A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation☆

Mikhail Chernov[a], Eric Ghysels[b,c]

[a]*Division of Economics and Finance, Columbia Business School, Uris Hall, 3027 Broadway, Room 7W, New York, NY 10027-6902, USA*
[b]*Department of Economics and Department of Finance, The Pennsylvania State University, 523 Kern Graduate Building, University Park, PA 16802-3305, USA*
[c]*Department of Economics and Kenan-Flagler Business School, University of North Carolina, Chapel Hill, NC 27599-3305, USA*

Received 1 October 1997; received in revised form 17 January 2000

Abstract

The purpose of this paper is to bridge two strands of the literature, one pertaining to the objective or physical measure used to model an underlying asset and the other

☆We would like to thank Torben Andersen, Mark Broadie, Stephen Brown, Charles Cao, Jérôme Detemple, Jens Jackwerth, Eric Jacquier, Frank Hatheway, Roger Lee, Kenneth Singleton, and especially Ron Gallant and George Tauchen for invaluable comments and insightful discussions. We also benefited greatly from the comments of Bill Schwert (the Editor) and the Referees who helped us improve the quality of our paper. In addition, we thank the participants of the NBER/NSF Time Series Conference held at the University of Chicago, the RISK Computational and Quantitative Finance conference, the Newton Institute Workshop on Econometrics and Financial Time Series, the Third Annual New England Finance Doctoral Students Symposium and Computational Finance 99 conference both held at NYU, the 9th Derivative Securities Conference at Boston University, and the 1999 Western Finance Association Meetings, as well as seminars at CIRANO, Michigan State, Penn State, Université Libre de Bruxelles, the University of Montreal, the University of Michigan, and the University of Virginia. We are solely responsible for all remaining errors. The material in this paper subsumes two earlier papers: 'What Data Should be Used to Price Options?' and 'Filtering Volatility and Pricing Options: A Comparison of Implied Volatility, GARCH and Stochastic Volatility'.

Published in: ***Journal of Financial Economics***

pertaining to the risk-neutral measure used to price derivatives. We propose a generic procedure using simultaneously the fundamental price, S_t, and a set of option contracts $[(\sigma^I_{it})_{i=1,m}]$ where $m \geqslant 1$ and σ^I_{it} is the Black–Scholes implied volatility. We use Heston's (1993. Review of Financial Studies 6, 327–343) model as an example, and appraise univariate and multivariate estimation of the model in terms of pricing and hedging performance. Our results, based on the S&P 500 index contract, show dominance of univariate approach, which relies solely on options data. A by-product of this finding is that we uncover a remarkably simple volatility extraction filter based on a polynomial lag structure of implied volatilities. The bivariate approach, involving both the fundamental security and an option contract, appears useful when the information from the cash market reflected in the conditional kurtosis provides support to price long term.

1. Introduction

According to modern asset pricing theory, the value of any asset can be computed as the expectation under the risk-neutral measure of the future cash flows discounted by the pricing kernel. The valuation of any contingent claim, like a European-style option contract, consists of specifying the pricing kernel and determining the appropriate risk-neutral measure transformation. These operations involve several critical steps. First, it should be noted that there is no completely 'model-free' way to proceed.[1] In particular, the characterization of the risk-neutral measure is intimately related to the price of market risk which in turn is determined by the model one adopts to describe the behavior of the fundamental asset underlying the option contract. This step is also inextricably linked to the estimation of parameters which select the data generating process among the class of models considered. There are different sources of data one could use for the purpose of estimating or calibrating parameters and there is certainly an abundant choice as there are many option contracts actively traded and long time series of the fundamental typically available. To further complicate matters it should be noted that many models describing the behavior of the fundamental process feature latent factors such as stochastic volatility. Therefore one faces also the task of using observations to not only estimate parameters but also to filter or extract the unobservable factors.

[1] Even the so-called nonparametric approaches (see for instance Aït-Sahalia and Lo, 1998; Broadie et al., 2000) either implicitly restrict the class of models by imposing regularity conditions to guarantee valid statistical inference or assume an explicit class of models (see for instance Rubinstein, 1994).

In recent years we have made considerable progress on various aspects of this research program. We know more about estimating diffusions particularly those involving stochastic volatility or other latent factors.[2] Parallel to this we witnessed the emergence of several studies suggesting schemes to extract risk-neutral measures from option prices (see inter alia references in footnote 2). Considerable effort was also devoted to filters for extracting the latent factors like volatility. These filters either involve options data or underlying fundamentals (but not both jointly).

The purpose of this paper is to bridge two strands of the literature, one pertaining to the objective or physical measure used to model the underlying asset and the other to the risk-neutral measure used to price derivatives. In fact, we start first with estimating both measures *jointly*. This poses several challenges as building a bridge between the objective and risk-neutral world prompts us to think about many new issues which need to be addressed. In addition, it also opens up new possibilities for comparing the information in the underlying fundamental and options data, a theme which has been the subject of many previous studies. Numerous papers have confronted empirical evidence obtained from derivative security markets with results from the underlying and vice versa. In particular, issues related to the informational content of option prices have been examined extensively in the literature (see for instance Christensen and Prabhala (1998) for the most recent example). Our attempt to model the price behavior of fundamental and derivative securities jointly is motivated by the very same issues hitherto raised in the literature. Namely, we want to learn more about the informational content of option prices. We also want to know how we can improve the statistical precision of diffusion parameters by incorporating options.

Our goal is to investigate these questions in a unifying framework. While we use the Heston (1993) model as specific example, it should be stressed at the outset that our analysis is not limited to any particular model. The choice of Heston's model is motivated by two important factors. First, it has closed-form option pricing formula, which represents a considerable computational advantage. Second, because Heston's model features analytic solutions it has received much attention in the literature (see Bakshi et al., 1997, for references), which makes our analysis directly comparable with results previously reported.

Financial theory also suggests that, for stochastic volatility models with two state variables, such as the models of Hull and White (1987), Scott (1987), Wiggins (1987), Heston (1993), and many others, one should consider the

[2] Numerous techniques have been proposed for the estimation of continuous time processes pertaining to the pricing of derivative securities. The literature on the estimation of diffusions with or without stochastic volatility and/or jumps, is summarized in a number of surveys and textbooks, including Bates (1996), Campbell et al. (1997), Ghysels et al. (1996), Melino (1994), Renault (1997) and Tauchen (1997).

fundamental and its derivative contracts jointly to estimate diffusion parameters and price options simultaneously. There are indeed appealing theoretical reasons to pursue this approach, as in a stochastic volatility economy options must be added to create a complete market model (Romano and Touzi, 1997). The complete market model guarantees the existence and uniqueness of the risk-neutral probability density used to price the option contracts. If done judiciously, this challenging task should dominate the use of a single source, whether options or fundamental.

Although no attempts have been made previously to estimate and appraise stochastic volatility models using the joint distribution of fundamentals and options, it is clear that much of the evidence in the literature suggests that we should gain from addressing this issue. For example, a recent paper by Gallant et al. (1998) adopts a strategy similar to ours, though not involving options. They consider the bivariate process of the fundamental and the daily high/low spread, which provides extra information about the course of volatility. In this paper we propose a generic procedure for estimating and pricing options using simultaneously the fundamental price, S_t, and a set of option contracts $[(\sigma^I_{it})_{i=1,m}]$, where $m \geqslant 1$ and σ^I_{it} is the Black–Scholes implied volatility. Please note that we can, in principle, manage a panel of options, i.e. a time series of cross-sections. The procedure we propose consists of two steps. First, we fit a seminonparametric (henceforth SNP) density of $[S_t, (\sigma^I_{it})_{i=1,m}]$, conditional on its own past $[S_\tau, (\sigma^I_{i\tau})_{i=1,m}]$ for $\tau < t$, using market data. Next we simulate the fundamental price and option prices, and calibrate the parameters of the diffusion and its associated option pricing model to fit the conditional density of the market data dynamics. The procedure coined by Gallant and Tauchen (1996) as Efficient Method of Moments (EMM), has been used primarily to estimate diffusions using only the fundamental price process S_t. We extend EMM to incorporate option prices and fundamentals simultaneously. The EMM procedure, which is a simulation-based estimation technique, allows estimating the model parameters under both objective and risk-neutral probability measures if we use simultaneously implied volatilities and the underlying asset data. Indeed, time series of the underlying asset provide parameters under the objective probability measure while risk-neutral parameters can be retrieved from options. Since the model we adopt has a closed-form option pricing formula, we can obtain the volatilities implied by the Black–Scholes formula from the simulated data and contrast them with their counterparts from the real data via the EMM framework. This procedure yields parameter estimates under the risk-neutral measure. Having estimated the risk-neutral and objective measures separately allows us to appraise the typical risk-neutral representations used in the literature. In particular, in order to obtain the closed-form solutions, the standard approach assumes that the linearity of the volatility drift is preserved. We are able to determine if this assumption is consistent with the data.

The task is challenging. Aside from specifying the methodology, we do not know in advance how well our procedure will work in producing better parameter estimates for diffusions. Further, we cannot predict what improvements in option pricing and hedging can be made. We compare univariate and multivariate models in terms of pricing and hedging performance. The univariate specifications consist of models only using the fundamental asset and models using only options data. It should be noted, however, that the knowledge of the estimated model parameters is not sufficient to compute an option price or a hedge ratio. We have to know the latent spot volatility as well. Previous studies treated spot volatility as a parameter and estimated it from the cross-section of options prices taken from the previous trading day. This approach introduces inconsistencies with the model. A recent extension of the SNP/EMM methodology introduced in Gallant and Tauchen (1998) allows us to address the problem. We filter spot volatilities via reprojection. That is, we compute the expected value of the latent volatility process using an SNP density conditioned on the observable processes, such as returns, or options data, or both.

Our results demonstrate the dominance of the univariate approach that relies solely on options. A by-product of this finding is that we uncover a remarkably simple volatility extraction filter based on a polynomial lag structure of implied volatilities. The bivariate approach appears useful when the information from the cash market provides support, through the conditional kurtosis, to price some long term options.

These findings prompt us to consider alternative volatility filters, such as for instance the one obtained from the GARCH class of models. We examine the quality of various filters through the window of the Black–Scholes (henceforth BS) option pricing model, which allows us to separate the effects of a particular pricing kernel from the filter's contribution. Interestingly, we find that the role of the pricing kernel is marginal compared to that of filtering spot volatility.

The remainder of the paper is organized as follows. Section 2 sets the stage for the analysis of the joint density function of fundamentals and options. We discuss first the issues addressed so far in the literature and present the model we estimate. In Section 2 we also provide a brief review of the EMM estimation and reprojection method. Section 3 reports the estimation results and examines the mapping from objective to risk-neutral measures, while Section 4 evaluates the performance of the estimated models. Section 5 studies the role of different volatility filters. The last section concludes. Technical material is covered in several appendices to the paper.

2. Joint estimation of the fundamental and option pricing processes

Numerous papers have confronted empirical evidence obtained from derivative security markets with results from the underlying and vice versa. In particular, issues pertaining to the informational content of option prices have been examined extensively in the literature (see Bates, 1996; Canina and Figlewski, 1993; Christensen and Prabhala, 1998; Day and Lewis, 1992; Fleming, 1998; Lamoureux and Lastrapes, 1993, among others). Ait-Sahalia et al. (1997) address essentially the same issue, comparing state-price densities (SPD) implied by time series of the S&P 500 index with the SPD implied by a cross-section of S&P 500 index options. They reject the hypothesis that the two state price densities are the same. As they examine models with volatility specified as a function of the stock price, one can view their result as a rejection of deterministic volatility models. Along the same lines, Dumas et al. (1998) examine the out-of-sample pricing and hedging performance of the same class of volatility models using also S&P 500 options data. Dumas et al. find that 'simpler is better', i.e. deterministic volatility models perform no better than ordinary implied volatility.

Our attempt to model the price behavior of fundamental and derivative securities jointly is motivated by the very same issues raised in the literature. Namely, we want to learn more about the informational content of option prices, as did Canina and Figlewski (1993) and many others. We also want to know how we can improve the statistical precision of diffusion parameters by incorporating options data as did Pastorello et al. (1994), who used at-the-money implied volatilities, replacing latent spot volatility, to estimate the Hull and White (1987) model.[3] Moreover, we also want to assess the advantages of multivariate schemes using financial criteria such as the out-of-sample pricing and hedging performance of models, like Bakshi et al. (1997), Dumas et al. (1998) and Jacquier and Jarrow (1998), among others.

We attempt to investigate these questions in a unifying framework. We use the stochastic volatility (SV) model specified by Heston (1993) for that purpose, though it should be stressed at the outset that our analysis is not limited to this particular model. The Heston model will be covered in a first subsection.

The joint modeling of returns and derivative security prices will present new challenges, which we will discuss in this section. First, we describe the data in Section 1.2. Since the EMM procedure is widely used and described elsewhere, notably in Gallant and Tauchen (1998), we will only summarize its major steps in a third subsection. A fourth subsection deals with reprojection methods, an extension of EMM to extract latent volatility processes. The empirical results are discussed in the last two subsections.

2.1. The Heston model

Following Heston (1993), we can write the model as

$$\mathrm{d}S(t)/S(t) = R\,\mathrm{d}t + \sqrt{V(t)}\,\mathrm{d}W_S^*(t) \tag{1}$$

[3] Pastorello et al. (1994) did not estimate the joint process as we propose to do in this paper.

and

$$\mathrm{d}V(t) = (\theta^* - \kappa^* V(t))\,\mathrm{d}t + \sigma_V \sqrt{1-\rho^2}\sqrt{V(t)}\,\mathrm{d}W_V^*(t)$$
$$+ \sigma_V \rho \sqrt{V(t)}\,\mathrm{d}W_S^*(t), \tag{2}$$

where the model is stated under the risk-neutral probability measure. Eq. (1) implies that the stock-price process $S(t)$ follows a geometric Brownian motion, with stochastic variance $V(t)$. Eq. (2) states that $V(t)$ follows a square-root mean-reverting process with the long-run mean θ^*/κ^*, speed of adjustment κ^* and variation coefficient σ_V. Imposing the restriction $\sigma_V^2 \leqslant 2\theta$ guarantees that $V(t)$ stays in the open interval $(0, \infty)$ almost surely (see, for instance, Cox et al., 1985). The Brownian motions $W_S^*(t)$ and $W_V^*(t)$ are assumed independent. Eqs. (1) and (2) imply, however, that $Corr_t(\mathrm{d}S(t)/S(t), \mathrm{d}V(t)) = \rho\,\mathrm{d}t$. Parameters with asterisks are those that change when the model is rewritten under the objective probability measure. Under the change of measure, the risk-free rate R is substituted by a drift parameter, μ_S, and all asterisks are removed. This model yields the following formula for a price of a call at time t, with time to maturity τ and strike K:

$$C(t, \tau, K) = S(t)\Pi_1(t, \tau, S, V) - K\mathrm{e}^{-R\tau}\Pi_2(t, \tau, S, V), \tag{3}$$

where the expressions for $\Pi_j, j = 1, 2$ are provided in Appendix A.

The common practice of estimating diffusions using the underlying asset and then relying on an option pricing formula has a number of drawbacks. Standard complete market asset pricing theory determines that one has to change the measure, from the objective measure to the risk neutral one.[4] This transformation is often ad hoc. Although continuous time general equilibrium preference-based asset pricing models readily deliver the direct connection from the objective measure to the risk-neutral measure, they often result in rather complex diffusion models for the underlying asset. The complex dynamics arise because the equilibrium asset price process is derived endogenously, based on the discounted flow of dividends, using an endogenously determined risk-neutral rate (see e.g. Broadie et al. 1997, for such a derivation). It is therefore common to use a simple diffusion for the asset return and volatility dynamics, and assume that the change of drift, which by Girsanov's theorem amounts to changing the measure, maintains the same type of processes.

[4] See Harrison and Kreps (1979) and Harrison and Pliska (1981) for further discussion. Arguments about completeness of markets are typically imposed to guarantee the existence of a unique risk-neutral measure.

To further examine the change of measure, let us consider the Radon-Nikodym derivative of the objective probability measure with respect to the risk-neutral one. This derivative can be computed as follows:

$$\xi_{0,t} = \exp\bigg(-\frac{1}{2}\int_0^t (\lambda_1^2(u) + \lambda_2^2(u))\,\mathrm{d}u - \int_0^t \lambda_1(u)\,\mathrm{d}W_S(u) - \int_0^t \lambda_2(u)\,\mathrm{d}W_V(u)\bigg), \tag{4}$$

where $\lambda(t) = (\lambda_1(t),\ \lambda_2(t))'$ is the vector with the market prices of risk, return and volatility risk, respectively. When we know the parameter values under both measures we can infer $\lambda(t)$. In particular, by Girsanov's theorem we have

$$\mu_S S(t) - \lambda_1(t)\sqrt{V(t)}S(t) = RS(t) \tag{5}$$

and

$$\theta - \kappa V(t) - \lambda_2(t)\sigma_V\sqrt{1-\rho^2}\sqrt{V(t)} - \lambda_1(t)\sigma_V\rho\sqrt{V(t)} = \theta^* - \kappa^* V(t). \tag{6}$$

Therefore,

$$\lambda_1(t) = \frac{\mu_S - R}{\sqrt{V(t)}} \tag{7}$$

and

$$\lambda_2(t) = \frac{C_1}{\sqrt{V(t)}} - C_2\sqrt{V(t)}, \tag{8}$$

where

$$C_1 = \frac{\theta - \theta^* - (\mu_S - R)\sigma_V\rho}{\sigma_V\sqrt{1-\rho^2}} \tag{9}$$

and

$$C_2 = \frac{\kappa - \kappa^*}{\sigma_V\sqrt{1-\rho^2}}. \tag{10}$$

Eq. (7) implies that the asset risk premium increases when volatility decreases. This counterintuitive effect of volatility on the risk premium implies arbitrage opportunities. In mathematical terms this effect means that the transformation to the risk neutral measure, described by the Girsanov theorem, may fail.

