Does Combining Models Help With Pricing of European Options?

Semin Ibišević

Thesis submitted for the degree of Master of Science in Quantitative Finance at Erasmus University Rotterdam

August, 2013

Student number: 319115
Supervisor: dr. Michel van der Wel
Co-reader: Prof. dr. Dick van Dijk

ABSTRACT

In this study I combine six individual models to price European call options on the S&P500 index. Each model relaxes one or more assumptions of the Black-Scholes (1973) framework. The combining methods considered include three parametric and three non-parametric approaches. Combining outperforms most individual models in terms of the Root Mean Squared Error during the overall period (January 2006-December 2011) and three sub-periods. A simple and easy-implementable way of combining based on the historical performance of the individual models yields the best results among all individual and combining models considered. This method attaches most weight to the Merton (1976) jump diffusion model.

Keywords: Combining; Empirical option pricing; Neural Networks
JEL Classification: C45; C5; E60

ACKNOWLEDGEMENTS

Mostly I would like to thank my supervisor dr. Michel van der Wel for his detailed and concerned attitude towards this study. I appreciate his availability, the good references, and the supplementary focus on the writing aspect of scientific papers. I thank the co-reader Prof. dr. Dick van Dijk for his well written and clear discussion points and useful comments on the combining aspect. Also, I've had the opportunity to gain additional experience in performing quantitative research by working with Prof. Robin Lumsdaine during an internship (2012), which I believe has also benefitted this study.

CONTENTS

1 INTRODUCTION

The focus of option pricing is to obtain model-based price estimates that approach the true market prices. Various papers analyze ways to improve and obtain better estimates than the pioneering option pricing model as proposed by Black and Scholes (1973) (*BLS* model). While this model suffers much from systematic biases[1], there is still not a consensus on which alternative is the best option pricing model. A reason could be that many studies focus on relaxing one or more assumptions of *BLS*, hence discriminating which assumptions are important and neglecting others. In this way, not all features of the market are always captured properly and one is faced with determining which generalizations are important before applying option pricing techniques. In this study I try to cope for these concerns by means of a combining framework that relaxes multiple assumptions of the *BLS* method. The implications of this study may reveal important aspects of the diverse option pricing models, such as the importance of specific assumptions throughout time.

The list of possible alternatives to the *BLS* model is exhaustive long. Most of these models can be categorized into two main classes, referred to as modern parametric option pricing and non-parametric option pricing. The *parametric methods* are easier to use and understand, and are usually an extended framework of the *BLS* method. Most option pricing models within this category are based on at least three assumptions: the underlying price process, the interest rate process, and the market price of factor risks. For every assumption, there are multiple choices to be made. Examples include methods with stochastic volatility (see e.g., Heston (1993), Heston and Nandi (2000)), jump diffusion (see e.g., Merton (1976), Cheang and Chiarella (2011)) and stochastic interest rates (see e.g., Bakshi et al. (1997)). *Non-parametric* approaches incorporate multiple features by means of relaxing multiple simplified assumptions through fitting the data instead of assuming a distribution. Again, there are multiple choices to be made on which method to use in order to achieve the desired objective. Examples include demand-based option pricing and option pricing by means of a (hybrid) Artificial Neural Network (see e.g., Andreou et al. (2008)). Many proponents of non-parametric techniques argue that these approaches allow for higher multivariate and more complex nonlinear relationships, which in the end leads to better out-of-sample pricing performance. While non-parametric approaches allow for higher complexity and nonlinear relationships, the models are difficult to implement, may be subject to overfitting and are also not always easily interpretable.

[1]These biases are the result of the simplified assumptions of the model (Andreou et al., 2008), such as a geometric Brownian motion of stock price movements, continuous trading, continuous share price, constant interest rates, constance variance of the underlying returns and no dividends. See e.g., Black and Scholes (1975), MacBeth and Merville (1980), Gultekin et al. (1982), Bates (1991), Bakshi et al. (1997), Bates (2003).

Hence, the decision of choosing a parametric or non-parametric approach is a trade off between pricing accuracy and complexity.

In search of the perfect option pricing model, we first need to determine which assumptions we want to relax and whether it is worth the additional effort in terms of complexity and implementation time. Several strings of related option pricing literature have therefore turned their attention to applications of the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model (Stentoft, 2011). Recent contributions have managed to relax many assumptions of this approach such that more flexible parametric specifications of the underlying distributions can be realized. This includes applications (i) of non-Gaussian frameworks (see e.g., Christoffersen et al. (2006) and Christoffersen et al. (2010)); (ii) where volatility has both long and short run components (see e.g., Christoffersen et al. (2008a)); (iii) where flexible mixture models are used (see e.g., Robouts and Stentoft (2010)); (iv) GARCH models with jump components (see e.g., Christoffersen et al. (2008b)). Still, while these complex models are more capable in capturing multiple dynamics at once, they are prone to larger parameter uncertainty or calibration issues, such as non-convergence of the likelihood function to a global optimum. In this way, the complex models may become misspecified, which holds true for most approaches that extend the BLS framework.

Choosing the 'perfect' model is therefore a choice among misspecified models, as Bakshi et al. (1997) notes, for which the model is chosen (i) that is the the least misspecified; (ii) has the lowest pricing error; (iii) that obtains the best performance in terms of trading or hedging. Discriminating between models however may not be the way to go as one may neglect important features from models that perform less during a specific period of time. It is imaginable that misspecified model producing high pricing errors during a recession, may actually be the better model during another business cycle with its individual characteristics. Also, in this way the focus lies on assessing the relative performance instead of actually testing the performance in absolute terms.

A way to overcome discriminating and neglecting models that may contain important information, is by means of utilizing features from multiple candidate models. If every model is misspecified, then we may get closer to the 'true' market price by modeling this price as a function of multiple estimates from different sources. In general, in the first step we calibrate every candidate model and produce an estimate, such as an option price. Then we construct a combined estimate by assigning weights to each individual candidate model. In the field of forecasting, Timmermann (2006) argues that combining actually may work in several situations. For instance, the risk of loss in neglecting important information is reduced by the combining frameworks which implicitly discriminate the 'bad' models when needed, by assigning a low weight. This is in particular an useful tool when there are structural breaks affecting the individual models. It also holds true for rapidly changing market environments such as the time-varying option pricing attitudes of market participants (see e.g., Rubinstein (1985)). In this way, the risk of negative effects due to model misidentifications can be reduced, in a similar fashion as with diversifying assets. This may lead to a smaller gap between the theoretical price and the desired 'true' market price, with appreciable consequences for many applications.

Despite the increasing popularity of combining in economics and finance (see e.g., Rapach et al. (2010)), applications in the option pricing literature are relatively rare. One of the fewer exceptions is a recent article by Andreou et al. (2008) who proposes a (hybrid) Artificial Neural Network model to incorporate information from both the Black-Scholes-Merton (Merton, 1973) model as the Corrado and Su (1996) framework. Andreou et al. (2008) find that with these frameworks, a significant gain can be achieved in terms of pricing, hedging and trading of the S&P500 index. This approach is motivated by the assumption that parametric models cannot adjust fast enough to changing market behavior. Neural networks are free of financial theory as they use the historical data to estimate the empirical option pricing functions. As a result, the approach can recognize and handle with any nonlinearity. Additionally if frequently trained, ANN's can adapt to various and dynamic market environments. These results are in line with other studies on non-parametric techniques that have shown to be capable of outperforming the *BLS* model and some extended alternatives, in terms of pricing, hedging and trading (see e.g., Yao et al. (2000), Heston and Nandi (2000), Lin and Yeh (2005), Andreou et al. (2008) and Wang (2009)).

However, these non-parametric (combining) frameworks also imply some crucial disadvantages due to the additional complexities involved. As Tu (1996) discusses, neural networks (i) are a black box[2], (ii) may be more difficult to use in the field, (iii) require greater computational resources, (iv) are prone to overfitting and (v) the development is empirical with many methodological issues remaining to be resolved. These concerns are especially pronounced in the more advanced NN training concepts, making it beyond the scope of a newcomer. Also, the greater computational time may have its clear disadvantages for intra-day traders who frequently use option pricing techniques to allocate their position. Similar arguments can be given for other non-parametric, extended parametric methods (such as Bakshi et al. (1997)) and other complex techniques in option pricing. Also, the performance of a given approach is strongly linked to the field of study (exchange rate pricing, asset pricing, interest rate hedging, trading, etc.). As a

[2] It may be a bit pity to treat the Neural Networks as a 'black box', as reasonably sound statistical procedures for their implementation are available. See e.g., Teräsvirta et al. (2006). I thank Dick van Dijk for this comment.

consequence, a complex or extended approach is not a very useful tool for the majority of practitioners, who seek for consistent, fast and preferably closed-form elegant solutions that can deal with the shortcomes of the *BLS* model.

In this study I apply a more pragmatic approach of combining by dealing with the shortcomes of the complex models, but without loosing its benefits. That is, I test a weighted combination combination of multiple option pricing methods that relax one or more assumptions. In this way, the combined model is a flexible and adaptable approach that fits various kind of data. Hence, as the combined model is a function of extended individual models, multiple relaxations are allowed and additional complexities are discarded. The importance of a given assumption can in the end also be expressed by the magnitude of the corresponding weight. The model is then also less sensitive to rapid changing market behavior as it can capture the changes by assigning different weights to the individual models.

My approach of combining option prices involves two steps. First I calibrate the six candidate models, following the literature of Black and Scholes (1973), Merton (1973), Corrado and Su (1996), Merton (1976), Cox and Rubinstein (1979) and Heston (1993). The motivation for the choice of these models can be mostly attributed to the relative 'simple' nature of these frameworks, for which a closed form solution is available and which are relative straightforward to implement[3]. As each model relaxes one or more assumptions of *BLS*, the method allows for higher complexities by capturing multiple features of the market at once. After calibration of these models, I obtain the in-sample pricing errors and analyze the parameter estimates. Next, I combine the six individual models by means of three parametric approaches where weights are assigned to each candidate model: (i) by means of the historical performance in terms of the Discounted Mean Squared Error; (ii) with equal weights; (iii) and the median of all six candidate model estimates. These frameworks are compared against three non-parametric benchmark combining methods that utilize information from the candidate models by means of Neural Networks. Additional robustness checks include the assessment of the performance of the considered models during different sub-periods and by discarding the best[4] performing model from the analysis.

The results of this study reveal that combining is an useful asset in the option pricing literature, producing lower pricing errors for the overall period (2006-2011) and various

[3]In contrast, it would also make sense to combine the more elaborate models. Yet they are harder to implement, may be more sensitive to overfitting and require greater computational time, making them not useful for a large part of practitioners.

[4]In one of the robustness checks, the best performing model is discarded as a candidate model for the combining frameworks. The motivation for this approach is to determine to which extent the low pricing errors of the combining frameworks is attributed to the performance of this particular method. This importance is also expressed by the size of the weight.

sub-periods. While best results are obtained with combining by means of Neural Networks during and before the recession, the combining framework that assigns weights to individual models based on the historical performance yields consistent results during all sub-periods considered. During the overall period, it outperforms all individual and combining models for most moneyness-maturity categories. An additional contribution of this study is an assessment of the empirical performance of the candidate methods throughout time. The most striking results is that relaxing the jump diffusion process assumption yields the best performance during the overall period.

The remainder of this paper is as follows. Section 2 discusses the individual option pricing models used as an input for the combining frameworks. This is followed by a short discussion on the data used in section 3 and the structural estimation of the parameters in section 4. Three parametric approaches of combining are proposed in section 5 along with the non-parametric frameworks that are used as an benchmark. The implied parameters are analyzed along with the in-sample pricing fit of all the models in section 6. To control for the possibility of overfitting, the out-of-sample pricing results and additional robustness checks are presented in section 7. The study concludes in section 8.

2 OPTION PRICING MODELS

In this study, all the models to be combined, from now on referred to as *individual models*, fall into the class of parametric approaches. For the combining schemes I both consider parametric as non-parametric methods in order to exploit the information content of the individual models as best as possible. In addition, this allows for testing the relative performance of the simple approaches against the more complex non-parametric methods. A good performing model, which is in its nature simple and intuitive, will be favored by the vast majority of practitioners.

The reminder of this section discusses the individual models, from which the obtained option prices will be used as an input for the combining methods. Moreover, in this study I only include the initial BLS model in addition to five option pricing methods that individually relax one or more assumptions of BLS. Table 1 gives an overview of the candidate models and the assumptions that are relaxed. Note that not all assumptions are considered, such as the assumption of no-arbitrage pricing and no taxes and transaction costs. The major reasons for neglecting are the non-availability of 'elegant' closed form solutions or a lack of popularity in literature. The implications of including these missing assumptions is therefore an avenue open for further investigation. Still, the considered catalog of models and assumptions included is expected to capture the main and most important features, such as the presence of dividends, time-varying interest rate and jump processes.

The selection of the individual models plays a crucial role in the combining frameworks. Moreover, when individual models and the corresponding estimated prices are strongly correlated (multi-collinearity), there may be a large fraction of estimation uncertainty in the assigned weights. For this reason, it is important to try to obtain a diverse selection of models which capture different characteristics of the evolution of the stock market. In particular for the option pricing literature, this may be a concern to take into account. Moreover, as most parametric methods are an extension of BLS, there is quite often an overlap between the particular methods. For instance, the Merton (1973) ($MER1$) framework can be seen as a restricted variant of Corrado and Su (1996) (CSU) for skewness $\mu_3 = 0$ and kurtosis $\mu_4 = 3$. For this reason, it may be questionable whether the selection of the individual models is 'diverse' enough. On the other hand, the benefits of reducing model uncertainty and coping for model misspecification, by means of combining, remains evaluated as a tradeoff against the weights uncertainty. Therefore, the pricing results of the combining methods in the end provide an answer to this concern.

Table 1: Overview of the individual option pricing models used in this study

This table gives an overview of the individual models (all parametric) used in this study for the purpose of combining. Additionally I report which assumption(s) of Black and Scholes (1973) (BLS) the models relax.

Assumptions *BLS*		
A1	Log normal distribution of stock returns	
A2	Continuous trading	
A3	Continuous evolution of the share price	
A4	Constant interest rates	
A5	Constant variance of the underlying returns	
A6	No dividends	
A7	Continuous diffusion of the underlying	
Individual models		*Relaxes assumption(s) BLS*
(i)	Black and Scholes (1973)	-
(ii)	Corrado and Su (1996)	A1, A6
(iii)	Merton (1976)	A2, A7
(iv)	Cox and Rubinstein (1979)	A3
(v)	Merton (1973)	A6
(vi)	Heston (1993)	A5, A4*

* While the original Heston model does not generalize the assumption of constant interest rates, it can be adapted to allow for time-varying rates. See Bakshi et al. (1997) for stochastic volatility models with an stochastic interest rate component (not included in this study).

