Generalized parameter functions for option pricing

Generalized parameter functions for option pricing

Panayiotis C. Andreou[a,*], Chris Charalambous[b], Spiros H. Martzoukos[b]

[a] *Durham Business School, Durham University, Mill Hill Lane, DH1 3LB, Durham, UK*
[b] *Dept. of Public and Business Administration, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus*

This version: 24 August 2009

Abstract

We extend the benchmark nonlinear deterministic volatility regression functions of Dumas et al. (1998) to provide a semi-parametric method where an enhancement of the implied parameter values is used in the parametric option pricing models. Besides volatility, skewness and kurtosis of the asset return distribution can also be enhanced. Empirical results, using closing prices of the S&P 500 index call options (in one day ahead out-of-sample pricing tests), strongly support our method that compares favorably with a model that admits stochastic volatility and random jumps. Moreover, it is found to be superior in various robustness tests. Our semi-parametric approach is an effective remedy to the curse of dimensionality presented in nonparametric estimation and its main advantage is that it delivers theoretically consistent option prices and hedging parameters. The economic significance of the approach is tested in terms of hedging, where the evaluation and estimation loss functions are aligned.

JEL classification: G13; G14

Keywords: Option pricing; Implied volatilities; Deterministic volatility functions; Delta-hedging; Semi-parametric approach

***** Corresponding author. Tel.: +44 (0)191 33 45 528; fax: +44 (0) 191 33 45201.
E-mail addresses: benz@pandreou.com (P.C. Andreou), bachris@ucy.ac.cy (C. Charalambous), baspiros@ucy.ac.cy (S.H. Martzoukos).

1. Introduction

In this paper, we extend the benchmark nonlinear Deterministic Volatility regression Functions (DVF) of Dumas et al. (1998) to provide a nonparametric enhancement of the implied parameter values used in Parametric Option Pricing Models (POPMs), such as the Black and Scholes (1973) (BS) and the Corrado and Su (1996) (CS). The method we propose enables the estimation of Generalized Parameter Functions (GPF) and allows the enhancement of parameters other than volatility (i.e., skewness and kurtosis). The resulting semi-parametric models, refer to as the enhanced Parametric Option Pricing Models (ePOPMs), outperform the counterpart DVF-based parametric ones by a large margin (in respect to out-of-sample pricing of S&P 500 index call options).

Systematic biases have been frequently documented (e.g., Rubinstein, 1994, Bakshi et al., 1997, Bates, 1996, Jackwerth, 2000, Derman, 1999) whereas the BS implied volatilities tend to vary across moneyness and times to maturity. The S&P 500 option-implied volatilities form a *smile* pattern prior to the October 1987 market crash, where out-the-money and in-the-money implied volatilities are higher than at-the-money ones. After the crash, a *smirk* pattern is most prominent, where implied volatilities decrease monotonically as the exercise price rises relative to the index level (a stylized fact that is apparent in our data as well). This is primarily due to investors' fear of large adverse price jumps (see Rubinstein, 1994, Derman, 1999, Pan, 2002). Implied volatilities also present a term structure pattern. However, regardless of the exact pattern, the term *volatility smile* is widely used to describe this anomaly. Recent POPMs that incorporate Stochastic Volatility (SV) or Stochastic Volatility and Jump (SVJ) risk factors (e.g., Bakshi et al., 1997, Bates, 1996 and 2000) mitigate[1] much of the bias associated with the original BS model. The former models result in better out-of-

[1] A similar effect is achieved indirectly with the Corrado and Su (1996) model (an expanded BS formula that incorporates option price adjustments for skewness and kurtosis), which is an important alternative to SV and SVJ, due to its simplicity when employed on a daily basis.

sample pricing performance compared with the simple BS model, with SVJ being superior to SV. As Bakshi et al. (1997) point out, though, both models are misspecified since, following the transformation from a risk-neutral to a real-world measure, their implied parameters are statistically inconsistent with those implicit in the time-series of S&P 500 returns. A discrepancy between risk-adjusted option implied parameters and the estimated parameters from the time-series can be attributed not only to a misspecification of the stochastic process of the asset returns, but also to micro-structure variables related to the characteristics of the option markets. Examples include liquidity proxied by options volume, transaction costs proxied by bid-ask spreads, and excess demand for out-the-money options (e.g., Bollen and Whaley, 2004, and Chang et al., 2009). According to Dumas et al. (1998) and Hull and Suo (2002), it is rather troublesome to apply SV and SVJ models on a daily basis (see also Crépey, 2004). On the other hand, BS has shown severe time endurance and is still widely used by practitioners since it generates reasonable prices for a wide spectrum of European options.

Dumas et al. (1998) estimate regression specifications of quadratic forms, which provide unique per contract volatility estimates and examine how well they predict option prices. They find that this interpolative regression approach produces more accurate option prices, compared with option prices obtained by solving a forward partial differential equation using the Crank-Nicholson finite-difference method (see also Derman, 1999). As discussed in Christoffersen and Jacobs (2004), Berkowitz (2009) and Christoffersen et al. (2009), although this model cannot be considered a replacement of structural POPMs, it is used as a benchmark in the existing literature because of its simplicity when employed on a daily basis. Dumas et al. (1998) conclude by suggesting that the DVF approach should be extended and generalized. Our method extends these benchmark regression-based models, whilst also retaining the intuition in Christoffersen and Jacobs (2004) according to which during implementing option valuation models, it is critical to align the estimation and evaluation loss functions.

Researchers have also considered the use of nonparametric techniques such as artificial neural networks that can be used for nonlinear regression. The main advantage of such methods is that they rely on fairly simple algorithms and that the underlying nonlinearity can be modeled from transactions data (see Duda et al., 2001). Standard applications of artificial neural networks do not involve any financial theory and can be used to estimate directly the unknown empirical option pricing function (hereafter called the standard neural network approach). Hutchinson et al. (1994) conclude that, although they do not constitute a substitute for the more traditional arbitrage-based pricing formulas, such methods may be more accurate and computationally more efficient alternatives than parametric models when the underlying asset's price dynamics are unknown, and when the pricing equation cannot be solved analytically. Evidence concerning the out-of-sample pricing performance of standard artificial neural networks is sometimes mixed (see discussions in Andreou et al., 2008), and their application for pricing options also has its limitations. First of all, standard neural networks are commonly interpreted as "black boxes" since they "learn" the empirical functions inductively from transactions data, without embedding any information related to the problem under scrutiny. Secondly, they are sensitive to time-varying input variables. This problem is exaggerated in option pricing, since key variables, such as volatility, can be very erratic (see Aït-Sahalia and Lo, 1998). Finally, the use of standard neural networks can deliver option prices that violate fundamental financial principles (e.g., irrational option values or hedging parameters).

The main contribution of this paper is the development of a nonparametric method to enhance the parameter values used in the POPMs, thus generalizing Dumas et al. (1998) regression-based DVF models. Our method is motivated by Christoffersen and Jacobs

(2004), according to whom the nonlinear regression-based DVF models represent a tough option pricing benchmark by which the performance of structural models can be measured. The GPF that we derive in this paper provide an enhancement of parameters beyond volatility, without specifying a-priori a deterministic parametric functional form, as in the DVF case. Specifically, the parameter enhancement provides the volatility to the BS and CS models. In addition, skewness, or skewness and kurtosis, can be enhanced for the CS model. The proposed semi-parametric method has the following advantages over the nonparametric ones. First, it retains the theoretical properties[2] of the parametric model being enhanced, in particular: *i*) non-negative option values (thus expecting satisfactory pricing performance at the boundary of option pricing areas, in both dense and sparse input areas), *ii*) the desire to conform to certain shape restrictions (see Aït-Sahalia and Duarte, 2003) as implied by the theory regarding the slope, convexity and monotonicity of the estimated option pricing function without the need to explicitly incorporate nonlinear constraints, *iii*) theoretically consistent hedging parameters (i.e., Greek letters), *iv*) non-negative implied state price densities, and *v*) no additional assumptions about investors' preferences for risk. Second, Bandler (1999), Aït-Sahalia and Lo (1998) and Aït-Sahalia and Duarte (2003) recognize that semi-parametric models can achieve the same degree of accuracy as the full nonparametric estimator with fewer regressors, fewer free parameters and considerably reduced sample size, thus being a remedy to the curse of dimensionality. Third, another advantage of our method compared with the Dumas et al. (1998) approach (see also discussions in Christoffersen and Jacobs, 2004) is that one can use long-term estimation (i.e., twelve months) of the GPF. At

[2] In the case of the CS model, in some extreme regions of skewness and kurtosis, the model may give negative option values and/or negative sensitivity of the call to the underlying asset (see discussions in Jondeau and Rockinger, 2001, as well as the models in Corrado, 2008, and León et al., 2009). In our sample, this is extremely rare (only in 0.1% of the sample are the negative values marginally below zero). Furthermore, in the case of data where such violations occur frequently, in our methodology we could easily constrain skewness and kurtosis to prohibit such inconsistencies.

the same time, we also capture the time-variation of the option valuation relationship by using daily implied parameters as inputs in the estimation of the GPF. This is in the same line of thought as Christoffersen et al. (2009) where the use of long-term data for the estimation of most parameters is complemented with the frequent re-estimation of implied spot volatilities. A final advantage of our semi-parametric option pricing model is the simplicity of the estimation of the GPF making it a computationally efficient task (calibrating the SVJ model requires at least three times more computational time for our testing period).

Another contribution of our study is to apply for the first time the regression-based DVF approach to the CS model. The SVJ model of Bates (1996) is included in the comparisons since it is an effective remedy to the BS biases (see also Bakshi et al., 1997). At the same time results for the SV sub-model are reported. In-sample, SVJ has the best fit while SV is inferior to the best DVF models. Out-of-sample, we find that extending the DVF approach to the case of CS can significantly improve the model's pricing performance, although SVJ dominates in the family of parametric models. We develop ePOPMs for both the BS and the CS models and we find that they outperform their DVF parametric counterparts in pricing S&P 500 index call options. The enhanced CS model is the overall best ePOPM and is competitive with the SVJ model in pricing performance. As one can see from the results of this study, the ePOPMs exhibit superior robustness compared with the SVJ model, in terms of pricing options: *i)* in moneyness regions that are not used in the estimation, and *ii)* in weekly instead of daily out-of-sample pricing tests. When models are optimized with a pricing criterion, the hedging performance of all models is rather poor and in line with the previous literature. Finally our study makes a contribution, through the estimation of the GPF for hedging, where optimization is based on a hedging criterion with significantly improved results, thus verifying the intuition in Christoffersen and Jacobs (2004) that the estimation and the evaluation loss functions should be aligned.

In the following sections, we review the parametric models and explain the implementation of the ePOPM structure via the GPF. This is followed by a discussion of the data, the filtering and the alternative versions of the models under comparison. We then discuss the pricing results, provide various pricing robustness checks and implement a single instrument delta-hedging strategy for the best models. Finally, we conclude. A comprehensive Appendix, which includes several results not shown in the main text due to space limitations, is available upon request.

