Journal of Financial Econometrics, 2023, 1–32
https://doi.org/10.1093/jjfinec/nbad001
Article

OXFORD

Score-Driven Modeling with Jumps: An Application to S&P500 Returns and Options

Luca Vincenzo Ballestra[1], Enzo D'Innocenzo [2,*], and Andrea Guizzardi [1]

[1]University of Bologna, Italy and [2]Vrije Universiteit Amsterdam, The Netherlands

*Address correspondence to Enzo D'Innocenzo, Department of Econometrics and Data Science, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, or e-mail: e.dinnocenzo@vu.nl.

Received November 16, 2021; revised December 30, 2022; editorial decision January 10, 2023; accepted January 11, 2023

Abstract

We introduce a novel score-driven model with two sources of shock, allowing for both time-varying volatility and jumps. A theoretical investigation is performed which yields sufficient conditions to ensure stationarity and ergodicity. We extend the model to consider a time-varying jump intensity. Both an in-sample and an out-of-sample analysis based on the S&P500 time series show that the proposed methodology provides excellent agreement with observed returns, outperforming more standard Generalized Autoregressive Contional Heteroskedasticity (GARCH) specifications with jumps. Finally, we apply our models to option pricing via risk neutralization. Results show this novel approach produces reliable implied volatility surfaces. Supplementary Materials including proofs, the derivation of the conditional Fisher information, and two figures showing additional empirical results are available online.

Key words: time-varying volatility, compound Poisson, observation-driven models, stationarity and ergodicity, option pricing, JEL Codes: C510, C530, C580

JEL classification: C510, C530, C580

Several empirical studies document that asset prices are affected by sharp and large discontinuities (jumps), due to unexpected news and events, see the recent works by Gürkaynak, Kisacikoğlu, and Wright (2020), Engle et al. (2021) and Jeon, McCurdy, and Zhao (2021). To account for these discontinuities, nonstandard GARCH approaches that also allow for jumps in the return process have been proposed by Vlaar and Palm (1993), Chan and Maheu (2002), Maheu and McCurdy (2004), Duan, Ritchken, and Sun (2006),

Christoffersen, Jacobs, and Ornthanalai (2012), Guégan, Ielpo, and Lalaharison (2013), and Ornthanalai (2014).

Importantly, GARCH models are not robust to misspecification or even small departures from the data-generating process (DGP) given that, in their original form, these models use squared-lagged innovations to update the value of the conditional variance. This makes them very sensitive to the presence of even a few outliers or extreme returns. A possible solution to this issue is provided by the so-called score-driven models, which were originally proposed by Creal, Koopman, and Lucas (2011) and Harvey (2013).

The key feature of score-driven models is that the dynamic of the time-varying parameters is described by an autoregressive process driven by a scaled version of the score function, that is, the derivative of the (postulated) conditional log-density. Moreover, the score-driven approach guarantees that the Kullback–Leibler divergence between the probability density function of the DGP and the model implied probability distribution diminishes at least locally, see Blasques, Koopman, and Lucas (2015). In addition to guaranteeing theoretical optimality, this property is crucial in model misspecifications, when the chosen modeling framework is different from the true economic dynamic of the empirical data. Score-driven models also provide a general framework that fully exploits shape of the observation conditional density as its characteristics can be specifically incorporated as driving forces for the time-varying parameters. As a final advantage, score-driven models belong to the class of observation-driven models, see Cox et al. (1981), since the evolution of the unobserved dynamic parameters depends only on historical data. Therefore, exactly like GARCH models, they can be easily filtered and estimated by maximum likelihood, which makes them appealing for practical applications.

Despite the strong advantages of this approach and the importance of accounting for jumps in asset price dynamics, literature on the score-driven approach with jumps is still lacking. To fill this gap, we develop a score-driven model with jumps (*SDJ*), where the conditional variance of the returns is assumed to follow an autoregressive process driven by the score of the predictive density, and the jumps are modeled by a compound Poisson process. This approach allows us to take into account the interaction between jumps and volatility, since the Poisson process used to specify the jumps is fully coupled with the dynamics of the variance. Moreover, from a theoretical standpoint, the proposed model offers the advantage of being strictly stationary and ergodic. In particular, based on the framework developed by Blasques, Koopman, and Lucas (2014), we establish one mild sufficient condition that ensures the ergodicity and strict stationarity of the return process.

We also develop two extensions of the *SDJ* model in which the conditional variance of the returns and the conditional jump intensity follow a bivariate (coupled) score-driven autoregression. These two approaches, which we label *SDSDJ-1* and *SDSDJ-2*, allow us to take into account the stochastic nature of the jump frequency, which depends on the contingent and continuously evolving macroeconomic conditions.

We test the empirical performances of the proposed *SDJ*, *SDSDJ-1*, and *SDSDJ-2* models, and we compare them against the GARCH models with jumps and both constant or time-varying intensity introduced by Christoffersen et al. (2008) and Christoffersen, Jacobs, and Ornthanalai (2012). We conduct both an in-sample and an out-of-sample exercise focusing on the S&P500 total return time series. The results reveal that the score-driven approaches provide a very good fitting of empirical data as they significantly outperform their GARCH counterparts. In a Monte Carlo exercise, we show the superiority of the

SDJ, *SDSDJ-1*, and *SDSDJ-2* models in filtering the unobserved variance process in a misspecified setting. Finally, we derive the risk-neutralized versions of the score-driven models and we apply them to pricing options written on the S&P500 index. Again, the empirical findings show that the proposed score-driven framework yields more accurate option prices than the GARCH benchmarks.

The contribution of the article is two-fold. First, we propose a new class of nonlinear models with jumps based on the score-driven approach; second, we apply the score-driven technology to value derivatives, which, to the best of our knowledge, is new to the literature.

The remainder of the article is organized as follows. In Section 1, we introduce the *SDJ* model and we establish sufficient conditions ensuring ergodicity and strict stationarity. In Section 2, we present the *SDSDJ-1* and *SDSDJ-2* extensions of the score-driven approach developed in Section 1. In Section 3, we briefly review the two GARCH models used as benchmarks. Section 4 shows the empirical performance of the proposed models and the benchmarks when dealing with the S&P500 total return time series. In Section 5, we present the Monte Carlo simulations showing the properties of our models that are relevant for practical applications. In Section 6, we risk-neutralize the models, whereas in Section 7, we test the abilities of the score-driven and GARCH models in pricing S&P500 options. Finally, Section 8 concludes.

1 Modeling Returns with Jumps

Let S_t denotes the price of a risky asset and let us consider the log-return $R_t = \ln\left(\frac{S_t+D_t}{S_{t-1}}\right)$, including the dividend D_t. Following Christoffersen, Jacobs, and Ornthanalai (2012), the return process is modeled as follows

$$R_t = r + \left(\lambda_z - \frac{1}{2}\right)h_t + (\lambda_y - \xi)\chi + \sqrt{h_t}Z_t + y_t. \tag{1}$$

The specification Equation (1) contains two sources of uncertainty. The first one, $\{Z_t\}_{t\in\mathbb{Z}}$, is a serially uncorrelated Gaussian process such that $Z_t \sim \mathcal{N}(0,1)$, whereas the second one, $\{y_t\}_{t\in\mathbb{Z}}$, is a compound Poisson process:

$$y_t = \sum_{j=0}^{N_t} X_t^{(j)}, \tag{2}$$

where $N_t \sim \mathcal{P}ois(\chi)$ is a Poisson random variable with constant intensity χ, that is,

$$\mathbb{P}(N_t = j) = \frac{e^{-\chi}\chi^j}{j!}, \quad j = 0, 1, 2, \ldots, \tag{3}$$

and $X_t^{(j)}$ is *IID* random variables such that

$$X_t^{(j)} \sim \mathcal{N}(\mu, \tau^2), \quad j = 0, 1, \ldots, N_t.$$

Moreover, in Equation (1), $\{h_t\}_{t\in\mathbb{Z}}$ is the conditional variance process of the normal innovation, whose dynamics will be specified in the following, r is the risk-free interest rate. The terms $\frac{1}{2}h_t$ and $\xi\chi$ in Equation (1), where $\xi = \exp\left\{\mu + \frac{\tau^2}{2}\right\} - 1$ are the so-called

convexity adjustment, which ensures that λ_z and λ_y are the market prices of risk for the normal and the compound Poisson innovations, respectively. In fact,

$$\mathbb{E}[\exp\{R_t\}|\mathcal{F}_{t-1}] = \exp\{r + \lambda_z h_t + \lambda_y \chi\},$$

where $\mathcal{F}_t = \{R_t, R_{t-1}, R_{t-2}, \ldots\}$ collects the past information. It is straightforward to retrieve the conditional moments of the return process. Specifically, the conditional mean, variance, skewness, and kurtosis are (see Das and Sundaram, 1997):

$$\begin{aligned}
&\mathbb{E}[R_t|\mathcal{F}_{t-1}] = r + \left(\lambda_z - \frac{1}{2}\right)h_t + (\lambda_y - \xi + \mu)\chi, \quad \mathbb{V}[R_t|\mathcal{F}_{t-1}] = h_t + \chi(\tau^2 + \mu^2),\\
&\text{Skew}[R_t|\mathcal{F}_{t-1}] = \frac{\chi(\mu^3 + 3\mu\tau^2)}{(h_t + \chi(\tau^2 + \mu^2))^{3/2}}, \quad \text{Kurt}[R_t|\mathcal{F}_{t-1}] = 3 + \frac{\chi(\mu^4 + 6\mu^2\tau^2 + 3\tau^4)}{(h_t + \chi(\tau^2 + \mu^2))^2}.
\end{aligned} \tag{4}$$

Given the information in $\mathcal{F}_{t-1}$, the conditional density of R_t is obtained by noting that

$$f(R_t) = \sum_{j=0}^{\infty} f(R_t|N_t = j)\mathbb{P}(N_t = j), \tag{5}$$

where the density of the returns conditional to $N_t = j$ and $\mathcal{F}_{t-1}$ is

$$f(R_t|N_t = j) = \frac{1}{\sqrt{2\pi(h_t + j\tau^2)}} e^{-\frac{(R_t - r - (\lambda_z - 1/2)h_t - (\lambda_y - \xi)\chi - \mu j)^2}{2(h_t + j\tau^2)}}. \tag{6}$$

Therefore, according to Equations (3) and (6), we have

$$f(R_t) = \sum_{j=0}^{\infty} f_j(R_t), \tag{7}$$

where

$$f_j(R_t) = \frac{1}{\sqrt{2\pi(h_t + j\tau^2)}} e^{-\frac{(R_t - r - (\lambda_z - 1/2)h_t - (\lambda_y - \xi)\chi - \mu j)^2}{2(h_t + j\tau^2)}} \frac{e^{-\chi}\chi^j}{j!}. \tag{8}$$

Note that, owing to Equation (7), we have

$$\sum_{j=0}^{\infty} \frac{f_j(R_t)}{f(R_t)} = 1. \tag{9}$$

Thus, the conditional log-density is defined as

$$\begin{aligned}
\ln f(R_t) = \ln\Bigg[\sum_{j=0}^{\infty} \exp\Bigg\{ &-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(h_t + j\tau^2)\\
&- \frac{(R_t - r - (\lambda_z - 1/2)h_t - (\lambda_y - \xi)\chi - \mu j)^2}{2(h_t + j\tau^2)} - \chi - \ln(j!) + j\ln(\chi)\Bigg\}\Bigg].
\end{aligned} \tag{10}$$

1.1 The Score-Driven Approach

We specify the conditional variance using a score-driven model (Creal, Koopman,and Lucas, 2011, 2013; Harvey, 2013).

First, in order to ensure positivity, we use the exponential link function:

$$h_t = \exp\{\psi_t\}. \tag{11}$$

Then, we specify the process $\{\psi_t\}_{t\in\mathbb{Z}}$ according to the recursion

$$\psi_{t+1} = \phi_t(\psi_t) = \omega_\psi + \alpha_\psi s^h(R_t, \psi_t) + \beta_\psi(\psi_t - \omega_\psi), \tag{12}$$

where $\omega_\psi \in \mathbb{R}$, $\alpha_\psi \neq 0$, and $\beta_\psi \in \mathbb{R}$ are the intercept, the score weight, and the autoregressive coefficients, respectively, and the driving force $s^h(R_t, \psi_t)$ will be defined in the following. Note that the condition $\alpha_\psi \neq 0$ is needed in order to keep the process in Equation (12) stochastic. Moreover, $\phi_t(\psi_t)$ will also depend on R_t, but, for the sake of simplicity, we omit this dependence.

