A New Class of Stochastic Volatility Models with Jumps: Theory and Estimation[1]

Mikhail Chernov

A. Ronald Gallant

Eric Ghysels

George Tauchen

First Draft: May 1999
This Revision: October 13, 1999

[1]The authors are respectively with the Department of Finance, The Pennsylvania State University, 609 Business Administration Building 1, University Park, PA 16802-3006, phone: 814-863-0486, fax: 814-865-3362, *e-mail*: `mchernov@psu.edu`, Department of Economics, University of North Carolina, CB# 3305, 6F Gardner Hall, Chapel Hill NC 27599-3305, phone: 919-966-5338, fax: 919-966-4986, *e-mail*: `gallant@unc.edu`, Departments of Economics and Finance, The Pennsylvania State University, 523 Kern Graduate Building, University Park, PA 16802-3305, phone: 814-865-7924, fax: 814-863-4775, *e-mail*: `eghysels@psu.edu` and Department of Economics, Duke University, 305 Social Sciences, Box 90097, Durham, NC 27708-0097, phone: 919-660-1812, fax: 919-684-8974, *e-mail*: `get@econ.duke.edu`.

Abstract

The purpose of this paper is to propose a new class of jump diffusions which feature both stochastic volatility and random intensity jumps. Previous studies have focused primarily on pure jump processes with constant intensity and log-normal jumps or constant jump intensity combined with a one factor stochastic volatility model. We introduce several generalizations which can better accommodate several empirical features of returns data. In their most general form we introduce a class of processes which nests jump-diffusions previously considered in empirical work and includes the affine class of random intensity models studied by Bates (1998) and Duffie, Pan and Singleton (1998) but also allows for non-affine random intensity jump components. We attain the generality of our specification through a generic Lévy process characterization of the jump component. The processes we introduce share the desirable feature with the affine class that they yield analytically tractable and explicit option pricing formula. The non-affine class of processes we study include specifications where the random intensity jump component depends on the size of the previous jump which represent an alternative to affine random intensity jump processes which feature correlation between the stochastic volatility and jump component. We also allow for and experiment with different empirical specifications of the jump size distributions. We use two types of data sets. One involves the S&P500 and the other comprises of 100 years of daily Dow Jones index. The former is a return series often used in the literature and allows us to compare our results with previous studies. The latter has the advantage to provide a long time series and enhances the possibility of estimating the jump component more precisely. The non-affine random intensity jump processes are more parsimonious than the affine class and appear to fit the data much better.

JEL classification: G13; C14, C52, C53

Key Words: Efficient Method of Moments, Poisson processes, Jump processes, Stochastic Volatility Models, Filtering

Introduction

Much research has been devoted to the estimation of continuous time processes to describe the evolution of fundamental asset prices. The failure of the Black-Scholes model has directed attention to continuous time processes more general than the standard geometric Brownian motion. Since asset returns feature a greater degree of kurtosis in the unconditional returns distributions than implied by normality, attempts at reconciling the theory with the empirical evidence have focused on extensions of the Black-Scholes model which include a Poisson-driven jump process.[1] Asset returns also display strong volatility clustering, a feature which can neither be reconciled with the Black-Scholes model nor with pure Poisson jump processes. Diffusions with stochastic volatility have therefore been introduced to capture persistence in volatility.[2] Clearly, jump-diffusion processes combining stochastic volatility and Poisson jumps processes have two appealing features to model stock returns. The continuous path stochastic volatility part accommodates the clustered random changes in the returns volatility whereas the discrete jump component accommodates the infrequent large price movements.[3] These models have received much attention recently. Various papers have examined the econometric estimation and/or derivative security pricing with such processes. Examples include Bates (1996a), Ho, Perraudin and Sørensen (1996), Scott (1997), Bakshi, Cao and Chen (1997), Scott (1997), Bates (1998), Andersen, Benzoni and Lund (1998), among others. Bakshi and Madan (1998a) and Duffie, Pan and Singleton (1998) provide very elegant general discussions of the particularly versatile class of affine jump-diffusions with stochastic volatility which yield analytic solutions to derivative security pricing.

Attempts to estimate jump-diffusions have focused almost exclusively on processes involving pure Poisson jump components. Empirical studies which combine stochastic volatility

[1]For work on jump-diffusions see Merton (1976), Cox and Ross (1976), Ball and Torous (1983, 1985), Jarrow and Rosenfeld (1984), Jorion (1988) among others. In particular see the survey of Bates (1996b) for further references.

[2]See for instance Hull and White (1987), Scott (1987), Wiggins (1987), Heston (1993) and many others and the survey by Ghysels *et al.* (1996) for further details.

[3]It is worth noting that a different model was proposed by Bakshi and Madan (1998b). It can be described as a purely discontinuous process which can capture the same empirical stylized facts as jump diffusions do. We will not consider this class of processes in this paper.

and jump processes, such as Bates (1996), Bakshi, Cao and Chen (1997) and Andersen, Benzoni and Lund (1998) assume the stochastic volatility and the jump components are mutually orthogonal. Moreover, since jumps are governed by a Poisson process, their intensity is assumed to be independently and identically distributed. The stochastic volatility component is also limited to a one factor model specification. While such models seem to fit (marginally) many aspects of the data it is clear that they can be improved upon. The stochastic structure of these models implies that the day after a stock market crash another crash is equally likely as the day before market nose-dived. In addition, the occurrence of jumps is independent of volatility. Bates (1998) estimated a class of jump-diffusions with random intensity for the jump process, more specifically where the intensity is an affine function of the stochastic volatility component. Duffie, Pan and Singleton (1998) generalize this class and Pan (1999) estimates multi-factor jump-diffusion models with affine stochastic jump intensity. This structure may not be entirely suitable either to accommodate some stylized facts. For instance the day before the 1987 stock market crash the volatility measured by the squared return on the S&P 500 index was roughly the same as the day after the crash. Therefore, in this case making the intensity of a crash a linear affine function of volatility would again result in the probability of a crash the day after Black Monday being the same as the trading day before the crash.[4]

The purpose of this paper is to introduce a class of stochastic volatility models with jumps which includes the affine class of random intensity models studied by Bates (1998), Duffie, Pan and Singleton (1998) and Pan (1999) but also allows for non-affine random intensity jump components. The class of models also feature multiple factor stochastic volatility processes. The non-affine random intensity jump processes appear to fit the data better compared to the constant and random affine intensity jump-diffusion models. Moreover, the processes we introduce share the desirable feature with the affine class that they yield analytically tractable and explicit option pricing formula. The non-affine class of processes we study include specifications where the random intensity jump component depends on the size of the previous jump. For instance large (negative) jumps reduce the intensity of jumps which is a realistic representation of crash behavior. It also represents an alternative to affine random intensity jump processes which feature correlation between the stochastic

[4]To be more precise the log of the return on the S&P 500 index was -0.0529 on Friday October 16, it was -0.2283 on Black Monday, namely October 19, 1987, and it was 0.0513 the day after.

volatility and jump component. We also allow for and experiment with different empirical specifications of the jump size distributions. The size of jumps is typically modeled as a random variable with a log-normal distribution since it yields analytically tractable option pricing models. However, other distributional specifications may fit the data better and still yield analytical solutions to option pricing. We have recently become aware of independent work by Johannes, Kumar and Polson (1999) who propose a class of jump diffusion processes which has some similarities with the ones introduced here. They analyze a process where the jump intensity is a function of past jump events and the absolute value of returns. Their analysis proceeds in discrete time and uses Bayesian MCMC methods to estimate the parameters and extract the latent volatility, jump intensity and size. While there are some similarities between the class of processes we study and their specification, there are also some important differences. One of the most notable differences is that we propose continuous time models which have option pricing formula which can be derived explicitly. For instance, some of the processes studied by Johannes *et al.* (1999), such as the probit specification with lagged jump time, do not have straightforward continuous time diffusion limits. In contrast, we study exclusively processes which yield readily derivative security prices. In particular, when we introduce several alternative distributional specifications for the jump size distribution, we select specifications which yield tractable solutions to option pricing and also yield an easy specification of the drift in the continuous time process.

In their most general form we introduce a class of processes which nests jump-diffusions previously considered in empirical work. In particular, special cases include the one-factor SV - Poisson jump processes studied by Andersen, Benzoni and Lund (1998), among others, and the affine class of multi-factor SV - affine random intensity jump processes studied by Bates (1998) and Pan (1999). They also include the non-affine models studied by Johannes *et al.* (1999). We attain the generality of our specification through a generic Lévy process characterization of the jump component.[5]

The empirical analysis in our paper tries to strike a balance between two objectives which are not easy to reconcile. Because the task of estimating jump-diffusions is challenging it is better to use time series covering a long history of market activity containing potentially several crashes. We consider two data sets. The first covers over 100 years of daily observations

[5]See also Bakshi and Madan (1998b) for a similar approach involving general Lévy measures to characterize jump processes.

on the Dow Jones Industrial Average (DJIA) index. Unfortunately, the Dow was until very recently an index without derivative contracts. The second data set is much shorter, namely contains daily observations from April 1986 until August 1997, and covers the widely used S&P500 index which features as the fundamental asset for a very liquid derivatives market, including the most actively traded set of SPX contracts. Since the class of models we consider readily yields option pricing formula, it could be used to estimate jump-diffusions using simultaneously S&P 500 data and SPX options contracts. Such endeavor would require the specification of both the risk neutral and objective measures of the process. In this paper we limit ourselves to the objective measure representation of the class of processes we introduce. In a companion paper, Chernov *et al.* (1999), we discuss the option pricing implications of the class of processes introduced here and estimate jump-diffusions with a panel of options and returns data using an approach similar to that in Chernov and Ghysels (1998). The purpose of examining a long Dow Jones series and a much shorter S&P 500 series is to make comparisons about the validity of the inference made across the two indices. This will then in turn help us with the sequel to this project, which consists of estimating jump-diffusions with S&P 500 and SPX options data jointly. Moreover, many of the existing empirical studies have used the S&P 500 index data as well which allows us to make direct comparisons with findings previously reported in the literature.

The paper is structured as follows. In the first section we discuss the generic class of processes. Then we briefly discuss the data and estimation procedure in a second section. A third section is devoted to the empirical results while a fourth section concludes.

