

CIRANO

Centre interuniversitaire de recherche
en analyse des organisations

Série Scientifique
Scientific Series

98s-22

**What Data Should Be Used
to Price Options?**

Mikhail Chernov, Eric Ghysels

Montréal
Juin 1998

CIRANO

Le CIRANO est une corporation privée à but non lucratif constituée en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du ministère de l'Industrie, du Commerce, de la Science et de la Technologie, de même que des subventions et mandats obtenus par ses équipes de recherche. La *Série Scientifique* est la réalisation d'une des missions que s'est données le CIRANO, soit de développer l'analyse scientifique des organisations et des comportements stratégiques.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de l'Industrie, du Commerce, de la Science et de la Technologie, and grants and research mandates obtained by its research teams. The Scientific Series fulfils one of the missions of CIRANO: to develop the scientific analysis of organizations and strategic behaviour.

Les organisations-partenaires / The Partner Organizations

- École des Hautes Études Commerciales
- École Polytechnique
- McGill University
- Université de Montréal
- Université du Québec à Montréal
- Université Laval
- MEQ
- MICST
- Banque Nationale du Canada
- Bell Québec
- Fédération des caisses populaires Desjardins de Montréal et de l'Ouest-du-Québec
- Hydro-Québec
- Industrie Canada
- Microcell Labs inc.
- Raymond, Chabot, Grant, Thornton
- Scetauroute
- Société d'électrolyse et de chimie Alcan Ltée
- Téléglobe Canada
- Ville de Montréal

Ce document est publié dans l'intention de rendre accessibles les résultats préliminaires de la recherche effectuée au CIRANO, afin de susciter des échanges et des suggestions. Les idées et les opinions émises sont sous l'unique responsabilité des auteurs, et ne représentent pas nécessairement les positions du CIRANO ou de ses partenaires.

This paper presents preliminary research carried out at CIRANO and aims to encourage discussion and comment. The observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of CIRANO or its partners.

ISSN 1198-8177

What Data Should Be Used to Price Options?*

Mikhail Chernov[†], Eric Ghysels[‡]

Résumé / Abstract

Nous présentons une procédure générique pour l'estimation et l'évaluation de modèles d'options avec volatilité stochastique où le sous-jacent et un ensemble de contrats d'options sont utilisés simultanément. Nos résultats démontrent qu'un modèle univarié avec seulement des données d'options domine en terme d'erreurs de prix hors-échantillon et en terme de couverture. Nous trouvons également un filtre d'extraction pour la volatilité latente qui est basé sur un polynôme de retards de volatilités implicites. Ayant simultanément la probabilité de risque neutre et la probabilité objective, nous pouvons vérifier, dans le contexte du modèle de Heston, si la transformation usuelle est empiriquement plausible. Nous rejetons le changement de mesure supposé dans ce modèle.

In this paper we propose a generic procedure for estimating and pricing options in the context of stochastic volatility models using simultaneously the fundamental price and a set of option contracts. We 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 that the univariate approach only involving options by and large dominates. A by-product of this finding is that we uncover a remarkably simple volatility extraction filter based on a polygonal lag structure of implied volatilities. The bivariate approach involving both the fundamental and an option appears useful when the information from the cash market provides support via the conditional kurtosis to price options. This is the case for some long-term options. Moreover, having estimated separately the risk-neutral and objective measures allows us to appraise the typical risk-neutral representations used in the literature. Using Heston's (1993) model as example we show that the usual transformation from objective to risk neutral density is not supported by the data.

Mots Clés : Titres dérivés, méthode de moments efficaces, prix d'états, filtrage, volatilité stochastique

Keywords : Derivative securities, efficient method of moments, state price densities, stochastic volatility models, filtering

JEL : G13, C14, C52, C53

* Corresponding Author: Eric Ghysels, Pennsylvania State University, 523 Kern Graduate Building, University Park, PA 16802-3305 USA tel: (814) 865-7924 fax: (814) 863-4775 e-mail: eghysels@psu.edu We would like to thank Charles Cao, Jérôme Detemple, and Eric Jacquier for invaluable comments and especially Ron Gallant and George Tauchen for insightful discussions.

[†] Pennsylvania State University

[‡] Pennsylvania State University and CIRANO

Introduction

Numerous techniques have been proposed for estimating continuous time processes and the pricing of derivative securities. They involve the estimation of diffusions with or without stochastic volatility and/or jumps. The literature is summarized in a number of surveys and textbooks, including Bates (1996), Campbell *et al.* (1996), Ghysels *et al.* (1996), Melino (1994), Renault (1997) and Tauchen (1997). Option pricing formula are subsequently used to price the derivative contracts written on the underlying fundamental. Hence the time series of the fundamental is used to estimate the diffusion in a first stage and a pricing model associated with the diffusion is used in a second to compute the option price. This can involve rather straightforward computations, as in the Black-Scholes options pricing model, or can require sophisticated estimation procedures when for instance diffusions with stochastic volatility are involved. Another approach tends to by-pass the first stage and computes the price of a new contract by extrapolating from a cross-section of listed options. Therefore, using only the cross-section of options results in ignoring any information in the fundamental that may not be present in the option prices.

Financial theory suggests that for stochastic volatility models with two state variables (such as the models of Hull and White (1987), Scott (1987), Heston (1993) and many others) one should consider the 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 we need to add for instance options to be able to complete the market.¹ The complete market setup 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 it is options or fundamental. The task is indeed challenging and besides knowing how to proceed with it we don't know a priori what gains can be made in terms of better parameter estimates for diffusions and in terms of pricing options and hedging performances. Although to the best of our knowledge no attempts were made 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

¹See for instance, Romano and Touzi (1993) who show that the combination of the fundamental and a European-type option contract complete the market in a stochastic volatility type economy.

addressing this issue.² 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 $[(\sigma_{it}^I)_{i=1,K}]$ where $K \geq 1$ and σ_{it}^I is the Black-Scholes implied volatility. Please note that we can in principle deal with a panel of options, i.e. a time series of cross-sections. The procedure we propose consists of two steps, first we fit the density (using so called SNP methods) of $[S_t, (\sigma_{it}^I)_{i=1,K}]$ conditional on its own past $[S_\tau, (\sigma_{i\tau}^I)_{i=1,K}]$ 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 S_t . We extend it to handle option prices and fundamentals simultaneously. We estimate the stochastic volatility model due to Heston (1993). 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 implied volatilities and the underlying asset data jointly. Indeed, times series of the underlying asset provide estimators under the objective probability measure while from the options we can retrieve risk neutral parameters. Since the model we adopt has a closed-form option pricing formula, we can obtain the BS implied volatilities from the simulated data and contrast them with their counterparts from the real data via the EMM framework. This leads to the parameter estimates under the risk-neutral measure.

We compare univariate and multivariate models in terms of pricing and hedging performance. The univariate specifications consist of models only using the fundamental (i.e. the usual setup) and models using only options data. It should be noted, however, that the knowledge of the estimated model parameters is not enough to compute an option price or a hedge ratio. We have to know the present day spot volatility. Previous studies treated the spot volatility as a parameter and estimated it from the previous day cross-section of options prices. This approach introduces inconsistencies with the model. Recent extension of the SNP/EMM methodology in Gallant and Tauchen (1998) allows us to address the problem. We find the spot volatilities via reprojection, i.e. we compute the expected value of the latent volatility process using the SNP density conditioned on the observable processes such as returns and/or options data.

²A 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.

Our results show that the univariate approach only involving options by and large dominates. 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 via the conditional kurtosis to price options. This is the case for some long term options. Moreover, having estimated *separately* the risk-neutral and objective measures 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. We introduce a test which is based on the Radon-Nikodym derivative of the objective measure with respect to the risk neutral one and uses the Novikov condition.

The remainder of the paper is organized as follows. Section 1 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 will estimate. Section 2 provides a brief review of the EMM estimation and reprojection method. Section 3 reports the estimation results. Section 4 evaluates the performance of the estimated models. The last section concludes. Technical material is covered in several Appendices to the paper.

1 Options and Fundamentals

It has long been recognized that there is no reason to focus only on either the cash or the options market prices separately. Numerous papers have confronted empirical evidence obtained from option markets with that of the cash market and vice versa. Questions about the informational content of the options prices, for instance, have often been raised in the literature (see Bates (1996), Canina and Figlewski (1993), Day and Lewis (1992), Fleming (1994), Lamoureux and Lastrapes (1993), among others). Aït-Sahalia *et al.* (1997) address essentially the same issue comparing state-price densities (SPD) implied by times series of the S&P 500 index and the SPD implied by a cross-section of SPX index options. They reject the hypothesis of the equality of the two SPD's. As they examine models with a deterministic volatility function of the stock price, one can view their evidence as a rejection of models with volatility as a deterministic function of the stock price. They attempt to reconcile the conflicting state-price densities by introducing jumps into the underlying asset model. Along the same lines, Dumas *et al.* (1997)

examine the predictive power and hedging performance of the deterministic volatility function model on the example of the S&P 500 options and find that "simpler is better", i.e. the model performs no better than ordinary implied volatility. Furthermore, Bakshi *et al.* (1997b), using the same index, show that any one-dimensional Markov model of underlying asset prices may lead to severe options mispricing. It is also common to argue that if stochastic volatility models better explain the behavior of underlying assets, they should also be able to reproduce the main stylized facts of implied volatility such as the volatility smile (the dependence of implied volatilities on an option moneyness) and the volatility term structure (the dependence on time and time to maturity).

The practice of estimating diffusions using the underlying asset and then relying on option pricing formula has a number of drawbacks. Standard complete market asset pricing theory determines that one has to change measure, from the objective to the risk neutral.³ This transformation is often somewhat ad hoc. Although continuous time general equilibrium preference-based asset pricing models readily yield the mapping from the objective to risk neutral measure, they often result in rather complex diffusion models for the underlying asset. This is because the equilibrium asset price process is derived endogenously, based on the discounted flow of dividends using an endogenously determined risk-free rate.⁴ It is therefore common to use a simple diffusion for the asset return and for the 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. In the case of the Heston (1993) model, which will be examined in greater detail later in the section, this means that the drift change for the volatility process is an affine function of volatility.⁵ Moreover, arguments about completeness of markets are typically imposed as an auxiliary assumption to guarantee the existence of a unique risk-free measure.

