Università Commerciale “Luigi Bocconi” Milano

Ph.D in Economics

XXI Cycle

Econometric Analysis of the Time-Changed Infinite Activity Lévy Option Pricing Models

JUNYE LI

Matr. No. 1094614

Thesis Committee:

Prof. Carlo Favero (Chair), Bocconi University

Prof. Fulvio Ortu, Bocconi University

Prof. Nicholas Polson, University of Chicago GSB

Academic Year 2008-2009

Econometric Analysis of the Time-Changed Infinite Activity Levy Option Pricing Models

Junye Li

M.Sc in Economics, Bocconi University, 2005

M.Eng in Systems Engineering, Beijing Jiaotong University, 2001

A Dissertation Submitted to the Department of Economics at Bocconi University

in Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy in Economics

January 2009

Acknowledgements

I am indebted to many people for help in preparation of this dissertation. First of all, I would like to thank my advisors Prof. Carlo Favero and Prof. Fulvio Ortu for their encouragement, patience and insight and knowledge imparted to me over the last four years. I am also grateful to Prof. Nick Polson from Chicago GSB for being my external advisor and carefully reading the first two chapters of this dissertation. Prof. Pietro Muliere is also highly appreciated for taking time to talk with me frequently.

Thanks are also given to my friends Weiwei Yin, Linlin Niu, Linyun Shang and my colleagues in departments of economics and finance for their support in fields of economics, finance, mathematics and statistics. I appreciate Gerald Abdesaken's help with my English in part of this work.

Finally, I am greatly indebted to my parents for their understanding, support and encouragement while I was pursuing my PhD at Bocconi University.

Executive Summary

This thesis consists of three essays, which have been completed during my doctoral study. Its focus is on econometric analysis of Continuous-Time Levy Stochastic Volatility Models.

The first chapter is *"Spectral Iterative Estimation of Tempered Stable Stochastic Volatility Models and Option Pricing"* which addresses alternative option pricing models and their estimation. The stock price dynamics is modeled by taking into account both stochastic volatility and jumps. Jumps are captured by the tempered stable process and stochastic volatility is introduced by time changing the stochastic processes. We propose a characteristic function based iterative estimation method, which overcomes the problem of non-tractable probability density functions of the models and eases computational difficulty related to other methods. An extension is also made to investigate double-jump model by introducing a jump component in the variance rate process.

The second chapter *"Sequential Bayesian Analysis of Time-Changed Infinite Activity Derivatives Pricing Models"* investigates the time-changed infinite activity derivatives pricing models from the sequential Bayesian perspective. Brownian subordination property of the infinite activity Levy process provides convenience to apply Bayesian filtering methods. I propose a sequential Monte Carlo method with the proposal density generated by the unscented Kalman filter. This approach overcomes to the large extent the problem of particle impoverishment inherent to the conventional particle filter and is more robust in real data applications. Simulation studies and real applications indicate that the underlying alone can't capture the dynamics of states and by including the derivatives in observations, the precision of state filtering gets improved dramatically, and the algorithm proposed can also effectively capture jumps realized by the infinite activity Levy process. The joint identification of the diffusion, stochastic volatility and infinite activity jumps can be achieved using both the underlying and derivatives data.

The last chapter *"Jump Dynamics, Volatility Components and Return-Volatility Relation"* studies the return-volatility relation by taking into account model specification problem. The stock price process is modeled by the time-changed Brownian motion and infinite activity Levy process, which introduces not only the stochastic diffusion volatility but also the stochastic jump intensity. The model

indicates that under the absence of leverage effects it becomes a variant of the Merton's ICAPM whereas under the existence of leverage effects, the return-volatility relation is determined by interactions between risk premia and leverage effects. It provides a theoretical justification for mixed empirical findings. Our empirical study finds a positive return-diffusion volatility relation and more interestingly a negative return-jump volatility relation.

Contents

Chapter 1

Spectral Iterative Estimation of Tempered Stable Stochastic Volatility Models and Option Pricing*

Abstract

This paper considers alternative option pricing models and their estimation. The stock price dynamics is modeled by taking into account both stochastic volatility and jumps. Jumps are captured by the tempered stable process and stochastic volatility is introduced via time changing the stochastic processes. We propose a characteristic function based iterative estimation method, which overcomes the problem of non-tractable probability density functions of the models and facilitates computation. Estimation results and option pricing performance indicate that the infinite activity stochastic volatility model dominates the finite activity model. We also provide an extension to investigate the double-jump model by introducing jumps in the variance rate process.

1 Introduction

Stochastic volatility and jumps in the stock price process are well documented. On the one hand, they are inherent components of the stock price dynamics (Bollerslev et al., 1994; Merton, 1976); on the other hand, they play an important role in the explanation of distributional characteristics of returns and of the implied volatility smile/skew of options. Bakshi et al.(1997) and Bates (2000) find that even though stochastic volatility alone could explain distributional skewness and leptokurtosis of stock returns to a certain extent, its ability to price short-maturity

*This is a joint work with Carlo Favero and Fulvio Ortu. We would like to thank the Editor Eric Renault and two anonymous referees for constructive comments. We also thank Nick Polson, Francesco Corielli, Pietro Muliere and participants at the Bocconi Finance Seminar, European Conference of Financial Management Association (2007), Asian Finance Association Annual Meeting (2007), EC2 Conference on Advances in Time Series Analysis (2007), 2nd European EIF Conference in Finance (2008), Workshop of Quantitative Finance (2008), and Conference in Mathematical and Statistical Methods for Actuarial Sciences and Finance (2008) for helpful comments.

options is limited. By introducing jumps in stock price modeling, this limitation is largely overcome. Jumps mainly affect short-maturity options, while stochastic volatility mainly affects long-maturity options.

Much work has been done on jump-diffusion stochastic volatility models (Bakshi et al., 1997; Bates, 1996, 2000; Pan, 2002; Andersen et al., 2002). These models regard jumps as rare events and use the compound Poisson process to capture them. Stochastic volatility is modeled by a mean-reverting square-root process (Heston, 1993). Even though these jump-diffusion stochastic volatility models perform acceptably in fitting stock price process and in pricing options, they feature a counterfactual assumption that jumps are rare events. By observing time-series evolution of stock prices, we find that the stock price process is accompanied not only by large jumps, but also by a lot of small jumps. Jumps are empirically not rare events. Based on this observation, alternative models are being developed. These models use infinite activity Lévy processes to capture large jumps as well as small jumps in stock price dynamics (Madan et al., 1998; Carr et al., 2002; Carr and Wu, 2003). Stochastic volatility is usually introduced by time-changing these Lévy processes (Carr et al., 2003; Carr and Wu, 2004).

In this paper, we introduce the tempered stable process, which is a Lévy process having at most six parameters in its Lévy density. Depending on different values of the stable index in its Lévy density, the tempered stable process can be an infinite activity process which generates an infinite number of (large and small) jumps, or a finite activity process which generates only a finite number of large jumps. The tempered stable process can also exhibit infinite or finite variation with different values of the stable index. As special cases, it includes many other stochastic processes, such as the compound Poisson process and the variance gamma process. Therefore, with this process we can unify the jump-diffusion model and infinite activity model into one framework with a rich structure. The first financial application of the tempered stable process has been introduced by Carr et al. (2002).

We apply the time-change approach to introduce stochastic volatility. Given a nonnegative right continuous with left limit (*Cadlag*) stochastic process υ_t, we define a stopping time as

$$T_t = \int_0^t \upsilon_{s-} ds,$$

which is finite almost surely. Intuitively, we could think of t as calender time and T_t as business time. The variable υ_t reflects the intensity of economic activity and we call it the variance

rate. Stochastic volatility is generated by replacing calender time t with business time T_t. For a stochastic process $X(t)$, its time-changed counterpart is defined by $X_{T_t} = X(T_t)$. If we assume independence between υ_t and $X(t)$, through iterated expectation we could obtain the conditional characteristic function of X_{T_t}

$$\begin{aligned} \phi_X(u,\tau) &= E[e^{iuX_{T_\tau}}|\mathcal{F}_t] = E[e^{-cT_\tau}|\mathcal{F}_t] \\ &= E[e^{-c\int_t^{t+\tau}\upsilon_s ds}|\mathcal{F}_t] \end{aligned} \tag{1}$$

and the joint conditional characteristic function of X_{T_τ} and $\upsilon_{t+\tau}$

$$\begin{aligned} \phi_{X,\upsilon}(u_1,u_2,\tau) &= E[e^{iu_1X_{T_\tau}+iu_2\upsilon_{t+\tau}}|\mathcal{F}_t] = E[e^{-cT_\tau}e^{iu_2\upsilon_{t+\tau}}|\mathcal{F}_t] \\ &= E[e^{-c\int_t^{t+\tau}\upsilon_s ds}e^{iu_2\upsilon_{t+\tau}}|\mathcal{F}_t], \end{aligned} \tag{2}$$

where, as we shall see in the text, c is related to the characteristic function of $X(t)$. Specifically, when computing option prices, we need the conditional characteristic function under the risk-neutral measure and when implementing estimation, we need the joint conditional characteristic function under the objective measure. Note that the calculation of (1) and (2) is equivalent to finding the transforms in the sense of Duffie, Pan and Singleton (2000) if υ_t falls into the class of the affine jump-diffusion processes.

We firstly build a general model with the exponential time-changed Brownian motion and tempered stable process. Time-changing is introduced via the square-root process (Cox et al., 1985). The variance rate is allowed to be correlated with the return process. The tempered stable process takes on different properties with respect to its stable index. We then investigate these properties by imposing different restrictions on the parameters. In particular, we want to see whether the tempered stable process acts as an infinite activity process or a finite activity process when both jump and stochastic volatility are taken into account in modeling stock price dynamics. We also want to empirically compare the performance of the jump-diffusion stochastic volatility model and the infinite activity stochastic volatility model in pricing options. Finally, we provide an extension of our general model by introducing a jump component in the variance rate process and investigate the double-jump model.

A major difficulty in continuous-time financial modeling is the lack of efficient tools for estimating and making inference with discretely observed samples, especially when models have

latent factors and jumps. This is particularly striking for the models studied in this paper. The frequently used simulation-based methods are difficult to implement since the models are hard to simulate and the traditional GMM is computationally demanding since high order derivatives need to be calculated. Fortunately, for most of the Lévy models, the analytical characteristic functions are obtainable. The characteristic function is equivalent to the probability density function and we could thus directly use it for estimation. Since our models contain the latent factors, we propose a characteristic function-based iterative method to jointly estimate the models by using information contained in both the stock and options markets. Given an initial parameter guess, we firstly back out the unobserved variance rates from options and regard them as if they are observable, and then implement the characteristic function based GMM with a continuum of moment conditions (Carrasco et al., 2007). With estimates obtained in the previous step, we repeat this fashion many times until a certain convergence criterion is reached. With this method, we not only obtain consistent estimates of model parameters, but also identify the market prices of risks as well as filter out a sequence of state variables which should be the best proxy for the true ones.

Estimation using the characteristic function is not novel. It has been investigated since as early as 1970's (Feuerverger and Mureika, 1977) and further discussed by Feuerverger and McDunnough (1981a, 1981b) and Feuerverger (1990). The method is recently redeveloped for the estimation of continuous-time financial models by Singleton (2001), Jiang and Knight (2002) and Chacko and Viceira (2003). Carrasco et al. (2007) extend the method by using a continuum of moment conditions and improve efficiency upon discrete moment conditions. These approaches are very useful for estimating models that do not contain unobserved state variables such as stochastic volatility. We overcome the problem of non-observability of stochastic volatility jointly using stock prices and options.

The investigation of joint estimation with time series data on stock prices and panel data on options is one of the frontiers in empirical work. Renault and Touzi (1996), Pastorello et al.(2000), and Chernov and Ghysels (2000) simply proxy unobserved volatility with the Black-Scholes implied volatility to estimate stochastic volatility models jointly. However, this approach becomes inappropriate when jumps are introduced to the stock price process. Garcia et al. (2006) jointly estimate stochastic volatility models using both option prices and high-frequency underlying prices based on series expansions of option prices and implied volatilities and on the method of moments which takes advantage of tractable moments of realized volatility. This

method suffers from the same problem as its predecessors. Pan (2002) advocates an implied-state GMM to focus directly on the joint dynamics of stock return and near-the-money short maturity options. Pastorello et al. (2003) propose a general iterative and recursive method on estimating structural nonadaptive models. This method actually encompasses a large set of implied state methodologies including the implied-state GMM. Our method is similar to those of Pan (2002) and Pastorello et al. (2003) except that we directly use the characteristic function. Direct use of the characteristic function in estimation makes the estimation of many Lévy models feasible and avoids the demanding tasks of simulation and computation of high-order derivatives.

The remainder of this paper is organized as follows. Section 2 provides a detailed derivation of our models. We derive the conditional characteristic function of return under the risk-neutral measure for option pricing and the joint conditional characteristic function of return and variance rate under the objective measure for estimation. Section 3 describes the characteristic function-based iterative joint estimation method used in this paper. Section 4 presents the data which include both stock prices and options. Section 5 discusses the results of our estimation and evaluates the models. Section 6 extends the model by introducing a jump component in the variance rate process. Lastly, section 7 concludes the paper. Proofs of propositions are provided in the appendix.

2 Models

In this section, we introduce our general model which is built on the time-changed Brownian motion and tempered stable process. Since we eventually aim at option pricing, we firstly specify the risk-neutral stock price dynamics in subsection 2.1 and then derive the objective one with the definition of the market prices of risks in subsection 2.2. We derive both the risk-neutral conditional characteristic function of return for option pricing and the objective joint conditional characteristic function of return and variance rate for model estimation.

2.1 Risk-Neutral Stock Price Dynamics

Under a given probability space $(\Omega, \mathcal{F}, Q)$ and the complete filtration $\{\mathcal{F}_t\}_{t\geq 0}$, we introduce the tempered stable process X_t, which is a Lévy process on R with the Lévy density defined as:

$$\mathrm{v}(x) = c\frac{e^{-\lambda_+ x}}{x^{1+\alpha}}1_{x>0} + c\frac{e^{-\lambda_- |x|}}{|x|^{1+\alpha}}1_{x<0} \tag{3}$$

where $c > 0$ and $\lambda_+, \lambda_- > 0$. To guarantee the finite quadratic variation, the stable index α should be less than 2. The Lévy density $\mathrm{v}(x)$ measures the arrival rate of jumps with size x defined on R^0 (real line without zero). Its characteristic function has the form:

$$\begin{aligned} \phi_X(u) &= E[e^{iuX(t)}] = e^{-t\psi(u)}, \\ \psi(u) &= -c\Gamma(-\alpha)\Big[(\lambda_+ - iu)^\alpha - \lambda_+^\alpha + (\lambda_- + iu)^\alpha - \lambda_-^\alpha\Big] \end{aligned} \tag{4}$$

with $\alpha \neq 1$ and $\alpha \neq 0$, where $\psi(u)$ is called the characteristic exponent, $u \subseteq R$ is the characteristic index and $\Gamma(\cdot)$ is the gamma function [1].

The parameters in the Lévy density (3) play different roles: c measures the overall and relative frequency of jumps; λ_+ and λ_- govern how fast the tails decay and lead to a skewed distribution when they are not the same; and the stable index α governs how the process evolves between big jumps. Specifically, if $\alpha < 0$, the tempered stable process becomes a compound poisson type finite activity process, while if $\alpha \geq 0$, it is an infinite activity process (in particular, when $\alpha = 0$, the tempered stable process becomes the well-know variance gamma process.). When $\alpha < 1$, the tempered stable process exhibits finite variation, whereas when $1 \leq \alpha \leq 2$, it has infinite variation.

The stock price process under the risk-neutral measure Q is modeled by an exponential time-changed Brownian motion and tempered stable process:

$$S_t = S_0 \exp\Big\{(r-q)t + \Big[W_{T_t} - k_W(1)T_t\Big] + \Big[X_{T_t} - k_X(1)T_t\Big]\Big\}, \tag{5}$$

$$T_t \equiv \int_0^t \upsilon_{s-} ds, \tag{6}$$

where r is a constant risk-free rate, q the dividend yield, W_t a standard Brownian motion, X_t the tempered stable process, T_t the stochastic business time, υ_t the variance rate, and $k_W(1)$ and $k_X(1)$ the convexity adjustments. For any stochastic process Y_t, the convexity adjustment could be derived from its cumulant exponent $k(s)$, which is defined as

$$k(s) \equiv \frac{1}{t} log(E[e^{sY_t}]) \equiv -\psi_Y(-is), \tag{7}$$

where $\psi_Y(\cdot)$ is the characteristic exponent of the process Y_t. Apparently, the convexity adjustment of Brownian motion is $k_W(1) = \frac{1}{2}$. The convexity adjustment of the tempered stable process $k_X(1)$ can be derived from (4).

The time-change approach is a standard technique to generate stochastic volatility (Carr et al., 2003). Randomly changed time can be regarded as business time or trading time. The randomness in business time generates the stochastic volatility. In fact, with the time-changing approach, we introduce not only stochastic volatility, but also stochastic higher moments such as skewness and kurtosis. The function $t :\mapsto T_t$ should be nonnegative and nondecreasing, requiring that the variance rate υ_t be a nonnegative process.

A well-known nonnegative process we can use for υ_t is the square-root process of Cox et al. (CIR process; 1985). Under the risk-neutral measure, this process has the following stochastic differential equation (SDE):

$$d\upsilon_t = \kappa(\theta - \upsilon_t)dt + \sigma\sqrt{\upsilon_t}dZ_t, \tag{8}$$

where if $\kappa > 0$, κ is the rate of mean-reversion; θ the long-run mean of the variance rate; σ a variation parameter; and Z_t another standard Brownian motion.

We allow W_t in (5) and Z_t in (8) to be correlated with the instantaneous correlation $[dW_t dZ_t] = \rho dt$, where $\rho \in [-1, 1]$. This is to accommodate the so-called leverage effect of the diffusion part. The leverage effect of jump is actually inherent in the time-changed model because during a time of high variance rate, business time flows faster and price jumps occur at an increased rate.

It is possible to time change two processes separately by using different variance rate processes. To keep parsimoneity of the model, in this paper we use the same variance rate process to time-change both the Brownian motion and the tempered stable process. Under these specifications, we obtain a model of stock price process capturing both the jumps and the stochastic volatility. In the following, we refer to this model as LTS-SV (Lévy Tempered Stable Stochastic Volatility Model). This general model can flexibly explain the negative skewness and leptokurtosis in the distribution of stock returns. Negative skewness can arise either from the difference in tail parameters of the tempered stable process or from negative correlation between the variance rate and return process. The positive excess kurtosis can arise either from a high jump frequency induced by the tempered stable process or from a volatile variance rate.

Since the return process is correlated with the variance rate process, we firstly internalize this correlation with the approach proposed by Carr and Wu (2004) and then derive the conditional characteristic function of log return with the transform approach in the sense of Duffie et al. (2000).

PROPOSITION 1: Define a new filtration $\mathcal{G}_t$ generated by the business time sigma algebra $\mathcal{F}_{T_t}$. The conditional characteristic function (CCF) of log return $R_{t+\tau} = \ln(S_{t+\tau}/S_t)$ in LTS-SV model with the variance rate process (8) under the risk-neutral measure Q is

$$\begin{aligned} \phi_R(u;\tau,\upsilon_t) &\equiv E^Q[e^{iuR_{t+\tau}}|\mathcal{G}_t] \\ &= e^{iu(r-q)\tau+A(u,\tau)+B(u,\tau)\upsilon_t}, \end{aligned} \tag{9}$$

where

$$\begin{aligned} A(u,\tau) &= -\frac{\kappa\theta}{\sigma^2}\Big[2\log\Big(1-\frac{(\gamma-\kappa^*)(1-e^{-\gamma\tau})}{2\gamma}\Big)+(\gamma-\kappa^*)\tau\Big], \\ B(u,\tau) &= \frac{2[\varphi_W(u)+\varphi_X(u)](1-e^{-\gamma\tau})}{(\gamma-\kappa^*)(1-e^{-\gamma\tau})-2\gamma}, \\ \varphi_W(u) &= \frac{1}{2}(iu+u^2), \\ \varphi_X(u) &= \psi_X(u)+iuk_X(1), \\ \kappa^* &= \kappa-iu\rho\sigma, \\ \gamma &= \sqrt{(\kappa^*)^2+2\sigma^2[\varphi_W(u)+\varphi_X(u)]}. \end{aligned}$$

Note that the characteristic function (9) depends on the unobserved variance rate υ_t. We also note that the information flow is now modeled by the complete filtration $(\mathcal{G}_t)_{t\geq 0}$ generated by the business time sigma algebra $\mathcal{F}_{T_t}$. With the above characteristic functions, we could use fast Fourier transform (FFT) to numerically compute option prices if we can observe the variance rate. Option pricing with FFT is proposed by Carr and Madan (1999). Chourdakis (2005) advocates the fractional Fourier transform (FRFT) in pricing options. It is demonstrated that FRFT is more efficient than FFT in the sense of computational precision by careful selection of the integration upper bound and grid sizes of the characteristic index and log strike. In this paper, we apply FRFT to option pricing.

2.2 Market Prices of Risks and Objective Joint CCF

By introducing stochastic volatility and jumps to the stock price process, the market is no longer complete. There may exist many equivalent martingale measures which can guarantee absence of arbitrage. This feature may produce extra difficulty and complexity in the change of measure since the objective dynamics could be extremely different from the risk-neutral one.

We are interested in the structure-preserving change of measure because it preserves tractability and the same structure under both measures. Under the objective measure P, which is assumed to be absolutely continuous with respect to Q, we propose the following stock price and variance rate dynamics,

$$\begin{aligned} S_t &= S_0 \exp\Big\{(r-q)t + \pi_W T_t + \Big[k_X^P(1) - k_X(1)\Big]T_t \\ &\quad + \Big[W_{T_t}^P - \frac{1}{2}T_t\Big] + \Big[X_{T_t}^P - k_X^P(1)T_t\Big]\Big\}, \end{aligned} \tag{10}$$

and

$$d\upsilon_t = [\kappa(\theta - \upsilon_t) + \pi_v \upsilon_t]dt + \sigma\sqrt{\upsilon_t}dZ_t^P \tag{11}$$

with $T_t = \int_0^t \upsilon_{s-}ds$. Define $\kappa^P \equiv \kappa - \pi_v$. In equations (10) and (11), the term $\pi_W T_t$ denotes the risk premium for the diffusion, the term $\pi_X \equiv (k_X^P(1) - k_X(1))T_t$ represents the risk premium for the jump process and $\pi_v \upsilon_t$ is the risk premium for the volatility. Under the change of measure, W_t^P and Z_t^P are still Brownian motions. To guarantee the absolute continuity between X_t and X_t^P, the coefficients α and c should remain unchanged and only tail parameters could be different (Sato, 1999; Cont and Tankov, 2004). Thus, under the objective measure, the tempered stable process has the Lévy density with the same structure as under the risk-neutral measure, but with different tail parameters. Furthermore, we assume that the risk-neutral measure is simply an exponential tilting of the objective measure. This is justified by the well-known Esscher transform. The Esscher transform is a minimum entropy change of the measure method (Chan, 1999), which indicates that there exists a constant ξ such that the objective Lévy density is related to the risk-neutral one through $\mathrm{v}^P(x) = e^{\xi x}\mathrm{v}(x)$. We thus have the following objective Lévy density of the tempered stable process X_t^P

$$\mathrm{v}^P(x) = c\frac{e^{-(\lambda_+ - \xi)x}}{x^{1+\alpha}}1_{x>0} + c\frac{e^{-(\lambda_- + \xi)|x|}}{|x|^{1+\alpha}}1_{x<0}. \tag{12}$$

The intuition behind this measure change is consistent with our understanding of financial market movements. Large jumps play very important roles in option pricing and risk management since they determine the tail behavior of the distribution of returns.

To estimate the model, we need the joint conditional characteristic function of return and the variance rate under the objective measure. The following proposition gives the tractable joint CCF of log return and the variance rate using the same method as before.

PROPOSITION 2: The joint conditional characteristic function of log return and variance rate with specifications of (10) and (11) under the objective measure is given by

$$\begin{aligned}\phi_{R,\upsilon}(u_1,u_2;\tau,\upsilon_t) &\equiv E^P[e^{iu_1R_{t+\tau}+iu_2\upsilon_{t+\tau}}|\mathcal{G}_t] \\ &= e^{iu_1(r-q)\tau+A(u_1,u_2,\tau)+B(u_1,u_2,\tau)\upsilon_t}, \end{aligned} \tag{13}$$

where

$$\begin{aligned}
A(u_1,u_2,\tau) &= \frac{\kappa\theta(ac-d)}{bcd}\log\left(\frac{c+de^{b\tau}}{c+d}\right)+\frac{\kappa\theta}{c}\tau, \\
B(u_1,u_2,\tau) &= \frac{1+ae^{b\tau}}{c+de^{b\tau}}, \\
a &= iu_2(d+c)-1, \\
b &= \frac{d(-\kappa^{P*}-2uc)+a(-\kappa^{P*}c+\sigma^2)}{ac-d}, \\
c &= -\frac{\kappa^{P*}+\sqrt{(\kappa^{P*})^2+2\sigma^2u}}{2u}, \\
d &= (1-iu_2c)\frac{-\kappa^{P*}+iu_2\sigma^2+\sqrt{(\kappa^{P*})^2+2\sigma^2u}}{-2iu_2\kappa^{P*}+(iu_2\sigma)^2-2u}, \\
u &= \varphi_W^P(u_1)+\varphi_X^P(u_1)-iu_1(\pi_W+\pi_X), \\
\varphi_W^P &= \frac{1}{2}(iu_1+u_1^2), \\
\varphi_X^P &= \psi_X^P(u_1)+iu_1k_X^P(1), \\
\kappa^{P*} &= \kappa^P-iu_1\rho\sigma,
\end{aligned}$$

This general model nests a number of specific models, obtained by imposing appropriate restrictions on the parameters. For example, the jump-diffusion stochastic volatility model can be obtained by imposing α to be negative. Therefore, estimation will naturally select the best-fitting specification.

3 Econometric Methodology

We assume that stock and options markets are fully integrated. It is well-known that the information content in the stock market differs from that in the options market. The stock market contains the historical information regarding stock price evolution, whereas the options market reflects information regarding the expectation of future stock prices. Parameter estimates should reflect both sources of information. We propose a characteristic function-based iterative

method which aims at making full use of the information contained in both markets and of the tractability of characteristic functions of the models.

3.1 CCF Based Iterative Estimation

For our models, we have two sets of parameters and a sequence of the state variable υ_t. The two sets of parameters are those of the risk-neutral parameters and the risk premium parameters. We denote them as

$$\Theta^{RN} = (\kappa, \theta, \sigma, \rho, c, \lambda_+, \lambda_-, \alpha)$$

and

$$\Theta^{RP} = (\pi_w, \xi, \pi_\upsilon),$$

respectively. According to our model specification, the risk-neutral parameters can be fully identified using the option price data alone. By using both the stock price data and option price data, we can not only identify the risk-premium parameters but also improve the estimation efficiency of the risk-neutral parameters. The main assumptions here are that there exists an one-to-one relationship between the observed option prices and the unobserved variance rates and a fixed point argument could ensure the convergence of the method (see Appendix B for the discussion of identification and convergence). Our iterative method works with the following steps:

Step 1: Given any initial guess of parameters $(\Theta^{RN})^{(0)}$, we use options data to back out variance rates. Thanks to the analytical characteristic function of return in our model, we could compute option price with the fractional fast Fourier transform. In principle, we could use any options traded on the market. However, in our estimation, we only choose at-the-money short maturity call options. This is because these options are the most liquid instruments and convey the most precise information about market fluctuation.

At this step, we obtain a sequence of variance rates $(\upsilon_t^{(1)})_{t=0}^T$.

Step 2: With the variance rate $(\upsilon_t^{(1)})_{t=0}^T$ obtained at Step 1 and stock price data, under the objective measure we implement the joint conditional characteristic function based GMM with a continuum of moment conditions (Joint CCF-CGMM) proposed by Carrasco et al. (2007), which is described in the following subsection. We note that the risk-premium parameters never appear in the option pricing model. It is then desirable to estimate the risk-neutral parameters and the risk-premium parameters iteratively in order to speed up the convergence and avoid the

local minima.

During this step we conduct an internal iterative loop to iteratively estimate these two sets of parameters, that is, firstly conditional on the risk premium parameters, we estimate the risk-neutral parameters; and then conditional on this estimated risk-neutral parameters, we estimate the risk premium parameters. Repeating a certain number of times, we get the risk-neutral parameter estimates $(\Theta^{RN})^{(1)}$ and the risk premium estimates $(\Theta^{RP})^{(1)}$. Here the convergence is very fast.

Step 3: With the risk-neutral parameter estimates $(\Theta^{RN})^{(1)}$ obtained in Step 2, we back out the variance rate $(\upsilon_t^{(2)})_{t=0}^T$ again as described in Step 1.

Step 4: Repeat Step 1, Step 2 and Step 3 many times until convergence is achieved.

Finally, we obtain the risk-neutral parameter estimates $(\Theta^{RN})^{(n)}$, the risk premium parameter estimates $(\Theta^{RP})^{(n)}$ and a sequence of variance rate $(\upsilon_t^{(n)})_{t=0}^T$. These estimates reflect information contained in both stock and options markets. Under certain regularity conditions, the parameter estimates $(\Theta^{RN})^{(n)}$ and $(\Theta^{RP})^{(n)}$ are consistent and normally distributed and the state variable sequence $(\upsilon_t^{(n)})_{t=0}^T$ is the best proxy for the true variance rates (Pan, 2002; Pastorello et al., 2003).

3.2 Joint CCF-CGMM

In this subsection, we summarize the recently developed conditional characteristic function based GMM with a continuum of moment conditions (CCF-CGMM; Carrasco et al., 2007) and show how this method could be used to estimate the models studied in this paper. The CCF-CGMM is computationally less demanding than the commonly used simulation-based method and traditional GMM. It also solves problems of singularity and instability induced by the discrete moment condition characteristic function based GMM (Singleton, 2001).

The models in this paper contain an unobserved state variable. The (log) stock price process is no longer Markovian. However, the full system $Y_t = (R_t, \upsilon_t)'$ is a Markov process and stationary. The joint conditional characteristic function has already been derived in Section 2. If we could observe the state variable υ_t, we could implement the (two-dimensional) characteristic function based GMM with a continuum of moment conditions. By the definition of conditional characteristic function, we have the following conditional moment conditions

$$E\Big[\exp\{iuY_{t+\tau}\} - \phi_{R,\upsilon}(u;\tau,\upsilon_t)\Big|\mathcal{G}_t\Big] = 0, \tag{14}$$

where $u = (u_1, u_2)$. The first term in (14) is the empirical joint characteristic function and the second is the theoretical joint conditional characteristic function of return and the variance rate derived in section 2. Since the joint conditional characteristic function of $R_{t+\tau}$ and $\upsilon_{t+\tau}$ depends only on υ_t, the conditional moments (14) could be transformed to the unconditional moments if there exists a set of instruments $Z(\cdot, \upsilon_t)$:

$$E\Big[Z(\cdot, \upsilon_t)\Big(\exp\{iuY_{t+\tau}\} - \phi_{R,\upsilon}(u; \tau, \upsilon_t)\Big)\Big] = 0. \tag{15}$$

υ_t is usually unobservable. However, since there are two markets (stock and options) based on the same stock price dynamics, we could back out υ_t from the option and regard it as if it is truly observable. Thus, we can apply CCF-CGMM with a continuum of moment conditions with respect to the characteristic index and spanning optimal instruments by the exponential functions. By doing so, the resulting CCF-CGMM solves the problems of efficiency and singularity of the covariance matrix. For our models, the instruments can be constructed by

$$Z(z, \upsilon_t) = e^{iz\upsilon_t}, \tag{16}$$

and accordingly we have moment functions

$$\epsilon(u, z; \upsilon_t, \beta) = e^{iz\upsilon_t}\Big[\exp\{iuY_{t+\tau}\} - \phi_{R,\upsilon}(u; \tau, \upsilon_t)\Big], \tag{17}$$

where β is a vector of parameters which we are going to estimate. Note that although the optimal instrument can not be obtained, it could be spanned by a set of basis functions (16). Under certain regularity conditions, CCF-CGMM estimation results in MLE efficiency if the state variable υ_t is observable through:

$$\hat{\beta}_T = \arg\min_{\beta \in B} \Big|\Big|\bar{\epsilon}_T(\beta)\Big|\Big|^2_{\mathcal{W}_T}, \tag{18}$$

where $\bar{\epsilon}_T(\beta) = \frac{1}{T}\sum_{t=1}^{T} \epsilon(u, z; \upsilon_t, \beta)$ is the sample counterpart of moment conditions; $\mathcal{W}_T$ is the weighting covariance operator; $||\cdot||$ stands for a norm defined in a Hilbert space of complex-valued functions [2] and B is a compact parameter space. The estimator $\hat{\beta}_T$ is asymptotically normal,

$$\sqrt{T}(\hat{\beta}_T - \beta_0) \xrightarrow{d} \mathcal{N}(0, V_T), \tag{19}$$

and

$$V_T = \left\langle E\Big(\frac{\partial \bar{\epsilon}_T(\beta)}{\partial \beta}\Big), E\Big(\frac{\partial \bar{\epsilon}_T(\beta)}{\partial \beta}\Big) \right\rangle_{\mathcal{W}_T}^{-1}, \tag{20}$$

where β_0 is true parameter and $\langle , \rangle_{\mathcal{W}_T}$ indicates the inner product with respect to $\mathcal{W}_T$ in the defined Hilbert space.

In their original paper, Carrasco et al. (2007) have detailedly discussed how to construct the weighting covariance operator $\mathcal{W}_T$ by introducing a regularization parameter and then simplify the optimization problem resorting on the fact that $\{\epsilon_t\}_{t\geq 0}$ forms a martingale difference sequence with respect to the filtration $\{\mathcal{G}_t\}_{t\geq 0}$. Here we suggest simply to use the identity matrix as the weighting covariance operator. In GMM, Cochrane (2005) advocates to use identity matrix as the weighting matrix, which produces more robust estimates. Appendix C conducts a Monte Carlo study with Heston stochastic volatility model, which shows that the loss of efficiency is not significant.