2. The parametric models used

The Black and Scholes (1973) formula is the first model that we consider. The dividend adjusted BS formula for the European call option is:

$$c^{BS} = Se^{-d_y T} N(d) - Xe^{-rT} N(d - \sigma\sqrt{T}), \quad (1)$$

$$d = \frac{ln(S / X) + (r - d_y)T + (\sigma\sqrt{T})^2 / 2}{\sigma\sqrt{T}}, \quad (1.a)$$

where c^{BS} is the price of the European call option; S is the spot price of the underlying asset; X is the exercise price of the call option; r is the continuously compounded risk free interest rate; d_y is the continuous dividend yield paid by the underlying asset; T is the time left until the option expiration date; σ^2 is the yearly variance of the rate of return for the underlying asset and $N(.)$ stands for the standard normal cumulative distribution.

The abundant empirical evidence regarding the smirk behaviour of the BS implied volatility is indicative of implied return distributions that are negatively skewed and heavier tailed than those implied by the BS log-normal distribution (see Bakshi et al., 1997 and Bates, 2000). For this reason, more general option pricing models are included in the parametric analysis. The Corrado and Su (1996) model, which is an extension of the BS model and can capture skewness and kurtosis of the underlying returns' distribution, is used. Specifically, Corrado and Su employ the Gram-Charlier series expansion[3] (Type A) of a normal density function, which is similar to the Taylor series expansions for analytic functions, and truncate the infinite series by excluding terms beyond the fourth moment. The resulting option pricing model is an expanded Black-Scholes formula that incorporates option price adjustments for skewness and kurtosis, as given below (see the correction in Brown and Robinson, 2002; see also Jondeau and Rockinger, 2001, Jurczenko et al., 2004, and León et al., 2009):

$$c^{CS} = c^{BS} + \mu_3 Q_3 + (\mu_4 - 3) Q_4, \tag{2}$$

where c^{BS} is the BS value for the European call option given in Eq. (1) and,

$$Q_3 = \frac{1}{3!} S e^{-d_y T} \sigma\sqrt{T}((2\sigma\sqrt{T} - d)n(d) + (\sigma\sqrt{T})^2 N(d)), \tag{2.a}$$

$$Q_4 = \frac{1}{4!} S e^{-d_y T} \sigma\sqrt{T}((d^2 - 1 - 3\sigma\sqrt{T}(d - \sigma\sqrt{T}))n(d) + (\sigma\sqrt{T})^3 N(d)), \tag{2.b}$$

where d is given by Eq. (1.a). In Eq. (2), Q_3 and Q_4 represent the marginal effect of non-normal skewness and kurtosis in the option price respectively, whereas μ_3 and μ_4 correspond to coefficients of skewness and kurtosis of the asset returns density function. In the above expressions, $n(.)$ is the standard normal probability density function. Option pricing based on the Gram-Charlier expansions has become popular because it satisfies crucial properties such as modelling flexibility and analytical tractability (see León et al., 2009 and references therein).

Motivated by the existing empirical evidence (e.g., Bakshi et al., 1997, Das and Sundaram, 1999), and unlike Christoffersen and Jacobs (2004) that concentrate on the SV model only, we consider the SVJ model of Bates (1996) as our benchmark (whilst also

[3] Backus et al. (2004) show that Gram-Charlier approximations can do a reasonable job in approximating generally used option pricing models like for instance the SVJ of Bates (1996) with reasonable (i.e., commonly estimated in the literature) parameter values (see also Jurczenko et al., 2004). In this paper, we find that the empirical pricing performance of such approximation can be further improved using the generalized parameter functions derived by our approach.

reporting results for SV). We must note that SV and SVJ models are not widely used by traders for pricing options (e.g., Hull and Suo, 2002) since they usually prefer more intuitive methodologies that are closer to the regression-based DVF approach (see also Brandt and Wu, 2002, and Crépey, 2004). Under a real-world measure, the instantaneous conditional variance, V_t, of the SVJ model follows a mean-reverting square root process:

$$\frac{dS}{S} = (\mu - \lambda\bar{\kappa})dt + \sqrt{V}dZ + \kappa dq, \tag{3}$$

$$dV = (\alpha - \beta V)dt + \sigma_v\sqrt{V}dZ_v, \tag{4}$$

with

$$cov(dZ, dZ_v) = \rho dt, \quad prob(dq = 1) = \lambda dt,$$

$$ln(1+\kappa) \sim N(ln(1+\bar{\kappa}) - 0.5\theta^2, \theta^2).$$

Here, μ is the instantaneous drift of the underlying asset; λ is the annual frequency of jumps; κ is the random percentage jump conditional on a jump occurring; q is a Poisson counting process with intensity λ; θ^2 is the jump variance; σ_v is the variation coefficient (volatility of volatility), and ρ is the correlation coefficient between the volatility shocks and the underlying asset movements. Moreover, β is the rate of mean reversion and α / β is the variance steady-state level (i.e., long-run mean). In the above diffusion specification, the correlation between the volatility and the returns of the underlying asset controls the level of skewness, whilst the variability of volatility allows for heavy-taildedness. Stochastic volatility models are not capable of generating high levels of skewness and kurtosis at short maturities under reasonable parameterizations and, as a consequence, they cannot generate implied volatility smiles as sharp as those typically observed (see Das and Sundaram, 1999, Pan, 2002). The addition of a jump component further enhances the distributional flexibility of the model, since it can internalize additional levels of negative skewness and excess kurtosis. This allows the model to better explain the heavy-tailedness of the asset return distribution and to generate realistic implied volatility smile patterns, especially at short maturities. The value of the European call option is given as a function of state variables and parameters, after transforming the true process given by Eqs. (3) and (4) into a risk-neutral one that incorporates the appropriate compensation for volatility risk and jump risk:

$$c^{SVJ} = e^{-rT}[E^*(S_T)\Pi_1 - X\Pi_2], \tag{5}$$

where $E^*(S_T)$ is the expectation with respect to a risk-neutral probability measure, and $E^*(S_T) = Se^{(r-d_y)T}$ is the forward price on the underlying asset. The risk-neutral probabilities Π_1 and Π_2 can be computed via the Fourier inversion of the complex-valued characteristic functions, associated with the moment generating functions of $ln(S_T / S)$ (see Bates, 1996, pg. 73-77 for more details). By constraining the jump component values equal to zero we get the European call prices for the SV model. We fit the POPMs to obtain daily the implied parameters that minimize the sum of squared pricing deviations from daily market prices, so these (risk-neutral) parameters indirectly account for any pricing of jump and volatility risks.

2.1. The deterministic volatility functions for BS and CS

The DVF specifications below mitigate the volatility smile anomaly, by allowing the implied volatilities to be a deterministic function of the option's strike price, X, and maturity, T. According to Bedendo and Hodges (2009), much attention has been devoted to understanding and modeling the dynamics of the implied volatility functions, which is a crucial task for both trading, pricing and risk-management of option positions[4]. For our

[4] For a discussion of studies that investigate the dynamics and predictability of implied volatility indices see Konstantinidi et al. (2008) and Becker et al. (2009).

analysis, we estimate the following three regression-based DVF specifications, as proposed[5] in Dumas et al. (1998):

$$\text{DVF\#1: } \sigma = max(0.01, a_0 + a_1 X + a_2 X^2), \quad (6)$$

$$\text{DVF\#2: } \sigma = max(0.01, a_0 + a_1 X + a_2 X^2 + a_3 T + a_4 XT), \quad (7)$$

$$\text{DVF\#3: } \sigma = max(0.01, a_0 + a_1 X + a_2 X^2 + a_3 T + a_4 XT + a_5 T^2). \quad (8)$$

DVF #1 attempts to capture the curvature of the implied volatility smile, using a quadratic function of the strike price. DVF #2 is quadratic in strike price and linear in maturity, which can explain the smile and the term effect. Motivated by the empirical presence of curvature (even humps) in the term structure of implied volatility[6], we also estimate DVF #3, which is a variation of DVF #2 and is also quadratic in time. The cross product XT in Eqs. (7) and (8) is important, since it allows the DVF slope with respect to strike to be time dependent and the slope with respect to time to depend on the strike level, thus capturing changes in the shape of the volatility smile over different maturities. Finally, following Dumas et al. (1998) a minimum value of 1% is imposed when estimating the regression functions to prevent negative values of volatility. The DVF parameters can be estimated using either Ordinary Least Squares (OLS) - where the loss function is the difference between the estimated and the actual contract-specific implied volatility - or Nonlinear Least Squares - where the loss function is the difference between the theoretical/model-based and the actual/market-based option prices. This methodological framework is conceptually similar to the one developed with the Space Mapping techniques in Bandler et al. (1999) where several parameter values

---

[5] Christoffersen and Jacobs (2004) limit their attention to DVF#3, which is the most general model (see also Christoffersen et al., 2009). We implement these models in the same way as with these authors but for completeness, we also estimate DVF#1 and DVF#2 as proposed in Dumas et al. (1998).

[6] Empirical evidence of such curvature in the time structure of implied volatility can be found in Bakshi et al. (1997), Christoffersen et al. (2009), and in Table 1 of our paper.

of an imperfect model are adjusted to make the imperfect model (for example a simple POPM like the BS) approximate the performance of a finer but more expensive or inaccessible one (for example, the market option prices). Aït-Sahalia and Lo (1998) examine a semi-parametric method, where they use the BS volatility loss function, but estimation is done through a nonparametric kernel regression, instead of OLS. More recently, Bedendo and Hodges (2009) propose a model based on a discrete and linear Kalman filter updating of the volatility skew, which involves numerical maximization of the relevant log-likelihood function. Christoffersen and Jacobs (2004) demonstrate that OLS estimates of the DVF parameters yield biased option pricing results and that a Nonlinear Least Squares pricing loss function should be used instead. In this paper, we implement the regression-based DVF approach not only for the BS but also - for the first time - for the CS model, using both loss functions (see section 4). The method that we propose is developed in section 3 and should be compared with the DVF-based BS and CS alternatives. For completeness, though, we also provide the results for SV and SVJ.

3. The ePOPM structure

In order to estimate the GPF nonparametrically, we employ feedforward artificial neural networks, which are universal function approximators. Cybenko (1989) has shown theoretically that these networks can approximate an unknown continuous function arbitrarily well. An artificial neural network is a collection of interconnected simple processing elements structured in successive layers and can be depicted as a network of links (*synapses*) and nodes (*neurons*) between layers (see also Andreou et al., 2008). A typical feedforward neural network has an input layer, one or more hidden layers and an output layer. Each interconnection corresponds to a modifiable parameter (weight or bias), which is adjusted according to the given problem via optimization (the training algorithm). Figure 1 depicts the general idea of the ePOPM structure we propose. For our analysis, inputs are set up in feature vectors, $\tilde{x}_p = [x_{1p}, x_{2p} ..., x_{Np}]$, for which there is an associated (known) target characterizing the problem, t_p, $p = 1,2,...,P$, where P is the number of the available sample feature vectors for a particular estimation sample and N is the number of input variables. To estimate the underlying relationships, the free parameters are adjusted in order to minimize the loss function of the error between the network output and the desired target values.

The network structure we propose has four layers. The first three are typical layers of a feedforward artificial neural network: an *input* layer with N variables, a *hidden* layer with H neurons, and an *output* layer with M neurons. For these layers, each node is connected with all neurons in the previous and the forward layer. Each connection is associated with a *weight*, $w_{in}^{(1)}$, and a *bias*, $w_{i0}^{(1)}$, in the input layer (*i=1,2,…,H, n=1,2,…N*) and a *weight*, $w_{ji}^{(2)}$, and a *bias*, $w_{j0}^{(2)}$, in the hidden layer (*j=1,2,…,M*). Each neuron behaves as a summing vessel that computes the weighted sum of its inputs to form a scalar term. With the use of some transfer/activation function (such as the log-sigmoid or the hyperbolic tangent sigmoid), it eventually works as a nonlinear mapping junction for the forward layer.

