GMM Estimation of Stochastic Volatility Models Using Transform-Based Moments of Derivatives Prices

Yannick Dillschneider* and Raimond Maurer

Goethe University Frankfurt

This version: May 20, 2021
First version: May 17, 2020

Abstract

Derivatives, especially equity and volatility options, contain valuable and oftentimes essential information for estimating stochastic volatility models. Absent strong assumptions, their typically highly nonlinear pricing dependence on the state vector prevents or at least severely impedes their inclusion into standard estimation approaches. This paper develops a novel and unified methodology to incorporate moments involving derivatives prices into a GMM estimation procedure. Invoking new results from generalized transform analysis, we derive analytically tractable expressions for exact moments and devise a computationally attractive approximation procedure. We exemplify our methodology with an estimation problem that *jointly* accounts for stock returns as well as prices of equity and volatility options. Finally, we provide numerical results that support the effectiveness of our methodology.

JEL classification: C32, C51, C58, G12, G13

Keywords: generalized transform analysis; stochastic volatility models; option pricing; GMM estimation

* Corresponding author.

Addresses: Dillschneider, Finance Department, Goethe University, Theodor-W.-Adorno-Platz 3, 60323 Frankfurt am Main, Germany, dillschneider@finance.uni-frankfurt.de; Maurer, Finance Department, Goethe University, Theodor-W.-Adorno-Platz 3, 60323 Frankfurt am Main, Germany, maurer@finance.uni-frankfurt.de.

Acknowledgments: We thank seminar participants at the EEA Annual Congress 2020 (online) and the Virtual Derivatives Workshop 2020 (online) for helpful comments and suggestions. This research was conducted with support from the German Investment and Asset Management Association (BVI). Opinions and any errors are solely those of the authors and not of the institutions with which the authors are affiliated, or any individual cited.

1 Introduction

On many occasions in finance, the researcher encounters a situation in which the estimation of a model featuring latent state variables becomes required or desired. The latent character of certain state variables can significantly complicate the estimation of model parameters. While latent state variables, by their very nature, are not directly observable, there are oftentimes derivatives contracts available, in particular options, whose observable prices depend on the latent state variables. Even if derivatives pricing itself is not the primary concern, relying only on the time series of observable state variables and completely neglecting the information contained in derivatives prices may have serious adverse consequences. Possible consequences range from inaccurate parameter estimates to even the failure to properly identify important parameters. Time series information generally allows to identify only those model parameters driving the real-world dynamics of the state vector. Parameters determining various types of risk premia, thereby linking the real-world to the risk-neutral state dynamics relevant for derivatives pricing, can remain unidentified. In that situation, observable derivatives prices may serve as surrogate for unobservable state variables to estimate the real-world dynamics and, beyond that, introduce information about risk premia. By contributing accurate information about otherwise poorly identified or unidentfied parameters, including derivatives prices into the estimation procedure can be expected to yield substantial statistical efficiency gains and overcome various identification issues. At the same time, one needs to be aware of potential "dark matter" issues in the sense of Chen et al. (2019) when asset prices dominate the estimation procedure relative to alternative information sources.

The case of stochastic volatility models, describing the joint evolution of an equity index and its instantaneous volatility, is an ideal candidate to exemplify the above-described situation. To capture important stylized facts of the data, state-of-the-art continuous-time stochastic volatility models, such as extensions of the classical Heston (1993) model, typically feature multiple latent state variables, one of which usually represents the instantaneous volatility level.[1] Not only are the dynamics sufficiently interesting, but also is a rich set of different derivatives contracts available, such as options on the equity index itself and on an associated volatility index, which represents an (equity) option-implied volatility measure.[2] What makes stochastic volatility models particularly appealing for our purposes is that each of the derivatives markets is found to contribute distinct information about state dynamics and risk premia. Pertaining to the equity derivatives, Bates (2000) and Eraker (2004), among others, document that estimates of some parameters differ significantly depending on whether options are included into the estimation procedure. Moreover, Barras and Malkhozov (2016) identify significant differences in variance risk premia when measured in equity underlying and option markets. Bardgett et al. (2019) further conclude that volatility derivatives contain incremental information about stock return volatility that is not already spanned by equity derivatives. In addition, Song and Xiu (2016) find evidence that standard model specifications capture risk premia reflected by equity options well, but fail to adequately account for risk premia embedded in volatility options. These findings raise the bar for stochastic volatility models to jointly capture the core features of all involved underlying and derivatives markets.

Despite the apparent benefits of including derivatives prices into the estimation process, they are often neglected in many estimation procedures. Besides potential data availability issues, the primary reason for this is the typically highly nonlinear functional dependence of derivatives prices on the latent state variables, which impedes their analytical tractability and inclusion into standard estimation procedures. In fact, as discussed in more detail below, available estimation approaches incorporating derivatives prices typically rely on computationally intensive techniques, such as extensive simulations and large-dimensional

[1] This is different for most discrete-time GARCH-type stochastic volatility specifications, in which the stochastic volatility is a function of observable realized returns.

[2] The most prominent volatility index is probably CBOE's VIX, which is derived from the prices of S&P 500 options.

optimizations, and may additionally impose strong and somewhat arbitrary assumptions on measurement errors of derivatives prices. Considerable data downsampling is required in order to actually implement these approaches, affecting usually both the time series (weekly or even monthly observations) and the cross section (few derivatives contracts at each point in time), thereby leading to a substantial economic information loss in the estimation process. The adverse effects of downsampling can be expected to turn out particularly severe for the case of stochastic volatility models, which are known to feature a rich multi-factor structure and a prominent short-term component.

This paper develops a novel and unified approach to incorporate a broad class of derivatives into a GMM estimation procedure, which is both computationally attractive without the need for data downsampling and compatible with realistic measurement error models.[3] As a starting point, we develop a general derivatives pricing formula, based on the generalized transform analysis introduced in Chen and Joslin (2012) and further developed in Dillschneider (2020). This unified formula is valid for a broad class of derivatives and, among others, covers equity options and volatility options. Using the general pricing formula, we then derive *exact* expressions for moments involving polynomials of derivatives prices. Our results rely on advanced tools from generalized transform analysis, allowing us to express the respective moments in analytically tractable form, assuming the availability of certain standard transforms of the state vector. To our knowledge, expressions of this kind are novel and may be interesting in their own right. However, practical computation of these exact moments generally requires numerical integration of dimensionality equal to the order of the polynomial. Without the use of sophisticated numerical integration techniques, which are beyond the scope of this paper, exact moments are computationally feasible only for low orders. To overcome these limitations, we proceed to derive *approximate* moments using polynomial expansion, requiring only the evaluation of first-order exact moments. Thus constructed approximate moments involving polynomials of derivatives prices can be computed efficiently using standard numerical integration techniques. Moreover, we theoretically verify convergence of the approximate moments to their exact counterparts under standard regularity conditions. In a numerical study, we further provide evidence that approximate moments are generally sufficiently accurate even for low approximation orders. Deriving exact and approximate moment conditions using our methodology, we devise a GMM estimation procedure that incorporates moments involving polynomials of derivatives prices, which is able to jointly account for equity and volatility derivatives.

As a concrete setting for illustrating our methodology, due to their high relevance and analytical tractability, we focus on stochastic volatility models in the affine jump diffusion class (e.g., Duffie et al. (2003, 2000) and Filipović and Mayerhofer (2009)). Extending previous results in the literature, we derive the required standard transforms of the state vector in an explicit multi-period setting, relying on recursive relations involving solutions of a system of generalized Riccati differential equations (cf. Dillschneider (2020)). In practical applications, these can be solved numerically in an efficient way by using vectorization techniques. As a special case, we additionally provide a method to arrive at closed-form expressions for polynomial moments of the state vector.

Despite presenting our methodological results in this particular setting, their scope extends much farther. Beyond affine jump diffusions, it suffices to consider models for which the required standard transforms of the state vector are sufficiently tractable. This covers, among others, discrete-time affine processes as well as certain Lévy-type processes (see also Chen and Joslin (2012) and Dillschneider (2020) for further examples). With this sort of tractability assured, various different model types apart from stochastic volatility models may be studied with our methodology. Indeed, a broad class of derivatives prices can be expressed in the required form. Examples include various interest rate derivatives, credit derivatives, dividend derivatives, and exchange rate derivatives.

[3]Explicitly accounting for measurement errors in derivatives prices has been advocated in several recent contributions, including Andersen et al. (2020) and Duarte et al. (2020).

Our methodological approach is naturally related to the strand of literature devoted to devising estimation procedures for stochastic volatility models. The existing literature comprises essentially three groups of estimation approaches, each incorporating a different granularity of the information conveyed by derivatives. A first group contains a large number of estimation approaches that do not directly account for derivatives as such.[4] Within these, latent state variables are generally proxied by observable variables or "integrated out," either numerically or through simulations. Natural candidates to proxy the instantaneous volatility level could be volatility indices or closely related quotes of instruments like variance swaps. Instead of degrading their role to proxy variables, a second group of approaches explicitly models these derivatives prices as mostly affine functions of the state vector and, thereby, incorporates a limited amount of the information available in derivatives markets.[5] Yet, most of the much finer information contained in the cross section of option prices is not directly accounted for. This is only achieved by a third group of approaches, which in fact incorporate individual option prices and are, therefore, most closely comparable to our approach in terms of capabilities.

Historically, the focus was initially on including equity options into existing estimation approaches, building on analytically tractable and computationally efficient transform-based pricing formulas (e.g., Bakshi and Madan (2000), Carr and Madan (1999), and Duffie et al. (2000)). In essence, the developed estimation approaches — either directly or indirectly — implement filtering procedures for latent state variables in various degrees of sophistication.

Without any simplifying assumptions regarding measurement errors of option prices, the exact filter for latent state variables is computationally infeasible. Instead, simulation-based methods can be relied upon to generate an approximation. Situated in a Bayesian framework, Eraker (2004) achieve this by Markov chain Monte Carlo methods, while Christoffersen et al. (2010) and Fulop and Li (2019) rely on particle filtering. Relatedly, Andersen et al. (2002) and Chernov and Ghysels (2000) extend the simulation-based efficient method of moments approach of Gallant and Tauchen (1996). While being versatile, the required extensive simulations create a huge computational burden for implementing simulation-based estimation procedures.

