Estimation and Filtration of Time-Changed Lévy Processes

David S. Bates

University of Iowa and NBER

2nd Risk Mangement Conference

Mont Tremblant, Québec

March 14, 2008

Objectives

- To estimate recently proposed models of time-changed Lévy processes using daily stock market excess returns over 1926 - 2006.

- To estimate (filter) the realizations of latent state variables

  -autocorrelations

  -volatility

  taking into account the fat-tailed properties of stock market returns.

Objectives

- To estimate recently proposed models of time-changed Lévy processes using daily stock market excess returns over 1926 - 2006.

- To estimate (filter) the realizations of latent state variables

  -autocorrelations

  -volatility

  taking into account the fat-tailed properties of stock market returns.

Methodological issues

- Density-based methods (MCMC, particle filters) are difficult with Lévy processes, given density functions are not typically known.

  -Li, Wells and Yu (2006): rely on special cases, and SV + i.i.d. jumps

- Approach here is based on ***characteristic functions***, which *are* known for Lévy processes.

Model

$$y_{t+1} = \rho_t y_t + \int_t^{t+\tau_t} ds_t$$

$$V_{t+1} = V_t + \int_t^{t+\tau_t} dV_t$$

$$\rho_{t+1} = \rho_t + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim N(0, \sigma_\rho^2) \ \textit{and} \ \textit{i.i.d.}$$

where y_{t+1} is the daily *excess* return on the CRSP value-weighted index, and

$$ds_t = (\mu_0 + \mu_1 V_t)dt + \left(\rho_{sv}\sqrt{V_t}\,dW_t - \tfrac{1}{2}\rho_{sv}^2 V_t\,dt\right) + (dL_t - \omega V_t\,dt)$$

$$dV_t = (\alpha - \beta V_t)dt + \sigma\sqrt{V_t}\,dW_t$$

Model

$$y_{t+1} = \rho_t y_t + \int_t^{t+\tau_t} ds_t$$

$$V_{t+1} = V_t + \int_t^{t+\tau_t} dV_t$$

$$\rho_{t+1} = \rho_t + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim N(0, \sigma_\rho^2) \ \textit{and} \ \textit{i.i.d.}$$

where y_{t+1} is the daily *excess* return on the CRSP value-weighted index, and

$$ds_t = (\mu_0 + \mu_1 V_t)dt + \left(\rho_{sv}\sqrt{V_t}\,dW_t - \tfrac{1}{2}\rho_{sv}^2 V_t\,dt\right) + (dL_t - \omega V_t\,dt)$$

$$dV_t = (\alpha - \beta V_t)dt + \sigma\sqrt{V_t}\,dW_t$$

Features

- ρ_t is (nonstationary) daily autocorrelation in returns (time-varying coefficients model)
- V_t is underlying conditional variance
- "Leverage effect" via $Corr_t(ds_t, dV_t) = \rho_{sv} < 0$
- Day-of-the-week effects & holidays captured by periodic variation in time horizon τ_t
- **Conditional fat tails** via compensated Lévy process $(dL_t - \omega V_t\,dt)$

  -with instantaneous variance $(1 - \rho_{sv}^2)V_t\,dt$, and

  -stochastic jump intensity: $Prob_t[dL_t = x] \propto k(x)V_t\,dt$

Focus of paper: alternative specifications of dL_t

In Lévy densities $Prob_t[dL_t = x] \propto k(x) V_t \, dt$, the functional form of $k(x)$ is substantially unconstrained.

Focus of paper: alternative specifications of dL_t

In Lévy densities $Prob_t[dL_t = x] \propto k(x) V_t \, dt$, the functional form of $k(x)$ is substantially unconstrained.

Finite-activity jumps (+ diffusion): $\int_{-\infty}^{\infty} k(x)\,dx < \infty$, $k(x) \propto p(x \mid jump)$ (compound Poisson)

SVJ1: diffusion + finite-activity normal jumps

SVJ2: diffusion + jumps from a mixture of normals

DEXP: diffusion + jumps from double exponential

Focus of paper: alternative specifications of dL_t

In Lévy densities $Prob_t[dL_t = x] \propto k(x) V_t \, dt$, the functional form of $k(x)$ is substantially unconstrained.

Finite-activity jumps (+ diffusion): $\int_{-\infty}^{\infty} k(x) dx < \infty$, $k(x) \propto p(x \mid jump)$ (compound Poisson)

SVJ1: diffusion + finite-activity normal jumps

SVJ2: diffusion + jumps from a mixture of normals

DEXP: diffusion + jumps from double exponential

Infinite-activity jumps

VG: diffusion + infinite-activity variance gamma

Y, YY pure-jump model of CGMY (2003)

LS pure-jump log-stable model of Carr & Wu (2003), with negative jumps only

Estimation Methodology: Bates (*RFS*, 2006)

This is a ***hidden Markov*** state space model.

Autocorrelation ρ_t and conditional variance V_t are ***latent*** state variables; not directly observed.

Estimation Methodology: Bates (*RFS*, 2006)

This is a ***hidden Markov*** state space model.
Autocorrelation ρ_t and conditional variance V_t are ***latent*** state variables; not directly observed.

Define

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

as the conditional characteristic function summarizing current knowledge of latent state variables (ρ_t, V_t), given past data $\boldsymbol{Y}_t = (y_1, \ldots, y_t)$.

Estimation Methodology: Bates (*RFS*, 2006)

This is a ***hidden Markov*** state space model.
Autocorrelation ρ_t and conditional variance V_t are ***latent*** state variables; not directly observed.

Define

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

as the conditional characteristic function summarizing current knowledge of latent state variables (ρ_t, V_t), given past data $\boldsymbol{Y}_t = (y_1, \ldots, y_t)$.

Filtration methodology of Bates (*RFS*, 2006) :

- Recursively update what is known about (ρ_t, V_t), as summarized by $G_{t|t}(\cdot)$, using the latest observation y_{t+1} and the equivalent of Bayes' law for characteristic functions.

- The log-likelihood function used in MLE is a by-product of the filtration.

- Approach does involve numerical integrations, and approximation issues.

Bayes' law for characteristic functions

If $F(i\Phi, i\psi) = E[e^{i\Phi y + i\psi x}]$ is the joint characteristic function (CF) of (observed) *y* and (latent) *x*, then the CF of *x conditional* upon observing *y* is

$$G_{x|y}(i\psi) = \frac{1}{2\pi p(y)} \int F(i\Phi, i\psi) e^{-i\Phi y} d\Phi \qquad \textbf{(1)}$$

where

$$p(y) = \frac{1}{2\pi} \int F(i\Phi, 0) e^{-i\Phi y} d\Phi \qquad \textbf{(2)}$$

is the marginal density of *y*.

