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

H. Peter Boswijk
Amsterdam School of Economics
University of Amsterdam
and Tinbergen Institute

Roger J. A. Laeven
Amsterdam School of Economics
University of Amsterdam, EURANDOM
and CentER

Evgenii Vladimirov
Amsterdam School of Economics
University of Amsterdam
and Tinbergen Institute

August 21, 2022

Abstract

We develop a novel filtering and estimation procedure for parametric option pricing models driven by general affine jump-diffusions. Our procedure is based on the comparison between an option-implied, model-free representation of the conditional log-characteristic function and the model-implied conditional log-characteristic function which is functionally affine in the model's state vector. We formally derive a linear state space representation and establish the asymptotic properties of the corresponding measurement errors. The state space representation allows us to use a suitably extended Kalman filter technique for the latent state vector, and an associated quasi-maximum likelihood estimator of the parameters, which brings important computational advantages. We analyze the finite-sample behavior of our procedure in Monte Carlo simulations. The applicability of our procedure is illustrated in two case studies that analyze S&P 500 option prices and the impact of exogenous state variables capturing Covid-19 reproduction and economic policy uncertainty.

Keywords: Options; Characteristic Function; Affine Jump-Diffusion; State Space Representation.
JEL Classification: Primary: C13; C58; G13; Secondary: C32; G01.

*We are very grateful to Torben Andersen, Kris Jacobs, Siem Jan Koopman, Olivier Scaillet, George Tauchen, Viktor Todorov, Fabio Trojani, and conference and seminar participants at the 2021 SoFiE Financial Econometrics Summer School at Northwestern University, the 2022 Quantitative Finance and Financial Econometrics (QFFE) Conference at Aix-Marseille University, the 2022 Annual SoFiE Conference at the University of Cambridge, the 2022 Dynstoch meeting at the Institut Henri Poincaré in Paris, the University of Amsterdam, Kellogg School of Management at Northwestern University, the Center for Econometrics and Business Analytics at St. Petersburg State University, and the Tinbergen Institute for helpful comments and suggestions. `Julia` code to implement the estimation procedure developed in this paper is available from `https://github.com/evladimirov/OptionModels-cKF-ccf`. This research was funded in part by the Netherlands Organization for Scientific Research (NWO) under grant NWO-Vici 2019/2020 (Laeven). Email addresses: `H.P.Boswijk@uva.nl`, `R.J.A.Laeven@uva.nl`, and `E.Vladimirov@uva.nl`.

1 Introduction

Over the past decades, explosive growth in the trading of option contracts has attracted the attention of academics and practitioners to the development and estimation of increasingly sophisticated option pricing models. The building blocks of many continuous-time option pricing models are semimartingale stochastic processes that govern the dynamics of the underlying asset. These processes are often latent with stochastic diffusive volatility as the prototypical example, as in the classical Heston (1993) model. The literature also suggests the need to allow for a discontinuous jump component, both in the asset price dynamics and in its volatility process, potentially with a time-varying stochastic jump intensity.

An important econometric challenge lies in estimating the parameters of these continuous-time models and in filtering their unobserved and time-varying components, since option prices are highly nonlinear functions of the state vector. This stands in contrast to, for instance, term structure models, where bond yields can be represented as a linear function of the states, at least within the affine framework (see, e.g., Piazzesi, 2010 for a review of the affine term structure literature). To evaluate option prices as a function of the state vector, one typically needs to apply either Fourier-based methods or simulation-based approaches, in both cases at a substantial computational cost. This is one of the reasons why in much of the empirical research on option pricing, only a subset of the available option price data is used, such as at-the-money contracts or weekly (typically Wednesday) options data.

In this paper, we develop a new latent state filtering and parameter estimation procedure for option pricing models governed by general affine jump-diffusion processes. Our procedure leverages the linear relationship between the logarithm of an option-implied, model-free spanning formula for the conditional characteristic function of the underlying asset return on the one hand, and the state vector induced by parametric model specification on the other hand. Linearity of the measurement and state updating equations of the resulting state space representation, with coefficient and variance matrices that are (semi-)closed-form functions of the parameters, allows us to exploit Kalman filtering techniques. The proposed estimation procedure is theoretically justified and moreover fast and easy to implement, circumventing the typical computational burden when conducting inference on option pricing models.

Exploiting the option-spanning formula of Carr and Madan (2001) for European-style payoff functions, we replicate the risk-neutral conditional characteristic function (CCF) of the underlying log-asset price at the expiration date in a completely model-independent way. In other words, we imply information about the CCF from the option prices without imposing any parametric assumptions on the underlying asset price dynamics. A similar option-spanning approach for the CCF is used by Todorov (2019) to develop an option-based nonparametric spot volatility estimator. On the other hand, a large stream of literature is devoted to parametric option pricing models belonging to the general affine jump-diffusion (AJD) family; notable examples are Heston (1993), Duffie, Pan, and Singleton (2000), Pan (2002), and Bates (2006).[1]

[1]See also, e.g., Broadie, Chernov, and Johannes (2007), Aït-Sahalia, Cacho-Diaz, and Laeven (2015), and Andersen, Fusari, and Todorov (2017).

The defining property of the AJD class is the exponential-affine joint CCF, which is available in semi-closed form. By comparing the two option pricing representations—model-free and model-implied—we can obtain a linear relation between the logarithm of the option-implied CCF and the model-dependent CCF within the affine framework.

The state vector in AJD option pricing models typically contains both observable processes and latent factors. We address the filtering of such latent factors by developing, in discrete time, a linear state space representation for this model class. This allows us to leverage the Kalman filter to learn the unobserved intrinsic components of the model and estimate the model parameters using quasi-maximum likelihood (QML). A similar approach is often used in the affine term structure literature, where the yields themselves are affine functions of the state vector (see, e.g., de Jong, 2000, Driessen, 2005, Duffee, 1999). Besides the possibility to exploit Kalman filtering and QML estimation techniques, another advantage of our approach is that, once the model-free CCF has been obtained from the data, no further numerical option pricing methods, such as the FFT approach of Carr and Madan (1999) or simulation-based methods, are needed. Therefore, our method reduces computational costs considerably relative to many existing approaches in the option pricing literature. We note that, whereas the parametric CCF is used to price options in Fourier-based methods, here we use the CCF to directly learn about the latent factors and model parameters.

We illustrate the developed estimation procedure in Monte Carlo simulations based on several AJD specifications. We consider a one-factor AJD option pricing model, with the stochastic volatility and jump intensity both being affine functions of a single latent process, and a two-factor AJD model specification with an observable exogenous factor. We find good finite-sample performance in both cases.

Finally, we demonstrate our new filtering and estimation approach in an empirical application to S&P 500 index options. In particular, we filter and estimate the latent volatility and jump intensity from a stochastic volatility model with double-exponential jumps. We also investigate the impact of the Covid-19 propagation rate on the stock market within this model by embedding the reproduction number into the volatility and jump intensity dynamics. Our results show that while the reproduction numbers have a mild effect on total volatility, they contribute substantially to the likelihood of jumps. By contrast, when we consider an Economic Policy Uncertainty index as exogenous factor, the jump intensity process is not affected, but the exogenous factor contributes significantly to diffusive volatility.

Various estimation and filtering strategies for option pricing models have been developed in the literature. These include the (penalized) nonlinear least squares methods in, e.g., Bakshi, Cao, and Chen (1997), Broadie et al. (2007), Andersen, Fusari, and Todorov (2015a); the efficient method of moments of Gallant and Tauchen (1996) as applied in Chernov and Ghysels (2000) and Andersen, Benzoni, and Lund (2002); the implied-state methods initiated by Pan (2002) and further analyzed by Santa-Clara and Yan (2010); the Markov Chain Monte Carlo method in Eraker (2004) and Eraker, Johannes, and Polson (2003); and the particle filtering method, see Johannes, Polson, and Stroud (2009) and Bardgett, Gourier, and Leippold (2019). Most of these estimation methods use as inputs option prices or a monotonic transformation

thereof, such as implied volatilities. By contrast, we propose an estimation procedure based on the prices of spanning option portfolios that contain *all* probabilistic information.

Estimation strategies based on the transform space of conditional characteristic functions are not new in the literature. For instance, Carrasco and Florens (2000) develop a generalized method of moments (GMM) estimator with a continuum of moment conditions based on the CCF; see also Singleton (2001), Carrasco, Chernov, Florens, and Ghysels (2007). In applications to option prices, Boswijk, Laeven, and Lalu (2015) and Boswijk, Laeven, Lalu, and Vladimirov (2021) propose to imply the latent state vector from a panel of options and then estimate the model via GMM with a continuum of moments. Bates (2006) develops maximum likelihood estimation and filtering using CCFs. In particular, he proposes a recursive likelihood evaluation by updating the CCF of a latent variable conditional upon observed data. However, unlike our approach, these methods suffer from a 'curse of dimensionality' since they require numerical integration over the dimension of the state vector.

Our work is also related to Feunou and Okou (2018), who exploit the linear relation between the four risk-neutral cumulants of the log-asset price and latent factors. They obtain these cumulants via a portfolio of options and employ the Kalman filter to estimate the latent factors. The main difference with our approach is that we use the CCF instead of the first four moments. The CCF contains much richer information, leading to more efficient inference. Another difference is in dimension reduction: Feunou and Okou (2018) use a two-step principal components analysis (PCA) to reduce the dimension of the risk-neutral cumulants observed at different maturities. Instead, we use a so-called collapsed Kalman filter approach, originally developed by Jungbacker and Koopman (2015), which does not suffer from information losses relative to the full-dimensional setting.

The paper is organized as follows. Section 2 provides the theoretical framework for aligning the option-implied and model-implied CCFs of the underlying asset price. In Section 3, we develop the state space representation, and state the main result about the orders of measurement and approximation errors under an in-fill asymptotic scheme. This allows us to develop the filtering approach and corresponding estimation procedure. Section 4 presents the Monte Carlo simulation results. We describe the data in Section 5 and demonstrate the empirical applications in Section 6. Conclusions are in Section 7. There are five appendices providing details on (*i*) the characteristic function replication, (*ii*) the proof of Proposition 1, (*iii*) the computation of conditional moments, (*iv*) the extrapolation scheme for option prices, and (*v*) additional simulation and empirical results.

2 Theoretical Framework

In this section we provide the theoretical framework for our approach. We start with extracting information about the conditional characteristic function from option prices allowing for general underlying dynamics. Next, we consider the conditional characteristic function within the generalized AJD class, which is exponentially affine in the model's state variables. Finally, we discuss how to align the two characteristic functions—option-implied model-free and

AJD model-implied—in order to make inferences about the model parameters and latent state variables.

2.1 Option-implied CCF

Throughout the paper, we fix a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. On this probability space, we consider the dynamics of an arbitrage-free financial market. The no-arbitrage assumption guarantees the existence of a risk-neutral probability measure $\mathbb{Q}$, locally equivalent to $\mathbb{P}$. Since we are interested in exploiting information from options, we formulate the model dynamics under $\mathbb{Q}$.

Let us denote by F_t the futures price at time t for a stock or an index futures contract with some fixed maturity. The absence of arbitrage implies that the futures price process is a semimartingale. In this subsection, we assume the following general dynamics for F_t under $\mathbb{Q}$:

$$\frac{\mathrm{d}F_t}{F_t} = v_t \mathrm{d}W_t + \int_{\mathbb{R}} x\tilde{\mu}(\mathrm{d}t, \mathrm{d}x), \tag{1}$$

where v_t is an adapted, locally bounded, but otherwise unspecified stochastic volatility process; W_t is a standard Brownian motion under $\mathbb{Q}$; μ is a counting random measure with compensator $\nu_t(\mathrm{d}x)\mathrm{d}t$ such that $\tilde{\mu}(\mathrm{d}t, \mathrm{d}x) := \mu(\mathrm{d}t, \mathrm{d}x) - \nu_t(\mathrm{d}x)\mathrm{d}t$ is the associated martingale measure and $\int (x^2 \wedge 1)\nu_t(\mathrm{d}x) < \infty$.

We further denote out-of-the-money (OTM) European-style option prices at time t with time-to-maturity $\tau > 0$ and strike price K by $O_t(\tau, K)$. Under the no-arbitrage assumption, the option prices equal the risk-neutral conditional expectations of the corresponding discounted payoff functions:

$$O_t(\tau, K) = \begin{cases} \mathbb{E}^{\mathbb{Q}}[e^{-r\tau}(F_{t+\tau} - K)^+ | \mathcal{F}_t], & \text{if } K > F_t, \\ \mathbb{E}^{\mathbb{Q}}[e^{-r\tau}(K - F_{t+\tau})^+ | \mathcal{F}_t], & \text{if } K \leq F_t. \end{cases}$$

The OTM price $O_t(\tau, K)$ is a call option price if $K > F_t$ and a put option price if $K \leq F_t$. For simplicity we assume a constant interest rate r.

Following Carr and Madan (2001), any twice continuously differentiable European-style payoff function $g(F_{t+\tau})$, with first and second derivatives $g^{(1)}$ and $g^{(2)}$, can be spanned via a position in risk-free bonds, futures (or stocks) and options with a continuum of strikes, as follows:

$$\begin{aligned} g(F_{t+\tau}) = g(x) + g^{(1)}(x)(F_{t+\tau} - x) \\ + \int_0^x g^{(2)}(K)(K - F_{t+\tau})^+ \mathrm{d}K + \int_x^\infty g^{(2)}(K)(F_{t+\tau} - K)^+ \mathrm{d}K. \end{aligned}$$

Here, $x \in \mathbb{R}^+$, the first and second terms on the right-hand side correspond to risk-free bonds and futures positions, and the third and fourth terms correspond to OTM options. Taking conditional expectations under the risk-neutral measure for $x = F_t$, we find that the price of a contingent claim with payoff function $g(F_{t+\tau})$ discounted at time t can be expressed as a weighted portfolio of a risk-free bond and OTM options:

$$\mathbb{E}^{\mathbb{Q}}[e^{-r\tau} g(F_{t+\tau}) | \mathcal{F}_t] = e^{-r\tau} g(F_t) + \int_0^\infty g^{(2)}(K) O_t(\tau, K) \mathrm{d}K. \tag{2}$$

This general spanning result lies behind the construction of one of most popular 'fear' indices—the VIX index, when $g(F_{t+\tau}) = \log(F_{t+\tau}/F_t)$. Some other applications of the spanning formula (2) include, for instance, the calculation of the option-implied skewness and kurtosis (Bakshi, Kapadia, & Madan, 2003) and the corridor implied volatility (Andersen & Bondarenko, 2007).

Applying this result to the complex-valued payoff function $g(x) = e^{\mathrm{i}u\log(x/F_t)}$ yields that the discounted conditional characteristic function (CCF) of log returns can be spanned as

$$\begin{aligned}\phi_t(u,\tau) &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[e^{\mathrm{i}u\log(F_{t+\tau}/F_t)}|\mathcal{F}_t] \\ &= e^{-r\tau} - (u^2+\mathrm{i}u)\int_0^\infty \frac{1}{K^2}e^{\mathrm{i}u(\log K - \log F_t)}\cdot O_t(\tau,K)\mathrm{d}K \\ &= e^{-r\tau} - (u^2+\mathrm{i}u)\frac{1}{F_t}\int_{\mathbb{R}} e^{(\mathrm{i}u-1)m}\cdot O_t(\tau,m)\mathrm{d}m, \end{aligned} \tag{3}$$

where $m = \log(K/F_t)$ is the log-moneyness of an option with the strike price K.[2]

It is important to emphasize that the spanning of the CCF in equation (3) is exact and is completely model independent akin to the VIX. Therefore, the conditional characteristic function of log returns over a particular horizon τ can be replicated in a model-free way given a single cross-sectional slice of liquid option prices with all strikes (and the same maturity τ). A similar approach of CCF spanning is taken by Todorov (2019) to nonparametrically estimate spot volatilities from option prices (taking the limit as $\tau \downarrow 0$).

The expression in (3) cannot be calculated analytically as we do not observe option prices for a continuum of strikes. Nevertheless, as we detail in Section 3.2, the expression for $\phi_t(u,\tau)$ in (3) is easy to approximate using a limited number of observable option prices. In our empirical applications, we further employ an interpolation-extrapolation scheme to improve the reliability of the approximation and take into account the observation errors in option prices. We denote by $\widehat{\phi}_t(u,\tau)$ the computationally feasible counterpart of the option-implied CCF.

2.2 Affine jump-diffusion CCF

The CCF of the log returns of the underlying asset is often considered under some parametric assumptions on the return dynamics. Hence, a model-implied CCF depends on the model parameters, that we generally do not know, and potentially on the dynamics of other latent processes, that affect the distribution of returns. Therefore, by aligning the model-free and parameter-dependent CCFs, we can learn model parameters and unobservable state dynamics.

We restrict our attention to a broad class of AJD models defined in Duffie et al. (2000). The main attraction of the AJD class is that the Laplace transform has a semi-closed-form expression and is of the exponential-affine form. Suppose X_t is a Markov process representing an d_X-dimensional state vector in $D \subset \mathbb{R}^{d_X}$ with the first component being the log price of an asset. We assume that under the physical and risk-neutral probability measures, the state

[2]With a slight abuse of notation, we use the same symbol O_t for the option value as function of (τ, K) or (τ, m).

vector X_t solves the following stochastic differential equation:

$$\mathrm{d}X_t = \mu(X_t;\theta)\mathrm{d}t + \sigma(X_t;\theta)\mathrm{d}W_t + \sum_{i=1}^{d_J} J_{i,t}\mathrm{d}N_{i,t}, \tag{4}$$

where W_t is a standard Brownian motion in $\mathbb{R}^{d_W}$; μ: $D \to \mathbb{R}^{d_X}$ and σ: $D \to \mathbb{R}^{d_X \times d_W}$ are the drift and diffusion function; $J_{i,t}\mathrm{d}N_{i,t}$ is a pure jump process with intensity $\{\lambda^i(X_t;\theta)\colon t \geq 0\}$ with λ^i: $D \to \mathbb{R}^+$ and jump size $J_{i,t}$ with conditional distribution ν^i on $\mathbb{R}^{d_X}$ for $i = 1, \ldots, d_J$; and θ is a vector of unknown parameters that governs the model for X_t. The specification (4) can be extended, e.g., to include infinite activity jumps and a time-dependent structure. See Duffie et al. (2000) and Duffie, Filipović, and Schachermayer (2003) for more details on the AJD class formulation.

Following Duffie et al. (2000), the drift $\mu(x)$, diffusive variance $\sigma(x)\sigma(x)'$ and jump intensities $\lambda^i(x)$ are assumed to be affine on D:

$$\begin{aligned}
\mu(x) &= K_0 + K_1 x, && K_0 \in \mathbb{R}^{d_X},\ K_1 \in \mathbb{R}^{d_X \times d_X},\\
\sigma(x)\sigma(x)' &= H_0 + \textstyle\sum_{j=1}^{d_X} x_j H_1^{(j)}, && H_0 \in \mathbb{R}^{d_X \times d_X},\ H_1^{(j)} \in \mathbb{R}^{d_X \times d_X},\ j = 1, \ldots, d_X,\\
\lambda^i(x) &= l_{i,0} + l_{i,1} \cdot x, && l_{i,0} \in \mathbb{R},\ l_{i,1} \in \mathbb{R}^{d_X},\ i = 1, \ldots, d_J,
\end{aligned}$$

where x_j is the j-th element of a vector x and $H_1^{(j)}$ for j=1, …, d_X form a $d_X \times d_X \times d_X$ tensor H_1 by stacking matrices along a new dimension. The joint regularity conditions on $(D, \mu, \sigma, \lambda, \nu)$, that guarantee a unique solution to the SDE (4), are discussed in Duffie and Kan (1996) and Dai and Singleton (2000). These joint conditions put constraints on the parameter vector θ. Therefore, we consider a model from the AJD class indexed by θ in a parameter space Θ containing such admissible parameter values, on which there is a unique solution to (4) that remains in D. For instance, in the case of the stochastic volatility component, the admissible parameter values in Θ ensure that the volatility process remains nonnegative, by satisfying the Feller condition. See also the discussion of the admissibility problem in Singleton (2006, Chapter 5).

Duffie et al. (2000) show that the affine dependence of the functions $\mu(x)$, $\sigma(x)\sigma(x)'$ and $\lambda(x)$ implies an exponential-affine form of the CCF of the state vector X_t. Specifically, the discounted joint CCF of $X_{t+\tau}$ conditional on $\mathcal{F}_t$ with $\tau > 0$ is given by

$$\psi_X(\mathbf{u}, \tau) = e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[e^{\mathrm{i}\mathbf{u}\cdot X_{t+\tau}}|\mathcal{F}_t] = e^{\alpha(\mathbf{u},\tau;\theta)+\beta(\mathbf{u},\tau;\theta)\cdot X_t}, \tag{5}$$

where $\mathbf{u} \in \mathbb{R}^{d_X}$ is an argument vector and $\alpha(\mathbf{u}, \tau; \theta)$ and $\beta(\mathbf{u}, \tau; \theta)$ are solutions to the following complex-valued system of ordinary differential equations (ODEs) in time:

$$\begin{cases}
\dot{\beta}(\mathbf{u}, s) = K_1'\beta(\mathbf{u}, s) + \frac{1}{2}\beta(\mathbf{u}, s)'H_1\beta(\mathbf{u}, s) + \sum_{i=1}^{d_J} l_1^i(\chi^i(\beta(\mathbf{u}, s)) - 1),\\
\dot{\alpha}(\mathbf{u}, s) = K_0'\beta(\mathbf{u}, s) + \frac{1}{2}\beta(\mathbf{u}, s)'H_0\beta(\mathbf{u}, s) + \sum_{i=1}^{d_J} l_0^i(\chi^i(\beta(\mathbf{u}, s)) - 1) - r,
\end{cases} \tag{6}$$

with initial conditions $\beta(\mathbf{u}, 0) = \mathrm{i}\mathbf{u}$ and $\alpha(\mathbf{u}, 0) = 0$. Here, $\chi^i(c) = \int_{\mathbb{R}^n} \exp(c \cdot z)\mathrm{d}\nu^i(z),\ c \in \mathbb{C}^{d_X}$ are jump transforms, that determine the conditional jump-size distributions. The ODE for β is known as the generalized Riccati equation, while the solution for the second ODE can be obtained by simply integrating the right-hand side expression over time.

The affine dependence of the characteristic exponent $\alpha(\mathbf{u}, \tau; \theta) + \beta(\mathbf{u}, \tau; \theta) \cdot X_t$ on the current state X_t is the defining property of the AJD class under some regularity conditions (see Duffie et al., 2003). In other words, the AJD can be defined as a class in which characteristic exponents of $X_{t+\tau}$ given X_t are affine functions of X_t. In fact, this is a key property in our estimation procedure. While it is also possible to obtain the CCF for some non-affine models, the exponential-affine form allows us to use the linear Kalman filtering technique in the estimation procedure. This is the main motivation why we restrict our attention to the parametric models of the AJD class.[3]

Unlike the option-implied CCF (3), the CCF in (5) is fully parametric, that is, it requires a parametric AJD model dynamics of the state vector X_t. Although the AJD class is more restrictive than the general dynamics of F_t in (1), it includes a myriad of popular option-pricing models such as in Heston (1993), Duffie et al. (2000), Pan (2002), Bates (2006), Broadie et al. (2007), Boswijk et al. (2015), and Andersen et al. (2017) among many others.

The state process X_t might include both observed and unobserved state variables that affect the dynamics of the log futures price $\log F_t$. In our empirical application, we consider the presence of both. Therefore, it is convenient to partition the state vector as $X_t' = (w_t', x_t')$, where w_t represents the observable component and x_t includes $d < d_X$ latent state variables. Then, the dynamics X_t, given by equation (4) can be rewritten as

$$\mathrm{d}w_t = \mu^w(w_t, x_t)\mathrm{d}t + \sigma^w(w_t, x_t)\mathrm{d}W_t + \sum_{i=1}^{d_J} J_{i,t}^w \mathrm{d}N_{i,t}, \tag{7}$$

$$\mathrm{d}x_t = \mu^x(w_t, x_t)\mathrm{d}t + \sigma^x(w_t, x_t)\mathrm{d}W_t + \sum_{i=1}^{d_J} J_{i,t}^x \mathrm{d}N_{i,t}, \tag{8}$$

where $\mu^w \colon D \to \mathbb{R}^{d_X - d}$, $\mu^x \colon D \to \mathbb{R}^d$, $\sigma^w \colon D \to \mathbb{R}^{(d_X - d) \times d_J}$, $\sigma^x \colon D \to \mathbb{R}^{d \times d_J}$ and $J_{i,t}^w$ and $J_{i,t}^x$ are marginal jump sizes of $J_{i,t}$ associated with w_t and x_t, respectively. In the simplest case, the observable component includes only the log futures prices, that is, $w_t = \log F_t$. In more general settings, the stochastic volatility is often the main latent driver of the log returns dynamics, as e.g., in Heston (1993).

2.3 Marrying the two CCFs

Given the two CCFs (3) and (5), we can now align them to conduct inferences about the parametric model fit and unobservable state variables. For that purpose, first note that the CCF in (5) is joint for the state vector X_t. Recall that we assume, without loss of generality, that the first component of the state vector X_t is the log futures price. Therefore, we can easily obtain its marginal CCF by plugging in an argument vector of the form $\mathbf{u_1} := (u, 0, \ldots, 0)' \in \mathbb{R}^{d_X}$ with $u \in \mathbb{R}$. To obtain the marginal CCF for the log returns, we further subtract the term $\mathrm{i}u \log F_t$ in the exponent. That is, the marginal CCF of log returns under the AJD specification is as

[3] The considered AJD class can be broadened further to the linear-quadratic jump-diffusion class by augmenting the state vector (see Cheng and Scaillet (2007) for more details).

follows:

$$\phi_t(u,\tau) = \psi_X(\mathbf{u_1},\tau)e^{-\mathrm{i}u\log F_t} = e^{\alpha(\mathbf{u_1},\tau;\theta)+\tilde{\beta}(\mathbf{u_1},\tau;\theta)\cdot X_t}, \tag{9}$$

where $\tilde{\beta}(\mathbf{u_1},\tau;\theta) := \beta(\mathbf{u_1},\tau;\theta) - \mathrm{i}\mathbf{u_1}$; i.e., $\tilde{\beta}(\mathbf{u_1},\tau;\theta)$ has only the first component different from $\beta(\mathbf{u_1},\tau;\theta)$ since we are interested in the CCF of log returns rather than log prices.