4 Data

The data used in this paper are S&P 500 index and index options traded in the Chicago Board Options Exchange (CBOE) during the period from January, 1996 to December, 1999. The data are in weekly frequency and there are totally 202 weeks. The dataset contains the following series on option Trading Date, Expiration Date, Strike Price, Last Price, Last Bid Price, Last Ask Price and Underlying Price [3]. The interest rates are proxied by the US 3-month Treasury bill rates, which, together with the dividend yields of S&P 500 index, are downloaded from *Datastream*.

Figure 1 plots the time-series of S&P 500 index and index returns, from which the characteristics of "jumps" and "time-varying/stochastic volatility" are clearly observable. For the purpose of model estimation, we use S&P 500 index prices and index at-the-money short maturity call options. The at-the-money short maturity (ATM-SM) calls are constructed as follows: among all call options, we choose those with moneyness [4] larger than 0.97 and less than 1.03 and with maturity greater than 15 days and less than 45 days. When there are more than one call option available at each time instant, we select that with moneyness closest to 1. The constructed ATM-SM calls have mean moneyness 1.000 with standard deviation 0.003 and mean maturity approximately 25 days with standard deviation 7.3 days. The Black-Scholes implied volatilities, maturities and moneyness of these constructed ATM-SM call options are depicted in Figure 2.

— Figure 1 around here —

— Figure 2 around here —

We also use call options from June, 1997 to December, 1999 to test our models. The following filters are applied to the dataset. First, we only consider call options. Second, we select call options with maturities greater than 6 days and less than 1.5 years [5] and with moneyness less than 1.06. Last, we exclude call options with last bid prices less than 3/8 dollar. By doing so, we obtain a cross sectional call options with 5,793 weekly observations. When the last price of an option in the dataset is zero, we proxy it with the midprice between the last bid price and the last ask price. Table 1 gives the descriptive statistics of S&P index returns, the constructed ATM-SM call options and the filtered call options.

— Table 1 around here —

5 Results and Discussion

In this section, we present the estimation results and discuss their implications. Subsection 5.1 presents parameter estimates and variance rate estimates and subsection 5.2 studies model comparison through pricing cross-sectional call options.

5.1 Iterative Joint CCF-CGMM Estimators

Table 2 reports the estimation results including estimates and standard deviations. Models are estimated by the iterative Joint CCF-CGMM described in Section 3. The standard errors, which are computed with the empirical counterpart of formula (20), are presented in brackets.

— Table 2 around here —

We first estimate the general model without any restriction (LTS-SV model). The iterative Joint CCF-CGMM is implemented with the number of total iterations 80 such that the norm of the difference in estimates between two successive iterations is sufficiently small. At each iteration, we again use an iterative approach to estimate the risk-neutral parameters and the risk-premium parameters. The convergence of this internal iteration is very fast. We set the number of iterations at 10. The initial values, which determine the success or failure of the algorithm, are carefully selected through trying different values. Figure 3 plots the convergence of our parameter estimation.

— Figure 3 around here —

Looking at jump related parameters, we find that the tempered stable process in this model acts as an infinite activity process with infinite variation since the estimate of α is 1.132 (positive and larger than one). The estimates of the risk-neutral tail parameters λ_+ and λ_- are respectively 22.635 and 1.635, indicating fast right tail dampening and left-skewed distribution. We have discussed in Section 2 that under the change of measure, only tail parameters change, and the other two (c and α) remain constant. The objective tail parameters are related to the risk-neutral ones through $\lambda_+^P = \lambda_+ - \xi$ and $\lambda_-^P = \lambda_- + \xi$. ξ is a risk-premium parameter. The positive estimate of ξ (8.783) implies that the risk-neutral distribution of the tempered stable process is more left-skewed than the objective one.

Turning to the variance rate parameters, we notice different findings. The estimate of κ (18.673) is large and indicates that the variance rate quickly reverts to its long-term mean θ (0.018) and the estimate of σ (0.929) is also relatively large. These are in stark contrast to previous empirical studies on the jump-diffusion stochastic volatility (JDSV) models (Bakshi et al., 1997; Bates, 1996, 2000; Pan, 2002; Andersen et al., 2002), which find the persistent volatility process and relatively small volatility of volatility parameter. The reason for these different findings is that in the JDSV models the stochastic volatility is only from the diffusion part, whereas in our model we use the same variance rate process to time change both the diffusion part and the jump part. The estimated κ and σ should reflect both the persistent diffusion volatility and the transient jump effect. In fact, if we use a different variance process to time change the jump part, the estimates of the mean-reverting parameter and the volatility of volatility parameter for this process are both huge (Javaheri, 2005; Li, 2008). Similar but less significant results have also been found by Huang and Wu (2004) with the different model specification using only option data. We have a negative risk premium of volatility π_v (-2.992), which is consistent with negative correlation between the return process and variance rate process (-0.934). We observe that the risk premium of diffusion is very tiny, indicating that the market doesn't take this risk factor into account and only jump and stochastic volatility are priced. The negative premium on the stochastic variance risk and the relative importance of the jump and volatility risk premia over the diffusion risk premium are also reported in Bates (2000), Pan (2002), Carr and Wu (2008) and others and in Pan (2002), respectively. Additional to the statistical significance of the jump risk premium, our estimate of the volatility risk premium is also highly significant. This is in contrast to Pan (2002) and Broadie et al. (2007), where they find the volatility risk

premium in JDSV models to be hardly significant. This is again because we time-change both the diffusion and jump parts with the same business time and the inherent feature of jumps in stock returns makes clear difference when estimating the volatility risk premium.

We now study another case in which the tempered stable process behaves like a compound Poisson process. Taking a negative value of α in our general model, we obtain a compound Poisson type jump-diffusion stochastic volatility model with an enriched structure. We refer to this model as LTS-SVJD, with which we could compare the infinite activity stochastic volatility model and the jump-diffusion stochastic volatility model. We assign α a value of -0.01. [6] The parameter estimates of the variance rate process do not change much in comparison to those of LTS-SV, but jump related parameter estimates are very different. The estimate of c, which reflects jump frequency, has a very large value (about 230). The negative value of α forces c to capture both large and small jumps and in this case, large jumps and small jumps are indistinguishable. The large c counteracts the effect of the small (negative) α such that LTS-SVJD could reflect stock price dynamics and results in similar values of skewness and kurtosis to those in LTS-SV [7]. The difference between tail parameters λ_+ and λ_- becomes smaller. The risk premium of diffusion in this model is still very small.

The tempered stable process can also take on infinite activity with finite variation when α lies in the interval $[0, 1)$. We study this case by simply taking α equal to 0 and thus the tempered stable process becomes the Variance Gamma process studied by Madan et al. (1998). We refer to this model as LTS-SVVG. Again, the variance rate parameter estimates are not very different from LTS-SV, but jump parameter estimates are different. The estimates are very similar between LTS-SVJD and LTS-SVVG since we select a very small absolute value for α in LTS-SVJD. We find that risk premia of diffusion and volatility are very robust over the models: the diffusion premium is negligible and the volatility premium is about -3.

We now look at variance rate estimates. Figure 4 plots the estimated variance rates of the three models studied. The general shapes are similar to that of the Black-Scholes implied volatility but the values are very different. Means of the square-root variance rates of LTS-SV, LTS-SVJD and LTS-SVVG are 12.4%, 11.2% and 11.2%, respectively. They are all lower than the mean of the Black-Scholes implied volatility. The implication of the variance rate is different from that of the implied volatility. The variance rate is related not only to the instantaneous variance of the diffusion part, but also to the jump arrival rate of the tempered stable process. Similiarity of variance rates is still observable between LTS-SVJD and LTS-SVVG.

— Figure 4 around here —

In GMM, J_T statistics could be applied to test of model's goodness of fit, but under the framework of CGMM, we lack appropriate statistics to conduct this test since we use an infinite number of moment conditions. Therefore we first make a qualitative test by checking the discretized moment conditions.

In the implementation of CCF-CGMM estimation, we approximate the (double) integral with sums in the objective function (18) by selecting 16 equally spaced values of each characteristic index (u_1 and u_2) from the range $[-\pi, \pi]$, and thus have 256 moment conditions in total. The mean of moment conditions is 1.537 for LTS-SV, whereas it is 1.785 for LTS-SVJD and 1.753 for LTS-SVVG [8]. As mentioned in Section 3, the full system $(\ln S_t, v_t)'$ is Markovian, so the sequence of moment functions $\{\epsilon_t\}_{t>0}$ forms a martingale difference sequence with respect to $\mathcal{G}_t$ and hence is theoretically uncorrelated. The means of the first-lag autocorrelation are nearly the same for these three models: they are -0.160. However the means of the second-lag autocorrelation are -0.011, -0.014 and -0.014 for LTS-SV, LTS-SVJD and LTS-SVVG, respectively.

The above discussion implies that the infinite activity stochastic volatility model is more appropriate than the jump-diffusion stochastic volatility model in fitting stock price evolution. We continue to discuss the models' superiority in the next subsection by their performance in pricing options.

5.2 Cross-Section Option Pricing

We begin to compute option prices with parameters and variance rates estimated by the iterative joint CCF-CGMM method for different models. In practice, this is more interesting. We test models according to their capacity in pricing options. The option prices are computed with FRFT and we apply two measures to evaluate the models. The first is the absolute pricing error, denoted as "*Aerr*" and the second is the relative pricing error, denoted as "*Rerr*". These two measures are defined respectively as

$$Aerr = \frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{n_t}\left|P_{ti}^{im} - P_{ti}\right|, \tag{21}$$

$$Rerr = \frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{n_t}\frac{\left|P_{ti}^{im} - P_{ti}\right|}{P_{ti}}, \tag{22}$$

where N is the total number of options we consider, T the number of weeks, n_t the number of options at date t, P_{ti}^{im} the model-implied option price and P_{ti} the market option price of the ith option at date t.

Table 3 presents the absolute and relative pricing errors for each model. The following findings are observed. First, for short maturity call options (maturity less than 60 days), we find that LTS-SV performs better than the other two models, no matter which moneyness we consider. All three models underprice options with moneyness less than 1.00, while they overprice options with moneyness larger than 1.00. Second, for medium maturity call options (maturity between 60-180 days), LTS-SV is better than the other two models for those options with moneyness less than 0.94 and greater than 1.03. LTS-SV underprices all options except those with moneyness larger than 1.03, whereas the other two overprice options with moneyness larger than 0.97. Third, for long maturity call options (maturity larger than 180 days), LTS-SV performs better than the other two only for options with moneyness in $[0.94, 0.97]$ and larger than 1.03. All models overprice the options. Fourth, LTS-SVVG and LTS-SVJD can't be empirically distinguished. They perform nearly the same for all options. Fifth, whenever LTS-SV underperforms the other two models, its inferiority is very small. Lastly, among all call options, these three models perform relatively better in pricing long maturity in-the-money options and worse in pricing deep out-of-the-money short maturity options.

— Table 3 around here —

6 Extension: Double-Jump Model

Recently, some empirical investigations with jump-diffusion stochastic volatility models indicate that a jump component is also necessary in the volatility process (Eraker et al., 2003; Broadie et al., 2007). Thus, a natural extension of our general model is to introduce a jump part in the variance rate process. To this end, we model the variance rate process under the risk-neutral measure with the following SDE,

$$d\upsilon_t = \kappa(\theta - \upsilon_t)dt + \sigma\sqrt{\upsilon_t}dZ_t + dJ_t, \tag{23}$$

where Z_t, as before, is a Brownian motion, which is correlated with W_t and independent of X_t. The new process J_t is a compound Poisson pure jump process, independent of W_t, Z_t and X_t, whose jump sizes are independent and exponentially distributed with mean μ_J and whose jump

times follow a Poisson process with jump intensity λ_J.[9] The variance rate process (23) is the so-called basic affine process (Duffie and Garleanu, 2001). With this specification, we now have jump components both in the return process and variance rate process.

We assume that under the change of measure, the jump process J_t does not change its parameters, that is, the risk premium for risk factor J_t is zero. We thus have the objective model as follows,

$$\begin{aligned} S_t &= S_0 \exp\Big\{(r-q)t + \pi_W T_t + \Big[k_X^P(1) - k_X(1)\Big]T_t \\ &\quad + \Big[W_{T_t}^P - \frac{1}{2}T_t\Big] + \Big[X_{T_t}^P - k^P(1)T_t\Big]\Big\}, \qquad (24) \\ d\upsilon_t &= [\kappa(\theta - \upsilon_t) + \pi_v \upsilon_t]dt + \sigma\sqrt{\upsilon_t}dZ_t^P + dJ_t, \qquad (25) \end{aligned}$$

with $T_t = \int_0^t \upsilon_{s-}ds$. Define $\kappa^P \equiv \kappa - \pi_v$ and $\pi_X \equiv (k_X^P(1) - k_X(1))T_t$. Following the same approach as in Proposition 2, we can derive the joint conditional characteristic function of the return and variance rate under the objective measure (also see Duffie and Garleanu, 2001).

PROPOSITION 3: The joint conditional characteristic function of log return and variance rate with specifications (24) and (25) under the objective measure is

$$\begin{aligned} \phi_{R,\upsilon}(u_1, u_2; \tau, \upsilon_t) &\equiv E^P[e^{iu_1 R_{t+\tau} + iu_2 \upsilon_{t+\tau}} | \mathcal{G}_t] \\ &= e^{iu_1(r-q)\tau + A(u_1,u_2,\tau) + B(u_1,u_2,\tau)\upsilon_t}, \qquad (26) \end{aligned}$$

where $A(u_1, u_2, \tau) = A_1(u_1, u_2, \tau) + A_2(u_1, u_2, \tau)$ and

$$\begin{aligned} A_1(u_1, u_2, \tau) &= \frac{\kappa\theta(a_1c_1 - d_1)}{b_1c_1d_1}\log\Big(\frac{c_1 + d_1e^{b_1\tau}}{c_1 + d_1}\Big) + \frac{\kappa\theta}{c_1}\tau, \\ A_2(u_1, u_2, \tau) &= \frac{\lambda_J(a_2c_2 - d_2)}{b_2c_2d_2}\log\Big(\frac{c_2 + d_2e^{b_2\tau}}{c_2 + d_2}\Big) + \frac{\lambda_J(1 - c_2)}{c_2}\tau, \\ B(u_1, u_2, \tau) &= \frac{1 + a_1e^{b_1\tau}}{c_1 + d_1e^{b_1\tau}}, \\ a_1 &= iu_2(d_1 + c_1) - 1, \\ b_1 &= \frac{d_1(-\kappa^{P*} - 2uc_1) + a_1(-\kappa^{P*}c_1 + \sigma^2)}{a_1c_1 - d_1}, \\ c_1 &= -\frac{\kappa^{P*} + \sqrt{(\kappa^{P*})^2 + 2\sigma^2u}}{2u}, \\ d_1 &= (1 - iu_2c_1)\frac{-\kappa^{P*} + iu_2\sigma^2 + \sqrt{(\kappa^{P*})^2 + 2\sigma^2u}}{-2iu_2\kappa^{P*} + (iu_2\sigma)^2 - 2u}, \\ a_2 &= \frac{d_1}{c_1}, \qquad b_2 = b_1, \end{aligned}$$

$$
\begin{aligned}
c_2 &= 1 - \frac{\mu_J}{c_1}, \qquad d_2 = \frac{d_1 - \mu_J a_1}{c_1}, \\
u &= \varphi_W^P(u_1) + \varphi_X^P(u_1) - iu_1(\pi_W + \pi_X), \\
\varphi_W^P &= \frac{1}{2}(iu_1 + u_1^2), \\
\varphi_X^P &= \psi_X^P(u_1) + iu_1 k_X^P(1), \\
\kappa^{P*} &= \kappa^P - iu_1\rho\sigma.
\end{aligned}
$$

Table 4 presents the parameter estimates of the double-jump model. Focusing on jump parameters of the variance rate process, the mean of jump size μ_J is 0.044 and the jump intensity λ_J is 0.383. These values indicate that the variance rate doesn't undergo large and frequent jumps. To compare with LTS-SV, we find that introducing the jump component in the variance rate process causes the large jump frequency, controlled by c, decreasing and small jump frequency, controlled by α, increasing. The smaller value of λ_- implies an even heavier left tail. The risk-premium parameters nearly remain the same as those of LTS-SV.

— Table 4 around here —

Looking at the moment conditions and autocorrelations of the moment functions, we find that the mean of moment conditions is 1.465 and means of the first two autocorrelation are -0.160 and -0.009. They are all smaller (in absolute term) than those of LTS-SV, indicating a better goodness-of-fit. As for other models, we investigate the performance of the double-jump model in pricing options. Table 3 reports option pricing errors of the double-jump model with model name LTS-SVDJ. The improvement in the absolute option pricing error over LTS-SV is observable for the short maturity options, but for the medium and long maturity options, it is negligible or even worse. We could conclude that the double-jump model we have studied only results in marginal improvement in option pricing and we also conclude that unlike the double-jump jump-diffusion stochastic volatility model (Duffie, Pan and Singleton, 2000), once the stock price process is modeled with the infinite activity Lévy process, Poisson jumps in the stochastic volatility are not critically important.

7 Conclusion

We have studied stock price dynamics by taking into account both stochastic volatility and jumps. Jumps are captured by the tempered stable process and stochastic volatility is intro-

duced by time changing the stochastic processes. For model estimation, we propose a characteristic function based iterative estimation method, which overcomes the problem of non-tractable probability density functions of the models and eases computational difficulty related to the simulation based estimation method and traditional GMM method. The empirical study indicates that the infinite activity stochastic volatility model is more preferable than the jump-diffusion stochastic volatility model. We also make an extension of the general model by introducing a jump component in the variance rate process. Direct use of the characteristic function makes it feasible to estimate this double-jump model. Our empirical study points to an only marginal improvement in option pricing by the adoption of the double-jump model.

In this paper, we estimate the models using stock price data and only at-the-money short maturity call options. All other call and put options are discarded in model estimation. Even though at-the-money short maturity options are the most liquid financial securities, there are many other options which are also very liquid and contain rich information about financial market movement. Therefore, it may be interesting to study models and their implications using not only at-the-money options, but also out-of-the-money and in-the-money options.

APPENDIX

A Proofs of Propositions

We first prove Proposition 2. Since the Brownian motions W_t^P and Z_t^P are correlated, we use the approach proposed by Carr and Wu (2004) to implement a change of measure in order to internalize this correlation. Define a new measure M, which is absolutely continuous with respect to the objective measure P

$$\frac{dM}{dP}\Big|_{\mathcal{G}_t} = \exp\left\{\left[iu(W_{T_t}^P - \frac{1}{2}T_t) + \varphi_W^P T_t\right] + \left[iu(X_{T_t}^P - k_X^P(1)T_t) + \varphi_X^P T_t\right]\right\}. \tag{27}$$

Under this new measure M, the variance rate process becomes,

$$\begin{aligned} d\upsilon_t &= (\kappa\theta - \kappa^P \upsilon_t + iu\rho\sigma\upsilon_t)dt + \sigma\sqrt{\upsilon_t}dZ_t^M \\ &= (\kappa\theta - \kappa^{P*}\upsilon_t)dt + \sigma\sqrt{\upsilon_t}dZ_t^M, \end{aligned} \tag{28}$$

where $\kappa^{P*} = \kappa^P - iu\rho\sigma$ and Z_t^M is now independent of W_t^P. The joint conditional characteristic function of $R_{t+\tau}$ and $\upsilon_{t+\tau}$ can then be calculated as follows with the approaches proposed by Duffie, Pan and Singleton (2000).

$$\begin{aligned}
\phi_{R,\upsilon}(u_1, u_2; \tau, \upsilon_t) &\equiv E^P[e^{iu_1 R_{t+\tau} + iu_2 \upsilon_{t+\tau}} | \mathcal{G}_t] \\
&= e^{iu_1(r-q)\tau} E^P\left[e^{iu_1[(\pi_W+\pi_X)T_\tau + (W_{T_\tau}^P - \frac{1}{2}T_\tau) + (X_{T_\tau}^P - k_X^P(1)T_\tau)] + iu_2\upsilon_{t+\tau}} \middle| \mathcal{G}_t\right] \\
&= e^{iu_1(r-q)\tau} E^P[M(\tau) e^{-[\varphi_W^P + \varphi_X^P - iu_1(\pi_W+\pi_X)]T_\tau + iu_2\upsilon_{t+\tau}} | \mathcal{G}_t] \\
&= e^{iu_1(r-q)\tau} E^M[e^{-\int_t^{t+\tau} u\upsilon_s ds} e^{iu_2\upsilon_{t+\tau}} | \mathcal{G}_t] \\
&= e^{iu_1(r-q)\tau + A(u_1,u_2,\tau) + B(u_1,u_2,\tau)\upsilon_t}, \qquad (29)
\end{aligned}$$

where $u = \varphi_W^P + \varphi_X^P - iu_1(\pi_W + \pi_X)$ and E^M is the expectation operator under the new measure M and $A(u_1, u_2, \tau) = A(t)$ and $B(u_1, u_2, \tau) = B(t)$ solve the following ODEs

$$\dot{B}(t) = u + \kappa^{P*}B(t) - \frac{1}{2}\sigma^2 B^2(t), \qquad (30)$$

$$\dot{A}(t) = -\kappa\theta B(t), \qquad (31)$$

with the boundary conditions $B(t+\tau) = iu_2$ and $A(t+\tau) = 0$. By solving these two ODEs, we obtain the result of Proposition 2.

Under the change of measure defined in the text, Proposition 1 can be easily proved by setting $u_2 = 0$ and suppressing the risk-premium parameters.

The proof of Proposition 3 is similar to that of Proposition 2 except that we now solve ODEs

$$\begin{aligned}
\dot{B}(t) &= u + \kappa^{P*}B(t) - \frac{1}{2}\sigma^2 B^2(t), \\
\dot{A}(t) &= -\kappa\theta B(t) - \lambda_J \frac{\mu_J B(t)}{1 - \mu_J B(t)},
\end{aligned}$$

with $A(t) = A(u_1, u_2, \tau)$ and $B(t) = B(u_1, u_2, \tau)$ as well as boundary conditions $B(t+\tau) = iu_2$ and $A(t+\tau) = 0$. By setting $u_2 = 0$ and suppressing the risk-premium parameters, we could obtain the risk-neutral conditional characteristic function of log return for the double-jump model easily. It can be used for option pricing.

B Identification and Convergence

According to our model specification, we estimate the risk-neutral parameters Θ^{RN}, the risk premium parameters Θ^{RP} and the latent state variable υ_t. For notational simplicity, we write our asset pricing model in the following form

$$Y_t = F(\upsilon_t, \Theta) \tag{32}$$

where $\Theta = \{\Theta^{RN}, \Theta^{RP}\}$. Since Y_t contains both the underlying stock price and the option price, (32) can be explicitly decomposed into two formulas

$$Y_t^1 = F_1(\upsilon_t, \Theta^{RN}), \tag{33}$$

$$Y_t^2 = F_2(\upsilon_t, \Theta^{RN}, \Theta^{RP}), \tag{34}$$

where Y_t^1 represents the option price, F_1 the risk-neutral model, Y_t^2 the stock price and F_2 the objective model.

We could then back out the variance rate υ_t from the option Y_t^1 given the risk-neutral parameters Θ^{RN} by assuming there is a one-to-one relationship between Y_t^1 and υ_t: $\upsilon_t = F_1^{-1}(Y_t^1, \Theta^{RN})$. Using the option price data $\{Y_t^1\}_{t\geq 0}$, the risk-neutral parameters Θ^{RN} could be fully identified, but we have problem identifying the risk premium parameters Θ^{RP} since these parameters never appear in the pricing formula (33). Fortunately, we also have the stock price data $\{Y_t^2\}_{t\geq 0}$ and the objective model F_2. Using both the stock price and option price data and the pricing formulas (33) and (34), we could not only identify the risk premium parameters Θ^{RP} but also improve the estimation efficiency of the risk-neutral parameters Θ^{RN}.

Assume that $Q_T\big(\Theta, \{Y_t\}_{0\leq t\leq T}\big)$ is the sample objective function and its population limit $Q_\infty\big(\Theta, \{Y_t\}_{-\infty\leq t\leq +\infty}\big)$ exists: $p\lim Q_T = Q_\infty$. Q_T satisfies the usual measurability and continuity conditions. The iterative method discussed in the text results in the consistent parameter estimates

$$\left[\hat{\Theta}_{n+1}^{RN}, \hat{\Theta}_{n+1}^{RP}\right] = \arg\min_{\Theta} Q_T\left[F_1^{-1}(Y_t^1, \Theta_n^{RN}), Y_t^2, \Theta\right] \tag{35}$$

for a sufficiently large n. Of course, to guarantee convergence we also assume that the mapping $\hat{\Theta} : \Theta \to \Theta$ has a fixed point at the true $\Theta = \Theta_0$ and is contracting on Θ. The consistency and root-T asymptotic normality of the estimates are well established. See Pastorello et al. (2003) and Pan (2002).

The parameters Θ^{RN} and Θ^{RP} could be estimated simultaneously at each iteration. However, we observe that Θ^{RP} never appears in the option pricing formula and the model is constructed through the no-arbitrage restriction. It is then desirable to estimate Θ^{RN} and Θ^{RP} iteratively in order to speed up the convergence and avoid the local minima.

In this paper, the objective function is constructed through the joint characteristic function of the return R_t and variance rate υ_t. The CCF-CGMM can produce the MLE efficient estimator if the variance rate υ_t is really observable and as long as the score belongs to the span of the moment conditions. The use of an infinite number of moment conditions can theoretically guarantee this latter condition. When υ_t is unobservable, the iterative methods adopted can still have consistent estimator. We refer readers to Carrasco et al. (2007) and to Pastorello et al. (2003) and Pan (2002) for further discussion of CCF-CGMM and iterative method, respectively.

C A Monte Carlo Study of Joint CCF-CGMM

In this appendix, we conduct a simple Monte Carlo study using Heston stochastic volatility model with the aim to manifest the estimation efficacy of the Joint CCF-CGMM. Heston stochastic volatility model can be obtained from our general model by simply suppressing the tempered stable process and related components. Under the objective measure, it has the form:

$$S_t = S_0 \exp\left\{\mu t + \left(W_{T_t} - \frac{1}{2}T_t\right)\right\}, \tag{36}$$

$$d\upsilon_t = \kappa(\theta - \upsilon_t)dt + \sigma\sqrt{\upsilon_t}dZ_t, \tag{37}$$

with $T_t \equiv \int_0^t \upsilon_s ds$. This time-changed form is equivalent in distribution to the original model (Heston, 1993).

By following the same procedure as in Proposition 2, we can derive the objective joint conditional characteristic function of the return and variance rate. With this joint CCF, we implement the Joint CCF-CGMM estimation discussed in the text. The identity matrix is selected as the weighting covariance operator.

The Monte Carlo study is based on 500 simulations with sample size 500 in weekly frequency. We simulate Heston model with an efficient scheme proposed by Andersen (2007). In this model we have five parameters $\Theta = (\mu, \kappa, \theta, \sigma, \rho)$ and the true values are given by $\Theta_0 = (0.150,\ 6.000,\ 0.025,\ 0.300,\ -0.600)$. We find that the optimization is less sensitive to the initial values when using the identity matrix as the weighting operator. Table 5 presents

our Monte Carlo study results, which indicate that the loss of efficiency is not significant.

— Table 5 around here —

References

[1] Andersen, L. (2007). "Efficient Simulation of the Heston Stochastic Volatility Model." Working Paper, Banc of America Securities.

[2] Andersen, T.G., L. Benzoni, and J. Lund. (2002). "An Empirical Investigation of Continuous-Time Equity Return Models." *The Journal of Finance* 57, 1239-1284.

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

[4] Bates, David S. (1996). "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options." *The Review of Financial Studies* 9, 69-107.

[5] Bates, David S. (2000). "Post-'87 Crash Fears in the S&P 500 Futures Option Market." *Journal of Econometrics* 94, 181-238.

[6] Bollerslev, T., R.F. Engle and D. Nelson. (1994). "ARCH Models." In R.F. Engle and D.L. McFadden (eds), *Handbook of Econometrics*, Vol. 4. Amsterdam, The Netherlands: Elsevier.

[7] Broadie, M., M. Chernov and and M. Johannes. (2007). "Model Specification and Risk Premia: Evidence from Futures Options." *Journal of Finance* 62, 1453-1490.

[8] Carr, P., H. Geman, D.B. Madan and M. Yor. (2002). "The Fine Structure of Asset Returns: An Empirical Investigation." *The Journal of Business* 75, 305-332.

[9] Carr, P., H. Geman, D.B. Madan and M. Yor. (2003). "Stochastic Volatility for Lévy Processes." *Mathematical Finance* 13, 345-382.

[10] Carr, P. and D.B. Madan. (1999). "Option Valuation Using the Fast Fourier Transform." *Journal of Computational Finance* 3, 61-73.

[11] Carr, P. and L. Wu. (2003). "Finite Moment Log Stable Process and Option Pricing." *Journal of Finance* 58, 753-777.

[12] Carr, P. and L. Wu. (2004). "Time-changed Lévy Processes and Option Pricing." *Journal of Financial Economics* 71, 113-141.

[13] Carr, P. and L. Wu. (2008). "Variance Risk Premiums." *Review of Financial Studies*, forthcoming.

[14] Carrasco, M. and J.P. Florens. (2000). "Generalization of GMM to a Continuum of Moment Conditions." *Econometric Theory* 16, 797-834.

[15] Carrasco, M., M. Chernov, J.P. Florens and E. Ghysels. (2007). "Efficient Estimation of General Dynamic Models with a Continuum of Moment Conditions." *Journal of Econometrics* 140, 529-573.

[16] Chacko, G. and L.M. Viceira. (2003). "Spectral GMM Estimation of Continuous-time Processes." *Journal of Econometrics* 116, 259-292.

[17] Chan, T. (1999). "Pricing Contingent Claims on Stocks Driven By Lévy Processes." *The Annals of Applied Probability* 9, 504-528.

[18] Chernov, M. and E. Ghysels. (2000). "A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation." *Journal of Financial Economics* 56, 407-458.

[19] Chourdakis, K. (2005). "Option Pricing Using the Fractional FFT." *Journal of Computational Finance* 8, 1-18.

[20] Cochrane, J. (2005). *Asset Pricing (Revised)*. Princeton: Princeton University Press.

[21] Cont, R. and P. Tankov. (2004). *Financial Modeling With Jump Processes*. London: Chapman & Hall/CRC.

[22] Cox, J. C., J.E. Ingersoll and S.A. Ross. (1985). "A Theory of the Term Structure of Interest Rates." *Econometrica* 53, 385-408.

[23] Duffie, D. and N. Garleanu. (2001). "Risk and Valuation of Collateralized Debt Obligations." *Financial Analysts Journal* 57, 41-59.

[24] Duffie, D., J. Pan and K. Singleton. (2000). "Transform Analysis and Asset Pricing for Affine Jump-Diffusions." *Econometrica* 68, 1343-1376.

[25] Eraker, B., M. Johannes and N. Polson. (2003). "The Impact of Jumps in Equity Index Volatility and Returns." *Journal of Finance* 58, 1269-1300.

[26] Feuerverger, A. (1990). "An Efficiency Result for the Empirical Characteristic Function in Stationary Time-series Models." *The Canadian Journal of Statistics* 18, 155-161.

[27] Feuerverger, A. and P. McDunnough. (1981a). "On the Efficiency of Empirical Characteristic Function Procedures." *Journal of the Royal Statistical Society* 43, 20-27.

[28] Feuerverger, A. and P. McDunnough. (1981b). "On Some Fourier Methods for Inference." *Journal of the American Statistical Association* 76, 379-387.

[29] Feuerverger, A. and R.A. Mureika. (1977). "The Empirical Characteristic Function and Its Application." *The Annals of Statistics* 5, 88-97.

[30] Gallant, A. R. and G. Tauchen. (1996). "Which Moments to Match?" *Econometric Theory* 12, 657-681.

[31] Garcia, R., M.A. Lewis, S. Pastorello and E. Renault. (2006). "Estimation of Objective and Risk-Neutral Distributions based on Moments of Integrated Volatility." Working Paper, CIRANO, University of Montreal.

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

[33] Huang, J. and L. Wu. (2004). "Specification Analysis of Option Pricing Models Based on Time-Changed Lévy Processes." *The Journal of Finance* 59, 1405-1439.

[34] Jiang, G.J. and J.L. Knight. (2002). "Estimation of Continuous-Time Processes via the Empirical Characteristic Function." *Journal of Business and Economic Statistics* 20, 198-212.

[35] Li, J. (2008). "Stochastic Jump Intensity, Stochastic Volatility and Stochastic Higher Moments in Asset Returns: An Empirical Investigation." Working Paper, Bocconi University, Milan.

[36] Madan, D., P. Carr and E. Chang. (1998). "The Variance Gamma Process and Option Pricing." *European Finance Review* 2, 79-105.

[37] Merton, R.C. (1976). "Option Pricing When Underlying Stock Returns Are Discontinuous." *Journal of Financial Economics* 3, 125-144.

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

[39] Pastorello, S., V. Patilea and E. Renault. (2003). "Iterative and Recursive Estimation in Structural Nonadaptive Models." *Journal of Business & Economic Statistics* 21, 449-509.

[40] Renault, E. and N. Touzi (1996). "Option Hedging and Implied Volatilities in a Stochastic Volatility Model." *Mathematical Finance* 6, 279-302.

[41] Sato, K. (1999). *Levy Processes and Infinitely Divisible Distributions*. Cambridge: Cambridge University Press.

[42] Singleton, K.J. (2001). "Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function." *Journal of Econometrics* 102, 111-141.

[43] Wu, L. (2006). "Dampened Power Law: Reconciling the Tail Behavior of Financial Security Returns." *The Journal of Business* 79, 1445-1473.

[44] Yu, J. (2004). "Empirical Characteristic Function Estimation and Its Applications." *Econometric Reviews* 23, 93-123.

