Models for S&P500 Dynamics: Evidence from Realized Volatility, Daily Returns, and Option Prices

Peter Christoffersen   Kris Jacobs   Karim Mimouni
Desautels Faculty of Management, McGill University*

July 31, 2007

Abstract

Most recent empirical option valuation studies build on the affine square root (SQR) stochastic volatility model. The SQR model is a convenient choice, because it yields closed-form solutions for option prices. However, relatively little is known about the resulting biases. We investigate alternatives to the SQR model, by comparing its empirical performance with that of five different but equally parsimonious stochastic volatility models. We provide empirical evidence from three different sources. We first use realized volatilities to assess the properties of the SQR model and to guide us in the search for alternative specifications. We then estimate the models using maximum likelihood on S&P500 returns. Finally, we employ nonlinear least squares on a panel of option data. In comparison with earlier studies that explicitly solve the filtering problem, we analyze a more comprehensive option data set. The scope of our analysis is feasible because of our use of the particle filter. The three sources of data we employ all point to the same conclusion: the SQR model is misspecified. Overall, the best of the alternative volatility specifications is a model with linear rather than square root diffusion for variance which we refer to as the VAR model. This model captures the stylized facts in realized volatilities, it performs well in fitting various samples of index returns, and it has the lowest option implied volatility mean squared errors in- and out-of-sample.

JEL Classification: G12

Keywords: Stochastic volatility; option valuation; particle filtering; skewness; kurtosis; mean reversion.

*Christoffersen (peter.christoffersen@mcgill.ca) and Jacobs (kris.jacobs@mcgill.ca) are also affiliated with CIRANO and CIREQ and want to thank FQRSC, IFM^2 and SSHRC for financial support. Mimouni (karim.mimouni@mail.mcgill.ca) was supported by IFM^2 and FQRSC. We are grateful for helpful comments from Yacine Aït-Sahalia (the Editor), Torben Andersen, Mikhail Chernov, Bjorn Eraker, Mark Kamstra, Nour Meddahi, Marcel Rindisbacher, Stephen Taylor, and two anonymous referees. Correspondence to: Peter Christoffersen, Desautels Faculty of Management, McGill University, 1001 Sherbrooke Street West, Montreal, Quebec, Canada, H3A 1G5; Tel: (514) 398-2869; Fax: (514) 398-3876.

1 Introduction

Following the finding that Black-Scholes (1973) model prices systematically differ from market prices, the literature on option valuation has formulated a number of theoretical models designed to capture these empirical biases. One particularly popular modeling approach has attempted to correct the Black-Scholes biases by modifying the assumption that volatility is constant across maturity and moneyness. Estimates from returns data and options data indicate that return volatility is time-varying, and modeling volatility clustering leads to significant improvements in the performance of option pricing models. It has also been demonstrated that it is necessary to model a leverage effect. The leverage effect captures the negative correlation between returns and volatility, and thus generates negative skewness in the distribution of the underlying asset return.[1]

The existing literature has almost exclusively modeled volatility clustering and the leverage effect within an affine or "square root" structure. In particular, the Heston (1993) model, which accounts for time-varying volatility and a leverage effect, has been implemented in a large number of empirical studies. Henceforth, we will refer to this model as the SQR model. The SQR model is convenient because it leads to (quasi) closed-form solutions for prices of European equity options. It is not surprising that the choice of model is partly driven by convenience: the estimation of option valuation models using large cross-sections of option contracts is computationally burdensome because of the need to filter the latent stochastic volatility. As the solution to the Heston (1993) SQR model relies on univariate numerical integration, it is relatively easier to estimate than many other models.

It is well recognized that the SQR model cannot capture important stylized facts. In order to address the limitations of the square root structure, the model is often combined with models of jumps in returns and/or volatility.[2] However, relatively few studies analyze non-affine stochastic volatility models, and therefore much less is known about the empirical biases that result from imposing the affine square root structure on the stochastic volatility dynamic in the first place. Notable exceptions that investigate option valuation using non-affine stochastic volatility models are Aït-Sahalia and Kimmel (2007), Jones (2003) and Benzoni (2002). The non-affine model in Benzoni (2002) does not improve on the performance of the Heston (1993) model. Jones (2003) analyzes the more general constant elasticity of variance (CEV) model using a bivariate time series of returns and an at-the-money short maturity option, and a number of his specification tests favor the non-affine constant elasticity of substitution model over the SQR model. Aït-Sahalia and Kimmel (2007) also estimate different models using joint time series of the underlying return and

[1] The leverage effect was first characterized in Black (1976). For empirical studies that emphasize the importance of volatility clustering and the leverage effect for option valuation see among others Benzoni (2002), Chernov and Ghysels (2000), Eraker (2004), Heston and Nandi (2000), and Pan (2002).

[2] For empirical studies that implement the Heston (1993) model by itself or in combination with different types of jump processes, see for example Andersen, Benzoni and Lund (2002), Bakshi, Cao and Chen (1997), Bates (1996, 2000), Broadie, Chernov and Johannes (2007), Benzoni (2002), Chernov and Ghysels (2000), Huang and Wu (2004), Pan (2002), Eraker (2004) and Eraker, Johannes and Polson (2003).

a short-dated at-the-money option price or implied volatility. They also find that the SQR model is misspecified, but the nature of the misspecification and the empirical evidence are different from Jones (2003).

A number of papers estimate the SQR and alternative stochastic volatility specifications on returns data. These studies find strong evidence against the SQR specification, see for example Chacko and Viceira (2003) and Chernov, Gallant, Ghysels and Tauchen (2003). Similarly, the realized volatility literature seems to strongly favor logarithmic, non-affine SV specifications. See for example, Andersen, Bollerslev, Diebold and Ebens (2001). Yet state-of-the-art option pricing papers, such as for example Broadie, Chernov and Johannes (2007), continue employing the SQR model as a building block.

This paper further investigates the empirical implications of adopting a square root stochastic volatility model for option valuation. We compare the empirical performance of the Heston (1993) square root stochastic volatility model with that of five alternatives. In contrast with Aït-Sahalia and Kimmel (2007) and Jones (2003), who compare the SQR model with more richly parameterized stochastic volatility models, we deliberately choose alternative specifications with the same number of parameters as the SQR model. This approach also differs from several other studies that enrich the SQR model by adding for example jumps to the basic model, which also increases the number of model parameters. We deliberately ask a very different question: is the SQR model the best possible stochastic volatility model to build upon? If not, what are the empirical deficiencies? Part of our motivation for keeping the number of parameters constant is our desire to make the in-sample comparison between models as meaningful as possible. It is well known that in-sample model comparisons favor relatively richly parameterized models, but that these in-sample model rankings will often be reversed in out-of-sample experiments.

We provide evidence on the six stochastic volatility specifications from three different sources. First, we use realized volatilities to assess the properties of the SQR model and to guide us in the search for alternative specifications. Subsequently, we estimate the model parameters using maximum likelihood on returns only, and finally we employ nonlinear least squares on a time series of cross-sections of option data with extensive cross-sectional variation across moneyness and maturity.

Our estimation on option contracts uses a rich panel of data. It therefore critically differs from Aït-Sahalia and Kimmel (2007) and Jones (2003), who estimate model parameters using a bivariate time series. Indeed, to the best of our knowledge our analysis of the SQR model uses substantially larger cross-sections of option contracts than any other available study that explicitly solve the filtering problem. Existing studies either solve the filtering problem and use a small cross-section of option data, or use a large cross-section and do not explicitly solves the filtering problem. For example, Aït-Sahalia and Kimmel (2007), Chernov and Ghysels (2000), Eraker (2004) and Jones (2003) explicitly solve the filtering problem using a relatively limited cross-section of option data. Studies that investigate a wider cross-section are for example Bakshi, Cao and Chen (1997), who estimate one cross-section at a time, and Bates (2000) and Huang and Wu (2004), who estimate a separate volatility parameter for each cross-section. Our estimation on a rich option data set while explicitly solving the filtering problem is a non-trivial contribution, because one can arguably only reliably estimate the parameters that determine the model smirk when using such a wide range of option contracts.

We are able to estimate a large number of stochastic volatility models using comprehensive cross-sections of options data thanks to our use of the particle filter. This methodology is extensively used in the engineering literature, and has also recently been used for financial applications.[3] Particle filtering provides a convenient tool for the analysis of latent factor models such as stochastic volatility models. It is easy to implement, and it can be adapted to provide the best possible fit to any objective function and any model of interest. In our opinion, the trade-off with other empirical methods is a very favorable one in the context of computationally intensive option valuation problems.

The three sources of data we employ all point to the same conclusion: the SQR model is misspecified. Overall, the best of the alternative volatility specifications is a model we refer to as the VAR model, which is of the GARCH diffusion type. This model captures the stylized facts in realized volatilities, it performs well in fitting a long sample of index returns, and it has the lowest option price mean squared errors in- and out-of-sample. However, when the crash of 1987 is included in the data, the VAR model is outperformed in return fitting by a model with a higher variance of variance.

The paper proceeds as follows. In Section 2 we discuss the benchmark Heston (1993) model in light of the evidence from realized volatilities. We also use the realized volatilities to help guide the search for alternative specifications. Section 3 discusses particle filter based estimation on index returns and index options. Section 4 presents the empirical results obtained using return and options data. Section 5 concludes.

2 Stochastic Volatility Specifications

This section discusses alternatives to the square root volatility specification in Heston (1993), which has become the standard building block for more elaborate models in the option valuation literature. Note that it is relatively straightforward to write down more heavily parameterized models that outperform the square root model. Trivially, a more heavily parameterized model will always outperform a simpler nested model in-sample. Moreover, Bates (2003) points out that even in short-horizon out-of-sample experiments, the richer models may outperform nested models because the "smile" pattern in option prices is very persistent. This finding is more likely if the models are re-estimated frequently, say daily as in the classic study by Bakshi, Cao and Chen (1997) for

[3] See Gordon, Salmond and Smith (1993) and Pitt and Shepherd (1999). See Johannes, Polson and Stroud (2002) for an application to returns data.

example.

Thus, we deliberately confine attention to models that have the same number of parameters as Heston's SQR model, and we estimate on a long sample of data. The models we consider can be thought of as alternative building blocks for more elaborate models containing jumps in returns and volatility as well as multiple volatility factors. However, such extensions are beyond the scope of this paper. In our paper, the objective is to establish a well-specified volatility dynamic, which in our opinion ought to precede the development of more elaborate versions of a possibly misspecified basic model.

2.1 Realized Variance: Implications for the Square Root Model

The Heston (1993) square root (SQR) model assumes that the instantaneous change in variance, V, has the following dynamics

$$\text{SQR: } dV = \kappa(\theta - V)dt + \sigma\sqrt{V}dw \tag{1}$$

where κ denotes the speed of mean reversion, θ is the unconditional variance, and σ determines the variance of variance. The variance innovation dw is a Brownian motion, and we have $corr(dz, dw) = \rho$, where dz is the innovation in the underlying spot price, given by

$$dS = \mu S dt + \sqrt{V} S dz. \tag{2}$$

We will use the return specification (2) in each of the volatility models considered below.

In order to explore the implications of the SQR specification, consider the instantaneous volatility dynamic implied by the model. Using Ito's lemma, we can write

$$\text{SQR: } d\sqrt{V} = \mu(V)dt + \frac{1}{2}\sigma dw \tag{3}$$

where the volatility drift $\mu(V)$ is a function of the variance level. Note that the SQR model implies that the instantaneous change in volatility should be Gaussian and homoskedastic: V does not show up in any way in the diffusion term for $d\sqrt{V}$. This is a strong implication, which can be quite easily evaluated empirically.

As a first piece of empirical evidence, we construct daily realized variances, RV_t, from intraday returns. We obtain intraday S&P500 quotes for 1996 through 2004 from which we construct a grid of two-minute returns. From these two-minute returns we construct robust realized variances using the two-scale estimator from Zhang, Mykland, and Aït-Sahalia (2005).[4] The two-scale estimator is defined as

$$RV_t^{TS} = RV_t^{Avr} - 0.0614 RV_t^{All}$$

where RV_t^{Avr} is the average of all the possible low-frequency RV estimates that use (in our case) 30-minute squared returns on the two-minute grid, and where RV_t^{All} is the high-frequency estimator

[4] We scale up the open-to-close realized variance estimates to match the overall variance in close-to-close daily returns for the sample period.

summing over all the available two-minute squared returns.[5] The coefficient on RV_t^{All} depends on the frequency of the sparse estimators (30 minutes here) and of the high-frequency estimator (2 minute here). Zhang, Mykland, and Aït-Sahalia (2005) show that even in the presence of market microstructure noise this estimator converges to the true integrated variance defined by $\int_{t-1}^{t} V_\tau d\tau$.