1 The Generic Class of Models

The purpose of this section is to present a general structure of jump/diffusion processes. Our discussion is based on the developments in Chernov (1999). The diffusion part of the generic class of models can have multiple factors like in Duffie, Pan and Singleton (1998). Since their analysis is very general and thorough we will not provide many details regarding the diffusion part of the processes. In particular we will assume that the stock (index) process satisfies $d\log S_t = dU_{1t} + dX_t$ where U_{1t} is the first element of the vector process U_t defined as:

$$dU_t = \mu(U_t, t)dt + \sigma(U_t, t)dW_t \tag{1.1}$$

where the process U_t is Markovian and takes values in an open subset D of $\mathbf{R}^N$, $\mu(y) = \Theta + \mathcal{K}y$ with $\mu : D \to \mathbf{R}^N$ and $\sigma(y)\sigma(y)' = h + \sum_{j=1}^N y_j H^{(j)}$ where $\sigma : D \to \mathbf{R}^{N\times N}$. Moreover, the vector Θ is $N\times 1$, the matrix $\mathcal{K}$ is $N \times N$ whereas h and H are all symmetric $N \times N$ matrices. The process W_t is a standard Brownian motion in $\mathbf{R}^N$. All further details regarding the regularity conditions pertaining to the U_t are discussed by Duffie, Pan and Singleton (1998) and therefore omitted. Instead, our main focus will be on the specification of the jump component X_t which is modeled as a Lévy process.

The remaining of this section is divided in three subsections. The first subsection deals with some desirable stylized features which we would like to have for jump processes. A first class of non-affine models involving specific Lévy processes is presented in a second subsection. The final subsection presents the general class which covers both affine and non-affine random intensity jump-diffusion processes. The Appendix describes some general properties of Lévy processes which will be used in the development of the models.

1.1 The Features of the Ideal Jump Process Specification

Particular assumptions about the Lévy measure will define a class of processes. The Lévy measure $\nu(\Lambda)$ can be interpreted as the probability of jumps into Λ, i.e. $\nu(\Lambda) = Pr\{\text{jump}|\text{jump size} \in \Lambda\} \cdot Pr\{\text{jump size} \in \Lambda\}$. The standard constant Poisson intensity representation can be written as:

$$\nu(dx) = \lambda(x) \cdot n(\mu_J, \sigma_J, x)dx, \text{ where} \tag{1.2}$$

$$\lambda(x) = \lambda \tag{1.3}$$

$$x \sim N(\mu_J, \sigma_J^2) \tag{1.4}$$

We noted that the stochastic structure of these models is not ideal. Somehow we would like to have time-varying intensity of the jump process so that for instance the day after a stock market crash another crash is very unlikely. Since there are not many extreme crash events we must prefer jump processes with a parsimonious parametric structure to make statistical inference meaningful. For the purpose of exposition we will at least for the moment assume we are not preoccupied with parsimony. Instead, we will try to define some ideal features a jump process should have to accommodate financial market data.

There are at least three ideal features. First and foremost we want jumps to be rare events. This is easy to assess when we examine constant intensity models which yield on average one jump every six months (see for instance Andersen, Benzoni and Lund (1998) who use S&P 500 data similar to ours). As already noted, we also want that the intensity decreases dramatically after a crash. Finally, we would also like the intensity to gradually increase as the memory of a crash fades away.

Bates (1998) and Duffie, Pan and Singleton (1998) accommodate some of these features by making the jump intensity an affine function of the stochastic volatility component. However, as we noted in the Introduction, making jump intensities a function of volatility may not always be successful at mimicking the ideal features particularly since stock markets tend to remain volatile shortly after a crash. The non-affine class of jump processes we present in the next section accommodate some, but not all of the ideal features, at least in their simplest specification. The class of jump processes we examine first has random jump intensity which depends on the size of the previous jump. Therefore, a large crash will dramatically reduce the probability of another jump. This can be viewed as an alternative to affine random intensity jump processes which feature correlation between the stochastic volatility and jump component. Unfortunately, the non-affine model assumes that the jump probability remains constant until the next jump. Hence, after a large crash one expects that it will take longer for a next jump to occur. A specification involving a more elaborate hazard model could be considered where for instance the intensity increases as the memory of the last crash evaporates.

Obviously, if we were not concerned with parsimony, we could simply combine the attributes of all the models, i.e. assume the intensity to be an affine function of volatility and of the size of the last jump with fading memory. The general class of models does indeed accommodate such a structure with a suitable specification of the Lévy measure.

1.2 A Class of Non-affine Random Intensity Jump-Diffusions

The affine class of processes studied by Duffie, Pan and Singleton (1998), and further discussed in the next section, assume that the Lévy measure depends on the latent factor process U_t. In this section we specify instead the jump component of bounded variation, namely we remove B_t in the Lévy decomposition (appearing in Appendix A), which distinguishes it from the affine class of processes which make the intensity a function of Brownian motions. The non-affine class of processes we study include specifications where the random intensity jump component depends on the size of the previous jump. As noted before, the jump component is defined as a Lévy process of bounded variation, i.e. we remove B_t in the Lévy decomposition (A.4) (or, equivalently, set $\sigma = 0$ in the Lévy-Khintchine formula (A.5)) and assume $\int \min(1, |x|)\nu(dx) < \infty$. In this case, $\int_{\{|x|<1\}} iux\nu(dx)$ is a well defined linear function, and the characteristic function defined in (A.5) can be expressed as:

$$\psi(u) = -i\delta u + \int_{\mathbf{R}} (1 - e^{iux})\nu(dx) \tag{1.5}$$

where δ is known as the drift coefficient.[6]

We will present three different specifications, each having the appealing feature of yielding analytical option pricing formula for European type contracts written on the stock price index process S_t appearing in (1.1).[7] The first specification for the Lévy measure ν is as follows:

$$\begin{aligned} \nu(dx) &= \lambda(x) \cdot n(\mu_J, \sigma_J, x)dx, \text{ where} && (1.6)\\ \lambda(x) &= \lambda \exp\left(-\frac{\gamma x^2}{2\sigma_J^2}\right), \gamma \geq 0 && (1.7)\\ x &\sim N(\mu_J, \sigma_J^2) && (1.8) \end{aligned}$$

This representation collapses to the standard one appearing in (1.2) - (1.4) when we set $\gamma = 0$. The functional form of ν in (1.6) - (1.8) suggests that large, in amplitude, jumps have a low probability of occurrence. Note that $2\sigma_J^2$ is not necessary in the specification of $\lambda(x)$. It is introduced for notational convenience in the subsequent analysis. We can compute the predictable component of the process $\exp(x) - 1$ using (A.3), namely:

$$\begin{aligned} E\{\int_{\mathbf{R}} (e^x - 1)N_t(dx)\} &= \lambda t \int_{\mathbf{R}} (e^x - 1)e^{-\frac{\gamma x^2}{2\sigma_J^2}} \frac{1}{\sqrt{2\pi\sigma_J^2}} e^{-\frac{(x-\mu_J)^2}{2\sigma_J^2}} dx \\ &= \frac{\lambda t}{\sqrt{1+\gamma}} e^{-\frac{\gamma\mu_J^2}{2(1+\gamma)\sigma_J^2}} \int_{\mathbf{R}} (e^x - 1)n\left(\frac{\mu_J}{1+\gamma}, \frac{\sigma_J}{\sqrt{1+\gamma}}, x\right) dx \\ &= \frac{\lambda}{\sqrt{1+\gamma}} \left(e^{\frac{\mu_J + 0.5\sigma_J^2}{1+\gamma}} - 1\right) e^{-\frac{\gamma\mu_J^2}{2(1+\gamma)\sigma_J^2}} t = \delta t && (1.9) \end{aligned}$$

This is the compensator of the process $\exp(X_t) - 1$ which affects the drift of the stock index process appearing in (1.1).

There are at least two alternatives to (1.6)-(1.7). The first is:

$$\nu(dx) = \lambda(x) \cdot C_\nu n(\mu_J, \sigma_J, x)dx, \text{ where} \tag{1.10}$$

[6] See Bertoin (1996) for further details

[7] Option pricing formula are derived and discussed in Chernov (1999).

$$
\begin{aligned}
C_\nu &= \Phi\left(-\frac{\mu_J}{\sigma_J}\right)^{-1} && (1.11)\\
\lambda(x) &= \lambda \exp\left(\frac{\gamma}{2\sigma_J^2}x\right), \gamma \geq 0, && (1.12)\\
x &\sim N^-(\mu_J, \sigma_J^2) && (1.13)
\end{aligned}
$$

In this case, we still have analytical tractability, but the jump component behaves quite differently. In particular, the jump size distribution appearing in (1.13) is the half-normal distribution which implies that jumps sizes have an upperbound μ_J. When $\mu_J = 0$ then we only allow for negative jumps, which may be a realistic presentation of the behavior observed in financial markets. To illustrate analytical tractability we compute again the predictable component of the exponentiated process:

$$
\begin{aligned}
E\{\int_{-\infty}^{0} e^x N_t(dx)\} &= \lambda t \int_{-\infty}^{0} e^x e^{\frac{\gamma}{2\sigma_J^2}x} C_\nu \frac{1}{\sqrt{2\pi\sigma_J^2}} e^{-\frac{(x-\mu_J)^2}{2\sigma_J^2}} dx\\
&= \lambda t C_\nu e^{(\gamma+2\sigma_J^2)(\frac{1}{4}(\gamma+2\sigma_J^2)+\mu_J)} \int_{-\infty}^{0} n(\mu_J + \frac{1}{2}(\gamma+2\sigma_J^2), \sigma_J, x) dx\\
&= \lambda t C_\nu e^{(\gamma+2\sigma_J^2)(\frac{1}{4}(\gamma+2\sigma_J^2)+\mu_J)} \Phi\left(-\frac{\gamma+2\sigma_J^2+2\mu_J}{2\sigma_J}\right) && (1.14)
\end{aligned}
$$

Another specification is based on the exponential distribution. This specification is close to the half-normal with $\mu_J = 0$, since jumps only take negative values, except that the shape of the distribution is different. The process is specified as:

$$
\begin{aligned}
\nu(dx) &= \lambda(x) e(\alpha, -x) dx, \text{ where} && (1.15)\\
\lambda(x) &= \lambda \exp\left(\frac{\gamma}{\alpha}x\right), \gamma \geq 0, && (1.16)\\
-x &\sim E(\alpha), \ \alpha > 0 && (1.17)
\end{aligned}
$$

where $E(\alpha)$ denotes the exponential distribution with parameter α. The compensator for such a process will be:

$$
\begin{aligned}
E\{\int_{-\infty}^{0} (e^x - 1) N_t(dx)\} &= \lambda t \int_{-\infty}^{0} (e^x - 1) e^{\frac{\gamma}{\alpha}x} \frac{1}{\alpha} e^{\frac{x}{\alpha}} dx\\
&\underset{y=-x}{=} \lambda t \int_{0}^{\infty} (e^{-y} - 1) e^{-\frac{\gamma}{\alpha}y} \frac{1}{\alpha} e^{-\frac{y}{\alpha}} dy\\
&= \lambda t \frac{1}{\gamma+1} \int_{0}^{\infty} (e^{-y} - 1) \frac{\gamma+1}{\alpha} e^{-\frac{y(\gamma+1)}{\alpha}}\\
&= \frac{\lambda t}{1+\gamma}\left(\frac{1}{1+\frac{\alpha}{1+\gamma}} - 1\right) = -\frac{\alpha\lambda t}{(1+\gamma)(1+\gamma+\alpha)}, && (1.18)
\end{aligned}
$$

$$\frac{\alpha}{1+\gamma} \quad > \quad -1 \tag{1.19}$$

The exponential specification is guided by one parameter (α) and the half-normal one by two (μ_J, σ_J). Hence we can assess parsimony versus a better fit in these two models. Also, the half-normal specification imposes an upper boundary on the jumps size, i.e. they should be lower than μ_J. This also adds flexibility regarding specification testing.