However, it is difficult to verify the conditions of the theorem in this particular case. Rydberg (1997) discusses these issues in more detail.

It is beyond the scope of the current paper to address these deficiencies of the Heston model, since our main focus is to study the joint estimation of risk neutral and objective measures. We employ the standard Heston model for that purpose. In practice, it is unlikely one would have to deal with extremely small volatility values. While the improvement of the Heston model is beyond the scope of the present paper, the condition presented in Eq. (7), above, should be considered by researchers who intend to extend the model for empirical or theoretical reasons. Finally, since our estimation strategy allows us to compute the prices of risk from Eqs. (7) and (8), we will revisit the issues related to the measure transformation in Section 2.2, where we discuss the estimation results.

2.2. The data

Like many previous studies, we examine the S&P 500 index and the SPX European option contract traded on the index. Our analysis requires both options and returns data. The source of our series is the Chicago Board of Options Exchange (CBOE). The data consist of daily last sale prices of options written on the S&P 500 index, as well as the closing price of the index. Specifically, the dataset contains the date, maturity month, option price, call/put flag, strike, open interest, volume, and the underlying index. There is a 15-min difference between the close of the AMEX, NASDAQ, and NYSE stock markets, where the 500 stocks included in the index are traded, and the Chicago options market. This difference leads to non-synchronicity biases. Harvey and Whaley (1991), and Bakshi et al. (1997), among others, suggest various schemes based on option price quotes or transactions around the 3 PM market close. We control for the possibility of such biases in our simulation procedure, which is discussed below and in Appendix C. The sample covers the time period from November 1985 until October 1994. We set aside the last year of data, November 1993 to October 1994, for the out-of-sample tests, and use the rest for estimation purpose. Plot (A) of Fig. 1 displays the S&P 500 series, showing the familiar pattern, including the negative return corresponding to the crash of October 1987. The dashed vertical line represents the end of the estimation sample and the beginning of the data used for the purpose of out-of-sample appraisals of the models.

Various schemes to extract implied volatilities from options have been suggested in the literature (see for instance Bates, 1996, for a survey). We concentrate our attention on the at-the-money (henceforth ATM) calls, where we define at-the-money as $S/K \in [0.97, 1.03]$, for instruments with short maturities, as these are the most liquid instruments. Because of the active trading, the implied volatilities of these contracts should convey the most precise information. Moreover, Harvey and Whaley (1991) note that the ATM volatilities are the

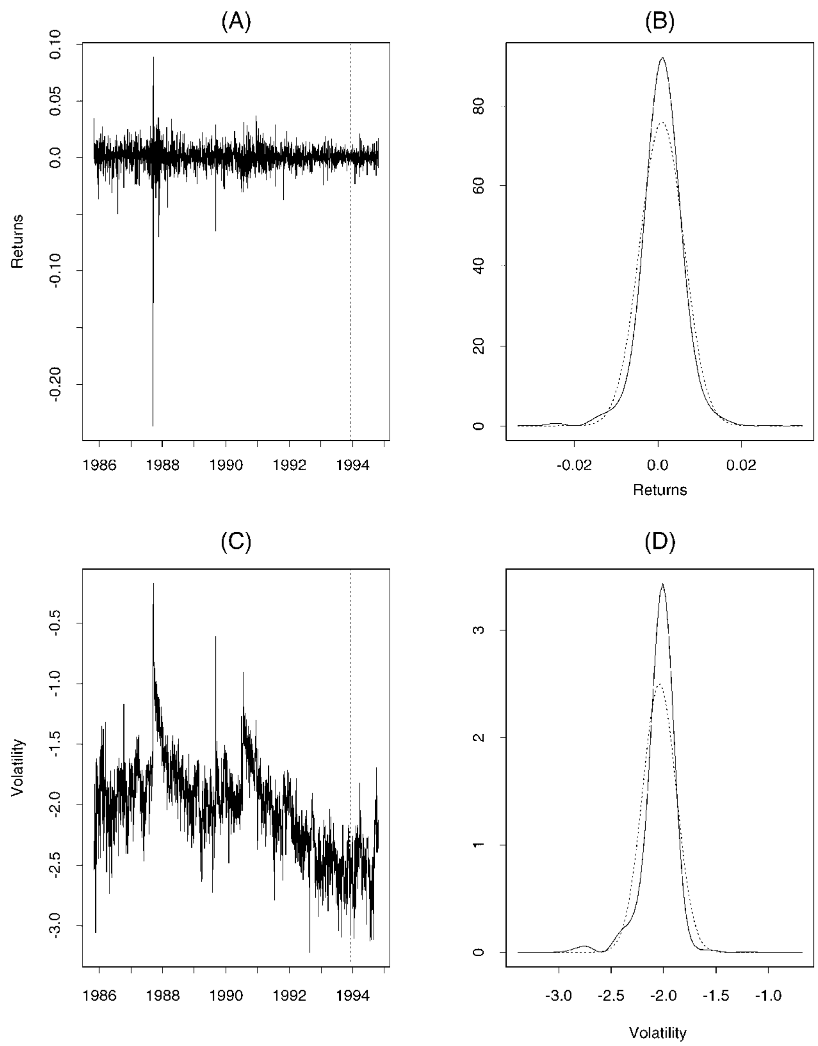

Fig. 1. The univariate series and SNP densities. We estimate the SNP density for the two univariate types of data: (i) the log-returns on the S&P500; (ii) the log of the BS implied volatilities of the closest to maturity and the money call options. The data are collected daily, and span the period from November 1985 to October 1993. The plots to the left are the time series of the data, the plots to the right are the estimated densities of the series. The solid line is a plot of an SNP fit, the dashed line is normal density with the same mean and variance.

most sensitive to changes in the spot volatility rate, since an option's vega is maximized when moneyness is close to 1. To be more precise, we select the calls with the shortest maturities and, in addition to being ATM, we require the option to have the strike as close to the index level as possible, such that $K^* = \text{argmin}_K |S/K - 1|$.[5]

In order to make the observed S&P 500 index and the simulated underlying fundamental data comparable, we adjusted the S&P 500 index for dividends. We took a constant dividend rate, continuously compounded, of 2% which is consistent with Broadie et al. (1997) for example. Therefore, we ignore the lumpiness of the dividend payments, and also conveniently avoid using the historically observed dividends, which would considerably complicate our simulation design. Moreover, since we are dealing with European index options, this issue matters little, as we are only interested in the total flow of dividends paid over the life of the contract. Finally, we used the monthly 3-month T-bill yield from CITIBASE as a proxy for the short-term interest rate. Since the stochastic volatility models we consider assume a constant interest rate, we take the average yield, which is equal to 5.81396%.

Finally, the estimation of SNP density requires the use of stationary and ergodic data. As noted before we will estimate SV models, using three types of data: (i) time series data on the S&P 500 index, (ii) BS volatilities implied by the closest to maturity and at the money calls on the index, and (iii) both data sets jointly. Therefore, the data entries to the SNP estimation routine are the log-returns on the index. Likewise, rather than using the implied volatilities, which are nonnegative, we will work with the log-volatilities. Despite the transformations, we will refer to these data series as S&P 500 and BS volatilities for convenience. Plot (C) of Fig. 1 displays the BS implied volatilities. We note that the highest volatility is, not surprisingly, observed at the time of the crash. Moreover, it is also important to note the downward trend in volatility beginning around 1991, and the reversal of this trend in the series coming after the estimation part of our data set. This reversal will make the out-of-sample exercise particularly interesting.[6]

2.3. *The efficient method of moments estimation procedure*

Several methods have been proposed to estimate the parameters of stochastic volatility models. These methods include the generalized method of moments

---

[5] Apart from the moneyness filter, several other filters were applied to the data. In particular, observations with a call price missing or equal to zero were dropped. The various filters applied to the data leave us with 1978 observations, which roughly corresponds to 247 observations per year.

[6] The more recent years, which are not covered by our sample, especially the summer and fall of 1998, indicate an upward trending volatility. In other words, the volatility process seems to revert to the mean as assumed in the Heston (1993) model we consider here. It should also be noted that the options contracts on S&P 500 were of American type during a very short period of our sample, namely prior to April 1986. Because we select short maturity contracts, for which the exercise premium is very close to zero, this change should not affect much of our results. Fig. 1 also shows that the American option part of our sample, prior to April 1986, does not introduce any abnormal patterns.

(GMM), quasi-maximum likelihood, and various simulation-based procedures, including Bayesian methods and efficient method of moments (see Ghysels et al., 1996, for a literature review). In this paper, we use the Efficient Method of Moments (henceforth EMM) procedure of Gallant and Tauchen (1996), which has already found many applications in the estimation of both continuous time and discrete time stochastic volatility models. Examples include Andersen and Lund (1997), Andersen et al. (1999), Gallant, Hsieh and Tauchen (1997), Gallant et al. (1999), Gallant and Tauchen (1998), Ghysels and Jasiak (1996), Jiang and van der Sluis (1998), and Tauchen (1997). So far, most of these applications only use a single data series, either a short rate process or a stock price (index). There are exceptions, notably Ghysels and Jasiak (1996), who consider the joint process of stock returns and trading volume, and Gallant, Hsu and Tauchen et al. (1999), who as noted before, consider the bivariate process of the fundamental asset and the daily high–low spread. EMM can be divided into two main parts, the estimation of the so-called score generator, which will be discussed first, and the estimation of the diffusion parameters, which will conclude this section.

Suppose the process of interest is denoted ι_t. In our application, this process can be univariate, bivariate, or involving a low-dimensional panel data set. If univariate, the process will include either asset returns or BS implied volatilities. If bivariate the process will include both returns and implied volatilities. In the general case a panel data set is considered which includes the fundamental asset returns series and M option contracts with different moneyness, different maturities, or both, represented by their BS implied volatilities $[(\sigma_{it}^I)_{i=1,M}]$. In a generic context, we assume that ι_t is a vector with L elements. It has a conditional distribution $p(\iota_t \mid I_t, \Theta)$, where I_t is the information set and Θ represents the parameters of the stochastic volatility model for ι_t. The asymptotically efficient method to estimate Θ is maximum likelihood (MLE), which involves maximizing the function

$$\frac{1}{T}\sum_{t=1}^{T} \log p(\iota_t \mid I_t, \Theta). \tag{11}$$

Maximizing Eq. (11) is equivalent to solving

$$\frac{1}{T}\sum_{t=1}^{T} \frac{\partial}{\partial\Theta}\log p(\iota_t \mid I_t, \hat{\Theta}) = 0, \tag{12}$$

where $\partial \log(p(\cdot \mid \cdot))/\partial\Theta$ is the score function. The above expression is the sample equivalent of

$$E\left(\frac{\partial}{\partial\Theta}\log p(\iota_t \mid I_t, \Theta)\right) = 0. \tag{13}$$

Unfortunately, it is very difficult to obtain the likelihood function for stochastic volatility models. Therefore, it is impossible to compute the score generator $\partial \log(p(\cdot \mid \cdot))/\partial \Theta$. To overcome this challenge, Gallant and Tauchen suggest computing instead a SNP density $f_k(\iota_t \mid X_t, \Xi)$, where X_t is the vector of M lags of ι_t and Ξ is the vector of parameters of the SNP model. The index k relates to the dimension of Ξ, and should expand at the appropriate rate as the sample size grows to accomplish MLE asymptotic efficiency (see Gallant and Long, 1997, for further discussion). We provide some specific details regarding the SNP density in Appendix B. Note that EMM can use any score generator which represents the data well. In this paper, we will use the SNP score generator, which is also required for the reprojection procedure described later.

As noted above, EMM has two parts. The first part is the estimation of the auxiliary score generator model. The estimated SNP density provides the input to the second stage of the EMM estimation procedure. More precisely, the SNP score function generates a set of moment conditions. In particular, Θ can be estimated through the moment conditions, or score function, similar to Eq. (13), which in this case will be:

$$m(\Theta, \hat{\Xi}) = \mathrm{E}\left(\frac{\partial}{\partial \Xi} \log f_k(\iota_t \mid X_t, \hat{\Xi})\right) = \int \frac{\partial}{\partial \Xi} \log f_k(\iota_t \mid X_t, \hat{\Xi})\, \mathrm{d}P(\iota, X, \Theta). \quad (14)$$

Since these moment conditions should have mean zero, they can be used as the basis for a GMM-type estimation procedure, which yields the desired estimate for the parameter vector Θ. The moment conditions are easier to compute by simulation instead of computing the integral in Eq. (14) numerically. Hence, we compute sample moments by simulating N observations of ι_t from the SV model.[7] With simulated time series of length N for ι_t, with candidate parameters Θ, the left-hand side of Eq. (12) translates into

$$m_N(\Theta, \hat{\Xi}) = \frac{1}{N} \sum_{t=1}^{N} \frac{\partial}{\partial \Xi} \log f_k(\iota_t(\Theta) \mid X_t(\Theta), \hat{\Xi}). \quad (15)$$

Then we can formulate the EMM estimator for the parameter vector Θ, using the following quadratic minimization criterion:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}}\; m_N(\Theta, \hat{\Xi})' W_T m_N(\Theta, \hat{\Xi}). \quad (16)$$

[7] Since the SV model is formulated in continuous time we need to discretize the process to generate simulated paths. We use the explicit order 2.0 weak stochastic differential equation (SDE) discretization scheme, described in Kloeden and Platen (1995 pp. 486–487), to simulate the processes defined in Eqs. (1) and (2). In Appendix C, we provide further details of the scheme employed to create discrete observations. The empirical results reported in the next section are based on simulated samples of size $N = 10{,}000$.

Because of the properties of the SNP model,

$$W_T = \frac{1}{T}\sum_{t=1}^{T} \frac{\partial \log f_k(\iota_t \mid X_t, \hat{\Xi})}{\partial \Xi}\frac{\partial \log f_k(\iota_t \mid X_t, \hat{\Xi})}{\partial \Xi'}. \tag{17}$$

Asymptotically, the EMM estimator is consistent and normal and has, under suitable regularity conditions, the same efficiency as MLE (see Gallant and Tauchen, 1996; Gallant and Long, 1997, for further discussion).

We will focus primarily on the case where ι_t represents a bivariate process of a stock return and a BS implied volatility. The distribution $p(\iota_t \mid I_t, \Theta)$ in this context is implicitly defined by Eqs. (1)–(3). Suppose now that the parameter vector Θ can be written as $(\mu_S, \Theta_c, \Theta_o, \Theta_n)$, where μ_S is the rate of return on the underlying asset under the objective probability measure, Θ_c contains the parameters common to the objective and risk-neutral measure, Θ_o represents the objective probability measure volatility drift parameters, and Θ_n the risk-neutral ones. In Eqs. (1) and (2), these parameter vectors correspond to $\Theta_c = (\sigma_V, \rho)$, $\Theta_o = (\theta, \kappa)$, and $\Theta_n = (\theta^*, \kappa^*)$. The first step consists of simulating the underlying processes from Eqs. (1) and (2) under the objective probability measure for a given set of parameters values for $(\mu_S, \Theta_c, \Theta_o)$. It should be noted that we simulate the latent volatility process $V(t)$, though it is not part of ι_t since we need to use $V(t)$ to compute the price, $S(t)$. Using the risk-neutral measure parameters (Θ_c, Θ_n), where the values of Θ_c remain the same, we compute the options price according to Eq. (3) and calculate the BS implied volatilities from these prices. Since we estimate Θ_o and Θ_n separately, we can test certain hypotheses about the transformation from objective to risk-neutral measures.

To obtain the BS volatilities, we need to apply the option pricing formula to the simulated data. This step requires knowing the parameters and choosing time to maturity and strike features of the contract. Obviously, time to maturity is not available in the simulated data. To make the simulated option prices comparable with the observed ones in the actual data, we replicate the maturities from the observed data. In order to decrease the simulation error, the sample size of the simulated data (recall $N = 10{,}000$) is much larger than the sample of actual data available. Therefore, we cycle through the sequence of observed maturities in the actual data set to cover the entire simulated data set. In particular, for the simulated ith observation, we use the maturity from the mod(i, T), where T is the sample size of the observed data. If the length of the simulated sample is N and a multiple of T, say $N = lT$, then this scheme amounts to replicating the number of times each maturity appears in the observed data by l.

We apply a similar strategy for the strike features of the contracts. In particular, we simulate moneyness instead of strikes, because the simulated sample path of $S(t)$, or price, can be quite different from the observed one. Matching the moneyness with the real data implies the existence of strikes not always observed in the real data sample. Using this strategy preserves the crucial properties of options. Since we can rewrite the option pricing formula in terms of moneyness, the dependence on strike features will be eliminated. We rotate moneyness from the observed data in exactly the same way as maturities to simulate BS implied volatilities.