Consider now the left-hand panels of Figure 1. Using daily realized volatilities from 1996 through 2004, the top-left panel of Figure 1 shows a quantile-quantile (QQ) plot of the daily realized volatility changes compared with the Gaussian distribution. The deviation of the data points from the straight line indicates that the Gaussian distribution is not a good assumption for daily changes in volatility. The observed tails (both left and right) are considerably fatter than the normal distribution would suggest. Now consider the middle-left panel in Figure 1, which scatter plots the daily volatility changes against the daily volatility level. According to the SQR model, this scatter plot should not display any systematic patterns. However, as the volatility level increases on the horizontal axis, a cone-shaped pattern in the daily volatility changes on the vertical axis is apparent. The bottom-left panel of Figure 1 confirms this finding: it scatter plots the absolute daily volatility changes against the daily volatility level. A simple OLS regression line is shown for reference. Notice the apparent positive relationship between the volatility level and the magnitude of the volatility changes. This pattern is in conflict with the homoskedastic volatility implication of the SQR model.

Using Ito's lemma, the dynamics for log variance in the SQR model can be written as

$$\text{SQR: } d\ln(V) = \mu(V)dt + \sigma \frac{1}{\sqrt{V}} dw \tag{4}$$

The right-hand panels of Figure 1 summarize the empirics of the logarithms of the realized variances. The top panel shows that daily changes in the realized log variances follow the Gaussian distribution quite closely. The result clearly differs from the corresponding top left panel for the daily volatility changes. Furthermore, the scatter of daily log variance changes against the log variance level in the middle-right panel of Figure 1 does not reveal a cone-shaped pattern. The bottom-right panel in Figure 1 confirms this result: it shows a virtually horizontal line when regressing absolute changes in log variance on log variance levels. The absence of a clear relationship between changes in the log variances and the log variance level in Figure 1 casts further doubt on the SQR specification, for which the instantaneous changes in log variances are heteroskedastic, as is evident from (4).

2.2 Alternative Specifications for Variance Dynamics

The evidence in Figure 1 suggests specifying the variance dynamic as follows

$$\text{VAR: } dV = \kappa(\theta - V)dt + \sigma V dw \tag{5}$$

[5] See Aït-Sahalia, Mykland and Zhang (2005) for a discussion of the optimal sampling frequency.

which has the property that $\ln(V)$ is homoskedastic.[6]

$$d\ln(V) = \mu(V)dt + \sigma dw, \text{ where}$$
$$\mu(V) = \left(\tfrac{1}{V}\kappa(\theta - V) - \tfrac{1}{2}\sigma^2\right)$$

Heston (1997) and Lewis (2000) suggest another alternative to the SQR model: the so-called "three halves" model defined by

$$3/2\text{N: } dV = \kappa V(\theta - V)dt + \sigma V^{3/2}dw \quad (6)$$

In addition to a different power on V in the diffusion term, the 3/2N model has the interesting implication that the variance drift is nonlinear in the level of the variance. We highlight this feature by denoting the model "3/2N". Another way to view the drift specification is that the variance mean reversion is a function of the level of V. The larger the level of variance in the market, the faster the speed of mean reversion κV to the unconditional level θ. This allows for a potentially fast mean-reversion following unusually large volatility spikes, such as October 1987, and slow mean-reversion during prolonged low volatility episodes, such as the mid 1990s. The 3/2N model also leads to a closed-form solution for European call option prices (see Lewis (2000)), even though to the best of our knowledge the model has not yet been implemented using this solution.

So far we have considered three different specifications of the diffusion and two different specifications for the drift of V. We can think of the three models considered so far as belonging to the class

$$dV = \kappa V^a(\theta - V)dt + \sigma V^b dw, \text{ for } a = \{0, 1\} \text{ and } b = \{1/2, 1, 3/2\} \quad (7)$$

Although we will focus our discussion on the SQR, VAR and 3/2N models, this framework encompasses a total of six models, which we denote

a	b	Name
0	1/2	SQR
1	1/2	SQRN
0	1	VAR
1	1	VARN
0	3/2	3/2
1	3/2	3/2N

All of these models will be estimated below.

The literature contains several good discussions of the properties of stochastic differential equations. See for instance Karlin and Taylor (1981). Aït-Sahalia (1996) contains a discussion of the properties of general interest rate processes and Lewis (2000) contains a treatment of certain variance processes. Jones (2000) provides an excellent discussion of the CEV stochastic variance

[6] Although it would seem natural from the realized variance analysis above, we do not consider the homoskedastic log SV model where $d\ln(V) = \kappa(\theta - \ln(V)) + \sigma dw$, because it typically performs very similar to–but slightly worse than–the VAR model.

process. He shows that for a model with $a = 0$ and $b > 1$, zero and plus infinity are unattainable values for the stochastic variance. He also shows that a solution to the variance SDE exists and is unique, and that the stationary distribution exists. Thus Jones (2003) covers the model we refer to as 3/2. Similar arguments can be used to show that these results hold also for the VAR model, the VARN model and the 3/2N model. For the SQRN model, all but the existence of the stationary distribution can be shown using similar arguments as well.

2.3 Volatility Risk Premia

In order to compute option prices, which depend on the price of volatility risk, we need to specify a volatility risk premium relationship for each of the above six models. Heston (1993) assumes that the volatility risk premium $\lambda(S, V, t)$ is equal to λV, so that the risk neutral dynamic for the SQR model is

$$\text{SQR: } dS = rSdt + \sqrt{V}Sz^* \tag{8}$$

$$dV = (\kappa(\theta - V) + \lambda V)\, dt + \sigma\sqrt{V}dw^* \tag{9}$$

$$= (\kappa - \lambda)(\kappa\theta/(\kappa - \lambda) - V)dt + \sigma\sqrt{V}dw^*$$

with $corr(dz^*, dw^*) = \rho$. Notice that the variance dynamic takes the same form under the physical and risk neutral measures.

By analogy with Heston's (1993) variance risk premium specification in the SQR model, we use the following functional form for the variance risk premium

$$\lambda(S, V, t) = \lambda V^{a+1} \tag{10}$$

This specification ensures that the variance dynamic takes the same form under the two measures, as is the case in the Heston (1993) SQR model. Under the risk neutral measure, we have

$$dS = rSdt + \sqrt{V}Sz^* \tag{11}$$

$$dV = (\kappa - \lambda)V^a(\kappa\theta/(\kappa - \lambda) - V)dt + \sigma V^b dw^*$$

with $corr(dz^*, dw^*) = \rho$.

In all six models the risk-neutralization can be obtained using a no-arbitrage argument. A utility-based argument behind the risk neutralization can be explicitly characterized for some cases when assuming log-utility. See Lewis (2000) for a thorough discussion of these issues.

2.4 Computing Option Prices

Heston (1993) demonstrates that the SQR model admits a closed form solution for option prices, which can be written as

$$C(V_t) = SP_{1,t} - Ke^{-r(T-t)}P_{2,t} \tag{12}$$

where $P_{1,t}$ and $P_{2,t}$ are computed using numerical integration of the conditional characteristic function.

Heston (1997) and Lewis (2000) show that for the 3/2N model a similar closed form solution is available, however, in this case the characteristic function involves the gamma and confluent hypergeometric function with complex arguments. To the best of our knowledge this solution is numerically very intensive and has not yet been implemented empirically. Barone-Adesi, Rasmussen and Ravanelli (2003) and Sabanis (2003) develop approximate analytical solutions for the VAR model, but an exact solution does not appear to have been found.

In order to ensure that differences across models are not driven by a particular numerical technique, we compute all model option prices using Monte Carlo. Thus the call option prices are computed via the Monte Carlo sample analogue of the discounted expectation

$$C(V_t) = e^{-r(T-t)} E_t^* \left[S_T - K, 0 \right] \tag{13}$$

where the expectation is calculated using the risk neutral measure.

In order to calculate Monte Carlo prices, we use 1,000 simulated paths and a number of numerical techniques to increase numerical efficiency, namely the empirical martingale method of Duan and Simonato (1998), stratified random numbers, antithetic variates and a Black-Scholes control variate technique.

To assess the accuracy of our Monte Carlo setup we compute a set of analytical option prices using Heston's (1993) model for various strike prices, moneyness and maturities. We then compute Monte Carlo option prices for the same options. Using the numerical example in Heston (1993), the risk-neutral variance process is parameterized as follows

$$dV = 2(.01 - V)dt + .2\sqrt{V} dw^* \tag{14}$$

The correlation ρ is set to -0.5, and the risk-free rate is set to zero.

Figure 2 reports the results. The Monte Carlo prices, denoted by '+', offer a very good approximation to the analytical prices, denoted by solid lines, in all cases. The figure contains two maturities: one month (left panels) and three months (right panels). The figure considers three spot variance levels: $V_0 = .005$, which is half the unconditional variance (top panels), $V_0 = .01$ which is equal to the unconditional variance (middle panels), and $V_0 = .02$ which is twice the unconditional variance (bottom panels).

3 Estimation Using the Particle Filter

The SQR model has been investigated empirically in a large number of studies. It is often used as a building block, together with models of jumps in return and volatility. For our purpose, it is important to note that when estimating the model on option data, a number of different techniques can be used. First, the model's parameters can be estimated using a single cross-section of option prices as in Bakshi, Cao and Chen (1997). A second type of implementation of the SQR model uses multiple cross-sections of option prices but does not use the information in the underlying asset returns. Instead, for every cross-section a different initial volatility is estimated, leading to a highly parameterized problem. This approach is taken for instance in Bates (2000) and Huang and Wu (2004). A third group of papers provide a likelihood-based analysis of the stochastic volatility model. See for example Aït-Sahalia and Kimmel (2007), Bates (2004), Jones (2003), and Eraker (2004), who provides a Markov Chain Monte Carlo analysis. A likelihood-based approach can combine information from the option data and the underlying returns, and impose consistency between the physical and risk neutral dynamics in estimation. Both the return data and the option data carry a certain weight in the objective function. Finally, Chernov and Ghysels (2000) use the efficient method of moments and Pan (2002) uses a method of moments technique as well. These methods can also combine information in option data with the information in underlying returns.

The empirical challenge in all the stochastic volatility models is that the spot volatility V_t is an unobserved latent factor. In order to estimate the parameters in the SV models we thus need to apply a filtering technique on observed index returns. The last two groups of papers discussed above explicitly solve the filtering problem. The filtering problem is also explicitly considered in some papers that use return data to estimate continuous-time stochastic volatility models, such as Chernov, Gallant, Ghysels and Tauchen (2003).

Our implementation solves the filtering problem using the particle filter (PF) algorithm for the six stochastic volatility models. As shown by Gordon, Salmond and Smith (1993) the PF offers a convenient filter for non-linear models such as the stochastic volatility models we consider here. Because the PF procedure is relatively new in finance, we discuss the implementation of this method in some detail.

We will be using the PF both for maximum likelihood (ML) estimation on returns and for nonlinear least squares (NLS) estimation on option prices. Below we first describe the PF, then the maximum likelihood importance sampling (MLIS) estimation methodology we use to estimate using return data, and finally the nonlinear least squares importance resampling (NLSIS) estimation methodology we use to estimate using option data.

3.1 Volatility Discretization

Working with log returns, we can write the generic SV process as

$$\begin{aligned} d\ln(S) &= \left(\mu - \frac{1}{2}V\right)dt + \sqrt{V}dz \\ dV &= \kappa V^a(\theta - V)dt + \sigma V^b dw \end{aligned} \tag{15}$$

Note that the equations in (15) specify how the unobserved state is linked to observed stock prices. This relationship allows us to infer the volatility path using the returns data. We first need to discretize (15). There are different discretization methods and every scheme has certain advantages and drawbacks. We use the Euler scheme, which is easy to implement and has been found to work well for this type of applications.[7] Discretizing (15) gives

$$\ln(S_{t+1}) = \ln(S_t) + \left(\mu - \frac{1}{2}V_t\right) + \sqrt{V_t} z_{t+1} \tag{16}$$

$$V_{t+1} = V_t + \kappa V_t^a (\theta - V_t) + \sigma V_t^b w_{t+1} \tag{17}$$

We implement the discretized model in (16) and (17) using daily returns, but all parameters will be expressed in annual units below. We now describe the volatility filtering step.

