Estimating Option Pricing Models Using a Characteristic Function-Based Linear State Space Representation

H. Peter Boswijk
(with Roger Laeven and Evgenii Vladimirov)

University of Amsterdam
Amsterdam School of Economics

tinbergen institute

EEA-ESEM Congress 2022, Bocconi University, Milano, August 24, 2022

Introduction

- Option price (panel) data contain valuable information about (risk-neutral) distribution and dynamics of underlying asset.
- Econometric methods for parametric option pricing models (with latent state vector):
  - (penalized) NLS: Bakshi, Chao & Chen (1997); Andersen, Fusari & Todorov (2015);
  - EMM: Chernov & Ghysels (2000); Andersen, Benzoni & Lund (2002);
  - implied-state GMM methods: Pan (2002); Boswijk, Laeven, Lalu & Vladimirov(2021);
  - MCMC: Eraker (2004); Eraker, Johannes & Polsen (2003);
  - particle filtering: Johannes, Polson & Stroud (2009); Bardgett, Gourier & Leippold (2019).
- Such methods typically compare observed (transformed) option prices with theoretical (parametric) counterparts, using Fourier- or simulation-based methods.
- This involves using cross-sectional information to back out latent states;
  - exceptions: Feunou & Okou (2018), Aït-Sahalia, Li & Li (2021).

Example: Heston (1993) model

$$\begin{aligned} \mathrm{d}S_t &= rS_t\mathrm{d}t + \sqrt{v_t}S_t\mathrm{d}W_{1t}, \\ \mathrm{d}v_t &= \kappa(\overline{v} - v_t)\mathrm{d}t + \sigma\sqrt{v_t}\mathrm{d}W_{2t}, \qquad \mathrm{d}W_{1t}\mathrm{d}W_{2t} = \rho\mathrm{d}t. \end{aligned}$$

- Calibration approach to get parameters $\theta = (\kappa, \overline{v}, \sigma, \rho)$ and latent v_t:

$$(\{\widehat{v}_t\}_{t=1,\dots,T}, \widehat{\theta}) = \underset{\theta, \{v_t\}_{t=1,\dots,T}}{\operatorname{argmin}} \sum_{t=1}^{T}\sum_{i=1}^{N}\left(O_t^{market}(K_i) - O_t^{model}(K_i;\theta, v_t)\right)^2.$$

- Evaluation $O_t^{model}(K_i;\theta)$ non-trivial: FFT/COS pricing methods in practice;
- $T+4$ parameters to calibrate;
- $\Rightarrow$ computationally intensive, time-series dependence not exploited.

This paper

Approach:

1. Transform option prices to (non-parametric) empirical characteristic function (based on Carr & Madan, 2001; Todorov, 2019).
2. Assume affine jump diffusion (AJD; Duffie, Pan & Singleton, 2000) for state vector, implying log-conditional characteristic function that is affine in state vector.
3. Couple steps 1 and 2 to obtain measurement / observations equation in state space model; state transition equation is time-discretised SDE of AJD.
4. Estimate model parameters by Gaussian QML based on (collapsed) Kalman filter.

Properties:

- Does not require numerical option pricing methods;
- Exploits both cross-sectional and time-series dependence;
- Based on the standard Kalman filter, computationally fast;
- Exploits all probabilistic information via characteristic function.

Payoff spanning

- Breeden & Litzenberger (1978): put option with strike K and time to expiration $\tau = T - t$:

$$P_t(K, T) = e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[\max(K - S_T, 0)|\mathcal{F}_t] = e^{-r\tau}\int_0^K (K - x)q_t(x, T)\mathrm{d}x,$$

implies risk-neutral density

$$q_t(x, T) = e^{r\tau}P_t''(x, T) = e^{r\tau}\left.\frac{\partial^2 P_t(K, T)}{\partial K^2}\right|_{K=x}.$$

- Carr and Madan (2001): European-style derivative with payoff function $g(S_T)$:

$$\begin{aligned} e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[g(S_T)|\mathcal{F}_t] &= e^{-r\tau}g(F_t) + \int_0^{F_t} g''(K)P_t(K, T)\mathrm{d}K + \int_{F_t}^{\infty} g''(K)C_t(K, T)\mathrm{d}K \\ &= e^{-r\tau}g(F_t) + \int_0^{\infty} g''(K)O_t(K, T)\mathrm{d}K. \end{aligned}$$

with futures price $F_t = \mathbb{E}^{\mathbb{Q}}[S_T|\mathcal{F}_t]$ and OTM option prices $O_t(K, T)$.

