content of Implied Volatilities and High-Frequency Index Returns', *Journal of Econometrics*, Vol. 105, pp. 5-26.

Bollerslev, T. and H. Mikkelsen (1996), 'Modelling and Pricing Long Memory in Stock Market Volatility', *Journal of Econometrics,* Vol. 73, pp. 151-184.

Breidt, F.J., N. Crato, and P. de Lima (1998), 'The detection and estimation of long memory in stochastic volatility', *Journal of Econometrics,* Vol. 83, pp. 325–348.

Brenner, M., and D. Galai (1989), 'New Financial Instruments for Hedging Changes in Volatility', *Financial Analysts Journal*, Vol. 4, pp. 61-65.

Brenner, M., and D. Galai (1993), 'Hedging Volatility in Foreign Currencies', *Journal of Derivatives*, Vol. 1, pp. 53-59.

Brenner, M., E. Ou, and J. Zhang (2006), 'Hedging Volatility Risk', *Journal of Banking and Finance*, Vol. 30, pp. 811-821.

Carr, P. and L. Wu, (2006), 'A Tale of Two Indices', *Journal of Derivatives,* Vol. 13, pp. 13-29.

Copeland, L., S. H. Poon, and R.C. Stapleton (2000), 'The Determinants of Implied Volatility: A Test Using LIFFE Option Prices', *Journal of Business Finance and Accounting*, Vol. 27, pp. 859-885.

Chan K.C., G.A. Karolyi, F.A. Longstaff,, and A.B. Sanders (1992), 'An Empirical Comparison of Alternative Models of the Short-Term Interest Rate', *Journal of Finance*, Vol. 47, pp. 1209-1227.

Corrado, C.J. and T. Miller (2005), 'The Forecast Quality of CBOE Implied Volatility Indexes', *Journal of Futures Markets*, Vol. 25, pp. 339-373.

Cox, J.C., J.E. Ingersoll and S.A. Ross (1985), 'A Theory of the Term Structure of Interest Rates', *Econometrica*, Vol. 53, pp. 385-408.

Christoffersen, P., K. Jacobs, and K. Mimouni (2006), 'Models for S&P500 Dynamics: Evidence from Realized Volatility, Daily Returns, and Option Prices', *Working Paper*, McGill University.

Daouk, H. and J.Q. Guo (2004), 'Switching Asymmetric GARCH and Options on a Volatility Index', *Journal of Futures Markets*, Vol. 24, pp. 251-282.

Das, R.S. and R. Sundaram (1999), 'Of Smiles and Smirks: A Term Structure Perspective', *Journal of Financial and Quantitative Analysis*, Vol. 34, pp. 211-240.

Detemple, J. and C. Osakwe (2000), 'The Valuation of Volatility Options', *European Finance Review*, Vol. 4, pp. 21-50.

Dickey, D.A. and W.A. Fuller (1979), ‘Distribution of the Estimators for Autoregressive Time Series with a Unit Root’, *Journal of the American Statistical Association*, Vol. 74, pp. 427-431.

Diebold, F. and A. Inoue (2001), ‘Long memory and regime switching’, *Journal of Econometrics*, Vol. 105, pp. 131–159.

Ding, Z., C.W.J. Granger and R. Engle (1993), ‘A long memory property of stock market returns and a new model’, *Journal of Empirical Finance,* Vol. 1, pp. 83–106.

Dittman, I. and C.W.J. Granger (2002), ‘Properties of nonlinear transformations of fractionally integrated processes’, *Journal of Econometrics*, Vol. 110, pp. 113–133.

Dotsis, G., D. Psychoyios and G. Skiadopoulos (2007), ‘An Empirical Comparison of Continuous-Time Models of Implied Volatility Indices’, *Journal of Banking and Finance* (forthcoming).

Duffie, D., J. Pan and K. Singleton (2000), ‘Transform analysis and asset pricing for affine jump-diffusions’, *Econometrica*, Vol. 68, pp. 1343-1376.

Engle, R.F. and V. K. Ng (1993), ‘Measuring and testing the impact of news and volatility’, *Journal of Finance*, Vol. 48, pp. 1749-1778.

Engle, R.F. and A.D. Smith (1999), ‘Stochastic permanent breaks’, *Review of Economics and Statistics*, Vol. 81, pp. 553–574.

Fleming, J., B. Ostdiek and R.E. Whaley (1995), ‘Predicting Stock Market Volatility: A New Measure’, *Journal of Futures Markets*, Vol. 15, pp. 265-302.