Notation Throughout the discussion of the individual models I use the same notation for the overlapping variables. This includes, S_t which is the stock price at time t, X the strike price, r the annualized continuously compounded risk-free rate (assuming 252 trading days in a year) and T the time to maturity of the option. All other variables and parameters are discussed in the corresponding sections. In section 4 I discuss ways to obtain the implied parameters of each candidate option pricing model, through solving the corresponding equations of each candidate model, iteratively for the implied parameters given the values of the observable market price.

2.1 Black-Scholes (1973)

The Black and Scholes (1973) for the price at time t of a European call option c_t^{BLS} is given by:

$$c_t^{BLS} = S_t N(d_1) - X e^{-rT} N(d_2), \tag{2.1}$$

with

$$d_1 = \frac{ln(S_t/X) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \tag{2.2}$$

$$d_2 = d_1 - \sigma\sqrt{T}. \tag{2.3}$$

Here σ is the stock price volatility which is the only variable that is unobserved. The function $N(x)$ is the cumulative probability distribution function of a standard normal distribution and hence expresses the probability that a variable following a standard normal distribution will be less than a particular value x. It can be easily seen that the call price increases for a greater distance between S_t and X, a higher volatility σ and higher risk free rate r. For a detailed description of this model and a derivation of the formula I refer to Hull (1999) and Black and Scholes (1972).

2.2 Merton (1973)

A generalization of the BLS model with respect to dividends is given by the Merton (1973) ($MER1$) framework. This method allows to price European options on stock or indices paying an observed dividend yield. The formula for the call option c_t^{MER1}, with dividends paid on the underlying stock, is given by:

$$c_t^{MER1} = S_t N(d_1^*) - X e^{-rT} N(d_2^*), \tag{2.4}$$

with

$$d_1^* = \frac{ln(S_t/X) + (r-\delta)T + (\sigma\sqrt{T})^2/2}{\sigma\sqrt{T}}, \quad (2.5)$$

$$d_2^* = d_1^* - \sigma\sqrt{T}. \quad (2.6)$$

Here δ is the continuously compounded annual dividend yield. As it is in the case for BLS, the volatility of the underlying is unobserved and needs to be estimated. Some of the shortcomings of this approach are the assumptions that dividends are paid out continuously and that it is a know constant. When pricing a stock index however, this problem is less profound as it averages outs the payments of multiple individual stocks. A higher dividend yield implies a lower call price, which is due its negative effect on the underlying stock price. Moreover, the stock price is adjusted downwards by a certain fraction of the dividend on the dividend payment date. This is because a dividend payment reflects a reduction in the company's market cap value.

2.3 Corrado and Su (1996)

Rubinstein (1985) and various others show that the implied volatility as a function of the moneyness ration (S/X) and time to expiration (T), derived via diverse option pricing models, typically exhibits a U shape. In literature this is refereed to as the *volatility smile*. This occurs when models do not account for returns which are negatively skewed and have excess kurtosis (Bates, 1991). As a correction for this concern, Corrado and Su (1996) propose a method to extend the $MER1$ model to account for these biases induced. Through a Gram-Charlier series expansion of the normal density function, the method provides skewness and kurtosis adjustment terms for the initial models. In this way, the method does not rely on specific assumptions about the underlying stochastic process and can be classified as a *semi-parametric* approach (Andreou et al., 2008). Following Andreou et al. (2008), I apply the correction for the $MER1$ model. The formula for the European call option c_t^{CSU} obtained with the Corrado and Su (CSU) method at time t is given by:

$$c_t^{CSU} = c_t^{MER1} + \mu_3 Q_3 + (\mu_4 - 3)Q_4, \quad (2.7)$$

with

$$Q_3 = \frac{1}{3!} S_t e^{-\delta T} \sigma\sqrt{T}((2\sigma\sqrt{T} - d_1^*)h(d_1^*) + \sigma^2 T N(d_1^*)), \quad (2.8)$$

$$Q_4 = \frac{1}{4!} S_t e^{-\delta T} \sigma\sqrt{T}([d_1^*]^2 - 1 - 3\sigma\sqrt{T}(d_1^* - \sigma\sqrt{T}))h(d_1^*) + \sigma^3 T^{3/2} N(d_1^*), \quad (2.9)$$

$$h(z) = \frac{1}{\sqrt{2\pi}} \exp(-z^2/2). \quad (2.10)$$

Here μ_3 and μ_4 are the coefficients of skewness and kurtosis of the returns. Like σ they are unobserved and need to be estimated. This can be achieved in a similar way as for the previous two models. For $\mu_3 = 0$ and $\mu_4 = 3$, the CSU framework is equivalent to $MER1$.

2.4 Merton (1976)

Another way to capture negative skewness and excess kurtosis of the underlying stock price density, and hence 'fat tails', is by including a Poisson jump component in the generation of the underlying stock returns as proposed by Merton (1976) ($MER2$). Merton assumes that in this way the extra randomness can be diversified away which is imposed by jumps in the underlying . As explained by Cheang and Chiarella (2011), assume that the jump sizes are normally distributed with mean α, variance ϑ^2 and jump-intensity λ, under the equivalent martingale measure $\mathbb{Q}$. For a stock paying continuous dividend at the rate δ and for which the stock price follows a geometric Brownian motion with an additional Poisson jump component, the European call option price c_t^{MER2} is given by:

$$c_t^{MER2} = \sum_{i=0}^{\infty} \frac{e^{-\overline{\lambda}T}(\overline{\lambda}T)^i}{i!} c_t^{BLS}(S_t, X, T, \sigma_i, r_i), \tag{2.11}$$

where $c_t^{BS}(.)$ is the Black-Scholes price without jumps with the following inputs: the spot price S_t, strike X, time-to-maturity T, the total variance with jumps σ_i and the adjusted risk free rate r_i. Hence $\overline{\lambda} = \lambda e^{\alpha + \frac{\vartheta^2}{2}}$ with

$$\sigma_i^2 = \sigma^2 + \frac{i\vartheta^2}{T}, \tag{2.12}$$

$$r_i = r - \lambda(e^{-\overline{\lambda}T} - 1) + \frac{i(\alpha + \frac{\vartheta^2}{2})}{T}. \tag{2.13}$$

The parameters to be estimated include the volatility of the underlying σ, and the three jump parameters λ, α and ϑ. The intensity is annualized and obviously a larger value of λ means that the jumps are expected to occur more frequently during a year. This jump diffusion model can be also seen as a weighted average of the Black-Scholes model, with the weights determined by the underlying asset price jumps dynamics.

2.5 Cox and Rubinstein (1979)

The binomial option pricing technique, as initialized by Cox and Ross (1976) and later extended as the Cox and Rubinstein (1979) approach, can be applied to calculate and price complex options. The model relaxes the assumption of continuous evolution of the share price by considering the evolution discrete and that it can only take two possible values each period. Moreover, the model assumes that the current stock price S_t can either move up or down with a specific proportion u and probability p. This implies that the underlying asset price can be seen as a binomial tree where each node represents a possible value of the asset price.

To price the call option, the model assumes that the underlying asset price rises by a factor $u = e^{\sigma\sqrt{h}}$ or decreases by a factor $d = u^{-1}$, with $h = T/n$ which is the size of the time interval between two successive jumps and n the number of time steps used in calculations. The probability of an upward jump is given by the risk-neutral probability $p = (e^{rh} - d)(u - d)$, which is chosen such that it simulates a geometric Brownian motion of the underlying stock.

For a given number of iterations n, the call price of an European call option is given by:

$$c_t^{COX} = \frac{\left[\sum_{j=a}^{n}\left(\frac{n!}{j!(n-j)!}\right)p^j(1-p)^{n-j}[u^j d^{n-j} S_t - X]\right]}{r^n}, \tag{2.14}$$

for a standing for a minimum number of upward moves which the stock must make over the next n periods for the call to finish in-the-money, equivalent to the smallest non-negative integer such that $u^a d^{n-a} S_t > K$. For $a > n$ this implies $c_t^{COX} = 0$ as call will finish out-of-the-money even if the stock moves upward every period. The value n is also referred to as the height of a binomial tree. For $n \to \infty$ the COX model converges to the continuous version of BLS. It may be optimal and intuitive to set n equivalent to T as it represents a time interval h of 1 between two successive jumps, corresponding to 1 day. However, this results in a long computational time. For this concern I set n equivalent to 25 which is approximately in the 0.18 quantile of all time-to-maturities for this dataset.

2.6 Heston (1993)

There is evidence that realized volatility of traded assets displays significant variability. Some explanations can be attributed to the economic effects that give rise to an equity skew, such as leverage effects and massive portfolio re-balancing in case of declining stock prices. By assuming nonconstant volatility of the underlying, Heston (1993) (HES) derives a closed-form solution to price European call options with stochastic volatility. This model also allows for correlation between volatility and spot asset returns and can be adapted to incorporate for stochastic interest rates (see e.g., Bakshi et al. (1997)). The HES method was introduced as an alternative for the at the moment stochastic volatility models which don't have a closed form solution (see e.g., Scott (1987) and Hull and White (1987)) or assumed that the volatility is uncorrelated with the spot asset (see e.g., Jarrow and Rudd (1982)). While the HES model may account for stochastic volatility and improve the pricing performance of call options under certain conditions, the model has five unknown parameters which can be difficult to calibrate[5]. The Heston formula is a solution in a form similar to the *BLS* model. Moreover, the call price c_t^{HES} on a non-dividend paying asset for the Heston model is given by:

$$c_t^{HES} = S_t P_1 - K e^{rT} P_2. \tag{2.15}$$

Here P_1 is also known as the delta of the call option, whereas P_2 is the conditional risk neutral probability that the asset price will be greater than the strike at maturity. The terms P_1,P_2 can be defined via the inverse Fourier transformation and are given by[6]:

$$P_j = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathbf{Re}\left[\frac{e^{-iuln(K)}\varphi_j(S_t,\sigma,T,u)}{iu}\right] du, \; j = 1,2. \tag{2.16}$$

The characteristic functions φ_1 and φ_2 are given in the form:

$$\varphi_j(S_t,\sigma,T,u) = \exp\left[C_j(T;\varphi) + D_j(T;\varphi)\sigma + i\varphi S_t\right], \; j = 1,2. \tag{2.17}$$

As described in Mikhailov and Nögel (2003), the characteristic functions can be substituted in the Garman equation to get the following ordinary differential equations for the unknown functions $C_j(T,\varphi)$ and $D_j(T,\varphi)$:

$$\begin{aligned} \frac{dC_j(T,\varphi)}{dT} - \kappa\varphi D_j(T,\varphi) - (r-d)\varphi i &= 0, \\ \frac{dD_j(T,\varphi)}{dT} - \frac{\nu^2\varphi D_j^2(T,\varphi)}{2} + (b_j - \rho\varphi i)D_j(T,\varphi) - u_j\varphi i + \frac{\varphi^2}{2} &= 0 \end{aligned}$$

With zero initial conditions $C_j(0,\varphi) = D_j(0,\varphi) = 0$, the solution of this system is given by:

$$C(T,\varphi) = (r-d)\varphi iT + \frac{k\varphi}{\nu^2}\left[(b_j - \rho\nu\varphi i + b)T - 2ln\left[\frac{1-ge^{br}}{1-g}\right]\right], \tag{2.18}$$

$$D(T,\varphi) = \frac{b_j - \rho\nu\varphi i + b}{\nu^2}\left[\frac{1-e^{bT}}{1-ge^{bT}}\right], \tag{2.19}$$

[5]Finding the global minimum can be difficult and depends on the optimization method used. Gradient based optimization methods can be useless and unique solutions need not necessarily to exist. Another danger when implementing the Heston framework using the Fast Fourier Transform approach is the 'Heston trap' (see e.g., Albrecher et al. (2007)) which can arise when computing the logarithm of complex numbers. Lord and Kahl (2010) discusses some of these complex discontinuities and how to avoid them in Heston-like models.

[6]**Re**[**x**] is the real part of a complex number **x**.

where:

$$
\begin{aligned}
g &= \frac{b_j - \rho v \varphi i + b}{b_j - \rho v \varphi i - b}, \\
b &= \sqrt{(\rho v \varphi i - b_j)^2 - v^2(2u_j \varphi i - \varphi^2)}, \\
u_1 &= 0.5,\ u_2 = -0.5, \\
a &= \kappa \varphi, \\
b_1 &= \kappa + \lambda - \rho v, \\
b_2 &= \kappa + \lambda.
\end{aligned}
$$

The parameters of interest, which are unobserved and need to be calibrated in this framework include σ the variance of the underlying, v the volatility of the variance, κ the mean reversion rate, φ the long run variance and ρ the correlation between the log returns and the volatility of the asset. I assume $\kappa > 0, v > 0$ and $|\rho| < 1$. For $v = 0$ this model converges to the *BLS* model with a time-dependent volatility. The remaining unknown is the market price of the volatility risk λ, which is in practice difficult to estimate. Yet it can be shown that under the martingale measure $\mathbb{Q}$, this parameter is eliminated. In addition, this parameter also eliminates when applying a Fast Fourier Transform (FFT) method to evaluate equation 2.15. Proposed by Carr and Madan (1999), this method is much quicker than using a numerical optimization for the mentioned integrals. This includes distinguishing between in-the-money (and at-the-money) and out-of-money options which depend differently on the intrinsic value of options[7]. The FFT method is discussed in more detail in Appendix A.1.

3 DATA

The models are evaluated and estimated using a data set on plain vanilla S&P500 European call options obtained from *OptionMetrics*, also available at the *WRDS* database. This is a popular index among researchers[8], especially due to its high liquidity. Additionally, the daily dividend distributions are available. The sample consists of daily bid-ask quotes on option contracts in the period January 3, 2006 - December 31, 2011. The call price is then assumed to be the mid price. The choice for the daily data makes this study useful for a majority of practitioners and institutions who, if needed, reallocate their position on a daily base for the purpose of hedging or trading.

Additionally, I obtain daily data on the zero-coupon interest rate and dividends on

[7]Out-of-money options don't have an intrinsic value. For this reason, Carr and Madan (1999) derive a FFT based on the time value of the option.

[8]See for example the references of Christoffersen et al. (2006).

the index, again from the OptionMetrics database. As the zero-coupon interest rates are given for only specific days-to-maturities, I interpolate the values for missing days such that rates for every day-to-maturity is available. By means of the Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) method, the resulting yield curve is smooth and I don't observe any anomalies. The S&P500 spot index series are obtained from Yahoo Finance, for which also the *dividend-exclusive series* are available. This is consistent with Bakshi et al. (1997) who adjust the series for discrete dividends by subtracting the present value of future dividends from the spot index level.