1.3 The General Class of Models

Recall that particular assumptions about the Lévy measure will define a class of processes. So far we examined random intensities for the jump process which are of the form:

$$\lambda(x) = \lambda \exp(G(x)) \tag{1.20}$$

where we considered functions $G(x)$ of the jump size x which were either quadratic, such as in (1.7), or linear as in (1.12) and (1.16). Common to these specifications is the constant scaling λ. In contrast, the models considered by Duffie, Pan and Singleton (1998) are:

$$\lambda(U_t) = \lambda_0(t) + \lambda_1(t)U_t \tag{1.21}$$

where the process U_t is of the affine class as specified in (1.1).[8] It is possible to combine the features of (1.20) and (1.21) and consider stochastic intensities for the jump process which are of the type:

$$\lambda(x, U) = \lambda_0(x,t) + \lambda_1(x,t)U_t \tag{1.22}$$

where for instance $\lambda_i(x,t) = \lambda_i(t)\exp(G(x))$. This specification yields a class of jump Lévy measures which combines the features of jump intensities depending on, say volatility, as well as the size of the previous jump. It relates to a case briefly mentioned in Duffie *et al.* (1999), namely one where m types of jumps, each having an intensity $\lambda_i(U_t) = \lambda_0^i(t) + \lambda_1^i(t)U_t$ for $i = 1, \ldots, m$ (see their Appendix F). At the end of Appendix F it is briefly mentioned that this case of finite mixture can be extended to an infinite number of jump types. In a sense the specification in (1.22) is such a generalization. Combining their results for option pricing with those discussed in our companion paper (Chernov *et al.* (1999)) yields options pricing formulas for jump diffusions with general affine features in a state vector process U_t and

[8] We are considering here the most general case considered in Duffie *et al.* (1999, equation A.5).

non-affine in jump size x. As mentioned in the Introduction, the analysis in our paper also relates to recent independent work by Johannes, Kumar and Polson (1999) who propose a class of jump diffusion processes which has a specification somewhat similar to (1.22). They analyze a jump intensity which is a function of past jump time (not size like we do), and the absolute value of returns. Using the absolute value of returns could be viewed as a non-affine specification. Since the absolute value of returns is a measure of return volatility it could easily be replaced by an affine specification involving latent volatility as in (1.21). The presence of the past jump time is more problematic. While it is suitable and well tailored to the discrete time specification of Johannes *et al.* (1999), with jump times extracted via MCMC methods, it does not yield continuous time diffusion limits. An alternative and potentially very useful specification would be to augment the stochastic intensity appearing in (1.22) with past *durations*, i.e. the time since the last jump, say $\tau(t)$ which is the time elapsed between the last jump and t where $\tau(t)$ belongs to $\mathbf{R}$ in a continuous time setting. Such a specification, which would involve:

$$\lambda(x, U, \tau) = (\lambda_0(t) + \lambda_1(t) U_t) \lambda(\tau(t)) \exp(G(x)) \tag{1.23}$$

and could accommodate increasing, decreasing or hump-shaped hazard functions and duration dependence of jump intensities.

2 Data and Estimation Procedures

As the task of estimating complex jump-diffusions is challenging it is certainly desirable to use time series covering potentially several crashes. The new class of models we proposed was kept analytically tractable for the purpose of option pricing. Unfortunately, exchange traded standardized derivatives contracts which are liquid and have reliably recorded prices are only of recent date, more precisely starting in the early seventies. This does not amount to a long data series with only a few extreme events. We consider therefore two data sets which are described in the first subsection. We use the Efficient Method of Moments (henceforth EMM) procedure of Gallant and Tauchen (1996) which has many applications in the estimation of stochastic volatility models and was also used by Andersen, Benzoni and Lund (1998) for the estimation of Poisson jump-diffusions. Since the EMM method has been discussed in detail elsewhere, see for instance Gallant and Tauchen (1998), we will only provide a brief description of it in the second subsection. The third subsection describes the details of the discretization scheme which will be used to simulate the processes.

2.1 The Data

We use two different data sets. The first covers roughly 100 years of daily observations on the Dow Jones, an index which was without exchange traded derivative contracts, at least until very recently. The estimation of SNP density requires the use of stationary and ergodic data. Therefore, the data entries to the SNP estimation routine are the log-returns on the index.

Like many previous studies we also examine the S&P 500 index, which is a much shorter data set, namely contains daily observations from April 1986 until August 1997. In order to make the observed S&P 500 index and the model specification for the underlying fundamental comparable, we adjusted the S&P 500 index for dividends. We took a constant continuously compounded dividend rate of 2% (this is consistent with for instance Broadie *et al.* (1997)). Therefore, we ignore the lumpiness of the dividend payments and also conveniently avoid using the historical observed dividends which would considerably complicate our simulation design.

It is worth noting that we did not prefilter the data, neither the S&P 500 series nor the DJIA daily returns. Our analysis is therefore different from Gallant, Rossi and Tauchen (1992) and in particular also different from Andersen, Benzoni and Lund (1998) who estimated jump-diffusions with filtered S&P 500 data. We refrained from filtering the data because the adjustments affect the features of the data that are relevant for option pricing. Since our ultimate goal is to price derivatives (see Chernov *et al.* (1999)) we use unfiltered data. In that regard our analysis is closer to the recent empirical studies of Johannes *et al.* (1999) and Pan (1999) who also used unfiltered data.

The stock market was closed at the outbreak of World War I from the end of trading on July 30, 1914, until December 12, 1914. There were 116 potential trading days during this period and the Dow Jones average declined 24.39 percent over the period. This gap is exceedingly small relative to the size (28,390) of the entire data set. We filled this period using a nonparametric bootstrap imputation strategy. To the log returns data we tried all possible insertions of 116 day segments from the available data. We judged the quality of an insertion by how well the 30 days of data at the beginning and end of a candidate insertion period matched the 30 days of data at the beginning and end of the missing data period. In addition the percentage change over the 116 days of candidate insertion data was compared to -24.39 percent. The criterion of quality was weighted squared error. The weight was 5 on the discrepancy from -24.39 percent and weights 1,1/2,...,1/30 on the 30 pre and post period log return squared deviations with points farthest from the join point receiving the lesser weight. We experimented with several departures from this weighting scheme, some of them rather extreme, and found that it made little difference. The strategy appears to be robust.

The three panels of Figure 1 report the (1) level of the DJIA Index, (2) the log of the index and (3) the return computed as 100 times the log first difference of the index. The data are daily and spanning the period from January 1896 to July 1999. The two vertical lines delineate the market closure period which was replaced by imputed data.

2.2 The EMM Estimation Procedure

Suppose the process of interest, namely the asset return process for S_t, is denoted ι_t. In its generic context we assume that ι_t is a vector with L elements (hence we can include besides returns also the daily high/low as in Gallant, Hsu and Tauchen (1999) or options data as in Chernov and Ghysels (1999)). It has a conditional distribution $p(\iota_t \mid I_t, \Theta)$, where I_t is the information set and Θ are the parameters of the jump-diffusion stochastic volatility model. If we knew how to compute the maximum likelihood (MLE) estimator for Θ we would take the sample equivalent of:

$$E\left(\frac{\partial}{\partial\Theta}\log p(\iota_t \mid I_t, \Theta)\right) = 0, \tag{2.1}$$

where $\partial \log(p(\cdot \mid \cdot))/\partial\Theta$ is the score function. Unfortunately, it is very difficult to obtain the likelihood function for jump-diffusions with stochastic volatility models and hence it is impossible to compute the score generator $\partial \log(p(\cdot \mid \cdot))/\partial\Theta$.

Gallant and Tauchen (1996) suggest a computationally tractable strategy for handling this situation. Their idea is to use the score function of an auxiliary model $f(\iota_t|I_t, \Xi)$, where f is a computationally tractable conditional density with parameter vector Ξ. The EMM estimator $\hat{\theta}$ is the solution to

$$\min_{\Theta} m(\theta, \hat{\Xi})' W m(\theta, \hat{\Xi})$$

where W is weighting matrix, $\hat{\Xi}$ is the QML estimator of Ξ based on f, and

$$m(\theta, \Xi) = E_\theta(\frac{\partial}{\partial \Xi} \log f(\iota_t | I_t, \Xi)$$

where $E_\theta(\cdot)$ is computed by simulating long realizations from the underlying model. EMM is effectively the Duffie and Singleton (1993) simulated method of moments approach with the score function of the auxiliary model used to define the moment conditions. From Gallant and Tauchen (1996), Tauchen (1997), and Gallant and Long (1997), it is known that the EMM estimator is consistent and asymptotically normal even if the auxiliary density $f(\iota_t | I_t, \Xi)$ is mispecificed and does not subsume the true density $f(\iota_t | I_t, \Xi)$. However, the closer the auxiliary density approximates the true density, then the closer EMM will be to achieving the full efficiency of MLE.

2.2.1 The SNP-GARCH Auxiliary Model

One strategy that is known to ensure full asymptotic efficiency (Gallant and Long, 1997) is to use the SNP expansion of Gallant and Tauchen (1989) to define the auxiliary model. The essence of the argument is that, under regularity conditions, SNP asymptotically estimates the true conditional density and hence the SNP scores span the true scores and full efficiency obtains. The use of SNP to define the auxiliary model results in useful diagnostics as well; see Tauchen (1997) and the references therein. We follow this strategy, and, in particular, we use a new version of SNP that includes a GARCH-type model (Bollerslev (1986)) as the base of the asymptotic expansion.