Let us consider the score of the predictive distribution:

$$\nabla_t^h = \frac{\partial \ln f(R_t)}{\partial h_t}\frac{\partial h_t}{\partial \psi_t}.$$

By differentiating Equation (10) with respect to h_t and by taking into account (Equation 11), we obtain

$$\nabla_t^h = \sum_{j=0}^{\infty}\frac{f_j(R_t)}{f(R_t)}\nabla_{j,t}^h, \tag{13}$$

where

$$\nabla_{j,t}^h = \frac{h_t}{2(h_t + j\tau^2)}\left(\frac{\left(R_t - r - (\lambda_z - 1/2)h_t - (\lambda_y - \xi)\chi - \mu j\right)^2}{h_t + j\tau^2} + 2(\lambda_z - 1/2)\left(R_t - r - (\lambda_z - 1/2)h_t - (\lambda_y - \xi)\chi - \mu j\right) - 1\right). \tag{14}$$

To specify the driving force, we use the approach in Catania (2019) and set $s^h(R_t, \psi_t)$ equal to the following function:

$$\breve{s}^{\,h}(R_t, \psi_t) = \sum_{j=0}^{\infty}\frac{f_j(R_t)}{f(R_t)}\frac{\nabla_{j,t}^h}{\mathbb{E}[-\nabla_{j,t}^{2,h}|N_t = j, \mathcal{F}_{t-1}]}. \tag{15}$$

We have the following result.

Proposition 1 *Formula (15) is equivalent to*

$$\breve{s}^{\,h}(R_t, \psi_t) = \sum_{j=0}^{\infty}\frac{f_j(R_t)}{f(R_t)}\frac{h_t + j\tau^2}{h_t[1 + 2(\lambda_z - 1/2)^2(h_t + j\tau^2)]} \times\left(\frac{\left(R_t - r - (\lambda_z - 1/2)h_t - (\lambda_y - \xi)\chi - \mu j\right)^2}{h_t + j\tau^2} + 2(\lambda_z - 1/2)\left(R_t - r - (\lambda_z - 1/2)h_t - (\lambda_y - \xi)\chi - \mu j\right) - 1\right), \tag{16}$$

with $\mathbb{E}[-\nabla_{j,t}^{2,h}|N_t = j, \mathcal{F}_{t-1}]$ *given by*

$$\mathbb{E}[-\nabla_{j,t}^{2,h}|N_t = j, \mathcal{F}_{t-1}] = h_t^2 \frac{1 + 2(\lambda_z - 1/2)^2(h_t + j\tau^2)}{2(h_t + j\tau^2)^2}.$$

Proof: See the Supplementary Material. □

However, in Figure 1, we plot the score $\breve{s}^{\,h}(R_t, \psi_t)$ as a function of ψ_t. As we see, for low values of ψ_t, $\breve{s}^{\,h}(R_t, \psi_t)$ is unbounded. This is due to the presence in the denominator of Equation (16) of the term h_t, which tends to zero as $\psi_t \to -\infty$, and the term $h_t + j\tau^2$, which tends to zero as $\psi_t \to -\infty$ when $j = 0$, since $h_t = \exp\{\psi_t\}$.

As we will show in Section 5 by a simulation exercise, the unboundedness of the score function may lead to explosive Monte Carlo trajectories, which would make the evaluation of option prices infeasible. From the theoretical standpoint, to ensure the existence of a nondegenerate stationarity and ergodicity region, we will require that the score function $\breve{s}^{\,h}(\cdot,\cdot)$ and its derivative w.r.t. the log-variance be bounded. Thus, we modify Equation (15) as follows:

$$s^h(R_t, \psi_t) = \sum_{j=0}^{\infty} \frac{f_j(R_t)}{f(R_t)} D_j^h(h_t) \frac{\nabla_{j,t}^h}{\mathbb{E}[-\nabla_{j,t}^{2,h}|N_t = j, \mathcal{F}_{t-1}]}, \tag{17}$$

where

$$D_j^h(h_t) = \frac{h_t^3}{(1 + h_t + \chi\tau^2)(h_t + \chi\tau^2)^2(h_t + j\tau^2)(1 + \mu^2 j^2)}. \tag{18}$$

Therefore, according to Equations (17) and (18), the final functional form of the score $s^h(R_t, \psi_t)$ is as follows

$$\begin{aligned} s^h(R_t, \psi_t) = {} & \frac{h_t^2}{(1 + h_t + \chi\tau^2)(h_t + \chi\tau^2)} \sum_{j=0}^{\infty} \frac{f_j(R_t)}{f(R_t)} \frac{1}{[1 + 2(\lambda_z - 1/2)^2(h_t + j\tau^2)]} \frac{1}{1 + \mu^2 j^2} \\ & \times \left(\frac{\left(R_t - r - (\lambda_z - 1/2)h_t - (\lambda_y - \xi)\chi - \mu j\right)^2}{h_t + j\tau^2} \right. \\ & \left. + 2(\lambda_z - 1/2)\left(R_t - r - (\lambda_z - 1/2)h_t - (\lambda_y - \xi)\chi - \mu j\right) - 1 \right), \end{aligned} \tag{19}$$

Remark 1 *The correction term* $D_j(h_t)$ *in* Equation (18) *has been chosen such that: (i) if* $h_t \to 0$ $D_j(h_t) = O(h_t^2)$ *if* $j = 0$ *and* $D_j(h_t) = O(h_t^3)$ *if* $j > 0$. *Thus,* $D_j(h_t)$ *allows us to offset the critical powers (tending to zero) of* h_t *in the denominator of both* Equation (15) *and its derivative (see the proof of Theorem 1 reported in the Supplementary Material); (ii) Both* $s^h(R_t, \psi_t)$ *and its derivative are bounded as* $h_t \to \infty$ *(see again the proof of Theorem 1); (iii) The factor* $1 + \mu^2 j^2$ *at the denominator of* Equation (18) *enhances the convergence of the infinite sum in both* Equation (16) *and its derivative w.r.t. the* log*-variance (see Equation S.5 in the proof of Theorem 1).*

In Figure 1, we show the scores $\breve{s}^{\,h}(R_t, \psi_t)$ and $s^h(R_t, \psi_t)$, as well as the score $\breve{s}^{\,h}(R_t, \psi_t)$ in the case with no jumps ($\chi = 0$ and $\lambda_y = 0$), as functions of the log-volatility ψ_t. This plot clearly highlights the unboundedness of $\breve{s}^{\,h}(R_t, \psi_t)$, and the boundedness and C^1 properties of $s^h(R_t, \psi_t)$ and the score function $\breve{s}^{\,h}(R_t, \psi_t)$ with no jumps.

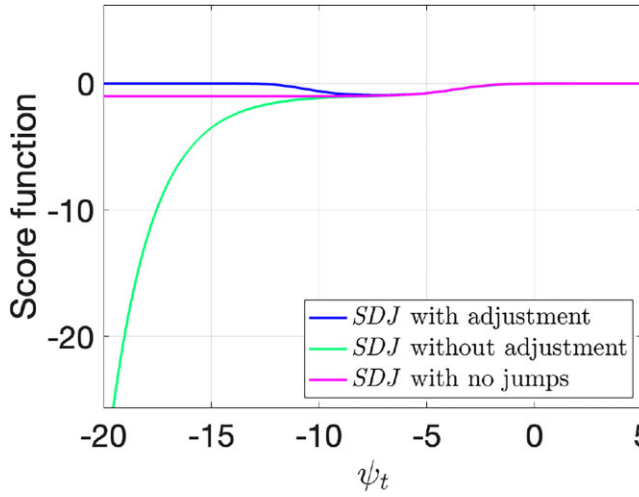

Figure 1 Score functions computed according to Equations (16) and (19). The return R_t is obtained using Equation (1) with $Z_t = 0$ and $y_t = 0$. The scores with and without adjustment are computed using Equations (19) and (16), respectively, with parameters as in Table 1. The score with no jumps is computed using Equation (16) with $\chi = 0$, $\lambda_Y = 0$ and remaining parameters as in Table 1.

1.2 Stationarity and Ergodicity of the *SDJ* Model

With the following theorem, we state a sufficient condition to ensure that the score-driven recursion Equation (12) admits a unique stationary and ergodic solution.

Theorem 1 *Consider the score-driven recursion in* Equation (12)*, with score function* $s^h(R_t, \psi_t)$ *as in* Equation (19)*, and assume that*

$$\mathbb{E}[g(Z_t, y_t)] < \frac{1 - |\beta_\psi|}{|\alpha_\psi|}, \tag{20}$$

where g_t *is defined as*

$$\begin{aligned} g(Z_t, y_t) = {} & \frac{5}{\chi\tau^2}\left(Z_t^2 + \frac{2y_t^2+2}{\chi\tau^2} + 1 + \frac{|Z_t|}{\sqrt{2}} + \frac{|y_t|}{\sqrt{2\chi\tau^2}} + \frac{1}{2\sqrt{2\chi\tau^2}}\right) \\ & + 2\left(Z_t^2 + \frac{y_t^2}{\chi\tau^2} + \frac{\mu^2}{2\tau^4} + \frac{|Z_t|}{2\sqrt{\chi\tau^2}}\left(\sqrt{Z_t^2 + \frac{y_t^2}{\chi\tau^2}} + \frac{|\mu|}{\tau^2}\right) + \frac{1}{2}\right) \\ & \times \left(Z_t^2 + \frac{2y_t^2+2}{\chi\tau^2} + \frac{1}{2\sqrt{2}}\sqrt{Z_t^2 + \frac{2y_t^2}{\chi\tau^2} + \frac{1}{2\chi\tau^2}} + 1\right) \\ & + \frac{1}{\chi\tau^2}\left[\frac{1}{4}\left(Z_t^2 + \frac{2y_t^2+2}{\chi\tau^2} + 1\right) + \frac{3\sqrt{3}}{16\sqrt{2}}\left(\frac{|Z_t|}{\sqrt{2}} + \frac{|y_t|}{\sqrt{2\chi\tau^2}} + \frac{1}{2\sqrt{2\chi\tau^2}}\right)\right] \\ & + |Z_t|\sqrt{Z_t^2 + \frac{2y_t^2+2}{\chi\tau^2} + \frac{1}{2\chi\tau^2}} + \left(Z_t^2 + \frac{2y_t^2+1}{\chi\tau^2} + 1\right) + \frac{1}{\chi\tau^2}\frac{|Z_t|}{2\sqrt{2}}. \end{aligned}$$

Then, for any $\{\psi_t\}_{t\geq 0}$, *such that* $\psi_{t+1} = \phi_t(\psi_t)$ *for all* $t \geq 0$, *there exists a unique stationary and ergodic process* $\{\tilde{\psi}_t\}_{t\in\mathbb{Z}}$, *solution of* Equation (12), *such that* $|\psi_t - \tilde{\psi}_t| \overset{\text{e.a.s.}}{\to} 0$ *as* $t \to \infty$. *Moreover, let* $\tilde{h}_t = \exp\{\tilde{\psi}_t\}$. *Then, the process* $\{R_t\}_{t\in\mathbb{Z}}$ *generated by* Equation (1) *with* $h_t = \tilde{h}_t$ *is stationary and ergodic.*

Proof: See the Supplementary Material. □

We note that condition (20) involves the expectation of a nonlinear function of Z_t and y_t which can be calculated by a Monte Carlo simulation or by numerical approximation.

2 Extensions to Time-Varying Jump Intensity

We now consider two extensions of the score-driven model introduced in Subsection 1.1 in which we allow for a time-varying jump intensity. Specifically, in the return process Equation (1), we substitute χ for χ_t, that is

$$R_t = r + \left(\lambda_z - \frac{1}{2}\right)h_t + (\lambda_y - \xi)\chi_t + \sqrt{h_t}Z_t + y_t. \tag{21}$$

Again, we model the evolution of χ_t by means of a score-driven approach. As done for the conditional variance Equation (11), we consider an exponential link function:

$$\chi_t = \exp\{\delta_t\}, \tag{22}$$

and then we directly model the vector process $\{(\psi_t, \delta_t)\}_{t\in\mathbb{Z}}$. As a first extension of the *SDJ*, we consider the following bivariate recursion:

$$\psi_{t+1} = \omega_\psi + \alpha_\psi s^h(R_t, \psi_t) + \gamma_\psi s^\chi(R_t, \delta_t) + \beta_\psi(\psi_t - \omega_\psi), \tag{23}$$

$$\delta_{t+1} = \omega_\delta + \alpha_\delta s^\chi(R_t, \delta_t) + \gamma_\delta s^h(R_t, \psi_t) + \beta_\delta(\delta_t - \omega_\delta), \tag{24}$$

where $\omega_\delta \in \mathbb{R}$, $\alpha_\delta \neq 0$, and $\beta_\delta \in \mathbb{R}$ have the same interpretation as in Equation (12). The additional parameters $\gamma_\psi \in \mathbb{R}$ and $\gamma_\delta \in \mathbb{R}$ allow us to couple the above recursions by means of the score innovations $s^h(R_t, \psi_t)$ and $s^\chi(R_t, \delta_t)$. We call the model formed by Equations (11), (21), (22), (23), and (24) as the score-driven model with separate dynamic jumps (*SDSDJ-1*).

