Journal of Financial Econometrics, 2021, 1–42
doi: 10.1093/jjfinec/nbab020
Article

Risk Premia and Lévy Jumps: Theory and Evidence*

Hasan Fallahgoul[1], Julien Hugonnier[2] and Loriano Mancini [3,*]

[1]Monash University, [2]Swiss Finance Institute, EPFL and CEPR and [3]Swiss Finance Institute, USI Lugano

Address correspondence to Loriano Mancini, Università della Svizzera italiana (USI), Institute of Finance, Via Buffi 13, CH-6900 Lugano, Switzerland, or e-mail: loriano.mancini@usi.ch.

Received August 24, 2020; revised June 28, 2021; editorial decision July 6, 2021; accepted July 23, 2021

Abstract

We develop a novel class of time-changed Lévy models, which are tractable and readily applicable, capture the leverage effect, and exhibit pure jump processes with finite or infinite activity. Our models feature four nested processes reflecting market, volatility and jump risks, and observation error of time changes. To operationalize the models, we use volume-based proxies of the unobservable time changes. To estimate risk premia, we derive the change of measure analytically. An extensive time series and option pricing analysis of sixteen time-changed Lévy models shows that infinite activity processes carry significant jump risk premia, and largely outperform many finite activity processes.

Key words: Lévy jumps, time changes, tempered stable law, time series, option pricing

JEL classification: C5, G12

Joint modeling of options and asset returns has proven to be a challenging task, as underscored by a voluminous literature that started in the 1970s. The non-normality and path dependency of asset returns affect option prices through complex risk premia, which require sophisticated models to describe their dynamics. Recently, time-changed Lévy processes have emerged as highly flexible models for options and asset returns.[1] While a number of

* We thank David Bates, Yan Dolinsky, Damir Filipovic, Elise Gourier, Dilip Madan, Chay Ornthanalai, Svetlozar Rachev, Stoyan Stoyanov, Fabio Trojani, and seminar participants at Monash University, University of St. Gallen, SoFiE and QMF for comments. Financial support from the Swiss Finance Institute, the Swiss National Science Foundation [154445], and the Centre for Quantitative Finance and Investment Strategies sponsored by BNP Paribas is gratefully acknowledged.

1 Lévy processes are characterized by independent increments over nonoverlapping time periods. Modeling asset returns as Lévy increments can easily generate non-normal returns, but cannot reproduce the stochastic volatility. The latter can be captured by stochastic time changes, namely by

Table 1. Specification of the Lévy subordinators

s_t	α	c	$\mathbb{V}[s_t]$	$t\psi(u)$
Gamma	0	ν	t/ν	$-tc\log(1-iu/\nu)$
Inverse Gaussian	$\frac{1}{2}$	$\sqrt{\nu/\pi}$	$t/(2\nu)$	$-tc2\sqrt{\pi}(\sqrt{\nu-iu}-\sqrt{\nu})$
Tempered stable	$(0, 1)$	$\frac{\nu^{1-\alpha}}{-\alpha\Gamma(-\alpha)}$	$\frac{(1-\alpha)}{\nu}t$	$tc\Gamma(-\alpha)((\nu-iu)^{\alpha}-\nu^{\alpha})$

The Lévy measure of s_t is $\Pi(dx) = c\exp(-\nu x)/x^{1+\alpha}dx$ in (4). The constant c is such that $\mathbb{E}[s_t] = t$. The variance of s_t is $\mathbb{V}[s_t]$. The characteristic exponent of s_t is $\psi(u) = \log(\mathbb{E}[e^{ius_t}])/t$. The Gamma function is denoted as $\Gamma(\cdot)$ and $i = \sqrt{-1}$.

studies analyze their theoretical properties, extensive empirical analysis of time-changed Lévy models are scant. This gap in the literature is significant because, given their generality, time-changed Lévy models can offer new insights on jump and volatility risk premia, and inform asset pricing.

The objective of this article is to develop a novel class of time-changed Lévy models, and to provide an extensive empirical analysis of these models and associated risk premia. Our models are characterized by four nested processes, capturing market, volatility, and jump risks, as well as observation error on the time change. Despite the apparent complexity, the models retain a high degree of analytical tractability because all processes are independent from each other. The key is to work with nested Lévy characteristic functions that have analytical forms. To operationalize the models, we use simple volume-based proxies of the time changes, exploiting the strong positive association between volume and volatility.[2] To study risk premia, we derive the change of measure explicitly. As it turns out, the conditional expectation of asset returns needs to feature a particular functional form of the time change to allow flexible specifications of risk premia.

We investigate empirically sixteen time-changed Lévy models. We obtain twelve models by combining three infinite activity Lévy subordinators (listed in Table 1) to generate jumps in returns; one-factor or two-factor processes to generate stochastic volatility; with or without observation error on our volume-based proxies of the time changes. We also consider four finite activity models in which the Lévy subordinator follows a Cox process, which is a Poisson-type process with stochastic intensity, combined with one-factor or two-factor volatility processes, and with or without observation error on the volume-based proxies of the time changes.

We begin the empirical analysis by fitting the time-changed Lévy models to daily Standard & Poor's 500 index returns from 1950 to 2019. Because time changes are not observable, time series fitting of time-changed Lévy models is challenging. Consequently, we view the time series analysis of the models as an important empirical contribution of this

stochastically changing the clock on which the Lévy process is run. Intuitively, stretching, and compressing the stochastic time change generates low and high volatility periods.

2 Our method is reminiscent of the approach adopted by Corsi et al. (2013). They use the realized volatility computed from high-frequency data as a proxy of the unobservable volatility driving their conditionally Gaussian discrete time models. We use readily available volume-based proxies of the time changes and embed these proxies in the characteristic function of time-changed Lévy processes.

article. We can summarize the main findings as follows. First, infinite activity models substantially outperform many finite activity and generalized autoregressive conditional heteroskedasticity (GARCH) models. The latter models are used as a benchmark, and known to provide a good fit to daily market returns. We trace the high performance of infinite activity models to the flexibility of their Lévy measure. Second, taking into account the observation error on the volume-based proxies uniformly improves the model fitting of market returns, irrespective of the Lévy model used. Third, among the sixteen time-changed Lévy models, the best performing model features a tempered stable subordinator and a two-factor volatility process, and allows for observation error on the volume-based proxies of the time change. In terms of likelihood contributions, the tempered stable subordinator and the observation error on the volume-based proxies are more important than the two-factor volatility process to achieve accurate fitting.

Next, we study the option pricing performance of the time-changed Lévy models. To do so, we derive a necessary and sufficient no-arbitrage condition, which appears to be novel in the literature. This condition places a joint restriction on Brownian risk premium and risk neutral Lévy measure, but does not involve volatility risk premia. To carry out the empirical option pricing analysis, we consider weekly cross-sections of index options spanning the sample period 1996–2019. Our main empirical findings from the option pricing analysis are as follows. First, time-changed Lévy models based on tempered stable subordinators and two-factor volatility processes substantially outperform several other finite and infinite activity models. Volume-based proxies of the time changes improve any model's option pricing performance, even though measurement errors have no risk premium. Second, infinite activity processes carry substantial jump risk premium (JRP). For most estimated models, the risk neutral Lévy measure is nearly twice the physical Lévy measure. This difference is economically large. For example, in the best performing model, a one-standard deviation increase in the Lévy jump under the physical and risk neutral measures induce an annual log-return of -0.88% and -3.68%, respectively. Third, although we allow for the *most* flexible specification of the volatility risk premium (VRP) that preserves the square-root feature of the volatility process, simple affine specifications appear to approximate well the estimated functional form of the VRP. This empirical finding suggests that jump risk, and not volatility risk, largely drives the complex relation between physical and risk neutral measures.

This article contributes to two strands of literature. There exists an extensive research on fitting time series models to asset returns. Prominent examples include Eraker et al. (2003) and Bates (2006). In this literature, there is nearly no time series study of time-changed Lévy models. A notable exception is Bates (2012) who fits various time-changed Lévy models to market returns to quantify crash risk. One impediment that may have restrained empirical applications of time-changed Lévy models is that the stochastic time change is not observable. We contribute to this literature by developing a novel class of time-changed Lévy models, and analyzing empirically both the time series fitting and the option pricing performance of our models.[3]

3 Carr et al. (2002) estimate Lévy models using the time series of stock returns, but their models do not feature stochastic volatility. Bakshi et al. (2008) and Bakshi and Wu (2010) estimate time-changed Lévy models using underlying asset prices and options. However, their analysis focuses on a specific model and covers a short time period.

Another strand of literature develops option pricing models and studies risk premia. Recent contributions include Christoffersen, Jacobs, and Ornthanalai (2012), Andersen, Fusari, and Todorov (2015b), Calvet et al. (2015), and Bardgett, Gourier, and Leippold (2019). Many of the proposed models belong to the class of affine jump diffusion models of Duffie, Pan, and Singleton (2000) in which jumps are modeled as Poisson-type processes. To allow for general jump structures and stochastic volatility, some theoretical studies have developed option-pricing models based on time-changed Lévy processes.[4] Carr et al. (2003) study several processes that feature explicit characteristic functions. Carr and Wu (2004) introduce a novel technique to capture leverage effects in time-changed Lévy processes. In their setting, Brownian increments and time changes can be correlated, significantly complicating the analysis of the resulting models.[5] We therefore follow a different approach and capture the leverage effect by a suitable specification of the return drift.

In the option pricing literature, the two most closely related studies to our work are Huang and Wu (2004) and Ornthanalai (2014), which we discuss in turn. Huang and Wu (2004) build on Carr and Wu (2004) and introduce a very general class of time-changed Lévy models in which different time changes can drive diffusive and jump components. However, their empirical option pricing analysis focuses only on the one-factor Heston model as stochastic volatility process; see their Subsection E in Section II. In addition, their calibrated risk neutral models are not informative about risk premia or time series dynamics. We complement their work by developing time-changed Lévy models that feature two-factor stochastic volatility processes, and we investigate their time series fitting, option pricing performance, and risk premia.

Ornthanalai (2014) uses a GARCH setup to develop models of asset returns driven by Brownian increments and Lévy jumps. His discrete time GARCH-Lévy models can generate the equivalent effect of random time change on Lévy processes. We use a different approach and fully specify time-changed Lévy models in continuous time. Importantly, time changes induced by our models do not follow discrete time GARCH dynamics. This means that using GARCH models to capture time changes would generate model inconsistencies in our setting. Notice that using realized volatilities as in Corsi, Fusari, and La Vecchia (2013) or Andersen, Fusari, and Todorov (2015a) would lead to similar problems. We therefore introduced volume-based proxies of time changes, and formally embed these proxies in the time-changed Lévy models via the characteristic functions. Finally, whereas Ornthanalai focuses his analysis on the equity risk premium (ERP), we center our analysis on volatility and jump risk premia.

The article is organized as follows. Section 1 introduces our novel class of time-changed Lévy models. Section 2 presents the time series analysis of time-changed Lévy models and other benchmark models. Section 3 studies risk neutralization and option pricing from a theoretical perspective. Section 4 presents the empirical option pricing results and estimated risk premia. Section 5 concludes.

4 Schoutens (2003) reviews derivative applications of classic time-changed Lévy models.

5 To develop their technique, Carr and Wu (2004) use the stopping time property of time changes. However, none of the processes they propose as time changes satisfy this assumption, which makes difficult the empirical application of their technique; see Fallahgoul and Nam (2020).

1 Model

This section introduces our class of time-changed Lévy models and discusses how to make the models operational. The price of the financial asset (e.g., index) under consideration evolves in continuous time and is given by

$$S_t = S_0 e^{(r-\delta)t+X_t}$$

where r and δ are constants that represent, respectively, the risk free rate of interest and the dividend yield of the asset, and

$$X_t = \log\left(e^{\delta t} S_t / S_0\right) - rt$$

models excess log-returns, under the usual assumption that dividends are continuously reinvested. To reduce notational complexity, we use interchangeably X_t or $X(t)$, and likewise for any other process that we encounter. When no confusion arises, we omit subscripts.

We model the log-return X_t as a time-changed Lévy process[6]

$$X_t = \gamma s(Y_t) + \sigma W(s(Y_t)) + \beta Y_t \tag{1}$$

where γ, σ, and β are constants, and $W(t)$ is a standard Brownian motion under the physical measure $\mathbb{P}$. The processes $s(t)$ and $Y(t)$ are $\mathbb{P}$-subordinators, that is, positive, non-decreasing, right-continuous, with left limit processes and initial value zero, and model the stochastic time changes. Intuitively, stochastic time changes reflect the "business time", which is influenced by trading activities and news arrivals, rather than the calendar time t. The two subordinators $s(t)$ and $Y(t)$ reproduce jumps and stochastic volatility of asset returns, respectively, which are ultimately a reflection of trading activities. As discussed below, the parameter β allows for a flexible change of measure, while negative values of γ induces the leverage effect and conditional asymmetries of the return distribution, as increases in $s(Y_t)$ tend to produce negative returns. All processes in (3) are independent from each other, which ensure a significant tractability of the model.

Time-changed Lévy processes are usually constructed by time changing once a Lévy process; see, for example, Carr et al. (2003). We take a different, more flexible approach, and use the two time changes $s(t)$ and $Y(t)$ to specify the dynamic of the log-return X_t. The process $s(t)$ is an infinite activity $\mathbb{P}$-subordinator, which means that its expected number of jumps is infinite in any finite time interval, induces jumps in X_t and is defined by its characteristic exponent[7]

$$\psi(u) = \frac{1}{t}\log \mathbb{E}[e^{ius(t)}] = \int_0^\infty (e^{iux} - 1)\Pi(dx) \tag{2}$$

6 A Lévy process is a continuous time stochastic process with independent increments over non-overlapping time intervals. Lévy processes in finance are typically used to model log asset prices, with the Brownian motion in the Black–Scholes model being a well-known example. Carr and Wu (2004) and Bates (2012) provide concise reviews of Lévy processes.

7 Bertoin (1996) introduces the terminology of characteristic exponent, Carr et al. (2003) call ψ the log characteristic function at unit time and Wu (2006) uses the terminology of cumulant exponent.

where $i = \sqrt{-1}$, u is real and $\Pi(dx)$ is its Lévy measure on $(0, \infty)$, that is, the expected number of jumps of size dx per unit of time. The process $Y(t)$ is an absolutely continuous $\mathbb{P}$-subordinator that we model as

$$Y(t) = \int_0^t y_s ds \tag{3}$$

for some non-negative, mean reverting process y_t that is interpreted as the rate of time change. The randomness of y_t reproduces the stochastic volatility of the asset, and the mean reversion of y_t induces volatility persistence, which are well-documented empirical features of asset returns. We follow the standard practice of normalizing the time change (Carr and Wu, 2004), and impose that the two subordinators have $\mathbb{P}$-expectation $\mathbb{E}[Y(t)] = \mathbb{E}[s(t)] = t$. That is, the stochastic time changes are an unbiased reflection of calendar time t. The normalization is achieved through parameter restrictions, reducing the number of parameters to estimate.

The model in (1) encompasses most models of asset returns in the literature. Varying the specification of $s(t)$ and $Y(t)$ generate a wide range of processes. Absent any time change, that is, imposing $s(t) = Y(t) = t$, the log-return X_t would be a Brownian motion with drift, and therefore S_t would follow a geometric Brownian motion as in the Black–Scholes model. This model would generate random returns, but neither jumps nor stochastic volatility, and would not be able to fit asset returns. Imposing that y_t in (3) follows a square-root process, $\gamma = -1/2$, $\sigma = 1$, $\beta = 0$ and $s(t) = t$, the model simplifies to the Heston model without leverage effect. Setting $\beta = 0$, $Y(t) = t$ and $s(t)$ to a Gamma process with mean rate t, the model reduces to the Variance Gamma model (Madan and Seneta, 1990). Under the same restrictions as above but letting $s(t)$ be an inverse Gaussian process, reduces the model to the Normal Inverse Gaussian model (Barndorff-Nielsen, 1997). Jump diffusion, stochastic volatility models, such as the Bates model, cannot be reached by imposing parameter restrictions to (1). These models can be easily obtained by changing the specification of the log-return to $X_t = \gamma s_t + \sigma W(Y_t) + \beta t$, where the three terms capture jump, diffusion, and drift component, respectively. In a straightforward extension of the model in (1) and similarly to Huang and Wu (2004), the terms $\gamma s(Y_t)$ and $W(s(Y_t))$ could be driven by two different subordinators. This extension would not hinder the analytical tractability of the models, and the theory developed below could be easily adapted to such models. However, this extension is not appealing for time series applications, as it could make the empirical identification of time changes rather weak.

To allow for a general specification of the jump component in X_t, we construct the $\mathbb{P}$-subordinator s_t from the tempered stable Lévy measure[8]

$$\Pi(dx) = c\frac{\exp(-\nu x)}{x^{1+\alpha}}dx, \quad x \in (0, \infty) \tag{4}$$

where $c, \nu > 0$, and $\alpha \in [0, 1)$. Because $\mathbb{E}[s_t] = t$ for all time t, we have that $\mathbb{E}[s_1] = \int_0^\infty x\Pi(dx) = 1$, which implies that $c = \nu^{1-\alpha}/\Gamma(1-\alpha)$, where $\Gamma(\cdot)$ is the Gamma function. The parameter α is the so-called tail exponent. As discussed in Carr et al. (2003), Bates (2012) and others, α controls the activity of "small jumps" and the speed at which the

8 Rachev et al. (2011) and Fallahgoul and Loeper (2021) provide overviews of tempered stable process.

Lévy measure approaches zero as the jump size x increases. If $\alpha < 1$, then s_t has infinite activity. When $\alpha \in \{0, 1/2\}$, the Lévy measure in (4) specializes to the Lévy measure of Gamma (Madan and Seneta, 1990) and Inverse Gaussian (Barndorff-Nielsen and Shephard, 2001) processes, respectively. The function $\exp(-\nu x)$ with $\nu > 0$ is the so-called tempering function and serves the purpose of dampening the Lévy measure to ensure that the moments of s_t are finite. In our empirical analysis below, we find strong evidence against infinite variation processes, which are characterized by $\nu = 0$ (and $\alpha > 1/2$), and therefore we do not consider such processes.