SNP-GARCH works as follows. We start with VAR model

$$\iota_t = b_0 + \sum_{j=1}^{M_u} B_j \iota_{t-j} + u_t$$

or simply

$$\iota_t = \mu_{t-1} + u_t$$

where b_0 and the B_j are parameters. Suppose u_t is conditionally Gaussian and so

$$R_{t-1}(u_t - \mu_t) \sim N(0, I)$$

where R_{t-1} is an upper triangular matrix such that $R_{t-1}R'_{t-1} = var_{t-1}(u_t)$. The GARCH-type parameterization of the conditional second moment structure is

$$vech(R_{t-1}) = \sum_{j=1}^{M_g} \rho_{gj} vech(R_{t-1-j}) + \sum_{j=1}^{M_a} \rho_{aj} abs(u_{t-j})$$

where ρ_{gj} and ρ_{aj} are conformable parameter matrices, and $abs(\cdot)$ is the element-wise absolute value function. Putting

$$f(\iota_t|I_t,\Xi) = \frac{\phi[R_{t-1}(u_t - \mu_t)]}{|R_{t-1}|}$$

where I_t contains lags of ι_t, Ξ contains the parameters of the VAR and the GARCH taken together, and ϕ is the multivariate standard Gaussian density, gives a basic multivariate GARCH model with Gaussian errors. The tuning parameters are the lag lengths M_u, M_g, and M_a.

Most financial data, however, are not conditionally Gaussian. Such data exhibit departures from Gaussianity such as skewness, thick tails (excess kurtosis), and other nonlinearities, and such departures can be time varying. The SNP corrects the basic model for the these departures from normality. It works as follows. Put

$$z_t = R_{t-1}(u_t - \mu_t)$$

which would be distributed as $\phi(\cdot)$ were the data Gaussian, and we could write

$$f(\iota_t|I_t,\Xi) = \frac{\phi(z_t)}{|R_{t-1}|}$$

The SNP expansion uses in place

$$f(\iota_t|I_t,\Xi) = \frac{[poly(z_t,I_t)]^2\phi(z_t)}{F_{t-1} \quad |R_{t-1}|}$$

where $poly(z_t, I_t)$ is a polynomial of degree K_1 in z_t whose coefficients are themselves polynomial functions of degree K_2 in M_p lags of ι_t, and F_{t-1} is the adjustment factor that ensures $\int f(\iota_t|I_t,\Xi)d\iota_t = 1$. The function $poly(z_t, I_t)$ corrects for departures from Gaussianity. It is squared in order to ensure positivity of the density, and the division by $F_{t-1} = \int poly(z_t,I_t)]^2\phi(z_t)dz_t/|R_{t-1}|$ ensures integration to unity. The dependence of the coefficients of the polynomial on M_p lags of ι_t accommodates state-dependent time varying departures from Gaussianity and other forms of nonlinearity to the extent these complications are features of the data. Under reasonable regularity conditions (smoothness conditions) the SNP density can approximate well any conditional density as K_1, K_2, and M_p grow.

To summarize, the tuning parameters of the SNP-GARCH are

M_u: Lag length of the VAR

M_g: Lag length of the GARCH portion

M_a: Lag length of the ARCH portion

K_1: Degree of the polynomial in z

K_2: Degree of the polynomial dependence on lags of ι_t

M_p: Lag length of the polynomial dependence on lags of ι_t

To determine the appropriate tuning parameters we follow a standard protocal that uses the Schwarz BIC criterion to move upwards along an expansion path. The path first determines M_u, then M_g and M_a. Then K_1 is increased from 0 until BIC reveals an acceptable value or until the upper limit $K_1 = 8$ is reached; the upper limit is imposed on K_1 as the computations become unstable for higher values, which sometimes occur due to over fitting small wiggles in the tails of the data. Lastly, M_p is set to $M_p = 1$ and K_2 is increased from 0 if BIC indicates thusly, and then M_p is increased as well if need be.

2.2.2 Details of EMM

The implementation thus has two parts. The first part is the estimation of the auxiliary score generator model, where here we use the SNP model determined by the model selection procedure just described. The estimated SNP density provides the input to the second stage of the EMM estimation procedure. More precisely, the SNP score function provides the underpinnings for constructing a set of moment conditions. In particular, Θ can be estimated through the moment conditions (score function) similar to (2.1), which in this case will be:

$$m(\Theta, \hat{\Xi}) = E\left(\frac{\partial}{\partial \Xi} \log f_k(\iota_t \mid X_t, \hat{\Xi})\right) = \int \frac{\partial}{\partial \Xi} \log f_k(\iota_t \mid X_t, \hat{\Xi}) dP(\iota, X, \Theta). \tag{2.2}$$

Since these moment conditions should have mean zero they can be used as the basis for the following GMM-type estimation procedure:

$$\hat{\Theta} = argmin_{\Theta}\, m(\Theta, \hat{\Xi})' W_T m(\Theta, \hat{\Xi}). \tag{2.3}$$

where

$$W_T = \frac{1}{T} \sum_{t=1}^{T} \frac{\partial \log f_k(\iota_t \mid X_t, \hat{\Xi})}{\partial \Xi} \frac{\partial \log f_k(\iota_t \mid X_t, \hat{\Xi})}{\partial \Xi'}. \tag{2.4}$$

The use of the outer-product of the gradient formulation for W_T is justified because the SNP provides a close approximation to the true conditional density, so its scores are thereby nearly serially uncorrelated. The moment conditions are easier to compute by simulation instead of computing numerically the integral in (2.2).[9] With simulated time series of length N for ι_t and with candidate parameters Θ, the left hand side of (2.2) translates into

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

Combining (2.3) with (2.5) yields the EMM estimator of Θ the vector containing parameters of the jump-diffusion equations (1.1) through (3.3) including those describing the Lévy measure jump component process.

2.3 Discretization Schemes

The EMM estimation procedure requires simulating equations (1.1) through (3.3). Such a task is now standard for the case of stochastic volatility models. We use the Euler-Maruyama SDE discretization, which is an explicit order 0.5 strong and order 1.0 weak scheme, to simulate the processes appearing in (1.1), (3.2) and (3.3) (see Kloeden and Platen (1995) pp. 486-487). We simulate 55,000 daily observations which are obtained by taking every fourth observation out of 220,000 simulated ones. The first 5,000 observations are dismissed to avoid the starting value problems. Since the case of continuous path diffusions has been covered extensively elsewhere we will focus exclusively here on the simulation of the jump components.

The simulation of jumps is based on the simulation of the time period till the next jump. For instance, for the constant intensity case, we initiate our sample at time 0 and the time of the first jump is determined by the mean duration between jump, which is $1/\lambda$.[10] The subsequent durations between jumps are determined by the draws from the exponential distribution with parameter λ. The jump size is determined simultaneously by drawing from the assumed jump size distribution, i.e $N(\mu_J, \sigma_J)$. In a more complicated case, when

[9] Since jump-diffusion models are formulated in continuous time we need to discretize the process to generate simulated paths. This will be discussed in the next subsection. The empirical results were obtained using simulated samples of size $N = 50,000$.

[10] It should be noted that the way we initiate the drawings is not essential because we dismiss first 5,000 observations.

the intensity depends on the jump size, we first determine the future jump size by drawing from the respective distribution. The randomly drawn value allows us to determine the new intensity value and, hence, the time to the next jump. Finally, when we have an intensity dependent on the realization of the stochastic factors, we can not assume the intensity to be constant for some period of time as was done previously. Therefore, we reevaluate the intensity value and hence the time to the next jump at each simulation step (4 times per day). Every step we obtain the simulated realization of the unobservable factors and substitute them into the expression for the intensity, i.e. $\lambda_t = \lambda + \gamma U_{2t}$, then we draw from the exponential distribution with this value of intensity as a parameter to determine the duration till the next jump. This scheme does not imply that the jump will occur in the obtained time, because at the next step the intensity may (and most likely does) change.

3 Empirical Results

The empirical results appear in Tables 1 through 3. The first two pertain to the estimation of jump diffusions for the S&P500 series and the last reports the finding based on DJIA data. To describe the empirical models it will be useful to introduce some notation. Since it is rather cumbersome to define a comprehensive notation covering all possibilities we do not attempt to be general. Instead, our notation will only aim at describing conveniently the models we will study. The models will be defined as $SVabcJde$ where a is the number of stochastic diffusion factors affecting the drift of the mean return equation (i.e. equation (1.1)) while b is the number of stochastic diffusion factors affecting the volatility of the mean return equation. Both a and b are no greater than c which counts the number of Brownian motions, i.e. the number of processes W_{it} for $i = 1, 2, \ldots$. We focus exclusively on processes with at most three factors, i.e. $c \leq 3$. The features of the jump component are described via d and e which follow J to indicate the presence of a jump process. When Jde is absent we focus exclusively on SV models not involving a jump process. The first letter pertains to the intensity, i.e. $d = C$ for a constant intensity, $d = EQ$ for an exponential intensity involving a quadratic function of the jump size, i.e. the specification appearing in (1.14), $d = EL$ for an exponential intensity involving a linear function of the jump size such as in equations (1.19) and (1.23) and $d = A$ will refer to affine specifications as in Duffie, Pan and Singleton (1998). Finally, e stands for the distributional assumptions regarding the size of the jump, namely $e = N$ is for a normal distribution, $e = E$ for the exponential and finally $e = HN$ for the half-normal. While this notation does not define all possible models unambiguously it will greatly simplify our discussion of the empirical results. We will not discuss the full array of all model specifications because many are not of practical interest. Instead, we start with processes previously considered in the literature and expand them via augmentations involving the new class of processes described in Section 2.

The remaining of this section is organized as follows. The first subsection pertains to the SNP density estimation. The parameter estimates of the jump-diffusion model are discussed in the next two subsections, one is devoted to the S&P 500 data and the other to the DJIA.