Following Andreou et al. (2008) and Bakshi et al. (1997), I employ several exclusion filters. This includes eliminating call contracts (i) for which the call contract price is greater than the asset value; (ii) with less than six days to expiration (as they may lead to liquidity-related biases); (iii) price quotes not satisfying the arbitrage restriction

$$c_t \geq max(0, S_t - X, S_t - \delta - X r_f),$$

with S_t the spot price, X the strike price and δ and r_f the dividend and risk-free rate respectively; (iv) price quotes lower than 3/8 (to discard to the impact of price discreteness on option valuation). By employing this criteria, a total of 155470 observations, approximately 14% percent of the total, are filtered out of the original sample that consisted of more than a million observations. This leaves 901593 observations to be included in the analysis which is larger than in any other study known. The exclusion of observations is mostly attributed to the filtering rules (ii) and (iv). The fraction deleted observations is similar as in Andreou et al. (2008).

Table 3 reports the sample properties of option contracts in the remaining dataset, categorized by the moneyness and time-to-expiration sets. There are more in-the-money (574802) than out-the-money (291677) options, which is also observed in the study of Bakshi et al. (1997). The average price call price is 211.03 with an average spread of 3.06. This is especially due to the high number of observations and high call price of the very deep in-the-money options ($S_t/X > 1.35$). Options with less than 60 days to maturity and that are out and at-the money are the most liquid, while options in-the-money and time-to-maturity more than 180 days the least.

4 PARAMETER ESTIMATION PROCEDURE

To apply option pricing models we need to find estimates for the volatility of the underlying and other corresponding parameters of the models. As these are usually unobservable, I minimize a price deviation function with respect to the unobserved parameters and spot volatility. The 'true' prices are assumed to be the market option prices (c_t^{mkt}).

TABLE 2: SAMPLE PROPERTIES OF THE S&P500 EUROPEAN CALL OPTIONS (2006-2011)

Reported are the average bid-ask mid-point quotes, the bid-ask spread (ask minus bid) shown in parenthesis () and the total number of observations shown in brackets []. The contracts are categorized by each moneyness-maturity categories and weighted to obtain an average value per set. The sample period covers January 3, 2006 until December 31, 2011. S_t is the spot price on the S&P500 and X the strike. The layout of this table is conform Bakshi et al. (1997) and the moneyness-maturity categories are inspired by Andreou et al. (2008).

Moneyness	**Days to expiration**			
S_t/X	< 60	$60-180$	> 180	*All*
< 0.85	1.62 (1.24) [7624]	2.64 (1.28) [21972]	13.95 (2.48) [71877]	10.58 (2.13) [101473]
$0.85-0.95$	4.38 (1.08) [33641]	11.24 (1.66) [43058]	52.53 (3.62) [44366]	24.47 (2.21) [121065]
$0.95-0.99$	12.34 (1.46) [29637]	33.07 (2.42) [21559]	92.82 (3.85) [17943]	39.69 (2.38) [69139]
$0.99-1.01$	27.30 (2.05) [15123]	51.75 (2.65) [10756]	113.32 (3.92) [9235]	57.41 (2.72) [35114]
$1.01-1.05$	52.13 (2.48) [27542]	74.89 (2.77) [19151]	133.96 (3.90) [17047]	80.85 (2.95) [63740]
$1.05-1.10$	94.72 (2.69) [30336]	111.28 (2.95) [19481]	166.05 (3.97) [19241]	119.27 (3.12) [69058]
$1.10-1.35$	205.15 (2.94) [88896]	214.85 (3.16) [57962]	256.73 (4.20) [60195]	222.86 (3.37) [207053]
> 1.35	461.61 (3.21) [75252]	507.58 (3.41) [65696]	574.75 (4.74) [94003]	519.73 (3.88) [234951]
All	189.00 (2.51) [308051]	197.25 (2.69) [259635]	242.60 (3.85) [333907]	211.23 (3.06) [901593]

At every point in time t, there are Q_t option contracts available, hence option transaction datapoints, with each different characteristics in terms of strike price and expiration date. Using all the option contracts from the past 10 days[9] up to today t, I split this in-sample dataset by the eight moneyness and three time-to-maturity categories[10], resulting in 24 sets of in-sample data. The following therefore holds for every model $k = 1,...,M$ and for every set of options data $1,...,N_{t,p}$ that share the same maturity and time-to-expiration:

$$\begin{aligned} c_t^k &= \{c_{t,p}^k\}_{p=1}^{24}, \\ c_{t,p}^k &= \{c_{t,p,j}^k\}_{j=1}^{N_{t,p}}, \\ Q_t &= \sum_{p=1}^{24} N_{t,p}. \end{aligned}$$

Then the difference

$$\varepsilon_{s,p,j}^k = c_{s,p,j}^{mkt} - c_{s,p,j}^k, \tag{4.1}$$

$$\text{for } j = 1,...,N_{s,p},\ s = t-10,...,t$$

between the market $c_{s,p,j}^{mkt}$ and the model value of a certain option $c_{s,p,j}^k$ is a function of the values taken by the unknown parameters. To find the optimal implied parameters, I solve at every point in time t the optimization problem in the following form:

$$SSE(t,p,k) = \min_{\theta_{t,p}^k} \sum_{s=t-10}^{t} \sum_{j=1}^{N_{s,p}} (\varepsilon_{s,p,j}^k)^2, \tag{4.2}$$

where $\theta_{t,p}^k$ is the vector of parameters corresponding to model k at time t and moneyness-maturity category p. This results for every time t and moneyness-maturity category p in $\theta^{BLS} = \{\sigma\}$, $\theta^{MER1} = \{\sigma\}$, $\theta^{csu} = \{\sigma, \mu_3, \mu_4\}$, $\theta^{MER2} = \{\sigma, \lambda, \alpha, \vartheta\}$, $\theta^{COX} = \{\sigma\}$ and $\theta^{HES} =$

[9] Andreou et al. (2008) and Bakshi et al. (1997) use only the current day's option contracts to the obtain the current day's implied parameters. The average number of option contacts in my dataset is approximately 596. In estimating the implied parameters by splitting this sample into the 24 moneyness-maturity categories, some categories (such as the very deep out-the-money and very deep in-the-money) are left without sufficient number of observations (less than 20). I therefore use all option contracts of the past 10 days to secure stable estimation of the implied parameters. In addition, this may have it's benefits for out-of-sample pricing as time-variation is taken into account as well. Taking into account only the current day's option contracts may be an additional robustness check open for further investigation.

[10] Andreou et al. (2008) analyzes the effect of inclusion of several sets of parameters, including historical parameters, a VIX volatility proxy derived by weighting implied volatilities and implied parameters which are obtained by minimizing the Sum of Squared Errors with (i) all available options data available during the in-sample period, (ii) options that share the same maturity, (iii) options that share the same maturity and expiration date and (iv) contract specific parameters. While the contract specific parameters yielded the best results in combination with a (hybrid) Black-Scholes based ANN model, the difference with the monyness-maturity approach, hence applied in this study, is negligible small. By this division, additional complexity is allowed to cope for the moneyness-maturity related biases as reported in many studies. In addition, this requires less computational time as with the contract specific approach.

$\{\sigma, \kappa, \nu, \varpi, \rho\}$. The SSE is minimized by means of nonlinear least squares based on the interior-point algorithm. To improve the calibration speed, I incorporate constraints for the spot volatility such that $0 < \sigma < 1$. To avoid obtaining local minima as much as possible I adapt various checks. That is, first I use three different starting values for all the unknown parameters; two based on reported averages in corresponding papers and one corresponding the obtained optimal parameters from the previous day $t-1$ with the same moneyness-maturity category p. Then I check for nearby optima by adding and subtracting a small number $\triangle$ to the converged parameter values. This is repeated iteratively for every point in time and for each of the twenty four categories.

5 COMBINING THE MODELS

Given the estimated parameters and option prices, this section describes a couple of ways to utilize the information from the individual models. The objective in combining is to find unique weights that are assigned to the individual estimates, which in the end leads to a closer estimate to the market prices and for which the moneyness and maturity-biases are minimized. These weights also reveal the importance of a specific model, and hence can be seen as a goodness of fit.

The motivation behind combining lies in the assumption that not a single model is capable of producing the 'true' price, and that the individual performance varies over time and by the different moneyness (spot S_t divided by the strike X) and time-to-maturity T categories. Combining has proven to be an useful asset in various other fields, such as equity premium forecasting and density forecasting. For instance, a recent article of Rapach et al. (2010) shows how combining delivers statistically and economically significant out-of-sample gains in equity premium forecasting, relative to the historical average consistently over time. Also, (i) combining forecasts incorporates information from numerous economic variables while substantially reducing forecast volatility; (ii) combination forecasts are linked to the real economy. Despite the increasing popularity of forecast combination in economics and finance, applications in the option pricing literature are relatively rare. One of the few exceptions is a recent article of Andreou et al. (2008) who utilize the information of the $MER1$ (Merton, 1973) and the CSU (Corrado and Su, 1996) framework by means of Neural Networks.

Another important aspect of combining individual models is the meaning of arbitrage-free. While the option pricing formulas for the individual models are obtained using a no-arbitrage condition, the combination of no-arbitrage models should not necessarily imply an arbitrage free combined framework. Some examples of arbitrage conditions that limit the degrees of freedom for a trader include are the put-call parity and the rule that bid-ask spread rates are not zero. I leave the implications of these conditions in a combining framework open for further investigation.

In the first part of this section I discuss parametric approaches, inspired by Rapach et al. (2010) and Stock and Watson (2004) in equity premium forecasting. In the second part I discuss a Artificial Neural Network approach that is expected to allow for higher complexity, following closely Andreou et al. (2008). This approach is therefore considered as a benchmark to the proposed parametric methodology which is a more pragmatic way of combining.

5.1 Parametric methods

For every individual model $k = 1,...,M$, at every point in time t there are multiple estimates of options prices with different strike and expiration date. To obtain the combined option prices, I first split the sample during the relevant time period into one of the eight moneyness and three time-to-maturity categories. This allows for flexible weights and dynamic behavior between the moneyness-maturity categories, resulting in twenty-four collections of option prices at every point in time t. From section 4 we know that $c_t^k = \left\{c_{t,p}^k\right\}_{p=1}^{24}$ is a collection of estimated option prices by individual model k corresponding to one of the twenty-four categories p.

In calculated the combined option price, again a distinction is made between out-of-sample $c_{t+1}^{comb,\text{oos}}$ and in-sample $c_t^{comb,\text{is}}$ combined option prices. For the out-of-sample combined option prices I use the forecasts of the option prices, hence using parameter estimates obtained with all data up to time t and then forecast the next day's $t+1$ option prices $c_{t+1}^{k,\text{oos}}$. In contrast, for in-sample combined option prices I use the fitted values $c_t^{k,\text{is}}$ of the option prices at time t. More specific, the combined option price in an in-sample framework is given by:

$$c_{t,p}^{comb,\text{is}} = \sum_{k=1}^{M} w_{t,p}^{k,\text{is}} c_{t,p}^{k,\text{is}} \quad \text{for} \quad p = 1,...,24. \tag{5.1}$$

The combined option price in an out-of-sample framework is given by:

$$c_{t+1,p}^{comb,\text{oos}} = \sum_{k=1}^{M} w_{t,p}^{k,\text{oos}} c_{t+1,p}^{k,\text{oos}} \quad \text{for} \quad p = 1,...,24. \tag{5.2}$$

In estimating the weights $w_{t,p}^{k,\Upsilon}$, for both approaches only the out-of-sample or in-sample prices of the past m days up to time t are used as given. Define $\Upsilon = \{\text{is},\text{oos}\}$. Then I evaluate the performance of the individual models during the hold-out period $t-m,...,t$ based on the discounted mean squared error (DMSE). Moreover, the weights are obtained by solving:

$$w_{t,p}^{k,\Upsilon} = \frac{\left(\phi_{t,p}^{k,\Upsilon}\right)^{-1}}{\sum_{l=1}^{M}\left(\phi_{t,p}^{l,\Upsilon}\right)^{-1}}, \quad \text{for} \ \ \Upsilon = \{\text{is}, \text{oos}\}, \tag{5.3}$$

where

$$\phi_{t,p}^{k,\Upsilon} = \sum_{s=0}^{m} \xi^{t-s}\left(c_{t,p}^{mkt} - c_{t,p}^{k,\Upsilon}\right)^2, \quad \text{for} \ \ \Upsilon = \{\text{is}, \text{oos}\}. \tag{5.4}$$

with ξ being the discounting factor and $c_{t,p}^{mkt}$ is the collection of true observed market price for the option with moneyness-maturity category p. Here a hold-out period is used of 10 past days, like for the estimation of the implied parameters. This method assigns higher weights to models with lower DMSE during the hold-out period. For $\xi < 1$ more importance is attached to the most recent observations and for $\xi = 1$ there is no discounting. In this study I set the discount factor to 0.9. The second class of weights includes averaging over the individual models (hence every model has the same weight) and taking the median of all the individual estimates. Moreover, for the mean combined option price the weights are given by $w_{t,p}^{k,\Upsilon} = 1/M$ for $k = 1, \ldots, M$ and the median combined option price is the median of $\left\{c_{t,p}^{k,\Upsilon}\right\}_{k=1}^{M}$.

There are numerous other ways to combine the candidate option prices. Geweke and Amisano (2011) for instance use a log scoring rule for predictive distributions to impose weights on individual candidates. These can on their turn again be based on the historical performance or fixed. Another idea is to estimate the weights and parameters *simultaneously* or to allow weights to switch between various regimes (Waggoner and Zha, 2012). Still, while the extended methods may allow for additional complexities and better pricing performance, linear pooling approaches remains a competitive substitute (see e.g., Timmermann (2006)). These avenues are open for further investigation.

5.2 Non-parametric methods

In this study I also pay attention to non-parametric approaches in order to provide a competitive framework against, the more simplistic, combining parametric models. By doing so, we can question whether the additional complexity of non-parametric approaches has significant benefits, and whether the models are a complement or substitute in the existing option pricing literature. Moreover, the non-parametric approaches, such as Artificial Neural Networks (ANN), may have their disadvantages such as higher computational time and be more difficult to use in practice. On the other hand, they have proven to be robust during rapidly changing market environments, making them useful during volatile periods such as the latest financial crisis. This section discusses the considered models in brief, with the choices and assumptions made.

