A simulation and empirical study of the behaviour of the maximum likelihood estimator for stochastic volatility jump-diffusion models*

Jean-François Bégin† Mathieu Boudreault‡

Accepted in Studies in Nonlinear Dynamics & Econometrics

Abstract

We investigate the behaviour of the maximum likelihood estimator (MLE) for stochastic volatility jump-diffusion models commonly used in financial risk management. A simulation study shows the practical conditions under which the MLE behaves according to theory. In an extensive empirical study based on nine indices and more than 6,000 individual stocks, we nonetheless find that the MLE is unable to replicate key higher moments. We then introduce a moment-targeted MLE—robust to model misspecification—and revisit both simulation and empirical studies. We find it performs better than the MLE, improving the management of financial risk.

Keywords: Maximum likelihood; Jump-diffusion model; Stochastic volatility; Discrete nonlinear filtering; Moment targeting.

JEL Classification: C13, C51, C58

*The authors would like to thank Louis Arsenault-Mahjoubi and Geneviève Gauthier for their helpful suggestions and comments. Bégin wishes to acknowledge the financial support of the Natural Science and Engineering Research Council of Canada (NSERC) and Simon Fraser University. Boudreault also thanks the financial support of NSERC. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

†Corresponding author. Department of Statistics and Actuarial Science, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6, telephone: 778-782-4478, e-mail: jbegin@sfu.ca.

‡Department of Mathematics, Université du Québec à Montréal, C.P. 8888, Succursale Centre-ville, Montréal, Québec, Canada H3C 3P8, telephone: 514-987-3000 x8560, e-mail: boudreault.mathieu@uqam.ca.

Supplementary Material

SM.A Discrete nonlinear filtering

Evaluating the likelihood function for jump-diffusion models as those presented in Section 2 is difficult because the volatility and jumps are latent variables. As such, this study relies on a deterministic DNF that uses a recursive prediction-update algorithm *à la* Kitagawa (1987). We use the method proposed by Bégin and Boudreault (2020); the latter is tailored for the jump-diffusion models investigated in this very study and is based on a discretization of the state-space continuous latent variables.

Let $\mathbf{x}_t = \begin{bmatrix} v_t & j_{y,t} & j_{v,t} \end{bmatrix} \in \mathcal{X}$ be the latent state variables at time t, where

$$j_{y,t} = \sum_{i=1}^{n_t} z_{y,t,i} \quad \text{and} \quad j_{v,t} = \sum_{i=1}^{n_t} z_{v,t,i}.$$

For a parameter set Θ, the likelihood function is

$$\mathcal{L}(\Theta \,|\, y_{1:T}) = f_\Theta(y_{1:T}) = \int_{\mathcal{X}^{T+1}} f_\Theta(\mathbf{x}_{0:T}, y_{1:T})\, d\mathbf{x}_{0:T},$$

which involves a multidimensional integration problem. Another way common way to write the likelihood function in econometrics is by conditioning on past returns,

$$\mathcal{L}(\Theta \,|\, y_{1:T}) = f_\Theta(y_1) \prod_{t=2}^{T} f_\Theta(y_t \,|\, y_{1:t-1}),$$

and therefore breaking down the high-dimensional integral into smaller problems. As explained in Bégin and Boudreault (2020), the likelihood contribution at time t can be written as the following integral:

$$\begin{aligned} f_\Theta(y_t \,|\, y_{1:t-1}) = \sum_{n=0}^{\infty} \iiint_{\mathbb{R}_+^3} & \mathbb{P}_\Theta(n_t = n)\, r_\Theta\left(y_t \,\middle|\, v_t, v_{t-1}, j_{v,t}, n_t = n\right) \\ & q_\Theta\left(v_t \,\middle|\, v_{t-1}, j_{v,t}\right) q_\Theta\left(j_{v,t} \,\middle|\, n_t = n\right) u(v_{t-1} \,|\, y_{1:t-1})\, dv_t\, dj_{v,t}\, dv_{t-1}, \end{aligned} \qquad \text{(SM.1)}$$

where

$$r_\Theta\left(y_t \,\middle|\, v_t, v_{t-1}, j_{v,t}, n_t = n\right) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(y_t - \mu_t)^2}{2\sigma_t^2}\right)$$

and

$$\begin{aligned} \mu_t &= \left(\mu - \frac{1}{2} v_{t-1} - \overline{\alpha}\omega\right) h + \frac{\rho}{\sigma}\left(v_t - v_{t-1} - \kappa(\theta - v_{t-1})h - j_{v,t}\right) + \alpha n + \rho_z j_{v,t}, \\ \sigma_t^2 &= v_{t-1}\left(1 - \rho^2\right) h + n\delta^2. \end{aligned}$$

We can obtain the integrand of Equation (SM.1) in closed form because $n_t \sim \mathcal{P}(\omega h)$, $v_t \,|\, v_{t-1}, j_{v,t} \sim \mathcal{N}\left(v_{t-1} + \kappa(\theta - v_{t-1})h + j_{v,t}, \sigma^2 v_{t-1} h\right)$, and $j_{v,t} \,|\, n_t \sim \Gamma(n_t, \nu)$, and the corresponding densities are known in closed form.[1] One can get the time-t posterior *density* of the latent states in a similar fashion by computing the following integral:

$$u(v_t \,|\, y_{1:t}) = \frac{1}{f_\Theta(y_t \,|\, y_{1:t-1})} \Bigg(\sum_{n=0}^{\infty} \iint_{\mathbb{R}_+^2} \mathbb{P}_\Theta(n_t = n)\, r_\Theta\left(y_t \,\middle|\, v_t, v_{t-1}, j_{v,t}, n_t = n\right)$$

[1] Note that $\mathcal{P}(m)$ denotes a Poisson distribution with mean m and $\Gamma(n, \nu)$ a gamma distribution with mean $n\nu$.

$$q_\Theta\left(v_t \,\middle|\, v_{t-1}, j_{v,t}\right) q_\Theta\left(j_{v,t} \,\middle|\, n_t = n\right) u\left(v_{t-1} \,|\, y_{1:t-1}\right) \, d j_{v,t} \, dv_{t-1}\Big).$$

Finally, Algorithm 1 summarizes the various steps of the DNF. The initial state density is set such that all the probability mass is at the long-run level of the variance process.

Algorithm 1 Discrete Nonlinear Filter

1: set $u(v_0) = p_\Theta(v_0)$, where p_Θ is the initial state density
2: **for** each $t \in \{1, ..., T\}$ **do**
3: get the likelihood contribution $f_\Theta(y_t \,|\, y_{1:t-1})$ using numerical integration
4: get the posterior distribution $u(v_t \,|\, y_{1:t})$ using numerical integration
5: **end for**
6: compute the likelihood function $\mathcal{L}(\Theta \,|\, y_{1:T})$ by taking the product of the likelihood contributions

SM.B Moments Derivations

In this section, we derive the first four moments for the SVCJ model. Recall that the latter nests all the models considered in this paper. To be able to get the moments, we first need to compute a few building blocks. We start from the discretization of Equations (3) and (4):

$$y_t = \left(\mu - \frac{1}{2}v_{t-1} - \overline{\alpha}\omega\right) h + \sqrt{v_{t-1}\, h}\varepsilon_{y,t} + \sum_{i=1}^{n_t} z_{y,t,i} = \left(\mu - \frac{1}{2}v_{t-1} - \overline{\alpha}\omega\right) h + \sqrt{v_{t-1}\, h}\varepsilon_{y,t} + j_{y,t},$$

$$v_t = v_{t-1} + \kappa(\theta - v_{t-1}) h + \sigma\sqrt{v_{t-1}\, h}\, \varepsilon_{v,t} + \sum_{i=1}^{n_t} z_{v,t,i} = v_{t-1} + \kappa(\theta - v_{t-1}) h + \sigma\sqrt{v_{t-1}\, h}\, \varepsilon_{v,t} + j_{v,t}.$$

Proposition 1. *The moment generating function (mgf) of $j_{v,t}$ evaluated at u is given by*

$$\varphi_{j_{v,t}}(u) = \exp\left(\omega h \left(\frac{1}{1 - u\, \nu} - 1\right)\right).$$