We estimate the Lévy subordinators that are most popular in the theoretical literature, namely Gamma ($\alpha = 0$), Inverse Gaussian ($\alpha = 1/2$), and Tempered Stable ($\alpha \in (0, 1)$). Table 1 summarizes their specifications. Column 3 reports the value of the scaling parameter c, as a function of the other parameters in the Lévy measure, to ensure that $\mathbb{E}[s_t] = t$. Column 4 reports the variance $\mathbb{V}[s_t]$, which is linear in time t. Column 5 reports the characteristic exponent in (2) of s_t, which takes a relatively simple form.

A well-known empirical regularity of index return distributions is that the left tail is higher than the right tail, that is, index returns exhibit negative skewness. Although we do not directly model positive and negative return jumps, the model in (1) generates negatively skewed return distributions when $\gamma < 0$. For example, when s_t follows a Tempered Stable subordinator, the centered third moment of X_t is

$$\frac{\partial^3}{\partial u^3} \log \mathbb{E}[\exp(uX_t)]|_{u=0} = \frac{(1-\alpha)((2-\alpha)\gamma^3 + 3\gamma\nu\sigma^2)}{\nu^2} t \tag{5}$$

which is indeed a negative quantity according to our parameter estimates, as γ is estimated to be negative. Because $\nu, \sigma > 0$ and $0 < \alpha < 1$, only when $\gamma = 0$ the return distribution is symmetric. Figure 5 shows the log density of X_t under $\mathbb{P}$ for a time-changed Lévy model based on the Tempered Stable subordinator. The negative skewness of the density is evident.

To determine the stochastic volatility of the log-return in (1), Appendix A proves the following proposition.

Proposition 1.1. *The instantaneous variance v_t is given by*

$$v_t = \lim_{\Delta \to 0} \frac{\mathbb{V}_t[X_{t+\Delta} - X_t]}{\Delta} = (\sigma^2 + \gamma^2 \mathbb{V}[s_1]) y_t \tag{6}$$

where $\mathbb{V}_t$ denotes the time-t conditional variance.

Modeling the rate of time change y_t is therefore equivalent to modeling the instantaneous variance v_t, up to the positive constant $(\sigma^2 + \gamma^2 \mathbb{V}[s_1])$. We consider two models for y_t, namely the one-factor and two-factor Heston models. The one-factor Heston model (SV1) is given by

$$dy_t = \kappa_y (1 - y_t) dt + \sigma_y \sqrt{y_t} dW_t^y \tag{7}$$

while the two-factor Heston model (SV2) is given by

$$dy_t = \kappa_y(m_t - y_t)dt + \sigma_y\sqrt{y_t}dW_t^y$$
$$dm_t = \kappa_m(1 - m_t)dt + \sigma_m\sqrt{m_t}dW_t^m \tag{8}$$

where κ_y, κ_m, σ_y, and σ_m are non-negative constant, and W_t^y and W_t^m are independent $\mathbb{P}$-Brownian motions. The SV2 model reduces to the SV1 model by imposing $\kappa_m = \sigma_m = 0$ and $m_0 = 1$. Because of its analytical tractability, SV1 is often used for modeling stochastic volatility, as in Huang and Wu (2004). However, various studies have shown that two factors (one fast moving and one slow moving) as in SV2 are necessary to capture volatility dynamics.[9] A straightforward model extension would be to add a jump component to (7) or (8), as in Eraker (2004) or Broadie, Chernov, and Johannes (2007). Because the jump component would not enter the characteristic function of the X_t as a conditioning variable, our method (developed below) could be directly applied to such models as well. Given that SV2 already features two factors and is in line with empirical evidence, we leave such a model extension to future work.

Because all processes in (1), namely $W(t)$, $s(t)$, and $Y(t)$, are independent, it is not immediately clear that the model would produce the so-called leverage effect, that is, the empirical phenomenon that asset returns and volatility changes are often negatively correlated.[10] Appendix A proves the following proposition.

Proposition 1.2. *The leverage effect in the model in* (1) *is given by the covariance between the asset return and the instantaneous variance change*

$$\mathrm{Cov}_0[X_t - X_0, v_t - v_0] = (\beta + \gamma)(\sigma^2 + \gamma^2\mathbb{V}[s_1])\mathrm{Cov}_0[Y_t, y_t]$$

where $\mathrm{Cov}_0[Y_t, y_t]$ *is positive for SV1 and SV2 models.*

Therefore, whenever $\beta + \gamma < 0$, the model in (1) reproduces the leverage effect.[11] In essence, the leverage effect is induced by the terms βY_t and $\gamma s(Y_t)$ in (1). No additional model requirements are necessary to reproduce the leverage effect. In the Carr and Wu (2004) setting, the leverage effect is captured by correlating $W(t)$ and $Y(t)$. To accommodate this correlation, they derive a complex-valued "leverage-neutral" measure to compute the characteristic function of log-returns. This theoretical result would require the time change

9 The consensus in the literature is that two factors are necessary (and probably sufficient) to model the volatility process; see Engle and Rangel (2008), Corradi et al. (2013), Filipović et al. (2016), and Aït-Sahalia et al. (2020).

10 The leverage effect was introduced by Black (1976), who suggested that a large negative return increases the financial and operating leverage, and raises equity return volatility; see also Christie (1982). Alternative economic interpretations based on risk premia and volatility feedback effects have been suggested. For example, an anticipated increase in volatility commands a higher rate of return from the asset, which is achieved by a fall in the asset price; French et al. (1987), and Campbell and Hentschel (1992). Bekaert and Wu (2000) provide a discussion of the leverage effect. More recently, Aït-Sahalia et al. (2013) estimate the magnitude of the effect using intraday data.

11 When $\beta + \gamma < 0$, the conditional expected excess return can be positive, $\mathbb{E}_0[e^{X_t} - 1] > 0$, which is the case according to our model estimates, even though the conditional expected excess log-return in (1) is negative, $\mathbb{E}_0[X_t] = (\beta + \gamma)\mathbb{E}_0[Y_t] < 0$.

$Y(t)$ to be a stopping time and adapted to the original filtration. However, none of the processes commonly used as time changes satisfy the assumption of being a stopping time, which significantly complicates the empirical application of leverage-neutral measures.

A key feature of the model in (1) is that the characteristic function of X_t can be computed in a relatively straightforward way. Denote by $\mathbb{E}_0$ the $\mathbb{P}$-expectation conditional on X_0 as well as y_0 or (y_0, m_0) depending on whether y_t follows the SV1 or SV2 model, respectively.

Proposition 1.3. *The characteristic function of X_t in* (1) *is given by*

$$\begin{aligned}\Phi(u; X_t) = \mathbb{E}_0[\exp(iuX_t)] &= \mathbb{E}_0\left[\exp\left(iu\beta Y_t + i\left(u\gamma + iu^2\frac{\sigma^2}{2}\right)s(Y_t)\right)\right] \\ &= \mathbb{E}_0\left[\exp\left(i\left(u\beta - i\psi\left(u\gamma + iu^2\frac{\sigma^2}{2}\right)\right)Y_t\right)\right] \\ &= \mathbb{E}_0[\exp\left(iq(u)\textstyle\int_0^t y_s ds\right)].\end{aligned} \tag{9}$$

The second equality follows by subconditioning on the whole path of the processes Y_t and $s(Y_t)$; the third equality follows from the second equality by subconditioning on the whole path of the process Y_t with ψ denoting the characteristic exponent of s_t; the fourth equality follows from (3) after setting

$$q(u) = u\beta - i\psi\left(u\gamma + iu^2\frac{\sigma^2}{2}\right). \tag{10}$$

To calculate the characteristic function in (9), we note that this expression is reminiscent of a bond-type pricing formula, where $iq(u)y_s$ plays the role of an instantaneous interest rate. To compute this expression, we use indeed tools developed in the fixed income literature. Because SV1 and SV2 are affine models (Duffie et al., 2000), the characteristic function of X_t is exponentially affine in the state variables, and is given by

$$\Phi(u; X_t) = \exp(A_j(t; iq(u)) + B_j(t; iq(u))y_0 + C_j(t; iq(u))m_0) \tag{11}$$

where $j = 1, 2$ for SV1 and SV2 models, respectively.

For the SV1 model, the functions $A_1(t; iq(u))$ and $B_1(t; iq(u))$ are available in closed form and $C_1(t; iq(u)) = 0$, which is a well-known result from the fixed income literature; see Appendix B. For the SV2 model, the characteristic function has no analytic solution (Grasselli and Tebaldi, 2008). In Appendix C, we develop an efficient method to compute the functions $A_2(t; iq(u))$, $B_2(t; iq(u))$ and $C_2(t; iq(u))$. Our method delivers the unknown functions as power series solutions of a system of first order differential equations. Interestingly, the coefficients of the power series have analytic expressions in terms of the model parameters and thus need to be derived only once, making computation time virtually zero in model estimation and option pricing calculations. In contrast, numerical solutions of differential equations have to be re-computed for each set of the parameter values during

a likelihood search. We conducted various numerical experiments that confirm the high accuracy of our method; see end of Appendix C.[12]

Unfortunately, despite the analytical tractability of time-changed Lévy processes, the model in (1) is not operational. The main reason is that the characteristic function of X_t in (11) depends on the rate of time change y_0 and possibly m_0, which are both unobservable. Furthermore, the models are spelled out in continuous time while in reality asset prices are only observed at discrete times. To operationalize the models, we first derive the discrete time dynamic of the log-returns induced by (1), and then introduce volume-based proxies for the unobservable time changes.

Let $\Delta > 0$ denote a fixed time step, for example one day that represents the frequency of observations. The discrete time dynamics of the excess log-return $X_{(k+1)\Delta} - X_{k\Delta}$, $k = 0, 1, 2, \ldots$, can be derived from (1) by noting that

$$X_{(k+1)\Delta} - X_{k\Delta} \overset{\mathcal{D}}{=} \gamma s(Y_{[k\Delta,(k+1)\Delta]}) + \sigma W(s(Y_{[k\Delta,(k+1)\Delta]})) + \beta Y_{[k\Delta,(k+1)\Delta]} \tag{12}$$

where $\overset{\mathcal{D}}{=}$ denotes an equality in distribution and $Y_{[k\Delta,(k+1)\Delta]} = \int_{k\Delta}^{(k+1)\Delta} y_s ds$. Intuitively, it is as if at each discrete date $k\Delta$, the process Y accumulating the stochastic time change would restart at zero because only the accumulated stochastic time change $Y_{[k\Delta,(k+1)\Delta]}$ over the time interval $[k\Delta, (k+1)\Delta]$ impacts the log-return volatility over that period.[13] We note that (12) can be easily extended to settings in which the discrete dates are not equally spaced in time or Δ is a positive random variable describing the random arrival of market prices, as is the case for intraday data; see Aït-Sahalia and Mykland (2003) for a treatment of random sampling in diffusive settings.

In discrete time, the characteristic function of $X_{(k+1)\Delta} - X_{k\Delta}$ in (11) is thus

$$\Phi(u; X_{(k+1)\Delta} - X_{k\Delta}) = \exp(A_j(\Delta; iq(u)) + B_j(\Delta; iq(u))y_{k\Delta} + C_j(\Delta; iq(u))m_{k\Delta}) \tag{13}$$

where $j = 1, 2$ for SV1 and SV2 models, respectively. Because $y_{k\Delta}$ and $m_{k\Delta}$ are not observable, this characteristic function is not operational. To avoid lengthy expressions, we discuss how to operationalize (13) for SV1 models in which $C_1(\Delta; iq(u)) = 0$. The case of SV2 models can be handled in a similar way.

We propose to construct proxies of y using trading volume, given that the rate of time change y controls the stochastic volatility as shown in (6). Trading activity is a natural proxy for stochastic time changes as it represents the "business" clock rather than the "calendar" clock. In fact, an extensive research has documented the strong positive

12 For example, the function $B_2(t; iq(u))$ in the SV2 model solves an autonomous differential equation, and has an analytic expression. For a wide range of model parameters, the analytic expression and the series approximation of $B_2(t; iq(u))$ based on three or more terms, are virtually indistinguishable. In our empirical analysis, we use five terms to implement the series solution of the functions $A_2(t; iq(u))$, $B_2(t; iq(u))$ and $C_2(t; iq(u))$.

13 The equality in distribution in (12) is based on the so-called infinite divisibility property of Lévy processes. Because Lévy processes have independent increments over non-overlapping intervals, it holds that $s_{(k+1)\Delta} \overset{\mathcal{D}}{=} s_{k\Delta} + s_{\Delta}$. Therefore, the distribution of $X_{(k+1)\Delta} - X_{k\Delta}$ depends on s_{Δ}, and not on past $s_{k\Delta}$.

association between volume and volatility.[14] Because trading volume is not stationary over long time spans, we use volume changes to proxy for the activity rate y of log-returns. Specifically, we fit a GARCH(1,1)-type model to daily changes of log trading volume, compute the GARCH variance to obtain the proxy $\tilde{y}_{k\Delta}$ of $y_{k\Delta}$ at the discrete dates $k\Delta$, $k = 0, 1, 2, \ldots$, and scale the $\tilde{y}_{k\Delta}$'s to have sample mean one because $\mathbb{E}[y_{k\Delta}] = 1$. The proxy $\tilde{y}_{k\Delta}$ is therefore a weighted average of past volume changes. We then assume that $\tilde{y}_{k\Delta}$ is related to $y_{k\Delta}$ through the equation $y_{k\Delta} = \tilde{y}_{k\Delta}\varepsilon_{k\Delta}$, where $\varepsilon_{k\Delta}$ is a non-negative multiplicative observation error. There are two reasons for using GARCH models. The first reason is to reduce the noise in the raw volume data. The second is to disentangle slow- and fast-moving components in the volume process to obtain proxies of y_t and m_t for SV2 models.

While volume-based proxies of time changes can be easily constructed, it is unclear which properties these proxies should have or how to use them in a theoretically consistent way to make time-changed Lévy models operational. Appendix D formally addresses these issues. Below we summarize the discussion. The observed filtration $\mathcal{O}_{k\Delta}$, collecting observable variables, is discrete and given by the σ-algebra generated by the random variables $\{(X_{j\Delta}, \tilde{y}_{j\Delta}) : j = 0, 1, \ldots, k\}$. Our key assumption is that the observation error satisfies the following Markovian-type property

$$\mathbb{P}[\{\varepsilon_{k\Delta} \in A\}|\mathcal{O}_{k\Delta}] = \mathbb{P}[\{\varepsilon_{k\Delta} \in A\}|\tilde{y}_{k\Delta}] \tag{14}$$

where $A \in \mathcal{B}(\mathbb{R}_+)$ and $\mathcal{B}(\mathbb{R}_+)$ is the Borel sets of $\mathbb{R}_+$. This assumption ensures that the distribution of the observation error $\varepsilon_{k\Delta}$ does not depend on the whole history of the proxy variable, which appears to be a reasonable condition to impose on observation errors. We can state the following proposition.

Proposition 1.4. *Using the assumption in (14) and substituting $y_{k\Delta} = \tilde{y}_{k\Delta}\varepsilon_{k\Delta}$ in (13), the characteristic function of the discrete time log-return $X_{(k+1)\Delta} - X_{k\Delta}$ is given by*

$$\begin{aligned}\mathbb{E}_{\mathcal{O}_{k\Delta}}[\Phi(u; X_{(k+1)\Delta} - X_{k\Delta})] &= \mathbb{E}_{\mathcal{O}_{k\Delta}}[\exp(A(\Delta; iq(u)) + B(\Delta; iq(u))y_{k\Delta})] \\ &= \exp(A(\Delta; iq(u))) \int_0^\infty \exp\left(B(\Delta; iq(u))\tilde{y}_{k\Delta}\varepsilon\right) f(\varepsilon|\tilde{y}_{k\Delta}) d\varepsilon\end{aligned} \tag{15}$$

where $f(\varepsilon|\tilde{y}_{k\Delta})$ is the conditional density of the observation error $\varepsilon_{k\Delta}$.

Specifying an observation error with known characteristic function, the integral in (15) can be readily computed by plugging $-iB(\Delta; iq(u))\tilde{y}_{k\Delta}$ in the characteristic function of $\varepsilon_{k\Delta}$ conditional on $\tilde{y}_{k\Delta}$. In our empirical analysis, we assume that the observation error $\varepsilon_{k\Delta}$ follows a Gamma distribution, $G(\delta_y, 1/(\delta_y - 1))$.[15] Two reasons motivate this choice. First, the error term $\varepsilon_{k\Delta}$ has to be positive to ensure that the time change is positive, and the Gamma distribution is a rather flexible positive distribution. Second, when $\varepsilon_{k\Delta}$ follows a Gamma distribution, $1/\varepsilon_{k\Delta}$ follows an Inverse Gamma distribution. This result is useful to impose $\mathbb{E}[1/\varepsilon_{k\Delta}] = 1$ via a parameter restriction in the Gamma distribution, which in turn

14 This line of research goes back at least to Gallant et al. (1992). Bollerslev et al. (2018) provide a recent empirical analysis of the volume-volatility relation using intraday data.