Notes

[1]When $\alpha = 1$ or $\alpha = 0$, the process has a different form of characteristic function. The characteristic exponent is

$$\psi_X(u) = c\left[iu/\lambda_+ + \log(1 - iu/\lambda_+)\right] + c\left[-iu/\lambda_- + \log(1 + iu/\lambda_-)\right], \tag{38}$$

when $\alpha = 0$ and it becomes

$$\psi_X(u) = -c(\lambda_+ - iu)\log(1 - iu/\lambda_+) - c(\lambda_- + iu)\log(1 + iu/\lambda_-). \tag{39}$$

when $\alpha = 1$. The characteristic function of a Lévy process could be derived from the Lévy-Kintchine theorem:

$$\psi_X(u) = -iu\mu + \frac{1}{2}u^2\sigma^2 + \int_{-\infty}^{\infty}(1 - e^{iux} + iux1_{|x|<1})\mathrm{v}(x)dx,$$

where ψ is the characteristic exponent and $(\mu, \sigma, \mathrm{v})$ the characteristic triplet.

[2]The estimation is implemented in the Hilbert space of complex-valued functions, which is defined as

$$\mathcal{L}^2(p) = \left\{f : \mathbf{R}^d \to \mathbf{C}; \int |f(u)|^2 p(u)du < \infty\right\},$$

where p is the reference pdf of a distribution and $p(u) > 0$ for all $u \in \mathbf{R}^d$. p dampens off all the oscillating behavior of integrands in estimation. As long as $p > 0$, the choice of p does not affect the estimation efficiency in large sample. The inner product on $\mathcal{L}^2(p)$ is $\langle f, g\rangle = \int f(u)\overline{g(u)}p(u)du$ and norm $||f|| = \sqrt{< f, f >}$, where the overline denotes complex conjugate.

[3]There is also information on Volume, High Price, Low Price and Open Price. In this paper we do not use this information

[4]Moneyness is defined as the ratio between the underlying price and strike, S/K.

[5]We adopt a 252-day/50-week year.

[6]In estimation α tends to zero if we do not assign a specific value. To make the model as efficient as possible, we simply take a value which is close to zero, say, -0.01.

[7]For the tempered stable process, its second, third and (excess) fourth central moments are respectively

$$\begin{aligned} m_2 &= c\Gamma(2-\alpha)(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2})t, \\ m_3 &= c\Gamma(3-\alpha)(\lambda_+^{\alpha-3} - \lambda_-^{\alpha-3})t, \\ m_4 &= c\Gamma(4-\alpha)(\lambda_+^{\alpha-4} + \lambda_-^{\alpha-4})t, \end{aligned}$$

from which we could calculate skewness $(m_3/m_2^{3/2})$ and kurtosis (m_4/m_2^2).

[8]All these values of mean moment conditions are scaled by 10^{-6}.

[9]The characteristic function of this jump process J_t is

$$\phi_J(z) = \exp\left\{-t\lambda_J \frac{iz\mu_J}{iz\mu_J - 1}\right\},$$

where $z \in R$ is the characteristic index.

Table 1: **Descriptive Statistics of Data**

A. S&P 500 Index Returns						
	T	Mean	St. Dev.	Skewness	Kurtosis	JB Test
Weekly	202	0.215	0.174	-0.837	6.668	1(<0.001)
B. Constructed ATM-SM Calls						
	Mean Mn.	Std Mn.	Mean Mt.	Std Mt.	Mean IV.	Std IV.
ATM-SM	1.000	0.003	25.114	7.328	0.187	0.054
C. S&P 500 Index Call Options						
		Moneyness S/K				
	Maturity	<0.94	0.94-0.97	0.97-1.00	1.00 -1.03	1.03 -1.06
Short	<60	3.531	14.502	25.895	43.515	68.664
		(449)	(391)	(1016)	(1112)	(778)
Medium	60-180	18.235	43.745	61.060	83.623	106.647
		(619)	(233)	(199)	(182)	(159)
Long	>180	32.670	88.464	109.020	127.213	147.568
		(247)	(118)	(100)	(98)	(92)

Note: Table presents the descriptive statistics of data we use for model estimation and option pricing. Data are from January, 1996 to December, 1999 in weekly frequency. There are 202 weeks in total. In panel A, JB Test is the Jarque-Bera normality test, where the value 1 indicates rejection of the null hypothesis of normality; In panel B, Mn stands for moneyness and Mt maturity (in days). For option pricing, we use option data from June, 1997 to December, 1999. There are 5,793 call options in total. Panel C presents the mean price and the number of call options (in brackets) of each group. We consider 15 groups.

Table 2: **Iterative Joint CCF-CGMM Estimation of Models**

	Risk-Neutral Parameters								Risk-Premium Parameters		
Model	κ	θ	σ	ρ	c	λ_+	λ_-	α	π_w	ξ	π_v
LTS-SV	18.673	0.018	0.929	-0.934	4.294	22.635	1.684	1.132	0.000	8.783	-2.992
	(0.792)	(0.002)	(0.042)	(0.028)	(0.235)	(2.397)	(0.319)	(0.078)	(0.000)	(0.735)	(0.000)
LTS-SVJD	18.956	0.015	0.837	-0.999	2.30e2	25.607	8.376	-0.010	0.000	7.733	-3.008
	(1.574)	(0.001)	(0.032)	(0.023)	(30.252)	(1.658)	(0.837)	(—)	(0.000)	(0.321)	(0.000)
LTS-SVVG	18.970	0.015	0.836	-0.999	2.21e2	25.462	8.272	0.000	0.000	7.717	-3.009
	(1.274)	(0.003)	(0.051)	(0.038)	(32.577)	(1.846)	(0.827)	(—)	(0.000)	(0.405)	(0.000)

Note: Models are estimated with the iterative Joint CCF-CGMM discussed in Section 3. Standard deviations are reported in brackets. LTS-SV is our general model; LTS-SVJD is the jump-diffusion type stochastic volatility model with the value of α in the tempered stable process being -0.01; LTS-SVVG is the model with α equal to 0 in the tempered stable process.

Table 3: **Absolute and Relative Option Pricing Errors**

		Absolute Error (Aerr)					Relative Error (Rerr)				
Maturity	Model	< 0.94	0.94-0.97	0.97-1.00	1.00-1.03	1.03-1.06	< 0.94	0.94-0.97	0.97-1.00	1.00-1.03	1.03-1.06
< 60	LTS-SV	0.442	1.148	1.317	1.576	1.707	0.150(-)	0.129(-)	0.069(-)	0.040(+)	0.027(+)
	LTS-SVJD	0.608	1.635	1.548	1.744	2.060	0.167(-)	0.137(-)	0.077(-)	0.044(+)	0.033(+)
	LTS-SVVG	0.607	1.635	1.548	1.745	2.065	0.167(-)	0.145(-)	0.082(-)	0.044(+)	0.033(+)
	LTS-SVDJ	0.435	1.139	1.308	1.554	1.659	0.146(-)	0.127(-)	0.069(-)	0.040(+)	0.026(+)
60-180	LTS-SV	1.341	1.932	2.070	2.232	2.048	0.112(-)	0.051(-)	0.036(-)	0.028(-)	0.021(+)
	LTS-SVJD	1.482	1.901	2.026	2.183	2.205	0.117(-)	0.051(-)	0.036(+)	0.028(+)	0.023(+)
	LTS-SVVG	1.478	1.898	2.026	2.183	2.216	0.117(-)	0.051(-)	0.036(+)	0.028(+)	0.023(+)
	LTS-SVDJ	1.337	1.944	2.081	2.237	2.093	0.108(-)	0.051(-)	0.036(-)	0.028(-)	0.021(+)
≥ 180	LTS-SV	1.857	2.388	2.309	2.370	1.891	0.086(+)	0.028(+)	0.022(+)	0.020(+)	0.014(+)
	LTS-SVJD	1.755	2.423	2.303	2.273	2.014	0.073(+)	0.029(+)	0.022(+)	0.019(+)	0.014(+)
	LTS-SVVG	1.758	2.420	2.316	2.280	2.000	0.074(+)	0.029(+)	0.022(+)	0.019(+)	0.014(+)
	LTS-SVDJ	1.846	2.380	2.443	2.667	2.296	0.086(+)	0.028(+)	0.023(+)	0.022(+)	0.016(+)

Note: Absolute pricing error and relative pricing error are calculated with formula (21) and (22), respectively. The data for option pricing are from June, 1997 to December, 1999 and are described in Section 4 and Table 1. Aerr is measured in US dollar. The signs "+" and "-" in brackets represent overpricing and underpricing on average of each model for each group of call options.

Table 4: **Parameter Estimates of Double-Jump Model**

VR CIR		VR Jump		Return Jump		Risk Premia	
κ	18.526	μ_J	0.044	c	1.811	π_w	0.000
	(0.662)		(0.005)		(0.683)		(0.000)
θ	0.017	λ_J	0.383	λ_+	22.172	ξ	8.694
	(0.007)		(0.037)		(2.112)		(0.660)
σ	0.863			λ_-	0.380	π_v	-2.993
	(0.028)				(0.171)		(0.000)
ρ	-0.999			α	1.364		
	(0.020)				(0.119)		

Note: The table presents parameter estimates of the double-jump model with the iterative Joint CCF-CGMM. VR CIR represents the parameters related to the CIR part of the variance rate process (23). VR Jump represents the parameters in the jump process J_t. Return Jump represents the parameters in the tempered stable process and Risk Premia represents the risk-premium parameters. Standard deviations are in brackets.

Table 5: **Monte Carlo Study of Joint CCF-CGMM**

Parameters	μ	κ	θ	σ	ρ
True Value	0.150	6.000	0.025	0.300	-0.600
Mean	0.152	6.351	0.026	0.290	-0.625
Median	0.154	6.424	0.025	0.295	-0.596
RMSE	0.050	1.824	0.008	0.061	0.152

Note: Monte Carlo Study is conducted with Heston stochastic volatility model. The number of simulations is 500 with sample size 500 in weekly frequency.

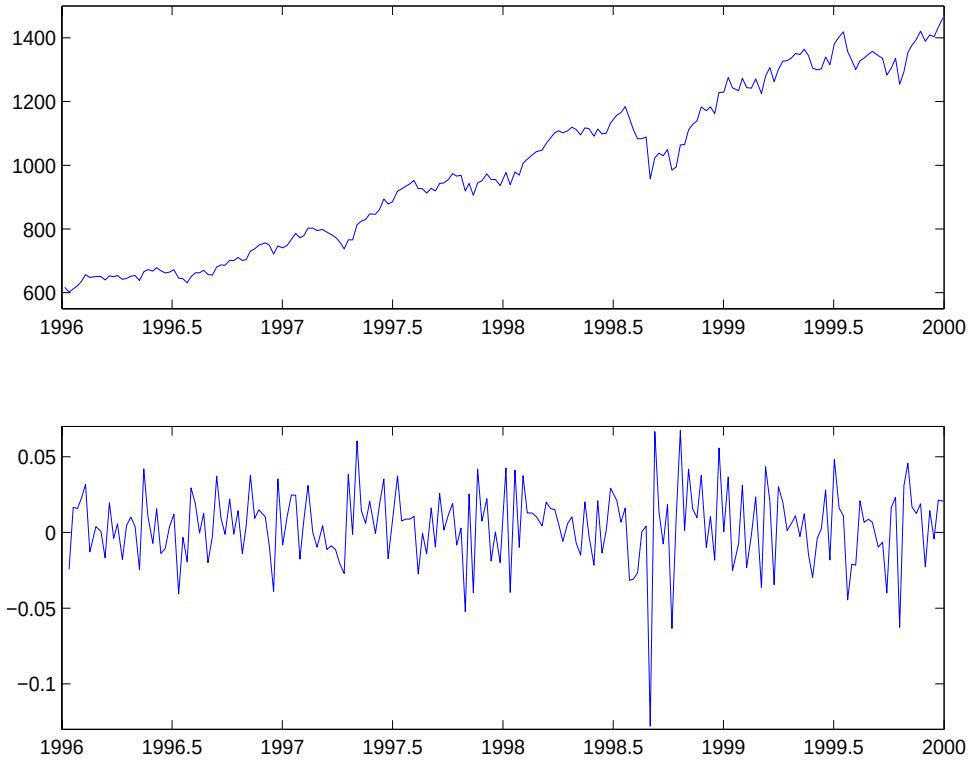

Figure 1: Time Series of S&P 500 Index and Index Returns

Note: The upper panel plots the time series of index prices and the lower panel plots the time series of index returns.

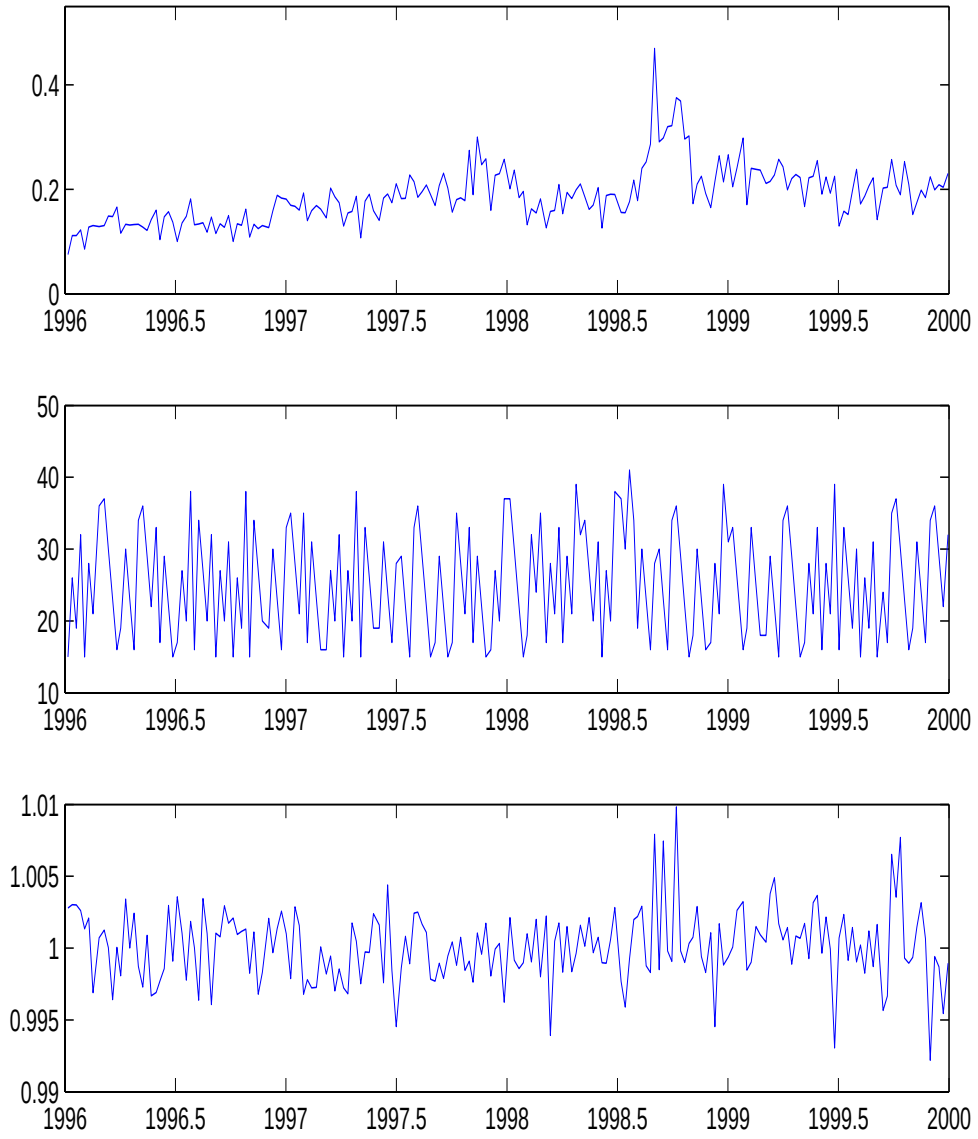

Figure 2: Implied Volatility, Maturity and Moneyness of ATM-SM Call Options

Note: The upper panel depicts the Black-Scholes implied volatilities of the constructed at-the-money short maturity call options, the middle and bottom panels plot the maturity (in days) and the moneyness (S/K) of the ATM-SM call options, respectively.

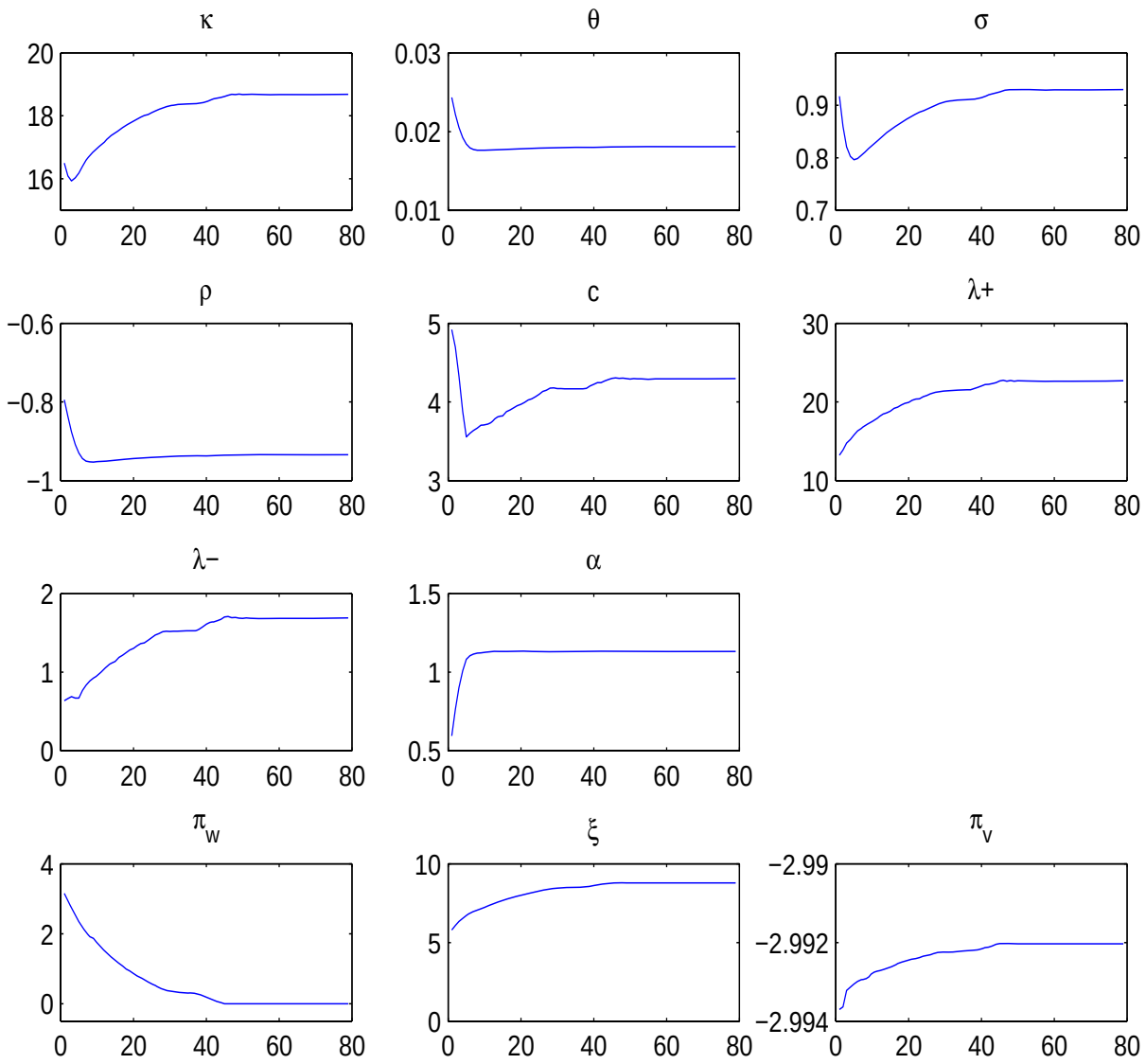

Figure 3: Convergence of Iterative Joint CCF-CGMM Estimation in LTS-SV

Note: Figure plots the convergence of parameter estimation in LTS-SV. The total number of iterations is 80. In each iteration, there are 10 internal iterations for estimation of the risk-neutral parameters and the risk-premium parameters alternately.

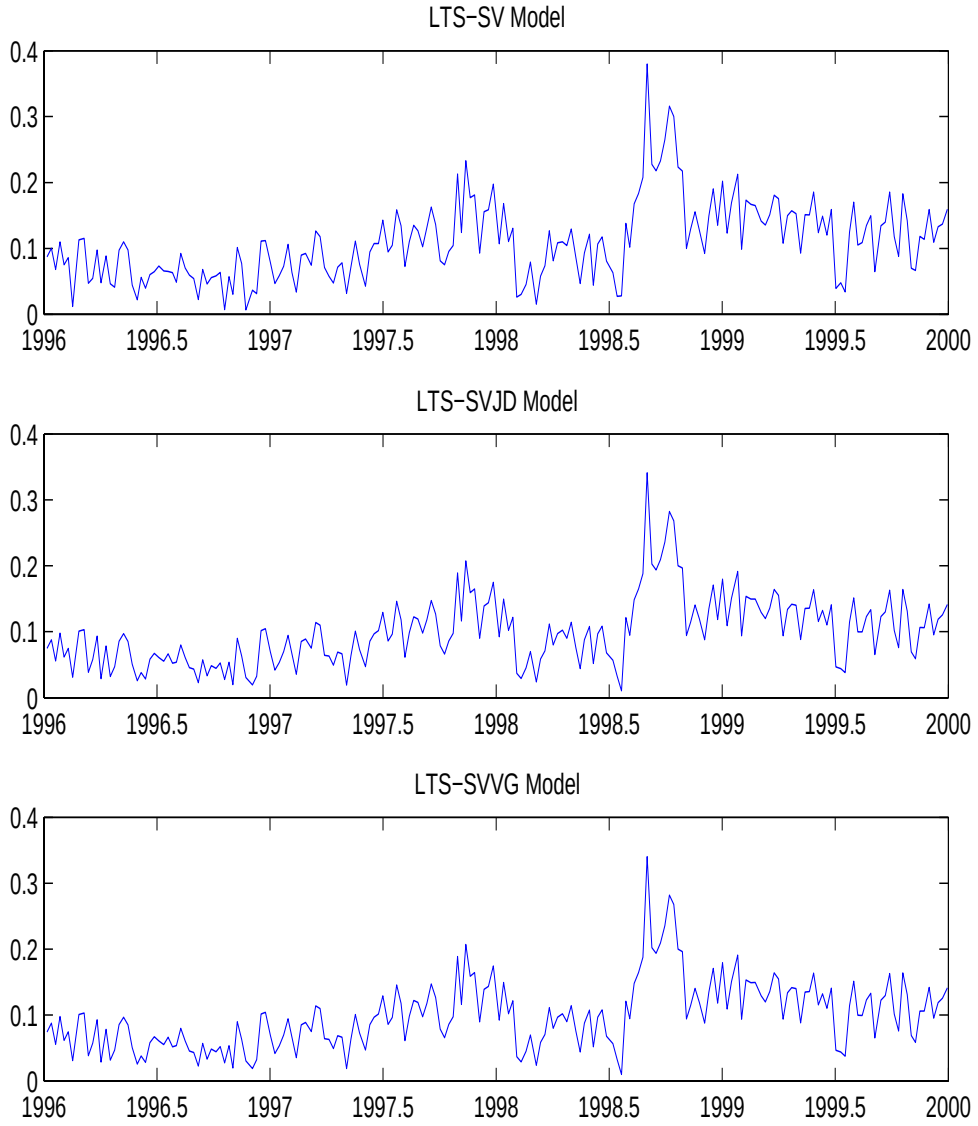

Figure 4: Estimated Square-Root Variance Rates

Note: Figure plots the estimated variance rates (square-root) for LTS-SV (upper panel), LTS-SVJD (middle panel) and LTS-SVVG (lower panel).

Chapter 2

Sequential Bayesian Analysis of Time-Changed Infinite Activity Derivatives Pricing Models*

Abstract

This paper investigates the time-changed infinite activity derivatives pricing models from the sequential Bayesian perspective. Brownian subordination property of the infinite activity Lévy process provides convenience to apply Bayesian filtering methods. I propose a sequential Monte Carlo method with the proposal density generated by the unscented Kalman filter. This approach overcomes to the large extent the problem of particle impoverishment inherent to the conventional particle filter and is more robust in real data applications. Simulation studies and real applications indicate that the underlying alone can't capture the dynamics of states and by including the derivatives in observations, the precision of state filtering gets improved dramatically, and the algorithm proposed can also effectively capture jumps realized by the infinite activity Lévy process. The joint identification of the diffusion, stochastic volatility and infinite activity jumps can be achieved using both the underlying and derivatives data.

1 Introduction

The infinite activity Lévy process can generate an infinite number of jumps within any finite time interval, capturing not only the rarely happened large jumps but also the frequently occurred small jumps. Because of this flexible feature, the infinite activity financial models have been increasingly attracting much attention and becoming popular both in academics and practice. However, the inference approaches for these models are still underdeveloped. Recently, a couple

*I am grateful to Nick Polson, Carlo Favero, Fulvio Ortu, Pietro Muliere, Claudio Tebaldi and the participants at Bocconi Finance and Statistics Seminars, the Fifth World Congress of Bachelier Finance Society (2008), Sveriges Riksbank State-Space Modeling Workshop for Financial and Economic Time Series, X Workshop on Quantitative Finance, and Conference on Recent Developments in Financial Econometrics for helpful comments.

of batch estimation methods have been developed and some of them have been approved to be powerful. To just name a few, Bakshi, Carr and Wu (2008) develop a maximum likelihood estimation method to study the time-changed Lévy model; Li, Favero and Ortu (2008) propose an iterative method combined with the characteristic function based continuous GMM also to study the time-changed Lévy models; Li, Wells and Yu (2008) apply MCMC method to estimate the Lévy jump stochastic volatility models; and Li (2008) investigates the double stochastic volatility models with Bayesian approaches.

However, in financial applications, we are more interested in sequentially estimating the models. People in the market have to estimate and forecast latent factors in real-time. We need to estimate and forecast volatility for option pricing and risk management whenever a new observation arrives and we also want to identify/disentangle different factors which influence the stock price dynamics. Furthermore, sequential estimation dramatically reduces computational cost with comparison to batch estimation, which has to restart estimation procedure every time new data arrive. There are already some works on sequential Bayesian analysis for the stochastic volatility and jump-diffusion models (Johannes, Polson and Stroud 2008; Stroud, Muller and Polson 2003). But for the infinite activity Lévy models, few works have been done yet.

If we can transform the infinite activity derivative pricing models into the dynamic state-space model (DSSM) framework, the sequential estimation via Bayesian filtering may be applied. DSSM can be written with the following form:

$$y_t = H(x_t, \Theta, w_t), \tag{1}$$

$$x_t = F(x_{t-1}, \Theta, v_t), \tag{2}$$

where the observation y_t is assumed to be conditionally independent given the state x_t with the distribution $p(y_t|x_t)$; the unobserved state x_t is modeled as a Markov process with the initial distribution $p(x_0)$ and transition law $p(x_t|x_{t-1})$; w_t and v_t are independent observation noise and state noise with mean 0 and variance R_w and R_v respectively and Θ is a set of static parameters. H and F are two vector-valued functions, which are potentially time-varying. Based on the past and current observations, filtering is a process of estimating system's current state, that is, finding the posterior distribution $p(x_t|y_{1:t})$, from which inference could be made. From Bayes'

rule, the posterior distribution $p(x_t|y_{1:t})$ could be computed through

$$p(x_t|y_{1:t}) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1})}. \tag{3}$$

The calculation and/or approximation of prior $p(x_t|y_{1:t-1})$, likelihood $p(y_t|x_t,)$ and evidence $p(y_t|y_{1:t-1})$ is the essence of Bayesian filtering and inference. If the parameter Θ is unknown, it needs to be estimated as well and we study the joint posterior distribution of state and parameter $p(x_t, \Theta|y_{1:t})$.

Lévy processes usually haven't analytically closed-form probability density functions and even if they have, often in very complicated forms. Fortunately, most of the infinite activity Lévy processes we meet in finance can be constructed through Brownian subordination, that is, time-changing a Brownian motion with drift using an independent positive increasing Lévy process. We could thus completely forget the complicated Lévy processes themselves and instead work with their subordinators. The representation of Brownian subordination for the infinite activity Lévy processes provides us convenience for simulation and computation of the likelihood and the Bayesian sequential Monte Carlo method then becomes feasible to be applied to the time-changed infinite activity models.

The previous study mainly focuses on how to use the underlying security to implement filtering. As a matter of fact, in the financial markets we usually observe two sets of financial securities: except the underlying security S_t, there also exist the derivatives O_t contingent on this underlying. The derivatives reflect the expectation of future market movements and many research find that they contain rich information about the evolution of state like volatility. Thus in this paper, I make an extension and let the observation y_t include these two sets of financial securities $y_t = \{S_t, O_t\}$. The simulation studies and real applications in section 4 and 5 show that the underlying alone can't capture the dynamics of states accurately and by including derivatives in observations, the precision of state filtering is improved dramatically.

If $H(\cdot)$ and $F(\cdot)$ are linear functions and if Gaussian distributions are assumed for x_t, w_t and v_t, the well-known Kalman filter could be applied and the optimal solution could be obtained. Unfortunately, the infinite activity derivative pricing models we meet in finance are all nonlinear and non-Gaussian. For recursive estimation of these models, the sequential Monte Carlo methods (also known as particle filters) are good choices. Particle filters are formally established by Gordon, Salmond and Smith (1993) and further developed by many others. The basic idea of

particle filters is to represent the distributions of all random variables with a number of particles (samples), directly drawn from the state space and update them using the sequential importance sampling and resampling method. One critical problem in implementation of particle filters is how to optimally design a proposal density in order to avoid sample degeneracy. The conventional particle filter (Gordon et al. 1993; Isard and Blake 1996) take the state transition law as the proposal density for its easiness of implementation. Since this proposal density doesn't take into account the latest observation, the algorithm is easy to be degenerate. The introduction of the leverage effect and jumps realized by the infinite activity process in continuous-time financial models makes the situation even worse. There are two ways to design a better proposal density. The first one improves the algorithm efficacy by using an auxiliary variable (Pitt and Shephard, 1999) and the other addresses the degeneracy problem directly by incorporating the latest observation in the proposal. In this paper I follow the second way and propose a sequential Monte Carlo method with the proposal density generated by the unscented Kalman filter based on the idea of van der Merwe, de Freitas, Doucet and Wan (2000).

The unscented Kalman filter (UKF), which is recently developed in the field of Engineering (Julier and Uhlman 1997; Wan and van der Merwe 2001) uses deterministic sampling approaches to approximate sufficient statistics. The idea behind this approach is that in order to estimate the state information after a nonlinear transformation, it is favorable to approximate the probability distribution directly instead of approximating nonlinear functions. The unscented Kalman filter overcomes to the large extent the pitfalls inherent to other nonlinear Kalman filters such as the extended Kalman filter and improves estimation accuracy and robustness without increasing computational cost. I apply UKF to each particle for update and prorogation. Even though the Gaussian assumption is not realistic for the infinite activity derivatives pricing models, it is less a problem to generate each individual particle with the distinct mean and variance. Furthermore, since UKF approximates the posterior mean and covariance up to the second order and the third for Gaussian prior, the nonlinearity of the system is well preserved.

I conduct simulation studies and real data applications to see the performance of the algorithm and at the same time answer the following questions. The first is whether it is helpful for the state filtering to incorporate the derivatives in the observation set. I investigate the information content contained in the underlying prices and derivative prices. The second to be addressed is whether the algorithm proposed in this paper can capture jumps (small and large) realized by the infinite activity Lévy processes and accommodate the leverage effect which usu-

ally render difficulty for other particle filters. And finally I investigate whether it is feasible to jointly identify the diffusion, stochastic volatility and infinite activity jumps. The joint identification is important for derivative pricing and risk management since the tail behavior of the underlying's distribution is mainly determined by jumps, especially by the large jumps.

This paper is the first to fully investigate the time-changed infinite activity derivatives pricing models from the sequential Bayesian perspective and also one of the first to apply the unscented filtering methods to finance. It is in parallel with Johannes et al. (2008) which studies the jump-diffusion stochastic volatility models with the conventional and auxiliary particle filters. The algorithm proposed in this paper performs better and is more robust than the conventional particle filter, which can not effectively capture the dynamics of states in the time-changed infinite activity Lévy models, especially in real data applications.

The rest of the paper is organized as follows. Section 2 builds the infinite activity derivatives pricing model and constructs corresponding state-space representation; Section 3 discusses the sequential Monte Carlo method with the proposal density generated by UKF; Section 4 conducts simulation studies with Heston Stochastic Volatility model, Variance Gamma model and Normal Inverse Gaussian model; Section 5 is applications using S&P 500 index and index options; And finally Section 6 concludes the paper. Proof and supplemental materials are given in Appendices.

2 Dynamic State-Space Framework

2.1 Infinite Activity Lévy Processes and Brownian Subordination

Under a given probability space $(\Omega, \mathcal{F}, P)$ and the complete filtration $\mathcal{F}_t$, a Lévy process X_t is the *cadlag* stochastic process with independent and stationary increments and $X_0 = 0$. By the Lévy-Kintchine theorem, the characteristic function of X_t has the following form

$$\phi_X(u) = E[e^{iuX_t}] = e^{-t\psi_X(u)}, \tag{4}$$

where $\psi_X(u)$ is called the characteristic exponent

$$\psi_X(u) = -iu\mu + \frac{1}{2}u^2\sigma^2 + \int_{-\infty}^{\infty}(1 - e^{iux} + iux1_{|x|<1})\mathrm{v}(x)dx, \tag{5}$$

with $(\mu, \sigma, \mathrm{v})$ being the characteristic triplet. The Lévy density $\mathrm{v}(x)$ measures the arrival rate of jumps with size x defined on R^0 (real line without zero). A Lévy process exhibits infinite activity when

$$\int_{R^0} \mathrm{v}(x)dx = \infty, \tag{6}$$

otherwise it is the finite activity process. The infinite activity Lévy process can generate an infinite number of jumps at any finite time interval, whereas the finite activity process can only generate a finite number of jumps within any finite time interval.

