

NBER WORKING PAPER SERIES

U.S. STOCK MARKET CRASH RISK, 1926-2006

David S. Bates

Working Paper 14913

<http://www.nber.org/papers/w14913>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

April 2009

I am grateful for comments on earlier versions of the paper from seminar participants at Iowa, Turin, Birkbeck, McGill's 2008 Risk Management Conference, the 18th Annual Derivatives Securities and Risk Management Conference in Arlington, VA, and Princeton's 2008 Conference on Implied Volatility Models. I also thank Ken French for providing the stock turnover data from French (2008). The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2009 by David S. Bates. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

U.S. Stock Market Crash Risk, 1926-2006  
David S. Bates  
NBER Working Paper No. 14913  
April 2009  
JEL No. C22,C46,G1,G13

**ABSTRACT**

This paper applies the Bates (RFS, 2006) methodology to the problem of estimating and filtering time-changed Lévy processes, using daily data on U.S. stock market excess returns over 1926-2006. In contrast to density-based filtration approaches, the methodology recursively updates the associated conditional characteristic functions of the latent variables. The paper examines how well time-changed Lévy specifications capture stochastic volatility, the “leverage” effect, and the substantial outliers occasionally observed in stock market returns. The paper also finds that the autocorrelation of stock market excess returns varies substantially over time, necessitating an additional latent variable when analyzing historical data on stock market returns. The paper explores option pricing implications, and compares the results with observed prices of options on S&P 500 futures.

David S. Bates  
Henry B. Tippie College of Business  
Department of Finance  
University of Iowa  
Iowa City, IA 52242-1000  
and NBER  
david-bates@uiowa.edu

What is the risk of stock market crashes? Answering this question is complicated by two features of stock market returns: the fact that conditional volatility evolves over time, and the fat-tailed nature of daily stock market returns. Each issue affects the other. What we identify as outliers depends upon that day's assessment of conditional volatility. Conversely, our estimates of current volatility from past returns can be disproportionately affected by outliers such as the 1987 crash. In standard GARCH specifications, for instance, a 10% daily change in the stock market has 100 times the impact on conditional variance revisions of a more typical 1% move.

This paper explores whether recently proposed continuous-time specifications of time-changed Lévy processes are a useful way to capture the twin properties of stochastic volatility and fat tails. The use of Lévy processes to capture outliers dates back at least to Mandelbrot's (1963) use of the stable Paretian distribution, and there have been many specifications proposed; e.g., Merton's (1976) jump-diffusion, Madan and Seneta's (1990) variance gamma; Eberlein, Keller and Prause's (1998) hyperbolic Lévy; and Carr, Madan, Geman and Yor's (2002) CGMY process. As all of these distributions assume identical and independently distributed returns, however, they are unable to capture stochastic volatility.

More recently, Carr, Geman, Madan and Yor (2003) and Carr and Wu (2004) have proposed combining Lévy processes with a subordinated time process. The idea of randomizing time dates back to at least to Clark (1973). Its appeal in conjunction with Lévy processes reflects the increasing focus in finance – especially in option pricing – on representing probability distributions by their associated characteristic functions. Lévy processes have log characteristic functions that are linear in time. If the time randomization depends on underlying variables that have an analytic conditional characteristic function, the resulting conditional characteristic function of time-changed Lévy processes is also analytic. Conditional probability densities, distributions, and option prices can then be numerically computed by Fourier inversion of simple functional transforms of this characteristic function.

Thus far, empirical research on the relevance of time-changed Lévy processes for stock market returns has largely been limited to the special cases of time-changed versions of Brownian

motion and Merton's (1976) jump-diffusion. Furthermore, there has been virtually no estimation of newly proposed time-changed Lévy processes solely from time series data.<sup>1</sup> Papers such as Carr et al (2003) and Carr and Wu (2004) have relied on option pricing evidence to provide empirical support for their approach, rather than providing direct time series evidence. The reliance on options data is understandable. Since the state variables driving the time randomization are not directly observable, time-changed Lévy processes are *hidden Markov* models – a challenging problem in time series econometrics. Using option prices potentially identifies realizations of those latent state variables, converting the estimation problem into the substantially more tractable problem of estimating state space models with observable state variables.

While options-influenced parameter and state variable estimates should be informative under the hypothesis of correct model specification, the objective of the paper is to provide estimates of crash risk based *solely* upon time series analysis. Such estimates are of interest in their own right, and can exploit a longer history of extreme stock market movements than can studies constrained by the availability of options data only since the 1980's. For instance, the **-20%** stock market crash of October 19, 1987 was the only daily stock market movement over 1945-2006 to exceed 10% in magnitude, whereas there were seven such movements over 1929-32. Furthermore, time-series based estimates can be relevant even for testing option pricing hypotheses. For instance, it has been asserted that deep OTM index put options appear overpriced, based on their surprisingly large negative returns since the '87 crash. But all such tests require reliable estimates of downside risk; and it can be difficult to establish whether puts are indeed overpriced based only on the limited amount of data since the 1987 crash.<sup>2</sup> Time series estimates can exploit a longer history of downside risk, and can be used to generate estimates of option prices that can be compared with observed option prices.

---

<sup>1</sup>Li, Wells and Yu (2008) use MCMC methods to estimate some models in which Lévy shocks are added to various stochastic volatility models. However, the additional Lévy shocks are i.i.d., rather than time-changed.

<sup>2</sup>See Broadie, Chernov, and Johannes (2006) for a Monte Carlo study of unhedged 1-month excess returns for puts on S&P 500 futures over August 1987 to June 2005. They find excess return estimates often lack statistical significance, especially when volatility is stochastic.

This paper provides direct time series estimates of some proposed time-changed Lévy processes, using the Bates (2006) *approximate maximum likelihood* (AML) methodology. AML is a filtration methodology that recursively updates conditional characteristic functions of latent variables over time given observed data. Filtered estimates of the latent variables are directly provided as a by-product, given the close link between moments and characteristic functions. The paper primarily focuses on the time-changed CGMY process, which nests various other processes as special cases. The approach will also be compared to the time-changed jump-diffusions previously estimated in Bates (2006).

A concern with any extended data set is the possibility that the data generating process may not be stable over time. Indeed, this paper identifies a major instability in the autocorrelation of daily stock market returns. Autocorrelation estimates appear to be nonstationary, and peaked at the extraordinarily high level of 35% in 1971, before trending downwards to the near-zero values observed since the 1980's. The instability is addressed directly, by treating the autocorrelation as another latent state variable to be estimated from observed stock market returns. The paper also finds apparent instabilities or specification issues in the 1-factor volatility process used, and explores the implications for volatility filtration and option pricing.

Overall, the time-changed CGMY process is found to be a slightly more parsimonious alternative to the Bates (2006) approach of using finite-activity stochastic-intensity jumps drawn from a mixture of normals, although the fits of the two approaches are not dramatically different. Interestingly, one cannot reject the hypothesis that stock market crash risk is adequately captured by a time-changed version of the Carr-Wu (2003) log-stable process. That model's implications for *upside* risk, however, are strongly rejected, with the model severely underpredicting the frequency of large positive outliers.

Section I of the paper progressively builds up the time series model used in estimation. Section I.1 discusses basic Lévy processes and describes the processes considered in this paper. Section I.2 discusses time changes and the equivalence with stochastic volatility. Section I.3 contains further modifications of the model to capture leverage effects, time-varying autocorrelations, and day-of-the-week effects. Section I.4 describes how the model is estimated,

using the Bates (2006) AML estimation methodology for hidden Markov models.

Section II describes the data on excess stock market returns over 1926-2006, and presents parameter estimates, diagnostics, and filtered estimates of latent autocorrelation and volatility. Section III examines option pricing implications, while Section IV concludes.

## I. Time-changed Lévy processes

### I.1 Lévy processes

A Lévy process  $L(t)$  is an infinitely divisible stochastic process; i.e., one that has independent and identically distributed increments over non-overlapping time intervals of equal length. The Lévy processes most commonly used in finance have been Brownian motion and the jump-diffusion process of Merton (1976), but there are many others. All Lévy processes without a Brownian motion component are pure jump processes. Such processes are characterized by their *Lévy density*  $k(x)$ , which gives the intensity (or frequency) of jumps of size  $x$ . Alternatively and equivalently, Lévy processes can be described by their generalized Fourier transform

$$F(u) \equiv E e^{uL(t)} = \exp[tf_{dL}(u)], \quad u \in D_u \subset \mathbb{C} \quad (1)$$

where  $u$  is a complex-valued element of the set  $D_u$  for which (1) is well-defined. If  $\Phi$  is real,  $F(i\Phi)$  is the characteristic function of  $L(t)$ , while  $tf_{dL}(\Phi)$  is the cumulant generating function of  $L(t)$ . Its linearity in time follows from the fact that Lévy processes have i.i.d. increments. Following Wu (2006), the function  $f_{dL}(u)$  will be called the *cumulant exponent* of  $L(t)$ .<sup>3</sup>

The Lévy-Khintchine formula gives the mapping between jump intensities  $k(x)$  and the cumulant exponent for arbitrary  $u \in D_u$ . Lévy processes in finance are typically specified for the log asset price, and then exponentiated:  $S(t) = \exp[L(t)]$ . For such specifications, it is convenient to write the Lévy-Khintchine formula in the form

$$f_{dL}(u) = u\mu + \int_{\mathbb{R} - \{0\}} [e^{ux} - 1 - u(e^x - 1)]k(x)dx, \quad (2)$$

where  $\mu = f_{dL}(1)$  is the continuously-compounded expected return on the asset:

---

<sup>3</sup>Carr et al (2003) call  $f_{dL}(i\Phi)$  the “unit time log characteristic function.” Bertoin (1996) uses the *characteristic exponent*, which takes the form  $\Psi(\Phi) \equiv -f_{dL}(i\Phi)$ .

$$ES(t) = Ee^{L(t)} = e^{\int \lambda k(x) dx} = e^{\mu t}. \quad (3)$$

Pure-jump Lévy processes can be thought of as a drift term plus an infinite sum  $L(t) = \int L_x(t) dx$  of independent point processes, each drift-adjusted to make  $\exp[L_x(t)]$  a martingale:

$$dL_x = x dN_x - (e^x - 1)k(x)dt, \quad (4)$$

where  $N_x$  is an integer-valued Poisson counter with intensity  $k(x)$  that counts the occurrence of jumps of fixed size  $x$ . The log characteristic function of a sum of independent point processes is the sum of the log characteristic functions of the point processes, yielding equation (2). Exponential martingale processes of the form  $L(t) = \int L_x(t) dx$  for  $L_x$  defined in (4) will be termed *compensated Lévy processes*, as will also diffusions of the form  $\sigma W_t - \frac{1}{2}\sigma^2 t$ .

As discussed in Carr et al (2002), Lévy processes are *finite-activity* if  $\int k(x) dx < \infty$ , and *infinite-activity* otherwise. Finite-activity jumps imply there is a non-zero probability that no jumps will be observed within a given time interval. Lévy processes are *finite-variation* if  $\int |x| k(x) dx < \infty$ , and infinite-variation otherwise. An infinite-variation process has sample paths of infinite length – a property also of Brownian motion. All Lévy processes must have finite  $\int \min(x^2, 1) k(x) dx$ , in order to be well-behaved, but need not have finite variance  $\int x^2 k(x) dx$  – the stable distribution being a counterexample. A priori, all financial prices must be finite-activity processes, since price changes reflect a finite (but large) number of market transactions. However, finite-activity processes can be well approximated by infinite-activity processes, and vice versa; e.g., the Cox, Ross and Rubinstein (1979) finite-activity binomial approximation to Brownian motion. Activity and variation will therefore be treated as empirical specification issues concerned with identifying which functional form  $k(x)$  for jump intensities best fits daily stock market excess returns.

I will consider two particular underlying Lévy processes for log asset prices. The first is Merton (1976)'s combination of a Brownian motion plus finite-activity normally distributed jumps:

$$d \ln S_t = \mu dt + (\sigma dW_t - \frac{1}{2}\sigma^2 dt) + (\gamma dN_t - \lambda \bar{k} dt) \quad (5)$$

where  $W_t$  is a Wiener process,

$N_t$  is a Poisson counter with intensity  $\lambda$ ,

$\gamma \sim N(\bar{\gamma}, \delta^2)$  is the normally distributed jump conditional upon a jump occurring, and

$\bar{k} = e^{\bar{\gamma} + \frac{1}{2}\delta^2} - 1$  is the expected percentage jump size conditional upon a jump.

The associated intensity of jumps of size  $x$  is

$$k(x) = \frac{\lambda}{\sqrt{2\pi\delta^2}} \exp\left[-\frac{(x - \bar{\gamma})^2}{2\delta^2}\right] \quad (6)$$

while the cumulant exponent takes the form

$$f_{Merton}(u) = \mu u + \frac{1}{2}\sigma^2(u^2 - u) + \lambda(e^{\bar{\gamma}u + \frac{1}{2}\delta^2 u^2} - 1 - u\bar{k}).$$

The approach can be generalized to allow alternate distributions for  $\gamma$  – in particular, a mixture of normals:

$$k(x) = \sum_{i=1}^2 \frac{\lambda_i}{\sqrt{2\pi\delta_i^2}} \exp\left[-\frac{(x - \bar{\gamma}_i)^2}{2\delta_i^2}\right]. \quad (7)$$

Second, I will consider the generalized CGMY process of Carr, Madan, Geman and Yor (2003), which has a Lévy density of the form

$$k(x) = \begin{cases} C_n e^{-G|x|} |x|^{-1-Y_n} & \text{for } x < 0 \\ C_p e^{-M|x|} |x|^{-1-Y_p} & \text{for } x > 0 \end{cases} \quad (8)$$

where  $C_n, C_p, G, M \geq 0$  and  $Y_p, Y_n < 2$ . The associated cumulant exponent is

$$f_{CGMY}(u) = (\mu - \omega)u + V \left\{ w_n \frac{(G - u)^{Y_n} - G^{Y_n}}{Y_n(Y_n - 1)G^{Y_n - 2}} + (1 - w_n) \frac{(M + u)^{Y_p} - M^{Y_p}}{Y_p(Y_p - 1)M^{Y_p - 2}} \right\} \quad (9)$$

where  $\omega$  is a mean-normalizing constant determined by  $f_{CGMY}(1) = \mu$ ;

$V$  is the variance per unit time, and

$w_n$  is the fraction of variance attributable to the downward-jump component.

The corresponding intensity parameters  $C_n, C_p$  in (8) are



$$C_n = \frac{w_n V}{\Gamma(2 - Y_n) G^{Y_n - 2}}, \quad C_p = \frac{(1 - w_n) V}{\Gamma(2 - Y_p) M^{Y_p - 2}} \quad (10)$$

where  $\Gamma(z)$  is the gamma function.

As discussed in Carr et al (2002), the  $Y$  parameters are key in controlling jump activity near 0, in addition to their influence over tail events. The process has finite activity for  $Y_p, Y_n < 0$ , finite variation for  $Y_p, Y_n < 1$ , but infinite activity or variation if  $\min(Y_p, Y_n)$  is greater or equal to 0 or 1, respectively. The model conveniently nests many models considered elsewhere. For instance,  $Y_n = Y_p = -1$  is the finite-activity double exponential jump model of Kou (2002), while  $Y_n = Y_p = 0$  includes the variance gamma model of Madan and Seneta (1990). As  $Y_p$  and  $Y_n$  approach 2, the CGMY process converges to a diffusion, and the cumulant exponent converges to the corresponding quadratic form

$$f_{CGMY}(u) = \mu u + \frac{1}{2} V (u^2 - u). \quad (11)$$

As  $G$  and  $M$  approach 0 (for arbitrary  $Y_p, Y_n$  and fixed  $C_n, C_p$ ), the Lévy density (8) approaches the infinite-variance log stable process advocated by Mandelbrot (1963), with a “power law” property for asymptotic tail probabilities. The log-stable special case proposed by Carr and Wu (2003) is the limiting case with only negative jumps ( $C_p = 0$ ). While infinite-variance for log returns, percentage returns have finite mean and variance under the log-stable specification. For daily stock market returns of less than 25% in magnitude, the log-stable process is well approximated by a finite-variance CGMY process with minimal exponential dampening; e.g.,  $G = .001$ .

The cumulant exponent of any finite-variance Lévy process can be written in the form

$$f_{dL}(u) = V g_{dL}(u) \quad (12)$$

where  $V = f_{dL}''(0)$  is variance per unit time and  $g_{dL}(u)$  is a standardized cumulant exponent with unitary variance. One can also combine Lévy processes, to nest alternative specifications within a broader specification. Any linear combination  $w_1 k_1(x) + w_2 k_2(x)$  of Lévy densities for nonnegative

weights that sum to one is also a valid Lévy density, and generates an associated standardized weighted cumulant exponent of the form  $w_1 g_1(\mathbf{u}) + w_2 g_2(\mathbf{u})$ , where  $g_i(\mathbf{u})$  is the standardized cumulant exponent associated with  $k_i(x)$  for  $i = 1, 2$ . The various  $g_{dL}(\mathbf{u})$  specifications that will be considered in this paper are listed in Table 1.

## I.2 Time-changed Lévy processes and stochastic volatility

Time-changed Lévy processes generate stochastic volatility by randomizing time in equation (1). Since the log transform (1) can be written as  $\ln F(\mathbf{u}) = g_{dL}(\mathbf{u}) V t$ , randomizing time is fundamentally equivalent to randomizing variance. As the connection between time changes and stochastic volatility becomes less transparent once “leverage” effects are added, I will use an explicit stochastic volatility (or stochastic intensity) representation of stochastic processes.

The leverage effect, or correlation between asset returns and conditional variance innovations, is captured by directly specifying shocks common to both. I will initially assume that the log asset price  $s_t \equiv \ln S_t$  follows a process of the form

$$\begin{aligned} ds_t &= (\mu_0 + \mu_1 V_t) dt + \left( \rho_{sv} \sqrt{V_t} dW_t - \frac{1}{2} \rho_{sv}^2 V_t dt \right) + dL_t \\ dV_t &= (\alpha - \beta V_t) dt + \sigma \sqrt{V_t} dW_t \end{aligned} \quad (13)$$

The log increment  $ds_t$  consists of the continuously-compounded return, plus increments to two exponential martingales.  $dW_t$  is a Wiener increment, while  $dL_t$  is the increment to a compensated Lévy process, with finite instantaneous variance  $(1 - \rho_{sv}^2) V_t dt$ . Further refinements will be added below, to match properties of stock market returns more closely.

This specification has various features or implicit assumptions. First, the approach allows considerable flexibility regarding the distribution of the instantaneous shock  $dL_t$  to asset returns, which can be Wiener, compound Poisson, or any other fat-tailed distribution. Three underlying Lévy processes are considered:

- 1) a second diffusion with variance  $(1 - \rho_{sv}^2) V_t$  that is independent of  $W_t$  (Heston, 1993);
- 2) finite-activity jumps drawn from a normal distribution or a mixture of normals; and
- 3) the generalized CGMY (2003) Lévy process from (8) above.

Combinations of these processes will also be considered, to nest the alternatives.

Second, the specification assumes a single underlying variance state variable  $V_t$  that follows an affine diffusion, and which directly determines the variance of diffusion *and* jump components. This approach generalizes the stochastic jump intensity model of Bates (2000, 2006) to arbitrary Lévy processes.