To specify the new driving force $s^\chi(R_t, \delta_t)$, we use the score of the predictive log-density as defined in Equation (10) with respect to χ_t, that is

$$\nabla_t^\chi = \sum_{j=0}^{\infty} \frac{f_j(R_t)}{f(R_t)} \nabla_{j,t}^\chi, \tag{25}$$

where

$$\nabla_{j,t}^\chi = \chi_t \left(\frac{(\lambda_y - \xi)(R_t - r - (\lambda_z - 1/2)h_t - (\lambda_y - \xi)\chi_t - \mu j)}{h_t + j\tau^2} + \frac{j}{\chi_t} - 1 \right).$$

Therefore, in line with what was done for Equation (17), we compute $s^\chi(R_t, \delta_t)$ according to

$$s^\chi(R_t, \delta_t) = \sum_{j=0}^{\infty} \frac{f_j(R_t)}{f(R_t)} D_j^\chi(\chi_t) \frac{\nabla_{j,t}^\chi}{\mathbb{E}[-\nabla_{j,t}^{2,\chi} | N_t = j, \mathcal{F}_{t-1}]}, \tag{26}$$

where

$$D_j^{\chi}(\chi_t) = \frac{h_t}{(h_t + j\tau^2)(1+\mu^2 j^2)},$$

and $\mathbb{E}[-\nabla_{j,t}^{2,\chi}|N_t = j, \mathcal{F}_{t-1}]$ is computed in Proposition 2.

As a second bivariate model, suggested by a Referee, we also consider

$$\begin{bmatrix} \psi_{t+1} \\ \delta_{t+1} \end{bmatrix} = \begin{bmatrix} \omega_\psi \\ \omega_\delta \end{bmatrix} + \begin{bmatrix} \beta_\psi & \beta_{\psi,\delta} \\ \beta_{\delta,\psi} & \beta_\delta \end{bmatrix} \begin{bmatrix} \psi_t - \omega_\psi \\ \delta_t - \omega_\delta \end{bmatrix} + \begin{bmatrix} \alpha_\psi & \gamma_\psi \\ \gamma_\delta & \alpha_\delta \end{bmatrix} \sum_{j=0}^{\infty} \frac{f_j(R_t)}{f(R_t)} \begin{bmatrix} \mathbb{E}[-\nabla_{j,t}^{2,h}|N_t = j, \mathcal{F}_{t-1}] & \mathbb{E}[-\nabla_{j,t}^{h,\chi}|N_t = j, \mathcal{F}_{t-1}] \\ \mathbb{E}[-\nabla_{j,t}^{\chi,h}|N_t = j, \mathcal{F}_{t-1}] & \mathbb{E}[-\nabla_{j,t}^{2,\chi}|N_t = j, \mathcal{F}_{t-1}] \end{bmatrix}^{-1} \begin{bmatrix} \nabla_{j,t}^{h} \\ \nabla_{j,t}^{\chi} \end{bmatrix}. \tag{27}$$

We can compute Equations (26) and (27) based on the following proposition.

Proposition 2 *Formula (26) is equivalent to*

$$s^{\chi}(R_t, \delta_t) = \sum_{j=0}^{\infty} \frac{f_j(R_t)}{f(R_t)} \frac{h_t}{[h_t + j\tau^2 + \chi_t(\lambda_y - \xi)^2]} \frac{1}{1+\mu^2 j^2} \times \left(\frac{(\lambda_y - \xi)(R_t - r - (\lambda_z - 1/2)h_t - (\lambda_y - \xi)\chi_t - \mu j)}{h_t + j\tau^2} + \frac{j}{\chi_t} - 1 \right). \tag{28}$$

Moreover, $\mathbb{E}[-\nabla_{j,t}^{2,h}|N_t = j, \mathcal{F}_{t-1}]$ *is computed as in Proposition 1, and*

$$\mathbb{E}[-\nabla_{j,t}^{2,\chi}|N_t = j, \mathcal{F}_{t-1}] = \chi_t \left(1 + \frac{\chi_t(\lambda_y - \xi)^2}{h_t + j\tau^2}\right),$$

$$\mathbb{E}[-\nabla_{j,t}^{h,\chi}|N_t = j, \mathcal{F}_{t-1}] = \mathbb{E}[-\nabla_{j,t}^{\chi,h}|N_t = j, \mathcal{F}_{t-1}] = h_t\chi_t \frac{(\lambda_z - 1/2)(\lambda_y - \xi)}{h_t + j\tau^2}.$$

Proof: See the Supplementary Material. □

We label model (27) as *SDSDJ-2*, which differs from *SDSDJ-1* as follows. First, in contrast to what was done in Equations (17) and (26), we do not apply adjustments to the score vector, following the specific suggestion of the Referee. Second, model *SDSDJ-2* employs the full Fisher information matrix to scale the score, whereas in *SDSDJ-1* the off-diagonal terms are neglected. Finally, *SDSDJ-1* assumes $\beta_{\psi,\delta} = \beta_{\delta,\psi} = 0$. Therefore, the *SDSDJ-1* is actually a parsimonious and modified version of the *SDSDJ-2*.

3 Benchmark Models

We will compare the proposed *SDJ*, *SDSDJ-1*, and *SDSDJ-2* models with the GARCH(1,1)-Jump approach (*GARCHJ*) developed by Christoffersen et al. (2008), and the dynamic volatility and separate dynamic jumps (*DVSDJ*) model introduced by Christoffersen, Jacobs, and Ornthanalai (2012).

3.1 The *GARCHJ* Model

The return process is specified according to Equation (1), whereas the dynamics of the conditional variance is as follows:

$$h_{t+1} = \omega_z + \alpha_z \left(Z_t + \frac{y_t}{\sqrt{h_t}} - c_z \sqrt{h_t} \right)^2 + \beta_z h_t.$$

We note that, with respect to the *SDJ* model (see Equations 11 and 12), the *GARCHJ* contains an additional parameter $c_z \in \mathbb{R}$.

3.2 The *DVSDJ* Model

The return process follows Equation (21), and the dynamics of the conditional variance and conditional intensity are described by the following GARCH recursions:

$$h_{t+1} = \omega_z + \beta_z h_t + \frac{\alpha_z}{h_t} \left(\sqrt{h_t} Z_t - c_z h_t \right)^2 + \gamma_z (y_t - e_z)^2, \tag{29}$$

$$\chi_{t+1} = \omega_y + \beta_y \chi_t + \frac{\alpha_y}{h_t} \left(\sqrt{h_t} Z_t - c_y h_t \right)^2 + \gamma_y (y_t - e_y)^2. \tag{30}$$

We note that, with respect to the more parsimonious *SDSDJ-1* model (see Equations 23 and 24), the *DVSDJ* has four additional parameters, namely $c_z, e_z, c_y, e_y \in \mathbb{R}$. The innovation processes $\{Z_t\}_{t\in\mathbb{Z}}$ and $\{y_t\}_{t\in\mathbb{Z}}$ cannot be directly observed based on the time series of the return R_t, but their computation requires an ad-hoc filtering procedure, which, in principle, could introduce an additional source of uncertainty in the parameter estimation.

4 Empirical Results for Returns

In this section, we present empirical results for the S&P500 daily total log-returns time series. Our data span from January 3, 1990 to December 31, 2019 (7559 observations), which are shown in Figure 2. It is well-known that jumps are rare events, so we chose a 30-year time window.

A basic list of descriptive statistics is reported in Table 2. As is clearly seen, the skewness, albeit negative (equal to -0.27), is not very different from zero. However, the kurtosis coefficient (11.85) is particularly high, which points to the advantage of using a model with jumps. Indeed, as documented by Eraker, Johannes, and Polson (2003), jumps yield more realistic tail patterns (with a high kurtosis in the returns) than models in which the volatility is the only stochastic component.

4.1 Maximum Likelihood Estimation

Employing an observation-driven approach allows us to estimate the model parameters by maximum likelihood. The log-likelihood function reads

$$\ell_{returns}(\boldsymbol{\theta}) = \sum_{t=1}^{T} \ln f(R_t), \tag{31}$$

where $f(R_t)$ is the conditional density Equation (7), and $\boldsymbol{\theta}$ is the vector of unknown parameters. For all the considered models, the initial variance h_1 is estimated by maximum likelihood along with all the other unknown parameters. Finally, for practical purposes, it is

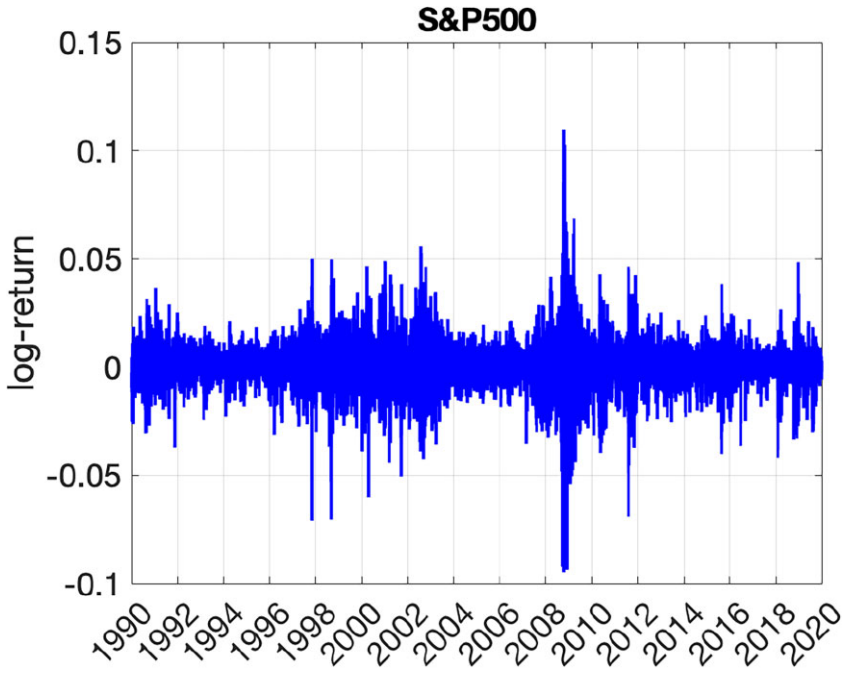

Figure 2 Daily time series of S&P500 total log -returns.

necessary to truncate the infinite sum in Equation (7). Thus, following Christoffersen, Jacobs, and Ornthanalai (2012), we consider only the first fifty jumps[1] for every $t = 1, 2, 3, \ldots$. The estimation results are presented in Table 1.

To validate the models, we focus on the maximized log -likelihood values, the Akaike, and the Bayesian information criteria (labeled *AIC* and *BIC*, respectively). If we consider the models with constant jump intensity, that is, the *SDJ* against the *GARCHJ*, the above three quantities clearly indicate that the *SDJ* model performs better than the *GARCHJ* (for instance, the *BIC* of the *SDJ* is approximately 300 points smaller than that of the *GARCHJ*). This suggests that the driving force $s^b(R_t, \psi_t)$ in Equation (12) is capable of improving the fit.

When we look at the models with time-varying jump intensity, the drastic improvement of the log -likelihood, of the *AIC* and of the *BIC* allows us to conclude that it is important to also take into account the time variations of the jump intensity. Moreover, by comparing the *SDSDJ-1* and *SDSDJ-2* with their GARCH-type counterpart, the *BICs* of the *SDSDJ-1* and *SDSDJ-2* are approximately 100 points lower than the *BIC* of the *DVSDJ*. Therefore, as the *BIC* is very conservative when selecting models with a large number of parameters, the more parsimonious *SDSDJ-1* and *SDSDJ-2* provide a better fit than the more

1 This value is conservative since if we consider a jump intensity equal to ten (approximately ten times the highest value obtained in our empirical analysis) the probability of having more than fifty jumps is almost null (3.62e-20).