3.1 SNP density estimation results

The first step is to determine the appropriate score generator for each of the two time series. For both the S&P 500 and DJIA data sets we estimated SNP-GARCH auxiliary models using as ι_t the one-day geometric return, i.e., 100 times the log first difference of the index. For the S&P 500 data set, the model selection protocol described in Subsection 2.2 gives $M_u = 1, M_g = 1, M_a = 1, K_1 = 8, K_2 = 0$. (With $K_2 = 0$, M_p is irrelevant). This auxiliary model is very similar to an AR(1) with a GARCH(1,1) conditional variance (except SNP-GARCH parameterizes the conditional standard deviation) with a time homogeneous thick-tailed error density; it is akin to a GARCH-t model. The autoregressive model in the leading term captures the modest serial correlation present in the index. In several cases, we need additional moments in the mean to identify all of the parameters of the model, so use an SNP-GARCH model with $M_u = 3$. BIC is known to be very conservative, so there is justification for adding these additional parameters; nonetheless, we expect some imprecision in estimation. For the long DJIA data set, the model selection protocol described in Subsection 2.2 gives $M_u = 1, M_g = 1, M_a = 1, K_1 = 8, M_p = 1, K_2 = 1$. This auxiliary model is an AR(1)-GARCH(1,1) model with a thick-tailed error density with time-varying shape characteristics, as reflected by the dependence of the Hermite coefficients on lagged returns ($K_1 > 0$).

3.2 Jump-diffusion model parameter estimates for S&P500 data

The processes we will consider for the S&P 500 index will be special cases of (1.1), and have the following specification for the stock (index) price S_t:

$$\begin{aligned} dS_t/S_t &= [\mu_1 + \mu_2 U_{2t} + \mu_3 U_{3t} - q - \delta]\, dt + \sqrt{\beta_2 U_{2t} + \beta_3 U_{3t}} dW_{1t} \\ &+ r_2 \sqrt{U_{2t}} dW_{2t} + r_3 \sqrt{U_{3t}} dW_{3t} + e^{\Delta X_t} - 1 \qquad (3.1) \\ dU_{2t} &= (\theta_2 - \kappa_2 U_{2t})dt + \sqrt{U_{2t}} dW_{2t} \qquad (3.2) \\ dU_{3t} &= (\theta_3 - \kappa_3 U_{3t})dt + \sqrt{U_{3t}} dW_{3t} \qquad (3.3) \end{aligned}$$

where W_{it} are independent Brownian motions for $i = 1, 2$ and 3. Because the model is driven by three Brownian motions we shall call it a three-factor SV model with jumps. The processes U_{it} for $i = 2, 3$ are latent volatility factors which are mutually orthogonal by construction. Since the Brownian motions W_{it} are uncorrelated we observe from (3.1) that the leverage effects are represented by r_2 and r_3. The drift in (3.1) includes, a constant μ_1, the continuously compounded dividend rate q and besides the processes U_{it} also δt which is the predictable component of $\exp(X_t) - 1$ where X_t is a Lévy process.[11] Without the Lévy process X_t we would have a stochastic volatility model with three factors, two of which are volatility factors. This class of models has been discussed in detail by Gallant, Hsu and Tauchen (1999). Using daily IBM stock data from January 1985 to January 1997 on close-to-close log price movements and the high/low range, they found that a three factor model does not fit all aspects of the data and in general more factors should be added. A priori we do not need to constraint the class of processes to three-factor SV models, as other factors could be included as well. Anticipating some of the empirical findings we simplified the exposition here by limiting the structure to a three-factor SV process. We start from a three-factor model and instead of augmenting it with a fourth continuous path volatility process we extend the class of processes with a Lévy jump process.

The empirical results appear in two tables, namely Tables 1 and 2. The former contains parameter estimates of three-factor SV models with various specifications for the jump process. Table 2 covers SV model specifications involving a single SV factor augmented with a jump process. The model appearing in the first column of Table 1 is a three-factor SV

[11]Note that application of the Itô formula to the function $\log S_t$ yields the representation in the generic form discussed in the section 1.

process $SV223$, hence without jump component, previously considered in the literature notably by Gallant, Hsu and Tauchen (1999). We note that the parameter estimates are fairly imprecise, i.e. the standard errors are large, and overall the model is rejected (based on the χ^2-statistics appearing at the end of Table 1). The next column in Table 1 is $SV223JCN$. Hence, the three-factor SV model is augmented with a pure jump diffusion Poisson process with a Gaussian distribution for the jump size. We observe that adding the basic constant intensity jump process improves the overall model fit and the parameter estimates $\hat{\mu}_J$, $\hat{\sigma}_J$ and $\hat{\lambda}$ are all very significant. Clearly, as noted by Andersen, Benzoni and Lund (1998), Bates (1998) and others, adding jump processes improves considerably the statistical fit of SV models. The parameter estimates are also close to those obtained by for instance Andersen *et al.* (1998) and Eraker *et al.* (1999), namely our parameter estimate of λ imply roughly three jumps a year.[12] At this point it is worth looking ahead briefly at Table 2, in particular to the model $SV112JCN$. Comparing $SV223$, $SV223JCN$ and $SV112JCN$ allows us to assess the necessity of a second volatility factor when a jump process is present. It is clear from the results pertaining to model $SV112JCN$ appearing in Table 2 that a single factor SV specification is more than adequate. The overall statistical fit of $SV112JCN$ with a χ^2-statistic of 18.56 (with 4 d.f.) is far superior to that of $SV223$ with an overall 48.76 (also with 4 d.f.) or $SV223JCN$ with 32.72 (with only 3 d.f.). In addition to the superior overall fit we also note that the parameter estimates are more precise though their numerical values are roughly the same across the different models.

Before continuing the discussion of two-factor SV models let us briefly explore further some alternative three-factor SV models. As noted before, we can not consider all possibilities since there are too many and only a few appear of interest. Besides $SV223$ and $SV223JCN$ we also report in Table 1 $SV113JCN$ and $SV113JA$ where the latter is an affine jump-diffusion which has three factors, yet only one factor appears in the SV diffusion whereas the second factor drives the random intensity, i.e. $\lambda_t = \lambda + \gamma U_{3t}$. This model might be viewed as

[12] We do not discuss in this paper the issue of modeling the jump intensity under the risk neutral measure. It is worth noting that our findings and those of Andersen *et al.* and Eraker *et al.* are significantly higher than the intensity of one jump every four years found by Bakshi *et al.* (1997) and one jump every two years found by Bates (1998) using exclusively options data. In the companion paper Chernov *et al.* (1999) we examine this issue by allowing the intensity to differ under the objective and risk neutral density. With distinct parameters we are able to better capture the price of jump risk and understand why such discrepancy in jump intensity is found.

a single factor SV model with random jump component and therefore is at the intersection of the models appearing in Table 1 and those appearing in Table 2. The $SV112JCN$ model is a constrained version of $SV222JCN$ where the SV factors affect only respectively the drift and volatility, i.e. $\mu_3 = \beta_2 = r_2 = 0$. Imposing these constraints improves the fit considerably. In particular, if we were to conduct a formal LR-type test we would not reject the null implied by the parameter constraints. The $SV113JCN$ model perform best in terms of overall fit across all the models appearing in Table 1. This includes the $SV113JA$ model which does not fare very well despite the fact it is a single factor model augmented with a jump process. It is worth noting that for the $SV113JA$ model the estimate $\hat{\gamma}$ is statistically insignificant, hence the data rejects the presence of a random affine factor in the jump intensity and the overall statistical fit of $SV113JA$ with 38.25 (with 4 d.f.) is far inferior to that of $SV112JCN$ with 18.56 (also with 4 d.f.). This evidence implies that among the class of affine random intensity two-factor SV models we would settle for the standard Poisson jump specification considered by Andersen, Benzoni and Lund (1998) and Bates (1998).

The results so far seem to lead to two conclusions: (1) a single factor SV model augmented with a Poisson jump process (i.e. $SV112JCN$) performs best and (2) adding a simple affine random intensity $\lambda_t = \lambda + \gamma U_{3t}$ does not seem to improve the statistical fit in comparison to the $SV112JCN$ specification. Hence, so far the constant intensity model appears the most attractive. We compare this model now with two-factor SV models with more general non-affine Lévy measures considered in Section 2. Models of this type will be labeled $d = EQ$ as indicated earlier. In Table 2 we consider various models, namely $SV112JEQN$, $SV112JELE$ and $SV112JLHN$. Hence we consider quadratic and linear dependence of the jump intensity on the jump size. For the linear case we also report two different distributional specifications for the jump size distribution, namely the normal $e = E$ and $e = HN$. Among the three models we find that $SV112JEQN$ with a χ^2 statistic of 11.40 (3 d.f.) is clearly the best and outperforms in terms of statistical fit considerably the standard Poisson process (though it has a χ^2 statistic with 4 d.f it has a lower p-value). It appears that the linear dependence of size is less successful. The model specification involving the half-normal jump size distribution yields an estimate of μ_J not significantly different from zero. Despite this finding it is clear that the alternative specification involving the exponential jump size distribution does not fare very well, indicating that the shape of the exponential distribution is not suitable.[13] The final model we consider in Table 2 is again an affine model, $SV112JA$

[13]Chacko and Viceira (1999) also consider a model involving an exponential jump size distribution, though

which only involves a single SV factor which both drives volatility and the jump intensity. This model does not fare very well, again, indeed it is again inferior to the standard Poisson jump-diffusion process. These findings seem to contradict those of Pan (1999) who rejects the constant intensity model against the random affine intensity model similar to $SV112JA$.[14] She considers a Lagrangian Multiplier test under the null $\lambda = 0$ for the specification $\lambda_t = \lambda + \gamma U_{it}$. The reported p-value of .12 is borderline and the relatively small sample in Pan's study, namely roughly eight years of weekly data or 400 observations, put in doubt the adequacy of the asymptotic distribution theory.[15]

The conclusions we can make at this point is that the new class of non-affine jump-diffusions does improve the statistical fit in comparison to the Poisson constant intensity model. Hence, there is a scope for looking at jump-diffusion processes with intensity structures which have features appearing to be more attractive for financial market behavior description.

3.3 Jump-diffusion model parameter estimates for DJIA data

Estimating jump diffusions with SV still remains a challenge. To the best of our knowledge, we do not know of any prior attempt to fit a diffusion to a data set of over 100 years of daily returns. It is not only the long span of data which is challenging but in particular the turbulent pre-WWII period which has prompted Schwert (1990) and many others to view the 30's as an exceptional period (see in particular Figure 1 which shows the familiar extremely volatile pattern of the 30's). We experimented with many different model specifications and do not report them all. Instead we focus on the most successful fits and alternatives which support hypotheses of interest. The relevant empirical results are summarized in Table 3.