Other suggested estimation approaches explicitly treat latent states as additional parameters that need to be estimated, such as Bates (2000), Boswijk et al. (2016), and Huang and Wu (2004). Thereby, they incorporate a time series dimension into traditional calibration exercises, in which only a risk-neutral pricing model is fitted to a cross section of options prices on a day-by-day basis (e.g., Bakshi et al. (1997)) or using option panels (e.g., Andersen et al. (2015, 2018)). Despite getting rid of the need to perform extensive simulations, the computational burden is simply relocated to the requirement of optimizing over a large-dimensional parameter space.

Imposing sufficiently strict assumptions on measurement errors simplifies the filtering problem up to the point where latent state variables can be exactly recovered from observed option prices by (numerically) inverting the pricing formula. Following this route, Pan (2000, 2002) proposes a so-called implied-state GMM approach, which Garcia et al. (2011) extend to additionally include moments of integrated volatility.[6] Equivalent assumptions in a maximum likelihood framework allow Aït-Sahalia and

[4]These include maximum likelihood (e.g., Aït-Sahalia and Kimmel (2007), Bakshi et al. (2006), and Bates (2006)), quasi-maximum likelihood (e.g., Harvey and Shephard (1996) and Ruiz (1994)), simulated maximum likelihood (e.g., Durham (2006) and Sandmann and Koopman (1998)), generalized method of moments (e.g., Aït-Sahalia et al. (2015b), Bollerslev and Zhou (2002), and Jiang and Oomen (2007)), simulated method of moments (e.g., Duffie and Singleton (1993)), efficient method of moments (e.g., Chernov et al. (2003) and Gallant and Tauchen (1996)), empirical characteristic function estimation (e.g., Carrasco et al. (2007), Chacko and Viceira (2003), and Singleton (2001)), and simulation-based Markov chain Monte Carlo methods (e.g., Eraker (2001), Eraker et al. (2003), and Jacquier et al. (1994)), among others.

[5]Examples for this approach include Bollerslev et al. (2011), Duan and Yeh (2010), Durham (2013), Egloff et al. (2010), Jones (2003), and Wu (2011) as well as, somewhat relatedly, Aït-Sahalia et al. (2015a). This approach can be extended to include portfolios having polynomial dependence on the state vector (e.g., Feunou and Okou (2018)).

[6]For single-factor stochastic volatility models, Aït-Sahalia et al. (2021b) develop a GMM approach based on closed-form implied volatility expansions for equity options. Gagliardini et al. (2011) also propose a GMM approach involving option prices, which, however, rests on the assumption that the state vector is observable.

Kimmel (2007) to obtain (approximate) transition densities involving option prices. In addition to the computational burden embedded in explicit pricing function inversions, these approaches maintain strict and somewhat arbitrary assumptions regarding measurement errors, which can lead to robustness issues as well as inherent inconsistencies when comparing different models.

With the advent of analytically tractable pricing formulas, such as those stemming from the generalized transform analysis of Chen and Joslin (2012), attention is increasingly devoted to investigating volatility options.[7] Methodologically, most estimation approaches previously invoked for incorporating equity derivatives can straightforwardly be extended to incorporate volatility derivatives, either on a stand-alone basis or jointly with equity derivatives. For the former, Branger et al. (2016) rely on quasi-maximum likelihood methods, while for the latter, empirical studies have primarily focused on calibration exercises (e.g., Carr and Madan (2014), Fouque and Saporito (2018), Kokholm and Stisen (2015), and Papanicolaou and Sircar (2014)). To our knowledge, only Bardgett et al. (2019) attempt a fully-fledged estimation, employing simulation-based Markov chain Monte Carlo techniques.

Since our approach partly relies on approximation techniques, this paper is also related to the strands of the literature developing approximation methods for parameter estimation or derivatives pricing. Beyond simulation-based approaches, which naturally involve stochastic approximations, several deterministic approximation methods are employed for the purpose of parameter estimation when exact expressions are unavailable or prohibitively costly to compute. These include likelihood expansions (e.g., Aït-Sahalia (2002, 2008), Bakshi and Ju (2005), Filipović et al. (2013), and Yu (2007)) as well as approximate moment conditions (e.g., Aït-Sahalia et al. (2015b) and Stanton (1997)).[8] A vast literature exists also on the use of various approximation techniques for the purpose of option pricing.[9] Predominantly, the intention of these methods is to simplify the pricing formula in a way that avoids numerical integration. Our primary intention is different, as we instead aim at simplifying the functional dependence of derivatives prices on the state vector.

The remainder of this paper is organized as follows. Section 2 introduces the general stochastic volatility model and some of its properties. Subsequently, section 3 presents a unified framework for pricing derivatives, which is then used in section 4 to derive moments involving derivatives prices. Building on these results, section 5 formulates our GMM estimation approach. Numerical results supporting our methodology are presented in section 6. Finally, section 7 concludes the paper. The appendix contains additional details, including derivations and proofs.

2 Affine stochastic volatility models

This section introduces the generic stochastic volatility model, for which we choose an affine jump diffusion framekwork. While there is a broad consensus about the necessity of multi-factor stochastic volatility models, less agreement is achieved with regard to the concrete factor structure. Yet, it is largely agreed upon that both diffusion and jump factors are required. Accounting for the large number of potential specifications, we present our model in a versatile setup that allows for multiple diffusive and jump risk sources.

[7]Instead of using a consistent model for the volatility index derived from stock price dynamics, a pragmatic approach is to model its dynamics on a stand-alone basis. This allows to resort to significantly simpler methods for derivatives pricing and parameter estimation (e.g., Dotsis et al. (2007) and Mencía and Sentana (2013)).

[8]See also Kristensen and Salanié (2017) for other types of approximations and potential improvements of the resulting estimation approaches.

[9]Existing approximation approaches for option prices include orthogonal polynomial expansion (e.g., Ackerer and Filipović (2019), Barletta and Nicolato (2017), Madan and Milne (1994), and Xiu (2014)), eigenfunction expansion (e.g., Davydov and Linetsky (2003), Lewis (1998), and Linetsky (2004, 2007)), Edgeworth expansion (e.g., Jarrow and Rudd (1982)), Fourier cosine expansion (e.g., Fang and Oosterlee (2009)), saddlepoint approximation (e.g., Glasserman and Kim (2009)), and auxiliary model approximation (e.g., Kristensen and Mele (2011)), among others. Relatedly, Aït-Sahalia et al. (2021a) propose a closed-form expansion method for option-implied volatilities.

The remainder of this section is organized as follows. Section 2.1 introduces the generic affine stochastic volatility model. A large number of state-of-the-art models are special cases of this model class. Some examples of such models are provided in section 2.2, without attempting an exhaustive enumeration. Subsequently, section 2.3 presents important results from standard transform analysis, which we will heavily draw upon in the remainder of this paper.

2.1 Generic affine model

Throughout, for each model considered, the state process $(X_t)_{t\geq 0}$ takes values in the state space $\mathcal{X} \subset \mathbb{R}^{n_X}$ and is defined on the real-world filtered probability space $(\Omega, \Sigma, \mathcal{F}, \mathbb{P})$, in which the sample space Ω is equipped with a σ-algebra Σ and the natural filtration $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$ of the respective state process, modeling the evolution of information. While the data is generated under the real-world probability measure $\mathbb{P}$, we assume that markets are arbitrage-free, which guarantees the existence of a risk-neutral probability measure $\mathbb{Q}$. On many occasions in this paper, to avoid redundancies, we will make statements under a generic probability measure $\mathbb{M}$, referring to either $\mathbb{P}$ or $\mathbb{Q}$.

The joint state vector $X_t = [\log S_t; Z_t]$ in our generic model, taking values in $\mathcal{X} = \mathbb{R} \times \mathcal{Z}$ for $\mathcal{Z} \subset \mathbb{R}^{n_Z}$, is composed of the stock price S_t and the n_Z-element vector Z_t containing additional state variables. In order to simplify the exposition and terminology, we assume that the stock price S_t is observable, while all state variables in Z_t are latent, i.e., not directly observable. Handling additional observable state variables results in rather straightforward modifications. Throughout, we will moreover assume that the latent state process $(Z_t)_{t\geq 0}$ is strictly stationary.

Under the generic probability measure $\mathbb{M}$, the state vector $X_t = [\log S_t; Z_t]$ is governed by the jump diffusion dynamics

$$\mathrm{d}\log S_t = \mu_S^{\mathbb{M}}(Z_{t-})\,\mathrm{d}t + \sigma_S(Z_{t-})\,\mathrm{d}W_t^{\mathbb{M}} + J_{S,t}\,\mathrm{d}N_t \tag{2.1a}$$

$$\mathrm{d}Z_t = \mu_Z^{\mathbb{M}}(Z_{t-})\,\mathrm{d}t + \sigma_Z(Z_{t-})\,\mathrm{d}W_t^{\mathbb{M}} + J_{Z,t}\,\mathrm{d}N_t\ , \tag{2.1b}$$

where $W_t^{\mathbb{M}}$ is an n_D-element vector standard Brownian motion and N_t is an n_J-element vector Poisson process with intensity $\lambda^{\mathbb{M}}(Z_{t-})$. Employing the definitions $\mu_X^{\mathbb{M}} = [\mu_S^{\mathbb{M}}; \mu_Z^{\mathbb{M}}]$, $\sigma_X = [\sigma_S; \sigma_Z]$, and $J_{X,t} = [J_{S,t}; J_{Z,t}]$, we can write the dynamics in equation (2.1) in the general form

$$\mathrm{d}X_t = \mu_X^{\mathbb{M}}(Z_{t-})\,\mathrm{d}t + \sigma_X(Z_{t-})\,\mathrm{d}W_t^{\mathbb{M}} + J_{X,t}\,\mathrm{d}N_t\ .$$

Analogous to Duffie et al. (2000), we impose the following affine restrictions on the drift vector $\mu_X^{\mathbb{M}}$, instantaneous diffusive covariance matrix $\Omega_X = \sigma_X \sigma_X^\top$, jump intensity vector $\lambda^{\mathbb{M}}$, and joint distribution $\nu^{\mathbb{M}}$ of jump sizes $J_{X,t}$:

- $\mu_X^{\mathbb{M}}(z) = A_{\mu,X}^{\mathbb{M}} + B_{\mu,X}^{\mathbb{M}} z$ with $A_{\mu,X}^{\mathbb{M}} \in \mathbb{R}^{n_X}$ and $B_{\mu,X}^{\mathbb{M}} \in \mathbb{R}^{n_X \times n_Z}$,
- $\mathrm{vec}[\Omega_X(z)] = A_{\Omega,X} + B_{\Omega,X} z$ with $A_{\Omega,X} \in \mathbb{R}^{n_X^2}$ and $B_{\Omega,X} \in \mathbb{R}^{n_X^2 \times n_Z}$,
- $\lambda^{\mathbb{M}}(z) = A_\lambda^{\mathbb{M}} + B_\lambda^{\mathbb{M}} z$ with $A_\lambda^{\mathbb{M}} \in \mathbb{R}^{n_J}$ and $B_\lambda^{\mathbb{M}} \in \mathbb{R}^{n_J \times n_Z}$, and
- $J_{X,t} \sim \nu^{\mathbb{M}}$ and i.i.d. over time.