Two alternate specifications are not considered, for different reasons. First, I do not consider the approach of Li, Wells and Yu (2008), who model log-differenced asset prices as the sum of a Heston (1993) stochastic volatility process and a constant-intensity fat-tailed Lévy process that captures outliers. Bates (2006, Table 7) found the stochastic-intensity jump model fits S&P returns better than the constant-intensity specification, when jumps are drawn from a finite-activity normal distribution or mixture of normals. Second, the diffusion assumption for  $V_t$  rules out volatility-jump models, such as the exponential-jump model proposed by Duffie, Pan and Singleton (2000) and estimated by Eraker, Johannes and Polson (2003). Estimation on simulated data indicates that the AML filtration methodology described below has difficulty identifying whether there are jumps in an underlying conditional variance state variable that is not directly observed.

Define  $y_T \equiv \int_{r=t}^T ds_r$  as the discrete-time return observed over horizon  $\tau = T - t$ , and define  $f_{dL}(\mathbf{u}) \equiv (1 - \rho_{sv}^2)V_t g_{dL}(\mathbf{u})$  as the cumulant exponent of  $dL_t$  conditional upon knowing  $V_t$ . By construction,  $g_{dL}(\mathbf{u})$  is a standardized cumulant exponent, with  $g_{dL}(\mathbf{1}) = \mathbf{0}$  and variance  $g_{dL}''(\mathbf{0}) = 1$ . A key property of affine models is the ability to compute the conditional generalized Fourier transform of  $(y_T, V_T)$ . This can be done by iterated expectations, conditioning initially on the future variance path:

$$\begin{aligned}
F(\Phi, \Psi \mid V_t, \tau) &\equiv E\left(e^{\Phi y_T + \Psi V_T} \mid V_t\right) \\
&= E\left\{E\left[e^{\Phi \mu_0 \tau + \int_{r=t}^T \Phi \left[\mu_1 - \frac{1}{2}\rho_{sv}^2\right] V_r dr + \rho_{sv} \sqrt{V_r} dW_r + dL_r} + \Psi V_T \mid \{V_r\}_{r=t}^T} \mid V_t\right]\right\} \\
&= E\left[e^{\Phi \mu_0 \tau + \int_{r=t}^T \left[\mu_1 \Phi + \frac{1}{2}\rho_{sv}^2(\Phi^2 - \Phi) + (1 - \rho_{sv}^2)g_{dL}(\Phi)\right] V_r dr + \Psi V_T} \mid V_t\right] \\
&\equiv E\left[e^{\Phi \mu_0 \tau + h(\Phi) \int_{r=t}^T V_r dr + \Psi V_T} \mid V_t\right]
\end{aligned} \tag{14}$$

for  $h(\Phi) \equiv \mu_1 \Phi + \frac{1}{2}\rho_{sv}^2(\Phi^2 - \Phi) + (1 - \rho_{sv}^2)g_{dL}(\Phi)$ . This is the generalized Fourier transform of the future spot variance  $V_T$  and the average future variance  $\bar{V}_{t \rightarrow T} \equiv \frac{1}{\tau} \int_{r=t}^T V_r dr$ . This is a well-known

problem (see, e.g., Bakshi and Madan, 2000), with an analytic solution if  $V_t$  follows an affine process. For the affine diffusion above,  $F(\bullet | V_t, \tau)$  solves the Feynman-Kac partial differential equation

$$-F_\tau + (\alpha - \beta V)F_V + \frac{1}{2}\sigma^2 VF_{VV} = -[\Phi\mu_0 + h(\Phi)V_t]F \quad (15)$$

subject to the boundary condition  $F(\Phi, \psi | V, 0) = \exp(\psi V)$ . The solution is

$$F(\Phi, \psi | V_t, \tau) = \exp[C(\tau; \Phi, \psi) + D(\tau; \Phi, \psi)V_t] \quad (16)$$

where

$$\begin{aligned} C(\tau; \Phi, \psi, \xi) &= \mu_0 \Phi \tau - \frac{\alpha \tau}{\sigma^2} (\rho_{sv} \sigma \Phi - \beta - \gamma) \\ &\quad - \frac{2\alpha}{\sigma^2} \ln \left[ 1 + \frac{1}{2} (\rho_{sv} \sigma \Phi - \beta - \gamma) \frac{1 - e^{\gamma \tau}}{\gamma} \right] - \frac{2\alpha}{\sigma^2} \ln[1 - \mathbf{K}(\Phi)\psi] \end{aligned} \quad (17)$$

$$D(\tau; \Phi, \psi, \xi) = \frac{2\mu_1 \Phi + \rho_{sv}^2 (\Phi^2 - \Phi) + 2(1 - \rho_{sv}^2)g_{dL}(\Phi)}{\gamma \frac{e^{\gamma \tau} + 1}{e^{\gamma \tau} - 1} + \beta - \rho_{sv} \sigma \Phi} + \frac{\Lambda(\Phi)\psi}{1 - \mathbf{K}(\Phi)\psi} \quad (18)$$

$$\gamma = \sqrt{(\rho_{sv} \sigma \Phi - \beta)^2 - 2\sigma^2 [\mu_1 \Phi + \frac{1}{2}\rho_{sv}^2 (\Phi^2 - \Phi) + (1 - \rho_{sv}^2)g_{dL}(\Phi)]} \quad (19)$$

$$\Lambda(\Phi) = \frac{\left( \frac{e^{\gamma \tau} + 1}{e^{\gamma \tau} - 1} \right)^2 - 1}{\left( \frac{e^{\gamma \tau} + 1}{e^{\gamma \tau} - 1} + \frac{\beta - \rho \sigma \Phi}{\gamma} \right)^2} \quad (20)$$

$$\mathbf{K}(\Phi) = \frac{\sigma^2}{\gamma \frac{e^{\gamma \tau} + 1}{e^{\gamma \tau} - 1} + \beta - \rho \sigma \Phi} \quad (21)$$

The specifications of  $g_{dL}(\cdot)$  considered in this paper are listed above in Table 1.

### I.3 Autocorrelations and other refinements

That stock indexes do not follow a random walk was recognized explicitly by Lo and MacKinlay (1988), and implicitly by earlier practices in variance and covariance estimation designed to cope

with autocorrelated returns; e.g., Dimson (1979)'s lead/lag approach to beta estimation. The positive autocorrelations typically estimated for stock index returns are commonly attributed to stale prices in the stocks underlying the index. A standard practice in time series analysis is to pre-filter the data by fitting an ARMA specification; see, e.g., Jukivuolle (1995). Andersen, Benzoni and Lund (2002), for instance, use a simple MA(1) specification to remove autocorrelations in S&P 500 returns over 1953-96; a data set subsequently used by Bates (2006).

Pre-filtering the data was considered unappealing in this study, for several reasons. First, the 1926-2006 interval used here is long, with considerable variation over time in market trading activity and transactions costs, and structural shifts in the data generating process are probable. Indeed, Andersen et al (2002, Table 1) find autocorrelation estimates from their full 1953-96 sample diverge from estimates for a 1980-96 subsample. Second, ARMA packages use a mean squared error criterion that is not robust to the fat tails observed in stock market returns. Finally, explicit consideration of autocorrelation is needed when assessing the variance of relevance to option pricing.

Consequently, autocorrelations were treated as an additional latent variable, to be estimated as part of the overall time series model. I will explore below two alternate models for daily log-differenced stock index excess returns  $y_t$ :

$$y_{t+1} = \rho_t y_t + \eta_{t+1} \quad (\text{Model 1}) \quad (22)$$

or

$$y_{t+1} = \rho_t y_t + (1 - \rho_t) \eta_{t+1} \quad (\text{Model 2}) \quad (23)$$

where

$$\begin{aligned} \eta_{t+1} &\equiv \int_{r=t}^{t+\tau_t} ds_r \\ V_{t+1} &= V_t + \int_{r=t}^{t+\tau_t} dV_r \\ \rho_{t+1} &= \rho_t + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim N(0, \sigma_\rho^2) \text{ and } i.i.d. \end{aligned} \quad (24)$$

$\tau_t$  is the effective length of a business day,  $\rho_t$  determines is the daily autocorrelation,  $ds_t$  is the instantaneous intradaily underlying shock to log asset prices, and  $V_t dt \equiv \text{Var}_t(ds_t)$  is the instantaneous conditional variance of  $ds_t$ . The intradaily shocks ( $ds_t, dV_t$ ) are given by (13) above.

Both models add an autocorrelation state variable  $\rho_t$  that captures the fact that autocorrelations of stock market returns are not constant over time.<sup>4</sup> Following the literature on time-varying coefficient models, the autocorrelation is modeled as a simple random walk, to avoid constraining estimates of  $\rho_t$ . Estimation of the autocorrelation volatility parameter  $\sigma_\rho$  endogenously determines the appropriate degree of smoothing to use when filtering the current autocorrelation value  $\rho_t$  from past data.

The two models differ in ease of use, in their implications for the interaction between volatility and autocorrelation, and in the pricing of risks. Model 1 assumes the stock market excess return residual  $\eta_{t+1} = y_{t+1} - \rho_t y_t$  is stationary (i.e., with a stationary conditional variance process), and that the current value of  $\rho_t$  affects only the conditional mean of  $y_{t+1}$ . Autocorrelation filtration in the model is consequently closer to standard autocorrelation estimation, and becomes identical when  $\eta_{t+1}$  is i.i.d. Gaussian and the autocorrelation is constant ( $\sigma_\rho = 0$ ). Model 1 is also somewhat more convenient for estimation, in that it has a “semi-affine” structure that can be directly estimated using the methodology of Bates (2006).

In Model 2,  $\eta_{t+1}$  is the *permanent impact* of daily shocks to stock index excess returns, and is again assumed stationary. The model assumes that infrequent trading in the component stocks (proxied by  $\rho_t$ ) slows the incorporation of such shocks into the observed stock index, but that the index ultimately responds fully once all stocks have traded.<sup>5</sup> Unlike Model 1, Model 2 is consistent with LeBaron’s (1992) observation that annual estimates of daily stock market volatility and autocorrelation appear inversely related. Higher autocorrelations smooth shocks across periods, reducing observed market volatility. Furthermore, the model is more suitable for pricing risks; i.e., identifying the equity premium, or the (affine) risk-neutral process underlying option prices. The current value of  $\rho_t$  affects both the conditional mean and higher moments of  $y_{t+1}$ , resulting in a

---

<sup>4</sup>See, e.g., Andersen, Benzoni and Lund (2002, Table I), who estimate different autocorrelations for 1953-96 and 1980-96.

<sup>5</sup>Jukivuolle (1995) distinguishes between the “observed” and “true” stock index when trading is infrequent, and proposes using a standard Beveridge-Nelson decomposition to identify the latter. This paper differs in assuming that the parameters of the ARIMA process for the observed stock index are not constant.

significantly different filtration procedure for estimating  $\rho_t$  from past excess returns. The time series model is not semi-affine, but I develop below a transformation of variables that makes filtration and estimation as tractable as for Model 1.

Both models build upon previous time series and market microstructure research into stock market returns. For instance, the effective length  $\tau_t$  of a business day is allowed to vary based upon various periodic effects. In particular, day-of-the-week effects, weekends, and holidays are accommodated by estimated time dummies that allow day-specific variation in  $\tau_t$ . In addition, time dummies were estimated for the Saturday morning trading available over 1926-52, and for the Wednesday exchange holidays in the second half of 1968 that are the focus of French and Roll (1986).<sup>6</sup> Finally, the stock market closings during the ‘‘Bank Holiday’’ of March 3-15, 1933 and following the September 11, 2001 attacks were treated as  $\frac{12}{365}$ - and  $\frac{7}{365}$ -year returns, respectively. Treating the 1933 Bank Holiday as a 12-day interval substantially reduces the influence of its +15.5% return on parameter estimation. September 17, 2001 saw a smaller movement, of -4.7%.

For Model 1, the cumulant generating function of future returns and state variable realizations *conditional* upon current values is analytic, and of the semi-affine form

$$\begin{aligned} \ln F(\Phi, \xi, \psi | y_t, \rho_t, V_t) &\equiv \ln E \left[ e^{\Phi y_{t+1} + \xi \rho_{t+1} + \psi V_{t+1}} | y_t, \rho_t, V_t \right] \\ &= C(\tau_t; \Phi, \xi, \psi) + (\xi + \Phi y_t) \rho_t + D(\tau_t; \Phi, \psi) V_t \end{aligned} \quad (25)$$

where  $C(\tau; \xi, \Phi, \psi) = C(\tau; \Phi, \psi) + \frac{1}{2} \sigma_p^2 \xi^2$ , and  $C(\tau; \Phi, \psi)$  and  $D(\tau; \Phi, \psi)$  are given in (17) and (18) above. For model 2, the conditional cumulant generating function is of the non-affine form

$$\ln F(\Phi, \xi, \psi | y_t, \rho_t, V_t) = C(\tau_t; (1 - \rho_t) \Phi, \xi, \psi) + (\xi + \Phi y_t) \rho_t + D(\tau_t; (1 - \rho_t) \Phi, \psi) V_t, \quad (26)$$

given the shocks to  $y_{t+1}$  are scaled by  $(1 - \rho_t)$ .<sup>7</sup>

---

<sup>6</sup>Gallant, Rossi and Tauchen (1992) use a similar approach, and also estimate monthly seasonals.

<sup>7</sup>Dilip Madan informs me that practitioners distinguish between *time-scaled* and *space-scaled* models of time-varying volatility. GARCH models are typically space-scaled, whereas Model 1 is a time-scaled model of stochastic volatility. Model 2 contains both (stationary) time scaling via  $V_t$  and the time dummies, and (non-stationary) space scaling via  $1 - \rho_t$ .

#### I.4 Filtration and maximum likelihood estimation

If the state variables  $(\rho_t, V_t)$  were observed along with returns, it would in principle be possible to evaluate the joint transition densities of the data and the state variable evolution by Fourier inversion of the joint conditional characteristic function  $F(i\Phi, i\xi, i\psi | y_t, \rho_t, V_t)$ , and to use this in a maximum likelihood procedure to estimate the parameters of the stochastic process. However, since  $(\rho_t, V_t)$  are latent rather than directly observed, this is a *hidden Markov* model that must be estimated by other means.

For Model 1, the assumption that the cumulant generating function (25) is affine in the latent state variables  $(\rho_t, V_t)$  implies that the hidden Markov model can be filtered and estimated using the approximate maximum likelihood (AML) methodology of Bates (2006). The AML procedure is a filtration methodology that recursively updates the conditional characteristic functions of the latent variables and future data conditional upon the latest datum. Define  $\mathbf{Y}_t \equiv \{y_1, y_2, \dots, y_t\}$  as the data observed up through period  $t$ , and define

$$G_{t|t}(i\xi, i\psi) \equiv E\left[e^{i\xi\rho_t + i\psi V_t} | \mathbf{Y}_t\right] \quad (27)$$

as the joint conditional characteristic function that summarizes what is known at time  $t$  about  $(\rho_t, V_t)$ . The density of the observation  $y_{t+1}$  conditional upon  $\mathbf{Y}_t$  can be computed by Fourier inversion of its conditional characteristic function:

$$p(y_{t+1} | \mathbf{Y}_t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{t|t}[i\Phi y_t, D(\tau_t; i\Phi, 0)] e^{C(\tau_t; i\Phi, 0, 0) - i\Phi y_{t+1}} d\Phi. \quad (28)$$

Conversely, the joint conditional characteristic function  $G_{t+1|t+1}(i\xi, i\psi)$  needed for the next observation can be updated given  $y_{t+1}$  by the characteristic-function equivalent of Bayes' rule:

$$G_{t+1|t+1}(i\xi, i\psi) = \frac{1}{2\pi p(y_{t+1} | \mathbf{Y}_t)} \int_{-\infty}^{\infty} G_{t|t}[i\xi + i\Phi y_t, D(\tau_t; i\Phi, i\psi)] e^{C(\tau_t; i\Phi, i\xi, i\psi) - i\Phi y_{t+1}} d\Phi. \quad (29)$$

The algorithm begins with an initial joint characteristic function  $G_{1|1}(\cdot)$  and proceeds recursively through the entire data set, generating the log likelihood function  $\sum \ln p(y_{t+1} | \mathbf{Y}_t)$  used in maximum likelihood estimation. Filtered estimates of the latent variables can be computed from derivatives of the joint conditional moment generating function, as can higher conditional moments:



$$E[\rho_{t+1}^m V_{t+1}^n | \mathbf{Y}_{t+1}] = \left. \frac{\partial^{m+n} G_{t+1|t+1}(\xi, \psi)}{\partial \xi^m \partial \psi^n} \right|_{\xi = \psi = 0}. \quad (30)$$

The above procedure, if implementable, would permit exact maximum likelihood function estimation of parameters. However, the procedure would require storing and updating the entire function  $G_{t|t}(\cdot)$  based on point-by-point univariate numerical integrations. As such a procedure would be slow, the AML methodology instead approximates  $G_{t|t}(\cdot)$  at each point in time by a moment-matching joint characteristic function, and updates the approximation based upon updated estimates of the moments of the latent variables. Given an approximate prior  $\hat{G}_{t|t}(\cdot)$  and a datum  $y_{t+1}$ , (30) is used to compute the posterior moments of  $(\rho_{t+1}, V_{t+1})$ , which are then used to create an approximate  $\hat{G}_{t+1|t+1}(\cdot)$ . The overall procedure is analogous to the Kalman filtration procedure of updating conditional means and variances of latent variables based upon observed data, under the assumption that those variables and the data have a conditional normal distribution. However, the equations (29) and (30) identify the optimal *nonlinear* moment updating rules for a given prior  $G_{t|t}(\cdot)$ , whereas standard Kalman filtration uses linear rules. It will be shown below that this modification in filtration rules is important when estimating latent autocorrelations and variances under fat-tailed Lévy processes. Furthermore, Bates (2006) proves that the iterative AML filtration is numerically stable, and shows that it performs well in estimating parameters and latent variable realizations.

Autocorrelations can be negative or positive, while conditional variance must be positive. Consequently, different two-parameter distributions were used to summarize conditional distributions of the two latent variables: Gaussian for autocorrelations, gamma for variances. Furthermore, since volatility estimates mean-revert within months whereas autocorrelation estimates evolve over years, realizations of the two latent variables were assumed conditionally independent. These assumptions resulted in an approximate conditional characteristic function of the form

$$\ln \hat{G}_{t|t}(\xi, \psi) = [\hat{\rho}_{t|t} \xi + \frac{1}{2} W_{t|t} \xi^2] - v_t \ln(1 - \kappa_t \psi). \quad (31)$$

The following summarizes key features of joint conditional distributions of the latent variables.

	Autocorrelation $\rho_t$	spot variance $V_t$
Distribution	$\rho_t   \mathbf{Y}_t \sim N(\hat{\rho}_{t t}, \mathcal{W}_{t t})$	$V_t   \mathbf{Y}_t \sim (\hat{V}_{t t}, P_{t t})$
conditional cumulant generating function	$\ln E[e^{\xi \rho_t}   \mathbf{Y}_t] = \hat{\rho}_{t t} \xi + \frac{1}{2} \mathcal{W}_{t t} \xi^2$	$\ln E[e^{\psi V_t}   \mathbf{Y}_t] = -\mathbf{v}_t \ln(1 - \kappa_t \psi)$ $\kappa_t \mathbf{v}_t = \hat{V}_{t t}, \kappa_t^2 \mathbf{v}_t = \hat{P}_{t t}$
initial CGF	$\rho_1   \mathbf{Y}_t \sim N(0, 10^2)$	$\kappa_1 = \frac{\sigma^2}{2\beta}, \mathbf{v}_1 = \frac{2\alpha}{\sigma^2}$

$\rho_t, V_t | \mathbf{Y}_t$  assumed independent for all  $t$ .