Table 2 Descriptive statistics of the daily S&P500 total log-returns

	Mean	Std deviation	Skewness	Kurtosis
S&P500	0.00	0.01	−0.27	11.85

parametrized *DVSDJ*. Finally, the *SDSDJ-1* has a slightly lower *BIC* than the *SDSDJ-2*, even though the latter outperforms the former if we only look at the *AIC* and the log-likelihood.

Passing to the analysis of the parameters, and focusing on the models with constant intensity, we can see that the coefficients of the volatility equation, that is, ω, α, β for the *SDJ* and, additionally, c for the *GARCHJ*, are in line with those that are typically retrieved in financial applications. In particular, consistent with Christoffersen et al. (2008) and Creal, Koopman, and Lucas (2013), we obtain low values for ω and α, and value close to one for the persistence parameter β. Turning to the parameters of the Poisson process, we note that the *SDJ* and the *GARCHJ* yield rather different values for the jump intensity (the former gives $\chi = 0.136$ whereas the latter gives $\chi = 0.043$). Therefore, the *SDJ* weights the jump term more than the *GARCHJ*, which supports other empirical studies, see Bakshi, Cao, and Chen (1997). Finally, the parameter μ in the *GARCHJ* is smaller than in the *SDJ*, whereas the τ^2s are quite similar in both models.

In Figure 3, we plot the conditional volatility $\sqrt{h_t}$ filtered with the models with constant jump intensity, as well as the conditional skewness and kurtosis, which are computed according to Equation (4). To better appreciate the differences in the five models, we zoom in on a specific time interval from January 2004 to December 2011. Similar figures for the entire period from January 1990 to December 2019 are included in the Supplementary Material. It can be clearly seen that *SDJ* and *GARCHJ* yield quite different results. In fact, *SDJ* appears more able to quickly adapt to clusters in the volatility level than the *GARCHJ* (c.f.r. Figure 3, top left and Figure 3, top right). Finally, the *GARCHJ* model implies a conditional skewness as low as −2.35, and a conditional kurtosis as high as 19.51, while the *SDJ* model implies substantially different values, namely a conditional skewness as low as −0.64, and a conditional kurtosis as high as 8.98.

Let us now consider the models with dynamic intensity. As expected, for the *SDSDJ-1*, *SDSDJ-2*, and *DVSDJ* models, the parameters of the dynamic equations show the presence of substantial persistence in both the volatility and the jump intensity processes. Quite interestingly, the *SDSDJ-1* and *SDSDJ-2* yield a higher γ for the Normal component, that is, the contribution of the scaled score innovation relative to the volatility in the log-volatility equation is greater than the contribution of the scaled score innovation relative to the jump intensity in the log-intensity equation.

Furthermore, we note that the jump intensity from the static *SDJ* model is equal to $\log(0.136) \approx -1.99$ which is close to the unconditional mean of the *SDSDJ-2*, that is, −2.317. In contrast, the restricted *SDSDJ-1* has a much higher unconditional jump intensity.

Similarly, in *DVSDJ*, the γ and, additionally, the e parameters (see Equations 29 and 30) confirm the importance of including the compound Poisson process in the dynamic equations for both the variance and the intensity, see Subsection 3.2. The remaining parameters

Table 1 Maximum likelihood estimation results

	SDJ		*SDSDJ-1*		*SDSDJ-2*		*GARCHJ*		*DVSDJ*	
	Normal	Jump	Normal	Jump	Normal	Jump	Normal	Jump	Normal	Jump
Parameters										
ω	−9.632		−7.721	−0.610	−7.579	−2.317	−2e-06		−3e-03	−3e-02
	(1e-02)		(1e-03)	(3e-04)	(2e-03)	(2e-04)	(4e-09)		(2e-08)	(2e-06)
α	0.135		0.011	0.255	−0.021	−0.016	2e-06		2e-06	2e-02
	(1e-03)		(3e-04)	(6e-04)	(3e-04)	(2e-05)	(3e-08)		(2e-09)	(2e-05)
β	0.992		0.909	0.890	0.870	0.982	0.986		0.829	0.985
	(2e-05)		(1e-04)	(2e-04)	(2e-04)	(1e-04)	(2e-04)		(3e-04)	(2e-05)
β_c					0.110	0.008				
					(1e-04)	(2e-04)				
γ			0.308	0.017	0.323	0.037			0.021	0.388
			(6e-04)	(3e-04)	(2e-04)	(6e-04)			(1e-06)	(1e-03)
c							3.085		193.339	70.065
							(2e-02)		(8e-02)	(8e-02)
e									0.117	0.702
									(7e-06)	(1e-05)
λ	−4.076	0.006	22.767	−0.009	22.667	−0.009	2.240	0.006	2.995	−0.001
	(9e-03)	(3e-03)	(3e-02)	(1e-04)	(3e-02)	(1e-04)	(6e-02)	(2e-04)	(1e-06)	(1e-03)
χ		0.136						0.043		
		(3e-04)						(1e-03)		

(continued)

Table 1 (continued)

	SDJ		*SDSDJ-1*		*SDSDJ-2*		*GARCHJ*		*DVSDJ*	
	Normal	Jump	Normal	Jump	Normal	Jump	Normal	Jump	Normal	Jump
μ		−0.002		−0.008		−0.006		−0.011		−0.011
		(3e-04)		(8e-04)		(8e-04)		(2e-04)		(2e-04)
τ^2		8e-05		1e-04		1e-04		2e-004		5e-005
		(1e-07)		(1e-07)		(1e-07)		(5e-07)		(8e-08)
h_1	9e-05		1e-05		6e-05		7e-06		1e-07	
	(8e-03)		(1e-02)		(2e-02)		(1e-08)		(1e-08)	
χ_1				1.110		0.069				4e-003
				(9e-03)		(8e-04)				(2e-04)
log-likelihood	25082.76		25242.62		25250.69		24940.07		25217.55	
AIC	−50147.52		−50457.25		−50469.39		−49860.14		−50399.11	
BIC	−50085.15		−50360.22		−50358.51		−49790.84		−50274.36	

The estimation period spans from January 3, 1990 to December 31, 2019 (7559 observations). The standard errors, reported in parenthesis, are computed by inverting the negative Hessian matrix evaluated at the optimum parameter values.

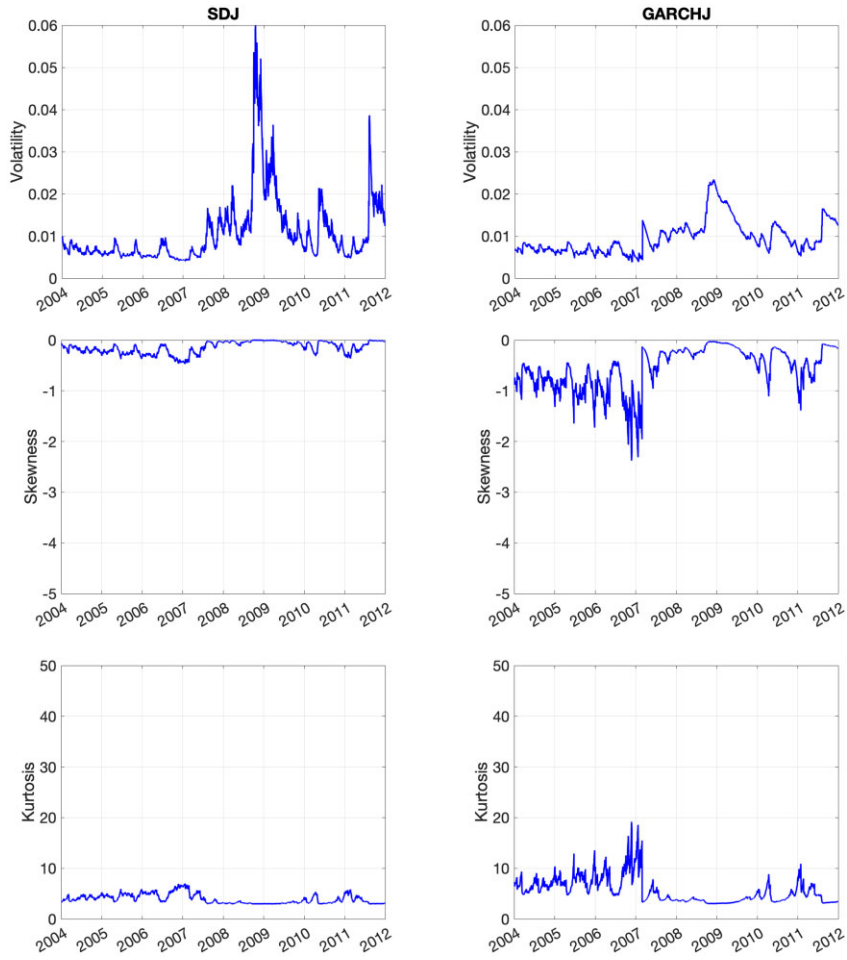

Figure 3 Filtered conditional volatility, skewness, and kurtosis, *SDJ* model (left) and *GARCHJ* model (right).

of the jump component, that is, μ and τ^2, are very similar among the three bivariate models.

We note that for the *SDSDJ-2* model, the estimated values of the α parameters are negative, which can be explained as follows. In Equation (24), we have two linear combinations of the two scores ∇_t^h and ∇_t^γ, which involve not only the α but also the γ parameters. Therefore, there is not a single direction associated with a single score, but two linear combinations of the two scores scaled by the inverse of the full Fisher's Information matrix. So, we have not experienced any issue with the negative α, as they are counterbalanced by the positive γ.

In Figure 4, we display the dynamics of the conditional volatility and the jump intensity, together with the higher conditional moments, filtered with the *SDSDJ-1*, *SDSDJ-2*, and *DVSDJ* models. Although the volatility shows a similar pattern between the three models,

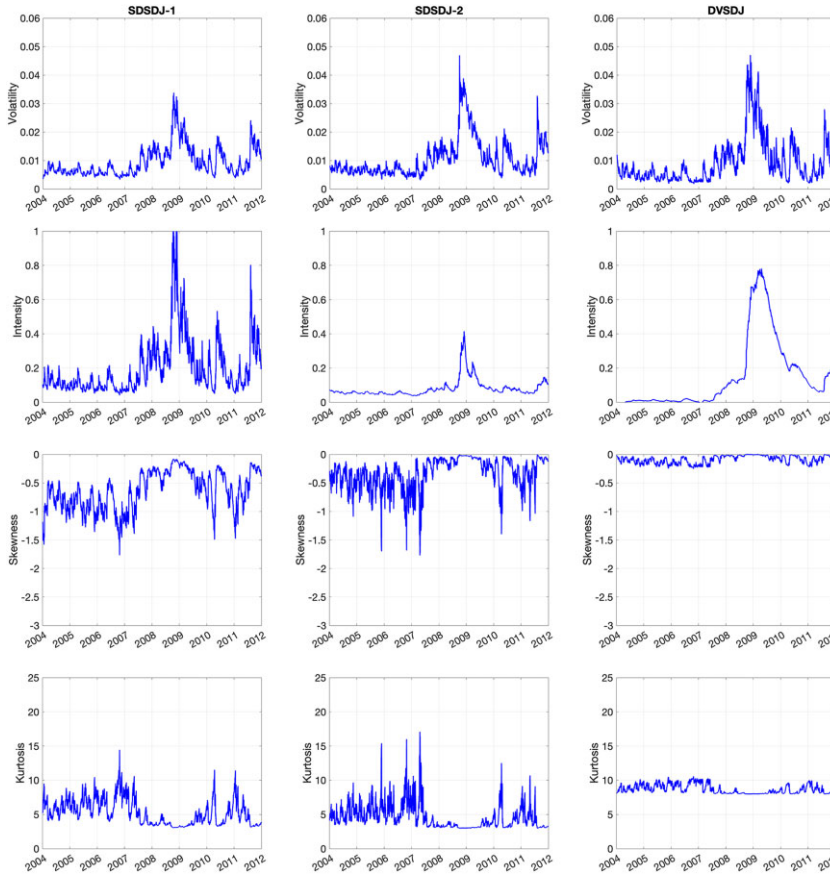

Figure 4 Filtered conditional volatility, jump intensity, skewness, and kurtosis, *SDSDJ-1* model (left), *SDSDJ-2* model (center), and *DVSDJ* model (right).

we note a substantial difference in the jump intensity. In addition, comparing Figures 3 and 4, we note that the volatility paths are similar among the *SDJ*, *SDSDJ-1*, and *SDSDJ-2* models, with the constant jump intensity path being slightly more noisy. This is due to the fact that the bivariate models account for the market risk through a higher jump intensity than that estimated with the monovariate model, assigning less weight to the volatility risk component. In this respect, in Table 1, we note that the *SDSDJ-1* and *SDSDJ-2* provide a jump mean parameter μ much higher in magnitude (and negative) than that of the *SDJ*. For the bivariate specifications, the risk premium λ_y is significantly lower than for the *SDJ*. In particular, in the models with time-varying intensity, the jump premia become negative ($\lambda_y = -0.009$ for both *SDSDJ-1* and *SDSDJ-2*) and, in turn, are compensated for by a higher (and positive) volatility premia ($\lambda_z = 22.767$ for *SDSDJ-1* and $\lambda_z = 22.667$ for *SDSDJ-2*). Finally, we note that jump premia changing sign when passing from the model with constant jump intensity to the model with time-varying jump intensity is also found in Christoffersen, Jacobs, and Ornthanalai (2012).