The fourth layer - which hereafter will be termed as the *enhanced layer* - allows a certain POPM to be part of the structure of the network. In this setting, the network structure embeds knowledge from the parametric model during estimation (thus resulting in a semi-parametric option pricing method). If we let $X_I = X_E \cup X_D$ denote the set of all input parameters that are necessary for the parametric model, then $X_E \subseteq X_I$ corresponds to the *enhanced* parameters provided nonparametrically via the GPF and $X_D \subset X_I$ corresponds to those that are passed to the parametric model directly (see Figure 1). The enhanced parameter set X_E is basically the researcher's choice and effectively manifests the number of neurons at the output layer and the type of activation function to be used at the enhanced layer. In addition, X_S represents the set of inputs to the nonparametric model. The estimated output, y, is given by:

$$y = f_{PM}(v, X_D), \qquad (9)$$

where

$$v = [f_{d_1}(d_1), f_{d_2}(d_2), ..., f_{d_M}(d_M)], \qquad (10)$$

with $f_{PM}(.,.)$ referring to the functional form of the parametric option pricing model. Moreover, $f_{d_j}(.)$ are smooth monotonically increasing activation functions and d_j are simply the descaled values of $y_j^{(2)}$ (as given by Eq. (11) below), where $j = 1,2,...,M$. We use the term *enhanced* for the v_j variables that are determined by the GPF mechanism. In our application, $v \equiv X_E$.

[Figure 1, here]

The elements of vector $[y_1^{(2)} \dots y_M^{(2)}]^T$ follow the functional form of a typical three-layer network[7]:

$$\begin{aligned} y_1^{(2)} &= f_M[w_{10}^{(2)} y_0 + \sum_{i=1}^{H} w_{1i}^{(2)} f_H(w_{i0}^{(1)} x_{s0} + \sum_{n=1}^{N} w_{in}^{(1)} x_{sn})] \\ \vdots \quad & \vdots \qquad\qquad\qquad \vdots \\ y_M^{(2)} &= f_M[w_{M0}^{(2)} y_0 + \sum_{i=1}^{H} w_{Mi}^{(2)} f_H(w_{i0}^{(1)} x_{s0} + \sum_{n=1}^{N} w_{in}^{(1)} x_{sn})] \end{aligned} , \qquad (11)$$

where $f_M(.)$ and $f_H(.)$ are smooth monotonically increasing activation functions associated with the output and the hidden layer respectively and x_{sn} (represented by X_S in Figure 1),

[7] It is also frequently termed as *one-hidden layer* network. Although additional hidden layer(s) can be used, Cybenko (1989) supports theoretically that the chosen structure that we employ in this paper can approximate any unknown continuous nonlinear function arbitrarily well. In the case of a discontinuous option pricing function, like for example before and after anticipated events, alternative network structures that involve more than one hidden layers may be necessary. This seems to be more relevant for stock options.

n=1,2,…,N, is just the scaled value of the input variables x_n. The left-hand side elements of Eq. (11) represent the enhanced parameters that are estimated simultaneously, using information propagated by the POPMs. The vector defined by the right-hand side of Eq. (11) is the GPF which produces the enhanced variables.

The estimation of the network model is formulated as a nonlinear optimization process, in which the network's free parameters are determined according to a loss function. The loss function (or the discrepancy between the estimated response, y_p, and the actual response[8], t_p) is defined as:

$$e_p(w) = y_p(w) - t_p, \qquad (12)$$

where w is an ξ-dimensional column vector with the weights and biases given by:

$$w = [w_{10}^{(1)},...,w_{1N}^{(1)},...,w_{H0}^{(1)},...,w_{HN}^{(1)},\ w_{10}^{(2)},...,w_{1H}^{(2)},...,w_{M0}^{(2)},...,w_{MH}^{(2)}]^T .$$

The backpropagation algorithm, which is based on the gradient descent vector is the most popular method for estimating feedforward artificial neural networks. This algorithm, though, is often unable to converge rapidly to the optimal solution. For this reason, in this paper we rely on the Levenberg-Marquardt algorithm, which is much more efficient in terms of computational time and convergence rate. Christoffersen and Jacobs (2004) have concluded that Nonlinear Least Squares estimation should be employed to avoid bias in option prices. In order to accommodate this, the free parameters of our network are updated to minimize the following sum of squared (pricing) error loss function[9]:

$$F(w) = \sum_{p=1}^{P} e_p^2(w) \equiv \sum_{p=1}^{P} (y_p - t_p)^2 . \qquad (13)$$

---

[8] For the application in this study, t_p refers to market prices of call options standardized by the strike price.

[9] The use of Sum of Squared Errors (SSE) is common in empirical option pricing studies and is also supported by Christoffersen and Jacobs (2004).

Then, at each iteration τ of the algorithm, the parameter vector w is updated as follows:

$$w_{\tau+1} = w_{\tau} + \left[J^T(w_{\tau})J(w_{\tau}) + l_{\tau}I\right]^{-1} J^T(w_{\tau})e(w_{\tau}), \qquad (14)$$

where $J(w_{\tau})$ is the $P \times \xi$ Jacobian matrix of the P*-dimensional* output error column vector at τ^{th} iteration and is given by:

$$J^T(w) = \left[\nabla e_1^T(w), ..., \nabla e_P^T(w)\right]. \qquad (15)$$

In Eq. (14), I is a $\xi \times \xi$ identity matrix and l_{τ} is a learning parameter. Large values of l_{τ} diminish the significance of the Hessian matrix, approximated by $J^T(w_{\tau})J(w_{\tau})$, thus leading to directions that approach the steepest descent, as dictated by the term $J^T(w_{\tau})e(w_{\tau})$. Small values of l_{τ} lead to directions that approach the Gauss-Newton algorithm, by allowing the influence of $J^T(w_{\tau})J(w_{\tau})$. Further technical details about the implementation of the Levenberg-Marquardt algorithm can be found in Hagan et al. (1996). Based on Eq. (14), the update of the free parameters takes place in a batch mode only after all input vectors have been presented to the network.

The quantity $\nabla e_p(w)$ is the gradient vector of $e_p(w)$ with respect to the optimized parameter vector w. The partial derivative of the error function in Eq. (13) with respect to the weight $w_{ji}^{(2)}$ at the hidden layer is:

$$\frac{\partial e_p}{\partial w_{ji}^{(2)}} = \delta_j^{(2)} y_i^{(1)}, \qquad (16)$$

where $y_i^{(1)}$ is the output of the i^{th} transfer function in the hidden layer and

$$\delta_j^{(2)} = \frac{\partial f_{PM}}{\partial v_j} f'_{d_j}(d_j) s_j f'_M(\psi_j^{(2)}), \qquad (17)$$

with $f'_M(\psi_j^{(2)})$ and $f'_{d_j}(d_j)$ the differentials at points $\psi_j^{(2)}$ and d_j respectively. Moreover, $\psi_j^{(2)}$ represents the input signal at the j^{th} transfer function in the output layer, that is,

$\psi_j^{(2)} = w_{j0}^{(2)} y_0 + \sum_{i=1}^{H} w_{ji}^{(2)} f_H(w_{i0}^{(1)} x_{s0} + \sum_{n=1}^{N} w_{in}^{(1)} x_{sn})$. Finally, s_j is the standard deviation of the enhanced variable, given that a *z*-score scaling has been applied to $y_j^{(2)}$ (as given by Eq. (11) above) at the enhanced layer.

The quantity $\frac{\partial f_{PM}}{\partial v_j}$ is the partial derivative of the parametric model that is used with respect to input v_j of Eq. (10), thus creating a semi-parametric method dedicated to pricing European call options. This quantity is very important during the estimation, because it incorporates the theoretical properties of the parametric model used in the enhanced layer. All partial derivatives (i.e., the Greek letters) necessary for the implementation of the models are provided in the Appendix. The partial derivative of the error function in Eq. (13) with respect to the weight $w_{in}^{(1)}$ at the input layer is:

$$\frac{\partial e_p}{\partial w_{in}^{(1)}} = \delta_i^{(1)} x_{sn}, \tag{18}$$

where

$$\delta_i^{(1)} = \varepsilon_i^{(1)} f'_H(\psi_i^{(1)}), \qquad \varepsilon_i^{(1)} = \sum_{j=1}^{M} w_{ji}^{(2)} \delta_j^{(2)}, \tag{19}$$

and x_{sn} is simply the *z*-score scaled value of x_n. In Eq. (19), $\psi_i^{(1)}$ represents the input signal at the i^{th} transfer function in the hidden layer, that is, $\psi_i^{(1)} = w_{i0}^{(1)} x_{s0} + \sum_{n=1}^{N} w_{in}^{(1)} x_{sn}$. The proposed network structure can accommodate a scaling scheme for both the inputs and the enhanced variables. This can be essential since it increases the effectiveness of the optimization algorithm and minimizes the significance of differing dimensions of the input/output signals (see Hagan et al., 1996 and Duda et al., 2001). In the current study we choose to apply a standard *z-score* scaling only for the input variables.

In our case, the smooth monotonically increasing activation functions are among the hyperbolic tangent sigmoid, $f(\eta)=\gamma\left[\dfrac{e^{b\eta}-e^{-b\eta}}{e^{b\eta}+e^{-b\eta}}\right]$, the logistic $f(\eta)=\dfrac{\gamma}{1+e^{-b\eta}}$ or the linear one, $f(\eta)=\eta$. In the aforementioned expressions, with $\gamma, b \in \Re$, γ controls the output range and b the slope of the activation function. Following the implications of the universal approximation theorem of Cybenko (1989), we always use the standard hyperbolic tangent sigmoid activation function for $f_H(.)$ in the hidden layer (γ and b equal unity), while we use a linear activation function for $f_M(.)$ in the output layer. The type of the activation function at the enhanced layer is dictated by the parametric model and the enhanced variable(s) chosen; therefore, it is possible for $f_{d_1}(.)$, $f_{d_2}(.),\ldots, f_{d_M}(.)$ to be different, depending on the case considered. These activation functions are necessary during the implementation of the method, in order to ensure that the value of each of the enhanced variables is within an acceptable range for use with the parametric model[10]. Specifically, in the cases of BS and CS, volatility is kept lower than 70% and, in the case of CS, skewness is confined in the [-15, 15] range and kurtosis is set to be smaller than 30. We select these values based on existing empirical evidence (i.e., Christoffersen and Jacobs, 2004, Aït-Sahalia and Lo, 1998, Corrado and Su, 1996, Bates, 1991).

[10] As it is explained in Duda et al. (2001), the overall range (defined by γ) and slope (defined by b) of the activation functions are not important and do not affect the estimation of the artificial neural networks, as long as we avoid values that permit too flat or too steep slopes, which would make the optimization algorithm too slow or oscillatory. When enhancing the BS or the CS volatility, the activation function at the enhanced layer is logistic with γ = 0.70 and b = 0.5 (allows only positive values). When enhancing the CS skewness, the activation function at the enhanced layer is the hyperbolic tangent with γ = 15 and b = 0.15 (allowing both positive and negative values). When enhancing the CS kurtosis, the activation function at the enhanced layer is again the logistic one with γ = 30 and b = 0.15. These bounds are wider than implied parameter values reported in Bates (1991) that includes the crash period of 1987. In our dataset, these parameter constraints have never been binding.