Initial variance was assumed drawn from its unconditional gamma distribution, with the parameters  $(\kappa_1, \mathbf{v}_1)$  given above. Since autocorrelations were assumed nonstationary, no unconditional distribution exists. Consequently, the AML algorithm for Model 1 was initiated using a relatively diffuse conditional distribution for the initial autocorrelation – one much wider than the plausible  $(-1, +1)$  range.

The parameters  $\theta_t \equiv (\hat{\rho}_{t|t}, \mathcal{W}_{t|t}; \kappa_t, \mathbf{v}_t)$  – or, equivalently the moments  $(\hat{\rho}_{t|t}, \mathcal{W}_{t|t}; \hat{V}_{t|t}, P_{t|t})$  – summarize what is known about the latent variables. These were updated daily using the latest observation  $y_{t+1}$  and equations (29) - (30). For each day, 5 univariate integrations were required: 1 for the density evaluation in (29), and 4 for the mean and variance evaluations in (30). An upper  $\Phi_{\max}$  was computed for each integral which upper truncation error would be less than  $10^{-10}$  in magnitude. The integrands were then integrated over  $(-\Phi_{\max}, \Phi_{\max})$  to a relative accuracy of  $10^{-9}$ , using IMSL's adaptive Gauss-Legendre quadrature routine DQDAG and exploiting the fact that the integrands for negative  $\Phi$  are the complex conjugates of the integrands evaluated at positive  $\Phi$ . On average between 234 and 448 evaluations of the integrand were required for each integration.<sup>8</sup>

The non-affine specification  $y_{t+1} = \rho_t y_t + (1 - \rho_t) \eta_{t+1}$  in Model 2 necessitates additional restrictions upon the distribution of latent  $\rho_t$ . In particular, it is desirable that the scaling factor  $1 - \rho_t$  be nonnegative, so that the lower tail properties of  $\eta_{t+1}$  originating in the underlying Lévy

---

<sup>8</sup>The FFT approach used in Carr et al (2002) uses 16,384 functional evaluations.

specifications do not influence the *upper* tail properties of  $y_{t+1}$ . Consequently, the distribution of latent  $1 - \rho_t$  for Model 2 is modeled as inverse Gaussian – a 2-parameter unimodal distribution with conditional mean  $1 - \hat{\rho}_{t|t}$  and variance  $\mathcal{W}_{t|t}$ . Appendix A derives the resultant filtration procedure for this model, exploiting a useful change of variables procedure. The filtration is initiated at  $\rho_1 \sim (0, .5^2)$ , and it is again assumed that  $\rho_t$  and  $V_t$  are conditionally independent for all  $t$ .

## II. Properties of U.S. stock market returns, 1926 - 2006

### II.1 Data

There are two readily available value-weighted measures of the U.S. stock market: the CRSP value-weighted index, and the S&P Composite Index. This paper will primarily focus upon the former for time series analysis, but will also consider the latter when assessing options on S&P 500 futures. The CRSP data consist of 21,519 daily cum-dividend returns over January 2, 1926 through December 29, 2006. CRSP daily returns for each month were converted to daily log *excess* returns using Ibbotson and Associates' data on monthly Treasury bill returns, and the formula

$$y_t = \ln(1 + R_t) - \frac{n_t}{N} \ln(1 + i) \quad (32)$$

where  $R_t$  is the daily CRSP cum-dividend return;

$i$  is that month's return on Treasury bills of at least 1 month to maturity;

$N$  is the number of calendar days spanned by the monthly Treasury bill return; and

$n_t$  is the number of calendar days spanned by the "daily" return  $R_t$ .

The monthly interest rate data were downloaded from Ken French's Web site, and extended backwards through 1926 using data in Ibbotson and Associates' *SBBI Yearbook*.

The Schwert (1990) data set of daily U.S. stock market returns provides cum-dividend returns on the S&P Composite Index from January 4, 1928 onwards.<sup>9</sup> The S&P index was based upon 90 stocks until March 4, 1957, and 500 stocks thereafter. I updated Schwert's data through 2006 using Schwert's data methodology: ex-dividend daily S&P 500 returns from CRSP were augmented by an average daily dividend yield computed from monthly S&P 500 dividend yields

---

<sup>9</sup> Schwert also has daily data extending back to 1885, based on the (price-weighted) Dow Jones Industrial Average.

from Bloomberg. Cum-dividend returns were then converted into log excess returns using (32). Furthermore, CRSP value-weighted returns were used *instead* of the S&P 90 returns prior to March 5, 1957, for two reasons. First, that delivers data over 1926 and 1927, which is important for volatility assessment prior to the 1929 stock market crash. Second, the S&P Composite Index is only reported to two decimal places, which creates significant rounding error issues for the low S&P index values (around 5) observed in the 1930's.

## II.2 Parameter estimates

Table 2 describes and provides estimates of the time dummies from the most general time-changed CGMY model,<sup>10</sup> with Wednesday returns (Tuesday close to Wednesday close) arbitrarily selected as the benchmark day. Daily variance tended to be highest at the beginning of the week and decline thereafter, but day-of-the-week effects do not appear to be especially pronounced. The major exception is the Saturday morning (10 AM to noon) trading generally available over 1926-52.<sup>11</sup> Saturdays were effectively 43% as long as the typical Wednesday. Total weekend variance (Friday close to Monday close) was  $(.43 + 1.05) / 1.10 - 1 = 34.5\%$  higher when Saturday trading was available (over 1926-52) than when it was not (over 1945-2006).<sup>12</sup> This is qualitatively similar to but less pronounced than the doubling of weekend variance found by Barclay, Litzenberger and Warner (1990) in Japanese markets when Saturday half-day trading was feasible. Barclay et al lucidly discuss market microstructure explanations for the increase in variance.

Holidays generally did not have a strong impact on the effective length of a business day – with the exception of holiday weekends spanning 4 calendar days. Consistent with French and Roll (1986), 2-day returns spanning the Wednesday exchange holidays in 1968 (Tuesday close to Thursday close) had a variance not statistically different from a typical 1-day Wednesday return, but

---

<sup>10</sup>Estimates from other specifications were virtually identical, with estimates typically within  $\pm 0.01$  of the YY model's estimates.

<sup>11</sup>Saturday trading was standard before 1945. Over 1945-51, it was increasingly eliminated in summer months, and was permanently eliminated on June 1, 1952.

<sup>12</sup>As the time dummy estimates are estimated jointly with the volatility and autocorrelation filtrations, the estimates of weekend variances with versus without Saturday trading control for divergences in volatility and autocorrelation levels in the two samples.

substantially less than the  $1 + .94 = 1.94$  two-day variance observed for returns from Tuesday close to Thursday close in other years. Overall, the common practice of ignoring day-of-the-week effects, weekends, and holidays when analyzing the time series properties of daily stock market returns appears to be a reasonable approximation, provided the data exclude Saturday trading.

Tables 3A and 3B report estimates for various specifications listed in Table 1, while Figure 1 presents associated normal probability plots for model 2. (The plots for Model 1 were similar.) As noted above, all specifications capture the leverage effect by a correlation  $\rho_{sv}$  with the diffusion shock to conditional variance. The specifications diverge in their modeling of the Lévy shocks  $dL_t$ , *orthogonal* to the variance innovation. SV is the Heston model, while SVJ1 and SVJ2 have a diffusion for small asset return shocks, plus finite-activity normally-distributed jumps to capture outliers. The other models examine the generalized time-changed CGMY model, along with specific parameter restrictions or relaxations.

Most specifications using either Model 1 or Model 2 have similar estimates for the parameters determining the conditional mean and stochastic variance evolution. The evidence for a variance-sensitive equity premium ( $\mu_1 > 0$ ) is stronger for Model 2 specifications, but  $\mu_1$  is not typically significantly different from zero for either model. Latent permanent variance in Model 2 mean-reverts towards an estimated average level around  $(.172)^2$ , with a half-life about 1.6 months. The SV and LS models are the outliers, with different estimates of the equity premia and variance process from other specifications. As discussed below in section II.6, this reflects these two specifications' substantially different approach to variance filtration, given different assessments of tail risk.

The various specifications primarily diverge in how they capture tail risk. The Merton-based SVJ1 and SVJ2 results in Table 3B largely replicate the jump risk results in Bates (2006). The SVJ1 model has symmetric normally-distributed jumps with standard deviation 3 - 3.4% and time-varying jump intensities that occur on average  $\lambda_1(\alpha/\beta) = 3.3 - 3.7$  jumps per year. As shown in Figure 1, this jump risk assessment fails to capture the substantial 1987 crash. By contrast, the SVJ2 model adds a second jump component that directly captures the 1987 outlier. The resulting increase in log likelihood from is statistically significant under a likelihood ratio test, with a marginal significance

level around 3% for Models 1 and 2.

The various CGMY models primarily diverge across the specification of the  $Y_p, Y_n$  parameters – whether they are set to specific levels, and whether they diverge for the intensities of positive versus negative jumps. The DEXP model with  $Y_p = Y_n = -1$  is conceptually similar to the jump-diffusion model SVJ1, but uses instead a finite-activity double exponential distribution for jumps. Despite the fatter-tailed specification, Figure 1 indicates the DEXP model has difficulties comparable to SVJ1 in capturing the 1987 crash. The VG model replaces the finite-activity double exponential distribution with the infinite-activity variance process ( $Y_p = Y_n = 0$ ), and does slightly better in fit. Both models include a diffusion component, which captures 73-74% of the variance of the orthogonal Lévy shock  $dL_t$ .

Specifications Y, YY, and LS involve pure-jump processes for the orthogonal Lévy process  $L_t$ , *without* a diffusion component. Overall, higher values of  $Y$  fit the data better – especially the 1987 crash, which ceases to be an outlier under these specifications. Relaxing the restriction  $Y_p = Y_n$  leads to some improvement in fit, with the increase in log likelihood (YY versus Y) having P-values of 1.8% and 0.8% for Models 1 and 2, respectively. Point estimates of the jump parameters ( $w_n, G, Y_n$ ) governing downward jump intensities diverge sharply from the parameters ( $1 - w_n, M, Y_p$ ) governing upward jump intensities when the  $Y_p = Y_n$  restriction is relaxed, although standard errors are large. The dampening coefficient  $G$  is not significantly different from zero, implying one cannot reject the hypothesis that the downward-jump intensity is from a stochastic-intensity version of the Carr-Wu (2003) log-stable process. By contrast, the upward intensity is estimated as a finite-activity jump process – which, however, still overestimates the frequency of big positive outliers (Figure 1, sixth panel).

Motivated by option pricing issues, Carr and Wu (2003) advocate using a log-stable distribution with purely downward jumps. An approximation to this model generated by setting  $G = .001$  and  $w_n = 1$  fits stock market returns very badly. The basic problem is that while the LS model does allow positive asset returns, it severely underestimates the frequency of large positive returns. This leads to a bad fit for the upper tail (Figure 1, last panel). However, the YY estimates indicate that the Carr-Wu specification can be a useful *component* of a model, provided the upward

jump intensity function is modeled separately.

Unrestricted CGMY models generate at least one  $Y$  parameter in the infinite-activity, infinite-variation range [1, 2], and typically near the diffusion value of 2. This suggests that the models may be trying to capture considerable near-zero activity. However, adding an additional diffusion component to the time-changed YY Lévy specification to capture that activity separately (specification YY\_D) led to no significant improvement in fit.

Overall, Figure 1 suggests the differences across the alternate fat-tailed specifications are relatively minor, and fit the data similarly over most of the data range ( $|z| < 3$ ). The models SV, SVJ1, DEXP, VG, and LS appear less desirable, given their failure to capture the largest outliers. The SVJ2, Y, and YY specifications appear to fit about the same. All models appear to have a small amount of specification error (deviations from linearity) in the  $z \in [-3.5, -2]$  range and in the upper tail ( $z > 3$ ).

### II.3 Unconditional distributions

A further diagnostic of model specification is the models' ability or inability to match the unconditional distribution of returns – in particular, the tail properties of unconditional distributions. Mandelbrot (1963) and Mandelbrot and Hudson (2004) argue that empirical tails satisfy a “power law:” tail probabilities plotted against absolute returns approach a straight line when plotted on a log-log graph. This empirical regularity underlies Mandelbrot's advocacy of the stable Paretian distribution, which possesses this property and is nested within the CGMY model for  $G = M = 0$ .

Mandelbrot's argument is premised upon i.i.d. returns, but the argument can be extended to time-changed Lévy processes. Conditional Lévy densities time-average; if the conditional intensity of moves of size  $x$  is  $(1 - \rho_{sv}^2) V_t k(x)$ , the *unconditional* frequency of moves of size  $x$  is  $(1 - \rho_{sv}^2) E(V_t) k(x)$ . Since unconditional probability density functions asymptotically approach the unconditional Lévy densities for large  $|x|$ , while unconditional tail probabilities approach the corresponding integrals of the unconditional Lévy densities, examining unconditional distributions may still be useful.

Figure 2a provides estimates of unconditional probability density functions of stock market excess return residuals  $\eta_{t+1} = y_{t+1} - \rho_t y_t$  for various specifications under Model 1, as well as data-based estimates from a histogram of filtered residuals  $\hat{\eta}_{t+1} = y_{t+1} - \hat{\rho}_{t|t} y_t$ . Given the day-of-the-week effects reported in Table 2, the unconditional density functions are a horizon-dependent mixture of densities, with mixing weights set equal to the empirical frequencies. (The two shocks spanning the market closings in 1933 and 2001 were omitted.) The substantial impact of the 1987 crash outlier upon parameter estimates is apparent. The SVJ2 estimates treat that observation as a unique outlier, while the CGMY class of models progressively fatten the lower tail as greater flexibility is permitted for the lower tail parameter  $Y_n$ . As noted above, the lower tail approaches the Carr-Wu (2003) log-stable (LS) estimate. However, the LS model is unable to capture the frequency of large positive outliers, and behaves similarly to the SV model in the upper tail. All models closely match the empirical unconditional density function in the  $\pm 3\%$  range where most observations occur; and all models underestimate the unconditional frequency of moves of 3% - 7% in magnitude.

Figure 2b provides similar estimates for unconditional lower and upper tail probabilities. In addition, 1000 sample paths of stock market excess return residuals over 1926-2006 were simulated via a Monte Carlo procedure using YY parameter estimates, in order to provide confidence intervals on tail probability estimates.<sup>13</sup> Unsurprisingly, the confidence intervals on extreme tail events are quite wide. The underestimation of moves of 3% - 7% in magnitude is again apparent, and is statistically significant. This rejection of the YY model does not appear attributable to misspecification of the Lévy density  $k(x)$ , which in Figure 1 captures conditional densities quite well. Rather, the poor unconditional fit in Figures 2a and 2b appears due to misspecification of volatility dynamics. Half of the 3-7% moves occurred over 1929 - 1935 – a prolonged high-volatility period that simulated volatility realizations from the 1-factor variance process of equation (13) generally do not match.

---

<sup>13</sup>Conditional variance sample paths were simulated using the approach of Bates (2006, Appendix A.6), while Lévy shocks  $\eta_{t+1}$  conditional upon intradaily average variance and data-based daily time horizons were generated via an inverse CDF methodology, with CDF's computed by Fourier inversion. The two shocks spanning the market closings in 1933 and 2001 were omitted.



Figure 3 plots model-specific tail probability estimates for the YY model on the log-log scales advocated by Mandelbrot, along with data-specific quantiles for 20,004 stock market residuals that have roughly a 1-day estimated time horizon ( $\pm 11\%$ ). The lower tail probability does indeed converge to the unconditional tail intensity

$$K(y) \equiv \int_{-\infty}^y k(x) dx = C_n G^{Y_n} \Gamma(-Y_n, G|y|) \quad (36)$$

where  $C_n = w_n(1 - \rho_{sv}^2) \tau \alpha / [\beta \Gamma(2 - Y_n) G^{Y_n - 2}]$  and  $\Gamma(a, z)$  is the incomplete gamma function. Furthermore, given  $G$  estimates near 0,  $K(y) \approx C_n y^{-Y_n}/Y_n$  is roughly a power function in  $y$ , implying near linearity when plotted on a log-log scale.

However, the graph indicates that the convergence of tail probabilities to the tail intensity  $K(y)$  occurs only for observations in excess of 5% in magnitude – roughly 5 standard deviations. As this is outside the range of almost all data, it does not appear that log-log scales provide a useful diagnostic of model specification and tail properties for daily data. This is partly due to stochastic volatility, which significantly slows the asymptotic convergence of unconditional tail probabilities to  $K(y)$  for large  $|y|$ . Absent stochastic volatility ( $\sigma = 0$ ), the tail probabilities of an i.i.d. YY Lévy process converge to  $K(y)$  for observations roughly in excess of 3% in magnitude (3 standard deviations).

No power law properties are observed for upper tail probabilities, given substantial estimated exponential dampening. The failure of both lower and upper unconditional tail probabilities to capture the frequency of moves of 3-7% in magnitude is again apparent, and statistically significant.

#### II.4 Subsample estimates

Table 4 provides estimates for data subsamples, as a test of the stability of the time series process. The mean, stochastic volatility and jump parameters were allowed to differ before and after March 5, 1957.<sup>14</sup> The time dummies (similar to those in Table 2) that capture day-of-the-week effects were

---

<sup>14</sup>The data split was chosen so that the second subsample's estimates could be compared with estimates from S&P 500 returns, as well as with other studies that use data starting in the 1950's (Andersen et al (2002), Bates (2006), Chernov et al (2003)).

kept common across subsamples; but some of those dummies also capture subsample-specific phenomena (Saturday trading before 1953; exchange holidays in 1968). The estimation and filtration over the two subsamples nest the full-sample estimates of Table 3, so that standard likelihood ratio tests can be used to test whether the divergence in subsample parameter estimates are statistically significant.

Parameter estimates diverge strongly across subsamples, with P-values less than  $10^{-16}$ , but in different fashions for the SVJ1 and YY models. For the SVJ1 model, the major divergence was clearly in the estimated volatility process. The 1926-57 period includes the highly volatile 1930's, yielding an overall average variance of  $(.202)^2$  over 1926-57, as opposed to  $(.149)^2$  over 1957-2006. The volatility dynamics also diverge, with volatility more volatile and with faster mean reversion over 1926-57 than over 1957-2006. Jump risk estimates also diverge, with more frequent but smaller jumps in the first half than in the second half. Progressively relaxing full-sample constraints on parameter categories (mean;  $\sigma_p$ ; stochastic volatility parameters; jump parameters) indicates that between 71% and 86% of the subsample improvement in log likelihood comes from using subsample stochastic volatility parameters. Between 8% and 22% of the change in log likelihood comes from using subsample jump parameters, depending on whether stochastic volatility or jump parameters are relaxed first.