Loosely speaking, these restrictions require affine functions of the latent state vector, subject to implicitly imposed coefficient restrictions assuring that all functions are well-defined. E.g., Duffie and Kan (1996) formulate a generalized Feller condition for affine diffusive covariance matrices.

Reflected in the dependence on $\mathbb{M}$ in drifts, jump intensities, and jump size distributions, the specification (2.1) allows for diffusive, jump intensity, and jump size risk premia, respectively. Specifically, for diffusive risk premia, we follow the general affine risk premium specification of Cheridito et al. (2007). Absence of arbitrage further dictates restrictions on the risk-neutral drift of the stock price process:

- $A_{\mu,S}^{\mathbb{Q}} = r - q - \frac{1}{2}A_{\Omega,S} - (\mathit{\Phi}_\nu^{\mathbb{Q}}([1;0]) - \iota)^\top A_\lambda^{\mathbb{Q}}$ and

- $B^{\mathbb{Q}}_{\mu,S} = -\frac{1}{2} B_{\Omega,S} - (\Phi^{\mathbb{Q}}_{\nu}([1;0]) - \iota)^\top B^{\mathbb{Q}}_{\lambda}$

for instantaneous diffusive stock price variance $\Omega_S(z) = A_{\Omega,S} + B_{\Omega,S} z$, denoting by ι the vector of ones and by $\Phi^{\mathbb{M}}_{\nu}$ the vector of marginal jump transforms of jump sizes under the law $\nu^{\mathbb{M}}$. Specifically, the i-th element of $\Phi^{\mathbb{M}}_{\nu}(\omega) \in \mathbb{C}^{n_J}$ for $\omega \in \mathbb{C}^{n_X}$ is given by $\int \exp(\omega \cdot J_{\bullet i,X}) \, \mathrm{d}\nu^{\mathbb{M}}$. Under standard integrability conditions, this assures that $\exp((q-r)t) \, S_t$ is a $\mathbb{Q}$-martingale. As is customary within stochastic volatility models, we assume that interest rates and dividend yields are constant, given by r and q, respectively. It should be noted that this assumption is *not* prerequisite for our methodology, but simplifies the exposition.[10]

2.2 Specification examples

The specification of the affine stochastic volatility model in equation (2.1) covers many state-of-the-art models advocated in the literature. In what follows, we briefly discuss several popular models, but by no means attempt to provide an exhaustive overview.

2.2.1 Heston model

As the basis for all further examples, consider a jump diffusion extension of the Heston (1993) stochastic volatility model. The single latent state variable $Z_{1,t}$ within this model is the instantaneous diffusion variance of the stock price process, which follows a square-root process such that the state vector $X_t = [\log S_t; Z_{1,t}]$ satisfies

$$\mathrm{d} \log S_t = (b^{\mathbb{M}}_0 + b^{\mathbb{M}}_1 Z_{1,t-}) \, \mathrm{d}t + Z^{1/2}_{1,t-} \, \mathrm{d}\tilde{W}^{\mathbb{M}}_{1,t} + J_{11,X,t} \, \mathrm{d}N_{1,t} \tag{2.2a}$$

$$\mathrm{d}Z_{1,t} = \kappa^{\mathbb{M}}_1 (\theta^{\mathbb{M}}_1 - Z_{1,t-}) \, \mathrm{d}t + \varsigma_1 Z^{1/2}_{1,t-} \, \mathrm{d}W^{\mathbb{M}}_{2,t} + J_{21,X,t} \, \mathrm{d}N_{1,t} \, , \tag{2.2b}$$

where $\mathrm{d}\tilde{W}^{\mathbb{M}}_{1,t} = (1 - \rho^2_1)^{1/2} \, \mathrm{d}W^{\mathbb{M}}_{1,t} + \rho_1 \, \mathrm{d}W^{\mathbb{M}}_{2,t}$ with constant correlation ρ_1 and $\lambda^{\mathbb{M}}(z) = \lambda^{\mathbb{M}}_0 + \lambda^{\mathbb{M}}_1 z_1$. Equation (2.2) takes into account that jumps in stock prices and volatility tend to occur simultaneously with intensity depending on the variance level (e.g., Eraker (2004)). The jump sizes $J_{11,X,t}$ and $J_{21,X,t}$ are usually assumed to be independent, since estimation of their dependencies is notoriously difficult (e.g., Branger et al. (2010)). Typically, price jumps are assumed to be normally distributed or follow some fat tailed distribution such as a double exponential (e.g., Kou and Wang (2004)), while variance jumps are usually assumed to be exponentially distributed.

2.2.2 Volatility components

The Heston model, even after introducing jumps, is typically found to be unable to adequately capture important properties of stock returns and option prices. For this reason, Bates (2000), among others, considers a multivariate extension of the Heston model, in which two independent square-root processes $Z_{1,t}$ and $Z_{2,t}$ determine the instantaneous diffusion variance $Z_{1,t} + Z_{2,t}$ of the stock price process. The idea can be further extended to an arbitrary number of components. In the bivariate case, the state vector $X_t = [\log S_t; Z_{1,t}; Z_{2,t}]$ follows dynamics of the form

$$\mathrm{d} \log S_t = (b^{\mathbb{M}}_0 + b^{\mathbb{M}}_1 Z_{1,t-} + b^{\mathbb{M}}_2 Z_{2,t-}) \, \mathrm{d}t + (Z_{1,t-} + Z_{2,t-})^{1/2} \, \mathrm{d}\tilde{W}^{\mathbb{M}}_{1,t} + J_{11,X,t} \, \mathrm{d}N_{1,t} \tag{2.3a}$$

$$\mathrm{d}Z_{1,t} = \kappa^{\mathbb{M}}_1 (\theta^{\mathbb{M}}_1 - Z_{1,t-}) \, \mathrm{d}t + \varsigma_1 Z^{1/2}_{1,t-} \, \mathrm{d}W^{\mathbb{M}}_{2,t} + J_{21,X,t} \, \mathrm{d}N_{1,t} \tag{2.3b}$$

$$\mathrm{d}Z_{2,t} = \kappa^{\mathbb{M}}_2 (\theta^{\mathbb{M}}_2 - Z_{2,t-}) \, \mathrm{d}t + \varsigma_2 Z^{1/2}_{2,t-} \, \mathrm{d}W^{\mathbb{M}}_{3,t} + J_{21,X,t} \, \mathrm{d}N_{1,t} \, , \tag{2.3c}$$

[10] Our methodology is fully compatible with general affine interest rate and dividend yield specifications. The established results carry over with mostly minor modifications. Merely the derivation of the affine relation in lemma 3.2, which is required for the pricing of volatility derivatives, necessitates stronger assumptions regarding the interest rate process.

where $\mathrm{d}\tilde{W}^{\mathbb{M}}_{1,t} = (1 - \varrho^2_{1,t-} - \varrho^2_{2,t-})^{1/2}\,\mathrm{d}W^{\mathbb{M}}_{1,t} + \varrho_{1,t-}\,\mathrm{d}W^{\mathbb{M}}_{2,t} + \varrho_{2,t-}\,\mathrm{d}W^{\mathbb{M}}_{3,t}$ with stochastic correlation $\varrho_{i,t} = \rho_i\, Z^{1/2}_{i,t}/(Z_{1,t}+Z_{2,t})^{1/2}$. Moreover, jump intensities are given by $\lambda^{\mathbb{M}}(z) = \lambda^{\mathbb{M}}_0 + \lambda^{\mathbb{M}}_1 z_1 + \lambda^{\mathbb{M}}_2 z_2$. For simplicity, jumps in the stock price and both components are assumed to be simultaneous in equation (2.3), which may easily be relaxed. The same choices for the jump size distributions can be made as in the univariate Heston case.

2.2.3 Stochastic mean reversion

Taking a different route, Duffie et al. (2000) propose an extension of the original Heston model by replacing the deterministic mean reversion level $\theta^{\mathbb{M}}_1$ by a stochastic one, driven by an autonomous process. In this model, $Z_{1,t}$ represents the instantaneous diffusion variance of the stock price, whereas $Z_{2,t}$ determines the stochastic mean reversion level of $Z_{1,t}$. Since $Z_{2,t}$ is expected to reflect slowly moving trends in the variance level, it is usually assumed that $Z_{2,t}$ is continuous and does not affect the jump intensity. The resulting dynamics of the state vector $X_t = [\log S_t; Z_{1,t}; Z_{2,t}]$ are

$$\begin{aligned}
\mathrm{d}\log S_t &= (b^{\mathbb{M}}_0 + b^{\mathbb{M}}_1 Z_{1,t-})\,\mathrm{d}t + Z^{1/2}_{1,t-}\,\mathrm{d}\tilde{W}^{\mathbb{M}}_{1,t} + J_{11,X,t}\,\mathrm{d}N_{1,t} && (2.4a)\\
\mathrm{d}Z_{1,t} &= \kappa^{\mathbb{M}}_1([\kappa^{\mathbb{M}}_1]^{-1}\kappa^{\mathbb{Q}}_1 Z_{2,t-} - Z_{1,t-})\,\mathrm{d}t + \varsigma_1 Z^{1/2}_{1,t-}\,\mathrm{d}W^{\mathbb{M}}_{2,t} + J_{21,X,t}\,\mathrm{d}N_{1,t} && (2.4b)\\
\mathrm{d}Z_{2,t} &= \kappa^{\mathbb{M}}_2(\theta^{\mathbb{M}}_2 - Z_{2,t-})\,\mathrm{d}t + \varsigma_2 Z^{1/2}_{2,t-}\,\mathrm{d}W^{\mathbb{M}}_{3,t}\,, && (2.4c)
\end{aligned}$$

where $\mathrm{d}\tilde{W}^{\mathbb{M}}_{1,t} = (1-\rho^2_1)^{1/2}\,\mathrm{d}W^{\mathbb{M}}_{1,t} + \rho_1\,\mathrm{d}W^{\mathbb{M}}_{2,t}$ with constant correlation ρ_1 and $\lambda^{\mathbb{M}}(z) = \lambda^{\mathbb{M}}_0 + \lambda^{\mathbb{M}}_1 z_1$.