Our attempt to model the fundamental and derivative securities jointly is motivated by the very same issues that have been raised in the literature so far. We want to learn about the informational content of options prices, like Canina and Figlewski (1993) and many others did. We want to know how we can improve the *statistical* precision of diffusion parameters by incorporating options. In a similar vein, Pastorello

³See Harrison and Kreps (1979) and Harrison and Pliska (1981) for further discussion.

⁴See for instance, Broadie *et al.* (1997) for such a derivation.

⁵Namely, both under the objective and risk-neutral measures the drift in volatility is linear. Any nonlinear functional drift transformation would yield a risk-neutral process not within the same class as the objective measure process.

et al. (1994) used time series of at-the-money (henceforth ATM) implied volatilities to estimate the Hull and White (1987) model.⁶ 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.* (1997a), Dumas *et al.* (1997) and Jacquier and Jarrow (1998), among others. We attempt to investigate these questions in a unifying framework. We use the SV model due to Heston (1993) for that purpose, though it should be stressed at the outset that our analysis is not limited to this particular model. The choice of Heston’s model was motivated by two important factors: (i) it has closed-form option pricing formula (Bakshi *et al.* (1997a) provide results more general than those provided by Heston), and (ii) because of the existence of analytic solutions it has received considerable attention in the literature (again see Bakshi *et al.* (1997a) and references therein). The latter makes our analysis directly comparable with results previously reported in the literature. Following Heston (1993) and Bakshi *et al.* (1997a) we can write the model as:

$$\frac{dS(t)}{S(t)} = Rdt + \sqrt{V(t)}dW_S^*(t) \quad (1.1)$$

$$dV(t) = (\theta^* - \kappa^*V(t))dt + \sigma_V \sqrt{1 - \rho^2} \sqrt{V(t)}dW_V^*(t) + \sigma_V \rho \sqrt{V(t)}dW_S^*(t), \quad (1.2)$$

where the model is stated under the risk-free probability measure. Equation (1.1) implies that the stock-price process $S(t)$ follows a geometric Brownian motion with stochastic variance $V(t)$. The second equation (1.2) 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 . The Brownian motions $W_S^*(t)$ and $W_V^*(t)$ are assumed independent. Equations (1.1) and (1.2) imply, however, that $Corr_t(dS(t)/S(t), dV(t)) = \rho dt$. Parameters with asterisks are those which change when we make a transition to the objective probability measure. With such a change the risk-free rate R would have to be substituted by μ_S and all asterisks removed. This model yields the following formula for a price of a call at time t , time to maturity τ and strike K :

$$C(t, \tau, K) = S(t)\Pi_1(t, \tau, S, V) - Ke^{-R\tau}\Pi_2(t, \tau, S, V), \quad (1.3)$$

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

⁶Pastorello *et al.* (1994) did not estimate the joint process as we propose to do in this paper. In addition, besides joint estimation we also cover filtering of the latent volatility process from options data.

Several methods have been proposed to estimate the parameters of stochastic volatility processes. These methods range from 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). The EMM method proposed by Gallant and Tauchen (1996) involves the simulation of discretization of (1.1) and (1.2). The usual way one proceeds is to take a time series sample of the underlying asset. This sample is summarized via a set of judiciously chosen sample moments. These moments are matched by moments obtained from simulating discretization of (1.1) and (1.2) under the objective probability measure (hence all the asterisks from parameters are removed). In the next section we will be more precise about the details of the procedure. The innovation in applying this procedure is the simulation of asset returns *and* option prices, or Black-Scholes implied volatilities, and match not only the moments of sample asset returns but also those of derivative securities data.

To conclude, it should be noted that we impose certain restrictions on the model parameters. These restrictions follow from the theory and previous empirical work. The first restriction follows from the requirement that the volatility can never take negative values. Since we have a square-root model, we can use the results of Cox *et al.* (1985), i.e. if $\sigma_V^2 \leq 2\theta$, then $V(t)$ in (1.2) does not reach 0 with probability 1. Obviously, $\sigma_V \geq 0$. Finally, the empirical literature, starting with Black (1976), finds negative correlation between volatility and the underlying asset (leverage effect). Hence we also imposed the restriction: $\rho \leq 0$.

2 EMM Estimation of the Joint Fundamental and Option Pricing Process

The Efficient Method of Moment procedure of Gallant and Tauchen (1996) has 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.* (1997), Gallant *et al.* (1997, 1998), Gallant and Tauchen (1998), Ghysels and Jasiak (1996) and Tauchen (1997). So far these applications only involve a single data series, either a short rate process or a stock price (index).⁷ Since the EMM procedure is widely used and described elsewhere, notably in Gal-

⁷There are exceptions, notably Ghysels and Jasiak (1996) who consider the joint process of stock returns and trading volume and Gallant *et al.* (1998), who as noted before, consider the bivariate process of the fundamental and the daily high/low spread.

lant and Tauchen (1998), we will only summarize the major steps here. EMM can be divided into two main parts. The first one is the estimation of the so called score generator. The second part yields the parameter estimates using simulations and GMM approach involving the score generator. A first subsection is devoted to the score generator. A second discusses the second stage, which is the moment matching principle. The third subsection deals with reprojection methods, an extension of EMM to extract latent volatility processes. A fourth and final subsection provides the data description.

2.1 The Score Generator

Suppose the process of interest is denoted ι_t . In our application this process can be univariate, only involving asset returns or BS implied volatilities, it can be bivariate involving both returns and implied volatilities, or in general a panel data series of returns and K option contracts with different moneyness and/or maturities represented by their BS implied volatilities, namely of $[(\sigma_{it}^I)_{i=1,K}]$. In a generic context we assume that ι_t is a vector with L elements. It has a conditional distribution $p(\iota_t | I_t, \Theta)$, where I_t is the information set and Θ are the parameters in the stochastic volatility model of ι_t . The asymptotically efficient method to estimate Θ is the maximum likelihood, which maximizes the function:

$$\frac{1}{T} \sum_{t=1}^T \log p(\iota_t | I_t, \Theta) \quad (2.1)$$

Maximization of (2.1) is equivalent to solving

$$\frac{1}{T} \sum_{t=1}^T \frac{\partial}{\partial \Theta} \log p(\iota_t | I_t, \hat{\Theta}) = 0 \quad (2.2)$$

which is the sample equivalent of:

$$E \left(\frac{\partial}{\partial \Theta} \log p(\iota_t | I_t, \hat{\Theta}) \right) = 0, \quad (2.3)$$

Unfortunately, it is very difficult to obtain the likelihood function for stochastic volatility models and hence it is impossible to compute the score generator $p(\cdot | \cdot)$. Gallant and Tauchen suggest to compute instead a SNP density $f_k(\iota_t | X_t, \Xi)$, 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.

2.2 Matching the Moments

The estimated SNP density provides the underpinnings for constructing a set of moment conditions, which in turn set the stage for a GMM-type estimation procedure for the parameter vector Θ . In particular, Θ can be estimated through the moment conditions (score function) similar to (2.3), which in this case will be:

$$m(\Theta, \hat{\Xi}) = E \left(\frac{\partial}{\partial \Xi} \log f_k(\iota_t | X_t, \hat{\Xi}) \right) = \int \frac{\partial}{\partial \Xi} \log f_k(\iota_t | X_t, \hat{\Xi}) dP(\iota, X, \Theta). \quad (2.4)$$

We compute sample moments by simulating N observations of ι_t from the SV model. In particular, in the previous section we noted that in a generic context we assume that ι_t is a vector with L elements. It has a distribution $p(\iota_t | I_t, \Theta)$, where I_t is the information set and Θ are the parameters in the stochastic volatility model of ι_t . We will focus primarily on the case where ι_t represents a bivariate process of a stock return and a BS implied volatility.⁸ The distribution $p(\iota_t | I_t, \Theta)$ in this context is implicitly defined by equations (1.1), (1.2) and (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 all the parameters common to the objective and risk neutral measure, whereas Θ_o represents the objective probability measure volatility drift parameters and Θ_n the risk neutral ones. Hence, in equations (1.1) and (1.2) these 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 (1.1) and (1.2) under the objective probability measure for a given set of parameters values for $(\mu_S, \Theta_c, \Theta_o)$. Please note that we simulate the latent volatility process $V(t)$, though it is not part of ι_t . Obviously we need to use $V(t)$ to compute $S(t)$ and therefore we simulate it. Then, 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 (1.3) and calculate the BS implied volatilities from these prices. It is important to note here that we do not impose any particular mapping between $(\Theta_o$ and $\Theta_n)$. This will allow us to test certain hypothesis about the transformation from objective to risk neutral

⁸We collected BS implied volatilities, as discussed in greater detail in the next section, and therefore will simulate those rather than call prices.

measures. We use the explicit order 2.0 weak SDE discretization scheme to simulate the processes (1.1) and (1.2) (Kloeden and Platen (1995) pp. 486-487). This scheme has a faster convergence than the Euler discretization and was also used in Gallant and Tauchen (1997). To obtain the BS volatilities, we need to apply the option pricing formula to the simulated data. This requires not only the parameters, but also 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. Since the sample size of the simulated data is much larger (in order to decrease the simulation error), we cycle through the sequence of observed maturities in the actual data to cover the entire simulated dataset. In particular, for the simulated i^{th} observation we use the maturity from the $\text{mod}(i, T)$ 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 l times each maturity appearing in the observed data. Moreover, we simulate moneyness rather than with strikes, because the simulated sample path of $S(t)$ can be quite different from the observed one. Matching the moneyness with the real data implies strikes not always observed in the real data sample. This strategy preserves, however, the crucial properties of options. Since we can rewrite the option pricing formula in terms of moneyness, the dependence on strike will be eliminated. We rotate moneyness from the observed data in exactly the same way as maturities to simulate BS implied volatilities.

Suppose we have simulated time series of length N for ι_t with candidate parameters Θ .⁹ Hence, the left hand side of (2.2) translates into

$$m_N(\Theta, \hat{\Xi}) = \frac{1}{N} \sum_{t=1}^N \frac{\partial}{\partial \Xi} \log f_k(\iota_t(\Theta) | X_t(\Theta), \hat{\Xi}). \quad (2.5)$$

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

$$\hat{\Theta} = \underset{\Theta}{\text{arg min}} m_N(\Theta, \hat{\Xi})' W_T m_N(\Theta, \hat{\Xi}). \quad (2.6)$$

Because of the properties of the SNP model,

$$W_T = \frac{1}{T} \sum_{t=1}^T \frac{\partial \log f_k(\iota_t | X_t, \hat{\Xi})}{\partial \Xi} \frac{\partial \log f_k(\iota_t | X_t, \hat{\Xi})}{\partial \Xi'}. \quad (2.7)$$

⁹The empirical results reported in the next section are based on $N = 10,000$.