15 We also experimented with an Inverse Gaussian distribution and the empirical results were unchanged.

ensures that $\mathbb{E}[\tilde{y}_{k\Delta}] = y_{k\Delta}$, that is, the volume-based proxy $\tilde{y}_{k\Delta}$ is an unbiased reflection of the rate of time change $y_{k\Delta}$. The integral in (15) is then given by

$$\int_0^\infty \exp\left(B(\Delta; iq(u))\tilde{y}_{k\Delta}\varepsilon\right) f(\varepsilon|\tilde{y}_{k\Delta})d\varepsilon = \left(1 - B(\Delta, iq(u))\tilde{y}_{k\Delta}/(\delta_y - 1)\right)^{-\delta_y}.$$

Using (15) a large class of time-changed Lévy models can be made operational. The characteristic function of the discrete time log-return $X_{(k+1)\Delta} - X_{k\Delta}$ now depends only on observable quantities and model parameters, and time-changed Lévy models can be readily implemented. When $y_{k\Delta}$ follows the SV2 process in (8), we fit the two-component GARCH model of Christoffersen et al. (2008) to daily changes of log trading volume, and compute the short-run and long-run GARCH variance components to obtain the proxies $\tilde{y}_{k\Delta}$ and $\tilde{m}_{k\Delta}$ of $y_{k\Delta}$ and $m_{k\Delta}$, respectively, assuming that the observation errors satisfy the Markovian-type property in (14).

In the next section, we begin the empirical analysis by fitting time-changed Lévy models to index returns. The objective is to assess the time series performance of the models, independently from their option pricing performance and the specification of risk premia. Proper time series fitting is important for example when computing risk measures of equity portfolios. Subsequently, we analyze option pricing and risk premia.[16]

2 Time Series Analysis

This section presents the time series data, the estimation method, and the empirical performance of time-changed Lévy models.

2.1 Time Series Data

For the time series analysis, we use the daily log-returns and trading volumes of the Standard & Poor's 500 index (S&P 500). Our sample consists of 17,538 daily data from February 28, 1950 to November 6, 2019. The trading volume is the total daily trading volume of stocks that constitute the S&P 500 index. The in-sample tests use 11,592 observations, until March 25, 1996, that is about 2/3 of the full sample. The remaining observations are used for out-of-sample tests. Several episodes of financial market turmoil, such as the global financial crisis (2007–2009) or the European debt crisis (2010–2012), belong to the out-of-sample period. It will be interesting to assess which time-changed Lévy models capture the large negative returns associated with these events, given that the model parameters are estimated without those returns.

2.2 Model Estimation

We estimate the time-changed Lévy models by maximum likelihood, using the volume-based proxy of the time change as in (15). Because the probability density of the log-return

16 An alternative approach would be to fit the models jointly to index returns and option prices. To keep the amount of empirical work manageable, we leave this exercise to future work. Notice, however, that this exercise would not allow to clearly assess the time series performance of the models.

in (1) is not available in closed-form, we first calculate the characteristic function and then recover the probability density via numerical Fourier inversion of the characteristic function.

There are various approaches to recover probability densities from characteristic functions, with the Fast Fourier transform (FFT) or fractional FFT being popular methods. However, probability densities of Lévy increments typically have large kurtosis, which may lead to numerical errors when applying classic Fourier inversion techniques. This issue is particularly severe for probability densities over short time horizons, as is the case in our time series analysis. The problem stems from the slow decay of the real part of the characteristic function, which is known as the Gibbs phenomenon. To achieve high accuracy in the computation of the probability density, and thereby overcome the Gibbs phenomenon, we enrich the Fourier-cosine series (COS) method of Fang and Oosterlee (2008) with a damping function. Appendix E explains our method in detail.

2.3 Empirical Findings

To label the time-changed Lévy models that we investigate empirically, we use three blocks of letters, B_1, B_2, B_3. The first block B_1 is F or empty depending on whether the model implementation takes into account or not the observation error in the volume-based proxy of the time change as in (15). When the implementation does not rely on (15), unobservable variables are set equal to their proxies, assuming that observation errors are absent. The second block B_2 is VG, NIG or NTS, depending on whether s_t is the Gamma, Inverse Gaussian or Tempered Stable subordinator. The third block B_3 is SV1 or SV2 depending on whether the stochastic volatility is the one-factor (7) or two-factor (8) process. For example, VGSV1 denotes the Variance Gamma model with the one-factor process driving the stochastic volatility, and FVGSV1 is the same model but implemented taking into account the observation error in the volume-based proxy as in (15).

We estimate sixteen time-changed Lévy models. We construct twelve models by combining the three Lévy subordinators s_t in Table 1; one-factor (SV1) or two-factor (SV2) process for the stochastic volatility; with or without taking into account the observation error in the volume-based proxy as in (15). All these twelve models are of infinite activity. As an illustrative example, Appendix B presents the VGSV1 model in detail. We also consider four finite activity models in which s_t follows a Poisson process time-changed with the SV1 or SV2 process, with or without using (15).[17] For the finite activity models, the block B_2 of letters in the model labels is given by P. We impose the Feller condition while estimating the models.

We consider various benchmarks. To assess the overall time series fitting of time-changed Lévy models, we consider the GJR GARCH model of Glosten, Jagannathan, and Runkle (1993) and the two-component GARCH model of Christoffersen et al. (2008), which are known to fit well daily index returns. To gage the usefulness of our volume-based proxy, we estimate eight models imposing that the time change is equal to the volume-based proxy, that is, the observation error is absent.

Tables 2 and 3 show the in-sample estimation results for time-changed Lévy models, when the stochastic volatility is driven by the one-factor (SV1) or two-factor (SV2) process,

17 In these models, s_t follows a so-called Cox process as its intensity is not constant and given by Y_t. Upon a jump occurring, the size of the return jump is normally distributed.

Table 2. Parameter estimates for time-changed Lévy models with one-factor stochastic volatility

Model	lnL	Jump		Drift		Bm	Stochastic Volatility		
		ν	α	γ	β	σ	κ_y	σ_y	δ_y
PSV1	33,492			–0.260	0.033	0.274	1.492	0.849	
	[14,025]			(0.007)	(0.000)	(0.000)	(0.027)	(0.011)	
FPSV1	33,715			–0.601	0.264	0.422	1.980	1.047	46.39
	[14,241]			(0.007)	(0.009)	(0.015)	(0.010)	(0.007)	(3.029)
VGSV1	39,726	110.23		–0.103	0.094	0.173	1.491	0.129	
	[18,102]	(1.991)		(0.000)	(0.000)	(0.000)	(0.209)	(0.000)	
FVGSV1	39,739	128.01		–0.112	0.101	0.186	2.509	0.48	12.91
	[18,133]	(3.083)		(0.000)	(0.007)	(0.000)	(0.000)	(0.006)	(2.073)
NIGSV1	39,729	52.84		–0.134	0.129	0.188	8.024	0.195	
	[18,099]	(3.095)		(0.000)	(0.006)	(0.000)	(0.068)	(0.000)	
FNIGSV1	39,741	40.87		–0.128	0.113	0.208	3.470	0.072	25.18
	[18,116]	(1.038)		(0.000)	(0.008)	(0.000)	(0.000)	(0.017)	(2.305)
NTSSV1	39,745	55.48	0.860	–0.131	0.114	0.198	1.206	0.166	
	[18,121]	(2.228)	(0.009)	(0.000)	(0.000)	(0.007)	(0.018)	(0.000)	
FNTSSV1	39,757	50.40	0.900	–0.109	0.099	0.182	1.850	0.387	7.19
	[18,135]	(4.007)	(0.005)	(0.000)	(0.000)	(0.000)	(0.014)	(0.000)	(1.085)

Time-changed Lévy models (1) are obtained from the Poisson (P) subordinator and the three Lévy subordinators in Table 1 (VG, NIG, NTS), combined with the one-factor (SV1) process (7) for the stochastic volatility, with (F) or without the volume based proxy of the time change as in (15). The observation error of the proxy of y follows a $G(\delta_y, 1/(\delta_y - 1))$. Model labels are described at the beginning of Section 2.3. Standard errors are in parenthesis. lnL is in-sample [out-of-sample] Log-likelihood. Bm is Brownian motion. Data are daily S&P 500 log-returns over 1950–1996, (in-sample period). Out-of-sample period is 1996–2019.

respectively. Various findings emerge from these estimates. First, going from finite to infinite activity models sharply improves the fitting of index returns. Vuong (1989) Likelihood Ratio (LR) tests soundly reject the null hypothesis that either FPSV1 or FPSV2 are equivalent to any infinite activity model. Second, using the volume-based proxy as in (15) uniformly improves time series fitting across all time-changed Lévy models. LR tests strongly reject the null hypothesis that models that allow for the observation error are equivalent to models that do not allow for the observation error, for one- and two-factor and for any Lévy measure. Third, the best performing model is FNTSSV2. This model features a flexible tempered stable Lévy subordinator and has by far the largest log-likelihood value across all the sixteen models that we estimate. Fourth, estimates of the tempering parameter ν in the Lévy measure (4) are largely different from zero, which provides statistical evidence against stable infinite variation subordinators, characterized by $\nu = 0$. Finally, estimates of the tail parameter α in (4) point to tempered stable subordinators for index returns. Estimates of α range between 0.83 and 0.90, and are statistically away from both 1/2 and 1 (see Table 1).

A closer inspection of the best performing model, FNTSSV2, allows us to gage the sources of its fitting accuracy. Taking the classic VGSV1 as a baseline time-changed Lévy model, we can decompose the log-likelihood increment from VGSV1 to FNTSSV2 as follows:

Table 3. Parameter estimates for time-changed Lévy models with two-factor stochastic volatility

Model	lnL	Jump		Drift		Bm	Stochastic volatility					
		ν	α	γ	β	σ	κ_y	σ_y	δ_y	κ_m	σ_m	δ_m
PSV2	38,112			–0.207	0.166	0.186	1.090	0.163		0.593	0.022	
	[17,008]			(0.000)	(0.002)	(0.000)	(0.007)	(0.009)		(0.000)	(0.000)	
FPSV2	38,126			–0.239	0.204	0.181	1.905	0.230	5.099	1.030	0.219	4.914
	[17,050]			(0.009)	(0.003)	(0.000)	(0.000)	(0.017)	(0.084)	(0.013)	(0.000)	(1.006)
VGSV2	39,746	109.44		–0.135	0.117	0.194	0.875	0.131		0.711	0.086	
	[18,148]	(6.309)		(0.000)	(0.005)	(0.009)	(0.001)	(0.000)		(0.000)	(0.000)	
FVGSV2	39,755	106.03		–0.117	0.103	0.188	0.883	0.139	6.395	0.591	0.082	3.110
	[18,153]	(3.073)		(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(1.083)	(0.000)	(0.000)	(0.558)
NIGSV2	39,749	37.03		–0.191	0.182	0.209	0.890	0.142		0.691	0.123	
	[18,149]	(4.728)		(0.006)	(0.000)	(0.009)	(0.022)	(0.001)		(0.007)	(0.000)	
FNIGSV2	39,761	43.28		–0.183	0.172	0.197	0.870	0.153	7.030	0.705	0.111	3.909
	[18,156]	(3.092)		(0.0024)	(0.000)	(0.005)	(0.000)	(0.000)	(1.044)	(0.000)	(0.000)	(0.274)
NTSSV2	39,760	52.00	0.871	–0.147	0.140	0.187	0.890	0.161		0.675	0.120	
	[18,150]	(2.840)	(0.008)	(0.000)	(0.000)	(0.000)	(0.009)	(0.000)		(0.005)	(0.000)	
FNTSSV2	39,770	46.07	0.830	–0.145	0.136	0.183	0.893	0.142	9.015	0.689	0.108	5.220
	[18,163]	(1.825)	(0.011)	(0.009)	(0.000)	(0.000)	(0.008)	(0.000)	(2.307)	(0.000)	(0.000)	(1.084)

See the notes to Table 2 for parameter definitions and sample data. Time-changed Lévy models (1) are based on the two-factor (SV2) process (8) for the stochastic volatility. The observation errors of the proxies of y and m follow a $G(\delta_y, 1/(\delta_y - 1))$ and $G(\delta_m, 1/(\delta_m - 1))$, respectively.

$$\text{FNTSSV2} - \text{VGSV1} = (\text{FNTSSV2} - \text{NTSSV2}) + (\text{NTSSV2} - \text{NTSSV1}) + (\text{NTSSV1} - \text{VGSV1})$$

where the equality is in log-likelihood units. The log-likelihood increment from VGSV1 to FNTSSV2 can be broken down as follows: 23% of the increment is due to the usage of the volume-based proxy that allows for observation error as in (15) (comparing FNTSSV2 and NTSSV2); 34% to the second volatility factor (comparing NTSSV2 and NTSSV1); 43% to the flexible tempered stable jump specification (comparing NTSSV1 and VGSV1). The above likelihood decomposition confirms that moving from one- to two-factor volatility process yields a sharp increase in the fitting of index returns, as is well-known from prior literature. Interestingly, the decomposition also indicates that allowing for a flexible Lévy subordinator is even more important. Using a suitable volume-based proxy that allows for observation error of the time change accounts for a sizable 1/4 of the log-likelihood increase from VGSV1 to FNTSSV2.

The benchmark GJR and two-component GARCH models substantially underperform any infinite activity model in terms of log-likelihood values.[18] This finding is relevant because GARCH models tend to fit well daily index returns. As an additional exercise, we treat the time change as a latent process, and jointly determine model parameters and time series trajectory of the time change using the Unscented Kalman Filter (UKF) and index returns only; see Appendix F. Although a thorough comparison of time-changed Lévy models estimated using the UKF or volume-based proxy is beyond the scope of this article, estimation results indicate that log-likelihood based on the UKF values are lower, which underscores the importance of using volume-based proxy.[19]

Tables 2 and 3 also report the out-of-sample log-likelihood values for the sixteen estimated models. The rankings of the models based on their in-sample and out-of-sample likelihood values are nearly identical. This finding indicates that the FNTSSV2 model is capturing intrinsic features of index returns, and not just overfitting in-sample data. Table 3 also confirms that the FNTSSV2 model is the most accurate model out-of-sample according to likelihood values.

To visualize the model fitting, Figure 1 shows the out-of-sample conditional normal probability plots of four selected models: the VGSV1 model, the best performing FNTSSV2 model, and the GJR and two-component GARCH models. The plots are based on the variable $z_{k\Delta} = F(X_{(k+1)\Delta} - X_{k\Delta}|\mathcal{O}_{k\Delta})$, $k = 1, 2, \ldots$, where $F(\cdot|\mathcal{O}_{k\Delta})$ is the conditional cumulative distribution function of the daily log-return $X_{(k+1)\Delta} - X_{k\Delta}$, estimated using a specific model. Comparing Panels (A) and (B) shows the benefit of going from a one-factor infinity activity process to a more general two-factor infinity activity process, which is estimated taking into account the observation error in the volume-based proxy as in (15). Major improvements are visible in the fitting of the left tail of the index return distribution. Comparing Panels (A) and (B) with Panels (C) and (D) confirms the substantially superior

18 The log-likelihood values of the GJR and two-component GARCH models are 16,326 and 16,407, respectively, and are about half the log-likelihood values of any infinite activity model in Tables 2 and 3.

19 For example, the log-likelihood values of VGSV1 and NTSSV2 are 39,158 and 39,627, respectively, and are well below the log-likelihood values of the corresponding models estimated using the volume-based proxy; see Tables 2 and 3.

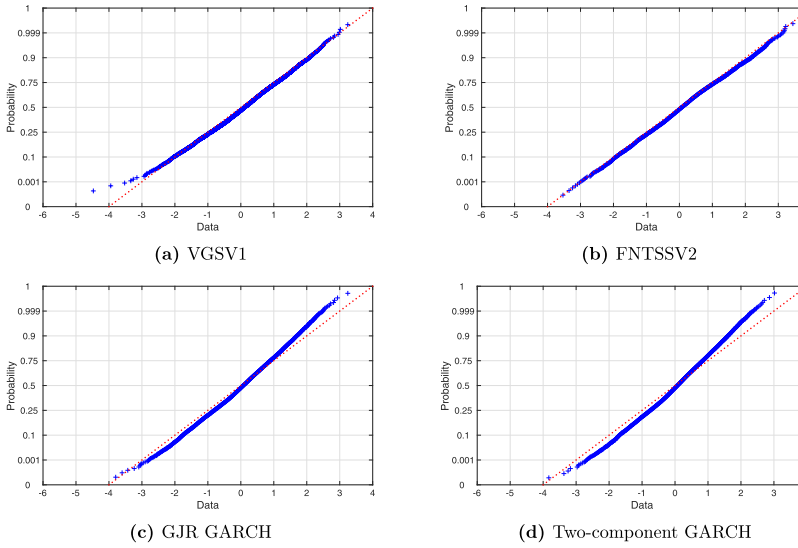

Figure 1. Out-of-sample normal probability plots for normalized returns for VGSV1, FNTSSV2, and benchmark GARCH models. Out-of-sample daily S&P 500 log-returns span 1996–2019. Labels of time-changed Lévy models are described at the beginning of Section 2.3. Diagonal dotted lines are the theoretical quantiles conditional upon correct specification. (**a**) VGSV1, (**b**) FNTSSV2, (**c**) GJR GARCH, (**d**) Two-component GARCH.

performance of infinite activity time-changed Lévy models relative to GARCH models. The lack of fit of the latter models is evident throughout the entire index return distribution.