Geweke, J. and S. Porter-Hudak (1983), ‘The estimation and application of long-memory time series models’, *Journal of Time Series Analysis,* Vol. 4, pp. 221-237

Giot, P. (2005), ‘Relationships between implied volatility indices and stock index returns’, *Journal of Portfolio Management*, Vol. 31, pp. 92-100.

Granger, C.W.J. (1980), ‘Long memory relationships and the aggregation of dynamic models’, *Journal of Econometrics*, Vol. 14, pp. 227–238.

Grünbichler, A. and F. Longstaff (1996), 'Valuing Futures and Options on Volatility', *Journal of Banking and Finance*, Vol. 20, pp. 985-1001.

Heston, S. (1993), 'A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options', *Review of Financial Studies*, Vol. 6, pp. 327–343.

Heston, S. and S. Nandi (2000a), 'Derivatives on Volatility: Some Simple Solutions Based on Observables', *Working Paper* , Federal Reserve Bank of Atlanta.

Heston, S. and S. Nandi (2000b), 'A Closed-Form GARCH Option Pricing Model', *Review of Financial Studies*, Vol. 13, pp. 585-626

Heynen, R., A.G.Z. Kemma and T. Vorst (1994), 'Analysis of the Term Structure of Implied Volatilities', *Journal of Financial and Quantitative Analysis*, Vol. 29, pp. 31-56.

Hull, J. C. and A. White (1987), 'Hedging the Risks from Writing Foreign Currency Options', *Journal of International Money and Finance*, Vol. 6, pp. 131-152.

Hull, J. C. and A. White (1988), 'An Analysis of the Bias in Option Pricing Caused by Stochastic Volatility', *Advances in Futures and Options Research*, Vol. 3, pp. 29-61.

Jones, C., (2003), 'The dynamics of stochastic volatility: evidence from underlying and options markets', *Journal of Econometrics*, Vol. 116, pp. 181–224.

Kou, S. G. and H. Wang (2004), 'Option Pricing Under a Double Exponential Jump Diffusion Model', *Management Science*, Vol. 50, pp. 1178-1192.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin (1992), 'Testing the null hypothesis of stationarity against the alternative of unit root', *Journal of Econometrics*, Vol. 54, pp. 159-178.

Lee, D. and P. Schmidt (1996), 'On the power of the KPSS test of stationarity against fractionally-integrated alternatives', *Journal of Econometrics*, Vol. 73, pp. 285–302.

Lo, A.W. (1991), 'Long Term Memory in Stock Market Prices', Vol. *Econometrica*, Vol. 59, pp. 1279-1313.

Mills, T.C. and R. N. Markellos (2007), *The Econometric Modelling of Financial Time Series*, 3rd edition, Cambridge: Cambridge University Press (forthcoming).

Moraux, F., P. Navatte and C. Villa (1999), 'The Predictive Power of the French Market Volatility Index: A Multi Horizons Study', *European Finance Review*, Vol.2, pp. 303-320.

Nelson, D. B. (1990), 'ARCH models as diffusion approximations', *Journal of Econometrics,* Vol. 45, pp. 7–39.

Pan, J., (2002), 'The jump-risk premia implicit in options: evidence from an integrated time-series study', *Journal of Financial Economics*, Vol. 63, pp. 3-50.

Phillips P.C.B. and P. Perron (1988), 'Testing for a unit root in time series regression', *Biometrika*, Vol. 75, pp. 335–46.

Phillips, P. C. B., and J. Yu, (2005), 'Jackknifing bond option prices' *Review of Financial Studies*, Vol. 18, 707–742.

Robinson, P. M. (1995), 'Log-periodogram regression of time series with long-range dependence', *Annals of Statistics*, Vol. 23, pp. 1048—1072.

Scott, L. O. (1987), 'Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application', *Journal of Financial and Quantitative Analysis*, Vol. 22, pp. 419-438.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, London: Chapman & Hall.

Simon, D.P. (2003), 'The Nasdaq Volatility Index During and After the Bubble', *Journal of Derivatives*, Vol. 11, pp. 9-24.

Stein, E. and J. Stein (1991), 'Stock Price Distributions with Stochastic Volatility: An Analytic Approach', *Review of Financial Studies*, Vol. 4, pp. 727-752.