3.2 Volatility Filtering

The particle filter algorithm relies on the approximation of the true density of the state V_{t+1} by a set of N discrete points or particles that are updated iteratively through equations (16) and (17). The filter is implemented by deriving the empirical distribution $\left\{V_{t+1}^j, W_{t+1}^j\right\}_{j=1}^N$ of V_{t+1}, conditional on knowledge of the empirical distribution $\left\{V_t^j, W_t^j\right\}_{j=1}^N$ of V_t. This implementation proceeds in three steps. Our particular implementation of the particle filter is referred to as the Sampling-Importance-Resampling (SIR) particle filter. The three steps are as follows

3.2.1 Step 1: Simulating the state forward: Sampling

This is done by computing V_{t+1}^j from the original set of particles $\{V_t^j\}_{j=1}^N$ assumed to be known at time t using equation (17) and taking the correlation into account.[8] First let

$$w_{t+1}^j = \rho z_{t+1}^j + \sqrt{1-\rho^2}\varepsilon_{t+1}^j \tag{18}$$

where $corr(z_{t+1}^j, \varepsilon_{t+1}^j) = 0$. Substituting (18) and (16) into (17) we get

$$V_{t+1}^j = V_t^j + \kappa \left(V_t^j\right)^a \left(\theta - V_t^j\right) + \sigma \left(V_t^j\right)^b \left(\rho \frac{\ln\left(\frac{S_{t+1}}{S_t}\right) - \left(\mu - \frac{1}{2}V_t^j\right)}{\sqrt{V_t^j}} + \sqrt{1-\rho^2}\varepsilon_{t+1}^j\right) \tag{19}$$

This simulates N particles and thus provides a set of possible values of V_{t+1}.

3.2.2 Step 2: Computing and normalizing the weights: Importance Sampling

At this point, we have a vector of N possible values of V_{t+1} and we know according to equation (16) that given the other available information, V_{t+1} is sufficient to generate $\ln(S_{t+2})$. Therefore, equation (16) offers a simple way to evaluate the likelihood that the observation S_{t+2} has been generated

[7] See for example Eraker (2001).

[8] We set the variance in the first period equal to the model-implied unconditional variance, that is, $V_0^j = \theta$, for all j. In the returns-based MLIS estimation $t = 0$ is simply the first day of observed returns. In the options-based NLSIS estimation, $t = 0$ is one year prior to the first available option quote.

by V_{t+1}. Hence, we are able to compute the weight given to each particle (or the likelihood or probability that the particle has generated S_{t+2}). The likelihood is computed as follows

$$W_{t+1}^{j} = \frac{1}{\sqrt{V_{t+1}^{j}}} \exp\left(-\frac{1}{2}\frac{\left(\ln\left(\frac{S_{t+2}}{S_{t+1}}\right) - \left(\mu - \frac{1}{2}V_{t+1}^{j}\right)\right)^{2}}{V_{t+1}^{j}}\right) \tag{20}$$

for $j = 1, .., N$. Finally, we have to normalize the weights by via $W_{t+1}^{j} = \frac{W_{t+1}^{j}}{\sum_{j=1}^{N} W_{t+1}^{j}}$.

3.2.3 Step 3: Resampling

The motivation for this step is that we want to propagate high probability particles often. We use a simple technique to resample the particles, eliminating the low probability particles and replicating the high probability particles. To do so we construct a set of integer variables $\{\iota_{t+1}^{j}\}_{j=1}^{N}$ which can be obtained in different ways. Our implementation uses the resampling scheme proposed by Pitt (2002) which yields an objective function that is smooth in the parameters.

First, the adjusted weights obtained in Step 2, W_{t+1}^{j}, are mapped into a set of integer variables $\{\iota_{t+1}^{j}\}_{j=1}^{N}$, using an algorithm that takes into account that the weights are not multiples of $1/N$. This algorithm is based on the empirical CDF of V smoothed using linear interpolation as suggested by Pitt (2002). The smoothing enables gradient based optimization and the computation of standard errors using conventional first-order techniques.

Next, we construct the new set of particles $\{V(\iota)_{t+1}^{j}\}_{j=1}^{N}$ by replicating each particle in the original set $\{V_{t+1}^{j}\}_{j=1}^{N}$ ι_{t+1}^{j} times. Therefore, the particles in the original set are either eliminated, or included one or multiple times according to their adjusted weights $\{W_{t+1}^{j}\}_{j=1}^{N}$. The higher the weight, W_{t+1}^{j}, the higher the integer variable ι_{t+1}^{j}, and the more often the original particle V_{t+1}^{j} is included in the resampled set $\{V(\iota)_{t+1}^{j}\}_{j=1}^{N}$.

We now have a new set of N particles and weights $\{V(\iota)_{t+1}^{j}, W(\iota)_{t+1}^{j}\}_{j=1}^{N}$ which are implicitly functions of the variable ι_{t+1} and which all have weights $1/N$. Steps 1, 2 and 3 are repeated for $t = 1, ...T$.

Once the particles and weights have been computed for each date, we are ready to construct the filtered volatility path by

$$\bar{V}_{t+1} = \sum_{j=1}^{N} W_{t+1}^{j} V_{t+1}^{j} \tag{21}$$

for each t. Conditional on a set of structural parameters, the PF thus delivers a time series of filtered volatilities.

3.3 Maximum Likelihood Estimation

Our first empirical strategy uses a long sample of daily S&P500 index returns to estimate each of the SV models by maximizing the likelihood. To this end, we need an estimation methodology which allows for models with a latent volatility factor.

Pitt (2002) builds on Gordon, Salmond and Smith (1993) to show that the parameters of latent factor models in general, and of the SV model in (16) and (17) in particular, can be estimated by maximizing the Maximum Likelihood Importance Sampling criterion, which is simply defined by

$$MLIS\,(\mu,\kappa,\theta,\rho,\sigma) = \sum_{t=1}^{T} \ln\left(\frac{1}{N}\sum_{j=1}^{N} W_t^j\right)$$

As described in the previous section, the particle weights, W_t^j, are determined via the conditional likelihood of particle j at time t. These individual likelihoods in turn are given from the model specification in (16) and (17), taking into account the assumption that z_t and w_t are correlated normal random variables. The MLIS criterion then simply averages the particle weights across particles, takes logs, and sums over time to create a log likelihood function.

A key challenge in the use of the MLIS for estimation and inference is that it is not generally smooth in the underlying parameters. However, as discussed above, Pitt's (2002) ingenious implementation of the particle filter, where the resampling in Step 3 is done in a smooth fashion, ensures that the MLIS criterion is smooth in the parameters. This smoothing drastically improves the numerical optimization performance and it enables us to compute reliable parameter standard errors using conventional first-order techniques.

3.4 Nonlinear Least Squares Estimation

Our second empirical strategy is to take a large panel of options traded on the S&P500 index and estimate each of the SV models by minimizing the option pricing errors on this sample. For all the SV models, our implementation uses the nonlinear least squares importance sampling (NLSIS) estimation technique, which minimizes the following mean squared implied volatility error

$$IVMSE\,(\mu,\kappa,\theta,\rho,\sigma,\lambda) = \frac{1}{N^T}\sum_{t,i}\left(IV_{i,t} - BS_i^{-1}\left\{C_i(\bar{V}_t)\right\}\right)^2 \qquad (22)$$

where $IV_{i,t}$ is the Black-Scholes implied volatility corresponding to the market price of option i quoted on day t. $C_i\left(\bar{V}_t\right)$ is the model price evaluated at the filtered volatility, $\bar{V}_t$ from (21), and $BS^{-1}\left\{C_i(\bar{V}_t)\right\}$ denotes the Black-Scholes inversion of the model option price.[9] The total number of options in the sample is denoted by $N^T = \sum_{t=1}^{T} N_t$, where T denotes the total number of days included in the options sample, and where N_t is the number of options with various strikes prices and maturities included in the sample at date t.

Our setup that minimizes (22) is different from existing studies that estimate SV models from option data while solving the filtering problem. Most other studies are likelihood-based. One advantage of our approach is that it is relatively straightforward and fast, which allows us to

[9] Notice that we could alternatively compute the model price as the weighted average of the option prices computed for each particle. However, such an approach would be computationally very costly. Note that if the distribution of particles is centered around the mean, the two approaches will yield very similar results.

estimate using much more extensive cross-sections of options. In our opinion, estimation using (22) has an additional advantage, because matching the objective function used in parameter estimation with the function subsequently used to evaluate the models ensures the best possible performance of the models in- and out-of-sample. This is motivated by the insights of Granger (1969), and Weiss and Andersen (1984) who demonstrate that the choice of objective function (also labeled loss function) is an integral part of model specification. It follows that estimating a model using one objective function and evaluating it using another one amounts to a suboptimal choice of objective function. Christoffersen and Jacobs (2004) demonstrate that this issue is empirically relevant for the estimation of the deterministic volatility functions in Dumas, Fleming and Whaley (1998). We thus choose to implement the SV models in a way that is consistent with these insights. Notice however that our use of the particle filtering algorithm is completely general: we can apply this technique to any volatility model and using any well-behaved loss function involving option prices and underlying returns, including a likelihood function.

Our optimization algorithm minimizes (22) using an iterative procedure on the structural parameters. At each iteration, the volatility is filtered through time using the information embedded in observed returns and the structural parameters. Using the filtered volatility and the structural parameters, option prices are computed and the IVMSE is calculated. The procedure searches in the structural parameter space until the optimum is reached.

3.5 Monte Carlo Experiment

In this section we assess the finite sample performance of the relatively new MLIS estimation procedure based on the particle filter. Following Bates (2006), who introduces a new approximate ML estimation procedure, we add the MLIS estimator to the large-scale Monte Carlo study in Andersen, Chung and Sørensen (1999), henceforth ACS. ACS use the following simple SV model as a data generating process

$$\ln(S_{t+1}) = \ln(S_t) + \sqrt{V_t} z_{t+1} \tag{23}$$

$$\ln(V_{t+1}) = \omega + \phi \ln(V_t) + \sigma w_{t+1} \tag{24}$$

where z_{t+1} and w_{t+1} are uncorrelated. They consider a large number of estimators including the QML method from Harvey, Ruiz and Shephard (1994), the GMM from Andersen and Sørensen (1996), MCMC from Jacquier, Polson and Rossi (1994), as well as the EMM implementation suggested by ACS themselves. Table 1 reproduces the results from ACS and includes the approximate maximum likelihood (AML) technique from Bates (2006). We of course also include the MLIS estimation procedure from Pitt (2002) which is used in this paper. The various estimators are compared to the benchmark (but of course unrealistic) case where the spot variance is observed and standard ML estimation is straightforward. We report results for the sample size $T = 2,000$ which is relevant for our subsequent empirical study.

Table 1 reports parameter bias and root mean squared error (RMSE) for each estimator using 500 Monte Carlo replications. The MCMC estimator has the lowest absolute bias for the ω parameter followed by the AML and the MLIS. The MLIS, MCMC and AML estimators have the lowest absolute biases for ϕ. The AML, EMM and MCMC estimators have the lowest biases for the σ parameter followed by the MLIS. The MCMC and MLIS have the lowest RMSE for ω, AML, MCMC and MLIS have the lowest RMSE for ϕ, and MCMC has the lowest RMSE for the σ parameter followed by AML and MLIS. Overall, the MLIS estimator seems to perform well. The fact that it is also easily implementable across a wide range of models and objective functions makes it a very suitable estimation method in our large-scale empirical study below.

4 Empirical Results

This section presents our empirical results obtained using returns and options data. First, we discuss the return-based MLIS estimation results and differences in volatility sample paths. Subsequently, we provide plots of conditional moments. We then introduce the option data set and estimate the models using NLSIS estimation. Finally we analyze the patterns in the option valuation errors.