2.2.4 Autonomous jump intensities

In the examples considered so far, jump intensities are determined as an affine function of the components driving the instantaneous diffusion variance of the stock price process. The latent state dynamics in equation (2.1) also allow to specify jump intensities by an autonomous process, so that $Z_{1,t}$ determines the instantaneous diffusion variance of the stock price, while $Z_{2,t}$ determines the jump intensity. For the latter, the literature usually considers either a pure diffusion (e.g., Wachter (2013)) or a pure jump process (e.g., Aït-Sahalia et al. (2015b)). Allowing the dynamics of the jump intensity process to be driven by a jump diffusion, the state vector $X_t = [\log S_t; Z_{1,t}; Z_{2,t}]$ is governed by

$$\begin{aligned}
\mathrm{d}\log S_t &= (b^{\mathbb{M}}_0 + b^{\mathbb{M}}_1 Z_{1,t-} + b^{\mathbb{Q}}_2 Z_{2,t-})\,\mathrm{d}t + Z^{1/2}_{1,t-}\,\mathrm{d}\tilde{W}^{\mathbb{M}}_{1,t} + J_{11,X,t}\,\mathrm{d}N_{1,t} && (2.5a)\\
\mathrm{d}Z_{1,t} &= \kappa^{\mathbb{M}}_1(\theta^{\mathbb{M}}_1 - Z_{1,t-})\,\mathrm{d}t + \varsigma_1 Z^{1/2}_{1,t-}\,\mathrm{d}W^{\mathbb{M}}_{2,t} + J_{21,X,t}\,\mathrm{d}N_{1,t} && (2.5b)\\
\mathrm{d}Z_{2,t} &= \kappa^{\mathbb{M}}_2(\theta^{\mathbb{M}}_2 - Z_{2,t-})\,\mathrm{d}t + \varsigma_2 Z^{1/2}_{2,t-}\,\mathrm{d}W^{\mathbb{M}}_{3,t} + J_{31,X,t}\,\mathrm{d}N_{1,t}\,, && (2.5c)
\end{aligned}$$

where $\mathrm{d}\tilde{W}^{\mathbb{M}}_{1,t} = (1-\rho^2_1)^{1/2}\,\mathrm{d}W^{\mathbb{M}}_{1,t} + \rho_1\,\mathrm{d}W^{\mathbb{M}}_{2,t}$ with constant correlation ρ_1 and $\lambda^{\mathbb{M}}(z) = \lambda^{\mathbb{M}}_0 + \lambda^{\mathbb{M}}_2 z_2$.

2.3 Standard transform analysis

The methodology developed in this paper relies on the tractability of certain classes of moments of the state vector. Variants of the standard transform analysis for affine jump diffusions yield the required expressions for the generic dynamics (2.1). To make these accessible, this section briefly reviews the main results from standard transform analysis required for the remainder of this paper. Technical details are delegated to appendix A.

Before proceeding, we introduce some further notation. For a non-decreasing time vector $\tilde{\tau} \in \mathbb{R}^{\tilde{n}}_+$, corresponding to a non-decreasing sequence of time points $\tilde{\tau}_i$ such that $\tilde{\tau}_{i+1} \geq \tilde{\tau}_i \geq 0$ with the convention that $\tilde{\tau}_0 = 0$, define the stacked vectors $Y_{t+\tilde{\tau}} = [\log S_{t+\tilde{\tau}_1} - \log S_{t+\tilde{\tau}_0}; \ldots; \log S_{t+\tilde{\tau}_{\tilde{n}}} - \log S_{t+\tilde{\tau}_{\tilde{n}-1}}]$ and $Z_{t+\tilde{\tau}} = [Z_{t+\tilde{\tau}_1}; \ldots; Z_{t+\tilde{\tau}_{\tilde{n}}}]$. From these, moreover define the vector $X_{t\oplus\tilde{\tau}} = [Y_{t+\tilde{\tau}}; Z_{t+\tilde{\tau}}]$. Economically, the elements of $Y_{t+\tilde{\tau}}$ thus correspond to log returns between consecutive time points in $\tilde{\tau}$. In what follows, we state expressions for certain moments of $X_{t\oplus\tilde{\tau}}$.

We start the discussion by considering exponential moments, which can be derived from the standard transform analysis of Duffie et al. (2000); details and derivations of the expressions are provided in appendix A.2. Under the regularity conditions established by Duffie et al. (2000), single-period exponential moments can be determined from the solution of the system of ODEs (A.8) of generalized Riccati type (cf. propositions A.2 and A.3). Except for few special cases possessing closed-form solutions, the ODEs need to be solved numerically. Efficient numerical solution schemes can be based on vectorization techniques.

Building on these single-period moments, we can iteratively derive multi-period exponential moments of $X_{t\oplus\tilde{\tau}}$ via recursive relations (cf. propositions A.4 and A.5). The following proposition summarizes the construction. It should be noted that under the dynamics specified by equation (2.1), the conditional exponential moments of $X_{t\oplus\tilde{\tau}}$ directly depend only on the initial value of the latent state variable Z_t, but not on the initial level of the stock price S_t. We use this property to characterize unconditional exponential moments of $X_{t\oplus\tilde{\tau}}$ by a limiting procedure.

Proposition 2.1. *Consider an argument* $\omega \in \mathbb{C}^{n_X \tilde{n}}$.

(i) Let assumption A.3 hold for $\tau = 0$. *Then we have*

$$\begin{aligned}\Phi^{\mathbb{M}}(\omega; \tilde{\tau}, 0, Z_t) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t\oplus\tilde{\tau}}) \,|\, \mathcal{F}_t] \\ &= \exp(A_\Phi(\omega; \tilde{\tau}, 0) + B_\Phi(\omega; \tilde{\tau}, 0) \cdot Z_t)\end{aligned} \tag{2.6}$$

with coefficients $A_\Phi(\omega; \tilde{\tau}, 0) \in \mathbb{C}$ *and* $B_\Phi(\omega; \tilde{\tau}, 0) \in \mathbb{C}^{n_Z}$ *given in equation (A.11).*

(ii) Let assumption A.4 hold. Then we have

$$\begin{aligned}\Phi^{\mathbb{M}}(\omega; \tilde{\tau}, \infty) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t\oplus\tilde{\tau}})] \\ &= \exp(A_\Phi(\omega; \tilde{\tau}, \infty))\end{aligned} \tag{2.7}$$

with coefficient $A_\Phi(\omega; \tilde{\tau}, \infty) \in \mathbb{C}$ *given in equation (A.14).*

Following the terminology of Chen and Joslin (2012), we proceed to study the so-called polynomial-log-linear (henceforth pl-linear) moments. The derivation of the expressions in appendix A.3 heavily relies on a version of the Faà di Bruno formula (cf. proposition A.1), which is stated in appendix A.1. Extending the regularity conditions of the exponential case analogous to Dillschneider (2020), single-period pl-linear moments can be determined by solving the augmented system of ODEs (A.18) of generalized Riccati type (cf. propositions A.6 and A.7). This augmented system jointly characterizes the derivatives of the coefficients of single-period exponential moments in equation (A.8). Again, solving this system generally calls for a numerical solution procedure in conjunction with vectorization techniques.

Exploiting these single-period expressions allows to iteratively derive multi-period pl-linear moments of $X_{t\oplus\tilde{\tau}}$ via recursive relations (cf. propositions A.8 and A.9). The following proposition summarizes the construction. In essence, under the imposed regularity conditions, the respective moment expressions are formed by differentiation of the expressions obtained in proposition 2.1, so that $\Phi^{\mathbb{M},[\alpha]} = \Phi^{\mathbb{M},(\alpha)} = \partial_\omega^\alpha \Phi^{\mathbb{M}}$ holds. It is therefore not surprising that the resulting conditional pl-linear moments of $X_{t\oplus\tilde{\tau}}$ directly depend only on the initial value of the latent state variable Z_t, which once again allows to form unconditional pl-linear moments of $X_{t\oplus\tilde{\tau}}$ by a limiting argument.

Proposition 2.2. *Consider an argument* $\omega \in \mathbb{C}^{n_X \tilde{n}}$ *and a multi-index* $\alpha \in \mathbb{N}^{n_X \tilde{n}}$.

(i) Let assumption A.7 hold for $\tau = 0$. Then we have

$$
\begin{aligned}
\Phi^{\mathbb{M},[\alpha]}(\omega;\tilde{\tau},0,Z_t) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t\oplus\tilde{\tau}})\,(X_{t\oplus\tilde{\tau}})^{\alpha}\,|\,\mathcal{F}_t] \\
&= \Phi^{\mathbb{M}}(\omega;\tilde{\tau},0,Z_t) \sum_{\tilde{\mathcal{Q}}(\alpha)} M^{\alpha}_{k,\ell}\,(A^{(\ell)}_{\Phi}(\omega;\tilde{\tau},0) + B^{(\ell)}_{\Phi}(\omega;\tilde{\tau},0)\cdot Z_t)^k
\end{aligned}
\tag{2.8}
$$

with coefficients $A^{(\beta)}_{\Phi}(\omega;\tilde{\tau},0) \in \mathbb{C}$ and $B^{(\beta)}_{\Phi}(\omega;\tilde{\tau},0) \in \mathbb{C}^{n_Z}$ for $\beta \leq \alpha$ given in equation (A.21).