Wagner, N. and A. Szimayer (2004), 'Local and Spillover Shocks in Implied Market Volatility: Evidence for the U.S. and Germany', *Research in International Business and Finance*, Vol. 18, pp. 237-251.

Whaley, R.E. (2000), 'The Investor Fear Gauge', *Journal of Portfolio Management*, Vol. 26, pp. 12-17.

Whaley, R.E. (1993), 'Derivatives on Market Volatility: Hedging Tools Long Overdue', *Journal of Derivatives*, Vol. 1, pp. 71-84.

Wiggins, J. (1987) 'Option Values under Stochastic Volatility: Theory and Empirical Estimates', *Journal of Financial Economics*, Vol. 19, pp. 351-372.

Table 1. Descriptive Statistics of VIX Index and first differences (ΔVIX)

	VIX	ΔVIX
Mean	0.1957	-1.23E-05
Median	0.1856	-0.0004
Maximum	0.4574	0.0992
Minimum	0.0931	-0.0780
Std. Dev.	0.0639	0.0122
Skewness	0.9382	0.5647
Kurtosis	3.7411	9.1172
Jarque-Bera	671.14**	6,378.52**
$\rho(1)$	0.981	-0.041
$\rho(2)$	0.964	-0.088
$\rho(3)$	0.950	-0.057
$p^2(1)$	0.975	0.201
$\rho^2(1)$	0.950	0.189
$\rho^2(1)$	0.932	0.204

ρ(q) and p^2(q) are autocorrelation and squared autocorrelation coefficients at lag *q* , respectively. Two (one) stars denote significance at the 1% (5%) level.

Table 2. Unit Root Test results of VIX

Test	Null Hypothesis	Test Statistic
Augmented Dickey-Fuller	Unit Root	-3.7892**
Phillips-Perron	Unit Root	-4.9068**
Kwiatkowski-Phillips-Schmidt-Shin	Stationarity	1.4366**

The Augmented Dickey-Fuller (Dickey and Fuller, 1979) and the Phillips-Perron (1988) test the null hypothesis of a unit root. The Kwiatkowski-Phillips-Schmidt-Shin (1992) tests the null hypothesis of stationarity. An intercept is included in all test regressions. Two (one) stars denote significance at the 1% (5%) level.

Table 3. Log Likelihood of Alternative Unconditional Distribution Models

	Parameters	VIX	ΔVIX
Normal	2	5,266.8	11,811.8
Log-normal	2	5,734.4	
t-student	1	5,295.3	12,221.3
Skewed *t*-student	2	5,841.5	12,225.6
Logistic	2	5,285.3	12,123.1
Exponential	2	5,051.4	
Non-central Chi-squared[9]	2	5,637.4	
Extreme (max)	2	5,647.4	
Pareto	2	4,142.9	
Weibull	2	5,267.1	

[9] The non-central Chi-squared distribution $\chi^2(k,\lambda)$ with k degrees of freedom and λ non-centrality parameter is a special case of the Generalized Gamma distribution $G(\gamma,\beta,\lambda)$ for $\gamma = k / 2$ and $\beta = 2,$ where γ is the shape parameter, β is the scale parameter and λ is the non-centrality parameter.

Table 4. Conditional tabulation of VIX *vs*. ΔVIX

		VIX					
		[0, 0.1)	**[0.1, 0.2)**	**[0.2, 0.3)**	**[0.3, 0.4)**	**[0.4, 0.5)**	**Total**
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	1,053	703	130	19	1,908
	[0.05, 0.1)	0	0	4	6	2	12
	Total	5	2,298	1,369	256	29	3,957

Table 5. Parameter estimates of diffusion and jump diffusion processes over complete sample (1/2/1990 to 9/13/2005)

Parameter	SR	SRJ	SRPJ	MRLP	MRLPJ
k	4.5496 (5.9778)	7.3800 (9.5121)	10.5004 (11.1326)	3.9598 (5.4873)	4.4887 (6.6083)
θ	0.1945 (19.9557)	0.1505 (21.7557)	0.1379 (24.0473)	-1.6853 (-29.8494)	-2.1326 (-19.4989)
σ	0.4048 (88.0705)	0.3502 (61.3238)	0.3294 (51.3363)	0.8857 (88.2150)	0.7504 (50.3114)
λ	-	19,4080 (4.5046)	263.8877 (9.1391)	-	41.9585 (3.1030)
$1/\eta$	-	0.0170 (8.2228)	0.0125 (4.5626)	-	0.068 (6.7492)
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 6. 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)