Note that the log of the (joint) CCF (also known as cumulant-generating function) is linear in the state vector X_t. Therefore, under the correctly specified AJD model we obtain a simple linear relation between the log of the option-spanned CCF[4] of log returns and the model's state vector:

$$\log\phi_t(u,\tau) = \alpha(\mathbf{u_1},\tau;\theta) + \tilde{\beta}(\mathbf{u_1},\tau;\theta)\cdot X_t. \tag{10}$$

Replacing the cumulant generating function on the left-hand side with its computationally feasible counterpart $\widehat{\phi}_t(u,\tau)$ (defined explicitly in Section 3.1), we obtain the following equation, which will play a central role in our estimation procedure:

$$\log\widehat{\phi}_t(u,\tau) = \alpha(\mathbf{u_1},\tau;\theta) + \tilde{\beta}(\mathbf{u_1},\tau;\theta)\cdot X_t + \xi_t(u,\tau), \qquad u\in\mathbb{R}, \tag{11}$$

where $\xi_t(u,\tau)$ is the measurement error, which is related to the observation, discretization and truncation errors in the option-spanning CCF calculations. We elaborate on the relation between the computable counterpart of the CCF and the source of the measurement errors in the next section.

Equation (11) is the key relation in our analysis and a few remarks shall be made here regarding it. First of all, (11) is essentially a *functional* linear model since this equation holds for any argument variable of the CCF, $u\in\mathbb{R}$. Furthermore, the functions $\alpha(\mathbf{u_1},\tau;\theta)$ and $\tilde{\beta}(\mathbf{u_1},\tau;\theta)$ are parameter-dependent and solutions to the system of Riccati ODEs (6). Therefore, if the state vector X_t is observable, then the model parameters can be obtained by solving a continuum version of a non-linear least-squares problem.

In the case when the state vector is latent, (11) represents a linear factor model with a continuum number of linear relations. The factors are the state components of the AJD model. Therefore, we could apply, for instance, a (functional) principal component analysis to learn the unobserved factors. In this paper, we utilize the classical Kalman filter to estimate both model parameters and latent factors.

In other words, (11) reveals that AJD models, often formulated in continuous time, may be estimated using well-studied approaches from the rich literature on linear factor and state space models. In a similar way, bond yields are linear functions of the state vector in affine term structure models (see Piazzesi (2010) for a review of this class of models). For instance, Duffee (1999), de Jong (2000), Driessen (2005) use the Kalman filter in their estimation of affine term structure models.

[4] Although the logarithm of a complex number is a multivalued function, here, the ambiguity is resolved given the fact that $\phi(0) = 1$ and the CCF is a continuous function. In fact, in practice we ensure that the logarithm of the CCF does not have 'jumps' by taking the logarithm sequentially with respect to u, starting from the origin.

Furthermore, another advantage of this approach is that it does not require evaluating option prices given a certain parametric model. This estimation approach is computationally more appealing than many alternative approaches, which often involve the Carr-Madan FFT pricer (Carr & Madan, 1999) or the COS method (Fang & Oosterlee, 2008) to price options. This also implies that the usage of the characteristic function is different: with the FFT or COS method one needs a model-dependent CCF only to evaluate option prices, while here we use the CCF to directly learn the latent factors and model parameters.

Finally, given the partition of the state vector into observable and unobservable components, the linear relation between the option-implied and model-implied CCFs can be rewritten as

$$\log \widehat{\phi}_t(u,\tau) = \alpha(\mathbf{u_1},\tau;\theta) + \beta^w(\mathbf{u_1},\tau;\theta)\cdot w_t + \beta^x(\mathbf{u_1},\tau;\theta)\cdot x_t + \xi_t(u,\tau), \qquad u\in\mathbb{R}, \tag{12}$$

where $\beta^w(\mathbf{u_1},\tau;\theta)\in\mathbb{C}^{d_X-d}$ and $\beta^x(\mathbf{u_1},\tau;\theta)\in\mathbb{C}^d$ are such that $\tilde{\beta}' = (\beta^{w'},\beta^{x'})$ is the solution to the ODE system (6). Representation (12) serves as a basis for an observation (or measurement) equation in our estimation procedure.

3 Estimation Procedure

In this section, we develop our filtering approach and corresponding estimation procedure for the general class of AJD models under consideration. First, we provide the state space representation for the defined class of models. Then, we detail our estimation strategy using the collapsed Kalman filter.

3.1 State space representation

As discussed in the previous section, we restrict our attention to the parametric models of the AJD class due to their exponential-affine form of the characteristic function. This form will allow us to exploit the linear Kalman filter in the estimation procedure. In the following, we summarize the assumptions we impose on the parametric model:

Assumption 1 *(i) The stochastic process X_t is Markov and affine, with finite second moments under both the physical and risk-neutral probability measures $\mathbb{P}$ and $\mathbb{Q}$. In particular, X_t is the unique solution to the SDE* (4) *and its characteristic function is of the exponential-affine form* (5)*;*

(ii) The true parameter vector θ_0 lies in the interior of a compact parameter space Θ containing admissible parameter values.

Assumption 1 jointly guarantees the existence of a unique solution to the SDE (4) within the AJD class. As discussed in Section 2.2, admissible values $\theta\in\Theta$ reflect the regularity conditions imposed on the model such that there is a unique solution to (4), with, e.g., non-negative volatilities and jump intensities. Such admissibility conditions will need to be checked in a case-by-case model analysis. Assumption $1(i)$ also presumes the technical conditions required to represent the AJD process, defined via the affine dependence of its drift, diffusive variance

and jump intensities on the state vector, via the exponential-affine characteristic function. For a detailed analysis of the AJD theory we refer to Duffie et al. (2000) and Duffie et al. (2003).

In our estimation procedure we discretize the continuous-time model along two dimensions: with respect to time and with respect to the argument of the CCF. The former naturally follows from the discrete observations of financial data, which we denote by the integer indices $t = 1, \ldots, T$. The latter allows us to rely on the existing literature about filtering techniques. For that, let us denote the collection of the discretely sampled argument as a set $\mathcal{U} \subseteq \mathbb{R}$ with cardinality $q \in \mathbb{N}$. We further consider options with $k \in \mathbb{N}$ different maturities τ and $n \in \mathbb{N}$ different log-moneyness values m on each day.

Since the input of our estimation procedure is a portfolio of option prices, we need to take into account the measurement errors in these option prices. For that purpose, we assume an observation error scheme on the option prices that constitute the portfolios. The measurement errors will be defined on the common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, but in what follows, the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is generated by the state process $\{X_t\}_{t\geq 0}$ only. Note that the theoretical option prices $O_t(\tau, m)$ are $\mathcal{F}_t$-measurable, and hence the same applies to functionals of the option prices such as the (theoretical) Black-Scholes implied volatility (BSIV) and vega.

Assumption 2 *Option prices are observed with an additive error term:*

$$\widehat{O}_t(\tau_i, m_j) = O_t(\tau_i, m_j) + \zeta_t(\tau_i, m_j), \qquad t = 1, \ldots, T, \quad i = 1, \ldots, k, \quad j = 1, \ldots, n, \tag{13}$$

where the observation errors $\zeta_t(\tau, m)$ are such that:

(i) $\zeta_t(\tau, m)$ *are $\mathcal{F}_t$-conditionally independent along tenors τ, moneyness m and time t;*

(ii) $\mathbb{E}[\zeta_t(\tau, m)|\mathcal{F}_t] = 0$*;*

(iii) $\mathbb{E}[\zeta_t(\tau, m)^2|\mathcal{F}_t] = \sigma_t^2(\tau, m) < \infty$ *with* $\sigma_t(\tau, m) := \sigma_\varkappa \kappa_t(\tau, m)\nu_t(\tau, m)$*, where* $\sigma_\varkappa \in \mathbb{R}^+$*, $\kappa_t(\tau, m)$ is the Black-Scholes implied volatility, and $\nu_t(\tau, m)$ is the Black-Scholes vega.*

The additive error assumption is commonly imposed in the option pricing literature. For instance, Andersen et al. (2015a) and Todorov (2019) use additive error assumptions for option prices quoted in terms of BSIV and dollar amount, respectively. Additive observation errors are also often implicitly assumed when calibrating an option pricing model to market-observed prices, since the calibration is often performed using non-linear least squares as in, e.g., Broadie et al. (2007).

Assumption 2(i) excludes dependence of the observation errors across strikes and is also often imposed in the literature (see, for instance, Christoffersen, Jacobs, & Mimouni, 2010, Andersen et al., 2015a and Todorov, 2019). This assumption can be relaxed by introducing a spatial dependence as in Andersen, Fusari, Todorov, and Varneskov (2021). This would, however, result in more complex expressions for the covariance terms in the measurement errors, that we derive below. Furthermore, Andersen et al. (2021) find evidence of limited dependence in the observation errors for S&P 500 index options. They also show that this dependence declined sharply for short-dated options in recent years due to improved liquidity. Since in our empirical

application we consider S&P 500 index options with short tenors focusing on the last three years, the independence assumption will play a secondary role for the estimation procedure.

The conditional mean zero Assumption 2(*ii*) is crucial for our main result. Assumption 2(*iii*) asserts the standard deviation of the observation errors to be proportional to the product of the option's BSIV and vega. The motivation for this structure is as follows. Let $\widehat{\kappa}(m_j)$ and $\kappa(m_j)$ denote the error-distorted and true BSIV of an option, and assume that the relative volatility errors $\varkappa_j = (\widehat{\kappa}(m_j) - \kappa(m_j))/\kappa(m_j)$ are homoskedastic across the strikes, such that $\mathbb{E}[\varkappa_j^2|\mathcal{F}_t] = \sigma_\varkappa^2$. A Taylor-series expansion of the Black-Scholes pricing function $O^{BS}(\widehat{\kappa}(m_j), m_j)$ around $\kappa(m_j)$ then gives $\widehat{O}(m_j) = O^{BS}(\hat{\kappa}(m_j), m_j) \approx O(m_j) + \nu(m_j)\kappa(m_j)\varkappa_j$, with $\nu(m_j) = \partial O^{BS}(\kappa(m_j), m_j)/\partial\kappa(m_j)$, the theoretical Black-Scholes vega. Homoskedastic errors in relative implied volatilities are also assumed by Christoffersen, Jacobs, and Ornthanalai (2012) and Du and Luo (2019) in their MLE based on the particle filter and the unscented Kalman filter, respectively.

Finally, to assess the CCF approximation error sizes, we impose the following assumption on the existence of moments for the underlying asset and on the log-moneyness grid that allows nonequidistant sampling:

Assumption 3 *(i) The underlying process and its reciprocal process have finite second moments under the risk-neutral measure:* $\mathbb{E}^{\mathbb{Q}}[F_{t+\tau}^2|\mathcal{F}_t] < \infty$ *and* $\mathbb{E}^{\mathbb{Q}}[F_{t+\tau}^{-2}|\mathcal{F}_t] < \infty$ *with* $\tau > 0$*;*

(ii) For the log-moneyness grid $\underline{m} = m_1 < \ldots < m_n = \overline{m}$*, there exists a deterministic sequence* Δm *such that* $\Delta m \to 0$ *as* $n \to \infty$ *and*

$$\eta\Delta m \leq \inf_{j=2,\ldots,n} \Delta m_j \leq \sup_{j=2,\ldots,n} \Delta m_j \leq \Delta m,$$

where $\Delta m_j := m_j - m_{j-1}$ *and* $\eta \in (0, 1]$ *is some constant.*

Using $n > 1$ observable option prices with time-to-maturity $\tau > 0$ and log-moneyness values $\{m_j\}_{j=1}^n$, we may approximate the CCF $\phi_t(u, \tau)$ given in (3) by replacing the theoretical option prices by their observed versions, and the integral by a Riemann sum:

$$\widehat{\phi}_t(u, \tau) = e^{-r\tau} - u_t \sum_{j=2}^{n} e^{(\mathrm{i}u-1)m_j} \cdot \widehat{O}_t(\tau, m_j)\Delta m_j, \tag{14}$$

where we use the notation $u_t := (u^2 + \mathrm{i}u)/F_t$, and where $\widehat{O}_t(\tau, m_j)$ satisfies Assumption 2. The deviation of the option-spanned CCF from its theoretical counterpart, $\zeta_t^{\phi}(u, \tau) := \widehat{\phi}_t(u, \tau) - \phi_t(u, \tau)$ stems from observation, truncation and discretization errors (where truncation refers to the fact that the integration interval $[\underline{m}, \overline{m}]$ does not cover the entire real line). The discretization and truncation errors also arise in VIX calculations and depend on the availability of option prices. However, they can be efficiently reduced by using an interpolation-extrapolation scheme (see, e.g., Jiang & Tian, 2005, 2007 and Chang, Christoffersen, Jacobs, & Vainberg, 2012), and taking the limit as $n \to \infty$. Appendix C.2 illustrates the impact of different measurement errors on the CCF approximation.

From the preceding analysis, the functional measurement equation (12) is obtained from the following log-linearization:

$$\xi_t(u,\tau) := \log \widehat{\phi}_t(u,\tau) - \log \phi_t(u,\tau) = \log\left(1 + \frac{\zeta_t^{\phi}(u,\tau)}{\phi_t(u,\tau)}\right) = \xi_t^{(1)}(u,\tau) + r_t(u,\tau), \tag{15}$$

where the observation errors $\xi_t^{(1)}(u,\tau)$ are defined by $\zeta_t^{(1)}(u,\tau)/\phi_t(u,\tau)$, with

$$\zeta_t^{(1)}(u,\tau) = -u_t \sum_{j=2}^{n} e^{(\mathrm{i}u-1)m_j} \cdot \zeta_t(\tau, m_j)\Delta m_j,$$

and where $r_t(u,\tau)$ is a remainder term that collects the log-linearized truncation and discretization errors and the higher-order terms from a Taylor-series expansion.

To formulate the main result, we turn the complex-valued functional measurement equation (12) into a real vector measurement equation, as usual in state space model formulations. First, we stack the log CCF and the corresponding measurement errors along q values $u_1, \ldots, u_q$ for the CCF argument $u \in \mathcal{U}$, for a fixed expiration period τ_i:

$$\log \widehat{\phi}_{t,i} := \begin{pmatrix} \log \widehat{\phi}_t(u_1,\tau_i) \\ \log \widehat{\phi}_t(u_2,\tau_i) \\ \vdots \\ \log \widehat{\phi}_t(u_q,\tau_i) \end{pmatrix}, \quad r_{t,i} := \begin{pmatrix} r_t(u_1,\tau_i) \\ r_t(u_2,\tau_i) \\ \vdots \\ r_t(u_q,\tau_i) \end{pmatrix}, \quad \xi_{t,i}^{(1)} := \begin{pmatrix} \xi_t^{(1)}(u_1,\tau_i) \\ \xi_t^{(1)}(u_2,\tau_i) \\ \vdots \\ \xi_t^{(1)}(u_q,\tau_i) \end{pmatrix}.$$

In a similar way, we denote by $a_{t,i}$, $b_{t,i}^w$ and $b_{t,i}^x$ the stacked outputs[5] of the functions $\alpha(u,\tau_i)$, $\beta^w(u,\tau_i)$ and $\beta^x(u,\tau_i)$, respectively. Next, to tackle the complex-valued observation equation (12), we stack the real and imaginary parts, as well as k maturities:

$$\underbrace{\begin{pmatrix} \Re(\log \widehat{\phi}_{t,1}) \\ \Im(\log \widehat{\phi}_{t,1}) \\ \vdots \\ \Re(\log \widehat{\phi}_{t,k}) \\ \Im(\log \widehat{\phi}_{t,k}) \end{pmatrix}}_{=:y_t \in \mathbb{R}^p} = \underbrace{\begin{pmatrix} \Re(a_{t,1}) \\ \Im(a_{t,1}) \\ \vdots \\ \Re(a_{t,k}) \\ \Im(a_{t,k}) \end{pmatrix}}_{=:\tilde{d}_t} + \underbrace{\begin{pmatrix} \Re(b_{t,1}^w) \\ \Im(b_{t,1}^w) \\ \vdots \\ \Re(b_{t,k}^w) \\ \Im(b_{t,k}^w) \end{pmatrix}}_{=:W_t} w_t + \underbrace{\begin{pmatrix} \Re(b_{t,1}^x) \\ \Im(b_{t,1}^x) \\ \vdots \\ \Re(b_{t,k}^x) \\ \Im(b_{t,k}^x) \end{pmatrix}}_{=:Z_t} x_t + \underbrace{\begin{pmatrix} \Re(r_{t,1}) \\ \Im(r_{t,1}) \\ \vdots \\ \Re(r_{t,k}) \\ \Im(r_{t,k}) \end{pmatrix}}_{=:r_{t,n}} + \underbrace{\begin{pmatrix} \Re(\xi_{t,1}^{(1)}) \\ \Im(\xi_{t,1}^{(1)}) \\ \vdots \\ \Re(\xi_{t,k}^{(1)}) \\ \Im(\xi_{t,k}^{(1)}) \end{pmatrix}}_{=:\varepsilon_t}, \tag{16}$$

where $p = 2qk$. Stacking the real and imaginary parts of the measurements is a natural approach when the state vector is real-valued;[6] a complex-valued state vector would have required a complex Kalman filter based on the so-called widely linear complex estimator, as in Dini and Mandic (2012). The stacked observation equation (16) links all available information from option prices with several tenors at time t to the state vectors w_t and x_t in a linear way.

To complete the state space model we need to augment the measurement equation (16) by a transition equation for the unobservable state vector x_t. This is a linear, discrete-time

[5]Here, we attribute these elements (and system matrices $\tilde{d}_t, W_t$ and Z_t in equation (16)) with additional time index although the coefficient functions are assumed to be time-invariant in the exposition. This is because in practice we can have different expiration periods for different days.

[6]See Singleton (2001) and Chacko and Viceira (2003), who use this approach in a GMM estimation setting based on the empirical characteristic function.

dynamic system, to be derived from the continuous-time stochastic differential equation. An Euler discretization of the state process (8) would converge to the true transition dynamics as the discretization step $\Delta t \to 0$. However, the maximum likelihood (ML) estimator based on the Euler discretization is, in general, inconsistent for fixed non-zero Δt (Lo, 1988), because the discretization has conditional moments different from those of the true process (Piazzesi, 2010). Fortunately, the AJD assumption under $\mathbb{P}$ implies that the first and second conditional moments of x_{t+1} given $\mathcal{F}_t$ are linear and available in semi-closed form (possibly requiring the solution of a system of ODEs):

$$\mathbb{E}[x_{t+1}|\mathcal{F}_t] = c_t + T_t x_t, \tag{17}$$

$$\mathrm{Var}(x_{t+1}|\mathcal{F}_t) = Q_t(x_t), \tag{18}$$

where Q_t: $\mathbb{R}^d \to \mathbb{R}^{d\times d}$ is an affine function in x_t. The finiteness of the conditional moments is ensured by Assumption 1(ii). Both conditional moments will in general be linear in both the observed state w_t and the latent state x_t; but because the former does not need filtering, we absorb its effect in the time-varying intercept c_t, and similarly in the intercept of the affine function Q_t.[7] In Appendix B we show how these transition coefficients can be computed for the AJD model. Using this approach, which will be model-dependent and hence has to be applied case by case, we obtain a discrete-time transition equation with the same conditional mean and variance as the true continuous-time process (but possible different higher moments). Quasi-maximum likelihood (QML) estimation based on conditionally normally distributed measurement and transition errors in the state space representation yields consistent estimation results (Fisher & Gilles, 1996). A similar approach has been adopted in the term structure literature (see, e.g., de Jong, 2000, Duffee, 2002).

We summarize the development of the state space representation, and analyze properties of the errors, in the following proposition. The main result contains a remainder term in the measurement equation that collects the truncation and discretization errors in the construction of $\log\widehat{\phi}(u,\tau)$ and higher-order terms in the log-linearization. This term vanishes under an asymptotic scheme, where $\overline{m} := \max_{1\le j\le n} m_j \to \infty$, $\underline{m} := \min_{1\le j\le n} m_j \to -\infty$ and $\Delta m \to 0$. We also denote the corresponding smallest and largest strike prices as $\underline{K}$ and $\overline{K}$, and express the asymptotic orders with respect to the number of option prices n with fixed maturity.

Proposition 1 *Suppose Assumptions 1, 2 and 3 hold, and in addition* $\underline{K} \asymp n^{-\underline{\alpha}}$ *and* $\overline{K} \asymp n^{\overline{\alpha}}$ *with* $\underline{\alpha} > 0$ *and* $\overline{\alpha} > 0$. *Then* $\{(y_t, x_t), t = 1, \ldots, T\}$ *satisfy the linear state space representation*

$$y_t = d_t + Z_t x_t + r_{t,n} + \varepsilon_t, \qquad \mathbb{E}[\varepsilon_t|\mathcal{F}_t] = 0, \qquad \mathbb{E}[\varepsilon_t\varepsilon_t'|\mathcal{F}_t] = H_t, \tag{19}$$

$$x_{t+1} = c_t + T_t x_t + \eta_{t+1}, \qquad \mathbb{E}[\eta_{t+1}|\mathcal{F}_t] = 0, \qquad \mathbb{E}[\eta_{t+1}\eta_{t+1}'|\mathcal{F}_t] = Q_t(x_t), \tag{20}$$

where $r_{t,n} = \mathcal{O}_p\left(n^{-2(\underline{\alpha}\wedge\overline{\alpha})} \vee \frac{\log n}{n}\right)$ *and* $\varepsilon_t = \mathcal{O}_p\left(\sqrt{\frac{\log n}{n}}\right)$; $d_t = \tilde{d}_t + W_t w_t$ *and* Z_t *are defined*

[7] The transition matrix T_t will not be time-varying in stationary AJD process with equidistant observations, but we do not impose this time-constancy in the notation, also to avoid confusion with the sample size T.

in (16) *and* c_t, T_t *and* Q_t *are as given in* (17)–(18)*; and* $H_t = \text{blkdiag}\{H_{t,1}, \ldots, H_{t,k}\}$, *with*

$$H_{t,i} = \sigma_\varkappa^2 \cdot \begin{pmatrix} \frac{1}{2}\Re(\widetilde{\Gamma}_{t,i} + \widetilde{C}_{t,i}) & \frac{1}{2}\Im(-\widetilde{\Gamma}_{t,i} + \widetilde{C}_{t,i}) \\ \frac{1}{2}\Im(\widetilde{\Gamma}_{t,i} + \widetilde{C}_{t,i}) & \frac{1}{2}\Re(\widetilde{\Gamma}_{t,i} - \widetilde{C}_{t,i}) \end{pmatrix}, \quad i = 1, \ldots, k, \tag{21}$$

where $\widetilde{\Gamma}_{t,i}$ *and* $\widetilde{C}_{t,i}$ *are covariance and pseudo-covariance matrices of* $\xi_{i,t}/\sigma_\varkappa$*, with elements*

$$(\widetilde{\Gamma}_{t,i})_{kl} = \frac{u_{k,t}\overline{u_{l,t}} \sum_{j=2}^{n} e^{(\mathrm{i}(u_k - u_l) - 2)m_j} \kappa_t^2(\tau_i, m_j)\nu_t^2(\tau_i, m_j)(\Delta m_j)^2}{\phi_t(u_k, \tau_i)\phi_t(-u_l, \tau_i)}, \quad k, l = 1, \ldots, q,$$
$$(\widetilde{C}_{t,i})_{kl} = \frac{u_{k,t}u_{l,t} \sum_{j=2}^{n} e^{(\mathrm{i}(u_k + u_l) - 2)m_j} \kappa_t^2(\tau_i, m_j)\nu_t^2(\tau_i, m_j)(\Delta m_j)^2}{\phi_t(u_k, \tau_i)\phi_t(u_l, \tau_i)}, \quad k, l = 1, \ldots, q.$$

Furthermore,

(i) $\mathbb{E}[\varepsilon_t \varepsilon_s'] = 0$ *and* $\mathbb{E}[\eta_t \eta_s'] = 0$ *for* $s \neq t = 1, \ldots, T$*;*

(ii) $\mathbb{E}[\varepsilon_t \eta_s'] = 0$ *for all* $s, t = 1, \ldots, T$*;*

(iii) $\mathbb{E}[\varepsilon_t x_1'] = 0$ *and* $\mathbb{E}[\eta_{t+1} x_1'] = 0$ *for* $t = 1, \ldots, T$.

The proof is given in Appendix A. The orders suggest that the remainder term goes to zero faster than the observation term given some minimum non-zero requirements for $\underline{\alpha}$ and $\overline{\alpha}$. In what follows, we assume that $(\underline{\alpha} \wedge \overline{\alpha}) > \frac{1}{4}$ and neglect the remainder term in the estimation and filtering procedures. The system matrices $Z_t, T_t, Q_t(x_t)$ and system vectors d_t and c_t are known up to a parameter vector θ, assumed to lie in the interior of a compact parameter space Θ by Assumption 1(iii). Similarly, the system matrix H_t depends on the data and θ (via u_t, ϕ_t, κ_t and ν_t), and an additional unknown parameter $\sigma_\varkappa^2$. Note that (d_t, Z_t, H_t) are derived from the $\mathbb{Q}$-dynamics of (4), whereas $(c_t, T_t, Q_t(\cdot))$ correspond to the $\mathbb{P}$-dynamics. Therefore, possible deviations between $\mathbb{P}$ and $\mathbb{Q}$, reflecting the presence of factor risk premia, will require an extension of the parameter vector; we discuss this possibility further in Section 4 and Appendix D. Estimation of θ and filtering of the latent state vector via (versions of) the Kalman filter is considered in the next sub-section.

3.2 Modified and Collapsed Kalman filter

Consider the state space representation (19)–(20), where from now on we will ignore the remainder term $r_{t,n}$, and hence assume that the set of strike prices $\{m_j\}_{j=1}^n$ on each day is rich enough to make this term negligible. Define $Y_t = \{y_1, \ldots, y_t\}$, and linear projections (denoted $\widehat{\mathbb{E}}$) of the state vector conditional on the data: $\widehat{x}_{t|t} = \widehat{\mathbb{E}}[x_t|Y_t]$ and $\widehat{x}_{t|t-1} = \widehat{\mathbb{E}}[x_t|Y_{t-1}]$, with corresponding mean square error matrices $P_{t|t} = \mathbb{E}[(x_t - x_{t|t})(x_t - x_{t|t})']$ and $P_{t|t-1} = \mathbb{E}[(x_t - x_{t|t-1})(x_t - x_{t|t-1})']$. Then a modified version of the Kalman filter reads as follows:

$$\begin{aligned} v_t &= y_t - (d_t + Z_t \widehat{x}_{t|t-1}), & F_t &= Z_t P_{t|t-1} Z_t' + H_t, \\ \widehat{x}_{t|t} &= \widehat{x}_{t|t-1} + P_{t|t-1} Z_t' F_t^{-1} v_t, & P_{t|t} &= P_{t|t-1} - P_{t|t-1} Z_t' F_t^{-1} Z_t P_{t|t-1}, \\ \widehat{x}_{t+1|t} &= c_t + T_t \widehat{x}_{t|t}, & P_{t+1|t} &= T_t P_{t|t} T_t' + Q_t(\widehat{x}_{t|t}), \end{aligned}$$

for $t = 1, \ldots, T$. Given the stationarity of the process, the initial conditions $x_{1|0}$ and $P_{1|0}$ for the filter can be set to the unconditional mean and variance, respectively.