Recursive filtration for hidden Markov models

For "semi-affine" processes such as the model above:

1) The transition CF of y_{t+1} and $x_{t+1} \equiv (\rho_{t+1}, V_{t+1})$ *conditional* on knowing (ρ_t, V_t) is analytic.

2) *Given* a joint CF $G_{t|t}(\cdot)$ summarizing current knowledge of (ρ_t, V_t) given past data $\boldsymbol{Y}_t$, the joint characteristic function $F(i\Phi, i\psi \mid \boldsymbol{Y}_t)$ of (y_{t+1}, x_{t+1}) conditional on $\boldsymbol{Y}_t$ is also analytic.

3) Given the latest datum y_{t+1}:
   - -Can numerically compute $p(y_{t+1} \mid \boldsymbol{Y}_t)$ from (2) (used in ML);
   - -Can update $G_{t|t}(\bullet) \rightarrow G_{t+1|t+1}(\bullet)$ using (1);
   - -Can compute updated moments of (ρ_{t+1}, V_{t+1}) from $G_{t+1|t+1}(\bullet)$.

4) Implementation (AML): moment-matching *approximate* $\hat{G}_{t+1|t+1}(\bullet)$:
   - -Start with priors: $\rho_t \mid \boldsymbol{Y}_t \sim N(\hat{\rho}_{t|t}, W_{t|t})$, $V_t \mid \boldsymbol{Y}_t \sim \Gamma(\hat{V}_{t|t}, P_{t|t})$ and independent
   - -Compute exact posterior moments given y_{t+1}
   - -Approximate posterior distributions by moment-matching normal & gamma

Generalizes the *robust Kalman filtration* approach of Masreliez (1975)

Bates (*RFS*, 2006): numerically stable, and performs well for estimation and filtration

Here: five numerical univariate integrations at each time step

Data: Daily CRSP value-weighted (cum-dividend) log *excess* returns, over January 2, 1926 - December 29, 2006.

October 19, 1987:	-20.0%	Nov. 6, 1931:	+10.2%
October 29, 1929:	-12.6%	Sept. 21, 1932:	+10.3%
October 28, 1929:	-11.9%	Oct. 30, 1929:	+11.3%
Nov. 6, 1929:	-10.1%		

Data: Daily CRSP value-weighted (cum-dividend) log *excess* returns, over January 2, 1926 - December 29, 2006.

October 19, 1987:	-20.0%	Nov. 6, 1931:	+10.2%
October 29, 1929:	-12.6%	Sept. 21, 1932:	+10.3%
October 28, 1929:	-11.9%	Oct. 30, 1929:	+11.3%
Nov. 6, 1929:	-10.1%		

Results

- Day-of-the-week effects
- Overall fit of various models
- Tail properties
- Autocorrelation filtration
- Volatility filtration

Day-of-the-week effects

Effective length of a business day, *relative* to 1-day Wednesday returns

		1926-2006			1957-2006		
#days	Description	NOBS	estimate	std. error	NOBS	estimate	std. error
1	Monday close → Tuesday close	3831	1.02	(.04)	2381	1.03	(.04)
1	**Tuesday → Wednesday**	**4037**	**1**		**2519**	**1**	
1	Wednesday → Thursday	3998	.94	(.03)	2479	.95	(.04)
1	Thursday → Friday	3924	.93	(.03)	2434	.90	(.04)
1	Friday → Saturday (1926-52)	1141	.43	(.02)			
2	Saturday → Monday (1926-52)	1120	1.05	(.05)			
2	Weekday holiday	341	1.25	(.11)	142	1.09	(.13)
2	Wednesday holiday in 1968	22	.73	(.33)	22	.79	(.33)
3	Weekend and/or holiday[a]	2755	1.10	(.04)	2322	1.15	(.05)
4	Holiday weekend	343	1.58	(.14)	281	1.61	(.15)
5	Holiday weekend	6	1.31	(1.00)	4	2.08	(2.61)
		21518			12584		

[a]Includes one weekday holiday (August 14 - 17, 1945)

Alternate specifications of jump intensities

$$Prob_t[dL_t = x] \propto k(x) V_t \, dt$$

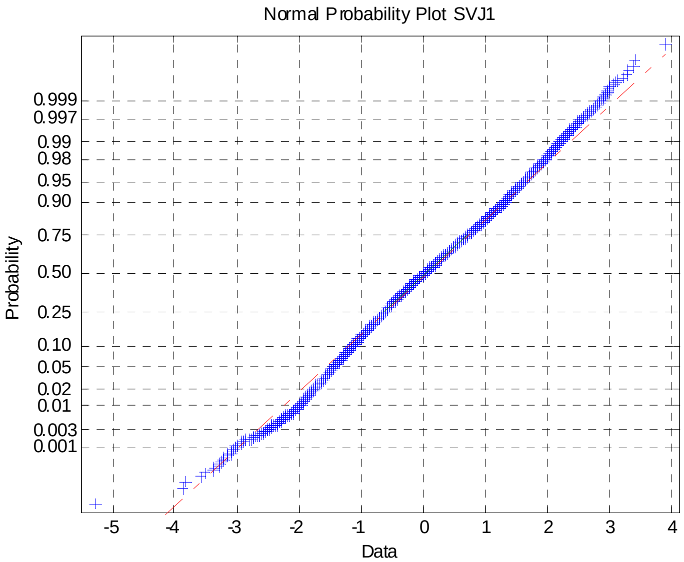
SVJ2: $k(x) \propto \sum_{i=1}^{2} \frac{\lambda_i}{\sqrt{2\pi\delta_i^2}} \exp\left[-\frac{(x-\bar{\gamma}_i)^2}{2\delta_i^2}\right]$ ln L = 75,049.07

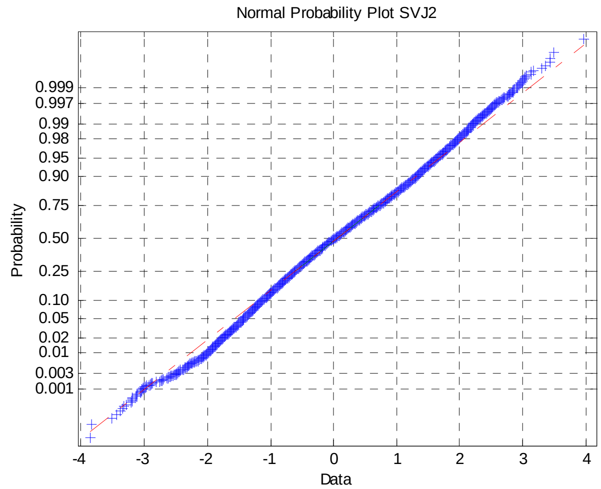
(SVJ1 if $\lambda_2 = 0$) ln L = 75,044.60