(ii) Let assumption A.8 hold. Then we have

$$
\begin{aligned}
\Phi^{\mathbb{M},[\alpha]}(\omega;\tilde{\tau},\infty) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t\oplus\tilde{\tau}})\,(X_{t\oplus\tilde{\tau}})^{\alpha}] \\
&= \Phi^{\mathbb{M}}(\omega;\tilde{\tau},\infty) \sum_{\tilde{\mathcal{Q}}(\alpha)} M^{\alpha}_{k,\ell}\,(A^{(\ell)}_{\Phi}(\omega;\tilde{\tau},\infty))^k
\end{aligned}
\tag{2.9}
$$

with coefficients $A^{(\beta)}_{\Phi}(\omega;\tilde{\tau},\infty)$ for $\beta \leq \alpha$ given in equation (A.24).

In equations (2.8) and (2.9), $\tilde{\mathcal{Q}}(\alpha) = \bigcup_{|\beta|\leq|\alpha|} \mathcal{Q}(\alpha,\beta)$ is a disjoint union and each $\mathcal{Q}(\alpha,\beta)$ is a set of multi-multi-indices k and ℓ, defined in equation (A.2). Moreover, $M^{\alpha}_{k,\ell}$ denotes the associated multi-multinomial coefficient, defined in equation (A.3). Finally, the tensor notation is defined in equation (A.4).

Evidently, polynomial moments may be computed as special cases of the pl-linear moments in proposition 2.2. This approach requires jointly solving systems of ODEs, which generally has to be performed numerically. Instead, closed-form expressions for polynomial moments can be obtained when treating affine jump diffusions as a particular instance of polynomial processes, which are formally introduced and studied in Cuchiero et al. (2012). Details of this approach are given in appendix A.4. Following Dillschneider (2020), single-period polynomial moments obtain from matrix expressions, which may be either determined analytically or numerically at virtually no computational cost (cf. propositions A.10 and A.11). Using these single-period expressions, multi-period polynomial moments of $X_{t\oplus\tilde{\tau}}$ result by running through recursive relations (cf. propositions A.12 and A.13). The upcoming proposition summarizes the construction.

Proposition 2.3. *Consider a multi-index $\alpha \in \mathbb{N}^{n_X\tilde{n}}$.*

(i) Let assumption A.11 hold for $\tau = 0$. Then we have

$$
\begin{aligned}
\Phi^{\mathbb{M},[\alpha]}(0;\tilde{\tau},0,Z_t) &= \mathrm{E}^{\mathbb{M}}[(X_{t\oplus\tilde{\tau}})^{\alpha}\,|\,\mathcal{F}_t] \\
&= \sum_{|\beta|\leq|\alpha|} b^{(\alpha)}_{\Phi,\beta}(\tilde{\tau},0)\,Z_t^{\beta}
\end{aligned}
\tag{2.10}
$$

with coefficients $b^{(\alpha)}_{\Phi,\beta}(\tilde{\tau},0) \in \mathbb{R}$ for $|\beta| \leq |\alpha|$ given in equation (A.33).

(ii) Let assumption A.12 hold. Then we have

$$
\begin{aligned}
\Phi^{\mathbb{M},[\alpha]}(0;\tilde{\tau},\infty) &= \mathrm{E}^{\mathbb{M}}[(X_{t\oplus\tilde{\tau}})^{\alpha}] \\
&= b^{(\alpha)}_{\Phi,0}(\tilde{\tau},\infty)
\end{aligned}
\tag{2.11}
$$

with coefficient $b^{(\alpha)}_{\Phi,0}(\tilde{\tau},\infty) \in \mathbb{R}$ given in equation (A.36).

3 Transform-based derivatives pricing

To present our methodology in a general form, we start with a unified theory for pricing a broad spectrum of derivatives. Our research agenda commands that the resultant pricing formula ought to be rather generic, but also analytically tractable and admitting a computationally efficient implementation. For this purpose, we rely on the generalized transform analysis introduced in Chen and Joslin (2012) and further studied in Dillschneider (2020), whose foundations are briefly reviewed in section 3.1; further details are discussed in Dillschneider (2020). To provide a common basis for the remainder of this paper, section 3.2 then derives a general transform-based derivatives pricing formula satisfying the requisite criteria. Subsequently, we specialize this formula to two important derivatives classes that occupy an exposed position within stochastic volatility modeling, namely equity derivatives in section 3.3 and volatility derivatives in section 3.4. Proofs of our results are given in appendix B. Further extensions are discussed in section 3.5, with details provided in appendix C.

3.1 Basic Schwartz distribution theory

By $\mathcal{S}(\mathbb{R}^m)$, referred to as the Schwartz space, we denote the space of rapidly decaying smooth functions,[11] regularly succinctly referred to as Schwartz functions. The associated continuous dual space $\mathcal{S}^*(\mathbb{R}^m)$, whose elements are called tempered distributions, contains all continuous linear functionals on $\mathcal{S}(\mathbb{R}^m)$. To denote the action of a tempered distribution $g \in \mathcal{S}^*(\mathbb{R}^m)$ on a Schwartz function $f \in \mathcal{S}(\mathbb{R}^m)$, we use the duality pairing notation $\langle g(y), f(y)\rangle$. A sufficiently well-behaved ordinary function g identifies with a regular tempered distribution via integration, $\langle g(y), f(y)\rangle = \int_{\mathbb{R}^m} g(y)\, f(y)\, \mathrm{d}y$. Using a regularization approach, this notion can be extended to a larger class of functions for which the ordinary integrals do not exist, yielding instances of a singular tempered distribution. Another prominent example in the class of singular tempered distributions is the Dirac delta functional δ, defined through the assignment $\langle \delta(y), f(y)\rangle = f(0)$.

Any Schwartz function $f \in \mathcal{S}(\mathbb{R}^m)$ has a Fourier transform $\hat{f} = \mathfrak{F}f \in \mathcal{S}(\mathbb{R}^m)$ with $\hat{f}(y) = \int_{\mathbb{R}^m} f(\tilde{y}) \exp(-\mathrm{i}y \cdot \tilde{y})\, \mathrm{d}\tilde{y}$. This allows to define the Fourier transform $\hat{g} = \mathfrak{F}g \in \mathcal{S}^*(\mathbb{R}^m)$ of the tempered distribution $g \in \mathcal{S}^*(\mathbb{R}^m)$ via the requirement that $\langle \hat{g}(y), f(y)\rangle = \langle g(y), \hat{f}(y)\rangle$ holds for all $f \in \mathcal{S}(\mathbb{R}^m)$.

For several of our applications, it is necessary to regard a Schwartz function as a function $f(y,z)$ of a pair of variables $y \in \mathbb{R}^m$, $z \in \mathbb{R}^n$. Such a function resides in the space $\mathcal{S}(\mathbb{R}^m \times \mathbb{R}^n)$. In that case, we may consistently define the distributional tensor product $((y,z) \mapsto g(y) \otimes h(z)) \in \mathcal{S}^*(\mathbb{R}^m \times \mathbb{R}^n)$ by

$$\langle g(y) \otimes h(z), f(y,z)\rangle = \langle g(y), \langle h(z), f(y,z)\rangle\rangle = \langle h(z), \langle g(y), f(y,z)\rangle\rangle\,. \tag{3.1}$$

This equation transports a distributional analogue to the classical Fubini integral theorem.

We extend the definitions above to subsets $\mathcal{Y} \subset \mathbb{R}^m$ as follows. The space $\mathcal{S}(\mathcal{Y})$ consists of all functions f such that there exists some $\tilde{f} \in \mathcal{S}(\mathbb{R}^m)$ coinciding with f on $\mathcal{Y}$. Likewise, the dual space $\mathcal{S}^*(\mathcal{Y})$ consists of those $g \in \mathcal{S}^*(\mathbb{R}^m)$ whose support is contained in $\mathcal{Y}$. Consequently, we can define $\langle g(y), f(y)\rangle = \langle g(y), \tilde{f}(y)\rangle$ for $g \in \mathcal{S}(\mathcal{Y})$ and $f \in \mathcal{S}^*(\mathcal{Y})$, where the choice of $\tilde{f} \in \mathcal{S}(\mathbb{R}^m)$ is inconsequential.

3.2 General derivatives

In order to compute derivatives prices corresponding to general payoff functions, we follow the standard risk-neutral pricing approach. For the ease of illustration, we suppose that all derivatives require premia to

[11] Formally, the requirement is that each semi-norm $\|\cdot\|_{\alpha,\beta}$ defined through $\|f\|_{\alpha,\beta} = \sup_{y\in\mathbb{R}^m} |y^\alpha\, \partial_y^\beta f(y)|$ is finite for any multi-indices $\alpha, \beta \in \mathbb{N}^m$.

be paid immediately at inception of the contract. Other empirically relevant features, such as futures-style margining, can be incorporated by minor modifications.[12]

To cover a broad class of relevant derivatives prices, fix a non-decreasing time vector $\tilde{T} \in \mathbb{R}^{\tilde{m}}$ and suppose that the derivative contract features a single payoff, which occurs $\tilde{T}_{\tilde{m}}$ periods ahead and is determined by $h(X_{t\oplus\tilde{T}}; K)$ for some payoff function h, where K is an additional parameter, e.g., representing the strike when considering an option. For the purpose of this paper, we restrict our attention to h of a particular form.

Assumption 3.1. *The payoff function h satisfies*

$$h(\tilde{x}; K) = \sum_{i=1}^{n_h} \exp(\bar{\omega}_i \cdot \tilde{x})\, g_i(\hat{\omega} \cdot \tilde{x}; K) \tag{3.2}$$

for $\bar{\omega}_i, \hat{\omega} \in \mathbb{R}^{n x \tilde{m}}$ and $(\tilde{y} \mapsto g_i(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$.

Within this setting, denote by $D_t(\tau) = \exp(-r\,\tau)$ the τ-period discount factor. Using $\Phi^{\mathbb{Q}}$ in the exponentially affine form of equation (2.6) under the conditions of proposition 2.1, we can then define the pricing transform for valuing derivatives written on $X_{t\oplus\tilde{T}}$ by

$$\begin{aligned} \Pi(\omega; \tilde{T}, Z_t) &= \mathrm{E}^{\mathbb{Q}}[D_t(\tilde{T}_{\tilde{m}})\, \exp(\omega \cdot X_{t\oplus\tilde{T}})\,|\,\mathcal{F}_t] \\ &= \exp(A_\Pi(\omega; \tilde{T}) + B_\Pi(\omega; \tilde{T}) \cdot Z_t)\;, \end{aligned} \tag{3.3}$$

where $A_\Pi(\omega; \tilde{T}) = A_\Phi^{\mathbb{Q}}(\omega; \tilde{T}) - r\,\tilde{T}_{\tilde{m}}$ and $B_\Pi(\omega; \tilde{T}) = B_\Phi^{\mathbb{Q}}(\omega; \tilde{T})$. In order to access the results of generalized transform analysis, we impose the following assumption on Π in equation (3.3).