4.1 MLIS Estimation Results from Index Returns

As mentioned above, our first empirical strategy uses a long sample of daily S&P500 index returns and estimates each of the SV models by maximizing the model fit for this sample. First, recall the SV model specifications we consider

$$dV = \kappa V^a(\theta - V)dt + \sigma V^b dw, \text{ for } a = \{0, 1\} \text{ and } b = \{1/2, 1, 3/2\}.$$

We use daily S&P500 returns from CRSP and estimate the physical parameters by maximizing

$$MLIS\,(\kappa, \theta, \rho, \sigma) = \sum_{t=1}^{T} \ln\left(\frac{1}{N}\sum_{j=1}^{N} W_t^j\right)$$

Prior to MLIS optimization the return drift, μ is fixed at the sample average daily return in all models. The optimal parameters as well as the MLIS optimum values and the volatility properties for each model are given in Table 2. Consider first the MLIS objective values. We report three sets of values. First, we present results for 1996-2004, which matches the sample period we will use in the subsequent option estimation analysis. Second, we present results for 1989-2004, which is a longer sample period that does not include the 1987 crash, and finally results for 1985-2004, which includes the crash. Note first that the ranking of models is very stable across the various samples. The VAR ($a = 0, b = 1$) model or the 3/2N ($a = 1, b = 3/2$) model is always best or second best. The SQR ($a = 0, b = 1/2$) model is ranked 4th or 5th and the SQRN specification ($a = 1, b = 1/2$) is always worst. While we do not have inference procedures for these MLIS objective values for non-nested models, recall that in standard LR tests, adding one parameter to a model is significant at the 5% level if the log-likelihood increases by approximately two points. The differences in objective values across these models that have identical number of parameters therefore appear to be quite large.

Table 2 also contains the parameter estimates for the 1996-2004 sample period. The steady-state annualized variance is given by θ which ranges from 0.0352 in the SQR model to 0.0837 in the 3/2N model. These numbers correspond to an annualized steady state volatility of 18.8% to 28.9%. The rightmost four columns of Table 2 provide the first four sample moments of the filtered volatility. Note that the mean filtered volatility across models is much closer together than the population values inferred from θ. The κ parameter captures annualized variance persistence in the models. Note that the κ parameters are not directly comparable between linear ($a = 0$) and nonlinear ($a = 1$) drift specifications. Figure 3 therefore plots the drift function for all models (solid lines) as a function of the level of variance. The linear drift specifications are given in the left-hand panels and the nonlinear drifts in the right-hand panels. The square root diffusions are in the top row, the linear diffusions in the middle row and the 3/2N diffusions in the bottom row. The differences between linear and nonlinear drift specification are most evident for large values of the spot variance where the mean-reversion in the nonlinear models is much stronger.

The diffusion parameter σ is comparable within a given diffusion specification (i.e. for a given value of b), but not across diffusion specifications. In order to facilitate comparisons of the models, Figure 3 shows the diffusion functions (dashed lines) plotted against the value of the spot variance V. Notice that the SQR specifications ($b = 1/2$) in the top row enable much less diffusion in the variance when the variance level is large, when comparing with the VAR diffusion specification ($b = 1$) in the middle row and the 3/2 specifications ($b = 3/2$) in the bottom row. The distribution of the annualized filtered volatility can also be gauged from the standard deviation, skewness, and kurtosis of the $\sqrt{\bar{V}_t}$ paths provided in Table 2. Note that not only is the standard deviation of volatility different across models, the VAR, 3/2 and 3/2N models also have much larger skewness and kurtosis of volatility than does the SQR model.

The final parameter estimate is that of ρ which captures the correlation between the shocks to return and variance. Ranging from –0.7876 to –0.7411, the estimate of ρ is stable across models and relatively large in magnitude. However, recent studies including Jones (2003) and Aït-Sahalia and Kimmel (2007) have obtained similarly large correlation estimates.

4.2 Conditional Moment Dynamics

We now present some more empirical evidence on the differences in models over time. We first plot the filtered volatility over time, subsequently we define and plot the conditional leverage effect which determines conditional skewness, and then we define and plot the conditional volatility of variance which drives conditional kurtosis.

Figure 4 plots the filtered volatility paths $\sqrt{\bar{V}_t}$ for each model during the 1996 to 2004 sample period. All volatility paths are shown on the same scale going form zero to 70 percent volatility in annual terms. Naturally, the overall pattern in volatility over time is similar across models. However, when volatility increases it tends to do so much more sharply in the 3/2 models and somewhat more sharply in the linear (VAR) diffusion models when compared with the square root diffusion models. The VAR and 3/2 diffusions thus exhibit more spikes in volatility when compared with the SQR model. Example of this include September 1998 (LTCM/Russia default), September 2001 (9/11), and July 2002 (bursting of the dot com bubble).

We define the conditional leverage effect in a given model as the conditional covariance between the index return and the variance in the model. We can approximate the conditional leverage in each model as

$$Cov_t(\ln(S_{t+1}), \bar{V}_{t+1}) = \rho\sigma\bar{V}_t^{b+1/2}$$

Notice that the differences in the specification of the variance diffusion term across models translate into different conditional leverage effects. In particular, the power on V varies between the square root, linear and 3/2 diffusion specifications.

Figure 5 shows the path for the annualized conditional leverage for each model. Again we use the same scale for all the plots so as to emphasize the differences across models. Notice how high-volatility episodes such as the September 1998 LTCM debacle and Russia default and the July 2002 stock market decline lead to sharp differences between the benchmark models, namely the SQR model in the top-left panel, the VAR model in the middle-left panel, and the 3/2N model in the bottom-right panel.

We can also compare the models in terms of their conditional variance of variance properties. We get

$$Var_t(V_{t+1}) = \sigma^2\bar{V}_t^{2b}. \tag{25}$$

Figure 6 shows the path for the annualized conditional volatility of variance for each model, i.e. the square root of the expression in (25). Notice again the sharp differences across models during the high volatility episodes. Again the differences across the benchmark SQR, VAR and 3/2N models are quite striking.

4.3 NLSIS Estimation on Option Prices and Returns

This section presents the results from the option-based NLSIS estimation. We first describe the option data and then present the parameter estimates and differences in fit between the models. For comparison we also report results on option fitting using the return-based MLIS estimates from Table 2.

We conduct our empirical option-based analysis using S&P500 index call options for the 1996-2004 period. We only use Wednesday and Thursday options data. For the in-sample analysis, we use the Wednesday data. Wednesday is the day of the week least likely to be a holiday. It is also less likely than other days such as Monday and Friday to be affected by day-of-the-week effects.

The decision to pick one day every week is to some extent motivated by computational constraints. The optimization problems are fairly time-intensive, and limiting the number of options reduces the computational burden. Using only Wednesday data allows us to study a long time-series, which is useful considering the highly persistent volatility processes. An additional motivation for only using Wednesday data is that following the work of Dumas, Fleming and Whaley (1998), several studies have used this setup.[10]

Panel A of Table 3 presents descriptive statistics for the options data for the 1996-2004 Wednesday in-sample data by moneyness and maturity. We estimate the models on a total of 16,506 contracts with an average call price of \$46.05 and average implied volatility of 20.32%. The implied volatility is largest for the in-the-money options reflecting the well-known volatility smirk in index options. The average implied volatility term structure is roughly flat during the period. Panel B of Table 3 shows that the Thursday sample used for out-of-sample valuation has 15,390 options with similar characteristics as the in-sample Wednesday data. We use the same calendar period for the out-of-sample study so as to avoid the impact of structural breaks in unconditional volatility. Because we are comparing equally parsimonious models estimated on a large sample, the Bates (2003) critique mentioned in Section 2, that more heavily parameterized models are favored in such out-of-sample experiments, does not apply.

Table 4 contains the parameter values obtained from minimizing the option implied volatility mean-squared-error defined above as

$$IVMSE\left(\kappa,\theta,\rho,\sigma,\lambda\right)=\frac{1}{N^{T}}\sum_{t,i}\left(IV_{i,t}-BS_{i}^{-1}\left\{C_{i}(\bar{V}_{t})\right\}\right)^{2} \tag{26}$$

Prior to NLSIS optimization the return drift, μ is again fixed at the sample average daily return in all models.

Consider first the in-sample root mean squared error (IVRMSE) column. We see that the SQR model performs the worst with an IVRMSE of 3.32% and the VAR model the best with an IVRMSE of 2.85%. The RMSE thus drops by about 14% (see the Ratio column) going from the SQR to the VAR model. Note again that this is without adding any parameters to the model. The 3/2N model is about 11% better than the SQR model and the VARN model with nonlinear drift and linear diffusion ($a=1,b=1$) is about 10% better than the SQR model. The remaining two models are only marginally better than the SQR model. Out of sample the overall IVRMSEs are slightly higher for all models but the relative differences are similar. The VAR model is about 17% better than the SQR model and the 3/2N model is about 11% better. The VARN model with ($a=1,b=1$) is about 9% better and the remaining two models improve upon the SQR model by less than 5%.

The improvement in IVRMSE of the VAR model over the benchmark SQR model is quite substantial, and consistent with the finding from the return-based estimation in Table 2. The most natural reference point is the rich literature that uses Poisson jumps in returns and/or volatility in

[10] See for instance Heston and Nandi (2000).

conjunction with an SQR stochastic volatility model. The evidence on the in-sample and especially the out-of-sample improvement provided by including the jump processes is inconclusive, with some studies finding moderate improvements, and others concluding that there is no improvement in fit. See for instance Bates (2000), Pan (2002) and Eraker (2004). In recent work, using a very different empirical setup, Broadie, Chernov and Johannes (2007) find large improvements in fit when including Poisson jumps. The improvement in fit from adopting the VAR model over the SQR model is roughly similar in-and out-of-sample. It is also consistent with the evidence from returns and realized variances discussed earlier.

The improvement in fit of the 3/2N model over the SQR model is less impressive but still substantial. Recall that the most impressive return-based performance of the 3/2N model in Table 2 was when the 1987 crash was included in the sample. Our evidence on the 3/2N model is interesting in light of the findings in Jones (2003) and Aït-Sahalia and Kimmel (2007). These papers both estimate a CEV model using a bivariate time series of index returns and options. However, they obtain very different results. Jones (2003) estimates a CEV parameter of 1.33 using a 1986-2000 sample, while Aït-Sahalia and Kimmel's (2007) estimate for the CEV parameter is 0.65. Our results suggest that one potential reason for these conflicting results is the different samples used in these studies: Aït-Sahalia and Kimmel's sample is 1990-2003, and therefore does not include the 1987 crash. However, it must be noted that Jones (2003) also estimates a CEV parameter of 1.17 using a 1988-2000 sample.

Comparing the option-based NLSIS estimates in Table 4 with the return-based estimates in Table 2 we find that the κ is generally smaller indicating a slower mean-reversion of variance when options are driving the parameters. The long-run variance θ is generally lower in Table 4 with the SQR model as the notable exception. The volatility of variance parameter σ is generally lower in Table 4, the notable exception again being the SQR model. The volatility risk premium parameter λ is generally small but positive for all models so that the unconditional variance is higher under the risk-neutral measure as expected. Finally, the correlation coefficient, ρ, is slightly smaller (in absolute value) when estimated using options.

In order to assess the economic difference between the NLSIS estimates and the MLIS estimates, the penultimate column of Table 4 reports the IVRMSEs from the Wednesday options when using the MLIS-based parameters in each model.[11] Not surprisingly, when compared with the NLSIS-based in-sample IVRMSEs, the MLIS-based errors are clearly considerably larger except for in the 3/2 model where the difference is negligible. Comparing across models in the MLIS-based column, it is clear that the SQR model is still worst. The VAR model is about 13% better and the 3/2N model 14% better than the SQR model. The 3/2 model now performs the best with an improvement of 20% over the SQR model. Somewhat surprisingly, the SQRN model performs relatively well in this experiment improving upon the SQR model by 15%.

[11] The volatility risk premium parameter, λ, is not identified when estimating the models on returns only. Furthermore, it is quantitatively small when estimated in Table 4. Thus, we simply set it to zero in this experiment.

Finally, the footnote to Table 4 reports the IVRMSE from two benchmark models. First, the simple Black-Scholes model where volatility is kept constant at the average implied volatility across all options and across the entire sample period. The resulting IVRMSE is 4.95% for the Wednesday sample and 5.22% for the Thursday sample. Note that, not surprisingly, all the SV models outperform the simple Black-Scholes model by a wide margin. The second benchmark is the so-called ad-hoc Black-Scholes model where the implied volatility for each option is set to the average implied volatility across options on the same weekday of the previous week. This ad-hoc model has an IVRMSE of 2.99% for the Wednesday sample and 3.06% for the Thursday sample. Thus only the VAR model estimated using NLSIS outperforms this benchmark in both samples. This result may be surprising but note that the ad-hoc model allows for the estimation of $9*52 = 468$ parameters versus only five parameters in the SV models. These results are not easily comparable with other recent studies as the weekly estimated ad-hoc benchmark is not often compared with SV models estimated on a long sample. Heston and Nandi (2000) provide a similar benchmark, but their SV models are estimated on much shorter samples.