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

Finally, it should be noted that the simulation approach described in this section also applies to situations only involving options data in which case the parameter vector consists of $(\Theta_c, \Theta_o, \Theta_n)$, or the more traditional case where stock returns are used in ι_t to estimate $(\mu_S, \Theta_c, \Theta_o)$. Hence, only cases with both fundamentals and derivatives will involve the full parameter vector $(\mu_S, \Theta_c, \Theta_o, \Theta_n)$.

2.3 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 (1.3). Several filters have been proposed in the literature, all involving exclusively the return process. Harvey *et al.* (1994) suggested to make use of the approximate Kalman filter based on a discrete time SV model. The exact filter was derived by Jacquier *et al.* (1994) in the context of a Bayesian analysis of the same model. Nelson and Foster (1994) showed how diffusion limit arguments applied to the class of EGARCH models provided a justification of EGARCH models as filters of the 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 extension all involve exclusively multivariate return series and cannot accommodate derivative security market information. Via a combination of the reprojection method and the joint density of returns and BS implied volatilities we rely on Gallant and Tauchen (1998) and propose a filtering method which in its generic form can be multivariate, (i) only involving a vector of returns (multiple series, something not considered in this paper but feasible), (ii) only involving a vector of options (more on this later), (iii) a mixture of the previous two. It is the latter strategy which will be adopted in this paper. Please note 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 aforementioned filtering methods of Harvey *et al.* (1994), Jacquier *et al.* (1994) and Nelson and Foster (1994). Univariate schemes only involving options have been proposed mostly informally and often, though not exclusively, rely on the Black-Scholes model. The key usually is that the options cross-section is used and today's volatility is treated as a parameter. The volatility is then backed

out from the cross-section via the minimization of the pricing error. Such a method appears in Bakshi *et al.* (1997a) for instance. It should be noted, however, that the reprojection approach applied to a vector of options, i.e. the aforementioned reprojection scheme (ii) is more general as it truly takes advantage of the panel structure (i.e. cross-section of time series). Even the univariate reprojection method only involving, say 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 not much exploited so far (and as we will see in the next section, remarkably good).

In the remainder of this section we will focus exclusively on the specific applications of filtering via reprojection which will be used in the empirical analysis of the paper. The specific applications can readily be extended, however, to the aforementioned generic specifications. The two most novel applications are the focus of our analysis here, one deals with a bivariate model of returns and an ATM option and the other deals with a univariate filter based on options only.¹⁰ The reprojection scheme relies on a one-step ahead forecast, which is an expectation with respect to the distribution of volatility conditional on the contemporaneous and lagged returns, denoted r_t , and lagged implied volatilities $p(V(t)|r_t, \sigma_{t-1}^I, r_{t-1}, \dots, \sigma_{t-M}^I, r_{t-M}, \hat{\Theta})$. The computation of the SNP density of $V(t)$ proceeds in several steps. First, using the estimates $\hat{\Theta}$ one simulates the processes (1.1) and (1.2) which produces the series $\{\hat{\sigma}_{t-1}^I, \hat{r}_t, \hat{V}_t\}_{t=1}^N$ in case of bivariate data and $\{\hat{\sigma}_{t-1}^I, \hat{V}_t\}_{t=1}^N$ in the univariate case.¹¹ Since k grows with the sample size N we select a simulation size of 10,000 observations. Denote the vector $(\hat{r}_t, \hat{\sigma}_{t-1}^I, \hat{r}_{t-1}, \dots, \hat{\sigma}_{t-M}^I, \hat{r}_{t-M})'$ (or $(\hat{\sigma}_{t-1}^I, \dots, \hat{\sigma}_{t-M}^I)'$ in the univariate case) by \hat{X}_t , using the notation from the previous section. Second, since the SNP density has a Gaussian lead term we need to transform volatility, a process which only takes positive values. While we need to filter volatility to plug in the SV call pricing formula we need to take a functional transformation to fit the Gaussian lead term SNP expansion. We consider a first order Taylor expansion of the logarithm of volatility denoted $L(V(t))$ and find the SNP density $f_k(L(\hat{V}_t)|\hat{X}_t)$.¹² The first order Taylor

¹⁰In a companion paper we discuss in greater detail filtering procedures comparable to the existing methods and based exclusively on return series. See Chernov and Ghysels (1998) for further details.

¹¹To streamline the notation we use \hat{V}_t for the simulated $V(t)$.

¹²Fitting SNP densities in a reprojection exercise requires a slightly different setup 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 variable would not be included in the conditioning set as in the standard SNP density estimation.

expansion of the logarithm of volatility facilitates recovery of \hat{V}_t . Indeed, Jensen's inequality yields that $E(C(V(t))) \leq C(E(V(t)))$, for any concave function C . This inequality becomes an equality only if C is linear, hence the use of $L(\cdot)$. The SNP density fitted to the simulated data therefore produces a mean of \hat{V}_t conditional on the observable vector \hat{X}_{t-1} , which yields the desired filtered values. The first order Taylor expansion can be taken around the mean of simulated \hat{V}_t 's. This approximation may be quite inaccurate in the tails, however. This is particularly important as the inverse transform may easily result in negative volatilities. This happens because the linear transformation imposes lower and upper limits on the range of a volatility. We therefore implement a procedure similar to a spline transformation.¹³

2.4 The data

The data we use consists of daily prices of options written on the S&P 500 index.¹⁴ The sample covers the time period from November 1985 until October 1994. We set aside the last year of data (i.e. 11:93 to 10:94) for the out-of-sample tests, and use the rest for estimation purpose. We concentrate our attention on the at-the-money calls. European ATM options with short maturities are the most liquid instruments. Hence, their implied volatilities should convey the most precise information. We define at-the-money as $S/K \in [0.97, 1.03]$. Several other data filters were applied to the data. For instance, observations with a call price missing or equal to zero were obviously dropped. In order to be able to easily compare the observed and simulated data, we adjusted the S&P 500 index for dividends. The continuously compounded rate of 2% was used for simplicity (this is consistent with Broadie *et al.* (1997)). We used the monthly 3-month T-bill yield from CITIBASE as a proxy for the short-term interest rate. Since stochastic volatility models assume a constant interest rate, we take the mean from these yields series, which is equal to 5.81396%.

The BS implied volatilities were computed using the Newton-Raphson method. Since in some situations (such as near to maturity out-of-the-money calls), implied volatilities can take very small, even negative,

¹³In particular, we take the following breakpoints 16th, 32nd, 68th and 84th percentiles and perform the first order Taylor expansions around the 8th, 24th, 50th, 76th and 92nd percentiles correspondingly. Our experiments showed that having no breakpoints resulted in 32% of negative volatilities, 2 breakpoints in 3.6% and 4 in just 1.6%. The remaining few negative values were substituted by 0.0001.

¹⁴Specifically, the dataset contains date, maturity month, option price, call/put flag, strike, open interest, volume and, finally, the underlying index.

values, we introduce a cut-off equal to .001. This filter removes only 4.84% of the data and, hence, seems to be reasonable.

As noted before we will estimate SV models, using the 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, (iii) both. To obtain the series of type (ii) 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:

$$K^* = \arg \min_K \left| \frac{S}{K} - 1 \right|.$$

This last filter leaves us with 1978 observations, which roughly corresponds to 247 observations per year, i.e. we have a week per year of missing observations.

Finally, the estimation of SNP density requires the use of stationary data. 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.¹⁵ However, we will refer to these data series as S&P500 and BS volatilities for convenience.

3 Empirical results

Following the setup of the EMM procedure we discuss first the estimation results of the SNP densities, which appear in a first subsection. Then the next subsection provides details of the model parameter estimates. The final subsection is devoted to testing the risk neutral measure transformation used in Heston's model.

3.1 The Empirical Score Generators

The SNP density estimation results are reported in Panel A of Table 1. 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 . To facilitate the interpretation of the results we supplement Panel A of Table 1 with a second Panel B which describes the generic features of SNP densities for different combinations of the tuning parameters. Comparing Panel A with the panel B, we note that all series are

¹⁵Contrary to the discussion of the previous section we do not need to worry here about the inverse transformation as we do not plug these volatilities in option pricing formula. Hence there is no need here to engage in the Taylor expansion approximations which make the inverse logarithmic operation simpler.

non-Gaussian ARCH-type with homogeneous innovations. It is worthwhile to examine the estimated density plots. Figure 1 reports plots of both raw data and estimated SNP densities for the univariate cases. Figure 2 does the same for the bivariate case involving returns and BS implied volatilities. Plot (a) in Figure 1 displays the S&P 500 series and shows the familiar pattern, including the negative return corresponding to the crash of October 1987. Plot (c) displays the corresponding BS implied volatilities. The highest volatility is also, not surprisingly, observed at the time of the crash. Plots (b) and (d) show the corresponding SNP densities with a standard normal p.d.f. superimposed (represented by the dashed line). The estimated densities show the familiar patterns, namely they are peaked, leptokurtic and slightly skewed.

Figure 2 reports the joint density in (a) and the contour plot in (b). The marginal densities appear in (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 and the restriction $\rho \leq 0$ we imposed.

3.2 The SV model estimation

We turn now to parameter estimates of Heston's SV model. Traditionally this model is estimated with returns only, so we report this configuration as a benchmark. Then we proceed to the estimation results exclusively relying 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. Table 2 reports all the estimation results. The first observation is that Heston's SV model is mostly rejected no matter what the data confirmation is (all z -statistics reported in the brackets are large). The rejections vary dramatically though across the different types of data. The returns data allow to approximate the process reasonably well, 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 example, Gallant *et al.* (1997)). Of course our major interest is not to appraise the model exclusively on the grounds of its statistical properties. Its pricing and hedging features, discussed in the next section, will be of prime interest.

Using the BS volatilities data improves the precision of the estimated parameters (the standard errors are of order 10^{-2} or less). The returns data do not allow us to identify the values of θ^* , κ^* of the drift in (1.2) under the risk-neutral measure. Hence, there are only two parameters in common between the first and second panel of rows in Table 2. They are the leverage effect ρ and the variance of the volatility equation σ_V . In

both cases we obtain roughly the same point estimates using returns and BS implied volatilities, yet the precision of the estimates are dramatically improved with the second data set. This is not so surprising as all estimated parameters are related to the volatility process (1.2). Therefore, looking at the process through the observed implied volatilities should give us more precision, an observation also made by Pastorello *et al.* (1994). The z -statistic is huge (247.447), however, indicating serious inconsistencies between the model and the data. In particular, it means that our model does not explain very well the information extracted from the options prices.