Proof. The mgf can be computed in the following way:

$$\begin{aligned}
\varphi_{j_{v,t}}(u) &= \mathbb{E}^{\mathbb{P}}\left[\exp\left(u\, j_{v,t}\right)\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\exp\left(u \sum_{i=1}^{n_t} z_{v,t,i}\right)\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[\exp\left(u \sum_{i=1}^{n_t} z_{v,t,i}\right) \,\middle|\, n_t\right]\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}^{\mathbb{P}}\left[\exp\left(u \sum_{i=1}^{n_t} z_{v,t,i}\right) \,\middle|\, n_t = n\right] \mathbb{P}\left(n_t = n\right) \\
&= \exp\left(\omega h \left(\mathbb{E}^{\mathbb{P}}\left[\exp\left(u\, z_{v,t,1}\right)\right] - 1\right)\right) \\
&= \exp\left(\omega h \left(\frac{1}{1 - u\, \nu} - 1\right)\right).
\end{aligned}$$

□

From this mgf, we can obtain the first four uncentred moments of $j_{v,t}$:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right] &= \left.\frac{\partial\, \varphi_{j_{v,t}}(u)}{\partial u}\right|_{u=0} = \omega h \nu, \\
\mathbb{E}^{\mathbb{P}}\left[j_{v,t}^2\right] &= \left.\frac{\partial^2 \varphi_{j_{v,t}}(u)}{\partial u^2}\right|_{u=0} = \omega h \nu^2 \left(2 + \omega h\right), \\
\mathbb{E}^{\mathbb{P}}\left[j_{v,t}^3\right] &= \left.\frac{\partial^3 \varphi_{j_{v,t}}(u)}{\partial u^3}\right|_{u=0} = \omega h \nu^3 \left(6 + 6\omega h + \omega^2 h^2\right), \\
\mathbb{E}^{\mathbb{P}}\left[j_{v,t}^4\right] &= \left.\frac{\partial^4 \varphi_{j_{v,t}}(u)}{\partial u^4}\right|_{u=0} = \omega h \nu^4 \left(24 + 36\omega h + 12\omega^2 h^2 + \omega^3 h^3\right),
\end{aligned}$$

Proposition 2. *The mgf of $j_{y,t}$ evaluated at u is given by*

$$\varphi_{j_{y,t}}(u) = \exp\left(\omega h \left(\frac{\exp\left(u\,\alpha + \frac{1}{2}\,u^2\,\delta^2\right)}{1 - u\,\rho_z\,\nu} - 1\right)\right).$$

Proof. The mgf can be computed in the following way:

$$\begin{aligned}
\varphi_{j_{y,t}}(u) &= \mathbb{E}^{\mathbb{P}}\left[\exp\left(u\, j_{y,t}\right)\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\exp\left(u \sum_{i=1}^{n_t} z_{y,t,i}\right)\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[\exp\left(u \sum_{i=1}^{n_t} z_{y,t,i}\right) \middle|\, n_t\right]\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}^{\mathbb{P}}\left[\exp\left(u \sum_{i=1}^{n_t} z_{y,t,i}\right) \middle|\, n_t = n\right] \mathbb{P}\left(n_t = n\right) \\
&= \exp\left(\omega h \left(\mathbb{E}^{\mathbb{P}}\left[\exp\left(u\, z_{y,t,1}\right)\right] - 1\right)\right) \\
&= \exp\left(\omega h \left(\mathbb{E}^{\mathbb{P}}\left[\exp\left(u\,\alpha + u\,\rho_z\, z_{v,t,1} + u\,\delta\, z\right)\right] - 1\right)\right) \\
&= \exp\left(\omega h \left(\mathbb{E}^{\mathbb{P}}\Big[\exp\left(u\,\alpha + u\,\delta\, z\right)\Big]\, \mathbb{E}^{\mathbb{P}}\left[\exp\left(u\,\rho_z\, z_{v,t,1}\right)\right] - 1\right)\right) \\
&= \exp\left(\omega h \left(\frac{\exp\left(u\,\alpha + \frac{1}{2}\,u^2\,\delta^2\right)}{1 - u\,\rho_z\,\nu} - 1\right)\right)
\end{aligned}$$

if $z \sim \mathcal{N}(0, 1)$. □

Similarly to $j_{v,t}$, we can easily get the first four uncentred moments of $j_{y,t}$ by using its mgf:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right] &= \left.\frac{\partial \varphi_{j_{y,t}}(u)}{\partial u}\right|_{u=0} = \omega h \left(\alpha + \rho_z \nu\right), \\
\mathbb{E}^{\mathbb{P}}\left[j_{y,t}^2\right] &= \left.\frac{\partial^2 \varphi_{j_{y,t}}(u)}{\partial u^2}\right|_{u=0} = \omega h \left(\alpha^2 + 2\alpha\rho_z\nu + \delta^2 + 2\rho_z^2\nu^2 + \omega h(\alpha + \rho_z\nu)^2\right), \\
\mathbb{E}^{\mathbb{P}}\left[j_{y,t}^3\right] &= \left.\frac{\partial^3 \varphi_{j_{y,t}}(u)}{\partial u^3}\right|_{u=0} = \omega h \Big(\alpha^3 + 3\alpha^2\rho_z\nu + 3\alpha\delta^2 + 6\alpha\rho_z^2\nu^2 + 3\delta^2\rho_z\nu + 6\rho_z^3\nu^3 + \omega^2 h^2(\alpha + \rho_z\nu)^3 \\
&\qquad + 3\omega h(\alpha + \rho_z\nu)\left(\alpha^2 + 2\alpha\rho_z\nu + \delta^2 + 2\rho_z^2\nu^2\right)\Big),
\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[j_{y,t}^4\right] = \left.\frac{\partial^4 \varphi_{j_{y,t}}(u)}{\partial u^4}\right|_{u=0} = \omega h \Big(& \alpha^4 + 4\alpha^3\rho_z\nu + 6\alpha^2\delta^2 + 12\alpha^2\rho_z^2\nu^2 + 12\alpha\delta^2\rho_z\nu \\
& + 24\alpha\rho_z^3\nu^3 + 3\delta^4 + 12\delta^2\rho_z^2\nu^2 + 24\rho_z^4\nu^4 + \omega^3 h^3(\alpha+\rho_z\nu)^4 \\
& + 6\omega^2h^2(\alpha+\rho_z\nu)^2\left(\alpha^2 + 2\alpha\rho_z\nu + \delta^2 + 2\rho_z^2\nu^2\right) \\
& + 3\omega h\left(\alpha^2 + 2\alpha\rho_z\nu + \delta^2 + 2\rho_z^2\nu^2\right)^2 \\
& + 4\omega h(\alpha+\rho_z\nu)\left(\alpha^3 + 3\alpha^2\rho_z\nu + 3\alpha\delta^2 + 6\alpha\rho_z^2\nu^2 + 3\delta^2\rho_z\nu + 6\rho_z^3\nu^3\right)\Big).
\end{aligned}
$$

The unconditional moments of the time-t returns depend on the first four unconditional moments of v_t; we therefore need to compute them.

Proposition 3. *The first four unconditional moments of v_t are given by*

$$
\mathbb{E}^{\mathbb{P}}\left[v_t\right] = \frac{\kappa\theta h + \mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]}{\kappa h},
$$

$$
\mathbb{E}^{\mathbb{P}}\left[v_t^2\right] = \frac{\kappa^2\theta^2h^2 + \mathbb{E}^{\mathbb{P}}\left[j_{v,t}^2\right] + 2\mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]\left(\mathbb{E}^{\mathbb{P}}\left[v_t\right](1-\kappa h) + \kappa\theta h\right) + \mathbb{E}^{\mathbb{P}}\left[v_t\right]\left(\sigma^2 h + 2\kappa\theta h - 2\kappa^2\theta h^2\right)}{2\kappa h - \kappa^2 h^2},
$$

$$
\mathbb{E}^{\mathbb{P}}\left[v_t^3\right] = \frac{\begin{pmatrix}
\kappa^3\theta^3h^3 + \mathbb{E}^{\mathbb{P}}\left[j_{v,t}^3\right] + \mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]\mathbb{E}^{\mathbb{P}}\left[v_t\right]\left(6\kappa\theta h - 6\kappa^2\theta h^2 + 3\sigma^2 h\right) \\
+\mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]\left(3\kappa^2h^2 - 6\kappa h + 3\right) + 3\mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]\kappa^2\theta^2h^2 \\
+3\mathbb{E}^{\mathbb{P}}\left[j_{v,t}^2\right]\left(\mathbb{E}^{\mathbb{P}}\left[v_t\right](1-\kappa h) + \kappa\theta h\right) + \mathbb{E}^{\mathbb{P}}\left[v_t\right]\left(3\kappa^2\theta^2h^2 - 3\kappa^3\theta^2h^3 + 3\kappa\theta\sigma^2h^2\right) \\
+\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]\left(3\kappa^3\theta h^3 - 6\kappa^2\theta h^2 - 3\kappa\sigma^2h^2 + 3\kappa\theta h + 3\sigma^2 h\right)
\end{pmatrix}}{3\kappa h - 3\kappa^2h^2 + \kappa^3h^3}
$$