	SR		SRJ		SRPJ		MRLP		MRLPJ	
Parameter	A	B	A	B	A	B	A	B	A	B
k	6.7809 (4.8646)	4.7561 (4.2697)	10.0246 (7.9152)	7.6445 (6.3798)	11.1900 (5.8252)	12.6767 (7.1709)	6.2675 (4.7258)	4.0656 (3.8960)	6.5474 (5.6811)	5.2050 (4.9772)
θ	0.1682 (20.0503)	0.2209 (14.8413)	0.1367 (22.6912)	0.1780 (14.0157)	0.1353 (20.9330)	0.1465 (12.2977)	-1.8188 (-34.8811)	-1.5585 (-20.5860)	-2.0848 (-28.1717)	-2.2330 (-7.6033)
σ	0.3843 (61.9047)	0.4254 (62.2626)	0.3167 (46.2252)	0.3840 (38.7842)	0.3122 (35.2948)	0.3446 (26.2871)	0.9143 (62.0300)	0.8589 (62.3677)	0.7402 (43.3197)	0.7444 (25.1828)
λ	-	-	18.7032 (3.5102)	21.8095 (2.3767)	149.9526 (1.9034)	478.5668 (2.5111)			26.1822 (3.1695)	132.5710 (1.6099)
$1/\eta$	-	-	0.0172 (5.7648)	0.01569 (5.0928)	0.0142 (3.5221)	0.0090 (5.7419)			0.072 (5.7153)	0.027 (4.2518)
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

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).

Figure 1. The VIX Index and first differences (ΔVIX)