Last but not least in Table 2 is the model estimated with the joint data. Here, we can fit both the objective and risk-neutral density parameters and therefore make direct comparisons with the previous two univariate setups. The bivariate model shows a great improvement in fit 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 all across the different configurations, while the point estimates roughly remain the same.

Since we are interested what type of data we should use to price and hedge options, we can conclude from Table 2, that the returns series only should not be used as we cannot 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: (i) BS implied volatilities, and (ii) joint returns and implied volatilities series. We will therefore focus exclusively on these two to appraise how options are priced and how well they perform for the purpose of hedging.

3.3 Examining the Affinity Restriction

Heston (1993) states his model under the objective probability measure. Assuming that the price of risk is proportional to the volatility, he makes a transition to the risk-neutral measure. Hence, under the Heston's setup $\theta = \theta^*$. Bakshi *et al.* (1997a) start out immediately in the risk-neutral world and, therefore, their results are slightly more general as they do not impose $\theta = \theta^*$. However, as mentioned earlier, both setups make a very strong assumption that the probability measure transformation preserves the linearity of the volatility drift. The bivariate setup yielded estimates of the entire parameter vector $\Theta = (\mu_S, \sigma_V, \rho, \theta, \kappa, \theta^*, \kappa^*)$, which allows us to examine whether the change of measure restrictions imposed in Heston's model are supported by the data. We propose two tests, one relatively easy to implement and a second which is computationally de-

manding. The former is based on the Novikov condition and is presented first.

The Radon-Nikodym derivative of the objective probability measure with respect to the risk-neutral one is computed as follows:

$$\begin{aligned} \xi_{t,\tau} = & \exp \left(-\frac{1}{2} \int_t^{t+\tau} (\lambda_1^2(u) + \lambda_2^2(u)) du \right. \\ & \left. - \int_t^{t+\tau} \lambda_1(u) dW_S(u) - \int_t^{t+\tau} \lambda_2(u) dW_V(u) \right), \end{aligned} \quad (3.1)$$

where $\lambda(t) = (\lambda_1(t), \lambda_2(t))'$ is the market price of risk. Since we know the parameter values under both measures we can infer $\lambda(t)$. By Girsanov's theorem we have:

$$\begin{aligned} \mu_S S(t) - \lambda_1(t) \sqrt{V(t)} S(t) &= RS(t), \quad (3.2) \\ \theta - \kappa V(t) - \lambda_1(t) \sigma_V \sqrt{1 - \rho^2} \sqrt{V(t)} - \lambda_2(t) \sigma_V \rho \sqrt{V(t)} &= \theta^* - \kappa^* V(t). \end{aligned} \quad (3.3)$$

Therefore,

$$\lambda_1(t) = \frac{\mu_S - R}{\sqrt{V(t)}}, \quad (3.4)$$

$$\lambda_2(t) = \frac{C_1}{\sqrt{V(t)}} - C_2 \sqrt{V(t)}, \quad (3.5)$$

where

$$C_1 = \frac{\theta - \theta^* - (\mu_S - R) \sigma_V \sqrt{1 - \rho^2}}{\sigma_V \rho}, \quad (3.6)$$

$$C_2 = \frac{\kappa - \kappa^*}{\sigma_V \rho}. \quad (3.7)$$

The aforementioned computations of the Radon-Nikodym derivative are legitimate provided that the Novikov condition is satisfied. This condition amounts to:

$$E_N(0, T) = E \left(\exp \left(\frac{1}{2} \int_0^T (\lambda_1^2(u) + \lambda_2^2(u)) du \right) \right) < \infty \quad (3.8)$$

One can derive a lower bound for E_N (see Appendix C), namely:

$$E_N(t, t + \tau) \geq \exp \left(\frac{1}{2} \left[\frac{\kappa}{\theta} ((\mu_S - R)^2 + C_1^2) + \frac{\theta}{\kappa} C_2^2 - 2C_1 C_2 \right] \tau \right)$$

$$\begin{aligned}
& + \frac{1}{2} \left[\frac{(\mu_S - R)^2 + C_1^2}{\theta} - C_2^2 \frac{\theta}{\kappa^2} \right] \log \left| 1 + \frac{\tau}{t} \right| \quad (3.9) \\
& \approx \exp \left(44.76\tau + 317.22 \log \left| 1 + \frac{\tau}{t} \right| \right),
\end{aligned}$$

where the estimated values of Θ and the formulas appearing in (3.6) and (3.7) were used to obtain the last expression. It then follows that as t approaches 0, E_N diverges to infinity. Hence the Novikov condition is violated and the Radon-Nikodym derivative does not exist for the mapping assumed in the Heston model. This implies that the data does not support the transformation from the objective to the risk neutral measure, which preserves the affine volatility drift in Heston's model. Of course, one could interpret the failure of the Novikov condition to hold as another indication that the entire model is rejected which is something we already reported in Table 2. However, the use of this test may not be redundant whenever a model is not rejected via z-statistics and one wants to examine the presumed change of measure.

The Novikov condition test shows that the Radon-Nikodym derivative does not exist. However, if it exists we may still not find the volatility drift assumption acceptable. Therefore the use of another test may be warranted. Unfortunately, such a test involves rather tedious computations and hence should be used only if the above approach is not feasible. In the remaining of this section we briefly elaborate on the construction of such a test, which like the Novikov condition evaluation, is to the best of our knowledge novel to the literature.

Breeden and Litzenberger (1978) show that the underlying asset state-price density (SPD) is a second derivative of a call option price with respect to strike evaluated at the assets price and normalized to integrate to one. This approach yields a formula for the fundamental SPD at time t and time to maturity τ as a function of the risk-neutral parameters only $\text{SPD}_1^S(S(t+\tau), t, \tau)$. Ait-Sahalia and Lo (1998) derive an expression for the SPD of $r_{t+\tau}$, namely:

$$\text{SPD}_1^r(r_{t+\tau}, t, \tau) = S(t)e^{r_{t+\tau}} \text{SPD}_1^S(S(t+\tau)e^{r_{t+\tau}}, t, \tau) \quad (3.10)$$

Since, we have the parameters estimated under both measures, we can derive the SPD from the pairs $\mu_S - R$, $\theta - \theta^*$, and $\kappa - \kappa^*$. In particular, one can compute the SPD in the following way:

$$\text{SPD}_2^r(r_{t+\tau}, t, \tau) = \xi_{t,\tau} \cdot p(r_{t+\tau}), \quad (3.11)$$

where $p(\cdot)$ is the unconditional density of the log-returns under the objective measure and $\xi_{t,\tau}$ is given in (3.1).

To implement the test one has to examine whether two SPD's are identical. Appraising the similarity of two probability densities is rather non-trivial, however. A simple strategy, for instance would consist of comparing moments and consider GMM overidentifying restrictions tests built on moment conditions derived from the alternative characterizations of the SPD's.¹⁶

4 Assessing Pricing and Hedging Performance

Statistical criteria for model selection are one thing, financial criteria such as pricing and hedging performance out-of-sample are obviously the ultimate scope of model selection for practitioners. Indeed, the application of complicated econometric techniques is warranted only if they do better than the models currently in use. Therefore, we investigate the performance of the alternative model specifications obtained from the previous section and make comparisons with the traditional Black-Scholes model estimated from daily options data. All the evaluations reported in this section are out-of-sample. Namely the estimation sample used in the preceding section covered November 1985 to October 1993, whereas the sample used to appraise all the models runs from November 1993 to October 1994.¹⁷ Hence, since we use out-of-sample data, our results will not be contaminated by in-sample data mining. A first subsection is devoted to pricing of options. A second looks at hedging and last but not least we devote a separate section to a remarkable simple volatility filter for option pricing which emerges from our analysis. The examination of this filter is indeed important as all the 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 the filtering.

4.1 Pricing reliability

The model parameter estimates, the reprojection filters to extract the latent volatility process and Heston's European call pricing formula give us all the ingredients to price any SPX contract with a particular time to maturity, strike and given cash price for the underlying index. Equipped with these tools we can compare the call prices predicted by the models

¹⁶This would be similar to forecast evaluation techniques based on moment comparisons, see e.g. Diebold and Mariano (1995).

¹⁷One could call this a genuine out-of-sample approach, as in Dumas *et al.* (1997) or Jacquier and Jarrow (1998), in comparison to some performance evaluations which use in-sample one step ahead forecasts (see for instance Bakshi *et al.* (1997a)).

with observed prices. In our appraisal we do not consider in-the-money calls with moneyness greater than 1.03, since there is a very thin market for such options. Furthermore, we separate the calls into twelve groups by 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_a^p , and has a dollar scale value. The second is a relative measure, denoted D_r^p , and reflects percentage deviations. The two measures are defined as follows:

$$D_a^p = \sqrt{\frac{1}{\sum_{t=1}^T n_t} \sum_{t=1}^T \sum_{i=1}^{n_t} (C_{it}^{observed} - C_{it}^{model})^2}, \quad (4.1)$$

$$D_r^p = \sqrt{\frac{1}{\sum_{t=1}^T n_t} \sum_{t=1}^T \sum_{i=1}^{n_t} \left(\frac{C_{it}^{observed} - C_{it}^{model}}{C_{it}^{observed}} \right)^2}. \quad (4.2)$$

The results, are divided in four different categories of moneyness, namely deep out-of-the-money (OTM with S/K less than 0.94), OTM (with $.94 < S/K < .97$), slightly OTM (with $.97 < S/K < 1$) and slightly in-the-money (with $1 < S/K < 1.03$). The last two categories are usually viewed as ATM options. Three maturity horizons are considered for all the moneyness categories, 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 out-of-the-money options we have roughly between 313 and 448 contracts to compute pricing errors and slightly less 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 smallest is 90. Table 4 contains the pricing errors D_a^p and D_r^p computed for three different model specifications. The first is “*BS*”, which involves pricing options with the Black-Scholes model and a volatility estimate based on the previous day ATM option contract (see for instance Bates (1996) for further discussion on this approach). 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 model. Table 4 reports besides the absolute and relative pricing errors

¹⁸The computation of hedging errors will be discussed in the next subsection. Suffice here to say that they involve less observations as contracts are required to be traded at least two consecutive days.

for each of the twelve moneyness/maturity groups also statistics which test whether there are any statistically significant differences between the pricing errors computed from *BS*, *Vol* and *Joint*. We observe that in general pricing errors can be large with dollar values (i.e. D_a^p) ranging from 5.23 dollars to 42 cents. Not surprisingly, large errors occur for long maturities. The relative errors (i.e. D_r^p) 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 based on the following set of moment conditions, namely:

$$\begin{pmatrix} (\sum_{t=1}^T n_t)^{-1} \sum_{t=1}^T \sum_{i=1}^{n_t} \left(\frac{(C_{it}^{observed} - C_{it}^{vol.model})}{C_{it}^{observed}} \right)^2 - m \\ (\sum_{t=1}^T n_t)^{-1} \sum_{t=1}^T \sum_{i=1}^{n_t} \left(\frac{(C_{it}^{observed} - C_{it}^{joint.model})}{C_{it}^{observed}} \right)^2 - m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (4.3)$$

where m is the common mean relative error under the null.¹⁹ The errors form a panel data set, because we have a time series of cross-sectional observations since within each cell a total of n_t contracts are traded. The sample data are therefore correlated which precludes us from taking a simple t -statistic. Instead, we use a GMM-based procedure involving a Newey and West estimator for the covariance matrix. The overidentifying restrictions test statistic based on the moment conditions in (4.3) is distributed χ^2 with one degree of freedom under the null hypothesis (see Hansen (1982)). The test statistics are reported in the left panel of Table 4.