If a Lévy process can be constructed through subordinating a Brownian, its characteristic function, probability density and Lévy density could be derived from those of subordinator. This property makes foundation for the simulation of a Lévy process and also provides convenience in construction of the state-space model representation.

LEMMA: Suppose that a Lévy process X_t can be represented by subordinating a Brownian motion with drift

$$X_t = \omega \mathcal{S}_t + \eta W(\mathcal{S}_t), \tag{7}$$

where W_t is a standard Brownian motion and $\mathcal{S}_t$ is an independent subordinator with the characteristic exponent ψ_S, the probability distribution μ at $t = 1$ and the Lévy density $\mathrm{v}_{\mathcal{S}}$. Then the characteristic function of X_t is given by

$$\phi_X(u) = E[e^{iuX_t}] = e^{-t\psi_S(u\omega + \frac{1}{2}iu^2\eta^2)}, \tag{8}$$

the probability density of X_t is

$$p(x) = \int_0^{+\infty} \frac{1}{\sqrt{2\pi s}} e^{-\frac{(x-\omega s)^2}{2\eta^2 s}} \mu^t(ds), \tag{9}$$

and the Lévy density $\mathrm{v}(x)$ of X_t has the form of

$$\mathrm{v}(x) = \int_0^{+\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-\omega t)^2}{2\eta^2 t}} \mathrm{v}_{\mathcal{S}}(dt). \tag{10}$$

PROOF: The Lemma can be proved through the application of Theorem 30.1 of Sato (1999) to Brownian subordination. ■

Most of infinite activity Lévy processes we meet in finance can be constructed through time-changing a Brownian motion with drift. For instance, the Generalized Hyperbolic process

(Eberlein 2001) can be represented as a time-changed Brownian motion with drift taking the generalized inverse Gaussian process (Barndorff-Nielsen 1977, 1981) as the subordinatior; the Normal Tempered Stable process (Barndorff-Nielsen and Shephard 2001) can be constructed by subordinating a Brownian motion with drift using the tempered stable process (Hougaard 1986); and as the two special cases of the Generalized Hyperbolic process and the Normal Tempered Stable process, the Variance Gamma process (Madan, Carr and Chang 1998) and Normal Inverse Gaussian process (Barndorff-Nielsen 1998) have the well-know Brownian subordination forms with subordinators being the Gamma process and inverse Gaussian process respectively. All the above subordinators have nice properties and are easy to work with. The recent work of Madan and Yor (2006) proves that CGMY process (Carr, Geman, Madan and Yor 2002) and Meixner process (Schoutens and Teugels 1998) could also be constructed by subordinating Brownian motions. Brownian subordination property provides us convenience to apply Bayesian filtering methods since we could just work on with the subordinator instead of the complicated subordinate itself.

2.2 Derivatives Pricing Model and State-Space Representation

The derivatives pricing model is built by taking into account both stochastic volatility and jumps in the dynamics of the underlying security. Jumps are modeled by the infinite activity Lévy process and stochastic volatility is introduced through time-changing the stochastic processes (Carr, Geman, Madan and Yor 2003).

Define a stopping time T_t given a nonnegative and right continuous with left limit stochastic process V_t under a probability space $(\Omega, \mathcal{F}, P)$ equipped with the complete filtration $\mathcal{F}_t$,

$$T_t = \int_0^t V_{s-} ds,$$

which is finite almost surely. Intuitively, we could think of t as calender time and T_t as business time. The variable V_t reflects the intensity of economic activity and we call it the variance rate. Stochastic volatility is generated by replacing calender time t with business time T_t. For a stochastic process $X(t)$, its time-changed counterpart is defined by

$$X_{T_t} = X(T_t)$$

A well-known positive stochastic process is the square-root CIR process of Cox, Ingersoll and Ross (1985), which is used in this paper as the variance rate process. The underlying securities are then modeled with the exponential time-changed Brownian motion and infinite activity Lévy process

$$
\begin{aligned}
S_t &= S_0 \exp\Big\{rt + \pi_W T_t^{(1)} + \Big[k_X^P(1) - k_X^Q(1)\Big] T_t^{(2)} \\
&\quad + \Big[W_{T_t^{(1)}} - \frac{1}{2} T_t^{(1)}\Big] + \Big[X_{T_t^{(2)}} - k_X^P(1) T_t^{(2)}\Big]\Big\}, && (11)\\
T_t^{(1)} &= \int_0^t V_{s-}^{(1)} ds, && (12)\\
T_t^{(2)} &= \int_0^t V_{s-}^{(2)} ds, && (13)\\
dV_t^{(1)} &= \kappa^{(1)}(\theta^{(1)} - V_t^{(1)})dt + \sigma^{(1)}\sqrt{V_t^{(1)}} dZ_t^{(1)}, && (14)\\
dV_t^{(2)} &= \kappa^{(2)}(\theta^{(2)} - V_t^{(2)})dt + \sigma^{(2)}\sqrt{V_t^{(2)}} dZ_t^{(2)}, && (15)
\end{aligned}
$$

where W_t is a Brownian motion; X_t is an infinite activity Lévy process which captures both the frequently happened small jumps and the rarely happen large jumps; $T_t^{(1)}$ and $T_t^{(2)}$ are two stochastic business time which are used to generate stochastic volatility from W_t and X_t respectively; $V_t^{(1)}$ and $V_t^{(2)}$ are variance rates to construct stochastic business time; $Z_t^{(1)}$ and $Z_t^{(2)}$ are two other Brownian motions and independent each other. $Z_t^{(1)}$ is allowed to be correlated to W_t with the correlation parameter ρt in order to accommodate the leverage effect and independent of X_t. $Z_t^{(2)}$ is independent of W_t and X_t. r is a constant risk-free interest rate; π_W is the diffusion risk premium; $k_X(1)$ is the convexity adjustment which is derived from the cumulant exponent of X_t: $k(s) \equiv \frac{1}{t} log(E[e^{sX_t}]) = -\psi_X(-is)$. The term $\pi_X \equiv k_X^P(1) - k_X^Q(1)$ captures the risk premium of jump.

Assume there exists an equivalent martingale measure Q, under which the risk-neutral model is defined as

$$
\begin{aligned}
S_t &= S_0 \exp\Big\{rt + \Big[W^Q_{T_t^{(1)}} - \frac{1}{2} T_t^{(1)}\Big] + \Big[X^Q_{T_t^{(2)}} - k_X^Q(1) T_t^{(2)}\Big]\Big\}, && (16)\\
dV_t^{(1)} &= \Big[\kappa^{(1)}(\theta^{(1)} - V_t^{(1)}) + \pi_{V^{(1)}} V_t^{(1)}\Big] dt + \sigma^{(1)}\sqrt{V_t^{(1)}} dZ_t^{(1)Q}, && (17)\\
dV_t^{(2)} &= \Big[\kappa^{(2)}(\theta^{(2)} - V_t^{(2)}) + \pi_{V^{(2)}} V_t^{(2)}\Big] dt + \sigma^{(2)}\sqrt{V_t^{(2)}} dZ_t^{(2)Q}, && (18)
\end{aligned}
$$

where $T_t^{(i)} = \int_0^t V_{s-}^{(i)} ds$ and $\pi_V^{(i)}$ are market prices of the volatility risks for $i = 1, 2$. Define $\kappa^{(i)Q} = \kappa^{(i)} - \pi_V^{(i)}$. When jumps and stochastic volatility are introduced into financial modeling,

the market is no longer complete. There may exist many different risk-neutral measures. We hope that the change of measure does not alter the structure of the stochastic processes and this could be done by means of the well-known Esscher transform, which is defined

$$\frac{dQ}{dP}\Big|_{\mathcal{F}_t} = \exp\{-\xi Y_t - k(-\xi)t\} \tag{19}$$

for a given Lévy process Y_t and a real number ξ. $k(\cdot)$ is the cumulant exponent of Y_t. Under the equivalent martingale measure, new Lévy density is just an exponential tilting of the objective one $\tilde{\mathrm{v}}(dy) = e^{-\xi y}\mathrm{v}(dy)$. Exponential tilting mainly affects big jumps and this is consistent with our understanding of financial market movements. It is large jumps that have dominant influence on option pricing and risk management. Esscher transform gives an equivalent martingale measure Q which has minimum relative entropy with respect to the objective measure P. Intuitively this means that Q is an equivalent martingale measure which is the closest to P in terms of information content (Chan 1999). Under Esscher transform, the risk premium is given by $\pi_Y = k(1) - k^Q(1)$, where $k^Q(s) = k(s-\xi) - k(-\xi)$ (Gerber and Shiu 1994).

Since the model does not admit an analytical derivative pricing formula, we have to rely on the numerical methods. Taking into account that the underlying process is correlated with the diffusion variance rate process, I firstly internalize this correlation with the approach proposed by Carr and Wu (2004) and then derive the conditional characteristic function of log return with the transform approach in the sense of Duffie, Pan and Singleton (2000).

PROPOSITION: Define a new filtration $\mathcal{G}_t$ generated by the business time sigma algebra $\mathcal{F}_{T_t}$. For the risk-neutral model defined above, the conditional characteristic function of log return $R_t = \ln(S_t/S_{t-\tau})$ has the following form:

$$\begin{aligned}\phi_R(u;t,V_{t-\tau}) &\equiv E^Q[e^{iuR_t}|\mathcal{G}_t] \\ &= e^{iurt - A(u,\tau) - B(u,\tau)V_{t-\tau}},\end{aligned} \tag{20}$$

where $A = A^{(1)} + A^{(2)}$, $B = [B^{(1)}, B^{(2)}]$, $V_{t-\tau} = [V^{(1)}_{t-\tau}, V^{(2)}_{t-\tau}]'$ and

$$\begin{aligned}A^{(i)}(u,\tau) &= \frac{\kappa^{(i)}\theta^{(i)}}{(\sigma^{(i)})^2}\Big[2\log\Big(1 - \frac{(\gamma^{(i)} - \kappa^{(i)*})(1-e^{-\gamma^{(i)}\tau})}{2\gamma^{(i)}}\Big) + (\gamma^{(i)} - \kappa^{(i)*})\tau\Big], \\ B^{(i)}(u,\tau) &= \frac{2\varphi^{(i)}(u)(1-e^{-\gamma^{(i)}\tau})}{2\gamma^{(i)} - (\gamma^{(i)} - \kappa^{(i)*})(1-e^{-\gamma^{(i)}\tau})},\end{aligned}$$

$$
\begin{aligned}
\gamma^{(i)} &= \sqrt{(\kappa^{(i)*})^2 + 2(\sigma^{(i)})^2\varphi^{(i)}(u)}, \\
\varphi^{(1)}(u) &\equiv \varphi_W(u) = \frac{1}{2}(iu + u^2), \\
\varphi^{(2)}(u) &\equiv \varphi_X(u) = \psi_X^Q(u) + iuk_X^Q(1), \\
\kappa^{(1)*} &= \kappa^{(1)Q} - iu\rho\sigma^{(1)}, \\
\kappa^{(2)*} &= \kappa^{(2)Q},
\end{aligned}
$$

where $i = 1, 2$, ψ_X^Q is the characteristic exponent of the infinite activity Lévy process under the risk-neutral measure and $k_X^Q(\cdot)$ is its cumulant exponent. Note that now the filtration is augmented by the stochastic business time and becomes $\mathcal{G}_t$.

PROOF: see Appendix A. ■

With this conditional characteristic function of log return, we could compute European type derivative price with the fast Fourier transform method (Carr and Madan 1997), especially the more efficient fractional fast Fourier transform approach (FRFT) proposed by Chourdakis (2005).

This time-changed infinite activity model is very general and contains many models we have met in literature. Bakshi et al. (2008) and Li et al. (2008) investigate this kind model with X_t being the tempered stable process from batch estimation perspective. Li et al. (2008) also investigate the double-jump model by introducing a jump part in the variance rate process and find that whenever the underlying is modeled with the infinite activity Lévy process, the introduction of jump in the variance rate process is not critically important. Therefore, in this paper I don't consider volatility jump.

Our aim is to estimate the model with Bayesian filtering methods. To this end, we need construct an appropriate state-space representation. As we see from the above, Brownian subordination property of the infinite activity Lévy process provides us extra convenience to construct state-space representation and as long as the subordinator is well-behaved, the sequential Monte Carlo method could be easily conducted.

First of all, we should de-correlate the underlying and variance rate for the applicability of Bayesian filtering. For the time-changed Brownian motion, we have

$$
W_{T_t^{(1)}} \stackrel{d}{=} \int_0^t \sqrt{V_t^{(1)}} dW_t, \tag{21}
$$

where d indicates that the equality holds in distribution. With this property and the fact $[dW_t, dZ_t^{(1)}] = \rho dt$, I rewrite the underlying process (11) and the variance rate process (14) into the following forms

$$\begin{aligned}
\ln S_t &= \ln S_0 + rt + \Big(\pi_W - \frac{1}{2}\Big)T_t^{(1)} + \int_0^t \sqrt{V_t^{(1)}}\Big(\sqrt{1-\rho^2}dW_t^* + \rho dZ_t^{(1)}\Big) \\
&\quad + \Big(\pi_X - k_X^P(1)\Big)T_t^{(2)} + X_{T_t^{(2)}}, \qquad (22) \\
\sqrt{V_t^{(1)}}dZ_t^{(1)} &= \frac{1}{\sigma^{(1)}}\Big[dV_t^{(1)} - \kappa^{(1)}(\theta^{(1)} - V_t^{(1)})dt\Big], \qquad (23)
\end{aligned}$$

where now W_t^* and $Z_t^{(1)}$ are independent. Note that the Brownian motion W_t^* in (22) is different from that in (11). Putting (23) into (22), we obtain

$$\begin{aligned}
\ln S_t &= \ln S_0 + rt + \Big(\pi_W - \frac{1}{2}\Big)T_t^{(1)} + \frac{\rho}{\sigma^{(1)}}\int_0^t \Big[dV_t^{(1)} - \kappa^{(1)}(\theta^{(1)} - V_t^{(1)})dt\Big] \\
&\quad + \int_0^t \sqrt{1-\rho^2}\sqrt{V_t^{(1)}}dW_t^* + \Big(\pi_X - k_X^P(1)\Big)T_t^{(2)} + X_{T_t^{(2)}}, \qquad (24)
\end{aligned}$$

which can be regarded as one of observation equations. Now we can see the observation noise in (24) and state noise in (14) are independent.

I then take advantage of Brownian subordination property and regard jump as one of states

$$X_{T_t^{(2)}} = \omega \mathcal{S}_{T_t^{(2)}} + \eta \tilde{W}(\mathcal{S}_{T_t^{(2)}}), \qquad (25)$$

where $\mathcal{S}_{T_t^{(2)}}$ is the subordinator under business time and also taken as one state $\mathcal{S}_t = g_{\mathcal{S}}(T_t^{(2)}; \Xi)$ with Ξ being parameters of the distribution $g_{\mathcal{S}}$.

Finally, discretizing the model with the small time interval τ and meanwhile taking into consideration the derivatives, we have the following state-space representation

Measurements:

$$\begin{aligned}
\ln S_t &= \ln S_{t-\tau} + \Big(r - \rho\frac{\kappa^{(1)}\theta^{(1)}}{\sigma^{(1)}}\Big)\tau + \Big[\frac{\rho}{\sigma^{(1)}}(\kappa^{(1)}\tau - 1) + (\pi_W - \frac{1}{2})\tau\Big]V_{t-\tau}^{(1)} \\
&\quad + \frac{\rho}{\sigma^{(1)}}V_t^{(1)} + \Big(\pi_X - k_X^P(1)\Big)\tau V_t^{(2)} + X_t + \sqrt{1-\rho^2}\sqrt{\tau V_{t-\tau}^{(1)}}w_t, \\
y_t^O &= f(S_t, V_t, \Theta) + \epsilon_t^O,
\end{aligned}$$

States:

$$
\begin{aligned}
\begin{pmatrix} V_t^{(1)} \\ V_{t-\tau}^{(1)} \end{pmatrix} &= \begin{pmatrix} \kappa^{(1)}\theta^{(1)}\tau \\ 0 \end{pmatrix} + \begin{pmatrix} 1-\kappa^{(1)}\tau & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} V_{t-\tau}^{(1)} \\ V_{t-2\tau}^{(1)} \end{pmatrix} + \begin{pmatrix} \sigma^{(1)}\sqrt{\tau V_{t-\tau}^{(1)}} \\ 0 \end{pmatrix} z_t^{(1)}, \\
V_t^{(2)} &= \kappa^{(2)}\theta^{(2)}\tau + \left(1-\kappa^{(2)}\tau\right)V_{t-\tau}^{(2)} + \sigma^{(2)}\sqrt{\tau V_{t-\tau}^{(2)}}z_t^{(2)}, \\
X_t &= \omega \mathcal{S}_t + \eta\sqrt{\mathcal{S}_t}\tilde{w}_t, \\
\mathcal{S}_t &= g_{\mathcal{S}}(\tau V_t^{(2)}; \Xi),
\end{aligned}
$$

where the derivatives are assumed to be collected with measurement errors $\epsilon_t^O \rightarrow \mathcal{N}(0, \sigma_O^2)$, independent of w_t, z_t and $\tilde{w}_t$, which are also independent standard normal white noises; S_t and y_t^O are the discretely observed underlying securities and derivative prices; $f(\cdot)$ is the theoretical option price computed from the model; the variance rates $V_t^{(1)}$, $V_t^{(2)}$, jump size X_t and jump time $\mathcal{S}_t$ are regarded as states. For the small time interval τ, like daily frequency or higher, discretization of the continuous-time models does not introduce significant bias in estimation (Eraker, Johannes and Polson 2003; Johannes et al. 2008). With this sate-space representation, we could apply sequential Monte Carlo method.

3 Unscented Sequential Bayesian Estimation

The infinite activity derivatives pricing model introduced in section 2 is obviously nonlinear and non-Gaussian. I propose the application of sequential Monte Carlo methods (particle filters) which approximate the posterior distribution of states by a set of weighted particles (samples) without making any explicit assumptions and thus can be applied to any nonlinear and non-Gaussian systems. Since introduced by Gordon et al. (1993), particle filters have been successfully used in many fields, but their application to financial problems has not attracted much attention. In order to improve efficiency upon the conventional particle filter and avoid the sample impoverishment, I propose a sequential Monte Carlo method with the proposal density designed by UKF. In the rest of this section, I firstly introduce the unscented Kalman filter and then discuss the unscented sequential Monte Carlo method.

3.1 Unscented Kalman Filter

The unscented Kalman filter is proposed to improve efficiency of Gaussian approximate inference algorithms. It is a straightforward application of the scaled unscented transformation, which uses the so-called sigma-points to cover and propagate the information of data (see Appendix B). The scaled unscented transformation can approximate posterior mean and covariance of a random variable undergoing a nonlinear transformation with accuracy up to the second order and the third order for Gaussian prior (Julier and Uhlman 1997). Based on the scaled unscented transformation, the unscented Kalman filter is derived. The unscented Kalman filter does not explicitly approximate or linearize the nonlinear observation and state models. It uses the true nonlinear models and updates state variables through a set of deterministic sigma points generated by the unscented transformation.

The unscented Kalman filter is implemented through the following procedure. Suppose for the moment that DSSM (1) (2) are nonlinear Gaussian. We firstly concatenate the state, observation noise and state noise at time $t-1$

$$x^e_{t-1} = \begin{bmatrix} x_{t-1} \\ w_{t-1} \\ v_{t-1} \end{bmatrix}, \tag{26}$$

whose dimension is $L = L_x + L_w + L_v$ with L_x, L_w and L_v being dimensions of state, observation noise and state noise respectively and whose mean and covariance are

$$\hat{x}^e_{t-1} = E\Big[x^e_{t-1}\Big], \tag{27}$$

$$P^e_{t-1} = \begin{bmatrix} P_{x_{t-1}} & 0 & 0 \\ 0 & R_w & 0 \\ 0 & 0 & R_v \end{bmatrix}. \tag{28}$$

We then use the scaled unscented transformation to form a set of $2L+1$ sigma points

$$\chi^e_{t-1} = \begin{bmatrix} \hat{x}^e_{t-1} & \hat{x}^e_{t-1} + \sqrt{(L+\lambda)P^e_{t-1}} & \hat{x}^e_{t-1} - \sqrt{(L+\lambda)P^e_{t-1}} \end{bmatrix} \tag{29}$$

and corresponding weights

$$
\begin{aligned}
w_0^{(m)} &= \frac{\lambda}{L+\lambda}, && (30)\\
w_0^{(c)} &= \frac{\lambda}{L+\lambda} + (1-\alpha^2+\beta), && (31)\\
w_i^{(m)} &= w_i^{(c)} = \frac{1}{2(L+\lambda)}, \quad i = 1,\ldots,2L && (32)
\end{aligned}
$$

where superscripts (m) and (c) indicate that the weights are for the construction of posterior mean and posterior covariance respectively; $\lambda = \alpha^2(L+\kappa) - L$ is a scaling parameter; κ, α and β are other unscented transformation parameters. Typical values for κ, α and β are 0, 10^{-3} and 2 respectively. These values should suffice for most purposes.

With sigma points constructed like this, we implement Kalman filter. For the time update

$$
\begin{aligned}
\chi^x_{t|t-1} &= F(\chi^x_{t-1}, \chi^v_{t-1}),\\
\hat{x}_t^- &= \sum_{i=0}^{2L} w_i^{(m)} \chi^x_{i,t|t-1},\\
P^-_{x_t} &= \sum_{i=0}^{2L} w_i^{(c)} (\chi^x_{i,t|t-1} - \hat{x}_t^-)(\chi^x_{i,t|t-1} - \hat{x}_t^-)',
\end{aligned}
$$

and for the measurement update

$$
\begin{aligned}
\mathcal{Y}_{t|t-1} &= H(\chi^x_{t|t-1}, \chi^w_{t|t-1}),\\
\hat{y}_t^- &= \sum_{i=0}^{2L} w_i^{(m)} \mathcal{Y}_{i,t|t-1},\\
P^-_{y_t} &= \sum_{i=0}^{2L} w_i^{(c)} (\mathcal{Y}_{i,t|t-1} - \hat{y}_t^-)(\mathcal{Y}_{i,t|t-1} - \hat{y}_t^-)',\\
P_{x_t y_t} &= \sum_{i=0}^{2L} w_i^{(c)} (\chi^x_{i,t|t-1} - \hat{x}_t^-)(\mathcal{Y}_{i,t|t-1} - \hat{y}_t^-)',\\
K_t &= P_{x_t y_t} (P^-_{y_t})^{-1},\\
\hat{x}_t &= \hat{x}_t^- + K_t(y_t - \hat{y}_t^-),\\
P_{x_t} &= P^-_{x_t} - K_t P^-_{y_t} K_t'.
\end{aligned}
$$

We then obtain the posterior mean $\hat{x}_t$ and posterior covariance P_{x_t} of the state x_t.

3.2 Unscented Sequential Monte Carlo Method

Our actual aim is to recursively estimate the posterior distribution $p(x_{0:t}|y_{1:t})$, the filtering density $p(x_t|y_{1:t})$ and/or expectations of any interests $E[f(x_{0:t})]$ given that DSSM (1) (2) are nonlinear and non-Gaussian. With Bayes' rule, we have the following recursive formula

$$p(x_{0:t}|y_{1:t}) = p(x_{0:t-1}|y_{1:t-1})\frac{p(y_t|x_t)p(x_t|x_{t-1})}{p(y_t|y_{1:t-1})}, \tag{33}$$

which is usually in a high non-analytical form because of the non-tractable denominator in (33). To make inference, we thus have to rely on approximation.

Under some regularity conditions, particle filters approximate the posterior state density $p(x_{0:t}|y_{1:t})$ with the empirical point-mass estimate $\hat{p}(x_{0:t}|y_{1:t})$,

$$\hat{p}(x_{0:t}|y_{1:t}) = \sum_{i=1}^{N} \tilde{w}_t^{(i)} \delta(x_{0:t} - x_{0:t}^{(i)}), \tag{34}$$

where $\tilde{w}_t^{(i)}$ is the normalized weight for each particle and $\delta(\cdot)$ denotes Dirac delta function. As the number of particles N increases, the accuracy of this approximation improves and the law of large number guarantees the convergence. Since it is usually impossible to directly sample from posterior density function, we could apply the importance sampling method and alternatively sample from a known and easy-to-sample proposal density function $\pi(x_{0:t}|y_{1:t})$. Given N samples $\{x_t^{(i)}, i = 1, 2, \cdots, N\}$ from the proposal density, the expectation $E[f(x_{0:t})]$ can be estimated by

$$\hat{E}[f(x_{0:t})] = \sum_{i=1}^{N} \tilde{w}_t^{(i)} f(x_{0:t}^{(i)}), \tag{35}$$

$$\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^{N} w_t^{(j)}}, \tag{36}$$

$$w_t^{(i)} = \frac{p(y_{1:t}|x_{0:t}^{(i)})p(x_{0:t}^{(i)})}{\pi(x_{0:t}^{(i)}|y_{1:t})}, \tag{37}$$

where w_t is the unnormalized weight which can be recursively updated

$$w_t = w_{t-1}\frac{p(y_t|x_t)p(x_t|x_{t-1})}{\pi(x_t|x_{0:t-1}, y_{1:t})}. \tag{38}$$

See Appendix C for a derivation.

The most frequently used proposal density is the transition density of the state. But in

implementation this particle filter usually meets a problem of sample depletion, that is, after a few steps of updating only a small number of particles have significant weights. To overcome this problem, a resampling stage is customarily conducted. However, we find that the main cause of sample depletion is the failure of moving particles to the high likelihood regions and this failure directly stems from the selection of the poor proposal density. Note that there are two critical places involving the proposal density: first, particles are drawn from the proposal density and second, the proposal density is used to calculate each particle's weight. Thus, a good choice of the proposal density $\pi(x_t|x_{0:t-1}, y_{1:t})$ plays paramount role in particle filters.

In principle, there are many choices of the proposal density as long as its support includes that of the posterior density and it is easy to sample from. However, it is proved that the optimal proposal density is

$$\pi(x_t|x_{0:t-1}, y_{1:t}) \equiv p(x_t|x_{t-1}, y_t), \tag{39}$$

which minimizes the variance of the importance weights (Doucet 2000). $p(x_t|x_{t-1}, y_t)$ doesn't usually have an analytical form and we don't know what the true value of x_{t-1} is as well. All we know about x_{t-1} is that the information reflected by x_{t-1} is summarized in $p(x_{t-1}|y_{1:t-1})$. Thus, the optimal proposal density could be approximated by

$$\int p(x_t|x_{t-1}, y_t)p(x_{t-1}|y_{1:t-1})dx_{t-1} = p(x_t|y_{1:t}), \tag{40}$$

to which the Kalman filter generated Gaussian approximation $\pi_N(x_t|y_{1:t})$ could be applied. The usually used nonlinear Kalman filter is the extended Kalman filter. But for the highly nonlinear system, the extended Kalman filter performs very poorly. An alternative is the unscented Kalman filter discussed in section 3.1, which is a derivative-free approach and proved to be more efficient than the extended Kalman filter. Therefore, in this paper I am going to use UKF to design the proposal density, that is, I approximate the optimal proposal density with a Gaussian density generated by UKF for each particle. This proposal density takes into account the latest observation and can capture the complicated structure such as multi-modalities, skewness and other high-order moments. Note that even though the Gaussian assumption is not realistic for the optimal density $p(x_t|x_{t-1}, y_t)$, it is less a problem to generate each individual particle with the distinct mean and variance. Furthermore, since UKF approximates the mean and covariance of the posterior up to the second order and the third for Gaussian prior, the nonlinearity of the system is well preserved. A complete algorithm is given in Appendix C.

4 Simulation Studies

In this section, I firstly conduct a simulation study with Heston stochastic volatility (SV) model in subsection 4.1. Heston SV model can be obtained through suppressing the jump part and related components from the general model discussed in section 2. I then move to derivative pricing models built on the Variance Gamma (VG) process and the Normal Inverse Gaussian (NIG) process respectively in subsection 4.2 and subsection 4.3. Both VG and NIG processes can be constructed through Brownian subordination using the gamma process $G(t;1,v)$ and the inverse Gaussian process $IG(t;1,v)$ respectively, which are reparameterized to have unit mean rate and variance rate v. Their characteristic exponents can be derived as follows from the Lemma

$$\psi_{VG}(u) = \frac{1}{v}\log(1 - iu\omega v + u^2\frac{v\eta^2}{2}), \tag{41}$$

$$\psi_{NIG}(u) = \frac{1}{v}(\sqrt{1 - 2iu\omega v + u^2\eta^2 v} - 1). \tag{42}$$

Indeed, they are two of the most popular infinite activity Lévy processes in finance and can suffice for many financial modeling purposes. Even though both of them are infinitely active, VG process is of finite variation whereas NIG process exhibits infinite variation.

For simplicity, in the simulation studies I assume zero risk premia and set variance rate $V_t^{(2)} = 1$ in the infinite activity jump models. Under these settings, for Heston model we have parameters $\Theta = \{\kappa^{(1)}, \theta^{(1)}, \sigma^{(1)}, \rho\}$ and states $x_t = V_t^{(1)}$; and for VG and NIG models we have parameters $\Theta = \{\omega, \eta, v, \kappa^{(1)}, \theta^{(1)}, \sigma^{(1)}, \rho\}$ and states $x_t = \{V_t^{(1)}, X_t, \mathcal{S}_t\}$. Extension to the general case as in subsection 2.2 is straightforward.

4.1 SV Model

I firstly study a stochastic volatility model without jump component: Heston Stochastic Volatility model. I simulate a sample of daily data with size 500 using true parameters $\Theta^* = \{6.00,\ 0.025,\ 0.30,\ -0.50\}$ and initial values $S_0 = 100$ and $V_0 = 0.025$. The simulated options data are those of at-the-money short maturity ones with strike equal to stock price and maturity equal to one month. I assume the risk-free interest rate is known and constant with value 6% and option prices are contaminated by the measurement noise $\epsilon_t^O \rightarrow N(0, 0.005)$. Figure 1 presents the simulated stock prices, returns, at-the-money options and corresponding

Black-Scholes implied volatility.

— Figure 1 around here —

A feature of this paper is that I use both the underlying and derivatives for state filtering. I apply the unscented sequential Monte Carlo method discussed in section 3 jointly using the stock price and option data with the number of particles 200, which is in stark contrast to the conventional particle filter and auxiliary particle filter where at least ten thousands of particles are usually used. Figure 2 upper panel presents the filtered and true volatility. The proposed algorithm can completely capture the volatility dynamics in Heston model even though the leverage effect is clearly introduced.

— Figure 2 around here —

We also want to compare the performance of our algorithm and the conventional particle filter. Figure 2 middle panel gives the result of joint filtering using the conventional particle filter. We find that its performance is nearly the same as that of the proposed algorithm except that it needs the large number of particles (2000). This is not surprise since the volatility in Heston model is modeled with the diffusion process and can not have large change at a short time. Using the transition law as the proposal density is hardly harmful for the filtering. But this is not the case in the jump models we will discuss in the following subsections.

It is desirable to investigate the information content contained in the underlying and derivatives. We would like to see whether it is helpful for state filtering to include derivatives in observations. For this purpose, I filter the model with the underlying data only and compare the results with those obtained above. Figure 2 lower panel presents the filtered volatility. We find that it obviously deviates from the true. This indicates that the derivatives contain richer information on volatility than the underlyings.

4.2 VG Model

We now move to the jump model. In this subsection, I study the time-changed variance-gamma derivatives pricing model. The model is obtained by specifically setting X_t in (11) being the variance gamma process. In the state-space representation, $g_{\mathcal{S}}$ is now a gamma random variable with mean τ and variance $\upsilon\tau$.

I simulate 500 daily data with initial values $S_0 = 100$, $V_0 = 0.025$, $X_0 = 0$, $\mathcal{S}_0 = 0$ and true parameters $\Theta^* = \{-0.04,\ 0.30,\ 0.05,\ 6.00,\ 0.025,\ 0.30,\ -0.50\}$. The simulated options

data are again those of at-the-money short maturity ones with strike equal to stock price and maturity equal to one month. The other assumptions on the risk-free rate and options are the same as before. Figure 3 presents the simulated stock prices, returns, at-the-money options and corresponding Black-Scholes implied volatility.

— Figure 3 around here —

Figure 4 presents the filtered square-root variance rate and jumps with the unscented sequential Monte Carlo method jointly using both the underlying and derivative prices. The top panel is the volatility filtering from which we find visually that the filtered volatility is very similar to the true. Johannes et al. (2008) find that for their filtering algorithm incorporation of the leverage effect in the jump-diffusion stochastic volatility model makes it difficult to accurately filter out states. For the algorithm proposed in this paper, I don't observe this difficulty when using both the underlying and derivative data. Furthermore, the algorithm can capture both the large and small jumps. The lower panel of Figure 4 presents the filtered jumps. Visually we can see that for the large jumps the algorithm achieves nearly perfect result and for the small jumps the result is also very satisfactory.

— Figure 4 around here —

I also run the algorithm with the underlying data alone. Figure 5 presents filtering results of the volatility and jumps. We find that the filtered jumps miss the large jump points and the filtered volatility is again away from the true. The complicated structure of the infinite activity Lévy process and the leverage effect make it very difficult to filter the volatility and jumps using the underlying data alone.

— Figure 5 around here —

As we know, in the variance gamma model, the parameter v mainly determines the jump structure. If v is large, the process generates many tiny jumps occasionally accompanied by the huge jumps which could be regarded as the event like market crash. Whereas if v is very small, most jumps happen with size around the mean τ. I study these cases and see whether the algorithm proposed can capture the rich jump structure generated by the variance gamma process. Figure 6 and Figure 7 present the true and filtered volatility and jumps for $v = 0.50$ and 0.005 respectively. In both figures, the top two panels are the simulated returns and the Black-Scholes implied volatility of the simulated options. The lower two panels are comparisons

of the filtered states to the true ones. We observe robustness of the algorithm. For both volatility and jumps filtering, the algorithm can capture them without difficulties in both cases. What's more striking is that in $v = 0.5$ case, the proposed method can successfully capture the huge jump which happens around point 390.