In a similar fashion as in Andreou et al. (2008), I implement the Neural Network model with as inputs the option prices from the CSU and $MER1$ models and target primarily the option market price c_t^{mkt}. The notation for this model is given by $ANN2$. In this way, the NN model is a function of parametric and non-parametric methodology. Moreover, $c_t^{ANN2} = f_2(c_t^{CSU}, c_t^{MER1})$ where $f_2(.)$ is the network that links information from the parametric models with the prices c_t^{CSU} Corrado and Su (1996) and c_t^{MER1} Merton (1973). As an additional benchmark I also include a Neural Network model that combines all the individual models considered ($ANNa$) through the neural network $f_a(.)$. Finally, I also model a conventional Artificial Neural Network option pricing model that uses as inputs the variables spot price S_t, strike X, dividend δ, zero-coupon interest rate r_f and time-to-maturity T, with notation $ANNx$. Hence, $c_t^{ANNx} = f_x(S_t, X, R_f, T, \delta)$. The targets for all the ANN models considered are the true and observed market prices, c_t^{mkt}. Andreou et al. (2008) considers two alternative target functions: (i) the option price divided by the strike price c_t^{mkt}/X; (ii) the residual between the actual call market price and the individual model based option price both divided by strike $c_t^{mkt}/X - c_t^k/X$. While this may lead to improved pricing performance, in this study however this would lead to unfair comparison between the parametric and non-parametric methods. Moreover, for the parametric methods the weights are only obtained by taking into account the residuals between the option market price and model-based price estimates. An avenue open for further investigation may be ways to obtain weights that also incorporate other information from the option contracts, such as the strike and time-to-expiration.

In implementing and estimating the neural networks, I again split the sample into 24 categories based on the moneyness and maturity of the option prices. Each network is estimated and optimized by the Mean Square Error (MSE) criterion, with a maximum of 200 iterations. Like in most cases in modeling an ANN model, the in-sample dataset is split into three subsets. First, the training or estimation set. Second the validation set where the optimal number of neurons and the weights are determined[11]. Third a testing set where the pricing performance is monitored. Finally, this calibrated neural network is used as as the resulting framework for pricing the option prices in-sample, at time t or out-of-sample, at $t+1$, with as inputs the corresponding prices or variables at that time. To train the ANN's, I utilize the modified LM algorithm. For a detailed discussion about the technical details and implementation of the NN models see A.2.

[11]For computational reasons I set the number of hidden layers equal to 10 (which is the default). This may be an avenue open for improvement as selecting the optimal number of neurons may be of significant importance; selecting too much leads to over-fitting yielding an unidentified model. See Balkin and Ord (2006) and Teräsvirta et al. (2006) for ways to determine the optimal number of hidden units.

6 IMPLIED PARAMETERS AND IN-SAMPLE PRICING FIT

The purpose of this section is to obtain a clear picture of how generalizing BLS can lead to improved performance, in terms of pricing error reduction, and to motivate the use of combining. Additional attention is given to the related moneyness- and maturity biases and the implications of model misspecification. Table 3 shows the weighted averages of the obtained parameters during the full period (2006-2011). Presented are the daily average and standard deviation of the estimated parameters with the daily root mean squared of in-sample pricing errors (RMSE). These statistics inform about the internal working of the individual models.

From the table, we can observe multiple features. The implied volatilities σ of BLS, COX and HES and the volatilities of the Merton models $MER1$ and $MER2$ seem to be relatively close to each other. It is interesting to see what impact this has on out-of-sample pricing and trading as small differences in implied volatility estimates can lead to significantly different performance. CSU is the only model with a remarkably higher implied volatility than the others. This could be due to a different influence of this parameter as opposed to the skewness μ_3 and kurtosis μ_4 parameters. Moreover, we observe an implied positive skewness of 0.46 and a lacking kurtosis of 2.46. This differs to some extent from the skewness and kurtosis of the spot rate during the time period 2005-2008, which is equivalent to -0.3607 and 2.605 respectively. The annualized jump frequency λ of the $MER2$ jump diffusion model is around 0.38 times a year, with an average jump size of -0.05 and jump size uncertainty 0.23. Allowing price jumps to occur can help to cope with more negative skewness and higher kurtosis without affecting the other parameters. The mean reversion parameter κ of the Heston model is about 17.46 (3.30). This parameter can be seen as the degree of volatility clustering. Hence, an increasing reversion parameter flattens the implied volatility smile. Yet in case of stochastic volatility models, capturing the skewness and kurtosis is mostly attributed to correlation ρ and volatility variation v parameters (Bakshi et al., 1997). A positive ρ of around 0.12 indicates a spread in the right tail and squeeze in the left tail, and therefore a right-tailed distribution, hence capturing the skewness. The volatility of the variance v has effect on the kurtosis of the distribution and is estimated to be 3.67, creating heavy tails on both sides. A higher v indicates a higher prominence of the volatility smile. In addition it reflects a higher probability of extreme movements.

Figure 1 shows the time-variation of the backed-out implied volatility parameters against the VIX index on the S&P500. The values are the average across all maturity-moneyness groups for each day in the sample. We observe for all models with exception of the CSU framework that the pattern is similar to that of the the VIX index.

Table 3: Implied parameters and in-sample fit

Each day in the sample (2006-2011), the structural parameters of a given model are estimated by minimizing the sum of squared pricing errors between the market price and the model-determined price for each option. The daily average of the estimated parameters is reported first, followed by its standard deviation in parentheses. The parameters are estimated by splitting the in-sample data into twenty-four categories, based on the moneyness (eight categories) and expiration date (three categories) of all the available options up to current day. The values presented are the weighted averages across the twenty-four categories and over time. σ is the implied volatility, μ_3 and μ_4 respectively the skewness and kurtosis of the underlying, λ the jump-frequency, α and ϑ the mean and standard deviation of the jump, κ the mean-reversion parameter φ the long run variance, ν the volatility of the volatility and ρ the correlation between the log-returns and volatility of the asset. BLS refers to the benchmark Black and Scholes (1973), MER1 to Merton (1973), CSU to Corrado and Su (1996), MER2 to the Merton (1976) jump diffusion, COX to Cox and Rubinstein (1979) and HES to the Heston (1993) stochastic volatility model. RMSE is the average daily Root Mean Squared Error of the relevant option prices during the sample period. T represent the time-to-maturity, hence indicating days to expiration date. The parameters in the groups under 'All Options', 'Short-Term Options', and 'At-The-Money Options' are obtained by respectively averaging across all moneyness-maturities, averaging across only options with maturity less than 60 days, and averaging across only options with moneyness 0.99-1.01 as input into the estimation

	All Options						**Short-Term Options**						**At-The-Money Options**					
Parameters	BLS	MER1	CSU	MER2	COX	HES	BLS	MER1	CSU	MER2	COX	HES	BLS	MER1	CSU	MER2	COX	HES
σ	0.13	0.18	0.34	0.17	0.14	0.14	0.14	0.18	0.34	0.18	0.15	0.10	0.15	0.18	0.38	0.17	0.17	0.13
	(0.08)	(0.08)	(0.05)	(0.08)	(0.08)	(0.07)	(0.09)	(0.10)	(0.08)	(0.10)	(0.09)	(0.07)	(0.08)	(0.07)	(0.12)	(0.10)	(0.08)	0.10
μ_3			0.46						-0.02						1.07			
			(0.30)						(0.25)						(0.56)			
$\mu4$			2.46						2.24						2.42			
			(0.37)						(0.41)						(1.11)			
λ				0.38						0.70						0.31		
				(0.22)						(0.51)						(0.41)		
α				-0.05						-0.02						-0.08		
				(0.11)						(0.10)						(0.10)		
ϑ				0.23						0.26						0.25		
				(0.03)						(0.06)						(0.07)		
κ						17.46						17.99						21.86
						(3.30)						(3.38)						(2.54)
ϖ						0.05						0.04						0.02
						(0.04)						(0.04)						(0.02)
ν						3.66						3.78						3.36
						(1.72)						(1.73)						(2.63)
ρ						0.12						0.07						0.05
						(0.19)						(0.19)						(0.16)
RMSE	7.43	4.66	5.17	4.05	7.74	8.27	3.01	2.49	2.84	2.49	3.21	4.31	5.79	4.04	4.49	5.01	5.73	7.18

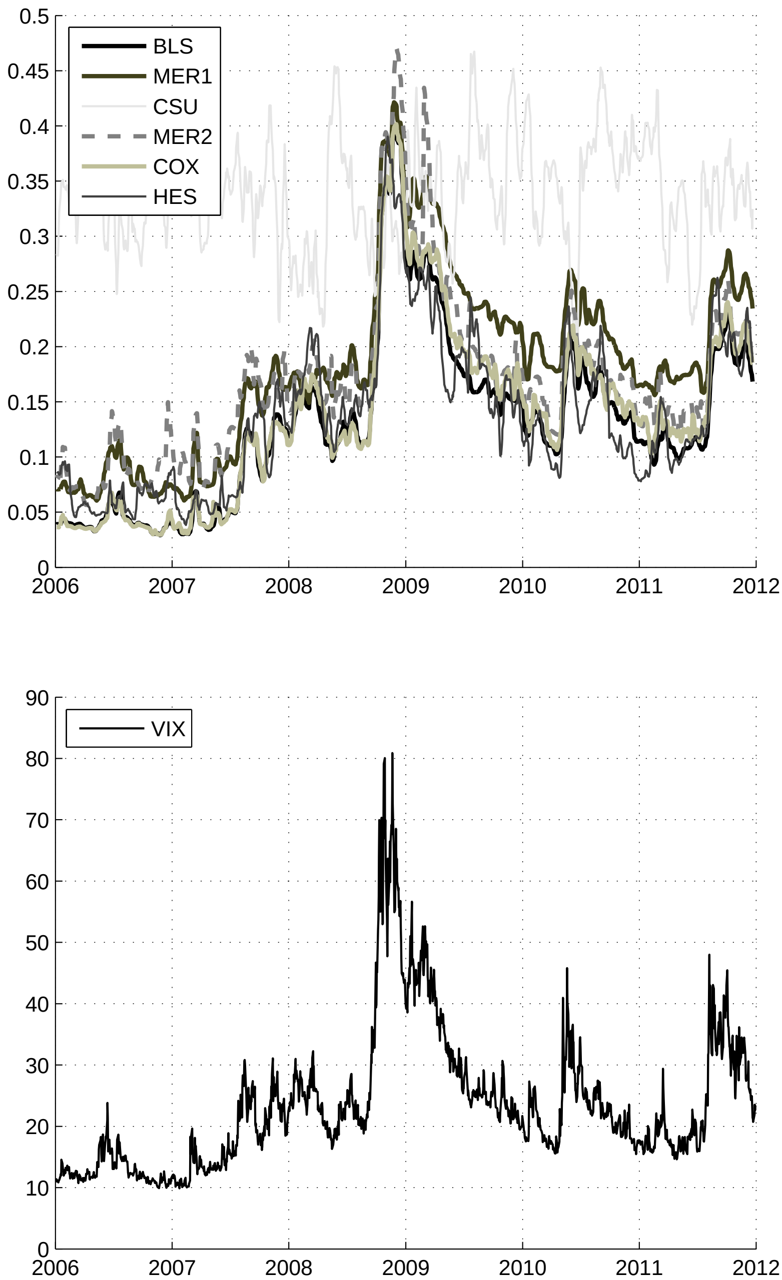
This lack of variation of the CSU method is probably compensated by the variation in the other parameters skewness μ_3 and kurtosis μ_4. The implications of these results are that the time-variation of the volatility seems to be captured for a large part by the estimated implied volatility parameters. For completeness, Table B.1 in the Appendix shows the implied volatilities averaged across time by the various moneyness-maturity categories and during different subperiods. The backed out volatilities show the typical strong U-shaped pattern or smile; when the call option goes from in-the-money to at-the-money and then again to out-to-money. This is the case for all time-to-expirations and the different sub-periods. This smile is the strongest for the short-term options which could be an indication that models are the most misspriced for this category. The moneyness- and maturity-related biases give an argument for relaxing the BLS assumptions and considering alternative option pricing models. As Bakshi et al. (1997) notes, to cope for the volatility smile, the better model should allow for negative skewness and excess kurtosis as the presence of a smile is an indication of negatively skewed returns with excess kurtosis.

From the pricing performance of the individual models, we see that especially Merton models $MER1$ and $MER2$ yield relatively small RMSE for all option contracts of 4.66 and 4.05 respectively. The generalization of the assumptions no-dividends and continuous diffusion of the underlying seems therefore to lead to better in-sample pricing performance. The result of the Heston stochastic volatility model is somewhat surprising with a RMSE of 8.27, being the worst performer. While this is mostly attributed to bad pricing performance for out-of-the money option contacts (see table B.2 in the Appendix), this holds true for other moneyness-maturity categories as well. It may imply that stochastic volatility plays a smaller role during the considered time period. Yet this does not seem plausible given the high volatile markets during 2007-2009. On the other hand, a high degree of volatility clustering could explain why assuming a constant volatility during a particular period of time may seem a reasonable assumption. Another explanation is the presence of calibration issues, as addressed by various papers (see e.g., Albrecher et al. (2007)). Finally, relaxing the assumption of continuous evolution of the share price doesn't seem to lead to improvement in in-sample pricing as well, as we can see from the pricing results of the COX model with a RMSE of 7.74.

As for the combining frameworks, it seems that only $cMSE$ which combines based on the historical performance yields a reduction in the RMSE for most moneyness and maturity categories. $cMEAN$, $cMED$ and the Neural Network model $cANN2$ fail to be of significant influence. As for $cMEAN$ and $cMED$ it may be that the individual forecasts are too much correlated (around 0.95-0.99) such that no significant benefits are achieved from combining. Hence, $cMSE$ takes into account possible biases due to model

Figure 1: Implied volatilities over time against the VIX rate

Presented are the average (across all moneyness-maturity categories) implied volatilities backed out by the various individual models against the VIX index. Each day in the sample (2006-2011), the structural parameters of a given model are estimated by minimizing the sum of squared pricing errors between the market price and the model-determined price for each option. The parameters are estimated by splitting the in-sample data into twenty-four categories, based on the moneyness (eight categories) and expiration date (three categories) of all the available options up to current day. The values presented are the weighted averages across the twenty-four categories and over time. BLS refers to the benchmark Black and Scholes (1973), MER1 to Merton (1973), CSU to Corrado and Su (1996), MER2 to the Merton (1976) jump diffusion, COX to Cox and Rubinstein (1979) and HES to the Heston (1993) stochastic volatility model.



misspecifications through the weights. As for $cANN2$, the Neural Networks may not be well calibrated. For instance, the number of neurons is not chosen properly. A closer look at the RMSE of this model over time (not shown here) reveals that in particular after 2010 the model produces large in-sample pricing errors. Hence, a structural break affecting the networks may require an additional calibration effort after a particular point in time.

Finally I take a closer look at the parameter estimates by averaging only across short-term, which are challenging to price, and at-the-money options, which are considered in many studies. From the table we can observe that the parameters are different among the various options with respect to maturity and moneyness. This holds especially true for the jump-intensity λ in $MER2$ and mean-reversion rate κ in HES. For the HES model we observe that the jump-intensity (jump-size) is the highest (lowest) for short-term-options. This implies that shocks are more likely to have an effect on short-term options, while the jumps have a smaller size and volatility. The volatility coefficient of short-term options is higher for BLS, $MER2$ and COX, implying that for the short-term options to be priced correctly by means of these models, the volatility needs to be more volatile than for all options of all maturities. This holds true for all the models estimated for at-the-money options with exception of HES. From the CSU parameters, we can see that short-term options are skewed less to none (-0.02), but that the at-the-money options exhibit a larger skew (1.07).