4.4 Decomposing the Option Valuation Errors

Tables 5 and 6 provide more detail on the option pricing fit of the various SV models. In Table 5 we report the in-sample RMSE by moneyness, maturity, and volatility level. The moneyness and maturity decompositions are as in Table 3 and the volatility level decomposition is according to the VIX index. We sort the data set into four bins using the quartiles of the CBOE VIX index for the sample period. The out-of-sample results are very similar and thus omitted from the tables.

Consider first the IVRMSE by moneyness in Panel A of Table 5. Note that the strong overall performance of the VAR model in Table 4 results from better than average performance in all moneyness categories. The VAR model is best at pricing deep-in-the money calls. The 3/2N model is also better than average for all categories and is particularly good at pricing out-of-the-money calls. The SQR model is below average in all moneyness categories and it is particular bad for in-the-money options.

Panel B of Table 5 reports the IVRMSE by maturity categories. The VAR model performs best for all but the longest maturities where the VARN model ($a = 1, b = 1$) is a bit better. The 3/2N model is better than average for all moneyness categories. The SQR model is below average for all moneyness categories and is the worst model in the two medium-term categories. Note in general that the nonlinear drift specifications do well for long maturity options.

Panel C of Table 5 reports the IVRMSE by VIX level categories. Rather than sorting by cross-sectional characteristics, we split the sample days into four categories sorted by the VIX level on each day of the sample. The VAR model is better than average for all VIX levels. The same is true for the 3/2N models. The SQR model is worst for the three categories with the highest volatility levels and better than average only for the lowest volatility category. Not surprisingly the IVRMSE tends to increase with the level of the VIX in all models.

To summarize Table 5: In the twelve categories considered, the VAR model performed better than average in all twelve and best in seven. The 3/2N model also performed better than average in all twelve and best in four. The SQR model performed below average in eleven and worst in seven categories.

Note in general across models that the IVRMSE is increasing in moneyness, which may be partly driven by the average implied volatility simply being increasing in moneyness (Table 3). Note also that the IVRMSE tends to decrease with maturity for all models. This is only partially explained by the average implied volatility being larger for short-term options (Table 3) as this effect is small. A likely explanation for both of these phenomena is that we have ruled out return jumps in the model specifications. Negative return jumps are likely to have most effect on short-term and deep-in-the-money call options.

Table 6 reports the bias (average market price less average model price) across moneyness, maturity and volatility level categories defined as in Table 5. Consider first the "All" column which shows that the SQR model has a large positive bias and thus underprices options on average. The overall biases are very small in all other models except for the 3/2 ($a = 0, b = 3/2$) model which tends to overprice options on average. Panel A reports the bias by moneyness. Note that the SQR model underprices all options and that all models underprice the deep in-the-money calls. All but the SQR tend to overprice the out-of-the money calls. The model errors tend to "smirk" and appear to leave a role for jumps.

Panel B of Table 6 reports the bias by maturity. The SQR model underprices all categories and the 3/2 model overprices all categories. The remaining models tend to overprice short-term options and underprice long-term options but the biases are generally small.

Panel C of Table 6 reports the bias by volatility level. The SQR model underprices all categories except the lowest volatility category. The VAR and 3/2N models both overprice slightly in the lowest volatility category and overprice in the highest volatility category. An even more flexible volatility specification may therefore be warranted.

5 Summary and Conclusions

This paper provides an empirical comparison of the affine SQR model of Heston (1993) with a range of non-affine but equally parsimonious option valuation models. An exploratory analysis using realized volatility data suggests that the SQR model is misspecified, and subsequent estimation on index returns and option prices confirms this conclusion. Based on the likelihood values, non-affine models in general are superior when estimated using a long sample of daily index returns. The VAR model consistently outperforms the SQR model based on an RMSE comparison when estimating the models using comprehensive panels of option contracts. While the focus of the option valuation literature on affine models is well motivated, because the resulting closed-form solutions are extremely convenient, our results suggest that this analytical convenience comes at a price, and non-affine models need to be studied more extensively.

At the methodological level, this paper uses a method to estimate continuous-time option valuation models that has not yet been applied to this particular problem. We use particle filtering, which is rather flexible and straightforward to implement. It can easily be used to investigate option data and underlying equity returns jointly for a wide range of objective functions as well as for a wide range of models. In our opinion, this method is extremely attractive compared to other statistical frameworks that have been used in this literature.

Our empirical results suggest a couple of extensions. First, it may prove interesting to investigate how the popular class of Poisson jump models can help improve the fit of non-affine stochastic volatility models. In particular, a comparison with available results for the SQR model may be worthwhile. An analysis of the Levy processes in Carr and Wu (2004) and Huang and Wu (2004) may also prove interesting. Second, the finding from the returns data that the 3/2N model is particularly useful in a sample that includes the 1987 crash, indicates that another interesting avenue for future work is to focus more explicitly on the CEV model, and particularly to investigate its performance when periodically re-estimating the CEV parameter.

References

Aït-Sahalia, Y. (1996), Testing Continuous-Time Models of the Spot Interest Rate, Review of Financial Studies, 9, 385-426.

Aït-Sahalia, Y. and R. Kimmel (2007), Maximum Likelihood Estimation of Stochastic Volatility Models, Journal of Financial Economics, 83, 413-452.

Aït-Sahalia, Y., P. Mykland, and L. Zhang (2005), How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise, Review of Financial Studies, 18, 351-416.

Andersen, T.G., L. Benzoni, and J. Lund (2002), Estimating Jump-Diffusions for Equity Returns, Journal of Finance, 57, 1239-1284.

Andersen, T.G., T. Bollerslev, F.X. Diebold, and H. Ebens (2001), The Distribution of Realized Stock Return Volatility, Journal of Financial Economics, 61, 43-76.

Andersen, T.G., H. Chung, and B.E. Sørensen (1999), Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study, Journal of Econometrics, 91, 61-87.

Andersen, T.G. and B.E. Sørensen (1996), GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study, Journal of Business and Economic Statistics, 14, 328-352.

Bakshi, G., C. Cao, and Z. Chen (1997), Empirical Performance of Alternative Option Pricing Models, Journal of Finance, 52, 2003-2049.

Barone-Adesi, G., H. Rasmussen, and C. Ravanelli (2003), An Option Pricing Formula for the GARCH Diffusion Model, Manuscript, University of Southern Switzerland.

Bates, D. (1996), Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options, Review of Financial Studies, 9, 69-107.

Bates, D. (2000), Post-'87 Crash Fears in the S&P 500 Futures Option Market, Journal of Econometrics, 94, 181-238.

Bates, D. (2003), Empirical Option Pricing: A Retrospection, Journal of Econometrics, 116, 387-404.

Bates, D. (2006), Maximum Likelihood Estimation of Latent Affine Processes, Review of Financial Studies, 19, 909-965.

Benzoni, L. (2002), Pricing Options Under Stochastic Volatility: An Empirical Investigation, Manuscript, University of Minnesota.

Black, F. (1976), Studies of Stock Price Volatility Changes, in: Proceedings of the 1976 Meetings of the Business and Economic Statistics Section, American Statistical Association, 177-181.

Black, F., and M. Scholes (1973), The Pricing of Options and Corporate Liabilities, Journal of Political Economy, 81, 637-659.

Broadie, M., Chernov, M. and M. Johannes (2007), Model Specification and Risk Premiums: The Evidence from the Futures Options, Journal of Finance, 62, 1453-1490.

Carr, P. and L. Wu (2004), Time-Changed Levy Processes and Option Pricing, Journal of Financial Economics, 17, 113–141.

Chacko, G. and L. Viceira (2003), Spectral GMM Estimation of Continuous-Time Processes, Journal of Econometrics, 116, 259-292.

Chernov, M., A.R. Gallant, E. Ghysels, and G. Tauchen (2003), Alternative Models for Stock Price Dynamics, Journal of Econometrics, 116, 225-257.

Chernov, M. and E. Ghysels (2000), A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Option Valuation, Journal of Financial Economics, 56, 407-458.

Christoffersen, P. and K. Jacobs (2004), The Importance of the Loss Function in Option Valuation, Journal of Financial Economics, 72, 291-318.

Duan, J.-C. and J.-G. Simonato (1998), Empirical Martingale Simulation for Asset Prices, Management Science, 44, 1218-1233.

Dumas, B., J. Fleming, and R. Whaley (1998), Implied Volatility Functions: Empirical Tests, Journal of Finance, 53, 2059-2106.

Eraker, B. (2001), MCMC Analysis of Diffusion Models with Application to Finance, Journal of Business and Economic Statistics, 19, 177-191.

Eraker, B. (2004), Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices, Journal of Finance, 59, 1367-1403.

Eraker, B., M. Johannes, and N. Polson (2003) The Role of Jumps in Returns and Volatility, Journal of Finance, 58, 1269-1300.

Gordon, N.J., D.J. Salmond, and A.F.M. Smith (1993), A Novel Approach to nonlinear/non-Gaussian Bayesian state estimation, IEE-Proceedings F, 140, 107-133.

Granger, C. (1969), Prediction with a generalized cost of error function, Operations Research Quarterly, 20, 199-207.

Harvey, A., E. Ruiz, and N. Shephard (1994), Multivariate Stochastic Variance Models, Review of Economic Studies, 61, 247-264.

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

Heston, S. (1997), A Simple New Formula for Options with Stochastic Volatility, Manuscript, John M. Olin School of Business, Washington University.

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

Huang, J.-Z. and L. Wu (2004), Specification Analysis of Option Pricing Models Based on Time-Changed Levy Processes, Journal of Finance, 59, 1405–1439.

Jacquier, E., N.G. Polson, and P.E. Rossi (1994), Bayesian Analysis of Stochastic Volatility Models, Journal of Business and Economic Statistics, 12, 1-19.

Johannes, M., N. Polson, N. and J. Stroud (2002), Nonlinear Filtering of Stochastic Differential Equations with Jumps, Manuscript, University of Chicago.

Jones, C. (2003), The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets, Journal of Econometrics, 116, 181-224.

Karlin, S. and H.M. Taylor (1981), A Second Course in Stochastic Processes, Academic Press, San Diego, CA.

Lewis, A. (2000), Option Valuation Under Stochastic Volatility, Finance Press, Newport Beach, CA.

Pan, J. (2002), The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study, Journal of Financial Economics, 63, 3-50.

Pitt, M. (2002), Smooth Particle Filters for Likelihood Evaluation and Maximization, Manuscript, University of Warwick.

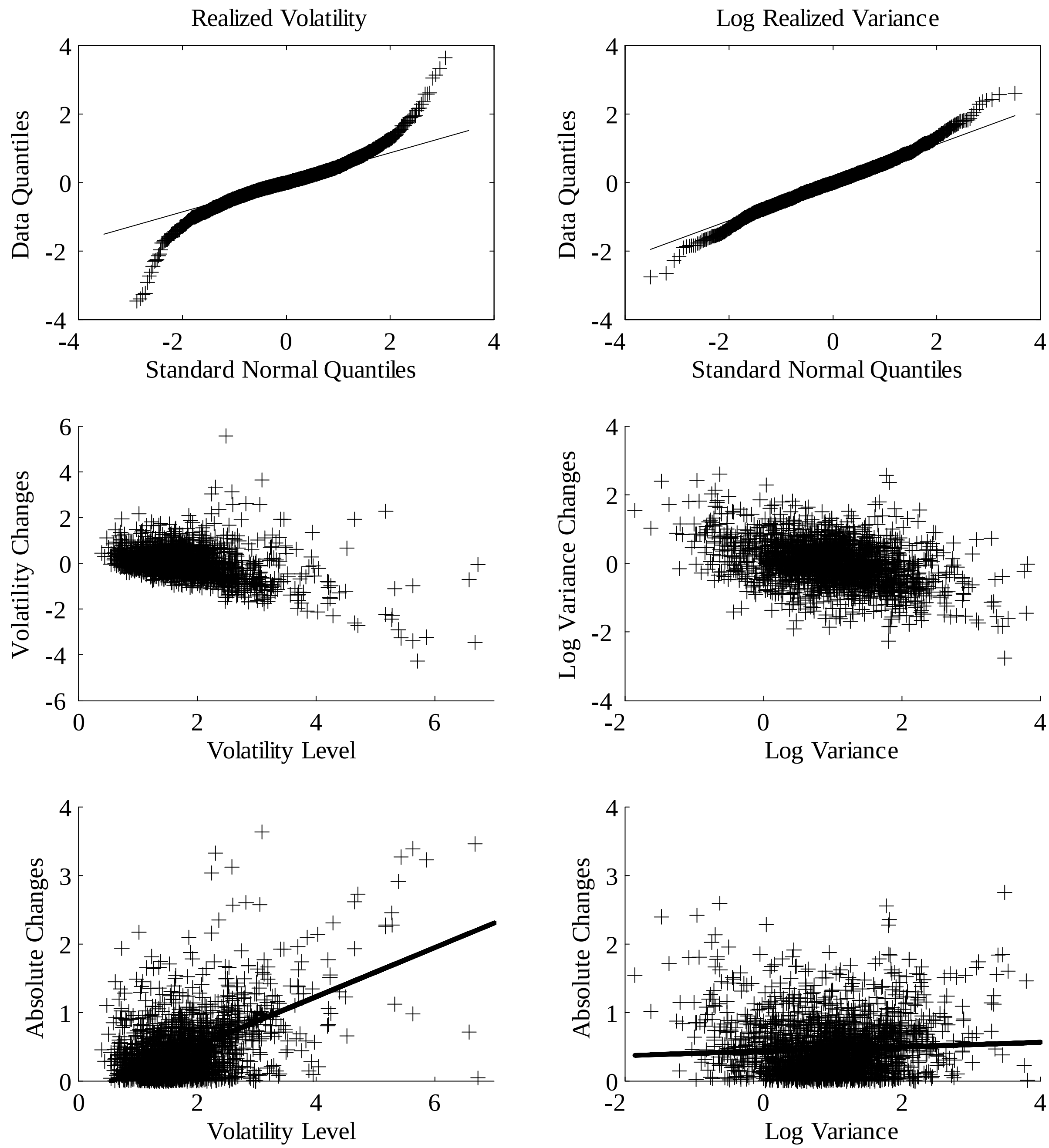
Pitt, M., and N. Shephard (1999), Filtering via Simulation: Auxiliary Particle Filters, Journal of the American Statistical Association, 94, 590-599.