Finally, it should be noted that the simulation approach described in this section also extends to situations involving only options data, or the more commonly used univariate setup based on returns series. In neither case is it possible to estimate the entire parameter vector Θ, which we wrote as $(\mu_S, \Theta_c, \Theta_o, \Theta_n)$. With options data, we cannot estimate the drift parameter μ_S under the objective measure. Hence, with options we can estimate the parameter vector $(\Theta_c, \Theta_o, \Theta_n)$.[8] When returns series are used we cannot recover the risk neutral volatility parameters, i.e. Θ_n. Hence, only cases with both fundamentals and derivatives will involve the full parameter vector $(\mu_S, \Theta_c, \Theta_o, \Theta_n)$.

2.4. Reprojection

Having obtained the EMM estimates of the model parameters $\hat{\Theta}$, we would like to extract the unobserved spot volatility $V(t)$ in order to price options according to Eq. (3). Several filters have been proposed in the literature, all involving the return process exclusively. Harvey et al. (1994) suggest employing the approximate Kalman filter based on a discrete time SV model. The exact filter is derived by Jacquier et al. (1994) in the context of a Bayesian analysis of the same model. Nelson and Foster (1994) demonstrate how diffusion limit arguments apply to the class of EGARCH models, and provide a justification for EGARCH models as filters for instantaneous volatility. Some attempts were made to extend these filters to a multivariate context (see, in particular, Harvey et al., 1994; Jacquier et al., 1995; and Nelson, 1996). These multivariate extensions exclusively use return series, and cannot accommodate derivative security market information. We propose a filtering method based on the reprojection procedure introduced by Gallant and Tauchen (1998). We briefly describe first the method, intuitively, in a generic context and focus on the specific applications we will consider.

Suppose we have a vector process consisting of observable and unobservable time series. For example, observables could be returns, or BS implied volatilities,

[8] Note that options data allow us to estimate the volatility parameter vector under both measures, i.e. both Θ_o and Θ_n. In particular, we proceed along the same lines as in the bivariate case, and simulate the latent volatility process $V(t)$ and the fundamental asset, to compute option prices and obtain implied volatilities. However, since the data is limited to options, and option pricing formulas omit the objective measure drift parameter, we cannot identify μ_S. All other parameters can be identified, since any variation in $(\Theta_c, \Theta_o, \Theta_n)$ affect option prices.

while the latent spot volatility would be unobservable. Let us denote the vector of contemporaneous and lagged observable variables by x_t, and the vector of contemporaneous unobservable variables by y_t. The filtering problem is equivalent to computing the following conditional expectation

$$\tilde{y}_t = \mathrm{E}(y_t \mid x_t) = \int y_t p(y_t \mid x_t, \Theta)\, \mathrm{d}y_t, \tag{18}$$

where Θ is the parameter vector. This expectation involves the conditional probability density of y_t given x_t. If we knew the density, we could estimate it by $\hat{p}(y_t \mid x_t) = p(y_t \mid x_t, \hat{\Theta})$. Unfortunately, for SV models there is no analytical expression for the conditional density available. Therefore, we need to estimate this density as $\hat{p}(y_t \mid x_t) = f_k(\hat{y}_t \mid \hat{x}_t)$, where $\hat{y}_t, \hat{x}_t$ are simulated from the SV model with parameters set equal to $\hat{\Theta}$, and where f_k is again a SNP density. Gallant and Long (1997) show that

$$\lim_{k \to \infty} f_k\ (\hat{y}_t \mid \hat{x}_t) = p(y_t \mid x_t, \hat{\Theta}). \tag{19}$$

Hence, the SNP density converges asymptotically to the true conditional probability density (where convergence is in terms of the Sobolev norm specified by Gallant and Long). Hence, reprojection provides an unbiased estimate of the latent process, namely spot volatility.

The reprojection filtering method takes a multivariate form under three scenarios. These scenarios are, (i) only involving a vector of multiple return series, an approach not considered in this paper but feasible, (ii) only involving a vector of options, as discussed further below and, (iii) a mixture of the previous two scenarios. The latter strategy will be our prime focus here. It should be noted that one can also consider univariate schemes which involve either the return series or the BS implied volatilities. The former univariate scheme would be comparable to the filtering methods of Harvey et al. (1994), Jacquier et al. (1994) and Nelson and Foster (1994). In Section 5, we will discuss such univariate volatility filters based on returns data. Univariate schemes only involving options have been proposed informally. Often, though not exclusively, such univariate filters rely on the Black–Scholes model and involve a cross-section of options, treating today's volatility as a parameter. Bakshi et al. (1997), for instance, employ this methodology and obtain volatility estimates which minimize the pricing error of the daily cross-section of options. The reprojection approach applied to a vector of options, such as the second reprojection scheme specified above, is more general since it takes fully advantage of the time series and cross-section data structure. Even the univariate reprojection method using solely an ATM option will result in a time series filter of implied volatilities, which to the best of our knowledge is a filtering scheme for instantaneous volatility which has not been fully exploited so far (and as we will see in the next section performs remarkably good).

In the remainder of this section we will focus exclusively on the applications of filtering using the reprojection method which contains novel features. The specific application can readily be extended to any of the aforementioned generic specifications. There are two novel applications, one of which deals with a bivariate model of returns and an ATM option, and the other which deals with a univariate filter based on options only. The reprojection scheme relies on a one-step ahead forecast, which is an expectation computed from the distribution of volatility conditional on the contemporaneous and lagged returns, denoted r_t, and lagged implied volatilities: $p(V(t)\,|\,r_t, \sigma^I_{t-1}, r_{t-1}, \ldots, \sigma^I_{t-M}, r_{t-M}, \hat{\Theta})$. Hence, $x_t = (r_t, \sigma^I_{t-1}, r_{t-1}, \ldots, \sigma^I_{t-M}, r_{t-M})$ and $y_t = V(t)$ in Eq. (18). The computation of the SNP density for $V(t)$ proceeds in several steps. First, using the estimates $\hat{\Theta}$ one simulates the processes specified in (1) and (2), which produces the series $\{\hat{\sigma}^I_{t-1}, \hat{r}_t, \hat{V}_t\}_{t=1}^N$, in the case of bivariate data, and $\{\hat{\sigma}^I_{t-1}, \hat{V}_t\}_{t=1}^N$, in the univariate case.[9] Let us again denote the vector $(\hat{r}_t, \hat{\sigma}^I_{t-1}, \hat{r}_{t-1}, \ldots, \hat{\sigma}^I_{t-M}, \hat{r}_{t-M})'$ (or $(\hat{\sigma}^I_{t-1}, \ldots, \hat{\sigma}^I_{t-M})'$ in the univariate case) by $\hat{x}_t$.

Second, since the SNP density has a Gaussian lead term, we need to transform the simulated volatility process which only takes positive values. It would be natural to consider the log-transformation of $\hat{V}_t$, as was done with the implied volatilities data (see Section 2.2). However, in this case we need an estimate of the untransformed spot volatility to substitute into the option pricing formula. We conclude this section with the description of a piecewise linear approximation to the log transformation which will allow us to recover spot volatility without any biases.

Consider a first order Taylor expansion of the logarithm of volatility, which we denote by $L(V(t))$. We find the SNP density $f_k(L(\hat{V}_t)\,|\,\hat{x}_t)$.[10] From Jensen's inequality, we have that $E(C(V(t))) \leqslant C(E(V(t)))$, for any concave function C. This inequality becomes an equality only if C is linear, hence the use of $L(\cdot)$. As a result, the SNP density fitted to the simulated data produces a mean of $\hat{V}_t$ conditional on the observable vector $\hat{x}_t$, which in turn yields the desired filtered values. The first order Taylor expansion is an approximation and, in theory, can be centered anywhere. One obvious starting point is to compute the Taylor expansion around the mean of simulated $\hat{V}_t$'s. This approximation may be quite inaccurate in the tails of the distribution, however. This drawback is particularly important as the inverse transform may easily result in negative

[9] To streamline the notation, we use $\hat{V}_t$ for the simulated $V(t)$. Furthermore, since k grows with the sample size N we select a simulation size of 10,000 observations.

[10] Fitting SNP densities in a reprojection exercise requires a slightly different specification since the explanatory variables are exogenous to the dependent variable. Hence, we set M_R equal to 0 (see Appendix B). The SNP code also requires other slight modifications, so that the lags of dependent variables would not be included in the conditioning set as in the standard SNP density specification.

volatilities.[11] Indeed, experiments showed that the first order Taylor expansion centered around the mean of simulated $\hat{V}_t$ resulted in 32% negative volatilities. We therefore took the first order Taylor expansion at different points, yielding a piecewise linear approximation similar to a spline transformation. Further experiments showed that a linear approximation centered around two points yielded roughly 3.6% negative volatilities. Centering the approximation around four points yielded roughly 1.6%. We therefore took a four-point linear spline and centered the first order Taylor expansions around the 8th, 24th, 50th, 76th, and 92nd percentiles of the simulated $\hat{V}_t$ marginal distribution, with break-points at the 16th, 32nd, 68th and 84th percentiles. The remaining few negative values were replaced by a small positive number, namely by 0.0001.

3. Empirical results

The discussion of the empirical results is divided into two subsections. We cover the two stages of the EMM procedure. The first stage, which pertains to the SNP density estimation, is covered in a first subsection. The parameter estimates of Heston's model are discussed in a second subsection, along with the computation of the market prices of risk.

3.1. SNP density estimation results

The SNP density estimation results are reported in Panel A of Table 1. We estimate the SNP density for the three types of data: (i) the log-returns on the S&P500, (ii) the log of the BS implied volatilities of the closest to maturity and ATM call options, and (iii) both series jointly. Rather than report the parameter estimates, we focus instead on the density structures as characterized by the tuning parameters K_z, K_X, M, M_μ, and M_R (see Appendix B for further details). To facilitate the interpretation of the results, Panel B describes the generic features of SNP densities for different combinations of the tuning parameters. In addition, Panel A reports the values of the objective function, s_n, based on the density in Eq. (B.1) found in Appendix B, the values of the Akaike information criterion (AIC), the Hannan and Quinn criterion (HQ), and the Schwarz Bayes information criterion (BIC). Each line reports the best BIC model within each class, as outlined in Panel B. The best model selected by BIC, is in boldface. Comparing Panels A and B of Table 1, we note that all series require models with $K_z > 0$, indicating non-Gaussian innovations. The fact that $M_R > 0$ means

[11] This outcome occurs because the linear transformation imposes lower and upper limits on the range of a volatility, unlike the log-transformation, which can potentially take any value on the real line.

Table 1
The SNP density estimation

We estimate the SNP density for the three types of data: (i) the log-returns on the S&P500; (ii) the log of the BS implied volatilities of the closest to maturity and the money call options; (iii) both series. Panel A reports the structure of the estimated densities and the values of the objective function s_n based on the density in (B.1), the values of the Akaike information criterion (AIC), the Hannan and Quinn criterion (HQ) and the Schwarz Bayes information criterion (BIC). Each line report the best BIC model within each class of density structures. The overall best model selected by BIC is in boldface. Panel B reports possible densities structures, as described in Gallant and Tauchen (1993). Panel C reports the reprojection SNP densities for spot volatility estimated with simulated data from Heston's model with the data of types (i)–(iii) in the conditioning set.

Panel A	Data type	K_z	K_X	M	M_μ	M_R	$s_n(\hat{\Xi})$	AIC	HQ	BIC
	S&P 500	0	0	1	3	0	1.4159	1.4185	1.4211	1.4256
		0	0	1	3	8	1.2140	1.2207	1.2275	1.2392
		9	**0**	**1**	**3**	**8**	1.1421	1.1534	1.1649	1.1848
		9	1	1	3	8	1.1371	1.1535	1.1703	1.1991
	BS vol's	0	0	1	9	0	1.0836	1.0893	1.0950	1.1050
		0	0	1	9	5	0.9510	0.9592	0.9676	0.9821
		6	**0**	**1**	**9**	**5**	0.8183	0.8296	0.8411	0.8610
		6	3	1	9	5	0.7889	0.8109	0.8335	0.8723
	Joint	0	0	1	4	0	2.4969	2.5077	2.5187	2.5376
		0	0	1	4	4	2.1626	2.1856	2.2092	2.2498
		4	**0**	**1**	**4**	**4**	1.9942	2.0214	2.0492	2.0970

Panel B	Parameter K, M_μ, M_R setting	Characterization of l_t
	$K_z = 0, K_X = 0, M \geqslant 0, M_\mu = 0, M_R = 0$	iid Gaussian
	$K_z = 0, K_X = 0, M \geqslant 0, M_\mu > 0, M_R = 0$	Gaussian VAR
	$K_z > 0, K_X = 0, M \geqslant 0, M_\mu > 0, M_R = 0$	non-Gaussian VAR, homogeneous innovations
	$K_z = 0, K_X = 0, M \geqslant 0, M_\mu \geqslant 0, M_R > 0$	Gaussian ARCH
	$K_z > 0, K_X = 0, M \geqslant 0, M_\mu \geqslant 0, M_R > 0$	non-Gaussian ARCH, homogeneous innovations
	$K_z > 0, K_X > 0, M > 0, M_\mu \geqslant 0, M_R \geqslant 0$	general non-linear process, heterogeneous innovations

Panel C	Data type	K_z	K_X	M	M_μ	M_R	$s_n(\hat{\Xi})$	AIC	HQ	BIC
	S&P 500	0	0	1	1	0	0.4601	0.4604	– 0.4607	0.4615
		4	0	1	1	0	0.4058	0.4065	0.4073	0.4090
		4	**1**	**1**	**1**	**0**	0.3859	0.3871	0.3886	0.3915
	BS vol's	**0**	**0**	**1**	**22**	**0**	– 0.8367	– 0.8343	– 0.8313	– 0.8256
		1	0	1	22	0	– 0.7889	– 0.7864	– 0.7833	– 0.7773
	Joint	0	0	1	7	0	– 25.2651	– 25.2618	– 25.2578	– 25.2498
		8	**0**	**1**	**7**	**0**	– 30.5975	– 30.5925	– 30.5865	– 30.5747
		8	6	1	7	0	– 30.6031	– 30.5775	– 30.5464	– 30.4855

that we also find ARCH effects, even with BS implied volatilities, in the Hermite polynomial expansions of the SNP densities. We also need autoregressive terms in the mean, since all expansions have $M_\mu > 0$. For BS implied volatilities we need the longest lags in the mean, namely M_μ is set equal to nine. The mean equation for BS volatilities is similar to a conditional second moment equation, which is comparable to the eight lags in the ARCH expansion of the return series (i.e. $M_R = 8$ for S&P 500 return series).[12]

It is worthwhile to examine the estimated density plots. In addition to the raw data, Fig. 1 also features plots of the estimated SNP densities for the univariate cases. Fig. 2 does the same for the bivariate case involving the joint process of returns and BS implied volatilities. The data section contains a discussion of Plots (A) and (C) in Fig. 1. These plots display the S&P 500 series and the corresponding BS implied volatilities. Plots (B) and (D) show the corresponding SNP densities, with a standard normal probability density function (p.d.f.) superimposed with a dashed line. The estimated densities show the familiar peaked, leptokurtic, and weakly skewed patterns. Plot (A) in Fig. 2 reports the joint density, Plot (B) gives the contour plot. The marginal densities appear in Plots (C) and (D). The contour plot suggests the presence of slight negative correlation between returns and volatility, which supports the presence of a leverage effect. Hence, in the estimation we imposed the restriction that $\rho \leqslant 0$.

3.2. SV model parameter estimates

We turn now to parameter estimates of Heston's SV model. Table 2 reports the estimation results. Traditionally, this model is estimated using only returns, so we report this configuration as a benchmark. Then we proceed to the estimation results relying exclusively on options, in the spirit of Pastorello et al. (1994). Next, we consider the structural parameters obtained from matching the moments dictated by the bivariate SNP score. The first observation is that Heston's SV model is rejected in every case, regardless of the data configuration. The standard Normal distribution is used to evaluate the z-statistics, which is an approximation to asymptotically χ^2 distributed test statistics, using GMM-type overidentifying restrictions (see Gallant and Tauchen, 1997, for further discussion).[13] All the z-statistics reported in Table 2 are large. The rejections vary dramatically across the rows in Table 2, depending on which data is used. In

[12] An alternative SNP density specification involving an AR-GARCH lead term has been suggested by Andersen et al. (1999). This alternative specification would be appropriate for discrete-time analogs to the continuous time diffusions, and potentially represents statistical efficiency gains. While these arguments clearly apply to univariate return series, it is not clear that they apply to BS volatilities and the bivariate models. We therefore use the original specification suggested by Gallant and Tauchen (1989).

[13] Pastorello et al. (1994) did not estimate the joint process as we propose to do in this paper.