In all, these findings suggest that the candidate option pricing models are probably not entirely correctly specified. Moreover, if it were the case, the 24 categories of option prices split by the moneyness-maturity categories, should not have resulted in different parameter estimates. Thus, this gives an argument for combining.

7 OUT-OF-SAMPLE PRICING

Results on the in-sample pricing performance show that relaxing the continuous diffusion of the underlying and no dividend assumptions yields the highest reduction in pricing errors. Also, combining based on the historical performance yields an improved fit. The increased fit however may be the consequence of having more parameters. In an out-of-sample framework on the other hand, having more parameters may lead to overfitting and hence a higher RMSE if the additional parameters does not have significant added value. I therefore next asses the performance of the methods in an out-of-sample framework, where the previous day's implied parameters and volatilities values as an input to compute the current days's option pricing models.

7.1 Pricing errors

Table 4 reports the RMSE of the various individual option pricing models, along with the combined framework. I again distinguish between the various moneyness and maturity categories in the presentation of the results. The relative performance of the models differs among the various moneyness-maturity categories, especially for maturities longer than 180 days. This can be for a part explained by the larger pricing errors that occur when the option prices are higher, as it is the case for longer maturities. Among the individual models, we can observe the same performance characteristics as for the in-sample results. That is, the $MER1$, $MER2$ and CSU models are among the best performers for most moneyness and maturity categories, whereas COX and HES fail to beat the BLS model. Non-dividend paying methods seem to be therefore in a disadvantage. The best performer is the $MER2$ jump diffusion model with an average RMSE of 4.37 over the whole dataset. These results imply that the increased fit for the individual models in an in-sample framework is in line with the performance of these methods in an out-of-sample framework. In addition, all models are capable of beating the random walk[12] (RW).

Combining based on the discounted historical performance ($cMSE$) is the overall best performer with a RMSE of 4.26 for the full dataset. This is consistent among almost all moneyness and maturity categories. Taking the median instead of the mean yields better results, but does not beat the $cMSE$. This is in contrast with the in-sample results where combining based on these schemes didn't had any added value. The NN ($cANN2$) that utilizes information from CSU and $MER1$ models performs worse compared to $cMSE$, whereas taking into account information from all the models ($cANNa$) leads to an improvement with a RSME of 7.68. Also, the 'two-step' approach similar as for Andreou et al. (2008) leads to better results as opposed to a pure Artificial NN method ($ANNx$). This implies that more benefits can be achieved and higher complexity captured, by considering a two-step approach that individually utilizes information from different frameworks instead of directly modeling the option prices by means of a Neural Network or complex extended framework. This can be somewhat compared to what we observe in the literature of equity premium forecasting. Timmermann (2006) and Rapach et al. (2010) show that such two-step way of modeling or forecasting, by utilizing information from individual models with one regressor, leads to better out-of-sample performance as opposed to one 'super' model with multiple regressors.

[12]One may suspect that the bad performance of the Random Walk (RW) is due to the unbalanced samples. More specific, when using previous day's option prices as an estimate for current day's there are no forecasts available for option contracts that are added at that specific day. In a separate analysis I match the sample of the RW forecasts with the other models and find that the results are qualitatively similar.

Table 4: Out-of-sample pricing performance of all the models (2006-2011)

Presented are the pricing errors of the various option pricing models considered in this study, in an out-of-sample framework. I compute the current's day option prices using the previous day's implied parameters and implied stock volatility. For each moneyness-maturity category, I report the Root Mean Squared Error (RMSE) of the difference between market price and model price for all option contacts. The sample period covers January 2006 - December 2011, with a total of 893708 observations. A total of 7885 observations are discarded from the analysis due to calibration issues or an non-feasible estimates. The individual models considered are the BLS (Black and Scholes, 1973), MER1 (Merton, 1973), CSU (Corrado and Su, 1996), MER2 (Merton, 1976), COX (Cox and Rubinstein, 1979) and HES Heston (1993). cMSE is the combined model which combines the individual option prices of all models, by assigning weights to each individual model based on the historical performance, measured by the Discounted Mean Squared Error (DMSE), during a hold-out period of 10 days. cMEAN and cMED refer to respectively the average and median of all model based estimates of the individual option pricing models on a given day. cANN2 is the neural network model, following Andreou et al. (2008) which utilizes information from the MER1 and CSU framework by means of an Artificial Neural Network. cANNa uses the same methodology as ANN2 but utilizes information from all individual models. cANNx is the conventional Artificial Neural Network model which uses the inputs of the BLS and targets the true market price. RW refers to the Random Walk option pricing model, which uses the past day true market price as an estimate for the current's day, for all contracts. Underlined values represent the lowest value for a given row/moneyness category.

Moneyness						Days to maturity: <60 days								
S_t/X	Obs.	BLS	MER1	CSU	MER2	COX	HES	cMSE	cMEAN	cMED	cANN2	cANNa	cANNx	RW
< 0.85	7624	8.99	7.00	7.93	7.40	9.42	12.98	7.30	8.20	7.65	7.32	6.63	9.48	17.46
0.85 - 0.95	33461	9.42	5.88	6.45	5.11	9.55	9.86	4.99	6.35	5.75	8.41	7.58	9.59	12.51
0.95 - 0.99	29312	9.86	5.15	5.79	4.31	9.95	9.19	4.24	6.57	5.24	7.74	7.02	9.76	10.85
0.99 - 1.01	14939	10.06	5.03	5.62	4.32	10.15	9.09	4.06	6.22	4.99	7.81	6.87	9.65	10.45
1.01 - 1.05	27219	9.78	5.08	5.63	4.15	9.88	9.28	4.15	6.16	4.99	7.78	7.01	9.66	10.40
1.05 - 1.10	30009	9.93	5.26	5.68	4.17	9.95	9.15	4.20	6.41	5.16	7.81	7.13	9.84	10.45
1.10 - 1.35	88190	10.04	5.63	6.05	4.38	10.12	9.53	4.33	6.06	5.27	8.33	7.35	9.70	10.69
> 1.35	74601	9.81	5.97	5.81	4.25	9.98	9.23	4.16	5.56	5.06	9.67	8.57	9.48	10.52
All	**305355**	**9.84**	**5.62**	**5.96**	**4.45**	**9.95**	**9.46**	**4.39**	**6.12**	**5.28**	**8.46**	**7.55**	**9.64**	**10.97**
Moneyness						Days to maturity: 60-180 days								
S_t/X	Obs.	BLS	MER1	CSU	MER2	COX	HES	cMSE	cMEAN	cMED	cANN2	cANNa	cANNx	RW
< 0.85	21932	8.16	5.98	6.57	5.58	8.48	10.27	5.56	6.32	6.04	8.23	7.32	9.65	15.07
0.85 - 0.95	42782	9.22	5.43	5.88	4.44	9.36	9.12	4.32	5.86	5.08	8.79	7.71	9.68	11.27
0.95 - 0.99	21363	9.66	4.93	5.64	4.25	9.73	8.87	4.06	6.07	4.91	8.33	7.38	9.60	10.63
0.99 - 1.01	10642	10.18	4.97	5.88	4.20	10.26	9.28	3.96	6.40	4.88	8.34	7.28	9.52	10.58
1.01 - 1.05	18957	10.13	5.02	6.00	4.22	10.14	8.98	4.04	6.22	4.97	8.13	7.32	9.64	10.35
1.05 - 1.10	19321	10.09	5.22	5.88	4.30	10.17	9.13	4.04	6.17	5.03	8.74	7.44	9.45	10.53
1.10 - 1.35	57432	9.71	5.25	5.66	4.18	9.80	8.97	4.03	6.14	5.11	8.87	7.83	9.49	10.95
> 1.35	65051	9.21	5.56	5.52	4.16	9.37	8.92	3.97	5.53	4.94	9.70	8.52	9.31	10.53
All	**257480**	**9.45**	**5.36**	**5.79**	**4.35**	**9.57**	**9.11**	**4.19**	**5.97**	**5.10**	**8.88**	**7.81**	**9.51**	**11.13**
Moneyness						Days to maturity: > 180 days								
S_t/X	Obs.	BLS	MER1	CSU	MER2	COX	HES	cMSE	cMEAN	cMED	cANN2	cANNa	cANNx	RW
< 0.85	71432	8.73	5.60	5.96	4.86	8.90	9.67	4.76	6.06	5.47	8.55	7.58	9.47	13.39
0.85 - 0.95	43920	10.16	5.16	5.58	4.33	10.19	9.27	4.24	6.55	5.06	8.23	7.45	9.72	11.13
0.95 - 0.99	17771	10.61	5.03	5.74	4.36	10.60	9.41	4.16	6.62	4.95	8.01	7.15	9.82	10.66
0.99 - 1.01	9127	10.30	5.02	5.70	4.32	10.34	9.59	4.10	6.89	5.06	8.00	7.12	10.02	10.67
1.01 - 1.05	16854	10.66	4.83	5.62	4.18	10.63	9.38	3.97	6.75	4.92	7.79	6.96	9.81	10.37
1.05 - 1.10	19052	11.10	5.06	5.64	4.29	11.07	9.51	4.13	7.13	5.01	7.94	7.13	9.70	10.43
1.10 - 1.35	59553	10.50	5.03	5.54	4.17	10.53	9.22	3.97	6.68	4.90	8.37	7.41	9.59	10.40
> 1.35	93164	9.81	5.39	5.40	3.98	9.89	8.90	3.99	5.98	4.85	9.50	8.53	9.49	10.11
All	**330873**	**9.92**	**5.26**	**5.62**	**4.31**	**9.99**	**9.28**	**4.21**	**6.36**	**5.04**	**8.62**	**7.71**	**9.60**	**11.08**
Overall	**893708**	**9.76**	**5.41**	**5.79**	**4.37**	**9.86**	**9.29**	**4.26**	**6.17**	**5.14**	**8.64**	**7.68**	**9.59**	**11.06**

The out-of-sample results of the combining frameworks are not entirely in line with the correlations between the individual models. These are relatively high (between 0.95 and 0.99); combining highly collinear forecasts would not per se give large reduction in the Mean Squared Error. Yet, we still see benefits arising from combining the individual frameworks. As in the case for the in-sample results, an explanation could be that the benefits of combining arise due to corrections for the misspecification of the individual models. If each model is to some extent misspecified, then by definition each price estimate is consistently biased. These biases are then corrected by the combining frameworks through the weights. Another possible reason could be that for the option pricing literature, the values of the correlations are not that informative as in the case for equity premium forecasting. Moreover, the option pricing forecasts represent the absolute value and not the returns, which are for their part more volatile and show higher dispersion.

Table 5 reports the Diebold-Mariano (DM) test statistics, here used to compare the pricing performance of the various models. In case the test-statistic follows a standard normal distribution, a positive value in the table larger than 1.645 (2.325) would imply that the model in the vertical heading has a larger pricing error than the model in the horizontal heading at a 5% (1%) significance level. However, this is not a valid assumption since this study compares linear and nonlinear nested models[13]. A way to solve for this concern is by bootstrapping the DM-statistic. However, due to the large number of parameters, variables and hence combinations possible, this is not an easy task to overcome due to the large computational time. Another alternative would be to consider the Model Confidence Set (MCS) procedure of Hansen et al. (2011). This method handles the limitations of the data by not assuming that a particular model is the true model. Instead, it is constructed such that it yields a best model with a given level of confidence. Still, the DM test statistic values presented can give some indication of the direction and to some extent the relative magnitude of the individual performance. For instance, we can observe that the $cMSE$ model outperforms all the other models as expected with a large DM-statistic. Exceptions are the DM-statistics against the $MER2$ and $cMED$ models, meaning that these three models yield close results. While we don't know the true distribution of the statistic, the values reveal to some extent how close these models are in terms of pricing performance during the considered time period.

[13]Comparing errors from nested models, as it is the case in this study, leads to tests with low power and a undersized statistic (McCracken (2007) and Clark and West (2007)). For this reason Clark and West (2007) develop an adjusted version of the Diebold and Mariano (1995) and West (1996) statistic that is line with the standard normal distribution and generates asymptotically valid inferences for comparing forecasts from nested linear models. This approach however is not valid as well, given that we combine nested nonlinear and linear models.

Table 5: Diebold-Mariano test for comparing pricing errors

Reported are the Diebold and Mariano (1995) test-statistics for comparing the pricing errors of the various models considered. Under the null hypothesis the expected squared errors of two competitive models are the same. Hence, a positive DM-value means that the model in the vertical heading has a larger pricing error than the model in the horizontal heading. Since we do not exclusively compare linear and non-nested models, the assumption of the statistic following a standard normal distribution is not valid. Therefore inferences made about the statistical significance of the values are dubious without first knowing the true distribution. The random walk (RW) method is discarded from this analysis due to an un-matching number of observations as opposed to the other methods.

	BLS	MER1	CSU	MER2	COX	HES	cMSE	cMEAN	cMED	cANN2	cANNall	cANNx
BLS		221.7	199.2	251.6	-30.0	24.9	253.9	196.9	236.0	47.4	87.3	8.2
MER1			-34.9	98.6	-234.6	-221.6	120.4	-61.2	26.4	-154.2	-107.7	-214.4
CSU				123.4	-211.2	-203.4	139.7	-30.2	57.2	-140.0	-91.3	-196.9
MER2					-264.9	-265.8	13.9	-155.2	-82.3	-195.7	-151.0	-259.0
COX						30.8	267.8	208.9	249.8	52.7	93.5	13.0
HES							269.6	183.2	243.2	28.8	72.7	-14.6
cMSE								-180.6	-109.8	-200.7	-155.1	-263.4
cMEAN									126.7	-122.1	-74.5	-181.5
cMED										-168.0	-120.8	-229.0
cANN2											48.0	-42.8
cANNall												-86.6
cANNx												

7.2 Sub-periods

The data set considered spans some remarkable years with major events such as the latest financial crisis. Due to the large fluctuations in the market and multiple events happening, the behavior and consistency of the various models may differ during different sub-periods. I therefore next asses the absolute performance of the models considered during three sub-periods: (i) before the latest recession, (ii) during the recession and (iii) after the recession. The start and end dates of the latest recession are obtained from the NBER website which also reports the historical business cycles.