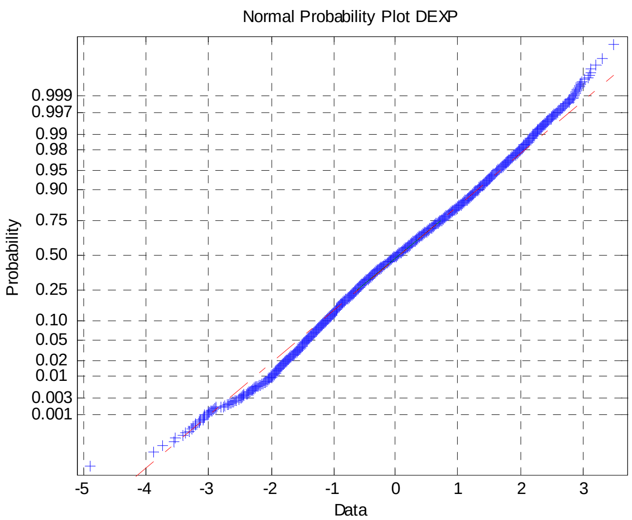
CGMY (2003): $k(x) = \begin{cases} C_n e^{-G|x|} |x|^{-1-Y_n} \; for \; x < 0 \\ C_p e^{-M|x|} |x|^{-1-Y_p} \; for \; x > 0 \end{cases}$

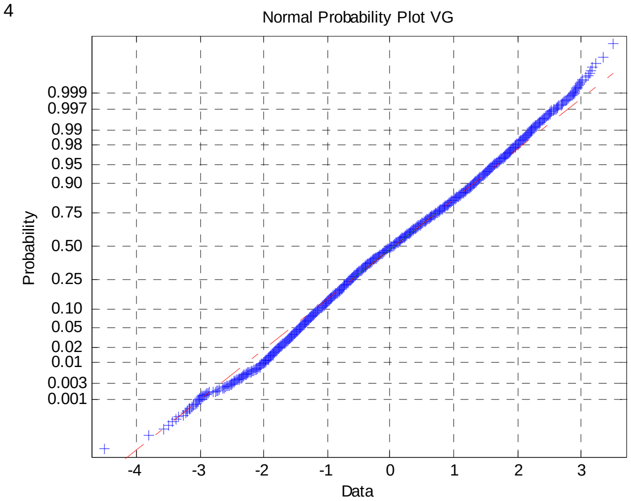
where C_n, C_p, G, $M \geq 0$ and $Y_p, Y_n \leq 2$.

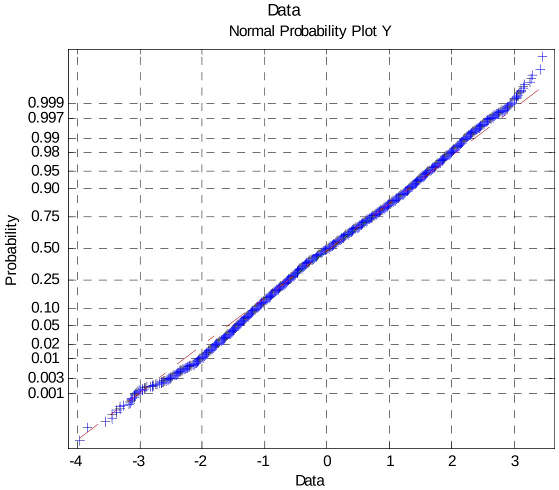
Special cases:			
	$Y_n = Y_p = -1$	double exponential	ln L = 75,047.62
	$Y_n = Y_p = 0$	variance gamma	ln L = 75,049.48
	$Y_n = Y_p = 2$	Brownian motion (Heston (1993))	
	$G = 0 = C_p$	Log-stable model (Carr-Wu (2003))	ln L = 75,005.53
	$Y_n = Y_p$		ln L = 75,050.12
		General YY model	**ln L = 75,052.12**