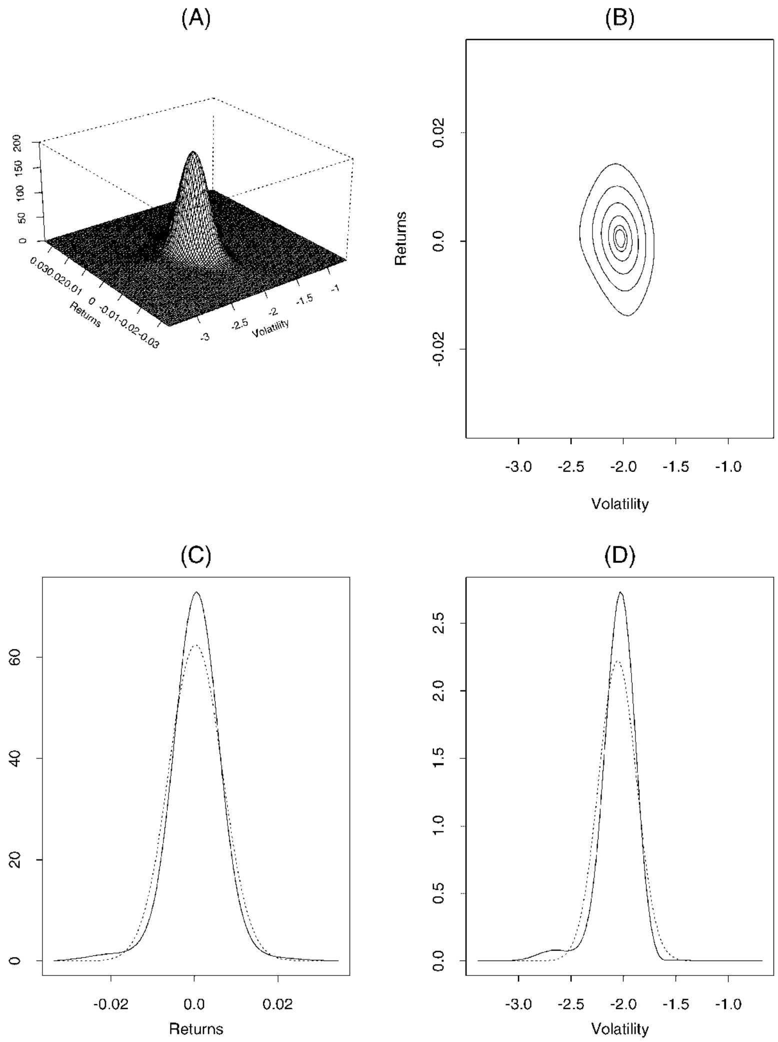

Fig. 2. The Bivariate SNP density. We estimate the joint SNP density for the following series: (i) the log-returns on the S&P500; (ii) the log of the BS implied volatilities of the closest to maturity and the money call options. The data are daily, and span the period from November 1985 to October 1993. Plot (A) shows the perspective plot of the estimated bivariate density; plot (B) is the contour plot at quantiles 10%, 25%, 50%, 75%, 90%, and 95%; plots (C) and (D) are the marginal densities of (i) and (ii) correspondingly, for which the solid line shows the plot of an SNP fit, and the dashed line shows the normal density with the same mean and variance.

particular, the returns data provide the weakest evidence against the model. However, the precision of the estimates is very poor. For example, the standard error for κ is 0.29. This result is consistent with previous findings (see, for

Table 2
EMM estimation results

We have estimated the stochastic volatility model

$$\frac{\mathrm{d}S(t)}{S(t)} = R\,\mathrm{d}t + \sqrt{V(t)}\,\mathrm{d}W_S^*(t),$$

and

$$\mathrm{d}V(t) = (\theta^* - \kappa^* V(t))\,\mathrm{d}t + \sigma_V\sqrt{1-\rho^2}\sqrt{V(t)}\,\mathrm{d}W_V^*(t) + \sigma_V\rho\sqrt{V(t)}\,\mathrm{d}W_S^*(t),$$

where the model is stated under the risk-free probability measure. We have denoted by asterisks only those parameters, which change when we make a transition to the objective probability measure. With such a change R would have to be substituted by μ_S and all asterisks removed. Three types of data were used in this estimation: (i) the log-returns on the S&P500; (ii) the log of the BS implied volatilities of the closest to maturity and the money call options; (iii) both series. Standard errors are reported in the parentheses, z-statistics testing the model are in the brackets, — denotes unidentifiable parameters. The reported parameter estimates are annualized.

Data type	$\hat{\mu}_S$	$\hat{\theta}$	$\hat{\theta}^*$	$\hat{\kappa}$	$\hat{\kappa}^*$	$\hat{\sigma}_V$	$\hat{\rho}$
S&P 500	0.12767392	0.01518484	—	0.92586310	—	0.06340828	– 0.01787568
[6.863]	(0.00000084)	(0.00514937)		(0.29340050)		(0.00814696)	(0.24603696)
Vol	—	0.01546074	0.00856834	0.92803334	0.70200229	0.06459151	– 0.01939315
[247.447]		(0.00017247)	(0.00001727)	(0.00003542)	(0.00019527)	(0.00028325)	(0.00032025)
Joint	0.10190398	0.01430366	0.00659546	0.93086335	0.69011763	0.06150001	– 0.01830000
[90.328]	(0.00007322)	(0.00005969)	(0.00007541)	(0.00009669)	(0.00026902)	(0.00007698)	(0.00009685)

example, Gallant et al., 1997). It is worth noting here that the estimated parameters are annualized, as is typically done in option pricing models. Hence the reported values are quite different from their GARCH counterparts, for instance, which are typically characterized on a daily basis. It is easy to provide the link between them, however, since the SV model allows for temporal aggregation. For example, the value of $\kappa \approx 0.93$ corresponds to $0.93/252 \approx 0.004$ on a daily scale, or roughly 0.996 as the corresponding persistence parameter in a GARCH(1,1) model.

Our major interest goes beyond appraising the model exclusively on the grounds of its statistical properties. Its pricing and hedging features, discussed in the next section, are the primary interest.

There are four parameters which overlap between the first row of Table 2, which pertains to parameter estimates obtain from returns data, and the second row, which reports the results obtained using BS implied volatilities. Recall from the discussion in Section 2.3 that, with options data, we can estimate the parameter vector $(\Theta_c,\ \Theta_o,\ \Theta_n)$ but not the drift μ_S under the objective measure. Therefore, the parameter vector $(\Theta_c,\ \Theta_o)$ is common to the first and second rows of Table 2. For all the parameters common to both data sets, we obtain roughly the same point estimates, yet the precision of the estimates are dramatically improved with the second data set. In theory, one would expect that options data should yield more precise diffusion parameter estimates, as almost all estimated parameters are related to the volatility process specified in Eq. (2). Therefore, looking at the process through the observed implied volatilities should give us more precision, an observation also made by Pastorello et al. (1994).[14] Nevertheless, the standard errors we find and report in Table 2 are extremely small. Pastorello et al. (1994) report a ratio of roughly 4 to 1 for the standard errors obtained from returns and options data. In Table 2, we find standard errors up to 10,000 times smaller, as in the case of κ. Further investigation has revealed to us that one should be very careful with the computation of standard errors. We have discovered that our results are affected by the numerical instabilities in the computation of standard errors, which are computed via the Jacobian equation $(\partial/\partial\Theta')m(\hat{\Theta},\ \hat{\Xi})$.[15] As there are no analytic expressions for the Jacobian, the derivatives are evaluated numerically. In the EMM code, as with any numerical gradient code, one can change the size of the lower bound on the differencing interval, which is denoted as h. Under normal circumstances, the choice of h should not substantially affect the resulting the standard error computations. We computed standard errors using different values of h. The

[14] Pastorello et al. (1994) consider the model of Hull and White (1987), and conduct a Monte Carlo study showing the dramatic improvement of parameter estimates when options data are used.

[15] We are grateful to the Referee and to George Tauchen, who helped us to find the source of the computational pitfalls.

default value for h in the EMM code is 10^{-7}, which yields the standard errors appearing in Table 2. We considered values for h of 10^{-i}, $i = 4, 5, 6, 7$, and 8. We found that deviations from the default value of h yielded in some cases smaller standard errors. For example, for θ we obtained standard errors ranging from 0.00003550 ($h = 10^{-8}$) to 0.00017247 reported as in Table 2. In other cases, we obtained substantially larger standard errors, such as the κ values ranging from 0.00003542, reported in Table 2 to 0.01496481 ($h = 10^{-5}$). Interestingly, the same sensitivity analysis of the numerical computations applied to the first row of Table 2, showed that the standard error computations were invariant to the choice of h when returns data are used. The instability of the computations can therefore only be explained by the considerably more complicated functions involved in computing $(\partial/\partial\Theta')m(\hat{\Theta}, \hat{\Xi})$ when options data, and hence an options formula, are used. At this point it is difficult to say which of the errors are correct when options data are used. This subject clearly requires a more thorough investigation. As we are sailing in unchartered waters, this subject surely will require further future research. Finally, to conclude the discussion of the parameter estimates with options data, note the large size of the z-statistic, 247.447, which indicates serious inconsistencies between the model and the data. Specifically, the z-statistic indicates that our model does not explain the information extracted from the options prices.

The model estimated with the joint data set of returns and options is also reported in Table 2. Here, we can fit both the objective and risk-neutral density parameters and make direct comparisons with the previous two univariate models. The bivariate model shows an improvement in fit when compared to the options-based approach, the z-statistic is greatly reduced, although the model is still strongly rejected. The precision of the estimates is the best of the different configurations, while the point estimates roughly remain the same.[16] The observation regarding the stability of point estimates is important as the bivariate estimation approach is affected by the non-synchronicity in options and returns data. This result confirms the observation made in Appendix C, that matching moments instead of sample paths appears to render any spurious non-synchronicity effects insignificant.

Since we are interested in which type of data we should use to price and hedge options, we can conclude from Table 2 that the returns series should not be used alone as we can with the univariate data sets, not directly infer the parameters under the risk-neutral probability measure, unless auxiliary assumptions are made. The two competing data sets which allow us to identify directly the necessary parameters are the BS implied volatilities and the joint returns and implied volatilities series. We will therefore focus exclusively on these two

[16] We have to again express caution, as the observation regarding the numerical complexities of computing standard errors when options data are used also applies to the bivariate setup.

alternatives in order to appraise how options are priced, and how well they perform for the purpose of hedging.

As noted in Section 1.1, the complete parameter vector $\hat{\Theta}$ allows us to compute the prices of risk. The estimated parameters also allow us to estimate the volatility $V(t)$ via the reprojection procedure. In the next section we will give the details of the actual implementation. Therefore, we can compute the sample paths for prices of risk appearing in Eqs. (7) and (8). These parameters are reported in Fig. 3, which has three panels. Plot (A) shows the values of $\lambda_1(t)$ computed with the reprojected volatilities, while Plot (B) does the same for $\lambda_2(t)$. Plot (C) reproduces the sample path of the actual reprojected volatilities $\hat{V}(t)$. The first two plots display the time series processes, which represent the risk adjustments to the Brownian motions W_S and W_V. Realizing that $\mathrm{d}W_V(t) = \mathrm{d}W_V^*(t) - \lambda_2(t)\,\mathrm{d}t$, it is clear that the risk-adjusted densities may differ substantially from the objective Brownian motions. Hence, as it is not possible to directly simulate processes under the risk neutral measure, one has to be careful when simulating fundamental processes.

4. Assessing pricing and hedging performance

While statistical criteria for model selection are important, financial criteria such as the out-of-sample pricing and hedging performance, form the basis of our model selection process. Therefore, we investigate the performance of the alternative model specifications obtained from the previous section. In order to assess the magnitude of the forecast errors, we use the Black–Scholes valuation model as a benchmark. All evaluations reported in this section are obtained using the post-estimation portion of our data set. The estimation sample used in the preceding section covered November 1985 to October 1993, whereas the sample used to appraise the models runs from November 1993 to October 1994.[17] Since we use out-of-sample data, our results will not be contaminated by in-sample data mining. In our discussion of Plot (B) of Fig. 1, which displays the BS implied volatilities, we noted that a downward trend in volatilities, which began around 1991, reversed in the post-estimation part of our data set. This reversal will make the out-of-sample exercise particularly interesting.

A first subsection is devoted to the pricing of options. A second subsection looks at hedging, and we devote a separate subsection to a simple volatility filter for option pricing which emerges from our analysis. The examination of this

[17] One could call this genuine out-of-sample approach, as in Dumas et al. (1998) or Jacquier and Jarrow (1998), unlike the pricing and hedging performance evaluations which rely on in-sample one-step ahead forecasts.

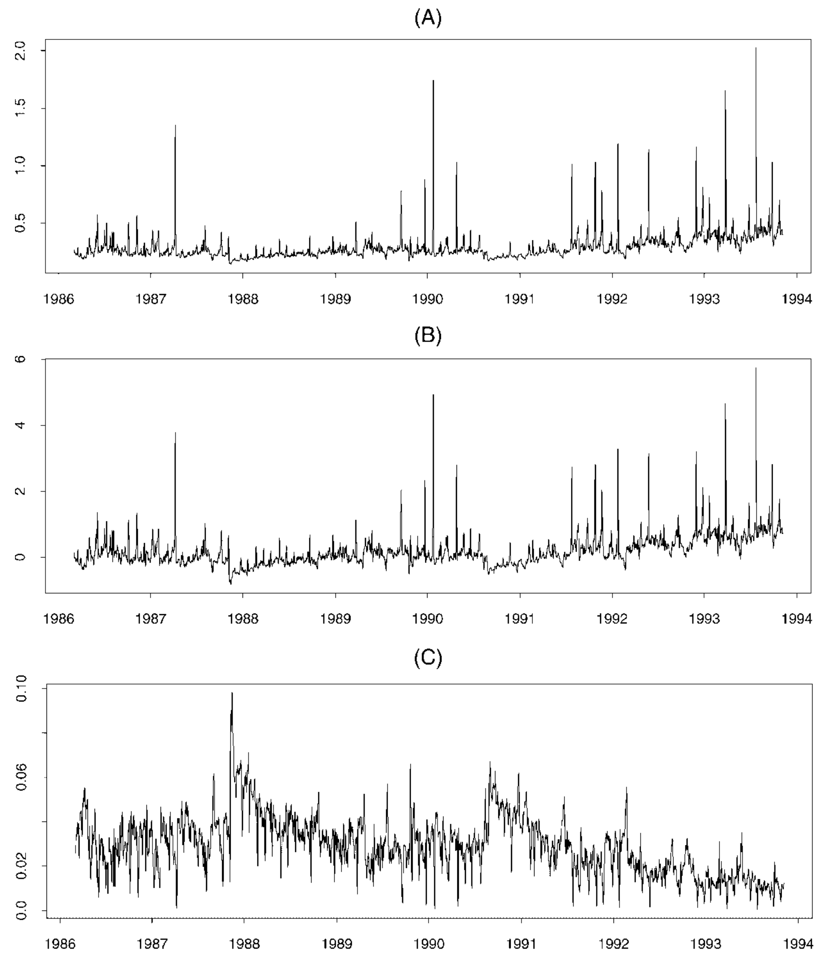

Fig. 3. The Market price of risk. The estimation of the Stochastic Volatility model based on bivariate data allows us to compute the market price of risk, $\lambda(t)$. We plot these prices for the entire estimation sample of daily data which span the period from November 1985 to October 1993. Panel (A) reports $\lambda_1(t)$, the market price of the asset risk, see Eq. (7). Panel (B) reports $\lambda_2(t)$, the market price of the volatility risk, see Eq. (8). Panel (C) is the time series of the reprojected volatility, $V(t)$, which was used to compute $\lambda_1(t)$ and $\lambda_2(t)$.

filter is important given that all model configurations yield roughly the same parameter estimates (recall the results in Table 2). Hence, whenever the same option pricing formula is used, the differences in hedging and pricing are mostly due to filtering.

4.1. Pricing reliability

The model parameter estimates, the reprojection filters, and Heston's European call pricing formula yield the ingredients needed to price any SPX contract with a particular time to maturity, strike, and cash price for the underlying index. Equipped with these tools, we can compare the call prices predicted by the models with observed prices. In our appraisal, we do not consider in-the-money calls with moneyness greater than 1.03, as there is a very thin market for such options. Furthermore, we separate the calls into twelve groups, according to moneyness and time to maturity. Since the number of contracts varies through time, we assume that each group contains n_t options at time t.

Two measures of pricing errors are considered. The first is an absolute measure, denoted $D_{\mathrm{a}}^{\mathrm{p}}$, and has a dollar value scale. The second is a relative measure, denoted $D_{\mathrm{r}}^{\mathrm{p}}$, and reflects percentage deviations. The two measures are defined as follows:

$$D_{\mathrm{a}}^{\mathrm{p}} = \sqrt{\frac{1}{\sum_{t=1}^{T} n_t} \sum_{t=1}^{T} \sum_{i=1}^{n_t} (C_{it}^{\mathrm{observed}} - C_{it}^{\mathrm{model}})^2} \tag{20}$$

and

$$D_{\mathrm{r}}^{\mathrm{p}} = \sqrt{\frac{1}{\sum_{t=1}^{T} n_t} \sum_{t=1}^{T} \sum_{i=1}^{n_t} \left(\frac{C_{it}^{\mathrm{observed}} - C_{it}^{\mathrm{model}}}{C_{it}^{\mathrm{observed}}} \right)^2}. \tag{21}$$

The results are divided in four different categories of moneyness. These are: deep out-of-the-money (OTM) with S/K less than 0.94, OTM with $0.94 < S/K < 0.97$, slightly OTM with $0.97 < S/K < 1$ and slightly in-the-money with $1 < S/K < 1.03$. The last two categories are usually viewed as at-the-money options. Three maturity horizons are considered for the moneyness categories, including short (less than 60 days), medium (between 60 and 180 days) and long (more than 180 days). The respective sample sizes for each of the twelve out-of-sample cells are reported in Table 3. For deep OTM options, we have between 313 and 448 contracts to use for computing pricing errors, and slightly fewer for computing hedging errors.[18] For the remaining three moneyness categories, there is a pronounced downward trend in the number of contracts as time-to-maturity increases. The maximum number of observations is 1377, and the smallest is 90. Table 4 contains the pricing errors, $D_{\mathrm{a}}^{\mathrm{p}}$ and $D_{\mathrm{r}}^{\mathrm{p}}$, computed for three different model specifications. The first specification, 'BS', prices options with the Black–Scholes model and a volatility estimate based on

[18] The computation of hedging errors, which require contracts with trading activity on two consecutive days, will be discussed in the next subsection.