The weighted RMSE averages, across all moneyness-maturity categories, of all models during the different sub-periods are given in table 6. Now we observe that while the two-step neural network models ($cANN2$ and $cANNa$) are not of significant meaning during the overall period, they are capable of outperforming most models during and before the recession. This result may be somewhat explained by the characteristic of the approach which is designed to cope for high complexity, which can arise during volatile periods such as a crisis. On the other hand, the pre-recession results suggest that two-step ANN's were also of influence during an expansion. The poor results after the recession may indicate a changed market behavior for which the two-step ANN's haven't adjusted rapidly enough or certain behavior that needs more attention in modeling. Like we observe for the overall period, including more individual models in the two-step ANN framework yields better results ($cANNa$ performs better than $cANN2$). The most striking result however is the simple parametric combining framework $cMSE$ shows consistent behavior across the various sub-periods. It consistently outperforms

Table 6: Pricing performance during sub-periods

Presented are the pricing errors of the various option pricing models considered in this study, in an out-of-sample framework. That is, I compute the current's day option prices using the previous day's implied parameters and implied stock volatility. I report the Root Mean Squared Error (RMSE) of the difference between market price and model price for all option contacts. The overall sample period covers January 2006 - December 2011, with a total of 893708 observations. The recession start and end date are obtained from the NBER website and indicate the latest recession, characterized by the recent financial crisis, which spans December 2007 - June 2009. Obs. stands for the number of observations.

Obs.	BLS	MER1	CSU	MER2	COX	HES	cMSE	cMEAN	cMED	cANN2	cANNa	cANNx	RW
Overall 893708	9.76	5.41	5.79	4.37	9.86	9.29	4.26	6.17	5.14	8.64	7.68	9.59	11.06
Recession 224071	7.85	5.15	6.20	5.13	8.15	9.78	4.89	5.92	5.66	4.83	3.94	9.47	13.30
Pre-recession 166910	13.45	3.83	5.70	4.56	13.24	9.99	3.84	8.68	4.74	3.65	2.88	10.07	8.68
Post-recession 500249	9.05	6.01	5.67	4.02	9.22	8.85	4.11	5.17	5.02	10.91	9.80	9.49	10.70

most models and shows also good results after the crisis, unlike the two-step ANN's. For a practitioner who is seeking for a consistent and relative easy approach to model option prices, $cMSE$ may be the best alternative.

7.3 Importance individual models

The performance of the $cMSE$ model is consistent and relative stable over time. This performance however depends much on the individual inputs; the choice of the individual models and their performance. As the weights are a function of the mean squared error, inferences made about the relative importance of the individual models is expected to be in line with the relative performance of the pricing errors.

A way to asses the importance of individual models is by looking at the magnitude of the assigned weights over time. I therefore calculate the weighted average weight by averaging over all the moneyness and maturity categories for each day during the sample. The outcome of these computations is given in figure 2. As expected, the models with lower $RMSE$ receive higher weights. The weight attributed to $MER2$ is almost always higher than 20%, with an increase from the end of 2009 through 2011. The BLS model also gains more importance during the recession, hence implying that relaxing the assumptions leads to a smaller improvement in pricing during this period of time. We can also observe an increase in importance of the COX framework during the same period, while CSU is especially more prevailing pre- and post the recession. The Heston framework is the least favored with an average assigned weight of 9% over the overall period. In all, every model is included and reasonably important during the full sample.

Figure 2: Weights attributed to the individual models by $cMSE$

Presented are the weighted average weights assigned by the combining method $cMSE$. The model assigns weights to each individual model by means of the past performance during a hold-out-period of 10 days, following closely the methodology of Stock and Watson (2004). The performance is based on the Discounted Mean Squared Error (DMSE) of the pricing errors of the individual models, with a discount factor of 0.9. The average weights are obtained by averaging across all maturities and moneyness categories. The grey thin line is the actual average weight, whereas the black thicker line represents a 20 day average, included for visional reasons. Over the full period, the average weight assigned to the models are as follows: 0.13 for BLS; 0.21 for $MER1$; 0.21 for CSU; 0.24 for $MER2$; 0.12 for COX; 0.09 for HES.

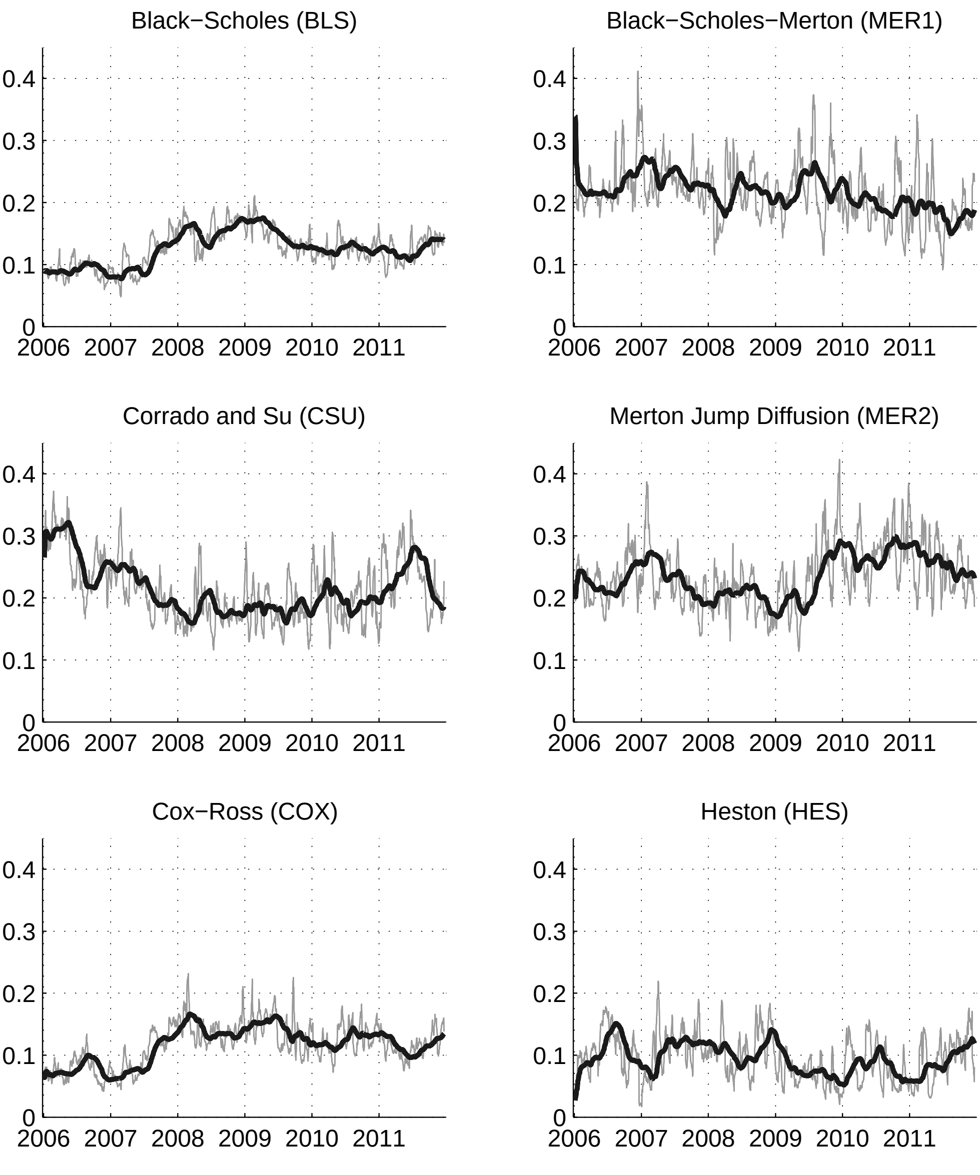


Table 7: Average weights by moneyness and maturity

Presented are the weighted average weights assigned by the combining method *cMSE* for the various moneyness and maturity categories. The model assigns weights to each individual model by means of the past performance during a hold-out-period of 10 days, following closely the methodology of Stock and Watson (2004). The performance is based on the Discounted Mean Squared Error (DMSE) of the pricing errors of the individual models, with a discount factor of 0.9. The average weights per moneyness category are obtained by averaging across all maturity categories and time; whereas the average weights per maturity category are obtained by averaging across all moneyness categories and time. The sample period covers Januari 2006 - December 2011.

	BLS	MER1	CSU	MER2	COX	HES
Moneyness						
< 0.85	<u>0.20</u>	0.18	0.19	0.17	0.14	0.12
0.85 - 0.95	<u>0.19</u>	0.19	0.18	0.16	0.18	0.09
0.95 - 0.99	0.19	<u>0.20</u>	0.18	0.17	0.18	0.07
0.99 - 1.01	0.17	<u>0.20</u>	0.18	0.15	0.17	0.12
1.01 - 1.05	0.16	<u>0.21</u>	0.18	0.17	0.17	0.11
1.05 - 1.10	0.15	<u>0.22</u>	0.19	0.18	0.15	0.11
1.10 - 1.35	0.13	0.21	0.19	<u>0.22</u>	0.13	0.10
> 1.35	0.03	0.23	0.26	<u>0.39</u>	0.03	0.06
Maturity						
< 60	0.13	0.20	0.20	<u>0.24</u>	0.13	0.09
60-180	0.13	0.20	0.22	<u>0.23</u>	0.12	0.10
> 180	0.13	0.22	0.21	<u>0.24</u>	0.11	0.09
All	0.13	0.21	0.21	<u>0.24</u>	0.12	0.09

The results also confirm that when choosing which assumptions relax, one should not neglect the characteristics of the time period.

Next I distinguish by the various moneyness and maturity categories in order to obtain a better understanding of how much the various type of options are explained by the different models. Table 7 presents the average weights, across the maturities, moneyness and time. The very deep out-the-money options are the best explained by the BLS model, hence receiving the highest weights. To explain the at-the-money options, most weight is attributed to the $MER1$ model, whereas the very deep in-the-money options are captured by the jump diffusion model $MER2$. A reason for the latter could be the magnitude of the jumps who become more prevailing for very deep in-the-money options; the jumps become larger when the option prices are higher. Across the maturities it are especially the $MER1$, CSU and $MER2$ who are receiving the highest weights, which holds true when averaging over all the categories. These results indicate that when combining, differentiating between the various moneyness categories may be more important than between the maturity categories. Hence, weight size is more dependent on the moneyness of options[14].

[14] An avenue open for further investigation may be to consider 'restricted' combination schemes, where the weights given to individual models are the across maturity and/or moneyness buckets. While this sacrifices the flexibility of the 'unrestricted' approach, it may reduce the estimation uncertainty in the

Table 8: Pricing performance of the parametric combining methods with and without Merton (1976)

Presented here are the results of the parametric combining frameworks with and without using the jump diffusion framework Merton (1976) or $MER2$ as individual input. The combining frameworks without Merton are a primarily a function of the methods BLS, CSU, $MER1$, COX and HES. To obtain the current's day option prices of the individual models I use the previous day's implied parameters and implied stock volatility. Then a combined option price is obtained by either taking the mean ($cMEAN$) or median ($cMED$), or by assigning weights to each individual model based on the performance of the individual models during the past 10 days ($cMSE$). This performance is based on the Discounted Mean Squared Errors (DMSE) with a discount factor of 0.9, and follows closely the methodology of Stock and Watson (2004). The overall sample period covers January 2006 - December 2011, with a total of 893708 observations. The recession start and end date are obtained from the NBER website and indicate the latest recession, characterized by the recent financial crisis, which spans December 2007 - June 2009. Obs. stands for the number of observations.

		Without Merton			**With Merton**		
	Obs.	cMSE*	cMEAN*	cMED*	cMSE	cMEAN	cMED
Overall	893708	4.70	6.96	6.31	<u>4.26</u>	6.17	5.14
Recession	224071	5.07	6.33	6.29	<u>4.89</u>	5.92	5.66
Pre-recession	166910	3.96	10.26	6.00	<u>3.84</u>	8.68	4.74
Post-recession	500249	4.75	5.77	6.43	<u>4.11</u>	5.17	5.02

The relative high weight assigned to the $MER2$ model indicates that much of the performance of the combining framework is attributed to this individual model. For this reason it is difficult to asses whether the performance of the proposed combining framework is the result of the good performance of the $MER2$ model or the result of combining. To differentiate between these two implications, I re-estimate the parametric combining frameworks without using the $MER2$ model as one of the individual inputs, and evaluate the performance during the various sub-periods. The outcome of this approach is presented in table 8. We observe that the difference in RMSE is relatively small for the $cMSE$ framework. The RMSE increases slightly from 4.26 to 4.70 for the overall period, which is mostly attributed to the deteriorated post-recession performance (increase of .64). The combining frameworks $cMEAN$ and $cMED$ suffers much more from excluding the $MER2$ individual model, which is as expected as the methods do not discriminate based on the historical performance. The results indicate that the proposed combining method $cMSE$ is less sensitive to the input of the individual models, and that the consistent and good results are not primarily attributed to the good performance of one of the individual models $MER2$. Moreover, the good performance of $cMSE$ is mostly attributed to the benefits of combining and not to the performance of primarily one of the individual models.

weights and thereby benefit the forecast accuracy. I thank Dick van Dijk for this insight.

8 CONCLUSION

This study tests a pragmatic approach of combining various parametric option pricing models to price European call options on the S&P500 index, using daily data from 2006 to 2011. The individual models considered include Black and Scholes (1973) (BLS), Merton (1973) ($MER1$), Corrado and Su (1996) (CSU), Merton (1976) ($MER2$), Cox and Rubinstein (1979) (COX) and Heston (1993) (HES). Each individual model relaxes one or several assumptions of the initial BLS framework. This includes a generalization of (i) log normal distribution of stock returns; (ii) continuous trading; (iii) continuous evolution of the share price; (iv) constant variance of the underlying returns; (v) no dividends; and (vi) continuous diffusion of the underlying.

The estimates of the structural parameters reveal the typical moneyness-maturity related biases, thus giving an argument for combining. The in-sample results reveal that combining based on the discounted historical performance ($cMSE$) is the only combining framework that consistently outperforms all individual models. A possible explanation could be that a consistent bias, imposed by the misspecification of the models, is corrected by means of the assigned weights of this method. In an out-of-sample framework, where previous day's parameter estimates and implied volatilities are used to price current day's option prices, I find that combining leads to gains in terms of pricing performance during the overall period (2006-2011) for most combining methods. The relative performance of the combining models is consistent among the various moneyness-maturity categories. Combining based on the historical performance is the overall best performer and outperforms the Neural Network (NN) models. On the other hand, during and before the recession (December 2007 - June 2009), combining by means of NN's yield the lowest pricing errors.

The weights of the $cMSE$ framework reveal that BLS (no relaxations) and COX (relaxation of continuous evolution of the share price) gain importance during the recession, while CSU (relaxation of log normal distribution of stock returns) loses. In line with the pricing results, the weights assigned to $MER2$ jump diffusion model is the highest across all the moneyness and maturity categories and during most part of the sample. A closer look at the weights assigned to the individual models across the various moneyness-maturity categories, reveals that when combining, differentiating between the various moneyness categories may be more important than between the maturity categories. In other words, weight size depends on the moneyness of options. Finally, a robustness check by discarding the best performing model ($MER2$) from the combining frameworks indicates that the good performance of the $cMSE$ method is the result of combining and not only to the good performance of this method. This is verified by the weights size of this method, which is on average 24%.