CCF option spanning

- Todorov (2019): Conditional characteristic function (CCF) of log-returns:

$$\phi_t(u,\tau) = e^{-r\tau}\mathbb{E}^{\mathbb{Q}}\left[e^{\mathrm{i}u\log(F_T/F_t)} \middle| \mathcal{F}_t \right]$$

corresponds to payoff function $g(x) = e^{\mathrm{i}u\log(x/F_t)}$, hence

$$\phi_t(u,\tau) = e^{-r\tau} - (u^2 + \mathrm{i}u)\int_0^\infty \frac{1}{K^2} e^{\mathrm{i}u(\log K - \log F_t)} O_t(T,K)\mathrm{d}K,$$

with $\phi_t(u,\tau) \in \mathbb{C}$ and $u \in \mathbb{R}$.

- Result is exact and model independent; requires continuum of strikes, but can be approximated by Riemann sum given observed option prices $\widehat{O}_t(K_j, T), j = 1, \ldots, n$, yielding:

$$\widehat{\phi}_t(u,\tau) = \phi_t(u,\tau) + \zeta_t(u,\tau).$$

$\zeta_t(u,\tau)$ reflects truncation, discretisation and measurement errors ($\widehat{O}_t - O_t$).

Affine jump diffusion

- State vector $X_t \in \mathbb{R}^d$ satisfies, under $\mathbb{Q}$:

$$\mathrm{d}X_t = \mu(X_t;\theta)\mathrm{d}t + \sigma(X_t;\theta)\mathrm{d}W_t + J_t\mathrm{d}N_t,$$

  where $\mu(\cdot;\theta)$, $\sigma(\cdot;\theta)\sigma(\cdot;\theta)'$ and jump intensity $\lambda(\cdot;\theta)$ are affine functions.
- X_t includes observable and latent variables; e.g., stochastic volatility: $X_t = (\log F_t, v_t)'$.
- Implies CCF

$$\phi_{X_t}(\mathbf{u},\tau) = e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[e^{\mathrm{i}\mathbf{u}\cdot X_{t+\tau}}|\mathcal{F}_t] = e^{\alpha(\mathbf{u},\tau;\theta)+\beta(\mathbf{u},\tau;\theta)\cdot X_t}, \qquad \mathbf{u}\in\mathbb{R}^d,$$

  with $\alpha(\mathbf{u},\tau;\theta)\in\mathbb{C},\ \beta(\mathbf{u},\tau;\theta)\in\mathbb{C}^d$ solutions to the ODE system.
- Hence marginal CCF for log-returns from $\mathbf{u_1} := (u,0,\ldots,0)'$.
- Large model class: Heston (1993), Pan (2002), Duffie *et al.* (2000), Andersen *et al.* (2017), *inter alia*.

Marrying two CCFs

- For true model with continuum of strikes:

$$\underbrace{\log \phi_t(u,\tau)}_{\text{model-free}} = \underbrace{\alpha(\mathbf{u_1},\tau;\theta) + \beta(\mathbf{u_1},\tau;\theta)\cdot X_t,}_{\text{model-dependent}} \quad u \in \mathbb{R}.$$

- Due to observation errors and approximation with discrete set of strikes:

$$\log \widehat{\phi}_t(u,\tau) = \alpha(\mathbf{u_1},\tau;\theta) + \beta(\mathbf{u_1},\tau;\theta)\cdot X_t + \xi_t(u,\tau), \quad u \in \mathbb{R}.$$

  Continuum of complex-valued linear measurement equations;
  - if X_t were observable, could estimate θ by (continuum) non-linear least squares;
  - since X_t is (partially) latent $\Rightarrow$ Kalman filter-based QML estimation.
- Note that $\xi_t(u,\tau) = \zeta_t(u,\tau)/\phi_t(u,\tau)$ + "log-linearisation error".