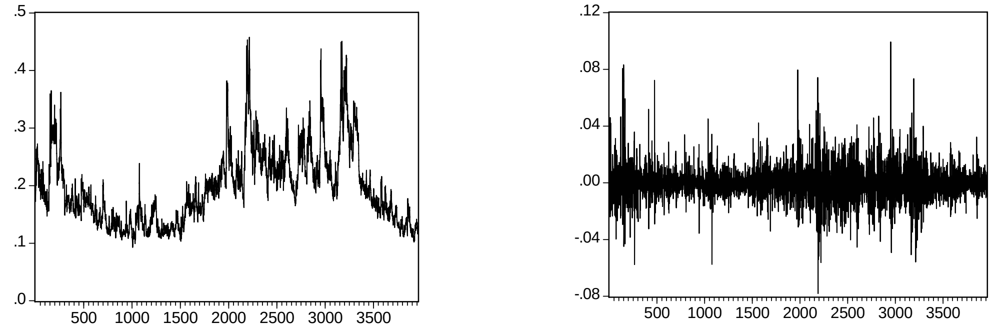
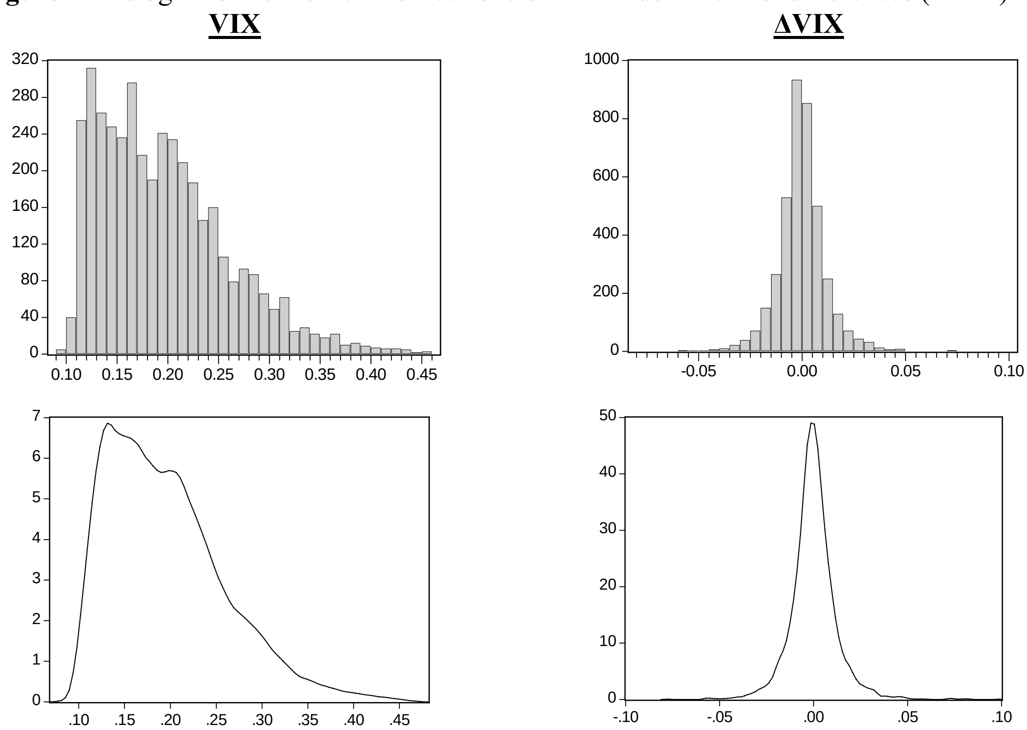
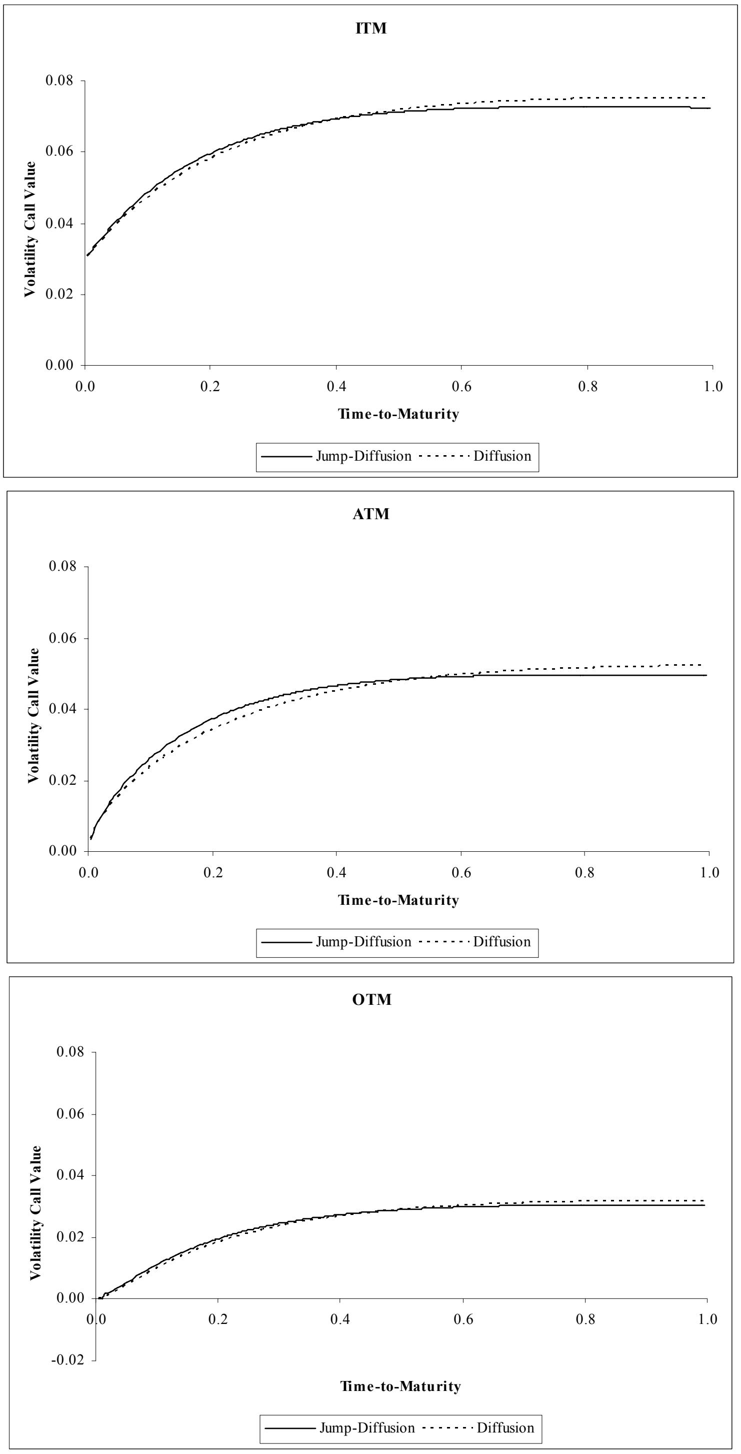


Figure 2. Histograms and Kernel Distributions of VIX Index and first differences (ΔVIX)