Because the error terms in the measurement and transition equations are not assumed to be Gaussian, the filtered state $\widehat{x}_{t|t}$ in conventional linear homoskedastic state space models is the linear projection (or minimum mean square error linear predictor) of the true process x_t, rather than its conditional expectation. This can be used to prove that quasi-maximum likelihood (QML) estimation based on the Gaussian likelihood yields consistent and asymptotically normal parameter estimates (Hamilton, 1994, Chapter 13). However, in general AJD models, the conditional variance $Q_t(x_t)$ is an affine function of the true latent state vector x_t. Therefore, the Kalman filter recursions have been modified by using $Q_t(\widehat{x}_{t|t})$ instead of the unobserved $Q_t(x_t)$. A similar modification is used in, e.g., de Jong (2000), Monfort, Pegoraro, Renne, and Roussellet (2017) and Feunou and Okou (2018). Although consistency of QML based on this modification has not been proved, Monte Carlo simulation results in these articles suggest that the method works well in practice.

Given the large dimension of the observation vector $p = 2qk$, the Kalman filter and its QML estimation will be computationally challenging if not infeasible. In fact, an important caveat with this approach is that one needs a non-singular innovation variance matrix F_t. Since our CCF approximation is based on common option price data for q different arguments u and fixed time-to-maturity τ, this matrix is likely to be (near-)singular for large q. Furthermore, with large cross-sectional dimension, the computation of the inverse matrix for each time t adds a significant computational burden to the estimation procedure. To overcome these issues we consider the collapsed Kalman filter, originally developed by Jungbacker and Koopman (2015), which we describe below. We extend their method to allow for a (near-)singular variance matrix H_t, using generalized inverses.

The idea of the collapsed Kalman filter is to transform the observation vector y_t into an uncorrelated pair of vectors y_t^* and y_t^+ such that the first vector y_t^* depends on the state vector x_t and has dimension $d \times 1$, while y_t^+ does not depend on x_t and has dimension $(p-d) \times 1$. Such a transformation can be done using, for instance, the projection matrices $A_t^* = (Z_t' H_t^- Z_t)^{-1} Z_t' H_t^-$ and $A_t^+ = L_t H_t^- (I_p - Z_t A_t^*)$, where L_t is chosen such that A_t^+ has full row rank and H^- is the generalized inverse of H. Since $A_t^* Z_t = I_p$ and $A_t^+ Z_t = 0$, the observation equation can be transformed into

$$\begin{pmatrix} y_t^* \\ y_t^+ \end{pmatrix} := \begin{bmatrix} A_t^* \\ A_t^+ \end{bmatrix} y_t = \begin{pmatrix} d_t^* \\ d_t^+ \end{pmatrix} + \begin{pmatrix} x_t \\ 0 \end{pmatrix} + \begin{pmatrix} \varepsilon_t^* \\ \varepsilon_t^+ \end{pmatrix}, \tag{22}$$

with $y_t^* = A_t^* y_t$, $y_t^+ = A_t^+ y_t$, $\varepsilon_t^* = A_t^* \varepsilon_t$, $\varepsilon_t^+ = A_t^+ \varepsilon_t$, $d_t^* = A_t^* d_t$ and $d_t^+ = A_t^+ d_t$. Using $H^- H H^- = H^-$, we have

$$\begin{aligned}
\mathrm{Var}(\varepsilon_t^*) &= A_t^* H_t A_t^{*\prime} = (Z_t' H_t^- Z_t)^{-1} =: H_t^*, \\
\mathrm{Var}(\varepsilon_t^+) &= A_t^+ H_t A_t^{+\prime} =: H_t^+, \\
\mathrm{Cov}(\varepsilon_t^*, \varepsilon_t^+) &= A_t^* H_t A_t^{+\prime} = A_t^* H_t (I_p - A_t^{*\prime} Z_t) H_t^- L' \\
&= A_t^* H_t H_t^- L' - (Z_t' H_t^- Z_t)^{-1} Z_t' H_t^- L' = A_t^* L' - A_t^* L' = 0.
\end{aligned}$$

In the preceding display, it has been assumed that $\text{rank}(Z_t' H_t^- Z_t) = d$; this is not very restrictive, given that the dimension d of the state vector will typically be much smaller than the dimension p of the observation vector. We also require that the matrix $A_t = [A_t^{*\prime}, A_t^{+\prime}]'$ is non-singular, such that the transformation $A_t y_t$ does not lead to a loss of information in the first two moments.

The representation (22) shows that information about the state vector x_t is contained in the observation equation for y_t^*; thus we may ignore the second equation with y_t^+ and focus on the collapsed state space model:

$$y_t^* = d_t^* + x_t + \varepsilon_t^*, \qquad \mathbb{E}[\varepsilon_t^* | \mathcal{F}_t] = 0, \ \text{Var}(\varepsilon_t^* | \mathcal{F}_t) = H_t^*, \tag{23}$$

$$x_{t+1} = c_t + T_t x_t + \eta_{t+1}, \qquad \mathbb{E}[\eta_{t+1} | \mathcal{F}_t] = 0, \ \text{Var}(\eta_{t+1} | \mathcal{F}_t) = Q_t(x_t). \tag{24}$$

Let us emphasize again that the collapsing transformation into a lower dimensional state space form is also valid for the Moore-Penrose inverse covariance matrix H_t^-. Therefore, we can collapse a high-dimensional data vector into a lower-dimensional vector even when the covariance system matrix of disturbances is (near-) singular.

The logarithm of the Gaussian likelihood function of the data vector $Y_T = (y_1', \ldots, y_T')'$ is

$$l(Y_T; \theta) = \sum_{t=1}^{T} \log p_\theta(y_t | Y_{t-1}),$$

where $p_\theta(y_t | Y_{t-1})$ is the (misspecified) Gaussian distribution of y_t conditional on Y_{t-1} (and $w_1, \ldots, w_{t-1}$), which can be evaluated via the prediction error decomposition based on the original state space representation (19)–(20). Given the assumption of a full rank transformation matrix $|A_t|$, the collapsed transformation allows to decompose the log-likelihood function $l(Y_T; \theta)$ into three parts to ease computation:

$$l(Y_T; \theta) = l(Y_T^*; \theta) + l(Y_T^+; \theta) + \sum_{t=1}^{T} \log |A_t|, \tag{25}$$

where Y_T^* and Y_T^+ are stacked vectors of y_t^* and y_t^+ over $t = 1, \ldots, T$, respectively.

The first term in (25) is the quasi-loglikelihood evaluated by the Kalman filter applied to the collapsed state space system (23)–(24):

$$l(Y_T^*; \theta) = -\frac{dT}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{T} \log |F_t^*| - \frac{1}{2} \sum_{t=1}^{T} v_t^{*\prime} F_t^{*-1} v_t^*,$$

where v_t^* are the prediction errors and F_t^* are their mean-square error (MSE) matrices from the Kalman filter.

Since y_t^+ does not depend on the state vector α_t and $|H_t^+| = 1$ may be imposed without loss of generality, see Jungbacker and Koopman (2015), the second term is given by

$$l(Y_T^+; \theta) = -\frac{(p-d)T}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{T} (y_t^+ - d_t^+)' (H_t^+)^{-1} (y_t^+ - d_t^+).$$

Fortunately, the last term in the expression above can be calculated without construction of the matrix A_t^+:

$$\begin{aligned}(y_t^+ - d_t^+)'(H_t^+)^{-1}(y_t^+ - d_t^+) &= (y_t - d_t)' A_t^{+'}(A_t^+ H_t A_t^+)^{-1} A_t^+ (y_t - d_t) \\ &= (y_t - d_t)' J_t^+ H_t^- (y_t - d_t) \\ &= (y_t - d_t)' J_t^+ H_t^- J_t^{+'} (y_t - d_t) \\ &= (y_t - d_t)' M_Z' H_t^- M_Z (y_t - d_t) \\ &= e_t' H_t^- e_t,\end{aligned}$$

where $M_Z = I - Z_t(Z_t' H_t^- Z_t)^{-1} Z_t' H_t^- = I - Z_t A_t^*$, $J_t^+ = A_t^{+'}(A_t^+ H_t A_t^+)^{-1} A_t^+ H_t$ and $e_t = M_Z(y_t - d_t)$, that is, these are the generalized least squares (GLS) residuals from the observation vector y_t with the covariate matrix Z_t and variance matrix H_t. For derivation details,[8] see Jungbacker and Koopman (2015).

Finally, the third term in (25), $|A_t|$, can be found from the relation

$$|A_t|^2 \cdot |H_t| = |A_t H_t A_t'| = |H_t^*| \cdot |H_t^+| = |H_t^*|, \tag{26}$$

which follows from the fact that the covariance matrix $A_t H_t A_t'$ is block diagonal given the uncorrelated error terms ε_t^* and ε_t^+ and using again $|H_t^+| = 1$.

Given the measurement error structure as implied by Proposition 1, the single scale parameter $\sigma_\varkappa^2$ of the covariance matrix can be factored out as $H_t = \sigma_\varkappa^2 \cdot \widetilde{H}_t$. The matrix $\widetilde{H}_t$ has a block-diagonal structure; although its blocks depend on the state vector and parameters via the theoretical BSIV $\kappa_t(\tau, m)$ and vega $\nu_t(\tau, m)$, we estimate these quantities directly from the data, hence they are not updated as we optimize over θ. Thus, we have from (26) that

$$\begin{aligned}\log|A_t| &= \frac{1}{2}\left(\log|H_t^*| - \log|H_t|\right) \\ &= \frac{1}{2}\left(\log|H_t^*| - \log\sigma_\varkappa^{2p} - \log|\widetilde{H}_t|\right) \\ &\propto \frac{1}{2}\log|H_t^*| - p\log\sigma_\varkappa.\end{aligned}$$

Therefore, the log-likelihood (25) is proportional to

$$l(Y_T;\theta) \propto \frac{1}{2}\sum_{t=1}^{T}\left(-\log|F_t^*| - v_t^{*'} F_t^{*-1} v_t^* - e_t' H_t^- e_t + \log|H_t^*|\right) - pT\log\sigma_\varkappa. \tag{27}$$

Note that the inversions and determinants of the matrices F_t^* and H_t^* can be computed efficiently since they have small dimensions $d \times d$. This eases maximization of the log-likelihood function (27) substantially.

The quasi maximum-likelihood parameter estimates $\widehat{\theta}$ are obtained by maximizing (27) over the model parameter space Θ (we implicitly assume that the parameter vector θ has been extended to include the additional parameter $\sigma_\varkappa^2$). Its asymptotic properties are analogous to QML estimation based on the (modified) Kalman filter, as discussed at the beginning of this

[8] The derivation in Jungbacker and Koopman (2015) is based on the invertible covariance matrix H_t, but the same result and the same derivation are valid when using the pseudo-inverse matrix H_t^-.

sub-section. In cases where the conditional covariance matrix Q_t does not depend on the latent state vector x_t, and the latent state process x_t is stationary, QML based on the Kalman filter will yield consistent and asymptotically normal estimators. When Q_t is affine in x_t, then QML based on the modified Kalman filter appears to have comparable properties in Monte Carlo simulations, but no formal consistency proof is available.

4 Monte Carlo Study

In this section we study the finite-sample performance of our estimation procedure. In particular, we consider two AJD specifications: a one-factor model and two versions of an option pricing model with two factors.

Before we turn to the Monte Carlo results, we note that our developed estimation and filtering approach uses information from option prices, but is agnostic about equity risk premia. The reason is that the measurements are constructed as a portfolios of options rather than the underlying asset. On the other hand, since the transition equation in the state space representation reflects the dynamics of the latent components (under $\mathbb{P}$), it is, in general, possible to learn about the risk premia associated with this latent process (for instance, the variance risk premium). However, our simulation results suggest (see Appendix D) that the $\mathbb{Q}$-information in the option prices largely dominates the $\mathbb{P}$-information, making the identification of risk premium parameters weak. A similar difficulty of identifying the physical dynamics arises in the term structure literature (see, e.g., Kim & Orphanides, 2012). Therefore, in what follows, we assume no variance (or state related) risk premia, that is, the latent components have the same dynamics under the both probability measures. Importantly, the results in Appendix D suggest that estimation of the $\mathbb{Q}$ parameters is hardly affected by imposing this (possibly invalid) restriction.

4.1 SVCDEJ

As a starting point, we illustrate the developed estimation approach based on the modification of the widely used 'double-jump' stochastic volatility model of Duffie et al. (2000). The modification is due to the double-exponential jump sizes in returns as in Kou (2002) and Andersen et al. (2015a), and a stochastic jump intensity that is a multiple of the stochastic variance, as in Pan (2002). We label this specification as 'SVCDEJ' for stochastic volatility model with co-jumps in volatility and double-exponential jumps in returns.

In particular, we assume the following data-generating process for the log forward price under the both $\mathbb{P}$ and $\mathbb{Q}$ probability measures:

$$\mathrm{d}\log F_t = (-\tfrac{1}{2}v_t - \mu\lambda_t)\mathrm{d}t + \sqrt{v_t}\mathrm{d}W_{1,t} + J_t\mathrm{d}N_t, \tag{28}$$

$$\mathrm{d}v_t = \kappa(\bar{v} - v_t)\mathrm{d}t + \sigma\sqrt{v_t}\mathrm{d}W_{2,t} + J_t^v\mathbf{1}_{\{J_t<0\}}\mathrm{d}N_t, \tag{29}$$

where two Brownian motions W_1 and W_2 are assumed to be correlated with coefficient ρ, and N_t is the Poisson jump process with jump intensity proportional to the stochastic variance, $\lambda_t = \delta v_t$. We further assume that J_t is the double-exponentially distributed jump size in

returns with probability density function

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

where p^+ and p^- are probabilities of positive and negative jumps, respectively, and η^+ and η^- are the corresponding conditional means of the jump sizes. We assume that all of these parameters are positive, $p^+ + p^- = 1$ and $\eta^+ < 1$. Given the jump size distribution, the expected relative jump size in returns is

$$\mu := \mathbb{E}[e^J - 1] = \frac{p^+}{1-\eta^+} + \frac{p^-}{1+\eta^-} - 1.$$

Finally, we allow the volatility to co-jump only with negative jumps in returns, with exponentially distributed jump sizes J_t^v with mean μ_v.

The model in (28)–(29) belongs to the AJD class and exhibits all important ingredients in option-pricing models: stochastic volatility, jump components in returns and volatility, time-varying jump intensity and a self-excitation feature (because a negative jump in returns is associated with a positive jump in volatility, which increases the volatility and hence jump intensity). Furthermore, this specification assumes a double-exponential jump size distribution in returns, which has been recently advocated in the literature (see, e.g., Kou, 2002, Andersen et al., 2015a and Bardgett et al., 2019).

The discounted marginal CCF of the log forward prices in the SVCDEJ model is given by

$$\psi_X(\mathbf{u}_1, \tau) = e^{-r\tau} \mathbb{E}^{\mathbb{Q}}[e^{\mathrm{i}u \log F_{t+\tau}} | \mathcal{F}_t] = e^{\alpha(u,\tau) + \beta_1(u,\tau) \log F_t + \beta_2(u,\tau) v_t}, \tag{30}$$

where $\alpha(u,\tau)$ and $\beta(u,\tau)$ are solutions to the complex-valued ODE system in time:

$$\begin{cases} \dot{\beta}_1(u,s) & = 0, \\ \dot{\beta}_2(u,s) & = -\left(\frac{1}{2} + \mu\delta\right)\beta_1(u,s) - \kappa\beta_2(u,s) + \frac{1}{2}\beta_1(u,s) + \rho\sigma\beta_1(u,s)\beta_2(u,s) \\ & \quad + \frac{1}{2}\sigma^2\beta_2^2(u,s) + \delta(\chi(\beta_1(u,s), \beta_2(u,s)) - 1), \\ \dot{\alpha}(u,s) & = \kappa\bar{v}\beta_2(u,s) - r, \end{cases}$$

with initial conditions $\beta_1(u,0) = \mathrm{i}u$, $\beta_2(u,0) = 0$ and $\alpha(u,0) = 0$. Here the 'jump transform' is

$$\chi(\beta_1, \beta_2) = \frac{p^+}{1-\beta_1\eta^+} + \frac{p^-}{(1+\beta_1\eta^-)(1-\beta_2\mu_v)}.$$

The CCF (30) of the log price is used to price options. For the state space representation, we turn it into the CCF of the log returns as described in Section 2.3. Using the fact that the solution to the ODE system has $\beta_1(u,\tau) = \mathrm{i}u$, the linear relation between the log of the option-spanned CCF and the state vector is given by

$$\log \widehat{\phi}_t(u,\tau) = \alpha(u,\tau) + \beta_2(u,\tau) v_t + \xi_t(u,\tau), \qquad u \in \mathbb{R},$$

where $\widehat{\phi}_t(u,\tau)$ is the option-implied CCF, $\tau > 0$ is the time-to-maturity of available options and $\xi_t(u,\tau)$ is the measurement error term due to observation and approximation errors in the

option-spanned CCF. We use this linear relation to construct the measurement equation as discussed in Section 3.1.

Following Appendix B, the conditional mean and variance of the latent stochastic volatility process are given by

$$\mathbb{E}[v_{t+1}|\mathcal{F}_t] = e^{g_1\Delta t}v_t + \frac{g_0}{g_1}\left(e^{g_1\Delta t} - 1\right), \tag{31}$$

$$\mathrm{Var}(v_{t+1}|\mathcal{F}_t) = -\frac{\sigma^2 + 2p^-\delta\mu_v^2}{2g_1^2}\left[2g_1\left(e^{g_1\Delta t} - e^{2g_1\Delta t}\right)v_t - g_0\left(1 - e^{g_1\Delta t}\right)^2\right], \tag{32}$$

with $g_0 = \kappa\bar{v}$ and $g_1 = -\kappa + p^-\delta\mu_v$. Equations (31)–(32) are then used to define the state updating equation:

$$v_{t+1} = c_t + T_t v_t + \eta_{t+1}, \tag{33}$$

where $c_t = \frac{g_0}{g_1}\left(e^{g_1\Delta t} - 1\right)$, $T_t = e^{g_1\Delta t}$ and $\mathrm{Var}(\eta_{t+1}|\mathcal{F}_t) = \mathrm{Var}(v_{t+1}|\mathcal{F}_t) =: Q_t(v_t)$. We also impose the Feller condition $2\kappa\bar{v} > \sigma^2$ and the covariance stationary condition $\kappa > p^-\delta\mu_v$.

The model specification has nine parameters of interest and a single parameter that characterizes the observation errors. We note that the parameter δ often enters as a multiple of p^-, which can possibly cause identification issue in the estimation procedure. Therefore, to avoid the identification issues we fix the probability of positive jumps to be $p^- = 0.7$. This is consistent with findings in Aït-Sahalia et al. (2015) and our empirical results for the unrestricted model.

In the simulation exercise we use $T = 500$ time points with $\Delta t = 1/250$. The time-series of the log prices and true spot volatilities are simulated using the Euler scheme from the specification (28)–(29). The options data are generated using the COS method of Fang and Oosterlee (2008) based on the CCF, specified in (30). The model parameters in the simulation are indicated in Table 1 along with the Monte Carlo results. The initial values are set to $F_0 = 100$ and $v_0 = 0.015$.

In the simulation we consider three tenors for options equal to 10, 30 and 60 days. For each tenor we simulate a finite number of options with log-moneyness between $\underline{m} = -10 \cdot \sigma_{ATM,\tau}\sqrt{\tau}$ and $\overline{m} = 4 \cdot \sigma_{ATM,\tau}\sqrt{\tau}$, where $\sigma_{ATM,\tau}$ is the BSIV of the ATM option with time-to-maturity τ. Furthermore, the strikes are generated equidistantly with $\Delta K = 0.01 \cdot F_t$. Finally, we distort the options data by adding the observation errors to the option prices for each tenor τ and each log-moneyness level m as specified in Assumption 2, i.e.,

$$\widehat{O}_t(\tau, m) = O_t(\tau, m) + \sigma_{\varkappa} \cdot \kappa_t(\tau, m)\nu_t(\tau, m) \cdot \epsilon$$

where ϵ is an i.i.d. standard normal random variable and $\sigma_{\varkappa} = 0.02$. The distorted option prices are then interpolated using the cubic spline and linearly extrapolated in total variance as discussed in Appendix C.1.

The covariance matrix of the errors in the measurement equation is calculated as given by equation (21). To calculate the pseudo-inverse of the $2q \times 2q$ covariance matrix $\widetilde{H}_{t,i}$ for each of the maturities $i = 1, \ldots, k$, we set the following level of the threshold for the singular values:

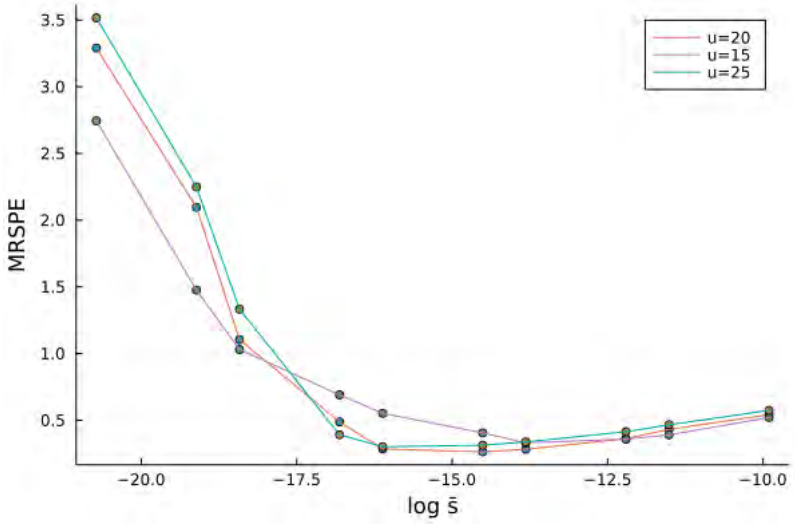
$$\mathrm{tol} := \bar{s} \cdot 2q \cdot \max_j s_j$$

Table 1: Monte Carlo results for the SVCDEJ model

parameter	σ	κ	$\bar{v}$	ρ	δ	η^+	η^-	μ_v	$\sigma_\varkappa$
				$u = 1, \ldots, 5$					
true value	0.450	8.000	0.0150	-0.9500	100.000	0.020	0.050	0.050	0.020
mean	0.498	8.213	0.0151	-0.8983	105.906	0.021	0.048	0.048	0.032
std dev	0.007	0.257	0.0005	0.0088	7.567	0.001	0.001	0.001	0.005
q10	0.494	7.965	0.0146	-0.9097	96.930	0.021	0.047	0.046	0.026
q50	0.498	8.198	0.0151	-0.8956	105.936	0.021	0.048	0.048	0.032
q90	0.504	8.503	0.0158	-0.8904	114.125	0.022	0.049	0.049	0.038
				$u = 1, \ldots, 10$					
true value	0.450	8.000	0.0150	-0.9500	100.000	0.020	0.050	0.050	0.020
mean	0.440	8.779	0.0136	-0.9968	136.200	0.023	0.045	0.043	0.035
std dev	0.017	0.310	0.0007	0.0119	13.316	0.001	0.002	0.002	0.006
q10	0.427	8.409	0.0130	-1.0000	118.640	0.022	0.043	0.041	0.027
q50	0.437	8.836	0.0134	-1.0000	139.079	0.023	0.045	0.043	0.036
q90	0.451	9.102	0.0142	-1.0000	150.730	0.024	0.047	0.046	0.042
				$u = 1, \ldots, 15$					
true value	0.450	8.000	0.0150	-0.9500	100.000	0.020	0.050	0.050	0.020
mean	0.440	8.723	0.0139	-0.9942	128.617	0.022	0.046	0.045	0.027
std dev	0.013	0.316	0.0005	0.0151	10.260	0.001	0.001	0.002	0.005
q10	0.428	8.330	0.0134	-1.0000	113.412	0.021	0.044	0.043	0.022
q50	0.438	8.741	0.0138	-1.0000	130.575	0.022	0.045	0.044	0.028
q90	0.455	9.102	0.0144	-0.9785	139.999	0.023	0.048	0.047	0.033
				$u = 1, \ldots, 20$					
true value	0.450	8.000	0.0150	-0.9500	100.000	0.020	0.050	0.050	0.020
mean	0.455	8.143	0.0147	-0.9558	110.734	0.022	0.048	0.046	0.023
std dev	0.007	0.205	0.0003	0.0126	4.847	0.001	0.001	0.001	0.006
q10	0.449	7.919	0.0144	-0.9707	105.192	0.021	0.047	0.045	0.018
q50	0.454	8.142	0.0147	-0.9563	110.689	0.021	0.048	0.046	0.022
q90	0.461	8.389	0.0150	-0.9430	116.719	0.022	0.049	0.047	0.027
				$u = 1, \ldots, 25$					
true value	0.450	8.000	0.0150	-0.9500	100.000	0.020	0.050	0.050	0.020
mean	0.460	7.918	0.0150	-0.9404	105.489	0.021	0.049	0.047	0.023
std dev	0.008	0.188	0.0003	0.0126	3.881	0.000	0.001	0.001	0.008
q10	0.453	7.697	0.0148	-0.9492	103.098	0.021	0.048	0.046	0.017
q50	0.459	7.958	0.0150	-0.9449	105.744	0.021	0.049	0.047	0.020
q90	0.468	8.082	0.0152	-0.9229	108.351	0.022	0.050	0.048	0.032
				$u = 1, \ldots, 30$					
true value	0.450	8.000	0.0150	-0.9500	100.000	0.020	0.050	0.050	0.020
mean	0.457	7.815	0.0149	-0.9432	111.115	0.022	0.049	0.044	0.026
std dev	0.011	0.209	0.0003	0.0190	5.794	0.001	0.001	0.002	0.010
q10	0.448	7.523	0.0147	-0.9596	104.771	0.021	0.048	0.042	0.019
q50	0.453	7.857	0.0148	-0.9514	111.854	0.022	0.048	0.044	0.023
q90	0.472	8.016	0.0151	-0.9105	116.199	0.023	0.049	0.046	0.038

Note: This table provides Monte Carlo simulation results for the SVCDEJ model, based on 300 replications. Six settings with different ranges of the argument are considered. Each panel lists, for each parameter, the true value, the Monte Carlo mean and standard deviation, and the 10th, 50th and 90th Monte Carlo percentile, respectively. We use $T = 500$ time points with $\Delta t = 1/250$. The initial values are set to $F_0 = 100$ and $v_0 = 0.015$. The threshold for singular values is set to $\bar{s} = 10^{-7}$. The probability of negative jumps is fixed to $p^- = 0.7$.