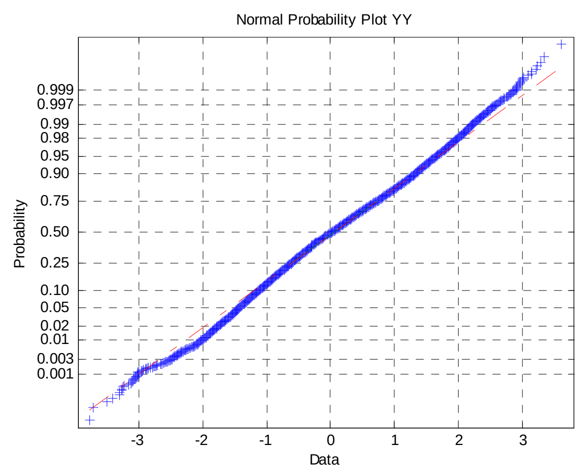












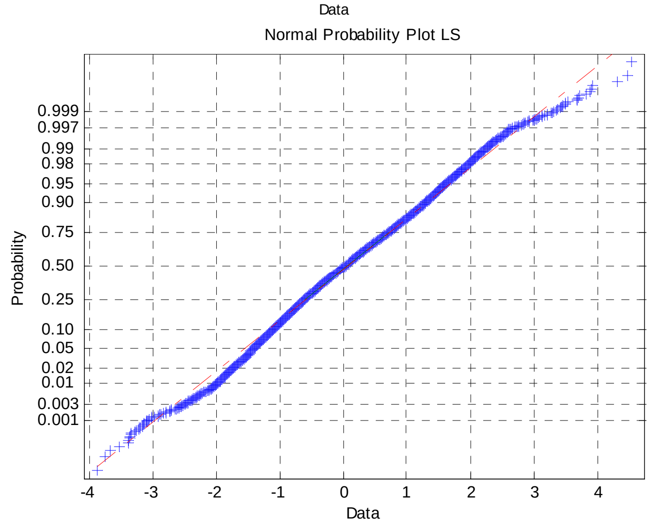
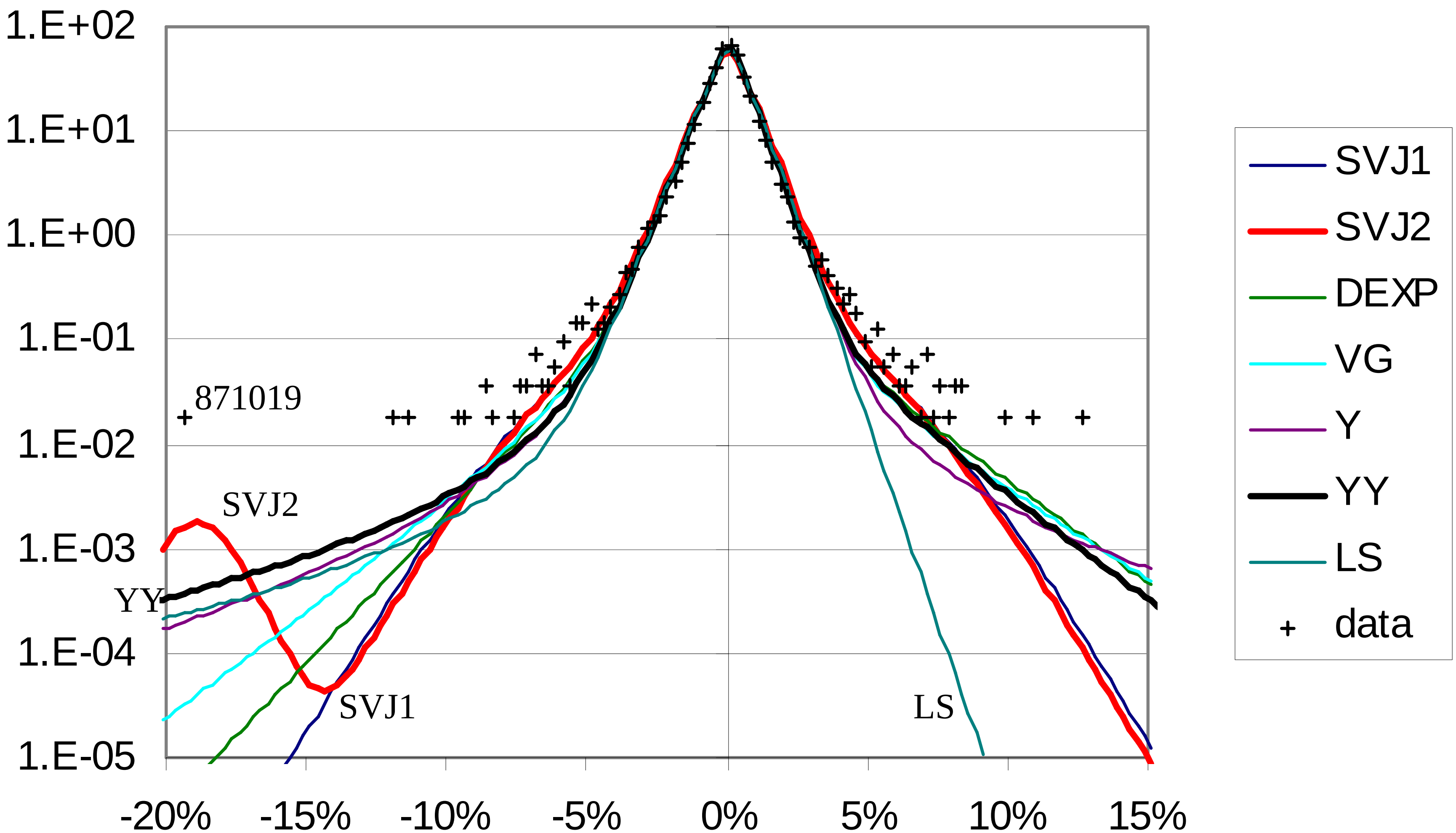


Figure 1. Normal probability plots for the normalized returns $z_{t+1} \equiv N^{-1}[CDF(y_{t+1} \mid \boldsymbol{Y}_t, \hat{\boldsymbol{\Theta}})]$, for different models.

Diagonal line: theoretical quantiles conditional upon correct specification

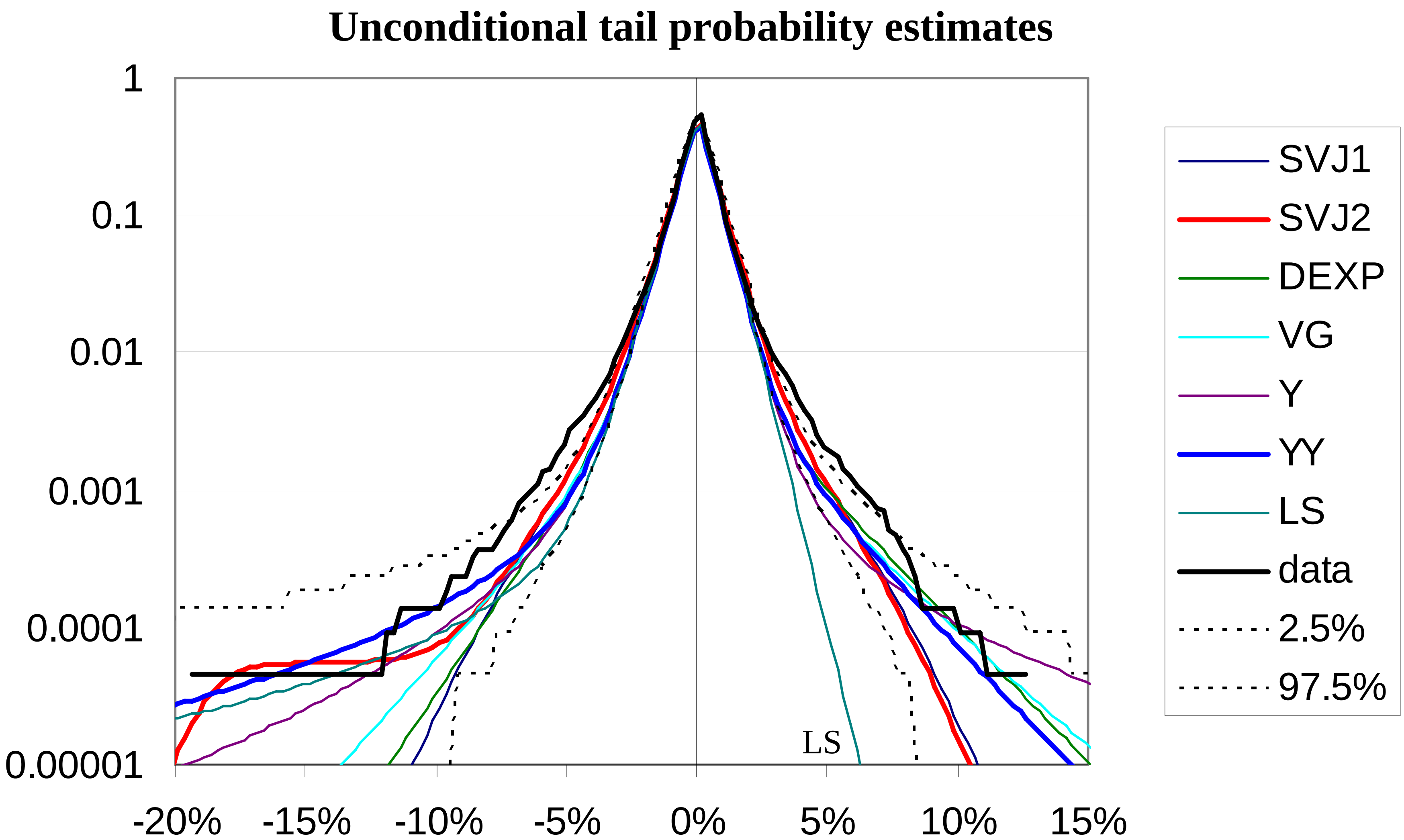
+: Empirical quantiles

Unconditional probability density function estimates



Data: from a histogram of daily excess returns' residuals (after autocorrelation correction)