Finally, we carry out the conditional density forecast evaluation in Diebold, Gunther, and Tay (1998) and Diebold, Hahn, and Tay (1999) of the time-changed Lévy models and GARCH models. To save space, we do not report detailed results, and only summarize the main findings. The out-of-sample autocorrelations of the first four moments of $z_{k\Delta}$ are approximately zero for most models, indicating that the models can accommodate the temporal dependence in index returns. However, only the most flexible models, in particular the FNTSSV2, capture well the conditional density of index returns, confirming the finding in Figure 1.

In sum, infinite activity time-changed Lévy models substantially outperform finite activity or GARCH models in the time series fitting of index returns. Besides making models operational, using the volume-based proxy of the time change as in (15) improves the time series fitting of any model. Among the sixteen time-changed Lévy models that we estimate, the best performing model features the tempered stable Lévy subordinator, the two-factor volatility process, and is estimated using the volume-based proxy allowing for observation error.

3 Option Pricing: Theoretical Analysis

This section presents the risk neutralization, the risk neutral characteristic function, and the option pricing formula for the time-changed Lévy models in (1).

3.1 Risk Neutralization

The discrete time market at dates $k\Delta, k = 0, 1, \ldots$, which consists in the risky asset and a riskless asset, is intrinsically incomplete. Therefore, the set of risk neutral measures consists in infinitely many elements. As usual with incomplete models, we pin down the risk neutral measure chosen by market participants by minimizing the distance between market and model-based option prices. We consider equivalent risk neutral probability measures $\mathbb{Q}$ with the following properties:

1. There exists a risk premium θ such that
$$W_t^{\mathbb{Q}} = W_t + \theta t \tag{16}$$
is a standard Brownian motion under $\mathbb{Q}$.
2. The $\mathbb{Q}$-subordinator s_t is a Lévy process in the same family of the $\mathbb{P}$-subordinator with characteristic exponent given by
$$\psi_{\mathbb{Q}}(u) = \frac{1}{t}\log \mathbb{E}^{\mathbb{Q}}[e^{ius_t}] = \int_0^\infty (e^{iux} - 1)\Pi_{\mathbb{Q}}(dx)$$
and Lévy measure $\Pi_{\mathbb{Q}}(dx) = c^{\mathbb{Q}} \exp(-\nu^{\mathbb{Q}} x)/x^{1+\alpha^{\mathbb{Q}}} dx$ for $x \in (0, \infty)$, $c^{\mathbb{Q}}, \nu^{\mathbb{Q}} > 0$, and $\alpha^{\mathbb{Q}} \in [0, 1)$.
3. The processes y_t and m_t remain affine under $\mathbb{Q}$, that is, either y_t follows the SV1 process or (y_t, m_t) follow the SV2 process.

No additional conditions are imposed on the risk neutral measures. For example, the tail exponents α and $\alpha^{\mathbb{Q}}$ are allowed to be different. Using the first two restrictions above, and arguments similar to those of Section 1, shows that the discrete time martingale condition

$$\mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_{k\Delta}}\left[e^{-(r-\delta)\Delta}\frac{S_{(k+1)\Delta}}{S_{k\Delta}}\right] = 1 \tag{17}$$

can be equivalently stated as

$$\begin{aligned} 1 &= \mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_{k\Delta}}[\exp(X_{(k+1)\Delta} - X_{k\Delta})] \\ &= \mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_{k\Delta}}[\exp((\beta - \phi_{\mathbb{Q}}(\sigma\theta - \sigma^2/2 - \gamma))\left(Y_{(k+1)\Delta} - Y_{k\Delta}\right))] \end{aligned} \tag{18}$$

where $\phi_{\mathbb{Q}}(\lambda) = -\psi_{\mathbb{Q}}(i\lambda)$ is the Laplace exponent of the non-decreasing Lévy process s_t under $\mathbb{Q}$ and $\lambda = \sigma\theta - \sigma^2/2 - \gamma$. The martingale condition in (17) is a discrete time condition in that it only considers the observed price process at discrete dates $k\Delta, k = 0, 1, \ldots$ Importantly, since $\left(Y_{(k+1)\Delta} - Y_{k\Delta}\right)$ is non-negative, from (18) we can state the following proposition.

Proposition 3.1. *A necessary and sufficient condition for $\mathbb{Q}$ to be a risk neutral measure is that*

$$\beta = \phi_{\mathbb{Q}}(\sigma\theta - \sigma^2/2 - \gamma). \tag{19}$$

The condition above places a joint constraint on the Brownian risk premium θ and the Lévy measure $\Pi_{\mathbb{Q}}(dx)$ determining the Laplace exponent $\phi_{\mathbb{Q}}$. We emphasize that if a time-changed Lévy model fails to satisfy the condition above, then the discrete time model admits arbitrage opportunities. Furthermore, condition (19) explains why the term βY_t

enters the model specification in (1). Setting $\beta = 0$ would impose a very tight constraint on the risk neutralization of X_t. In fact, the Laplace transform of $s(t)$ under $\mathbb{Q}$ is $\mathbb{E}^{\mathbb{Q}}[e^{-\lambda s(t)}] = e^{-\phi_{\mathbb{Q}}(\lambda)t} = e^{-\beta t}$, where the last equality is due to (19) when $\lambda = \sigma\theta - \sigma^2/2 - \gamma$. If $\beta = 0$ like in traditional time-changed Lévy models, then it must be the case that $\lambda s(t) = 0$ and therefore $\lambda = 0$, given that $s_t \geq 0$ with probability one under $\mathbb{Q}$. That is, if $\beta = 0$ then $\sigma\theta - \sigma^2/2 - \gamma = 0$. Because σ and γ are estimated from the time series of index returns, the last equality would determine the Brownian risk premium θ, which would impair the ability of time-changed Lévy models to fit option prices.

Beyond the requirement of equivalence in (19), the risk neutralization of time-changed Lévy processes in (1) does not require any additional constraint on the risk premia carried by the activity rate y_t or the observation error $\varepsilon_{k\Delta}$. For simplicity, we assume that observation errors carry no risk premia, and hence follow the same distribution under $\mathbb{Q}$ and satisfy the same Markovian-type property in (14) as under $\mathbb{P}$.

To specify the VRP, we use the most general form of risk premium that preserves the affine property of y_t in SV1, and (y_t, m_t) in SV2, when changing measure from $\mathbb{P}$ to $\mathbb{Q}$. This form of risk premium is called the extended affine market price of risk specification; see Cheridito, Filipovic, and Kimmel (2007). For the SV1 process in (7), the VRP is such that

$$dW_t^y = dW_t^{\mathbb{Q},y} + a_y\sqrt{y_t}dt + \frac{b_y}{\sqrt{y_t}}dt \tag{20}$$

where a_y and b_y are constants, and $W_t^{\mathbb{Q},y}$ is a $\mathbb{Q}$-Brownian motion. The SV1 process under $\mathbb{Q}$ is therefore

$$dy_t = \kappa_{\mathbb{Q},y}(\theta_{\mathbb{Q},y} - y_t)dt + \sigma_y\sqrt{y_t}dW_t^{\mathbb{Q},y}$$

where $\kappa_{\mathbb{Q},y} = \kappa_y - \sigma_y a_y$ and $\theta_{\mathbb{Q},y} = (\kappa_y + \sigma_y b_y)/(\kappa_y - \sigma_y a_y)$. The risk premium can alter the speed of mean reversion $\kappa_{\mathbb{Q},y}$ and the long run mean $\theta_{\mathbb{Q},y}$ of y_t under $\mathbb{Q}$.[20]

For the SV2 model in (8), the extended affine risk premium specification is such that

$$\begin{aligned} dW_t^y &= dW_t^{\mathbb{Q},y} + a_y\sqrt{y_t}dt + \frac{b_y}{\sqrt{y_t}}dt \\ dW_t^m &= dW_t^{\mathbb{Q},m} + a_m\sqrt{m_t}dt + \frac{b_m}{\sqrt{m_t}}dt \end{aligned} \tag{21}$$

where $W^{\mathbb{Q},m}$ is a $\mathbb{Q}$-Brownian motion. The SV2 model under $\mathbb{Q}$ is thus given by

$$\begin{aligned} dy_t &= \kappa_{\mathbb{Q},y}(\theta_{\mathbb{Q},y,c} + \theta_{\mathbb{Q},y,d}m_t - y_t)dt + \sigma_y\sqrt{y_t}dW_t^{\mathbb{Q},y} \\ dm_t &= \kappa_{\mathbb{Q},m}(\theta_{\mathbb{Q},m} - m_t)dt + \sigma_m\sqrt{m_t}dW_t^{\mathbb{Q},m} \end{aligned} \tag{22}$$

where $\kappa_{\mathbb{Q},m} = \kappa_m - \sigma_m a_m$,

20 Because we impose the Feller condition when fitting SV1 models to index returns under $\mathbb{P}$, we impose this condition under $\mathbb{Q}$ as well when pricing options.

$$\theta_{\mathbb{Q},m} = \frac{\kappa_m + \sigma_m b_m}{\kappa_m - \sigma_m a_m}, \quad \theta_{\mathbb{Q},y,c} = \frac{\sigma_y b_y}{\kappa_y - \sigma_y a_y}, \quad \theta_{\mathbb{Q},y,d} = \frac{\kappa_y}{\kappa_y - \sigma_y a_y}.$$

The extended affine specification provides a high degree of flexibility for modeling the VRP. This specification is rarely estimated in practice, perhaps because accurate estimation is challenging in small datasets. It is therefore interesting to assess the shape of the VRP in a large panel of options, which is one of the empirical contributions of our option pricing analysis in Section 4.

3.2 The $\mathbb{Q}$-Characteristic Function and Option Pricing Formula

To compute option prices, we use the risk neutral characteristic function

$$\mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_{k\Delta}}[\Phi_{\mathbb{Q}}(u; X_{n\Delta} - X_{k\Delta})] = \mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_{k\Delta}}[\exp(iu(X_{n\Delta} - X_{k\Delta}))]$$

where $(n-k)\Delta$ is the time to maturity of the option. The same arguments as in the computation of the $\mathbb{P}$-characteristic function show that

$$\mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_{k\Delta}}[\Phi_{\mathbb{Q}}(u; X_{n\Delta} - X_{k\Delta})] = \mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_{k\Delta}}[\exp\Big(iq_{\mathbb{Q}}(u)\int_{k\Delta}^{n\Delta} y_s ds\Big)] \tag{23}$$

where

$$q_{\mathbb{Q}}(u) = u\beta - i\psi_{\mathbb{Q}}\Big(u(\gamma - \sigma\theta) + iu^2\frac{\sigma^2}{2}\Big).$$

To calculate the above $\mathbb{Q}$-characteristic function, we use the volume-based proxy of the time change as in (15), obtaining $\tilde{y}_{k\Delta}$ for SV1 models and $(\tilde{y}_{k\Delta}, \tilde{y}_{k\Delta})$ for SV2 models, and assuming that the observation errors carry no risk premium. As usual, when pricing options via Fourier inversion, we rewrite the time-$k\Delta$ forward price of a call option with strike price K and expiring at time $n\Delta$, as

$$\mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_{k\Delta}}[\max(S_{n\Delta} - K, 0)] = S_{k\Delta}\mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_{k\Delta}}[\max(\exp((r-\delta)(n-k)\Delta + X_{n\Delta} - X_{k\Delta}) - K/S_{k\Delta}, 0)]. \tag{24}$$

Using the COS method with the damping function in Appendix E, we invert the $\mathbb{Q}$-characteristic function of $X_{n\Delta} - X_{k\Delta}$ in (23) to compute the option price in (24).

We close this section by discussing the leverage effect under $\mathbb{Q}$. Calculations similar to those we performed under $\mathbb{P}$ in Section 1 show that the $\mathbb{Q}$-instantaneous variance is given by

$$v_t^{\mathbb{Q}} = \lim_{\Delta\to 0}\frac{\mathbb{V}^{\mathbb{Q}}_t[X_{t+\Delta} - X_t]}{\Delta} = \Big(\sigma^2\mathbb{E}^{\mathbb{Q}}[s_1] + (\gamma - \sigma\theta)^2\mathbb{V}^{\mathbb{Q}}[s_1]\Big)y_t.$$

Therefore, the leverage effect under $\mathbb{Q}$ depends on the sign of

$$\mathrm{Cov}^{\mathbb{Q}}_0[X_t - X_0, v_t^{\mathbb{Q}} - v_0^{\mathbb{Q}}] = \Big(\beta + (\gamma - \sigma\theta)\mathbb{E}^{\mathbb{Q}}[s_1]\Big)\Big(\sigma^2\mathbb{E}^{\mathbb{Q}}[s_1] + (\gamma - \sigma\theta)^2\mathbb{V}^{\mathbb{Q}}[s_1]\Big)\mathrm{Cov}^{\mathbb{Q}}_0[Y_t, y_t].$$

The same calculations as under $\mathbb{P}$ in Appendix A show that in the SV1 model, and in the SV2 model when $\theta_{\mathbb{Q},y,d} > 0$ in (22), the sign of the constant term $\beta + (\gamma - \sigma\theta)\mathbb{E}^{\mathbb{Q}}[s_1]$

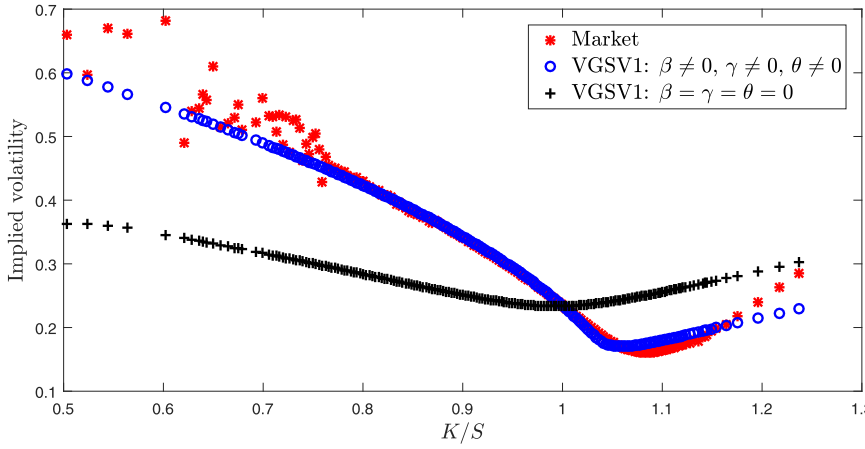

Figure 2. Implied volatilities of 30-day SPX options on August 31, 2015 (Market). Model-based implied volatilities of Variance Gamma (Madan and Seneta, 1990) time-changed with a Heston process, with leverage effect (VGSV1: $\beta \neq 0$, $\gamma \neq 0$, $\theta \neq 0$) and without leverage effect (VGSV1: $\beta = \gamma = \theta = 0$).

determines the sign of the $\mathbb{Q}$-leverage. The primary reason for accommodating the leverage effect in an option pricing model is to capture skewed implied volatilities. To illustrate this point, Figure 2 shows thirty-day market implied volatilities of European options on the S&P 500 index observed on August 31, 2015, and model-based implied volatilities generated by two VGSV1 models. The first VGSV1 model features the leverage effect as discussed above, and captures the skewed implied volatilities quite well. The second VGSV1 model has no leverage effect because we set $\beta = \gamma = \theta = 0$. As a result, this model can only produce a symmetric implied volatility smile, and cannot fit observed volatilities.

4 Option Pricing: Empirical Results

This section describes the option data, the empirical pricing performance, and the estimated risk premia of time-changed Lévy models.

4.1 Option Data

We consider European options on the S&P 500 index (symbol: SPX) which are among the most actively traded options, and have been investigated in a number of empirical studies including Christoffersen, Jacobs, and Mimouni (2010), Ornthanalai (2014), Andersen, Fusari, and Todorov (2017), and Bardgett et al. (2019). Since out-of-the-money (OTM) options are more actively traded than in-the-money options, we consider mid closing prices of OTM puts and calls on each Wednesday from January 1996 to December 2019. We apply the usual screening criteria on raw option contracts; see, for example, Barone-Adesi et al. (2008).[21] All option data are downloaded from OptionMetrics. The final dataset

21 Specifically, options with time to maturity less than 10 days or more than 360 days, prices less than \$0.05, or zero trading volume are discarded. We also compute the implied volatility of each

Table 4. S&P 500 index option data

Option characteristics	$10 \leq T < 80$	$80 \leq T < 180$	$180 \leq T$
Panel A: Number of option contracts			
$K/S < 0.95$	498,464	198,821	126,153
$0.95 \leq K/S < 1$	133,580	34,008	14,107
$1 \leq K/S < 1.10$	208,095	59,697	27,249
$1.10 \leq K/S$	48,605	38,832	40,741
Panel B: Average option prices			
$K/S < 0.95$	3.64	11.25	20.19
$0.95 \leq K/S < 1$	21.46	56.11	88.38
$1 \leq K/S < 1.10$	10.10	30.96	61.38
$1.10 \leq K/S$	0.73	2.51	9.35
Panel C: Standard deviation option prices			
$K/S < 0.95$	5.51	13.67	23.20
$0.95 \leq K/S < 1$	14.92	21.70	30.85
$1 \leq K/S < 1.10$	13.36	25.32	3.27
$1.10 \leq K/S$	1.96	4.69	12.83
Panel D: Average implied volatility			
$K/S < 0.95$	0.31	0.31	0.30
$0.95 \leq K/S < 1$	0.15	0.16	0.18
$1 \leq K/S < 1.10$	0.12	0.13	0.15
$1.10 \leq K/S$	0.21	0.16	0.16
Panel E: Standard deviation implied volatility			
$K/S < 0.95$	0.14	0.13	0.11
$0.95 \leq K/S < 1$	0.05	0.04	0.04
$1 \leq K/S < 1.10$	0.04	0.04	0.04
$1.10 \leq K/S$	0.13	0.07	0.05

Out-of-the-money put and call options on the S&P 500 index from January 1996 to December 2019. Moneyness is defined as the ratio of the strike price over the index price, that is, K/S. Time to maturity T is in days.

contains 1,428,352 option contracts, with the number of put and call options being 1,005,133 (70.37%) and 423,219 (29.63%), respectively. The large number of put options reflects the increased demand of those options during and after the 2007–2009 financial crisis.