State transition equation

- $X_t = (w_t', x_t')'$, where w_t is observable and x_t contains latent states.
- If model is also affine under $\mathbb{P}$, then $\mathbb{E}^{\mathbb{P}}[x_{t+1}|\mathcal{F}_t]$ and $\text{Var}^{\mathbb{P}}(x_{t+1}|\mathcal{F}_t)$ are affine in x_t:

  $$\mathbb{E}^{\mathbb{P}}[x_{t+1}|\mathcal{F}_t] = c_t + Tx_t, \qquad \text{Var}^{\mathbb{P}}(x_{t+1}|\mathcal{F}_t) = Q(x_t),$$

  with effect of w_t absorbed in c_t. Hence transition equation

  $$x_{t+1} = c_t + Tx_t + \eta_{t+1}, \qquad \mathbb{E}^{\mathbb{P}}[\eta_{t+1}|\mathcal{F}_t] = 0, \quad \text{Var}^{\mathbb{P}}(\eta_{t+1}|\mathcal{F}_t) = Q(x_t).$$

- Functional form of $c_t(\theta)$, $T(\theta)$, $Q(\cdot;\theta)$ can sometimes be obtained in closed form; in general, by numerical methods.
- Different parameter values under $\mathbb{P}$ and $\mathbb{Q}$ $\Rightarrow$ need to extend θ by risk-premia.

Measurement equation

- Turn complex function-valued measurement equation

$$\log \widehat{\phi}_t(u,\tau) = \alpha(\mathbf{u_1},\tau;\theta) + \beta_w(\mathbf{u_1},\tau;\theta)\cdot w_t + \beta_x(\mathbf{u_1},\tau;\theta)\cdot x_t + \xi_t(u,\tau),$$

  into real vector-valued measurement equation by:
  - stacking real and imaginary parts;
  - stacking these 2×1 vector equations for (u_i,τ_j), $i=1,\ldots,q; j=1,\ldots,k$.
- Hence measurement equation is

$$y_t = d_t + Z_t x_t + r_{t,n} + \varepsilon_t,$$

  where $r_{t,n}$ is a remainder term (discussed below), and, e.g.,

$$y_t = \begin{pmatrix} y_{1t} \\ \vdots \\ y_{kt} \end{pmatrix}, \qquad y_{jt} = \begin{pmatrix} y_{1j,t} \\ \vdots \\ y_{qj,t} \end{pmatrix}, \qquad y_{ij,t} = \begin{pmatrix} \mathrm{Re}\log\widehat{\phi}_t(u_i,\tau_j) \\ \mathrm{Im}\log\widehat{\phi}_t(u_i,\tau_j) \end{pmatrix}.$$

  Effect of observed states w_t absorbed in d_t.

State space representation

Assumptions

1. X_t is Markov, affine jump diffusion, with "admissible" $\theta \in \Theta$ under $\mathbb{P}$ and $\mathbb{Q}$.
2. Observation errors $\zeta_t(\tau, m) = \widehat{O}_t(\tau, m) - O_t(\tau, m)$ are $\mathcal{F}_t$-conditionally independent along τ, $m = \log(K/F_t)$ and t,
$$\mathbb{E}[\zeta_t(\tau, m)|\mathcal{F}_t] = 0, \qquad \mathrm{Var}(\zeta_t(\tau, m)|\mathcal{F}_t) = \sigma_{\varkappa}^2 \kappa_t^2(\tau, m)\nu_t^2(\tau, m),$$
with $\kappa_t(\tau, m)$ and $\nu_t(\tau, m)$ the Black-Scholes implied volatility and vega, resp.
3. $\mathbb{E}^{\mathbb{Q}}[F_{t+\tau}^2|\mathcal{F}_t] < \infty$ and $\mathbb{E}^{\mathbb{Q}}[F_{t+\tau}^{-2}|\mathcal{F}_t] < \infty$ for $\tau > 0$.
4. Log-moneyness grid $\underline{m} = m_1 < \ldots < m_n = \overline{m}$ satisfies
$$e^{m_1} = \mathcal{O}(n^{-\underline{\alpha}}), \quad e^{m_n} = \mathcal{O}(n^{\overline{\alpha}}), \qquad \underline{\alpha} > 0, \overline{\alpha} > 0,$$
$$\eta \Delta m \leq \inf_{2 \leq j \leq n} (m_j - m_{j-1}) \leq \sup_{2 \leq j \leq n} (m_j - m_{j-1}) \leq \Delta m,$$
where $\lim_{n \to \infty} \Delta m = 0$ and $\eta \in (0, 1)$.