Regarding the time-varying jump intensity χ_t, we note that the curve provided by model *SDSDJ-2* is less noisy and has lower peaks than in model *SDSDJ-1*. Furthermore, if we look at the whole time interval from 1990 to 2020 (Figure S2 in the Supplementary Material) we see that model *DVSDJ* yields quite a persistent jump intensity in the time interval from 1998 to 2004, whereas *SDSDJ-1* and *SDSDJ-2* exhibit lower persistence. Figure 4 shows that the conditional skewness is quite similar among the three models. On the other hand, the kurtosis reveals that the *SDSDJ-1* and *SDSDJ-2* are able to retrieve a conditional distribution of the S&P500 returns with fatter tails than *DVSDJ*. In particular, the *SDSDJ-1* model provides a conditional kurtosis as high as 15.72, whereas the *DVSDJ* model is as high as 10.04.

4.2 Out-of-Sample Analysis

To evaluate the predictive performances of the rival models, we forecast the 5% and 1% Value-at-Risk (VaR) from one to five days ahead. We test the models during different periods of crisis by splitting the dataset into three subperiods and running separate out-of-sample analyses: From January 1, 1996 to December 31, 2003 (to include the September 11, 2001 crisis), from January 1, 2004 to December 31, 2011 (to include the 2008 financial crisis), and finally, from January 1, 2012 to December 31, 2019, a noncrisis period in the United States.

For each period, we estimate the model parameters using a rolling window equal to 20% of the days in the period. We start by using the first 20% of the daily S&P500 total returns to calibrate the models, leaving the remaining data for the out-of-sample analysis. Then, for each day in the out-of-sample, we re-estimate the model parameters using the past 20% daily observations and compute forecasts from one to five days ahead.

In particular, for the *SDJ* model, at each forecasting time t^*, we simulate 100,000 sample paths of R_{t^*+l}, with $l = 1, 2, \ldots, 5$, according to the return Equation (1), or, equivalently, Equation (21). This is done by randomly generating the innovations $\sqrt{h_{t^*+l}}Z_{t^*+l}$, where h_t follows Equations (11) and (12), $Z_t \sim \mathcal{N}(0,1)$ and y_{t^*+l} is the compound Poisson process in Equation (2), $l = 1, 2, \ldots, 5$. A similar procedure is also used for the *GARCHJ*, *DVSDJ*, *SDSDJ-1*, and *SDSDJ-2* models. Then, we estimate the VaR at both the $\alpha = 5\%$ and $\alpha = 1\%$ levels by computing the empirical quantiles of the 100,000 simulated sample paths of R_{t^*+l}.

In line with Creal, Koopman, and Lucas (2011), we assess the statistical accuracy of the VaR forecasts by performing the proportion of failures test proposed by Kupiec (1995) and the correct conditional coverage test from Christoffersen (1998). The results obtained for the proposed *SDJ*, *SDSDJ-1*, and *SDSDJ-2* models, together with the competing *GARCHJ* and *DVSDJ* models, are reported in Tables 3 and 4.

Focusing on the models with constant intensity, we note that overall the *SDJ* specification shows a lower number of failures than the *GARCHJ*. Both tests show the superiority of the *SDJ* in the period from January 1, 1996 to December 31, 2003 while in the period from January 1, 2012 to December 31, 2019, only the Kupiec (1995) test provides similar evidence. However, the two models perform similarly over the period from January 1, 2004 to December 31, 2011.

The bivariate models, *SDSDJ-1* and *SDSDJ-2*, also perform substantially the same. If we consider the three periods together, with *SDSDJ-1*, the Kupiec (1995) test is only rejected one time and the Christoffersen (1998) test is rejected two times, whereas with *SDSDJ-2*, the Kupiec (1995) test never rejects and the Christoffersen (1998) test rejects one time. In contrast, the performance of the *DVSDJ* is considerably worse, as the Kupiec (1995) and Christoffersen (1998) tests reject seventeen and ten times, respectively (on all periods).

Table 3 Likelihood ratio test statistic of Kupiec (1995), *p*-values in parentheses

	January 1996 to December 2003				
	$\alpha = 5\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	17.980*(1e-05)	1.507(0.219)	0.265(0.606)	0.102(0.749)	0.100(0.748)
SDSDJ-1	0.034(0.853)	0.838(0.359)	1.978(0.656)	0.065(0.797)	0.074(0.792)
SDSDJ-2	0.034(0.853)	0.888(0.327)	1.977(0.656)	0.397(0.528)	0.197(0.656)
GARCHJ	19.372*(1e-05)	2.349(0.117)	6.602*(0.010)	7.492*(0.006)	7.468*(0.006)
DVSDJ	19.192*(1e-05)	2.456(0.107)	1.388(0.2875)	1.842(0.174)	0.995(0.318)
	$\alpha = 1\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	21.823*(2e-06)	16.979*(1e-05)	3.274(0.070)	3.274(0.070)	3.274(0.070)
SDSDJ-1	0.654(0.418)	0.626(0.428)	2.645(0.103)	1.620(0.203)	1.617(0.203)
SDSDJ-2	0.656(0.419)	0.428(0.388)	2.644(0.103)	1.617(0.203)	1.624(0.202)
GARCHJ	19.877*(2e-06)	17.980*(1e-05)	42.051*(1e-11)	38.910*(4e-10)	27.184*(2e-07)
DVSDJ	18.640*(2e-06)	3.274(0.070)	11.518*(6e-04)	12.351*(6e-04)	15.687*(7e-04)
	January 2004 to December 2011				
	$\alpha = 5\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	1.962(0.161)	3.157(0.075)	2.527(0.111)	2.577(0.112)	1.464(0.226)
SDSDJ-1	1.464(0.226)	1.035(0.309)	0.676(0.411)	0.676(0.408)	0.391(0.532)
SDSDJ-2	1.468(0.223)	1.963(0.161)	1.464(0.226)	1.464(0.226)	0.676(0.410)
GARCHJ	5.926*(0.015)	4.801*(0.028)	4.801*(0.028)	4.801*(0.028)	5.811*(0.011)
DVSDJ	4.801*(0.028)	3.801(0.051)	2.954(0.085)	2.954(0.086)	2.955(0.085)
	$\alpha = 1\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	9.038*(0.003)	13.244*(3e-4)	15.557*(8e-05)	13.244*(0.047)	11.068*(0.001)
SDSDJ-1	2.645(0.114)	1.562(0.211)	2.645(0.104)	3.954*(0.046)	2.645(0.103)
SDSDJ-2	2.656(0.101)	1.563(0.211)	2.655(0.102)	1.563(0.211)	2.645(0.103)
GARCHJ	5.468*(0.019)	7.167*(0.007)	7.167*(0.007)	7.167*(0.007)	7.176*(0.008)
DVSDJ	7.167*(0.007)	11.068*(0.001)	11.068*(0.001)	11.068*(0.001)	11.700*(8e-4)
	January 2012 to December 2019				
	$\alpha = 5\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	1.035(0.309)	0.391(0.532)	0.676(0.411)	0.391(0.532)	6.903*(9e-03)
SDSDJ-1	1.033(0.330)	0.050(0.822)	1.451(0.228)	3.592(0.060)	0.391(0.532)
SDSDJ-2	1.035(0.309)	0.059(0.801)	1.926(0.159)	1.592(0.302)	0.398(0.531)
GARCHJ	14.011*(2e-04)	10.231*(1e-03)	12.019*(5e-03)	12.019*(5e-03)	14.011*(2e-03)
DVSDJ	14.684*(1e-04)	2.232(0.135)	0.004(0.949)	1.168(0.279)	2.167(0.141)

(continued)

Table 3 (continued)

	$\alpha = 1\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	0.208(0.647)	0.209(0.647)	0.209(0.648)	2.329(0.127)	4.781*(0.029)
SDSDJ-1	0.198(0.656)	0.208(0.656)	2.745(0.097)	0.224(0.636)	0.224(0.636)
SDSDJ-2	0.207(0.650)	0.205(0.658)	2.328(0.127)	0.647(0.208)	2.328(0.127)
GARCHJ	0.209(0.647)	23.245*(1e-04)	0.080(0.799)	0.209(0.648)	0.927(0.337)
DVSDJ	40.055*(2e-10)	7.855*(5e-03)	11.988*(5e-03)	9.857*(1e-03)	9.858*(2e-03)

The asterisk denotes significance at 95% level.

Overall, results suggest that the *GARCHJ* and the *DVSDJ* models do not provide a very satisfactory representation of the left (and fat) tail of the conditional distribution of the returns. By contrast, the proposed *SDJ*, *SDSDJ-1*, and *SDSDJ-2* models yield a much better fit for extreme losses and thus should be more useful for risk management.

In addition, we compare the accuracy of the VaR forecast using the Model Confidence Set (*MCS*) of Hansen, Lunde, and Nason (2011), and as in González-Rivera, Lee, and Mishra (2004), we measure performance by means of the quantile loss function used in the quantile regression, see Koenker and Bassett (1978), defined as

$$QL_{t^*+l}^{VaR} = \left(\alpha - \mathbb{1}_{\{R_{t^*+l} < VaR_{t^*+l}(\alpha)\}}\right)\left(R_{t^*+l} - VaR_{t^*+l}(\alpha)\right),$$

where $VaR_{t^*+l}(\alpha)$ denotes the VaR on day t at level α. Results are reported in Tables 5–7. As we see, the three score-driven models perform much better than the GARCH counterparts in the period from January 1, 1996 to December 31, 2003, while in the period from January 1, 2004 to December 31, 2011, all five models perform approximately the same. In the period from January 1, 2012 to December 31, 2019, the models with constant intensity do not provide accurate VaR forecasts, while the models with dynamic jump intensity yield satisfactory and similar performances.

5 Monte Carlo Simulations

5.1 Accuracy of the Filters under Model Misspecification

To demonstrate that the score-driven approach is particularly suitable in the case of model misspecification, we perform a Monte Carlo experiment to test the proposed models against their GARCH counterparts. Precisely, we consider yet another model with both stochastic volatility and jumps, namely, the Bates model, where the price S and the variance v satisfy the following stochastic differential equations:

$$dS(t) = \mu S(t) + \sqrt{v(t)}S(t)dW_1(t) + (\eta - 1)S(t)dJ(t), \tag{32}$$

$$dv(t) = \gamma(\theta - v(t)) + \sigma\sqrt{v(t)}dW_2(t), \tag{33}$$

where μ, γ, θ, and σ are constant parameters and W_1 and W_2 are standard Wiener processes, with constant correlation ρ. Moreover, J is a Poisson process with constant intensity parameter λ, independent of W_1 and W_2, and η is a random variable measuring the jump size. Following Chiarella et al. (2009), we assume that η is distributed according to a log-Normal distribution, with probability density given by

Table 4 Likelihood ratio test statistic of Christoffersen (1998), *p*-values in parentheses