We define moneyness as the ratio of the strike price over the index price, that is, K/S. A put option is said to be deep OTM if $K/S < 0.95$, and OTM if $0.95 \leq K/S < 1$. A call option is said to be OTM if $1 \leq K/S < 1.10$, and deep OTM if $1.10 \leq K/S$. An option contract with time to maturity T days has short maturity if $10 \leq T < 80$, intermediate maturity if $80 \leq T < 180$, and long maturity if $180 \leq T$.

Table 4 reports descriptive statistics for the 1,428,352 option contracts. The number of deep OTM put (call) options is 823,438 (128,178), which corresponds to 57.65% (8.97%) of the option contracts (Panel A). Short and long maturity options account for 62.22% and

option, and eliminate those options whose implied volatility differs by >5% from the implied volatility reported in OptionMetrics.

14.58%, respectively, of the total sample. The average put (call) price ranges from \$3.64 (\$0.73) for short maturity, deep OTM options to \$88.34 (\$61.38) for long maturity, OTM options (Panel B). Call option prices tend be more volatile than put options in relative terms (Panel C). For each maturity bucket, the implied volatility smile across moneyness is evident (Panel D). When the time to maturity increases, the implied volatility smile tend to become flatter, but the standard deviation of implied volatilities remain high especially for deep OTM puts (Panel E).

4.2 Option Pricing and Risk Premia

We split the option sample into an in-sample part from January 1996 to mid-January 2013 and an out-of-sample part from mid-January 2013 to December 2019. For each of the sixteen time-changed Lévy models in Section 2.3, we use in-sample option data to calibrate risk premia parameters, enforcing the no-arbitrage condition (19). Specifically, we calibrate the VRP parameters a_y, b_y in (20) and a_y, b_y, a_m, b_m in (21) for SV1 and SV2 models, respectively, the Lévy measure parameter $\nu^{\mathbb{Q}}$ and for NTS models $\alpha^{\mathbb{Q}}$ as well. To restrict the number of parameters needed to calibrate for the Lévy measure, we impose $\mathbb{E}^{\mathbb{Q}}[s_t] = t$. Throughout the calibration procedure, we set the value of θ such that the no-arbitrage condition (19) holds, given the model parameters estimated from index returns.[22] We achieve parameter calibration as usual, namely by minimizing the sum of squared pricing errors across all Wednesdays in our in-sample period, and treat this minimization as a nonlinear least squares problem to obtain standard errors for the calibrated parameters. The procedure above pins down the risk neutral measure implicit in market option prices.

4.2.1 Jump and volatility risk premia.

Table 5 shows the calibrated parameters for the infinite activity models.[23] The Lévy measure under $\mathbb{Q}$ is substantially different from the Lévy measure under $\mathbb{P}$ across all models. Values of $\nu^{\mathbb{Q}}$ and $\alpha^{\mathbb{Q}}$ are systematically lower than their counterparts under $\mathbb{P}$. Low values of $\nu^{\mathbb{Q}}$ mean that the tempering function $\exp(-\nu^{\mathbb{Q}}x)$ in the Lévy measure decays slowly to zero when the jump size increases, and implies that large jumps are likely to occur under $\mathbb{Q}$. Low values of $\alpha^{\mathbb{Q}}$ indicate high jump activity especially near zero. For example, in the best performing FNTSSV2 model, $\mathbb{V}^{\mathbb{Q}}[s_1] = 0.069$ whereas $\mathbb{V}^{\mathbb{P}}[s_1] = 0.004$.[24] To gage the economic magnitude of these values, consider the following. In (1), set $Y_t = t$ and $W_t = 0$, which are their respective $\mathbb{P}$-expectations, and $t = 1$ year. A one-standard deviation increase in s_1 under $\mathbb{P}$ would induce a $\gamma\sqrt{\mathbb{V}^{\mathbb{P}}[s_1]} = -0.88\%$ drop in the annual log-return. A corresponding increase in s_1 under $\mathbb{Q}$ would induce a $\gamma\sqrt{\mathbb{V}^{\mathbb{Q}}[s_1]} = -3.68\%$ drop in the log-return, which is a far more negative return. These values indicate that going from $\mathbb{P}$ to $\mathbb{Q}$ a substantial amount of probability mass is shifted to the left tail of the return distribution

22 Equation (19) can be solved for θ explicitly. For example, when s_t is a Gamma subordinator, $\theta = (\exp(\beta/\nu^{\mathbb{Q}}) - 1 + \gamma/\nu^{\mathbb{Q}} + \sigma^2/(2\nu^{\mathbb{Q}}))\nu^{\mathbb{Q}}/\sigma$.

23 To save space and given their inferior time series performance, we do not report the calibrated parameters for the finite activity models.

24 From Supplementary Table 1, $\mathbb{V}[s_1] = (1-\alpha)/\nu$ for tempered stable subordinators.

Table 5. Parameters and option pricing performance of time-changed Lévy models.

Model	IS	$\nu^{\mathbb{Q}}$	$\alpha^{\mathbb{Q}}$	a_y	b_y	a_m	b_m	θ	OOS
Panel A									
VGSV1	[0.009]	15.74		3.130	0.890			0.036	[0.010]
		(1.009)		(0.382)	(0.000)				
FVGSV1	0.904	12.09		1.830	0.831			0.036	0.938
		(1.388)		(0.375)	(0.007)				
NIGSV1	0.932	6.10		10.929	1.077			0.071	0.947
		(0.194)		(1.099)	(0.005)				
FNIGSV1	0.899	3.84		7.555	1.040			0.036	0.939
		(0.153)		(1.004)	(0.022)				
NTSSV1	0.883	5.08	0.792	2.940	0.303			0.014	0.911
		(0.137)	(0.048)	(0.064)	(0.000)				
FNTSSV1	0.871	3.99	0.691	3.574	0.215			0.038	0.903
		(0.050)	(0.006)	(0.404)	(0.000)				

Model	IS	$\nu^{\mathbb{Q}}$	$\alpha^{\mathbb{Q}}$	a_y	b_y	a_m	b_m	θ	OOS
Panel B									
VGSV2	0.874	11.04		3.60	0.301	0.250	0.091	0.007	0.909
		(1.073)		(0.022)	(0.009)	(0.006)	(0.000)		
FVGSV2	0.870	13.72		2.990	0.308	1.700	0.097	0.022	0.899
		(1.374)		(0.055)	(0.004)	(0.000)	(0.000)		
NIGSV2	0.894	7.95		4.088	0.319	0.987	0.250	0.066	0.918
		(1.105)		(0.329)	(0.011)	(0.009)	(0.005)		
FNIGSV2	0.879	9.07		5.704	0.470	0.848	0.150	0.047	0.908
		(2.004)		(0.922)	(0.019)	(0.001)	(0.000)		
NTSSV2	0.875	3.84	0.681	3.330	0.107	1.739	0.082	0.083	0.903
		(0.008)	(0.000)	(0.818)	(0.000)	(0.004)	(0.000)		
FNTSSV2	0.806	3.74	0.759	4.505	0.273	1.900	0.088	0.046	0.894
		(0.099)	(0.005)	(0.138)	(0.000)	(0.000)	(0.000)		

Risk neutral and risk premium parameters are calibrated to in-sample option data as described in Section 4.2. Risk-neutral parameters $\nu_{\mathbb{Q}}$ and $\alpha_{\mathbb{Q}}$ determine the Lévy jump measure (Table 1); volatility risk-premiums a_y, b_y for SV1 models, and a_y, b_y, a_m, b_m for SV2 models are defined in (20) and (21), respectively; Brownian risk premium θ is defined in (16). Model labels are described at the beginning of Section 2.3. Model pricing performance is measured as the price RMSE. IS (OOS) is the in-sample (out-of-sample) RMSE of the corresponding model divided by the RMSE of VGSV1 (benchmark) model. For VGSV1 IS and OOS in square brackets are the RMSE. Option data spans January 1996 to December 2019. In-sample period is from January 1996 to mid-January 2013. Out-of-sample period is from mid-January 2013 to December 2019.

because of the infinite activity process $s(t)$. In other words, the risks captured by the Lévy process $s(t)$ carry a significant risk premium.

To further illustrate the JRP in $s(t)$, Figure 3 shows the Lévy measure in (4) estimated under $\mathbb{Q}$ and under $\mathbb{P}$ for the tempered stable processes in the FNTSSV2 model, for a select range of jump sizes. The $\mathbb{Q}$-Lévy measure is systematically higher than the $\mathbb{P}$-Lévy measure. For example, jumps of size larger than 0.01 are expected to occur about six times per year under $\mathbb{P}$, while their frequency doubles under $\mathbb{Q}$.[25] Also, jumps larger than 0.05 have nearly zero measure under $\mathbb{P}$, but their frequency is still around 2.5 jumps per year under $\mathbb{Q}$.

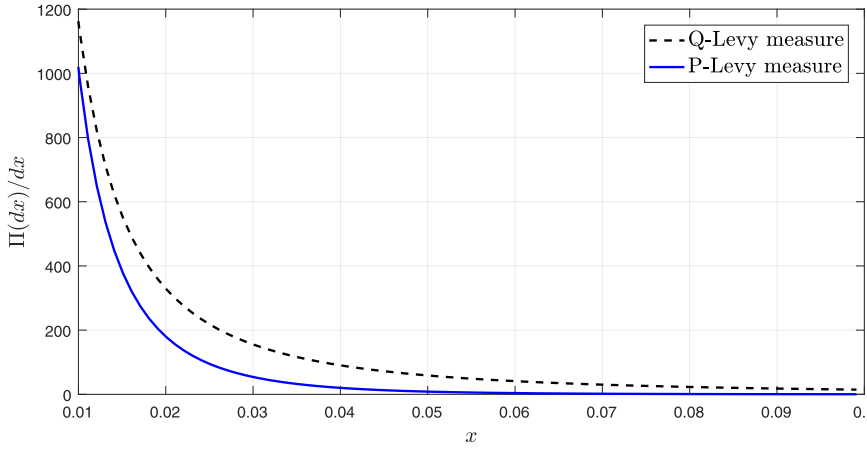

Figure 3. Lévy measures of tempered stable processes. The Lévy measure $\Pi(dx)$ in (4) is the expected number, per unit time, of jumps of size dx. The graph shows the Lévy densities under $\mathbb{Q}$ and $\mathbb{P}$ estimated in the FNTSSV2 model. Tables 3 and 5 show the estimates of the parameters in the Lévy measures.

Higher-order moments of the Lévy subordinator under $\mathbb{P}$ and $\mathbb{Q}$ are also quite different.[26] In the FNTSSV2 model, the skewness of $s(t)$ under $\mathbb{P}$ is 0.42, while under $\mathbb{Q}$ is 1.31, that is three times larger. For the other models, risk premia attached to Lévy subordinators are somewhat different but always substantial. For example, in the VGSV1 model, $\mathbb{V}^{\mathbb{P}}[s_1] = 0.009$ and $\mathbb{V}^{\mathbb{Q}}[s_1] = 0.064$, which correspond to a -0.98% and -2.60% drop in the annual log-return when $s(t)$ increases by a one-standard deviation under $\mathbb{P}$ and $\mathbb{Q}$, respectively.

Turning our attention to volatility risk premia, Table 5 reports estimates of the extended affine risk premium specifications. The coefficients a_y, b_y, a_m, and b_m are estimated to be positive across all models, meaning that volatility risk premia are negative, which is in line with the extant literature.[27] In the SV2 model, the level of risk premia is controlled by a_y and a_m when y_t and m_t increase. Because estimates of a_y are substantially larger than estimates of a_m, the fast moving factor y_t commands a larger risk premium than the slow moving factor m_t. Aït-Sahalia, Karaman, and Mancini (2020) reach a similar conclusion analyzing variance swap data.

Most studies in the literature consider the affine VRP specification ($b_y = b_m = 0$), whereas the extended affine specification is rarely estimated. An advantage of extended affine over affine specifications is the large flexibility afforded by the former and thus its ability to uncover potential non-linearities in risk premia. A relevant question is therefore to what extent such non-linearities are actually present in the data. Figure 4 shows the estimated extended affine risk premia of y_t and m_t in the FNTSSV2 model. Non-linearities in risk premia become very pronounced when y_t and m_t are close to zero. However, inspecting

25 Recall that $\int_A \Pi(dx)$ is the expected number of jumps, per unit time, whose size lies in A.

26 The j-th centered moment of s_1 can be computed differentiating j times with respect to $\tilde{u}$ the log moment generating function $\psi(-i\tilde{u})$ from Table 1 and evaluating it at $\tilde{u} = 0$.

27 Recall that the risk premium per unit of volatility of y_t and m_t is given by $-a_y\sqrt{y_t} - b_y/\sqrt{y_t}$ and $-a_m\sqrt{m_t} - b_m/\sqrt{m_t}$, respectively.

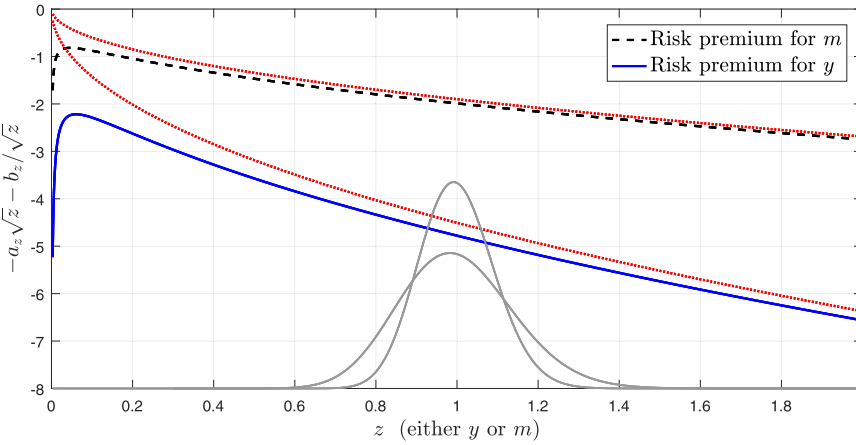

Figure 4. VRP. Extended affine market price of risk specification in (21) for the SV2 process with $\mathbb{Q}$ dynamic in (22) in the FNTSSV2 model. Table 5 reports estimates of the risk premia parameters in (21). Dotted lines are affine approximations of the extended affine specifications. Probability densities of y and m are shown at the bottom of the graph (density of y is more spread out than density of m).

the estimated parameters controlling the dynamic of the SV2 process reveals that the realizations of y_t and m_t virtually never approach zero. To illustrate this point, superimposed in Figure 4 are the estimated unconditional $\mathbb{P}$ probability densities of y_t and m_t.[28] While y_t is more volatile than m_t, their probability densities are essentially zero for values of y_t and m_t in the range from 0 to 0.2, where non-linearities of risk premia are pronounced.

Figure 4 also displays the VRP obtained by setting $b_y = b_m = 0$ in the extended risk premium specifications, which can be interpreted as the restricted affine specifications of the risk premium. These affine specifications appear to approximate quite closely the shape of the extended specifications for a wide range of values of y_t and m_t, where their probability densities are concentrated, and non-linearities are almost absent.

Figure 5 visualizes the impact of risk premia on the log-return density. In the FNTSSV2 model, we set $v_0 = m_0 = 1$, which are their unconditional values, and we consider the horizon of one week. As expected, the density under $\mathbb{P}$ is negatively skewed, reflecting the negative skewness of log-returns as in (5). When we allow for the VRP only, the density changes marginally relative to the $\mathbb{P}$ density, meaning that the VRP is modest over the one-week horizon. When instead we allow for the JRP only, the resulting density is close to the $\mathbb{Q}$ density, especially in the left tail, underscoring the relevance of the JRP over the one-week horizon.

Finally, we consider the conditional ERP defined as

28 Because m_t follows an autonomous SV1 process, from the Cox–Ingersoll–Ross model, it is known that the unconditional density of m_t is a Gamma distribution with probability density $f(m) = \omega^\nu/\Gamma(\nu)m^{\nu-1}\exp(-\omega m)$, where $\Gamma(\nu)$ is the Gamma function evaluated at ν, and $\omega = \nu = 2\kappa_m/\sigma_m^2$. The unconditional density of y_t in the SV2 model can be computed by numerical integration $f(y) = \int_0^\infty f(y|m)f(m)dm$, where $f(y|m) = \tilde{\omega}^{\tilde{\nu}}/\Gamma(\tilde{\nu})y^{\tilde{\nu}-1}\exp(-\tilde{\omega}y)$ with parameters $\tilde{\omega} = 2\kappa_y/\sigma_y^2$ and $\tilde{\nu} = 2\kappa_y m/\sigma_y^2$.