The most striking result which emerges from the table is dominance of the *Vol* approach. Hence, the estimation of SV models *only* involving ATM options outperforms *BS* and the bivariate approach denoted *Joint*. Besides yielding statistically significant superior pricing performance, we observe from the results reported in Table 4 that the improvements over *BS* range from a 1.21 dollar reduction (the relative error dropping from 0.45 with *BS* to 0.33) for OTM long maturity, to a small eight cents reduction with relative error dropping from 5.78 to 4.46 for slightly OTM short maturity. The *Vol* specification outperforms the *Joint* with the largest improvement for the deep OTM long maturity category, though this time the improvement is only 48 cents. The smallest one is also in the same moneyness category but has short maturity and is equal to

¹⁹Computations with the absolute errors could be performed as well but are not reported here.

only 2 cents. It is worth re-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 difference 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* by the *Vol* model specification is somewhat surprising we conducted some further tests which are not reported in the table. These additional tests required a departure from the reprojection procedure described in section 2.3. It is not clear from the fitted SNP density in the reprojection routine whether certain individual moments of returns may provide information that is enhancing 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.²⁰ This specific filter, therefore investigates how much of the leverage effect in the stock index which is not already incorporated in the implied volatilities improves 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. Indeed, we found that the conditional kurtosis, i.e. 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 using Heston's call price model with a relatively simple volatility filtering scheme using past options data yields the most desirable outcome across all the three model specifications considered. By and large, all the models perform relatively well at short maturities, of course. The discrepancies really play out 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 incorporating in the estimation sample contracts similar in nature to the ones priced out-of-sample would certainly improve the record across the *Vol* and *Joint* specifications. We will further elaborate on this in the concluding section and delegate this to future research.

²⁰Hence, we ignore the Hermite polynomial expansion terms in the SNP density and isolate the Gaussian kernel.

4.2 Hedging performance

It is common to hedge the market risk via a combination of stocks and options. There are numerous hedging techniques available in theory, some are quite sophisticated and involve several instruments simultaneously. However, since options are in zero net supply (i.e. there always has to be a party holding an opposite position in an option contract), many of the sophisticated hedging strategies are hard to implement in practice (see Figlewski (1998) for further discussion). We will therefore concentrate on a simple minimum variance hedging which uses just one option. In particular, if we want to hedge our position in one call with $X_S(t)$ shares of stock, we choose the number of shares in such a way, that the remaining cash position

$$X_0(t) = C(t, \tau, K) - X_S(t)S(t) \quad (4.4)$$

has minimum variance. This is achieved by taking

$$X_S(t) = \Delta_S(t, \tau, K) + \frac{\rho\sigma_V}{S(t)}\Delta_V(t, \tau, K), \quad (4.5)$$

where

$$\Delta_S(t, \tau, K) = \frac{\partial C(t, \tau, K)}{\partial S} = \Pi_1 + S(t)\frac{\partial \Pi_1}{\partial S} - Ke^{-R\tau}\frac{\partial \Pi_2}{\partial S} \quad (4.6)$$

and

$$\Delta_V(t, \tau, K) = \frac{\partial C(t, \tau, K)}{\partial V} = S(t)\frac{\partial \Pi_1}{\partial V} - Ke^{-R\tau}\frac{\partial \Pi_2}{\partial V}, \quad (4.7)$$

Since the SV model does not account for all sources of risk (for example, the interest rate is assumed to be constant), we will not end up with zero, if we try to unwind our position the next day. The hedging error will therefore be:

$$H(t + \Delta t) = X_S(t)S(t + \Delta t) + X_0(t)e^{R\Delta t} - C(t + \Delta t, \tau - \Delta t, K), \quad (4.8)$$

where Δt is equal to one day.

We use summary statistics similar to those in the previous subsection to report the model hedging performance. In particular, we construct absolute (D_a^h) defined as:

$$D_a^h = \sqrt{\frac{1}{\sum_{t=1}^T n_t} \sum_{t=1}^T \sum_{i=1}^{n_t} H_i^2(t)}. \quad (4.9)$$

and comparable relative (D_r^h) measures. To obtain the latter, let us rewrite (4.8) as:

$$H(t + \Delta t) = X_0(t)e^{R\Delta t} - \tilde{X}_0(t + \Delta t) = \text{Cash}_{t+\Delta t}^{model} - \text{Cash}_{t+\Delta t}^{observed} \quad (4.10)$$

where \tilde{X}_0 is different from X_0 because the number of shares was computed in the previous period, not as defined in (4.4). This yields the relative measure of the hedging error which is defined as:

$$\begin{aligned} D_r^h &= \sqrt{\frac{1}{\sum_{t=1}^T n_t} \sum_{t=1}^T \sum_{i=1}^{n_t} \left(\frac{\text{Cash}_{it}^{observed} - \text{Cash}_{it}^{model}}{\text{Cash}_{it}^{observed}} \right)^2} \\ &= \sqrt{\frac{1}{\sum_{t=1}^T n_t} \sum_{t=1}^T \sum_{i=1}^{n_t} \left(\frac{H_i(t)}{\tilde{X}_{i0}(t)} \right)^2} \quad (4.11) \end{aligned}$$

The hedging performance comparisons for the *BS*, *Vol* and *Joint* are reported in Table 5, again using the classification of twelve moneyness/time-to-maturity cells. The absolute and relative errors are complemented with statistical tests built on moment conditions similar to (4.3) which yield an overidentifying restrictions test. 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. Therefore the hedging strategies, unlike the pricing errors, appear to be rather insensitive to model specification.

4.3 Filtering latent spot volatility with options data

Since the univariate *Vol* specification involving options is the most successful and novel, some further digression on its practical implementation is warranted. The most innovative aspect of the pricing procedure is the construction of the volatility extraction filter. The filter is remarkably simple. Constructing a reprojection filter involves a model selection procedure for the tuning parameters K_z , K_X , M , M_μ and M_R of the SNP density (see in particular Panel B of Table 1 for their interpretation). These tuning parameters are selected via the BIC criterion. The optimal model specification which minimizes the BIC criterion is a simple twenty two lags linear operator. Hence the ATM implied volatilities are combined in a weighted historical moving average with a 22 day window. Table 6 lists the filter weights which extract spot volatility from *BS* implieds.²¹ The weights range from 0.24514 to 0.03277. The intercept is

²¹The reprojection software we used does not produce standard errors for the coefficients, hence only point estimates are reported.

negative and equal to -0.16320 . The largest weights in the filter appear at the longer lags and the increase in weights is roughly linear in the lags. Hence, the extraction scheme puts more weight on observations from the options markets which 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), involve 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 unfortunately not as simple as running a linear regression model. Since spot volatility is a latent process one can only recover the filter weights via the reprojection procedure. To clarify this, suppose we would consider the much simpler task of regressing *implied* rather than *spot* volatilities on the same window of past implieds. Such a linear regression model is also reported in Table 6, with its lag coefficients appearing side-by-side with the reprojection filter. The OLS parameter estimates show absolutely no resemblance, neither numerically nor statistically, to extraction filter weights. The difference between the two lag polynomials is rather easy to explain when one realizes that implied volatility is not like spot volatility but rather relates to the expected volatility over the remaining time to maturity of a contract (see for instance Hull and White (1987)). 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 is relatively straightforward.

Conclusion

The basic idea that both the derivatives and cash markets should contain information to price and hedge option contracts was implemented. Its execution involved the use of multivariate simulation-based estimation procedures for stochastic volatility processes. 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 $K \geq 1$ and σ_{it}^I is the Black-Scholes implied volatility. The econometric methods involved are due to Gallant and Tauchen (1996, 1998). We focused primarily on two cases: (i) using an ATM option contract, hence a univariate setup without cash market information and (ii) a bivariate specification with the underlying fundamental and an ATM European contract. The bivariate approach involving both the fundamental and an option appears useful when the information from the cash market provides support via the conditional kurtosis to price options. This is the case for some long term options but the effect is only marginal. Our results, based on the S&P 500 index contract, showed that the univariate approach only involving options by and large dominated. 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 the same diffusion parameter estimates, we found that the filtering of spot volatility is the key ingredient in our procedure. The univariate filter with options appears to excel at this task. Perhaps one could label this the “simple is better” result for SV models, to paraphrase the findings of Dumas *et al.* (1997). An important qualification has nevertheless to be made regarding simplicity. The analysis in this paper rests on a class of option pricing models which is far more sophisticated than the Black-Scholes model. Within this class of SV models we showed that a simple approach works best *and* outperforms the commonly used strategy which exploits a cross-section of the previous day options plugged into the Black-Scholes formula. Furthermore, since our analysis relied on out-of-sample pricing performance we can deduce from the results of Dumas *et al.* (1997), who used the same SPX contract and examined in detail models with deterministic volatility functions, that SV models like Heston’s combined with option price filters should outperform the binomial tree type models as well.