Figure 1: MRSPE for different levels of $\bar{s}$



Note: This figure plots the MRSPE for different levels of $\bar{s}$ for three different ranges of arguments. In particular, for each dot we estimate $N = 300$ simulations and consider $\bar{s} = \{10^{-9}, 5{\cdot}10^{-9}, 10^{-8}, 5{\cdot}10^{-8}, 10^{-7}, 5{\cdot}10^{-7}, 10^{-6}, 5{\cdot}10^{-6}, 10^{-5}, 5{\cdot}10^{-5}\}$ and plot it in the log-scale.

with $\bar{s} = 10^{-7}$ and where s_j are all singular values of $\widetilde{H}_{t,i}$. We also consider robustness to the choice of $\bar{s}$.

Table 1 provides the Monte Carlo results for the SVCDEJ model for six different ranges of the argument sets $\mathcal{U}$. The results in general show a good finite-sample performance of the QMLE. We notice that for smaller ranges of the argument of the CCF, the estimates exhibit biases for some model parameters. This is expected since the larger ranges provide larger information on which we build the filtering and parameter estimation procedures. On the other hand, we also notice that the variance of some parameter estimates starts increasing for a very large range of arguments ($u = 1, \ldots, 30$). This is likely due to an increased variance in the CCF approximation for large arguments u.

To explore the robustness to the choice of the truncation level in the pseudo-inversion of the covariance matrix, we also consider another values for $\bar{s}$. In particular, we run $N = 300$ simulations for each level of $\bar{s}$ with the same parameters as in Table 1 and construct the mean root square percentage error (MRSPE) metrics, defined as $\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\hat{\theta}_i - \theta_0}{\theta_0}\right\|$. Figure 1 plots the MRSPE for different levels of $\bar{s}$ and three different ranges of arguments. As we can see, the levels $\bar{s}$ in between 10^{-7} and 10^{-6} yield the smallest MRSPE. In the following simulations and empirical part we therefore set $\bar{s} = 10^{-7}$.

We end this subsection by noting that we have also performed simulation exercises for some other alternative one-factor specifications. In Appendix D, we include the simulation results for the SVCJ model with a Gaussian jump size distribution, including a version with a variance risk premium, and the model with two separate counting processes for positive and negative jumps. The former shows a very good finite-sample performance, while the latter, a richer specification, shows reasonable results, reaching the identification limits.

4.2 SVCDEJ with external factors

Now we extend the one-factor specification by adding an external factor. This modification can be seen as a two-factor specification, but we will assume that the second factor is observable. The motivation comes from the fact that in some situations we might have an understanding of possible drivers of the risks in the market. Therefore, we would like to embed and quantify the impact of exogenous variables to the model's risk.

In particular, next to the stochastic volatility component we introduce the exogenous factor h_t, which affects the intensity of jumps and the diffusive component. The model reads as follows:

$$\mathrm{d}\log F_t = (-\tfrac{1}{2}V_t - \mu\lambda_t)\mathrm{d}t + \sqrt{v_t}\mathrm{d}W_{1,t} + q\sqrt{h_t}\mathrm{d}W_{3,t} + J_t\mathrm{d}N_t, \tag{34}$$

$$\mathrm{d}v_t = \kappa(\bar{v} - v_t)\mathrm{d}t + \sigma\sqrt{v_t}\mathrm{d}W_{2,t} + J_t^v\mathbf{1}_{\{J_t<0\}}\mathrm{d}N_t, \tag{35}$$

$$\mathrm{d}h_t = \kappa_h(\bar{h} - h_t)\mathrm{d}t + \sigma_h\sqrt{h_t}\mathrm{d}W_{4,t}, \tag{36}$$

where $V_t = v_t + q^2 h_t$ is the total diffusive variance of the process and the intensity processes λ is now also affected by h_t, i.e., $\lambda_t = \delta v_t + \gamma h_t$. We assume that $W_{3,t}$ and $W_{4,t}$ are independent Brownian motions, jointly independent of $(W_{1,t}, W_{2,t})$. Therefore, the process h_t is exogenous to the SVCDEJ dynamics, meaning that the dynamics of y_t and v_t do not affect the dynamics of h_t. In turn, the exogenous factor h_t affects the intensity of jumps and the diffusive component of the log return dynamics. This specification is similar to the two-factor model in Andersen et al. (2015a), which includes short- and long-term stochastic volatility components. The difference is that the exogenous process h_t is observable, although its parameters are unknown.

In the Monte Carlo simulation we consider two possible estimation approaches. In the first approach, we assume the correct specification of the dynamics for h_t with known true parameters κ_h, $\bar{h}$ and σ_h. In practice, these parameters can be pre-estimated given the observed path of the exogenous process. In the second approach we estimate the misspecified model, in which the contribution of h_t is constant throughout the maturity of an option. In other words, in this case we ignore the dynamics of h_t when pricing options, but h_t still affects the level of intensity and of total variance. The motivation is that when the exogenous process is persistent and smooth relative to v_t, its dynamics can be neglected when pricing options with short expiration periods. In a similar way, the interest rates are often assumed to enter the option pricing in a deterministic way. Moreover, the true parametric specification for an exogenous variable is likely unknown in practice. However, if its dynamics is persistent and smooth, we can find its effect on option prices via this specification. Therefore, in the Monte Carlo experiment, we simulate h_t with lower mean-reversion rate than in v_t, mimicking what we will use in the empirical application.

The Monte Carlo setting for the SVCDEJ model with external factor is the same as in the SVCDEJ specification. The parameters of the external factor are set to be $\kappa_h = 1$, $\bar{h} = 1$ and $\sigma_h = 0.1$. The simulation results are provided in Table 2. The parameters of the SVCDEJ model exhibit similar performance in the both estimation settings. Importantly, the parameters related to the external factors γ and q also show similar performance in the misspecified setting as in the correctly specified model. We emphasize that the latter is achieved due to simulating

relatively smooth and persistent exogenous process h_t and using short-dated options in the estimation procedure.

Table 2: Monte Carlo results for the SVCDEJ model with external factor

parameter	σ	κ	$\bar{v}$	ρ	δ	η^+	η^-	μ_v	γ	q	$\sigma_\varkappa$
(a) estimation with fixed parameters for h_t											
true value	0.450	8.00	0.0150	-0.950	100.00	0.0200	0.050	0.050	1.500	0.050	0.020
mean	0.452	8.15	0.0150	-0.939	116.21	0.0211	0.049	0.046	1.480	0.046	0.024
std dev	0.034	0.36	0.0019	0.047	11.37	0.0008	0.001	0.002	0.167	0.005	0.011
q10	0.423	7.65	0.0140	-0.977	102.67	0.0206	0.048	0.044	1.348	0.042	0.016
q50	0.442	8.22	0.0145	-0.954	117.07	0.0209	0.048	0.045	1.531	0.048	0.019
q90	0.502	8.52	0.0167	-0.857	128.02	0.0218	0.050	0.047	1.565	0.049	0.043
(b) estimation of misspecified model											
true value	0.450	8.00	0.0150	-0.950	100.00	0.0200	0.050	0.050	1.500	0.050	0.020
mean	0.450	8.16	0.0150	-0.939	116.93	0.0210	0.049	0.046	1.473	0.046	0.024
std dev	0.034	0.37	0.0022	0.044	11.47	0.0016	0.001	0.002	0.165	0.006	0.012
q10	0.425	7.70	0.0140	-0.973	106.71	0.0206	0.048	0.044	1.380	0.043	0.016
q50	0.441	8.23	0.0146	-0.954	117.78	0.0209	0.048	0.045	1.516	0.047	0.019
q90	0.492	8.55	0.0159	-0.861	128.68	0.0214	0.049	0.047	1.546	0.048	0.043

Note: This table provides Monte Carlo simulation results for the SVCDEJ model with an exogenous factor, based on 300 replications. Each panel lists, for each parameter, the true value, the Monte Carlo mean and standard deviation, and the 10th, 50th and 90th Monte Carlo percentile, respectively. We use $T = 500$ time points with $\Delta t = 1/250$. The range for the arguments is set to $u = 1, \ldots, 20$ and the threshold $\bar{s} = 10^{-7}$. The initial values are set to $F_0 = 100$ and $v_0 = 0.015$. The probability of negative jumps is fixed to $p^- = 0.7$. The parameters of the external factor are set to be $\kappa_h = 1$, $\bar{h} = 1$ and $\sigma_h = 0.1$.

5 Data

This section provides detail on the data, and its selection process, that we use in our empirical application. Since our estimation procedure utilizes the option-implied CCFs, we also pay careful attention in this section to the construction of these characteristics.

5.1 Data description

In this paper we use options data on the S&P 500 stock market index obtained from the Chicago Board Options Exchange (CBOE). We focus on the period from May 1, 2017 to April 1, 2021, covering the turbulent period in the stock market due to the Covid-19 pandemic. The CBOE provides end-of-day option quotes and a snapshot at 3:45 pm ET, 15 minutes prior the market closure. We use the latter to calculate mid-quotes since it is considered to be a more accurate representation of market liquidity than the former. The data contain both the 'standard' AM-settled SPX options and Weeklys and End-of-Months PM-settled SPXW products. The settlement value for the SPX options is based on the open level of the S&P 500 index on settlement day, while for the SPXW options it is based on the closing prices of the index.

Table 3: Descriptive statistics for S&P 500 index options

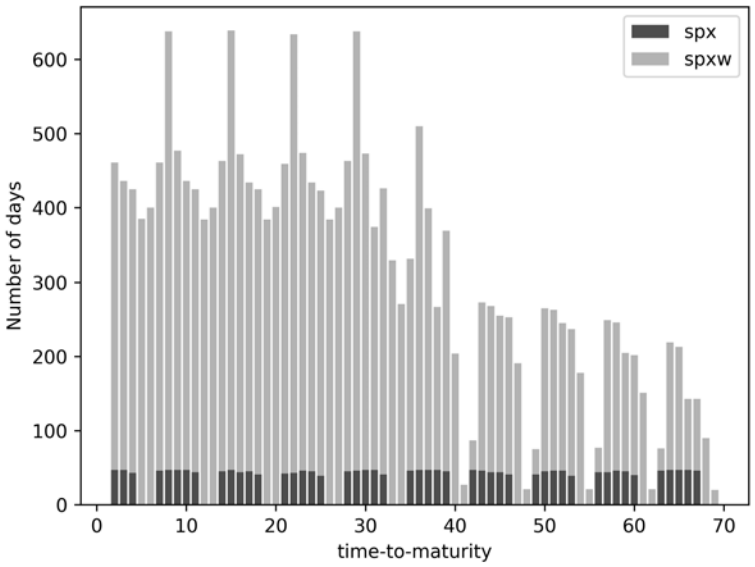
	$k \leq 0.8$	$0.8 < k \leq 0.95$	$0.95 < k \leq 1.03$	$1.03 < k \leq 1.1$	$1.1 < k$	Total
			Panel A: Total volume of option contracts (in millions)			
$\tau \in (2, 9]$	4.52	52.61	167.09	13.27	1.25	238.73
$\tau \in (9, 30]$	25.47	74.58	168.06	29.93	3.96	302.00
$\tau \in (30, 60]$	25.85	61.97	109.53	28.80	4.22	230.38
$\tau \in (60, 90]$	11.68	21.70	40.82	11.05	3.05	88.30
$\tau \in (90, 180]$	19.93	22.83	27.23	10.59	4.57	85.15
$\tau \in (180, 365]$	10.74	10.91	10.82	4.54	5.12	42.12
Total	98.19	244.60	523.55	98.18	22.16	986.68
			Panel B: Volume of OTM option contracts (in millions)			
$\tau \in (2, 9]$	4.12	52.03	139.33	11.94	0.96	208.37
$\tau \in (9, 30]$	24.04	73.54	133.21	27.87	3.36	262.03
$\tau \in (30, 60]$	23.82	61.11	83.72	27.28	3.85	199.79
$\tau \in (60, 90]$	10.62	21.21	26.72	10.44	2.78	71.77
$\tau \in (90, 180]$	18.77	22.15	18.23	9.92	4.30	73.35
$\tau \in (180, 365]$	10.34	10.29	6.96	4.06	4.88	36.53
Total	91.71	240.32	408.17	91.51	20.14	851.85
			Panel C: Average OTM option price ($)			
$\tau \in (2, 9]$	0.43	1.38	8.69	2.69	1.73	4.32
$\tau \in (9, 30]$	0.92	5.53	23.53	5.47	2.68	10.07
$\tau \in (30, 60]$	2.42	13.28	40.10	11.84	4.68	16.73
$\tau \in (60, 90]$	5.05	24.89	63.56	22.00	6.17	24.93
$\tau \in (90, 180]$	10.97	50.74	104.11	47.11	11.76	43.14
$\tau \in (180, 365]$	18.80	90.42	155.53	85.29	19.67	53.05
Total	7.97	22.29	47.73	22.35	9.64	24.02

Note: Descriptive statistics for filtered option data on the S&P 500 stock market index. The sample contains daily option data from 1 May 2017 to 1 April 2021. Observations are bucketed into six categories based on the time-to-maturity, τ, and into five categories with respect to the moneyness level, defined as strike-to-forward ratio $k = K/F$.

Given that we need a reliable and wide coverage of option prices for each tenor, we use a fairly generous set of filters. In particular, we retain option observations that satisfy the following criteria: (i) bid price is strictly positive and ask-to-bid ratio is less than 10; (ii) the maturity is larger than or equal to 2 calendar days, but less than or equal to 365 calendar days; (iii) it is not an early-closure day. The first criterion filters out illiquid observations and the second one limits our consideration in terms of options' maturity. The third criterion rules out shortened trading sessions, which in total constitute 10 days in our sample.

For each tenor we determine the moneyness based on the forward index level, $F_t(\tau)$. For that, we use the put-call parity to calculate the forward price for close to at-the-money (ATM) options. Specifically, we use up to 5 option pairs with the smallest absolute difference between the call and put prices. The median of their forward-implied prices is taken as the forward index level for the corresponding tenor. The risk-free rates are obtained by interpolating the LIBOR rates to a particular tenor. Finally, given the calculated forward prices and moneyness levels, we

Figure 2: Stacked bar chart of time-to-maturity frequency



Note: This figure plots a stacked bar chart for the frequency of tenors in S&P 500 index options. The sample contains daily option data from 1 May 2017 to 1 April 2021. The indicators 'spx' and 'spxw' correspond to AM-settled 'standard' and PM-settled 'weeklys' and end-of-month contracts, respectively.

retain only out-of-the-money (OTM) options for further exploitation. The descriptive statistics of the S&P 500 index options data sample is provided in Table 3. We observe that the largest portion of trading volume is due to trades of OTM contracts and options with time-to-maturity less than 60 calendar days. Figure 2 plots frequency of tenors up to 70 calendar days.

5.2 CCF-spanning option portfolios

The construction of the CCF-spanning option portfolios requires reliable option slices with wide coverage of strikes. Given that most of the trading volume is concentrated in the option contracts with time-to-maturity less than 60 days, in our empirical analysis we use several short-dated options with expiration period no more than 2 months. In particular, on each trading day we keep six tenors closest to 8, 15, 22, 29, 36 and 61 days[9] from the left with the largest trading volume and number of quoted OTM option contracts. Specifically, starting with option slices closest to the indicated tenors, we compare them with every next further distant from the left option slice and prefer the next one if it has larger trading volume and larger number of quoted contracts for OTM options. Table 4 provides the descriptive statistics for each of the six selected tenors over the considered time span. We notice the wide coverage of strikes since the average minimum put and call prices are close to the tick size of \$0.05, especially for very short-dated options. We also notice that in the selected option sample, each option slice at each trading day contains at least 55 different quoted contracts. Therefore, no additional filters on the minimum number of contracts are imposed. In total, we have 978 trading days, with six

[9]The first five of these tenors are the most representative in the sample, see Figure 2.

Table 4: Descriptive statistics for the selected sample

Number	1	2	3	4	5	6	Total
avg. tenor	6.97	12.10	18.95	26.08	33.95	53.67	25.29
avg. min put	0.08	0.09	0.10	0.11	0.11	0.14	0.10
avg. min call	0.09	0.09	0.10	0.11	0.12	0.14	0.11
avg. max price	22.48	31.65	40.24	47.56	54.83	71.05	44.64
avg. # options	133.75	183.64	204.47	206.90	216.50	238.84	197.35
avg. min K/F	0.79	0.69	0.60	0.54	0.49	0.42	0.59
avg. max K/F	1.06	1.09	1.12	1.14	1.16	1.21	1.13
avg. atm bsiv	0.147	0.151	0.151	0.152	0.154	0.158	0.152

Note: Descriptive statistics for the selected sample data of options on S&P 500 stock market index. The sample contains daily option data from 1 May 2017 to 1 April 2021. For each trading day, we select six option tenors closest to 8, 15, 22, 29, 36 and 61 days from the left and with the largest trading volume and number of quoted OTM option contracts. The table provides the descriptive statistics for each of the six tenors over the sample.

different tenors at each of them, resulting in 1,158,059 number of contracts in the sample.

The inputs of our estimation procedure are option portfolios representing CCFs rather than BSIVs, as is common in the literature. Therefore, we pay careful attention to the construction of the option-implied CCF. As discussed in Section 3.2, we use the Riemann sum approximation to obtain a computable counterpart of the CCF spanning (3). However, in order to reduce the discretization and truncation errors, we employ an interpolation-extrapolation technique. In particular, we interpolate option prices using cubic splines with carefully selected knot sequences and extrapolate beyond the observable range of strike prices using a parametrization that satisfies the asymptotic result of Lee (2004). The details of the interpolation-extrapolation scheme are provided in Appendix C.1.

The calculation of the option-implied CCF uses the Riemann sum approximation (14) applied to the result of the interpolation-extrapolation scheme. The construction is conducted for each day and for each maturity separately. In particular, for equation (14) we set $\Delta m = 0.0001$ with sufficiently wide range of log-moneyness between $\underline{m} = -6$ and $\overline{m} = 2$.

To conclude this section we emphasize again that the option prices or a monotonic transformation thereof are not used as inputs in our developed estimation procedure as is common in many existing approaches. Instead, we use the option portfolios that replicate the CCF of the log returns. Furthermore, unlike in many other papers, our option dataset is daily and utilizes the information from the short-dated options with maturities between two days and two months.

6 Empirical Applications

After having constructed the dataset of the option-implied CCF for the S&P 500 index options, we now turn to the empirical application of our estimation approach.

6.1 SVCDEJ

We start with estimating the SVCDEJ model, specified in Section 4.1, (28)–(29). Table 5 provides the results of the parameter estimates for the SVCDEJ model based on the range of arguments $u = 1, \ldots, 20$ and $\bar{s} = 10^{-7}$. In agreement with the Monte Carlo design in the previous section, and following results in Aït-Sahalia et al. (2015), we fix the parameter $p^- = -0.7$ to avoid identification issues. Standard errors are calculated using the sandwich form covariance matrix.

Table 5: SVCDEJ estimation results

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

Note: This table provides the parameter estimates and standard errors for the SVCDEJ model. The model is estimated based on $u = 1, \ldots, 20$ and $\bar{s} = 10^{-7}$ with fixed $p^- = 0.7$.

The parameter estimates in Table 5 are consistent with the literature and intuitive. For instance, Andersen et al. (2015a) found the mean jump sizes to be 1.71% and -5.33% for positive and negative jumps in their three-factor model specification. They, however, use only the Wednesday options with different time period from 1996 to 2010.

We note that the leverage parameter is estimated close to the boundary. Although the empirical literature suggests that the correlation ρ is negative and large in absolute value, almost perfectly correlated diffusive parts in the return and volatility are very unlikely. This might be possibly due to exploitation of the short-dated options that typically exhibit steeper implied volatility slopes. Andersen et al. (2017) also found the correlation to be close to -1 in their dataset dominated by the option contracts with maturities less than 2 months.

We also note that the estimated measurement standard error $\sigma_\varkappa$ corresponds to a standard deviation of about 25% of the implied volatility. This is larger than what one might expect of measurement errors in option prices only, and may be interpreted to indicate missing state variables or functional form misspecification. In agreement with this, some of the extensions of the model considered below and in Appendix D show a lower estimate of $\sigma_\varkappa$.

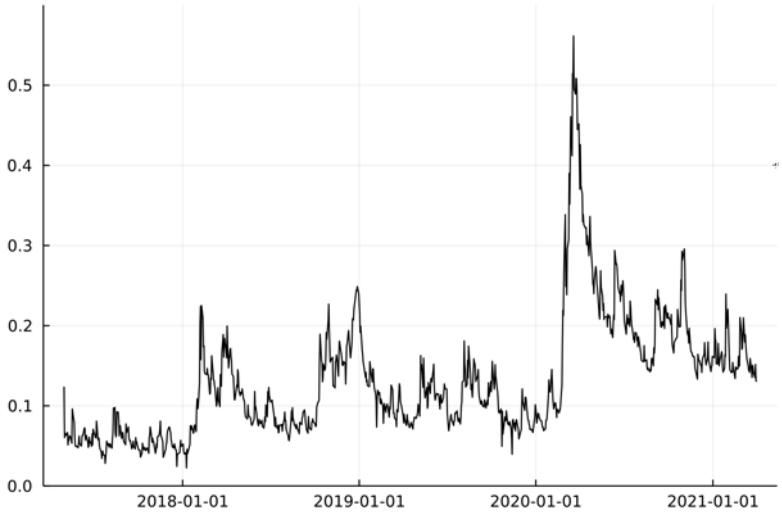
Figure 3 plots the square root of the filtered variance $\widehat{x}_{t+1|t}$ given the parameter estimates of the SVCDEJ model. As expected, the filtered volatility exhibits a relatively stable volatility regime prior to 2020 and jumps up in March 2020 at the beginning of the Covid-19 pandemic.

6.2 SVCDEJ with external factor

Now we turn to the specification with embedded external factors. The motivation comes from the fact that in some situations we might have an understanding of possible drivers of the risks in the market, and thus, we would like to quantify their effect.

An example is the recent Covid-19 crisis. The Covid-19 pandemic has dramatically affected our lives. It has also had a tremendous impact on the world's economy and financial markets. The beginning of the pandemic was especially associated with a spike in uncertainty. This

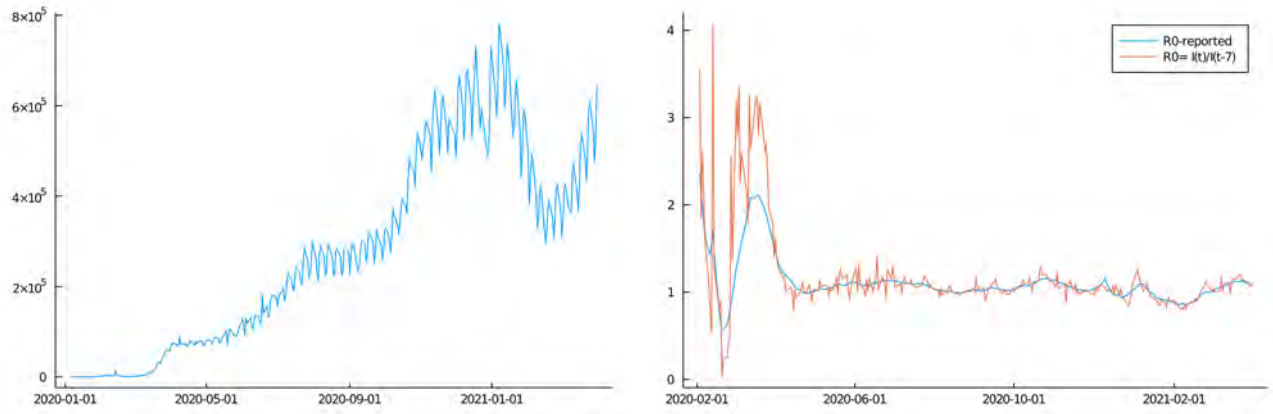
Figure 3: SVCDEJ filtered volatility



Note: This figure plots the filtered volatility (square root of the filtered state $\hat{x}_{t+1|t}$) given the parameter estimates of the SVCDEJ model. The Kalman filter recursions and the model parameter estimates are used for the filtering.

uncertainty surrounded many aspects of the beginning of the Covid-19 pandemic: the contagiousness and lethality of the virus, the time needed to develop vaccines, the effectiveness of measures, the work-from-home policies, travel bans, etc. In this paper we analyse the impact of the Covid-19 pandemic on the stock market through the lens of option prices. In particular, we consider how the spread of the virus affected the likelihood of jump events and volatility in the U.S. stock market.

Figure 4: Covid-19 world daily cases and reproduction numbers.



(a) Daily cases (b) Reproduction numbers

Note: This figure plots the world daily Covid-19 cases and the two reproduction numbers: one is taken from the web-site 'Our world in Data' (uses methodology of Arroyo-Marioli, Bullano, Kucinskas, and Rondón-Moreno (2021)); the second is calculated as the ratio I_t/I_{t-7}, where I_t is number of infected people in day t and 7 is the reported serial interval for Covid-19.

Figure 4, Panel (a), plots the daily cases of the Covid-19 infections around the world obtained from the World Health Organization (WHO). As we can see from the figure, the reported number of daily cases does not represent well the contagiousness of the virus. Therefore, Panel (b) of

Figure 4 provides the reproduction number dynamics taken from the web-site 'Our world in Data' and calculated as the ratio $R_t = I_t/I_{t-7}$, where I_t is number of infected people in day t and 7 is the reported serial interval for Covid-19. The former is based on the methodology of Arroyo-Marioli et al. (2021) and is already smoothed in time.[10] The latter is unsmoothed and based on the assumption that the serial interval is 7 days, which is consistent with the recent epidemiology literature (see, e.g., Maier & Brockmann, 2020, Prem et al., 2020, Flaxman et al., 2020, Arroyo-Marioli et al., 2021). We will use the latter as the reproduction number for our analysis below.

To quantify the effect of the Covid-19 pandemic on the financial market, we simply embed the reproduction numbers to the (time-varying) levels of the volatility and jump intensity processes as described in Section 4.2. In other words, given that the reproduction number is relatively persistent process, we will treat it as a deterministic process when pricing the options. In a similar way, the risk-free rate and dividend yields are often assumed to be deterministic in the option pricing literature. This allows us to be agnostic about the parametric dynamics of the reproduction number. Furthermore, given the short-dated options in the consideration, the errors due to the deterministic assumption will likely to be negligible.[11]

Table 6: SVCDEJ estimation results with reproduction number as external factor

	σ	κ	$\bar{v}$	ρ	δ	η^+	η^-	μ_v	γ	q	$\sigma_\varkappa$
$\hat{\theta}$	0.5678	11.549	0.0140	-1.000	130.12	0.0181	0.0413	0.0667	2.64	0.0003	0.245
s.e.	0.0176	0.646	0.0007	0.023	11.93	0.0007	0.0012	0.0029	0.25	0.0001	0.004

Note: This table provides the parameter estimates and standard errors for the SVCDEJ model with the reproduction number as external factor. The model is estimated based on $u = 1, \ldots, 20$ and $\bar{s} = 10^{-7}$ with fixed $p^- = 0.7$.