	January 1996 to December 2003				
	$\alpha = 5\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	1.521(0.217)	2.232(0.135)	2.232(0.130)	3.838(0.051)	3.691(0.055)
SDSDJ-1	1.281(0.257)	1.650(0.199)	2.069(0.145)	3.102(0.078)	0.806(0.369)
SDSDJ-2	2.518(0.112)	1.650(0.198)	2.209(0.136)	3.100(0.080)	0.809(0.307)
GARCHJ	0.611(0.434)	6.766*(0.009)	6.201*(0.012)	5.870*(0.015)	7.379*(0.007)
DVSDJ	1.901(0.167)	2.044(0.152)	3.290(0.069)	4.553*(0.033)	5.069*(0.024)
	$\alpha = 1\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	0.021(0.884)	0.014(0.905)	0.014(0.906)	2.362(0.124)	2.362(0.126)
SDSDJ-1	2.588(0.108)	2.362(0.124)	0.015(0.905)	0.014(0.906)	0.014(0.905)
SDSDJ-2	2.612(0.106)	1.978(0.159)	0.014(0.905)	0.015(0.905)	0.010(0.905)
GARCHJ	8.355*(0.003)	9.143*(0.002)	7.701*(0.004)	8.399*(0.002)	7.177*(0.008)
DVSDJ	0.031(0.869)	9.133*(0.002)	7.266*(0.005)	8.295*(0.006)	2.328(0.127)
	January 2004 to December 2011				
	$\alpha = 5\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	3.603(0.057)	7.580*(0.006)	3.179(0.074)	3.179(0.074)	1.993(0.158)
SDSDJ-1	0.705(0.401)	2.322(0.127)	2.680(0.102)	2.680(0.102)	3.070(0.079)
SDSDJ-2	0.532(0.465)	3.603(0.057)	1.993(0.158)	4.062*(0.044)	2.680(0.102)
GARCHJ	4.060*(0.043)	3.194(0.074)	3.195(0.073)	3.195(0.073)	3.194(0.073)
DVSDJ	4.060*(0.044)	2.755(0.096)	5.771*(0.016)	2.361(0.124)	5.771*(0.016)
	$\alpha = 1\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	0.908(0.340)	3.196(0.073)	2.755(0.097)	3.195(0.074)	3.685(0.054)
SDSDJ-1	2.119(0.144)	0.262(0.608)	2.119(0.145)	6.300*(0.012)	2.119(0.150)
SDSDJ-2	2.115(0.145)	0.261(0.609)	2.115(0.145)	0.262(0.609)	2.115(0.145)
GARCHJ	1.422(0.232)	4.839*(0.027)	4.839*(0.027)	4.839*(0.028)	4.839*(0.027)
DVSDJ	1.146(0.284)	3.683(0.055)	3.683(0.055)	3.683(0.055)	3.683(0.059)
	January 2012 to December 2019				
	$\alpha = 5\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	7.258*(0.007)	7.155*(0.007)	7.484*(0.006)	9.212*(0.002)	8.149*(0.002)
SDSDJ-1	4.602*(0.043)	1.993(0.158)	2.322(0.115)	2.321(0.127)	2.680(0.102)
SDSDJ-2	3.698(0.054)	1.990(0.159)	0.146(0.701)	0.146(0.701)	0.102(0.749)
GARCHJ	13.382*(0.004)	6.918*(0.008)	6.295*(0.012)	4.321*(0.035)	8.159*(0.003)
DVSDJ	0.405(0.524)	1.466(0.170)	2.233(0.127)	4.321*(0.035)	2.786(0.095)

Table 4 (continued)

	$\alpha = 1\%$				
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
SDJ	4.984*(0.025)	11.880*(5e-04)	8.359*(0.003)	7.573*(0.006)	23.416*(6e-06)
SDSDJ-1	2.119(0.145)	3.683(0.050)	3.194(0.073)	3.194(0.073)	3.638(0.054)
SDSDJ-2	0.405(0.524)	3.195(0.073)	3.194(0.073)	3.196(0.073)	3.680(0.050)
GARCHJ	4.696*(0.028)	9.976*(0.002)	8.356*(0.004)	7.574*(0.006)	20.445*(6e-06)
DVSDJ	4.565*(0.032)	3.195(0.074)	3.195(0.073)	3.195(0.073)	3.635(0.059)

The asterisk denotes significance at 95% level.

Table 5 Left columns: $QL^{VaR}_{t^*+l}$ ratios for *SDSDJ-1*, *SDSDJ-2*, *GARCHJ*, and *DVSDJ* over *SDJ*; right columns: models included in the *MCS* of Hansen, Lunde, and Nason (2011) using $QL^{VaR}_{t^*+l}$ loss function

	January 1996 to December 2003									
	$\alpha = 5\%$									
	$l = 1$		$l = 2$		$l = 3$		$l = 4$		$l = 5$	
	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*
SDJ	1.00	✓	1.00	✓	1.00	✓	1.00	✓	1.00	✓
SDSDJ-1	1.08	✓	1.16	✓	1.05	✓	1.03	✓	1.07	✓
SDSDJ-2	1.08	✓	1.17	✓	1.06	✓	1.02	✓	1.07	✓
GARCHJ	1.18	✓	0.63		0.82		0.90		0.84	
DVSDJ	0.98		1.02	✓	0.98	✓	0.91		0.88	
	$\alpha = 1\%$									
	$l = 1$		$l = 2$		$l = 3$		$l = 4$		$l = 5$	
	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*
SDJ	1.00	✓	1.00	✓	1.00	✓	1.00		1.00	
SDSDJ-1	1.18	✓	1.00	✓	1.15	✓	1.01	✓	1.02	✓
SDSDJ-2	1.18	✓	1.00	✓	1.18	✓	0.98		0.99	
GARCHJ	0.99		0.65		0.86		0.88		0.96	
DVSDJ	0.99		0.75		0.85		0.96		1.02	✓

Values greater than 1 indicate the better out-of-sample performance.

$$f(\eta) = \frac{1}{\sqrt{2\pi}\sigma_j \eta} e^{-\frac{(\ln \eta - \mu_j)^2}{2\sigma_j^2}},$$

where σ_j and μ_j are fixed parameters.

We conduct a simulation exercise where we assume that models Equations (32) and (33) are the DGP, and we test the performance of the misspecified score-driven and GARCH models in

Table 6 Left columns: $QL^{VaR}_{t^*+l}$ ratios for *SDSDJ-1*, *SDSDJ-2*, *GARCHJ*, and *DVSDJ* over *SDJ*; right columns: models included in the *MCS* of Hansen, Lunde, and Nason (2011) using $QL^{VaR}_{t^*+l}$ loss function

	January 2004 to December 2011									
	$\alpha = 5\%$									
	$l = 1$		$l = 2$		$l = 3$		$l = 4$		$l = 5$	
	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*
SDJ	1.00	✓	1.00		1.00	✓	1.00	✓	1.00	✓
SDSDJ-1	1.03	✓	1.01	✓	1.03	✓	1.01	✓	1.01	✓
SDSDJ-2	1.05	✓	1.02	✓	1.02	✓	1.00	✓	1.03	✓
GARCHJ	0.88		1.01	✓	1.01	✓	1.01	✓	1.02	✓
DVSDJ	0.89		1.01	✓	0.99		1.01	✓	0.99	
	$\alpha = 1\%$									
	$l = 1$		$l = 2$		$l = 3$		$l = 4$		$l = 5$	
	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*
SDJ	1.00	✓	1.00	✓	1.00	✓	1.00	✓	1.00	✓
SDSDJ-1	1.03	✓	1.01	✓	1.03	✓	0.95		1.05	✓
SDSDJ-2	1.05	✓	1.01	✓	0.99		0.98		0.98	
GARCHJ	1.08	✓	0.82		0.98		0.99		0.99	
DVSDJ	1.03	✓	0.98	✓	1.03	✓	1.04	✓	1.01	✓

Values greater than one indicate the better out-of-sample performance.

filtering the (unobserved) variance process. We simulate 1000 sample paths with $T = 1000$ observations each by using an Euler time-discretized version of Equations (32) and (33), where the parameters are chosen as in Table 8 and are taken from Ballestra and Sgarra (2010).

For each of these trajectories, we estimate the five competing models by maximum likelihood and then we filter the volatility process. To measure the accuracy in filtering the volatility paths, we adopt the mean absolute error (*MAE*) and the mean squared error (*MSE*) which, for each trajectory, are given by

$$MAE = \frac{1}{T}\sum_{t=1}^{T}|\hat{v}(t) - v(t)|, \quad MSE = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{v}\left(t\right) - v(t)\right)^2,$$

where $\hat{v}(t)$ and $v(t)$ denote the filtered conditional variance and the true variance processes. The results are reported in Table 9.

As we can see, the score-driven models provide a more accurate filtering than the GARCH counterparts, according to both the *MAE* and *MSE*. Moreover, the *SDSDJ-1* and *SDSDJ-2* perform better than the *SDJ*. For a visual interpretation of the results, we report in Figures 5 and 6 the true and the filtered volatility, averaged over the 1000 simulated trajectories.

Table 7 Left columns: $QL^{VaR}_{t^*+l}$ ratios for *SDSDJ-1*, *SDSDJ-2*, *GARCHJ*, and *DVSDJ* over *SDJ*; right columns: models included in the *MCS* of Hansen, Lunde, and Nason (2011) using $QL^{VaR}_{t^*+l}$ loss function

	January 2012 to December 2019									
	$\alpha = 5\%$									
	$l = 1$		$l = 2$		$l = 3$		$l = 4$		$l = 5$	
	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*	$QL^{VaR}_{t^*+l}$	*MCS*
SDJ	1.00		1.00		1.00		1.00		1.00	
SDSDJ-1	0.98		1.05	✓	1.07	✓	1.06	✓	1.00	✓
SDSDJ-2	1.01	✓	1.05	✓	1.07	✓	1.05	✓	1.01	✓
GARCHJ	0.99		0.97		0.93		0.91		0.92	
DVSDJ	1.01	✓	1.00	✓	1.01	✓	0.95		1.00	✓
	$\alpha = 1\%$									
	$l = 1$		$l = 2$		$l = 3$		$l = 4$		$l = 5$	
	QL^{VaR}	*MCS*	QL^{VaR}	*MCS*	QL^{VaR}	*MCS*	QL^{VaR}	*MCS*	QL^{VaR}	*MCS*
SDJ	1.00		1.00		1.00		1.00		1.00	
SDSDJ-1	1.03	✓	1.05	✓	1.03	✓	1.01	✓	1.08	✓
SDSDJ-2	1.02	✓	1.05	✓	1.03	✓	1.02	✓	1.07	✓
GARCHJ	0.98		0.97		0.92		0.97		0.94	
DVSDJ	0.98		0.98		1.02	✓	1.03	✓	1.00	✓

Values greater than one indicate the better out-of-sample performance.

Table 8 DGP model parameters

μ	γ	θ	σ	ρ	λ	σ_j	μ_j
0.03	2	0.04	0.25	−0.5	0.2	0.4	−0.58

Table 9 Average *MAE* and *MSE* relative to *SDJ* over 1000 Monte Carlo trajectories

	SDJ	*SDSDJ-1*	*SDSDJ-2*	*GARCHJ*	*DVSDJ*
Avg. MAE	1.0000	0.9762	0.9833	1.0252	1.0197
Avg. MSE	1.0000	0.9867	0.9625	1.0354	1.0049

5.2 Unboundedness of the Scaled Score without Adjustment

Finally, in Equation (19), we discuss the use of the modified scaled scores in the jump sum. As we previously noted, in models *SDJ* and *SDSDJ-1*, we made this adjustment because otherwise the scaled score is unbounded for small values of the variance. To see why this modification is needed, we perform the following Monte Carlo exercise. For each of the

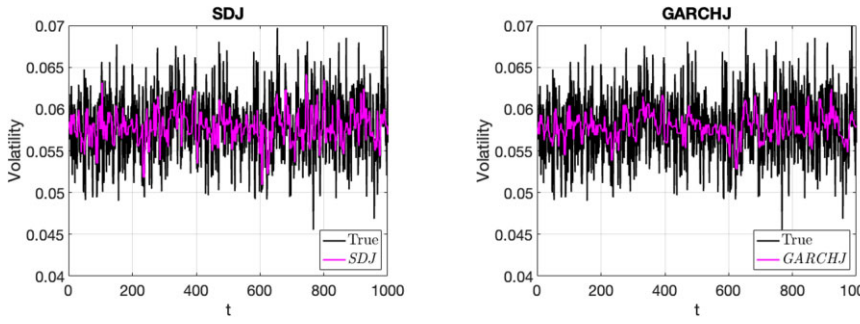

Figure 5 Volatilities computed averaging over 1000 simulated trajectories for the *SDJ* and *GARCHJ*.