The first two columns of Table 3 report the empirical results for $SV112JCN$ and $SV113JCN$ specifications which fitted the S&P500 data well. Unfortunately both models strongly fail at fitting the DJIA series, judging for instance by the χ^2 statistics. We

they distinguish negative from positive jumps. Their empirical findings are comparable to ours. Namely, they find a negative jump intensity similar to our findings and find a zero positive intensity.

[14] To be precise Pan's model is a $SV314JA$ model.

[15] Note also that the null hypothesis of interest is not really the one considered by Pan, namely $\lambda = 0$ (with $\gamma \neq 0$) (which is the null of the LM test) but rather $\gamma = 0$ (with $\lambda > 0$ for constant intensity) for the specification $\lambda_t = \lambda + \gamma U_{it}$. It is the latter hypothesis which we test here (via a Wald test).

found remarkable improvements by considering a different drift specification for the DJIA data. In particular, we used a drift specification which corresponds to the class of models studied by Lo and Wang (1995), namely:

$$\begin{aligned} dS_t/S_t &= [\mu_1 + \mu_3 U_{3t} - q - \delta]\, dt + \sqrt{\beta_2 U_{2t}} dW_{1t} + r_2 \sqrt{U_{2t}} dW_{2t} + r_3 dW_{3t} \\ &+ e^{\Delta X_t} - 1 \qquad (3.4) \\ dU_{2t} &= (\theta_2 - \kappa_2 U_{2t})dt + \sqrt{U_{2t}} dW_{2t} \qquad (3.5) \\ dU_{3t} &= -\kappa_3 U_{3t} dt + dW_{3t} \qquad (3.6) \end{aligned}$$

where W_{it} are again independent Brownian motions for $i = 1, 2$ and 3. Please note that the U_{3t} is an Ornstein-Uhlenbeck (henceforth OU) process which captures the predictable component of the drift instead of the square root process in (3.1). We will therefore denote this class of processes as $SV - OUabcJcd$.

The model appearing in the third column of Table 3 is $SV - OU113JCN$ which is comparable to $SV113JCN$ and allows us to appraise the improvement in fit due to the change in the specification of the drift. The overall fit, measured by the χ^2 statistic indicates that there is more than a fourfold decrease in the objective function. Since we are dealing with a recalcitrant series we should note that the standard errors of the parameters still remain large. For instance, the parameter estimates of the drift process U_{3t} are insignificant both under the square root and OU specification. However, under the OU drift specification, we observe a very different volatility process (as one would expect from the analysis of Lo and Wang (1995)). On the one hand, under the square root specification the volatility process is highly persistent whereas it is not under the OU drift. The long run mean as measured by $\beta_2\theta_2/\kappa_2$ is higher for the OU specification. On the other hand, it appears that the jump intensity in the $SV113JCN$ model is smaller than when a OU drift is used. Moreover, for the $SV - OU113JCN$ model the average jump size is larger with a smaller standard deviation. Though not reported we also tried to improve the fit by adding an additional volatility factor, i.e. $SV - OU124JCN$ and $SV124JCN$ processes, and found this type of specification not fruitful. These results prompt us to consider refinements of the jump intensity only for the $SV - OU113Jde$ specification.

The last two columns of Table 3 contain the results of the $SV - OU113JQN$ and $SV - OU113JA$. These two columns allow us to appraise the random intensity specification, in particular the affine class of processes and the non-affine class introduced in Section 2. The

$SV - OU113JQN$ specification improves upon the corresponding constant intensity model, whereas the $SV - OU113JA$ does not. Hence, we find results similar to those obtained from the S&P 500 series which indicate that the random non-affine intensity does seem to be a successful specification. One plausible explanation is that the random affine intensity model would require more data to be accurately estimated, despite the fact that we used over 100 years of daily data. In particular, it seems that there are two strategies to fit random affine intensity. We consider both, namely one consists of choosing the intensity as $\lambda_t = \lambda + \gamma U_{2t}$ where U_{2t} also drives latent volatility, the other consists of using $\lambda_t = \lambda + \gamma U_{3t}$ where U_{3t} is correlated with U_{2t}, where the latter drives volatility. The former specification is parsimonious, but creates a tension between the desire to fit volatility well and the need to come up with a reasonable jump intensity. The latter specification resolves this tensions since it disentangles the latent diffusion U_{2t} and the latent intensity U_{3t}. Unfortunately, this affine specification involves many additional parameters, in particular compared to the non-affine class we proposed.

4 Conclusions

In this paper we proposed a new class of jump diffusions which feature both stochastic volatility and random intensity jumps. The non-affine class of processes we study include specifications where the random intensity jump component depends on the size of the previous jump. This class represents an alternative to affine random intensity jump processes which feature correlation between the stochastic volatility and jump component. We also allow for and experiment with different empirical specifications of the jump size distributions. In their most general form we introduced a class of processes which nests jump-diffusions previously considered in empirical work and includes the affine class of random intensity models studied by Bates (1998) and Duffie, Pan and Singleton (1998) but also allows for non-affine random intensity jump components. We used two types of data sets. One involves the S&P500 and the other comprises of 100 years of daily Dow Jones index. The former is a return series often used in the literature and allowed us to compare our results with previous studies. The latter has the advantage to provide a long time series which enhances the possibility of estimating the jump component more precisely. For the DJIA we also imputed a short patch of data which corresponds to the WWI four month market closure. We fitted the processes for the newly extended data set. The non-affine random intensity jump processes are typically more parsimonious than the affine class and appear to fit the data much better compared to the constant and random affine intensity jump diffusion processes. Both for the S&P500 and DJIA series we found that non-affine models fit better than constant intensity Poisson jump diffusions. We also found that random affine intensity models deteriorate the fit when compared to the Poisson specification. It should also be noted that we described several desirable features for jump processes, which are better addressed by the new class of processes we introduced. It is therefore comforting to find that our empirical results confirm the intuition gained from the theoretical specifications.

The processes we introduced share the desirable feature with the affine class that they yield analytically tractable and explicit option pricing formula. In Chernov *et al.* (1999) we take advantage of this and fit the models introduced in this paper to a panel of options and returns data. This not only requires to specify the objective measure process (of the S&P 500 series) but also the risk neutral process and compute option prices.

A A Digression on Lévy Processes

An adapted process $X = (X_t)_{0\leq t<\infty}$ with $X_0 = 0$ a.s. is a Lévy process if (a) X has increments independent of the past; (b) X has stationary increments; and (c) X is continuous in probability.[16] Poisson processes and Brownian motion are special cases of Lévy processes. The jump component of X will be of special interest in our analysis and will be denoted by $\Delta X_t = X_t - X_{t-}$, where $X_{t-} = \lim_{s\uparrow t} X_s$. For any Borel set Λ in $\mathbf{R}$ which is bounded away from 0 we can characterize stopping times $T^n_\Lambda = \inf\{t > T^{n-1}_\Lambda : \Delta X_t \in \Lambda\}$ (where $T^0_\Lambda = 0$) and define a new process as follows:

$$N_t(\Lambda) = \sum_{0<s\leq t} 1_\Lambda(\Delta X_s) := \sum_{n=1}^{\infty} 1_{\{T^n_\Lambda \leq t\}} \tag{A.1}$$

$N_t(\Lambda) - N_s(\Lambda)$ is the number of jumps which the process $X_{s+u} - X_s$ has in $\Lambda, 0 \leq u \leq t-s$. It can be shown that $N(\Lambda)$ is a Poisson process and is the building block for what is known as the Lévy measure. In particular, the Lévy measure $\nu(\Lambda)$ of the Lévy process X for any Borel set Λ is defined as the parameter associated with this Poisson process, namely $\nu(\Lambda) = E\{N_1(\Lambda)\}$. A number of results pertaining to the Lévy measure will be useful to enumerate.

If f is Borel function which is finite on Λ, then the following integral is also a Lévy process:

$$\int_\Lambda f(x)N_t(dx) = \sum_{0<s\leq t} f(\Delta X_s)1_\Lambda(\Delta X_s) \tag{A.2}$$

Moreover, for Borel functions satisfying $f1_\Lambda \in L^2(d\nu)$ we have:

$$E\{\int_\Lambda f(x)N_t(dx)\} = t\int_\Lambda f(x)\nu(dx). \tag{A.3}$$

There are two fundamental results which we will use in the next subsections. The first is known as the Lévy decomposition and the second as the Lévy-Khintchine formula.

(Lévy decomposition). Let X be a Lévy process. Then X has a decomposition:

$$X_t = B_t + \int_{\{|x|<1\}} x(N_t(dx) - t\nu(dx)) + \alpha t + \sum_{0<s\leq t} \Delta X_s 1_\Lambda(\Delta X_s) \tag{A.4}$$

where B is a Brownian motion and $\nu(dx)$ is a Lévy measure on $\mathbf{R}\setminus\{0\}$ such that $\int \min(1, x^2)\nu(dx) < \infty$.

[16] The discussion in this section is based on Protter (1990).

(Lévy-Khintchine formula). The characteristic function of the Lévy process X_t is equal to $exp(-t\psi(u))$, where:

$$\psi(u) = -i\alpha u + \frac{1}{2}\sigma^2 u^2 + \int_{\{|x|\geq 1\}} (1 - e^{iux})\nu(dx) + \int_{\{|x|<1\}} (1 - e^{iux} + iux)\nu(dx). \qquad \text{(A.5)}$$

Moreover, given ν, σ, α, the corresponding Lévy process is unique in distribution.

Both fundamental results will be used in the remaining of this section to construct and analyze the jump component of the different classes of jump-diffusion processes.

References

[1] Andersen, T., L. Benzoni and J. Lund (1998): Estimating Jump-Diffusions for Equity Returns. *Discussion Paper*, Northwestern University.

[2] Bakshi, G. and D. Madan (1998a): Spanning and Derivative-Security Valuation, *Journal of Financial Economics*, (forthcoming).

[3] Bakshi, G. and D. Madan (1998b): What is the Probability of a Stock Market Crash?, *Discussion paper*, University of Maryland.

[4] Ball, C. and W. Torous (1983): A Simplified Jump Process for Common Stock Returns, *Journal of Financial and Quantitative Analysis*, **18**, 53-61.

[5] Ball, C. and W. Torous (1985): On Jumps in Common Stock Prices and Their Impact on Call Option Pricing, *Journal of Finance*, **40**, 155-173.

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

[7] Bates, D.S. (1996b): Testing Option Pricing Models. In Maddala, G.S. and C.R. Rao (eds.), *Handbook of Statistics, Vol. 14*, Elsevier Science B.V.