Assumption 3.2. *$(y \mapsto \Pi(\mathfrak{b}(y); \tilde{T}, z)) \in \mathcal{S}(\mathcal{Y})$ for $\mathfrak{b}([\omega; \tilde{y}]) = \omega + \mathrm{i}\tilde{y}\hat{\omega}$, $\mathcal{Y} = \bigcup_{i=1}^{n_h}\{\bar{\omega}_i\} \times \mathbb{R}$, and all $z \in \mathcal{Z}$.*

With the general payoff function (3.2) and the pricing transform in equation (3.3), we are ready to approach derivatives pricing using the results from generalized transform analysis. The following proposition states a compact form of the associated price function $\mathcal{V}$ in terms of distributional tensor products.

Proposition 3.1. *Let assumptions 3.1 and 3.2 hold. Then we have*

$$\begin{aligned} \mathcal{V}(Z_t; K, \tilde{T}) &= \mathrm{E}^{\mathbb{Q}}[D_t(\tilde{T}_{\tilde{m}})\, h(X_{t\oplus\tilde{T}}; K)\,|\,\mathcal{F}_t] \\ &= \langle w(y; K), \Pi(\mathfrak{b}(y); \tilde{T}, Z_t)\rangle\;, \end{aligned} \tag{3.4}$$

where $y = [\tilde{\omega}; \tilde{y}]$ and $\mathfrak{b}([\tilde{\omega}; \tilde{y}]) = \tilde{\omega} + \mathrm{i}\tilde{y}\hat{\omega}$. Moreover, $(y \mapsto w(y; K)) \in \mathcal{S}^*(\mathcal{Y})$ is given by the distributional tensor product

$$w([\tilde{\omega}; \tilde{y}]; K) = \frac{1}{2\pi} \sum_{i=1}^{n_h} \delta(\tilde{\omega} - \bar{\omega}_i) \otimes \hat{g}_i(\tilde{y}; K)\;, \tag{3.5}$$

in terms of the distributional Fourier transforms $(\tilde{y} \mapsto \hat{g}_i(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$.

Expressing $\mathcal{V}$ in terms of a single tempered distribution w instead of a sum appears to be a purely cosmetic manipulation for the purpose of derivatives pricing, but will enormously simplify the analysis in subsequent sections when considering moments of derivatives prices. In the cases relevant for this paper, it will be possible to represent the action of the tempered distributions $\hat{g}_i$ in proposition 3.1 by regularized integrals. Since Π needs to be determined numerically in most cases, evaluation of the price

[12] E.g., for futures-style margining, the results hold with the risk-neutral transform $\Phi^{\mathbb{Q}}$ replacing the pricing transform Π defined in equation (3.3) below (e.g., Cox et al. (1981)).

function $\mathcal{V}$ in equation (3.4) will, therefore, generally require solving a series of one-dimensional numerical integration problems.

3.3 Equity derivatives

Equity derivatives written on the stock price constitute an important class of derivatives that is covered as a special case of section 3.2. To be precise, consider a plain-vanilla European option on the stock price $S_{t+\tau}$ for some $\tau \in \mathbb{R}_+$, whose price is normalized by the current stock price S_t in order to achieve stationarity.[13] Fixing the time vector $\tilde{T} = [\tau]$ as well as the log moneyness strike K, the call and put payoff of this option are given by $h^C_{\text{stock}}(X_{t\oplus\tilde{T}}; K)$ and $h^P_{\text{stock}}(X_{t\oplus\tilde{T}}; K)$, respectively, where

$$h^C_{\text{stock}}(\tilde{x}; K) = (\exp([1;0]\cdot\tilde{x}) - \exp(K))\,\mathfrak{U}([1;0]\cdot\tilde{x} - K) \tag{3.6a}$$

$$h^P_{\text{stock}}(\tilde{x}; K) = (\exp(K) - \exp([1;0]\cdot\tilde{x}))\,\mathfrak{U}(K - [1;0]\cdot\tilde{x})\,. \tag{3.6b}$$

Here, $\mathfrak{U}$ denotes the Heaviside step function. Each of the payoff functions in equation (3.6) satisfies the conditions of assumption 3.1 and can straightforwardly be expressed in the form of equation (3.2).

Denote the derivatives price associated to h^O_{stock} in equation (3.6) by $\mathcal{V}^O_{\text{stock}}$ for option type $O \in \{C, P\}$. The following corollary to proposition 3.1 yields an expression for $\mathcal{V}^O_{\text{stock}}$ as a special case of equation (3.4).

Corollary 3.1. *Let h^O_{stock} be as in equation (3.6). Moreover, let assumption 3.2 hold for $\bar{\omega}_1 = [1;0]$, $\bar{\omega}_2 = [0;0]$, $\hat{\omega} = [1;0]$. Then we have*

$$\mathcal{V}^O_{\text{stock}}(Z_t; K, \tilde{T}) = \langle w^O_{\text{stock}}(y; K), \Pi(\mathfrak{b}_{\text{stock}}(y); \tilde{T}, Z_t)\rangle\,, \tag{3.7}$$

where $y = [\tilde{\omega}; \tilde{y}]$ and $\mathfrak{b}_{\text{stock}}([\tilde{\omega}; \tilde{y}]) = \tilde{\omega} + \mathrm{i}\tilde{y}[1;0]$. The associated $(y \mapsto w^O_{\text{stock}}(y; K)) \in \mathcal{S}^*(\mathcal{Y})$ are given by

$$w^C_{\text{stock}}([\tilde{\omega}; \tilde{y}]; K) = (\delta(\tilde{\omega} - [1;0]) - \exp(K)\,\delta(\tilde{\omega})) \otimes (\tfrac{1}{2}\,\delta(\tilde{y}) + F^C_{\text{stock}}(\tilde{y}; K)\,(\mathrm{i}\tilde{y})^{-1}) \tag{3.8a}$$

$$w^P_{\text{stock}}([\tilde{\omega}; \tilde{y}]; K) = (\exp(K)\,\delta(\tilde{\omega}) - \delta(\tilde{\omega} - [1;0])) \otimes (\tfrac{1}{2}\,\delta(\tilde{y}) - F^P_{\text{stock}}(\tilde{y}; K)\,(\mathrm{i}\tilde{y})^{-1}) \tag{3.8b}$$

with $F^O_{\text{stock}}(\tilde{y}; K) = \frac{1}{2\pi}\exp(-\mathrm{i}K\tilde{y})$.

The statement of the preceding corollary 3.1 extends to the (prepaid) forward contract with payoff $h^C_{\text{stock}}(X_{t\oplus\tilde{T}}; -\infty)$. As a limiting case of the call price in corollary 3.1 when letting $K \to -\infty$, its price function is given by $\mathcal{V}^C_{\text{stock}}(Z_t; -\infty, \tilde{T}) = \Pi([1;0]; \tilde{T}, Z_t)$, with associated tempered distribution $w^C_{\text{stock}}([\tilde{\omega}; \tilde{y}]; -\infty) = \delta(\tilde{\omega} - [1;0]) \otimes \delta(\tilde{y})$.

For practical implementation, we give an integral representation of the tempered distribution w^O_{stock} in corollary 3.1. Applications arising in the further course of this paper not only require evaluation for Π as in equation (3.7), but also for other transforms. Therefore, the following lemma treats a generic transform Υ, which covers Π as a special case. When applied to Π, the integral representation of w^O_{stock} in equation (3.9) essentially recovers well-known transform-based option pricing formulas in, e.g., Bakshi and Madan (2000) and Duffie et al. (2000) (see also Chen and Joslin (2012) for further discussion).

Lemma 3.1. *Let $(y \mapsto \Upsilon(\mathfrak{b}_{\text{stock}}(y))) \in \mathcal{S}(\mathcal{Y})$. Then w^O_{stock} in corollary 3.1 can be represented in integral*

[13] This normalization is possible whenever the non-normalized option price is homogeneous of degree one in the initial stock price, which is the case for the specification in equation (2.1). Merton (1973) advocates such homogeneity as a natural property of option prices.

form as

$$\langle w^O_{\text{stock}}(y;K), \Upsilon(\mathfrak{b}_{\text{stock}}(y))\rangle = \left(\frac{c^O}{2}\Upsilon([1;0]) + \int_{\mathbb{R}_+} \frac{\Delta^{(1)}_{\tilde{y}}(F^O_{\text{stock}}(\tilde{y};K)\,\Upsilon([1;0]+\mathrm{i}\tilde{y}[1;0]))}{\mathrm{i}\tilde{y}}\,\mathrm{d}\tilde{y}\right) - \exp(K)\left(\frac{c^O}{2}\Upsilon([0;0]) + \int_{\mathbb{R}_+} \frac{\Delta^{(1)}_{\tilde{y}}(F^O_{\text{stock}}(\tilde{y};K)\,\Upsilon(\mathrm{i}\tilde{y}[1;0]))}{\mathrm{i}\tilde{y}}\,\mathrm{d}\tilde{y}\right), \tag{3.9}$$

with regularization $\Delta^{(1)}_{\tilde{y}} f(\tilde{y}) = f(\tilde{y}) - f(-\tilde{y})$ as well as option indicators $c^C = +1$ and $c^P = -1$.

Lemma 3.1 states the integral representation of $w^O_{\text{stock}}(y;K)$ for arbitrary generic transforms. The associated regularization requires evaluation of the transform along the whole real line. A computationally more efficient regularization may be employed whenever $\Upsilon(\mathfrak{b}_{\text{stock}}([\tilde{\omega};\tilde{y}]))$ is Hermitian as a function of $\tilde{y}$, which will generally be the case for the transforms considered in this paper. In that situation, equation (3.9) continues to hold when replacing $\Delta^{(1)}_{\tilde{y}}$ with $\tilde{\Delta}^{(1)}_{\tilde{y}}$ defined by $\tilde{\Delta}^{(1)}_{\tilde{y}} f(\tilde{y}) = 2\mathrm{i}\,\Im f(\tilde{y})$. The latter requires evaluation of the transform only along the positive real half-line.