Model-based estimates: mixture of horizon-dependent unconditional PDF's



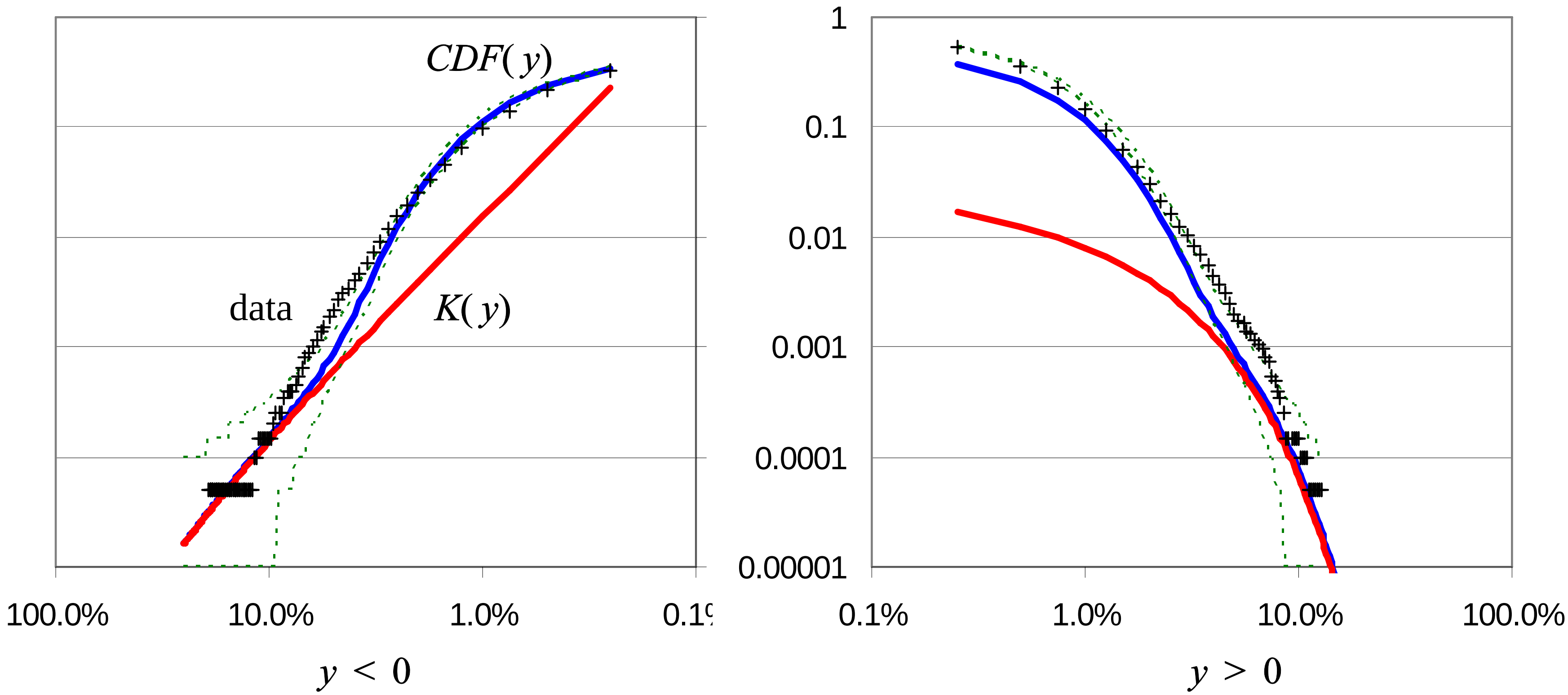
Parametric models:

- Capture extreme tails, and $|y| < 2\%$ (2 SD's)
- Underestimate frequency of returns of [3%, 7%] in magnitude.

(95% confidence interval based on 1000 simulated sample paths, using YY parameter estimates)

A power law?

Distribution functions approach *unconditional* intensity functions $K(y) \equiv (1-\rho_{sv}^2)\frac{\alpha}{\beta}\tau\int_{-\infty}^{y}k(x)\,dx$ as $y \to -\infty$; but only for $|y| > 5\%$ (5 standard deviations).



Unconditional tail probabilities and intensity functions versus $|y|$; log scales on both axes

Data: Excess returns' residuals with estimated time horizons of ≈ 1 day (20,004 observations).

Autocorrelation filtration

Given conditional normal approximation, autocorrelation filtration is the *robust Kalman filtration* of Masreliez (1975):