The 1957-2006 subsample estimates for the YY model are more heavily affected by the 1987 crash than are the full-sample estimates. The parameter  $G$  approaches its lower bound of zero, implying that the lower tail density is approaching the time-changed version of the infinite-variance log-stable distribution. Correspondingly, the subsample estimate of unconditional variance  $\alpha/\beta = .365^2$  becomes substantially meaningless, and cannot be compared with estimates from other models or other periods. By contrast, the estimates over 1926-57 are strictly finite-variance. Given strong interactions between stochastic volatility and jump parameters, it is not clear which is more responsible for the strong rejections of parameter stability across subsamples.

## II.5 Autocorrelation filtration

Given that the prior distribution  $\rho_t | \mathbf{Y}_t$  is assumed  $N(\hat{\rho}_{t|t}, \mathbb{W}_{t|t})$ , it can be shown that the autocorrelation filtration algorithm (30) for Model 1 updates conditional moments via the robust

Kalman filtration approach of Masreliez (1975):

$$\hat{\rho}_{t+1|t+1} = \hat{\rho}_{t|t} - y_t W_{t|t} \frac{\partial \ln p(y_{t+1} | \mathbf{Y}_t)}{\partial y_{t+1}} \quad (33)$$

$$W_{t+1|t+1} = \sigma_\rho^2 + (y_t W_{t|t})^2 \frac{\partial^2 \ln p(y_{t+1} | \mathbf{Y}_t)}{\partial y_{t+1}^2} \quad (34)$$

If  $y_{t+1} | \mathbf{Y}_t$  were conditionally normal, the log density would be quadratic in  $y_{t+1}$ , and (33) would be the linear updating of standard Kalman filtration. More generally, the conditionally fat-tailed properties of  $y_{t+1}$  are explicitly recognized in the filtration.<sup>15</sup> The partial derivatives of log densities can be computed numerically by Fourier inversion.

Figure 4 illustrates the autocorrelation filtrations estimated under various models. For model 1, the autocorrelation revision is fairly similar to standard Kalman filtration for observations within a  $\pm 2\%$  range – which captures most observations, given a unconditional daily standard deviation around 1%. However, the optimal filtration for fat-tailed distributions is to *downweight* the information from returns larger than 2% in magnitude. The exceptions are the stochastic volatility (SV) and Carr-Wu log-stable (LS) specifications. Those specifications do not particularly downweight outliers occurring in non-fat tails: in both tails for SV, in the upper tail for LS.

The autocorrelation filtration under Model 2 is different. Since  $y_{t+1} = \rho_t + (1 - \rho_t)\eta_{t+1}$  in that model, large observations of  $y_{t+1}$  are attributable either to large values of  $1 - \rho_t$  (small values of  $\rho_t$ ), or to large values of the Lévy shocks captured by  $\eta_{t+1}$ . The resulting filtration illustrated in the lower panels of Figure 2 is consequently sensitive to medium-size movements in a fashion substantially different from the Model 1 specifications.

Figure 5 presents filtered estimates of the daily autocorrelation from the YY model. The most striking result is the extraordinarily pronounced increase in autocorrelation estimates from 1941 - 1971, with a peak of 35% reached in June 1971. Estimates from other models give

---

<sup>15</sup>See Schick and Mitter (1994) for a literature review of robust Kalman filtration.

comparable results, as do crude sample autocorrelation estimates using a 1- or 2-year moving window.<sup>16</sup> After 1971, autocorrelation estimates fell steadily, and became insignificantly different from zero after 2002. This broad pattern is observed both for Models 1 and 2, although the precise estimates diverge given the different filtration methodologies. Filtered autocorrelation estimates appear inversely related to measures of annual stock turnover computed by French (2008), attaining values closer to zero in the high-turnover periods before 1933 and after 1982. This is consistent with the standard stale-price explanation of autocorrelation in stock index returns.

Figure 5 also indicates that the estimates of daily autocorrelation are virtually nonstationary, indicating that fitting ARMA processes with time-invariant parameters to stock market excess returns is fundamentally pointless. The conditional standard deviation asymptotes at about 4½%, implying a 95% confidence interval of ±9% for the autocorrelation estimates.

## II.6 Volatility filtration

When returns follow an autocorrelated process with i.i.d. shocks of the form

$$y_{t+1} = \rho y_t + u_{t+1}, \quad u_{t+1} \sim (\bar{u}, \sigma_u^2) \quad (35)$$

there are various ways of measuring variance:

$$\text{Conditional (or residual) variance: } \text{Var}[y_{t+1} | y_t] = \sigma_u^2$$

$$\text{Unconditional variance of returns: } \text{Var}[y_{t+1}] = \text{Var}\left[\frac{u_{t+1}}{1 - \rho L}\right] = \frac{\sigma_u^2}{1 - \rho^2}$$

$$\text{Conditional permanent variance: } \text{Var}\left[\sum_{\tau=1}^{\infty} y_{t+\tau} | y_t\right] = \text{Var}\left[\frac{u_{t+1}}{1 - \rho}\right] = \frac{\sigma_u^2}{(1 - \rho)^2}$$

where  $L$  is the lag operator. These measure of variance are also approximately relevant in the above models with stochastic conditional volatility and slow-moving autocorrelation. The  $\alpha / \beta$  values in Table 3A are estimates of the average level of *residual* variance for model 1, but estimates of average *permanent* variance for model 2; hence the higher estimates for the latter. Furthermore, the

---

<sup>16</sup>See LeBaron (1992, Figure 1) for annual estimates of the daily autocorrelation of S&P composite index returns over 1928-1990.

ratio of return variance to permanent variance is  $(1 - \rho)^2 / (1 - \rho^2) = (1 - \rho) / (1 + \rho)$ , which is less than 1 for  $\rho > 0$  and is monotonically decreasing in  $\rho$ . If permanent conditional variance is stationary and autocorrelation evolves independently of permanent variance, as is assumed in model 2, periods of high autocorrelation will generate periods of low observed variance of returns – a property consistent with the inverse relationship between annual estimates of daily autocorrelation and volatility over 1928-1990 reported in LeBaron (1992).

The left panel of Figure 6 illustrates how the estimated conditional volatility  $E_{t+1} \sqrt{V_{t+1}}$  is updated for the various specifications under model 1. The conditional volatility revisions use median parameter values  $(\kappa_t, \nu_t) = (.00294, 5.85)$  for the prior gamma distribution of  $V_t$ , implying a conditional mean  $\kappa_t \nu_t = (.131)^2$  that is close to the  $(.129)^2$  median value observed for  $\hat{V}_{t|t}$  estimates from the YY model.<sup>17</sup> For comparability with GARCH analyses such as Hentschel (1995), Figure 4 shows the “news impact curve,” or revision in conditional volatility estimates upon observing a given excess return, using the methodology of Bates (2006, pp.931-2).

All news impact curves are tilted, with negative returns having a larger impact on volatility assessments than positive returns. This reflects the leverage effect, or estimated negative correlation between asset returns and volatility shocks. All specifications process the information in small-magnitude asset returns similarly. Furthermore, almost all specifications truncate the information from returns larger than 3 standard deviations. This was also found in Bates (2006, Figure 1) for the SVJ1 model, indicating such truncation appears to be generally optimal for arbitrary fat-tailed Lévy processes. The SV and LS exceptions support this rule. The LS model has a fat lower tail but not an especially fat upper tail, and truncates the volatility impact of large negative returns but not of large positive returns. The fact that volatility revisions are not monotonic in the magnitude of asset returns is perhaps the greatest divergence of these models from GARCH models, which almost invariably specify a monotonic relationship.<sup>18</sup> However, since moves in excess of  $\pm 3$  standard

---

<sup>17</sup>As  $\hat{V}_{t|t}$  estimates have substantial positive skewness, the median is substantially below the mean estimate of  $(.159)^2$  reported in Table 3A.

<sup>18</sup>An exception is Maheu and McCurdy (2004), who put a jump filter sensitive to outliers into a GARCH model. They find that the sensitivity of variance updating to the latest squared return should be reduced for outliers, for both stock and stock index returns.

deviations are rare, all specifications will generate similar volatility estimates most of the time. The volatility filtrations for model 2 shown in the right panel of Figure 6 for median parameters  $(\kappa_t, \nu_t) = (.00385, 6.01)$  are qualitatively similar to those for model 1.

Figure 7 presents the filtered estimates of conditional annualized permanent volatility over 1926-2006 from the YY model 2, as well as the associated conditional standard deviation.<sup>19</sup> Volatility estimates from other models (*except* SV and LS) are similar – as, indeed, is to be expected from the similar volatility updating rules in Figure 4. The conditional standard deviation is about 2.8%, indicating a 95% confidence interval of roughly  $\pm 5\frac{1}{2}\%$  in the annualized volatility estimates. Because of the 81-year time scale, the graph actually shows the longer-term volatility dynamics *not* captured by the model, as opposed to the intra-year volatility mean reversion with 2-month half-life that *is* captured by the model. Most striking is, of course, the turbulent market conditions of the 1930's, unmatched by any comparable volatility in the post-1945 era. The graph indicates the 1-factor stochastic variance model is too simple, and suggests that multifactor specifications of variance evolution are worth exploring.<sup>20</sup>

The inset to Figure 5 compares adjusted filtered volatility estimates  $.8655 E_t \sqrt{V_t \tau_t}$  over 1987-89 with realized volatility estimates computed daily from intradaily 15-minute log-differenced S&P 500 futures prices. (Open-to-close futures returns were 86.55% as volatile as close-to-close futures returns over 1982-2001.) The inset shows that the AML filtration methodology using daily data generally tracks realized intradaily volatility quite closely. The filtered estimates do not capture major realized volatility spikes – especially over October 19-28, 1987. The models estimated in this paper interpret such spikes and the accompanying large daily stock market movements as stock market jumps. However, the clustering of high intradaily volatility values is important time series

---

<sup>19</sup>“Annualized” volatility refers to the choice of units. Since time is measured in years,  $V_t$  is variance per year, while the daily volatility estimate for a Wednesday return with an estimated length of  $1/260$  years (from Table 2) is approximately  $E_t \sqrt{V_t/260}$ . Since variance mean-reverts with an estimated half-life of roughly 2 months, it is not appropriate to interpret Figure 5 as showing the volatility estimate for a 1-year investment horizon.

<sup>20</sup>The inadequacies of AR(1) representations of conditional variance are already well-known in volatility research, and have motivated research into long-memory processes.

evidence *against* the models' diffusive-volatility assumption, and supports an alternate volatility-jump specification.

Filtered volatility estimates do appear sensitive to the data interval used in estimation, via the underlying parameter estimates. For instance, the subsample SVJ1 estimates in Table 4 yield filtered annualized  $\sqrt{\hat{V}_{t|t}}$ 's that are 1.86% higher on average over 1926-57 than the full-sample estimates, and 1.29% lower over 1957-2006. A significant underlying factor is the estimate of unconditional volatility  $\sqrt{\alpha/\beta}$  in Table 4, which is higher in the first than in the second subsample, and which significantly influences volatility filtration. A similar influence of unconditional variance estimates upon conditional volatility estimates is observed in GARCH models.<sup>21</sup>

### III. Option pricing implications

Do these alternative models imply different option prices? Exploring this issue requires identifying the appropriate pricing of equity, jump, and stochastic volatility risks. Furthermore, the presence of substantial and stochastic autocorrelation raises issues not previously considered when pricing options. In particular, the observed stock index level underlying option prices can be stale, while the relevant volatility measure over the option's lifetime is also affected. The variance of the sum of future stock market returns is not the sum of the variances when returns are autocorrelated.

To examine these issues, I will focus upon Model 2, with its interpretation in equations (23) - (24) of  $ds_t$  as the permanent shock to the log stock market level. Furthermore, I will use the myopic power utility pricing kernel of Bates (2006) to price the various risks:

$$d\ln M = \mu_m^* dt - R ds_t. \quad (36)$$

This pricing kernel constrains both the equity premium estimated under the objective time series model, and the transformations of those estimates into the risk-neutral process appropriate for

---

<sup>21</sup>See, e.g., Andersen et al (2005, p.7), who note that GARCH(1,1) models diverge from the RiskMetrics approach in taking into account mean reversion of conditional variance towards the unconditional variance. This implies that GARCH conditional volatility estimates are affected by sample-specific estimates of the unconditional variance.

pricing options. In particular, the instantaneous equity premium is

$$(\mu_0 + \mu_1 V_t) dt = -E_t \left[ \left( e^{ds_t} - 1 \right) \left( e^{-Rds_t} - 1 \right) \right] \quad (37)$$

which implies

$$\begin{aligned} \mu_0 &= 0 \\ \mu_1 V_t &= [R\rho_{sv}^2 + R(1 - \rho_{sv}^2)(1 - f_{jump})] V_t - \int_{-\infty}^{\infty} (e^x - 1)(e^{-Rx} - 1)k(x) dx \\ &\approx R V_t \end{aligned} \quad (38)$$

where  $1 - f_{jump}$  is the fraction of variance attributable to an orthogonal diffusion term. The approximation follows from first-order Taylor expansions, and from the fact that jumps account for a fraction  $f_{jump}(1 - \rho_{sv}^2)$  of overall variance  $V_t$ . The equity premium (37) is well-defined for the SVJ1 and SVJ2 models. For the CGMY models, the restriction  $G \geq R$  is required for a finite equity premium; the intensity of downward jumps must fall off faster than an investor's risk aversion to such jumps. The log-stable process is inconsistent with a finite equity premium for  $R > 0$ .<sup>22</sup>

The change of measure from objective to risk-neutral jump intensities takes the form

$$k^*(x) = k(x)e^{-Rx} \quad (39)$$

under a myopic power utility pricing kernel. This has assorted implications for parameter transformations that depend upon the precise specification of the Lévy density  $k(x)$ . For the SVJ models, as discussed in Bates (2006), this modified jump intensity shifts the mean jump size  $\bar{\gamma}_i$  by an amount  $-R\delta_i^2$ , while leaving the jump standard deviation  $\delta_i$  unchanged. For the CGMY model, the risk adjustment replaces the downward and upward exponential dampening parameters  $G$  and  $M$  by  $G - R$  and  $M + R$ , respectively, while leaving the  $C$  and  $Y$  parameters unchanged.<sup>23</sup> These risk adjustments alter the  $g_{dL}(\Phi)$  and  $h(\Phi)$  functions in equation (14). Table 5 summarizes the parameter transformations under the various models.

---

<sup>22</sup>Carr and Wu (2003) specify a log-stable process for the *risk-neutral* process underlying option prices. This can be generated from a CGMY process for the actual process with only downward jumps, and with  $G = R$ .

<sup>23</sup>Wu (2006) discusses this transformation.



The key risk aversion parameter  $R$  used for change of probability measure was estimated by imposing the equity premium restrictions (37), and re-estimating all times series models. The additional parameter restriction  $G \geq R$  was also imposed upon all CGMY models, and was binding for the YY model.<sup>24</sup> Parameter estimates reported in Table 6 changed little relative to those in Table 3B, while risk aversion was estimated at roughly 2.5 for all models. Furthermore, the restriction of a purely variance-sensitive equity premium ( $\mu_0 = 0$ ) was not rejected for any model.

I address the potential impact of autocorrelations upon option prices by examining prices of options on S&P 500 *futures*. I assume that stock index futures prices respond instantaneously and fully to the arrival of news, whereas lack of trading in the underlying stocks delays the incorporation of that news into the reported S&P 500 stock index levels. Furthermore, I assume that index arbitrageurs effectively eliminate any stale prices in the component stocks on days when futures contracts expire, so that stale stock prices do not affect the cash settlement of stock index futures. MacKinlay and Ramaswamy (1988) provide evidence supportive of both assumptions.

These assumptions have the following implications under Model 2 (equations 23 and 24):

1. the observed futures price  $F_t$  underlying options on S&P 500 futures is not stale;
2. log futures price innovations equal the intradaily innovations  $ds_t$  of equation (13):

$$d\ln F_t = ds_t. \quad (40)$$

Consequently, European options on stock index futures can be priced directly using a risk-neutral version of (40) – which is affine, simplifying option evaluation considerably. Furthermore, option prices do not depend upon  $\rho_t$ , except indirectly through the impact of autocorrelation filtration upon the filtration of latent permanent variance  $V_t$ . Following Bates (2006), European call prices on an S&P 500 futures contract can be priced as

$$\begin{aligned} c(F_t, T_t, \tau_t; X | \mathbf{Y}_t) &= E[c(F_t, V_t, T_t, \tau_t; X) | \mathbf{Y}_t, F_t] \\ &= e^{-r_t T_t} F_t - e^{-r_t T_t} X \left[ \frac{1}{2} + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{C^*(i\Phi, 0; \tau_t) + g_{t|t}^Y [D^*(i\Phi, 0; \tau_t)] - i\Phi \ln(X/F_t)}}{i\Phi(1 - i\Phi)} d\Phi \right], \end{aligned} \quad (41)$$

---

<sup>24</sup>Wu (2006) proposes an alternate pricing kernel with negative risk aversion for downside risk, thereby automatically imposing  $G \geq R$ .

where  $C^*(\cdot)$  and  $D^*(\cdot)$  are risk-neutral variants<sup>25</sup> of those in equations (17)-(21);

$\tau_t$  is the effective maturity of the option, given individual days' length from Table 2;

$T_t = n_t/365$  is the maturity associated with the continuously compounded Treasury bill yield

$r_t$ , given  $n_t$  calendar days until option maturity; and

$g_{t|t}^V[\Psi] = -v_t \ln(1 - \kappa_t \Psi)$  is the filtered cumulant generating function of  $V_t$  that summarizes what is known about  $V_t$  given past data  $\mathbf{Y}_t$ .

The associated implicit standard deviations (ISD's) for standardized maturity  $T_t$  can then be computed using the Black (1976) formula for European options on futures.

The ISD's from the various models are graphed in Figure 8, and are compared with observed ISD's computed from settlement prices for American options on S&P 500 futures with non-zero trading volume on December 29, 2006. Figure 8 also shows 95% confidence intervals, computed as in Bates (2006) based on parameter uncertainty alone, and on parameter and state uncertainty associated with  $V_t$  estimation. All models, including the SV model, generate virtually identical option prices at end-2006 over a range of  $\pm 2$  standard deviations – a range that contains the most actively traded options. The estimated level of the ATM ISD was roughly the same across all specifications, reflecting an absence of recent major outliers that would induce divergences in estimated volatility from different specifications. The tilt of the volatility smirk for near-the-money options appears to be driven primarily by the “leverage effect,” or correlation between shocks to variance and stock market returns. Only for deep out-of-the-money put options do the divergences in estimated tail properties generate substantially different ISD patterns across models.<sup>26</sup> Furthermore, those divergences across models generally decrease at longer maturities, as the impact of jumps falls in importance relative to the projected dynamics of stochastic volatility. The maturity profile of the YY model is the exception, reflecting the fact that the risk-neutral distribution is almost infinite-variance.

---

<sup>25</sup>The parameters  $(\mu_0, \mu_1, \alpha, \beta)$  are replaced by  $(0, 0, \alpha, \beta^*)$ , while the function  $g_{\mathcal{A}}(\Phi)$  is replaced by a risk-adjusted function  $g_{\mathcal{A}}^*(\Phi)$  given in Table 5 for specific models that captures the risk-adjusted jump intensities from (39).