— Figure 6 around here —

— Figure 7 around here —

To see the method proposed in this paper is really superior to the conventional particle filter in the infinite activity jump model. I also run the conventional particle filter with the proposal density being the state transition law for the three sets of data generated before using both the underlying and derivative data. I find that there is no hope to obtain satisfied results if we use the same number of particles as before (200). I then increase the number of particles to 2000. The volatility filtering is acceptable (not reported) because of the reason we discussed before in SV model but not the jump filtering. Figure 8 presents the filtered result of jumps using the conventional particle filter. We find that it is very difficult for the conventional particle filter to capture the large jumps, especially those like market crash. For example, in $v = 0.5$ case, the conventional particle filter completely misses the largest jump around point 390 and in $v = 0.05$ case it also misses the large jumps. The inability to capture the large jumps may result in disastrous errors in derivative pricing and risk management. This is because these large jumps determine the tail behavior of the underlying distribution and therefore play a paramount role in pricing options and risk management. Furthermore, I find that in real data applications in section 5 the conventional particle filter completely loses its capacity to capture both volatility and jumps. This again proves that the unscented sequential Monte Carlo method is more robust than the conventional particle filter.

— Figure 8 around here —

4.3 NIG Model

Unlike VG process, which is of finite variation, NIG process takes on infinite variation. The time-changed NIG model is obtained by setting X_t in (11) to be the normal inverse Gaussian process. In the state-space representation, g_S is now an inverse Gaussian random variable with mean τ and variance $v\tau$. The inverse Gaussian random variable can be sampled efficiently with the approach of Michael, Schucany and Haas (1976).

As in VG model, I also simulate 500 daily data with initial values $S_0 = 100$, $V_0 = 0.025$, $X_0 = 0$, $\mathcal{S}_0 = 0$ and true parameters $\Theta^* = \{-0.04,\ 0.3,\ 0.10,\ 6.00,\ 0.025,\ 0.3,\ -0.5\}$. With these data, I implement joint state filtering with the stock price and option data. Figure 9 presents the filtered results. The top panel is the volatility filtering and the lower panel is the jumps filtering. From this figure, we again find that the unscented sequential Monte Carlo method performs very well for both volatility and jump filtering. The leverage effect and infinite activity jumps do not introduce difficulties to the algorithm.

— Figure 9 around here —

The same procedures could be applied to NIG model as we have done for VG model. Even though the results are not reported here for the save of space, the conclusion is nearly the same. Using the underlying only, we can not effectively filter states out. For the different jump structures induced by the parameter v, the proposed method can capture them satisfactorily whereas the conventional particle filter can not.

5 Applications

I continue to investigate the unscented sequential Bayesian estimation approach with the real data S&P 500 index and index options. The data are those traded in the Chicago Board Options Exchange (CBOE) over the period from January 1997 to June 2003 in daily frequency. There are totally 1633 trading days. The dataset contains the following series on option Trading Date, Expiration Date, Strike Price, Last Price, Last Bid Price, Last Ask Price and Underlying Price. The interest rates are proxied by the US 3-month Treasury bill rates, which are downloaded from *Datastream*.

I construct two sets of options for the convenience of application. They are at-the-money options (ATM) with moneyness (S/K) in the range $[0.97, 1.03]$ and out-of-the-money options (OTM) with moneyness in $[0.94, 0.97]$. They are all short-maturity options with maturity larger than 15 days but less than 45 days. Whenever at each time instant there are more than one call option in each set, I select those with moneyness closest to 1 for the at-the-money set and with moneyness closest to 0.97 for the out-of-the-money set. Now we have three sets of observations: one set of stock prices and two sets of option prices. Mid-prices between Bid and Ask of option prices are used when implementing filtering. Table 1 presents the descriptive statistics of S&P 500 index returns and the constructed call options.

— Table 1 around here —

With these data, we can implement the unscented sequential Monte Carlo method discussed in Section 3 to the infinite activity Lévy models built in Section 2. I take the Variance Gamma and Normal Inverse Gaussian models as examples again and still assume $V_t^{(2)} = 1$. I obtain parameter estimates with Markov Chain Monte Carlo (MCMC) methods (Li 2008). There remains a problem that the risk-premium parameters are not available. I thus assume that the parameters of the jump process and volatility process don't change under the change of measure. This is an unrealistic assumption. Many studies find that the jump and volatility have significant risk premia. But for our methodologically illustrative purpose, this assumption could be acceptable. The risk-neutral drifts of the models are substituted with the risk-free interest rates, indicating that only diffusion risk premium is taken into account. Table 2 presents parameter estimates of VG and NIG models.

— Table 2 around here —

Because of its relevance to option pricing, here only the volatility filtering results are provided. Figure 10 presents the filtered volatility using S&P 500 index and index options data for VG and NIG modelsin the second and third panels respectively. We can see that the filtered results are reasonably well with comparison to the evolution of returns (first panel) in both models even though we haven't taken into account the risk premia of the volatility and jump. The high volatility periods are apparently captured by the proposed method. The lowest panel presents the filtered volatility difference between these two models.

— Figure 10 around here —

We now empirically check how good the filtered volatility is in option pricing. Figure 11 plots the BS implied volatility from the market prices and the model-implied prices. The left panels are for VG model and the right panels for NIG model. The options are calls with the short maturity 28 days (upper panels) and the medium maturity 53 days (lower panels) and the strikes $[950, 975, 995, 1005, 1025, 1050]$ on June 30, 2003; the spot price at this day is 974.5; the interest rate has a very low value 0.84% and the filtered volatility which is 13.8% and 15.6% respectively for VG and NIG models are used for pricing options. We see that the fits are fair and for the short maturity options VG model performs better and for the medium maturity options NIG model is superior. We could deduce if we take the risk premia of the volatility and jump into account, these fits would be improved.

— Figure 11 around here —

I also conduct the unscented sequential Monte Carlo method using S&P 500 index alone and implement the conventional particle filter with the proposal density being the transition law of states jointly using S&P 500 index and index options data to the VG and NIG models. I find nearly the same results as we have seen in simulation studies: S&P 500 index alone can't effectively capture the dynamics of states, especially the stochastic volatility and even though for the simulated data, the conventional particle filter performs acceptably in filtering the volatility, for the real data it completely loses its ability to capture both the volatility and jumps (results are not reported and available by request.).

For the infinite activity derivatives pricing models with the leverage effect, in order to effectively capture the dynamics of states with Bayesian filtering methods, both the underlying and derivatives had better be used and simply taking the state transition law as the proposal density in particle filtering is of inferiority.

6 Concluding Remarks

In this paper, I investigate the time-changed infinite activity derivatives pricing models from the sequential Bayesian perspective relying on Brownian subordination property of the infinite activity Lévy processes. I propose a sequential Monte Carlo method with the proposal density generated by the unscented Kalman filter. This approach overcomes to the large extent the problem of particle impoverishment inherent to the conventional particle filters and is more robust in real data applications. The joint identification of the diffusion, stochastic volatility and infinite activity jumps can be achieved using both the underlying and derivatives data.

One of most important issues I haven't studied is parameter estimation. In this paper I assume the static parameters are known and not taken into account in the analysis. However, the parameters have to be estimated. The batch estimation methods mentioned in Section 1 could be applied. In practice, we don't need to estimate parameters every time new data arrive since they are by definition static. We could update them periodically in order to improve the efficiency. However, batch estimation is time-consuming, especially when we consider the joint estimation using both the underlying and derivative data. Alternatively we could sequentially estimate the parameters as we have done for states. On-line parameter estimation is still an open issue. There are some methods which have been proposed and applied in some particu-

lar cases such as Liu and West (2001), Storvik (2002), Andrieu, Doucet and Tadic (2005) and Johannes and Polson (2007). But they all have their own limitations. For the time-changed infinite activity models with the leverage effect, these methods can not do anything. It thus is interesting to investigate this issue.

APPENDIX

A Proof of Proposition

In the risk-neutral model, Brownian motion W_t^Q and $Z_t^{(1)Q}$ are correlated and others independent. I use the approach proposed by Carr and Wu (2004) to implement a change of measure and internalize this correlation.

Define a new measure M, which is absolutely continuous with respect to the risk-neutral measure Q,

$$\frac{dM}{dQ}\Big|_{\mathcal{G}_t} = \exp\left\{\left[iu(W_{T_t}^Q - \frac{1}{2}T_t) + \varphi_W^Q T_t\right]\right\}. \tag{43}$$

Under this new measure M, the variance rate process $V_t^{(1)}$ becomes,

$$\begin{aligned} dV_t^{(1)} &= (\kappa^{(1)}\theta^{(1)} - \kappa^{(1)Q}V_t^{(1)} + iu\rho\sigma^{(1)}V_t^{(1)})dt + \sigma^{(1)}\sqrt{V_t^{(1)}}dZ_t^{(1)M} \\ &= (\kappa^{(1)}\theta^{(1)} - \kappa^{(1)*}V_t^{(1)})dt + \sigma^{(1)}\sqrt{V_t^{(1)}}dZ_t^{(1)M}, \end{aligned} \tag{44}$$

where $\kappa^{(1)*} = \kappa^{(1)Q} - iu\rho\sigma^{(1)}$ and $Z_t^{(1)M}$ is now independent of W_t^Q. The conditional characteristic function of R_t then can be calculated as follows with the approaches proposed by Duffie, Pan and Singleton (2000).

$$\begin{aligned} \phi_R(u) &\equiv E^Q[e^{iuR_t}|\mathcal{G}_t] \\ &= e^{iur\tau}E^Q\left[e^{iu\left[\left(W^Q_{T_\tau^{(1)}} - \frac{1}{2}T_\tau^{(1)}\right) + \left(X^Q_{T_\tau^{(2)}} - k_X^Q(1)T_\tau^{(2)}\right)\right]}\Big|\mathcal{G}_t\right] \\ &= e^{iur\tau}E^Q\left[M(\tau)e^{-\varphi_W^Q T_\tau^{(1)}}|\mathcal{G}_t\right]E^Q\left[e^{-\varphi_X^Q T_\tau^{(2)}}|\mathcal{G}_t\right] \\ &= e^{iur\tau}E^M\left[e^{-\int_{t-\tau}^t \varphi_W^Q V_s^{(1)}ds}|\mathcal{G}_t\right]E^Q\left[e^{-\int_{t-\tau}^t \varphi_X^Q V_s^{(2)}ds}|\mathcal{G}_t\right] \\ &= e^{iur\tau}e^{A^{(1)}(u,\tau)+B^{(1)}(u,\tau)V_t^{(1)}}e^{A^{(2)}(u,\tau)+B^{(2)}(u,\tau)V_t^{(2)}}, \end{aligned} \tag{45}$$

where E^M is the expectation operator under the new measure M and $A^{(1)}(u,\tau) = A^{(1)}(t)$, $B^{(1)}(u,\tau) = B^{(1)}(t)$ as well as $A^{(2)}(u,\tau) = A^{(2)}(t)$, $B^{(2)}(u,\tau) = B^{(2)}(t)$ solve the following two sets of ODEs

$$\begin{cases} \dot{B}^{(1)}(t) &= u + \kappa^{(1)*}B^{(1)}(t) - \frac{1}{2}(\sigma^{(1)})^2(B^{(1)})^2(t) \\ \dot{A}^{(1)}(t) &= -\kappa^{(1)}\theta^{(1)}B^{(1)}(t) \end{cases}$$

$$\begin{cases} \dot{B}^{(2)}(t) &= u + \kappa^{(2)*}B^{(2)}(t) - \frac{1}{2}(\sigma^{(2)})^2(B^{(2)})^2(t) \\ \dot{A}^{(2)}(t) &= -\kappa^{(2)}\theta^{(2)}B^{(2)}(t) \end{cases}$$

where $\kappa^{(2)*} = \kappa^{(2)Q}$ and boundary conditions $B^{(i)}(t) = 0$ and $A^{(i)}(t) = 0$ for $i = 1, 2$. Solving these two sets of ODEs, we obtain the result of the Proposition 3 .

B The Scaled Unscented Transformation

The scaled unscented transformation is a method for calculating the statistics of a random variable which undergoes a nonlinear transformation (Julier and Uhlman 1997). For a nonlinear function

$$y = f(x), \tag{46}$$

assume the mean and covariance of x (with dimension L) are $\bar{x}$ and P_x. The mean and covariance of y could be calculated through forming a set of $2L+1$ sigma points χ:

$$\chi_0 = \bar{x}, \tag{47}$$

$$\chi_i = \bar{x} + (\sqrt{(L+\lambda)P_x})_i, \qquad i = 1, \ldots, L, \tag{48}$$

$$\chi_i = \bar{x} - (\sqrt{(L+\lambda)P_x})_{i-L}, \quad i = L+1, \ldots, 2L, \tag{49}$$

where $\lambda = \alpha^2(L+\kappa) - L$ is a scaling parameter; the constant α determines the spread of the sigma points around $\bar{x}$ and is usually set to be a small positive value and κ is a second scaling parameter with value set to 0 or $3-L$. These sigma points are propagated through the nonlinear function f,

$$\mathcal{Y}_i = f(\chi_i), \quad i = 0, 1, \ldots, 2L, \tag{50}$$

and the mean and covariance of y are approximated with a weighted sample mean and covariance of posterior sigma points,

$$\bar{y} = \sum_{i=0}^{2L} w_i^{(m)} \mathcal{Y}_i, \tag{51}$$

$$P_y = \sum_{i=0}^{2L} w_i^{(c)} (\mathcal{Y}_i - \bar{y})(\mathcal{Y}_i - \bar{y})', \tag{52}$$

where the weights are given by

$$w_0^{(m)} = \frac{\lambda}{L+\lambda}, \tag{53}$$

$$w_0^{(c)} = \frac{\lambda}{L+\lambda} + (1 - \alpha^2 + \beta), \tag{54}$$

$$w_i^{(m)} = w_i^{(c)} = \frac{1}{2(L+\lambda)}, \quad i = 1 : 2L, \tag{55}$$

where β is a covariance correction parameter and used to incorporate prior knowledge of the distribution of x. The scaled unscented transformation can approximate posterior mean and covariance with accuracy up to 3rd order for Gaussian inputs and for non-Gaussian inputs, the accuracy can be reached to at least 2nd order, with accuracy of third and higher order determined by parameters α and β.

C The Unscented Sequential Monte Carlo Method

For any function $f(x_{0:t})$, its expectation can be computed through importance sampling with the proposal density π by

$$\begin{aligned}
E[f(x_{0:t})] &= \int f(x_{0:t}) p(x_{0:t}|y_{1:t}) dx_{0:t} \\
&= \int f(x_{0:t}) \frac{p(x_{0:t}|y_{1:t})}{\pi(x_{0:t}|y_{1:t})} \pi(x_{0:t}|y_{1:t}) dx_{0:t} \\
&= \int f(x_{0:t}) \frac{p(y_{1:t}|x_{0:t}) p(x_{0:t})}{p(y_{1:t}) \pi(x_{0:t}|y_{1:t})} \pi(x_{0:t}|y_{1:t}) dx_{0:t} \\
&= \frac{\int f(x_{0:t}) \omega_t \pi(x_{0:t}|y_{1:t}) dx_{0:t}}{\int p(y_{1:t}|x_{0:t}) p(x_{0:t}) dx_{0:t}} \\
&= \frac{\int f(x_{0:t}) \omega_t \pi(x_{0:t}|y_{1:t}) dx_{0:t}}{\int \omega_t \pi(x_{0:t}|y_{1:t}) dx_{0:t}} \\
&= \frac{E_\pi[\omega_t f(x_{0:t})]}{E_\pi[\omega_t]},
\end{aligned} \tag{56}$$

where the unnormalized weight

$$\omega_t = \frac{p(y_{1:t}|x_{0:t})p(x_{0:t})}{\pi(x_{0:t}|y_{1:t})} \tag{57}$$

can be updated recursively

$$\begin{aligned}
\omega_t &= \frac{p(y_{1:t}|x_{0:t})p(x_{0:t})}{\pi(x_{0:t-1}|y_{1:t-1})\pi(x_t|x_{0:t-1},y_{1:t})} \\
&= \frac{p(y_{1:t}|x_{0:t})p(x_{0:t})}{\pi(x_{0:t-1}|y_{1:t-1})\pi(x_t|x_{0:t-1},y_{1:t})} \times \frac{p(y_{1:t-1}|x_{0:t-1})p(x_{0:t-1})}{p(y_{1:t-1}|x_{0:t-1})p(x_{0:t-1})} \\
&= \frac{p(y_{1:t-1}|x_{0:t-1})p(x_{0:t-1})p(y_{1:t}|x_{0:t})p(x_{0:t})}{\pi(x_{0:t-1}|y_{1:t-1})p(y_{1:t-1}|x_{0:t-1})p(x_{0:t-1})\pi(x_t|x_{0:t-1},y_{1:t})} \\
&= \omega_{t-1}\frac{p(y_{1:t}|x_{0:t})p(x_{0:t})}{p(y_{1:t-1}|x_{0:t-1})p(x_{0:t-1})\pi(x_t|x_{0:t-1},y_{1:t})} \\
&= \omega_{t-1}\frac{p(y_t|x_t)p(x_t|x_{t-1})}{\pi(x_t|x_{0:t-1},y_{1:t})}.
\end{aligned}$$

Suppose we have N samples $\{x_t^{(i)}, i = 1, 2, \cdots, N\}$ from the proposal density. Formula (56) can be estimated by

$$\begin{aligned}
\hat{E}[f(x_{0:t})] &= \sum_{i=1}^{N} \tilde{w}_t^{(i)} f(x_{0:t}^{(i)}), \\
\tilde{w}_t^{(i)} &= \frac{w_t^{(i)}}{\sum_{j=1}^{N} w_t^{(j)}}.
\end{aligned}$$

In this paper, I apply the unscented Kalman filter to design the proposal density π. Meanwhile a resampling step is also introduced. The resulted Unscented Sequential Monte Carlo method is then implemented through the following steps:

- **Step 1:** initializing at $t = 0$: Draw a set of particles $\{x_0^{(i)}, i = 1, \ldots, N\}$ from the prior $p(x_0)$ and give each particle a weight $\frac{1}{N}$;
- **Step 2:** for $t = 1, 2, \ldots$
  - update the Gaussian prior distribution for each particle $x_{t-1}^{(i)}$ with UKF and we obtain posterior mean $\bar{x}_t^{(i)}$ and variance $P_t^{(i)}$;
  - sample $x_t^{(i)} \longrightarrow \tilde{\pi}(x_t^{(i)}|y_{1:t}) \equiv \mathcal{N}(\bar{x}_t^{(i)}, P_t^{(i)})$;
  - update the weight for each particle with

$$w_t^{(i)} = w_{t-1}^{(i)} \frac{p(y_t|x_t^{(i)})p(x_t^{(i)}|x_{t-1}^{(i)})}{\tilde{\pi}(x_t^{(i)}|y_{1:t})};$$

- normalize the weights: $\tilde{w}_t^{(i)} = w_t^{(i)} / \sum_j^N w_t^{(j)}$;

- **Step 3:** Residual Resampling
  - retain $N_i' = \left\lfloor N\tilde{w}_t^{(i)} \right\rfloor$ copies of $x_t^{(i)}$;
  - uniformly sample the remaining $N_t'' = N - \sum_{i=1}^N N_i'$ with new weights $(N\tilde{w}_t^{(i)} - N_i')/N_t''$;
  - reset the weights to $1/N$;

- **Step 4:** output a set of new particles $\{x_t^{(i)}, i = 1, \ldots, N\}$ with equal weight $1/N$.

References

[1] Ait-Sahalia, Y. (2004), "Disentangling Diffusion from Jumps", *Journal of Financial Economics*, 74, 487-528.

[2] Andrieu, C., Doucet, A., and Tadic, V.B. (2005), "On-Line Parameter Estimation in General State-Space Models", Working Paper, University of Bristol.

[3] Bakshi, G., Carr, P., and Wu, L. (2008), "Stochastic Risk Premiums, Stochastic Skewness in Currency Options and Stochastic Discount Factors in International Economics", Forthcoming, *Journal of Financial Economics*.

[4] Barndorff-Nielson, O.E. (1977), "Exponentially Decreasing Distributions for the Logarithm of Particle Size", *Proceedings of the Royal Society*, A353, 401-419.

[5] — (1981), "Hyperbolic Distributions and Ramifications: Contributions to Theory and Applications", In C. Taillie, G. Patil and B. Baldessari (eds), *Statistical Distributions in Scientific Work*, Vol 4, Reidel, Dordrecht.

[6] — (1998), "Processes of Normal Inverse Gaussian Type", *Finance and Stochastics*, 2, 41-68.

[7] —, and Shephard, N. (2001), "Normal Modified Stable Processes", *Theory of Probability and Mathematical Statistics*, 65, 1-19.

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

[9] — (2000), "Post-'87 Crash Fears in the S&P 500 Futures Option Market", *Journal of Econometrics*, 94, 181-238.

[10] Carr, P., Geman, H., Madan, D.B., and Yor, M. (2002), "The Fine Structure of Asset Returns: An Empirical Investigation", *Journal of Business*, 75, 305-332.

[11] —, —, —, and — (2003), "Stochastic Volatility for Lévy Processes", *Mathematical Finance*, 13, 345-382.

[12] —, and Madan, D.B. (1999), "Option Valuation Using the Fast Fourier Transform", *Journal of Computational Finance*, 3, 61-73.

[13] —, and Wu, L. (2004), "Time-changed Lévy Processes and Option Pricing", *Journal of Financial Economics*, 71, 113-141.

[14] Chan, T. (1999), "Pricing Contingent Claims on Stocks Driven By Lévy Processes", *The Annals of Applied Probability*, 9, 504-528.

[15] Chourdakis, K. (2005), "Option Pricing Using the Fractional FFT", *Journal of Computational Finance*, 8, 1-18.

[16] Cont, R., and Tankov, P. (2004), *Financial Modeling With Jump Processes*, London: Chapman & Hall/CRC.

[17] Cox, J. C., Ingersoll, J.E., and Ross, S.A. (1985), "A Theory of the Term Structure of Interest Rates", *Econometrica*, 53, 385-408.

[18] Doucet, A., de Freitas, N., and Gordon, N. (2001), *Sequential Monte Carlo Methods in Practice*, New York: Springer.

[19] —, Godsill, S., and Andrieu, C. (2000), "On Sequential Monte Carlo Sampling Methods for Bayesian Filtering", *Statistics and Computing*, 10, 197-208.

[20] Duffie, D., Pan, J., and Singleton, K. (2000), "Transform Analysis and Asset Pricing for Affine Jump-Diffusions", *Econometrica*, 68, 1343-1376.

[21] Eberlein, E. (2001), "Application of Generalized Hyperbolic Lévy Motions to Finance", In O.E. Barndorff-Nielson, T. Mikosch and S. Resnick (eds), *Lévy Processes: Theory and Applications*. Boston: Birkhauser.

[22] Eraker, B., Johannes, M., and Polson, N. (2003), "The Impact of Jumps in Equity Index Volatility and Returns", *Journal of Finance*, 58, 1269-1300.

[23] Gerber, H. U., and Shiu, E. S.W. (1994), "Option Pricing by Esscher Transforms", *Transactions of Society of Actuaries*, 46, 99-191.

[24] Gordon, N., Salmond, D., and Smith, A. (1993), "Novel Approach to Nonlinear and Non-Gaussian Bayesian State Estimation", *IEEE Proceedings-F*, 140, 107-113.

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

[26] Hougaard, P. (1986), "Survival Models for Heterogeneous Populations Derived from Stable Distributions", *Biometrika*, 73, 387-396.

[27] Huang, J., and Wu, L. (2004), "Specification Analysis of Option Pricing Models Based on Time-Changed Lévy Processes", *The Journal of Finance* 59, 1405-1439.

[28] Isard, M., and Blake, A. (1996), "Visual Tracking by Stochastic Propagation of conditional density", *Proceedings of 4th European Conference on Computer Vision*, 343-356.

[29] Johannes, M., and Polson. N. (2008), "Particle Filtering and Parameter Learning", Working Paper, University of Chicago.

[30] —, —, and Stroud, J. (2008), "Optimal Filtering of Jump-Diffusions: Extracting Latent States from Asset Prices", Forthcoming, *Review of Financial Studies*.

[31] Julier, S.J., and Uhlmann, J.K. (1997), "A New Extension of the Kalman Filter to Nonlinear Systems", In *Proceedings of AeroSense: the 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*.

[32] Li, J. (2008), "Stochastic Jump Intensity, Stochastic Volatility and Stochastic Higher Moments in Asset Returns: An Empirical Investigation", Working Paper, Bocconi University, Milan.

[33] —,Favero, C., and Ortu, F. (2008), "Spectral Iterative Estimation of Tempered Stable Stochastic Volatility Models and Option Pricing", Working Paper, Bocconi University, Milan.

[34] Li, H., Wells, M.T., and Yu, C.L. (2008), "A Bayesian Analysis of Return Dynamics with Lévy Jumps", *Review of Financial Studies*, 21, 2345-2378.

[35] Liu, J., and West, M. (2001), "Combined Parameter and State Estimation in Simulation-Based Filtering", In A. Doucet, N. de Freitas and N. Gordon (eds), *Sequential Monte Carlo Methods in Practice*, New York: Springer.

[36] Madan, D., Carr, P., and Chang, E. (1998), "The Variance Gamma Process and Option Pricing", *European Finance Review*, 2, 79-105.

[37] —, and Yor, M. (2006), "Representing the CGMY and Meixner Lévy Processes as Time-Changed Brownian Motions", Working Paper, University of Maryland.

[38] Michael, J., Schucany, W., and Haas, R. (1976). "Generating Random Variates Using Transformations with Multiple Roots", *The American Statistician*, 30, 88-90.

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

[40] Sato, K. (1999), *Levy Processes and Infinitely Divisible Distributions*, Cambridge: Cambridge University Press.

[41] Storvik, G. (2002), "Particle Filters for State-Space Models with the Presence of Unknown Static Parameters", *IEEE Transactions on Signal Processing*, 50, 281-289.

[42] Schoutens, W., and Teugels, J.L. (1998), "Lévy Processes, Plynomials and Martingales", *Communications in Statistics: Stochastic Models*, 14, 335-349.

[43] Stroud, J., Muller, P., and Polson, N. (2003), "Nonlinear State-Space Models with State-Dependent Variances',' *Journal of the American Statistical Association*, 98, 377-386.

[44] van der Merwe, R., de Freitas, J., Doucet, A., and Wan, E. (2000), "The Unscented Particle Filter", Technical Report, University of Cambridge.

[45] Wan, E., and van der Merwe, R. (2001), "The Unscented Kalman Filter", In S. Haykin (ed), *Kalman Filtering and Neural Networks*, New York: John Wiley & Sons.

Table 1: **Descriptive Statistics of the Data**

A. S&P 500 Index Returns						
	T	Mean	St. Dev.	Skewness	Kurtosis	JB Test
Daily	1633	0.044	0.211	-0.054	4.968	1($<$ 0.001)
B. Constructed Calls						
	Mean Mn.	Std Mn.	Mean Mt.	Std Mt.	Mean IV.	Std IV.
OTM	0.962	0.003	22.414	9.756	0.194	0.051
ATM	1.000	0.004	25.268	7.400	0.213	0.051

Note: Table presents the descriptive statistics of S&P 500 index and index option data used in the paper. Data are from January 1997 to June 2003 in daily frequency. There are totally 1633 business days. In panel A, mean and standard deviation are annualized and JB Test is the Jarque-Bera normality test, whose value 1 (p-value in bracket) means the test rejects the hypothesis of normality; In panel B, Mn stands for moneyness, Mt maturity (in days) and IV BS implied volatility.

Table 2: **Parameter Estimates with MCMC**

	μ	ω	η	v	$\kappa^{(1)}$	$\theta^{(1)}$	$\sigma^{(1)}$	ρ
VG	0.066	-0.040	0.104	0.004	3.227	0.031	0.262	-0.974
	(0.019)	(0.012)	(0.011)	(0.001)	(0.780)	(0.006)	(0.040)	(0.018)
NIG	0.067	-0.060	0.092	0.009	4.329	0.035	0.360	-0.977
	(0.023)	(0.018)	(0.007)	(0.003)	(0.941)	(0.006)	(0.049)	(0.014)

Note: Models are estimated with MCMC method after discretization with time interval τ

$$\begin{aligned} y_t &= y_{t-\tau} + \mu\tau + X_\tau + \sqrt{\tau V_{t-\tau}^{(1)}} w_t, \\ V_t^{(1)} &= V_{t-\tau}^{(1)} + \kappa^{(1)}\left(\theta^{(1)} - V_{t-\tau}^{(1)}\right)\tau + \sigma^{(1)}\sqrt{\tau V_{t-\tau}^{(1)}} z_t^{(1)}, \\ X_\tau &= \omega \mathcal{S}_\tau + \eta\sqrt{\mathcal{S}_\tau}\tilde{w}_t, \\ \mathcal{S}_\tau &= g_{\mathcal{S}}(\tau; 1, v), \end{aligned}$$

where $g_{\mathcal{S}}$ is the Gamma process for VG model and inverse Gaussian process for NIG model. Both subordinators are constructed to have unit mean rate and variance rate v. MCMC simulation is run with the number of iterations 50,000, the first 20,000 of which is discarded and the last 30,000 is kept for inference. Sample means and standard deviations (in brackets) are presented.

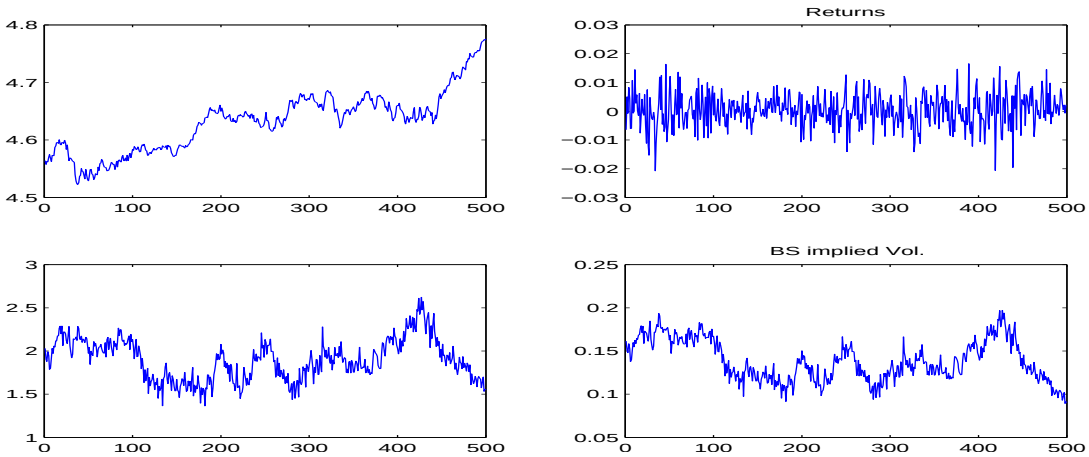

Figure 1: **Simulated Data from SV Model**

Note: Data are simulated from the Heston SV Model. The parameters $\{\kappa^{(1)}, \theta^{(1)}, \sigma^{(1)}, \rho\}$ are set to be $\{6.00,\ 0.025,\ 0.30,\ -0.50\}$. The at-the-money short maturity options are those with strike equal to the underlying and maturity 1 month.

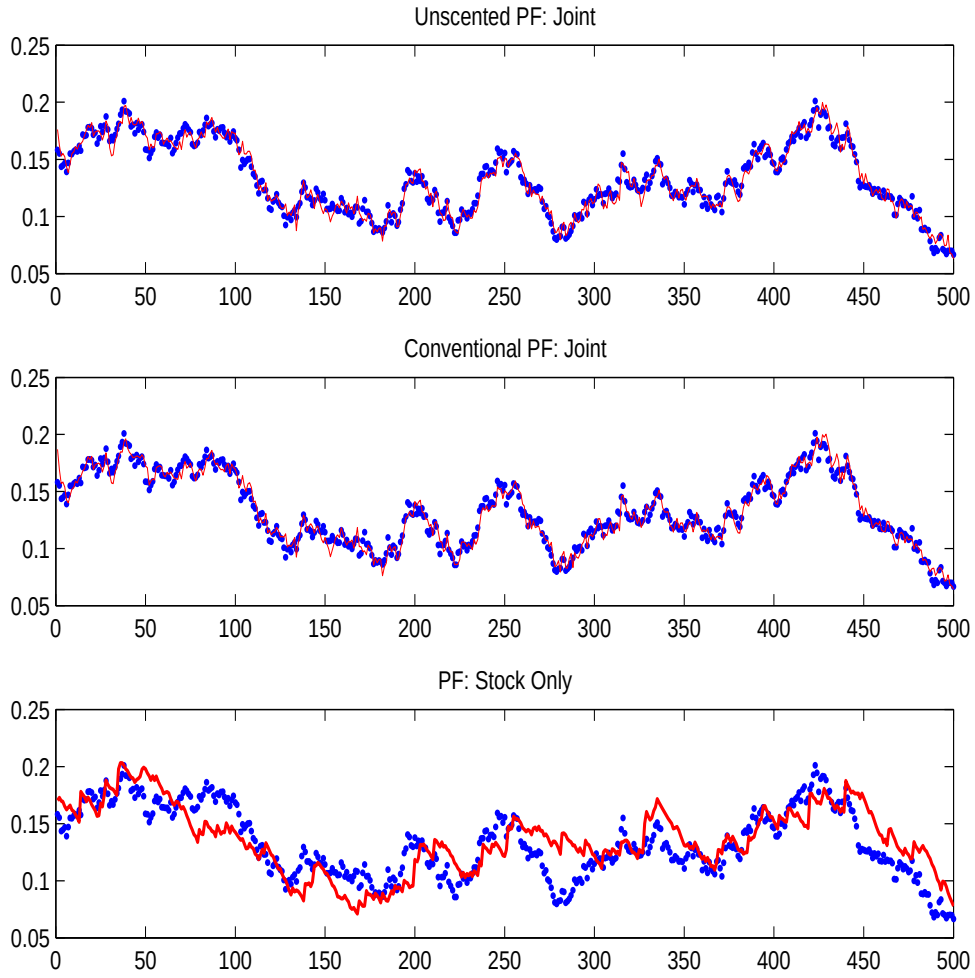

Figure 2: **SV State Filtering**

Note: The volatility is filtered with the stock prices and option prices using the unscented sequential Monte Carlo method and the conventional particle filter respectively in the upper and middle panels. The lower panels presents the filtered results using stock price data alone. The number of particles in the unscented sequential Monte Carlo method is 200 and in the conventional particle filter 2000.