Table 6 provides the results for the SVCDEJ model with the reproduction number as an exogenous factor. The results suggest that the reproduction number dynamics has no significant effect on the total volatility and the likelihood of the positive jumps. However, a one unit increase in the reproduction number leads to an increase of the intensity of negative jumps by 2.64. In other words, the reproduction number partly explains the source of the negative jump intensity dynamics. Figure 5 illustrates the dynamics of the negative jump intensity with and without the added effect of the reproduction number R_t.

Table 7: SVCDEJ estimation results with EPU as external factor

	σ	κ	$\bar{v}$	ρ	δ	η^+	η^-	μ_v	γ	q	$\sigma_\varkappa$
$\hat{\theta}$	0.4887	10.254	0.0116	-1.000	223.94	0.0144	0.0410	0.0462	0.00	0.0369	0.249
s.e.	0.0186	0.783	0.0006	0.037	7.96	0.0011	0.0008	0.0024	0.13	0.0027	0.003

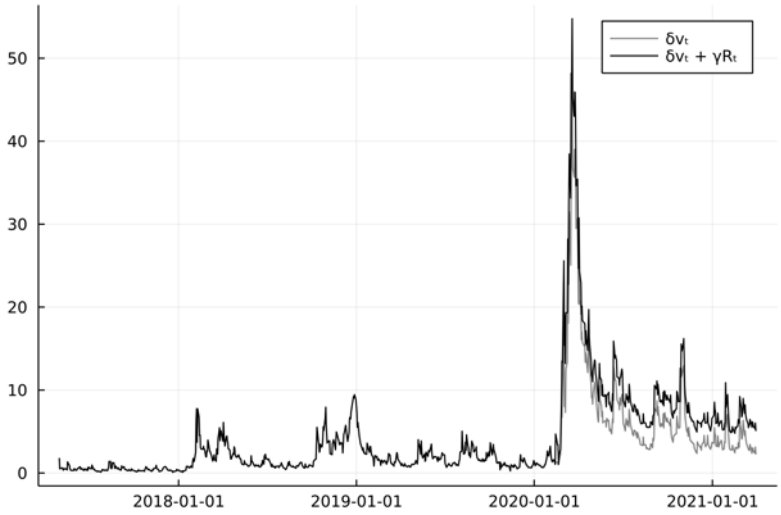
Note: This table provides the parameter estimates and standard errors for the SVCDEJ model with the EPU as external factor. The model is estimated based on $u = 1, \ldots, 20$ and $\bar{s} = 10^{-7}$ with fixed $p^- = 0.7$.

It is also possible to investigate the contribution of some other external factors on the

[10] In fact, Arroyo-Marioli et al. (2021) use the Kalman smoother.

[11] Similarly, Andersen et al. (2017) and Boswijk et al. (2021) consider approximation of the return process with 'freezed' spot volatility when estimating their option pricing models with short-dated options.

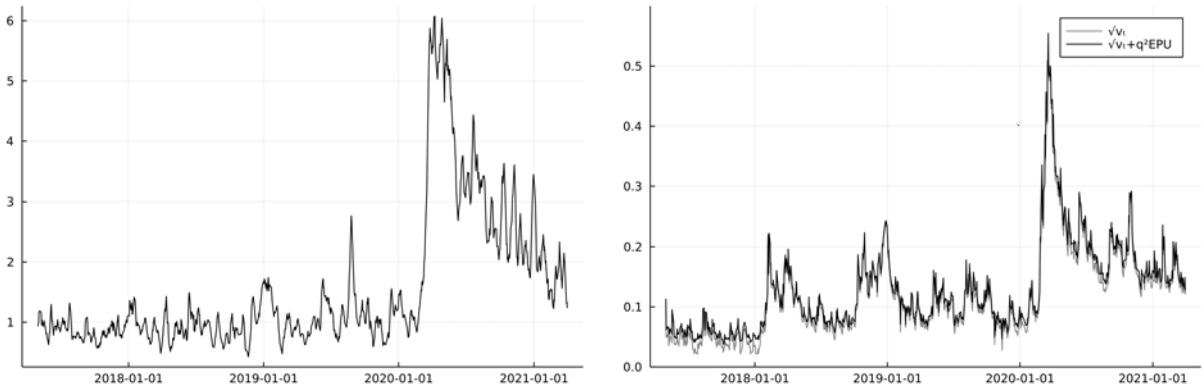
Figure 5: SVCDEJ jump intensity



Note: This figure plots the filtered jump intensity given the parameter estimates of the SVCDEJ model with reproduction number dynamics.

diffusive volatility and jump intensity. For instance, in Table 7 we provide the estimation results for the SVCDEJ model with the Economic Policy Uncertainty (EPU) index as the external factor. The EPU index, developed by Baker, Bloom, and Davis (2016), reflects the policy-related economic uncertainty based on newspaper coverage frequency. The estimation results suggests that, unlike the reproduction number, the EPU index shows no effect on the jump intensity process, but exhibits larger contribution to the diffusive volatility of the model; see Figure 6.

Figure 6: EPU index and filtered diffusive volatility



(a) EPU index (b) Total and filtered diffusive volatility

Note: Panel (a) plots the EPU index of Baker et al. (2016). The data is divided by 100 and a one-week moving average is applied. Panel (b) plots the total and filtered diffusive volatility from the estimation result of the SVCDEJ with the EPU index. The total volatility is given as $\sqrt{v_t + q^2 h_t}$, where h_t is the external factor.

7 Conclusion

In this paper, we have proposed a novel state filtering and parameter estimation procedure for option pricing models that belong to the affine jump-diffusion class. Our procedure utilizes the log of the option-implied and model-free conditional characteristic function as well as the model-implied conditional log-characteristic function, which is functionally affine in the model's state vector. We have provided a corresponding linear state space representation for the considered class of option pricing models, which allows us to exploit suitably adapted collapsed Kalman filtering techniques. Our novel estimation procedure is fast and easy to implement, circumventing the typical computational burden when working with option pricing models. We have demonstrated the applicability of our procedure in two empirical illustrations that analyze S&P 500 index options and the impact of exogenous variables capturing Covid-19 reproduction and economic policy uncertainty data.

Although we have focused on Gaussian QML estimation based on the Kalman filter, which delivers good results in our Monte Carlo simulations, the same state space formulation can also be analyzed by more refined methods such as the particle filter, see Johannes et al. (2009), Christoffersen, Dorion, Jacobs, and Karoui (2014) and Bardgett et al. (2019). Such methods could exploit the non-Gaussianity and heteroskedasticity in the data to obtain more efficient estimates, at the cost of some increased computational complexity. We note, however, that such extensions would still not require option price evaluation by the FFT of COS methods, and thus retain an important advantage of our approach.

Our proposed procedure allows for identification and estimation of factor risk premium parameters, by combining the risk-neutral parameters entering the measurement equation with the objective parameters entering the transition equation. Monte Carlo simulation results suggest that option price data are not very informative about such risk premia, which is why we have concentrated on the case where the objective and risk-neutral measures coincide. Fortunately, the simulation results suggest that inference on the risk-neutral parameters is quite robust with respect to deviations from this assumption. For more accurate inference on (volatility) risk premium parameters, it may be possible to combine the information in daily option prices as considered in this paper with realised measures based on high-frequency returns on the underlying. We will explore this possibility in future research.

Appendix A Proofs

In this appendix, we provide the proof of Proposition 1. First, we state and prove some preliminary results.

A.1 Preliminary results

We start by formally defining the measurement errors in the CCF approximation. Under the observation error structure specified in Assumption 2 and the CCF approximation given by equation (14), the total measurement error in the option-spanned CCF is defined by its three components:

$$
\begin{aligned}
\zeta_t^{\phi}(u,\tau) &:= \widehat{\phi}_t(u,\tau) - \phi_t(u,\tau) \\
&= -u_t \sum_{j=2}^{n} e^{(\mathrm{i}u-1)m_j} \cdot \widehat{O}_t(\tau,m_j)\Delta m_j + u_t \int_{-\infty}^{\infty} e^{(\mathrm{i}u-1)m} \cdot O_t(\tau,m)\mathrm{d}m \\
&= -u_t \left[\sum_{j=2}^{n} e^{(\mathrm{i}u-1)m_j} \cdot O_t(\tau,m_j)\Delta m_j + \sum_{j=2}^{n} e^{(\mathrm{i}u-1)m_j} \cdot \zeta_t(\tau,m_j)\Delta m_j \right] \\
&\quad + u_t \int_{-\infty}^{\infty} e^{(\mathrm{i}u-1)m} \cdot O_t(\tau,m)\mathrm{d}m \\
&= \underbrace{-u_t \sum_{j=2}^{n} e^{(\mathrm{i}u-1)m_j} \cdot \zeta_t(\tau,m_j)\Delta m_j}_{=:\zeta_t^{(1)}(u,\tau)} \\
&\quad + \underbrace{u_t \int_{-\infty}^{m_1} e^{(\mathrm{i}u-1)m} \cdot O_t(\tau,m)\mathrm{d}m + u_t \int_{m_n}^{\infty} e^{(\mathrm{i}u-1)m} \cdot O_t(\tau,m)\mathrm{d}m}_{=:\zeta_t^{(2)}(u,\tau)} \\
&\quad + \underbrace{u_t \sum_{j=2}^{n} \int_{m_{j-1}}^{m_j} \left[e^{(\mathrm{i}u-1)m} \cdot O_t(\tau,m) - e^{(\mathrm{i}u-1)m_j} \cdot O_t(\tau,m_j) \right] \mathrm{d}m}_{=:\zeta_t^{(3)}(u,\tau)} \\
&= \zeta_t^{(1)}(u,\tau) + \zeta_t^{(2)}(u,\tau) + \zeta_t^{(3)}(u,\tau).
\end{aligned}
$$

The error terms $\zeta_t^{(1)}(u,\tau)$, $\zeta_t^{(2)}(u,\tau)$ and $\zeta_t^{(3)}(u,\tau)$ represent observation, truncation and discretization errors, respectively. In order to characterize the asymptotic orders of these errors, we make use of the following auxiliary result:

Lemma 1 *Let $f_{t+\tau} = \frac{F_{t+\tau}}{F_t}$ be the normalized to value at time t futures price for $\tau > 0$. For each $p > 0$ we have for all $m > 0$ the call price bounds*

$$
\frac{O_t(\tau,m)}{F_t} \leq \left(\frac{p}{p+1}\right)^p \frac{e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[f_{t+\tau}^{p+1}|\mathcal{F}_t]}{p+1} e^{-pm}. \tag{A.1}
$$

For each $q > 0$ we have for all $m < 0$ the put price bounds

$$
\frac{O_t(\tau,m)}{F_t} \leq \left(\frac{q}{q+1}\right)^q \frac{e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[f_{t+\tau}^{-q}|\mathcal{F}_t]}{q+1} e^{(1+q)m}. \tag{A.2}
$$

Proof: The result is the straightforward adaptation of Theorem 2.1 in Lee (2004). □

Lemma 1 relates the moments of the underlying process and of its reciprocal to the bounds of the option prices. Similar to Qin and Todorov (2019) and Todorov (2019), we assume the existence of at least the second order moment of the underlying process and of its reciprocal, formulated in Assumption 3. If higher moments exist, then we can obtain even tighter bounds for the remainder term in Proposition 1 due to (A.1) and (A.2).

The following lemma establishes the order of magnitude of the truncation and discretization errors under the joint asymptotic scheme, expressed with respect to the number of option prices n with fixed maturity. As in the main text, we denote the smallest and largest log-moneyness as $\underline{m} = \min_{1\leq j\leq n} m_j$ and $\overline{m} = \max_{1\leq j\leq n} m_j$, and the corresponding strike prices as $\underline{K}$ and $\overline{K}$. In the proofs we denote with C_t an $\mathcal{F}_t$-adapted random variable which does not depend on m and might change from line to line.

Lemma 2 *Suppose* $\mathbb{E}^{\mathbb{Q}}[F_{t+\tau}^{p+1}|\mathcal{F}_t] < \infty$ *and* $\mathbb{E}^{\mathbb{Q}}[F_{t+\tau}^{-q}|\mathcal{F}_t] < \infty$ *for some* $p > 0$ *and* $q > 0$*, Assumption 3(ii) holds, and in addition* $\underline{K} \asymp n^{-\underline{\alpha}}$ *and* $\overline{K} \asymp n^{\overline{\alpha}}$ *with* $\underline{\alpha} > 0$ *and* $\overline{\alpha} > 0$*. Then for* $n \to \infty$ *we have*

$$\zeta_t^{(2)}(u,\tau) = \mathcal{O}_p\left(n^{-(q\underline{\alpha}\wedge(1+p)\overline{\alpha})}\right), \tag{A.3}$$

and

$$\zeta_t^{(3)}(u,\tau) = \mathcal{O}_p\left(\frac{\log n}{n^{1+q\underline{\alpha}\wedge(p+1)\overline{\alpha}}}\right) \tag{A.4}$$

for a fixed $u \in \mathcal{U}$ *and* $\tau > 0$*.*

Proof: We start with the truncation error. For $m_n > 0$ and $m_1 < 0$ and using Lemma 1 we can bound the upper and lower truncation parts as follows:

$$\begin{aligned}
\left|\frac{1}{F_t}\int_{m_n}^{\infty} e^{(\mathrm{i}u-1)m}\cdot O_t(\tau,m)\mathrm{d}m\right| &\leq \int_{m_n}^{\infty}\left|e^{(\mathrm{i}u-1)m}\right|\cdot\left|\frac{O_t(\tau,m)}{F_t}\right|\mathrm{d}m \leq C_t e^{-(1+p)m_n},\\
\left|\frac{1}{F_t}\int_{-\infty}^{m_1} e^{(\mathrm{i}u-1)m}\cdot O_t(\tau,m)\mathrm{d}m\right| &\leq \int_{-\infty}^{m_1}\left|e^{(\mathrm{i}u-1)m}\right|\cdot\left|\frac{O_t(\tau,m)}{F_t}\right|\mathrm{d}m \leq C_t e^{qm_1},
\end{aligned}$$

where, as mentioned before, C_t does not depend on m and might change from line to line. Therefore, for $\underline{m} \to -\infty$ and $\overline{m} \to \infty$ we have

$$\begin{aligned}
\zeta_t^{(2)}(u,\tau) &= u_t\int_{-\infty}^{\underline{m}} e^{(\mathrm{i}u-1)m}\cdot O_t(\tau,m)\mathrm{d}m + u_t\int_{\overline{m}}^{\infty} e^{(\mathrm{i}u-1)m}\cdot O_t(\tau,m)\mathrm{d}m\\
&= \mathcal{O}_p\left(e^{-q|\underline{m}|}\right) + \mathcal{O}_p\left(e^{-(1+p)|\overline{m}|}\right) = \mathcal{O}_p\left(e^{-(q|\underline{m}|\wedge(1+p)|\overline{m}|)}\right) = \mathcal{O}_p\left(n^{-(q\underline{\alpha}\wedge(1+p)\overline{\alpha})}\right).
\end{aligned}$$

For the discretization errors we use the following decomposition:

$$\begin{aligned}
&\int_{m_{j-1}}^{m_j}\left[e^{(\mathrm{i}u-1)m}\cdot O_t(\tau,m) - e^{(\mathrm{i}u-1)m_j}\cdot O_t(\tau,m_j)\right]\mathrm{d}m\\
&= \int_{m_{j-1}}^{m_j}\left[\left(e^{(\mathrm{i}u-1)m} - e^{(\mathrm{i}u-1)m_j}\right)\cdot O_t(\tau,m_j) + e^{(\mathrm{i}u-1)m}\left(O_t(\tau,m) - O_t(\tau,m_j)\right)\right]\mathrm{d}m.
\end{aligned}$$

By applying the mean value theorem we have

$$\left|e^{(\mathrm{i}u-1)m}-e^{(\mathrm{i}u-1)m_j}\right|\leq|\mathrm{i}u-1|\left|e^{(\mathrm{i}u-1)\tilde{m}}\right|\Delta m_j\leq e^{-m_{j-1}}(|u|\vee 1)\Delta m_j,$$

and

$$\begin{aligned}\left|e^{(\mathrm{i}u-1)m}\left(O_t(\tau,m)-O_t(\tau,m_j)\right)\right|&\leq e^{-m_{j-1}}\left|\frac{\partial O_t(\tau,m)}{\partial m}\Big|_{m=\tilde{m}}\right|\Delta m_j\\&\leq e^{-m_{j-1}}\left|\frac{\partial O_t(\tau,m)}{\partial K}\Big|_{K=\tilde{K}}\right|e^{\tilde{m}}F_t\Delta m_j\\&\leq\left|\frac{\partial O_t(\tau,m)}{\partial K}\Big|_{\tilde{K}}\right|e^{\Delta m_j}F_t\Delta m_j,\end{aligned}$$

where $\tilde{m}=\log\frac{\tilde{K}}{F_t}$ is between m and m_j.

Lemma 1 implies that for all m

$$\frac{O_t(\tau,m)}{F_t}\leq C_t e^{-(pm\vee-(1+q)m)}.$$

Furthermore, the derivative with respect to the strike price is the risk neutral distribution or survival functions, which can be bounded using the Markov inequality. In particular, for $m>0$:

$$\begin{aligned}\left|\frac{\partial O_t(\tau,m)}{\partial K}\right|&=e^{-r\tau}\mathbb{Q}(F_{t+\tau}>K)=e^{-r\tau}\mathbb{Q}\left(f_{t+\tau}>e^m\right)\\&\leq e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[f_{t+\tau}^{p+1}|\mathcal{F}_t]e^{-(p+1)m},\end{aligned}$$

and for $m<0$

$$\begin{aligned}\left|\frac{\partial O_t(\tau,m)}{\partial K}\right|&=e^{-r\tau}\mathbb{Q}(F_{t+\tau}<K)=e^{-r\tau}\mathbb{Q}\left(f_{t+\tau}^{-1}>e^{-m}\right)\\&\leq e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[f_{t+\tau}^{-q}|\mathcal{F}_t]e^{qm}.\end{aligned}$$

Therefore

$$\left|\frac{\partial O_t(\tau,m)}{\partial K}\right|\leq C_t e^{-((p+1)m\vee-qm)}.$$

Combining all these inequalities together we obtain

$$\begin{aligned}&\left|\frac{1}{F_t}\int_{m_{j-1}}^{m_j}\left[e^{(\mathrm{i}u-1)m}\cdot O_t(\tau,m)-e^{(\mathrm{i}u-1)m_j}\cdot O_t(\tau,m_j)\right]\mathrm{d}m\right|\\&=\left|\frac{1}{F_t}\int_{m_{j-1}}^{m_j}\left[\left(e^{(\mathrm{i}u-1)m}-e^{(\mathrm{i}u-1)m_j}\right)\cdot O_t(\tau,m_j)+e^{(\mathrm{i}u-1)m}\left(O_t(\tau,m)-O_t(\tau,m_j)\right)\right]\mathrm{d}m\right|\\&\leq\left(C_t(|u|\vee 1)\Delta m_j e^{-m_{j-1}}e^{-[pm_{j-1}\vee-(1+q)m_{j-1}]}+C_t e^{\Delta m}e^{-[(p+1)m_{j-1}\vee-qm_{j-1}]}\Delta m_j\right)\Delta m_j\\&\leq C_t e^{-[(p+1)m_{j-1}\vee-qm_{j-1}]}(\Delta m_j)^2.\end{aligned}$$

Then, for fixed $\overline{m}$ and $\underline{m}$, and $\Delta m\to 0$, we have

$$\zeta_t^{(3)}(u,\tau)=u_t\sum_{j=2}^{n}\int_{m_{j-1}}^{m_j}\left[e^{(\mathrm{i}u-1)m}\cdot O_t(\tau,m)-e^{(\mathrm{i}u-1)m_j}\cdot O_t(\tau,m_j)\right]\mathrm{d}m=\mathcal{O}_p(\Delta m).\tag{A.5}$$

The result (A.5) needs to be adapted for the joint asymptotic scheme, where $\overline{m}$, $\underline{m}$ and Δm all depend on $n \to \infty$. First,

$$\begin{aligned}
&\Big|\sum_{j=2}^{n}\frac{1}{F_t}\int_{m_{j-1}}^{m_j}\Big[e^{(\mathrm{i}u-1)m}\cdot O_t(\tau,m)-e^{(\mathrm{i}u-1)m_j}\cdot O_t(\tau,m_j)\Big]\,\mathrm{d}m\Big| \\
&\qquad\leq C_t\sum_{j=2}^{n}e^{-[(p+1)m_{j-1}\vee -qm_{j-1}]}(\Delta m_j)^2 \\
&\qquad\leq C_t\Delta m\sum_{j=2}^{n}e^{-[(p+1)m_{j-1}\vee -qm_{j-1}]}\Delta m_j.
\end{aligned}$$

The sum on the right hand side of the inequality is the Riemann approximation that converges to the following integral:

$$\begin{aligned}
\sum_{j=2}^{n}e^{-[(p+1)m_{j-1}\vee -qm_{j-1}]}\Delta m_j \longrightarrow \int_{\underline{m}}^{\overline{m}}e^{-[(p+1)m\vee -qm]}\mathrm{d}m &= \int_{\underline{m}}^{0}e^{qm}\mathrm{d}m+\int_{0}^{\overline{m}}e^{-(p+1)m}\mathrm{d}m \\
&= \mathcal{O}_p\left(n^{-q\underline{\alpha}}\right)+\mathcal{O}_p\left(n^{-(p+1)\overline{\alpha}}\right) \\
&= \mathcal{O}_p\left(n^{-(q\underline{\alpha}\wedge(p+1)\overline{\alpha})}\right).
\end{aligned}$$

Given Assumption 3*(ii)* on the log-moneyness grid, we can bound Δm as

$$\frac{\overline{m}-\underline{m}}{\eta n}\geq\Delta m\geq\frac{\overline{m}-\underline{m}}{n}.$$

Hence $\Delta m=\mathcal{O}_p\left(\frac{\log n}{n}\right)$. Then the order of the discretization error under the joint asymptotic scheme is

$$\zeta_t^{(3)}(u,\tau)=\mathcal{O}_p\left(n^{-(q\underline{\alpha}\wedge(p+1)\overline{\alpha})}\right)\mathcal{O}_p\left(\frac{\log n}{n}\right)=\mathcal{O}_p\left(\frac{\log n}{n^{1+q\underline{\alpha}\wedge(p+1)\overline{\alpha}}}\right).$$

□

A.2 Proof of Proposition 1

Using Lemma 2, Assumption 3 on the moments of the underlying process and observation error Assumption 2, we can decompose the CCF approximation as

$$\begin{aligned}
\widehat{\phi}_t(u,\tau)-\phi_t(u,\tau) &= \zeta_t^{(1)}(u,\tau)+\mathcal{O}_p\left(\frac{\log n}{n^{1+2(\underline{\alpha}\wedge\overline{\alpha})}}\vee n^{-2(\underline{\alpha}\wedge\overline{\alpha})}\right) &\text{(A.6)}\\
&= \zeta_t^{(1)}(u,\tau)+\mathcal{O}_p\left(n^{-2(\underline{\alpha}\wedge\overline{\alpha})}\right) &\text{(A.7)}
\end{aligned}$$

with

$$\zeta_t^{(1)}(u,\tau)=-u_t\sum_{j=2}^{n}e^{(\mathrm{i}u-1)m_j}\cdot\zeta_t(\tau,m_j)\Delta m_j. \tag{A.8}$$

We now show that $\zeta_t^{(1)}(u,\tau)=\mathcal{O}_p\left(\sqrt{\frac{\log n}{n}}\right)$. In fact, the standard deviation of the observation errors is proportional to the Black-Scholes vega, which decreases with $|m|\to\infty$. More

specifically, the vega is given by

$$\nu_t(\tau, m) = F_t\sqrt{\tau}N'(d_+)$$
$$d_+ = -mv^{-1/2} + \frac{1}{2}v^{1/2},$$

where $N'(x)$ is the standard normal pdf and $v(m) := \kappa^2(\tau, m)\tau$. Hence

$$\nu_t^2(\tau, m) = F_t^2\tau\frac{1}{2\pi}e^{-d_+^2} = F_t^2\tau\frac{1}{2\pi}e^{-(v^{-1}m^2 - m + \frac{1}{4}v)}.$$

Therefore, given Assumption 2, we get

$$\begin{aligned}
\mathbb{E}\left[\left|\zeta_t^{(1)}(u,\tau)\right|^2\middle|\mathcal{F}_t\right] &\le |u_t|^2\sum_{j=2}^{n} e^{-2m_j}\cdot\mathbb{E}\left[\zeta_t(\tau, m_j)^2|\mathcal{F}_t\right](\Delta m_j)^2 \\
&\le |u_t|^2\sum_{j=2}^{n} e^{-2m_j}\cdot\sigma_{\varkappa}^2\kappa^2(\tau, m_j)F_t^2\tau\frac{1}{2\pi}e^{-(v_j^{-1}m_j^2 - m_j + \frac{1}{4}v_j)}(\Delta m_j)^2 \\
&\le C_t\Delta m\sum_{j=2}^{n} v_j e^{-v_j^{-1}m_j^2 - m_j - \frac{1}{4}v_j}\Delta m_j \\
&\le C_t\Delta m\sum_{j=2}^{n} v_j e^{-d_-^2(m_j)}\Delta m_j,
\end{aligned}$$

where $d_-(m) := -mv^{-1/2}(m) - \frac{1}{2}v^{1/2}(m)$. Then the right-hand side summation converges to

$$\int_{-\infty}^{\infty} v(m)\exp\left(-d_-^2(m)\right)\mathrm{d}m =: \int_{-\infty}^{\infty} h(m)\mathrm{d}m,$$

provided that the function $h(m)$ is integrable. To show the latter, we focus on the tail behavior since h is continuous and bounded on the definite interval $(\underline{m}, \overline{m})$. For that, we will make use of the asymptotic results of Lee (2004):

$$\limsup_{m\to-\infty}\frac{v(m)}{|m|} = \underline{\beta}^* \text{ with } \underline{\beta}^* \in [0,2] \text{ and } \frac{1}{2\underline{\beta}^*} + \frac{\underline{\beta}^*}{8} - \frac{1}{2} = \sup\{q : \mathbb{E}[F_{t+\tau}^{-q}|\mathcal{F}_t] < \infty\},$$
$$\limsup_{m\to\infty}\frac{v(m)}{|m|} = \overline{\beta}^* \text{ with } \overline{\beta}^* \in [0,2] \text{ and } \frac{1}{2\overline{\beta}^*} + \frac{\overline{\beta}^*}{8} - \frac{1}{2} = \sup\{p : \mathbb{E}[F_{t+\tau}^{1+p}|\mathcal{F}_t] < \infty\}.$$

That is, for $m < 0$, the implied variance $v(m)$ grows at most as fast as $-\underline{\beta}^* m$ with some $\underline{\beta}^*$. Given Assumption 3 on moments of the reciprocal process, we have $\underline{\beta}^* < 1$, which implies that

$$h(m) = v\exp\left(-v^{-1}m^2 - m - \frac{1}{4}v\right) \le -\underline{\beta}^* m\exp\left(\frac{m}{\underline{\beta}^*} - m\right)$$

is integrable on the left tail. The integrability on the right tail is achieved even without the use of moment conditions. Therefore, since the summation converges under the joint asymptotic scheme to the integral above, we have that $\zeta_t^{(1)}(u,\tau) = \mathcal{O}_p\left(\sqrt{\Delta m}\right) = \mathcal{O}_p\left(\sqrt{\frac{\log n}{n}}\right)$.