Sabanis, S.(2003), Stochastic Volatility and the Mean Reverting Process, The Journal of Futures Markets, 23, 33–47.

Weiss, A., and A. Andersen (1984), Estimating time series models using the relevant forecast evaluation criterion, Journal of the Royal Statistical Society, Series A, 147, 484-487

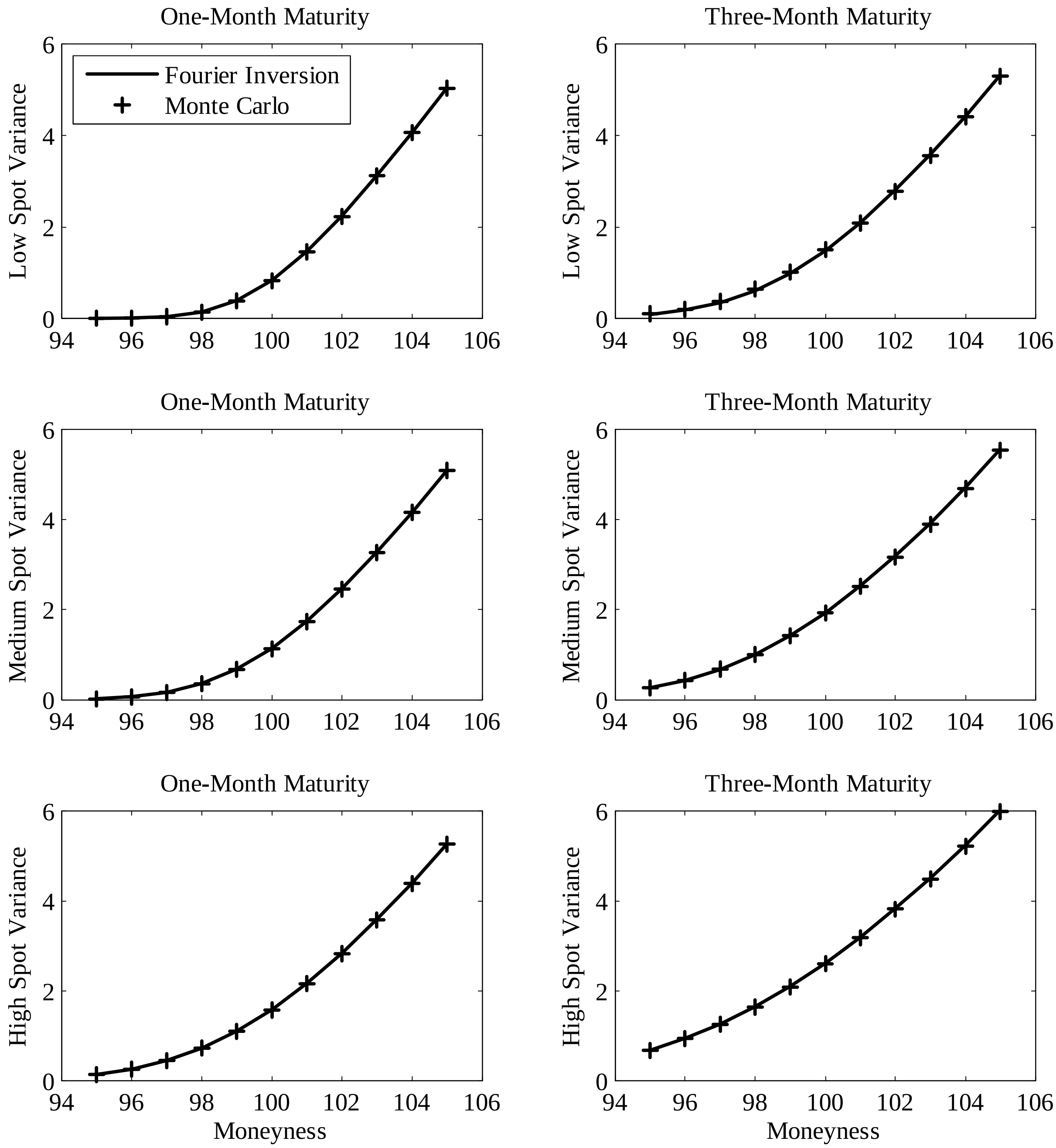
Zhang, L., P.A. Mykland, and Y. Aït-Sahalia (2005), A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data, Journal of the American Statistical Association, 100, 1394-1411.

Figure 1: Diagnostics of Realized Volatility and Log Realized Variance. 1996-2004.



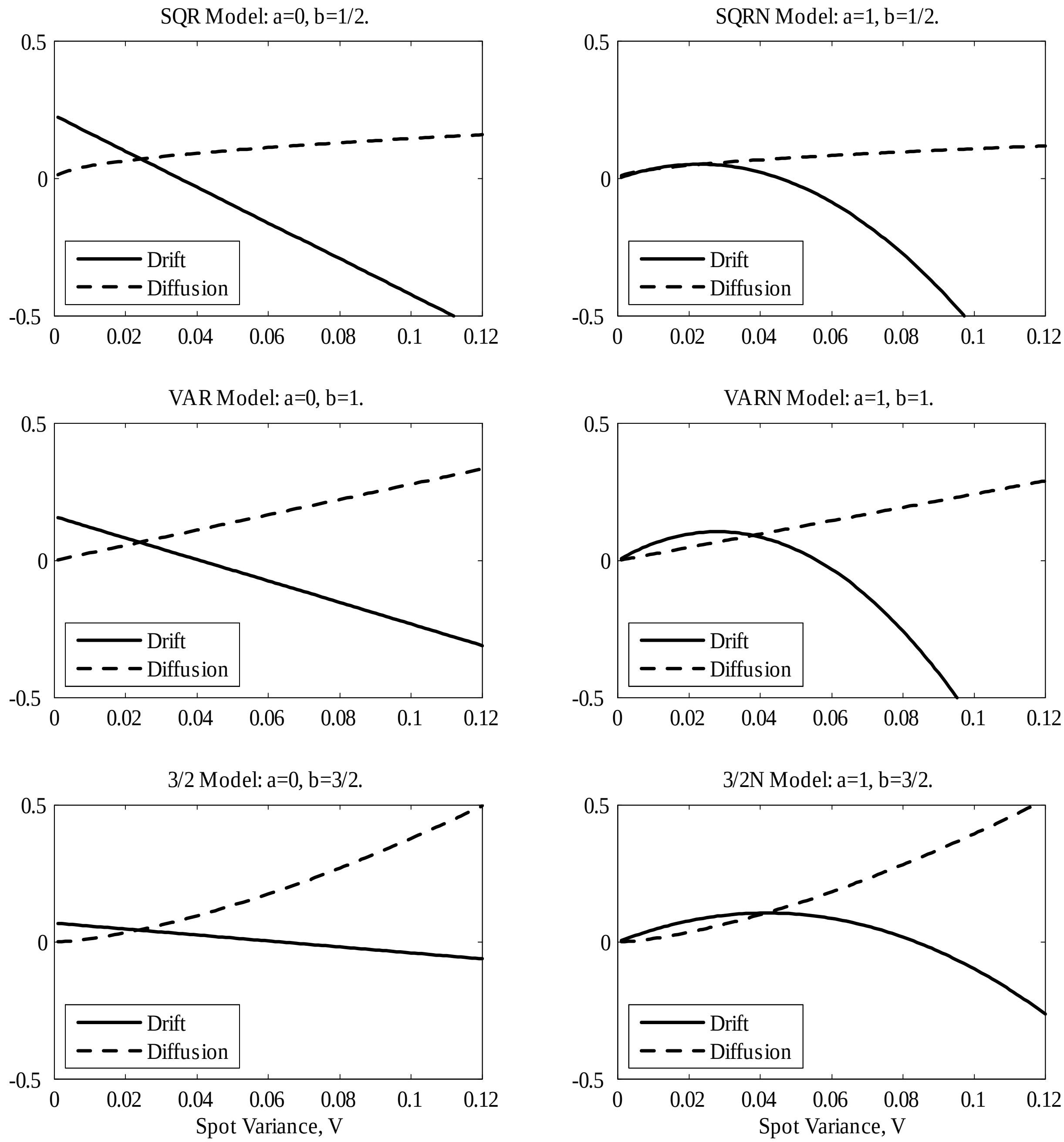
Notes to figure: Using realized volatilities (left panels) and log realized variances (right panels) from 1996 through 2004, the top panel plots the quantiles of the daily changes against quantiles from the normal distribution. The middle panels scatter plot the daily changes against the daily levels. The bottom panels scatter plot the absolute daily changes against the daily levels and show an OLS regression line for reference.

Figure 2: Analytical (Fourier Inversion) and Monte Carlo Prices for the Heston (1993) Model.



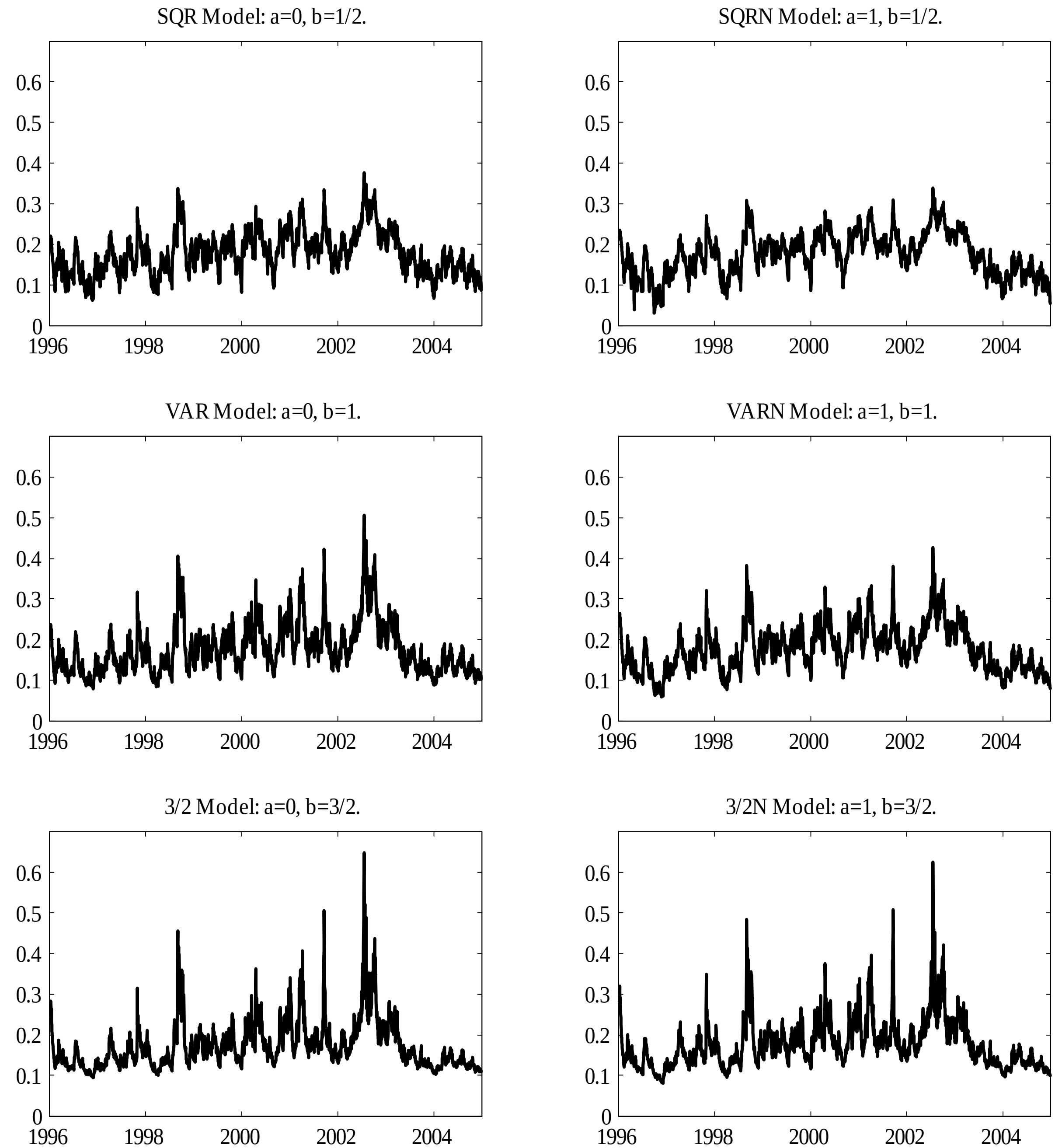
Notes to figure: We first compute option prices from Heston's (1993) model using the Fourier inversion technique (solid) and then using Monte Carlo simulation ('+'). We consider two maturities: one month (left panels) and three months (right panels). We consider three spot variance levels: Half the unconditional variance (top panels), equal to the unconditional variance (middle panels) and twice the unconditional variance (bottom panels).