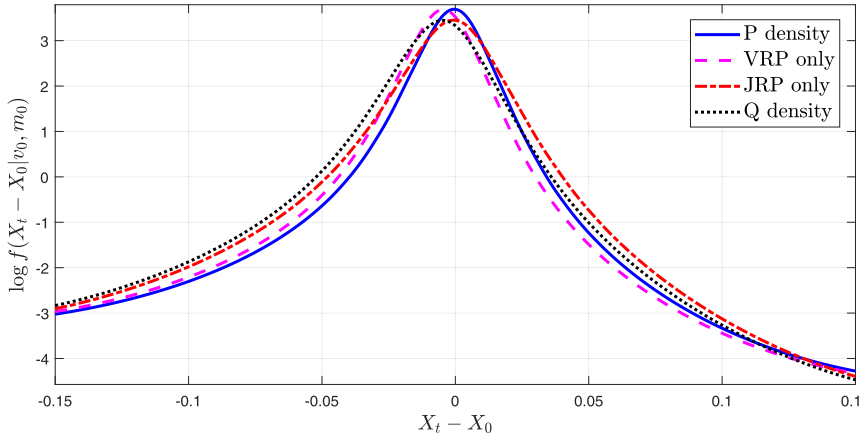

Figure 5. Log-return density. Log density of the log-return in the FNTSSV2 model (i) under the $\mathbb{P}$ measure, (ii) when the VRP only is present, (iii) when the JRP only is present, and (iv) under the $\mathbb{Q}$ measure. The time horizon t is one week and $v_0 = m_0 = 1$.

$$\mathrm{ERP}_{k\Delta} = \mathbb{E}^{\mathbb{P}}_{\mathcal{O}_{k\Delta}}\left[\frac{S_{n\Delta} - S_{k\Delta}}{S_{k\Delta}}\right] - \mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_{k\Delta}}\left[\frac{S_{n\Delta} - S_{k\Delta}}{S_{k\Delta}}\right] \tag{25}$$

where $(n-k)\Delta$ is one year. The dynamic of the ERP is determined by the rate of time changes $y_{k\Delta}$ and $m_{k\Delta}$ through their volume-based proxy $\tilde{y}_{k\Delta}$ and $\tilde{m}_{k\Delta}$. The time-change Lévy models produce an ERP of about 5%, which is in line with historical estimates; see, for example, Mehra (2006). To assess the impact of jump and volatility risks on the ERP, we define two "hybrid measures", $\mathbb{P}^J$ and $\mathbb{P}^{SV}$. The first measure $\mathbb{P}^J$ is equal to $\mathbb{P}$ except that the Lévy jump parameters are the risk neutral ones. This measure reflects the risk premium due to Lévy jumps only, and we define $\mathrm{ERP}^J_{k\Delta}$ replacing $\mathbb{P}$ by $\mathbb{P}^J$ in (25). The second measure $\mathbb{P}^{SV}$ is equal to $\mathbb{P}$ but the parameters controlling the stochastic volatility process are the risk neutral parameters. The measure $\mathbb{P}^{SV}$ reflects the risk premium due to stochastic volatility only, and we define $\mathrm{ERP}^{SV}_{k\Delta}$ replacing $\mathbb{P}$ by $\mathbb{P}^{SV}$ in (25). We compute $\mathrm{ERP}^J_{k\Delta}$ and $\mathrm{ERP}^{SV}_{k\Delta}$ throughout the entire sample period, from 1950 to 2019, using the estimated time-change Lévy models. In line with the findings in Ornthanalai (2014), for most models the relative magnitude of $\mathrm{ERP}^J_{k\Delta}$ and $\mathrm{ERP}^{SV}_{k\Delta}$ indicates that almost 50% of the ERP is due to infinite activity Lévy jumps.

In sum, infinite activity processes appear to carry a large JRP, while the VRP is well approximated by affine specifications at least over the bulk of the volatility distribution. These findings are based on the FNTSSV2 model, which outperforms many competing models in fitting index returns and, as discussed next, in pricing options.

4.2.2 Option pricing performance.

As is customary in the literature, we evaluate the option pricing performance of each model using the root mean square error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\tilde{P}_i^{mkt} - \tilde{P}_i^{mod})^2}$$

where N is the number of options, $\tilde{P}^{mkt} = P^{mkt}/S$ and $\tilde{P}^{mod} = P^{mod}/S$ are the observed and model-based option prices divided by the level of the S&P 500 index, respectively. Prior empirical option pricing applications of time-changed Lévy models, such as Huang and Wu (2004), have focused on the VGSV1 model, and therefore we use this model as a benchmark. In these applications, unobservable variables driving the time change are recovered from option prices as if they were model parameters, every time the model is calibrated. This approach simplifies the application of time-changed Lévy models but can generate time series trajectories of time changes that are inconsistent with the assumed model. We do not follow this approach and use instead the volume-based proxy of the time change as in (15).

Table 5 reports the RMSE, averaging across all moneyness and maturity options, for each model relative to the RMSE of the VGSV1 model. We can summarize the empirical findings as follows. First, the FNTSSV2 model consistently outperforms all other models and often by a large extent. The FNTSSV2 has an in-sample RMSE almost 20% lower than VGSV1 and a similar performance out-of-sample. Second, models that take into account the observation error in the volume-based proxy of the time change as in (15) always outperform models that do not. This finding indicates that our volume-based proxy of the time change is effective also for pricing options, even though observation errors carry no risk premium. Third, comparing the RMSE of models with the same number of volatility factors allows us to gage the importance of infinite activity $\mathbb{Q}$-subordinators for pricing options. For example, in-sample RMSE of FVGSV2, FNIGSV2, and FNTSSV2 relative to VGSV1 are 0.870, 0.879, and 0.806, respectively. While Gamma and Inverse Gaussian subordinators tend to perform similarly, the Normal Tempered Stable subordinator stands out and induces a sizable 20% improvement in RMSE.

Table 6 shows the in-sample pricing performance of selected models across moneyness and maturities. To save space we consider only four models, namely PSV1, VGSV1, FVGSV1, and FNTSSV2. Panel A reports the RMSE of the benchmark VGSV1 model. The pricing performance of this model is rather uniform across moneyness but deteriorates when the time to maturity increases. As the VGSV1 model has only one volatility factor, it cannot accommodate the term structure of option prices. The ratios of RMSE of PSV1 to VGSV1 (Panel B) are nearly always larger than one, underscoring the benefits of using infinite rather than finite activity models. Given that improvements in RMSE tend to be largest for short maturity and out-of-the-money put options, infinity activity subordinators appear to be particularly important for modeling the left tail of the risk neutral return distribution. The ratios of RMSE of FVGSV1 to VGSV1 (Panel C) are nearly always smaller than one, which highlight the benefits of allowing for the observation error in the volume-based proxy of time changes as in (15). Fitting improvements are largest for options around the moneyness as these options are most sensitive to volatility. The ratios of FNTSSV2 to VGSV1 (Panel D) are substantially smaller than one in most cases. Using a flexible Lévy measure and a two-factor stochastic volatility process appear to be key for pricing options.

Table 6. In-sample option pricing performance across moneyness and maturitiesPanel A: VGSV1

	$T < 30$	$30 \leq T < 80$	$80 \leq T < 180$	$180 \leq T < 250$	$250 \leq T$
$K/S < 0.95$	0.005	0.008	0.016	0.015	0.019
$0.95 \leq K/S < 0.975$	0.005	0.008	0.016	0.018	0.017
$0.975 \leq K/S < 1$	0.004	0.008	0.014	0.018	0.018
$1 \leq K/S < 1.025$	0.004	0.007	0.012	0.016	0.017
$1.025 \leq K/S < 1.05$	0.003	0.007	0.011	0.016	0.017
$1.05 \leq K/S < 1.10$	0.003	0.008	0.013	0.015	0.019
$1.10 \leq K/S$	0.001	0.007	0.015	0.014	0.020
Panel B: PSV1/VGSV1					
$K/S < 0.95$	1.115	1.109	1.118	1.094	1.001
$0.95 \leq K/S < 0.975$	1.092	1.095	1.094	1.080	1.055
$0.975 \leq K/S < 1$	1.081	1.088	1.079	1.084	1.080
$1 \leq K/S < 1.025$	1.062	1.093	1.073	1.060	1.029
$1.025 \leq K/S < 1.05$	1.094	1.075	1.066	1.038	0.990
$1.05 \leq K/S < 1.10$	1.097	1.090	1.082	1.071	1.041
$1.10 \leq K/S$	1.090	1.085	1.086	1.079	1.063
Panel C: FVGSV1/VGSV1					
$K/S < 0.95$	0.984	0.992	0.963	0.997	1.052
$0.95 \leq K/S < 0.975$	0.976	0.944	0.947	0.981	1.009
$0.975 \leq K/S < 1$	0.925	0.991	0.919	0.942	0.955
$1 \leq K/S < 1.025$	0.911	0.871	0.895	0.905	0.939
$1.025 \leq K/S < 1.05$	0.920	0.954	0.940	0.937	0.964
$1.05 \leq K/S < 1.10$	0.951	0.968	0.955	0.959	0.986
$1.10 \leq K/S$	0.998	0.983	0.959	0.981	1.003
Panel D: FNTSSV2/VGSV1					
$K/S \leq 0.95$	0.805	0.811	0.819	0.827	0.991
$0.95 \leq K/S < 0.975$	0.795	0.796	0.804	0.821	0.988
$0.975 \leq K/S < 1$	0.784	0.775	0.793	0.796	0.990
$1 \leq K/S < 1.025$	0.770	0.742	0.758	0.772	0.960
$1.025 \leq K/S < 1.05$	0.779	0.804	0.781	0.785	1.004
$1.05 \leq K/S < 1.10$	0.802	0.817	0.794	0.791	1.011
$1.10 \leq K/S$	0.814	0.839	0.827	0.849	1.016

Panel A: Price RMSEs of VGSV1 across moneyness and maturity of out-of-the-money options on the S&P 500 index. Panel B: Ratio of RMSEs of PSV1 to VGSV1. Panel C: Ratio of RMSEs of FVGSV1 to VGSV1. Panel D: Ratio of RMSEs of FNTSSV2 to VGSV1. Maturity T is in days. Moneyness is K/S, where K and S are the strike and underlying price, respectively. Model parameters are from Tables 2, 3 and 5. In-sample period is from January 1996 to mid-January 2013.

Improvements in RMSE can be as large as 25% when pricing around the moneyness options with intermediate maturities.[29]

We now compare the pricing performance of the time-changed Lévy models to existing models in the literature. The comparison should be interpreted cautiously as different studies rely on different SPX option data and sample periods. We consider the RMSE of implied volatilities in percentage from in-sample analyses, as some studies do not provide out-of-sample analyses. Gruber, Tebaldi, and Trojani (2021; Table 3 in their online Appendix) report RMSE from 0.68% to 1.32% for their matrix affine jump diffusion models, but their (in-sample) option data is relatively short and spans only January 1996 to December 2002. Andersen et al. (2015b; their Table 6) report a RMSE of 1.71% for their best performing three-factor model, and RMSE up to 3.14% for their benchmark models. Their option sample spans January 1996 to July 2010. Corsi et al. (2013; their Table 4) report a RMSE of 3.82% for their HARGL model, for an option sample spanning January 1996 to December 2004. The RMSE of the VGSV1 model and the best performing model (FNTSSV2) are 2.83% and 1.97%, respectively. The latter RMSE is of similar magnitude to the lowest RMSE reported in the literature. We note that our (in-sample) data is largest and spans January 1996 to January 2013, challenging the time-changed Lévy models to fit such a large sample.

As a robustness check, to understand how "erratic" OTM puts could impact the calibration exercise, we re-calibrated selected models excluding deep OTM options. The model parameters were almost unaffected and the pricing performance of most models improved somewhat relative to the RMSE reported in Tables 5 and 6. The model ranking and overall pricing performance remain the same.

We close this section with an additional out-of-sample analysis. Israelov and Kelly (2017) report that affine models cannot predict future option returns. To understand whether time-changed Lévy models have any forecasting power for option returns, we carry out the following simple exercise. In the out-of-sample period (January 2013 to December 2019), we consider OTM options with maturity between one week and one month, about 203,000 options. At each date $k\Delta$, we forecast the terminal payoff of each option. For example, we compute $\mathbb{E}^{\mathbb{P}}_{\mathcal{O}_{k\Delta}}[\max(S_{n\Delta} - K, 0)]$ for a call option with strike price K and maturity $n\Delta > k\Delta$. The forecast is straightforward to compute and obtained via Fourier inversion. As in Israelov and Kelly (2017), to evaluate the forecast accuracy we run Mincer–Zarnowitz regressions. That is, we regress the realized payoff $\max(S_{n\Delta} - K, 0)$ on a constant and the forecast $\mathbb{E}^{\mathbb{P}}_{\mathcal{O}_{k\Delta}}[\max(S_{n\Delta} - K, 0)]$. If the forecast is accurate, the intercept is zero and the slope is one, in a statistical sense. The R^2 of the regression provides an overall measure of the forecast accuracy. Across the various time-changed Lévy models, regression results indicate that the intercepts are around –0.00004 and slopes around 1.0001, and are only occasionally statistically away from zero and one, respectively. The models have a tendency to overestimate the (zero) payoff of options that expire OTM, and to underestimate the payoff of options that expire in-the-money. These two effects tend to cancel out, which produces slopes of the Mincer–Zarnowitz regressions close to one. The R^2's are however quite large and around 20%, indicating that time-changed Lévy models have forecasting power of option payoffs.

29 Out-of-sample pricing errors across moneyness and maturities largely confirm the in-sample pricing results and are not reported.

5 Conclusion

We develop a novel class of time-changed Lévy models that are characterized by four nested processes, capturing market, volatility, and jump risks as well as observation error of time change. To study risk premia, we derive the change of measure analytically. To operationalize the models, we propose to use volume-based proxies for the unobservable time changes, taking advantage of the strong positive relation between volume and volatility. An extensive time series and option pricing analysis shows that infinite activity models largely outperforms many finite activity models. The best performing model exhibits a flexible tempered stable Lévy subordinator, a two-factor volatility process, and allows for observation error of the volume-based proxy. Infinite activity processes appear to carry substantial JRP, while the VRP seems to follow relatively simple affine dynamics.

In future work, our time-changed Lévy models could be used to investigate whether option and volatility markets are truly segmented or instead flexible models can actually reconcile the empirical regularities of both markets, following a similar analysis as in Bardgett et al. (2019). Moreover, because the characteristic function of asset returns is readily available, our models could be used to study exotic derivatives, such as caps and swaptions as in Leippold and Stromberg (2014). Another direction for future research would be to study the link between time changes and different trading frequencies of various agents.

Appendix A: The Leverage Effect

To facilitate readability of lengthy expressions we use a light notation, omitting subscripts whenever possible. The SV2 family of models in (8) is defined by the system of stochastic differential equations

$$\begin{aligned} dY_t &= y_t dt \\ dy_t &= \kappa_y(m_t - y_t)dt + \sigma_y\sqrt{y_t}dW_t^y \\ dm_t &= \kappa_m(1 - m_t)dt + \sigma_m\sqrt{m_t}dW_t^m \end{aligned}$$

where W^y and W^m are independent Brownian motions.

Define the instantaneous variance of the return process by

$$v_t \equiv \lim_{\Delta\to 0}\frac{\mathbb{V}[X_{t+\Delta} - X_t|\mathcal{F}_t]}{\Delta}$$

where $\mathcal{F}_t$ is the time-t information set. Irrespective of the model used for the process y_t in (5) we have that

$$\mathbb{V}[X_{t+\Delta} - X_t|\mathcal{F}_t] = (\beta+\gamma)^2\mathbb{V}[Y_{t+\Delta} - Y_t|\mathcal{F}_t] + (\sigma^2 + \gamma^2\mathbb{V}[s_1])\mathbb{E}[Y_{t+\Delta} - Y_t|\mathcal{F}_t]$$

and it immediately follows that

$$v_t = (\sigma^2 + \gamma^2\mathbb{V}[s_1])y_t.$$

To prove this relation we start by observing that

$$\begin{aligned}\mathbb{V}[X_{t+\Delta} - X_t|\mathcal{F}_t] = \beta^2\mathbb{V}[Y_{t+\Delta} - Y_t|\mathcal{F}_t] + \gamma^2\mathbb{V}[s_{Y_{t+\Delta}} - s_{Y_t}|\mathcal{F}_t] \\ +\sigma^2\mathbb{V}[W(s_{Y_{t+\Delta}}) - W(s_{Y_t})|\mathcal{F}_t] \\ +2\sigma\beta\mathrm{Cov}[Y_{t+\Delta} - Y_t, W(s_{Y_{t+\Delta}}) - W(s_{Y_t})|\mathcal{F}_t] \\ +2\gamma\beta\mathrm{Cov}[Y_{t+\Delta} - Y_t, s_{Y_{t+\Delta}} - s_{Y_t}|\mathcal{F}_t] \\ +2\gamma\sigma\mathrm{Cov}[s_{Y_{t+\Delta}} - s_{Y_t}, W(s_{Y_{t+\Delta}}) - W(s_{Y_t})|\mathcal{F}_t].\end{aligned}$$

Then, using the properties of the Lévy process s_t and the fact that W_t is a Brownian motion we deduce that

$$\begin{aligned}\mathbb{V}[s_{Y_{t+\Delta}} - s_{Y_t}|\mathcal{F}_t] &= \mathbb{V}[Y_{t+\Delta} - Y_t|\mathcal{F}_t] + \mathbb{V}[s_1]\mathbb{E}[Y_{t+\Delta} - Y_t|\mathcal{F}_t] \\ \mathbb{V}[W(s_{Y_{t+\Delta}}) - W(s_{Y_t})|\mathcal{F}_t] &= \mathbb{E}[s_{Y_{t+\Delta}} - s_{Y_t}|\mathcal{F}_t] = \mathbb{E}[Y_{t+\Delta} - Y_t|\mathcal{F}_t] \\ \mathrm{Cov}[Y_{t+\Delta} - Y_t, W(s_{Y_{t+\Delta}}) - W(s_{Y_t})|\mathcal{F}_t] &= 0 \\ \mathrm{Cov}[Y_{t+\Delta} - Y_t, s_{Y_{t+\Delta}} - s_{Y_t}|\mathcal{F}_t] &= \mathbb{V}[Y_{t+\Delta} - Y_t|\mathcal{F}_t] \\ \mathrm{Cov}[s_{Y_{t+\Delta}} - s_{Y_t}, W(s_{Y_{t+\Delta}}) - W(s_{Y_t})|\mathcal{F}_t] &= 0\end{aligned}$$

and the result follows.