<sup>26</sup>These results differ from Bates (2000, Table 3), who reports substantial divergences between the SV and SVJ models based substantially upon implicit parameter estimation. It would appear that implicit parameter estimates are strongly affected by the prices of deep OTM options.

Figure 9 chronicles estimated and observed at-the-money ISD's over 1983-2006 for the short-term options with maturities of 14 days or more. The overall evolution is broadly comparable to the estimates in Bates (2006, Figure 7). However, the ISD estimates are substantially closer to observed ISD's, with average values over 1988-2006 of 15.8% and 16.8%, respectively. The results differ from the larger divergences in Bates (2006), for two reasons. First, the earlier study used Anderson, Benzoni and Lund's (2002) data, who prefiltered S&P returns over 1953-1996 by estimating an MA(1), and then rescaled estimated residuals to match the mean and variance of the original data. Prefiltration removes the autocorrelation structure of the data, and consequently underestimates the average level  $\alpha/\beta$  of permanent variance relevant for pricing options. Re-estimating SVJ2 Model 2 on the raw S&P returns underlying the Anderson et al data over 1953-96 raises  $\alpha/\beta$  estimates from  $(.127)^2$  to  $(.143)^2$ .

Second, as shown above in Table 6,  $\alpha/\beta$  estimates are substantially higher when data from the volatile 1930's are included:  $(.174)^2$  over 1926-2006 for the SVJ2 model, as opposed to  $(.149)^2$  over 1957-2006. A higher  $\alpha/\beta$  estimate affects ISD estimates at all maturities, through its impact on the filtration algorithm for estimating spot variance  $V_t$  as well as through its impact on forecasts of future variance. For instance, some resulting filtered measures of risk-neutral volatility over 1988-2006 using the two sets of SVJ2 parameter estimates are:

Estimation	$\sqrt{\alpha/\beta^*}$	$\sqrt{\hat{V}_{t t}^*}$	Average values over 1988-2006		
			1-month $ISD_{ATM}$	1-month $ISD_{ATM}^{options}$	difference in ISD's
1926-2006	18.7%	15.9%	15.8%	16.8%	<b>1.0%</b>
1957-2006	16.3%	14.4%	14.4%	16.8%	<b>2.4%</b>
difference	2.4%	1.5%	1.4%		

Nevertheless, the broad assessment of previous studies appears unchanged. Observed ISD's from options on index futures do appear higher on average over the post-1987 period than is justified by risk-adjusted valuations based upon time series analysis, even when volatility assessments include data from the 1930's.

#### IV. Summary and Conclusions

This paper provides time series estimates of the time-changed CGMY (2003) Lévy process, and compares them to the time-changed finite-activity jump-diffusions previously considered by Bates (2006). Overall, while it is important to use adequately fat-tailed distributions when filtering volatility and other latent variables, it does not seem especially important *which* fat-tailed distribution one uses. Estimates of the volatility process and realizations are substantially unchanged across most specifications, while the option pricing implications are virtually identical for all but the deepest out-of-the-money options. The exceptions are Heston's (1993) stochastic volatility model, which underestimates upper and lower tail risk, and the log-stable model of Carr and Wu (2003), which underestimates upper tail risk. This underestimation of tail risk makes volatility estimates excessively sensitive to outliers, and also affects estimates of the volatility process. *Conditional* upon similar volatility estimates, however, even the Heston model fits option prices similarly to the fat-tailed distributions for all but deep OTM options. For these stochastic volatility/stochastic intensity models, the tilt of the volatility smirk for near-the-money options ( $\pm 2$  standard deviations) appears primarily driven by the "leverage" effect.

The paper also documents some structural shifts over time in the data generating process. Most striking is the apparently nonstationary evolution of the first-order autocorrelation of daily stock market returns, which rose from near-zero in the 1930's to 35% in 1971, before drifting down again to near-zero values after 1987. Autocorrelation estimates are inversely related to stock turnover, and are of considerable importance when assessing stock market volatility. The paper develops methods of dealing with time-varying autocorrelation, by treating it as an additional latent state variable to be filtered from observed data. Longer-term trends in volatility are also apparent in the filtered estimates, suggesting a need for multifactor models of conditional variance.

### References

- Andersen, Torben G., Luca Benzoni, and Jesper Lund (2002). "An Empirical Investigation of Continuous-Time Equity Return Models." *Journal of Finance* **57**, 1239-1284.
- Andersen, Torben G., Tim Bollerslev, Peter F. Christoffersen, and Francis X. Diebold (2005). "Practical Volatility and Correlation Modeling for Financial Market Risk Management." NBER working paper 11069, January.
- Bakshi, Gurdip and Dilip B. Madan (2000). "Spanning and Derivative-Security Valuation." *Journal of Financial Economics* **55**, 205-238.
- Barclay, Michael J., Robert H. Litzenberger, and Jerold B. Warner (1990). "Private Information, Trading Volume, and Stock-Return Variances." *Review of Financial Studies* **3**, 233-254.
- Bates, David S. (2000). "Post-'87 Crash Fears in the S&P 500 Futures Option Market." *Journal of Econometrics* **94**, 181-238.
- Bates, David S. (2006). "Maximum Likelihood Estimation of Latent Affine Processes." *Review of Financial Studies* **19**, 909-965.
- Bertoin, Jean (1996). *Lévy Processes*, Cambridge: Cambridge University Press.
- Black, Fischer (1976). "The Pricing of Commodity Contracts." *Journal of Financial Economics* **3**, 167-179.
- Broadie, Mark N., Mikhail Chernov, and Michael Johannes (2006). "Understanding Index Option Returns," Columbia University working paper.
- Carr, Peter, Hélyette Geman, Dilip B. Madan, and Marc Yor (2002). "The Fine Structure of Asset Returns: An Empirical Investigation." *Journal of Business* **75**, 305-332.
- Carr, Peter, Hélyette Geman, Dilip B. Madan, and Marc Yor (2003). "Stochastic Volatility for Lévy Processes." *Mathematical Finance* **13**, 345-382.
- Carr, Peter and Liuren Wu (2003). "The Finite Moment Log Stable Process and Option Pricing." *Journal of Finance* **58**, 753-777.
- Carr, Peter and Liuren Wu (2004). "Time-changed Lévy Processes and Option Pricing." *Journal of Financial Economics* **71**, 113-141.
- Chernov, Mikhail, A. Ronald Gallant, Eric Ghysels, and George Tauchen (2003). "Alternative Models for Stock Price Dynamics." *Journal of Econometrics* **116**, 225-257.
- Clark, Peter K. (1973). "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices." *Econometrica* **41**, 135-155.

- Cox, John C., Stephen A. Ross, and Mark Rubinstein (1979). "Option Pricing: A Simplified Approach." *Journal of Financial Economics* **7**, 229-263.
- Dimson, Elroy (1979). "Risk Measurement When Shares are Subject to Infrequent Trading." *Journal of Financial Economics* **7**, 197-226.
- Duffie, Darrell, Jun Pan, and Kenneth J. Singleton (2000). "Transform Analysis and Asset Pricing for Affine Jump-Diffusions." *Econometrica* **68**, 1343-1376.
- Eberlein, Ernst, Ulrich Keller, and Karsten Prause (1998). "New Insights into Smile, Mispricing, and Value at Risk: The Hyperbolic Model." *Journal of Business* **71**, 371-405.
- Eraker, Bjorn, Michael Johannes, and Nicholas G. Polson (2003). "The Impact of Jumps in Volatility and Returns." *Journal of Finance* **58**, 1269-1300.
- French, Kenneth R. (2008). "The Cost of Active Investing." *Journal of Finance* **63**, 1537-1573.
- French, Kenneth R. and Richard Roll (1986). "Stock Return Variances: The Arrival of Information and the Reaction of Traders." *Journal of Financial Economics* **17**, 5-26.
- Gallant, A. Ronald, Peter E. Rossi, and George Tauchen (1992). "Stock Prices and Volume." *Review of Financial Studies* **5**, 199-242.
- Hentschel, Ludger (1995). "All in the Family: Nesting Symmetric and Asymmetric GARCH Models." *Journal of Financial Economics* **39**, 71-104.
- Heston, Steve L. (1993). "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." *Review of Financial Studies* **6**, 327-344.
- Jukivuolle, Esa (1995). "Measuring True Stock Index Value in the Presence of Infrequent Trading." *Journal of Financial and Quantitative Analysis* **30**, 455-464.
- Kou, Steve (2002). "A Jump Diffusion Model for Option Pricing." *Management Science* **48**, 1086-1101.
- LeBaron, Blake D. (1992). "Some Relations between Volatility and Serial Correlations in Stock Returns." *Journal of Business* **65**, 199-219.
- Li, Haitao, Martin T. Wells, and Cindy L. Yu (2008). "A Bayesian Analysis of Return Dynamics with Stochastic Volatility and Lévy Jumps." *Review of Financial Studies* **21**, 2345-2378.
- Lo, Andrew W. and A. Craig MacKinlay (1988). "Stock Market Prices Do Not Follow Random Walks: Evidence from a New Specification Test." *Review of Financial Studies* **1**, 41-66.
- MacKinlay, A. Craig and Krishna Ramaswamy (1988). "Index-Futures Arbitrage and the Behavior of Stock Index Futures Prices." *Review of Financial Studies* **1**, 137-158.

Madan, Dilip B. and Eugene Seneta (1990). "The Variance Gamma (V.G.) Model for Share Market Returns." *Journal of Business* **63**, 511-525.

Maheu, John M. And Thomas H. McCurdy (2004). "News Arrival, Jump Dynamics and Volatility Components for Individual Stock Returns." *Journal of Finance* **59**, 755-793.

Mandelbrot, Benoit B. (1963). "The Variation of Certain Speculative Prices." *Journal of Business* **36**, 394-419.

Mandelbrot, Benoit B. and Richard L. Hudson (2004). *The (mis)Behavior of Markets: A Fractal View of Risk, Ruin, and Reward*. New York: Basic Books.

Masreliez, C. J. (1975). "Approximate Non-Gaussian Filtering with Linear State and Observation Relations." *IEEE Transactions on Automatic Control* **20**, 107-110.

Merton, Robert C. (1976). "Option Pricing When Underlying Stock Returns are Discontinuous." *Journal of Financial Economics* **3**, 125-144.

*SBBI Yearbook, 2006*. Chicago: R. G. Ibbotson Associates.

Schick, Irvin C. And Sanjoy K. Mitter (1994). "Robust Recursive Estimation in the Presence of Heavy-tailed Observation Noise." *The Annals of Statistics* **22**, 1045-1080.

Schwert, G. William (1990). "Indexes of U.S. Stock Prices from 1802 to 1987." *Journal of Business* **63**, 399-426.

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

## Appendix A. Filtration under Model 2

From equation (26), the cumulant generating function (CGF) for future  $\{y_{t+1}, \rho_{t+1}, V_{t+1}\}$  conditional upon knowing  $\{y_t, \rho_t, V_t\}$  is

$$\ln F(\Phi, \xi, \psi | y_t, \rho_t, V_t) = C(\tau_t; (1 - \rho_t)\Phi, \xi, \psi) + (\xi + \Phi y_t)\rho_t + D(\tau_t; (1 - \rho_t)\Phi, \psi)V_t. \quad (\text{A.1})$$

The filtered CGF conditional upon only observing past data  $\mathbf{Y}_t$  can be computed by integrating over the independent conditional distributions of the latent variables  $\{\rho_t, V_t\}$ :

$$\begin{aligned} F(\Phi, \xi, \psi | \mathbf{Y}_t) &= \int \int e^{C((1 - \rho_t)\Phi, \cdot) + (\xi + \Phi y_t)\rho_t + D((1 - \rho_t)\Phi, \cdot)V_t} p(V_t | \mathbf{Y}_t) p(\rho_t | \mathbf{Y}_t) dV_t d\rho_t \\ &= \int e^{C((1 - \rho_t)\Phi, \cdot) + (\xi + \Phi y_t)\rho_t + g_{t|t}[D((1 - \rho_t)\Phi, \cdot)]} p(\rho_t | \mathbf{Y}_t) d\rho_t \end{aligned} \quad (\text{A.2})$$

where  $g_{t|t}(\psi) \equiv -v_t \ln(1 - \kappa_t \psi)$  is the gamma conditional CGF for latent  $V_t$ . Under the change of variables  $\{z, x\} \equiv \{(1 - \rho_t)\Phi, 1 - \rho_t\}$ , and under the assumption that the scaling term  $x = 1 - \rho_t > 0$ , the Fourier inversion used in evaluating  $p(y_{t+1} | \mathbf{Y}_t)$  from (A.2) becomes

$$\begin{aligned} p(y_{t+1} | \mathbf{Y}_t) &= \frac{1}{\pi} \text{Re} \left\{ \int_{\Phi=0}^{\infty} \int_{\rho_t=-\infty}^1 e^{C((1 - \rho_t)i\Phi, \cdot) + i\Phi y_t \rho_t + g_{t|t}[D((1 - \rho_t)i\Phi, \cdot)] - i\Phi y_{t+1}} p(\rho_t | \mathbf{Y}_t) d\rho_t d\Phi \right\} \\ &= \frac{1}{\pi} \text{Re} \left\{ \int_{z=0}^{\infty} e^{C(iz, 0, 0) + g_{t|t}[D(iz, 0, 0)] - iy_t z} \left( \int_{x=0}^{\infty} \frac{1}{x} e^{\frac{-iz(y_{t+1} - y_t)}{x}} p(x | \mathbf{Y}_t) dx \right) dz \right\} \end{aligned} \quad (\text{A.3})$$

where  $\text{Re}\{c\}$  denotes the real component of complex-valued  $c$ , and the  $1/x$  term in the integrand reflects the Jacobean from the change of variables.

It is convenient to use the two-parameter inverse Gaussian distribution to approximate  $p(x | \mathbf{Y}_t)$ :

$$p(x | \mathbf{Y}_t) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[-\frac{\lambda(x - \mu)^2}{2\mu^2 x}\right], \quad x > 0 \quad (\text{A.4})$$

where  $\mu = E[x | \mathbf{Y}_t]$  and  $\lambda = \mu^3 / \text{Var}[x | \mathbf{Y}_t]$  are  $t$ -dependent parameters that summarize what is known about  $x$  (and about  $\rho_t$ ) at time  $t$ . Under this specification, the inner integration inside (A.3) can be replaced by the analytic function



$$\begin{aligned}
M_{-1}(a) &\equiv \int_{x=0}^{\infty} x^{-1} e^{\frac{a}{x}} p(x | \mathbf{Y}_t) dx \\
&= \sqrt{\frac{\lambda}{(\lambda - 2a)^3}} \left[ 1 + \frac{\sqrt{\lambda(\lambda - 2a)}}{\mu} \right] \exp\left[ \frac{\lambda - \sqrt{\lambda(\lambda - 2a)}}{\mu} \right]
\end{aligned} \tag{A.5}$$

for  $a = -iz(y_{t+1} - y_t) \equiv -iz\Delta y$ . Consequently, evaluating (A.3) involves only univariate numerical integration.<sup>1</sup>

Similar univariate integrations are used for filtering  $V_{t+1}$  and  $\rho_{t+1}$  conditional upon observing  $y_{t+1}$ . The noncentral posterior moments of  $V_{t+1}$  are given by

$$E(V_{t+1}^m | \mathbf{Y}_{t+1}) = \frac{1}{\pi p(y_{t+1} | \mathbf{Y}_t)} \operatorname{Re} \left\{ \int_{z=0}^{\infty} \frac{\partial^m \left[ e^{C(z, 0, \Psi) + g_{it}[D(z, 0, \Psi)] - iy_t z} \right]}{\partial \Psi^m} \Bigg|_{\Psi=0} M_{-1}(-iz\Delta y) dz \right\} \tag{A.6}$$

where the derivatives with respect to  $\Psi$  inside the integrand can be easily evaluated from the specifications for  $C(\bullet)$  and  $D(\bullet)$  in equations (17) - (18). The posterior moments of  $\rho_{t+1}$  can be computed by taking partials of (A.2) with respect to  $\xi$ , and then again using change of variables to reduce the Fourier inversion to a univariate integration. The resulting posterior mean and variance of  $\rho_{t+1}$  are

$$\hat{\rho}_{t+1|t+1} = 1 - \frac{1}{\pi p(y_{t+1} | \mathbf{Y}_t)} \operatorname{Re} \left\{ \int_{z=0}^{\infty} e^{C(z, 0, 0) + g_{it}[D(z, 0, 0)] - iy_t z} M_0(-iz\Delta y) dz \right\} \tag{A.7}$$

$$W_{t+1|t+1} = \sigma_{\varepsilon}^2 + \frac{1}{\pi p(y_{t+1} | \mathbf{Y}_t)} \operatorname{Re} \left\{ \int_{z=0}^{\infty} e^{C(z, 0, 0) + g_{it}[D(z, 0, 0)] - iy_t z} M_1(-iz\Delta y) dz \right\} \tag{A.8}$$

where

---

<sup>1</sup>A more “natural” choice would be to represent  $x$  by a beta distribution over the range  $[0, 2]$ . That would constrain  $|\rho_t| < 1$ , and results in an  $M_{-1}(a)$  term that involves the confluent hypergeometric U-function. However, I could not find a method for evaluating that function that was fast, accurate, and robust to all parameter values.

$$M_0(a) \equiv \int_{x=0}^{\infty} e^{\frac{a}{x}} p(x | Y_t) dx = \sqrt{\frac{\lambda}{\lambda - 2a}} \exp\left[\frac{\lambda - \sqrt{\lambda(\lambda - 2a)}}{\mu}\right] \quad (\text{A.9})$$

and

$$M_1(a) \equiv \int_{x=0}^{\infty} x e^{\frac{a}{x}} p(x | Y_t) dx = \mu \exp\left[\frac{\lambda - \sqrt{\lambda(\lambda - 2a)}}{\mu}\right]. \quad (\text{A.10})$$

Finally, the conditional distribution function  $P(y) \equiv \text{Prob}[y_{t+1} < y | Y_t]$  that is used in QQ plots takes the form

$$P(y) = \frac{1}{2} - \frac{1}{\pi} \text{Re} \left\{ \int_{z=0}^{\infty} \frac{\exp[C(z, 0, 0) + g_{t|t}[D(z, 0, 0)] - iy_t z]}{iz} M_0(-iz \Delta y) dz \right\}. \quad (\text{A.11})$$

**Table 1****Standardized cumulant exponents (with unitary variance) for various compensated Lévy specifications**

diffusion:  $g_{SV}(u) = \frac{1}{2}(u^2 - u)$

Normally distributed jumps:  $g_{Ji}(u) = \frac{1}{\bar{\gamma}_i^2 + \delta_i^2} \left[ e^{u\bar{\gamma}_i + \frac{1}{2}u^2\delta_i^2} - 1 - u(e^{\bar{\gamma}_i + \frac{1}{2}\delta_i^2} - 1) \right]$

CGMY jump process:  $g_{CGMY}(u) = w_n \frac{(G - u)^{Y_n} - G^{Y_n}}{Y_n(Y_n - 1)G^{Y_n - 2}} + (1 - w_n) \frac{(M + u)^{Y_p} - M^{Y_p}}{Y_p(Y_p - 1)M^{Y_p - 2}} - \omega u$

with  $\omega$  such that  $g_{CGMY}(1) = 0$

General specification:  $g_{dL}(u) = w_{SV} g_{SV}(u) + \sum_{i=1}^2 w_{Ji} g_{Ji}(u) + w_{CGMY} g_{CGMY}(u)$

Model	Weights			Parameter restrictions
	$w_{SV}$	$w_{Ji}$	$w_{CGMY}$	
SV	1			
SVJ1	$\frac{1}{1 + \lambda_1(\bar{\gamma}_1^2 + \delta_1^2)}$	$w_{J1} = \frac{\lambda_1(\bar{\gamma}_1^2 + \delta_1^2)}{1 + \lambda_1(\bar{\gamma}_1^2 + \delta_1^2)}$		$w_{J2} = 0$
SVJ2	$\frac{1}{1 + \sum_{i=1}^2 \lambda_i(\bar{\gamma}_i^2 + \delta_i^2)}$	$\frac{\lambda_i(\bar{\gamma}_i^2 + \delta_i^2)}{1 + \sum_{i=1}^2 \lambda_i(\bar{\gamma}_i^2 + \delta_i^2)}$		
DEXP	$1 - f_{jump}$		$f_{jump}$	$Y_n = Y_p = -1$
VG	$1 - f_{jump}$		$f_{jump}$	$Y_n = Y_p = 0$
Y			1	$Y_n = Y_p$
YY			1	
LS			1	$w_n = 1, G = .001$
YY_D	$1 - f_{jump}$		$f_{jump}$	

For SVJ1 and SVJ2, the fraction  $f_{jump}$  of variance attributable to jumps is  $w_{J1}$  and  $w_{J1} + w_{J2}$ , respectively.