The optimal number of hidden neurons is chosen via a cross-validation procedure. Cross-validation is a simple, very popular and straightforward technique, requiring only a minor computational burden (negligible in our case). Its purpose is to mitigate the overfitting problem and it is effective as long as the data generating mechanism of the datasets involved is relatively stable (because of this we select small validation and testing datasets). All ePOPM structures with one-to-ten hidden neurons are estimated and the one that performs the best in the validation period is selected. Using only the in-sample data the model is initialized, estimated and cross-validated with twenty different initializations (using thirty initializations does not improve the results). We employ the Nguyen and Windrow network initialization technique as it is explained in Hagan et al. (1996). This technique generates initial weight and bias values for a nonlinear activation function, so that the active regions of the layer's neurons are distributed (roughly) evenly over the input space. After defining the optimal network structure in-sample (i.e., known data), its weights are frozen and its pricing capability is tested out-of-sample in a third separate testing dataset (i.e., unknown data).

Finally, it must be noted that at first the implementation of our approach may appear more complex than that of the SVJ model. Nevertheless, in both cases the main difficulty is the need to minimize the pricing error, using a (numerical) Nonlinear Least Squares optimization algorithm. Moreover, each iteration during the SVJ estimation involves numerical integration of the complex-valued characteristic functions, associated with the moment generating functions of $ln(S_T / S)$. In general, estimating the GPF can be a simple and computationally efficient task, since calibrating the SVJ model requires at least three times more computational time for our testing period.

4. Data and methodology

Our dataset has been obtained from Commodity Systems Inc. and covers the period from January 2002 to August 2004 for a total of 671 trading days. The S&P 500 index call options are used, because this option market is extremely liquid and data synchronicity is a minimal issue (see also Garcia and Gencay, 2000, and Aït-Sahalia and Lo, 1998). For each trading day, the last bid and ask call prices for all contracts are available, along with the strike price[11], expiration date[12], volume and open interest. Similar to Jackwerth (2000) and Chernov and Ghysels (2000), we use a daily dividend yield, d_y, which in our case is collected from Datastream. In our analysis we use the midpoint of the call option bid-ask spread, since, as noted by Dumas et al. (1998), using bid-ask midpoints rather than trade prices reduces noise in the cross sectional estimation of implied parameters. Each day, the midpoint of the call option bid-ask spread at the close of the market, c^{mrk}, is matched with the closing value of the S&P 500 index. Time to maturity is computed assuming 252 days per year. We also use nonlinear cubic spline interpolation for matching each option contract with a continuous interest rate, r, that corresponds to the option's maturity. For this purpose, 1, 3, 6, and 12 months constant maturity T-bills rates were considered (collected daily from the U.S. Federal Reserve Bank Statistical Releases).

To create our dataset we first implement filtering rules that are similar to those in Bakshi et al. (1997). In addition, we require at least four data-points per maturity to ensure

---

[11] For the purposes of this study we use the following moneyness categories: *deep out-the-money* (DOTM) when $S/X \leq 0.90$, *out-the-money* (OTM) when $0.90 < S/X \leq 0.95$, *just out-the-money* (JOTM) when $0.95 < S/X \leq 0.99$, *at-the-money* (ATM) when $0.99 < S/X \leq 1.01$, *just in-the-money* (JITM) when $1.01 < S/X \leq 1.05$, *in-the-money* (ITM) when $1.05 < S/X \leq 1.10$, *deep in-the-money* (DITM) when $S/X > 1.10$.

[12] In terms of maturity, an option contract is classified as *short-term maturity* (where maturity ≤ 60 calendar days), as *medium-term maturity* (where maturity is between 61 and 180 calendar days) and as *long-term maturity* (where maturity > 180 calendar days).

that every maturity is satisfactorily represented. The final dataset has a total of 37202 observations and compares favourably with previous studies that test nonparametric methods (e.g., Hutchinson et al., 1994, Aït-Sahalia and Lo, 1988, Andreou et al., 2008). Sample characteristics for the dataset can be found in Table 1, where the average (contract-specific) implied volatility parameters are reported as well (see further discussions in section 5). The volatility anomaly is obvious both for the BS and the CS models.

[Table 1, here]

To estimate the GPF we use a chronological data partitioning via a rolling-forward procedure in order to have a better simulation of the actual option trading conditions. The data is divided into eighteen different training/estimation (*trn*) and validation (*vld*) sets, each followed by a separate and *non-overlapping* testing (*tst*) set. Each *trn*, *vld* and *tst* dataset has a 12, 2 and 1 month spanning period respectively and on average each dataset includes 14072, 2374 and 1202 observations[13]. In order to carry out the analysis, we create an aggregate testing period with 21644 data-points by pooling together the pricing estimates of all eighteen non-overlapping *tst* periods. For the aggregate testing period, we compute and tabulate the following: the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), the Median Absolute Error (MdAE), the 5th Percentile of Absolute Error (P_5AE), and the 95th Percentile of Absolute Error (P_{95}AE). Most of the analysis is based on the RMSE, since according to Christoffersen and Jacobs (2004) estimation and evaluation of a model should be based on the same error measure. In addition, they conclude that RMSE estimates perform the best among different loss functions. Finally, Bates (2000) points out that RMSE is a relatively intuitive error measure and is useful for comparison purposes.

[13] Keeping the model's free parameters the same for a one-month period is consistent with the reasoning of Bates (2000). Daily recalibration of the free parameters would imply that the enhanced models are never to be taken seriously as a genuine data generating mechanism.

4.1. Implied parameters

Our methodology for extracting the daily overall average implied parameters is similar to that in previous studies (for example, see Bates, 1991, Bakshi et al., 1997, Christoffersen et al. 2009, and León et al., 2009) that follow a simultaneous equation procedure to minimize a price deviation function with respect to the unknown parameters. Market option prices, c^{mrk}, are assumed to be the corresponding POPM prices, c^{k}, plus a random additive disturbance term, ε^{k}, where *k* = BS, CS, or SVJ. To find the implied parameter values per model, we solve an optimization problem that has the following form:

$$SSE(t) = \min_{\xi^k} \sum_{j=1}^{P_t} (\varepsilon_j^k)^2 = \min_{\xi^k} \sum_{j=1}^{P_t} (c_j^{mrk} - c_j^k)^2 , \tag{20}$$

where P_t refers to the number of different call option transaction data-points available in day *t*, and ξ^k refers to the unknown parameters associated with a specific POPM (*k* = BS, CS or SVJ). The SSE in Eq. (20) is minimized via Nonlinear Least Squares with a subspace trust region method, based on the Newton approach offered by the MATLAB® Optimization Toolbox. In order to minimize the possibility of obtaining implied parameters that correspond to a local minimum of the error surface with each model, we use several starting values for the unknown parameters based on daily average values reported by previous literature (more details are included in our Appendix which is available upon request). We obtain the following sets of daily overall average, *av*, implied parameters: *a*) daily overall average BS implied volatility estimates $\xi^{BS}=\{\sigma_{av}^{BS}\}$, *b*) daily overall average CS implied estimates $\xi^{CS}=\{\sigma_{av}^{CS}, \mu_3, \mu_4\}$, and *c*) daily overall average SVJ implied estimates $\xi^{SVJ}=\{\sigma_{av}^{SVJ}, \lambda, \bar{\kappa}, \theta, \alpha, \beta, \sigma_v, \rho\}$.

In order to have a pure unconstrained optimization problem and to avoid implausible implied values, we enforce certain transformations to each model parameters via strictly increasing and differentiable functions. Specifically: *i*) via log transformations we constrain σ_{av}^{BS}, σ_{av}^{CS}, and σ_{av}^{SVJ} to be positive[14], μ_4 to be smaller than 30, λ to be smaller than 10, θ, α and σ_v to be smaller than 2 and β to be smaller than 20, and *ii*) via the hyperbolic tangent sigmoid functions we constrain μ_3 to lie between -15 and +15, $\bar{\kappa}$ to lie between -0.99 and 0.99 and ρ to be between -1 and +1. For similar treatments of the optimization phase, see Bates (1991 and 2000) and Jondeau and Rockinger (2001).

We also estimate daily the three DVF models defined earlier. For BS, this is straightforward. For CS, this is done in two steps. Firstly, we fit the CS model to market option prices, to obtain the daily overall average implied parameter values ξ^{CS}. Then, skewness and kurtosis are fixed to those obtained earlier and the coefficients for the three different DVF models are estimated each day, using OLS[15] (*Lc*) and Nonlinear Least Squares (*NLc*), as in the BS case. For the Nonlinear Least Squares, we use several initializations to minimize the risk of estimating coefficients based on a local minimum of the optimization function (more details are provided in our Appendix). We denote the regression-based DVF models by using appropriate subscripts in the volatility variable: σ_{NL1}^{k}, σ_{NL2}^{k} and σ_{NL3}^{k} for the Nonlinear Least Squares estimation and σ_{L1}^{k}, σ_{L2}^{k} and σ_{L3}^{k} for OLS estimation (*k* = BS or CS).

For pricing and hedging positions at time instant *t*, the implied structural parameters derived at day *t*-1 are used together with all other information required. Daily recalibration of

[14] Similarly with the SVJ, we also calibrate the SV model. In addition, like Bakshi et al. (1997) and Christoffersen and Jacobs (2004), the instantaneous conditional variance, V_t, is calibrated daily as well (where for consistency, its square root is denoted as σ_{av}^{SVJ}).

[15] For OLS estimation in the CS case, we also need to recalibrate the model to obtain a daily contract-specific implied volatility value (after fixing the values for skewness and kurtosis).

the implied parameters (DVF and overall average) for POPMs is also adopted by Bakshi et al. (1997) and Christoffersen and Jacobs (2004) (see also Hull and Suo, 2002, Crépey, 2004, and Berkowitz, 2009).

4.2. The alternative models

With the BS_j models, we use as input *S, X, T,* d_y, *r*, and any of the following seven volatility estimates: σ_j^{BS}, where $j \in \{av, L1, L2, L3, NL1, NL2, NL3\}$. Similarly, we denote the seven parametric CS_j alternatives. Finally, for the SV and SVJ models we use the overall average parameter estimates. The notation for the enhanced models depends on the parametric model considered. We use eBS_{av} to denote the enhanced model where the BS volatility σ_{av}^{BS} is enhanced. In the same manner, we use eCS_{av}^1 to denote the ePOPM where the CS volatility σ_{av}^{CS} is enhanced, eCS_{av}^2 to denote the case where σ_{av}^{CS} and μ_3 are enhanced and finally, eCS_{av}^3 to denote the case where σ_{av}^{CS}, μ_3 and μ_4 are enhanced. In addition, the dividend adjusted moneyness ratio $(Se^{-d_y T})/X$ and the time to maturity, T, are also used as inputs to estimate the GPF.