The results imply that more benefits and higher complexity can be achieved by considering a two-step approach that utilizes information from different frameworks instead of considering a larger 'super' model, which is in line with the literature on equity premium forecasting. In addition, a simple pragmatic approach can already lead to lower pricing errors. It consistently outperforms most models and shows also good results after the crisis, unlike the NN's. For a practitioner who is seeking for a consistent and relative easy approach to model option prices, combining may be a competitive alternative.

While this study takes into account various approaches to utilize information as best as possible, there are still avenues open for further investigation. First one may look at the implications of combining on trading and hedging and other applications that test the economic significance. A second focus may lie in the choice of different weighting schemes and the implications of estimating the weights and implied parameters simultaneously. Determining under which circumstances combining works for option pricing may reveal important characteristics too. Third, combining can be also analyzed for other types of options, such as exotic trading instruments and foreign exchange option pricing. Fourth, attention can be given to arbitrage-free question: is a combination of arbitrage-free models still arbitrage free and what does this imply? Fifth, the study can be extended by additional robustness checks and by inclusion of additional individual models. For instance, the Heston model can be extended to allow for stochastic interest rate variation and the inclusion of dividends. The combining NN's can be also re-estimated using similar target functions as in Andreou et al. (2008), even though this does not allow for fair comparison with the parametric methods.

REFERENCES

Albrecher, H., Mayer, P., Schoutens, W., and Tistaert, J. The little Heston trap. *Wilmott Magazine*, pages 83–92, 2007.

Andreou, P., Charalambous, C., and Martzoukos, S. H. Pricing and trading European options by combining Artificial Neural Networks and parametric models with implied parameters. *European Journal of Operational Research*, 185(3):1415–1433, 2008.

Bakshi, G., Cao, C., and Chen, Z. Empirical performance of alternative options pricing models. *Journal of Finance*, 52(5):2003–2049, 1997.

Balkin, S. and Ord, J. Automatic Neural Network modeling for univariate time series. *International Journal of Forecasting*, 16:509–515, 2006.

Bates, D. The crash of '87: Was it expected? The evidence from options markets. *Journal of Finance*, 46(3):1009–1044, 1991.

Bates, D. Empirical option pricing: A retrospection. *Journal of Econometrics*, 116(1): 387–404, 2003.

Black, F. and Scholes, M. The valuation of option contracts and a test of market efficiency. *Journal of Finance*, 27(2):399–417, 1972.

Black, F. and Scholes, M. The pricing of options and corporate liabilities. *Journal of Political Economy*, 1973.

Black, F. and Scholes, M. Fact and fantasy in the use of options. *The Financial Analysts Journal*, 31:36–41 and 61–72, 1975.

Carr, P. and Madan, D. B. Option evaluation using the Fast Fourier transform. *Journal of Computational Finance*, 2(4):61–73, 1999.

Cheang, G. H. L. and Chiarella, C. A modern view on Merton's jump-diffusion model. Research paper 287, Quantitative Finance Research Centre of the University of Technology Sydney, 2011.

Christoffersen, P., Heston, S., and Jacobs, K. Option valuation with conditional skewness. *Journal of Econometrics*, 131:253–284, 2006.

Christoffersen, P., Dorion, C., Jacobs, K., and Wang, Y. Option valuation with long-run and short-run volatility components. *Journal of Financial Economics*, 90:272–297, 2008a.

Christoffersen, P., Jacobs, K., and Ornthanalai, C. *Exploring Time-Varying Jump Intensities: Evidence from S&P500 Returns and Options*. Manuscript, McGill University, 2008b.

Christoffersen, P., Dorion, C., Jacobs, K., and Wang, Y. Volatility components, affine restrictions and non-normal innovations. *Journal of Business and Economic Statistics*, 28(4):483–502, 2010.

Clark, T. E. and West, K. D. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138:291–311, 2007.

Corrado, C. J. and Su, T. Skewness and kurtosis in S&P500 index returns implied by option prices. *Journal of Financial Research*, 19(2):175–192, 1996.

Cox, J. C. and Ross, S. A. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3(1-2):145–166, 1976.

Cox, J. C., R. S. A. and Rubinstein, M. Option pricing: A simplified approach. *Journal of Financial Economics*, 7:229–263, 1979.

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

Diebold, F. X. and Mariano, R. S. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):134–143, 1995.

Geweke, J. and Amisano, G. Optimal prediction pools. *Journal of Econometrics*, Vol. 164: pp. 130–141, 2011.

Gultekin, N., Rogalski, R., and Tinic, S. Option pricing model estimates: Some empirical results. *Financial Management*, 11:58–69, 1982.

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

Hagan, M. and Menhaj, M. Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, 5(6):989–993, 1994.

Hansen, P., Lunde, A., and Nason, J. The model confidence set. *Econometrica*, 79(2): 453–497, 2011.

Heston, S. L. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993.

Heston, S. and Nandi, S. A closed-form GARCH option valuation model. *The Review of Financial Studies*, 13:585–625, 2000.

Hull, J. *Options, Futures, and Other Derivatives*. Prentice Hall, 1999.

Hull, J. and White, A. The pricing of options on assets with stochastic volatilities. *Journal of finance*, 42:281–300, 1987.

Jarrow, R. and Rudd, A. Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics*, 10(3):347–369, 1982.

Lin, C. T. and Yeh, H. Y. The valuation of Taiwan stock index option prices: comparison of performances between Black-Scholes and Neural Network model. *Journal of Statistics & Management Systems*, 8:355–367, 2005.

Lord, R. and Kahl, C. Complex logarithms in Heston-like models. *Mathematical Finance*, 20(4), 2010.

MacBeth, J. and Merville, L. Tests of the Black-Scholes and Cox call option valuation models. *Journal of Finance*, 35(2):285–301, 1980.

McCracken, M. W. Asymptotics for out of sample tests of Granger causality. *Journal of Econometrics*, 140:719–52, 2007.

Merton, R. C. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

Merton, R. C. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144, 1976.

Mikhailov, S. and Nögel. Heston's stochastic volatility model implementation, calibration and some extensions. Wilmott magazine, Fraunhofer Institute for Industrial Mathematics, Kaiserslautern, Germany, 2003.

Moodley, N. The Heston model: A practical approach. Technical report, University of the Witwatersrand, 2005.

Rapach, D. E., Strauss, J. K., and Zhou, G. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies*, 23(2):821–862, 2010.

Robouts, J. and Stentoft, L. Option pricing with assymetric heteroskedastic normal mixture models. Scientific publications no. 2010s-38, CIRANO, 2010.

Rubinstein, M. Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from august 23, 1976 through august 31, 1978. *Journal of Finance*, 40(2):455–480, 1985.

Scott, L. O. Option pricing when the variance changes randomly: Theory, estimation and application. *Journal of financial and quantitative analysis*, 22:419–438, 1987.

Stentoft, L. What we can learn from pricing 139,879 individual stock options. Research Paper 52, CREATES, 2011.

Stock, J. H. and Watson, M. W. Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23:405–30, 2004.

Teräsvirta, T., Medeiros, M. C., and Rech, G. Building Neural Network models for time series: a statistical approach. *Journal of Forecasting*, 25(1):49–75, 2006.

Timmermann, A. *Handbook of Economic Forecasting*, chapter Forecast combinations. Elsevier, Amsterdam, The Netherlands, 2006.

Tu, J. Advantages and disadvantages of using Artificial Neural Networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, 49(11): 1225–1231, 1996.

Waggoner, D. and Zha, T. Confronting model misspecification in macroeconomics. *Journal of Econometrics*, 171(2):167–184, 2012.

Wang, Y. Nonlinear Neural Network forecasting model for stock index option price: Hybrid GJR-GARCH approach. *Expert Systems with Applications*, 36(1):564–570, 2009.

West, K. D. Asymptotic inference about predictive ability. *Econometrica*, 64:1067–84, 1996.

Yao, J., Li, Y., and Tan, C. Option price forecasting using Neural Networks. *The International Journal of Management*, 28(4):455–466, 2000.

A TECHNICAL DETAILS

A.1 Fast Fourier Transform Heston

Following Moodley (2005), I briefly summarize the Fast Fourier Transform (FFT) method to evaluate the closed-form solution of an European call option for the Heston model. This approach has been proposed by Carr and Madan (1999) to overcome the numerical difficulties that arise when evaluating the integrals in the Heston framework, see for instance equation 2.16. The formula derived by Carr and Madan (1999) depends on the intrinsic value of an option. However, out-the-money options don't have an intrinsic value. For this reason the authors distinguish between options with and without an intrinsic value, which need to be applied for the corresponding option contracts.

FFT for options with an intrinsic value The call price c_t^{HES} for at-the-money and in-the-money options, on a non-dividend paying asset for the Heston model is given by:

$$c_t^{HES} \approx \frac{e^{-\alpha\kappa_u}}{\pi} \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(u-1)} e^{ibv_j} F_c(v_j) \frac{\eta}{3}(3+(-1)^j - \Delta_{j-1}) \tag{A.1}$$

for which

$$\begin{aligned} b &= \pi/\eta, \\ cc &= 600, \\ N &= 4096, \\ \eta &= cc/N, \\ v_j &= \eta(j-1), \\ \kappa_u &= -b + (2b/N)(u-1),\ u = 1,2,...,N+1. \end{aligned}$$

These parameters are given by Carr and Madan (1999) such that there exist a balance between optimization time and pricing accuracy. The characteristic function F_c of $ln(S_t)$ under the martingale measure $\mathbb{Q}$ can be expressed as:

$$\begin{aligned} F_c(\phi) &= e^{A(\phi)+B(\phi)+C(\phi)}, \\ A(\phi) &= i\phi(ln(S_t)+rT), \\ B(\phi) &= \frac{2\zeta(\phi)(1-e^{\psi(\phi)T})\sigma}{2\psi(\phi)-(\psi(\phi)-\gamma(\phi))(1-e^{\psi(\phi)T})}, \\ C(\phi) &= -\frac{\kappa\theta}{\nu^2}\left[2log\left(\frac{2\psi(\phi)-(\psi(\phi)-\gamma(\phi))(1-e^{-\psi(\phi)T})}{2\psi(\phi)}\right)+(\psi(\phi)-\gamma(\phi))T\right], \end{aligned}$$

$$\begin{aligned}\zeta(\phi) &= -\frac{1}{2}(\phi^2 - i\phi),\\ \psi(\phi) &= \sqrt{\gamma(\phi)^2 - 2\nu^2\zeta(\phi)},\\ \gamma(\phi) &= \kappa - \rho\nu\phi i.\end{aligned}$$

The parameters of interest, which are unobserved and need to be calibrated in this framework include σ the variance of the underlying, ν the volatility of the variance, κ the mean reversion rate, θ the long run variance and ρ the correlation parameter. The remaining parameter of interest is α which is usually chosen such that it satisfies the following constraints:

$$\begin{aligned}\mathbb{E}[S_T^{\alpha+1}] &< \infty,\\ F_c(-(\alpha+1)i) &< \infty.\end{aligned}$$

Following Carr and Madan (1999) I choose an α of 1.25.

FFT for options without an intrinsic value The call price c_t^{HES} for out-the-money options, on a non-dividend paying asset for the Heston model is given by:

$$c_t^{HES} \approx \frac{1}{\pi \sinh(\alpha k_u)} \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(u-1)} e^{ibv_j} \gamma_T(v_j) \frac{\eta}{3}(3 + (-1)^j - \Delta_{j-1}) \tag{A.2}$$

with:

$$\gamma_T(v_j) = \frac{\chi_T(v_j - i\alpha) - \chi_T(v_j + i\alpha)}{2}, \tag{A.3}$$

$$\chi_T(v_j) = e^{-rT}\left[\frac{1}{1+v_j} - \frac{e^{rT}}{iv_j} - \frac{F_c(v_j - i)}{v_j^2 - iv_j}\right]. \tag{A.4}$$

For a more detailed description of the derivation, implementation and code to evaluate these terms see also Moodley (2005) and Mikhailov and Nögel (2003).

A.2 Neural Networks

This section discusses the Neural Network combining frameworks in more detail. Two of the three benchmark combining frameworks considered in this study are inspired by Andreou et al. (2008). These authors propose a method to utilize information from the individual models by means of an Artificial Neural Network (ANN). In the following I summarize their methodology in brief.

Neural Networks are a collection of systems of interconnected 'neurons' or nodes structured in successive layers. This resulting network consists of an *input* layer with N input variables; a *hidden* layer with H neurons and a single neuron *output* layer. The connections between the layers are formed through weights $w_{k,i}$ with $k = 1,...,H$ and $i = 1,...,N$ and a bias v_0 for the output layer. In the end, this results in neurons that consists of four elements: (i) a vector of input signals; (ii) vector weights and the bias; (iii) a neuron that sums the product of the input signals with the corresponding weights and bias; (iv) and a neuron *transfer* function. A convenient way to look at ANN's is to see them as a nonlinear regression tool in the form:

$$Y = G(\tilde{x}) + \varepsilon_{ANN},$$

that connect the input variable vector $\tilde{x} = [x_1, x_2, ..., x_N]$ with the target function Y through the unknown function $f(.)$ and an error term ε_{ANN}. These inputs are prepared in terms of feature vectors $\tilde{x_q} = [x_{1,q}, x_{2,q}, ..., x_{N,q}]$ and are assumed to have a corresponding or associated known target $Y = t_q$. In this study the known target is set to the option market price $t_q = c_q^{mkt}$ with $q = 1,...,P$ with P the number of available features. While Andreou et al. (2008) considers different target functions that may lead to better pricing performance in terms of Root Mean Squared Error, setting the option market price as the target functions allows for fair comparison with the parametric methods. Furthermore, this study considers three non-parametric approaches that distinguish themselves by the choice of the input variables. The first Neural Network combining framework ($cANN2$) uses as inputs the model-based estimated option prices from the Corrado and Su (1996) and Merton (1973) frameworks. The second non-parametric combining framework ($cANNa$) utilizes information from all the individual models by setting the inputs to the model-based estimated option prices of all the models, resulting in six input variables. The third non-parametric combining approach ($cANNx$) uses as inputs the spot price S_t, strike X, dividend δ, risk free rate r and the time-to-maturity T of the option contracts. To estimate the output y or combined option price this involves computing y_q for all the features:

$$y_q = f_0 \left[v_0 + \sum_{k=1}^{H} v_k f_H (b_k + \sum_{i=1}^{N} w_{k,i} x_{i,q} \right]. \tag{A.5}$$

We can observe that there are some similarities with the parametric combining frameworks a proposed in this study; the output or combined option price y is a function of individual option pricing model estimates. Following Andreou et al. (2008), the hidden layer uses a *hyperbolic tangent sigmoid transfer function* and the output layer a *linear transfer function*. I also use one hidden layer such that the ANN framework operates as a nonlinear regression tool that can be trained by most functions, see also Cybenko (1989).