[8] Bates, D. (1998): Post-'87 Crash Fears in the S&P 500 Futures Option Market, *Journal of Econometrics*, (forthcoming).

[9] Bertoin, J. (1996): *Lévy Processes*, Cambridge University Press

[10] Broadie, M., J. Detemple, E. Ghysels and O. Tòrres (1997): American Options with Stochastic Volatility and Dividends: A Nonparametric Approach, *Journal of Econometrics*, (forthcoming).

[11] Chacko, G. and L. Viceira (1999): Spectral GMM Estimation of Continuous-Time Processes, *Working Paper*, Harvard University

[12] Chernov, M. (1999): Essays in Financial Econometrics, *Ph.D. dissertation*, Pennsylvania State University

[13] Chernov, M. and E. Ghysels (1998): 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* (forthcoming).

[14] Chernov, M., A.R. Gallant, E. Ghysels and G. Tauchen (1999): Estimating Jump-Diffusions with a Panel of Options and Returns, Work in Progress.

[15] Das, S. and R. Sundaram (1997): Of Smiles and Smirks: A Term-Structure Perspective, *Working Paper*, Harvard University

[16] Day, T. and C. Lewis (1992): Stock Market Volatility And the Information Content of Stock Index Options, *Journal of Econometrics*, **52**, 267-287

[17] Duffie, D., J. Pan and K. Singleton (1998): Transform Analysis and Option Pricing for Affine Jump-Diffusions, Working Paper, Stanford University.

[18] Duffie, D. and K. Singleton (1993): Simulated Moments Estimation of Markov Models of Asset Prices, *Econometrica*, **61**, 929-952.

[19] Eraker, B., M. Johannes and N. Polson (1999): Return Dynamics with Jumps to Stochastic Volatility and Returns, *Discussion Paper*, Graduate School of Business, University of Chicago.

[20] Fenton, V.M. and A.R. Gallant (1996): Qualitative and Asymptotic Performance of SNP Density Estimators, *Journal of Econometrics* **74**, 77-118.

[21] Gallant, A.R., D. Hsieh and G. Tauchen (1997): Estimation of stochastic volatility models with diagnostics, *Journal of Econometrics*, **81**, 159-192.

[22] Gallant, A.R., C.-T. Hsu and G.Tauchen (1999): Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance, *Review of Economics and Statistics*, (forthcoming).

[23] Gallant, A.R. and J. R. Long (1997): Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Square, *Biometrika*, **84**, 125-141.

[24] Gallant, A.R., P.E. Rossi and G. Tauchen (1992): Stock Prices and Volume, *Review of Financial Studies* **5**, 199-242.

[25] Gallant, A.R. and G. Tauchen (1989): Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications, *Econometrica*, **57**, 1091-1120.

[26] Gallant, A.R. and G. Tauchen (1993): SNP: A Program for Nonparametric Time Series Analysis, version 8.3, User's Guide, *Discussion paper*, University of North Carolina at Chapel Hill.

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

[28] Gallant, A.R. and G. Tauchen (1997): EMM: a program for efficient method of moments estimation, Version 1.4, User's guide, *Discussion paper*, University of North Carolina at Chapel Hill.

[29] Gallant, A.R. and G. Tauchen (1998): Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions, *Journal of American Statistical Association*, **93**, 10-24.

[30] Ghysels, E., A. Harvey and E. Renault (1996): Stochastic Volatility. In Maddala, G.S. and C.R. Rao (eds.): *Handbook of Statistics, Vol. 14*, Elsevier Science B.V.

[31] Ho, M., W. Perraudin and B. Sørensen (1996): A Continuous-Time Arbitrage-Pricing Model With Stochastic Volatility and Jumps, *Journal of Business and Economic Statistics*, **14**, 31-43.

[32] Jarrow, R.A. and E. Rosenfeld (1984): Jump Risks and the Intertemporal Capital Asset Pricing Model, *Journal of Business*, **57**, 337-351.

[33] Johannes, M., R. Kumar and N.G. Polson (1999): State Dependent Jump Models: How do US Equity Indices Jump?, *Discussion Paper*, Graduate School of Business, University of Chicago.

[34] Jorion, P. (1988): On Jump Processes in the Foreign Exchange and Stock Markets, *Review of Financial Studies*, **1**, 427-445.

[35] Kloeden, P.E. and E. Platen (1995): *Numerical Solution of Stochastic Differential Equations*, Springer Verlag.

[36] Lo, A. W. and J. Wang (1995): Implementing Option Pricing Models when Asset Returns are Predictable, *Journal of Finance* **50**, 87-129.

[37] Pan, J. (1999): "Integrated" Time-Series Analysis of Spot and Option Prices, *Discussion Paper*, Graduate School of Business, Stanford University.

[38] Protter, P. (1990): *Stochastic Integration and Differential Equations*, Springer-Verlag

[39] Schwert, G. W. (1990): Indexes of U.S. Stock Prices from 1802 to 1987, *Journal of Business*, **63**, 399-426.

[40] Scott, L. (1997): Pricing Stock Options in a Jump-Diffusion Model with Stochastic Volatility and Interest Rates: Applications of Fourier Inversion Methods, *Mathematical Finance*, **7**, 413-426.

[41] Tauchen, G. (1997): New Minimum chi-square Methods in Empirical Finance, in Kreps, D. and K. Wallis (ed.) *Advances in Economics and Econometrics: Theory and Applications - Volume III*, Cambridge University Press.

Figure 1. The Dow Jones Industrial Average Index - Daily data 1896 - 1999

The three panels report the (1) level of the DJIA Index, (2) the log of the index and (3) the return computed as 100 times the log first difference of the index. The data are daily and spanning the period from January 1896 to July 1999. The two vertical lines delineate the market closure period which was replaced by imputed data.

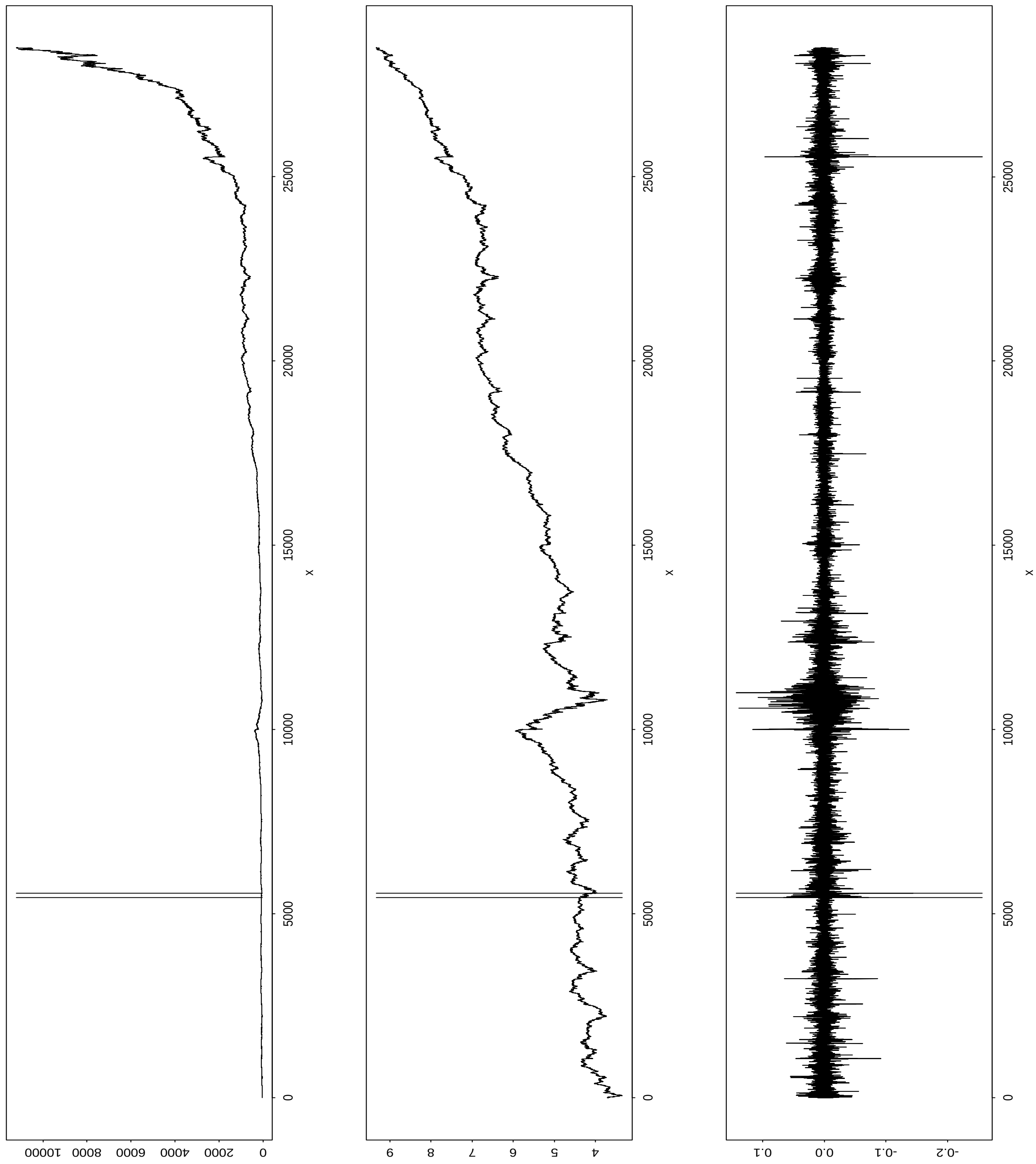

Table 1 : EMM estimation of Two-Factor SV Models with Jumps for S&P 500 data

We have estimated the jump/stochastic volatility diffusion models:

$$
\begin{aligned}
\frac{dS_t}{S_t} &= [\mu_1 + \mu_2 U_{2t} + \mu_3 U_{3t} - q - \delta]\, dt + \sqrt{\beta_2 U_{2t} + \beta_3 U_{3t}} dW_{1t} \\
&+ r_2 \sqrt{U_{2t}} dW_{2t} + r_3 \sqrt{U_{3t}} dW_{3t} + e^{\Delta X_t} - 1 \\
dU_{it} &= (\theta_i - \kappa_i U_{it})dt + \sqrt{U_{it}} dW_{it}, \qquad i = 2, 3 \\
\nu(dx) &= \lambda_t \cdot n(\mu_J, \sigma_J, x)dx \\
x &\sim N(\mu_J, \sigma_J^2)
\end{aligned}
$$

Models are defined as $SVabcJde$ where a is the number of stochastic diffusion factors affecting the drift, b the number of factors affecting the volatility of the mean return equation, c counts the number of Brownian motions. Models with J to indicate the presence of a jump process, with the first letter pertaining to the intensity, i.e. $d = C$ for constant and $d = A$ refers to affine specifications whereas e stands for the jump size distributional assumptions with $e = N$ for normal. The affine model SV113JA has the intensity: $\lambda_t = \lambda + \gamma U_{3t}$ while for all other models $\lambda_t = \lambda$. All models were estimated using the 3111a000 score except for SV113JA which was estimated with the 11118000 score.