While the point estimates of the diffusion parameters were roughly the same we also found that the use of options in estimation greatly improves the precision of diffusion parameter estimates. This finding confirms the insights of Pastorello *et al.* (1994) and Renault (1997). More

over, having estimated separately the risk-neutral and objective measures allows us to appraise the typical risk-neutral representations used in the literature. Using Heston's (1993) model as example we showed that the usual transformation from objective to risk neutral density is not supported by the data. To examine this we introduced a test which is based on the Radon-Nikodym derivative of the objective measure with respect to the risk neutral and uses the Novikov condition. We also found that the hedging performance is rather insensitive to model specification. In contrast, pricing options is the most prone to model error.

We know that 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 setup where more than one option is used, namely a set of option contracts $[(\sigma_{it}^I)_{i=1,K}]$ 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 involving exclusively options do 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 and other exotic option contracts for such an exploration. Finally, throughout the paper we used Heston's model since it yields closed form pricing formula. Alternative SV model specifications should also be entertained such as the models due to Hull and White (1987) and Scott (1987). The techniques proposed in this paper would also allow us to consider option pricing models with jumps.

Technical Appendices

A The call options pricing using SV model

Here we provide the details of the call options pricing formula under the SV model from Bakshi *et al.* (1997a). The call price in (1.3) is expressed through $\Pi_j, j = 1, 2$, which are equal to:

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

and

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

where:

$$\xi = \sqrt{[\kappa^* - (1 + i\phi)\rho\sigma_V]^2 - i\phi(i\phi + 1)\sigma_V^2}$$

$$\begin{aligned} & f_2(t, \tau, S(t), V(t); \phi) = \\ \exp \left\{ i\phi R\tau - \frac{\theta^*}{\sigma_V^2} \left[2 \ln \left(1 - \frac{[\psi - \kappa^* + i\phi\rho\sigma_V](1 - e^{-\psi\tau})}{2\psi} \right) \right] - \right. \\ & \quad \frac{\theta^*}{\sigma_V^2} [\psi - \kappa^* + i\phi\rho\sigma_V]\tau + \\ & \quad \left. i\phi \ln S(t) + \frac{i\phi(i\phi - 1)(1 - e^{-\psi\tau})}{2\psi - [\psi - \kappa^* + i\phi\rho\sigma_V](1 - e^{-\psi\tau})} V(t) \right\} \quad (\text{A.3}) \end{aligned}$$

where:

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

B SNP density estimation

SNP is a method of nonparametric time series analysis. The code can be downloaded from George Tauchen's website <http://www.econ.duke.edu/~get/snp.html>. The method 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 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 $f_k(z_t | X_t, \Xi) \propto [P_K(z_t, X_t)]^2 \phi(z_t)$, where $P_K(\cdot)$ denotes a multivariate polynomial of degree K_z and $\phi(z)$ denotes the standard normal (possibly multivariate) p.d.f. The process X_t is the vector of M lags of the process of interest and the index k denotes the dimension of Ξ which expands as the sample size increases (see discussion in sections 2.2 and 2.3). Since $f_k(z)$, as the distribution proxy, has to integrate to 1, the constant of proportionality is the inverse of $N_K(X_t) = \int [P_K(s, X_t)]^2 \phi(s) ds$. To achieve a unique representation, the constant term of the polynomial part is put equal to one.

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

$$f_k(\iota_t | X_t, \Xi) = \frac{1}{N_K(X_t)} \frac{[P_K(z_t, X_t)]^2 \phi(z_t)}{|det(R)|}, \quad (\text{B.4})$$

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

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

$$P_K(z_t, X_t) = \sum_{i=0}^{K_z} a_i(X_t) z^i = \sum_{i=0}^{K_z} \left(\sum_{j=0}^{K_X} a_{ij} X_t^j \right) z^i, \quad (\text{B.5})$$

and, 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 put to zero, one gets normal distribution. When K_z is positive and $K_X = 0$, one gets a density that can approximate over a large class whose shape is constant with respect to variation in x , i.e. conditionally homogeneous class. Finally, when both K_z and K_X are positive, we approximate over conditionally heterogeneous class.

The optimal values of K (as well as M_μ, M_R) are chosen based on standard criterions and tests, such as AIC, BIC, Wald and LR tests. Gallant, Tauchen (1993) provide a table which matches value of the parameters K, M_μ, M_R and popular time series models. We reproduce this table here (table 1B).

C The Lower Bound for the Novikov Condition

In this appendix we provide the derivation of the lower bound on E_N which we used in (3.9) to test whether the data supports the assumption that the linearity of the volatility drift is preserved under the probability measure transformation.

First, we consider an expected value of the spot volatility from (1.2):

$$\begin{aligned}\bar{V} \equiv E(V(t)) &= \int_0^t E(\theta - \kappa V(u)) du = \int_0^t (\theta - \kappa \bar{V}) du \\ &= \theta t - \kappa \bar{V} t \Rightarrow \bar{V} = \frac{\theta t}{1 + \kappa t}\end{aligned}\tag{C.1}$$

Then for the Novikov condition (3.8), we have:

$$\begin{aligned}E_N(t, t + \tau) &= E\left(\exp\left(\frac{1}{2} \int_t^{t+\tau} (\lambda_1^2(u) + \lambda_2^2(u)) du\right)\right) \\ &\stackrel{\geq 1}{\geq} \exp\left(\frac{1}{2} \int_t^{t+\tau} E(\lambda_1^2(u) + \lambda_2^2(u)) du\right) \\ &\stackrel{\geq 2}{\geq} \exp\left(\frac{1}{2} \int_t^{t+\tau} \left[\frac{(\mu_S - R)^2 + C_1^2}{\bar{V}} + C_2^2 \bar{V} - 2C_1 C_2\right] du\right) \\ &\stackrel{= 3}{=} \exp\left(\frac{1}{2} \int_t^{t+\tau} \left[\frac{((\mu_S - R)^2 + C_1^2)(1 + \kappa u)}{\theta u} \right. \right. \\ &\quad \left. \left. + C_2^2 \frac{\theta u}{1 + \kappa u} - 2C_1 C_2\right] du\right) \\ &\stackrel{= 4}{=} \exp\left(\frac{1}{2} \left[\frac{(\mu_S - R)^2 + C_1^2}{\theta} \int_t^{t+\tau} \left(\frac{1}{u} + \kappa\right) du \right. \right. \\ &\quad \left. \left. + C_2^2 \frac{\theta}{\kappa^2} \int_{1+\kappa t}^{1+\kappa(t+\tau)} \left(1 - \frac{1}{w}\right) dw - 2C_1 C_2 \int_t^{t+\tau} du\right]\right) \\ &= \exp\left(\frac{1}{2} \left[\frac{\kappa}{\theta} ((\mu_S - R)^2 + C_1^2) + \frac{\theta}{\kappa} C_2^2 - 2C_1 C_2\right] \tau \right. \\ &\quad \left. - C_2^2 \frac{\theta}{\kappa^2} \log\left|1 + \frac{\tau}{1/\kappa + t}\right| + \frac{(\mu_S - R)^2 + C_1^2}{\theta} \log\left|1 + \frac{\tau}{t}\right|\right)\end{aligned}$$

The inequality 1 follows from Jensen's inequality for convex functions. The inequality 2 is obtained by substituting in the values of $\lambda_1(t)$, $\lambda_2(t)$ provided in (3.4), (3.5) first and by applying Jensen's inequality to the function $1/x$ second. The equality 3 is obtained from (C.1) and equality

4 by making a substitution $w = 1 + \kappa u$ in the middle integral. If we have $\kappa \in [0, 1]$, then substituting $1/\kappa$ in the last line by 1 we obtain an inequality which leads us to the desired result in (3.9).

References

- [1] Aït-Sahalia, Y. and A. Lo (1998): Nonparametric Estimation of State-Price Densities Implicit in Financial Prices, *Journal of Finance*, **53**, 499-548.
- [2] Aït-Sahalia, Y., Y. Wang and F. Yared (1997): Do options markets correctly price the probabilities of movements of the underlying asset?, *Discussion paper*, University of Chicago.
- [3] Andersen, T.G., H.-J. Chung and B.E. Sørensen (1997): Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study, *Discussion Paper*, Northwestern University.
- [4] Andersen, T. and J. Lund (1996): The Short Rate Diffusion Revisited: An Investigation Guided by the Efficient Method of Moments. *Discussion Paper*, Northwestern University.
- [5] Andersen, T.G. and J. Lund (1997): Estimating continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics*, **77**, 343-377.
- [6] Bakshi, G., C. Cao and Z. Chen (1997a): Empirical performance of alternative option pricing models, *Journal of Finance*, **52**, 2003-2049.
- [7] Bakshi, G., C. Cao and Z. Chen (1997b): Can Markovian models explain option price dynamics? Lessons from high-frequency option data, *Discussion paper*, Penn State University.
- [8] Bates, D.S. (1996): Testing Option Pricing Models. In Maddala, G.S. and C.R. Rao (eds.), *Handbook of Statistics, Vol. 14*, Elsevier Science B.V.
- [9] Black, F. (1976): Studies in Stock Price Volatility Changes, *Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section, American Statistical Association*, 177-181.
- [10] Black, F. and M.S. Scholes (1973): The pricing of options and corporate liabilities, *Journal of Political Economy*, **81**, 637-659.
- [11] Breeden, D. and R. Litzenberger (1978): Prices of State Contingent Claims Implicit in Option Prices, *Journal of Business*, **51**, 621-652.
- [12] Broadie, M., J. Detemple, E. Ghysels and O. Torres (1997): American Options with Stochastic Volatility and Dividends: A Nonparametric Approach, *Journal of Econometrics*, (forthcoming).

- [13] Campbell, J., A. Lo and C. MacKinlay (1997): *The Econometrics of Financial Markets*, Princeton University Press.
- [14] Canina, L. and S. Figlewski (1993): The informational content of implied volatility, *Review of Financial Studies*, **6**, 659-681.
- [15] Chernov, M. and E. Ghysels (1998): Filtering Stochastic Volatility via Reprojection, Work in Progress.
- [16] Cox, J., J. Ingersoll and S. Ross (1985): A theory of the term structure of interest rates, *Econometrica*, **53**, pp. 385-408.
- [17] Day, T. and C. Lewis (1992): Stock Market Volatility And the Information Content of Stock Index Options, *Journal of Econometrics*, **52**, 267-287
- [18] Diebold, F.X. and R.S. Mariano (1995): Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, **13**, 253-263.
- [19] Dumas, B., J. Fleming and R.E. Whaley (1997): Implied volatility functions: Empirical Tests, *Journal of Finance*, (forthcoming).
- [20] Fenton, V.M. and A.R. Gallant (1996): Qualitative and Asymptotic Performance of SNP Density Estimators, *Journal of Econometrics* **74**, 77-118.
- [21] Fleming, J. (1994): The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices, *Discussion Paper*, Rice University
- [22] Figlewski, S. (1998): Derivatives Risks, Old and New, *Discussion Paper*, New York University.
- [23] Gallant, A.R., D. Hsieh and G. Tauchen (1997): Estimation of stochastic volatility models with diagnostics, *Journal of Econometrics*, **81**, 159-192.
- [24] Gallant, A.R., C.-T. Hsu and G. Tauchen (1998): Calibrating Volatility Diffusions and Extracting Integrated Volatility, *Discussion Paper*, Duke University.
- [25] Gallant, A.R. and Jonathan R. Long (1997): Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Square, *Biometrika*, **84**, 125-141.