Table 3
Options groups used for evaluatiing model performance

To assess pricing and hedging performance of alternative model specifications, we separate the call options into twelve groups, sorted by moneyness and time to maturity. Moneyness is computed as the ratio of the current S&P 500 index level and the contract's strike price. The table reports the number of observations in each group. The total number of observations is reported in the headers.

Moneyness	Pricing sample ($n = 7424$) Days to expiration			Hedging sample ($n = 7132$) Days to expiration		
	< 60	60–180	$\geqslant 180$	< 60	60–180	$\geqslant 180$
< 0.94	337	448	313	299	437	279
0.94–0.97	1080	828	168	1037	818	159
0.97–1.00	1377	938	195	1324	930	187
1.00–1.03	1038	612	90	975	604	83

the previous day ATM option contract (see for instance Bates, 1996, for further discussion on this approach).

The aim of the present exercise is to compare the pricing performance of different models, giving each an equal chance to succeed. Since the implementation of BS is ad hoc, we make every attempt to implement SV in a consistent way, although this task is very difficult. The discussion in Bates (1996) allows us to view the ATM implied volatility as a way to introduce some consistency into the BS implementation. Furthermore, if we price a deep OTM contract based on its volatility implied from the previous day, we will obtain a smaller pricing error. However, this practice does not give an equal chance to results based on an SV formula, which could have been estimated based on OTM contracts as well. Since the focus of the present study is information extraction from ATM options and asset returns, we believe that the proposed scheme allows for a balanced approach to the task.

Next is the specification, denoted 'Vol', calculated using Heston's call price model estimated with a univariate options data set and a univariate reprojection filter. The specification denoted 'Joint' corresponds to the bivariate model. Besides the absolute and relative pricing errors for each of the twelve moneyness/maturity groups, Table 4 also reports whether there are any statistically significant differences between the pricing errors computed from these three specifications. We observe that, in general, pricing errors can be large, with absolute dollar values ranging from 5.23 to 0.42. Not surprisingly, large errors occur for longer maturities. The relative errors range from 20.71 to 0.14, with the same pattern.

For our purpose, a comparison across different model specifications is more important. As noted before, to appraise the differences we compute formal tests

Table 4
The SV model pricing errors

We consider post estimation sample data from November 1993 to October 1994 to compute out-of-sample pricing errors. We evaluate three model specifications, according to moneyness and time to maturity. The first is '*BS*', which involves pricing options with the Black–Scholes model and a volatility estimate based on the previous day at-the-money option contract. Next is the specification denoted 'Vol' in the table, which involves Heston's call price model estimated using a univariate specification involving options. The specification denoted 'Joint' corresponds to the bivariate specification. The first column is the moneyness category. The next column lists the model specifcations. The third column reports the absolute standard error $D_{\mathrm{a}}^{\mathrm{p}}$ (see Eq. (20)) and the relative standard error $D_{\mathrm{r}}^{\mathrm{p}}$ (see Eq. (21)) by each maturity group. The remaining columns report the χ_1^2-statistics from the GMM overidentifying restrictions tests. These tests are aimed to determine whether the means of $((C_{it}^{\text{observed}} - C_{it}^{\text{model}})/(C_{it}^{\text{observed}}))^2$ computed from a pair of model estimates are significantly different. We report results of the tests within each group and overall. * indicates significance at the 1% level, ** indicates significance at the 5% level, and *** indicates significance at the 10% level.

Moneyness	Model	Pricing Errors Days to expiration						Models	Tests Days to expiration			All groups test
		< 60		60–180		⩾ 180			< 60	60–180	⩾ 180	
		$D_{\mathrm{a}}^{\mathrm{p}}$	$D_{\mathrm{r}}^{\mathrm{p}}$	$D_{\mathrm{a}}^{\mathrm{p}}$	$D_{\mathrm{r}}^{\mathrm{p}}$	$D_{\mathrm{a}}^{\mathrm{p}}$	$D_{\mathrm{r}}^{\mathrm{p}}$					
< 0.94	BS	0.54	4.62	1.56	20.71	3.12	4.93	Vol-BS	3.61***	5.79**	0.29	16.81*
	Vol	0.42	2.01	0.92	12.58	2.26	4.06	Vol-Joint	2.69	10.40*	1.32	24.01*
	Joint	0.44	2.42	1.23	17.35	2.74	4.80	Joint-BS	2.81***	1.30	0.01	5.93**
0.94–0.97	BS	0.82	12.12	1.88	12.35	4.62	0.45	Vol-BS	10.95*	4.17**	17.64*	16.81*
	Vol	0.58	8.39	1.38	7.63	3.41	0.33	Vol-Joint	15.90*	5.42**	11.33*	24.01*
	Joint	0.73	10.79	1.66	9.91	3.86	0.37	Joint-BS	2.10	2.27	8.20*	5.93**
0.97–1.00	BS	1.13	5.78	2.48	0.38	5.23	0.31	Vol-BS	5.72**	19.63*	12.28*	16.81*
	Vol	1.05	4.46	2.20	0.30	4.24	0.24	Vol-Joint	4.93**	38.20*	18.07*	24.01*
	Joint	1.22	4.93	2.48	0.36	4.71	0.27	Joint-BS	3.08***	1.38	5.78**	5.93**
1.00–1.03	BS	1.25	0.14	2.31	0.17	4.63	0.18	Vol-BS	0.07	3.04***	8.16*	16.81*
	Vol	1.35	0.15	2.33	0.16	3.85	0.14	Vol-Joint	11.31*	13.49*	4.19**	24.01*
	Joint	1.40	0.16	2.43	0.17	4.21	0.16	Joint-BS	4.42**	0.07	3.24***	5.93**

based on the following set of moment conditions, namely:

$$\begin{pmatrix} \left(\sum_{t=1}^{T} n_t\right)^{-1} \sum_{t=1}^{T} \sum_{i=1}^{n_t} ((C_{it}^{\text{observed}} - C_{it}^{\text{vol.model}})/(C_{it}^{\text{observed}}))^2 - m \\ \left(\sum_{t=1}^{T} n_t\right)^{-1} \sum_{t=1}^{T} \sum_{i=1}^{n_t} ((C_{it}^{\text{observed}} - C_{it}^{\text{joint.model}})/(C_{it}^{\text{observed}}))^2 - m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{22}$$

where m is the common mean relative error under the null hypothesis. Computations using the absolute errors could be performed as well, but are not reported here. The common parameter m in Eq. (22) implies that the two pricing errors have the same mean. Rejecting the null hypothesis of a common mean implies that one model is on average superior in producing price estimates. We can proceed in several ways to test the null hypothesis. The errors form a panel data set. On a daily basis, each maturity/moneyness cell contains a total of n_t contracts. The sample data are therefore correlated, which prevents us from considering a simple t-statistic. Instead, we use a GMM-based procedure involving a Newey and West estimator for the covariance matrix. The common mean, m, entails one overidentifying restriction on the moment conditions in (22). Therefore, the overidentifying restrictions test statistic is distributed as a χ^2, with one degree of freedom under the null hypothesis (see Hansen, 1982). The test statistics are reported in Table 4. We report overidentifying restrictions test statistics for pairwise comparisons between Vol and BS, Vol and Joint, and finally Joint versus BS. These tests are reported for the three maturity horizons and for each of the moneyness categories. In addition, we also report one joint statistic per moneyness category for pricing errors over all three maturity horizons combined.

The most striking result which emerges from the table is the dominance of the Vol approach. The comparisons of Vol and BS, Vol and Joint are always significant at the 1% level for all moneyness categories, with the three maturities combined. In each of the individual maturity-moneyness cells, we also observe the statistical significance of Vol, though there are exceptions, and most tests are only at the 5% or 10% level. Moreover, in each of the cases in which the overidentification test involving pairwise comparisons of Vol with either BS or Joint are significant, we find that the former has the smallest pricing error. Hence, the estimation of SV models involving only ATM options outperforms BS and the bivariate approach, denoted Joint. Besides yielding superior pricing performance, we observe from the results reported in Table 4 that the improvements over Black–Scholes volatility filter range from a 1.21 dollar reduction for OTM long maturity, to a small eight cents for slightly OTM short maturity. The Vol specification outperforms the Joint specification with the largest improvement for the deep OTM long maturity category, although the improvement is only 48 cents. The smallest difference is also in the same moneyness category but has short maturity and is equal to only 2 cents. It is worth emphasizing here that the comparisons between Joint and Vol involved the same call price formula of Heston and, basically, the same parameters (see Table 2). Hence, the differences are very much due to the filtering procedures, i.e. the procedure to extract the latent volatility process from the data.

Since the outperformance of the joint reprojection filter by the implied volatility reprojection filter model specification is somewhat surprising, we conducted further tests which are not reported in the table. These additional tests required a departure from the reprojection procedure described in Section 2.4. Since the details of the reprojection filter underlying Vol will be discussed in the next section, we focus only on the augmentation from Vol to Joint filtering, where the bivariate filter supplements the options data with returns series. It is not clear from the fitted SNP density in the reprojection routine whether certain individual moments of returns may provide information that enhances the pricing of options. To separate the potential contribution of individual moments, we computed reprojection filters only involving a particular moment of returns in addition to implied volatilities. For instance, to determine whether the mean return has any pricing information, we consider a bivariate filter with implied volatilities and past and concurrent returns. Hence, we ignore the Hermite polynomial expansion terms in the SNP density. This specific filter investigates the improvements brought by the leverage effect in the stock index, not already incorporated in the implied volatilities, to the pricing of European-type contracts. The same strategy can be applied to isolate the informational content of the second, third, and fourth power of returns. These computations reinforced our findings with the SNP reprojection filter, with one potentially important exception. We found that the conditional kurtosis, relying on a filter using a lag operator in the fourth power of returns in addition to implied volatilities, helps to improve the pricing of long maturity slightly in-the-money option contracts. The statistically significant improvement decreased the pricing error from 3.85 to 3.03 dollars.

The general conclusion we can draw so far is that Heston's call price model with a relatively simple volatility filtering scheme, using past options data, yields the most desirable outcome across all three model specifications considered. All models perform relatively well at short maturities, of course. The discrepancies emerge in the cells involving long maturities and OTM contracts. It should be noted that the estimation of the SV models was confined to ATM options. Therefore, one would expect that fitting and filtering with contracts similar to those priced out-of-sample, like OTM contracts, would certainly reduce the pricing errors of the Vol and Joint specifications. We will elaborate on this idea in the concluding section, and delegate this exploration to future research. In Section 4, we revisit some of these questions when we consider separating the filtering effects from the pricing kernel effects.

4.2. Hedging performance

A common practice for hedging market risk relies on a combination of stocks and options. Many hedging techniques exist, some of which are quite sophisticated and require the simultaneous consideration of several instruments. Many of the sophisticated hedging strategies are difficult to implement in practice (see Figlewski, 1998, for further discussion). We therefore concentrate on a simple minimum variance hedging strategy which uses just one option. In particular, if we want to hedge our position in one call with $N_S(t)$ shares of stock, we choose the number of shares in such a way that the remaining cash position

$$CP(t) = C(t,\ \tau, K) - N_S(t)S(t) \tag{23}$$

has minimum variance, which is achieved by taking

$$N_S(t) = \Delta_S(t,\ \tau, K) + \frac{\rho\sigma_V}{S(t)}\Delta_V(t,\ \tau, K), \tag{24}$$

where

$$\Delta_S(t,\ \tau, K) = \frac{\partial C(t,\ \tau, K)}{\partial S} = \Pi_1. \tag{25}$$

Note that this simple formula can be obtained because the SV option price in Eq. (3) is homogeneous of degree one in $S(t)$ and K (for further details, see Nandi, 1998). Finally, we also have that

$$\Delta_V(t,\ \tau, K) = \frac{\partial C(t,\ \tau, K)}{\partial V} = S(t)\frac{\partial \Pi_1}{\partial V} - K\mathrm{e}^{-R\tau}\frac{\partial \Pi_2}{\partial V}, \tag{26}$$

The Heston SV model does not account for all sources of risk. For example, this model assumes that the interest rate is constant. Therefore, if we try to unload our position the next day, we will not end up with a zero balance. The hedging error will therefore be

$$H(t + \Delta t) = N_S(t)S(t + \Delta t) + CP(t)\mathrm{e}^{R\Delta t} - C(t + \Delta t,\ \tau - \Delta t, K), \tag{27}$$

where Δt is equal to one day.

We use summary statistics similar to those in the previous subsection to report the hedging performance for each model. In particular, we construct an absolute error ($D^{\mathrm{h}}_{\mathrm{a}}$) defined as

$$D^{\mathrm{h}}_{\mathrm{a}} = \sqrt{\frac{1}{\sum_{t=1}^{T} n_t}\sum_{t=1}^{T}\sum_{i=1}^{n_t} H_i^2(t)} \tag{28}$$

and comparable relative ($D_{\mathrm{r}}^{\mathrm{h}}$) measures. To obtain the latter, let us rewrite Eq. (27) as

$$H(t+\Delta t) = CP(t)\mathrm{e}^{R\Delta t} - \overline{CP}(t+\Delta t) = \mathrm{Cash}_{t+\Delta t}^{\mathrm{model}} - \mathrm{Cash}_{t+\Delta t}^{\mathrm{observed}}, \tag{29}$$

where $\overline{CP}$ is different from CP because the number of shares was computed in the previous period, not as defined Eq. (23). This yields the relative measure of the hedging error, which is defined as

$$\begin{aligned} D_{\mathrm{r}}^{\mathrm{h}} &= \sqrt{\frac{1}{\sum_{t=1}^{T} n_t} \sum_{t=1}^{T} \sum_{i=1}^{n_t} \left(\frac{\mathrm{Cash}_{it}^{\mathrm{observed}} - \mathrm{Cash}_{it}^{\mathrm{model}}}{\mathrm{Cash}_{it}^{\mathrm{observed}}} \right)^2} \\ &= \sqrt{\frac{1}{\sum_{t=1}^{T} n_t} \sum_{t=1}^{T} \sum_{i=1}^{n_t} \left(\frac{H_i(t)}{\overline{CP}_i(t)} \right)^2}. \end{aligned} \tag{30}$$

The hedging performance comparisons for the BS, Vol and Joint are reported in Table 5, again using the classification of 12 moneyness/time-to-maturity cells. The absolute and relative errors are complemented with statistical tests built on moment conditions, similar to Eq. (22), which yield an overidentifying restrictions test. We compute the overidentifying restrictions tests for all individual cells as in Table 4, including the grouped maturity tests shown in the last column of Table 5. We observe that the Vol specification is again the dominant one, yet not significantly different from both alternatives. In fact, all three specifications perform rather well, and hardly any of the overidentifying restrictions tests are significant, except in four individual cells suggesting some very weak improvements arising from the use of the Vol specification. Overall, we conclude from the results in Table 5 that hedging strategies, unlike the pricing errors, appear to be insensitive to model specification.

5. Disentangling filtering and pricing kernel effects

To determine the advantages of various filters, we need to construct a situation in which we can separate filtering effects from the pricing formula. In Table 4, for instance, the comparison of the Black–Scholes model pricing BS with the Heston model Vol, involves two components, one due to the filtering and the other due to the different pricing formula. In the first subsection, we examine all models through the window of the BS option pricing model, which allows us to separate and appraise the effect of the alternative volatility filters on the pricing of options. In the second subsection, we compare pricing performance of SV, BS, and GARCH option pricing models with their proper pricing formulas. For GARCH models, we follow Duan (1995), who developed a framework for option pricing under a local risk-neutral probability measure. This

Table 5
The SV model hedging errors

We consider post estimation sample data from November 1993 to October 1994 to compute out-of-sample hedging errors. We evaluate three model specifications by moneyness and time to maturity. The first is 'BS', which involves pricing options with the Black–Scholes model and a volatility estimate based on the previous day at-the-money option contract. Next is the specification denoted 'Vol' in the table, which involves Heston's call price model estimated using a univariate specification involving options. The specification denoted 'Joint' corresponds to the bivariate specification. The first column is the moneyness category. The next column lists the model specifcations. The third column reports the absolute standard error D_a^h (see Eq. (28)) and the relative standard error D_r^h (see Eq. (30)) by each maturity group. The remaining columns report the χ_1^2-statistics from the GMM overidentifying restrictions tests. These tests are aimed to determine whether the means of $((C_{it}^{\text{observed}} - C_{it}^{\text{model}})/(C_{it}^{\text{observed}}))^2$ computed from a pair of model estimates are significantly different. We report results of the tests within each group and overall. * indicates significance at the 1% level, ** indicates significance at the 5% level, and *** indicates significance at the 10% level.