$$
\mathbb{E}^{\mathbb{P}}\left[v_t^4\right] = \frac{\begin{pmatrix}
\kappa^4\theta^4h^4 + \mathbb{E}^{\mathbb{P}}\left[j_{v,t}^4\right] + 4\mathbb{E}^{\mathbb{P}}\left[j_{v,t}^3\right]\left(\mathbb{E}^{\mathbb{P}}\left[v_t\right]\left(1-\kappa h\right) + \kappa\theta h\right) \\
+\mathbb{E}^{\mathbb{P}}\left[j_{v,t}^2\right]\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]\left(6\kappa^2h^2 - 12\kappa h + 6\right) \\
+\mathbb{E}^{\mathbb{P}}\left[j_{v,t}^2\right]\mathbb{E}^{\mathbb{P}}\left[v_t\right]\left(12\kappa\theta h - 12\kappa^2\theta h^2 + 6\sigma^2 h\right) \\
+6\mathbb{E}^{\mathbb{P}}\left[j_{v,t}^2\right]\kappa^2\theta^2h^2 + \mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]\mathbb{E}^{\mathbb{P}}\left[v_t^3\right]\left(4 - 12\kappa h + 12\kappa^2h^2 - 4\kappa^3h^3\right) \\
+\mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]\left(12\kappa\theta h - 24\kappa^2\theta h^2 + 12\kappa^3\theta h^3 + 12\sigma^2 h - 12\kappa\sigma^2h^2\right) \\
+\mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]\mathbb{E}^{\mathbb{P}}\left[v_t\right]\left(12\kappa^2\theta^2h^2 - 12\kappa^3\theta^2h^3 + 12\kappa\theta\sigma^2h^2\right) + 4\mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]\kappa^3\theta^3h^3 \\
+\mathbb{E}^{\mathbb{P}}\left[v_t^3\right]\left(4\kappa\theta h - 12\kappa^2\theta h^2 + 12\kappa^3\theta h^3 - 4\kappa^4\theta h^4 + 6\sigma^2 h - 12\kappa\sigma^2h^2 + 6\kappa^2\sigma^2h^3\right) \\
+\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]\left(6\kappa^2\theta^2h^2 - 12\kappa^3\theta^2h^3 + 6\kappa^4\theta^2h^4 + 12\kappa\theta\sigma^2h^2 - 12\kappa^2\theta\sigma^2h^3 + 3\sigma^4h^2\right) \\
+\mathbb{E}^{\mathbb{P}}\left[v_t\right]\left(4\kappa^3\theta^3h^3 - 4\kappa^4\theta^3h^4 + 6\kappa^2\theta^2\sigma^2h^3\right)
\end{pmatrix}}{4\kappa h - 6\kappa^2h^2 + 4\kappa^3h^3 - \kappa^4h^4}
$$

Proof. For the first expectation, we have that

$$
\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[v_t\right] &= \mathbb{E}^{\mathbb{P}}\left[v_{t-1} + \kappa\left(\theta - v_{t-1}\right)h + \sigma\sqrt{v_{t-1}\,h}\,\varepsilon_{v,t} + j_{v,t}\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[v_{t-1}\right] + \kappa\theta h - \kappa\mathbb{E}^{\mathbb{P}}\left[v_{t-1}\right]h + \sigma\mathbb{E}^{\mathbb{P}}\left[\sqrt{v_{t-1}\,h}\,\mathbb{E}^{\mathbb{P}}\left[\varepsilon_{v,t}\,\middle|\,\mathcal{F}_t\right]\right] + \mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[v_t\right] + \kappa\theta h - \kappa\mathbb{E}^{\mathbb{P}}\left[v_t\right]h + \mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]
\end{aligned}
$$

because the unconditional expectation $\mathbb{E}^{\mathbb{P}}[v_t]$ is equal to $\mathbb{E}^{\mathbb{P}}[v_{t-1}]$. This leads to

$$\Rightarrow \mathbb{E}^{\mathbb{P}}[v_t] = \frac{\kappa\theta h + \mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]}{\kappa h}.$$

For the second expectation, we use a similar technique:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[v_t^2\right] &= \mathbb{E}^{\mathbb{P}}\left[\left(v_{t-1} + \kappa(\theta - v_{t-1})h + \sigma\sqrt{v_{t-1}\,h}\,\varepsilon_{v,t} + j_{v,t}\right)^2\right] \\
&= \mathbb{E}^{\mathbb{P}}\Big[\kappa^2\theta^2h^2 + \kappa^2v_{t-1}^2h^2 - 2\kappa^2\theta v_{t-1}h^2 + 2\kappa\theta j_{v,t}h + 2\sigma\sqrt{v_{t-1}\,h}\varepsilon_{t,v}j_{v,t} - 2\kappa v_{t-1}j_{v,t}h \\
&\qquad - 2\kappa v_{t-1}^2h + \sigma^2v_{t-1}\varepsilon_{v,t}^2h + 2\kappa\theta\sigma\sqrt{v_{t-1}\,h}\varepsilon_{t,v}h - 2\kappa\sigma v_{t-1}\sqrt{v_{t-1}\,h}\varepsilon_{t,v}h \\
&\qquad + 2\sigma v_{t-1}\sqrt{v_{t-1}\,h}\varepsilon_{t,v} + 2\kappa\theta v_{t-1}h + j_{v,t}^2 + 2v_{t-1}j_{v,t} + v_{t-1}^2\Big] \\
&= \mathbb{E}^{\mathbb{P}}\Big[\mathbb{E}^{\mathbb{P}}\Big[\kappa^2\theta^2h^2 + \kappa^2v_{t-1}^2h^2 - 2\kappa^2\theta v_{t-1}h^2 + 2\kappa\theta j_{v,t}h + 2\sigma\sqrt{v_{t-1}\,h}\varepsilon_{t,v}j_{v,t} - 2\kappa v_{t-1}j_{v,t}h \\
&\qquad - 2\kappa v_{t-1}^2h + \sigma^2v_{t-1}\varepsilon_{v,t}^2h + 2\kappa\theta\sigma\sqrt{v_{t-1}\,h}\varepsilon_{t,v}h - 2\kappa\sigma v_{t-1}\sqrt{v_{t-1}\,h}\varepsilon_{t,v}h \\
&\qquad + 2\sigma v_{t-1}\sqrt{v_{t-1}\,h}\varepsilon_{t,v} + 2\kappa\theta v_{t-1}h + j_{v,t}^2 + 2v_{t-1}j_{v,t} + v_{t-1}^2 \,\Big|\, \mathcal{F}_{t-1}\Big]\Big] \\
&= \mathbb{E}^{\mathbb{P}}\Big[\kappa^2\theta^2h^2 + \kappa^2v_{t-1}^2h^2 - 2\kappa^2\theta v_{t-1}h^2 + 2\kappa\theta j_{v,t}h - 2\kappa v_{t-1}j_{v,t}h \\
&\qquad - 2\kappa v_{t-1}^2h + \sigma^2v_{t-1}h + 2\kappa\theta v_{t-1}h + j_{v,t}^2 + 2v_{t-1}j_{v,t} + v_{t-1}^2\Big], \\
&= \kappa^2\theta^2h^2 + \mathbb{E}^{\mathbb{P}}\left[v_t^2\right]\left(1 + \kappa^2h^2 - 2\kappa h\right) \\
&\quad + \mathbb{E}^{\mathbb{P}}\left[-2\kappa^2\theta v_th^2 + 2\kappa\theta j_{v,t}h - 2\kappa v_tj_{v,t}h + \sigma^2v_th + 2\kappa\theta v_th + j_{v,t}^2 + 2v_tj_{v,t}\right],
\end{aligned}$$

which leads to

$$\Rightarrow \mathbb{E}^{\mathbb{P}}\left[v_t^2\right] = \frac{\mathbb{E}^{\mathbb{P}}\left[j_{v,t}^2\right] + 2\mathbb{E}^{\mathbb{P}}\left[j_{v,t}\right]\left(\mathbb{E}^{\mathbb{P}}\left[v_t\right](1-\kappa h) + \kappa\theta h\right) + \mathbb{E}^{\mathbb{P}}\left[v_t\right]\left(\sigma^2 + 2\kappa\theta(1-\kappa h)\right)h + \kappa^2\theta^2h^2}{2\kappa h - \kappa^2h^2}.$$

The third and fourth moments are obtained similarly and their derivations are therefore omitted. □

We now have all the required building blocks to compute the time-t return moments.