- [26] Gallant, A.R. and G. Tauchen (1989): Semiparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications, *Econometrica*, **57**, 1091-1120.
- [27] Gallant, A.R. and G. Tauchen (1993): SNP: A Program for Nonparametric Time Series Analysis, version 8.3, User's Guide, *Discussion paper*, University of North Carolina at Chapel Hill.
- [28] Gallant, A.R. and G. Tauchen (1996): Which Moments to Match? *Econometric Theory*, **12**, 657-681.
- [29] Gallant, A.R. and G. Tauchen (1997): EMM: a program for efficient method of moments estimation, Version 1.4, User's guide, *Discussion paper*, University of North Carolina at Chapel Hill.
- [30] Gallant, A.R. and G. Tauchen (1998): Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions, *Journal of American Statistical Association*, (forthcoming).
- [31] Ghysels, E., A. Harvey and E. Renault (1996): Stochastic Volatility. In Maddala, G.S. and C.R. Rao (eds.): *Handbook of Statistics, Vol. 14*, Elsevier Science B.V.
- [32] Ghysels, E. and J. Jasiak (1996): Stochastic Volatility and Time Deformation, *Discussion Paper CIRANO*.
- [33] Hansen, L. (1982): Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica*, **50**, 1029-1054.
- [34] Hansen, L. and R. Jagannathan (1991): Implications of Security Market Data for Models of Dynamic Economies, *Journal of Political Economy*, **99**, pp. 225-262.
- [35] Harrison, M. and D. Kreps (1979): Martingales and Arbitrage in Multiperiod Securities Markets, *Journal of Economic Theory*, **20**, 381-408.
- [36] Harrison, M. and S. Pliska (1981): Martingales and Stochastic Integrals in the Theory of Continuous Trading, *Stochastic Processes and Their Applications*, **11**, 215-260.
- [37] Harvey, A.C., E. Ruiz and N. Shephard (1994): Multivariate Stochastic Variance Models, *Review of Economic Studies* **61**, 247-264.

- [38] 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.
- [39] Hull, J. and A. White (1987): The pricing of options on assets with stochastic volatilities, *Journal of Finance*, **42**, 281-300.
- [40] Jacquier, E. and R. Jarrow (1998): Model Error in Contingent Claim Models: Dynamic Evaluation, *Journal of Econometrics*, (forthcoming).
- [41] Jacquier, E., N.G. Polson and P.E. Rossi (1994): Bayesian Analysis of Stochastic Volatility Models (with discussion): *Journal of Business and Economic Statistics* **12**, 371-417.
- [42] Jacquier, E., N.G. Polson and P.E. Rossi (1995): Stochastic Volatility: Univariate and Multivariate Extensions, Discussion Paper, Boston College.
- [43] Kloeden, P.E. and E. Platen (1995): *Numerical Solution of Stochastic Differential Equations*, Springer Verlag.
- [44] Lamoureux, C.G. and W.D. Lastrapes (1993): Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities, *Review of Financial Studies*, **6**, 293-326.
- [45] Melino, A. (1994): Estimation of Continuous-time Models in Finance. In Sims, C.A. (eds.): *Advances in Econometrics*, Cambridge University Press.
- [46] Nelson, D.B. and D.P. Foster (1994): Asymptotic Filtering Theory for Univariate ARCH Models, *Econometrica* **62**, 1-41.
- [47] Nelson, D.B. (1996): Asymptotic Filtering Theory for Multivariate ARCH Models, *Journal of Econometrics* **62**, 1-41.
- [48] Pastorello, S., E. Renault and N. Touzi (1994): Statistical inference for random variance option pricing, *Unpublished manuscript*.
- [49] Renault, E. (1997): Econometric models of option pricing errors. In Kreps, D.M. and K.F. Wallis (eds.): *Advances in economics and econometrics: Theory and applications, Seventh World Congress, Vol. 3*, Cambridge University Press.
- [50] Renault, E. and N. Touzi (1996): Option Hedging and Implied Volatilities in a Stochastic Volatility Model, *Mathematical Finance*, **6**, 279-302.

- [51] Romano, M. and N. Touzi (1993): Contingent Claims and Market Completeness in a Stochastic Volatility Model, *Discussion paper*, CREST, Paris.
- [52] Schwarz, G. (1978): Estimating the Dimension of the Model, *Annals of Statistics*, **6**, 461-464.
- [53] Scott, L. (1987): Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application, *Journal of Financial and Quantitative Analysis*, **22**, pp. 419-438.
- [54] Tauchen, G. (1997): Chi-square methods in empirical finance. In Kreps, D.M. and K.F. Wallis (eds.): *Advances in economics and econometrics: Theory and applications, Seventh World Congress, Vol. 3*, Cambridge University Press.

Figures and Tables

Figure 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 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.

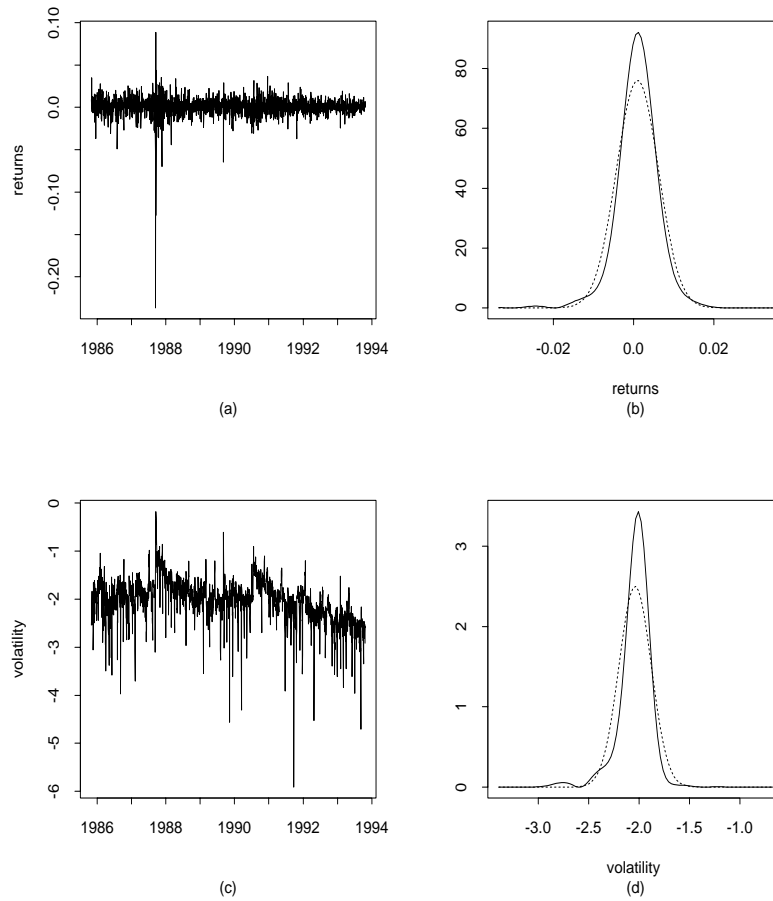


Figure 2: The bivariate SNP density

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. (a) is the perspective plot of the estimated bivariate density; (b) is the contour plot at quantiles 10%, 25%, 50%, 75%, 90%, and 95%; (c) and (d) are the marginal densities of (i) and (ii) correspondingly, the solid line is a plot of an SNP fit, the dashed line is normal density with the same mean and variance.

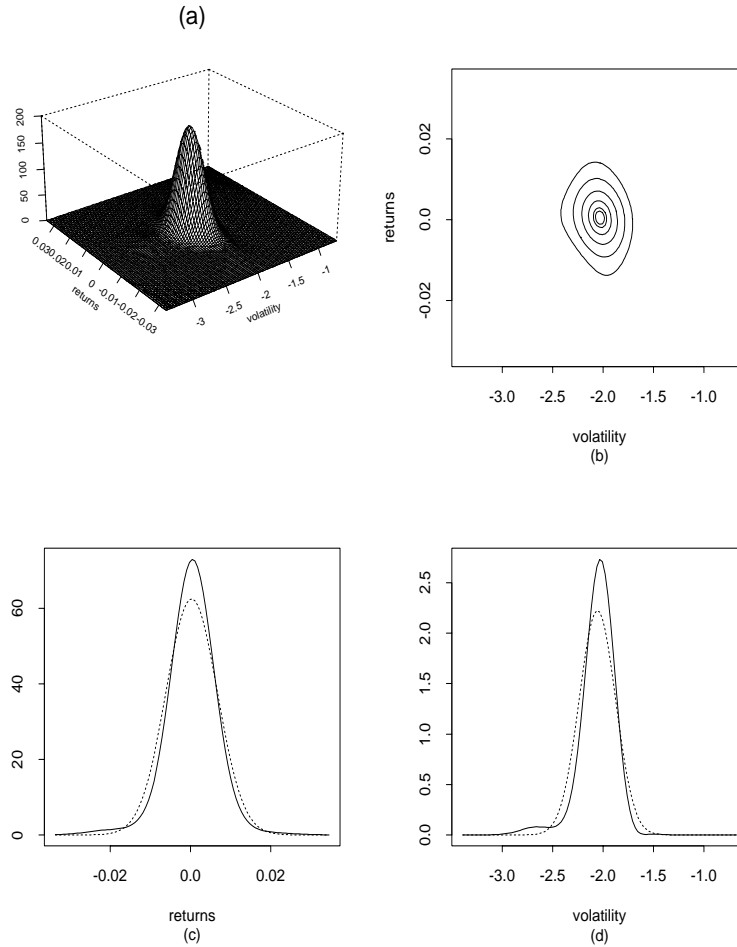


Table 1
The SNP density estimation

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 density. Panel B reports possible densities structures, as described in Gallant and Tauchen (1993).