Moneyness	Model	Hedging errors Days to expiration						Models	Tests Days to expiration			All groups test
		< 60		60–180		⩾ 180			< 60	60–180	⩾ 180	
		D_a^h	D_r^h	D_a^h	D_r^h	D_a^h	D_r^h					
< 0.94	BS	2.35	0.968	1.33	0.875	1.35	0.067	Vol-BS	0.34	0.01	3.44***	0.89
	Vol	2.43	0.751	1.33	0.881	1.37	0.029	Vol-Joint	0.81	2.24	1.02	1.07
	Joint	2.44	1.525	1.37	1.396	1.36	0.092	Joint-BS	0.47	2.18	0.30	1.10
0.94–0.97	BS	2.04	1.202	1.10	0.804	1.51	0.024	Vol-BS	0.05	0.03	1.76	0.89
	Vol	2.06	1.064	1.07	0.718	1.62	0.009	Vol-Joint	1.24	0.99	3.75***	1.07
	Joint	2.07	0.470	1.07	0.053	1.61	0.010	Joint-BS	1.65	1.10	1.67	1.10
0.97–1.00	BS	2.26	5.368	1.30	0.013	2.37	0.012	Vol-BS	0.87	1.99	1.07	0.89
	Vol	2.25	1.570	1.30	0.007	2.34	0.009	Vol-Joint	1.00	4.81**	1.77	1.07
	Joint	2.22	0.066	1.30	0.008	2.35	0.009	Joint-BS	1.04	1.69	1.12	1.10
1.00–1.03	BS	1.75	0.005	1.67	0.005	3.97	0.012	Vol-BS	0.01	2.38	1.74	0.89
	Vol	1.77	0.005	1.66	0.005	3.38	0.011	Vol-Joint	1.01	7.20*	1.45	1.07
	Joint	1.77	0.005	1.67	0.005	3.54	0.011	Joint-BS	0.29	0.19	1.39	1.10

comparison will allow us to see whether any pricing formula more advanced than the Black–Scholes model adds any improvement in pricing performance. Moreover, this two-tier comparison allows us to separate the role of volatility filtering from that of the pricing kernel in pricing derivative contracts.

5.1. Filtering latent spot volatility

We propose to use BS, GARCH, and SV models first as volatility filters, and then, evaluate them via the BS formula. To facilitate the discussion, we introduce the following notation for the filtering schemes (estimators of the $\sqrt{V(t)}$): σ_t^I is the estimate of today's volatility by the previous day's BS implied ATM option, σ_t^G is the estimate of instantaneous volatility using a univariate GARCH(1, 1) model (see below for details), σ_{rt}^R is the estimate of instantaneous volatility via reprojection using a univariate scheme based on returns, and σ_{vt}^R is the estimate of instantaneous volatility via reprojection using a univariate scheme based on BS implied volatitility.

Each of these filtering schemes warrant further discussion. Since the univariate Vol specification of Table 4 represents the most successful and novel approach, we first discuss the practical implementation of σ_{vt}^R. Constructing a reprojection filter involves a model selection procedure for the tuning parameters of the SNP density, namely K_z, K_X, M, M_μ and M_R. This time, the density is estimated this time with simulated data. To facilitate comparison, we report the reprojection model selection results in Panel C of Table 1 along side with the empirical sample results, which appear in Panel A of Table 1. Note that the SNP densities applied to the sample data are not comparable to those used in the reprojection. Recall from equation (18) that we are computing a conditional distribution of latent spot volatility given BS implied volatilities. Hence, in the case of Vol the densities reported in Panel A are univariate conditional densities of BS implied volatilities given their own past, whereas in Panel C we fit a conditional reprojection density of latent spot volatilities conditional on implied volatilities. Therefore, we do not expect the SNP densities to coincide. Moreover, the densities in Panel C cannot be autoregressive, involving past latent volatilities. Hence, no AR or ARCH parts appear in the fitted reprojection densities. The tuning parameters are again selected using the BIC criterion. For the reprojection density involving BS implied volatilities, the model specification that minimizes the BIC criterion is a simple linear operator with 22 lags. In Panel C of Table 1, we note that M_μ equals 22. Aside from a constant, no other tuning parameters are set to values greater than zero. Hence, the ATM implied volatilities are combined in a weighted historical moving average with a $22-$ day window.

The filter is remarkably simple, given that we started out with a general specification of a SNP conditional distribution. Table 6 lists the filter weights

Table 6
Reprojection model with implied volatilities

The coefficients of the reprojection model used to filter the latent volatility process from the observed past implied volatilities are listed in the table. We also report parameters of AR(22) model for the logarithm of implied volatilities.

Lags	Reprojection		AR(22) model	
	Parameters	Std. err.	Parameters	Std.err.
0	− 0.16320	0.01043	− 1.00038	0.05046
1	0.24514	0.00149	0.45085	0.06300
2	0.20591	0.00144	0.15847	0.06863
3	0.18804	0.00147	− 0.03892	0.06936
4	0.16293	0.00154	0.10117	0.06895
5	0.14055	0.00167	0.00499	0.06893
6	0.11489	0.00184	0.00697	0.06887
7	0.09978	0.00203	0.04363	0.06875
8	0.08792	0.00225	− 0.04505	0.06827
9	0.07836	0.00253	− 0.01029	0.06835
10	0.06675	0.00285	0.01011	0.06821
11	0.05675	0.00321	0.02365	0.06805
12	0.05069	0.00337	− 0.08067	0.06818
13	0.04506	0.00365	0.07710	0.06840
14	0.04715	0.00384	0.01280	0.06865
15	0.04298	0.00405	− 0.13689	0.06861
16	0.04179	0.00449	0.07321	0.06911
17	0.04882	0.00468	0.05775	0.06927
18	0.04386	0.00503	− 0.10836	0.06954
19	0.03987	0.00529	0.12760	0.06964
20	0.03976	0.00554	0.01582	0.07011
21	0.03277	0.00587	− 0.13449	0.06957
22	0.03280	0.00629	− 0.00279	0.06478

which extract spot volatility from BS implied volatilities. The weights range from 0.25 to 0.03. The intercept is negative, and equal to − 0.16. The largest weights in the filter appear at the shorter lags, and the decrease in weights is roughly linear in the lags. Hence, the extraction scheme puts less weight on observations from the options markets that date back about one month, or 22 trading days.

The simplicity of this scheme is rather surprising if one thinks of the complexity of the task. Indeed, alternative filters, such as those proposed by Harvey et al. (1994), Jacquier et al. (1994), and Nelson and Foster (1994), require highly nonlinear functions of returns. Hence, the virtue of using volatility data to predict future spot volatility is that one can limit the filter to a linear structure. It should be noted, however, that the construction of the filter is not as simple as running a linear regression model, as spot volatility is a latent process. Therefore, the calculation of the filter weights remains a nontrivial task which cannot be performed by simple regression methods. However, once diffusion parameters are available the task can be completed easily. To clarify this point, suppose that we would consider the simple task of regressing implied rather than spot volatilities on the same window of past implied volatilities. Such a linear regression model is also reported in Table 6, with its lag coefficients appearing along side the reprojection filter. The OLS parameter estimates show no resemblance, either numerically or statistically, to extraction filter weights. The difference between the two lag polynomials arises because implied volatility is, unlike spot volatility, related to the expected volatility over the remaining time to maturity of a contract (see, for instance, Hull and White, 1987 or Ghysels et al., 1996).[19]

We noted before that to construct volatility estimates based on returns, σ_{rt}^R, Harvey et al. (1994) suggest using the approximate Kalman filter based on a discrete time SV model. The exact filter is derived by Jacquier et al. (1994) in the context of a Bayesian analysis of the same model. Along the same lines, one can use the reprojection approach using EMM applied to returns data. The parameter estimates are reported in Table 2, first line. This set of estmiates form the first of two return-based filters we will consider. The reprojection SNP tuning parameters for returns are again reported in Panel C of Table 1. Panel C of Table 1 also displays the reprojection density of the joint specification. We did not include this filter in the analysis of this section because the SV model using options data was shown to dominate the joint model in terms of pricing.

The next filter, namely σ_t^G, is based on the work of Nelson and Foster (1994) who show that diffusion limit arguments applied to the class of ARCH models provide a justification of ARCH models as filters for instantaneous volatility. To implement this filter, let us consider the GARCH option pricing model specified in Duan (1995). This model is described under the local risk-neutral probability measure by

$$\log S_t/S_{t-1} = R - \tfrac{1}{2}V(t) + \sqrt{V(t)}\varepsilon_t^* \tag{31}$$

and

$$V(t) = \alpha_0 + \alpha_1 V(t-1)(\varepsilon_{t-1}^* - c - \lambda)^2 + \beta_1 V(t-1), \tag{32}$$

---

[19] Furthermore, this volatility filtering scheme provides insights into the literature that evaluates implied volatilities as estimates of spot volatility. The motivation for the work in this area is the implication of the Hull and White (1997) model that the BS volatility is equal to the average integrated volatility over the remaining life of an option. The conclusions of the studies are somewhat mixed (see Bates, 1996, for a review). The major findings are that implied volatilities are positively biased estimates of the actual volatility, but still contain information about future volatility values.

where $\varepsilon_t^* \sim$ N(0, 1). Volatility in this case follows a NGARCH(1, 1), which is the conventional GARCH(1, 1) process of Bollerslev (1986) adjusted by the unit risk premium λ. The leverage effect is represented by c, which has an effect similar to that of ρ in Eq. (2).[20] To streamline the presentation we will focus on a GARCH model without leverage effects, and set c equal to zero. However, we did consider both GARCH with and without leverage in our empirical work. Finally, for the GARCH model, Eqs. (31) and (32) suggest to filter volatility, denoted as σ_t^G, using the following recursive relationship

$$\sigma_t^G = \left(\hat{\alpha}_0 + \hat{\alpha}_1 \left(\log \frac{S_{t-1}}{S_{t-2}} - R + \tfrac{1}{2} \sigma^{G2}_{t-1} - \hat{\lambda} \sigma^G_{t-1} \right)^2 + \hat{\beta}_1 \sigma^{G2}_{t-1} \right)^{1/2}, \quad (33)$$

where the maximum likelihood estimates are: $\hat{\lambda} = 0.0599$,$\hat{\alpha}_0 = 0.556 \times 10^{-5}$, $\hat{\alpha}_1 = 0.127$, $\hat{\beta}_1 = 0.828$. As is typically the case for these models, the parameters are statistically significant and the models are not rejected by the data using standard diagnostics. The residuals and the squared residuals of the GARCH model are uncorrelated, and Portmanteau tests with up to 24 lags confirm the absence of residual autocorrelation. In particular, the p-values for residuals are equal to 0.212 (up to 6 lags), 0.574 (up to 12 lags), 0.743 (up to 18 lags), and 0.776 (up to 24 lags). The p-values for the squared residuals are equal to 0.928, 0.991, 0.998, and 0.999 respectively.

Fig. 4 presents the plots of the filtered volatilities. Panel (A) contrasts the BS implied volatilities and volatilities obtained from the GARCH model according to Eq. (33). Since the GARCH model is specified for daily observations, all the parameters pertain to the daily frequency, and the filtered volatility has daily units of measurement. To make the GARCH filtered volatility comparable to the BS volatilities and the SV-based volatilities, which have yearly units of measurement, we have multiplied the GARCH volatility by $\sqrt{250}$.[21] We see that the two filters of σ_t (σ_t^I and σ_t^G) provide very different forecasts. Panel (B) allows us to evaluate the volatility reprojected from the SV model estimated based on the returns data, σ_{rt}^R. In order to have a common benchmark of comparison with σ_t^G, we also plot σ_t^I in this panel. We note that σ_{rt}^R has a much smaller amplitude in comparison to both σ_t^I and σ_t^G. Finally, panel (B) of Fig. 4 also reports the reprojection filter from the SV model estimated with implied volatilities data, σ_{vt}^R. Although the filters in panel (B) are obtained from the same model specification with similar parameters estimates, σ_{rt}^R and σ_{vt}^R yield very different results, indicating that the key feature is not the model itself but the volatility filter it offers. This observation motivates us to consider all of the

[20] Note that, we have only one source of uncertainty in the GARCH model. Therefore, c is not a correlation coefficient.

[21] We know that GARCH models do not temporally aggregate. Therefore, this operation gives us only a rough estimate for the yearly data, which is sufficient for visual comparison.

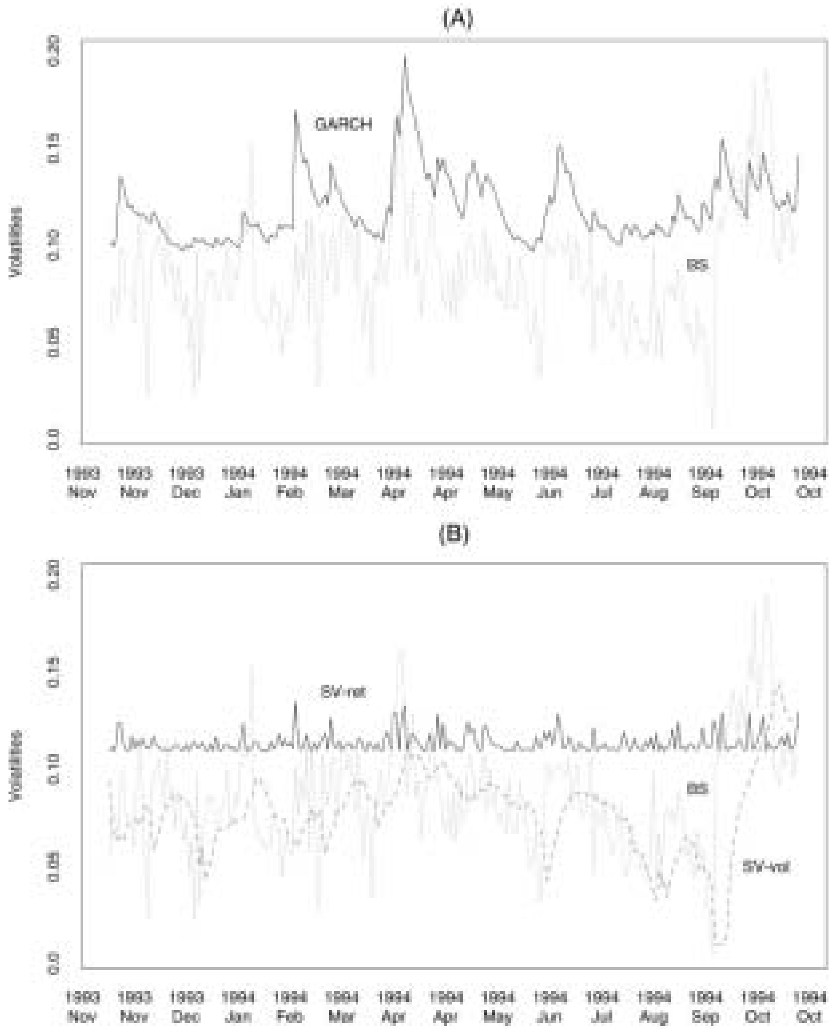

Fig. 4. The Filtered volatilities. We plot the volatilities $\hat{\sigma}_t$ which were estimated from different models. Panel (A) shows the volatility obtained from the GARCH model and the previous day Black–Scholes implied volatility. Panel (B) shows the volatilty reprojected from the SV model based on the returns series and on the implied volatilities series. It also shows the implied volatilities to have a common benchmark with the GARCH volatility plot in (A).

obtained filters in the framework of a single simple model, namely BS. The next section proceeds with the evaluation of the option pricing quality of the filters.

Table 7, Panel A, contains the pricing errors D_a^p and D_r^p. Obviously, GARCH and SV filters are suboptimal schemes here, as they are not used with their corresponding option pricing formula. We consider primarily the relative error term D_r^p, to appraise the differences, though Table 7 reports both types of errors. We observe that the GARCH filter σ_t^G is dominated by σ_t^I in every case except for the two long maturity ATM ($0.97 < S/K < 1.03$) groups. The best improvement of the GARCH filter over the implied volatilities is 1.79 dollars for the slightly OTM group, with relative error decreasing from 0.31 to 0.19. The best case of the implied volatility domination over GARCH is OTM medium maturity, where the price improvement is 95 cents with relative error decreasing from 19.62 to 12.35.

We also studied the performance of the GARCH model with leverage effects. The results are not reported in the table as they are qualitatively similar to the ones without leverage, though slightly and uniformly worse than the results obtained from GARCH without leverage. This observation is yet another example of conflicts between statistical and financial criteria. The leverage parameter in the GARCH model is statistically, though marginally, significant. Yet the resulting filter through the eyes of the Black–Scholes formula worsens the pricing performance. The same observation also applies to SV models, which are rejected on statistical grounds, while GARCH models are not rejected.

Table 7 also shows that the reprojection filter based on the model estimated with returns data σ_{rt}^R outperforms σ_t^I only for short maturity deep OTM and for long maturity OTM and ATM calls. The best improvement occurs again for the slightly OTM group, and is equal to 2.90 dollars with the relative error decreasing from 0.31 to 0.14. The best case where BS implied volatility dominates the returns reprojected volatility is the slightly OTM short maturity group with a 44 cents gain, for which the relative error changes from 12.29 to 5.78. The BS – σ_{rt}^R specification uniformly dominates BS – σ_t^G with the highest gain for the long maturity deep OTM calls equal to \$ 1.48, and the relative error dropping from 16.49 to 13.63. The lowest gain is 21 cents in the short maturity slightly ITM category. BS – σ_{vt}^R dominates BS – σ_t^I almost uniformly. The exceptions are slightly ITM short and medium maturities, and all long maturities except for deep OTM. However, in these cases, the differences displayed, are not statistically significant. The best improvement, which is equal to 55 cents, occurs in the medium maturity deep OTM group. These observations allow us to make some additional conclusions regarding the relative pricing performance of the four specifications. In particular, BS – σ_t^G, and therefore BS – σ_{rt}^R, clearly outperforms BS – σ_{vt}^R for long maturity ATM calls. BS – σ_{rt}^R also does better for the long maturity OTM group.