3.4 Volatility derivatives

An important class of volatility derivatives is written on volatility indices constructed by the methodology of CBOE's VIX. To consistently model the evolution of the VIX associated to the state dynamics (2.1), we employ the usual theoretical representation of the squared VIX by a static portfolio comprising a continuum of out-of-the-money equity options (e.g., Carr and Madan (2001)). Fixing the reference period for the VIX at τ_{vix} equal to 30 calendar days, we obtain the representation

$$VIX^2_t = \frac{2\exp(r\tau_{\text{vix}})}{\tau_{\text{vix}}}\left(\int_{-\infty}^0 \exp(-K)\,\mathcal{V}^P_{\text{stock}}(Z_t;K,\tau_{\text{vix}})\,\mathrm{d}K + \int_0^\infty \exp(-K)\,\mathcal{V}^C_{\text{stock}}(Z_t;K,\tau_{\text{vix}})\,\mathrm{d}K\right). \tag{3.10}$$

This theoretical construction forms the conceptual basis for practical VIX-type indices, which aim at approximating the right-hand side of equation (3.10) using a finite number of observed option quotes.[14]

For the purpose of derivatives pricing, the thus constructed VIX is hardly tractable in the form of equation (3.10). As articulated in Carr and Wu (2009), however, it can be expressed in terms of a jump-adjusted quadratic variation of the (forward) stock price. Exploiting this insight, the following lemma establishes an affine dependence of VIX^2_t on the latent state vector Z_t for the state dynamics in equation (2.1).

Lemma 3.2. *Let VIX^2_t be given as in equation (3.10). It holds that*

$$VIX^2_t = a_{\text{vix}} + b_{\text{vix}} \cdot Z_t \tag{3.11}$$

for coefficients $a_{\text{vix}} \in \mathbb{R}$ and $b_{\text{vix}} \in \mathbb{R}^{nz}$ given in equation (B.26).

The affine relation in lemma 3.2 allows us to study the pricing of options written on the VIX as a special case of the results in section 3.2. Specifically, consider a plain-vanilla European option on the volatility index $VIX_{t+\tau}$ for some $\tau \in \mathbb{R}_+$. Fixing the time vector $\tilde{T} = [\tau]$ and the squared strike $K \geq 0$, we denote the call and put payoff of this option by $h^C_{\text{vix}}(X_{t\oplus\tilde{T}};K)$ and $h^P_{\text{vix}}(X_{t\oplus\tilde{T}};K)$, respectively. Using

[14] For an analysis of the approximation errors incurred in the practical realization of equation (3.10), the interested reader is referred to, e.g., Jiang and Tian (2005, 2007).

the affine expression for VIX_t^2 in equation (3.11), we have

$$h^C_{\mathrm{vix}}(\tilde{x}; K) = ((a_{\mathrm{vix}} + [0; b_{\mathrm{vix}}] \cdot \tilde{x})^{1/2} - K^{1/2})\, \mathfrak{U}((a_{\mathrm{vix}} + [0; b_{\mathrm{vix}}] \cdot \tilde{x}) - K) \tag{3.12a}$$

$$h^P_{\mathrm{vix}}(\tilde{x}; K) = (K^{1/2} - (a_{\mathrm{vix}} + [0; b_{\mathrm{vix}}] \cdot \tilde{x})^{1/2})\, \mathfrak{U}(K - (a_{\mathrm{vix}} + [0; b_{\mathrm{vix}}] \cdot \tilde{x}))\,. \tag{3.12b}$$

As before, $\mathfrak{U}$ denotes the Heaviside step function. Each of the payoff functions in equation (3.12) satisfies the conditions of assumption 3.1 and constitutes a special case of equation (3.2).

Denote the derivatives price associated to h^O_{vix} in equation (3.12) by $\mathcal{V}^O_{\mathrm{vix}}$ for option type $O \in \{C, P\}$. The following corollary to proposition 3.1 states an expression for $\mathcal{V}^O_{\mathrm{vix}}$ as a special case of equation (3.4).

Corollary 3.2. *Let h^O_{vix} be as in equation (3.12). Moreover, let assumption 3.2 hold for $\bar{\omega}_1 = [0; 0]$, $\hat{\omega} = [0; b_{\mathrm{vix}}]$. Then we have*

$$\mathcal{V}^O_{\mathrm{vix}}(Z_t; K, \tilde{T}) = \langle w^O_{\mathrm{vix}}(y; K), \Pi(\mathfrak{b}_{\mathrm{vix}}(y); \tilde{T}, Z_t)\rangle\,, \tag{3.13}$$

where $y = [\tilde{\omega}; \tilde{y}]$ and $\mathfrak{b}_{\mathrm{vix}}([\tilde{\omega}; \tilde{y}]) = \tilde{\omega} + \mathrm{i}\tilde{y}[0; b_{\mathrm{vix}}]$. The associated $(y \mapsto w^O_{\mathrm{vix}}(y; K)) \in \mathcal{S}^*(\mathcal{Y})$ are given by

$$w^C_{\mathrm{vix}}([\tilde{\omega}; \tilde{y}]; K) = \delta(\tilde{\omega}) \otimes F^C_{\mathrm{vix}}(\tilde{y}; K)\, (\mathrm{i}\tilde{y})^{-3/2} \tag{3.14a}$$

$$w^P_{\mathrm{vix}}([\tilde{\omega}; \tilde{y}]; K) = \delta(\tilde{\omega}) \otimes (K^{1/2}\, \delta(\tilde{y}) + F^P_{\mathrm{vix}}(\tilde{y}; K)\, (\mathrm{i}\tilde{y})^{-3/2}) \tag{3.14b}$$

with $F^C_{\mathrm{vix}}(\tilde{y}; K) = +\frac{1}{4\pi} \exp(\mathrm{i}a_{\mathrm{vix}}\tilde{y})\, \Gamma(1/2, \mathrm{i}K\tilde{y})$ and $F^P_{\mathrm{vix}}(\tilde{y}; K) = -\frac{1}{4\pi} \exp(\mathrm{i}a_{\mathrm{vix}}\tilde{y})\, \gamma(1/2, \mathrm{i}K\tilde{y})$. Here, Γ and γ denote the upper and lower incomplete Gamma function, respectively.

The (prepaid) forward contract with payoff $h^C_{\mathrm{vix}}(X_{t \oplus \tilde{T}}; 0)$ has the price price function $\mathcal{V}^C_{\mathrm{vix}}(Z_t; 0, \tilde{T})$, which results as a special case of the call price in corollary 3.2 when setting $K = 0$.

For practical implementation, we present an integral representation of the tempered distribution w^O_{vix} in corollary 3.2. Analogous to lemma 3.1, the upcoming lemma gives such a representation for a generic transform Υ, which covers Π in equation (3.13) as a special case. When applied to Π, the integral representation of w^O_{vix} in equation (3.15) yields a similar pricing formula as those in, e.g., Branger et al. (2016), Lian and Zhu (2013), Pacati et al. (2018), and Sepp (2008b).[15]

Lemma 3.3. *Let $(y \mapsto \Upsilon(\mathfrak{b}_{\mathrm{vix}}(y))) \in \mathcal{S}(\mathcal{Y})$. Then w^O_{vix} in corollary 3.2 can be represented in integral form as*

$$\langle w^O_{\mathrm{vix}}(y; K), \Upsilon(\mathfrak{b}_{\mathrm{vix}}(y))\rangle = \frac{1 - c^O}{2} K^{1/2}\, \Upsilon([0; 0]) + \int_{\mathbb{R}_+} \frac{\Delta^{(3/2)}_{\tilde{y}}(F^O_{\mathrm{vix}}(\tilde{y}; K)\, \Upsilon(\mathrm{i}\tilde{y}[0; b_{\mathrm{vix}}]))}{(\mathrm{i}\tilde{y})^{3/2}}\, \mathrm{d}\tilde{y}\,, \tag{3.15}$$

with regularization $\Delta^{(3/2)}_{\tilde{y}} f(\tilde{y}) = f(\tilde{y}) - \mathrm{i}\, f(-\tilde{y}) - (1 - \mathrm{i})\, f(0)$ as well as option indicators $c^C = +1$ and $c^P = -1$.

The integral representation of w^O_{vix} in lemma 3.3 may be employed for arbitrary generic transforms. For the special case of $\Upsilon(\mathfrak{b}_{\mathrm{vix}}([\tilde{\omega}; \tilde{y}]))$ being Hermitian as a function of $\tilde{y}$, we may alternatively devise a computationally more efficient regularization. Specifically, in that case, equation (3.15) continues to hold with $\tilde{\Delta}^{(3/2)}_{\tilde{y}}$ replacing $\Delta^{(3/2)}_{\tilde{y}}$, where we define $\tilde{\Delta}^{(3/2)}_{\tilde{y}} f(\tilde{y}) = (1 - \mathrm{i})\, (\Re f(\tilde{y}) - \Im f(\tilde{y}) - f(0))$.

[15] Related pricing formulas are also derived for the case of derivatives on quadratic variation (e.g., Broadie and Jain (2008), Carr and Lee (2009), Friz and Gatheral (2005), and Sepp (2008a)).

3.5 Extensions using complex Fourier theory

The derivatives pricing formula established in proposition 3.1 is given in terms of distributional Fourier transforms. For important special cases, such as the equity and volatility derivatives treated in corollaries 3.1 and 3.2, respectively, these distributional Fourier transforms may admit regularized integral representations, as in lemmas 3.1 and 3.3. In certain cases, the derivatives pricing formula can even be rewritten in terms of ordinary (square-integrable) Fourier transforms, admitting a regular integral representation. Appendix C discusses the case of functions exhibiting exponential growth using so-called complex Fourier theory.