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

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

Autocorrelation filtration

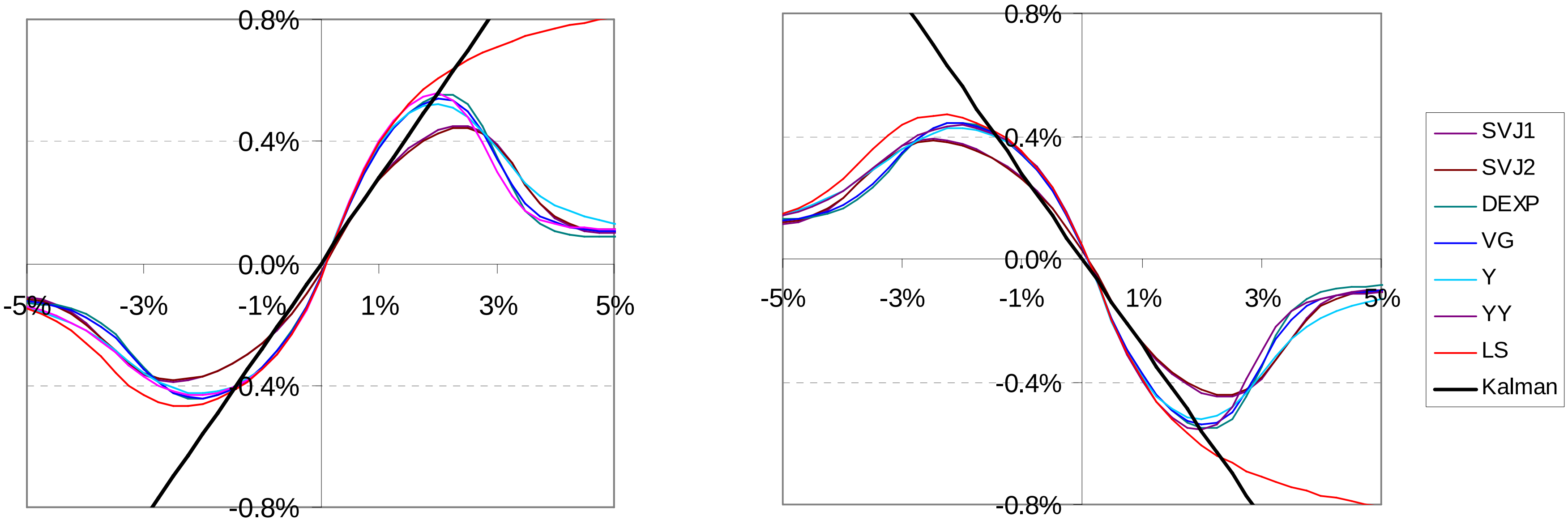
Given conditional normal approximation, autocorrelation filtration is the *robust Kalman filtration* of Masreliez (1975):

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

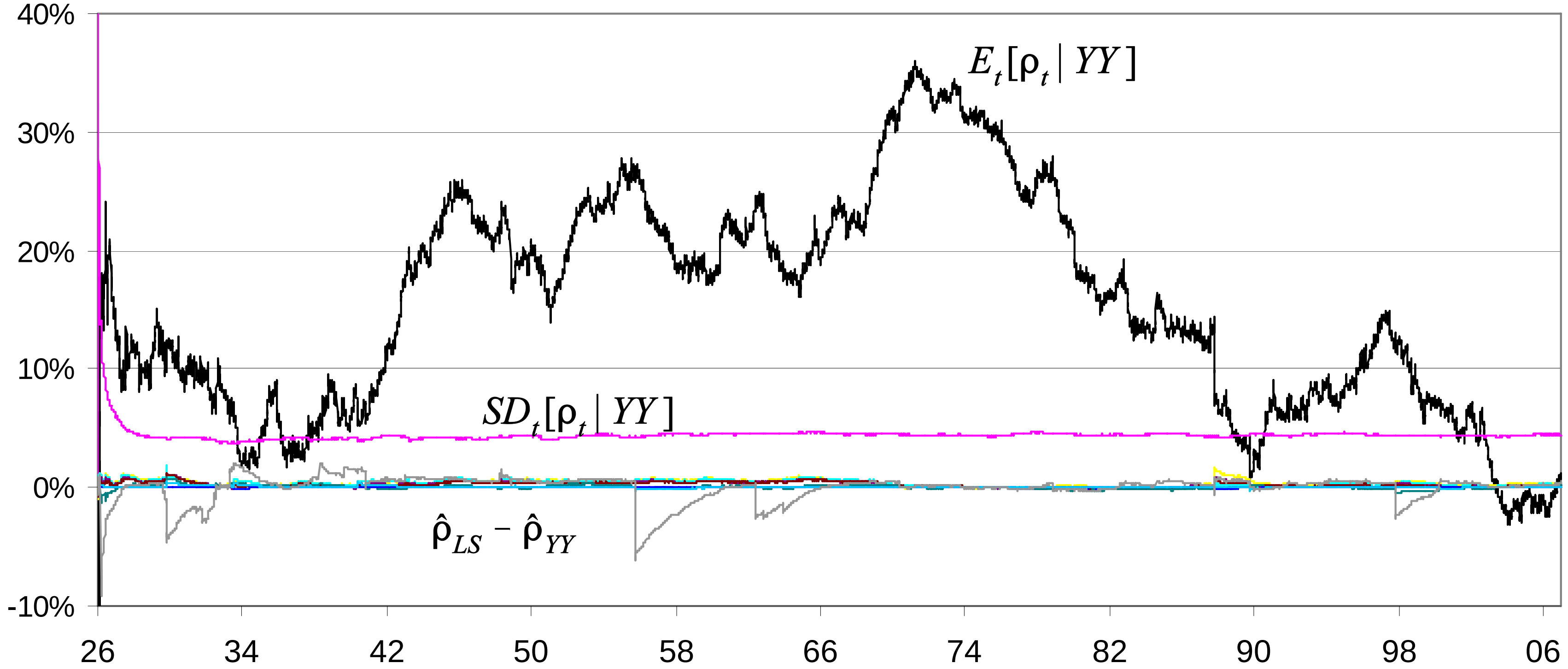
$$W_{t+1|t+1} = \sigma_\rho^2 + (y_t W_{t|t})^2 \frac{\partial^2 \ln p(y_{t+1} | \boldsymbol{Y}_t)}{\partial y_{t+1}^2}$$