5. Model calibration and analysis of the pricing results

As mentioned earlier, to obtain the best in-sample daily overall average implied parameters for each model, we try several alternative starting values[16]. As we have verified,

[16] Trying many starting values for estimation (in-sample) is necessary although not discussed much in the literature (see also discussions in Jondeau and Rockinger, 2000). All results we report are out-of-sample avoiding any data-snooping bias. Our effort is consistent with other authors like Bakshi et al. (1997), Christoffersen and Jacobs (2004), etc. In our rolling-forward procedure, after trying many initializations each day we select the best model "in-sample" and then test the selected model "out-of-sample".

this ensures that the parametric models are calibrated properly in-sample, avoiding thereby local minima, especially for the SV, SVJ, and the nonlinear DVF models. In Table 2, we report the mean daily estimates for the overall average parameter values of the POPMs along with their standard errors in parentheses (defined as the standard deviation of the daily parameter values, divided by the square root of the number of observations). Similarly to Corrado and Su (1996) and León et al. (2009), the results for CS demonstrate that the implied index return distributions are negatively skewed with higher kurtosis than permitted by the BS assumptions. Thus, the BS model with daily overall average volatility, *av*, is expected to perform poorly, compared with other models that allow for more flexible distributions (see also Bakshi et al., 1997). Regarding the volatility process parameters, similar to Bakshi et al. (1997) we observe that the implied volatilities of BS, SV and SVJ are quite close to each other and that the implied long-run mean volatility $\sqrt{\alpha / \beta}$ for SV equals 0.238 and is higher than the 0.144 for SVJ. In addition, the variation coefficient (volatility of volatility) σ_v for SV is almost double, compared with the one obtained for SVJ. In the SVJ model, additional volatility is induced by the random arrival and the random size of the jump. Thus, when calibrating the SVJ model from market data, it is reasonable to expect the long-run mean value of volatility and the volatility of volatility to be lower than that of the SV model. Our empirical results are consistent with the above-mentioned intuition and with results reported in Bakshi et al. (1997) - see also Bates (1996 and 2000). In addition, we find that the magnitude of the correlation coefficient, ρ, in SV is higher in relation to SVJ, indicating that SVJ captures part of the excess kurtosis and negative skewness with the jump component. According to Bates (1996), the negative correlation coefficient in SV is not enough to generate sufficiently negative implicit skewness. Therefore, a very high volatility of volatility (implausibly high relatively to the time series properties of asset prices) may be necessary to match the observed option values[17] (similar conclusions are obtained by Bakshi et al., 1997). Regarding the jump components of SVJ, we find that the average yearly frequency of jumps is 0.679, the average jump size is -14.9% and the jump size volatility is 19.2%. The jump size parameter values are higher compared with those of Bakshi[18] et al. (1997) and closer to those of Bates (2000). For the period of the study, the option implied values for the presence and magnitude of jumps are much higher compared with those implicit in the time series of the S&P 500 returns. They possibly reflect investors' fear of jumps that may not materialize in practice (see discussions of the peso-problem for index options in Aït-Sahalia et al., 2001).

[Table 2, here]

Table 3 tabulates the in-sample coefficient estimates for the DVF#2 model of BS and CS (mean values and below in parenthesis the *t*-statistics). The first observation is that the sign of the average coefficient values for BS coincides with those obtained by Dumas et al. (1998) and Christoffersen and Jacobs (2004): negative for X and T, whilst positive for X^2 and XT. Interestingly, the sign of the T and XT coefficients for CS are opposite to those for BS. This can be corroborated from Table 1, where we see that CS implied volatility term structure is different from the term structure implied by the BS model (especially for the long-term options). Second, similar to Christoffersen and Jacobs (2004), we observe in all DVF#2 parameters (both for BS and CS) that the average coefficient values of *NLc* are significantly smaller, compared with *Lc*, and more importantly less volatile[19]. According to

[17] Bates (2000) favours SVJ by arguing that in the presence of a jump component, the model provides option prices more compatible with market prices and generates more plausible implied stochastic volatility parameter values.

[18] In the case of Bakshi et al. (1997), the S&P 500 index exhibits a major uptrend move in the whole period, while in our case it exhibits a major downward-trend move during the first 15 months and a steady upward movement afterwards.

[19] In general, we have verified that the DVF#2 BS coefficient values plotted across the 671 days in our sample resemble, closely (for the common coefficients), the plots presented in Christoffersen and Jacobs (2004).

Christoffersen and Jacobs (2004), this manifests their better out-of-sample performance (as we also show below).

[Table 3, here]

5.1. Pricing results

Table 4A provides the in-sample RMSE and Table 4B demonstrates the out-of-sample pricing performance of all parametric models considered, in terms of RMSE, MAE, RMeSE, P_5AE and P_{95}AE for the aggregate testing period.

[Tables 4A, 4B, here]

We first focus our attention on Table 4A (in-sample) for the parametric BS, CS, SV and SVJ models. Before the alternative DVF versions are considered, the more complex parametric models with daily overall average values (*av*) exhibit superior performance and, thus, SVJ is the best, followed by SV. The regression-based DVF approach improves the pricing performance of BS and CS models considerably (especially with the nonlinear DVF#3); yet, the SVJ is the overall best model in-sample.

Concentrating on Table 4B (out-of-sample performance of the parametric models)[20], we see that the DVF-based CS models provide better performance than the corresponding DVF-based BS ones, in terms of RMSE. The best model is CS_{NL2}, which improves the RMSE performance over BS_{NL2} by 14%. The nonlinear DVF#2 model provides the best out-of-sample performance for both BS (consistent with the results in Dumas et al., 1998) and CS (although for the BS case the nonlinear DVF#1 was equally good in terms of the RMSE, but inferior by far in terms of the other error measures). The slightly inferior performance of the nonlinear DVF#3 model is not surprising (see also Dumas et al., 1998 who argue in favor of

[20] Our RMSE results are larger compared for instance to those of Bakshi et al. (1997) and Dumas et al. (1998), since for the period we examine the average index option prices are most of the time double or triple theirs.

the greater parsimony in the volatility function provided by DVF#2). Similar to Christoffersen and Jacobs (2004), we find that SV[21] underperforms BS_{NL2}. Despite this, among parametric models, the SVJ model is the best performer in all error metrics.

We then look at Table 5 for the out-of-sample pricing performance of the ePOPMs, and we see that all enhanced models have excellent performance. The out-of-sample RMSE for our models is in the range of 1.50 to 1.75, which is quite reasonable, since the bid-ask spread of the at-the-money options is in the range of 1.65 to 2.10. The BS enhanced version, eBS_{av}, shows an improvement of about 15% over the BS_{NL2} model. The best enhanced CS version is eCS_{av}^{2} in terms of RMSE, which enhances two parameters of CS (i.e., volatility and skewness) and improves RMSE by 15% over CS_{NL2}. The eCS_{av}^{3} model is also a good performer but it does not outperform eCS_{av}^{2} indicating that most of the pricing improvement comes from enhancing skewness. This is in the same line of thought as León et al. (2009), where negative skewness plays a more fundamental role in determining the S&P 500 index option prices than excess kurtosis; this is probably due to the smirk behavior that dominates the S&P 500 index implied volatility functions. It must be noted that the enhanced models outperform by far the equivalent parametric models with overall average parameters[22]. A

[21] We notice that in some cases the stochastic volatility model produces large out-of-sample mispricings (see also discussions in Jondeau and Rockinger, 2000). The problematic observations include options with maturities longer than 180 days. Table 4B presents two sets of pricing results for the stochastic volatility model. $SV*$ results are the original ones before replacing extreme mispricing observations. SV results are obtained after replacing, with BS_{NL2}, $SV*$ values that differ by more than 50% compared with BS_{NL2}. In total, 747 observations are replaced for the out-of-sample period. Christoffersen and Jacobs (2004, p. 307) mention that they have removed from their out-of-sample results certain problematic SV observations.

[22] For completeness, we also test the counterpart enhanced models with the nonlinear DVF#2 volatility input and we observe only a marginal impact on the results (see Appendix).

final observation is that the best enhanced models[23] are also competitive with the SVJ model, which is rather expensive to calibrate properly on a daily basis.

[Table 5, here]

In order to derive pair-wise comparisons regarding the pricing performance in terms of Mean Squared Error (MSE), we use symmetric and asymmetric *t*-tests (details are available in the Appendix). From the tests we conclude that the ePOPMs outperform the equivalent POPMs (both with overall average and DVF parameter estimates), with the difference to be statistically significant at the 1% level. The best ePOPM model is eCS_{av}^{2}, which is competitive with SVJ - any difference in performance is not statistically significant - and is much easier to estimate, since GPF is estimated once a month[24].

We use graphical diagnostics to investigate whether our semi-parametric method imposes any discipline on the models and to preclude the possibility that the enhanced parameters are just moving around excessively through time (plots are not displayed for brevity). Specifically, for each of the 379 out-of-sample days of the period from March 3rd, 2003 to August 31st, 2004, we use the estimated GPF and DVF models and get predictions for the daily volatility values for moneyness equal to 0.90, 1.00 and 1.10 and time to maturity

---

[23] We also check the performance of standard neural networks like the ones used in previous studies (i.e. Hutchinson et al., 1994, Garcia and Gencay, 2000). The optimization/training methodology is similar to the one employed for the enhanced models. The results for standard feedforward artificial neural networks are always inferior to those of the enhanced models. Specifically, their RMSE is between 2.013 and 2.743, which is quite large compared to the enhanced models whose RMSE is consistently below 1.754.

[24] In the Appendix, we tabulate the RMSE of the best performing models in terms of moneyness and time to maturity (7x3=21 classes). In summary, although SVJ has the best RMSE performance on aggregate, still it does not produce the least RMSE in all moneyness and time to maturity classes. By comparing eCS_{av}^{2} with eCS_{av}^{3}, we see that the enhancement of kurtosis in the latter model helps to improve the deep out-the-money (DOTM) and deep in-the-money (DITM) options but does not offer any improvement in other cases. This result is intuitive, since kurtosis affects the tails. Furthermore, eCS_{av}^{2} is slightly inferior to SVJ for short term options and marginally better for medium and long-term options.

equal to 21, 63, 126, 189 and 252 trading days (in total 15 combinations per day, per model). We can conclude that the enhanced volatility estimates, derived by the GPF, are in general less volatile/erratic compared with the DVF volatility estimates and they track better the evolution of the actual daily implied volatility[25]. A representative (numerical) summary of the graphical diagnostics is exhibited in Table 6, where we report the RMSE between the daily actual contract-specific volatility, implied by market option prices, and the volatility estimates obtained by the DVF-based BS_{NL2} and CS_{NL2} models and the enhanced models eBS_{av} and eCS^1_{av}. The first three rows of numerical results show RMSE for three specific moneyness cases (always aggregating time to maturity), while the last row of numerical results exhibits the aggregate RMSE values. The general conclusion is that the enhanced models provide more accurate predictions of the implied volatility surface, compared with the regression-based DVF counterparts.

[Table 6, here]

Table 7 shows the mean values of the enhanced parameters (i.e., volatility, skewness and kurtosis) for the eBS_{av}, eCS^1_{av}, eCS^2_{av} and eCS^3_{av} models for different maturity and moneyness classes for the aggregate testing period. These parameters are provided by the GPF (see Eqs. (10) and (11) and also X_E variables in the enhanced layer in Figure 1). For eCS^3_{av}, the enhanced volatilities preserve a smile effect in the short and medium term options, the enhanced skewness is increasing in moneyness and decreasing in maturity and the enhanced kurtosis exhibits a hump shape in moneyness. For eCS^2_{av}, the enhanced volatilities preserve a similar smile effect in the short and medium-term options and skewness exhibits a hump shape in moneyness for short and medium maturity options,

[25] The actual implied volatility is extracted daily from market option prices that are closest to each combination of moneyness and time to maturity.

similarly decreasing in maturity. Like the empirical evidence reported in León et al. (2009), we find that the enhanced skewness and kurtosis tend to be lower when the time to expiration is longer, especially for the JOTM, ATM and JITM options. Das and Sundaram (1999) compare the stochastic volatility model with jump-diffusion and conclude that "*it is less obvious whether the theoretical predictions of either class of models are – or can be made – consistent with the observed term structures of these deviations*". The authors wonder whether SV and Jump models can fit the data well by capturing the level of skewness and kurtosis, implied by the data for all maturities. They state that, compared with empirical observations, Jump models allow too rapid a decay in skewness and kurtosis and stochastic volatility models exhibit a hump shape overly pronounced for intermediate maturities. By contrast, the results for the enhanced CS model show that it allows more flexibility in all estimated parameters, not only in terms of maturity but also in terms of moneyness.