We can again conclude that the best volatility filter is σ_{vt}^R, with the exception of pricing long maturity and $0.94 < S/K < 1.03$ calls, for which σ_{rt}^R is the best. Both specifications are SV filters. We can provide several explanations for this result. The first is that long maturity contracts may reflect long memory, which is not taken into an account by the models we evaluate in this paper. One would have to consider the long memory models (for SV see Comte and Renault, 1995, and Harvey, 1993; for GARCH models see Baillie et al., 1995). Nevertheless, one could speculate that the SV filters considered here have greater flexibility and capture some of the long memory characteristics. The fact that σ_{rt}^R outperforms σ_{vt}^R at long maturities may be exploited by relying on Backus et al. (1997) and Das and Sundaram (1999), who show that in the SV model kurtosis of returns increases with the time interval, or time to maturity in the option pricing context. Therefore, the volatilities implied from the long maturity contracts have a more pronounced smile effect. This suggests that, for shorter maturities, volatilities implied from the ATM options, as used in our computation of σ_{vt}^R, are a good proxy for all BS implied volatilities. However, this result does not apply to longer maturities. The ATM implied volatilities may be a good reflection of the returns kurtosis for short maturities, but they are not adequate for long maturities. Hence we have to use the fourth power of returns to adequately reflect the kurtosis in pricing long term options. Our results support this theoretical conjecture, especially because the SNP density involved in the returns based reprojection required all the moments of returns up to the fourth moment. This finding also confirms the results reported in the previous section which showed that bivariate returns and options filters dominate σ_{vt}^R only at long maturities. This outcome is due to the kurtosis effect captured by the fourth moment of returns, and is not present in the ATM options at short maturities underlying σ_{vt}^R.

5.2. The contribution of the pricing formula

We consider now whether a model generating a particular volatility filter is important for option pricing. In particular, we address the question whether there is any pricing improvement when we use the original pricing formula, relying on SV and GARCH modeling, rather than the Black–Scholes pricing formula. Panel B of Table 7 reports the results. Surprisingly, rather than the BS one. Panel B of Table 7 reports the results. Surprisingly, we observe that GARCH option pricing errors are dramatically and uniformly dominated by their BS – σ_t^G counterparts. The GARCH call option value is determined as the expected value of normalized payoff at maturity under the local risk-neutral measure, and can be computed via Monte-Carlo simulations. For our purpose, we adopt paper the empirical martingale simulation strategy proposed by Duan and Simonato (1998). The GARCH with leverage model, not reported in the table, performs roughly the same way. Overall, we need to conclude that the GARCH pricing formula does not perform very well for pricing options. The σ_{rt}^R filter combined with the SV option pricing formulas shows a very different and opposite picture. The SV formula uniformly improves upon its BS counterpart. Contrary to the σ_{rt}^R case, the volatilities-based SV filter substituted into the

Table 7

Volatility filter and model contributions to option pricing

In this table, the role of volatility filter and particular model specification in option pricing is evaluated separately. In Panel A we consider the pricing performance of various volatility filters by substituting them into the Black–Scholes option valuation formula. In panel B we assess whether the pricing formula associated with a particular volatility filter improves option pricing as compared to the Black–Scholes formula reported in Panel A. We consider three models. The first is 'BS', which is the Black-Scholes model, the second is 'GARCH' model as specified in Duan (1995), and the third, 'SV', is the stochastic volatility model of Heston (1993). The volatility filters considered are: σ_t^I – the estimate of today's volatility by the previous day's 'BS' implied ATM option; σ_t^G – the estimate of instantaneous volatility using the 'GARCH' model; σ_{rt}^R – the estimate of instantaneous volatility via reprojection from 'SV' using a univariate scheme based on returns; σ_{vt}^R is the estimate of instantaneous volatility via reprojection from 'SV' using a univariate scheme based on BS implied volatilities. In Panels A and B, the first column is the moneyness category. The next column in Panel A lists the filters which are used in conjuction with the BS formula, while the second column in Panel B lists the filter/pricing formula combinations. The remaining columns are similar to the ones in Table 4. Namely, they report the absolute, D_a^p, and relative, D_r^p, errors by moneyness and maturity categories. GMM overidentifying restrictions tests, which determine whether the means of the relative pricing errors computed from a pair of models are significantly different, are reported next. In Panel A, pairwise tests within each moneyness-maturity group for all possible filters combinations (omitting repetitions) and overall groups tests are reported via the χ_1^2-statistics. The tests in Panel B compare the performance of the filter and pricing formula combinations and their BS formula counterparts from Panel A. The percentiles of the χ_1^2 distribution are: 90th – 2.71, 95th – 3.84, 99th – 6.63.

Panel A

Moneyness	$\hat{\sigma}_t$	Pricing errors Days to expiration						Pairwise tests Days to expiration									All groups		
		< 60		60–180		⩾ 180		< 60			60–180			⩾ 180			tests		
		D_a^p	D_r^p	D_a^p	D_r^p	D_a^p	D_r^p	σ_t^I	σ_t^G	σ_{rt}^R	σ_t^I	σ_t^G	σ_{rt}^R	σ_t^I	σ_t^G	σ_{rt}^R	σ_t^I	σ_t^G	σ_{rt}^R
< 0.94	σ_t^I	0.54	4.62	1.56	20.71	3.12	4.93												
	σ_t^G	0.67	6.71	2.35	33.15	3.77	16.49	3.6			10.2			4.8			42.8		
	σ_{rt}^R	0.44	3.42	1.51	22.89	2.29	13.63	2.3	5.0		0.5	12.6		4.5	3.2		27.2	18.9	
	σ_{vt}^R	0.42	2.03	1.01	12.61	3.03	3.65	3.5	4.5	2.2	5.6	18.6	13.4	0.4	4.4	3.9	17.5	53.6	51.1
0.94–0.97	σ_t^I	0.82	12.12	1.88	12.35	4.62	0.45												
	σ_t^G	1.39	21.57	2.83	19.62	4.09	0.45	30.4			9.3			0.0			42.8		
	σ_{rt}^R	1.08	18.34	2.13	18.93	2.93	0.34	22.8	14.6		7.6	0.4		11.9	8.8		27.2	18.9	
	σ_{vt}^R	0.58	8.17	1.57	7.27	4.65	0.46	11.7	37.9	38.1	4.5	15.4	15.8	0.1	0.0	11.0	17.5	53.6	51.1

0.97–1.00	σ_t^I	1.13	5.78	2.48	0.38	5.23	0.31												
	σ_t^G	1.84	12.78	2.66	0.53	3.44	0.19	22.0			20.3			18.0			42.8		
	σ_{rt}^R	1.57	12.29	2.19	0.51	2.33	0.14	18.4	1.7		8.1	1.3		28.2	10.0		27.2	18.9	
	σ_{vt}^R	1.10	4.37	2.60	0.35	5.79	0.33	6.0	22.1	18.5	2.9	20.4	9.0	0.6	19.4	26.3	17.5	53.6	51.1
1.00–1.03	σ_t^I	1.25	0.14	2.31	0.17	4.63	0.18												
	σ_t^G	1.56	0.23	2.29	0.18	2.87	0.11	29.2			1.8			11.0			42.8		
	σ_{rt}^R	1.35	0.21	1.95	0.16	2.18	0.09	19.2	9.8		0.7	8.12		16.8	2.7		27.2	18.9	
	σ_{vt}^R	1.39	0.15	2.61	0.18	4.72	0.18	0.7	22.5	13.2	1.1	0.2	2.2	0.0	11.3	16.8	17.5	53.6	51.1

Panel B

Moneyness	Model – $\hat{\sigma}_t$	Pricing errors Days to expiration						Tests: model vs. BS Days to expiration			All groups test
		< 60		60–180		⩾ 180		< 60	60–180	⩾ 180	
		D_a^p	D_r^p	D_a^p	D_r^p	D_a^p	D_r^p				
< 0.94	BS – σ_t^I	0.54	4.62	1.56	20.71	3.12	4.93				
	GARCH – σ_t^G	0.98	9.68	5.20	65.58	11.11	47.85	11.02	24.67	5.78	38.75
	SV– σ_{rt}^R	0.43	3.30	1.41	21.86	1.91	12.42	4.36	22.87	4.49	37.16
	SV – σ_{vt}^R	0.42	2.01	0.92	12.58	2.26	4.06	0.35	0.01	1.61	3.92
0.94–0.97	BS – σ_t^I	0.82	12.12	1.88	12.35	4.62	0.45				
	GARCH – σ_t^G	2.36	36.34	6.28	39.39	12.88	1.37	42.03	17.57	41.35	38.75
	SV– σ_{rt}^R	1.06	17.99	2.01	18.43	2.54	0.29	39.78	17.23	28.57	37.16
	SV – σ_{vt}^R	0.58	8.39	1.38	7.63	3.41	0.33	22.20	8.89	30.49	3.92
0.97–1.00	BS – σ_t^I	1.13	5.78	2.48	0.38	5.23	0.31				
	GARCH – σ_t^G	3.19	18.03	6.79	1.23	12.33	0.73	13.59	59.65	50.69	38.75
	SV– σ_{rt}^R	1.54	12.16	2.07	0.49	2.12	0.13	15.81	54.50	24.49	37.16
	SV – σ_{vt}^R	1.05	4.46	2.20	0.30	4.24	0.24	4.88	85.57	27.24	3.92
1.00–1.03	BS – σ_t^I	1.25	0.14	2.31	0.17	4.63	0.18				
	GARCH – σ_t^G	2.93	0.37	6.45	0.48	11.59	0.47	110.64	104.59	30.67	38.75
	SV–σ_{rt}^R	1.33	0.21	1.87	0.15	2.11	0.08	18.68	69.35	3.86	37.16
	SV–σ_{vt}^R	1.35	0.15	2.33	0.16	3.85	0.14	6.61	63.22	25.95	3.92

SV pricing formula shows a mixed picture. In particular, the BS – σ_{vt}^{R} combination is significantly better in three groups: OTM short maturity, OTM medium maturity, and slightly OTM short maturity. These findings prompt the question of whether the suboptimal BS – σ_t^G dominates GARCH option pricing, and why this result also happens in some of the BS – σ_{vt}^{R} cases. One possible interpretation is that the parsimony of BS comes into play. The Black–Scholes model appears to be so robust that, when used here with the GARCH volatility filter σ_t^G, for example, it smoothes out the GARCH deficiencies and prices options better.

The general conclusion we can draw from Table 7 is that the Heston SV model almost always provides both superior volatility filtering and option pricing. Within the Heston model, the filter specification based on implied volatilities is the best performer, with the exception of long maturity options, which are better priced with σ_{rt}^{R}. On the other hand, one can notice that the volatility filtering becomes more important than the option pricing formula in which the volatiltiy is subsequently substituted. For example, the differences in pricing errors between BS $-$ σ_{vt}^{R} and SV $-$ σ_{vt}^{R} are minimal if compared with pricing errors obtained from other volatility filtering schemes for short and medium maturity options.

6. Conclusion

We considered a generic procedure for estimating and pricing options in the context of stochastic volatility models using simultaneously the fundamental price, S_t, and a set of option contracts $[(\sigma_{it}^{I})_{i=1,K}]$ where $K \geqslant 1$ and σ_{it}^{I} is the Black–Scholes implied volatility. The joint estimation enabled us to recover objective and risk neutral measures simultaneously. This approach, bridging two strands of the literature, opens new possibilities for comparing the information in the underlying fundamental and options data, for conducting tests examining the transformation between the two measures, and investigating the general equilibrium underpinnings of models. Moreover, the bivariate approach involving both the underlying fundamental and an option appears useful for pricing derivatives when the information from the cash market, through the conditional kurtosis of returns, provides support to price options. This result holds for some long term options but the effect is only marginal. However, our results based on the S&P 500 index contract, showed that the univariate approach only involving options dominated for the purpose of pricing. A by-product of this finding is that we uncover a remarkably simple volatility extraction filter based on a polynomial lag structure of implied volatilities. Since the competing model specifications yielded roughly similar diffusion parameter estimates, we found that the filtering of spot volatility is the key ingredient in our procedure. In fact, by comparing various volatility filters, we found that the filter based on implied volatilities performs very well even if it is used in combination with the Black–Scholes formula.

At the general level, our results demonstrate how volatility filters obtained from the same model, and the same parameter estimates, can be very different because alternative filters exploit the information in the panel structure of options and returns differently. Our analysis also shows that financial criteria, like pricing and hedging performance, represent the most valuable selection methods. Models which are rejected using statistical goodness of fit criteria, in our case Heston's model, can be the basis for a relatively effective filter, when effectiveness is measured with respect to pricing. On the contrary, models that fit the data well, like GARCH, might turn out to be quite ineffective when it comes to valuing derivative securities.

Ultimately, financial and statistical criteria need to be reconciled. Heston's model has its obvious limitations, since it assumes that interest and dividend rates are constant. Several extensions of the Heston model have been suggested, including the addition of state variables which determine stochastic interest rates, dividends, and jump components. In particular, Bakshi and Madan (2000) and Duffie et al. (1998) discuss a general framework of jump-diffusions of the affine class which yields analytical option pricing formulas. The generic framework imposes a certain uniformity on the empirical studies in option pricing because the models are easier to compare, given that they are nested in the general model specification.

The methods proposed in this paper can be improved upon in several ways. First, we estimated the continuous time processes using only ATM contracts. It would be worth investigating a specification where more than one option is used, namely a set of option contracts $[(\sigma^I_{it})_{i=1,K}]$ where some are ATM and others are OTM. This setup will surely improve the pricing of long term OTM options. It would also be intriguing to find out how multivariate filters using several options contracts perform in comparison to the univariate option-based filter. Last but not least, it would be worth exploring the result suggesting that conditional kurtosis information in the cash market improves pricing. One could consider using LEAPS option contracts for such an exploration.

Appendix A. The call options pricing using SV model

We provide the details of the call options pricing formula under the SV model from Bakshi et al. (1997). The call price in Eq. (3) is expressed through Π_j, $j = 1, 2$, which are equal to:

$$\Pi_j = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\frac{\mathrm{e}^{-\mathrm{i}\phi \ln K} f_j(t,\ \tau, S(t), V(t);\ \phi)}{i\phi}\,\mathrm{d}\phi \tag{A.1}$$

and

$$
\begin{aligned}
&f_1(t,\ \tau, S(t), V(t);\ \phi)\\
&\quad = \exp\left\{\mathrm{i}\phi R\tau - \frac{\theta^*}{\sigma_V^2}\left[2\ln\left(1 - \frac{[\xi - \kappa^* + (\mathrm{i}\phi + 1)\rho\sigma_V](1 - \mathrm{e}^{-\xi\tau})}{2\xi}\right)\right]\right.\\
&\qquad - \frac{\theta^*}{\sigma_V^2}[\xi - \kappa^* + (\mathrm{i}\phi + 1)\rho\sigma_V]\tau + \mathrm{i}\phi \ln S(t)\\
&\qquad \left. + \frac{\mathrm{i}\phi(\mathrm{i}\phi + 1)(1 - \mathrm{e}^{-\xi\tau})V(t)}{2\xi - [\xi - \kappa^* + (1 + \mathrm{i}\phi)\rho\sigma_V](1 - \mathrm{e}^{-\xi\tau})}\right\},
\end{aligned} \tag{A.2}
$$

where

$$
\xi = \sqrt{[\kappa^* - (1 + \mathrm{i}\phi)\rho\sigma_V]^2 - \mathrm{i}\phi(\mathrm{i}\phi + 1)\sigma_V^2}, \tag{A.3}
$$

and

$$
\begin{aligned}
&f_2(t,\ \tau, S(t), V(t);\ \phi)\\
&\quad = \exp\left\{\mathrm{i}\phi R\tau - \frac{\theta^*}{\sigma_V^2}\left[2\ln\left(1 - \frac{[\psi - \kappa^* + \mathrm{i}\phi\rho\sigma_V](1 - \mathrm{e}^{-\psi\tau})}{2\psi}\right)\right]\right.\\
&\qquad - \frac{\theta^*}{\sigma_V^2}[\psi - \kappa^* + \mathrm{i}\phi\rho\sigma_V]\tau + \mathrm{i}\phi \ln S(t)\\
&\qquad \left. + \frac{\mathrm{i}\phi(\mathrm{i}\phi - 1)(1 - \mathrm{e}^{-\psi\tau})}{2\psi - [\psi - \kappa^* + \mathrm{i}\phi\rho\sigma_V](1 - \mathrm{e}^{-\psi\tau})}V(t)\right\},
\end{aligned} \tag{A.4}
$$

and

$$
\psi = \sqrt{[\kappa^* - \mathrm{i}\phi\rho\sigma_V]^2 - \mathrm{i}\phi(\mathrm{i}\phi - 1)\sigma_V^2} \tag{A.5}
$$

Appendix B. SNP density estimation

SNP is a method of nonparametric time series analysis which employs a Hermite polynomial series expansion to approximate the conditional density of a multivariate process. An appealing feature of this expansion is that it is a nonlinear nonparametric model that directly nests the Gaussian VAR model, the semiparametric VAR model, the Gaussian ARCH model, and the semiparametric ARCH model. The SNP model is fitted using conventional maximum likelihood methods, together with a model selection strategy that determines the appropriate degree of the polynomial.