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

With $y_t = +1\%$ With $y_t = -1\%$

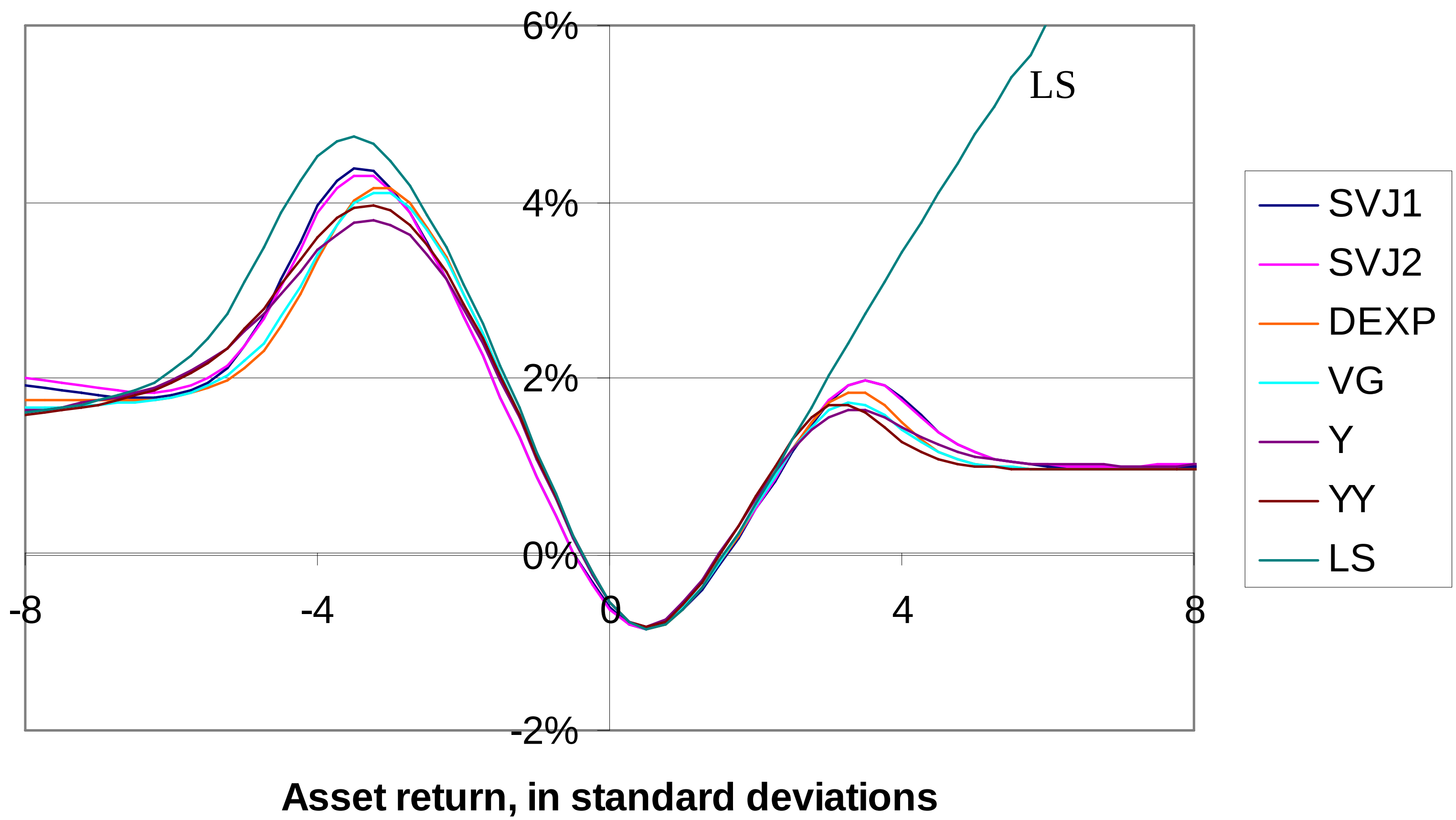


Daily autocorrelation estimates



Autocorrelation estimates $\hat{\boldsymbol{\rho}}_{t|t}$ from YY model, conditional standard deviations, and autocorrelation estimates' divergences from YY estimates for other models

Volatility filtration



News impact curves: Revision in volatility estimate ***conditional*** upon observing a standardized return.

-Quite different from standard GARCH news impact curves:

(exception: Maheu & McCurdy, *JF* 2004)

Hentschel (1995): News impact curves from various GARCH models

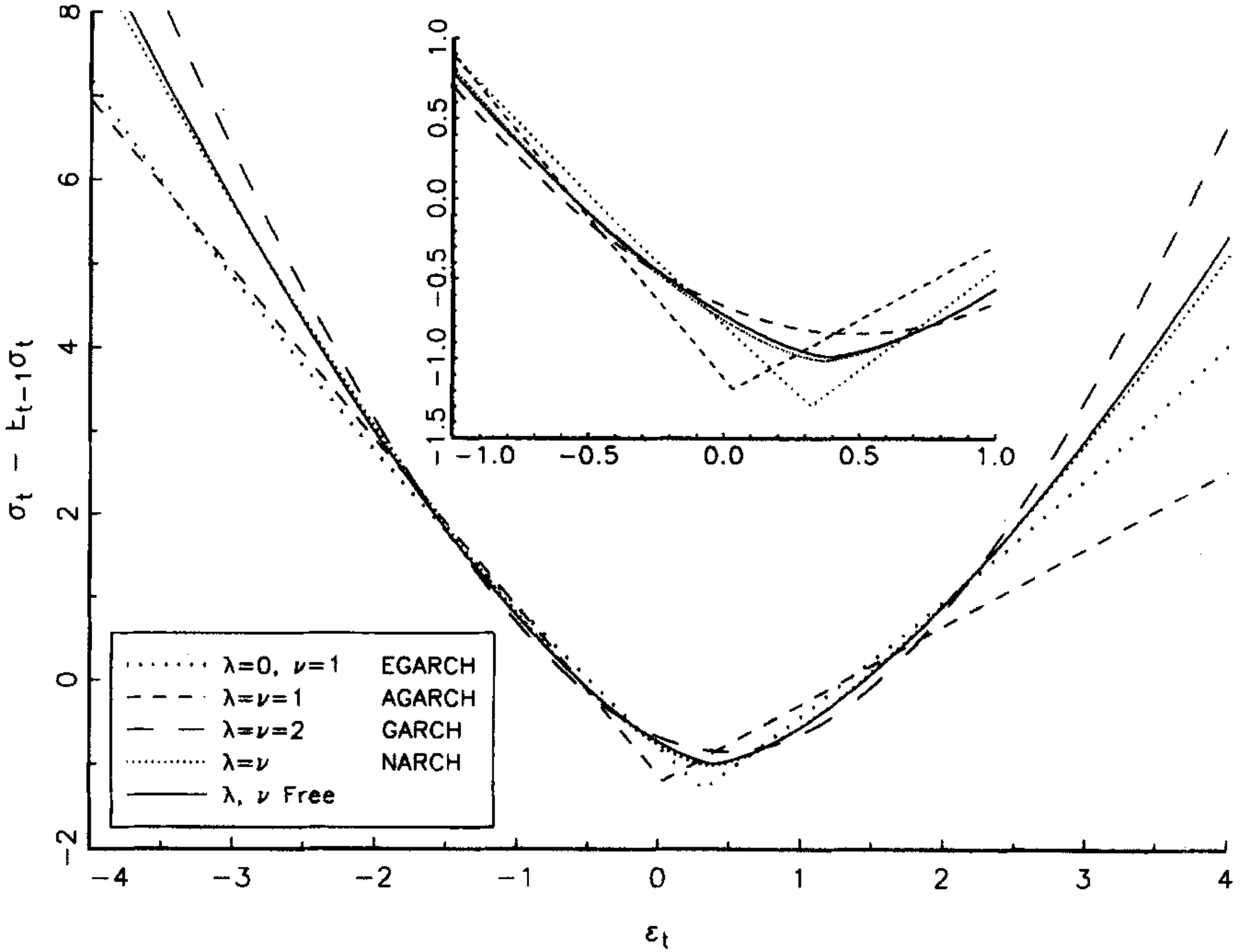


Fig. 3. News impact curves.