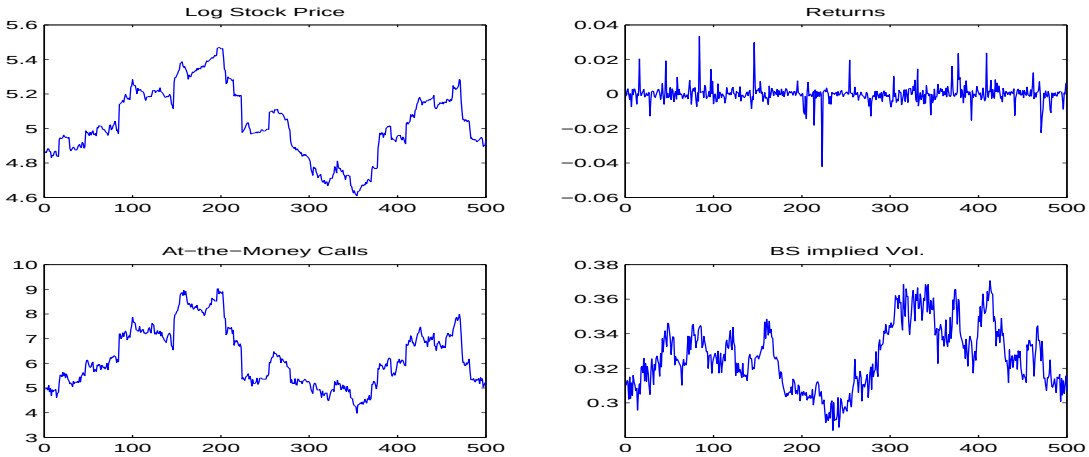

Figure 3: **Simulated Data from VG Model** ($v = 0.05$)

Note: Data are simulated from the Variance Gamma Model. The parameters $\{\omega, \eta, v, \kappa^{(1)}, \theta^{(1)}, \sigma^{(1)}, \rho\}$ are set to be $\{-0.04,\ 0.30,\ 0.05,\ 6.00,\ 0.025,\ 0.30,\ -0.50\}$. The at-the-money short maturity options are those with strike equal to the underlying and maturity 1 month. They are computed using FRFT.

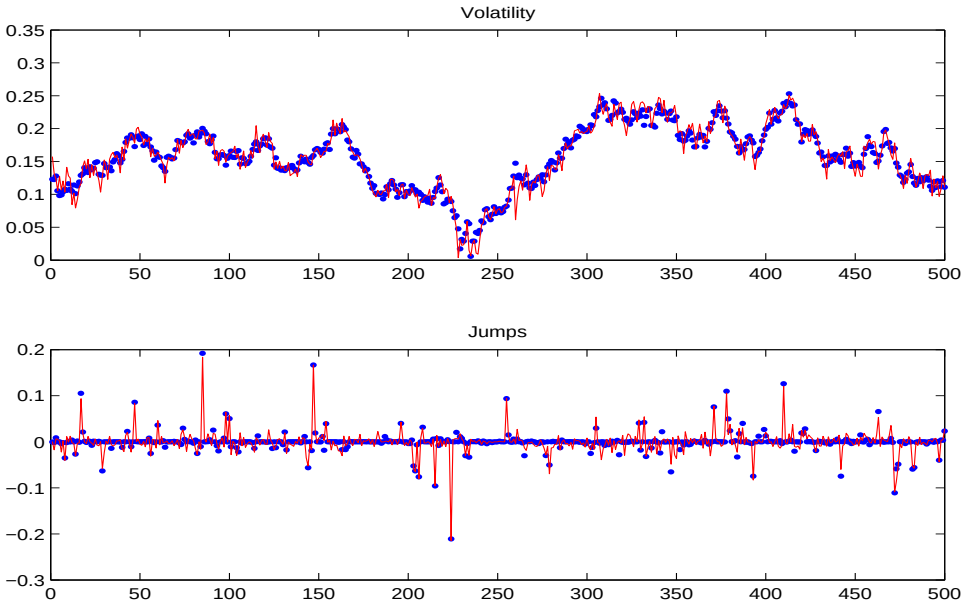

Figure 4: **VG Joint State Filtering**

Note: The volatility and jumps are filtered with the stock prices and option prices using the unscented sequential Monte Carlo method. The upper panel presents the filtered volatility and the lower panel the filtered jumps. The dots represent the true values and the lines the estimated values. The number of particles is 200.

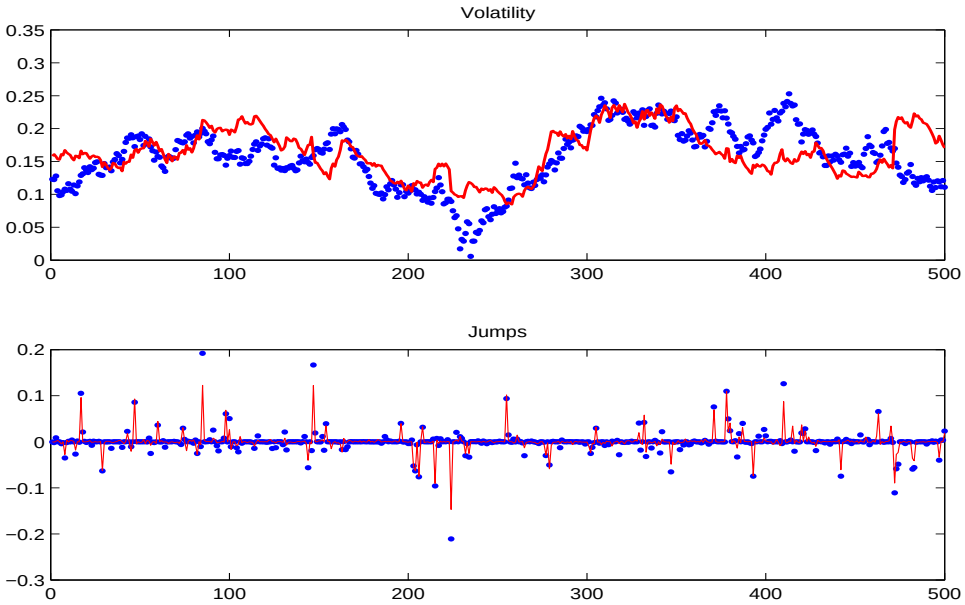

Figure 5: **VG State Filtering with Stock Prices Only**

Note: The volatility and jumps are filtered with the stock prices only using the unscented sequential Monte Carlo method. The upper panel presents the filtered volatility and the lower panel the filtered jumps. The dots represent the true values and the lines the estimated values. The number of particles is 2000.

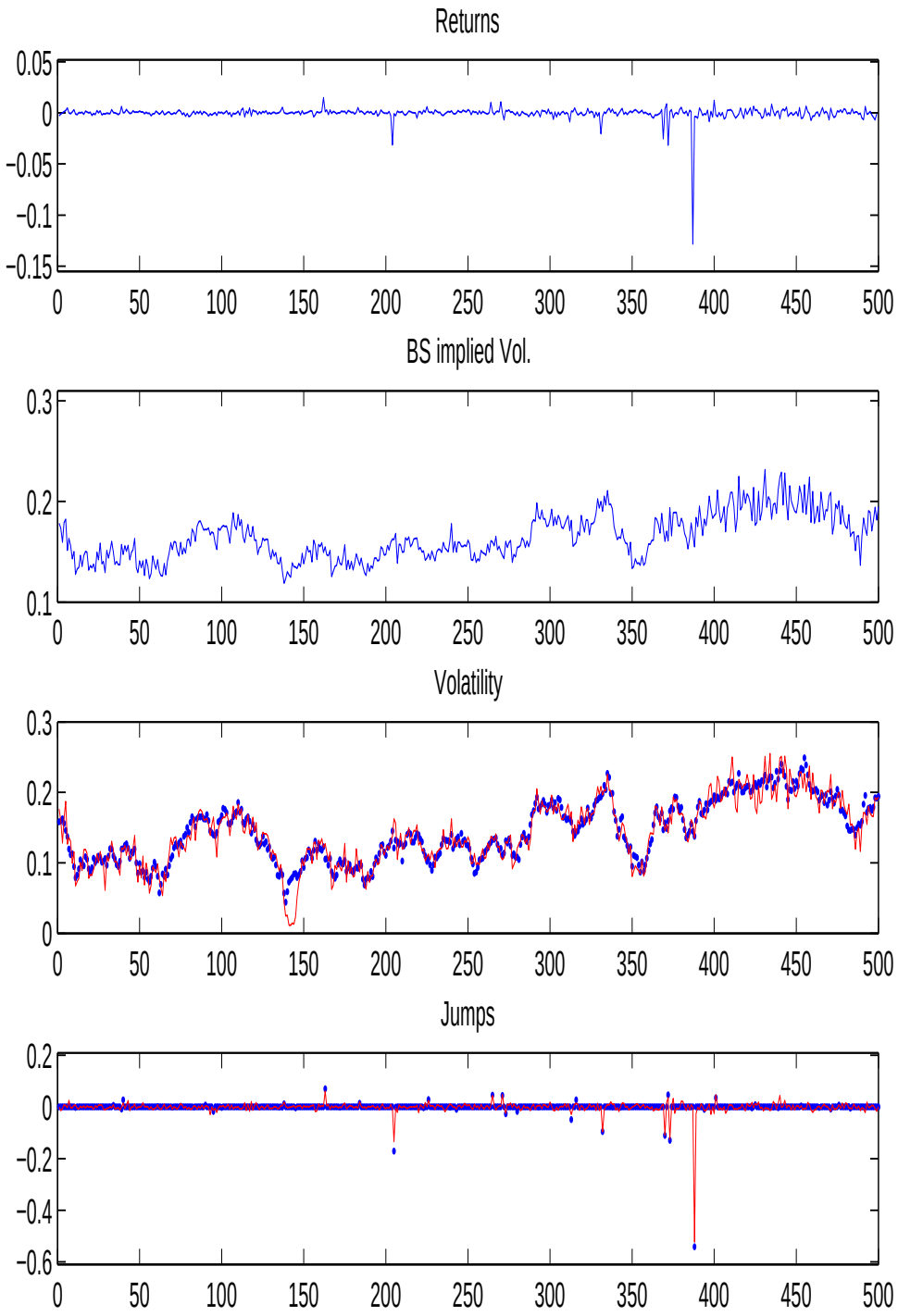

Figure 6: **Simulated Data and Joint State Filtering for VG Model** ($v = 0.5$)

Note: Data are simulated with the same parameters as before except parameter $v = 0.50$ from the variance gamma model. The top two panels present simulated returns and BS implied volatility. The lower two panels present the true (dots) and filtered (lines) volatility and jumps. The number of particles is 200.

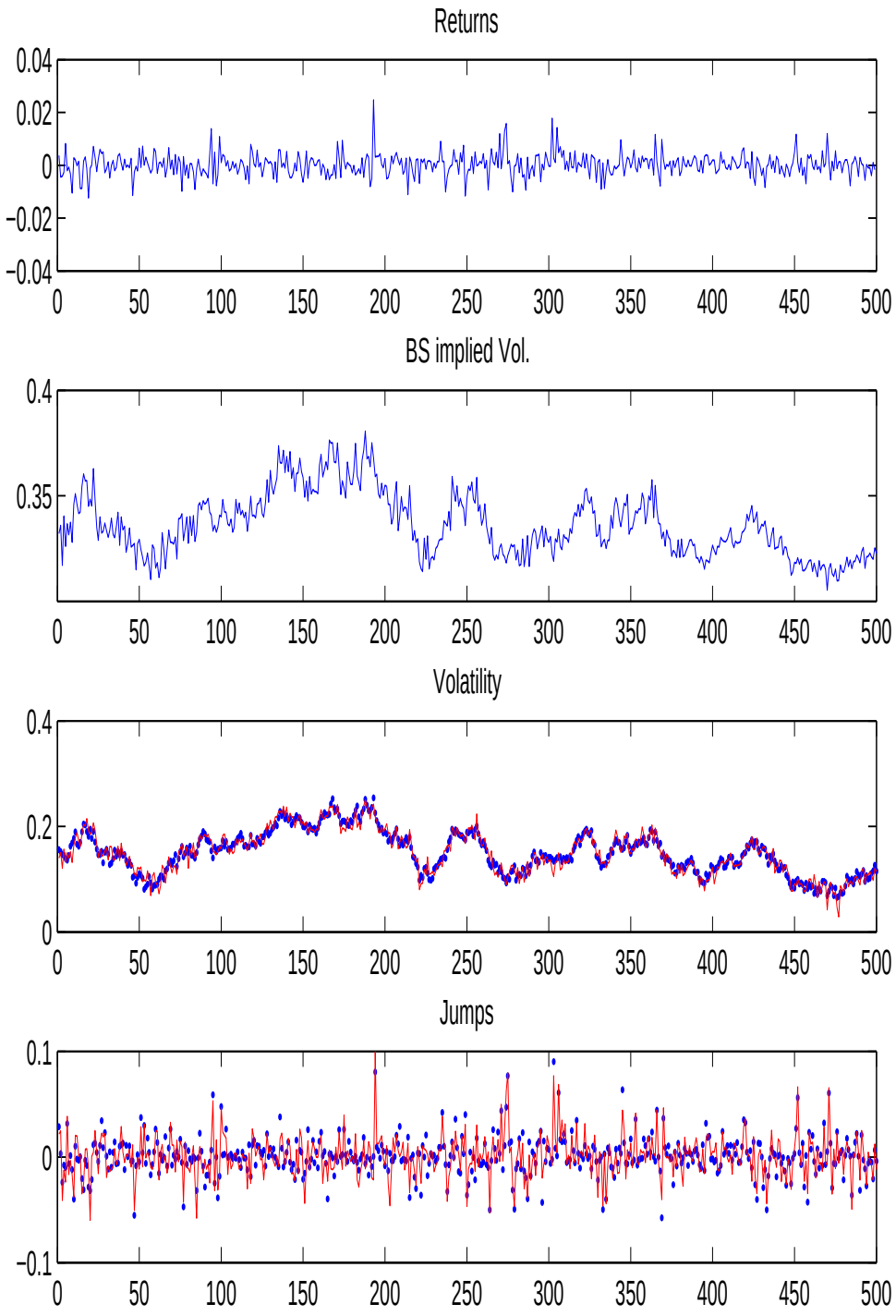

Figure 7: **Simulated Data and Joint State Filtering for VG Model** ($v = 0.005$)

Note: Data are simulated with the same parameters as before except parameter $v = 0.005$ from the variance gamma model. The top two panels present simulated returns and BS implied volatility. The lower two panels present the true (dots) and filtered (lines) volatility and jumps. The number of particles is 200.

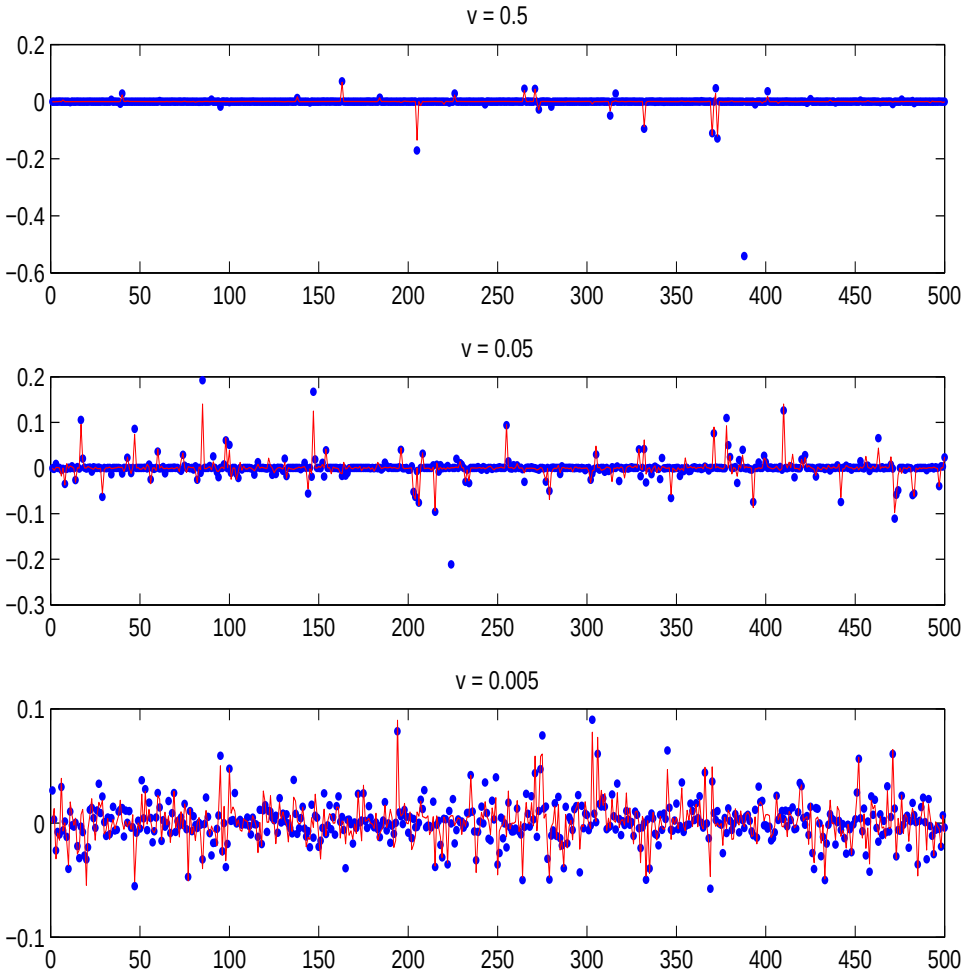

Figure 8: **VG Joint State Filtering with the Conventional Particle Filter**

Note: The states are filtered using the stock prices and option prices with the conventional particle filter which takes state transition law as the proposal density. The figure presents the filtered jumps for the different jump structure models. The dots are the true values and the lines filtered values. The number of particles is 2000.

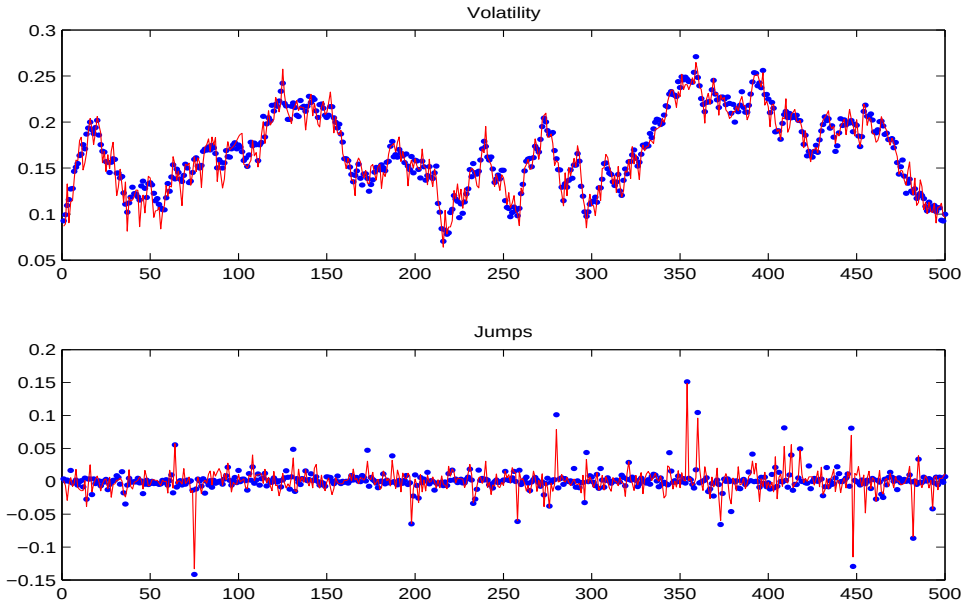

Figure 9: **NIG Joint State Filtering**

Note: The volatility and jumps are filtered with the stock prices and option prices using the unscented sequential Monte Carlo method. The upper panel presents the filtered volatility and the lower panel the filtered jumps. The dots represent the true values and the lines the estimated values. The number of particles is 200.

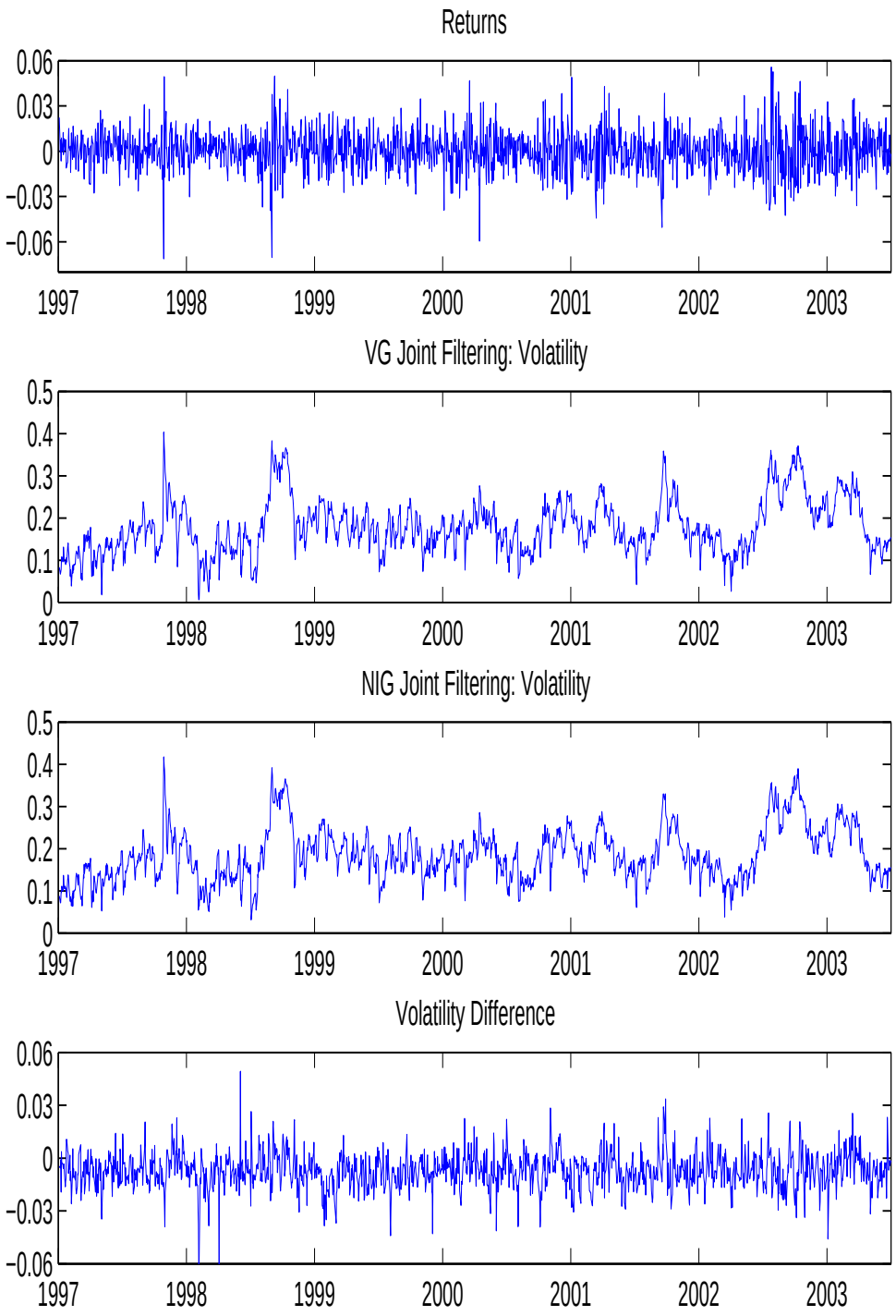

Figure 10: **Jointly Filtered Volatility**

Note: The unscented sequential Monte Carlo method is applied to VG and NIG models with S&P 500 index and index options data. The first panel is the return time series; the second and third panels are the filtered volatility in VG and NIG models respectively; and the last panel is the filtered volatility difference of two models studied. The number of particles is 300.

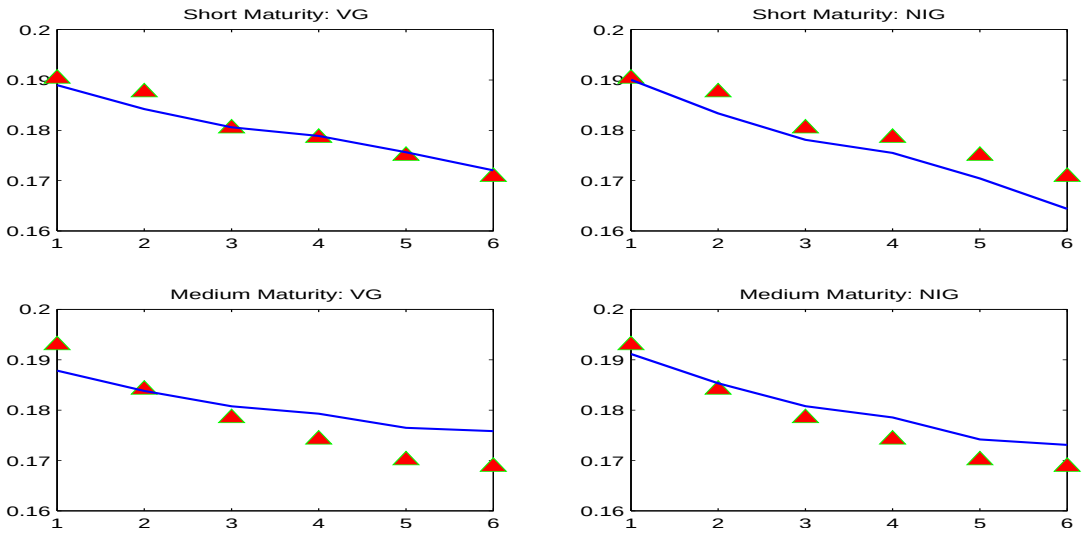

Figure 11: **BS Implied Volatilities from Market Prices and Model-Implied Prices**

Note: The figure plots BS implied volatilities from the market prices (Triangles) and the model-implied prices (Lines). The left panels are for VG model and the right panels for NIG model. The options are those with maturity 28 days (upper panels) and 53 days (lower panels) and strikes $[950, 975, 995, 1005, 1025, 1050]$ on June 30, 2003; the spot price is 974.5; the interest rate 0.84% and the filtered volatilities are used for option pricing.

Chapter 3

Jump Dynamics, Volatility Components and Return-Volatility Relation

Abstract

This paper investigates the return-volatility relation by taking into account model specification problem. The stock price process is modeled by the time-changed Brownian motion and infinite activity Lévy process, which introduces not only the stochastic diffusion volatility but also the stochastic jump intensity. The model indicates that under the absence of leverage effects it becomes a variant of the Merton's ICAPM whereas under the existence of leverage effects, the return-volatility relation is determined by interactions between risk premia and leverage effects. It provides a theoretical justification for mixed empirical findings. Our empirical study finds a positive return-diffusion volatility relation and more interestingly a negative return-jump volatility relation.

1 Introduction

The return-volatility relation is a fundamental issue in asset pricing theory. Merton (1973, 1980) argue with the intertemporal capital asset pricing model (ICAPM) that the conditional expected excess return on the market portfolio should be positively related to its conditional volatility

$$E_{t-1}[R_t] = \alpha + \gamma E_{t-1}[V_t], \tag{1}$$

where R_t and V_t are the excess return and variance of the market portfolio, respectively; γ is a risk aversion parameter of the representative agent in the market and α is a constant term, which should be zero if the model is correctly specified. The intuition behind this positive relation is that if the volatility is really priced by the market, the expected volatility increase

would require a higher return on the portfolio. This is often referred to as the volatility feedback effect (Pindyck, 1984). Whereas the ICAPM and the volatility feedback effect imply a positive relation, the leverage effect, which is firstly discussed by Black (1976) and Christie (1982), indicates that the decrease of the stock price causes an increase in debt-to-equity ratio, which makes the stock riskier and increases its volatility. The investigation in this strand generally finds that the volatility responds asymmetrically to the negative and positive returns.

Empirical studies on the return-volatility relation have resulted in mixed findings. On the one hand, French et al. (1987), Chou (1988), Campbell and Hentsche (1992), Ghysels et al. (2004), Bali and Peng (2006) and Guo and Whitelaw (2006), have reported the positive and significant estimate of the risk aversion parameter. On the other hand, others (Campbell, 1987; Breen et al., 1989; Turner et al., 1989; Chou et al., 1992; Glosten et al., 1993; Brandt and Kang, 2004; Lettau and Ludvigson, 2004) have documented a negative and/or insignificant relation between the return and volatility. Furthermore, the return-volatility relation is sensitive to the model, exogenous predictors and the length of the return horizon (Harvey, 2001; Harrison and Zhang, 1999).

Since the stock price volatility is not directly observable, most of the above studies apply either the nonparametric methods or the parametric GARCH models to estimating the volatility. But these methods suffer at least from one serious problem: volatility estimate bias. For instance, the frequently used nonparametric volatility estimate is the realized volatility, which is consistent only if the data generating process follows a diffusion process and GARCH models explicitly indicate that the volatility is deterministic and can be completely inferred from the past information. Moreover, these studies barely provide any theoretical justifications for their results.

In this paper, I closely study the stock price dynamics and the time-series relationship between the return and volatility by taking into consideration the model specification problem with reference to the burgeoning literature in continuous-time financial modeling. We want to build a model which can not only accommodate the stylized facts observed in the markets, but also provide a theoretical explanation for mixed empirical findings.

First, stochastic volatility and jump have been becoming the fundamental factors in studying continuous-time asset pricing models. In the latest decade, investigations have nearly exclusively focused on modeling stochastic volatility using the square-root process (Cox, In-

gersoll and Ross, 1985; Heston, 1993) and modeling return jumps using the compound Poisson process. With different datasets in asset returns and/or options, these studies (Bakshi, Cao and Chen, 1997; Bates, 2000; Pan, 2002; Andersen, Benzoni and Lund, 2002; Eraker Johannes and Polson, 2003; Eraker, 2004; Broadie, Chernov and Johannes, 2007) have reached almost the same results: stochastic volatility alone can not capture the distributional characteristics of asset returns and explain the implied volatility skew/smile of options and a jump component in the return (and the volatility) process is indispensable.

Second, having noticed that to assume return jump a rare event is implausible and also inconsistent with the discretely observed sample data (Madan, 2001; Carr et al., 2002; Geman, 2002), I thus rely on the infinite activity Lévy process. A fundamental difference between the infinite activity Lévy process and the finite activity compound Poisson process is that the former can produce an infinite number of jumps at any finite time interval and regards return jump as a common event. The superiority of the infinite activity Lévy models have already been reported in Huang and Wu (2004), Li, Wells and Liu (2008), Li, Favero and Ortu (2008) and others.

Third, I model asset price dynamics with the exponential Brownian motion and infinite activity Lévy process, both of which are time-changed by the different independent variance rates. By time-changing both the Brownian motion and the infinite activity Lévy process, we introduce not only the stochastic diffusion volatility but also the stochastic jump arrival rate. It also implies stochastic higher moments (skewness and kurtosis) in asset returns. There is evidence that the jump arrives at the different rates and tends to be clustered (Maheu and McCurdy, 2004). This is in contrast to the jump-diffusion stochastic volatility model, where the jump intensity is usually assumed to be constant. Brownian motion and Lévy process can be interpreted as systematic (market) risk and idiosyncratic (credit) risk respectively and stochastic diffusion volatility and stochastic jump arrival rate indicate that the information flow of these two risks is time varying with the different speed and may cause rapid price fluctuations. Now the stochastic return volatility has two sources: one is from the diffusion part and the other from the infinite activity jump part. Thus, we can explicitly study the relationship between the return and the component of the return volatility.

Finally and most importantly, the model sheds new light on the return-volatility relation. It indicates that under the absence of leverage effects the model becomes a variant of

the Merton's ICAPM whereas under the existence of leverage effects, the return-volatility relation is determined by interactions between risk premia and leverage effects. The model contains French, Schwert and Stambaugh (1987) as a special case. Our model provides a theoretical justification for mixed empirical findings of the return-volatility relation obtained by the previous studies without relying on any exogenous variables.

Many Lévy processes have already been investigated and applied in financial modeling. In this paper, I mainly focus on two very popular infinite activity Lévy processes: the Variance Gamma process (VG; Madan et al., 1998) and the Normal Inverse Gaussian process (NIG; Barndoff-Nielsen, 1998). Both processes have analytical characteristic functions, probability densities and Lévy densities and thus are more mathematically tractable. Indeed, they are two of the most popular infinite activity Lévy processes in finance and can suffice for many financial modeling purposes. Even though they are both infinitely active, they represent two types of infinite activity Lévy processes. The Variance Gamma process is of finite variation whereas the Normal Inverse Gaussian process takes on infinite variation.

In order to empirically investigate the return-volatility relation, I estimate the model with the Bayesian method. Contrary to the classical methods, Bayesian estimation regards all the parameters Θ and the states H in a model as random variables and tries to find their posterior distribution $p(\Theta, H|Y)$ conditional on the observations Y. Markov Chain Monte Carlo (MCMC) methods are usually applied to sample from this posterior distribution for parameter estimation, state estimation and model comparison. Bayesian estimation with MCMC is particularly suitable to the continuous-time financial models (Johannes and Polson, 2003). It simultaneously estimates parameters and latent states through computing their posterior distributions and delivers exact finite sample inferences. At the same time, it saves the computational time dramatically. In fact, for our stochastic jump intensity stochastic volatility models, the Bayesian method is the only efficient and easy-to-implement way to carry out estimation.

Ideally, the joint estimation with the return and option data is desirable in order to obtain both the historical and risk-neutral parameters and in order that the volatility estimate reflects information contained both in the stock market and options market (Pan, 2002; Eraker, 2004; Li et al., 2008). With the historical and risk-neutral parameters, we could explicitly test the return-volatility relation. But this method is usually computationally intensive and

is only feasible for the small dataset. I thus adopt a two-stage procedure. I firstly estimate the objective model with the return data alone which are long enough to contain the typical market behaviors: market crash, volatile market and tranquil market. This estimation results in direct estimates of both the diffusion volatility and the jump intensity (volatility), with which I then investigate the relationship between the return-volatility.