[Table 7, here]

6. Robustness analysis for the pricing results

We employ several robustness tests to check the performance of the enhanced models, which can be seen as tests of overfitting. We first check the pricing performance robustness of the enhanced models (in-sample and out-of-sample) by using fewer hidden neurons in the validation phase (the results are not reported for brevity). We use one-to-eight and one-to-six hidden neurons and the results show significant robustness compared with the case of one-to-ten hidden neurons that is used in our analysis. Specifically, with one-to-eight hidden neurons, the out-of-sample RMSE deteriorates 1% at most, while for one-to-six hidden neurons RMSE deteriorates 4% at most[26].

[26] We also calculate (not reported for the sake of brevity) out-of-sample RMSE by fixing the number of hidden neurons during estimation across periods. BS-based enhanced models exhibit similar performance for seven-to-ten hidden neurons

In the next test, we disaggregate each model's RMSE into two components: the RMSE of contract observations (in total 16609) that are common both in day *t*-1 (day used to extract the implied parameters) and in day *t* (out-of-sample day) and the RMSE of observations (in total 5035) that can be thought of as unseen data (they exist in day *t*, but not in day *t*-1). Results are presented in Table 8 (columns 2-3). It is important to see whether the models' performance in the unseen data is close to the performance in the common data or if it deviates significantly. As expected, the RMSE for the common (unseen) contract observations is always lower (higher) than the RMSE in the aggregate testing period. We observe that DVF models are less accurate when are used to price unseen observations, when compared with the BS and CS with overall average implied parameters. We also confirm that SV, compared with SVJ, is highly inaccurate when is used to price unseen data (especially of long maturity). We also observe that the enhanced models are robust with unseen data and some even outperform the SVJ model (for example, the RMSE of eCS^2_{av} in the unseen dataset is 1.693, quite smaller than 1.910 of SVJ). The robustness of the enhanced models is important, since unseen data comprises a considerable part of the contracts traded every day; about 23% of out-of-sample observations in our dataset are contracts not traded the day before[27].

[Table 8, here]

We also examine the models' performance to price totally unseen contracts (outside the moneyness range used in estimation). Specifically, for the period between January 2002 and August 2004, there are 784 observations with moneyness in the range of 0.70-0.80 and

---

with RMSE deterioration around 3.7%-6.7%, while between three and seven hidden neurons, RMSE is still below 1.90. CS-based enhanced models with five or more hidden neurons exhibit quite similar out-of-sample RMSE, which are close to the optimal ones, shown in Table 5.

[27] According to Bedendo and Hodges (2009), new trades incorporate new information; thus, robustness of the option pricing method to such trades is very important.

453 observations with moneyness in the range of 1.20-1.30 (in total 1237 new observations). Results are reported in the last column of Table 8**.** We note that CS-based parametric models outperform significantly the equivalent BS ones. The SV model (contrary to SVJ) performs very poorly (we verify again that large errors are mostly produced from long-term options). The enhanced models exhibit superior performance. In fact, the RMSE of eCS_{av}^{3} is 1.341, outperforming by far that of SVJ, which is 1.709; from unreported results, we also find this performance to be consistent in all moneyness and maturity ranges of the new observations.

One must remember that the enhanced models are estimated and then are used for a whole month, whereas the parametric ones (overall average and DVF) are estimated on a daily basis. Thus, as a final check we calculate the out-of-sample RMSE of the parametric models five days ahead. In such a case, the out-of-sample RMSE deteriorates for BS_{av} by 7.7%, for BS_{NL2} by 34.5%, for CS_{av} by 19.2%, for CS_{NL2} by 30.9%, for SV by 23.4% and for SVJ by 43.3%. The RMSE for five days ahead deteriorates a lot, making the enhanced models superior to the parametric ones (since they now outperform the SVJ considerably). The RMSE for ten days ahead (not shown for brevity) deteriorates even further. These results are consistent with the arguments in Christoffersen and Jacobs (2004), Bollen and Whaley (2004), and Hull and Suo (2002) that implied volatility functions are not persistent across time.

7. Single instrument delta-hedging analysis

We now investigate the hedging performance of the models and for consistency we use a single instrument delta-hedging analysis (see discussions in Bakshi et al., 1997). Bollen and Whaley (2004) find that the slope of the daily implied volatility functions, in terms of moneyness, is erratic, which may explain the poor performance of pricing models when used for hedging (see also discussions in Crépey, 2004). As suggested by Christoffersen and

Jacobs (2004), the above ambiguity may be due to the inappropriate choice of the loss function. They suggest that "*the best possible parameter estimates for a hedging exercise will likely be obtained using a hedging based loss function*". For the enhanced models, in order to align the estimation and evaluation loss functions, we employ the following methodology (see also Garcia and Gencay, 2000): enhanced parameters are estimated by minimizing the pricing RMSE but monitoring the hedging RMSE in the validation sample, so that for each period the model with the lowest hedging error is chosen. The results are compared with hedging results obtained by the parametric models and with those obtained by enhanced models optimized and chosen, based on a pricing loss function.

Single instrument delta-hedging results for the parametric models are reported in Table 9A. The single instrument analysis we employ is more appropriate for the parametric BS and CS and the respective enhanced models (thus the discussion is focused mainly on the comparison of those models). The BS models are the best performers among the parametric ones with respect to hedging, in contrast to pricing, where the CS models are superior. Moreover, the results here coincide with the propositions in Crépey (2004), according to which, in negatively skewed markets, the DVF-based BS models should provide better overall hedging performance than the BS_{av} on daily rebalancing intervals (see also discussions in Derman, 1999). Similar to the results of Dumas et al. (1998), most of the DVF-based BS models we consider are indistinguishable with respect to hedging performance. Similarly indistinguishable among themselves are the parametric CS-based models[28]. The hedging performance results of the enhanced models in Table 9B, are first given for models chosen for pricing and then for models chosen explicitly for hedging. In the first case, we compare results with those of the parametric models, with only the CS-based

[28] Evidence that parametric models, which can handle negative skewness and excess kurtosis can underperform the BS model for single instrument hedging, is also documented in Capelle-Blancard et al. (2001) and Jurczenko et al. (2004).

ones (eCS_{av}^{2} , and eCS_{av}^{3}) showing a noticeable improvement over the respective parametrics. The improvement is present when skewness (i.e., eCS_{av}^{2}) or both skewness and kurtosis (i.e., eCS_{av}^{3}) are enhanced, but not in the case when only volatility (i.e., eCS_{av}^{1}) is enhanced. Finally, when models are estimated based on a hedging criterion, we see that eCS_{av}^{2} and eCS_{av}^{3} improve hedging performance considerably, confirming the intuition in Christoffersen and Jacobs (2004). We note that the benchmark SVJ model (and SV) underperforms, as expected, the parametric BS and the enhanced models. The obtained enhanced parameters for models chosen on the hedging criterion capture efforts to hedge against exposures to the risks of the underlying asset[29]. From unreported results, we conjecture that the hedged positions seem to have two properties: skewness is positive, immunizing against the downsize risk and kurtosis is high, immunizing against the prospect of extreme returns (i.e., fat tails).

[Tables 9A and 9B, here]

We must remark that Bakshi et al. (1997) (see also Dumas et al., 1998, and Chernov and Ghysels, 2000) find that their models' hedging performance is virtually indistinguishable and that the hedging-based rankings of the models are in sharp contrast to their out-of-sample pricing performance. We reach similar conclusions for the case of the parametric models. However, as a significant variation from previous literature, we note that this is not the case for the enhanced models, especially the CS-based ones.

8. Summary and conclusions

In this study, we extend the benchmark nonlinear deterministic volatility regression functions of Dumas et al. (1998) for option pricing, with a nonparametric method to estimate

[29] The enhanced models chosen for hedging have poor performance in pricing, with RMSEs well above 2.

generalized parameter functions. The resulting enhanced parametric models have many desirable properties, compared with the standard implementation of artificial neural networks, such as theory consistent option values and hedging parameters. In general, this semi-parametric method is proposed as a way to alleviate deficiencies in the modern parametric option models and in standard nonparametric approaches. For pricing and hedging purposes, we use the S&P 500 index call options for the period from January 2002 to August 2004. We compare the GPF method with parametric models using both daily overall average implied parameters (for all parametric models considered) and daily contract-specific implied parameters, derived by the DVF regression method (for the BS and CS models). The SVJ and SV models are also included in the analysis.

We discuss the calibration of the parametric models, first for the overall average implied parameters of BS, CS, SV and SVJ and then for the DVF-based BS and CS models. We find that a careful estimation/optimization search is needed to obtain good implied parameters. The results obtained out-of-sample support the method that we propose. The first finding is that the DVF approach, when applied to CS, provides results superior to both CS (with overall average parameter estimates) and BS (with either overall average or DVF estimates). The second finding is that the SVJ model is the best model among the parametric ones, whilst the SV is inferior to the DVF-based BS and CS models. The third finding is that the increase in the pricing accuracy of the enhanced BS and CS models over the best performing BS and CS parametric ones is considerable and statistically significant. In general, the best enhanced models (with daily implied parameters but monthly estimation of the GPF) are comparable to the daily estimated SVJ. In addition, in respect to the complexity of the GPF and in pricing of contracts not used during estimation, the enhanced method is found to be robust.

In accordance with Christoffersen and Jacobs (2004), we observe that single instrument delta-hedging results for ePOPMs, chosen using a hedging criterion, outperform all of the parametric models as well as those ePOPMs chosen using a pricing criterion. The method we propose can also be employed to other studies, such as that of Brandt and Wu (2002), where option parameters are estimated from liquid European options and then applied to price illiquid or exotic derivatives. In addition, our method allows estimation of the term and the moneyness structure of skewness and kurtosis, which is essential for value-at-risk analysis (see Das and Sundaram, 1999).

Acknowledgements

This work is an extension of the third chapter of the PhD dissertation of P. Andreou. We are grateful to Ike Mathur (the Editor) and one anonymous reviewer for providing us with thorough comments and constructive suggestions. We are thankful to J. Bandler for motivation and discussions, to E. Derman, S. Perrakis, D. Paxton, R. Stapleton, L. Nordén, S. Staikouras and seminar participants at the U. of Cyprus, the U. of Manchester, the EFMA Annual Meetings in Madrid, Spain (June 2006), the Int. Conference on Computing in Economics and Finance in Limassol, Cyprus (June 2006), the FMA European Annual Meeting in Barcelona, Spain (May 2007), the EFMA Annual Meetings in Vienna, Austria (June 2007) and the CMS Conference Annual Meetings at the Imperial College in London, UK (March 2008) for useful discussions and comments. We also wish to acknowledge financial support from a U. of Cyprus research grant on contingent claims valuation.