Figure 5: Volatility estimates (YY model), associated conditional standard deviations, and deviations from YY estimates for other models

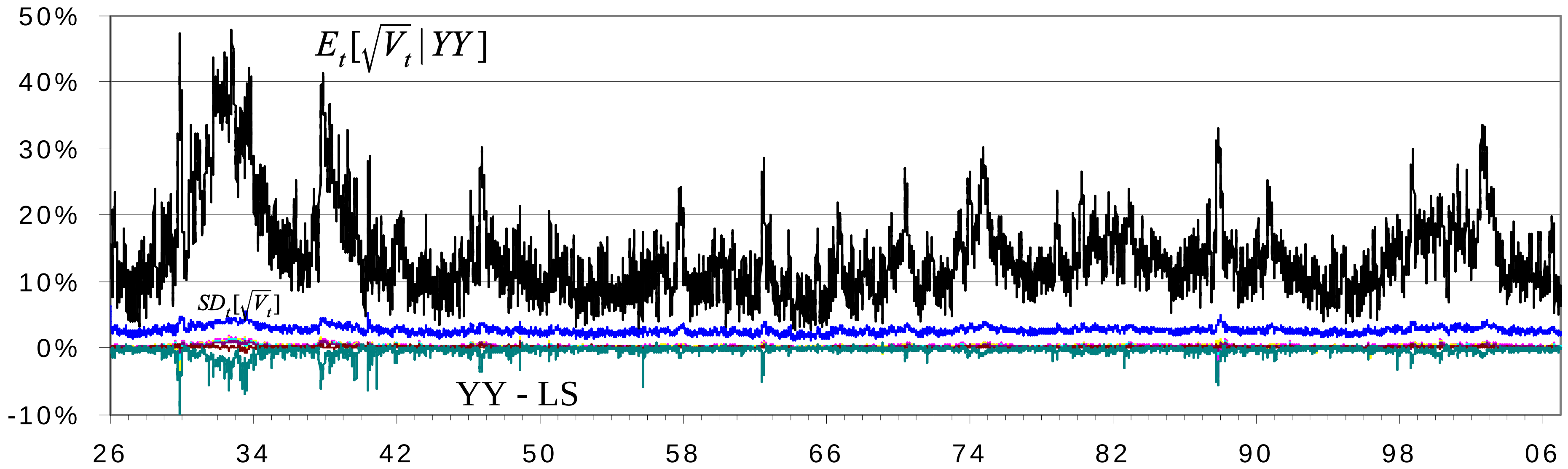
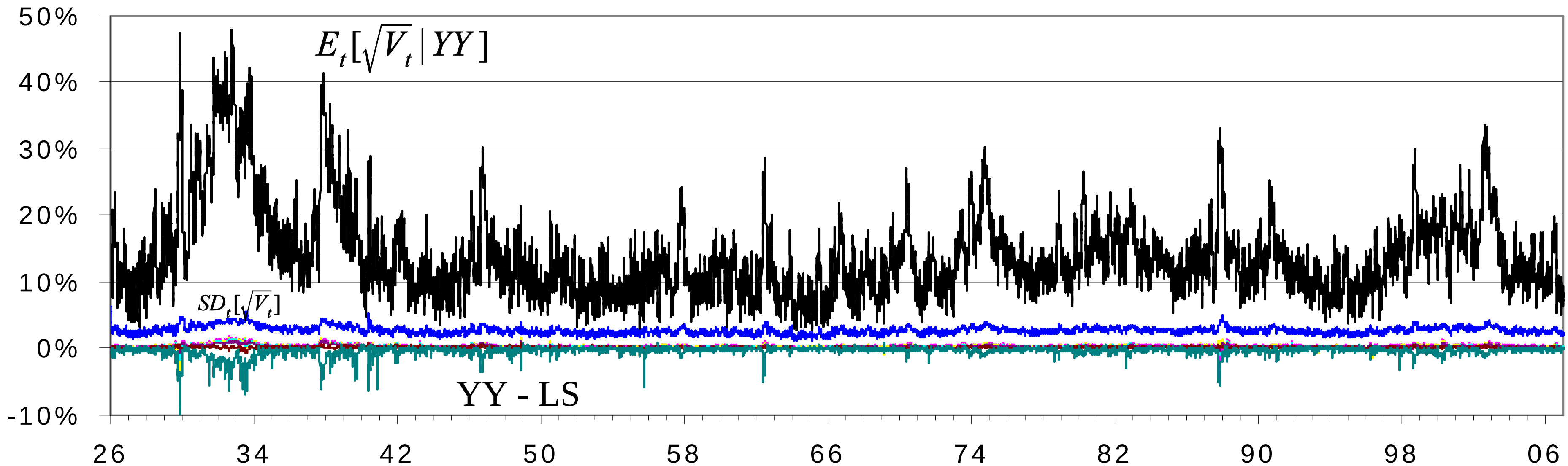


Figure 5: Volatility estimates (YY model), associated conditional standard deviations, and deviations from YY estimates for other models



Graph illustrates longer-term volatility dynamics *not* captured by the 1-factor SV model

Volatility process: $dV_t = (\alpha - \beta V_t)dt + \sigma\sqrt{V_t}\,dW_t$

	$\sqrt{\alpha/\beta}$	HL (mths)	σ	ρ_{sv}
1926-2006	.159 (.008)	2.1 (.2)	.362 (.019)	-.572 (.031)
1957-2006	.129 (.006)	1.9 (.2)	.275 (.014)	-.653 (.030)

Summary and conclusions

- AML filtration/estimation methodology useful for time-changed Lévy processes

- Substantial and nonstationary autocorrelation needs to be taken into account in time series models of stock market returns

- Alternate fat-tailed models fit about the same
  - -SVJ2: ‘87 crash as a unique outlier
  - -CGMY: More parsimonious; but nearly log-stable model of downside risk

- *All* of these fat-tailed distributions imply that outliers should be down-weighted, when updating volatility and autocorrelation estimates
  - -Not standard practice with GARCH models

- 1-factor SV model clearly inadequate

Summary and conclusions

- AML filtration/estimation methodology useful for time-changed Lévy processes

- Substantial and nonstationary autocorrelation needs to be taken into account in time series models of stock market returns

- Alternate fat-tailed models fit about the same
  - -SVJ2: ‘87 crash as a unique outlier
  - -CGMY: More parsimonious; but nearly log-stable model of downside risk

- *All* of these fat-tailed distributions imply that outliers should be down-weighted, when updating volatility and autocorrelation estimates
  - -Not standard practice with GARCH models

- 1-factor SV model clearly inadequate

Extensions/work in progress

- 2-factor SV specifications

- Spline-based “semi-nonparametric” specifications of Lévy density $k(x)V_t$

- Option pricing implications