From Assumption 2 it also follows that $\mathbb{E}[\zeta_t^{(1)}(u,\tau)|\mathcal{F}_t] = 0$, while the discretization and truncation errors $\zeta_t^{(2)}(u,\tau)$ and $\zeta_t^{(3)}(u,\tau)$ are $\mathcal{F}_t$-measurable. Therefore, the covariance and the

pseudo-covariance terms of the CCF approximation are given by the second moments of the observation errors $\zeta_t^{(1)}(u,\tau)$, that is,

$$\begin{aligned}
\mathrm{Cov}(\zeta_t^{\phi}(u_i,\tau),\zeta_t^{\phi}(u_j,\tau)) :&= \mathbb{E}\left[\left(\zeta_t^{\phi}(u_i,\tau)-\mathbb{E}[\zeta_t^{\phi}(u_i,\tau)]\right)\overline{\left(\zeta_t^{\phi}(u_j,\tau)-\mathbb{E}[\zeta_t^{\phi}(u_j,\tau)]\right)}\middle|\mathcal{F}_t\right]\\
&= \mathbb{E}\left[\zeta_t^{(1)}(u_i,\tau)\zeta_t^{(1)}(-u_j,\tau)\middle|\mathcal{F}_t\right]\\
&= u_{i,t}\overline{u_{j,t}}\sum_{j=2}^{n} e^{(\mathrm{i}(u_i-u_j)-2)m_j}\cdot\sigma_t^2(\tau,m_j)(\Delta m_j)^2\\
&= \sigma_{\varkappa}^2\cdot\underbrace{u_{i,t}\overline{u_{j,t}}\sum_{j=2}^{n} e^{(\mathrm{i}(u_i-u_j)-2)m_j}\cdot\kappa_t^2(\tau,m_j)\nu_t^2(\tau,m_j)(\Delta m_j)^2}_{=:\gamma_t(u_i,u_j,\tau)}\\
&= \sigma_{\varkappa}^2\cdot\gamma_t(u_i,u_j,\tau),
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{PCov}(\zeta_t^{\phi}(u_i,\tau),\zeta_t^{\phi}(u_j,\tau)) :&= \mathbb{E}\left[\left(\zeta_t^{\phi}(u_i,\tau)-\mathbb{E}[\zeta_t^{\phi}(u_i,\tau)]\right)\left(\zeta_t^{\phi}(u_j,\tau)-\mathbb{E}[\zeta_t^{\phi}(u_j,\tau)]\right)\middle|\mathcal{F}_t\right]\\
&= \mathbb{E}\left[\zeta_t^{(1)}(u_i,\tau)\zeta_t^{(1)}(u_j,\tau)\middle|\mathcal{F}_t\right]\\
&= \sigma_{\varkappa}^2\cdot\underbrace{u_{i,t}u_{j,t}\sum_{j=2}^{n} e^{(\mathrm{i}(u_i+u_j)-2)m_j}\cdot\kappa_t^2(\tau,m_j)\nu_t^2(\tau,m_j)(\Delta m_j)^2}_{=:c_t(u_i,u_j,\tau)}\\
&= \sigma_{\varkappa}^2\cdot c_t(u_i,u_j,\tau),
\end{aligned}$$

for any $u_i,u_j\in\mathcal{U}$, where $\overline{z}$ denotes the complex conjugate of a complex number $z\in\mathbb{C}$. In other words, the covariances of the total measurement errors in the CCF approximation are determined by the properties of the observation errors in option prices only. Note that the terms $\gamma_t(u_i,u_j,\tau)$ and $c_t(u_i,u_j,\tau)$ depend only on option's characteristics such as BSIV, BS vega and moneyness levels. That is, the covariance terms are parametrized using only a single parameter $\sigma_{\varkappa}$ that reflects the variance of the observation errors in option prices.

The measurement equation for the filtering problem is given in terms of the log CCF. Therefore, by applying a Taylor-series expansion to the difference of the logs and using the error decomposition of the CCF approximation we have

$$\begin{aligned}
\xi_t(u,\tau) := \log\widehat{\phi}_t(u,\tau)-\log\phi_t(u,\tau) &= \log\left(1+\frac{\zeta_t^{(1)}(u,\tau)+\zeta_t^{(2)}(u,\tau)+\zeta_t^{(3)}(u,\tau)}{\phi_t(u,\tau)}\right)\\
&= \xi_t^{(1)}(u,\tau)+r_t(u,\tau),
\end{aligned}$$

where

$$\xi_t^{(1)}(u,\tau) := \frac{\zeta_t^{(1)}(u,\tau)}{\phi_t(u,\tau)} = \mathcal{O}_p\left(\sqrt{\frac{\log n}{n}}\right)\ \text{and}\ r_t(u,\tau)=\mathcal{O}_p\left(n^{-2(\underline{\alpha}\wedge\overline{\alpha})}\vee\frac{\log n}{n}\right).$$

We note that the remainder term collects the log-linearization of the truncation and discretization errors and higher-order terms from a Taylor-series expansion.

After stacking each component of the measurement equation (12) as well as the observation errors $\xi_t^{(1)}(u,\tau)$ and remainder $r_t(u,\tau)$ along arguments, real and imaginary parts, and maturity, we get state space measurement equation (19) in Proposition 1.

To derive the covariance matrix of the measurement errors, we first consider the covariance and pseudo-covariance matrices of the stacked vector $\xi_{t,\tau}^{(1)} = \left(\xi_t^{(1)}(u_1,\tau),\ldots,\xi_t^{(1)}(u_q,\tau)\right)'$ for a fixed time t and time-to-maturity τ are given by

$$\begin{aligned}\Gamma_{t,\tau} :&= \mathbb{E}\left[\xi_{t,\tau}^{(1)}\overline{\xi_{t,\tau}^{(1)}}'\right] = \left[\mathbb{E}[\xi_t^{(1)}(u_i,\tau)\xi_t^{(1)}(-u_j,\tau)]\right]_{1\le i,j\le q} \qquad \text{(A.9)}\\ &= \sigma_\varkappa^2\cdot\left[\frac{\gamma_t(u_i,u_j,\tau)}{\phi_t(u_i,\tau)\phi_t(-u_j,\tau)}\right]_{1\le i,j\le q} =: \sigma_\varkappa^2\cdot\widetilde{\Gamma}_{t,\tau},\\ C_{t,\tau} :&= \mathbb{E}\left[\xi_{t,\tau}^{(1)}\xi_{t,\tau}^{(1)\prime}\right] = \left[\mathbb{E}[\xi_t^{(1)}(u_i,\tau)\xi_t^{(1)}(u_j,\tau)]\right]_{1\le i,j\le q} \qquad \text{(A.10)}\\ &= \sigma_\varkappa^2\cdot\left[\frac{c_t(u_i,u_j,\tau)}{\phi_t(u_i,\tau)\phi_t(u_j,\tau)}\right]_{1\le i,j\le q} =: \sigma_\varkappa^2\cdot\widetilde{C}_{t,\tau}.\end{aligned}$$

Since the $\xi_{t,\tau}^{(1)}$ is the complex-valued random vector, the covariance matrix of the stacked real and imaginary parts of $\xi_{t,\tau}^{(1)}$ is of the following form:

$$\begin{aligned}H_{t,\tau} = \mathrm{Var}\left[\begin{pmatrix}\Re(\xi_{t,\tau}^{(1)})\\ \Im(\xi_{t,\tau}^{(1)})\end{pmatrix}\right] &= \begin{pmatrix}\frac{1}{2}\Re(\Gamma_{t,\tau}+C_{t,\tau}) & \frac{1}{2}\Im(-\Gamma_{t,\tau}+C_{t,\tau})\\ \frac{1}{2}\Im(\Gamma_{t,\tau}+C_{t,\tau}) & \frac{1}{2}\Re(\Gamma_{t,\tau}-C_{t,\tau})\end{pmatrix}\\ &= \sigma_\varkappa^2\cdot\begin{pmatrix}\frac{1}{2}\Re(\widetilde{\Gamma}_{t,\tau}+\widetilde{C}_{t,\tau}) & \frac{1}{2}\Im(-\widetilde{\Gamma}_{t,\tau}+\widetilde{C}_{t,\tau})\\ \frac{1}{2}\Im(\widetilde{\Gamma}_{t,\tau}+\widetilde{C}_{t,\tau}) & \frac{1}{2}\Re(\widetilde{\Gamma}_{t,\tau}-\widetilde{C}_{t,\tau})\end{pmatrix}\\ &= \sigma_\varkappa^2\cdot\widetilde{H}_{t,\tau}.\end{aligned}$$

Given Assumption 2, the error terms $\zeta_{t,\tau}^{(1)}$ and $\xi_{t,\tau}^{(1)}$ are conditionally independent along maturity and time. This implies that the stacked along maturities measurement errors ε_t are also conditionally independent, thus $\mathbb{E}[\varepsilon_t\varepsilon_s'] = 0$ for $s \neq t$, and their covariance matrix has a blog-diagonal form: $H_t = \mathrm{blkdiag}\{H_{t,1},\ldots,H_{t,k}\}$.

The disturbance term in the state updating equation is given by $\eta_{t+1} = x_{t+1} - \mathbb{E}[x_{t+1}|\mathcal{F}_t]$. Therefore, η_t is martingale difference sequences, thus $\mathbb{E}[\eta_t\eta_s'] = 0$ for $s \neq t = 1,\ldots,T$.

Since the measurement errors ε_t have zero mean conditional on filtration $\mathcal{F}_t$, we also have that $\mathbb{E}[\varepsilon_t x_t] = 0$. Given the state process is stationary and the initial condition is the unconditional mean, $\mathbb{E}[\varepsilon_t x_1'] = 0$ and $\mathbb{E}[\eta_{t+1}x_1'] = 0$. All of these imply that $\mathbb{E}[\varepsilon_t\eta_s'] = 0$ for all $s,t = 1,\ldots,T$. □

Appendix B Conditional Moments

In this appendix, we show how the conditional mean and variance can be computed for the AJD class of models. In particular, we derive the closed-form expressions for the conditional mean and variance in the univariate case and discuss how these moments can be obtained in the multivariate setting numerically with little computational costs. While semi-closed expressions

are also available in the multivariate setting, they are more cumbersome to work with in practice since they typically require matrix exponential and integrations.

We start with considering the univariate version of the AJD process in (4), denoted as x_t:

$$\mathrm{d}x_t = \mu(x_t)\mathrm{dt} + \sigma(x_t)\mathrm{d}W_t + J_t\mathrm{d}N_t, \tag{B.1}$$

with $\mu(x_t) = k_0 + k_1 x_t$, $\sigma^2(x_t) = h_0 + h_1 x_t$, $\lambda(x_t) = l_0 + l_1 x_t$, where all coefficients are real-valued numbers and the standard Brownian motion W_t and the counting process N_t are univariate processes. The jump size distribution ν on $\mathbb{R}$ is independent of time and of any form of randomness in the model. We further assume that the SDE (B.1) has a strong unique solution and the first two moments are well defined. For more notational details see Section 2.2 and Duffie et al. (2000). Furthermore, for notational simplicity, let us introduce the following terms:

$$\mu_J := \mathbb{E}[J], \quad \mu_{J2} := \mathbb{E}[J^2], \quad g_0 := k_0 + l_0\mu_J, \quad g_1 := k_1 + l_1\mu_J.$$

An infinitesimal generator $\mathcal{D}$ defined at a bounded C^2 function $f\colon D \to \mathbb{R}$ is

$$\mathcal{D}f(x) = f_x(x)\mu(x) + \frac{1}{2}f_{xx}(x)\sigma(x)^2 + \lambda(x)\int_{\mathbb{R}}[f(x+z) - f(x)]\mathrm{d}\nu(z),$$

where f_x and f_{xx} are the bounded first and second derivatives, respectively. Dynkin's formula yields

$$\mathbb{E}[f(x_T)|\mathcal{F}_t] = f(x_t) + \mathbb{E}\left[\int_t^T \mathcal{D}f(x_s)\mathrm{d}s\middle|\mathcal{F}_t\right].$$

Therefore, we can find the conditional moments by applying Dynkin's formula for $f(x) = x$:

$$\begin{aligned}
\mathbb{E}[x_T|\mathcal{F}_t] &= x_t + \mathbb{E}\left[\int_t^T \left(\mu(x_s) + \lambda(x_s)\int_{\mathbb{R}} z\mathrm{d}\nu(z)\right)\mathrm{d}s\middle|\mathcal{F}_t\right] \\
&= x_t + \mathbb{E}\left[\int_t^T (k_0 + k_1 x_s + (l_0 + l_1 x_s)\mu_J)\,\mathrm{d}s\middle|\mathcal{F}_t\right] \\
&= x_t + \int_t^T \left(k_0 + l_0\mu_J + (k_1 + l_1\mu_J)\mathbb{E}[x_s|\mathcal{F}_t]\right)\mathrm{d}s \\
&= x_t + \int_t^T \left(g_0 + g_1\mathbb{E}[x_s|\mathcal{F}_t]\right)\mathrm{d}s,
\end{aligned}$$

where Fubini's theorem is used in the third line. Hence, we can obtain the first conditional moment by solving the ODE:

$$\mathrm{d}\mathbb{E}[x_s|\mathcal{F}_t] = \left(g_0 + g_1\mathbb{E}[x_s|\mathcal{F}_t]\right)\mathrm{d}s,$$

with the initial condition $\mathbb{E}[x_t|\mathcal{F}_t] = x_t$. Therefore, the conditional expectation is given by

$$m_t(T) := \mathbb{E}[x_T|\mathcal{F}_t] = e^{g_1(T-t)}x_t + \frac{g_0}{g_1}\left(e^{g_1(T-t)} - 1\right). \tag{B.2}$$

Now we are interested in deriving the conditional variance:

$$\mathrm{Var}(x_T|\mathcal{F}_t) = \mathbb{E}[(x_T - \mathbb{E}[x_T|\mathcal{F}_t])^2|\mathcal{F}_t].$$

Note that

$$x_T - \mathbb{E}[x_T|\mathcal{F}_t] = \mathbb{E}[x_T|\mathcal{F}_T] - \mathbb{E}[x_T|\mathcal{F}_t] = \int_t^T \mathrm{d}\mathbb{E}[x_T|\mathcal{F}_s] = \int_t^T \mathrm{d}m_s(T).$$

The dynamics of the conditional mean for fixed $T > t$ can be obtained by using Ito's lemma:

$$\begin{aligned}\mathrm{d}m_t(T) &= \left[-g_1 e^{g_1(T-t)} x_t - g_0 e^{g_1(T-t)}\right]\mathrm{dt} + e^{g_1(T-t)}(\mu(x_t)\mathrm{dt} + \sigma(x_t)\mathrm{d}W_t) + e^{g_1(T-t)} J_t \mathrm{d}N_t \\ &= e^{g_1(T-t)}\left[-(l_0 + l_1 x_t)\mu_J \mathrm{dt} + \sigma(x_t)\mathrm{d}W_t + J_t \mathrm{d}N_t\right].\end{aligned}$$

Note that the process $m_t(T)$ for fixed T is a local martingale. Thus, we can use the Ito isometry to get the conditional variance:

$$\begin{aligned}\mathrm{Var}(x_T|\mathcal{F}_t) &= \mathbb{E}\left[\left(\int_t^T \mathrm{d}m_s(T)\right)^2 \Big|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[\int_t^T e^{2g_1(T-s)}\sigma^2(x_s)\mathrm{ds}\Big|\mathcal{F}_t\right] + \mu_{J2}\cdot\mathbb{E}\left[\int_t^T e^{2g_1(T-s)}\lambda(x_s)\mathrm{ds}\Big|\mathcal{F}_t\right] \\ &= \int_t^T e^{2g_1(T-s)}(h_0 + h_1\mathbb{E}[x_s|\mathcal{F}_t])\mathrm{ds} + \mu_{J2}\cdot\int_t^T e^{2g_1(T-s)}(l_0 + l_1\mathbb{E}[x_s|\mathcal{F}_t])\mathrm{ds} \\ &= (h_0 + l_0\mu_{J2})\int_t^T e^{2g_1(T-s)}\mathrm{ds} + (h_1 + l_1\mu_{J2})\cdot\int_t^T e^{2g_1(T-s)}\mathbb{E}[x_s|\mathcal{F}_t]\mathrm{ds},\end{aligned}$$

where we have used Fubini's theorem to interchange the expectation and integration. Given the conditional expectation, the second integral can be further simplified:

$$\begin{aligned}\int_t^T e^{2g_1(T-s)}\mathbb{E}[x_s|\mathcal{F}_t]\mathrm{ds} &= \int_t^T e^{2g_1(T-s)}\left[e^{g_1(s-t)}x_t + \frac{g_0}{g_1}\left(e^{g_1(s-t)} - 1\right)\right]\mathrm{ds} \\ &= e^{2g_1T}\left[\int_t^T e^{-g_1(s+t)}x_t + \frac{g_0}{g_1}\left(e^{-g_1(s+t)} - e^{-2g_1 s}\right)\mathrm{ds}\right] \\ &= e^{2g_1T}\left[-\frac{1}{g_1}\left(e^{-g_1(T+t)} - e^{-2g_1 t}\right)x_t - \frac{g_0}{g_1^2}\left(e^{-g_1(T+t)} - e^{-2g_1 t}\right) + \frac{g_0}{2g_1^2}\left(e^{-2g_1T} - e^{-2g_1 t}\right)\right] \\ &= -\frac{1}{g_1}\left(e^{g_1(T-t)} - e^{2g_1(T-t)}\right)x_t + \frac{g_0}{2g_1^2}\left(-2e^{g_1(T-t)} + 2e^{2g_1(T-t)} + 1 - e^{2g_1(T-t)}\right) \\ &= -\frac{1}{g_1}\left(e^{g_1(T-t)} - e^{2g_1(T-t)}\right)x_t + \frac{g_0}{2g_1^2}\left(1 - e^{g_1(T-t)}\right)^2.\end{aligned}$$

Therefore, the conditional variance in the univariate case is given by

$$\begin{aligned}\mathrm{Var}(x_T|\mathcal{F}_t) = &-\frac{1}{2g_1}(h_0 + l_0\mu_{J2})\left(1 - e^{2g_1(T-t)}\right) \\ &-\frac{1}{2g_1^2}(h_1 + l_1\mu_{J2})\left[2g_1\left(e^{g_1(T-t)} - e^{2g_1(T-t)}\right)x_t - g_0\left(1 - e^{g_1(T-t)}\right)^2\right]. \quad \text{(B.3)}\end{aligned}$$

Equations (B.2) and (B.3) are the basis for the formulation of the transition equation (20) as discussed in Section 3.1. It is crucial for our application to note that the conditional mean (B.2) and conditional variance (B.3) of the univariate AJD process x_T at time $T > t$ conditional on information at time t are affine functions in x_t. The affinity of the conditional moments yields the linear state updating equation, which, in turn, allows us to use the linear Kalman filtering technique.

Using the same reasoning, it is also possible to derive the analogues of equations (B.2) and (B.3) for the multivariate AJD process. However, these expressions involve matrix exponential and their integrations, which makes it cumbersome to work with analytically. Fortunately, the conditional moments can be easily obtained by differentiating the CCF, which is known in a semi-closed form for the AJD class (see Section 2.2). The finite difference approximations of the first and second derivatives around the origin yield very precise results with little additional computational costs. One can also easily verify that the affine property of the conditional moments holds in the multivariate case by differentiating the exponentially-affine CCF.

Appendix C Interpolation-Extrapolation Scheme and CCF Replication

In this appendix, we discuss in detail the option interpolation-extrapolation scheme we adopt and illustrate the CCF 'payoff' replication given a portfolio of option prices.

C.1 Interpolation-Extrapolation Scheme

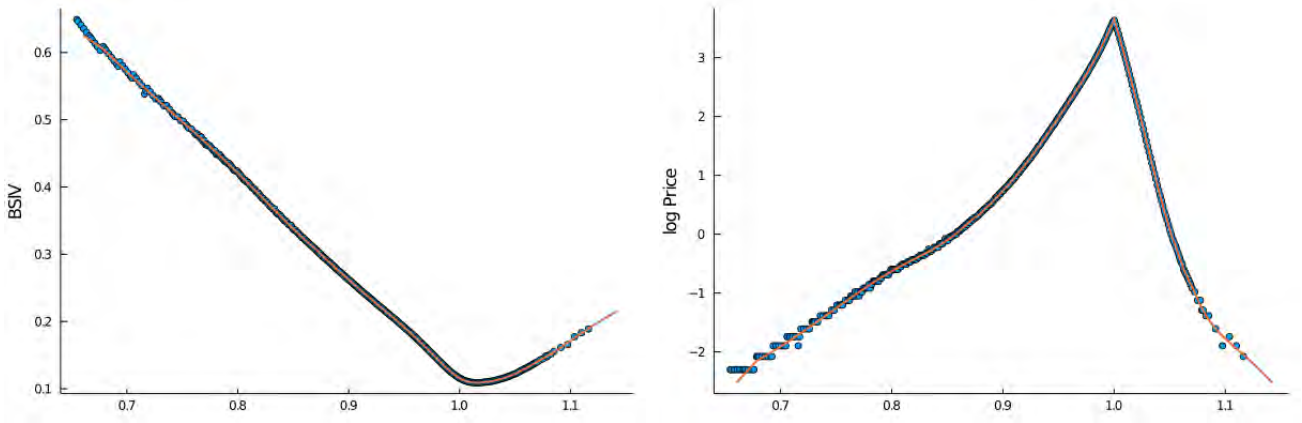
C.1.1 Interpolation

For each trading day and for each tenor, we interpolate option prices using cubic splines. We interpolate option data, expressed in terms of their total implied variance, defined as $v(m,\tau) = \kappa^2(m,\tau) \cdot \tau$, where $\kappa(m,\tau)$ is the Black-Scholes implied volatility for an option with log-moneyness m and tenor τ. This is similar to interpolating on the implied volatility domain, but it will provide us further advantages when we go to the extrapolation scheme.

Cubic splines provide a great tool for the interpolation of options data and are commonly used in the literature, see, for instance, Jiang and Tian (2007), Malz (2014) among many others. Furthermore, they are also used as an approximation method, that allows to smooth out noise in the data; see, for instance, Bliss and Panigirtzoglou (2002), Fengler (2009). For the latter it is common to penalize the squared second derivative of the spline. This, however, might induce a loss of flexibility of the spline leading to larger approximation errors, especially for short-dated options, which are of great importance in our analysis. In this paper, we use the standard cubic spline, but instead of providing all data as knot points, we explicitly specify which data points shall be used as knots for spline interpolation. This allows us to interpolate in some domains and smooth out in others, taking the best out of the interpolation and approximation spline schemes.

In fact, close to ATM options are more liquid than very deep OTM counterparts. Thus, intuitively, information in these options is more reliable, and we would not like to distort it, by imposing smoothing constraints. Very deep OTM options, on the other hand, may be quite illiquid. Furthermore, the tick size for deep OTM options becomes larger relative to their value. This might lead to observing a sequence of the same midpoint quote prices in the data. Figure C.1 provides an example of such 'flat' prices for put options, visible on the right panel for very deep OTM options. These prices clearly violate arbitrage-free assumptions. However,

Figure C.1: Spline interpolation-extrapolation example. April 1, 2021, 15 days-to-maturity



Note: This figure plots an example of the interpolation-extrapolation scheme for options traded on April 1, 2021 with 15 days-to-maturity. The option data (blue dots) are interpolated using the cubic spline (orange line). Interpolation is conducted on the total implied variance domain. The left panel plots the data in terms of Black-Scholes implied volatility, while the right panel plots the data in terms of log prices. Moneyness $k = K/F_t$ is on the horizontal axis.

throwing them away would reduce available information, needed to extract CCF; these prices are not uninformative, but the tick size distorts their information. Therefore, instead of eliminating 'flat' prices, we will just not include them as knot points in our spline interpolation scheme. In other words, we would not require the spline function to go through all data points for deep OTM options, but rather let it approximate information in them.[12]

More formally, first, we include the closest to ATM put option, m_{n^*}, to the knot sequence and then iteratively include put options to the left m_i for $i = 2, \ldots, n^* - 1$ such that all following conditions are satisfied: (i) $P(m_i) < P(m_{i-1})$ and $C(m_i) > C(m_{i-1})$; (ii) $P(m_{i+1}) < P(m_i)$ and $C(m_{i+1}) > C(m_i)$; (iii) daily trading volume for $P(m_i)$ is larger than one. The first two criteria check for arbitrage conditions. The third one filters out possible stale prices from being a knot point. Similar mirrored conditions are applied to OTM call prices. Such a constructed knot sequence will likely contain more closer to ATM options and less deep OTM options, resulting in more interpolation in the former range and more approximation in the later one.

We note that we do not filter out option data that violate arbitrage conditions. Instead, we do not include these points into the knot sequence, thus we do not require the spline to go exactly through these points. Furthermore, another reason is that we use the option-implied CCFs rather than option prices themselves as inputs in our estimation procedure. Similarly, the CBOE does not impose any arbitrage filters in the calculation of the VIX index, except eliminating zero-bid quotes (CBOE, 2015). Figure C.1 provides an example for interpolation-extrapolation scheme for option slice traded on April 1, 2020 with 15 days to maturity.

C.1.2 Extrapolation

Truncation errors are more challenging to address than the discretization errors, since one needs to make assumptions on the dynamics of option prices (either in dollar or volatility terms)

[12]Recall that we interpolate/approximate data on the total implied variance domain, not in terms of implied volatility, option prices or log prices as illustrated in Figure C.1.

beyond the observable range of strikes. On the other hand, as the OTM option prices decrease with $|m|$, the impact of the truncation error is expected to be small for very liquid options (such as index options) that cover a wide range of strike prices. However, the truncation might deteriorate the CCF approximation even for small argument values. This is especially might be a serious issue after a sudden market movements, since options with smaller (or larger) strikes might simply not be issued immediately to cover a new range of strikes.

It is common in the literature to use flat extrapolation; e.g., see again Bliss and Panigirtzoglou (2002), Jiang and Tian (2005) and Malz (2014). Under the flat extrapolation scheme the implied volatility is simply set equal to the volatility of the observable extreme-strike options, i.e., $\kappa(\underline{m}, \tau)$ for the left hand-side and $\kappa(\overline{m}, \tau)$ for the right hand-side of volatility smile. This approach is very simple to implement, and allows a significant reduction of the truncation error. However, the main caveat of flat extrapolation is that it assumes Black-Scholes log-normal model to apply in the tails, beyond the observable range of strikes.