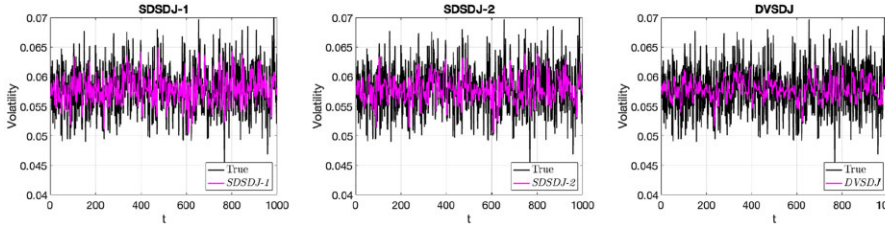

Figure 6 Volatilities computed averaging over 1000 simulated trajectories for the *SDSDJ-1*, *SDSDJ-2*, and *DVSDJ*.

five models, we simulate 1000 trajectories of $T = 365$ daily observations each based on the parameters reported in Table 1. Then, for each of these paths, we check if the variance explodes, that is, if $h_t > 100000$. The results are reported in Table 10.

We see that model *SDSDJ-2*, which does not employ the score adjustment, generates explosive trajectories, whereas the remaining models do not. In the following analysis, as we compute option prices based on simulated trajectories of the S&P500 returns, the *SDSDJ-2* model is excluded.

6 Option Valuation

Jumps play an important role in option modeling, as emphasized by the rich literature that follows the work of Bates (1991). Option pricing is therefore an important context for assessing models with jumps, and thus we evaluate the performances of the *SDJ*, *SDSDJ-1*, *GARCHJ*, and *DVSDJ* approaches in computing the prices of a wide number of options with several different strikes and maturities. To this aim, following Christoffersen et al. (2008), Christoffersen, Jacobs, and Ornthanalai (2012, 2013), and Bormetti et al. (2020), we first perform the risk neutralization of the models based on a suitable pricing kernel. For the sake of brevity, we only describe this for *SDSDJ-1*, since the risk-neutralized *SDJ* is readily obtained as a special case. Instead, for the risk neutralization of the *GARCHJ* and the *DVSDJ* models, the interested reader is referred to Christoffersen et al. (2008) and Christoffersen, Jacobs, and Ornthanalai (2012).

Table 10 Ratio of cases in which the variance process h_t explodes (#explsions), that is, $\eta\tau >$ 100,000, and average explosion time

	SDJ	*SDSDJ-1*	*SDSDJ-2*	*GARCHJ*	*DVSDJ*
#explosions	0	0	0.61	0	0
Explosion time	*n.e.*	*n.e.*	186	*n.e.*	*n.e.*

The notation *n.e.* indicates that no explosion was detected.

6.1 Risk Neutralization

Following Christoffersen, Jacobs, and Ornthanalai (2012), we consider the pricing kernel with affine dynamics

$$M_{t+1} = \frac{\exp\{-r - \gamma V_{t+1} - (\gamma + \gamma_j) y_{t+1}\}}{\mathbb{E}[\exp\{-\gamma V_{t+1} - (\gamma + \gamma_j) y_{t+1}\} | \mathcal{F}_t]}, \tag{34}$$

where $V_{t+1} = \sqrt{h_{t+1}} Z_{t+1}$ is the nonstandardized normal innovation. We guess the conditional Radon–Nikodym derivative in the following form

$$\frac{\mathrm{d}Q_{t+1}}{\mathrm{d}P_{t+1}} = \frac{\exp\{\Lambda_V V_{t+1} + \Lambda_y y_{t+1}\}}{\mathbb{E}[\exp\{\Lambda_V V_{t+1} + \Lambda_y y_{t+1}\} | \mathcal{F}_t]}, \tag{35}$$

and we impose

$$\frac{\mathrm{d}Q_{t+1}}{\mathrm{d}P_{t+1}} = \exp\{r\} M_{t+1}, \tag{36}$$

so that, upon Equations (34)–(36), we obtain $\Lambda_V = -\gamma$ and $\Lambda_y = -\gamma - \gamma_j$.

Therefore, we have the following result.

Proposition 3 *Let us consider the return process specified by* Equation (21) *and the Radon–Nikodym derivative* (Equation 35). *Then, the equivalent martingale measure Q exists if and only if the coefficients* Λ_V *and* Λ_y *satisfy*

$$\lambda_z \quad + \Lambda_V = 0, \tag{37}$$

$$\lambda_y - \xi + \exp\left\{\Lambda_y \mu + \frac{\Lambda_y^2 \tau^2}{2}\right\}\left(\exp\left\{\mu + \Lambda_y \tau^2 + \frac{\tau^2}{2}\right\} - 1\right) = 0. \tag{38}$$

Proof: See the Supplementary Material. □

Finally, the following proposition yields the dynamics of the risk-neutralized processes $\{R_t\}_{t\in\mathbb{Z}}$ and $\{h_t\}_{t\in\mathbb{Z}}$.

Proposition 4 *Under the risk-neutral measure Q, the dynamics of the return process is as follows*

$$
\begin{aligned}
R_t &= \ln\left(\frac{S_t}{S_{t-1}}\right) = r - \frac{1}{2}h_t + (\lambda_y - \xi)\chi_t + V_t^* + y_t^*, \\
h_t &= \exp\{\psi_t\}, \quad \chi_t = \exp\{\delta_t\} \\
\psi_{t+1} &= \omega_\psi + \alpha_\psi s^{h,*}(V_t^*, y_t^*, \psi_t) + \gamma_\psi s^{\chi,*}(V_t^*, y_t^*, \delta_t) + \beta_\psi(\psi_t - \omega_\psi), \\
\delta_{t+1} &= \omega_\delta + \alpha_\delta s^{\chi,*}(V_t^*, y_t^*, \delta_t) + \gamma_\delta s^{h,*}(V_t^*, y_t^*, \psi_t) + \beta_\delta(\delta_t - \omega_\delta),
\end{aligned}
$$

where (under the measure Q) $V_t^* \sim \mathcal{N}(0, h_t)$, y_t^* *is a compound Poisson process with intensity* $\chi_t^* = \chi_t \exp\{\Lambda_y \mu + \frac{\Lambda_y^2 \tau^2}{2}\}$, *jump size mean* $\mu^* = \mu + \Lambda_y \tau^2$, *and jump size variance* τ^2, *whereas* $s^*(V_t^*, y_t^*, \psi_t)$ *and* $s^*(V_t^*, y_t^*, \delta_t)$ *are defined as in Supplementary Equations (S.23) and (S.24), respectively.*

Proof: See the Supplementary Material. □

6.2 Computing Option Prices

Let us consider a European option (either Call or Put) on the underlying asset S_t with maturity T, and let $\Pi(S_T)$ denote its payoff. The option price at a generic time $t < T$, which we denote by O_t, is computed as follows, see Heston and Nandi (2000), Christoffersen, Jacobs, and Ornthanalai (2013):

$$
O_t = \exp\{-r(T-t)\}\mathbb{E}^{\mathbb{Q}}[\Pi(S_T)|\mathcal{F}_t]. \tag{39}
$$

As in Christoffersen, Jacobs, and Ornthanalai (2012, 2013), the conditional expectation in Equation (39) is evaluated by a Monte Carlo simulation, since it cannot be obtained in closed form. Precisely, for each option price, we simulate 1000 trajectories.

7 Option Valuation Empirics

We consider European Put and Call options on the S&P500 index, with data retrieved from Thomson Reuters Eikon Datastream, which reports the prices of options with maturities starting from 2015. Thus, we select the S&P500 options with maturities ranging from 2015 to 2019, and we consider the time series of their daily prices from September 18, 2013 to November 27, 2019.

As a common practice, we apply several exclusion filters to obtain the final panel of Put and Call option contracts. Specifically, following Bormetti et al. (2020), we keep only the options with time-to-maturity between 14 and 365 days. Following Christoffersen, Jacobs, and Ornthanalai (2012), we select only out-of-the-money Put and Call options (we compute the moneyness as K/S_t, where K is the strike price and S_t is the underlying index level), and we filter out illiquid quotes by selecting only the six most liquid strikes at each maturity, and we consider option quotes only on Wednesday. Finally, as in Bakshi, Cao, and Chen (1997), we further remove price quotes lower than 3.8$. In Table 11, we report the resulting number of option prices on the S&P500 index sorted by moneyness and maturity.

7.1 Joint Maximum Likelihood Estimation

As argued by Ornthanalai (2014), fitting and evaluating a dynamic model to a panel of option prices could be really challenging, especially if the goal is to estimate the parameters also using returns. In fact, several studies have documented the importance of taking into

Table 11 Number of S&P500 option prices considered, sorted by moneyness and day-to-maturity (*DTM*)

	Number of option prices			
	Maturity			
Moneyness	$14 \leq DTM \leq 50$	$50 < DTM \leq 150$	$150 < DTM \leq 365$	All
$0.8 \leq K/S_t \leq 0.9$	194	957	535	1686
$0.9 < K/S_t \leq 1.02$	429	1329	1328	3086
$1.02 < K/S_t \leq 1.2$	284	855	1162	2301
All	907	3141	3025	7073

account both the sources of information, see Chernov and Ghysels (2000), Pan (2002), Santa-Clara and Yan (2010). However, we note that the joint estimation method proposed by Ornthanalai (2014) can be easily adapted for our score-driven models, and provides a feasible and efficient procedure for gaining more insights about option pricing performances.

To apply this method, we need to define a suitably weighted log-likelihood function that considers both the daily time series of log-returns used in Section 4 and the panel of option contracts considered in this section. While for the daily returns the log-likelihood, $\ell_{returns}(\boldsymbol{\theta})$, is already available in Equation (31), for the option pricing error structure further assumptions are required. Formally, if we have N option prices, the option pricing error for the i-th contract, u_i, for $i = 1, \ldots, N$, is defined as the relative implied volatility loss function:

$$u_i = \frac{IV_i^{MKT} - IV_i^{MOD}}{IV_i^{MKT}}, \tag{40}$$

where IV_i^{MKT} and IV_i^{MOD} denote the market and the model implied volatilities of the i-th option, respectively, computed according to the popular Black-Scholes model, see Black and Scholes (1973). Then, following Christoffersen, Jacobs, and Ornthanalai (2012) and Ornthanalai (2014), the likelihood associated with the error on implied volatility is computed by using a Gaussian specification:

$$\ell_{options}(\boldsymbol{\theta}) = -\frac{N}{2}\ln\left(2\pi\sigma_u^2\right) - \frac{1}{2}\sum_{i=1}^{N}\frac{u_i^2}{\sigma_u^2}.$$

It is important to note that the number N of data points available from the option panel is different from the number T of daily returns, see Section 4. Therefore, according to Christoffersen, Jacobs, and Ornthanalai (2012) and Ornthanalai (2014), to assign an equal weight to returns and option prices, we consider the following weighted joint (total) log-likelihood:

$$\ell_{joint}(\boldsymbol{\theta}) = \frac{T+N}{2}\frac{\ell_{returns}(\boldsymbol{\theta})}{T} + \frac{T+N}{2}\frac{\ell_{options}(\boldsymbol{\theta})}{N}. \tag{41}$$

The joint estimation results are presented in Table 12. As we can see, the score-driven approaches, namely the *SDJ* and *SDSDJ-1*, provide a higher joint log-likelihood than the *GARCHJ* and the *DVSDJ*.

Table 12 Joint maximum likelihood estimation results

	SDJ		*SDSDJ-1*		*GARCHJ*		*DVSDJ*	
	Normal	Jump	Normal	Jump	Normal	Jump	Normal	Jump
Parameters								
ω	−11.142		−9.194	−2.946	−1e-06		1e-06	−7e-07
	(1e-02)		(8e-03)	(7e-04)	(3e-09)		(2e-08)	(2e-06)
α	0.150		0.023	0.070	2e-06		2e-06	1e-03
	(1e-03)		(8e-04)	(8e-04)	(3e-08)		(4e-09)	(3e-05)
β	0.993		0.969	0.963	0.980		0.560	0.585
	(2e-05)		(8e-04)	(8e-04)	(3e-04)		(1e-04)	(3e-05)
γ			0.010	0.009			0.010	0.780
			(9e-04)	(6e-04)			(2e-06)	(3e-03)
c					3.093		116.380	70.065
					(2e-02)		(8e-02)	(8e-02)
e							0.004	0.885
							(7e-06)	(1e-05)
λ	−4.732	0.002	5.804	0.005	2.241	0.005	0.885	0.002
	(9e-03)	(3e-03)	(3e-02)	(5e-04)	(6e-02)	(2e-04)	(4e-06)	(7e-03)
χ		0.136				0.047		
		(3e-04)				(1e-03)		
μ		−0.001		−0.008		−0.012		−0.024
		(3e-04)		(8e-04)		(2e-04)		(5e-04)
τ^2		8e-05		6e-04		2e-04		1e-05
		(1e-07)		(1e-07)		(5e-07)		(8e-08)
h_1	9e-05		4e-05		7e-06		1e-07	
	(8e-03)		(1e-02)		(1e-08)		(1e-08)	
χ_1				0.137				1e-03
				(9e-03)				(2e-04)
Λ	4.732	−11.570	−5.804	−77.575	−2.241	−14.013	−0.885	−3.117
log-likelihood	21719.76		24686.62		20480.07		20810.55	

The estimation period spans from January 3, 1990 to December 31, 2019 (7559 observations) for the returns, and spans from September 18, 2013 to November 27, 2019 (7073 prices) for the option prices. The standard errors reported in parenthesis are computed by inverting the negative Hessian matrix evaluated at the optimum parameter values.