State space representation

Proposition

Under Assumptions 1–4,

$$y_t = d_t + Z_t x_t + r_{t,n} + \varepsilon_t, \qquad \mathbb{E}[\varepsilon_t \varepsilon_t' | \mathcal{F}_t] = H_t,$$

$$x_{t+1} = c_t + T x_t + \eta_{t+1}, \qquad \mathbb{E}[\eta_{t+1} \eta_{t+1}' | \mathcal{F}_t] = Q(x_t),$$

where:

- $r_{t,n} = \mathcal{O}_P(n^{-2(\underline{\alpha} \wedge \overline{\alpha})} \vee n^{-1} \log n)$ and $\varepsilon_t = \mathcal{O}_P\left(\sqrt{n^{-1} \log n}\right)$.
- The errors satisfy
  - $\mathbb{E}[\varepsilon_t \varepsilon_s'] = 0$, $\mathbb{E}[\eta_t \eta_s'] = 0$ for all $t \neq s$;
  - $\mathbb{E}[\varepsilon_t \eta_s'] = 0$ for all t and s;
  - $\mathbb{E}[\varepsilon_t x_t'] = 0$, $\mathbb{E}[\eta_{t+1} x_t'] = 0$ for all t.
- H_t is $\mathcal{F}_t$-measurable and estimable, involving a single unknown parameter $\sigma_{\varkappa}^2$.

Collapsed Kalman filter

- Recall

$$y_t = d_t + Z_t x_t + \varepsilon_t, \qquad \mathbb{E}[\varepsilon_t \varepsilon_t' | \mathcal{F}_t] = H_t,$$

  where $\dim(y_t) \gg \dim(x_t) = d$.

- Collapsed Kalman filter of Jungbacker & Koopman (2015):

$$y_t \quad \mapsto \quad A_t y_t = \begin{bmatrix} A_t^* \\ A_t^+ \end{bmatrix} y_t = \begin{pmatrix} y_t^* \\ y_t^+ \end{pmatrix} = \begin{pmatrix} d_t^* \\ d_t^+ \end{pmatrix} + \begin{pmatrix} x_t \\ 0 \end{pmatrix} + \begin{pmatrix} \varepsilon_t^* \\ \varepsilon_t^+ \end{pmatrix},$$

  with $A_t^* = (Z_t' H_t^- Z_t)^{-1} Z_t' H_t^-$, $A_t^+ = L_t H_t^- (I_p - Z_t A_t^*)$ and $\mathrm{cov}(\varepsilon_t^*, \varepsilon_t^+) = 0$.

- H_t might be singular, hence generalised inverse H_t^-.
- If A_t is non-singular, all information about the state vector x_t is contained in $y_t^* \in \mathbb{R}^d$ (same dimension as x); hence proceed with

$$\begin{aligned} y_t^* &= d_t^* + x_t + \varepsilon_t^*, & \mathbb{E}[\varepsilon_t^* \varepsilon_t^{*\prime} | \mathcal{F}_t] &= H_t^*, \\ x_{t+1} &= c_t + T x_t + \eta_{t+1}, & \mathbb{E}[\eta_{t+1} \eta_{t+1}' | \mathcal{F}_t] &= Q(x_t). \end{aligned}$$

QML estimation

- If $Q(\cdot)$ does not vary with x_t, Kalman filter delivers minimum MSE predictor $\widehat{x}_{t|t-1}$ of x_t.
- This is the basis for consistency of the Gaussian QML estimator.
- For the heteroskedastic case, de Jong (2000) and others have suggested a modified Kalman filter, replacing $Q(x_t)$ by $Q(\widehat{x}_{t|t-1})$ in the Kalman filter recursions.
- Athough this methods invariably works well in Monte Carlo simulations, no formal asymptotic justification is available.
- Alternatively, we may apply particle filtering or MCMC / importance sampling, sacrificing the computational advantage of the approach.