All expected values, variances, and covariances below are conditional on time-0 information set. For simplicity, we omit such a dependence. To compute the leverage effect implied by the model, we need to calculate

$$\begin{aligned}\mathrm{Cov}[\nu_t, X_t] &= (\sigma^2 + \gamma^2\mathbb{V}[s_1])\mathrm{Cov}[y_t, X_t] \\ &= (\beta + \gamma)(\sigma^2 + \gamma^2\mathbb{V}[s_1])\mathrm{Cov}[y_t, Y_t]\end{aligned}$$

where the second equality follows from the fact that

$$\begin{aligned}\mathbb{E}[y_t X_t] &= \mathbb{E}[\beta y_t Y_t + \gamma y_t s_{Y_t} + \sigma y_t W(s_{Y_t})] \\ &= \mathbb{E}[\beta y_t Y_t + \gamma y_t Y_t] = (\beta + \gamma)\mathbb{E}[y_t Y_t] \\ \mathbb{E}[y_t]\mathbb{E}[X_t] &= \mathbb{E}[y_t]\mathbb{E}[\beta Y_t + \gamma s_{Y_t} + \sigma W(s_{Y_t})] = \mathbb{E}[y_t]\mathbb{E}[\beta Y_t + \gamma Y_t] \\ &= (\beta + \gamma)\mathbb{E}[y_t]\mathbb{E}[Y_t].\end{aligned}$$

Since

$$\begin{aligned}\mathrm{Cov}[y_t, Y_t] &= \mathbb{E}[y_t \textstyle\int_0^t y_u du] - \mathbb{E}[y_t]\mathbb{E}[\int_0^t y_u du] \\ &= \textstyle\int_0^t \mathbb{E}[y_u y_t] du - \int_0^t \mathbb{E}[y_u]\mathbb{E}[y_t] du = \int_0^t \mathrm{Cov}[y_u, y_t] du\end{aligned}$$

we have that

$$\mathrm{Cov}[\nu_t, X_t] = (\beta + \gamma)(\sigma^2 + \gamma^2\mathbb{V}[s_1]) \int_0^t \mathrm{Cov}[y_u, y_t] du$$

and it follows that, irrespective of the model that is used for y_t, the leverage effect is determined by the sign of

$$(\beta + \gamma)\mathrm{Cov}[y_u, y_t].$$

In the SV1 model, we have that

$$(\beta+\gamma)\text{Cov}[y_u, y_t] = (\beta+\gamma)e^{-\kappa_y(t-u)}\mathbb{V}[y_u]$$

so that the leverage effect is entirely determined by the sign of $\beta+\gamma$.

In the SV2 model, the situation is slightly more complex. Using the fact that

$$y_t = e^{-\kappa_y(t-u)}y_u + \kappa_y\int_u^t e^{-\kappa_y(t-s)}m_s ds + \int_u^t e^{-\kappa_y(t-s)}\sigma_y\sqrt{y_s}dW_s^y$$

it can be shown that

$$\text{Cov}[y_u, y_t] = e^{-\kappa_y(t-u)}\mathbb{V}[y_u] + \kappa_y\int_u^t e^{-\kappa_y(t-s)}\text{Cov}[y_u, m_s]ds$$

and a further calculation gives

$$\begin{aligned}\text{Cov}[y_u, m_s] &= \text{Cov}[y_u, \mathbb{E}[m_s|\mathcal{F}_u]]\\ &= \text{Cov}[y_u, e^{-\kappa_m(s-u)}m_u + \left(1-e^{-\kappa_m(s-u)}\right)]\\ &= e^{-\kappa_m(s-u)}\text{Cov}[y_u, m_u]\end{aligned}$$

for all $u \leq s$ so that the same result as in SV1 will hold provided that $\text{Cov}[y_u, m_u] \geq 0$. The specification of the model implies that

$$\begin{aligned}y_u - \mathbb{E}[y_u] &= e^{-\kappa_y u}(M_u^y + A_u - \mathbb{E}[A_u])\\ m_u - \mathbb{E}[m_u] &= e^{-\kappa_m u}M_u^m = m_u - e^{-\kappa_m u}m_0 - \kappa_m\textstyle\int_0^u e^{-\kappa_m(u-s)}ds\end{aligned}$$

where

$$A_u = \kappa_y\int_0^u e^{\kappa_y x}m_x dx$$

and the processes (M_u^y, M_u^m) are orthogonal martingales with initial value zero. Using these expressions to compute the covariance, we obtain that

$$\begin{aligned}\text{Cov}[y_u, m_u] &= \mathbb{E}[(y_u - \mathbb{E}[y_u])(m_u - \mathbb{E}[m_u])]\\ &= \mathbb{E}[e^{-(\kappa_y+\kappa_m)u}M_u^m(M_u^y + A_u - \mathbb{E}[A_u])]\\ &= e^{-\kappa_y u}\mathbb{E}[e^{-\kappa_m u}M_u^m A_u] = e^{-\kappa_y u}\mathbb{E}[A_u(m_u - \mathbb{E}[m_u])]\\ &= \kappa_y\mathbb{E}[\textstyle\int_0^u e^{-\kappa_y(u-x)}m_x(m_u - \mathbb{E}[m_u])dx]\\ &= \kappa_y\mathbb{E}[\textstyle\int_0^u e^{-\kappa_y(u-x)}(m_x - \mathbb{E}[m_x])(m_u - \mathbb{E}[m_u])dx]\\ &= \kappa_y\textstyle\int_0^u e^{-\kappa_y(u-x)}\text{Cov}[m_x, m_u]dx = \kappa_y\int_0^u e^{-(\kappa_y+\kappa_m)(u-x)}\mathbb{V}[m_x]dx\end{aligned}$$

where the last equality follows from the fact that

$$\begin{aligned}\text{Cov}[m_x, m_u] &= \text{Cov}[m_x, \mathbb{E}[m_u|\mathcal{F}_x]]\\ &= \text{Cov}[m_x, e^{-\kappa_m(u-x)}m_x + \left(1-e^{-\kappa_m(u-x)}\right)]\\ &= e^{-\kappa_m(u-x)}\text{Cov}[m_x, m_x] = e^{-\kappa_m(u-x)}\mathbb{V}[m_x]\end{aligned}$$

for all $x \leq u$.

Appendix B —Example: Variance Gamma SV1 Model

As an illustrative example of the time-changed Lévy model in (1), we describe the Variance Gamma SV1 model. In this model, s_t is a Gamma subordinator (see Table 1) and y_t follows the SV1 process in (7). The time-t instantaneous variance is given by

$$v_t = (\sigma^2 + \gamma^2 \mathbb{V}[s_1])y_t = (\sigma^2 + \gamma^2/\nu)y_t$$

as $\mathbb{V}[s_1] = 1/\nu$ from Table 1. In contrast to models based on SV2 processes, the characteristic function of X_t can be computed explicitly, and is given by

$$\Phi(u; X_t) = \exp(A_1(t; iq(u)) + B_1(t; iq(u))y_0) \tag{26}$$

with the functions

$$A_1(t; iq) = -\frac{\kappa_y}{\sigma_y^2}\left[2\log\left(\frac{2\eta - (\eta - \kappa_y)(1 - e^{-\eta t})}{2\eta}\right) + (\eta - \kappa_y)t\right]$$

$$B_1(t; iq) = \frac{2iq(1 - e^{-\eta t})}{2\eta - (\eta - \kappa_y)(1 - e^{-\eta t})}$$

and $\eta = \sqrt{\kappa_y^2 - 2iq\sigma_y^2}$. The characteristic exponent $\psi(u)$ of the Gamma subordinator s_t is given in Table 1 and the function $q(u)$ in (10), which enters $A_1(t; iq)$ and $B_1(t; iq)$ above, is given by

$$q(u) = u\beta + i\nu \log\left(1 - i(u\gamma + iu^2\frac{\sigma^2}{2})/\nu\right).$$

Appendix C: Series Solution for the Characteristic Function

We use the light notation in Appendix A and write the SV2 model as

$$dy_t = \kappa_y(\theta_{y,c} + \theta_{y,d}m_t - y_t)dt + \sigma_y\sqrt{y_t}dW_t^y$$
$$dm_t = \kappa_m(\theta_m - m_t)dt + \sigma_m\sqrt{m_t}dW_t^m.$$

The specification above encompasses both $\mathbb{P}$- and $\mathbb{Q}$-dynamics. For example, $\theta_{y,c} = 0$ and $\theta_{y,d} = \theta_m = 1$ under $\mathbb{P}$. The parameter specification under $\mathbb{Q}$ is in (22).

The characteristic function of the time change $(T - t) \mapsto \int_t^T y_s ds$ is defined by

$$f(t, y_t, m_t; u) = \mathbb{E}_t[e^{u\int_t^T y_s ds}] = e^{A(T-t;u) + B(T-t;u)y_t + C(T-t;u)m_t}$$

for a given $u \in \mathbb{C}$ and some unknown functions (omitting subscripts) A, B, and C such that

$$A(0; u) = B(0; u) = C(0; u) = 0. \tag{27}$$

Combining the fact that the process

$$e^{u\int_0^t y_s ds} f(t, y_t, m_t; u)$$

is a martingale with a standard separation of variables argument shows that these functions must solve the system of ordinary differential equations given by

$$A'(\tau;u) = \kappa_y\theta_{y,c}B(\tau;u) + \kappa_m\theta_m C(\tau;u)$$
$$B'(\tau;u) = u - \kappa_y B(\tau;u) + \frac{1}{2}\sigma_y^2 B(\tau;u)^2$$
$$C'(\tau;u) = \kappa_y\theta_{y,d}B(\tau;u) - \kappa_m C(\tau;u) + \frac{1}{2}\sigma_m^2 C(\tau;u)^2$$

subject to (27). With the exception of the function $B(\tau;u)$, which actually solves an autonomous equation, it is not possible to derive an explicit solution to this system. Instead, we construct a power series solution by postulating that

$$H(\tau;u) = \sum_{k=0}^{\infty} h_k(u)\tau^k, \quad (h,H) \in \{(a,A),(b,B),(c,C)\}.$$

To determine the sequence $(a_k(u), b_k(u), c_k(u))_{k=0}^{\infty}$ of unknown coefficients, we start by observing that the boundary condition (27) implies

$$a_0(u) = b_0(u) = c_0(u) = 0, \quad u \in \mathbb{C}.$$

Substituting the conjectured series solution into the system of differential equations, using Cauchy's product formula

$$\Big(\sum_{k=0}^{\infty} h_k(u)\tau^k\Big)^2 = \sum_{k=0}^{\infty}\Big(\sum_{\ell=0}^{k} h_\ell(u)h_{k-\ell}(u)\Big)\tau^k$$

to compute the squares, and matching terms shows that the unknown coefficients can be computed recursively as follows:

$$a_1(u) = u - b_1(u) = c_1(u) = 0$$
$$b_{k+1}(u) = \frac{1}{1+k}\left[-\kappa_y b_k(u) + \frac{1}{2}\sigma_y^2\sum_{\ell=0}^{k} b_\ell(u)b_{k-\ell}(u)\right]$$
$$c_{k+1}(u) = \frac{1}{1+k}\left[\kappa_y\theta_{y,d}b_k(u) - \kappa_m c_k(u) + \frac{1}{2}\sigma_m^2\sum_{\ell=0}^{k} c_\ell(u)c_{k-\ell}(u)\right]$$

and

$$a_{k+1}(u) = \frac{1}{1+k}\left[\kappa_y\theta_{y,c}b_k(u) + \kappa_m\theta_m c_k(u)\right].$$

We confirm the accuracy of our series solution by running two main sanity checks. First, we consider the function $B(\tau;u)$ that solves an autonomous differential equation and has an analytic solution. This differential equation appears in both SV1 and SV2 models. Second, we consider the characteristic function of SV1 models that has the analytic solution given by (26). We compare the two analytic solutions with their corresponding series solutions. For the range of estimated or calibrated parameter values, the analytic solutions and the series solutions are virtually identical when the number of terms in the series solution is at least three and the time horizon τ is less than one year. In our empirical analysis, we implemented the series solution using five terms. We also experimented with ten and fifteen terms, and results were unchanged.

Appendix D: Theoretical Aspects of the Proxy of the Time Change

Let $(X_t, y_t)_{t\geq 0}$ be a Markov process, modeling excess log-return and rate of time change. Assume that we observe $(X_{k\Delta}, \tilde{y}_{k\Delta})$ at discrete dates $k\Delta, k = 0, 1, \ldots$, where $\tilde{y}_{k\Delta} = \mathcal{S}(\varepsilon_{k\Delta}, y_{k\Delta})$ is a signal for the unobservable $y_{k\Delta}$ and $\mathcal{S}$ is some function invertible in $y_{k\Delta}$.[30] Denote the σ-algebras as $\mathcal{F}_t = \sigma\{(X_s, y_s) : s \leq t\}$ and $\mathcal{O}_{k\Delta} = \sigma\{(X_{j\Delta}, \tilde{y}_{k\Delta}) : j = 0, 1, \ldots, k\}$.

Assumption. The observation error $\varepsilon_{k\Delta}$ is independent of $\{X_{(k+j)\Delta} : j > 0\}$ given $\mathcal{F}_{k\Delta}$.

Lemma. For any bounded function $g : \mathbb{R} \to \mathbb{C}$, we have that

$$\mathbb{E}[g(X_{(k+j)\Delta})|\mathcal{O}_{k\Delta}] = \mathbb{E}[\phi_g(j, X_{k\Delta}, y_{k\Delta})|\mathcal{O}_{k\Delta}] = \int \phi_g(j, X_{k\Delta}, \mathcal{S}^{-1}(\varepsilon, \tilde{y}_{k\Delta}))\mathbb{P}[\varepsilon_{k\Delta} \in d\varepsilon|\mathcal{O}_{k\Delta}]$$

for some function $\phi_g : \mathbb{N} \times \mathbb{R} \times \mathbb{R}^+ \to \mathbb{C}$.

Proof. Let g be a bounded function. Then

$$\begin{aligned}
\mathbb{E}[g(X_{(k+j)\Delta})|\mathcal{O}_{k\Delta}] &= \mathbb{E}[\mathbb{E}[g(X_{(k+j)\Delta})|\mathcal{F}_{k\Delta} \vee \sigma\{\varepsilon_{u\Delta} : u = 0, \ldots, k\}]|\mathcal{O}_{k\Delta}] \\
&= \mathbb{E}[\mathbb{E}[g(X_{(k+j)\Delta})|\mathcal{F}_{k\Delta}]|\mathcal{O}_{k\Delta}] \\
&= \mathbb{E}[\phi_g(j, X_{k\Delta}, y_{k\Delta})|\mathcal{O}_{k\Delta}] \\
&= \mathbb{E}[\phi_g(j, X_{k\Delta}, \mathcal{S}^{-1}(\varepsilon_{k\Delta}, \tilde{y}_{k\Delta}))|\mathcal{O}_{k\Delta}]
\end{aligned}$$

where the first equality follows from the tower law; the second equality follows from our assumption on $\varepsilon_{k\Delta}$; the third equality follows from the Markov property of (X_t, y_t); the last equality follows from the invertibility of the signal $\mathcal{S}(\varepsilon, y)$ with respect to y.

When the rate of time change depends on two unobservable variables, y and m, the lemma above can be extended to cover this situation by introducing an additional signal for m.

Appendix E: Fourier Inversion: COS Method with Damping Function

This section presents the method we use to recover probability density functions of time-changed Lévy processes from their characteristic functions. To achieve high accuracy and overcome the so-called Gibbs phenomenon, namely the slow decay of the real part of the characteristic function, we enrich the COS Method (Fang and Oosterlee, 2008) with a damping function.