References

Aït-Sahalia, Y., Duarte, J., 2003. Nonparametric option pricing under shape restrictions. Journal of Econometrics 116, 9-47.

Aït-Sahalia, Y., Lo, A., 1998. Nonparametric estimation of state-price densities implicit in financial asset prices. Journal of Finance 53, 499-547.

Aït-Sahalia, Y., Wang, Y., Yared, F., 2001. Do option markets correctly price the probabilities of movement of the underlying asset? Journal of Econometrics 102, 67-110.

Andreou, P.C., Charalambous, C., Martzoukos, S.H., 2008. Pricing and trading European options by combining artificial neural networks and parametric models with implied parameters. European Journal of Operational Research 185, 1415-1433.

Backus, D., Foresi, S., Li, K., Wu, L., 2004. Accounting for biases in Black-Scholes. Available at SSRN: http://ssrn.com/abstract=585623.

Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative options pricing models. Journal of Finance 52, 2003-2049.

Bandler, W.J., Ismail, A.M., Sanchez E.R.J., Zhang, J.Q., 1999. Neuromodeling of microwave circuits exploiting space-mapping technology. IEEE Transactions on Microwave Theory and Techniques 47, 2417-2427.

Bates, D.S., 1991. The crash of ’87: Was it expected? The evidence from options markets. Journal of Finance 46, 1009-1044.

Bates, D.S., 1996. Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options. The Review of Financial Studies 9, 69-107.

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

Becker, R., Clements, A.E., McClelland, A., 2009. The jump component of S&P 500 volatility and the VIX index. Journal of Banking and Finance 33, 1033-1038.

Bedendo, M., Hodges, S.D., 2009. The dynamics of the volatility skew: A Kalman filter approach. Journal of Banking and Finance 33, 1156-1165.

Berkowitz, J., 2009. On justifications for the ad hoc Black-Scholes method of option pricing. *Forthcoming in Studies in Nonlinear Dynamics and Econometrics* (available at: www.uh.edu/~jberkowi/pbs.pdf).

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637-654.

Bollen, N.P.B, Whaley, R.E., 2004. Does net buying pressure affect the shape of implied volatility functions? Journal of Finance 59, 711-753.

Brandt, M.W., Wu, T., 2002. Cross-sectional tests of deterministic volatility functions. Journal of Empirical Finance 9, 525-550.

Brown, C.A, Robinson, D.M, 2002. Skewness and kurtosis implied by option prices: A correction. Journal of Financial Research 25, 279-282.

Capelle-Blancard, G., Jurczenko, E., Maillet, B., 2001. The approximate option pricing model: Performances and dynamic properties. Journal of Multinational Financial Management 11, 427-443.

Chang, E.C., Ren, J., Shi, Q., 2009. Effects of the volatility smile on exchange settlement practises: The Hong Kong case. Journal of Banking and Finance 33, 98-112.

Chernov, M., Ghysels, E., 2000. A study towards a unified approach to the joint estimation objective and risk neutral measures for the purpose of option valuation. Journal of Financial Economics 56, 407-458.

Christoffersen, P., Heston, S.L, Jacobs, K., 2009. The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. Available at SSRN: http://ssrn.com/abstract=961037.

Christoffersen, P., Jacobs, K., 2004. The importance of the loss function in option valuation. Journal of Financial Economics 72, 291-318.

Corrado, C.J., 2008. The hidden Martingale restriction in Gram-Charlier option prices. Journal of Futures Markets 27, 517-534.

Corrado, C.J., Su, T., 1996. Skewness and kurtosis in S&P 500 index returns implied by option prices. Journal of Financial Research 19, 175-192.

Crépey, S., 2004. Delta-hedging vega risk? Quantitative Finance 4, 559-579.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signal and Systems 2, 303-314.

Das, S.R., Sundaram, R.K., 1999. Of smiles and smirks: A term structure perspective. Journal of Financial and Quantitative Analysis 34, 211-239.

Derman, E., 1999. Regimes of volatility. Risk 12, 55-59.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification. John Wiley and Sons, Inc.

Dumas, B., Fleming, J., Whaley, R.E, 1998. Implied volatility functions: Empirical tests. Journal of Finance 53, 2059-2106.

Garcia, R., Gencay, R., 2000. Pricing and hedging derivative securities with neural networks and a homogeneity hint. Journal of Econometrics 94, 93-115.

Hagan, M., Demuth, H., Beale, M., 1996. Neural Network Design. PWS Publishing Company.

Hull, J., Suo, W., 2002. A methodology for assessing model risk and its application to the implied volatility function model. Journal of Financial and Quantitative Analysis 37, 297-318.

Hutchinson, J.M., Lo, A., Poggio, T., 1994. A nonparametric approach to pricing and hedging derivative securities via learning networks. Journal of Finance 49, 851-889.

Jackwerth, J.C., 2000. Recovering risk aversion from option prices and realized returns. The Review of Financial Studies 13, 433-451.

Jondeau, E., Rockinger, M., 2000. Reading the smile: The message conveyed by methods which infer risk neutral densities. Journal of International Money and Finance 19, 885-915.

Jondeau, E., Rockinger, M., 2001. Gram-Charlier densities. Journal of Economic Dynamics and Control 25, 1457-1483.

Jurczenko, E., Maillet, B., Negrea, B., 2004. A note on skewness and kurtosis adjusted option pricing models under the Martingale restriction. Quantitative Finance 4, 479-488.

Konstantinidi, E., Skiadopoulos, G., Tzagkaraki, E., 2008. Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices. Journal of Banking and Finance 32, 2401-2411.

León, Á., Mencía, J., Sentana, E., 2009. Parametric properties of semi-nonparametric distributions, with applications to option valuation. Journal of Business & Economic Statistics 27, 176-192.

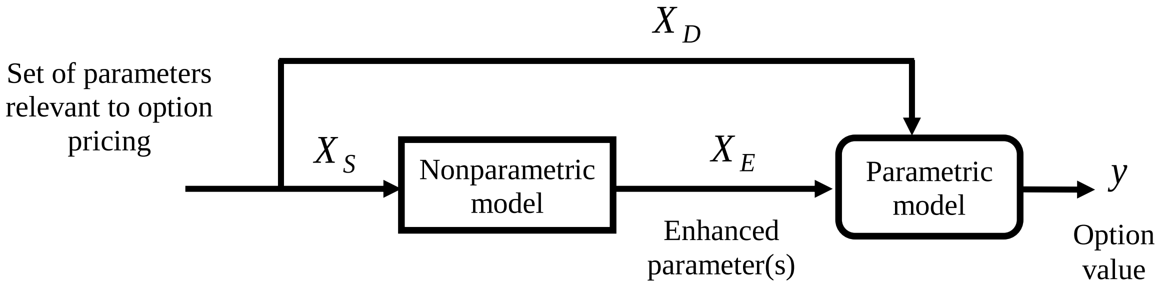
Pan, J., 2002. The jump-risk premia implicit in options: Evidence from an integrated time-series study. Journal of Financial Economics 63, 3-50.

Rubinstein, M., 1994. Implied binomial trees. Journal of Finance 49, 771-818.

Figures

Figure 1

Schematic description of the semi-parametric enhanced models (ePOPMs)



In the proposed semi-parametric option pricing model, the option value is provided by a parametric model. $X_I = X_E \cup X_D$ denotes the set of all input parameters that are necessary for the parametric model. $X_E \subseteq X_I$ corresponds to the enhanced parameters estimated nonparametrically and $X_D \subset X_I$ to those that are passed directly to the parametric model. In addition, X_S denotes the set of inputs to the nonparametric model.

Tables

Table 1
Sample characteristics

	DOTM	OTM	JOTM	ATM	JITM	ITM	DITM
S/X	**<0.90**	**0.90-0.95**	**0.95-0.99**	**0.99-1.01**	**1.01-1.05**	**1.05-1.10**	**≥1.10**
Short-Term Options ≤ 60 Days							
Call Price	2.749	4.742	9.589	20.483	38.920	72.966	119.305
BS Implied Volatility	0.258	0.212	0.179	0.184	0.206	0.254	0.339
CS Implied Volatility	0.429	0.290	0.214	0.200	0.202	0.219	0.266
# total sample obs	633	2525	5338	3080	4070	2172	1088
# out of sample obs	87	926	3308	2111	2745	1400	573
Medium -Term Options 61-180 Days							
Call Price	5.913	12.585	26.234	40.451	57.694	87.608	132.080
BS Implied Volatility	0.206	0.183	0.190	0.198	0.214	0.229	0.259
CS Implied Volatility	0.304	0.233	0.223	0.219	0.229	0.226	0.233
# total sample obs	2100	2802	2618	1474	1628	1014	701
# out of sample obs	684	1580	1569	927	1021	710	524
Long-Term Options > 180 Days							
Call Price	14.034	31.002	50.150	64.452	80.207	105.852	147.541
BS Implied Volatility	0.183	0.185	0.193	0.196	0.209	0.217	0.234
CS Implied Volatility	0.253	0.231	0.230	0.225	0.236	0.233	0.246
# total sample obs	1606	1273	1114	633	630	375	328
# out of sample obs	761	743	667	429	391	259	229

Figures refer to average market values of call options (first line), Black and Scholes implied volatility (second line), Corrado and Su implied volatility (third line), the total number of observations for the period from January 2nd, 2002 to August 31st, 2004 (fourth line) and the total observations used in the aggregate out-of-sample for the period from March 3rd, 2003 to August 31st, 2004 (fifth line).

Table 2
Daily average implied parameters for the parametric models

	σ	μ_3	μ_4	λ	$\bar{\kappa}$	θ	α	β	σ_v	ρ
BS	0.198 (0.002)									
CS	0.232 (0.003)	-1.099 (0.013)	4.179 (0.048)							
SV	0.212 (0.003)						0.186 (0.003)	3.274 (0.067)	0.809 (0.011)	-0.639 (0.003)
SVJ	0.183 (0.003)			0.679 (0.018)	-0.149 (0.008)	0.192 (0.013)	0.057 (0.002)	2.762 (0.085)	0.400 (0.013)	-0.614 (0.013)

Daily average implied parameter values are obtained by minimizing the sum of squared pricing deviations between a parametric model's estimates and the actual market value of the call options, for the period from January 2nd, 2002 to August 31st, 2004. The standard error of each parameter is reported in parenthesis.

Table 3

Summary statistics of coefficient estimates for the DVF#2 model

	Intercept	*X*	X^2	*T*	*XT*
	Black and Scholes				
Lc	2.105	-0.003	1.307E-06	-0.372	3.668E-04
	(60.921)	(-53.965)	(49.310)	(-18.200)	(19.411)
NLc	1.404	-0.002	7.451E-07	-0.228	2.304E-04
	(79.634)	(-61.423)	(49.912)	(-19.056)	(20.649)
	Corrado and Su				
Lc	1.081	-0.002	8.312E-07	0.403	-3.917E-04
	(25.984)	(-21.527)	(24.414)	(25.037)	(-22.795)
NLc	0.571	-0.001	3.366E-07	0.249	-2.331E-04
	(24.906)	(-14.728)	(16.332)	(26.957)	(-23.436)

Daily average coefficients obtained from fitting DVF#2 for the Black and Scholes and for the Corrado and Su models for the period from January 2nd, 2002 to August 31st, 2004. The *t*-statistics of each average coefficient value (computed from the daily values of the estimated coefficients) are reported below in parenthesis. *Lc* (*NLc*) is the ordinary least squares (nonlinear least squares) coefficient vector taken by regressing implied volatility on the variables included in DVF#2 specifications. All *t*-values indicate statistical significance at 1%.