Model	SV223	SV223JCN	SV113JCN	SV113JA
$\hat{\mu}_1$	0.132870	0.072948	0.001580*1e-12	0.046320
	(0.213147)	(0.080304)	unstable	(0.075570)
$\hat{\mu}_2$	0.079053	0.079139	0.411781	0.102597
	(0.251951)	(0.094494)	(0.116036)	(0.188348)
$\hat{\mu}_3$	0.010516*1e-4	0.269996*1e-5	-	-
	unstable	unstable		
$\hat{\theta}_2$	0.843691	0.844504	1.508225	1.029086
	(1.991426)	(0.302993)	(0.299325)	(0.971498)
$\hat{\kappa}_2$	0.917504	0.916678	3.549144	0.825234
	(2.762832)	(0.147274)	(0.861993)	(1.052742)
$\hat{\theta}_3$	2.067381	2.067410	1.529808	0.979910
	(30.989558)	(1.958765)	(0.201505)	(35.345242)
$\hat{\kappa}_3$	1.729627	1.729604	2.157226	1.409948
	(28.789018)	(1.595787)	(0.530902)	(50.244581)
$\hat{\beta}_2$	0.003114	0.003849	-	0.005578
	(0.044275)	(0.017498)		(0.023430)
$\hat{\beta}_3$	0.000162	0.0000000	0.016623	-
	(0.051173)	(0.018935)	(0.020961)	
$\hat{r}_2$	-0.096263	-0.097060	-	-0.097655
	(0.000513)	(0.000545)		(0.001850)
$\hat{r}_3$	-0.055639	-0.055137	-0.049610	0.005432*1e-4
	unstable	unstable	(0.000442)	(12.699546)
$\hat{\mu}_J$	-	-0.020521	-0.023427	-0.021676
		(0.000118)	(0.002031)	(0.003932)
$\hat{\sigma}_J$	-	0.009520	0.007059	0.014013
		(0.002201)	(0.002175)	(0.006577)
$\hat{\lambda}$	-	3.000001	2.998357	2.797726
		(0.550854)	(0.600289)	(1.456101)
$\hat{\gamma}$	-	-	-	0.081575
				(0.594263)
χ^2	48.76	32.72	15.51	38.25
d.f.	4	3	2	4

Table 2 : EMM estimation of One-Factor SV Models with Jumps for S&P 500 data

We have estimated the stochastic volatility model

$$\begin{aligned}
\frac{dS_t}{S_t} &= [\mu_1 + \mu_2 U_{2t} - q - \delta]\, dt + \sqrt{\beta_2 U_{2t}} dW_{1t} + r_2 \sqrt{U_{2t}} dW_{2t} + e^{\Delta X_t} - 1 \\
dU_{2t} &= (\theta_2 - \kappa_2 U_{2t}) dt + \sqrt{U_{2t}} dW_{2t} \\
\nu(dx) &= \lambda(x) \cdot n(\mu_J, \sigma_J, x) dx \\
\lambda(x) &= \lambda \exp\left(-\gamma x^2 / 2\sigma_J^2\right) \quad (EQ) \quad or \quad \lambda(x) = \lambda \quad (C) \quad or \quad \lambda(x) = \lambda \exp\left(\gamma x / 2\sigma_J^2\right) \quad (EL) \\
x &\sim N(\mu_J, \sigma_J^2) \quad (N) \quad or \quad x \sim N^-(\mu_J, \sigma_J^2) \quad (HN) \quad or \quad -x \sim E(\mu_J) \quad (E)
\end{aligned}$$

Models are defined as $SVabcJde$ where a is the number of stochastic diffusion factors affecting the drift, b the number of factors affecting the volatility of the mean return equation, c counts the number of Brownian motions. Models with J to indicate the presence of a jump process, with the first letter pertaining to the intensity, i.e. $d = C$ for constant, $d = EQ$ for an exponential intensity involving a quadratic function of the jump size, $d = EL$ for an exponential intensity involving a linear function and $d = A$ refers to affine specifications whereas e stands for the jump size distributional assumptions, $e = N$ for normal, $e = E$ for the exponential and $e = HN$ for the half-normal. The affine model SV113JA has the intensity: $\lambda_t = \lambda + \gamma U_{2t}$. All models were estimated using the 11118000 score.

Model	SV112JEQN	SV112JELHN	SV112JELE	SV112JCN	SV112JA
$\hat{\mu}_1$	0.071972	0.064941	0.161791	0.073498	0.166911
	(0.033751)	(0.070531)	(0.076028)	(0.023511)	(0.035861)
$\hat{\mu}_2$	0.075369	0.098765	0.021839	0.083416	0.172390*1e-4
	(0.056346)	(0.054543)	(0.070181)	(0.049293)	(57.825991)
$\hat{\theta}_2$	0.967770	1.032835	0.947224	1.033178	1.148968
	(0.105565)	(0.060379)	(0.369838)	(0.052287)	(0.133005)
$\hat{\kappa}_2$	0.670705	0.742851	0.686430	0.732717	0.836631
	(0.157884)	(0.050690)	(0.369936)	(0.101032)	(0.018785)
$\hat{\beta}_2$	0.004082	0.004500	0.006209	0.003862	0.005417
	(0.013646)	(0.012423)	(0.005735)	(0.009888)	(0.008309)
$\hat{r}_2$	-0.096373	-0.097367	-0.089315	-0.097716	-0.089973
	(0.000692)	(0.000795)	(0.000605)	(0.000475)	(0.000452)
$\hat{\mu}_J$	-0.021436	0.010035	0.012657	-0.021770	-0.018917
	(0.000686)	(0.041816)	(0.000619)	(0.000762)	(0.001681)
$\hat{\sigma}_J$	0.010403	0.018826	-	0.010169	0.009608
	(0.002154)	(0.037472)		(0.002381)	(0.002761)
$\hat{\lambda}$	3.000031	2.936123	3.278344	2.999988	2.814294
	(0.273190)	(0.029980)	(0.321760)	(0.161400)	(0.176918)
$\hat{\gamma}$	0.115995*1e-4	0.099782	0.008445*1e-4	-	0.088978
	(0.083194)	(0.059141)	(7.943110)	-	(0.064554)
χ^2	11.40	16.48	25.55	18.56	19.61
d.f.	3	3	4	4	3

Table 3 : EMM estimation of Two-Factor SV Models with Jumps for DJIA data

In addition to the $SVabcJed$ specification (see Tables 1 and 2) we have estimated the jump/stochastic volatility diffusion models, denoted $SV-OUabcJde$:

$$\begin{aligned}
\frac{dS_t}{S_t} &= [\mu_1 + \mu_3 U_{3t} - q - \delta]\, dt + \sqrt{\beta_2 U_{2t}} dW_{1t} + r_2 \sqrt{U_{2t}} dW_{2t} + r_3 dW_{3t} + e^{\Delta X_t} - 1 \\
dU_{2t} &= (\theta_2 - \kappa_2 U_{2t})dt + \sqrt{U_{2t}} dW_{2t} \\
dU_{3t} &= -\kappa_3 U_{3t} dt + dW_{3t} \\
\nu(dx) &= \lambda_t \cdot n(\mu_J, \sigma_J, x)dx \\
x &\sim N(\mu_J, \sigma_J^2)
\end{aligned}$$

We used the 11118010 score and simulated 100,000 observations, with q set equal to 0.02 as with S&P 500. For the affine specification, we used $\lambda_t = \lambda + \gamma U_{2t}$

Model	$SV112JCN$	$SV113JCN$	$SV-OU113JCN$	$SV-OU113JQN$	$SV-OU113JA$
$\hat{\mu}_1$	0.046559	0.157500	0.087286	0.092479	0.087532
	(0.017538)	(0.607279)	(0.022674)	(0.019105)	(0.020465)
$\hat{\mu}_3$	0.120493	-0.048095	1.045668	1.413880	1.006125
	(0.023567)	(0.7222656)	(0.982440)	(1.225785)	(0.732928)
$\hat{\theta}_2$	0.857881	0.884507	1.056946	1.056889	1.056959
	(0.011708)	(0.219369)	(0.151645)	(0.145666)	(0.116551)
$\hat{\kappa}_2$	1.547089	1.177652	5.101415	5.100381	5.101508
	(0.106633)	(0.686901)	(1.444110)	(1.583616)	(1.259693)
$\hat{\theta}_3$	-	1.703226	-	-	-
		(9.328473)			
$\hat{\kappa}_3$	-	1.986907	2.231755	3.849306	2.073362
		(11.768044)	(4.220407)	(6.876322)	(3.057748)
$\hat{\beta}_2$	0.002223	0.011461	0.048722	0.050548	0.048503
	(0.006258)	(0.006447)	(0.014084)	(0.013840)	(0.013933)
$\hat{r}_2$	-0.193783	-0.111399	-0.125386	-0.121842	-0.124842
	(0.009200)	(0.026412)	(0.031573)	(0.037093)	(0.023539)
$\hat{r}_3$	-	0.021566	-0.004370	-0.004093	-0.003031
		(0.011901)	(0.045553)	(0.039771)	(0.035396)
$\hat{\mu}_J$	-0.021106	-0.014417	-0.019792	-0.019364	-0.019445
	(0.001765)	(0.002048)	(0.001352)	(0.002222)	(0.001375)
$\hat{\sigma}_J$	0.009554	0.013148	0.004024	0.006076	0.002783
	(0.002666)	(0.002194)	(0.001387)	(0.001590)	(0.001706)
$\hat{\lambda}$	2.933500	2.920929	3.505758	3.506019	2.979023
	(0.020266)	(0.023720)	(0.014583)	(0.021218)	(0.128849)
$\hat{\gamma}$	-	-	-	0.094175*1e-4	2.146769
				(0.020821)	(0.458639)
χ^2	330.87	329.19	76.62	72.24	87.17
d.f.	13	10	11	10	10