Proposition 4. *The first four unconditional moments of y_t are given by*

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[y_t\right] &= K - \tfrac{1}{2}E^{\mathbb{P}}\left[v_t\right]h + \mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right], \\
\mathbb{E}^{\mathbb{P}}\left[y_t^2\right] &= K^2 + 2K\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right] + \mathbb{E}^{\mathbb{P}}\left[j_{y,t}^2\right] + \mathbb{E}^{\mathbb{P}}\left[v_t\right]h - \mathbb{E}^{\mathbb{P}}\left[v_t\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right]h - K\mathbb{E}^{\mathbb{P}}\left[v_t\right]h + \tfrac{1}{4}\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]h^2, \\
\mathbb{E}^{\mathbb{P}}\left[y_t^3\right] &= K^3 + 3K^2\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right] + 3K\mathbb{E}^{\mathbb{P}}\left[j_{y,t}^2\right] + \mathbb{E}^{\mathbb{P}}\left[j_{y,t}^3\right] + 3\mathbb{E}^{\mathbb{P}}\left[v_t\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right]h - \tfrac{3}{2}\mathbb{E}^{\mathbb{P}}\left[v_t\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}^2\right]h \\
&\quad + 3K\mathbb{E}^{\mathbb{P}}\left[v_t\right]h - 3K\mathbb{E}^{\mathbb{P}}\left[v_t\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right]h - \tfrac{3}{2}K^2\mathbb{E}^{\mathbb{P}}\left[v_t\right]h - \tfrac{3}{2}\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]h^2 + \tfrac{3}{4}\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right]h^2 \\
&\quad + \tfrac{3}{4}K\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]h^2 - \tfrac{1}{8}\mathbb{E}^{\mathbb{P}}\left[v_t^3\right]h^3, \\
\mathbb{E}^{\mathbb{P}}\left[y_t^4\right] &= K^4 + 4K^3\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right] + 6K^2\mathbb{E}^{\mathbb{P}}\left[j_{y,t}^2\right] + 4K\mathbb{E}^{\mathbb{P}}\left[j_{y,t}^3\right] + \mathbb{E}^{\mathbb{P}}\left[j_{y,t}^4\right] \\
&\quad + 6\mathbb{E}^{\mathbb{P}}\left[v_t\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}^2\right]h - 2\mathbb{E}^{\mathbb{P}}\left[v_t\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}^3\right]h + 12K\mathbb{E}^{\mathbb{P}}\left[v_t\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right]h - 6K\mathbb{E}^{\mathbb{P}}\left[v_t\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}^2\right]h \\
&\quad + 6K^2\mathbb{E}^{\mathbb{P}}\left[v_t\right]h - 6K^2\mathbb{E}^{\mathbb{P}}\left[v_t\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right]h - 2K^3\mathbb{E}^{\mathbb{P}}\left[v_t\right]h + 3\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]h^2 - 6\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right]h^2 \\
&\quad + \tfrac{3}{2}\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}^2\right]h^2 - 6K\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]h^2 + 3K\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right]h^2 + \tfrac{3}{2}K^2\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]h^2 + \tfrac{3}{2}\mathbb{E}^{\mathbb{P}}\left[v_t^3\right]h^3 \\
&\quad - \tfrac{1}{2}\mathbb{E}^{\mathbb{P}}\left[v_t^3\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right]h^3 - \tfrac{1}{2}K\mathbb{E}^{\mathbb{P}}\left[v_t^3\right]h^3 + \tfrac{1}{16}\mathbb{E}^{\mathbb{P}}\left[v_t^4\right]h^4,
\end{aligned}$$

where $K = (\mu - \overline{\alpha}\omega)\,h$.

Proof. For the first moment, we have that

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[y_t\right] &= \mathbb{E}^{\mathbb{P}}\left[\left(\mu - \frac{1}{2}v_{t-1} - \overline{\alpha}\omega\right)h + \sqrt{v_{t-1}\,h}\varepsilon_{y,t} + j_{y,t}\right] \\
&= K - \tfrac{1}{2}\mathbb{E}^{\mathbb{P}}\left[v_{t-1}\right]h + \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[\sqrt{v_{t-1}\,h}\varepsilon_{y,t}\,\middle|\,\mathcal{F}_t\right]\right] + \mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right] \\
&= K - \tfrac{1}{2}\mathbb{E}^{\mathbb{P}}\left[v_t\right]h + \mathbb{E}^{\mathbb{P}}\left[\sqrt{v_{t-1}\,h}\mathbb{E}^{\mathbb{P}}\left[\varepsilon_{y,t}\,|\,\mathcal{F}_t\right]\right] + \mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right] \\
&= K - \tfrac{1}{2}E^{\mathbb{P}}\left[v_t\right]h + \mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right],
\end{aligned}$$

given that $\mathbb{E}^{\mathbb{P}}\left[v_t\right] = \mathbb{E}^{\mathbb{P}}\left[v_{t-1}\right]$. The second moment is computed in a similar fashion:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[y_t^2\right] &= \mathbb{E}^{\mathbb{P}}\left[\left(\left(\mu - \frac{1}{2}v_{t-1} - \overline{\alpha}\omega\right)h + \sqrt{v_{t-1}\,h}\varepsilon_{y,t} + j_{y,t}\right)^2\right] \\
&= \mathbb{E}^{\mathbb{P}}\Big[K^2 + 2Kj_{y,t} + j_{y,t}^2 + v_{t-1}^2 h\varepsilon_{y,t}^2 - j_{y,t}v_{t-1}h - Kv_{t-1}h + \tfrac{1}{4}v_{t-1}^2h^2 \\
&\qquad + 2\sqrt{v_{t-1}\,h}\varepsilon_{y,t}j_{y,t} + 2K\sqrt{v_{t-1}\,h}\varepsilon_{y,t} - v_{t-1}h\sqrt{v_{t-1}\,h}\varepsilon_{y,t}\Big] \\
&= K^2 + 2K\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right] + \mathbb{E}^{\mathbb{P}}\left[j_{y,t}^2\right] + \mathbb{E}^{\mathbb{P}}\left[v_t\right]h - \mathbb{E}^{\mathbb{P}}\left[v_t\right]\mathbb{E}^{\mathbb{P}}\left[j_{y,t}\right]h - K\mathbb{E}^{\mathbb{P}}\left[v_t\right]h + \tfrac{1}{4}\mathbb{E}^{\mathbb{P}}\left[v_t^2\right]h^2.
\end{aligned}$$

The third and fourth moments are obtained similarly and their derivations are therefore omitted. □

SM.C Additional results: Empirical study of the MLE

Table SM.1 reports MLE parameters, log-likelihood values, and CvM statistics for the nine stock indices: Panel A contains the parameters for SV, Panel B for SVYJ, and Panel C for SVCJ. For a given index, the log-likelihood value also increases as a function of model complexity, which is reassuring.

Table SM.2 shows the distribution (mean, standard deviation and five percentiles) of the parameter estimates across the 6,019 stocks of the CRSP universe.

SM.D Model misspecification with a regime-switching model

We consider model misspecification as a potential source of the issue by simulating paths from a different data generating process and assessing the closeness of the MLE-implied distributions with the observed ones.