Panel A

Data type	K_z	K_X	M	M_μ	M_R
S&P 500	9	0	1	3	8
BS vol's	6	0	1	9	5
Joint	4	0	1	4	4

Panel B

Parameter K, M_μ, M_R setting	Characterization of ι_t
$K_z = 0, K_X = 0, M \geq 0, M_\mu = 0, M_R = 0$	iid Gaussian
$K_z = 0, K_X = 0, M \geq 0, M_\mu > 0, M_R = 0$	Gaussian VAR
$K_z > 0, K_X = 0, M \geq 0, M_\mu > 0, M_R = 0$	non-Gaussian VAR, homog. innov.
$K_z = 0, K_X = 0, M \geq 0, M_\mu \geq 0, M_R > 0$	Gaussian ARCH
$K_z > 0, K_X = 0, M \geq 0, M_\mu \geq 0, M_R > 0$	non-Gaussian ARCH, homog. innov.
$K_z > 0, K_X > 0, M > 0, M_\mu \geq 0, M_R \geq 0$	gen. non-lin. process, heterog. innov.

Table 2
EMM estimation results

We have estimated the stochastic volatility model

$$\begin{aligned} \frac{dS(t)}{S(t)} &= Rdt + \sqrt{V(t)}dW_S^*(t), \\ dV(t) &= (\theta^* - \kappa^*V(t))dt + \sigma_V\sqrt{1 - \rho^2}\sqrt{V(t)}dW_V^*(t) + \sigma_V\rho\sqrt{V(t)}dW_S^*(t), \end{aligned}$$

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.

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

Table 3
The number of observations in the options groups used for pricing/hedging performance evaluations

To assess pricing and hedging performance of alternative model specifications we separate the calls into twelve groups by moneyness and time to maturity. The table reports number of observations in each group. The total number of observations is reported in the headers.

Moneyness	Pricing (7424)			Hedging (7132)		
	Days to expiration			Days to expiration		
	<60	60-180	≥180	<60	60-180	≥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

Table 4: The SV model pricing errors

To compute out-of-sample pricing errors, we consider post estimation sample data from 11:93 to 10:94. We report the absolute standard error D_a^p (see (4.1)) and the relative standard error D_r^p (see (4.2)) by moneyness and time to maturity groups and for three model specifications. 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.

We also perform the GMM overidentifying restrictions test to determine whether the means of

$$\left(\frac{C^{observed} - C^{model}}{C^{observed}} \right)^2$$

computed from a pair of model estimates are significantly different. The left panel of the table reports the results of the test by groups and overall, i.e. the values of the χ_1^2 -statistics. The statistics in boldface indicate that the second model does better than the first.

Moneyness	Model	Pricing Errors									Models			Tests			All groups test						
		Days to expiration									Days to expiration			Days to expiration									
		<60			60-180			≥180			<60			60-180				≥180					
D_a^p	D_r^p	D_p^p	D_a^p	D_r^p	D_p^p	D_a^p	D_r^p	D_p^p	D_a^p	D_r^p	D_p^p	Vol-BS	Vol-Joint	Joint-BS	Vol-BS	Vol-Joint	Joint-BS	Vol-BS	Vol-Joint	Joint-BS			
<0.94	BS	0.54	4.62	1.56	20.71	3.12	4.93	3.12	4.93	3.12	4.93	3.12	4.93	3.61***	5.79**	0.29	16.81*						
	Vol	0.42	2.01	0.92	12.58	2.26	4.06	2.26	4.06	2.26	4.06	2.26	4.06	2.69	10.40*	1.32	24.01*						
	Joint	0.44	2.42	1.23	17.35	2.74	4.80	2.74	4.80	2.74	4.80	2.74	4.80	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	4.62	0.45	4.62	0.45	4.62	0.45	10.95*	4.17**	17.64*	16.81*						
	Vol	0.58	8.39	1.38	7.63	3.41	0.33	3.41	0.33	3.41	0.33	3.41	0.33	15.90*	5.42**	11.33*	24.01*						
	Joint	0.73	10.79	1.66	9.91	3.86	0.37	3.86	0.37	3.86	0.37	3.86	0.37	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	5.23	0.31	5.23	0.31	5.23	0.31	5.72**	19.63*	12.28*	16.81*						
	Vol	1.05	4.46	2.20	0.30	4.24	0.24	4.24	0.24	4.24	0.24	4.24	0.24	4.93**	38.20*	18.07*	24.01*						
	Joint	1.22	4.93	2.48	0.36	4.71	0.27	4.71	0.27	4.71	0.27	4.71	0.27	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	4.63	0.18	4.63	0.18	4.63	0.18	0.07	3.04***	8.16*	16.81*						
	Vol	1.35	0.15	2.33	0.16	3.85	0.14	3.85	0.14	3.85	0.14	3.85	0.14	11.31*	13.49*	4.19**	24.01*						
	Joint	1.40	0.16	2.43	0.17	4.21	0.16	4.21	0.16	4.21	0.16	4.21	0.16	4.42**	0.07	3.24***	5.93**						

* indicates significance at the 1% level; ** indicates significance at the 5% level; *** indicates significance at the 10% level.

Table 5: The SV model hedging errors

To compute out-of-sample hedging errors, we consider post estimation sample data from 11:93 to 10:94. We report the absolute standard error D_a^h (see (4.9)) and the relative standard error D_r^h (see (4.11)) by moneyness and time to maturity groups and for three model specifications. 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”, which involves Heston’s call price model estimated using a univariate specification involving options. The specification denoted “Joint” corresponds to the bivariate specification.

We also perform the GMM overidentifying restrictions test to determine whether the means of

$$\left(\frac{\text{Cash}^{observed} - \text{Cash}^{model}}{\text{Cash}^{observed}} \right)^2$$

computed from a pair of model estimates are significantly different. The left panel of the table reports the results of the test by groups and overall, i.e. the values of the χ^2 -statistics. The statistics in boldface indicate that the second model does better than the first.

Moneyness	Model	Hedging Errors						Models	Tests			All groups test
		Days to expiration							Days to expiration			
		<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	<60	60-180	≥180				
<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

* indicates significance at the 1% level; ** indicates significance at the 5% level; *** indicates significance at the 10% level.

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 and concurrent returns are listed in the table. The model selected by the BIC criterion involves 22 lags of implied Black-Scholes volatilities and no lags of returns. We also report parameters and standard errors of AR(22) model for the implied volatilities.

Lags	Coefficients from		Std. err. for AR (22)
	reprojection	AR(22) model	
0	-0.16320	0.02365	0.00431
1	0.03280	0.53407	0.06307
2	0.03277	0.17519	0.07077
3	0.03976	-0.04310	0.07160
4	0.03987	0.13626	0.07155
5	0.04386	-0.02329	0.07190
6	0.04882	0.06550	0.07203
7	0.04179	-0.01373	0.07195
8	0.04298	-0.05188	0.07112
9	0.04715	-0.02546	0.07133
10	0.04506	0.01824	0.07088
11	0.05069	0.05437	0.07101
12	0.05675	-0.08946	0.07137
13	0.06675	0.14030	0.07175
14	0.07836	-0.01184	0.07261
15	0.08792	-0.17026	0.07262
16	0.09978	0.11176	0.07349
17	0.11489	0.01261	0.07401
18	0.14055	-0.08485	0.07479
19	0.16293	0.07815	0.07461
20	0.18804	0.06746	0.07499
21	0.20591	-0.17115	0.07446
22	0.24514	0.00677	0.06807

Liste des publications au CIRANO *

Cahiers CIRANO / *CIRANO Papers* (ISSN 1198-8169)

- 96c-1 Peut-on créer des emplois en réglementant le temps de travail ? / Robert Lacroix
- 95c-2 Anomalies de marché et sélection des titres au Canada / Richard Guay, Jean-François L'Her et Jean-Marc Suret
- 95c-1 La réglementation incitative / Marcel Boyer
- 94c-3 L'importance relative des gouvernements : causes, conséquences et organisations alternative / Claude Montmarquette
- 94c-2 Commercial Bankruptcy and Financial Reorganization in Canada / Jocelyn Martel
- 94c-1 Faire ou faire faire : La perspective de l'économie des organisations / Michel Patry

Série Scientifique / *Scientific Series* (ISSN 1198-8177)

- 98s-21 Using a Financial Training Criterion Rather than a Prediction Criterion / Yoshua Bengio
- 98s-20 Inférence fondée sur les statistiques des rendements à long terme / Cosme Vodounou
- 98s-19 Structural Change Tests for Simulated Method of Moments / Eric Ghysels et Alain Guay
- 98s-18 Managing the Risk of IT Outsourcing / Benoit A. Aubert, Sylvie Dussault, Michel Patry et Suzanne Rivard
- 98s-17 Learning Conditions in the Context of R&D and Development Projects: Empirical Evidence from a Research Centre / Mario Bourgault et Hélène Sicotte
- 98s-16 Assessing the Risk of IT Outsourcing / Benoit A. Aubert, Michel Patry et Suzanne Rivard
- 98s-15 Determinants of Desired Career Paths among Canadian Engineers / Michel Tremblay, Thierry Wils et Caroline Proulx
- 98s-14 Why Is the Bid Price Greater than the Ask? Price Discovery during the Nasdaq Pre-Opening / Charles Cao, Eric Ghysels et Frank Hatheway
- 98s-13 "Informal Thinkering": How Is It Important? / Pierre J. Tremblay
- 98s-12 Organize to Generate Innovation: Tools to Measure Technical Change, Productivity and Change-Generating Efforts / Pierre J. Tremblay
- 98s-11 Worksharing in Québec : Five Case Studies / Michael Huberman et Paul Lanoie
- 98s-10 Predictors of Hierarchical Success for Male and Female Canadian Managers / Denis Chênevert et Michel Tremblay
- 98s-09 Managerial Career Success in Canadian Organizations: Is Gender a Determinant? / Denis Chênevert et Michel Tremblay
- 98s-08 Job Characteristics and the Form of Compensation / W. Bentley MacLeod et Daniel Parent
- 98s-07 Technological Capability and Productivity Growth: An Industrialized / Industrializing Country Comparison / Pierre J. Tremblay
- 98s-06 L'évolution du taux d'activité des femmes au Canada, 1976-1994 : Une analyse de cohortes / Paul Beaudry et Thomas Lemieux

* Vous pouvez consulter la liste complète des publications du CIRANO et les publications elles-mêmes sur notre site World Wide Web à l'adresse suivante :

<http://www.cirano.umontreal.ca/publication/page1.html>