Figure 3: Drift and Diffusion Functions for Various SV Models.



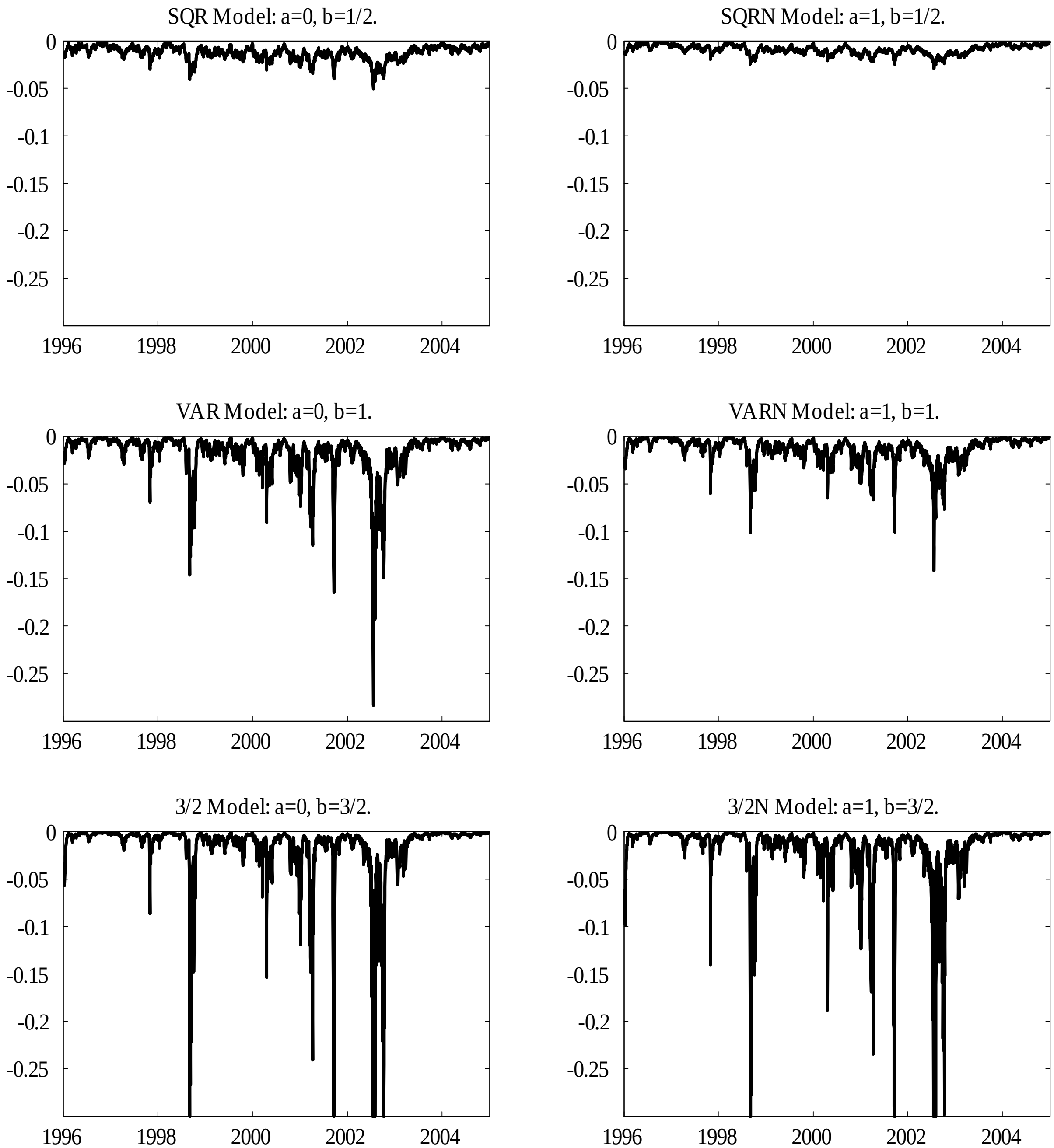
Notes to figure: The solid lines denote the drift function $\kappa V^a (\theta - V)$ plotted against the level of spot variance, V. The dashed lines denote the diffusion function σV^b plotted against the level of spot variance. The parameter estimates are from Table 2 using daily returns from 1996 through 2004.

Figure 4: Spot Volatility Paths for Various SV Models. 1996-2004.



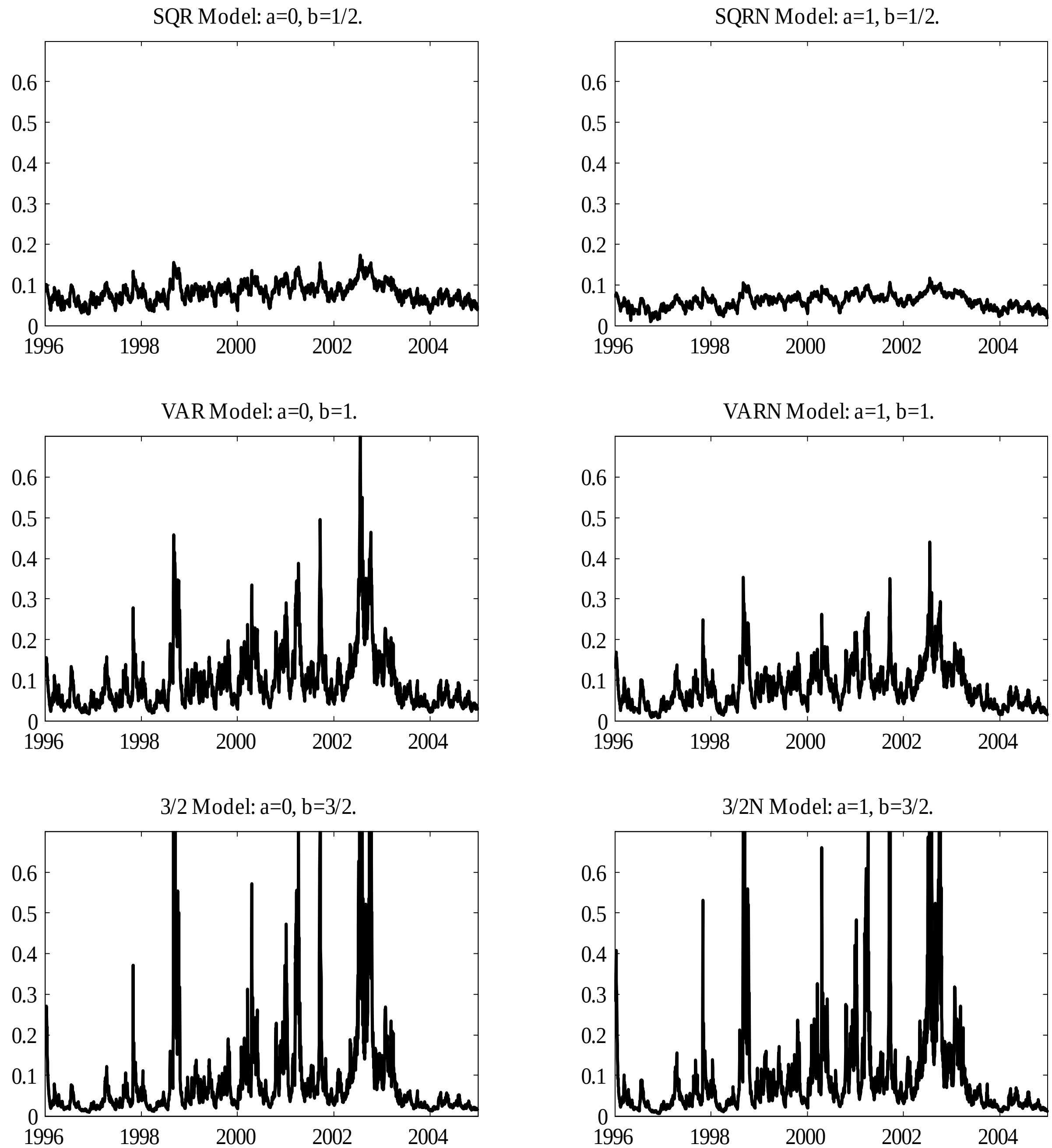
Notes to figure: For each SV model of the form $dV = \kappa V^a(\theta - V)dt + \sigma V^b dw$, we plot the annualized daily filtered spot volatility path, $\sqrt{\bar{V}_t}$ during 1996-2004. The parameters are from the MLIS estimation on daily S&P500 returns from 1996 through 2004 as reported in Table 2.

Figure 5 Conditional Leverage Paths for Various SV Models. 1996-2004.



Notes to figure: For each SV model of the form $dV = \kappa V^a(\theta - V)dt + \sigma V^b dw$, we plot the annualized daily conditional leverage path defined as the conditional covariance between shocks to returns and shocks to variance, $\rho\sigma\bar{V}_t^{b+1/2}$. The parameters are from the MLIS estimation on daily S&P500 returns from 1996 through 2004 as reported in Table 2. The vertical axes have been truncated for the 3/2 diffusions in order to facilitate comparisons with the other models.

Figure 6: Conditional Volatility of Variance Paths for Various SV Models. 1996-2004.



Notes to figure: For each SV model of the form $dV = \kappa V^a(\theta - V)dt + \sigma V^b dw$, we plot the annualized daily conditional volatility of variance path defined as the square root of the conditional variance of the variance of returns, $\sigma \bar{V}_t^b$. The parameters are from the MLIS estimation on daily S&P500 returns from 1996 through 2004 as reported in Table 2. The vertical axes have been truncated for the 3/2 diffusions in order to facilitate comparisons with the other models.

Table 1: Monte Carlo Study of MLIS and Alternative Estimation Methods.

Estimators	
ML|V	ML conditional on observing the true spot variance
QML	Quasi ML as in Harvey, Ruiz, and Shephard (1994)
GMM	Andersen and Sørensen (1996)
EMM	Efficient method of moments as in Andersen, Chung and Sørensen (1999)
AML	Approximate ML of Bates (2006)
MCMC	Markov Chain Monte Carlo as in Jacquier, Polson and Rossi (1994)
MLIS	Method used in this paper. See Pitt (2002)

Parameter	ω	ϕ	σ
True values	-0.736	0.9	0.363
Estimators		**Bias**	
ML|V	-0.015	-0.002	0.000
QML	-0.117	-0.020	0.020
GMM	0.150	0.020	-0.080
EMM	-0.057	-0.007	-0.004
AML	-0.039	0.005	0.005
MCMC	-0.026	-0.004	-0.004
MLIS	**0.049**	**0.004**	**-0.017**
Estimators		**RMSE**	
ML|V	0.076	0.010	0.006
QML	0.460	0.060	0.110
GMM	0.310	0.040	0.120
EMM	0.224	0.030	0.049
AML	0.173	0.023	0.043
MCMC	0.150	0.020	0.034
MLIS	**0.157**	**0.020**	**0.046**

Notes: Following Bates (2006), we compare the MLIS estimator to those considered in the large scale Monte Carlo study by Andersen, Chung and Sørensen (1999). We generate 500 Monte Carlo samples of 2,000 returns with zero drift and stochastic volatility following the logaritmic SV model in (23) and (24) with constant term ω, persistence ϕ, and diffusion parameter σ. We then estimate the model using MLIS and compare the bias and RMSE from MLIS to the bias and RMSE reported in Bates (2006) and Andersen, Chung and Sørensen (1999).