To illustrate this assumption, we generate 20-year return paths using a regime-switching (RS) model in the spirit of Hamilton (1989). Specifically, we consider the following time-t return:

$$y_t = \left(\mu_{s_t} - \frac{v_{s_t}}{2}\right)h + \sqrt{v_{s_t}\,h}\,\varepsilon_{y,t},$$

where s_t is a Markov chain that corresponds to one of the regimes or states; i.e., $s_t \in \{1,2,3\}$. The regimes are governed by the following transition probabilities:

$$p_{ij} = \mathbb{P}\left(s_t = j\,|\,s_{t-1} = i\right), \quad i \in \{1,2,3\} \text{ and } j \in \{1,2,3\}.$$

The model parameters are set to realistic values.[2] We do not consider more models at this stage as

[2]The RS parameters used in this study are: $\mu_1 = 0.2$, $\mu_2 = 0$, $\mu_3 = -0.2$, $v_1 = 0.01$, $v_2 = 0.04$, $v_3 = 0.16$, $p_{12} = 0.02$,

Table SM.1: **Maximum likelihood parameter estimates for major stock indices.**

Panel A: SV												
	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	LL	CvM
GSPC	0.048	5.762	0.031	0.485	−0.716	0.000	0.000	0.000	0.000	0.000	25220	0.0410
	(0.017)	(0.378)	(0.002)	(0.011)	(0.017)	–	–	–	–	–		
FTSE	0.014	3.724	0.031	0.356	−0.682	0.000	0.000	0.000	0.000	0.000	24898	0.0325
	(0.020)	(0.312)	(0.002)	(0.012)	(0.030)	–	–	–	–	–		
DJIA	0.067	5.846	0.029	0.451	−0.662	0.000	0.000	0.000	0.000	0.000	25351	0.0337
	(0.018)	(0.436)	(0.002)	(0.013)	(0.019)	–	–	–	–	–		
SX5E	0.046	5.255	0.042	0.512	−0.621	0.000	0.000	0.000	0.000	0.000	23569	0.0210
	(0.022)	(0.440)	(0.001)	(0.016)	(0.030)	–	–	–	–	–		
TSX	0.096	3.017	0.024	0.286	−0.373	0.000	0.000	0.000	0.000	0.000	26297	0.0518
	(0.018)	(0.447)	(0.001)	(0.013)	(0.040)	–	–	–	–	–		
CAC	0.048	5.813	0.045	0.536	−0.700	0.000	0.000	0.000	0.000	0.000	23042	0.0114
	(0.023)	(0.344)	(0.002)	(0.013)	(0.017)	–	–	–	–	–		
DAX	0.075	5.890	0.050	0.585	−0.601	0.000	0.000	0.000	0.000	0.000	23019	0.0291
	(0.025)	(0.449)	(0.003)	(0.022)	(0.030)	–	–	–	–	–		
NI225	0.024	12.556	0.052	0.875	−0.575	0.000	0.000	0.000	0.000	0.000	22266	0.0096
	(0.025)	(0.626)	(0.002)	(0.025)	(0.023)	–	–	–	–	–		
HSI	0.053	6.504	0.064	0.705	−0.428	0.000	0.000	0.000	0.000	0.000	22321	0.0736
	(0.007)	(0.573)	(0.004)	(0.026)	(0.029)	–	–	–	–	–		
Panel B: SVYJ												
	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	LL	CvM
GSPC	0.013	5.061	0.031	0.452	−0.774	50.507	−0.002	0.005	0.000	0.000	25240	0.0576
	(0.001)	(0.399)	(0.002)	(0.019)	(0.029)	(23.479)	(0.001)	(0.000)	–	–		
FTSE	0.004	3.399	0.030	0.335	−0.728	7.115	−0.011	0.000	0.000	0.000	24904	0.0335
	(0.020)	(0.275)	(0.002)	(0.002)	(0.032)	(2.582)	(0.001)	(0.010)	–	–		
DJIA	0.053	4.629	0.029	0.410	−0.703	4.116	−0.013	0.006	0.000	0.000	25367	0.0373
	(0.020)	(0.394)	(0.002)	(0.013)	(0.021)	(2.423)	(0.005)	(0.003)	–	–		
SX5E	0.026	3.985	0.040	0.447	−0.654	4.937	−0.015	0.008	0.000	0.000	23589	0.0197
	(0.024)	(0.414)	(0.003)	(0.020)	(0.033)	(0.789)	(0.003)	(0.001)	–	–		
TSX	0.057	2.281	0.020	0.231	−0.507	14.606	−0.008	0.006	0.000	0.000	26335	0.0317
	(0.016)	(0.341)	(0.002)	(0.015)	(0.043)	(6.315)	(0.003)	(0.001)	–	–		
CAC	0.031	5.142	0.043	0.508	−0.752	4.763	−0.017	0.000	0.000	0.000	23064	0.0095
	(0.008)	(0.405)	(0.002)	(0.021)	(0.026)	(1.323)	(0.001)	(0.011)	–	–		
DAX	0.005	4.287	0.053	0.552	−0.665	68.766	−0.003	0.007	0.000	0.000	23050	0.0647
	(0.028)	(0.491)	(0.004)	(0.027)	(0.038)	(14.502)	(0.000)	(0.001)	–	–		
NI225	−0.004	7.528	0.049	0.649	−0.660	3.331	−0.007	0.025	0.000	0.000	22284	0.0067
	(0.026)	(0.552)	(0.003)	(0.025)	(0.024)	(1.115)	(0.002)	(0.003)	–	–		
HSI	0.092	3.422	0.057	0.430	−0.467	2.177	−0.013	0.046	0.000	0.000	22350	0.0568
	(0.034)	(0.456)	(0.004)	(0.026)	(0.036)	(0.654)	(0.008)	(0.006)	–	–		

Note: This table reports the maximum likelihood estimates along with their robust standard errors (in parentheses) calculated from the outer product of the gradient at the optimum parameter value (for more details on this method, see Appendix B.5 of Rémillard, 2013). The parameters are estimated using the maximum likelihood estimation (MLE) methods. We consider nine major stock indices: the S&P 500 index (GSPC), the Financial Times Stock Exchange 100 (FTSE), the Dow Jones Industrial Average (DJIA), the S&P/TSX Composite index (TSX), the Euro Stoxx 50 index (SXS5), the Cotation Assistée en Continu 40 (CAC), the Deutscher Aktienindex Performance Index (DAX), the Nikkei 225 index (NI225), and the Hang Seng index (HSI). LL stands for log-likelihood and CvM for Cramér-von Mises statistics. The CvM statistics are multiplied by 100.

this is not the main object of this study—we simply want to illustrate the issue in this section. In the

$p_{13} = 0$, $p_{21} = 0.02$, $p_{23} = 0.01$, $p_{31} = 0.01$, and $p_{32} = 0.03$.

Table SM.1: **Maximum likelihood parameter estimates for major stock indices, continued.**

Panel C: SVCJ	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	LL	CvM
GSPC	0.007	5.040	0.026	0.413	−0.751	2.503	−0.013	0.000	0.012	−0.875	25251	0.0575
	(0.021)	(0.368)	(0.002)	(0.018)	(0.025)	(0.714)	(0.002)	(0.004)	(0.002)	(0.195)		
FTSE	0.003	4.070	0.024	0.293	−0.720	1.054	−0.022	0.000	0.026	−0.103	24915	0.0226
	(0.023)	(0.422)	(0.002)	(0.017)	(0.034)	(0.334)	(0.003)	(0.031)	(0.005)	(0.095)		
DJIA	0.024	5.493	0.022	0.378	−0.700	2.691	−0.011	0.000	0.013	−0.958	25379	0.0381
	(0.022)	(0.496)	(0.002)	(0.018)	(0.029)	(0.833)	(0.003)	(0.003)	(0.003)	(0.254)		
SX5E	0.062	5.198	0.029	0.459	−0.656	2.141	−0.010	0.006	0.011	−1.075	23594	0.0026
	(0.023)	(0.311)	(0.002)	(0.016)	(0.031)	(1.131)	(0.008)	(0.008)	(0.005)	(0.398)		
TSX	0.046	4.377	0.011	0.189	−0.455	2.235	−0.015	0.000	0.019	−0.740	26350	0.0147
	(0.022)	(0.382)	(0.001)	(0.013)	(0.047)	(0.513)	(0.003)	(0.021)	(0.002)	(0.124)		
CAC	0.016	5.574	0.039	0.501	−0.741	2.993	−0.014	0.000	0.008	−1.129	23077	0.0096
	(0.028)	(0.428)	(0.002)	(0.021)	(0.026)	(1.093)	(0.003)	(0.007)	(0.002)	(0.265)		
DAX	0.005	2.818	0.059	0.486	−0.726	58.395	−0.002	0.008	0.000	0.012	23052	0.1037
	(0.037)	(0.276)	(0.013)	(0.009)	(0.037)	(0.935)	(0.045)	(0.001)	(0.001)	(0.228)		
NI225	0.007	8.980	0.038	0.565	−0.715	2.515	−0.003	0.036	0.041	−0.118	22295	0.0096
	(0.032)	(0.652)	(0.003)	(0.030)	(0.032)	(0.832)	(0.013)	(0.006)	(0.010)	(0.251)		
HSI	0.039	3.180	0.027	0.195	−0.699	2.073	−0.017	0.049	0.038	−0.101	22404	0.0156
	(0.033)	(0.254)	(0.002)	(0.010)	(0.017)	(0.440)	(0.013)	(0.006)	(0.008)	(0.308)		

next sections, we consider an extensive empirical exercise with more than 6,000 stocks that will serve as a better, more realistic test in the context of model misspecification.