Monte Carlo experiment: SVCDEJ

$$\begin{aligned} \mathrm{d}\log F_t &= (-\tfrac{1}{2}v_t - \mu\lambda_t)\mathrm{d}t + \sqrt{v_t}\mathrm{d}W_{1t} + Z_t\mathrm{d}N_t, \\ \mathrm{d}v_t &= \kappa(\overline{v} - v_t)\mathrm{d}t + \sigma\sqrt{v_t}\mathrm{d}W_{2t} + Z_t^v\mathbf{1}_{\{Z_t<0\}}\mathrm{d}N_t, \end{aligned}$$

where:

- $\mathrm{d}W_{1t}\mathrm{d}W_{2t} = \rho\mathrm{d}t$;
- N_t is jump counter, with intensity $\lambda_t = \delta v_t$;
- simultaneous jumps in volatility with negative jumps in returns, with $\exp(1/\mu_v)$ distribution;
- distribution of jump size Z_t is double exponential with pdf

$$f_Z(z) = \frac{1-p^-}{\eta^+}e^{-z/\eta^+}\mathbf{1}_{\{z\geq 0\}} + \frac{p^-}{\eta^-}e^{z/\eta^-}\mathbf{1}_{\{z<0\}},$$

with p^- probability of negative jump, and η^+ and η^- the means of positive and negative jump sizes.

Monte Carlo experiment: SVCDEJ

Simulation set-up:

- Initial values: $F_0 = 100$ and $v_0 = 0.015$;
- $T = 500, \ \Delta t = 1/250$;
- p^- fixed at 0.7 (not estimated);
- Times to maturity: $\tau = 10$, 30 and 60 days;
- Finite set of strikes with $\Delta K = 0.01 F_t$;
- Add observation errors $\widehat{O}_t(\tau, m) - O_t(\tau, m) \sim N(0, 0.02^2 \kappa_t^2(\tau, m) \nu_t^2(\tau, m))$;
- Interpolation-extrapolation scheme applied to $\widehat{O}_t(\tau, m)$.

Monte Carlo results: SVCDEJ

parameter	σ	κ	$\overline{v}$	ρ	δ	η^+	η^-	μ_v	$\sigma_\varkappa$
				$u = 1, \ldots, 10$					
true value	0.450	8.000	0.015	-0.950	100.00	0.020	0.050	0.050	0.020
mean	0.440	8.779	0.014	-0.997	136.20	0.023	0.045	0.043	0.035
std dev	0.017	0.310	0.001	0.012	13.32	0.001	0.002	0.002	0.006
q10	0.427	8.409	0.013	-1.000	118.64	0.022	0.043	0.041	0.027
q50	0.437	8.836	0.013	-1.000	139.08	0.023	0.045	0.043	0.036
q90	0.451	9.102	0.014	-1.000	150.73	0.024	0.047	0.046	0.042
				$u = 1, \ldots, 20$					
true value	0.450	8.000	0.015	-0.950	100.00	0.020	0.050	0.050	0.020
mean	0.455	8.143	0.015	-0.956	110.73	0.022	0.048	0.046	0.023
std dev	0.007	0.205	0.000	0.013	4.85	0.001	0.001	0.001	0.006
q10	0.449	7.919	0.014	-0.971	105.19	0.021	0.047	0.045	0.018
q50	0.454	8.142	0.015	-0.956	110.69	0.021	0.048	0.046	0.022
q90	0.461	8.389	0.015	-0.943	116.72	0.022	0.049	0.047	0.027

Empirical application

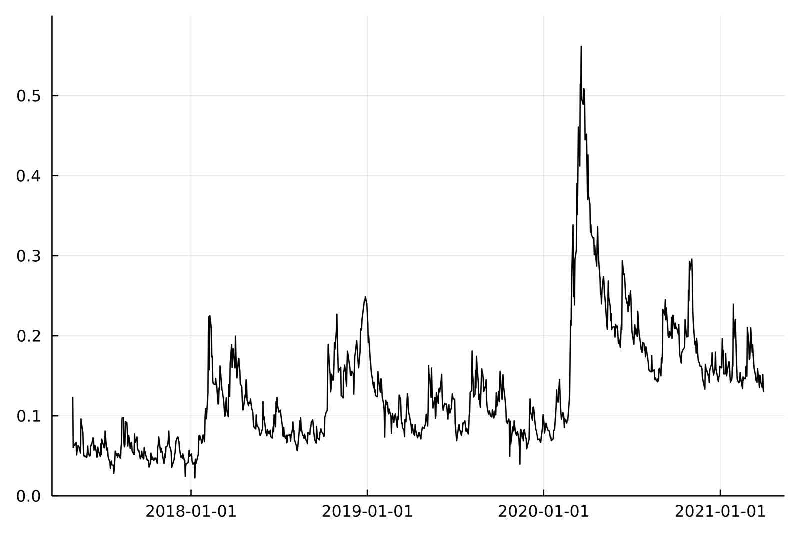
Data:

- S&P 500 index options from the CBOE;
- From May 1, 2017 to April 1, 2021 (978 trading days);
- At each day keep option with 6 different tenors less than 60 days;
- Data filters and extrapolation-interpolation scheme are applied.