Therefore, in this paper, we extrapolate the total implied variances $v(m, \tau)$ linearly in the log-moneyness m beyond the observable range of strikes. This particular linear parametrization is motivated by the asymptotic results of Lee (2004), who analyzed the behaviour of the implied volatility smile as strikes tend to infinity. Another example of the parametrization that satisfies Lee's asymptotic result is the SVI model, commonly used among practitioners (Gatheral & Jacquier, 2014). However, it is well known that the SVI approach might not provide a good fit for short-dated options. Thus, we use the more flexible cubic spline for interpolation within the observable range of strikes, as detailed above, and, similar to SVI, extrapolate implied variance linearly in log-moneyness.

In fact, the asymptotic result of Lee (2004) for extreme strikes states that the implied volatility wings should not grow faster than $|m|^{1/2}$ and should not grow slower than $|m|^{1/2}$, unless the underlying asset has finite moments of all orders. In particular, Lee (2004) shows that

$$\limsup_{m \to -\infty} \frac{\kappa^2(m, \tau) \cdot \tau}{|m|} = \underline{\beta}^* \text{ with } \underline{\beta}^* \in [0, 2] \quad \text{and} \tag{C.1}$$

$$\limsup_{m \to \infty} \frac{\kappa^2(m, \tau) \cdot \tau}{|m|} = \overline{\beta}^* \text{ with } \overline{\beta}^* \in [0, 2]. \tag{C.2}$$

Furthermore, he establishes that there is a one-to-one correspondence between $\overline{\beta}^*$ and the number of moments of the underlying process F_τ, and $\underline{\beta}^*$ and the number of moments of $1/F_\tau$. For instance, for the right tail the moment formula for implied volatility is

$$\frac{1}{2\overline{\beta}^*} + \frac{\overline{\beta}^*}{8} - \frac{1}{2} = \sup\{p : \mathbb{E}[F_\tau^{1+p}] < \infty\}.$$

This result allows us to conjecture the asymptotically valid parametrization to extrapolate the implied volatility beyond the observable range of strikes. In other words, we assume that the total variance $v(m, \tau) = \kappa^2(m, \tau) \cdot \tau$ is linear in log-moneyness:

$$v(m, \tau) = c + \beta m.$$

An intercept coefficient is introduced to guarantee continuity between interpolation and extrapolation domains. These coefficients are exactly determined by the smallest and largest observable strike prices (or corresponding log-moneyness levels, denoted by $\underline{m}$ and $\overline{m}$, respectively) given the slopes $\underline{\beta}$ and $\overline{\beta}$ for the left and right tails, respectively:

$$\underline{c} = v(\underline{m}, \tau) - \underline{\beta}\underline{m} \quad \text{and} \quad \overline{c} = v(\overline{m}, \tau) - \overline{\beta}\overline{m}.$$

Therefore, what we have left to do is to establish the choice of the coefficients $\underline{\beta}$ and $\overline{\beta}$. Note that the moment formulas are devoted for the upper bounds (C.1) and (C.2). Furthermore, finding the number of finite moments of the underlying would require parametrizing the dynamics of F_τ.[13] The latter is not desirable in our application since we want to fit another parametric model afterwards. Instead, we simply use the derivative of the fitted cubic spline at the last observable strikes to determine the slopes:

$$\underline{\beta} = -\frac{\partial v(m,\tau)}{\partial m}\bigg|_{m=\underline{m}} \text{ and } \overline{\beta} = \frac{\partial v(m,\tau)}{\partial m}\bigg|_{m=\overline{m}}.$$

Note that Lee's bounds for the slopes are asymptotic result. Therefore, the slopes $\underline{\beta}$ and $\overline{\beta}$ should also satisfy these bounds, which translates in $\underline{\beta} \in [-2, 0]$ for the left slope. However, arbitrage-free conditions for our parametrization might possibly be tighter since we are in a finite setting. To obtain these conditions, we closely follow the derivation sketched in Jäckel (2014).

First, similar to Gatheral and Jacquier (2014) we introduce a function $w : \mathbb{R} \to \mathbb{R}$:

$$w(m) := \left(1 - \frac{m}{2} \cdot \frac{v'(m)}{v(m)}\right)^2 - \frac{v'(m)^2}{4}\left(\frac{1}{4} + \frac{1}{v(m)}\right) + \frac{1}{2}v''(m).$$

Then the second derivative of an option with respect to the strike price can be expressed via this function as

$$\frac{\partial^2 O(\tau, K)}{\partial K^2} = e^{-r\tau}\frac{N'(d_-)}{\sqrt{v(m)K}}w(m),$$

where $N'(x)$ is the pdf of the standard normal distribution and $d_- = -\frac{m}{\sqrt{v}} - \frac{\sqrt{v}}{2}$. Since the second derivative of on option yields the risk-neutral density, to guarantee the arbitrage-freeness of the slice any extrapolation (and interpolation) methods should satisfy $w(m) \geq 0$ for all $m \in \mathbb{R}$. See also Gatheral and Jacquier (2014) for more details. In the following, we investigate this inequality for our chosen parametrization.

Substituting $v(m) = c + \beta \cdot m$ into the inequality $w(m) \geq 0$ we have

$$w(m) = \left(1 - \frac{m}{2} \cdot \frac{\beta}{v(m)}\right)^2 - \frac{\beta^2}{4}\left(\frac{1}{4} + \frac{1}{v(m)}\right) \geq 0.$$

[13] Note that the flat-extrapolation assumes the log-normality of the underlying asset in the tails. Since all moments of the log-normal distribution exist, it means that the slope has to be indeed zero.

Multiplying this inequality by non-negative factor $4 \cdot v(m)^2$ results in

$$\begin{aligned}
(2v(m) - m\beta)^2 - \beta^2 \left(\frac{1}{4}v(m)^2 + v(m)\right) &\geq 0 \\
(2c + \beta m)^2 - \frac{\beta^2}{4}(c^2 + 4c + (2c\beta + 4\beta)m + \beta^2 m^2) &\geq 0 \\
\beta^2 \left(1 - \frac{\beta^2}{4}\right) m^2 + \left(4c - \frac{\beta^2}{4}(2c+4)\right)\beta m + 4c^2 - \frac{\beta^2}{4}(c^2 + 4c) &\geq 0 \\
\frac{1}{4}\beta^2(4-\beta^2)m^2 + (4c - \frac{1}{2}\beta^2 c - \beta^2)\beta m + 4c^2 - \frac{1}{4}\beta^2 c^2 - c\beta^2 &\geq 0 \qquad \text{(C.3)}
\end{aligned}$$

The left-hand side is the quadratic function in m with non-negative quadratic coefficient $\frac{1}{4}\beta^2(4-\beta^2)$ since $|\beta| \leq 2$. The discriminant of this quadratic polynomial is $\Delta = \beta^4(c^2 - 4c + \beta^2)$. Thus, if $\Delta < 0$, no real roots exist, and hence, $w(m) \geq 0$ for all β and m. Therefore, for the right tail we have

$$\begin{aligned}
\Delta < 0 \iff c^2 - 4c + \beta^2 &< 0 \\
(\overline{v} - \beta\overline{m})^2 - 4(\overline{v} - \beta\overline{m}) + \beta^2 &< 0 \\
(\overline{m}^2 + 1)\beta^2 - 2\overline{m}(\overline{v} - 2)\beta + \overline{v}^2 - 4\overline{v} &< 0
\end{aligned}$$

where we have substituted values for c, determined by the largest extreme strike price with the moneyness $\overline{m}$ and the total variance $\overline{v} := v(\overline{m})$. The discriminant of the quadratic equation on the left-hand side is given by $\widetilde{\Delta} = 4(4\overline{m}^2 - \overline{v}^2 + 4\overline{v})$. Assuming $\overline{v} < 4$ we get that the real roots of this polynomial have different signs. Therefore, we get a condition:

$$0 < \overline{\beta} < \frac{\overline{m}(\overline{v} - 2) + \sqrt{\overline{\Delta}}}{\overline{m}^2 + 1} \quad \text{if } \overline{\Delta} := 4\overline{m}^2 - \overline{v}^2 + 4\overline{v} > 0. \qquad \text{(C.4)}$$

If $\overline{\Delta} \leq 0$, then $\Delta > 0$. Assuming the positive density at the largest strike, that is, $w(\overline{m}) > 0$, we only need to ensure that the global minimum of the quadratic function on the left-hand side of (C.3) is to the left of $\overline{m}$:

$$\begin{aligned}
\frac{\beta^2 c + 2\beta^2 - 8c}{\beta(4-\beta^2)} &< \overline{m} \\
(\overline{v} - \beta\overline{m} + 2)\beta^2 - 8(\overline{v} - \beta\overline{m}) &< \overline{m}\beta(4-\beta^2) \\
(\overline{v} + 2)\beta^2 + 4\overline{m}\beta - 8\overline{v} &< 0.
\end{aligned}$$

Since the discriminant of this quadratic function is $\Delta_1 = 16(\overline{m}^2 + 2\overline{v}^2 + 4\overline{v}) > 0$ we get another condition on $\overline{\beta}$:

$$0 < \overline{\beta} < \frac{-2\overline{m} + 2\sqrt{\overline{m}^2 + 2\overline{v}^2 + 4\overline{v}}}{\overline{v} + 2}. \qquad \text{(C.5)}$$

Combining (C.4) and (C.5) we get the arbitrage-free condition for the right tail slope $\overline{\beta}$:

$$0 \leq \overline{\beta} < \min(\beta_{max}, 2)$$

with

$$\beta_{max} = \begin{cases} \max\left(\frac{\overline{m}(\overline{v}-2)+\sqrt{\overline{\Delta}}}{\overline{m}^2+1}, \frac{-2\overline{m}+2\sqrt{\overline{m}^2+2\overline{v}^2+4\overline{v}}}{\overline{v}+2}\right) & \text{if } \overline{\Delta} > 0 \\ \frac{-2\overline{m}+2\sqrt{\overline{m}^2+2\overline{v}^2+4\overline{v}}}{\overline{v}+2} & \text{if } \overline{\Delta} \leq 0. \end{cases}$$

Note that since the right boundary in (C.5) is always positive, $\beta_{max} > 0$ regardless of whether $\overline{v} < 4$ or not. Thus, the assumption imposed above on $\overline{v} < 4$ is redundant.

The left-tail slope boundaries can be obtained in a similar way by substituting the observable option price $\underline{v} := v(\underline{m})$ with the smallest strike price $\underline{m}$ and taking the negative roots of quadratic polynomials where necessary:

$$\max(\beta_{min}, -2) < \underline{\beta} \leq 0$$

with

$$\beta_{min} = \begin{cases} \max\left(\frac{\underline{m}(\underline{v}-2)-\sqrt{\underline{\Delta}}}{\underline{m}^2+1}, \frac{-2\underline{m}-2\sqrt{\underline{m}^2+2\underline{v}^2+4\underline{v}}}{\underline{v}+2}\right) & \text{if } \underline{\Delta} > 0 \\ \frac{-2\underline{m}+2\sqrt{\underline{m}^2+2\underline{v}^2+4\underline{v}}}{\underline{v}+2} & \text{if } \underline{\Delta} \leq 0. \end{cases}$$

where $\underline{\Delta} := 4\underline{m}^2 - \underline{v}^2 + 4\underline{v}$.

C.2 Conditional Characteristic Function Replication

This appendix illustrates the CCF 'payoff' replication given a portfolio of option prices.

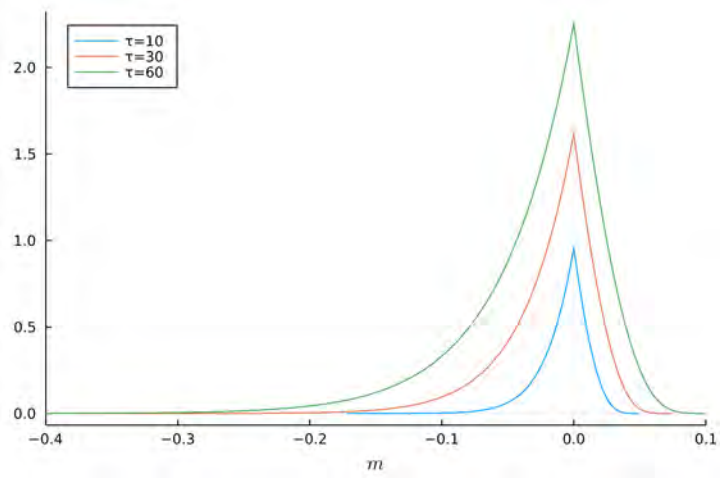
Figure C.2 provides an illustration of OTM option prices and corresponding spanning, option portfolio weights for simulated data from the stochastic volatility (SV) model of Heston (1993). The option portfolio weights are scaled by the factor in front of the integral in the CCF spanning equation (3). We notice that the spanning option portfolios require both short and long positions in OTM options and the weights start oscillating for larger values of arguments u.

Figure C.3 further provides an example of the real and imaginary parts of the integrands $e^{(iu-1)m} \cdot O_t(\tau, m)$ against the log-moneyness level $m = \log(K/F_t)$ for four different values of the arguments u with fixed time-to-maturity $\tau = 30$ days. The corresponding integrals need to be calculated to replicate the CCF 'payoff' function.

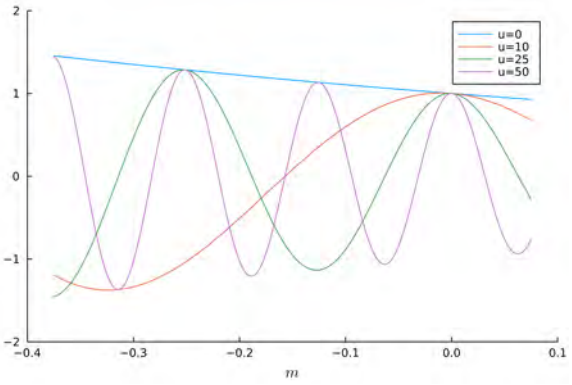
As discussed in Section 3.1, we replicate the CCF 'payoff' using the Riemann sum approximation based on an interpolation-extrapolation scheme applied to the set of observable option prices. Figure C.4 provides an illustration of the impact of different measurement errors on the option-implied CCF. For the illustration, we simulate option prices from the SVCJ model using a similar setup as described in Section 4. In particular, at each time point we have a discrete set of strikes and additive observation errors in the observed option prices. We fix the time-to-maturity to $\tau = 10$ days, $u = 20$ and focus only on the real part of the CCF. These values are chosen to emphasize the impact of the measurement errors. The impact of the discretization errors, for instance, is typically smaller for larger maturities and smaller argument values.

Figure C.4 (a) plots the measurement errors when we use a finite set of observed option prices $\widehat{O}_t(\tau, m)$. That is, it shows the total measurement errors $\zeta_t^{\phi}(u, \tau)$ that aggregate observation errors $\zeta_t^{(1)}(u, \tau)$, truncation errors $\zeta_t^{(2)}(u, \tau)$, and discretization errors $\zeta_t^{(3)}(u, \tau)$, formally defined in Appendix A. As we can see, the discretization and truncation errors yield a small bias in the CCF approximation. In Figure C.4 (b) we turn off the impact of observation errors, that is, we use a finite set of *true* option prices $O_t(\tau, m)$. Thus, the errors in the option-spanned CCF contain only the discretization and truncation errors, $\zeta_t^{(2)}(u, \tau) + \zeta_t^{(3)}(u, \tau)$. We overlay this plot

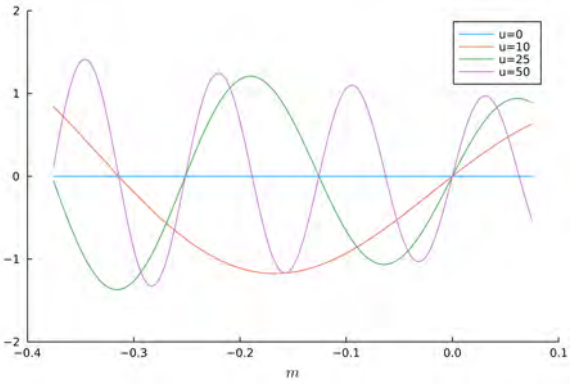
Figure C.2: OTM option prices and corresponding (scaled) option portfolio weights for the SV model of Heston (1993).



(a) OTM option prices



(b) Re, option portfolio weights



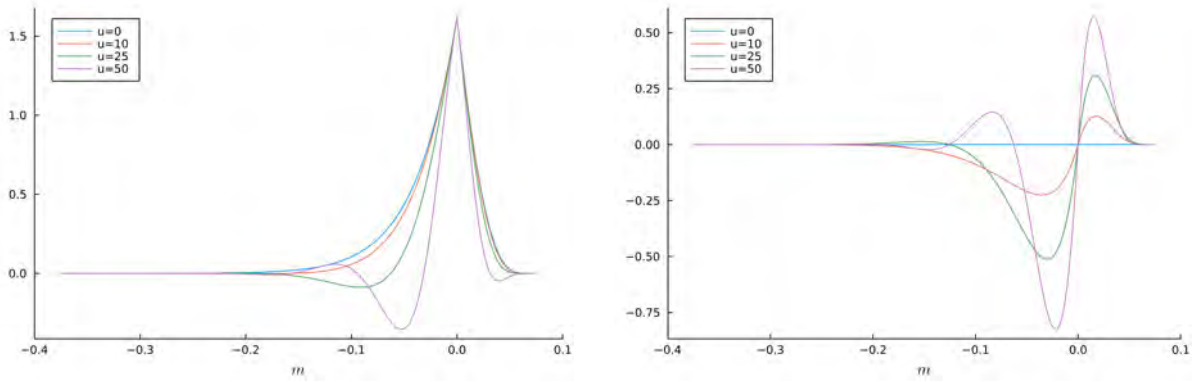
(c) Im, option portfolio weights

Note: Panel (a) of this figure plots OTM option prices for the simulated SV model of Heston (1993) against the log-moneyness level $m = \log(K/F_t)$ for three different levels of maturity τ. Panels (b) and (c) provide the (scaled) option portfolio weights $e^{(iu-1)m}$ to replicate the CCF for different arguments u and fixed time-to-maturity $\tau = 30$ days. The following parameters are used: $\kappa = 8,\ \sigma_v = 0.45,\ \bar{v} = 0.015,\ \rho = -0.95,\ F_0 = 100$ and $v_0 = 0.015$.

with the ATM BSIV to illustrate that these errors are due to the value of options, that is, they are driven by the model's dynamics. Figure C.4 (c) plots the sum of the first two figures, that is, $\zeta_t^{(1)}(u, \tau)$. This figure is already centered around zero.

For Figures C.4 (d) and (e), we use a cubic spline interpolation and extrapolate option prices outside of the observed range as described in Appendix C.1. Figure C.4 (d) plots errors in the CCF approximation when we use the true finite set of option prices, interpolated with the cubic spline to reduce the discreteness of the observed strikes. As we can see, the discretization and truncation errors are largely reduced (note the scale of the y-axis). Finally, Figure C.4 (e) illustrates the errors in the CCF approximation when options are observed with errors and the cubic spline is also applied to $\widehat{O}_t(\tau, m)$. As we can see, the errors in this case are mostly coming from the observation errors. Hence, this figure largely resembles Figure C.4 (c). In other words, the interpolation-extrapolation scheme kills off the level effect that comes from the impact of the discretization and truncation errors, but not the observation errors.

Figure C.3: Integrands $e^{(iu-1)m} \cdot O_t(\tau, m)$ for the SV model of Heston (1993).



(a) $\Re\big(e^{(iu-1)m} \cdot O_t(\tau, m)\big)$ (b) $\Im\big(e^{(iu-1)m} \cdot O_t(\tau, m)\big)$

Note: This figure plots the real and imaginary parts of four integrands $e^{(iu-1)m} \cdot O_t(\tau, m)$ for the simulated SV model of Heston (1993) against the log-moneyness level $m = \log(K/F_t)$ for different arguments u and fixed time-to-maturity $\tau = 30$ days. The following parameters are used: $\kappa = 8,\ \sigma_v = 0.45,\ \bar{v} = 0.015,\ \rho = -0.95,\ F_0 = 100$ and $v_0 = 0.015$.

Appendix D Additional Simulation and Empirical Results

In this appendix we provide the Monte Carlo results for another two option pricing specifications and a specification with the variance risk premia. We also provide additional empirical results.

D.1 Additional simulation results

D.1.1 SVCJ

We additionally illustrate the developed estimation approach based on the 'double-jump' stochastic volatility model of Duffie et al. (2000) with Gaussian jump size distribution. In particular, we assume the following process for the log forward price under both probability measures $\mathbb{P}$ and $\mathbb{Q}$:

$$\mathrm{d}\log F_t = (-\tfrac{1}{2}v_t - \mu\lambda_t)\mathrm{d}t + \sqrt{v_t}\mathrm{d}W_{1,t} + J_t\mathrm{d}N_t, \tag{D.1}$$

$$\mathrm{d}v_t = \kappa(\bar{v} - v_t)\mathrm{d}t + \sigma\sqrt{v_t}\mathrm{d}W_{2,t} + J_t^v\mathrm{d}N_t, \tag{D.2}$$

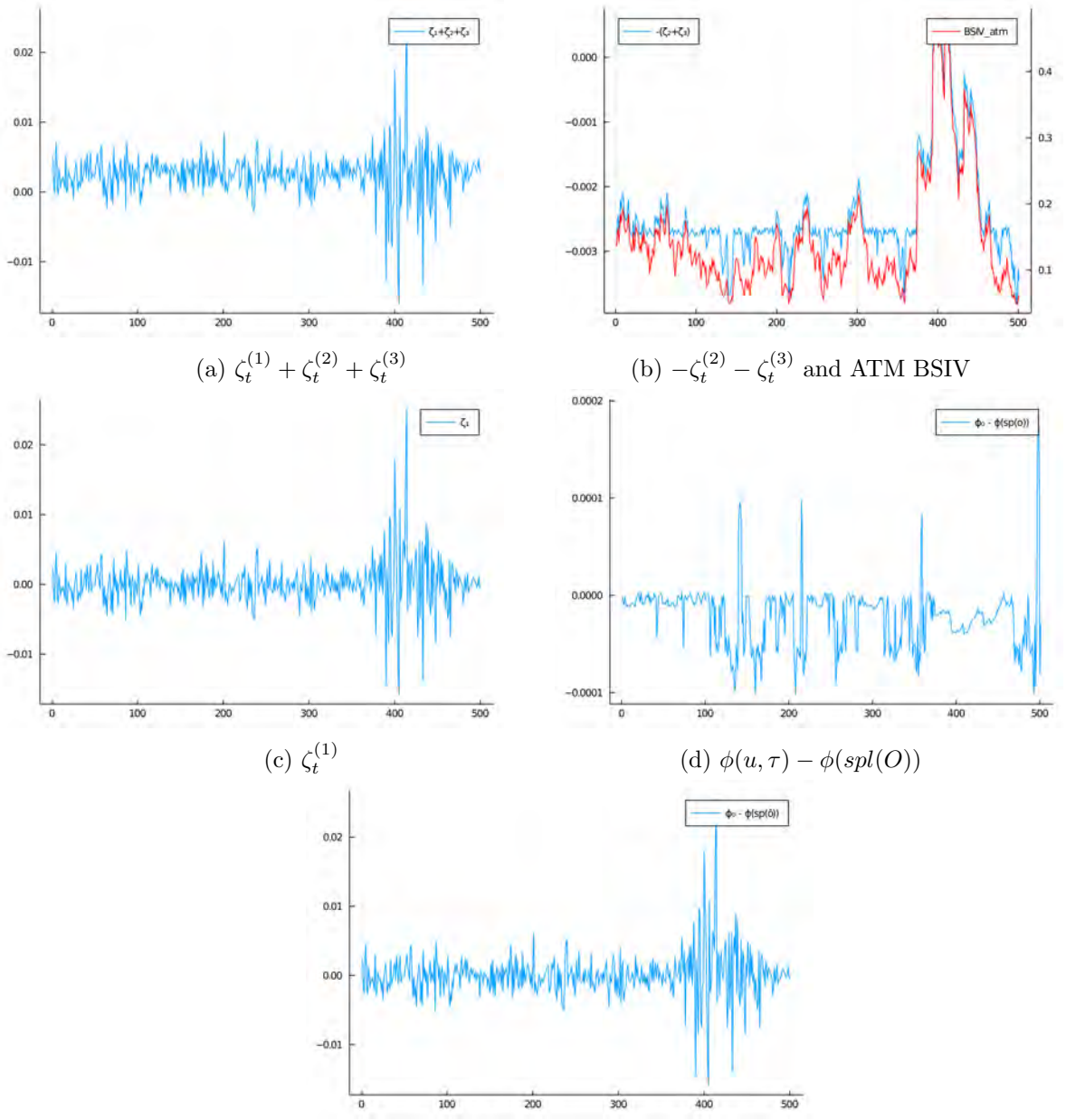
where almost all ingredients are the same as in the SVCDEJ specification in Section 4.1 except the distribution of the jump sizes. In particular, here we assume that the jump sizes in returns are Gaussian $J \sim \mathcal{N}(\mu_J, \sigma_J^2)$ and the jump sizes in volatility are independent from jump sizes in returns with $J^v \sim \exp(1/\mu_v)$.

Similar to the main specification, this model belongs to the AJD class and the log of the option-spanned CCF is linear in the latent state process v_t. The conditional mean and variance of the latent stochastic volatility process are given by

$$\mathbb{E}[v_{t+1}|\mathcal{F}_t] = e^{g_1\Delta t}v_t + \frac{g_0}{g_1}\left(e^{g_1\Delta t} - 1\right), \tag{D.3}$$

$$\mathrm{Var}(v_{t+1}|\mathcal{F}_t) = -\frac{\sigma^2 + 2\delta\mu_v^2}{2g_1^2}\left[2g_1\left(e^{g_1\Delta t} - e^{2g_1\Delta t}\right)v_t - g_0\left(1 - e^{g_1\Delta t}\right)^2\right], \tag{D.4}$$

Figure C.4: Impact of the three types of measurement errors: Illustration.



(a) $\zeta_t^{(1)} + \zeta_t^{(2)} + \zeta_t^{(3)}$

(b) $-\zeta_t^{(2)} - \zeta_t^{(3)}$ and ATM BSIV

(c) $\zeta_t^{(1)}$

(d) $\phi(u,\tau) - \phi(spl(O))$

(e) $\phi(u,\tau) - \phi(spl(\widehat{O}))$

Note: This figure plots the impact of the three types of measurement errors on the option-implied CCF. The figures illustrate the approximation for the real part of the CCF with $\tau = 10$, $u = 20$. The same Monte Carlo simulation setup as described in Section 4 is used here to simulate data from the SVCJ model.

with $g_0 = \kappa\bar{v}$ and $g_1 = -\kappa + \delta\mu_v$. Equations (D.3) and (D.4) are used to define the state updating equation:

$$v_{t+1} = c_t + T_t v_t + \eta_{t+1} \tag{D.5}$$

where $c_t = \frac{g_0}{g_1}\left(e^{g_1\Delta t} - 1\right)$, $T_t = e^{g_1\Delta t}$ and $\text{Var}(\eta_{t+1}|\mathcal{F}_t) = \text{Var}(v_{t+1}|\mathcal{F}_t)$. We also impose the Feller condition $2\kappa\bar{v} > \sigma^2$ and the covariance stationary condition $\kappa > \delta\mu_v$.