With the S&P 500 index daily data ranging from January 1986 to December 2000, I find that the infinite activity infinite variation NIG models perform better than the infinite activity finite variation VG models and furthermore the stochastic jump arrival rate models have larger flexibility. More importantly, I obtain the positive relationship between the conditional expected excess return and the diffusion volatility, which is only statistically significant in NIG model and the negative relationship between the conditional expected excess return and the jump volatility, which is statistically significant both in NIG and VG models. Our theoretical and empirical study verifies and improves the indirect evidence of French et al. (1987). I also observe that not only the individual volatility but also the aggregate volatility respond more to negative returns than to positive returns, indicating the existence of the strong diffusion and jump leverage effects.

The rest of paper is organized as follows. Section 2 builds the asset pricing model through time-changing both a Brownian motion and an infinite activity Lévy process and theoretically studies the return-volatility relation; Section 3 discusses the Bayesian inference with MCMC. Section 4 implements empirical study with the S&P 500 index data. Section 5 empirically investigates the return-volatility relation relying on the estimated volatility. Finally Section 6 concludes the paper.

2 Models

2.1 Infinite Activity Lévy Processes

Under a given probability space $(\Omega, \mathcal{F}, P)$ and the complete filtration $(\mathcal{F}_t)_{t\geq 0}$, a Lévy process X_t is the *cadlag* stochastic process with the independent and stationary increments and $X_0 = 0$. By the Lévy-Kintchine theorem, the characteristic function of X_t has the following form

$$\phi_X(u) = E[e^{iuX_t}] = e^{-t\psi_X(u)}, \tag{2}$$

where $\psi_X(u)$ is called the characteristic exponent

$$\psi_X(u) = -iu\mu + \frac{1}{2}u^2\sigma^2 + \int_{-\infty}^{\infty}(1 - e^{iux} + iux1_{|x|<1})\mathrm{v}(x)dx, \tag{3}$$

with $(\mu, \sigma, \mathrm{v})$ being the characteristic triplet and $u \in R$ being the characteristic index. The Lévy density $\mathrm{v}(x)$ measures the arrival rate of jumps with size x defined on R^0 (real line without zero). A Lévy process exhibits infinite activity when

$$\int_{R^0} \mathrm{v}(x)dx = \infty, \tag{4}$$

otherwise it is a finite activity process. The infinite activity Lévy process can generate an infinite number of jumps at any finite time interval, whereas the finite activity process can only generate a finite number of jumps within any finite time interval. If a Lévy process has

$$\int_{R^0} (|x| \wedge 1)\mathrm{v}(x)dx = \infty, \tag{5}$$

we say it is of infinite variation, otherwise it is of finite variation.

This paper applies two very popular infinite activity Lévy processes: the Variance Gamma process (Madan et al., 1998) and the Normal Inverse Gaussian process (Barndoff-Nielsen, 1998). Even though they are both infinitely active, they represent two types of infinite activity Lévy processes: the Variance Gamma process is of finite variation whereas the Normal Inverse Gaussian process takes on infinite variation. Both of them can be constructed through subordinating a Brownian motion with drift using an independent subordinator

$$X_t = \omega \mathcal{S}_t + \eta W^X(\mathcal{S}_t), \tag{6}$$

where W_t is a standard Brownian motion and $\mathcal{S}_t$ is a subordinator which is constructed to have unit mean rate and variance rate v. The subordinator is the Gamma process $\mathcal{S}_t = G(t; 1, v)$ with the probability density

$$f_G(x) = x^{\frac{t}{v}-1}\frac{1}{v^{\frac{t}{v}}\Gamma(\frac{t}{v})}e^{-\frac{x}{v}} \tag{7}$$

in the Variance Gamma process and the Inverse Gaussian process $\mathcal{S}_t = IG(t; 1, v)$

$$f_{IG}(x) = \frac{t}{\sqrt{2\pi v}} x^{-\frac{3}{2}} e^{-\frac{(x-t)^2}{2vx}} \tag{8}$$

in the Normal Inverse Gaussian process. With the Brownian subordination representation (6), their characteristic functions can be easily derived through following the Theorem 30.1 of Sato (1999)

$$\psi_{VG}(u) = \frac{1}{v} \log\left(1 - iu\omega v + u^2 \frac{\eta^2 v}{2}\right), \tag{9}$$

$$\psi_{NIG}(u) = \frac{1}{v}\left(\sqrt{1 - 2iu\omega v + u^2\eta^2 v} - 1\right), \tag{10}$$

and the Lévy densities and probability density functions can also be derived with the same theorem. Table 1 presents the first four cumulants of the Variance Gamma process and the Normal Inverse Gaussian process. We note that both of them have the same mean and variance, but the Normal Inverse Gaussian process can generate even skewer and fatter-tailed distribution. We also see that parameter ω mainly determines skewness, parameter v controls the kurtosis and parameter σ affects the variance level.

— Table 1 around here —

2.2 Asset Price Dynamics

Define a stopping time T_t given a nonnegative right continuous with left limit stochastic process V_t

$$T_t = \int_0^t V_{s-} ds,$$

which is finite almost surely. Intuitively, we could think of t as the calender time and T_t as the stochastic business time. The variable V_t is called the variance rate, which reflects the intensity of economic activity and the speed of information flow. For a stochastic process $X_t = X(t)$, its time-changed counterpart is defined by

$$X_{T_t} = X(T_t).$$

A well-known positive stochastic process is the square-root process (Cox et al., 1985), which is used in this paper as the variance rate process. Asset price dynamics is then modeled

though time-changing a Brownian motion and an infinite activity Lévy process

$$
\begin{aligned}
S_t &= S_0 \exp\Big\{ rt + \pi_W T_t^{(1)} + \pi_X T_t^{(2)} + \Big[W_{T_t^{(1)}} - \frac{1}{2} T_t^{(1)} \Big] \\
&\quad + \Big[X_{T_t^{(2)}} - k_X(1) T_t^{(2)} \Big] \Big\}, && (11) \\
T_t^{(1)} &= \int_0^t V_{s-}^{(1)} ds, && (12) \\
T_t^{(2)} &= \int_0^t V_{s-}^{(2)} ds, && (13) \\
dV_t^{(1)} &= \kappa^{(1)} (\theta^{(1)} - V_t^{(1)}) dt + \sigma^{(1)} \sqrt{V_t^{(1)}} dZ_t^{(1)}, && (14) \\
dV_t^{(2)} &= \kappa^{(2)} (1 - V_t^{(2)}) dt + \sigma^{(2)} \sqrt{V_t^{(2)}} dZ_t^{(2)}, && (15)
\end{aligned}
$$

where

- r is the deterministic risk-free interest rate, possibly time-varying;
- W_t is the Brownian motion and X_t is the infinite activity Lévy process which in this paper is the Variance Gamma process or the Normal Inverse Gaussian process, capturing both the frequently happened small jumps and the rarely happen large jumps;
- $T_t^{(1)}$ and $T_t^{(2)}$ are two stochastic business time which are used to generate the stochastic volatility for W_t and the stochastic jump arrival rate for X_t and $V_t^{(1)}$ and $V_t^{(2)}$ are two variance rates to construct the stochastic business time;
- π_W and π_X are the risk-premium rates for the diffusion and the jump and $k_X(1)$ is the convexity adjustment of X_t which is derived from its cumulant generating function: $k(s) \equiv \frac{1}{t} \log(E[e^{sX_t}]) = -\psi_X(-is)$;
- $Z_t^{(1)}$ and $Z_t^{(2)}$ are two other Brownian motions and independent each other. $Z_t^{(1)}$ is allowed to be correlated to W_t with the correlation parameter ρ and independent of X_t. $Z_t^{(2)}$ is independent of W_t and X_t.

Note that the long-run mean of $V_t^{(2)}$ is normalized to 1 in order to be consistent to the construction of the subordinator. It is constructed to have unit mean rate, but when its mean becomes stochastic, the long-run mean is still to be assumed unit.

The above model takes into account the stochastic diffusion volatility and the jump with stochastic arrival rate. The introduction of stochastic jump arrival rate makes the model more

flexible and can extend the jump effect to the relatively long horizon. This is consistent to the observation that the jumps are indeed autocorrelated and the market crashes are actually realized in a series of jumps over time. Furthermore, even under the lack of correlation between the return and diffusion volatility, the model can generate the time-varying/stochastic skewness and kurtosis through the time-changed jump process. This can be see from Table 1 with the calendar time replaced by the stochastic business time.

The conditional return variance V_t now have two sources: one is from the diffusion part (referred to as the diffusion volatility) and the other from the jump part (referred to as the jump volatility)

$$V_t = V_t^{(1)} + (\omega^2 v + \eta^2) V_t^{(2)}, \tag{16}$$

from which we can see that at market crash, the large jump intensity could contribute to the abrupt move of the return volatility. If we set $V_t^{(2)} = 1$ and $\sigma^{(2)} = 0$, we obtain the frequently applied constant jump arrival rate model.

We explicitly introduce the correlation between Brownian motions W_t and Z_t, which accommodates the so-called leverage effect from the diffusion part. The leverage effect of the jump is actually inherent in the above time-changed model because during a time of high variance rate, business time flows faster and price jumps occur at an increased rate. If the parameters ρ and ω are negative, both the diffusion and jump volatility react more to negative returns than to positive returns. In effect, empirical studies have always found negative values of ρ and ω.

In addition to the leverage effect, the model also implies the volatility feedback effect. This can be intuitively seen from the diffusion and jump risk premium parts $\pi_W T_t^{(1)}$ and $\pi_X T_t^{(2)}$ in the asset price process (12). If the volatility is really priced by the market, the parameters π_W and π_X should be positive. The market requires a higher return for the riskier asset. Empirical investigations using both the stock and options data find that the diffusion risk premium is usually small but positive and the jump claims a high positive risk premium (Pan, 2002; Li et al., 2008).

2.3 Return-Volatility Relation

To further investigate the interdependence between the return and volatility, we rewrite our model. For the time-changed Brownian motion, we have

$$W_{T_t^{(1)}} \stackrel{d}{=} \int_0^t \sqrt{V_t^{(1)}} dW_t, \tag{17}$$

where d indicates that the equality holds in distribution. With this property and the fact $[dW_t, dZ_t^{(1)}] = \rho dt$, the asset price process (12) and the variance rate process (14) can be reformulated into the following forms

$$\begin{aligned} \ln S_t &= \ln S_0 + rt + \Big(\pi_W - \frac{1}{2}\Big) T_t^{(1)} + \int_0^t \sqrt{V_t^{(1)}} \Big(\sqrt{1-\rho^2} dW_t^* + \rho dZ_t^{(1)}\Big) \\ &\quad + \Big(\pi_X - k_X(1)\Big) T_t^{(2)} + X_{T_t^{(2)}}, \end{aligned} \tag{18}$$

$$\sqrt{V_t^{(1)}} dZ_t^{(1)} = \frac{1}{\sigma^{(1)}} \Big[dV_t^{(1)} - \kappa^{(1)}(\theta^{(1)} - V_t^{(1)}) dt\Big], \tag{19}$$

where now W_t^* and $Z_t^{(1)}$ are independent. Note that the Brownian motion W_t^* in (18) is different from that in (12). Putting (19) into (18), we obtain

$$\begin{aligned} \ln S_t &= \ln S_0 + rt + \Big(\pi_W - \frac{1}{2}\Big) \int_0^t V_s^{(1)} ds + \frac{\rho}{\sigma^{(1)}} \int_0^t \Big[dV_s^{(1)} - \kappa^{(1)}(\theta^{(1)} - V_s^{(1)}) ds\Big] \\ &\quad + \sqrt{1-\rho^2} \int_0^t \sqrt{V_s^{(1)}} dW_s^* + \Big(\pi_X - k_X(1)\Big) \int_0^t V_s^{(2)} ds + X_{T_t^{(2)}}, \end{aligned} \tag{20}$$

from which we can derive the relationship between the excess return and volatility within a time interval $\big[t-\tau, t\big]$,

$$\begin{aligned} R_{t-\tau,t} - r\tau &= -\rho \frac{\kappa^{(1)}\theta^{(1)}}{\sigma^{(1)}} \tau + \frac{\rho}{\sigma^{(1)}} (V_t^{(1)} - V_{t-\tau}^{(1)}) + \Big(\pi_W - \frac{1}{2} + \rho \frac{\kappa^{(1)}}{\sigma^{(1)}}\Big) \int_{t-\tau}^t V_s^{(1)} ds \\ &\quad + \Big(\pi_X + \omega - k_X(1)\Big) \int_{t-\tau}^t V_s^{(2)} ds + v_t, \end{aligned} \tag{21}$$

where $v_t = \sqrt{1-\rho^2} \int_0^t \sqrt{V_s^{(1)}} dW_s^* + \eta \int_0^t \sqrt{V_s^{(2)}} dW_s^X$. It thus becomes clear that the not only the current but also the lagged diffusion volatility affect the the excess return.

The interdependence between the return and the current and lagged volatility differs from most of studies mentioned in Introduction which mainly focus on one relationship. Brandt

and Kang (2004) formally studies the non-contemporaneous interdependence using a latent VAR approach. Whitelaw (1994) and Lettau and Ludvigson (2001) also document the lead-lag phenomenon.

PROPOSITION: Under the model specification in Subsection 2.2 (and equation (21)) and the standard restrictions on ρ and ω: $\rho \leq 0$, $\omega \leq 0$, the conditional expected excess return has the following relationship with its conditional volatility,

- **Case I:** without leverage effects $\rho = 0$, $\omega = 0$

$$E_{t-\tau}\Big[R_{t-\tau,t} - r\tau\Big] = \Big(\pi_W - \frac{1}{2}\Big)\tau V_{t-\tau}^{(1)} + \Big(\pi_X - k_X(1)\Big)\tau V_{t-\tau}^{(2)}. \tag{22}$$

  This case is actually a variant of Merton's ICAPM (1). The constant term is zero; the excess return is positively related to the diffusion volatility and the jump volatility if the risks are really priced by the market.

- **Case II:** with leverage effects $\rho < 0$, $\omega < 0$

$$\begin{aligned} E_{t-\tau}\Big[R_{t-\tau,t} - r\tau\Big] &= -\rho\frac{\kappa^{(1)}\theta^{(1)}}{\sigma^{(1)}}\tau + \Big(\pi_W - \frac{1}{2} + \rho\frac{\kappa^{(1)}}{\sigma^{(1)}}\Big)\tau V_{t-\tau}^{(1)} \\ &\quad + \Big(\pi_X - k_X(1) + \omega\Big)\tau V_{t-\tau}^{(2)}, \end{aligned} \tag{23}$$

  where I use a property of the square-root process: $E_{t-\tau}[V_t^{(1)}] = V_{t-\tau}^{(1)} e^{-\kappa^{(1)}\tau} + \theta^{(1)}(1 - e^{-\kappa^{(1)}\tau})$, which is approximately equal to $V_{t-\tau}^{(1)}$ for the small time interval τ. In this case, the constant term is now non zero and should be positive. The relationships between the conditional expected excess return and the conditional diffusion and jump volatility are determined by interactions between risk premia and leverage effects.

- **Case III:** French, Schwert and Stambaugh (1987)
  From the Case II, if we define the unexpected diffusion volatility as $V_t^{u(1)} = V_t^{(1)} - V_{t-\tau}^{(1)}$, we have

$$\begin{aligned} R_{t-\tau,t} - r\tau &= -\rho\frac{\kappa^{(1)}\theta^{(1)}}{\sigma^{(1)}}\tau + \frac{\rho}{\sigma^{(1)}}V_t^{u(1)} + \Big(\pi_W - \frac{1}{2} + \rho\frac{\kappa^{(1)}}{\sigma^{(1)}}\Big)\tau V_{t-\tau}^{(1)} \\ &\quad + \Big(\pi_X - k_X(1) + \omega\Big)\tau V_{t-\tau}^{(2)} + v_t. \end{aligned} \tag{24}$$

  This is a variant of French, Schwert and Stambaugh (1987) except that we have a jump

volatility term. The excess return now has a negative relation to the unexpected change in the (diffusion) volatility and this negative relation provides indirect evidence of a positive relation between the excess return and the predicted (diffusion) volatility.■

The model (Case II) explains why the empirical investigations have obtained mixed findings since risk premia and leverage effects play opposite roles in determining the relationships of the return and the diffusion and jump volatility. Clearly, different data, volatility estimate, or the addition of exogenous variables may result in different findings. Bollerslev and Zhou (2006) also find that the existence of the leverage effect makes it difficulty to reveal positive relationship between the return and volatility. Our model not only provides a theoretical foundation but also introduces an extra term of the jump volatility, which is not studied by nearly every work mentioned above.

3 Bayesian Estimation

In this section, I firstly discuss the discretization of the model and construct the state-space model representation in subsection 3.1; then briefly describe the implementation of MCMC in subsection 3.2 and finally the model comparison under Bayesian framework is introduced in subsection 3.3.

3.1 Model Discretization

We have noted that for the time-changed Brownian motion, we have the distributional equivalence (17) and for the time-changed infinite activity Lévy process, with its Brownian subordination property we have

$$X_{T_t^{(2)}} = \omega \mathcal{S}_{T_t^{(2)}} + \eta W(\mathcal{S}_{T_t^{(2)}}), \tag{25}$$

$$\mathcal{S}_{T_t^{(2)}} = \mathcal{S}(T_t^{(2)}; 1, v). \tag{26}$$

With these features, we could write the model into a state-space representation. Since using the stock price data alone, we have a identification problem for estimating the risk-premium parameters. I thus merge them with the drift and convexity adjustment terms and estimate one parameter μ^*. Having defined log stock price as $y_t = \ln S_t$ and discretized the

model with small time interval τ, we get the following state-space representation of the model

$$
\begin{aligned}
y_t &= y_{t-\tau} + \mu^*\tau + X_\tau + \sqrt{\tau V_{t-\tau}^{(1)}} w_t, && (27)\\
X_\tau &= \omega \mathcal{S}_\tau + \eta\sqrt{\mathcal{S}_\tau}\tilde{w}_t, && (28)\\
\mathcal{S}_\tau &= \mathcal{S}(\tau V_{t-\tau}^{(2)}; 1, v), && (29)\\
V_t^{(1)} &= V_{t-\tau}^{(1)} + \kappa^{(1)}\left(\theta^{(1)} - V_{t-\tau}^{(1)}\right)\tau + \sigma^{(1)}\sqrt{\tau V_{t-\tau}^{(1)}} z_t^{(1)}, && (30)\\
V_t^{(2)} &= V_{t-\tau}^{(2)} + \kappa^{(2)}\left(1 - V_{t-\tau}^{(2)}\right)\tau + \sigma^{(2)}\sqrt{\tau V_{t-\tau}^{(2)}} z_t^{(2)}, && (31)
\end{aligned}
$$

where the variance rates $V_t^{(1)}$, $V_t^{(2)}$, jump size X_t and jump time $\mathcal{S}_t$ are unobservable and regarded as states $H = \{V_t^{(1)}, V_t^{(2)}, X_t, \mathcal{S}_t\}_{t\geq 0}$. We also have the following parameters to be estimated $\Theta = \{\mu^*, \omega, \eta, v, \alpha, \kappa^{(1)}, \theta^{(1)}, \sigma^{(1)}, \rho, \kappa^{(2)}, \sigma^{(2)}\}$. Eraker et al. (2003) find that for the small time interval τ, like daily frequency or higher, Euler discretization of the continuous-time models does not introduce significant bias in estimation.

3.2 MCMC Implementation

Bayesian estimation tries to find the posterior distribution of parameters and states given the whole set of observations, that is, $p(\Theta, R|Y)$ where $Y = \{y_t\}_{t\geq 0}$. With Bayes' rule, this posterior can be derived with the likelihood and the prior

$$p(\Theta, H|Y) \propto p(Y|H, \Theta)p(H|\Theta)p(\Theta), \tag{32}$$

where $p(Y|H, \Theta)$ is the likelihood given states and parameters; $p(H|\Theta)$ is the probability distribution of states conditional on parameters and $p(\Theta)$ is the prior distribution of parameters.

In most cases, direct sampling from the posterior distribution $p(\Theta, H|Y)$ is impossible because of its high dimension and complicated form. We could then iteratively draw from its full conditionals $p(\Theta|H, Y)$ and $p(H|\Theta, Y)$ using Gibbs sampling method. Parameter set Θ and state set H can further be broken into smaller blocks.

Since the diffusion volatility and returns are correlated, for Bayesian estimation of the correlation parameter ρ and the volatility of volatility parameter $\sigma^{(1)}$, I follow Jacquier, Polson and Rossi (2004) and reparameterize $(\rho, \sigma^{(1)})$ to (ξ, h) with $\xi = \rho\sigma^{(1)}$ and $h = (1-\rho^2)(\sigma^{(1)})^2$. It is not difficult to derive that all the parameters except v have conjugate priors in both VG and NIG cases, that is, μ^*, $\kappa^{(1)}$, $\theta^{(1)}$, $\kappa^{(2)}$, $\theta^{(2)}$, ω and ξ are normals and h, $(\sigma^{(2)})^2$ and η^2 are

inverse Gammas. Therefore, their posteriors have the same distributions as those of the priors and the standard methods can be used to sample from these posteriors. For the parameter v, in VG case it has a non-standard posterior distribution and MCMC sampling methods have to be used and in NIG case, v has the inverse Gamma conjugate prior.

For Bayesian estimation of states, I follow the single-move approach. We can derive that the jump size X_t has a normal posterior and the volatility states $V_t^{(1)}$ and $V_t^{(2)}$ have non-standard posteriors in both VG and NIG cases and again to sample from them MCMC methods have to be applied.

I also find that the posterior distribution of $\mathcal{S}_t$ has a nice form in NIG case. It has the generalized inverse Gaussian form $GIG(\lambda, \psi, \chi)$

$$p_{NIG}(\mathcal{S}_t|Y, H_{-\mathcal{S}_t}, \Theta) \propto \mathcal{S}_t^{-2} \exp\left\{-\frac{1}{2}\left[\left(\frac{1}{v}+\frac{\omega^2}{\eta^2}\right)\mathcal{S}_t + \left(\frac{(\tau V_{t-1}^{(2)})^2}{v}+\frac{X_t^2}{\eta^2}\right)\mathcal{S}_t^{-1}\right]\right\}, \tag{33}$$

where $\lambda = -1$, $\psi = 1/v + \omega^2/\eta^2$ and $\chi = (\tau V_{t-1}^{(2)})^2/v + X_t^2/\eta^2$. I thus sample from (33) using the algorithm proposed by Dagpunar (1989) which gives independent samples. In VG model, the posterior distribution of $\mathcal{S}_t$ is also in the generalized inverse Gaussian form

$$p_{VG}(\mathcal{S}_t|Y, H_{-\mathcal{S}_t}, \Theta) \propto \mathcal{S}_t^{\frac{\tau V_{t-1}^{(2)}}{v}-\frac{3}{2}} \exp\left\{-\frac{1}{2}\left[\left(\frac{2}{v}+\frac{\omega^2}{\eta^2}\right)\mathcal{S}_t + \frac{X_t^2}{\eta^2}\mathcal{S}_t^{-1}\right]\right\}. \tag{34}$$

But the algorithm of Dagpunar (1989) can't efficiently draw samples since GIG parameter $\chi = \frac{X_t^2}{\eta^2}$ is usually very small in this case and I thus rely on MCMC methods.

To sample from the non-standard posterior distributions, I mainly use the slice sampling method (Neal, 2003). Slice sampler works with the following steps for a posterior distribution $p(x)$ of interest:

Step 1: Starting from an initial value x_0;

Step 2: Draw a real value y, uniformly from $(0, p(x_0))$;

Step 3: Find an interval around x_0, that contains all, or much, of the slice $S = \{x : y < p(x)\}$;

Step 4: Draw the new point, x, uniformly from the part of the slice within this interval as a sample from the distribution.

Slice sampling method can adaptively change the scale in choosing slice, which makes it easier to tune than Metropolis-Hastings algorithm and also avoids problems that arise when

the appropriate scale of changes varies over the time (Neal, 2003). This adaptive property is particularly suitable to draw samples from the posterior distributions of states.

3.3 DIC and Model Comparison

In Bayesian framework, ideally we should use Bayes factor to compare the different models. It is based on the computation of the *evidence*, the probability of getting the data given the model. However, it is not easy to obtain an accurate estimation of the evidence using MCMC sampling. Kim et al. (1998) and Chib et al. (2002) apply particle filters to compute Bayes factor, but for the infinite activity stochastic volatility models, Li (2008) shows that with the underlying security alone, particle filters can be very inefficient. Instead of computing the evidence, Spiegelhalter et al. (2002) propose using Deviance Information Criterion (DIC), which tackles the issues of goodness-of-fit and model complexity using an approximate decision-theoretic justification. Indeed, DIC can be shown to be equivalent to the evidence when the deviance is Gaussian. It is particularly convenient for the complex hierarchical models.

The deviance is defined as the posterior distribution of the log likelihood

$$D(\vartheta) = -2 \ln f(y|\vartheta) + 2 \ln g(y), \tag{35}$$

where ϑ is the model parameters, probably augmented by the states and $g(y)$ is a standardizing term that doesn't affect model comparison and will be ignored. The goodness-of-fit is then summarized by the posterior expectation of the deviance

$$\bar{D} = E_{\vartheta|y}[D(\vartheta)] = E_{\vartheta|y}[-2 \ln f(y|\vartheta)] \tag{36}$$

and the complexity is given by the expected deviance minus the deviance evaluated at the posterior mean $\bar{\vartheta}$ of the parameters

$$p_D = E_{\vartheta|y}[D(\vartheta)] - D(E_{\vartheta|y}[\vartheta]). \tag{37}$$

p_D can be interpreted as the effective number of parameters in the model. DIC is then defined by combining (36) and (37) as

$$\text{DIC} \quad = \quad \bar{D} + p_D. \tag{38}$$

The advantage of this measure is that it is trivial to compute relevant quantities when performing MCMC on the models. DIC could be regards as a Bayesian version or generalization of the Akaike information criterion (AIC). Recently it is successfully used to compare stochastic volatility models (Berg et al., 2004) and its use in missing data models is fully investigated by Celeux et al. (2006).

4 Estimation Results

4.1 Data

The data used to estimate the statistical parameters are S&P 500 stock index from January 1986 to December 2000 with daily frequency, totally 3791 observations and is downloaded from *Datastream*. This period is long enough to contain the typical market behaviors we could observe: market crash on October 19, 1987 (-22.9%), the volatile market after 1997 and relatively tranquil period between 1993 and 1996. Figure 1 plots the time-series evolution of the index and index returns and Table 4 first column gives the descriptive statistics of the index returns.

— Figure 1 around here —

We find that the annualized mean value of returns in this period is around 12.2% and the historical volatility is 16.9%. A striking feature of data is the high non-normality of the return distribution with the skewness -2.96 and kurtosis 64.55. The Jarque-Bera test easily rejects the null hypothesis of normality of returns with the very small p-value.

4.2 Bayesian Parameter and State Estimation

Two groups, totally four models are estimated with the Bayesian method discussed in Section 3. They are VG/NIG constant jump arrival rate models (referring to as VGSV and NIGSV) and VG/NIG stochastic jump arrival rate models (referring to as VG2SV and NIG2SV).

Table 2 presents the parameter estimates. Posterior mean, standard deviation (in bracket) and 95% credible interval (in square brackets) are reported.

— Table 2 around here —

We first look at the constant jump arrival rate models. The estimates of the diffusion volatility processes in VGSV and NIGSV models are very similar. The long-run mean param-

eter θ in each model is around 15.5%, smaller than that of the historical volatility (16.9%) as expected; the volatility of volatility parameters σ are 0.263 and 0.253 respectively and the mean-reverting parameter κ of VGSV model is a little bit smaller than that of NIGSV model with values 2.241 and 2.338 respectively. However, the jump parameters are different in both models, especially the jump structure parameter v, indicating that the VG process and NIG process are two different Lévy processes. As we know from the Section 2 that VG process is of finite variantion whereas NIG process takes on infinite variation.

Bayesian estimation can also give us state estimates. Figure 2 plots the posterior means of the diffusion volatility and jumps. We find that the general shapes of the diffusion volatility and jumps are similar in VGSV and NIGSV models, but in NIGSV model, the jump component plays more important role than in VGSV model. This can be seen from the estimated diffusion volatility and jump at market crash or volatile market. We know that at market crash on October 19, 1987, there is a 22.9% downward of S&P 500 index return. The estimated jump size in VGSV is less than 19% but it is larger than 21% in NIGSV; correspondingly, the estimated diffusion volatility is larger in VGSV than in NIGSV.

— Figure 2 around here —

We now move to the stochastic jump arrival rate models. First of all, we observe that the jump structure parameters in VG2SV and NIG2SV models are nearly the same as those in VGSV and NIGSV models respectively. This is not surprising since in constant jump arrival rate models the subordinators have the unit mean and variance rate v and in stochastic jump arrival rate models even though the subordinators have stochastic mean and variance rate, their long-run means are still unit and v. Secondly, the mean-reverting parameters and the volatility of volatility parameters become smaller, but the long-run means of the diffusion volatility nearly keep unchanged. These facts imply the diffusion volatility processes in stochastic jump arrival rate models are more persistent and the large part of return variation is contributed by the jump component. Finally and more interestingly, the parameters of the stochastic jump arrival rate processes in both models are all highly significant. We note that the mean-reverting parameter $\kappa^{(2)}$ and the volatility of volatility parameter $\sigma^{(2)}$ in each of these two models are very large, 65.5 and 9.1 respectively in VG2SV model and 71.9 and 9.7 respectively in NIG2SV model.

From the estimated parameters of the diffusion volatility processes and the jump arrival

rate processes, we could deduce that the diffusion volatility is very persistent and displays the long-memory effect and it plays role over the long horizon; however, for the jump arrival rate processes, on the one hand the fast mean reversion implies that its effect is not as persistent as the diffusion volatility and hence dominate the return behavior in the short horizon and on the other hand, the large volatility of volatility can extend its effect to relatively long horizon. Furthermore, the large mean-reverting parameter and volatility of volatility parameter of the jump arrival rate process together imply that the return volatility could be abruptly moved up through a sudden increase of the jump intensity.

The estimated jump arrival rates in Figure 3 really reveal the stochastic/time-varying jump arrival rate. But we note that its behavior in VG2SV and NIG2SV is very different, implying from the other aspect the different jump structures between VG process and NIG process. Now the return variance has two sources: one is from the diffusion part and the other from the jump part. We can see from Figure 3 that the contribution of the diffusion volatility is decreased since the estimated diffusion volatility is uniformly lower than those in the constant jump arrival rate models. The jump component in NIG2SV model still plays larger role than that in VG2SV model, same as in the constant jump arrival rate models.

— Figure 3 around here —

The estimates of ρ and ω are both negative and significant, indicating the existence of the diffusion and jump leverage effects. We thus can see from the Proposition of Section 2 that the return-volatility relation holds for the Case II (and Case III). The return should have a negative relation with the unexpected diffusion volatility. Since we have problems to identify the risk-premium parameters in estimation with the stock price data alone, we can not explicitly determine the relationships between the conditional expected excess return and the conditional diffusion volatility as well as the conditional expected excess return and the conditional jump volatility. I will investigate this issue more detailedly in the next section.

4.3 Model Comparison

A couple of questions naturally arise: in modeling S&P 500 index returns and uncovering the return-volatility relation, which model/jump process is more suitable? and whether is it necessary to introduce the stochastic jump arrival rate?

We firstly check return residuals which are defined as

$$\epsilon_t^y = \frac{y_{t+\tau} - y_t - \mu^* \tau - X_{t+\tau}}{\sqrt{V_t^{(1)} \tau}} \sim N(0,1) \tag{39}$$

for each of four models. Intuitively, if a model is correctly specified, the residuals ϵ_t^y should be asymptotically normally distributed with mean zero and variance 1. Table 3 upper panel presents the first four moments of the return residuals for four models: mean, standard deviation, skewness and kurtosis. We can see, especially from the skewness and kurtosis that NIG models perform better than VG models and the stochastic jump arrival rate models are superior to the constant jump arrival rate models. All in all, NIG2SV is the best model.

— Table 3 around here —

Recently, some empirical studies argue that a jump component is also necessary in the volatility process (Eraker et al., 2004; Broadie et al., 2007). Table 3 middle panel gives the skewness and kurtosis of the diffusion volatility residuals, which is

$$\epsilon_t^V = \frac{V_{t+\tau}^{(1)} - V_t^{(1)} - \kappa^{(1)}(\theta^{(1)} - V_t^{(1)})\tau}{\sigma^{(1)} \sqrt{V_t^{(1)} \tau}}. \tag{40}$$

If there really needs a jump component, the volatility residuals should have positive skewness and large excess kurtosis (Pan, 2002; Eraker, 2004; Broadie et al., 2007). We see from the table that the skewness is negative and small and the kurtosis is not far away from 3 for each of these four models. These results indicate that when we model the asset price dynamics with the infinite activity Lévy processes, we do not need a jump component in the volatility process. Why could this happen? To compare our results with those in Eraker et al. (2003) and Eraker (2004), we find that the jump component plays much more important roles in the infinite activity jump models than in the finite activity Poisson models, especially at market crash and volatile markets and that the large values of $\kappa^{(2)}$ and $\sigma^{(2)}$ can easily increase the return volatility to a high level in short time.

To formally conduct model comparison, Bayes factors are desirable. But for the stochastic volatility infinite activity jump models, the computation of Bayes factors is usually very difficult. I thus rely on Deviance Information Criterion (DIC) discussed in Subsection 3.3. A recent paper by Berg et al. (2004) has illustrated the potential advantages of this information

criterion in determining the appropriate (discrete-time) stochastic volatility models. Table 3 lower panel presents the criteria of goodness-of-fit ($\bar{D}$) , model complexity (p_D) and Deviance Information Criterion (DIC) for all models. We find that in fitting data, VG models are better than NIG models, but they also have larger values of model complexity. When taking into account both goodness of fit and model parsimoneity, we see that the values of DIC in NIG models are smaller than those in VG models, indicating NIG models perform better than VG models. NIG2SV model is again the best one from this criterion.