Because the probability density function $f(x)$ of X has not an analytic form, it is approximated by a discrete series that involves its characteristic function $\Phi(\cdot; X)$. The approximation of $f(x)$ via the COS method using N terms is given by

$$f(x) = \sum_{k=0}^{N-1'} F_k \cos\left(k\pi \frac{x-a}{b-a}\right) \tag{28}$$

30 In our empirical investigation, we set $\mathcal{S}(\varepsilon_{k\Delta}, y_{k\Delta}) = y_{k\Delta}/\varepsilon_{k\Delta}$.

where $\sum'$ indicates that the first term in the summation is weighted by one-half, $[a, b] \in \mathbb{R}$ is the support over which $f(x)$ is approximated, and denoting with Re the real part of a complex number

$$F_k = \frac{2}{b-a} Re\left\{ \Phi\left(\frac{k\pi}{b-a}; X\right) \exp\left(-i\frac{ka\pi}{b-a}\right)\right\}.$$

To overcome the Gibbs phenomenon, we scale the terms in the discrete series via the damping function $g(z) = (1 + \cos(\pi z))/2$ and obtain

$$F_k = g\left(\frac{k}{N}\right)\frac{2}{b-a} Re\left\{ \Phi\left(\frac{k\pi}{b-a}; X\right) \exp\left(-i\frac{ka\pi}{b-a}\right)\right\}.$$

These new F_k terms are plugged in the discrete series (28), which is used for maximum likelihood estimation of time-changed Lévy models.

To price a European option, the probability density function of the log-price is approximated via the COS method. Let $C_0(K, T)$ denote the time-0 price of a call option with strike price K and maturity T, and $\tilde{K} = e^{-(r-\delta)T}K/S_0$. Then,

$$\begin{aligned} C_0(K,T) &= e^{-rT}\mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_0}[\max(S_T - K, 0)] = e^{-\delta T}S_0\mathbb{E}^{\mathbb{Q}}_{\mathcal{O}_0}[\max(e^{X_T} - \tilde{K}, 0)] \\ &= e^{-\delta T}S_0 \int_{-\infty}^{+\infty} \max(e^x - \tilde{K}, 0) f(x)dx \\ &\approx e^{-\delta T}S_0 \sum_{k=0}^{N-1'} Re\Phi\left(\frac{k\pi}{b-a}; X\right)\exp\left(-i\frac{k\pi a}{b-a}\right)\Bigg\} \end{aligned} \qquad V_k$$

where the approximation $\approx$ uses the discrete series (28) and

$$V_k = g\left(\frac{k}{N}\right)\frac{2}{b-a}\int_a^b \max(e^x - \tilde{K}, 0)\cos(k\pi\frac{x-a}{b-a})dx.$$

Appendix F: Estimation of Time-changed Lévy Models with UKF

This section describes our implementation of the UKF to estimate time-changed Lévy models and to recover the time change from index returns without using any volume-based proxy. To spare lengthy expression, we focus on the VGSV1 model. The other models can be estimated in a similar way.

We cast the VGSV1 model in state space form, which consists of a transition equation and a measurement equation. The transition equation describes the discrete time dynamic of the state process x which is given by the SV1 process y. The transition equation is obtained from an Euler discretization of (7) at daily frequency

$$x_{k+1} = \phi_0 + \phi_1 x_k + w_k$$

where $\phi_0 = \kappa_y\Delta$, $\phi_1 = (1 - \kappa_y\Delta)$, $w_k \sim \mathcal{N}(0, Q_k)$, $Q_k = \sigma_y^2\Delta x_k$, and $\Delta = 1/252$.[31] The measurement equation provides the relation between index returns and the state process, and is given by

31 In SV2 models, the state process is two-dimensional $x = [y\ m]'$, where $'$ denotes transposition, and the transition equation is obtained from discretizing (8).

$$z_k = h(x_k) + v_k$$

where $v_k \sim \mathcal{N}(0, R)$. We obtain the function $h(x_k)$ from the first two conditional moments of daily log-returns

$$h(x_k) = \begin{bmatrix} (\beta + \gamma)\Delta x_k \\ (\beta + \gamma)^2(\Delta x_k)^2 + (\gamma^2/\nu + \sigma^2)\Delta x_k \end{bmatrix}$$

which gives a nonlinear function h, hence the necessity to use the UKF. To simplify the functional form of $h(x_k)$, we approximate the first two moments of the log-returns up to Δ^2 which is of the order 10^{-5}, resulting in an accurate approximation at daily frequency. Below we provide a brief discussion of the UKF, starting with the classic Kalman filter.

If the function $h(x_k)$ were linear, that is, $h(x_k) = h_0 + h_1 x_k$, the Kalman filter would provide efficient estimates of the conditional mean and variance of the state vector. Let $\hat{x}_{k|k-1} = \mathbb{E}_{k-1}[x_k]$ and $\hat{z}_{k|k-1} = \mathbb{E}_{k-1}[z_k]$ denote the expectation of x_k and z_k, respectively, using information up to and including time k–1, and let $P_{k|k-1}$ and $F_{k|k-1}$ denote the corresponding error covariance matrices. Furthermore, let $\hat{x}_k = \mathbb{E}_k[x_k]$ denote the expectation of x_k including information at time k, and let P_k denote the corresponding error covariance matrix. The Kalman filter consists of two steps: prediction and update. In the prediction step, $\hat{x}_{k|k-1}$ and $P_{k|k-1}$ are given by

$$\hat{x}_{k|k-1} = \phi_0 + \phi_1 \hat{x}_{k-1}$$
$$P_{k|k-1} = \phi_1 P_{k-1} \phi_1' + Q_k$$

where $'$ denotes transposition, and $\hat{z}_{k|k-1}$ and $F_{k|k-1}$ are given by

$$\hat{z}_{k|k-1} = h_0 + h_1 \hat{x}_{k|k-1} \tag{29}$$

$$F_{k|k-1} = h_1 P_{k|k-1} h_1' + R. \tag{30}$$

In the update step, the estimate of the state vector is refined based on the difference between observed and predicted quantities, with $\hat{x}_k = \mathbb{E}_k[x_k]$ and P_k given by

$$\hat{x}_k = \hat{x}_{k|k-1} + K_k(z_k - \hat{z}_{k|k-1}) \tag{31}$$

$$P_k = P_{k|k-1} - K_k F_{k|k-1} K_k' \tag{32}$$

where the so-called Kalman gain $K_k = P_{k|k-1} h_1' F_{k|k-1}^{-1}$.

In our setting, the function $h(x_k)$ is nonlinear, and the Kalman filter has to be modified. Nonlinear state space models have traditionally been handled with the extended Kalman filter, which effectively linearizes the measure equation around the predicted state. In recent years, the UKF has emerged as a superior alternative. Rather than approximating the measurement equation, it uses the true nonlinear measurement equation and approximates the distribution of the state vector with a deterministically chosen set of sample points, called "sigma points" that capture the true mean and covariance of the state vector. When propagated through the nonlinear function $h(x_k)$, the sigma points capture the mean and covariance of the data accurately to the second order (third order for Gaussian states) for any nonlinearity.

Specifically, a set of $2L + 1$ sigma points and associated weights are selected according to the following scheme

$$\begin{aligned}
\hat{\chi}^0_{k|k-1} &= \hat{x}_{k|k-1}, & \omega^0 &= \frac{\kappa}{L+\kappa} \\
\hat{\chi}^i_{k|k-1} &= \hat{x}_{k|k-1} + \left(\sqrt{(L+\kappa)P_{k|k-1}}\right)_i, & \omega^i &= \frac{1}{2(L+\kappa)}, \quad i = 1, \ldots, L \\
\hat{\chi}^i_{k|k-1} &= \hat{x}_{k|k-1} - \left(\sqrt{(L+\kappa)P_{k|k-1}}\right)_i, & \omega^i &= \frac{1}{2(L+\kappa)}, \quad i = L+1, \ldots, 2L
\end{aligned}$$

where L is the dimension of $\hat{x}_{k|k-1}$, κ is a scaling parameter, ω^i is the weight associated with the ith sigma point, and $\left(\sqrt{(L+\kappa)P_{k|k-1}}\right)_i$ is the ith column of the matrix square root. Then, in the prediction step, (29) and (30) are replaced by

$$\hat{z}_{k|k-1} = \sum_{i=0}^{2L} \omega^i h(\hat{\chi}^i_{k|k-1})$$

$$F_{k|k-1} = \sum_{i=0}^{2L} \omega^i (h(\hat{\chi}^i_{k|k-1}) - \hat{z}_{k|k-1})(h(\hat{\chi}^i_{k|k-1}) - \hat{z}_{k|k-1})' + R.$$

The update step is still given by (31) and (32), but with the Kalman gain K_k computed as

$$K_k = \sum_{i=0}^{2L} \omega^i (\hat{\chi}^i_{k|k-1} - \hat{x}_{k|k-1})(h(\hat{\chi}^i_{k|k-1}) - \hat{z}_{k|k-1})' F^{-1}_{k|k-1}.$$

Finally, the log-likelihood function is given by

$$\sum_{k=1}^{n} -\frac{1}{2}\left[2\log(2\pi) + \log|F_{k|k-1}| + (z_k - \hat{z}_{k|k-1})' F^{-1}_{k|k-1}(z_k - \hat{z}_{k|k-1})\right]$$

where n is the sample size of daily log-returns. We maximize the log-likelihood with respect to the model parameters except ν, which is not well estimated. In sum, the procedure above jointly returns parameter estimates and the filtered trajectory of the latent time change $\hat{x}_k$ using solely index returns.

References

Aït-Sahalia, Y., J. Fan, and Y. Li. 2013. The Leverage Effect Puzzle: Disentangling Sources of Bias at High Frequency. *Journal of Financial Economics* 109: 224–249.

Aït-Sahalia, Y., M. Karaman, and L. Mancini. 2020. The Term Structure of Variance Swaps and Risk Premia. *Journal of Econometrics* 219: 204–230.

Aït-Sahalia, Y., and P. A. Mykland. 2003. The Effects of Random and Discrete Sampling When Estimating Continuous-Time Diffusions. *Econometrica* 71: 483–549.

Andersen, T. G., N. Fusari, and V. Todorov. 2015a. Parametric Inference and Dynamic State Recovery from Option Panels. *Econometrica* 83: 1081–1145.

Andersen, T. G., N. Fusari, and V. Todorov. 2015b. The Risk Premia Embedded in Index Options. *Journal of Financial Economics* 117: 558–584.

Andersen, T. G., N. Fusari, and V. Todorov. 2017. Short-Term Market Risks Implied by Weekly Options. *The Journal of Finance* 72: 1335–1386.

Bakshi, G., P. Carr, and L. Wu. 2008. Stochastic Risk Premiums, Stochastic Skewness in Currency Options, and Stochastic Discount Factors in International Economies. *Journal of Financial Economics* 87: 132–156.

Bakshi, G., and L. Wu. 2010. The Behavior of Risk and Market Prices of Risk over the Nasdaq Bubble Period. *Management Science* 56: 2251–2264.

Bardgett, C., E. Gourier, and M. Leippold. 2019. Inferring Volatility Dynamics and Risk Premia from the S&P500 and VIX Markets. *Journal of Financial Economics* 131: 593–618.

Barndorff-Nielsen, O. E. 1997. Processes of Normal Inverse Gaussian Type. *Finance and Stochastics* 2: 41–68.

Barndorff-Nielsen, O. E., and N. Shephard. 2001. Non-Gaussian Ornstein–Uhlenbeck-Based Models and Some of Their Uses in Financial Economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63: 167–241.

Barone-Adesi, G., R. F. Engle, and L. Mancini. 2008. A GARCH Option Pricing Model with Filtered Historical Simulation. *Review of Financial Studies* 21: 1223–1258.

Bates, D. S. 2006. Maximum Likelihood Estimation of Latent Affine Processes. *Review of Financial Studies* 19: 909–965.

Bates, D. S. 2012. US Stock Market Crash Risk, 1926–2010. *Journal of Financial Economics* 105: 229–259.

Bekaert, G., and G. Wu. 2000. Asymmetric Volatility and Risk in Equity Markets. *Review of Financial Studies* 13: 1–42.

Bertoin, J. 1996. *Lévy Processes*. Cambridge: Cambridge University Press.

Black, F. 1976. "Studies of Stock Price Volatility Changes." *Proceedings of the 1976 Meetings of the American Statistical Association*, Washington DC, pp. 171–181.

Bollerslev, T., J. Li, and Y. Xue. 2018. Volatility and Public News Announcements. *The Review of Economic Studies* 85: 2005–2041.

Broadie, M., M. Chernov, and M. Johannes. 2007. Model Specification and Risk Premia: Evidence from Futures Options. *The Journal of Finance* 62: 1453–1490.

Calvet, L., M. Fearnley, A. Fisher, and M. Leippold. 2015. What's beneath the Surface? Option Pricing with Multifrequency Latent States. *Journal of Econometrics* 187: 498–511.

Campbell, J. Y., and L. Hentschel. 1992. No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns. *Journal of Financial Economics* 31: 281–318.

Carr, P., H. Geman, D. B. Madan, and M. Yor. 2002. The Fine Structure of Asset Returns: An Empirical Investigation. *The Journal of Business* 75: 305–332.

Carr, P., H. Geman, D. B. Madan, and M. Yor. 2003. Stochastic Volatility for Lévy Processes. *Mathematical Finance* 13: 345–382.

Carr, P., and L. Wu. 2004. Time-Changed Lévy Processes and Option Pricing. *Journal of Financial Economics* 71: 113–141.

Cheridito, P., D. Filipović, and R. Kimmel. 2007. Market Price of Risk Specifications for Affine Models: Theory and Evidence. *Journal of Financial Economics* 83: 123–170.

Christie, A. A. 1982. The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects. *Journal of Financial Economics* 10: 407–432.

Christoffersen, P., K. Jacobs, and K. Mimouni. 2010. Volatility Dynamics for the S&P500: Evidence from Realized Volatility, Daily Returns, and Option Prices. *Review of Financial Studies* 23: 3141–3189.

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

Corradi, V., W. Distaso, and A. Mele. 2013. Macroeconomic Determinants of Stock Volatility and Volatility Premiums. *Journal of Monetary Economics* 60: 203–220.

Corsi, F., N. Fusari, and D. L. Vecchia. 2013. Realizing Smiles: Options Pricing with Realized Volatility. *Journal of Financial Economics* 107: 284–304.

Diebold, F. X., T. A. Gunther, and A. S. Tay. 1998. Evaluating Density Forecasts with Applications to Financial Risk Management. *International Economic Review* 39: 863–883.

Diebold, F. X., J. Hahn, and A. S. Tay. 1999. Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High-Frequency Returns on Foreign Exchange. *Review of Economics and Statistics* 81: 661–673.

Duffie, D., J. Pan, and K. Singleton. 2000. Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica* 68: 1343–1376.

Engle, R. F., and J. G. Rangel. 2008. The spline-GARCH Model for Low-Frequency Volatility and Its Global Macroeconomics Causes. *Review of Financial Studies* 21: 1187–1222.

Eraker, B. 2004. Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices. *The Journal of Finance* 59: 1367–1404.

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

Fallahgoul, H., and G. Loeper. 2021. Modeling Tail Risk with Tempered Stable Distributions: An Overview. *Annals of Operations Research* 299: 1253–1280.

Fallahgoul, H., and K. Nam. 2020. Correlated time-changed Lévy Processes. Technical report, Monash CQFIS working paper No. 2018–8.

Fang, F., and C. W. Oosterlee. 2008. A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions. *SIAM Journal on Scientific Computing* 31: 826–848.

Filipović, D., E. Gourier, and L. Mancini. 2016. Quadratic Variance Swap Models. *Journal of Financial Economics* 119: 44–68.

French, K. R., G. W. Schwert, and R. F. Stambaugh. 1987. Expected Stock Returns and Volatility. *Journal of Financial Economics* 19: 3–29.

Gallant, A. R., P. E. Rossi, and G. Tauchen. 1992. Stock Prices and Volume. *Review of Financial Studies* 5: 199–242.

Glosten, L. R., R. Jagannathan, and D. E. Runkle. 1993. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance* 48: 1779–1801.

Grasselli, M., and C. Tebaldi. 2008. Solvable Affine Term Structure Models. *Mathematical Finance* 18: 135–153.

Gruber, P., C. Tebaldi, and F. Trojani. 2021. The Price of the Smile and Variance Risk Premia. *Management Science*, forthcoming.

Huang, J. Z., and L. Wu. 2004. Specification Analysis of Option Pricing Models Based on Time-Changed Lévy Processes. *The Journal of Finance* 59: 1405–1439.

Israelov, R., and B. T. Kelly. 2017. Forecasting the Distribution of Option Returns. *SSRN Electronic Journal*. Available at SSRN: https://ssrn.com/abstract=3033242.

Leippold, M., and J. Stromberg. 2014. Time-Changed Lévy LIBOR Market Model: Pricing and Joint Estimation of the Cap Surface and Swaption Cube. *Journal of Financial Economics* 111: 224–250.

Madan, D. B., and E. Seneta. 1990. The Variance Gamma Model for Share Market Returns. *The Journal of Business* 63: 511–524.

Mehra, R. 2006. The Equity Premium Puzzle: A Review. *Foundations and Trends® in Finance* 2: 1–81.

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

Rachev, S. T., Y. S. Kim, M. L. Bianchi, and F. J. Fabozzi. 2011. *Financial Models with Levy Processes and Volatility Clustering*. USA: John Wiley & Sons.

Schoutens, W. 2003. *Lévy Processes in Finance: Pricing Financial Derivatives*. USA: John Wiley & Sons.

Vuong, Q. H. 1989. Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica* 57: 307–333.

Wu, L. 2006. Dampened Power Law: Reconciling the Tail Behavior of Financial Security Returns. *The Journal of Business* 79: 1445–1473.