We use the simulation setting as in Section 4.1. The simulation results are provided in Table 8. As the richer specification in Section 4.1, the results of the SVCJ model also have

Table 8: Monte Carlo results for the SVCJ model

parameter	σ	κ	$\bar{v}$	ρ	δ	μ_J	σ_J	μ_v	$\sigma_\varkappa$
				$u = 1, \ldots, 15$					
true value	0.400	5.000	0.02	-0.95	20.000	-0.100	0.04	0.05	0.02
mean	0.410	4.869	0.0207	-0.9382	17.013	-0.110	0.0343	0.0520	0.0215
std dev	0.012	0.115	0.0007	0.0151	3.171	0.014	0.0124	0.0026	0.0044
q10	0.400	4.681	0.0201	-0.9516	11.209	-0.136	0.0100	0.0501	0.0166
q50	0.405	4.905	0.0204	-0.9445	18.537	-0.103	0.0404	0.0507	0.0206
q90	0.433	4.988	0.0219	-0.9114	19.573	-0.100	0.0431	0.0570	0.0274
				$u = 1, \ldots, 20$					
true value	0.400	5.000	0.02	-0.95	20.000	-0.100	0.04	0.05	0.02
mean	0.403	4.913	0.0202	-0.9444	18.866	-0.103	0.0397	0.0502	0.0198
std dev	0.006	0.091	0.0003	0.0085	1.420	0.006	0.0050	0.0010	0.0062
q10	0.396	4.830	0.0199	-0.9524	17.903	-0.105	0.0375	0.0496	0.0156
q50	0.403	4.930	0.0202	-0.9456	19.048	-0.102	0.0408	0.0501	0.0185
q90	0.410	5.000	0.0206	-0.9373	20.088	-0.100	0.0426	0.0506	0.0234
				$u = 1, \ldots, 25$					
true value	0.400	5.000	0.02	-0.95	20.000	-0.100	0.04	0.05	0.02
mean	0.395	4.907	0.0200	-0.9555	20.424	-0.097	0.0435	0.0495	0.0207
std dev	0.009	0.129	0.0003	0.0146	1.297	0.004	0.0029	0.0005	0.0097
q10	0.384	4.776	0.0196	-0.9710	19.209	-0.101	0.0405	0.0489	0.0152
q50	0.395	4.934	0.0200	-0.9564	20.392	-0.097	0.0436	0.0495	0.0178
q90	0.404	5.010	0.0203	-0.9409	22.029	-0.093	0.0473	0.0500	0.0284

Note: This table provides Monte Carlo simulation results for the SVCJ model, based on 500 replications. Three settings with different ranges of the argument are considered. Each panel lists, for each parameter, the true value, the Monte Carlo mean and standard deviation, and the 10th, 50th and 90th Monte Carlo percentile, respectively. We use $T = 500$ time points with $\Delta t = 1/250$. The initial values are set to $F_0 = 100$ and $v_0 = 0.02$. The threshold for singular values is set to $\bar{s} = 10^{-7}$.

good finite-sample properties. We also note that the 'double-jump' specification includes other widely used option pricing models such us the stochastic volatility model of Heston (1993).

D.1.2 SVCJ with a variance risk premium

Since the transition equation in the state space representation reflects the $\mathbb{P}$-dynamics of the latent components, it is, in general, possible to learn about the risk premia associated with this latent process. In this subsection, we provide Monte Carlo simulation results in the SVCJ model with a variance risk premium (VRP). In particular, we model the VRP π_v as the difference between the mean-reversion parameters under the $\mathbb{P}$ and $\mathbb{Q}$ measures, that is, in the state transition equation (D.5) we change the mean-reversion parameter to $\kappa^{\mathbb{P}} = \kappa + \pi_v$.

Table 9 provides the Monte Carlo simulation results for the SVCJ model with a VRP. We consider three estimation approaches. First, we fix the VRP parameter to its true value $\pi_v = 1$. Second, we assume no VRP when estimating the model, although the true model is simulated with non-zero VRP, that is, we fix $\pi_v = 0$ in the estimation procedure. Finally, we estimate the VRP along with all model parameters.

As the results suggest, it is hard to identify the VRP in this setting (see the third panel

Table 9: Monte Carlo results for the SVCJ model with variance risk premium

parameter	σ	κ	$\bar{v}$	ρ	δ	μ_J	σ_J	μ_v	$\sigma_\varkappa$	π_v
constrained, $\pi_v = 1.0$										
true value	0.400	5.000	0.2000	-0.950	20.000	-0.100	0.0400	0.0500	0.0200	1.000
mean	0.402	4.924	0.0202	-0.946	18.882	-0.103	0.0396	0.0503	0.0158	-
std dev	0.007	0.073	0.0003	0.008	1.515	0.006	0.0051	0.0010	0.0031	-
q10	0.396	4.841	0.0199	-0.953	18.012	-0.105	0.0374	0.0496	0.0127	-
q50	0.401	4.931	0.0202	-0.947	19.136	-0.101	0.0408	0.0501	0.0155	-
q90	0.409	5.003	0.0205	-0.939	20.182	-0.099	0.0427	0.0508	0.0189	-
constrained, $\pi_v = 0.0$										
true value	0.400	5.000	0.2000	-0.950	20.000	-0.100	0.0400	0.0500	0.0200	1.000
mean	0.402	4.924	0.0202	-0.946	18.882	-0.103	0.0396	0.0503	0.0158	-
std dev	0.007	0.073	0.0003	0.008	1.515	0.006	0.0051	0.0010	0.0031	-
q10	0.396	4.841	0.0199	-0.953	18.012	-0.105	0.0374	0.0496	0.0127	-
q50	0.401	4.931	0.0202	-0.947	19.136	-0.101	0.0408	0.0501	0.0155	-
q90	0.409	5.003	0.0205	-0.939	20.182	-0.099	0.0427	0.0508	0.0189	-
unconstrained										
true value	0.400	5.000	0.2000	-0.950	20.000	-0.100	0.0400	0.0500	0.0200	1.000
mean	0.402	4.923	0.0202	-0.945	18.865	-0.103	0.0395	0.0503	0.0158	4.029
std dev	0.007	0.074	0.0003	0.008	1.546	0.006	0.0052	0.0011	0.0031	4.942
q10	0.396	4.835	0.0199	-0.953	17.985	-0.105	0.0373	0.0496	0.0127	-4.338
q50	0.401	4.931	0.0202	-0.947	19.136	-0.101	0.0408	0.0502	0.0155	4.558
q90	0.409	5.003	0.0206	-0.939	20.182	-0.099	0.0427	0.0508	0.0189	9.883

Note: This table provides Monte Carlo simulation results for the SVCJ model with a variance risk premium, based on 500 replications. Each panel lists, for each parameter, the true value, the Monte Carlo mean and standard deviation, and the 10th, 50th and 90th Monte Carlo percentile, respectively. The argument range is set to $u = 1, \ldots, 20$. We use $T = 500$ time points with $\Delta t = 1/250$. The initial values are set to $F_0 = 100$ and $v_0 = 0.02$. The threshold for singular values is set to $\bar{s} = 10^{-7}$.

in Table 9). It appears that the $\mathbb{Q}$-information in the option prices largely dominates the $\mathbb{P}$-information, making the identification of the VRP relatively weak. A similar issue arises in the term structure literature, where often calibrated bond prices imply unrealistic $\mathbb{P}$-dynamics (see the discussion in Kim & Orphanides, 2012). However, we also notice that in all three estimation settings, the identification of the $\mathbb{Q}$-parameters barely changes. That is, even in the model with the VRP parameter fixed to zero (the second panel in Table 9), the parameter estimates have good finite-sample properties. In the empirical application, we also find that introducing the VRP parameter does not have a significant effect on the estimates of the model's $\mathbb{Q}$-parameters. However, supported by these simulation results, we do not give an economic interpretation to the VRP estimate. For VRP estimation, one can use, e.g., a non-parametric approach based on high-frequency data as in Bollerslev and Todorov (2011) and Andersen, Fusari, and Todorov (2015b).

D.1.3 SVCEJ

Instead of the double-exponential jump size distribution considered in Section 4.1, or the Gaussian distribution considered above, we may consider separate exponential distributions for pos-

itive and negative jumps. Following Bardgett et al. (2019), we consider two separate counting processes for negative and positive jumps N_t^- and N_t^+, respectively, and modify the SVCJ specification as follows:

$$\mathrm{d}\log F_t = (-\frac{1}{2}v_t - \mu^-\lambda_t^+ - \mu^-\lambda_t^+)\mathrm{d}t + \sqrt{v_t}\mathrm{d}W_{1,t} + J_t^-\,\mathrm{d}N_t^- + J_t^+\mathrm{d}N_t^+, \tag{D.6}$$

$$\mathrm{d}v_t = \kappa(\bar{v} - v_t)\mathrm{d}t + \sigma\sqrt{v_t}\mathrm{d}W_{2,t} + J_t^v\mathrm{d}N_t^-, \tag{D.7}$$

where λ_t^- and λ_t^+ are the corresponding jump intensities for negative and positive jumps, and $-J_t^-$ and J_t^+ are exponentially distributed negative and positive jump sizes in log returns with means η^- and η^+, respectively. Note that the negative jump sizes have negative support, that is, J_t^- is negative exponential. Given the jump size distributions, the expected relative jump sizes in returns are $\mu^- = \mathbb{E}[e^{J^-}-1] = \frac{1}{1+\eta^-}-1$ and $\mu^+ = \mathbb{E}[e^{J^+}-1] = \frac{1}{1-\eta^+}-1$. We further let the intensities be affine functions of the stochastic volatility, that is, $\lambda_t^- = \delta_0^- + \delta_1^- v_t$ and $\lambda_t^+ = \delta_0^+ + \delta_1^+ v_t$. However, to keep a moderate number of parameters, we set $\delta_1^+ = 0$ and $\delta_0^- = 0$.

Table 10: Monte Carlo results for the SVCEJ model

parameter	σ	κ	$\bar{v}$	ρ	δ_0^+	δ_1^-	η^+	η^-	μ_v	$\sigma_\varkappa$
					$u = 1,\ldots,15$					
true value	0.450	8.000	0.015	-0.95	2.000	100.000	0.01	0.05	0.05	0.02
mean	0.483	8.063	0.0159	-0.9242	0.512	99.475	0.0355	0.0527	0.0529	0.0626
std dev	0.069	1.506	0.0026	0.0796	1.690	21.104	0.0211	0.0196	0.0351	0.1820
q10	0.459	7.642	0.0150	-0.9577	0.028	90.701	0.0191	0.0485	0.0473	0.0170
q50	0.486	7.914	0.0157	-0.9261	0.125	99.718	0.0320	0.0495	0.0482	0.0229
q90	0.511	8.306	0.0168	-0.9049	0.578	106.121	0.0524	0.0505	0.0504	0.0302
					$u = 1,\ldots,20$					
true value	0.450	8.000	0.015	-0.95	2.000	100.000	0.01	0.05	0.05	0.02
mean	0.462	8.178	0.0149	-0.9551	1.309	110.631	0.0166	0.0510	0.0464	0.0482
std dev	0.042	0.804	0.0020	0.0266	1.566	18.690	0.0130	0.0156	0.0056	0.1425
q10	0.443	7.849	0.0144	-0.9851	0.531	101.950	0.0121	0.0478	0.0452	0.0143
q50	0.454	8.099	0.0148	-0.9588	0.964	107.556	0.0155	0.0489	0.0468	0.0184
q90	0.488	8.410	0.0154	-0.9165	1.742	115.366	0.0190	0.0498	0.0484	0.0368
					$u = 1,\ldots,25$					
true value	0.450	8.000	0.015	-0.95	2.000	100.000	0.01	0.05	0.05	0.02
mean	0.466	7.966	0.0154	-0.9417	1.526	104.783	0.0157	0.0515	0.0479	0.0381
std dev	0.048	0.615	0.0022	0.0321	1.322	17.732	0.0203	0.0140	0.0114	0.1101
q10	0.442	7.671	0.0146	-0.9726	0.602	97.221	0.0106	0.0488	0.0462	0.0129
q50	0.454	7.963	0.0150	-0.9484	1.394	103.412	0.0126	0.0496	0.0476	0.0152
q90	0.494	8.171	0.0159	-0.8935	2.130	109.523	0.0167	0.0505	0.0492	0.0417

Note: This table provides Monte Carlo simulation results for the SVCEJ model, based on 300 replications. Three settings with different ranges of the argument are considered. Each panel lists, for each parameter, the true value, the Monte Carlo mean and standard deviation, and the 10th, 50th and 90th Monte Carlo percentile, respectively. We use $T = 500$ time points with $\Delta t = 1/250$. The initial values are set to $F_0 = 100$ and $v_0 = 0.015$. The threshold for singular values is set to $\bar{s} = 10^{-7}$.

This specification is somewhat richer than the SVCDEJ considered in Section 4.1 since the positive jumps are modeled by a separate counting process with its own positive jump intensity λ_t^+. Nevertheless, this specification also belongs to the AJD class and its CCF of the log forward

prices has a semi-closed form. The state updating equation is defined in a similar way as in the other specifications.

The Monte Carlo simulation results for the SVCEJ model are provided in Table 10. We notice that most of the parameters exhibit good finite-sample performance. However, the parameters related to the positive jumps are biased and have large standard deviation.

D.2 Additional empirical results

Table 11 provides additional empirical results for the same model specification as in Section 4.1. Next to the empirical results with fixed $p^- = 0.7$ in Section 6, we provide the estimates with unrestricted probability of negative jumps and with different fixed values $p^- = 0.65$ and $p^- = 0.75$. Overall, the results indicate similar parameter estimates as in Table 5, suggesting the robustness of our empirical results. We also note larger standard errors of the parameter estimates in the unrestricted model, specifically of the parameter δ which enters the model as a multiple of p^-. This is in line with our simulation results for the unrestricted model (not provided here), which shows the identification limits of the considered unrestricted model. Therefore, in the main empirical results we focus on the restricted model, complementing it with Table 11.

Table 11: SVCDEJ estimation results

	σ	κ	$\bar{v}$	ρ	δ	p^-	η^+	η^-	μ_v	$\sigma_\varkappa$
					unconstrained					
$\hat{\theta}$	0.505	8.368	0.0152	-1.000	167.68	0.6619	0.0195	0.0424	0.0516	0.253
s.e.	0.071	0.762	0.0018	0.054	17.04	0.0155	0.0014	0.0007	0.0039	0.004
					constrained, $p^- = 0.75$					
$\hat{\theta}$	0.503	8.259	0.0153	-1.000	148.01	0.75	0.0218	0.0422	0.0517	0.253
s.e.	0.006	0.091	0.0004	0.011	2.66		0.0005	0.0006	0.0004	0.004
					constrained, $p^- = 0.65$					
$\hat{\theta}$	0.506	8.007	0.0160	-1.000	162.09	0.65	0.0196	0.0432	0.0523	0.253
s.e.	0.006	0.178	0.0003	0.021	1.36		0.0003	0.0006	0.0009	0.004

Note: This table provides the parameter estimates and standard errors for the SVCDEJ model. The model is estimated based on $u = 1, \ldots, 20$ and $\bar{s} = 10^{-7}$.

Table 12 provides empirical results for the alternative model specification SVCEJ detailed in subsection D.1.3, with or without exogenous state variables. Positive and negative jumps are modeled by separate counting processes with their own jump intensities λ_t^+ and λ_t^-, depending on the exogenous variable with coefficients γ^+ and γ^-, respectively. We observe a similar magnitude of most of the parameter estimates, suggesting robustness of our empirical results.

Table 12: SVCEJ estimation results with exogenous state variables

	σ	κ	$\bar{v}$	ρ	δ_0^+	δ_1^-	η^+	η^-	μ_v	γ^+	γ^-	q	$\sigma_\varkappa$
							no exogenous						
$\hat{\theta}$	0.481	8.31	0.0139	-1.00	3.76	107.3	0.0100	0.0445	0.061	-	-	-	0.221
s.e.	0.007	0.26	0.0002	0.01	0.08	3.99	0.0002	0.0006	0.002				0.004
							R_0						
$\hat{\theta}$	0.547	11.28	0.0133	-1.00	0.999	83.11	0.0150	0.0439	0.079	0.036	1.587	0.016	0.213
s.e.	0.005	0.25	0.0003	0.01	0.045	2.97	0.0004	0.0006	0.002	0.300	0.142	0.015	0.004
							EPU						
$\hat{\theta}$	0.489	8.40	0.0142	-1.00	2.74	102.5	0.0112	0.0448	0.063	0.0008	0.0001	0.0002	0.221
s.e.	0.010	0.41	0.0007	0.01	0.22	3.91	0.0004	0.0011	0.001	0.2903	0.0291	0.0300	0.004

Note: This table provides the parameter estimates and standard errors for the SVCEJ model. The model is estimated based on $u = 1, \ldots, 20$ and $\bar{s} = 10^{-7}$.

References

Aït-Sahalia, Y., Cacho-Diaz, J., & Laeven, R. J. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, *117*(3), 585–606.

Andersen, T. G., Benzoni, L., & Lund, J. (2002). An empirical investigation of continuous-time equity return models. *The Journal of Finance*, *57*(3), 1239–1284.

Andersen, T. G., & Bondarenko, O. (2007). Construction and interpretation of model-free implied volatility. *NBER Working Paper No. 13449*.

Andersen, T. G., Fusari, N., & Todorov, V. (2015a). Parametric inference and dynamic state recovery from option panels. *Econometrica*, *83*(3), 1081–1145.

Andersen, T. G., Fusari, N., & Todorov, V. (2015b). The risk premia embedded in index options. *Journal of Financial Economics*, *117*(3), 558–584.

Andersen, T. G., Fusari, N., & Todorov, V. (2017). Short-term market risks implied by weekly options. *The Journal of Finance*, *72*(3), 1335–1386.

Andersen, T. G., Fusari, N., Todorov, V., & Varneskov, R. T. (2021). Spatial dependence in option observation errors. *Econometric Theory*, *37*(2), 205–247.

Arroyo-Marioli, F., Bullano, F., Kucinskas, S., & Rondón-Moreno, C. (2021). Tracking R of COVID-19: A new real-time estimation using the Kalman filter. *PLoS One*, *16*(1), e0244474.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, *131*(4), 1593–1636.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of Finance*, *52*(5), 2003–2049.

Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies*, *16*(1), 101–143.

Bardgett, C., Gourier, E., & Leippold, M. (2019). Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets. *Journal of Financial Economics*, *131*(3), 593–618.

Bates, D. S. (2006). Maximum likelihood estimation of latent affine processes. *The Review of Financial Studies*, *19*(3), 909–965.

Bliss, R. R., & Panigirtzoglou, N. (2002). Testing the stability of implied probability density functions. *Journal of Banking & Finance*, *26*(2-3), 381–422.

Bollerslev, T., & Todorov, V. (2011). Tails, fears, and risk premia. *The Journal of Finance*, *66*(6), 2165–2211.

Boswijk, H. P., Laeven, R. J., & Lalu, A. (2015). Asset returns with self-exciting jumps: Option pricing and estimation with a continuum of moments. Working paper, University of Amsterdam and Tinbergen Institute.

Boswijk, H. P., Laeven, R. J., Lalu, A., & Vladimirov, E. (2021). Jump contagion among stock market indices: Evidence from option markets. *Tinbergen Institute Discussion Paper 2021-086/III*.

Broadie, M., Chernov, M., & Johannes, M. (2007). Model specification and risk premia: Evidence from futures options. *The Journal of Finance*, *62*(3), 1453–1490.

Carr, P., & Madan, D. (1999). Option valuation using the fast Fourier transform. *Journal of Computational Finance*, *2*(4), 61–73.

Carr, P., & Madan, D. (2001). Optimal positioning in derivative securities. *Quantitative Finance*, *1*(1), 19–37.

Carrasco, M., Chernov, M., Florens, J.-P., & Ghysels, E. (2007). Efficient estimation of general dynamic models with a continuum of moment conditions. *Journal of Econometrics*, *140*(2), 529–573.

Carrasco, M., & Florens, J.-P. (2000). Generalization of GMM to a continuum of moment conditions. *Econometric Theory*, *16*(6), 797–834.

CBOE. (2015). *VIX white paper.* `https://cdn.cboe.com/resources/vix/vixwhite.pdf`.

Chacko, G., & Viceira, L. M. (2003). Spectral GMM estimation of continuous-time processes. *Journal of Econometrics*, *116*(1-2), 259–292.

Chang, B.-Y., Christoffersen, P., Jacobs, K., & Vainberg, G. (2012). Option-implied measures of equity risk. *Review of Finance*, *16*(2), 385–428.

Cheng, P., & Scaillet, O. (2007). Linear-quadratic jump-diffusion modeling. *Mathematical Finance*, *17*(4), 575–598.

Chernov, M., & Ghysels, E. (2000). A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *Journal of Financial Economics*, *56*(3), 407–458.

Christoffersen, P., Dorion, C., Jacobs, K., & Karoui, L. (2014). Nonlinear Kalman filtering in affine term structure models. *Management Science*, *60*(9), 2248–2268.

Christoffersen, P., Jacobs, K., & Mimouni, K. (2010). Volatility dynamics for the S&P500: Evidence from realized volatility, daily returns, and option prices. *The Review of Financial Studies*, *23*(8), 3141–3189.

Christoffersen, P., Jacobs, K., & Ornthanalai, C. (2012). Dynamic jump intensities and risk premiums: Evidence from S&P500 returns and options. *Journal of Financial Economics*, *106*(3), 447–472.

Dai, Q., & Singleton, K. J. (2000). Specification analysis of affine term structure models. *The Journal of Finance*, *55*(5), 1943–1978.

de Jong, F. (2000). Time series and cross-section information in affine term-structure models. *Journal of Business & Economic Statistics*, *18*(3), 300–314.

Dini, D. H., & Mandic, D. P. (2012). Class of widely linear complex Kalman filters. *IEEE Transactions on Neural Networks and Learning Systems*, *23*(5), 775–786.

Driessen, J. (2005). Is default event risk priced in corporate bonds? *The Review of Financial Studies*, *18*(1), 165–195.

Du, D., & Luo, D. (2019). The pricing of jump propagation: Evidence from spot and options markets. *Management Science*, *65*(5), 2360–2387.

Duffee, G. R. (1999). Estimating the price of default risk. *The Review of Financial Studies*, *12*(1), 197–226.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *The Journal of Finance*, *57*(1), 405–443.

Duffie, D., Filipović, D., & Schachermayer, W. (2003). Affine processes and applications in finance. *The Annals of Applied Probability*, *13*(3), 984–1053.

Duffie, D., & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, *6*(4), 379–406.

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, *68*(6), 1343–1376.

Eraker, B. (2004). Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *The Journal of Finance*, *59*(3), 1367–1403.

Eraker, B., Johannes, M., & Polson, N. (2003). The impact of jumps in volatility and returns. *The Journal of Finance*, *58*(3), 1269–1300.

Fang, F., & Oosterlee, C. W. (2008). A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, *31*(2), 826–848.

Fengler, M. R. (2009). Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance*, *9*(4), 417–428.

Feunou, B., & Okou, C. (2018). Risk-neutral moment-based estimation of affine option pricing models. *Journal of Applied Econometrics*, *33*(7), 1007–1025.

Fisher, M., & Gilles, C. (1996). Estimating exponential-affine models of the term structure. *Unpublished working paper. Federal Reserve Bank of Atlanta*.

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., ... others (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, *584*(7820), 257–261.

Gallant, A. R., & Tauchen, G. (1996). Which moments to match? *Econometric Theory*, 657–681.

Gatheral, J., & Jacquier, A. (2014). Arbitrage-free SVI volatility surfaces. *Quantitative Finance*, *14*(1), 59–71.

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applica-

tions to bond and currency options. *The Review of Financial Studies*, *6*(2), 327–343.

Jäckel, P. (2014). Clamping down on arbitrage. *Wilmott*, *2014*(71), 54–69.

Jiang, G. J., & Tian, Y. S. (2005). The model-free implied volatility and its information content. *The Review of Financial Studies*, *18*(4), 1305–1342.

Jiang, G. J., & Tian, Y. S. (2007). Extracting model-free volatility from option prices: An examination of the VIX index. *The Journal of Derivatives*, *14*(3), 35–60.

Johannes, M. S., Polson, N. G., & Stroud, J. R. (2009). Optimal filtering of jump diffusions: Extracting latent states from asset prices. *The Review of Financial Studies*, *22*(7), 2759–2799.

Jungbacker, B., & Koopman, S. J. (2015). Likelihood-based dynamic factor analysis for measurement and forecasting. *The Econometrics Journal*, *18*(2), 1–21.

Kim, D. H., & Orphanides, A. (2012). Term structure estimation with survey data on interest rate forecasts. *Journal of Financial and Quantitative Analysis*, *47*(1), 241–272.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, *48*(8), 1086–1101.

Lee, R. W. (2004). The moment formula for implied volatility at extreme strikes. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, *14*(3), 469–480.

Lo, A. W. (1988). Maximum likelihood estimation of generalized Itô processes with discretely sampled data. *Econometric Theory*, 231–247.

Maier, B. F., & Brockmann, D. (2020). Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*, *368*(6492), 742–746.

Malz, A. M. (2014). A simple and reliable way to compute option-based risk-neutral distributions. *FRB of New York Staff Report*(677).

Monfort, A., Pegoraro, F., Renne, J.-P., & Roussellet, G. (2017). Staying at zero with affine processes: An application to term structure modelling. *Journal of Econometrics*, *201*(2), 348–366.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics*, *63*(1), 3–50.

Piazzesi, M. (2010). Affine term structure models. In *Handbook of Financial Econometrics: Tools and Techniques* (pp. 691–766). Elsevier.

Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M., Davies, N., ... others (2020). The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study. *The Lancet Public Health*, *5*(5), e261–e270.

Qin, L., & Todorov, V. (2019). Nonparametric implied Lévy densities. *The Annals of Statistics*, *47*(2), 1025–1060.

Santa-Clara, P., & Yan, S. (2010). Crashes, volatility, and the equity premium: Lessons from S&P 500 options. *The Review of Economics and Statistics*, *92*(2), 435–451.

Singleton, K. J. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, *102*(1), 111–141.

Singleton, K. J. (2006). *Empirical dynamic asset pricing: Model specification and econometric assessment*. Princeton University Press.

Todorov, V. (2019). Nonparametric spot volatility from options. *The Annals of Applied Probability*, *29*(6), 3590–3636.