Table 2: Parameter Estimates from Maximum Likelihood Importance Sampling on Returns.

Model			Parameter Estimates for 1996-2004				MLIS Objective Values			Filtered Volatility, 1996-2004			
Name	a	b	κ	θ	σ	ρ	1996-2004	1989-2004	1985-2004	Mean	StdDev	Skewness	Kurtosis
SQR	0	1/2	6.5200	0.0352	0.4601	-0.7710	7,064.7	13,359.3	16,595.0	17.698	5.166	0.487	0.086
Standard Errors:			1.1096	0.0026	0.0309	0.0375							
SQRN	1	1/2	100.0291	0.0457	0.3425	-0.7527	7,045.1	13,328.9	16,532.2	17.495	5.172	0.043	-0.411
Standard Errors:			14.4534	0.0038	0.0193	0.0234							
VAR	0	1	3.9248	0.0408	2.7790	-0.7876	7,074.5	13,372.5	16,631.7	17.673	6.146	1.302	2.120
Standard Errors:			1.1392	0.0067	0.1949	0.0345							
VARN	1	1	133.9347	0.0560	2.4188	-0.7559	7,066.1	13,362.5	16,615.4	17.708	5.569	0.581	0.316
Standard Errors:			25.3311	0.0053	0.1654	0.0417							
3/2	0	3/2	1.0852	0.0633	11.9534	-0.7411	7,064.9	13,352.9	16,625.8	17.294	5.863	1.989	6.341
Standard Errors:			0.8260	0.0351	0.9125	0.0432							
3/2N	1	3/2	60.1040	0.0837	12.4989	-0.7591	7,068.8	13,362.9	16,638.4	17.683	5.893	1.616	4.659
Standard Errors:			23.9651	0.0247	0.8575	0.0426							

Notes: For each model, we estimate the parameters using Maximum Likelihood Importance Sampling (MLIS) on daily returns. The reported parameter estimates use data from January 4, 1996 to December 31, 2004. The return drift parameter μ is fixed at the sample average return of 0.091 for all models. The parameters are reported in annual units. Standard errors are computed using the outer product of the gradient at the optimal parameter values. The MLIS objective value is reported for estimations on longer samples as well. The first four moments of the distribution of the annualized filtered volatility is reported in the last four columns of the table.

Table 3: S&P500 Index Call Option Data. 1996-2004. In-and Out-of-Sample.

Panel A. Option Data Charateristics by Moneyness and Maturity. In-Sample

	S/X<0.975	0.975<S/X<1	1<S/X<1.025	S/X>1.025	All
Number of Contracts	5,761	3,859	3,077	3,809	16,506
Average Call Price	29.73	37.22	47.05	78.85	46.05
Average IV	19.35	19.37	20.13	22.67	20.32

	DTM<30	30<DTM<90	90<DTM<180	DTM>180	All
Number of Contracts	2,552	7,829	2,992	3,133	16,506
Average Call Price	34.82	40.09	50.83	65.51	46.05
Average IV	20.60	20.33	20.21	20.15	20.32

Panel B. Option Data Charateristics by Moneyness and Maturity. Out-of-Sample

	S/X<0.975	0.975<S/X<1	1<S/X<1.025	S/X>1.025	All
Number of Contracts	4,695	3,775	3,072	3,848	15,390
Average Call Price	30.29	37.29	46.59	78.52	47.32
Average IV	19.26	19.34	20.05	22.78	20.38

	DTM<30	30<DTM<90	90<DTM<180	DTM>180	All
Number of Contracts	3,782	6,541	2,717	2,350	15,390
Average Call Price	35.86	40.32	53.99	77.52	47.32
Average IV	20.62	20.20	20.35	20.51	20.38

Notes: We use Wednesday closing call option data (in-sample) and Thursday closing call option data (out-of-sample) from OptionMetrics from January 1, 1996 through December 31, 2004.

Table 4: NLSIS Estimates and Implied Volatility Root Mean Squared Error (IVRMSE). 1996-2004.

Model			NLSIS Estimation on Option Implied Volatilities					In-Sample		Out-of-Sample		MLIS-Based	
Name	a	b	κ	θ	σ	λ	ρ	IVRMSE	Ratio	IVRMSE	Ratio	IVRMSE	Ratio
SQR	0	1/2	3.1146	0.0523	0.5826	9.00E-05	-0.6520	3.32%	1.0000	3.57%	1.0000	4.11%	1.0000
Standard Errors:			1.50E-03	3.13E-05	1.27E-04	2.16E-03	2.41E-04						
SQRN	1	1/2	67.8964	0.0340	0.1747	1.14E-03	-0.6069	3.17%	0.9565	3.42%	0.9580	3.49%	0.8479
Standard Errors:			3.94E-02	6.26E-06	7.42E-05	1.47E-03	3.16E-04						
VAR	0	1	2.4730	0.0272	1.1884	4.37E-03	-0.7116	2.85%	0.8602	2.98%	0.8338	3.57%	0.8677
Standard Errors:			1.12E-03	6.84E-06	3.20E-04	1.17E-03	2.06E-04						
VARN	1	1	64.4378	0.0367	1.1214	1.93E-03	-0.6749	3.00%	0.9041	3.25%	0.9089	3.63%	0.8823
Standard Errors:			3.95E-02	1.91E-05	4.36E-04	4.94E-02	2.50E-04						
3/2	0	3/2	1.5284	0.0336	7.9501	7.91E-04	-0.7169	3.27%	0.9864	3.42%	0.9585	3.27%	0.7962
Standard Errors:			2.22E-03	3.65E-05	2.94E-03	2.15E-03	2.10E-04						
3/2N	1	3/2	50.9140	0.0388	6.2593	3.36E-04	-0.6854	2.96%	0.8923	3.19%	0.8933	3.52%	0.8559
Standard Errors:			4.08E-02	2.68E-05	2.41E-03	5.13E-02	2.35E-04						

Notes: For each model, we estimate the parameters using NLSIS on the 16,506 Wednesday closing option quotes observed from January 4, 1996 to December 31, 2004. The return drift parameter μ is fixed at the sample average return of 0.091 for all models. Standard errors are computed using the outer product of the gradient at the optimal parameter values. Out-of-Sample refers to the fit of the 15,390 Thursday closing option prices observed from January 4, 1996 to December 31, 2004. The MLIS-Based IVRMSE refers to using the return-based parameters from Table 2 (1996-2004) when computing option prices on the Wednesday data. The Black-Scholes benchmark yields an RMSE of 4.95% for the Wednesday sample and 5.22% for the Thursday sample. The ad-hoc Black-Scholes benchmark uses the average IV for the same weekday of the previous week and has an RMSE of 2.99% for the Wednesday sample and 3.06% for the Thursday sample.

Table 5: In-Sample IVRMSE (%) by Moneyness, Maturity, and VIX Level. 1996-2004.

Model			Panel A. IVRMSE by Moneyness				
Name	a	b	S/X<0.975	0.975<S/X<1	1<S/X<1.025	S/X>1.025	All
SQR	0	1/2	2.9326	3.0455	3.1786	4.1422	3.3186
SQRN	1	1/2	2.8124	2.9454	2.8974	4.0159	3.1742
VAR	0	1	2.6774	2.7946	2.7627	3.2240	2.8547
VARN	1	1	2.6516	2.7400	2.7433	3.8329	3.0004
3/2	0	3/2	3.1133	3.1284	3.0787	3.7687	3.2735
3/2N	1	3/2	2.6274	2.6839	2.7105	3.7848	2.9612
Average			2.8024	2.8896	2.8952	3.7948	

Model			Panel B. IVRMSE by Maturity				
Name	a	b	DTM<30	30<DTM<90	90<DTM<180	DTM>180	All
SQR	0	1/2	3.9390	3.4192	2.9164	2.8364	3.3186
SQRN	1	1/2	4.0059	3.1942	2.7791	2.6717	3.1742
VAR	0	1	3.3337	2.8626	2.5713	2.6628	2.8547
VARN	1	1	3.7828	3.0117	2.5873	2.5909	3.0004
3/2	0	3/2	3.9129	3.2968	2.7903	3.0640	3.2735
3/2N	1	3/2	3.6968	2.9422	2.5717	2.6679	2.9612
Average			3.7785	3.1211	2.7027	2.7490	

Model			Panel C. IVRMSE by VIX Level				
Name	a	b	VIX<18.6	18.6<VIX<21.5	21.5<VIX<25.5	VIX>25.5	All
SQR	0	1/2	2.6782	2.9798	3.3293	4.0926	3.3186
SQRN	1	1/2	2.8225	2.8591	3.2171	3.7164	3.1742
VAR	0	1	2.7453	2.4309	2.9197	3.2599	2.8547
VARN	1	1	2.5974	2.7144	3.1236	3.4840	3.0004
3/2	0	3/2	3.4262	2.6110	3.1364	3.8039	3.2735
3/2N	1	3/2	2.5246	2.6044	3.0750	3.5302	2.9612
Average			2.7990	2.6999	3.1335	3.6478	

Notes: We use the NLSIS estimates from Table 4 to compute the option implied volatility root mean squared error (IVRMSE) in percent for various moneyness, maturity, and volatility (VIX Level) bins for each model. The contracts used in the table are for the 1996-2004 in-sample period, which consists of Wednesday closing call option quotes.

Table 6: In-Sample IV Bias (%) by Moneyness, Maturity, and VIX Level. 1996-2004.

Model			Panel A. IV Bias by Moneyness				
Name	a	b	S/X<0.975	0.975<S/X<1	1<S/X<1.025	S/X>1.025	All
SQR	0	1/2	0.7068	0.7870	1.0238	1.9194	1.0645
SQRN	1	1/2	-0.6771	-0.6152	0.0043	1.8672	0.0515
VAR	0	1	-0.6094	-0.6958	-0.1805	1.2372	-0.1235
VARN	1	1	-0.6775	-0.5778	-0.0131	1.6782	0.0133
3/2	0	3/2	-1.5553	-1.4361	-0.8849	0.6143	-0.9018
3/2N	1	3/2	-0.6893	-0.6079	-0.0538	1.6291	-0.0168

Model			Panel B. IV Bias by Maturity				
Name	a	b	DTM<30	30<DTM<90	90<DTM<180	DTM>180	All
SQR	0	1/2	1.2380	1.2778	0.9723	0.4781	1.0645
SQRN	1	1/2	-0.1901	-0.2085	0.4193	0.5470	0.0515
VAR	0	1	-0.5755	-0.3722	0.4333	0.3342	-0.1235
VARN	1	1	-0.2453	-0.1111	0.1765	0.3788	0.0133
3/2	0	3/2	-0.8248	-0.9782	-0.5592	-1.1008	-0.9018
3/2N	1	3/2	-0.3405	-0.1458	0.2103	0.3523	-0.0168

Model			Panel C. IV Bias by VIX Level				
Name	a	b	VIX<18.6	18.6<VIX<21.5	21.5<VIX<25.5	VIX>25.5	All
SQR	0	1/2	-0.3329	1.2439	1.4689	1.8733	1.0645
SQRN	1	1/2	-0.2572	-0.5301	-0.4701	1.4638	0.0515
VAR	0	1	-1.5854	0.1206	0.2370	0.7340	-0.1235
VARN	1	1	-0.5825	-0.2336	-0.2830	1.1524	0.0133
3/2	0	3/2	-2.4271	-0.5256	-0.4660	-0.1883	-0.9018
3/2N	1	3/2	-0.9767	-0.0477	-0.0751	1.0325	-0.0168

Notes: We use the NLSIS estimates from Table 4 to compute the option implied volatility bias (in percent) for various moneyness, maturity, and volatility (VIX Level) bins for each model. The contracts used in the table are for the 1996-2004 in-sample period, which consists of Wednesday closing call option quotes.