**Table 2: Effective length of a business day, relative to 1-day Wednesday returns: 1926-2006.** Estimates from YY model. Estimates from other models are almost identical.

#days	Description	NOBS	Model 1		Model 2	
			estimate	std. error	estimate	std. error
1	Monday close → Tuesday close	3831	1.02	(.04)	1.03	(.03)
1	<b>Tuesday close → Wednesday close</b>	<b>4037</b>	<b>1</b>		<b>1</b>	
1	Wednesday → Thursday	3998	.94	(.03)	.94	(.03)
1	Thursday → Friday	3924	.93	(.03)	.92	(.03)
1	Friday → Saturday (1926-52)	1141	.43	(.02)	.44	(.02)
2	Saturday close → Monday close (1926-52)	1120	1.05	(.05)	1.07	(.05)
2	Weekday holiday	341	1.25	(.11)	1.26	(.10)
2	Wednesday exchange holiday in 1968	22	.73	(.33)	.81	(.35)
3	Weekend and/or holiday <sup>a</sup>	2755	1.10	(.04)	1.10	(.04)
4	Holiday weekend	343	1.58	(.14)	1.56	(.13)
5	Holiday weekend	<u>6</u>	1.31	(1.00)	1.25	(.93)
		21518				
	Annualization factor: Wednesday → yearly		259.8	(5.6)	260.3	(5.5)

<sup>a</sup>Includes one weekday holiday (August 14 - 17, 1945)

**Table 3A: Estimates of parameters affecting the conditional means and volatilities.** Data: daily CRSP value-weighted excess returns, 1926-2006. See equations (6) - (10), (13), and (25) for definitions of parameters. Models with  $f_{jump} < 1$  combine Lévy jump processes with an additional independent diffusion, with variance proportions  $(f_{jump}, 1 - f_{jump})$ , respectively. Standard errors are in parentheses.

Model 1: $y_{t+1} = \rho_t y_t + \eta_{t+1}$									
Model	ln $L$	Conditional mean			Stochastic volatility				
		$\mu_0$	$\mu_1$	$\sigma_\rho \sqrt{252}$	$\sqrt{\alpha/\beta}$	$\beta$	$\sigma$	$\rho_{sv}$	HL (mths)
SV	74,940.85	.013 (.015)	2.16 (.90)	.029 (.007)	.153 (.004)	5.83 (.44)	.452 (.010)	.625 (.018)	1.4 (.1)
SVJ1	75,043.90	.042 (.015)	.91 (.91)	.030 (.006)	.155 (.005)	4.39 (.40)	.374 (.011)	-.641 (.020)	1.9 (.2)
SVJ2	75,048.49	.042 (.003)	.87 (.76)	.030 (.007)	.155 (.007)	4.34 (.37)	.371 (.015)	-.642 (.018)	1.9 (.2)
DEXP	75,047.33	.043 (.015)	.87 (.90)	.031 (.007)	.155 (.005)	4.23 (.38)	.368 (.012)	-.587 (.020)	2.0 (.2)
VG	75,049.09	.043 (.015)	.92 (.91)	.030 (.006)	.155 (.005)	4.22 (.39)	.366 (.012)	-.586 (.020)	2.0 (.2)
Y	75,049.63	.042 (.015)	.90 (.92)	.030 (.006)	.156 (.009)	3.89 (.38)	.351 (.020)	-.576 (.032)	2.1 (.2)
<b>YY</b>	<b>75,052.56</b>	<b>.041 (.015)</b>	<b>.87 (.92)</b>	<b>.030 (.006)</b>	<b>.158 (.009)</b>	<b>4.00 (.38)</b>	<b>.360 (.019)</b>	<b>-.571 (.031)</b>	<b>2.1 (.2)</b>
YY_D	75,052.81	.042 (.015)	.93 (.91)	.030 (.006)	.156 (.006)	3.99 (.38)	.355 (.013)	-.579 (.021)	2.1 (.2)
LS	75,007.86	.018 (.015)	1.50 (.73)	.031 (.007)	.171 (.006)	4.60 (.40)	.431 (.015)	-.541 (.020)	1.8 (.2)
Model 2: $y_{t+1} = \rho_t y_t + (1 - \rho_t) \eta_{t+1}$									
SV	74,999.87	-.014 (.020)	3.04 (.90)	.043 (.005)	.170 (.004)	8.01 (.57)	.562 (.015)	-.658 (.017)	1.0 (.1)
SVJ1	75,092.10	.033 (.020)	1.69 (1.04)	.036 (.005)	.171 (.004)	5.80 (.49)	.457 (.015)	-.674 (.018)	1.4 (.1)
SVJ2	75,096.68	.037 (.020)	1.25 (.89)	.036 (.005)	.172 (.004)	5.71 (.49)	.456 (.015)	-.673 (.018)	1.4 (.1)
DEXP	75,094.20	.034 (.020)	1.44 (.90)	.036 (.005)	.171 (.004)	5.67 (.49)	.452 (.015)	-.625 (.018)	1.5 (.1)
VG	75,094.70	.034 (.020)	1.42 (.90)	.037 (.005)	.171 (.004)	5.56 (.48)	.447 (.016)	-.623 (.018)	1.6 (.1)
Y	75,093.68	.036 (.021)	1.35 (.90)	.036 (.005)	.172 (.007)	5.18 (.46)	.432 (.021)	-.613 (.027)	1.6 (.1)
<b>YY</b>	<b>75,097.20</b>	<b>.033 (.020)</b>	<b>1.44 (.90)</b>	<b>.036 (.005)</b>	<b>.172 (.006)</b>	<b>5.23 (.47)</b>	<b>.437 (.018)</b>	<b>-.613 (.022)</b>	<b>1.6 (.1)</b>
YY_D	75,097.49	.035 (.020)	1.36 (.90)	.036 (.005)	.172 (.005)	5.23 (.47)	.436 (.016)	-.616 (.020)	1.6 (.1)
LS	75,045.48	.053 (.019)	1.50 (.76)	.031 (.003)	.174 (.005)	4.68 (.41)	.436 (.015)	-.576 (.019)	1.8 (.2)

**Table 3B: Estimates of jump parameters.** Standard errors in parentheses.

$$\text{Model 1: } y_{t+1} = \rho_t y_t + \eta_{t+1}$$

Model	$f_{jump}$	CGMY parameters					Merton parameters		
		$w_n$	$G$	$M$	$Y_n$	$Y_p$	$\lambda_i$	$\bar{\gamma}$	$\delta$
SVJ1	.150 (.017)						142.7 (22.7)	.000 (.002)	.032 (.002)
SVJ2	.156 (.054)						162.9 (30.9)	.000 (.002)	.029 (.002)
DEXP	.253 (.027)	.49 (.01)	66.1 (6.0)	45.4 ( 8.4)		-1	0.5 (1.6)	-.189 (.094)	.005 (.189)
VG	.272 (.030)	.52 (.07)	41.1 (5.4)	31.6 (9.1)		0			
Y	1	.59 (.06)	7.0 (4.6)	2.3 (7.3)		1.87 (.03)			
<b>YY</b>	<b>1</b>	<b>.88 (.03)</b>	<b>1.6 (4.5)</b>	<b>40.1 (31.3)</b>		<b>1.94 (.01)</b>			<b>-.24 (1.36)</b>
YY_D	.436	.72 (.15)	6.9 (9.0)	49.2 (34.9)		1.71 (.35)			-.72 (1.57)
LS	1	1	.001			1.96 (.01)			

$$\text{Model 2: } y_{t+1} = \rho_t y_t + (1 - \rho_t) \eta_{t+1}$$

SVJ1	.133 (.015)						114.1 (19.2)	-.001 (.003)	.034 (.002)
SVJ2	.140 (.015)						126.4 (23.7)	.000 (.002)	.031 (.002)
DEXP	.236 (.026)	.53 (.06)	55.4 (5.4)	50.0 (4.8)		-1	0.41 (.04)	-.198 (.022)	.010 (.046)
VG	.257 (.030)	.54 (.07)	41.1 (4.9)	31.6 (10.7)		0			
Y	1	.59 (.05)	6.8 (4.2)	3.2 (8.2)		1.87 (.03)			
<b>YY</b>	<b>1</b>	<b>.89 (.03)</b>	<b>2.6 (4.1)</b>	<b>71.1 (57.8)</b>		<b>1.935 (.014)</b>			<b>-1.96 (2.62)</b>
YY_D	.380 (.158)	.90 (.30)	8.1 (8.7)	51.8 (51.7)		1.619 (.403)			-1.08 (2.36)
LS	1	1	.001			1.965 (.006)			

**Table 4: Subsample estimates for Model 2.** Estimates “w/o Oct ’87” exclude daily data observed in October 1987, but include the full month’s return of -20.6%. Split-sample estimates involve different parameter values before/after March 5, 1957, apart from time dummies. Standard errors are in parentheses.

Model	Period	ln $L$	Conditional mean			Stochastic volatility				
			$\mu_0$	$\mu_1$	$\sigma_\rho \sqrt{252}$	$\sqrt{\alpha/\beta}$	$\beta$	$\sigma$	$\rho_{sv}$	HL (mths)
SVJ1	full	75,092.10	.033 (.020)	1.69 (1.04)	.036 (.005)	.171 (.004)	5.80 (.49)	.457 (.015)	-.674 (.018)	1.4 (.1)
SVJ1	1926 - 1957	}75,183.99	.051 (.034)	1.35 (1.38)	.050 (.009)	.202 (.007)	8.82 (1.03)	.678 (.035)	-.661 (.027)	0.9 (.1)
	1957 - 2006		.003 (.027)	2.90 (1.61)	.027 (.005)	.149 (.005)	4.93 (0.60)	.314 (.015)	-.725 (.023)	1.7 (.2)
YY	full	75,097.20	.033 (.020)	1.44 (.90)	.036 (.005)	.172 (.006)	5.23 (.47)	.437 (.018)	-.613 (.022)	1.6 (.1)
YY	w/o Oct ’87	75,065.01	.036 (.020)	1.37 (.91)	.035 (.005)	.170 (.005)	5.12 (.46)	.427 (.016)	-.620 (.019)	1.6 (.1)
YY	1926 - 1957	}75,196.14	.056 (.034)	1.03 (1.15)	.051 (.009)	.201 (.008)	6.81 (.88)	.657 (.033)	-.585 (.026)	1.2 (.2)
	1957 - 2006		.012 (.027)	.35 (.67)	.025 (.005)	.365 (.320)	4.89 (.56)	.404 (.026)	-.281 (.247)	1.7 (.2)
Model	$f_{jump}$	CGMY parameters					Merton parameters			
		$w_n$	$G$	$M$	$Y_n$	$Y_p$	$\lambda_i$	$\bar{\gamma}$	$\delta$	
SVJ1	full	.133 (.015)					114.1 (19.2)	-.001 (.003)	.034 (.002)	
SVJ1	1926 - 1956	.167 (.015)					216.8 (54.9)	.000 (.003)	.028 (.003)	
	1957 - 2006	.093 (.020)					49.5 (12.0)	-.003 (.007)	.043 (.004)	
YY	full	1	.89 (.03)	2.6 (4.1)	71.1 (57.8)	1.94 (.01)	-1.96 (2.6)			
YY	w/o Oct ’87	1	.89 (.03)	5.6 (5.3)	72.0 (60.8)	1.93 (.02)	-1.76 (2.6)			
YY	1926 - 1957	1	.86 (.04)	20.7 (7.9)	97.8	1.82 (.05)	-3.1 (4.5)			
	1957 - 2006	1	.92 (.15)	0.0 (0.0)	(107.2)	1.54 (.41)	1.94 (.03)			
					6.0 (13.4)					

**Table 5: Change of parameters (objective versus risk-neutral) under a myopic power utility pricing kernel  $d\ln M = \mu_m^* dt - R ds$ .**

	<u>Objective</u>	<u>Risk-neutral</u>
Equity premium $\mu_0 + \mu_1 V_t$	$\mu_0 = 0, \mu_1 \approx R$	0
General jump intensity	$k(x) V_t$	$k^*(x) V_t = [k(x)e^{-Rx}] V_t$
<b>Variance process <math>V_t</math></b>		
mean reversion	$\beta$	$\beta^* \equiv \beta + R\sigma\rho$
UC mean	$\alpha/\beta$	$\alpha/\beta^*$
<b>Merton jump process</b>		
mean jump size	$\bar{\gamma}_i$	$\bar{\gamma}_i^* = \bar{\gamma}_i - R\delta_i^2$
jump SD	$\delta_i$	$\delta_i$
jump intensity	$\lambda_i V_t$	$\lambda_i^* = \lambda_i \exp(-R\bar{\gamma}_i + \frac{1}{2}R^2\delta_i^2)$

$$g_{SVJ2}^*(\Phi) = \frac{\frac{1}{2}(\Phi^2 - \Phi) + \sum_{i=1}^2 \lambda_i^* \left[ e^{\Phi\bar{\gamma}_i^* + \frac{1}{2}\Phi^2\delta_i^2} - 1 - \Phi \left( e^{\bar{\gamma}_i^* + \frac{1}{2}\delta_i^2} - 1 \right) \right]}{1 + \sum_{i=1}^2 \lambda_i (\bar{\gamma}_i^2 + \delta_i^2)}$$

**CGMY jump process**

G	$G^* = G - R$ (must be $\geq 0$ )
M	$M^* = M + R$
$Y_n, Y_p$	$Y_n, Y_p$

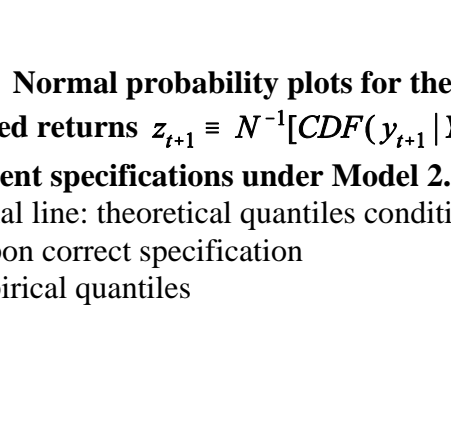
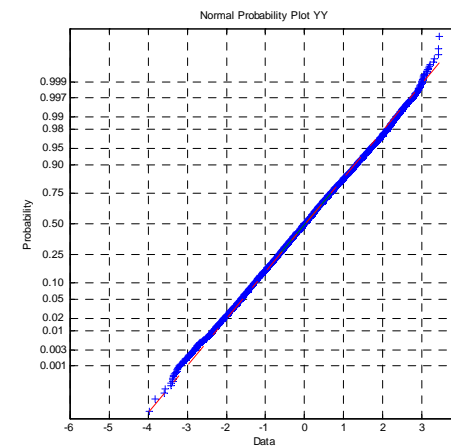
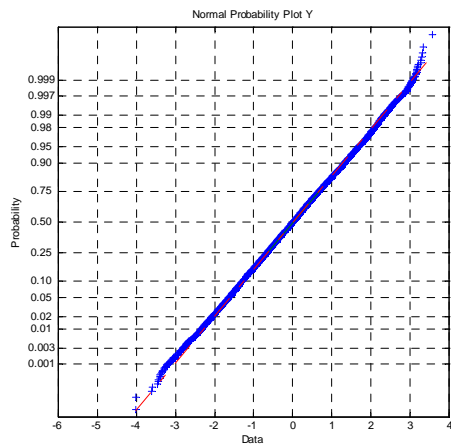
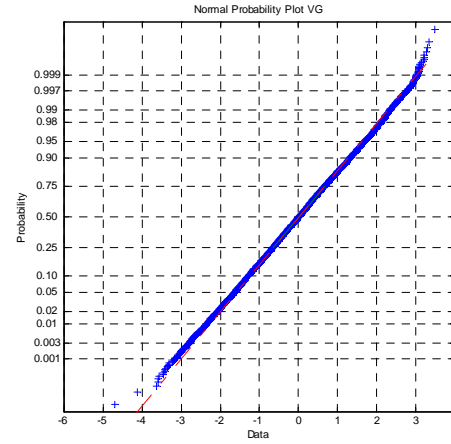
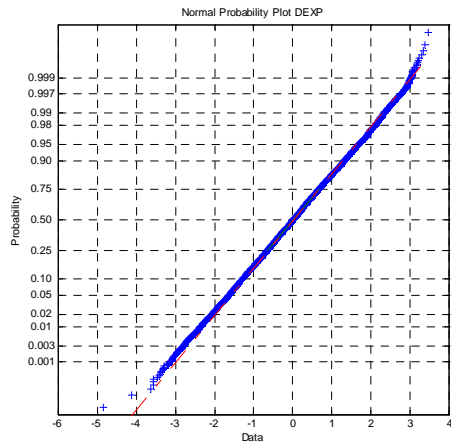
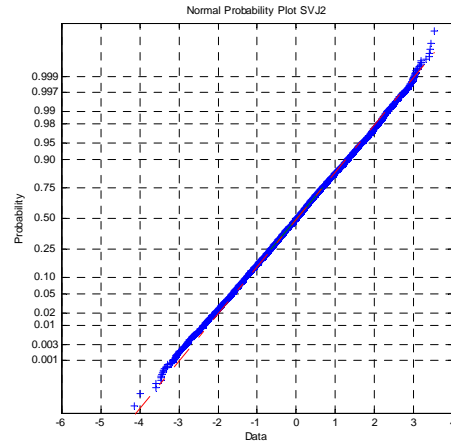
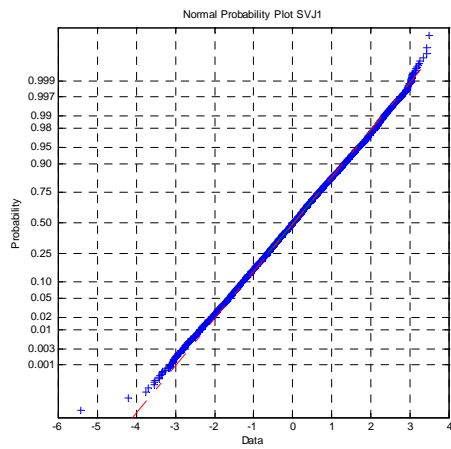
$$g_{CGMY}^*(\Phi) = \frac{1}{2}(1 - f_{jump})(\Phi^2 - \Phi) + f_{jump} w_n \frac{(G^* - \Phi)^{Y_n} - \Phi(G^* - 1)^{Y_n} + (\Phi - 1)(G^*)^{Y_n}}{Y_n(Y_n - 1)G^{Y_n - 2}} + f_{jump} w_p \frac{(M^* + \Phi)^{Y_p} - \Phi(M^* + 1)^{Y_p} + (\Phi - 1)(M^*)^{Y_p}}{Y_p(Y_p - 1)M^{Y_p - 2}}$$