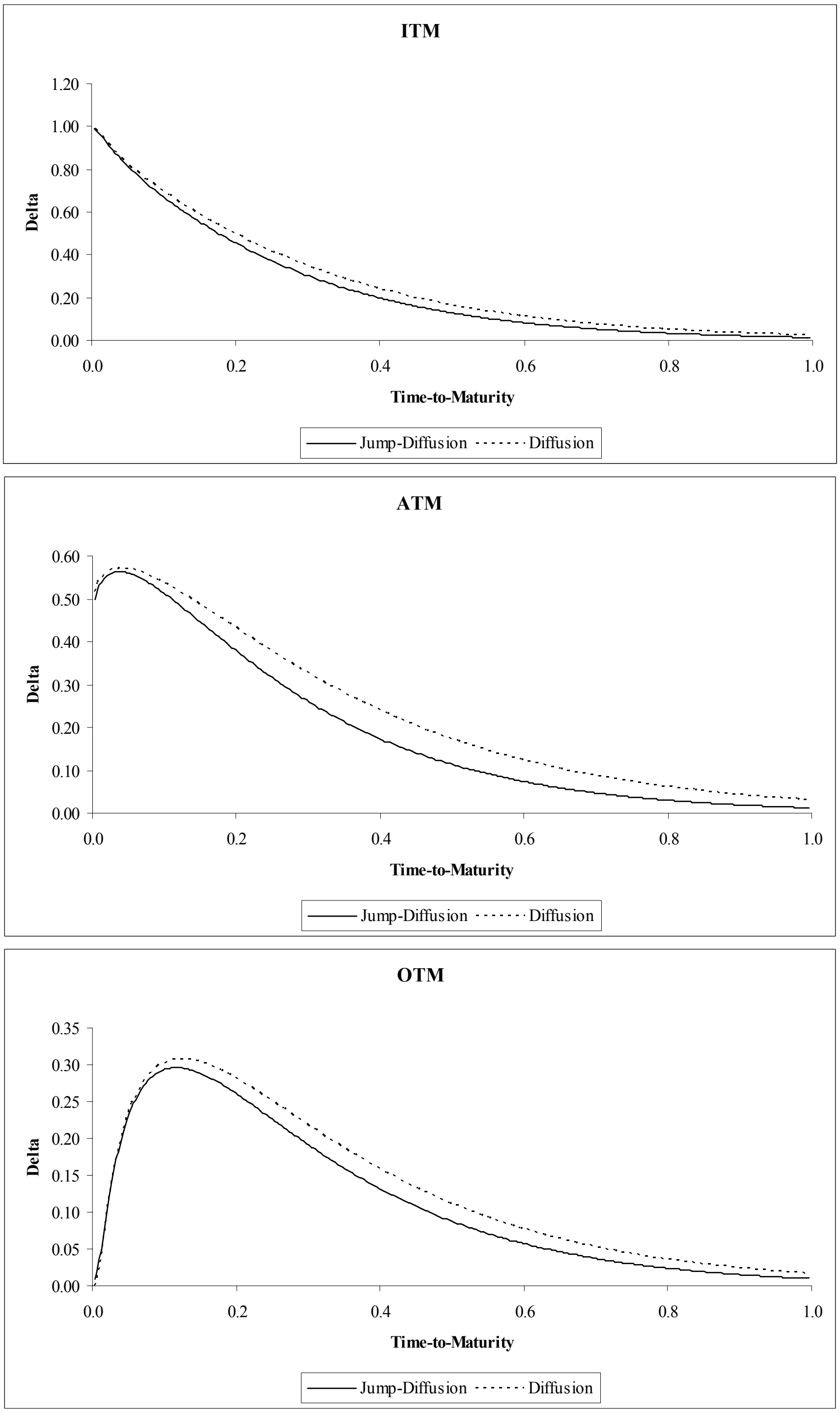
Densities were estimated with Epanechikov kernel functions over 100 points. The bandwidth was determined according to the method suggested by Silverman (1986).