7.2 Option Valuation Results

To assess the performance of the four competing models, we follow Majewski, Bormetti, and Corsi (2015) and Alitab et al. (2019), and employ the (percentage) implied volatility root mean square error:

$$IVRMSE(\%) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(IV_i^{MKT} - IV_i^{MOD}\right)^2} \times 100, \tag{42}$$

where, as in Subsection 7.1, N is the total number of option prices considered (7073) while IV_i^{MKT} and IV_i^{MOD} denote the market and the model implied volatilities of the i-th option, respectively.

Table 13 Implied volatility root *MSE* in percentage points (*IVRMSE* [%])

Panel A: *IVRMSE* (%) overall			
SDJ	*SDSDJ-1*	*GARCHJ*	*DVSDJ*
4.24	3.89	5.90	5.39

Panel B: *IVRMSE* (%) sorted by moneyness				
Moneyness	*SDJ*	*SDSDJ-1*	*GARCHJ*	*DVSDJ*
$0.8 \leq K/S_t \leq 0.9$	2.81	2.47	4.85	3.12
$0.9 < K/S_t \leq 1.02$	3.78	3.48	6.93	5.87
$1.02 < K/S_t \leq 1.2$	6.95	4.45	8.88	8.67

Panel C: *IVRMSE* (%) sorted by day-to-maturity (*DTM*)				
Maturity	*SDJ*	*SDSDJ-1*	*GARCHJ*	*DVSDJ*
$14 \leq DTM \leq 50$	5.43	4.60	5.77	5.78
$50 < DTM \leq 150$	4.20	3.64	5.59	5.31
$150 < DTM \leq 365$	4.69	2.65	6.27	5.12

Panel D: *IVRMSE* (%) sorted by *VIX* level				
Level	*SDJ*	*SDSDJ-1*	*GARCHJ*	*DVSDJ*
$VIX \leq 22$	1.65	1.60	1.76	1.78
$22 < VIX \leq 30$	4.74	4.46	5.27	5.51
$30 < VIX$	9.73	9.68	11.81	9.77

The option time series spans from September 18, 2013 to November 27, 2019.

The results obtained, shown in Table 13, reveal that the score-driven specifications yield a better fitting than the GARCH approaches. In fact, for the *SDJ* and the *SDSDJ-1*, the overall pricing errors are equal to 4.24% and 3.89%, respectively, whereas for the *GARCHJ* and the *DVSDJ*, the overall pricing errors are equal to 5.90% and 5.39%, respectively. The picture does not change if we sort the options for moneyness or maturity. In particular, the *IVRMSE* obtained using the *SDSDJ-1* is always smaller than or equal to 4.45%, whereas the *GARCHJ* and the *DVSDJ* provide errors also equal to 8.88% and 8.67%, respectively.

Furthermore, to investigate the model performances at different volatility scenarios, for every day in the sample period, we sort the option price according to the value of the *VIX* on that day, see Panel D of Table 13. The results obtained reveal that score-driven models perform better than the GARCH counterparts for every level of *VIX* considered.

8 Conclusions

To better cope with the empirical features of stock prices, and to incorporate relevant information about the probability distribution of returns, this article presents a novel approach

that couples the predicting power and the flexibility of score-driven models with a jump process. In particular, the proposed *SDJ* model assumes that returns are affected by two sources of randomness, namely a Gaussian process and a compound Poisson process, and employs a score-driven framework for the dynamics of the conditional variance of the former. The resulting specification benefits from the fact that the process used to incorporate the jumps is fully integrated with the dynamics of the conditional variance, and that it has a straightforward estimation by maximum likelihood. In doing so, we contribute by extending the score-driven literature to better cope with financial stylized facts.

We have performed a theoretical investigation of the *SDJ* model, establishing a sufficient condition for its stationarity and ergodicity. In addition, we also present two extensions of the *SDJ* approach, namely, the *SDSDJ-1* and *SDSDJ-2* models, which take into account the time-varying features of the jump intensity.

An empirical analysis focusing on the S&P500 time series is conducted, to test the *SDJ*, *SDSDJ-1*, and *SDSDJ-2* against the *GARCHJ* and *DVSDJ* approaches of Christoffersen et al. (2008) and Christoffersen, Jacobs, and Ornthanalai (2012). Overall, our score-driven models turn out to be superior to their GARCH counterparts both in terms of goodness-of-fit and out-of-sample VaR prediction, confirming that the proposed specifications are particularly suitable for capturing the non-Gaussian behavior of returns due to outliers and occasional bursts.

Finally, we test how our approach works in a challenging empirical setting such as option pricing. We find that our score-driven models provide a smaller pricing error and more reliable implied volatility surfaces than the GARCH benchmarks. To the best of our knowledge, this is the first time that the score-driven methodology is used for valuing derivatives. Thus, the present article opens the way to possible future applications of score-driven models in option pricing.

Supplementary Material

Supplementary data are available at *Journal of Financial Econometrics* online.

The Supplementary Material contains proofs, the derivation of the conditional Fisher information and two additional figures showing additional empirical results.

References

Alitab, D., G. Bormetti, F. Corsi, and A. A. Majewski. 2019. A Jump and Smile Ride: Jump and Variance Risk Premia in Option Pricing. *Journal of Financial Econometrics* 18: 121–157.

Bakshi, G., C. Cao, and Z. Chen. 1997. Empirical Performance of Alternative Option Pricing Models. *The Journal of Finance* 52: 2003–2049.

Ballestra, L. V., and C. Sgarra. 2010. The Evaluation of American Options in a Stochastic Volatility Model with Jumps: An Efficient Finite Element Approach. *Computers & Mathematics with Applications* 60: 1571–1590.

Bates, D. S. 1991. The Crash of '87: Was It Expected? The Evidence from Options Markets. *The Journal of Finance* 46: 1009–1044.

Black, F., and M. Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81: 637–654.

Blasques, F., S. J. Koopman, and A. Lucas. 2014. Stationarity and Ergodicity of Univariate Generalized Autoregressive Score Processes. *Electronic Journal of Statistics* 8: 1088–1112.

Blasques, F., S. J. Koopman, and A. Lucas. 2015. Information-Theoretic Optimality of Observation-Driven Time Series Models for Continuous Responses. *Biometrika* 102: 325–343.

Bormetti, G., R. Casarin, F. Corsi, and G. Livieri. 2020. A stochastic Volatility Model with Realized Measures for Option Pricing. *Journal of Business & Economic Statistics* 38: 856–871.

Catania, L. 2019. Dynamic Adaptive Mixture Models with an Application to Volatility and Risk. *Journal of Financial Econometrics* 19: 531–564.

Chan, W. H., and J. M. Maheu. 2002. Conditional Jump Dynamics in Stock Market Returns. *Journal of Business & Economic Statistics* 20: 377–389.

Chernov, M., and E. Ghysels. 2000. A Study towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation. *Journal of Financial Economics* 56: 407–458.

Chiarella, C., B. Kang, G. H. Meyer, and A. Ziogas. 2009. The Evaluation of American Option Prices under Stochastic Volatility and Jump-Diffusion Dynamics Using the Method of Lines. *International Journal of Theoretical and Applied Finance* 12: 393–425.

Christoffersen, P., K. Jacobs, and C. Ornthanalai. 2012. Dynamic Jump Intensities and Risk Premiums: Evidence from S&P500 Returns and Options. *Journal of Financial Economics* 106: 447–472.

Christoffersen, P., K. Jacobs, and C. Ornthanalai. 2013. GARCH Option Valuation: Theory and Evidence. *The Journal of Derivatives* 21: 8–41.

Christoffersen, P., K. Jacobs, C. Ornthanalai, and Y. Wang. 2008. Option Valuation with Long-Run and Short-Run Volatility Components. *Journal of Financial Economics* 90: 272–297.

Christoffersen, P. F. 1998. Evaluating Interval Forecasts. *International Economic Review* 39: 841–862.

Cox, D. R., G. Gudmundsson, G. Lindgren, L. Bondesson, E. Harsaae, P. Laake, K. Juselius, and S. L. Lauritzen. 1981. Statistical Analysis of Time Series: Some Recent Developments. *Scandinavian Journal of Statistics* 8: 93–115.

Creal, D., S. J. Koopman, and A. Lucas. 2011. A Dynamic Multivariate Heavy-Tailed Model for Time-Varying Volatilities and Correlations. *Journal of Business & Economic Statistics* 29: 552–563.

Creal, D., S. J. Koopman, and A. Lucas. 2013. Generalized Autoregressive Score Models with Applications. *Journal of Applied Econometrics* 28: 777–795.

Das, S. R., and R. K. Sundaram. 1997. "Taming the Skew: Higher-Order Moments in Modeling Asset Price Processes in Finance." Working paper 5976, National Bureau of Economic Research.

Duan, J. C., P. Ritchken, and Z. Sun. 2006. Approximating GARCH-Jump Models, Jump-Diffusion Processes, and Option Pricing. *Mathematical Finance* 16: 21–52.

Engle, R. F., M. K. Hansen, A. K. Karagozoglu, and A. Lunde. 2021. News and Idiosyncratic Volatility: The Public Information Processing Hypothesis. *Journal of Financial Econometrics* 19: 1–38.

Eraker, B., M. Johannes, and N. Polson. 2003. The Impact of Jumps in Volatility and Returns. *The Journal of Finance* 58: 1269–1300.

González-Rivera, G., T.-H. Lee, and S. Mishra. 2004. Forecasting Volatility: A Reality Check Based on Option Pricing, Utility Function, Value-at-Risk, and Predictive Likelihood. *International Journal of Forecasting* 20: 629–645.

Guégan, D., F. Ielpo, and H. Lalaharison. 2013. Option Pricing with Discrete Time Jump Processes. *Journal of Economic Dynamics and Control* 37: 2417–2445.

Gürkaynak, R. S., B. Kisacikoğlu, and J. H. Wright. 2020. Missing Events in Event Studies: Identifying the Effects of Partially Measured News Surprises. *American Economic Review* 110: 3871–3912.

Hansen, P. R., A. Lunde, and J. M. Nason. 2011. The Model Confidence Set. *Econometrica* 79: 453–497.

Harvey, A. C. 2013. *Dynamic Models for Volatility and Heavy Tails*. Cambridge: Econometric Society Monograph - Cambridge University Press.

Heston, S. L., and S. Nandi. 2000. A Closed-Form GARCH Option Valuation Model. *The Review of Financial Studies* 13: 585–625.

Jeon, Y., T. H. McCurdy, and X. Zhao. 2021. News as Sources of Jumps in Stock Returns: Evidence from 21 Million News Articles for 9000 Companies. *Journal of Financial Economics* (in press).

Koenker, R., and G. Bassett. 1978. Regression Quantiles. *Econometrica* 46: 33–50.

Kupiec, P. H. 1995. Techniques for Verifying the Accuracy of Risk Measurement Models. *The Journal of Derivatives* 3: 73–84.

Maheu, J. M., and T. H. McCurdy. 2004. News Arrival, Jump Dynamics, and Volatility Components for Individual Stock Returns. *The Journal of Finance* 59: 755–793.

Majewski, A. A., G. Bormetti, and F. Corsi. 2015. Smile from the Past: A General Option Pricing Framework with Multiple Volatility and Leverage Components. *Journal of Econometrics* 187: 521–531.

Ornthanalai, C. 2014. Lévy Jump Risk: Evidence from Options and Returns. *Journal of Financial Economics* 112: 69–90.

Pan, J. 2002. The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study. *Journal of Financial Economics* 63: 3–50.

Santa-Clara, P., and S. Yan. 2010. Crashes, Volatility, and the Equity Premium: Lessons from S&P 500 Options. *The Review of Economics and Statistics* 92: 435–451.

Vlaar, P. J. G., and F. C. Palm. 1993. The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroscedasticity, and Jumps. *Journal of Business & Economic Statistics* 11: 351–360.