SM.D.1 Impact of model misspecification on the MLE

Based on 500 RS paths, we estimate the SV, SVYJ, and SVCJ parameters using MLE, similarly to the experiment in Section 3. Then, using these parameters, we compute moments and CvM statistics.

Table SM.3 reports the average moments and CvM statistics for the three models. For all models, MLE yields biased mean and kurtosis values. The MLE-based kurtosis estimates are also very variable from one run to the other. The average standard deviation obtained via MLE is closer to the true value for SVYJ and SVCJ, whereas the MLE-implied skewness is correctly captured only for the SVCJ model. In addition to the moment mismatch, the MLE is incapable of reproducing the return distribution: the average CvM statistic tends to be somewhat large, generally speaking. It significantly decreases as more complex models are considered (i.e., there is a decrease of about 0.0007 from SV to SVYJ and of about 0.0002 from SVYJ to SVCJ).

In sum, even though SV, SVYJ, and SVCJ are capable of replicating the return distribution generated by a different model (RS in this case), the parameters estimated via MLE do not reproduce the return distribution fully.

SM.D.2 Impact of model misspecification on the mtMLE

We now consider model misspecification in the context of mtMLE. Again, based on 500 RS paths, we estimate the SV, SVYJ, and SVCJ parameters using mtMLE and use these parameters to compute moments and CvM statistics.

Table SM.4 reports the average moments and CvM statistics for the three models and the mtMLE method. For the SV model, all moments (even those not included in the moment targeting step) are

Table SM.2: **Distribution of maximum likelihood parameter estimates for individual stocks.**

Panel A: SV											
	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	CvM
Mean	0.147	8.734	0.244	1.245	−0.306	–	–	–	–	–	0.4098
Std Dev	0.151	11.343	0.242	0.985	0.232	–	–	–	–	–	0.8917
P_5	−0.061	0.037	0.026	0.192	−0.700	–	–	–	–	–	0.0076
P_{25}	0.065	0.943	0.085	0.600	−0.465	–	–	–	–	–	0.0314
Median	0.141	4.389	0.167	1.087	−0.261	–	–	–	–	–	0.0841
P_{75}	0.226	11.332	0.327	1.601	−0.159	–	–	–	–	–	0.3156
P_{95}	0.391	36.378	0.696	2.806	−0.003	–	–	–	–	–	2.2480
Panel B: SVYJ											
	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	CvM
Mean	0.146	6.731	0.169	0.936	−0.335	6.023	−0.041	0.120	–	–	0.2080
Std Dev	0.154	10.576	0.158	0.922	0.222	9.216	0.272	0.111	–	–	0.5219
P_5	−0.065	0.227	0.021	0.206	−0.699	0.057	−0.185	0.005	–	–	0.0032
P_{25}	0.065	0.995	0.066	0.439	−0.496	1.423	−0.029	0.059	–	–	0.0154
Median	0.141	2.890	0.123	0.669	−0.312	3.717	−0.006	0.096	–	–	0.0445
P_{75}	0.226	6.785	0.220	1.079	−0.193	7.000	0.009	0.149	–	–	0.1429
P_{95}	0.391	31.992	0.471	2.648	−0.018	19.125	0.061	0.303	–	–	0.9751
Panel C: SVCJ											
	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	CvM
Mean	0.128	8.680	0.133	1.087	−0.331	5.534	0.022	0.086	0.044	−1.628	0.8102
Std Dev	0.150	11.713	0.143	1.147	0.219	5.121	0.263	0.097	0.074	3.160	1.2116
P_5	−0.085	0.761	0.004	0.208	−0.675	0.182	−0.108	0.000	0.000	−9.025	0.0034
P_{25}	0.052	2.093	0.041	0.441	−0.495	1.794	0.001	0.032	0.007	−2.208	0.0214
Median	0.126	3.811	0.091	0.699	−0.323	4.185	0.033	0.064	0.018	−1.194	0.1393
P_{75}	0.204	9.203	0.173	1.232	−0.180	7.486	0.081	0.110	0.048	−0.626	1.1726
P_{95}	0.355	40.056	0.400	3.401	−0.002	17.732	0.193	0.252	0.175	1.769	3.6540

Note: This table reports sample statistics for MLE estimated parameters for SV, SVYJ, and SVCJ as well as sample statistics for Cramér-von Mises statistics. We consider 6,019 post-1990 return series of at least five years of consecutive data. Std Dev represents the standard deviation, P_q the q^{th} percentile of the distribution, and CvM the Cramér-von Mises statistics (multiplied by 100).

Table SM.3: **Moments and Cramér-von Mises computed from maximum likelihood parameter estimates for a misspecified model.**

	True	SV		SVYJ		SVCJ	
Mean	2.263	3.628	(2.264)	4.374	(1.905)	3.981	(1.633)
Std Dev	1.242	1.468	(0.606)	1.255	(0.287)	1.234	(0.108)
Skewness	−0.133	−0.026	(0.062)	0.001	(0.024)	−0.162	(0.327)
Kurtosis	6.894	7.853	(18.844)	5.495	(5.122)	8.119	(9.812)
CvM		0.1083		0.0377		0.0168	

Note: This table reports average model moments and Cramér-von Mises (CvM) statistics obtained for parameter estimates obtained with MLE out of 500 20-year return paths when the model is misspecified. The data generating process is a regime-switching (RS) model with three regimes. The RS parameters used in this study are: $\mu_1 = 0.2$, $\mu_2 = 0$, $\mu_3 = -0.2$, $v_1 = 0.01$, $v_2 = 0.04$, $v_3 = 0.16$, $p_{12} = 0.02$, $p_{13} = 0$, $p_{21} = 0.02$, $p_{23} = 0.01$, $p_{31} = 0.01$, $p_{32} = 0.03$. We report the mean (multiplied by 10,000), the standard deviation (multiplied by 100), the skewness, and the kurtosis of daily returns. The CvM statistics are multiplied by 100.

more precisely estimated with mtMLE than with MLE.

For SVYJ and SVCJ, we see similar results: the mtMLE average model moments are closer than those obtained with MLE. The standard deviations for mtMLE parameters also tend to be smaller than those found with the MLE method.

Table SM.4: **Moments and Cramér-von Mises computed from moment-targeting maximum likelihood parameter estimates for a misspecified model.**

	True	SV		SVYJ		SVCJ	
Mean	2.263	2.177	(1.915)	2.136	(1.845)	2.046	(1.827)
Std Dev	1.242	1.232	(0.075)	1.227	(0.107)	1.235	(0.075)
Skewness	−0.133	−0.021	(0.033)	−0.147	(0.133)	−0.153	(0.130)
Kurtosis	6.894	6.437	(3.579)	6.842	(0.731)	6.864	(0.602)
CvM		0.0080		0.0093		0.0106	

Note: This table reports average model moments and Cramér-von Mises (CvM) statistics obtained for parameter estimates obtained with mtMLE out of 500 20-year return paths when the model is misspecified. For more details, see the caption of Table SM.3.

In addition to the moment mismatch mentioned above, the MLE method reproduces the return distribution poorly: the average CvM statistic is 13.6 times larger for SV, 4.1 times larger for SVYJ, and 1.6 times larger for SVCJ than the results obtained with mtMLE. This implies that the mtMLE fares well even in the context of model misspecification.

SM.E Parameter Choices for Moment Targeting

This section of the supplementary material investigates different moment constraints. In the core of the paper, we used $\Theta_{\text{mt}} = \{\mu, \theta, \alpha, \delta\}$. This specification works well, but is not unique. Here, we look at two different specifications. We only consider SVYJ in this exercise for the sake of conciseness.

Table SM.5 reports biases and root-mean-squared errors (RMSEs) when $\Theta_{\text{mt}} = \{\mu, \theta, \omega, \alpha\}$, whereas Table SM.6 reports the same thing but when $\Theta_{\text{mt}} = \{\mu, \theta, \omega, \delta\}$. Overall, the results from these two tables are similar to those presented in Tables 1 and 5 of the article: the parameter are accurately estimated, on average, for both methods. It does not seem that using a different Θ_{mt} impacts materially the results for mtMLE.