The SNP method is based on the notion that a Hermite expansion can be used as a general purpose approximation to a density function. Here, it is employed in the form $f_k(z_t \mid X_t, \Xi) \propto [P_K(z_t, X_t)^2]\phi(z_t)$, where $P_K(\cdot)$ denotes a multivariate polynomial of degree K_z and $\phi(z)$ denotes the standard normal, possibly multivariate, probability density function (p.d.f.). The process X_t is the vector of M lags of the process of interest. The index k denotes the dimension of Ξ, which expands as the sample size increases (see discussion in Section 2.1). Since $f_k(z)$, as the distribution proxy, has to integrate to 1, the constant of proportionality is the inverse of $N_K(X_t) = \int [P_K(s, X_t)^2]\phi(s)\,\mathrm{d}s$. To achieve a unique representation, the constant term of the polynomial part is set equal to one.

First of all, let us note that if the process of interest $\imath \sim N(\mu, \Sigma = RR')$, where R is an upper triangular matrix, then $\imath = Rz + \mu$ and

$$f_k(\imath \mid X_t, \Xi) = \frac{1}{N_K(X_t)} \frac{[P_K(z_t, X_t)^2]\phi(z_t)}{|det(R)|}, \tag{B.1}$$

where $\Xi = (a_{ij}, \mu, R)$ and is estimated by QML.

The mean μ at time t depends on M_μ lags of $\imath$: $\mu_t^X = b_0 + BX_t$. R_t^X also depends on M_R lags of the process of interest, when an ARCH leading term is present. A Hermite polynomial is used to expand the density around the leading Gaussian density, namely

$$P_K(z_t, X_t) = \sum_{i=0}^{K_z} a_i(X_t) z^i = \sum_{i=0}^{K_z} \left(\sum_{j=0}^{K_X} a_{ij} X_t^j \right) z^i. \tag{B.2}$$

As discussed above, we set a_{00} equal to 1 for the uniqueness of the representation. The subscript K denotes the vector (K_X, K_z, M). According to Andersen and Lund (1996), if the leading term is specified carefully, $M = 1$ generally suffices in the univariate SNP setting.

When K_z is set equal to zero, one obtains normal distributions. When K_z is positive and $K_X = 0$, one obtains a density which is conditionally homogeneous with respect to X. Finally, when both K_z and K_X are positive, we obtain conditionally heterogeneous classes of densities.

The optimal values of K, (as well as M_μ and M_R), are chosen based on standard criterions and tests, such as AIC, BIC and HQ model selection criteria. Gallant and Tauchen (1993) provide a table which matches values of the parameters K, M_μ, M_R, to popular time series models. We reproduce this table here (Table 8). Typically, BIC is used to select the model specification. We follow this strategy in Table 1, Panels A and C.

Appendix C. Details of the SDE discretization scheme

The discussion in this appendix is based on Kloeden and Platen (1995), to which we refer the reader for additional details.

The EMM estimation procedure requires the process under study to be simulated. It is impossible to simulate the actual process if it does not have an analytical solution, which is the case for Heston's model. Therefore, to simulate the process at hand, we need to discretize the process specified in Eqs. (1) and (2). The discretizations are based on the truncation of the Ito–Taylor expansion formula. The more terms we use, the more accurate is the expansion. The Ito–Taylor formula also involves the derivatives of the drift and diffusion coefficients with respect to the process of interest. Hence, an approximation of these derivatives may be advisable to improve computational efficiency. The discretization schemes mainly differ in the number of terms that one keeps and the way one approximates, or fails to approximate, the above-mentioned derivatives. For the following reasons, we choose to work with the explicit order 2.0 weak discretization scheme.

Each strong scheme of order p is a weak scheme of order $2p$. Schemes are classified as weak if they approximate moments rather than sample paths of a diffusion. Strong schemes, on the ohterh hand, approximate sample paths. Moreover, strong schemes require more time to simulate, and since we match moments instead of sample paths in our estimation, we use a weak scheme. The order of the scheme determines the quality of an approximation. We say that a discrete approximation $Y(\Delta t) = (S_n, V_n)'$ of the process $X = (S(t), V(t))'$ converges weakly with order $p > 0$ to X at time T as $\Delta t \to 0$ if, for each moment M, there exists a positive constant C which does not depend on Δt, and a finite $\Delta_0 t$ such that

$$|M(X_T) - M(Y_T(\Delta t))| \leqslant C\Delta t^p \tag{C.1}$$

for any $\Delta t \in (0, \Delta_0 t)$. In our case $p = 2$, which yields a faster convergence than the commonly used Euler discretization, a weak scheme of order 0.5. Note that, following the tradition in the option pricing literature, we estimate the annualized parmeters, taking $\Delta t = 1/252$. Kloeden and Platen show that, under standard regularity conditions, the moments of the diffusions exist provided the initial value is a constant (see Kloeden and Platen, 1995, Theorem 4.5.4). The discretization scheme guarantees that the moments of the discretized processes also exist. Therefore, it is legitimate to implement the moments matching procedure.

The term "explicit" in the discretization scheme refers to the derivatives approximation method. In case of deterministic differential equations, this method works similar to the Runge–Kutta methods. Since this method imposes fewer computational burdens than the implicit scheme, it is often preferable to use.

This approach also allows us to control for the non-synchronicity bias discussed in the data section. We can show that by matching moments instead of sample paths, the error introduced by the differential in stock prices over the fifteen minute closing interval has a negligible impact on the estimation of the moments. We simulate the fundamental process as being observed at 4:00 pm, or time T, and we use these simulated values to substitute in the implied volatility function, $I(S_t, V_t)$, implicitly based on the option price, which is observed at 4:15 pm, or time $T + \delta t$. We will use the Ito-Taylor expansion formula to compute the error introduced by this operation. The EMM procedure matches the moments of the function $I(\cdot)$. Therefore, we have to assess the size of the error for all possible moments $E(I^n(\cdot))$. Since it is a sufficiently smooth function, we can approximate it with polynomials based on $(S_t, V_t)'$ to any degree of accuracy. Hence, in our discussion, we will consider $E(f(S_t, V_t))$ for any smooth function f. In particular, according to Theorem 5.5.1 of Kloeden and Platen (1995), we can write:

$$\begin{aligned}
f(S_{T+\delta t}, V_{T+\delta t}) = {} & f(S_T, V_T) + f_{(0)}(S_T, V_T)\int_T^{T+\delta t} \mathrm{d}u \\
& + f_{(1)}(S_T, V_T)\int_T^{T+\delta t} \mathrm{d}W_S(u) \\
& + f_{(2)}(S_T, V_T)\int_T^{T+\delta t} \mathrm{d}W_V(u) \\
& + f_{(1,1)}(S_T, V_T)\int_T^{T+\delta t}\int_T^{v} \mathrm{d}W_S(u)\,\mathrm{d}W_S(v) \\
& + f_{(2,2)}(S_T, V_T)\int_T^{T+\delta t}\int_T^{v} \mathrm{d}W_V(u)\,\mathrm{d}W_V(v) \\
& + f_{(1,2)}(S_T, V_T)\int_T^{T+\delta t}\int_T^{v} \mathrm{d}W_S(u)\,\mathrm{d}W_V(v) \\
& + f_{(2,1)}(S_T, V_T)\int_T^{T+\delta t}\int_T^{v} \mathrm{d}W_S(u)\,\mathrm{d}W_V(v) + \mathrm{o}(\delta t), \qquad \text{(C.2)}
\end{aligned}$$

where $f_{(\cdot)}$ are the Ito coefficient functions (Kloeden and Platen, 1995, 5.3). As we already mentioned, we are interested in the accuracy of the moments. Therefore, we have to compute $E(f(S_{T+\delta t}, V_{T+\delta t}))$:

$$\begin{aligned}
& E(f(S_{T+\delta t}, V_{T+\delta t})) \\
& \quad = E(f(S_T, V_T)) + E(f_{(0)}(S_T, V_T))\delta t + \mathrm{o}(\delta t) \\
& \quad = E(f(S_T, V_T)) + E\bigg(\frac{\partial}{\partial t} + \mu_S S\frac{\partial}{\partial S} + (\theta - \kappa V)\frac{\partial}{\partial V} \\
& \qquad + \tfrac{1}{2}VS^2\frac{\partial^2}{\partial S^2} + \sigma_V \rho V S\frac{\partial^2}{\partial S \partial V} + \tfrac{1}{2}\sigma_V^2 V\frac{\partial^2}{\partial V}\bigg)f(S_T, V_T)\delta t + \mathrm{o}(\delta t) \\
& \quad = E(f(S_T, V_T)) + \mathrm{o}(\delta t). \qquad \text{(C.3)}
\end{aligned}$$

The last equality follows from the fact that the expected value of the infinitesimal generator $f_{(0)}(S_T, V_T)$ of the process specified in Eqs. (1) and (2) is equal to zero (see, for instance, Hansen and Scheinkman, 1995). Since the interval δt is relatively small, at 1/96th of the simulation interval Δt, substituting $f(S_{T+\delta t}, V_{T+\delta t})$ by $f(S_T, V_T)$ in the implied volatility moments matching procedure introduces only a negligable error.

References

Aït-Sahalia, Y., Lo, A., 1998. Nonparametric estimation of state-price densities implicit in financial prices. Journal of Finance 53, 499–548.

Ait-Sahalia, Y., Wang, Y., Yared, F., 1997. Do options markets correctly price the probabilities of movements of the underlying asset? Unpublished discussion paper. University of Chicago, IL.

Andersen, T.G., Chung, H.-J., Sørensen, B.E., 1999. Efficient method of moments estimation of a stochastic volatility model: a Monte Carlo study. Journal of Econometrics 91, 61–87.

Andersen, T.G., Lund, J., 1996. The short rate diffusion revisited: an investigation guided by the efficient method of moments. Unpublished discussion paper, Northwestern University, Evanston, IL.

Andersen, T.G., Lund, J., 1997. Estimating continuous-time stochastic volatility models of the short-term interest rate. Journal of Econometrics 77, 343–377.

Backus, D., Foresi, S., Li, K., Wu, L., 1997. Accounting for biases in Black–Scholes. Unpublished discussion paper, New York University.

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

Bakshi, G., Madan, D., 2000. Spanning and derivative-security valuation. Journal of Financial Economics 55.

Bates, D.S., 1996. Testing option pricing models. In: Maddala, G.S., Rao, C.R. (Eds.),, Handbook of Statistics, Vol. 14. Elsevier, Amsterdam.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics 31, 307–327.

Broadie, M., Detemple, J., Ghysels, E., Tòrres, O., 1997. American options with stochastic volatility and dividends: a nonparametric approach. Journal of Econometrics, forthcoming.

Campbell, J., Lo, A., MacKinlay, C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton, NJ.

Canina, L., Figlewski, S., 1993. The informational content of implied volatility. Review of Financial Studies 6, 659–681.

Comte, F., Renault, E., 1995. Long memory continuous time stochastic volatility models. Paper presented at the HFDF-I Conference, Zurich.

Cox, J., Ingersoll, J., Ross, S., 1985. A theory of the term structure of interest rates. Econometrica 53, 385–408.

Christensen, B.J., Prabhala, N.R., 1998. The relation between implied and realized volatility. Journal of Financial Economics 50, 125–150.

Das, S., Sundaram, R., 1999. Of smiles and smirks: a term-structure perspective. Journal of Financial and Quantitative Analysis 34, 211–240.

Day, T., Lewis, C., 1992. Stock market volatility and the information content of stock index options. Journal of Econometrics 52, 267–287.

Duan, J.C., 1995. The GARCH option pricing model. Mathematical Finance 5, 13–32.

Duan, J.C., Simonato, J.G., 1998. Empirical martingale simulation for asset prices. Management Science 44, 1218–1233.
Duffie, D., Pan, J., Singleton, K., 1998. Transform analysis and option pricing for affine jump-diffusions. Econometrica, forthcoming.
Dumas, B., Fleming, J., Whaley, R.E., 1998. Implied volatility functions: empirical tests. Journal of Finance 53, 2059–2106.
Figlewski, S., 1998. Derivatives risks, old and new. Wharton-Brookings Papers on Financial Services I.
Fleming, J., 1998. The quality of market volatility forecasts implied by S&P 100 index option prices. Journal of Empirical Finance 5, 317–345.
Gallant, A.R., Hsieh, D., Tauchen, G., 1997. Estimation of stochastic volatility models with diagnostics. Journal of Econometrics 81, 159–192.
Gallant, A.R., Hsu, C.-T., Tauchen, G., 1999. Using daily range data to calibrate volatility diffusions and extract the forward integrated variance, Review of Economics and Statistics, forthcoming.
Gallant, A.R., Long, J.R., 1997. Estimating stochastic differential equations efficiently by minimum chi-square. Biometrika 84, 125–141.
Gallant, A.R., Tauchen, G., 1989. Seminonparametric estimation of conditionally constrained heterogeneous processes: asset pricing applications. Econometrica 57, 1091–1120.
Gallant, A.R., Tauchen, G., 1993. SNP: A program for nonparametric time series analysis. Version 8.3, User's Guide. University of North Carolina, Chapel Hill, unpublished discussion paper.
Gallant, A.R., Tauchen, G., 1996. Which moments to match?. Econometric Theory 12, 657–681.
Gallant, A.R., Tauchen, G., 1997. EMM: A program for Efficient Method of Moments estimation. Version 1.4, User's guide. Unpublished discussion paper. Duke University, Durham, NC.
Gallant, A.R., Tauchen, G., 1998. Reprojecting partially observed systems with application to interest rate diffusions. Journal of American Statistical Association 93, 10–24.
Ghysels, E., Harvey, A., Renault, E., 1996. Stochastic volatility. In: Maddala, G.S., Rao, C.R. (Eds.), Handbook of Statistics, Vol. 14. Elsevier, Amsterdam.
Ghysels, E., Jasiak, J., 1996. Stochastic volatility and time deformation. Unpublished discussion paper, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), Montreal.
Hansen, L., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.
Hansen, L., Scheinkman, J., 1995. Back to the future: generating moment implications for continuous-time Markov processes. Econometrica 63, 767–804.
Harrison, M., Kreps, D., 1979. Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory 20, 381–408.
Harrison, M., Pliska, S., 1981. Martingales and stochastic integrals in the theory of continuous trading. Stochastic Processes and Their Applications 11, 215–260.
Harvey, A.C., 1993. Long memory in stochastic volatility. Unpublished discussion paper. London School of Economics.
Harvey, A.C., Ruiz, E., Shephard, N., 1994. Multivariate stochastic variance models. Review of Economic Studies 61, 247–264.
Harvey, C.R., Whaley, R.E., 1991. S&P 100 index option volatility. Journal of Finance 46, 1551–1561.
Heston, S.L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies 6, 327–343.
Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatilities. Journal of Finance 42, 281–300.
Jacquier, E., Jarrow, R., 1998. Model error in contingent claim models: dynamic evaluation. Journal of Econometrics forthcoming.
Jacquier, E., Polson, N.G., Rossi, P.E., 1994. Bayesian analysis of stochastic volatility models (with discussion). Journal of Business and Economic Statistics 12, 371–417.

Jacquier, E., Polson, N.G., Rossi, P.E., 1995. Stochastic volatility: univariate and multivariate extensions. Unpublished discussion paper, Boston College.

Jiang, G., van der Sluis, P., 1998. Pricing stock options under stochastic volatility and interest rates with efficient method of moments estimation. Unpublished discussion paper, University of Groningen, Netherlands.

Kloeden, P.E., Platen, E., 1995. Numerical Solution of Stochastic Differential Equations. Springer, New York, NY.

Lamoureux, C.G., Lastrapes, W.D., 1993. Forecasting stock-return variance: toward an understanding of stochastic implied volatilities. Review of Financial Studies 6, 293–326.

Melino, A., 1994. Estimation of continuous-time models in finance. In: Sims, C.A. (Ed.), Advances in Econometrics. Cambridge University Press, Cambridge.

Nandi, S., 1998. How important is the correlation between returns and volatility in a stochastic volatility model? Emprical evidence from pricing and hedging in the S&P 500 index options market. Journal of Banking and Finance 22, 589–610.

Nelson, D.B., Foster, D.P., 1994. Asymptotic filtering theory for univariate ARCH models. Econometrica 62, 1–41.

Nelson, D.B., 1996. Asymptotic filtering theory for multivariate ARCH models. Journal of Econometrics 62, 1–41.

Pastorello, S., Renault, E., Touzi, N., 1994. Statistical inference for random variance option pricing. Jornal of Business and Economic Statistics, forthcoming.

Renault, E., 1997. Econometric models of option pricing errors. In: Kreps, D.M., Wallis, K.F. (Eds.), Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress, Vol. 3. Cambridge University Press, Cambridge.

Romano, M., Touzi, N., 1997. Contingent claims and market completeness in a stochastic volatility model. Mathematical Finance 7, 399–412.

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

Rydberg, T.H., 1997. A note on the existence of unique equivalent martingale measures in a Markovian setting. Finance and Stochastics 1, 251–257.

Scott, L., 1987. Option pricing when the variance changes randomly: theory, estimation, and an application. Journal of Financial and Quantitative Analysis 22, 419–438.

Tauchen, G., 1997. New minimum chi-square methods in empirical finance. In: Kreps, D.M., Wallis, K.F. (Eds.), Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress, Vol. 3. Cambridge University Press, Cambridge.

Wiggins, J.B., 1987. Option values under stochastic volatility: theory and empirical estimates. Journal of Financial Economics 19, 351–372.