**Table 6: Parameter estimates over 1926-2006 on spliced CRSP/S&P 500 data with constrained equity premium:  $\mu_0 = 0$ ,  $\mu_1 \approx R$ .**

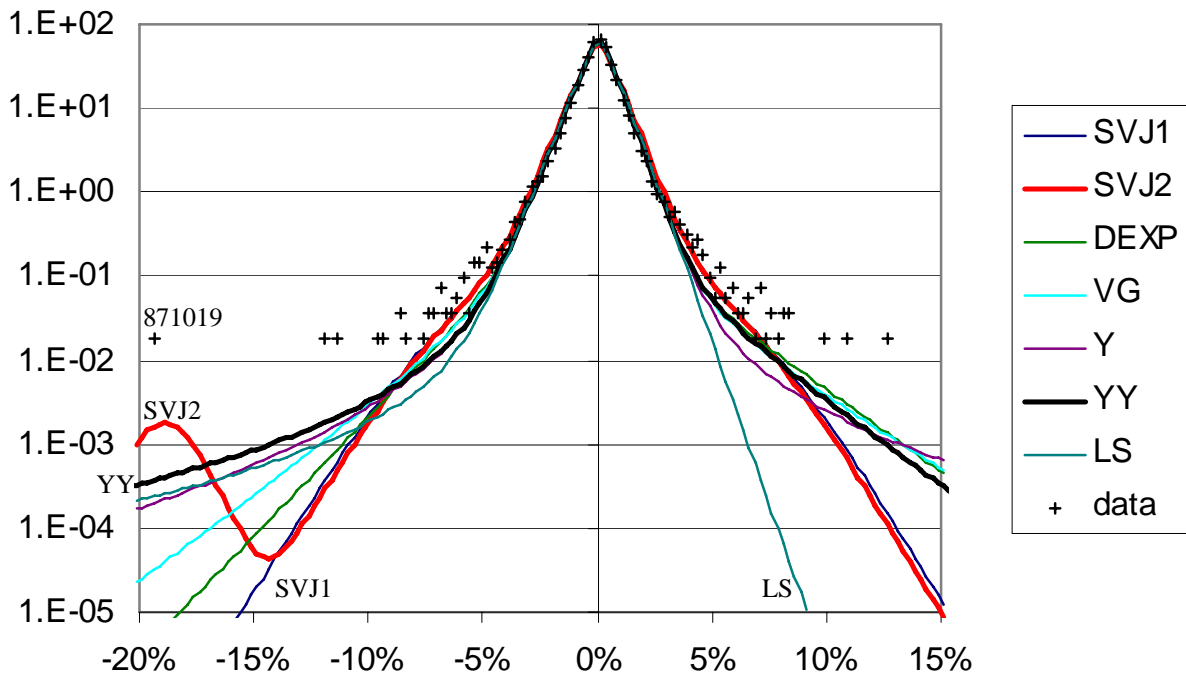
Model	ln $L$	LR test of $\mu_0 = 0$ (p-value)	Conditional mean			Stochastic volatility				
			$R$	$\mu_1$	$\sigma_\rho \sqrt{252}$	$\sqrt{\alpha/\beta}$	$\beta$	$\sigma$	$\rho_{sv}$	HL (mths)
SV	74,028.53	.383	2.49 (.62)	2.49 (.62)	.043 (.005)	.172 (.004)	7.18 (.48)	.534 (.014)	-.649 (.016)	1.2 (.1)
SVJ1	74,119.26	.265	2.44 (.61)	2.44 (.61)	.038 (.005)	.173 (.004)	5.76 (.43)	.448 (.015)	-.678 (.017)	1.4 (.1)
SVJ2	74,125.33	.507	2.43 (.57)	2.44 (.58)	.037 (.005)	.174 (.004)	5.76 (.42)	.449 (.015)	-.679 (.017)	1.4 (.1)
DEXP	75,121.73	.247	2.44 (.61)	2.44 (.61)	.037 (.005)	.174 (.004)	5.68 (.43)	.444 (.015)	-.632 (.017)	1.5 (.1)
VG	74,122.51	.206	2.50 (.61)	2.50 (.61)	.037 (.005)	.174 (.004)	5.62 (.43)	.441 (.015)	-.631 (.017)	1.5 (.1)
Y	74,122.19	.185	2.42 (.61)	2.42 (.61)	.037 (.005)	.174 (.006)	5.29 (.41)	.427 (.018)	-.623 (.022)	1.6 (.1)
YY <sup>a</sup>	74,124.33	.212	2.38	2.42	.037	.175	5.26	.429	-.621	1.6
Estimates on S&P 500 data over 1957-2006										
SVJ2	43,707.25		2.97 (.88)	2.98 (.88)	.026 (.006)	.149 (.006)	3.97 (.44)	.289 (.014)	-.721 (.024)	2.1 (.2)
Model	$f_{jump}$	CGMY parameters					Merton parameters			
		$w_n$	$G$	$M$	$Y_n$	$Y_p$	$\lambda_i$	$\bar{\gamma}$	$\delta$	
SVJ1	.126 (.015)						108.7 (18.2)	-.001 (.003)	.034 (.002)	
SVJ2	.138 (.021)						122.7 (23.1)	.000 (.002)	.031 (.002)	
							0.43 (.38)	-.219 (.027)	.003 (.150)	
DEXP	.228 (.026)	.55 (.06)	51.6 (5.0)	53.9 (12.7)		-1				
VG	.247 (.028)	.53 (.07)	35.9 (4.6)	34.4 (11.1)		0				
Y	1	.58 (.05)	5.4 (4.0)	4.5 (8.7)		1.87 (.03)				
YY <sup>a</sup>	1	.90	2.4	64.3		1.94	-1.29			
Estimates on S&P 500 data over 1957-2006										
SVJ2	.096 (.050)						81.0 (30.1)	.002 (.004)	.030 (.005)	
							.55 (1.32)	-.213 (.035)	.003 (.020)	

<sup>a</sup>Parameter constraint  $G \geq R$  was binding for the YY model; standard errors could not be computed.

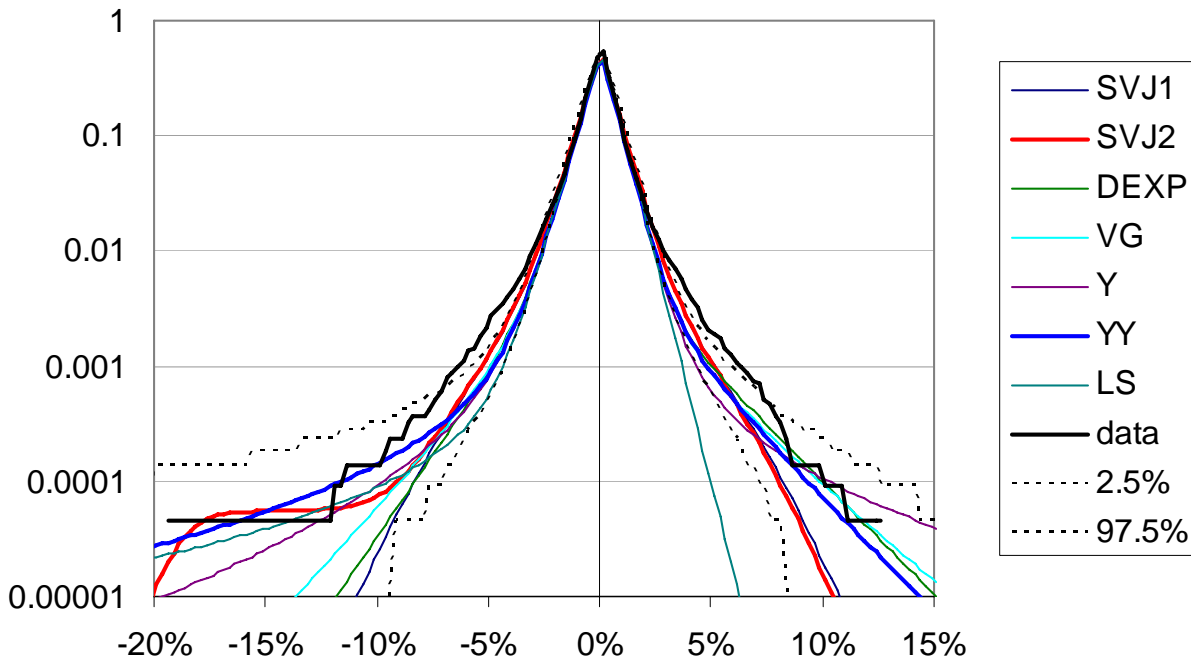


**Figure 1. Normal probability plots for the normalized returns  $z_{t+1} \equiv N^{-1}[CDF(y_{t+1} | Y_t, \hat{\theta})]$ , for different specifications under Model 2.**  
 Diagonal line: theoretical quantiles conditional upon correct specification  
 +: Empirical quantiles

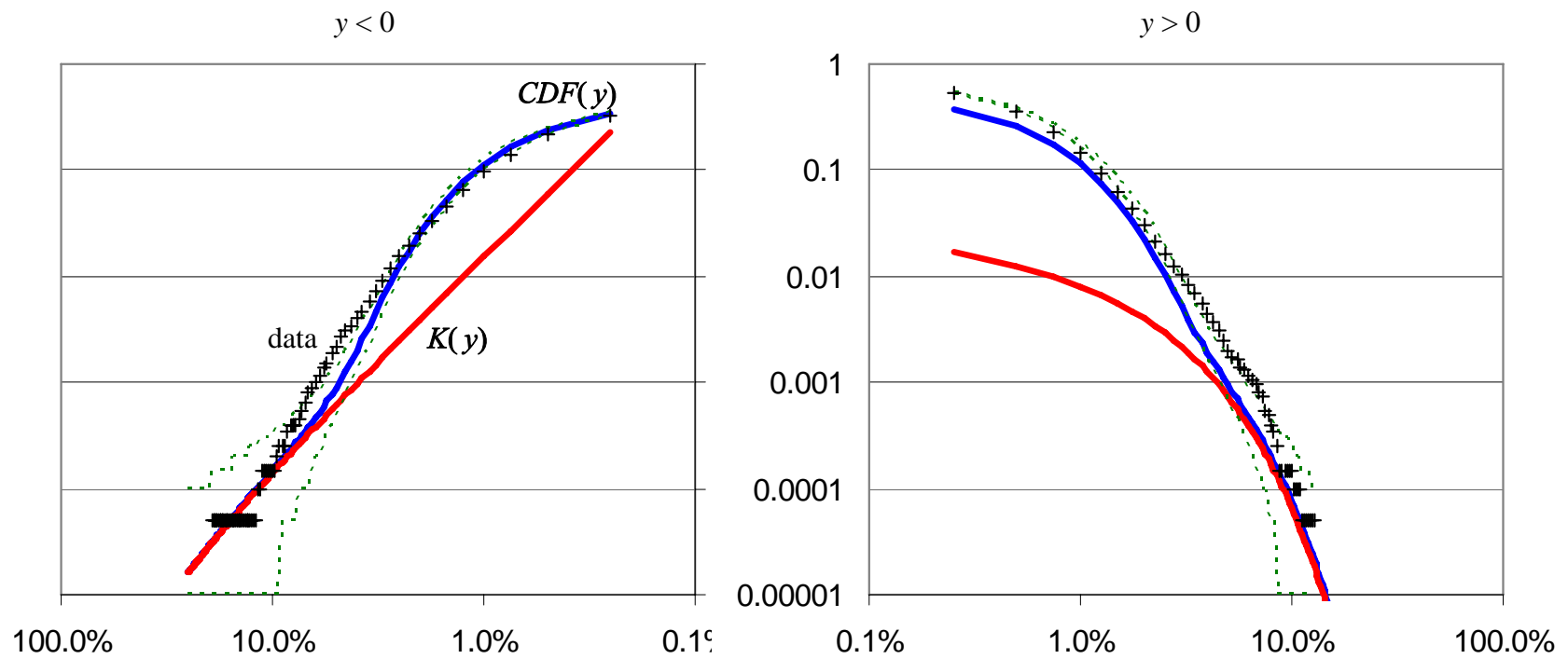
**Figure 2a. Unconditional probability density functions from Model 1 specifications.** Data-based estimates are from a histogram of residuals (.25% cell width).



**Figure 2b. Unconditional tail probability estimates.** The dotted lines give 95% confidence intervals, based upon 1000 simulations of the 1926-2006 data set under YY parameter estimates.



**Figure 3. Unconditional tail probabilities and tail intensity functions versus  $|y|$ .** Log scales on both axes. Data-based estimates from excess returns' residuals for 20,004 business days with estimated time horizons of approximately 1 day ( $\pm 11\%$ ). Dotted lines give 95% confidence intervals, based upon 1000 simulated sample paths under YY parameter estimates.



**Figure 4: Autocorrelation revision  $\hat{\rho}_{t+1|t+1} - \hat{\rho}_{t|t}$  conditional on observing  $y_{t+1}$ , and conditional on  $y_t = \pm 1\%$**

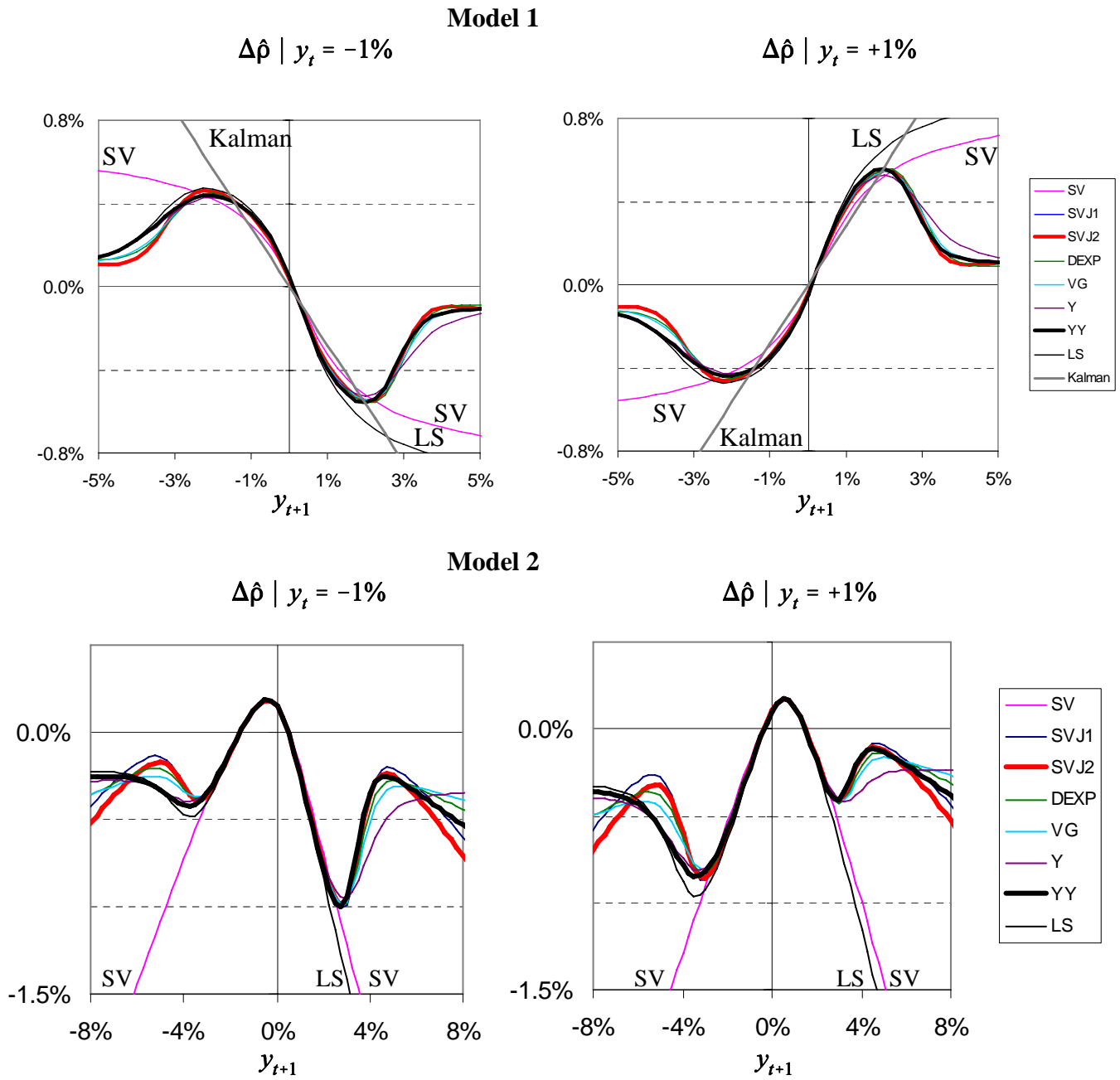
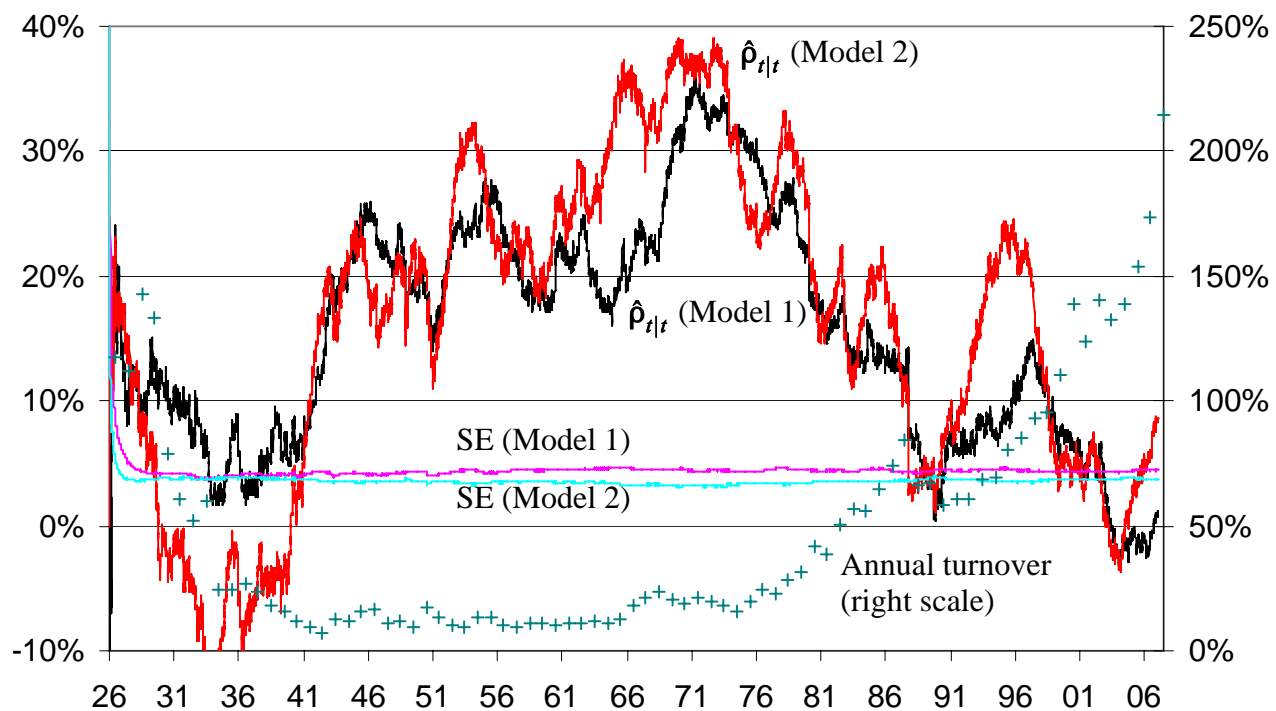
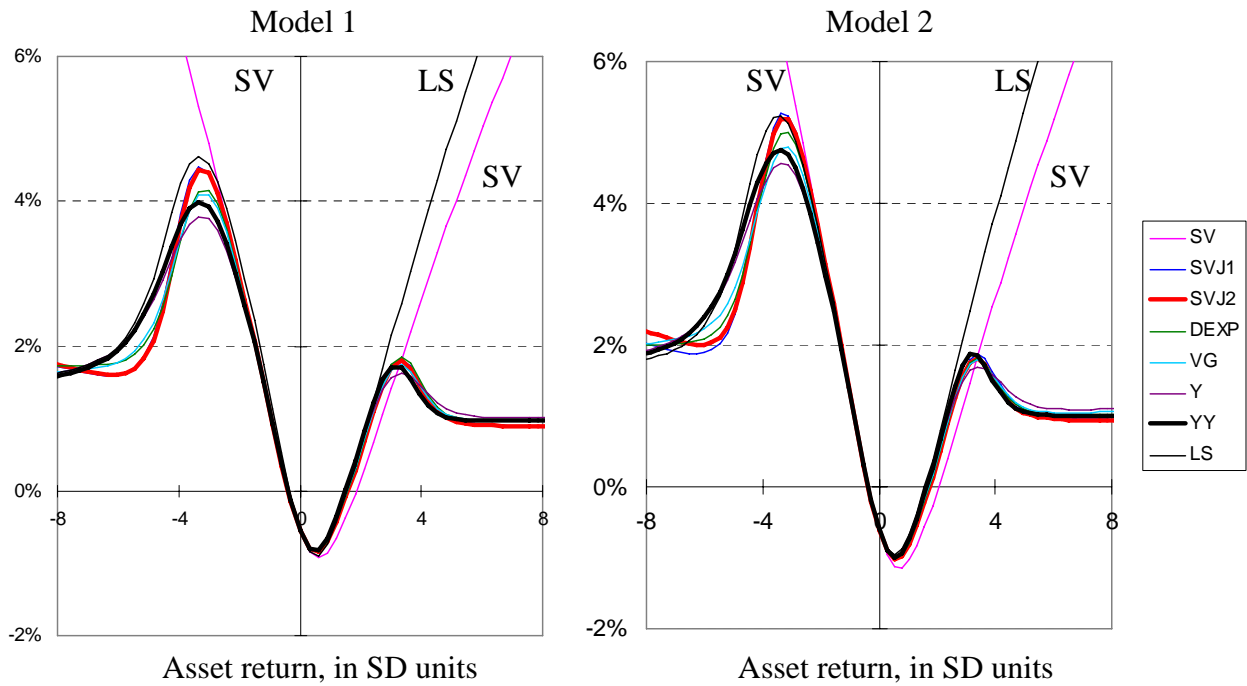