SM.F Additional results: Empirical study of the mtMLE

Table SM.7 reports mtMLE parameters, log-likelihood values, and CvM statistics for the nine stock indices; this table mirrors the results reported in Table SM.1 but for the mtMLE instead of the MLE methodology.

Table SM.8 mirrors Table SM.2 and reports the mtMLE individual stock results.

Table SM.5: **Maximum likelihood and moment-targeting maximum likelihood parameter estimates with ω instead of δ in the moment targeting.**

Panel A: SVYJ, ZM-SSD											
		μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z
True		0.060	3.000	0.015	0.300	−0.750	1.000	0.000	0.010	0.000	0.000
MLE	Bias	−0.002	0.114	0.000	0.000	−0.008	−0.293	0.001	0.002	–	–
	RMSE	(0.022)	(0.363)	(0.002)	(0.016)	(0.030)	(1.032)	(0.026)	(0.016)	–	–
mtMLE	Bias	0.002	0.071	0.000	−0.001	−0.004	−0.229	0.000	−0.003	–	–
	RMSE	(0.028)	(0.425)	(0.003)	(0.016)	(0.023)	(0.604)	(0.001)	(0.007)	–	–
Panel B: SVYJ, NM-SSD											
		μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z
True		0.060	3.000	0.015	0.300	−0.750	1.000	−0.050	0.010	0.000	0.000
MLE	Bias	−0.002	0.077	0.000	−0.001	−0.010	0.003	0.000	0.000	–	–
	RMSE	(0.025)	(0.347)	(0.002)	(0.016)	(0.028)	(0.306)	(0.005)	(0.004)	–	–
mtMLE	Bias	0.002	0.178	0.000	−0.003	−0.004	0.013	0.000	−0.001	–	–
	RMSE	(0.030)	(0.472)	(0.003)	(0.017)	(0.025)	(0.327)	(0.004)	(0.006)	–	–
Panel C: SVYJ, ZM-LSD											
		μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z
True		0.060	3.000	0.015	0.300	−0.750	1.000	0.000	0.050	0.000	0.000
MLE	Bias	−0.003	0.091	0.000	−0.001	−0.011	0.030	0.000	−0.003	–	–
	RMSE	(0.025)	(0.382)	(0.002)	(0.016)	(0.028)	(0.333)	(0.016)	(0.010)	–	–
mtMLE	Bias	0.002	−0.026	0.000	−0.004	−0.008	0.052	0.000	0.006	–	–
	RMSE	(0.031)	(0.283)	(0.003)	(0.014)	(0.014)	(1.215)	(0.000)	(0.194)	–	–
Panel D: SVYJ, NM-LSD											
		μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z
True		0.060	3.000	0.015	0.300	−0.750	1.000	−0.050	0.050	0.000	0.000
MLE	Bias	−0.003	0.101	0.000	0.000	−0.011	−0.059	−0.006	−0.005	–	–
	RMSE	(0.026)	(0.378)	(0.002)	(0.015)	(0.027)	(0.316)	(0.017)	(0.011)	–	–
mtMLE	Bias	0.002	0.129	0.000	0.000	−0.004	0.014	−0.004	−0.006	–	–
	RMSE	(0.032)	(0.512)	(0.003)	(0.017)	(0.029)	(0.325)	(0.013)	(0.010)	–	–

Note: This table reports biases and root-mean-square errors (RMSEs) for parameter estimates obtained with maximum likelihood estimation (MLE) and moment-targeting maximum likelihood estimation (mtMLE). The mtMLE method targets parameters μ, θ, ω, and α. The parameter estimation problem is repeated 500 times for each parameter set and each model. For each problem, we generate 20 years of daily data. ZM stands for zero jump mean ($\alpha = 0$), NM for negative jump mean ($\alpha = -0.05$), SSD for small jump standard deviation ($\delta = 0.01$), and LSD for large jump standard deviation ($\delta = 0.05$).

Table SM.6: **Maximum likelihood and moment-targeting maximum likelihood parameter estimates with ω instead of α in the moment targeting.**

Panel A: SVYJ, ZM-SSD											
		μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z
True		0.060	3.000	0.015	0.300	−0.750	1.000	0.000	0.010	0.000	0.000
MLE	Bias	−0.002	0.114	0.000	0.000	−0.008	−0.293	0.001	0.002	–	–
	RMSE	(0.022)	(0.363)	(0.002)	(0.016)	(0.030)	(1.032)	(0.026)	(0.016)	–	–
mtMLE	Bias	0.002	0.129	0.000	−0.003	−0.010	−0.132	−0.003	0.005	–	–
	RMSE	(0.028)	(0.458)	(0.003)	(0.016)	(0.022)	(0.656)	(0.067)	(0.015)	–	–
Panel B: SVYJ, NM-SSD											
		μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z
True		0.060	3.000	0.015	0.300	−0.750	1.000	−0.050	0.010	0.000	0.000
MLE	Bias	−0.002	0.077	0.000	−0.001	−0.010	0.003	0.000	0.000	–	–
	RMSE	(0.025)	(0.347)	(0.002)	(0.016)	(0.028)	(0.306)	(0.005)	(0.004)	–	–
mtMLE	Bias	0.002	0.085	0.000	−0.004	−0.006	−0.002	0.000	−0.001	–	–
	RMSE	(0.030)	(0.357)	(0.003)	(0.016)	(0.017)	(0.249)	(0.004)	(0.002)	–	–
Panel C: SVYJ, ZM-LSD											
		μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z
True		0.060	3.000	0.015	0.300	−0.750	1.000	0.000	0.050	0.000	0.000
MLE	Bias	−0.003	0.091	0.000	−0.001	−0.011	0.030	0.000	−0.003	–	–
	RMSE	(0.025)	(0.382)	(0.002)	(0.016)	(0.028)	(0.333)	(0.016)	(0.010)	–	–
mtMLE	Bias	0.002	0.121	0.000	0.000	−0.005	0.017	0.001	−0.004	–	–
	RMSE	(0.031)	(0.450)	(0.003)	(0.016)	(0.025)	(0.340)	(0.016)	(0.009)	–	–
Panel D: SVYJ, NM-LSD											
		μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z
True		0.060	3.000	0.015	0.300	−0.750	1.000	−0.050	0.050	0.000	0.000
MLE	Bias	−0.003	0.101	0.000	0.000	−0.011	−0.059	−0.006	−0.005	–	–
	RMSE	(0.026)	(0.378)	(0.002)	(0.015)	(0.027)	(0.316)	(0.017)	(0.011)	–	–
mtMLE	Bias	0.002	0.137	0.000	−0.001	−0.007	−0.002	0.001	0.001	–	–
	RMSE	(0.032)	(0.420)	(0.003)	(0.016)	(0.023)	(0.479)	(0.019)	(0.008)	–	–

Note: See the caption of Table SM.5 for more details.

Table SM.7: **Moment-targeting maximum likelihood parameter estimates for major stock indices.**

Panel A: SV

	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	LL	CvM
GSPC	0.089	5.712	0.030	0.485	−0.689	0.000	0.000	0.000	0.000	0.000	25218	0.0283
	–	(0.038)	–	(0.013)	(0.026)	–	–	–	–	–		
FTSE	0.053	3.795	0.029	0.348	−0.660	0.000	0.000	0.000	0.000	0.000	24896	0.0169
	–	(0.121)	–	(0.012)	(0.031)	–	–	–	–	–		
DJIA	0.092	5.718	0.028	0.442	−0.639	0.000	0.000	0.000	0.000	0.000	25350	0.0244
	–	(0.129)	–	(0.012)	(0.029)	–	–	–	–	–		
SX5E	0.063	4.907	0.044	0.517	−0.611	0.000	0.000	0.000	0.000	0.000	23566	0.0268
	–	(0.387)	–	(0.021)	(0.031)	–	–	–	–	–		
TSX	0.060	3.198	0.023	0.283	−0.397	0.000	0.000	0.000	0.000	0.000	26295	0.0499
	–	(0.039)	–	(0.009)	(0.037)	–	–	–	–	–		
CAC	0.060	5.720	0.046	0.542	−0.693	0.000	0.000	0.000	0.000	0.000	23041	0.0147
	–	(0.039)	–	(0.016)	(0.026)	–	–	–	–	–		
DAX	0.091	5.810	0.049	0.579	−0.597	0.000	0.000	0.000	0.000	0.000	23019	0.0224
	–	(0.407)	–	(0.022)	(0.029)	–	–	–	–	–		
NI225	0.011	12.878	0.055	0.907	−0.572	0.000	0.000	0.000	0.000	0.000	22265	0.0162
	–	(0.475)	–	(0.023)	(0.022)	–	–	–	–	–		
HSI	0.107	6.025	0.060	0.653	−0.424	0.000	0.000	0.000	0.000	0.000	22320	0.0480
	–	(0.074)	–	(0.015)	(0.030)	–	–	–	–	–		