SVCDEJ parameter estimates

	σ	κ	$\overline{v}$	ρ	δ	η^{+}	η^{-}	μ_v	$\sigma_{\varkappa}$
$\hat{\theta}$	0.5051	8.325	0.0153	-0.997	157.51	0.0204	0.0424	0.0519	0.253
s.e.	0.0075	0.207	0.0005	0.012	7.28	0.0005	0.0007	0.0009	0.004

SVCDEJ filtered volatility



SVCDEJ with external factor

- The 2020 volatility increase may by driven by external factor: the Covid-19 pandemic.
- Does the spread of the virus affect the volatility and the likelihood of jumps in the stock market?
- Embed the reproduction number R_0 into the SVCDEJ model.
- Assume no prior knowledge on the dynamics of R_0: assumed constant throughout option life-time;
  - for a persistent process, this is a reasonable assumption;
  - similar to treatment of interest rates in option prices.

SVCDEJ with external factor

$$
\begin{aligned}
\mathrm{d}\log F_t &= (-\tfrac{1}{2}V_t - \mu\lambda_t)\mathrm{d}t + \sqrt{v_t}\mathrm{d}W_{1,t} + q\sqrt{R_t}\mathrm{d}W_{3,t} + Z_t\mathrm{d}N_t, \\
\mathrm{d}v_t &= \kappa(\overline{v} - v_t)\mathrm{d}t + \sigma\sqrt{v_t}\mathrm{d}W_{2,t} + Z_t^v \mathbf{1}_{\{Z_t<0\}}\mathrm{d}N_t,
\end{aligned}
$$

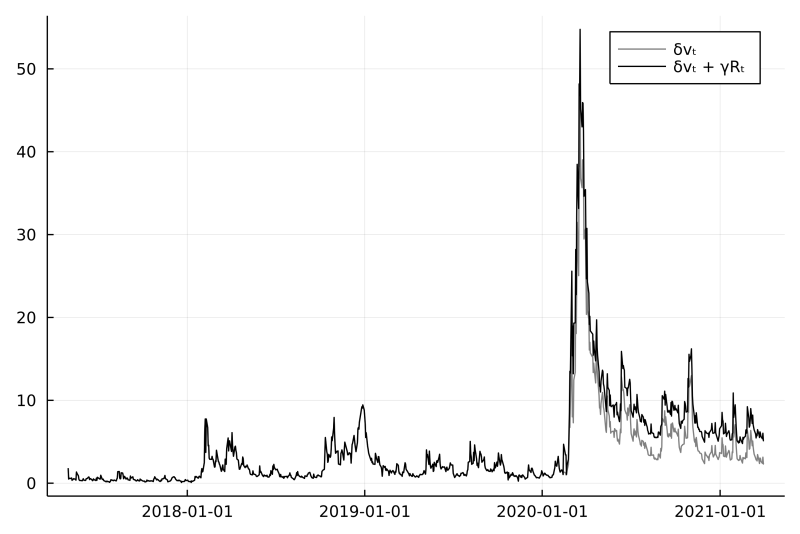
with

$$
V_t = v_t + q^2 R_t, \qquad \lambda_t = \delta v_t + \gamma R_t.
$$

SVCDEJ estimation with reproduction number as external factor

	σ	κ	$\overline{v}$	ρ	δ	η^+	η^-	μ_v	γ	q	$\sigma_\varkappa$
$\hat{\theta}$	0.568	11.55	0.014	-1.00	130.1	0.018	0.041	0.067	2.64	0.0003	0.245
s.e.	0.018	0.646	0.001	0.02	11.93	0.001	0.001	0.003	0.25	0.0001	0.004

Intensity for SVCDEJ with reproduction number



Conclusion

- Develop a new estimation and filtering procedure for AJD option pricing models.
- Based on linear relationship between option-implied log-CCF and state vector, following from AJD assumption.
- Linearity of measurement and transition equations allows the use of Kalman filter.
- Estimation procedure is fast and easy to implement.
- Monte-Carlo results suggest good finite-sample performance for different AJD models;
  - stronger identification of risk premia may require bringing in high-frequency information.
- Application to S&P 500 index options during the pandemic.