5 Evidence of Return-Volatility Relation

The Bayesian estimation results in both parameter and volatility estimates. Since the risk-premium parameters can not be identified using the stock price data alone, the relationships between the conditional expected excess return and the conditional volatility can not be completely determined using parameter estimates. In this section, I investigate these relationships with the volatility estimates obtained from the previous section. This is also a backtesting procedure. If our model is correctly specified, we could use the estimated volatility to recover the relationship indicated by the Proposition of Section 2.

Since the stochastic jump arrival rate models are preferred from Section 4, in this section I mainly focus on NIG2SV and VG2SV models. Table 4 presents the summary statistics of returns and the estimated volatility (both variance and standard deviation). As we discussed before, returns are negatively skewed with very fat tails. They are approximately serially uncorrelated. However, the estimated diffusion volatility $V_t^{(1)}$ is highly persistent. We see that the first six autocorrelations in both models are all larger than 0.96. The estimated jump volatility also shows persistence, but not as strong as the estimated diffusion volatility. The sixth autocorrelations are 0.46 and 0.41 in VG2SV and NIG2SV models, respectively. We also observe that the estimated jump volatility of NIG2SV model is more positively skewed and has fatter tails than VG2SV model.

— Table 4 around here —

I firstly investigate the return-volatility relation implied by the Proposition. As indicated by Case II and the formula (23), the relationships are determined by interactions between risk

premia and leverage effects. I thus estimate the empirical equivalence to (23),

$$ER_t = \alpha_0 + \alpha_1 V_{t-1}^{(1)} + \alpha_2 V_{t-1}^{(2)} + \xi_t \tag{41}$$

where ER_t denotes the excess returns between S&P 500 index returns and the risk-free rates which are proxied by the yields of US 3-month treasury bill. In order to check the robustness, I also estimate the standard deviation version

$$ER_t = \alpha_0 + \alpha_1 \sqrt{V_{t-1}^{(1)}} + \alpha_2 \sqrt{V_{t-1}^{(2)}} + \xi_t. \tag{42}$$

We would also like to investigate the indirect evidence of French et al. (1987) in Case III. Thus, the following two regressions are estimated

$$ER_t = \alpha_0 + \alpha_1 V_t^{u(1)} + \alpha_2 V_{t-1}^{(1)} + \alpha_3 V_{t-1}^{(2)} + \xi_t, \tag{43}$$

$$ER_t = \alpha_0 + \alpha_1 \sqrt{V_t^{u(1)}} + \alpha_2 \sqrt{V_{t-1}^{(1)}} + \alpha_3 \sqrt{V_{t-1}^{(2)}} + \xi_t, \tag{44}$$

where $V_t^{u(1)} = V_t^{(1)} - V_{t-1}^{(1)}$ is defined as the *ex post* unexpected diffusion volatility.

The parameter estimates, along with their standard errors based on the Newey-West method with 18 lags are reported in Table 5. In the left panel, I use the variance as the regressors and in the right panel, standard deviation is the regressors. We first look at the Case II. The constant coefficients are all significantly positive and small, consistent with the theory in the Proposition. We obtain the positive estimate of α_1 in both models, indicating that the excess return is positively related to the conditional diffusion volatility. But this estimate is statistically significant only in NIG2SV model. We get a significant negative estimate of α_2 in both models, implying that the excess return is negatively related to the conditional jump volatility. The goodness-of-fit (R^2) is small, about 1.00% in VG2SV model and a litter bit biger in NIG2SV model (2.92%). All these results hold in the standard deviation case.

— Table 5 around here —

We now investigate French et al. (1987) model (Case III). Looking at the left panel, we have findings which are consistent to the theory. First, goodness-of-fit is improved dramatically. For example, R^2 of NIG2SV model is increased to 82.5%. Second, again we have

significant small but positive constant estimates. Third, the estimate of α_1 in both models is negative and significant, indicating the existence of the indirect evidence. We have negative estimate of α_2, which is statistically insignificant in NIG2SV model or marginally significant in VG2SV model. Finally, the estimate of α_3 is significantly negative, the same as before. The results are robust in the standard deviation regressions.

The negative estimate of α_2 provides an indirect evidence of the positive relationship between the conditional excess return and the conditional diffusion volatility as explained in French et al. (1987). If the current diffusion volatility is larger than the predicted, the predicted diffusion volatility will be revised upward for all future periods because of its property of persistence/long memory. If the excess return is positively related to the conditional diffusion volatility, the discount rate for future cash flows will increase. The higher discount rate reduces both their present value and the current stock price if the cash flows are not affected.

We have a completely new finding, that is, the conditional excess return is negatively related to the conditional jump volatility. Before explaining this finding, we make an extension of Case III regressions by incorporating the unpredicted jump volatility even though our model doesn't imply this point,

$$ER_t = \alpha_0 + \alpha_1 V_t^{u(1)} + \alpha_2 V_{t-1}^{(1)} + \alpha_3 V_t^{u(2)} + \alpha_4 V_{t-1}^{(2)} + \xi_t, \tag{45}$$

$$ER_t = \alpha_0 + \alpha_1 \sqrt{V_t^{u(1)}} + \alpha_2 \sqrt{V_{t-1}^{(1)}} + \alpha_3 \sqrt{V_t^{u(2)}} + \alpha_4 \sqrt{V_{t-1}^{(2)}} + \xi_t, \tag{46}$$

where $V_t^{u(2)} = V_t^{(2)} - V_{t-1}^{(2)}$ is defined as the *ex post* unexpected jump volatility.

Table 5 panel "Extension" presents parameter estimates and Newey-West standard errors. We find positive and significant estimate of α_3 in both models no matter using the variance or standard deviation. The estimate of α_4 is negative and also significant. To compare with the Case III, the diffusion volatility related parameters and R^2s do not change much. This clearly indicates that the diffusion volatility and the jump play very different roles in controlling the return dynamics.

One explanation for this positive α_3 and negative α_4 is very similar with that for diffusion volatility but in opposite way. If the current jump volatility is larger than the predicted, the predicted jump volatility will be revised not upward but downward for all future periods because of its property of non-persistence/short memory. This means that in the future, the market gives a smaller probability that the stock price jumps downward and this will make

the current stock price higher. This also implies that the market does not price the expected jumps.

We now empirically investigate the leverage effects with our diffusion and jump volatility estimates. Even though both the leverage effects and the volatility feedback effect can cause the asymmetric volatility. The causality process for these two effect are different. While the volatility feedback effect asserts that the volatility increase enhances the required return on asset, the leverage effect claims that return shocks lead to changes in the conditional volatility. Thus, here I investigate the leverage effects with the following regressions

$$V_t = \alpha_0 + \alpha_1 ER_t + v_t, \tag{47}$$

$$\sqrt{V_t} = \alpha_0 + \alpha_1 ER_t + v_t, \tag{48}$$

where V_t can be the diffusion volatility, the jump volatility or the return volatility which is defined in equation (16) as $V_t = V_t^{(1)} + (\omega^2 v + \eta^2) V_t^{(2)}$ for both models.

Table 6 presents the parameter estimates and their standard errors computed with the Newey-West method (18 lags). In both the variance and standard deviation cases, the estimate of constant α_0 is positive and significant and the slope estimate α_1 is significantly negative in each of these two models, indicating the existence of the strong diffusion and jump leverage effects.

— Table 6 around here —

6 Concluding Remarks

This paper investigates the return-volatility relation by taking into account model specification problem. The stock price process is modeled by the time-changed Brownian motion and infinite activity Lévy process, which introduces not only the stochastic diffusion volatility but also the stochastic jump intensity. The model sheds new light on the return-volatility relation. It indicates that under the absence of leverage effects it becomes a variant of the Merton's ICAPM whereas under the existence of leverage effects, the return-volatility relation is determined by interactions between risk premia and leverage effects parameters. The model contains French, Schwert and Stambaugh (1987) as a special case. Our model provides a theoretical justification for mixed empirical findings of the return-volatility relation obtained

by the previous studies without relying on any exogenous variables. The empirical study finds a positive return-diffusion volatility relation and more interestingly a negative return-jump volatility relation. The study also find that not only the individual volatility but also the aggregate volatility respond more to negative returns than to positive returns, indicating the existence of the strong diffusion and jump leverage effects.

References

[1] Andersen, T.G., L. Benzoni, and J. Lund. (2002). "An Empirical Investigation of Continuous-Time Equity Return Models." *Jornal of Finance* 57, 1239-1284.

[2] Bakshi, G., C. Cao and Z. Chen. (1997). "Empirical Performance of Alternative Option Pricing Models." *Jornal of Finance* 52, 2003-2049.

[3] Bali, T. and L. Peng. (2006). "Is There a Risk-Return Trade-Off? Evidence From High-Frequency Data." *Journal of Applied Econometrics* 21, 1169-1198.

[4] Barndorff-Nielson, O.E. (1998). "Processes of Normal Inverse Gaussian Type." *Finance and Stochastics* 2, 41-68.

[5] Bates, David S. (2000). "Post-'87 Crash Fears in the S&P 500 Futures Option Market." *Jornal of Econometrics* 94, 181-238.

[6] Black, F. (1976). "Studies in Stock Price Volatility Changes." *Proceedings of the 1976 Business Meeting of the Business and Economics Section*, 177-181.

[7] Berg, A., R. Meyer and J. Yu. (2004). "Deviance Information Criterion for Comparing Stochastic Volatility Models." *Journal of Business & Economic Statistics* 22, 107-120.

[8] Bollerslev, T. and H. Zhou. (2006). "Volatility Puzzles: A Simple Framework for Gauging Return-Volatility Regressions." *Journal of Econometrics* 131, 123-150.

[9] Brandt, M.W. and Q. Kang (2004). "On the Relationship between the Conditional Mean and Volatility of Stock Returns: A Latent VAR Approach." *Journal of Financial Economics* 72, 217-257.

[10] Breen, W., L.R. Glosten and R. Jagannathan. (1989). "Economic Significance of Predictable Variations in Stock Index Returns." *Journal of Finance* 44, 1177-1189.

[11] Broadie, M., M. Chernov and M. Johannes. (2007). "Model Specification and Risk Premia: Evidence from Futures Options." *Journal of Finance* 62, 1453-1490.

[12] Campbell, J.Y. (1987). "Stock Returns and the Term Structure." *Journal of Financial Economics* 18, 373-399.

[13] Campbell, J.Y. and L. Hentschel. (1992). "No News Is Good News: An Asymmetric Model of Changing Volatility in Stock Returns." *Journal of Financial Economics* 31, 281-318.

[14] Carr, P., H. Geman, D.B. Madan and M. Yor. (2002). "The Fine Structure of Asset Returns: An Empirical Investigation." *The Journal of Business* 75, 305-332.

[15] Carr, P., H. Geman, D.B. Madan and M. Yor. (2003). "Stochastic Volatility for Lévy Processes." *Mathematical Finance* 13, 345-382.

[16] Celeux, C., F. Forbes, C.P. Robert and D.M. Titteringto. (2006). "Deviance Information Criteria for Missing Data Models." *Bayesian Analysis* 1, 651-674.

[17] Chib, S., F. Nardari and N. Shephard. (2002). "Markov Chain Monte Carlo Methods for Stochastic Volatility Models." *Journal of Econometrics* 108, 281-316.

[18] Chou, R.Y. (1988). "Volatility Persistence and Stock Evaluations: Some Empirical Evidence Using GARCH." *Journal of Econometrics* 3, 279-294.

[19] Chou, R.Y., R.F. Engle and A. Kane. (1992). "Measuring Risk Aversion from Excess Returns on a Stock Index." *Journal of Econometrics* 52, 201-224.

[20] Christie, A.A.. (1982). "The Stochastic Behavior of Common Stock Variances–Value, Leverage and Interest Rate Effect." *Journal of Financial Economics* 10, 407-432.

[21] Cont, R. and P. Tankov. (2004). *Financial Modeling With Jump Processes*. London: Chapman & Hall/CRC.

[22] Cox, J. C., J.E. Ingersoll and S.A. Ross. (1985). "A Theory of the Term Structure of Interest Rates." *Econometrica* 53, 385-408.

[23] Dagpunar, J. (1989). "An Easily Implemented Generalized Inverse Gaussian Generator." *Communications in Statistics-Simulations* 18, 703-710.

[24] Eraker, B. (2004). "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices." *Journal of Finance* 59, 1367-1404.

[25] Eraker, B., M. Johannes and N. Polson. (2003). "The Impact of Jumps in Equity Index Volatility and Returns." *Journal of Finance* 58, 1269-1300.

[26] French, K.R., G.W. Schwert and R.F. Stambaugh. (1987). "Expected Stock Returns and Volatility." *Journal of Financial Economics* 19, 3-30.

[27] Geman, H. (2002). "Pure Jump Lévy Processes for Asset Price Modeling." *Journal of Banking and Finance* 26, 1297-1316.

[28] Ghysels, E., P. Santa-Clara and R. Valkanov. (2004). "There Is a Risk-Return Tradeoff After All." *Journal of Financial Economics* 76, 509-548.

[29] Glosten, L.R., R. Jagannathan and D. Runkle. (1993). "On the Relation between the Expected Value and the Volatility of the Nomial Excess Return on Stocks." *Journal of Finance* 48, 1779-1801.

[30] Guo, H. and R.F. Whitelaw. (2006). "Uncovering the Risk-Return Relation in the Stock Market." *Journal of Finance* 61, 1433-1463.

[31] Harrison, P. and H.H. Zhang. (1999). "An Investigation of the Risk and Return Relation at Long Horizons." *Review of Economics and Statistics* 81, 399-408.

[32] Harvey, C.R. (2001). "The Specification of Conditional Expectations." *Journal of Empirical Finance* 8, 573-637.

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

[34] Huang, J. and L. Wu. (2004). "Specification Analysis of Option Pricing Models Based on Time-Changed Lévy Processes." *The Journal of Finance* 59, 1405-1439.

[35] Jacquier, E., N. Polson and P.E. Rossi. (1994). "Bayesian Analysis of Stochastic Volatility Models." *Journal of Business & Economic Statistics* 12, 371-417.

[36] Jacquier, E., N. Polson and P.E. Rossi. (2004). "Bayesian Analysis of Stochastic Volatility Models with Fat-tails and Correlated Errors." *Journal of Econometrics* 122, 185-212.

[37] Johannes, M. and N. Polson. (2003). "MCMC Methods for Continuous-Time Financial Econometrics." In L.P. Hansen and Y. Ait-Sahalia (eds), *Handbook of Financial Econometrics*. Elsevier: Amsterdam. Forthcoming.

[38] Kim, S., N. Shephard and S. Chib. (1998). "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models." *Review of Economic Studies* 65, 361-393.

[39] Lettau, M. and S. Ludvigson. (2004). "Measuring and Modeling Variation in the Risk-Return Tradeoff." In L.P. Hansen and Y. Ait-Sahalia (eds), *Handbook of Financial Econometrics*. Elsevier: Amsterdam. Forthcoming.

[40] Li, H., M.T. Wells and C.L. Yu. (2006). "A Bayesian Analysis of Return Dynamics with Lévy Jumps." *Review of Financial Studies*, 21, 2345-2378.

[41] Madan, D. (2001). "Financial Modeling with Discontinuous Price Processes." In O.E. Barndorff-Nielson, T. Mikosch and S. Resnick (eds), *Lévy Processes: Theory and Applications*. Boston: Birkhauser.

[42] Madan, D., P. Carr and E. Chang. (1998). "The Variance Gamma Process and Option Pricing." *European Finance Review* 2, 79-105.

[43] Neal, R.M. (2003). "Slice Sampling (with Discussions)." *The Annals of Statistics* 31, 705-767.

[44] Maheu, J.M. and T.H. McCurdy. (2004). "News Arrival, Jump Dynamics and Volatility Components for Individual Stock Returns." *Journal of Finance* 64, 755-793.

[45] Merton, R.C. (1973). "An Intertemporal Capital Asset Pricing Model." *Econometrica* 41, 867-887.

[46] Merton, R.C. (1980). "On Estimating the Expected Return on the Market: An Exploratory Investigation." *Journal of Financial Economics* 8, 323-361.

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

[48] Pindyck, R. (1984). "Risk, Inflation and the Stock Market." *American Economic Review* 74, 335-351.

[49] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge: Cambridge University Press.

[50] Spiegelhalter, D.J., N.G. Best, B.P. Carlin and A. van der Linde. (2002). "Bayesian Measures of Model Complexity and Fit (with Discussions)." *Journal of the Royal Statistical Society, Series B* 64, 583-639.

[51] Turner, C.M., R. Startz and R. Nelson. (1989). "A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market." *Journal of Financial Economics* 25, 3-22.

[52] Whitelaw, R. (1994). "Time Variation and Covariations in the Expectation and Volatility of Stock Market Returns." *Journal of Finance* 49, 515-541.

Table 1: **Cumulants of VG and NIG**

	VG	NIG
c_1	ωt	ωt
c_2	$(\omega^2 v + \eta^2)t$	$(\omega^2 v + \eta^2)t$
c_3	$3\omega(\frac{2}{3}\omega^2 v^2 + \eta^2 v)t$	$3\omega(\omega^2 v^2 + \eta^2 v)t$
c_4	$3v(\eta^4 + 2\omega^4 v + 4\eta^2\omega^2 v)t$	$3v(\eta^4 + 5\omega^4 v + 6\eta^2\omega^2 v)t$

Note: The cumulants can be derived from the Characteristic functions or Cumulant generating functions. They are closely related to the moments. The first two cumulants are mean and variance respectively and the third and fourth cumulants are nonnormalized skewness and excess kurtosis. The skewness and excess kurtosis are then calculated as $skw = c_3/c_2^{3/2}$ and $ekt = c_4/c_2^2$.

Table 2: **Bayesian Parameter Estimates**

	μ^*	ω	η	v	$\kappa^{(1)}$	$\theta^{(1)}$	$\sigma^{(1)}$	ρ	$\kappa^{(2)}$	$\sigma^{(2)}$
A. Constant Jump Arrival Rate Models										
VGSV	0.169	-0.109	0.079	0.032	2.241	0.025	0.263	-0.655	—	—
	(0.035)	(0.036)	(0.005)	(0.001)	(0.700)	(0.007)	(0.030)	(0.062)	—	—
NIGSV	0.143	-0.080	0.091	0.210	2.338	0.024	0.253	-0.627	—	—
	(0.033)	(0.037)	(0.008)	(0.072)	(0.766)	(0.009)	(0.026)	(0.064)	—	—
B. Stochastic Jump Arrival Rate Models										
VG2SV	0.154	-0.069	0.077	0.036	1.763	0.026	0.232	-0.705	65.545	9.099
	(0.034)	(0.028)	(0.005)	(0.002)	(0.635)	(0.011)	(0.026)	(0.064)	(2.964)	(0.253)
NIG2SV	0.146	-0.067	0.085	0.223	2.102	0.024	0.245	-0.611	71.930	9.746
	(0.031)	(0.029)	(0.007)	(0.054)	(0.668)	(0.006)	(0.024)	(0.063)	(5.690)	(0.414)

Note: Models are estimated with Bayesian methods discussed in Section 3. Totally 50,000 simulations are implemented: the first 30,000 simulations are discarded as burn-in and the last 20,000 simulations are kept for inference. Posterior mean and standard error (in brackets) are presented.

Table 3: **Statistics for Model Comparison**

	VGSV	NIGSV	VG2SV	NIG2SV
A. Return Residuals				
Mean	0.025	0.020	0.025	0.020
Std.	0.839	0.871	0.831	0.865
Skewness	0.102	0.033	0.061	0.035
Kurtosis	2.612	2.711	2.698	2.747
B. Volatility Residuals				
Skewness	-0.127	-0.096	-0.096	-0.116
Kurtosis	2.742	2.673	2.708	2.737
C. Deviance Information Criterion				
$\bar{D}$	-26,906.6	-26,720.0	-26,947.1	-26,751.7
p_D	835.3	606.7	873.5	629.2
DIC	-26,071.3	-26,113.3	-26,073.6	-26,122.5

Note: The top panel reports the empirical mean, standard deviation, skewness and kurtosis of return residuals defined in (39); the middle panel presents the skewness and kurtosis of volatility residuals defined in (40) respectively; and lower panel gives the statistics of Deviance Information Criterion. Posterior means of parameters and states are used.

Table 4: **Summary Statistics of Returns and Volatility**

		VG2SV				NIG2SV			
	R_t	$V_t^{(1)}$	$V_t^{(2)}$	$\sqrt{V_t^{(1)}}$	$\sqrt{V_t^{(2)}}$	$V_t^{(1)}$	$V_t^{(2)}$	$\sqrt{V_t^{(1)}}$	$\sqrt{V_t^{(2)}}$
Mean	0.122	0.016	1.295	0.119	1.122	0.018	1.279	0.125	1.128
Std. Dev.	0.169	0.013	0.395	0.046	0.187	0.014	0.197	0.048	0.079
Skewness	-2.962	1.630	0.352	0.798	0.196	1.569	3.827	0.748	2.596
Kurtosis	64.551	5.942	7.456	3.195	6.682	5.544	34.705	3.141	18.893
Minimum	-0.229	0.002	0.033	0.040	0.181	0.002	0.913	0.043	0.956
Maximum	0.087	0.081	4.302	0.284	2.074	0.080	3.756	0.282	1.938
ρ_1	0.019	0.995	0.943	0.995	0.907	0.996	0.950	0.995	0.949
ρ_2	-0.055	0.989	0.848	0.989	0.791	0.990	0.853	0.990	0.851
ρ_3	-0.045	0.983	0.741	0.983	0.675	0.985	0.738	0.985	0.735
ρ_4	-0.025	0.977	0.636	0.978	0.568	0.980	0.620	0.980	0.616
ρ_5	-0.019	0.972	0.540	0.973	0.476	0.974	0.509	0.974	0.505
ρ_6	-0.010	0.965	0.458	0.967	0.400	0.967	0.407	0.969	0.404

Note: The table presents the summary statistics of Daily S&P 500 index returns ranging from January 1986 to December 2000 and the estimated diffusion volatility and jump volatility (both in variance and standard deviation forms). All of them are annualized.

Table 5: **Return Volatility Relation**

	Variance						Standard Deviation					
	α_0	α_1	α_2	α_3	α_4	$R^2(\%)$	α_0	α_1	α_2	α_3	α_4	$R^2(\%)$
A. Case II												
VG2SV	0.004	0.008	-0.003	—	—	1.00	0.005	0.002	-0.004	—	—	0.57
	(0.001)	(0.018)	(0.001)	—	—		(0.002)	(0.005)	(0.002)	—	—	
NIG2SV	0.012	0.026	-0.010	—	—	2.92	0.025	0.008	-0.023	—	—	2.63
	(0.003)	(0.013)	(0.002)	—	—		(0.005)	(0.004)	(0.005)	—	—	
B. Case III												
VG2SV	0.003	-7.741	-0.027	-0.002	—	80.9	0.005	-2.027	-0.007	-0.003	—	71.1
	(0.001)	(0.246)	(0.013)	(0.001)	—		(0.002)	(0.080)	(0.003)	(0.001)	—	
NIG2SV	0.008	-7.526	-0.012	-0.006	—	82.5	0.017	-2.033	-0.002	-0.015	—	72.6
	(0.002)	(0.145)	(0.013)	(0.001)	—		(0.005)	(0.062)	(0.003)	(0.004)	—	
C. Extension												
VG2SV	0.003	-7.761	-0.028	0.003	-0.002	81.0	0.004	-2.030	-0.008	0.004	-0.002	71.1
	(0.001)	(0.251)	(0.013)	(0.001)	(0.001)		(0.002)	(0.080)	(0.003)	(0.001)	(0.001)	
NIG2SV	0.007	-7.618	-0.013	0.012	-0.005	82.9	0.015	-2.069	-0.003	0.035	-0.013	72.7
	(0.002)	(0.142)	(0.013)	(0.003)	(0.002)		(0.005)	(0.060)	(0.003)	(0.008)	(0.005)	

Table 6: **Leverage Effect**

	VG2SV			NIG2SV		
	Diffusion	Jump	Return	Diffusion	Jump	Return
A. Variance						
α_0	0.016	1.297	0.024	0.018	1.281	0.028
	(0.001)	(0.020)	(0.001)	(0.001)	(0.003)	(0.001)
α_1	-0.116	-3.584	-0.138	-0.115	-3.319	-0.143
	(0.029)	(1.217)	(0.034)	(0.029)	(0.844)	(0.034)
$R^2(\%)$	1.00	1.00	1.16	0.78	3.24	1.18
B. Standard Deviation						
α_0	0.119	1.129	0.150	0.125	1.129	0.164
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)
α_1	-0.366	-1.311	-0.362	-0.353	-1.261	-0.350
	(0.078)	(0.414)	(0.075)	(0.078)	(0.273)	(0.072)
$R^2(\%)$	0.72	0.56	0.94	0.62	2.87	1.03

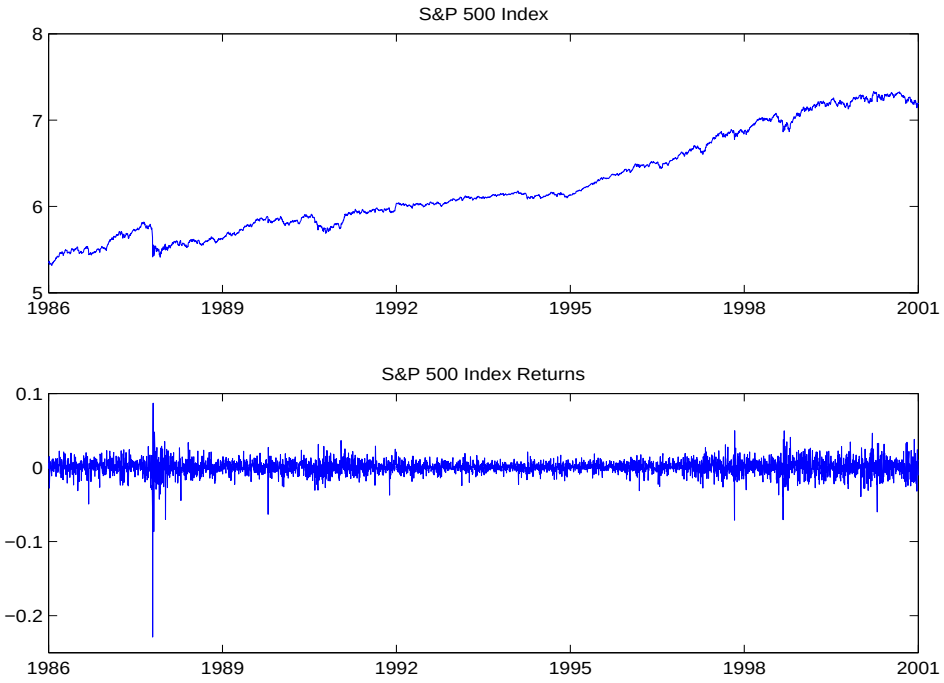

Figure 1: **S&P 500 Index and Trading Volume**

Note: The Figure plots the time series of S&P 500 index and index returns. Data are from January 1986 to December 2000 in daily frequency. There are totally 3791 business days.

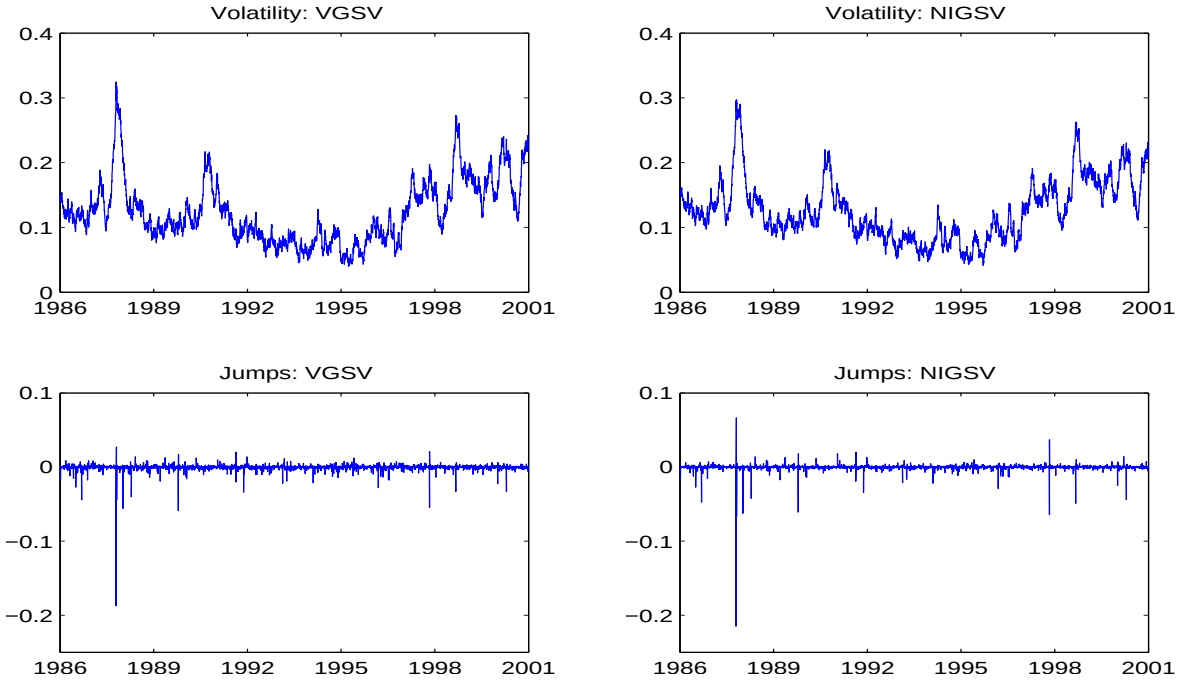

Figure 2: **Estimated Volatility and Jumps of VGSV and NIGSV Models**

Note: The figure plots the posterior means of the diffusion volatility and the jumps of VGSV and NIGSV models. The left panels are for VGSV model and the right panels for NIGSV model.

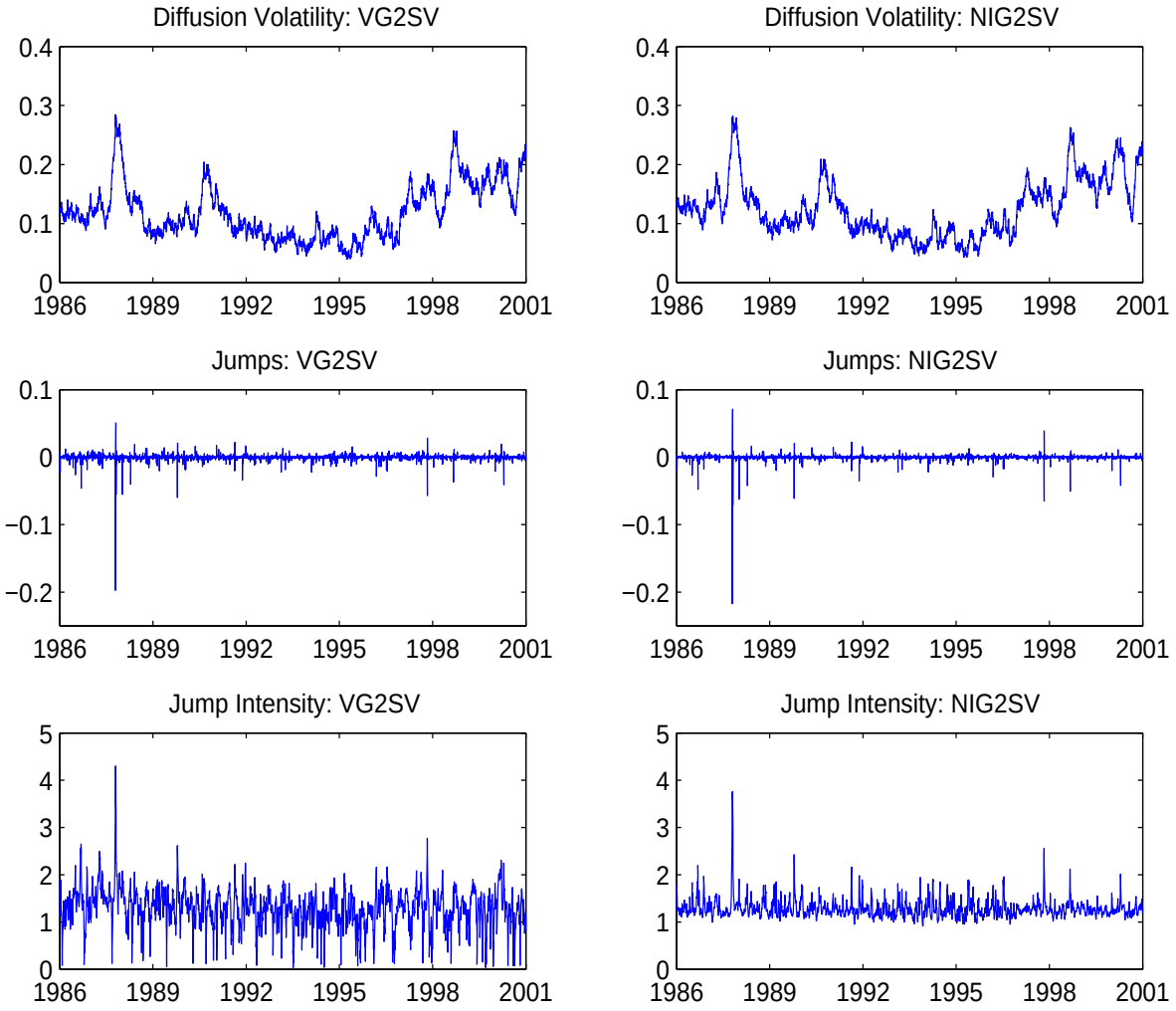

Figure 3: **Estimated Diffusion Volatility, Jumps and Jump Intensity of VG2SV and NIG2SV Model**

Note: The figure plots the posterior means of the diffusion volatility, the jumps and the jump arrival rates of VG2SV and NIG2SV models. The left panels are for VG2SV model and the right panels for NIG2SV model.