Panel B: SVYJ

	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	LL	CvM
GSPC	0.089	5.552	0.029	0.459	−0.698	0.338	−0.017	0.069	0.000	0.000	25218	0.0200
	–	(0.025)	–	(0.012)	(0.026)	(0.123)	–	–	–	–		
FTSE	0.053	3.662	0.028	0.334	−0.675	0.322	−0.007	0.063	0.000	0.000	24891	0.0118
	–	(0.115)	–	(0.011)	(0.029)	(0.147)	–	–	–	–		
DJIA	0.092	6.112	0.026	0.437	−0.668	0.463	−0.010	0.061	0.000	0.000	25352	0.0157
	–	(0.329)	–	(0.008)	(0.019)	(0.187)	–	–	–	–		
SX5E	0.063	4.074	0.042	0.452	−0.635	0.531	−0.009	0.065	0.000	0.000	23581	0.0210
	–	(0.250)	–	(0.002)	(0.033)	(0.225)	–	–	–	–		
TSX	0.060	2.892	0.022	0.258	−0.426	0.430	−0.036	0.055	0.000	0.000	26312	0.0317
	–	(0.072)	–	(0.009)	(0.038)	(0.156)	–	–	–	–		
CAC	0.060	5.546	0.045	0.522	−0.712	0.362	−0.007	0.071	0.000	0.000	23052	0.0093
	–	(0.210)	–	(0.018)	(0.027)	(0.186)	–	–	–	–		
DAX	0.091	5.459	0.047	0.535	−0.608	0.239	−0.019	0.080	0.000	0.000	23031	0.0201
	–	(0.095)	–	(0.016)	(0.030)	(0.119)	–	–	–	–		
NI225	0.011	8.870	0.052	0.714	−0.654	0.509	−0.015	0.076	0.000	0.000	22276	0.0117
	–	(0.293)	–	(0.021)	(0.023)	(0.186)	–	–	–	–		
HSI	0.107	4.093	0.052	0.445	−0.464	1.468	−0.003	0.075	0.000	0.000	22347	0.0319
	–	(0.468)	–	(0.025)	(0.034)	(0.362)	–	–	–	–		

Note: This table reports the moment targeting-based parameters along with their robust standard errors (in parentheses) calculated from the outer product of the gradient at the optimum parameter value. For more details, see the caption of Table SM.1.

Table SM.7: **Moment-targeting maximum likelihood parameter estimates for major stock indices, continued.**

Panel C: SVCJ											
	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	LL	CvM
GSPC	0.089	7.610	0.017	0.368	−0.730	2.456	0.032	0.026	0.030	−0.969	25223	0.0123
	–	(0.530)	–	(0.008)	(0.030)	(0.514)	–	–	(0.006)	(0.193)	–	
FTSE	0.053	3.820	0.020	0.246	−0.726	0.759	0.034	0.036	0.037	−0.841	24909	0.0050
	–	(0.330)	–	(0.015)	(0.037)	(0.202)	–	–	(0.010)	(0.277)	–	
DJIA	0.092	6.310	0.018	0.324	−0.713	1.609	0.030	0.031	0.029	−0.947	25362	0.0101
	–	(0.417)	–	(0.001)	(0.030)	(0.189)	–	–	(0.004)	(0.138)	–	
SX5E	0.063	5.037	0.033	0.436	−0.671	0.798	−0.047	0.046	0.052	0.612	23593	0.0125
	–	(0.425)	–	(0.021)	(0.033)	(0.289)	–	–	(0.018)	(0.029)	–	
TSX	0.060	4.140	0.012	0.188	−0.513	1.348	0.013	0.029	0.026	−1.056	26336	0.0116
	–	(0.427)	–	(0.013)	(0.045)	(0.232)	–	–	(0.005)	(0.164)	–	
CAC	0.060	6.160	0.031	0.436	−0.798	1.575	0.018	0.045	0.044	−0.442	23065	0.0039
	–	(0.414)	–	(0.006)	(0.027)	(0.445)	–	–	(0.012)	(0.315)	–	
DAX	0.091	5.225	0.033	0.383	−0.691	1.490	0.025	0.042	0.045	−0.613	23044	0.0118
	–	(0.074)	–	(0.001)	(0.036)	(0.367)	–	–	(0.011)	(0.233)	–	
NI225	0.011	6.878	0.036	0.486	−0.719	2.905	0.017	0.045	0.028	−0.725	22289	0.0062
	–	(0.554)	–	(0.026)	(0.031)	(0.666)	–	–	(0.006)	(0.372)	–	
HSI	0.107	3.430	0.024	0.182	−0.702	2.194	0.009	0.067	0.040	−0.298	22403	0.0083
	–	(0.256)	–	(0.010)	(0.017)	(0.380)	–	–	(0.008)	(0.455)	–	

Table SM.8: **Distribution of moment-targeting maximum likelihood parameter estimates for individual stocks.**

Panel A: SV											
	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	CvM
Mean	0.126	10.801	0.480	1.757	−0.279	–	–	–	–	–	0.3119
Std Dev	0.643	13.201	1.534	1.647	0.254	–	–	–	–	–	0.7150
P_5	−0.057	0.034	0.029	0.226	−0.710	–	–	–	–	–	0.0057
P_{25}	0.003	1.096	0.096	0.647	−0.439	–	–	–	–	–	0.0218
Median	0.037	5.811	0.207	1.323	−0.223	–	–	–	–	–	0.0531
P_{75}	0.131	14.760	0.434	2.223	−0.112	–	–	–	–	–	0.2267
P_{95}	0.383	44.531	1.255	5.029	0.002	–	–	–	–	–	1.8579
Panel B: SVYJ											
	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	CvM
Mean	0.083	7.453	0.260	1.229	−0.345	12.593	−0.012	0.110	–	–	0.0676
Std Dev	0.155	11.699	0.563	1.472	0.237	21.551	0.146	0.206	–	–	0.2105
P_5	−0.106	0.167	0.018	0.219	−0.711	0.090	−0.088	0.000	–	–	0.0012
P_{25}	0.007	0.926	0.071	0.442	−0.530	2.225	−0.013	0.035	–	–	0.0059
Median	0.065	2.772	0.143	0.684	−0.333	5.848	−0.001	0.071	–	–	0.0164
P_{75}	0.145	7.617	0.284	1.331	−0.177	10.940	0.008	0.122	–	–	0.0501
P_{95}	0.329	39.038	0.768	4.371	−0.002	72.541	0.058	0.286	–	–	0.2773
Panel C: SVCJ											
	μ	κ	θ	σ	ρ	ω	α	δ	ν	ρ_z	CvM
Mean	0.057	5.343	0.215	1.000	−0.406	6.455	−0.007	0.110	0.035	−0.952	0.2137
Std Dev	0.082	9.152	0.928	1.362	0.227	5.690	0.036	0.203	0.065	1.836	0.5380
P_5	−0.056	0.192	0.003	0.178	−0.714	0.255	−0.038	0.000	0.000	−3.814	0.0015
P_{25}	0.017	0.773	0.022	0.335	−0.584	2.412	−0.020	0.029	0.002	−1.309	0.0066
Median	0.058	1.889	0.054	0.523	−0.442	4.665	−0.010	0.067	0.014	−0.800	0.0230
P_{75}	0.090	5.066	0.137	0.977	−0.247	8.618	0.002	0.118	0.038	−0.253	0.1170
P_{95}	0.182	25.176	0.661	3.905	−0.014	21.204	0.037	0.313	0.139	0.941	1.2228

Note: This table reports sample statistics for mtMLE estimated parameters for SV, SVYJ, and SVCJ as well as sample statistics for Cramér-von Mises statistics. We consider 6,019 post-1990 return series of at least five years of consecutive data. Std Dev represents the standard deviation, P_q the q^{th} percentile of the distribution, and CvM the Cramér-von Mises statistics (multiplied by 100).