Table 4A

In-sample pricing performance of the parametric models

	BS_{av}	BS_{L1}	BS_{NL1}	BS_{L2}	BS_{NL2}	BS_{L3}	BS_{NL3}
RMSE	3.365	2.964	1.783	1.492	0.861	1.305	0.690
	CS_{av}	CS_{L1}	CS_{NL1}	CS_{L2}	CS_{NL2}	CS_{L3}	CS_{NL3}
RMSE	1.572	2.098	1.467	1.119	0.782	0.954	0.636
	SVJ	SV					
RMSE	0.437	0.696					

Root Mean Squared Error (RMSE) values regarding the in-sample pricing performance for all parametric (overall average and DVF) models are obtained by minimizing the sum of squared pricing deviations between a model's estimates and the actual market values of the call options for the period from January 2nd, 2002 to August 31st, 2004.

Table 4B

Out-of-sample pricing performance of the parametric models

	BS_{av}	BS_{L1}	BS_{NL1}	BS_{L2}	BS_{NL2}	BS_{L3}	BS_{NL3}
RMSE	3.285	3.128	1.984	2.921	2.008	3.260	2.382
MAE	2.579	1.908	1.509	1.530	1.186	1.468	1.139
MeAE	2.172	1.164	1.213	0.962	0.833	0.834	0.739
P_5AE	0.242	0.091	0.115	0.082	0.078	0.072	0.067
P_{95}AE	6.396	6.440	3.796	4.364	3.100	4.161	2.983
	CS_{av}	CS_{L1}	CS_{NL1}	CS_{L2}	CS_{NL2}	CS_{L3}	CS_{NL3}
RMSE	2.245	2.794	2.110	2.248	1.766	2.667	2.189
MAE	1.709	1.890	1.609	1.451	1.257	1.438	1.252
MeAE	1.358	1.233	1.276	1.002	0.929	0.945	0.881
P_5AE	0.118	0.106	0.115	0.085	0.085	0.085	0.085
P_{95}AE	4.370	6.107	4.144	3.997	3.422	3.879	3.328
	SVJ	SV^*	SV				
RMSE	1.498	4.541	2.488				
MAE	1.071	1.551	1.318				
MeAE	0.796	0.900	0.904				
P_5AE	0.065	0.077	0.078				
P_{95}AE	2.996	3.413	3.362				

Error performance results (out-of-sample pricing) for all parametric models for the aggregate period from March 3rd, 2003 to August 31st, 2004. SV^* results are the original ones before replacing extreme mispricing observations. SV results are obtained after replacing, with BS_{NL2}, SV^* values that differ by more than 50% compared to BS_{NL2}. In total, 747 observations are replaced. RMSE is the Root Mean Squared Error, MAE is the Mean Absolute Error, MeAE is the Median Absolute Error and P_5AE (P_{95}AE) is the 5th (95th) Percentile of Absolute Errors.

Table 5

Out-of-sample pricing performance of the enhanced models (ePOPMs)

	eBS_{av}	eCS_{av}^{1}	eCS_{av}^{2}	eCS_{av}^{3}
RMSE	1.754	1.646	1.532	1.568
MAE	1.327	1.243	1.167	1.176
MeAE	1.046	0.957	0.925	0.908
P_5AE	0.097	0.084	0.079	0.079
$P_{95}AE$	3.473	3.345	3.081	3.169

Error performance results (out-of-sample pricing) for enhanced parametric models for the aggregate period from March 3rd, 2003 to August 31st, 2004. RMSE is the Root Mean Squared Error, MAE is the Mean Absolute Error, MeAE is the Median Absolute Error and P_5AE ($P_{95}AE$) is the 5th (95th) Percentile of Absolute Errors.

Table 6

Out-of-sample RMSE of volatility predicted by selected models

	BS_{NL2}	eBS_{av}	CS_{NL2}	eCS_{av}^{1}
	RMSE accuracy across specific moneyness cases			
S/X=0.9	0.021369	0.009754	0.020567	0.018049
S/X=1.0	0.010938	0.010813	0.020244	0.017309
S/X=1.1	0.02933	0.026033	0.026925	0.020244
	RMSE accuracy (aggregate)			
	0.0219	0.0172	0.0228	0.0186

Root Mean Squared Error (RMSE) values regarding the out-of-sample volatility prediction for selected models for the aggregate period from March 3rd, 2003 to August 31st, 2004. Volatility estimates for each day for specific moneyness (for five representative maturities) are compared with contract-specific implied volatility, taken using the market call option prices closest to each moneyness/maturity combination. Contract-specific implied volatility for the Corrado and Su model is computed after fixing the skewness and kurtosis coefficients to their daily overall average values.

Table 7

Summary statistics regarding the enhanced parameters for models (ePOPMs) optimized and selected using a pricing criterion

	DOTM	OTM	JOTM	ATM	JITM	ITM	DITM
S/X	<0.90	0.90-0.95	0.95-0.99	0.99-1.01	1.01-1.05	1.05-1.10	≥1.10
Short-Term Options ≤ 60 Days							
eBS_{av} **volatility**	0.220	0.171	0.158	0.167	0.180	0.207	0.254
eCS_{av}^{1} **volatility**	0.314	0.211	0.177	0.175	0.177	0.188	0.211
eCS_{av}^{2} **volatility**	0.236	0.182	0.166	0.172	0.181	0.198	0.228
eCS_{av}^{2} **skewness**	-0.550	-0.374	-0.347	-0.396	-0.432	-0.489	-0.496
eCS_{av}^{3} **volatility**	0.252	0.193	0.179	0.186	0.194	0.212	0.237
eCS_{av}^{3} **skewness**	-0.784	-0.651	-0.511	-0.391	-0.287	-0.194	-0.105
eCS_{av}^{3} **kurtosis**	5.008	5.517	6.089	6.030	5.854	5.742	5.635
Medium -Term Options 61-180 Days							
eBS_{av} **volatility**	0.173	0.158	0.166	0.176	0.187	0.204	0.234
eCS_{av}^{1} **volatility**	0.226	0.190	0.187	0.190	0.193	0.196	0.206
eCS_{av}^{2} **volatility**	0.187	0.170	0.177	0.185	0.192	0.200	0.215
eCS_{av}^{2} **skewness**	-0.499	-0.449	-0.487	-0.532	-0.547	-0.604	-0.626
eCS_{av}^{3} **volatility**	0.202	0.185	0.193	0.202	0.207	0.216	0.231
eCS_{av}^{3} **skewness**	-0.849	-0.747	-0.623	-0.522	-0.475	-0.360	-0.208
eCS_{av}^{3} **kurtosis**	5.488	5.894	5.933	5.873	5.694	5.711	5.722
Long-Term Options > 180 Days							
eBS_{av} **volatility**	0.164	0.168	0.174	0.181	0.192	0.201	0.220
eCS_{av}^{1} **volatility**	0.210	0.199	0.197	0.200	0.206	0.207	0.217
eCS_{av}^{2} **volatility**	0.185	0.187	0.190	0.195	0.201	0.203	0.214
eCS_{av}^{2} **skewness**	-0.627	-0.655	-0.710	-0.758	-0.686	-0.717	-0.757
eCS_{av}^{3} **volatility**	0.195	0.199	0.203	0.207	0.214	0.218	0.229
eCS_{av}^{3} **skewness**	-0.881	-0.801	-0.741	-0.678	-0.645	-0.543	-0.453
eCS_{av}^{3} **kurtosis**	5.451	5.596	5.622	5.557	5.475	5.433	5.429

Moneyness and time to maturity tabulation of enhanced parameters implied by the enhanced models for the aggregate testing period from March 3rd, 2003 to August 31st, 2004.

Table 8

Robustness analysis - RMSE for common/unseen and totally new observations

		Common in *t*-1 and *t*	**Unseen in *t*-1**	**Extra (new) observations**
BS_{av}	3.285	2.825	4.478	5.967
BS_{NL2}	2.008	1.298	3.432	4.802
CS_{av}	2.245	2.056	2.779	2.660
CS_{NL2}	1.766	1.482	2.483	3.089
SV	2.488	1.451	4.436	7.112
SVJ	1.498	1.348	1.910	1.709
eBS_{av}	1.754	1.680	1.978	1.838
eCS_{av}^{2}	1.532	1.480	1.693	1.836
eCS_{av}^{3}	1.568	1.497	1.785	1.341
#obs		16609	5035	1237

The results in the second and third (numerical) columns disaggregate each model's Root Mean Squared Error (RMSE) into two components: *i)* the RMSE of contract observations (in total 16609) that are common both in day *t*-1 (day used to extract the implied parameters) and in day *t* (out-of-sample day), and *ii)* the RMSE of contract observations (in total 5035) that can be thought of as unseen data (they exist in *t*, but not in *t*-1). The RMSE in the last column refers to the out-of-sample pricing performance of each model for 1237 totally unseen contract observations (outside the moneyness range used in estimation). The first column of results repeats RMSE for the aggregate dataset for comparison purposes.

Table 9A

Out-of-sample hedging performance of parametric models

	BS_{av}	BS_{L1}	BS_{NL1}	BS_{L2}	BS_{NL2}	BS_{L3}	BS_{NL3}
RMSE	1.180	1.116	1.135	1.114	1.118	1.114	1.116
MAE	0.900	0.835	0.857	0.829	0.835	0.828	0.831
MeAE	0.710	0.635	0.663	0.621	0.633	0.620	0.624

	CS_{av}	CS_{L1}	CS_{NL1}	CS_{L2}	CS_{NL2}	CS_{L3}	CS_{NL3}
RMSE	1.369	1.355	1.363	1.356	1.354	1.356	1.352
MAE	1.038	1.021	1.033	1.019	1.018	1.018	1.014
MeAE	0.805	0.779	0.800	0.776	0.778	0.774	0.771

	SVJ	SV
RMSE	1.359	1.378
MAE	1.010	1.020
MeAE	0.750	0.758

Error measures (out-of-sample) for single instrument delta-hedging performance of all parametric models for the aggregate period from March 3rd, 2003 to August 31st, 2004. RMSE is the Root Mean Squared Error, MAE is the Mean Absolute Error and MeAE is the Median Absolute Error.

Table 9B

Out-of-sample hedging performance of the enhanced models (ePOPMs)

	eBS_{av}	eCS_{av}^{1}	eCS_{av}^{2}	eCS_{av}^{3}
	Hedging performance for enhanced parametric models selected using a pricing criterion			
RMSE	1.123	1.353	1.222	1.192
MAE	0.843	1.020	0.903	0.869
MeAE	0.643	0.783	0.671	0.622
	Hedging performance for enhanced parametric models selected using a hedging criterion			
RMSE	1.117	1.301	1.093	1.080
MAE	0.826	0.960	0.815	0.800
MeAE	0.613	0.701	0.615	0.593

Error measures (out-of-sample) for single instrument delta-hedging performance of the enhanced parametric models (ePOPMs) for the aggregate testing period from March 3rd, 2003 to August 31st, 2004. The upper panel of results presents hedging performance of enhanced models that are optimized and selected using a pricing criterion, while the lower panel presents hedging performance of enhanced models selected using a hedging criterion. RMSE is the Root Mean Squared Error, MAE is the Mean Absolute Error and MeAE is the Median Absolute Error.