Figure 3. Value of the volatility call option as a function of time-to-maturity



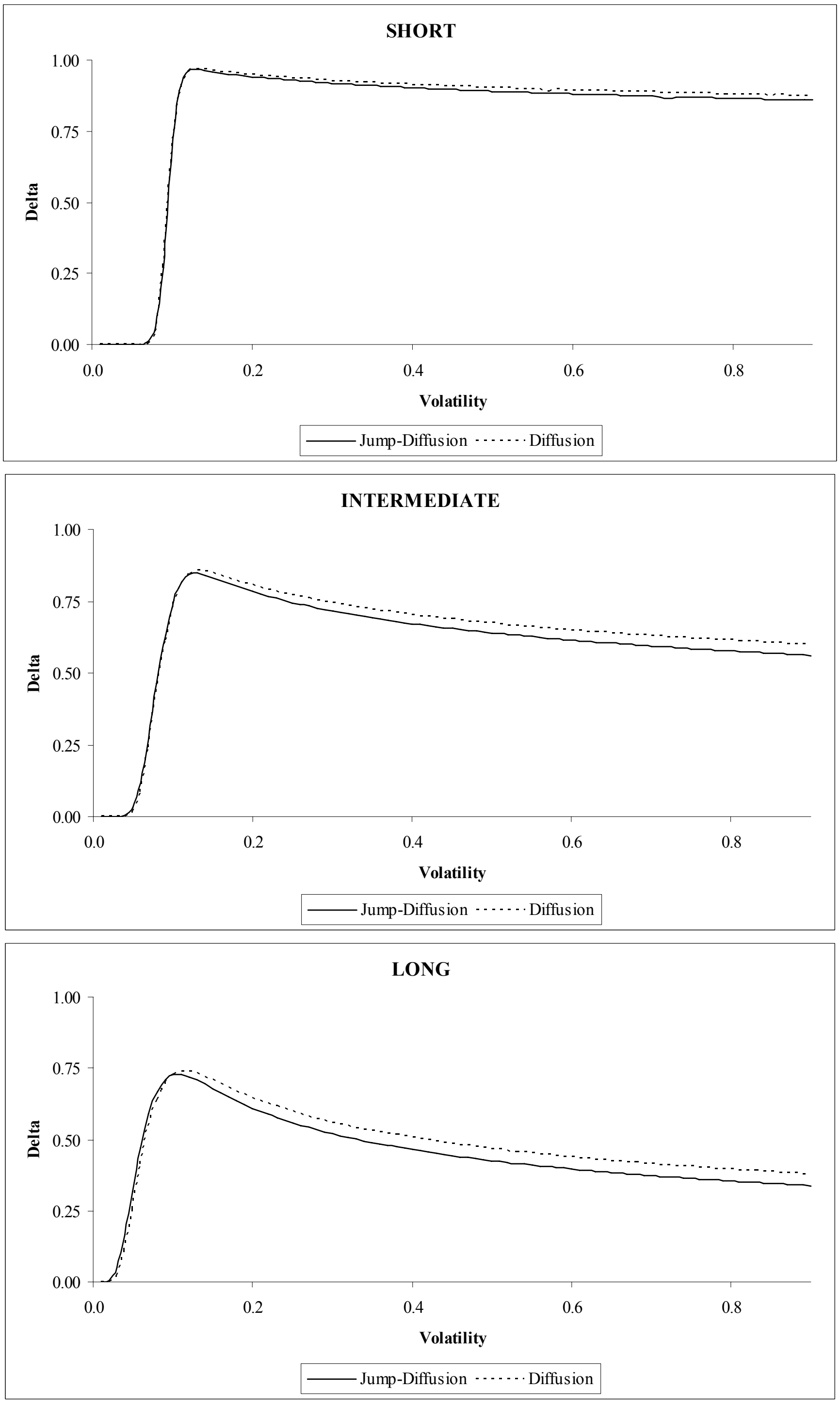
Estimated for three different moneyness levels: 20% in-the-money (ITM), at-the-money (ATM) and 20% out-of-the-money (OTM). The solid line corresponds to the case where there are no jumps in the volatility process (i.e., model of Detemple and Osakwe, 2000) using the estimated parameters k, θ, and σ from Table 5, fifth column. The dotted 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 5, sixth column. We assume that $r = 5\%$ and $V_t = 15\%$.

Figure 4. Delta of the volatility call option as a function of time-to-maturity



Estimated for three different moneyness levels: 20% in-the-money (ITM), at-the-money (ATM) and 20% out-of-the-money (OTM). The solid 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 5, fifth column. The dotted 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 5, sixth column. We assume that $r = 5\%$ and $V_t = 15\%$.

Figure 5. 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 solid 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 5, fifth column. The dotted 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 5, sixth column. We assume that $r = 5\%$ and $V_t = 15\%$.