In the ANN framework, the first step in optimization involves training the networks. I use the modified LM algorithm for this purpose which updates the weights and biases in such a way that minimizes the following performance function:

$$F(W) = \sum_{q=1}^{P} e_q^2 = \sum_{q=1}^{P} (y_g - t_q)^2.$$

Here W is the N-dimensional column vector that contains the weights and biases $W = [b_1, ..., b_H, w_{1,1}, ..., w_{H,N}, v_o, ..., v_H]^T$. At each operation τ of the LM algorithm, the weight vector is updated by the following equation:

$$W_{\tau+1} = W_\tau + \left[J^T(W_\tau)J(W_\tau) + \mu_\tau I \right]^{-1} J^T(W_\tau)e(W_\tau),$$

with I an $N \times N$ identity matrix, $J(W)$ the $P \times N$ Jacobian matrix of the P-dimensional output error column vector $e(W)$. The learning parameters μ_τ is adjusted in each iteration such that convergence is obtained. Further discussion of the LM algorithm is given by Hagan and Menhaj (1994) and Hagan et al. (1996). This algorithm along with the neural networks is also implemented and available in the MATLAB toolbox 'Neural Network Toolbox'[15]. Further discussion on splitting up the dataset is given in section 5.2.

B COMPLEMENTARY RESULTS

[15]See also: `http://www.mathworks.com/products/neural-network/`

Table B.1: Implied volatilities by various moneyness-maturity categories and periods

This table shows the implied volatilities backed out by the various individual models for different moneyness and maturity categories. Each day in the sample (2006-2011), the implied volatility of a given model is estimated by minimizing the sum of squared pricing errors between the market price and the model-determined price for each option. The parameters are estimated by splitting the in-sample data into twenty-four categories, based on the moneyness (eight categories) and expiration date (three categories) of all the available options up to current day. The values presented are the weighted averages across different sub-periods. The overall sample period covers January 2006 - December 2011, with a total of 893708 observations. The recession start and end date are obtained from the NBER website and indicate the latest recession, characterized by the recent financial crisis, which spans December 2007 - June 2009. BLS refers to the benchmark Black and Scholes (1973), MER1 to Merton (1973), CSU to Corrado and Su (1996), MER2 to the Merton (1976) jump diffusion, COX to Cox and Rubinstein (1979) and HES to the Heston (1993) stochastic volatility model.

Moneyness (S/X)	Days to maturity < 60 days						Days to maturity 60-180 days						Days to maturity > 180 days					
Overall period	BLS	MER1	CSU	MER2	COX	HES	BLS	MER1	CSU	MER2	COX	HES	BLS	MER1	CSU	MER2	COX	HES
< 0.85	0.218	0.222	0.413	0.544	0.219	0.178	0.146	0.154	0.306	0.255	0.172	0.224	0.123	0.148	0.177	0.136	0.137	0.190
0.85 - 0.95	0.139	0.144	0.196	0.154	0.146	0.068	0.140	0.150	0.311	0.130	0.139	0.082	0.133	0.165	0.259	0.144	0.132	0.141
0.95 - 0.99	0.151	0.158	0.352	0.147	0.145	0.086	0.153	0.167	0.523	0.151	0.151	0.124	0.137	0.178	0.263	0.150	0.144	0.142
0.99 - 1.01	0.161	0.172	0.410	0.172	0.174	0.080	0.160	0.178	0.469	0.179	0.175	0.111	0.130	0.185	0.262	0.158	0.155	0.206
1.01 - 1.05	0.170	0.188	0.298	0.170	0.163	0.102	0.161	0.189	0.324	0.174	0.161	0.120	0.130	0.190	0.259	0.159	0.134	0.190
1.05 - 1.10	0.163	0.211	0.323	0.177	0.152	0.093	0.155	0.203	0.330	0.169	0.141	0.133	0.129	0.198	0.215	0.157	0.125	0.196
1.10 - 1.35	0.145	0.261	0.559	0.199	0.134	0.114	0.152	0.228	0.406	0.183	0.146	0.152	0.112	0.211	0.282	0.168	0.102	0.207
> 1.35	0.250	0.238	0.397	0.174	0.175	0.138	0.179	0.243	0.574	0.174	0.113	0.196	0.108	0.306	0.492	0.166	0.035	0.144
Recession	BLS	MER1	CSU	MER2	COX	HES	BLS	MER1	CSU	MER2	COX	HES	BLS	MER1	CSU	MER2	COX	HES
< 0.85	0.240	0.243	0.407	0.475	0.230	0.192	0.183	0.190	0.330	0.207	0.202	0.248	0.167	0.189	0.161	0.169	0.175	0.245
0.85 - 0.95	0.211	0.216	0.239	0.196	0.208	0.108	0.206	0.216	0.247	0.186	0.195	0.119	0.186	0.216	0.255	0.197	0.178	0.142
0.95 - 0.99	0.231	0.239	0.315	0.241	0.225	0.122	0.223	0.238	0.439	0.248	0.225	0.142	0.193	0.230	0.272	0.206	0.203	0.142
0.99 - 1.01	0.244	0.253	0.301	0.280	0.266	0.131	0.232	0.249	0.340	0.308	0.257	0.192	0.196	0.236	0.347	0.224	0.220	0.326
1.01 - 1.05	0.254	0.267	0.261	0.258	0.249	0.184	0.238	0.258	0.263	0.276	0.243	0.206	0.197	0.241	0.269	0.216	0.210	0.273
1.05 - 1.10	0.264	0.286	0.320	0.284	0.247	0.149	0.244	0.272	0.188	0.258	0.229	0.183	0.197	0.249	0.179	0.222	0.195	0.248
1.10 - 1.35	0.240	0.311	0.557	0.273	0.316	0.193	0.239	0.296	0.389	0.244	0.230	0.239	0.157	0.265	0.317	0.220	0.201	0.272
> 1.35	0.334	0.281	0.389	0.238	0.359	0.202	0.214	0.306	0.564	0.253	0.294	0.378	0.135	0.313	0.401	0.211	0.154	0.266
Before recession	BLS	MER1	CSU	MER2	COX	HES	BLS	MER1	CSU	MER2	COX	HES	BLS	MER1	CSU	MER2	COX	HES
< 0.85	0.248	0.251	0.505	0.586	0.253	0.110	0.102	0.108	0.228	0.498	0.159	0.315	0.069	0.098	0.214	0.106	0.103	0.099
0.85 - 0.95	0.080	0.085	0.165	0.146	0.100	0.028	0.081	0.090	0.303	0.083	0.092	0.036	0.069	0.104	0.296	0.083	0.084	0.138
0.95 - 0.99	0.086	0.093	0.192	0.081	0.084	0.050	0.086	0.101	0.531	0.085	0.084	0.093	0.063	0.112	0.298	0.092	0.070	0.142
0.99 - 1.01	0.093	0.104	0.370	0.093	0.097	0.053	0.088	0.109	0.470	0.096	0.096	0.079	0.078	0.117	0.314	0.091	0.069	0.134
1.01 - 1.05	0.095	0.117	0.296	0.104	0.090	0.063	0.086	0.117	0.331	0.103	0.083	0.069	0.035	0.120	0.266	0.102	0.117	0.100
1.05 - 1.10	0.053	0.129	0.362	0.100	0.158	0.050	0.047	0.125	0.444	0.093	0.130	0.079	0.030	0.123	0.262	0.080	0.098	0.134
1.10 - 1.35	0.032	0.159	0.490	0.128	0.142	0.053	0.031	0.106	0.301	0.112	0.147	0.049	0.016	0.115	0.201	0.107	0.102	0.082
> 1.35	0.051	0.065	0.666	0.112	0.275	0.061	0.050	0.068	0.641	0.116	0.142	0.062	0.064	0.166	0.571	0.111	0.142	0.074
After recession	BLS	MER1	CSU	MER2	COX	HES	BLS	MER1	CSU	MER2	COX	HES	BLS	MER1	CSU	MER2	COX	HES
< 0.85	0.191	0.195	0.391	0.589	0.199	0.200	0.136	0.143	0.313	0.215	0.156	0.195	0.135	0.158	0.160	0.137	0.137	0.224
0.85 - 0.95	0.137	0.142	0.192	0.134	0.142	0.081	0.144	0.154	0.358	0.131	0.139	0.096	0.149	0.180	0.232	0.156	0.139	0.142
0.95 - 0.99	0.149	0.157	0.499	0.139	0.140	0.091	0.159	0.173	0.569	0.140	0.154	0.135	0.158	0.196	0.231	0.160	0.164	0.142
0.99 - 1.01	0.162	0.172	0.509	0.164	0.174	0.068	0.168	0.186	0.550	0.161	0.184	0.085	0.162	0.204	0.168	0.167	0.180	0.215
1.01 - 1.05	0.176	0.192	0.324	0.165	0.168	0.088	0.177	0.199	0.361	0.165	0.175	0.106	0.165	0.211	0.248	0.169	0.173	0.208
1.05 - 1.10	0.193	0.222	0.295	0.177	0.183	0.095	0.189	0.219	0.333	0.173	0.173	0.145	0.168	0.222	0.221	0.175	0.163	0.219
1.10 - 1.35	0.178	0.275	0.613	0.209	0.199	0.120	0.195	0.265	0.497	0.198	0.194	0.174	0.165	0.251	0.325	0.182	0.155	0.264
> 1.35	0.139	0.344	0.197	0.185	0.273	0.160	0.142	0.338	0.529	0.170	0.214	0.192	0.091	0.343	0.489	0.180	0.113	0.139

Table B.2: In-sample pricing performance of all the models (2006-2011)

Presented are the pricing errors of the various option pricing models considered in this study, in an in-sample framework. I compute the current's day option prices using the current day's implied parameters and implied stock volatility. For each moneyness-maturity category, I report the Root Mean Squared Error (RMSE) of the difference between market price and model price for all option contacts. The sample period covers January 2006 - December 2011, with a total of 894305 observations. A total of 7288 observations are discarded from the analysis due to calibration issues or an non-feasible estimates. The individual models considered are the BLS (Black and Scholes, 1973), MER1 (Merton, 1973), CSU (Corrado and Su, 1996), MER2 (Merton, 1976), COX (Cox and Rubinstein, 1979) and HES Heston (1993). cMSE is the combined model which combines the individual option prices of all models, by assigning weights to each individual model based on the historical performance, measured by the Discounted Mean Squared Error (DMSE), during a hold-out period of 10 days. cMEAN and cMED refer to respectively the average and median of all model based estimates of the individual option pricing models on a given day. cANN2 is the neural network model, following Andreou et al. (2008) which utilizes information from the MER1 and CSU framework by means of an Artificial Neural Network. Underlined values represent the lowest value for a given row/moneyness category.

Moneyness		Days to maturity <60 days									
S_t/X	Obs.	BLS	MER1	CSU	MER2	COX	HES	cMSE	cMEAN	cMED	cANN2
< 0.85	7624	1.66	1.67	1.71	1.63	2.43	1.81	1.59	1.63	1.64	9.13
0.85 - 0.95	33467	2.41	2.39	2.45	2.65	2.84	3.07	2.41	2.43	2.43	17.20
0.95 - 0.99	29346	3.07	3.11	3.13	3.41	3.33	7.30	3.14	3.33	3.12	15.40
0.99 - 1.01	14959	3.42	3.46	3.43	4.02	3.51	5.60	3.46	3.57	3.45	10.78
1.01 - 1.05	27256	3.34	3.36	3.30	3.68	3.29	5.40	3.25	3.35	3.29	7.51
1.05 - 1.10	30048	3.21	2.92	2.95	3.14	3.46	4.98	2.83	2.99	2.93	4.96
1.10 - 1.35	88236	3.12	2.57	2.72	2.17	3.40	4.09	2.11	2.39	2.35	4.50
> 1.35	74619	2.97	1.58	2.82	1.55	2.98	3.26	1.33	1.85	1.89	6.15
All	**305555**	**3.01**	**2.49**	**2.84**	**2.49**	**3.21**	**4.31**	**2.28**	**2.54**	**2.50**	**8.08**
Moneyness		Days to maturity: 60-180 days									
S_t/X	Obs.	BLS	MER1	CSU	MER2	COX	HES	cMSE	cMEAN	cMED	cANN2
< 0.85	21932	1.88	1.89	1.96	2.22	3.12	2.56	1.91	1.96	1.92	17.03
0.85 - 0.95	42795	3.27	3.40	3.53	3.53	3.66	8.49	3.29	3.51	3.31	15.77
0.95 - 0.99	21379	4.15	4.34	4.26	4.76	4.81	13.28	4.27	4.72	4.25	11.10
0.99 - 1.01	10654	4.14	4.20	4.05	5.41	4.20	5.31	4.09	4.19	4.13	7.96
1.01 - 1.05	18975	4.15	4.12	4.11	4.80	4.18	5.54	3.97	4.07	4.05	5.97
1.05 - 1.10	19333	4.51	3.83	3.75	4.13	4.44	5.50	3.67	3.86	3.77	4.70
1.10 - 1.35	57470	6.02	4.25	3.49	3.07	6.41	4.86	3.02	3.56	3.27	4.47
> 1.35	65089	7.69	3.63	4.27	2.46	7.66	6.68	2.03	3.56	4.09	5.44
All	**257627**	**5.15**	**3.72**	**3.72**	**3.36**	**5.45**	**6.54**	**2.98**	**3.60**	**3.58**	**8.49**
Moneyness		Days to maturity: >180 days									
S_t/X	Obs.	BLS	MER1	CSU	MER2	COX	HES	cMSE	cMEAN	cMED	cANN2
< 0.85	71460	4.42	5.04	6.80	5.76	7.14	6.41	4.04	3.92	4.05	17.05
0.85 - 0.95	43958	6.46	6.68	9.35	7.55	7.13	24.79	5.80	7.18	6.09	10.16
0.95 - 0.99	17785	8.28	5.97	7.65	7.13	7.63	29.52	6.02	8.44	6.75	7.42
0.99 - 1.01	9139	11.09	5.44	6.36	8.89	10.58	9.60	5.76	7.49	7.20	6.74
1.01 - 1.05	16872	12.63	5.23	7.46	6.58	11.98	10.08	5.59	7.38	6.35	6.58
1.05 - 1.10	19074	14.54	5.04	7.75	7.05	13.18	9.01	5.41	7.95	6.00	6.32
1.10 - 1.35	59619	16.23	7.20	8.48	5.69	16.02	9.59	5.58	9.51	6.96	7.12
> 1.35	93216	22.43	11.00	9.95	4.86	22.27	14.14	6.23	12.32	9.39	9.87
All	**331123**	**13.28**	**7.39**	**8.45**	**6.01**	**13.71**	**13.27**	**5.48**	**8.48**	**6.81**	**10.37**
Overall	**894305**	**7.43**	**4.66**	**5.17**	**4.05**	**7.74**	**8.27**	**3.67**	**5.04**	**4.41**	**9.05**