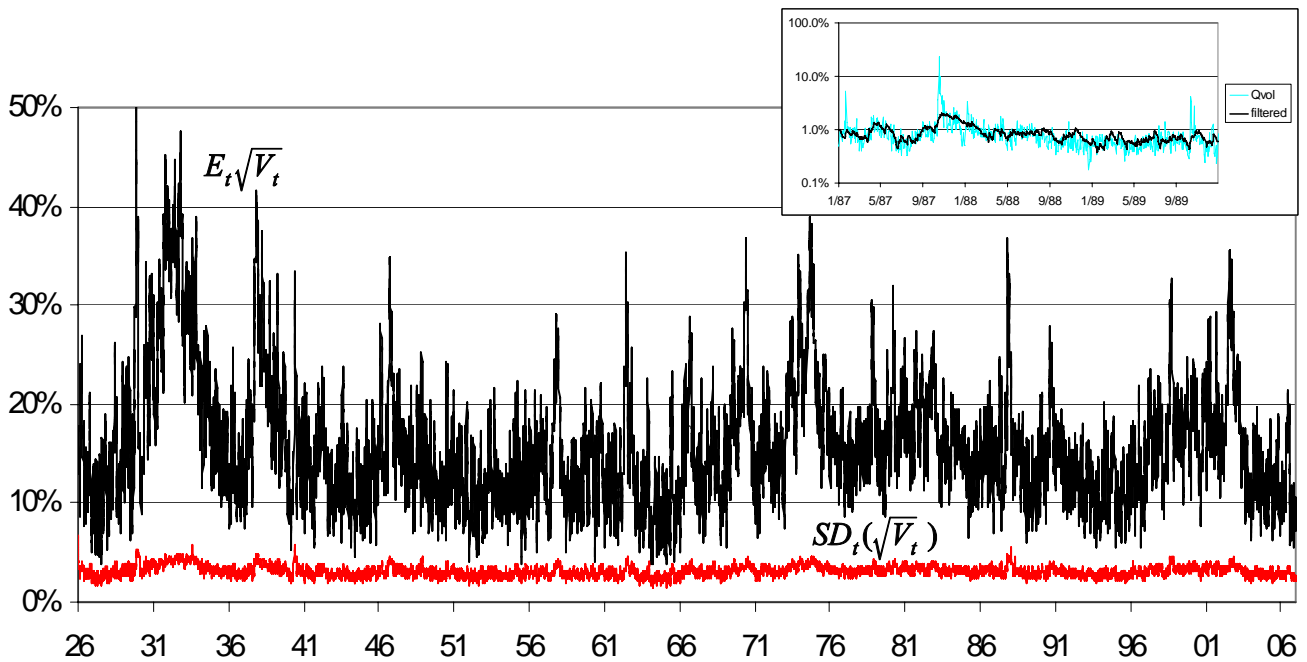
Figure 5: Autocorrelation estimates and standard errors from YY model, and stocks' annual turnover from French (2008).



**Figure 6: News impact curves for various models**



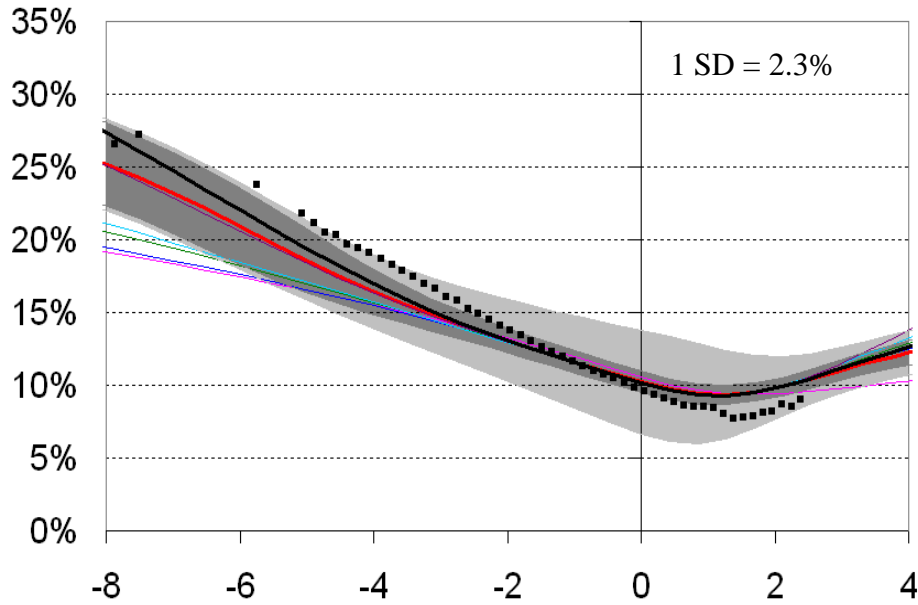
The graph show the revision in estimated annualized standard deviation  $(E_{t+1} - E_t)\sqrt{V_{t+1}}$  conditional upon observing a standardized return of magnitude  $y_{t+1}/\sqrt{\hat{V}_{t|t}}/252$ .



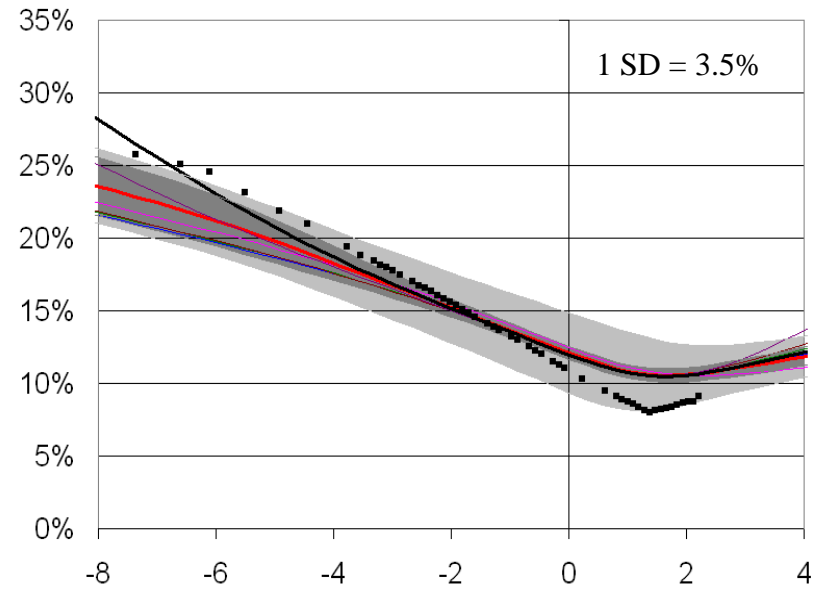
**Figure 7: Estimates of annualized permanent volatility (YY model 2) and standard errors. Inset compares *daily* (open to close) filtered volatility estimates over 1987-89 with the realized volatilities computed daily from 15-minute log-differenced S&P 500 futures prices, on a log scale.**

**Figure 8: Estimated and observed ISD's for options on S&P 500 futures on December 29, 2006.** Moneyness is measured in standard deviation units, given the maturity-specific at-the-money ISD from options. 95% confidence intervals from the SVJ2 model are shown for parameter uncertainty (dark grey), and parameter and state uncertainty (light grey).

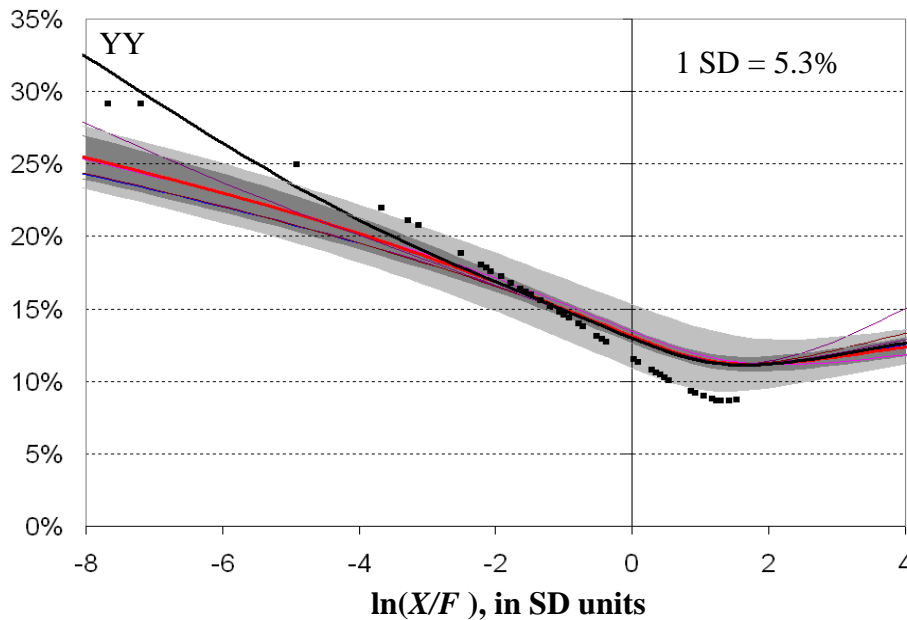
**Jan '07: 21 days maturity**



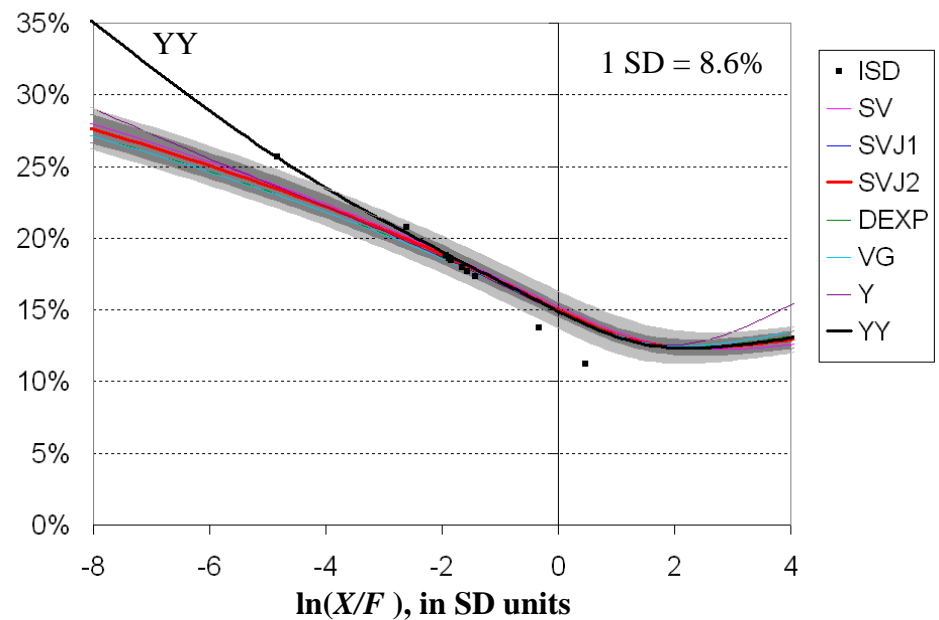
**Feb. '07: 49 days maturity**



**March '07: 77 days maturity**

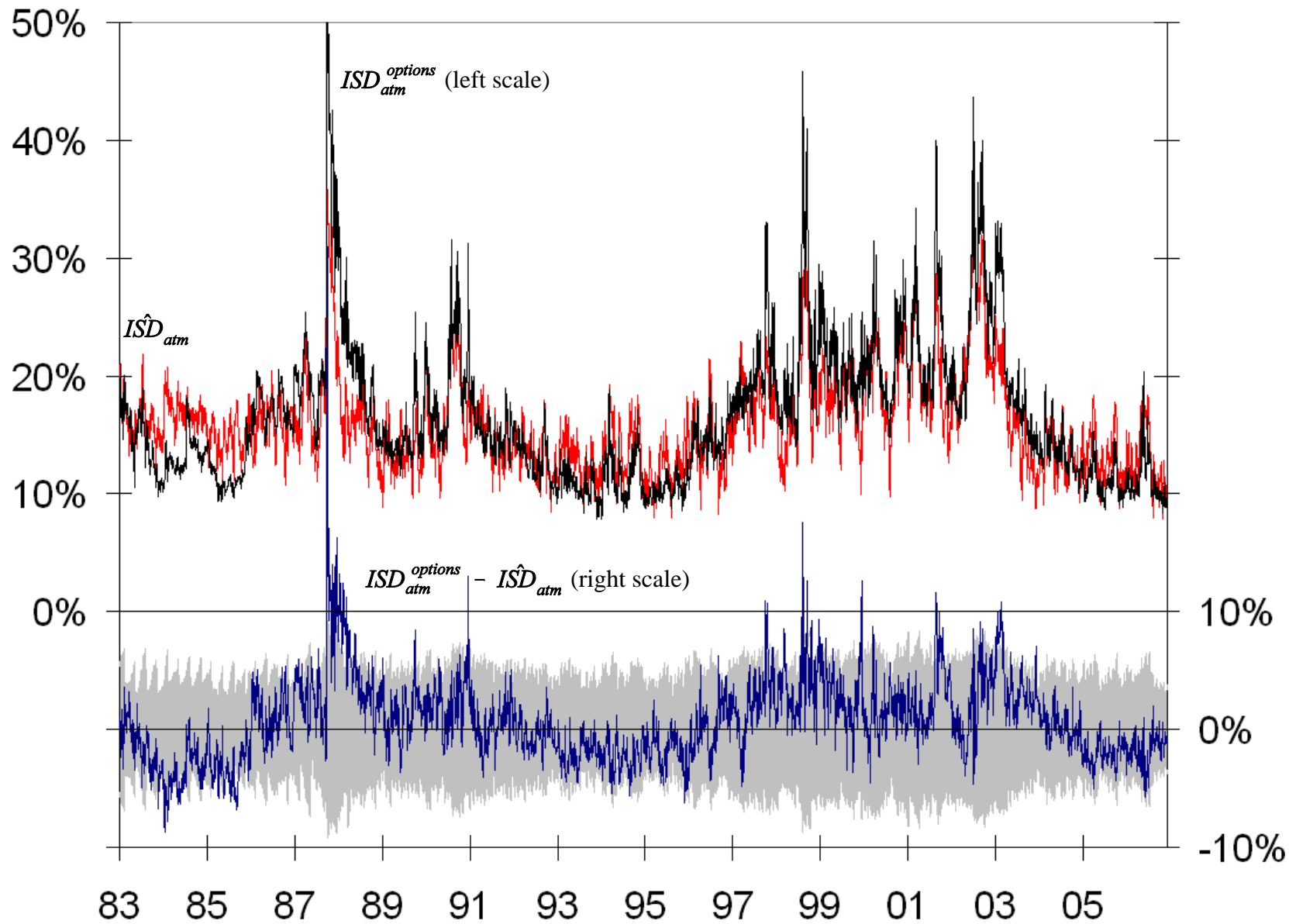


**June '07: 168 days maturity**



- ISD
- SV
- SVJ1
- SVJ2
- DEXP
- VG
- Y
- YY





**Figure 9: Estimated and observed at-the-money ISD's, 1983-2006.** Observed ISD's are from short-term options on S&P 500 futures with at least 14 days to maturity, while estimated ISD's are based on the SVJ2 specification's parameter and volatility estimates. The grey area is the 95% confidence interval for the difference, given both parameter and state uncertainty.