Specifically, suppose that the payoff function h in the form of equation (3.2) may be written in terms of functions g_i exhibiting exponential growth. By exponential scaling, we can easily construct another representation of h in terms of square-integrable functions in accordance with assumption 3.1. Given that the pricing transform satisfies assumption 3.2 after scaling, proposition 3.1 justifies a derivatives pricing formula (3.4) in the integral form

$$\mathcal{V}(Z_t; K, \tilde{T}) = \int_{\mathcal{Y}_\varepsilon} w_\varepsilon(y; K)\, \Pi(\mathfrak{b}(y); \tilde{T}, Z_t)\, \mathrm{d}y\,, \tag{3.16}$$

with w_ε given analogous to equation (3.5) and employing the functional definition of the Dirac delta. Here, ε reflects the exponential scaling. By construction, w_ε is now given in terms of the (square-integrable) complex Fourier transforms of g_i, whereas the respective ordinary Fourier transforms may not exist. In essence, the derivatives price in equation (3.16) can thus be represented as a sum of ordinary integrals.

Particular examples of functions exhibiting exponential growth arise for the equity and volatility derivatives considered in sections 3.3 and 3.4, respectively. Following an exponential scaling procedure that is thoroughly outlined in appendix C, we can establish regular transform-based pricing formulas in the form of equation (3.16). We provide these formulas in corollaries C.1 and C.2, leading to the ordinary integral representations in lemmas C.3 and C.4, respectively. Intuitively, one may think of the distributional pricing formulas and their regularized integral representations derived in sections 3.3 and 3.4 as limiting cases (as $\varepsilon \to 0$) of the corresponding ordinary ones.

In principle, a derivation of the thus discussed results can rely on the distribution-based theory of this section. What appendix C conveys in addition is that these results may as well be derived by largely relying on distribution-free techniques from ordinary Fourier and integration theory. When limiting the attention to payoff functions of this particular form, one may even completely abandon Schwartz distribution theory and instead rely on ordinary Fourier and integration techniques, at the expense of lower generality with respect to the structure of payoff functions and more intricate regularity conditions. Details of this are discussed in Dillschneider (2020).

4 Moments involving derivatives prices

Based on the unified derivatives pricing theory established in section 3, this section develops expressions for moments involving polynomials of derivatives prices, which will form the basis for section 5, where we devise our GMM-type estimation approach. Section 4.1 describes the basic setup for studying moments of derivatives prices. For determining these moments, section 4.2 introduces an extension of Schwartz distribution theory. Section 4.3 then derives expressions for exact moments that will be shown to be analytically tractable, but computationally feasible only for low orders. Nevertheless, the derived expressions can be used to develop an effective approximation procedure in section 4.4. Finally, to make our results more easily accessible, section 4.5 studies several concrete examples. While not the focus of our presentation, section 4.6 shows that our methodology can straightforwardly be extended to include realistic measurement errors in derivatives prices. Derivations and proofs are contained in appendix D.

4.1 Basic setup

To cover a broad class of interesting moments, we now turn to a general setting in which multiple derivatives prices are available at each given date. Specifically, consider a vector of derivatives prices at time t, denoted by V_t and taking values in $\mathbb{R}^{n_V}$. Each of its elements may correspond to a different underlying and contract specification. To preserve generality of our results, we merely require that all derivatives prices are determined according to the general formula in proposition 3.1. By equation (3.4), each element of V_t can thus be expressed as

$$V_{i,t} = \mathcal{V}_i(Z_t; K_i, \tilde{T}_i) = \langle w_i(y_i; K_i), \Pi(\mathfrak{b}_i(y_i); \tilde{T}_i, Z_t)\rangle \tag{4.1}$$

in terms of a tempered distribution $(y_i \mapsto w_i(y_i; K_i)) \in \mathcal{S}^*(\mathcal{Y}_i)$ and the pricing transform Π in equation (3.3). For the expression in equation (4.1) to be well-defined, we suppose that assumption 3.2 holds accordingly, so that $(y_i \mapsto \Pi(\mathfrak{b}_i(y_i); \tilde{T}_i, z)) \in \mathcal{S}(\mathcal{Y}_i)$ for any $z \in \mathcal{Z}$. As before, for a non-decreasing time vector $\tilde{\tau} \in \mathbb{R}^{\tilde{n}}$, we moreover construct the stacked vector $V_{t+\tilde{\tau}} = [V_{t+\tilde{\tau}_1}; \ldots; V_{t+\tilde{\tau}_{\tilde{n}}}]$.

4.2 Extended Schwartz distribution theory

Expressions for moments involving derivatives prices may be derived under ordinary Schwartz distribution theory. However, it turns out that we may further relax the required conditions. As in Dillschneider (2020), we therefore extend the notion of a Schwartz space to include also functions that are Schwartz only after appropriate regularization.

Formally, we denote by $\tilde{\mathcal{S}}(\mathbb{R}^m; v)$ an extended Schwartz space as follows. Take some positive weighting function $v \in \mathcal{C}^\infty(\mathbb{R}^m)$. Construct the space $\tilde{\mathcal{S}}(\mathbb{R}^m; v)$ to contain all smooth functions $f \in \mathcal{C}^\infty(\mathbb{R}^m)$ such that $\tilde{f} = vf \in \mathcal{S}(\mathbb{R}^m)$. When v is (at most) slowly increasing, we indeed have that $\tilde{\mathcal{S}}(\mathbb{R}^m; v) \supset \mathcal{S}(\mathbb{R}^m)$. The associated continuous dual space $\tilde{\mathcal{S}}^*(\mathbb{R}^m; v)$ yields the space of extended tempered distributions. For elements $g \in \tilde{\mathcal{S}}^*(\mathbb{R}^m; v)$ and $f \in \tilde{\mathcal{S}}(\mathbb{R}^m; v)$, we define the action of g on f via the identity $\langle g(y), f(y)\rangle = \langle \tilde{g}(y), \tilde{f}(y)\rangle$, where $\tilde{g} = v^{-1}g \in \mathcal{S}^*(\mathbb{R}^m)$ and $\tilde{f} = vf \in \mathcal{S}(\mathbb{R}^m)$. When v is (at most) slowly increasing, we have that $\tilde{\mathcal{S}}^*(\mathbb{R}^m; v) \subset \mathcal{S}^*(\mathbb{R}^m)$.

The above construction carries over to tensor products on extended Schwartz spaces $\tilde{\mathcal{S}}(\mathbb{R}^m \times \mathbb{R}^n; v \otimes u)$ with positive weighting functions $v \in \mathcal{C}^\infty(\mathbb{R}^m)$ and $u \in \mathcal{C}^\infty(\mathbb{R}^n)$. Specifically, the tensor product then satisfies $\langle g(y) \otimes h(z), f(y, z)\rangle = \langle \tilde{g}(y) \otimes \tilde{h}(z), \tilde{f}(y, z)\rangle$, where the right-hand-side is given by equation (3.1) with $\tilde{g} = v^{-1}g \in \mathcal{S}^*(\mathbb{R}^m)$, $\tilde{h} = u^{-1}h \in \mathcal{S}^*(\mathbb{R}^n)$, and $\tilde{f} = (v \otimes u)f \in \mathcal{S}(\mathbb{R}^m \times \mathbb{R}^n)$. It follows by this construction that an analogue of equation (3.1) holds in extended Schwartz spaces.

The definitions extend naturally to subsets $\mathcal{Y} \subset \mathbb{R}^m$. In that case, the space $\tilde{\mathcal{S}}(\mathcal{Y}; v)$ consists of all functions f such that there exists some $\tilde{f} \in \mathcal{S}(\mathbb{R}^m)$ coinciding with f on $\mathcal{Y}$. Likewise, the dual space $\tilde{\mathcal{S}}^*(\mathcal{Y}; v)$ consists of the $g \in \tilde{\mathcal{S}}^*(\mathbb{R}^m; v)$ whose support is contained in $\mathcal{Y}$. Consequently, we can define $\langle g(y), f(y)\rangle = \langle g(y), \tilde{f}(y)\rangle$ for $g \in \tilde{\mathcal{S}}^*(\mathcal{Y}; v)$ and $f \in \tilde{\mathcal{S}}(\mathcal{Y}; v)$, where the concrete choice of $\tilde{f} \in \tilde{\mathcal{S}}(\mathbb{R}^m; v)$ is inconsequential.

4.3 Exact moments

Within the setting presented in section 4.1, our first result establishes that monomials of $V_{t+\tilde{\tau}}$ can again be written in terms of a tempered distribution applied to some associated pricing transform.

Lemma 4.1. *For any multi-index $\beta \in \mathbb{N}^{n_V}$, we have that*

$$(V_{t+\tilde{\tau}})^{\beta} = \langle w^{\beta}(y; K), \Pi^{\beta}(\mathfrak{b}(y); \tilde{T}, Z_{t+\tilde{\tau}}) \rangle \,, \tag{4.2}$$

where the tempered distribution $(y \mapsto w^{\beta}(y; K)) \in \mathcal{S}^*(\mathcal{Y}^{\beta})$ is given in equation (D.7). Moreover, the associated pricing transform $(y \mapsto \Pi^{\beta}(\mathfrak{b}(y); \tilde{T}, z)) \in \mathcal{S}(\mathcal{Y}^{\beta})$ for any $z \in \mathcal{Z}^{\tilde{n}}$ has the form

$$\Pi^{\beta}(\mathfrak{b}(y); \tilde{T}, Z_{t+\tilde{\tau}}) = \exp(A^{\beta}_{\Pi}(\mathfrak{b}(y); \tilde{T}) + B^{\beta}_{\Pi}(\mathfrak{b}(y); \tilde{T}) \cdot Z_{t+\tilde{\tau}}) \,, \tag{4.3}$$

where A^{β}_{Π} and B^{β}_{Π} are given in equation (D.9).

The tempered distribution w^{β} in lemma 4.1 essentially equals a tensor product of the tempered distributions w_i, with multiplicities being determined by the multi-index β. As the action of each w_i can generally be represented by a one-dimensional integral, the action of w^{β} can accordingly be expressed in terms of a $|\beta|$-dimensional integral.

Using lemma 4.1, we are now in place to determine joint moments of state variables $X_{t\oplus\tilde{\tau}}$ and derivatives prices $V_{t+\tilde{\tau}}$. Thereby, we extend the class of analytically tractable moments beyond those introduced in section 2.3. To arrive at the desired result, we rely on the extended Schwartz theory in the sense of section 4.2, suggesting the following regularity conditions.

Assumption 4.1. *There exists positive $q_{\beta} \in \mathcal{C}^{\infty}(\mathcal{Z}^{\tilde{n}})$ satisfying the following conditions:*

(i) $((y, z) \mapsto \Pi^{\beta}(\mathfrak{b}(y); \tilde{T}, z)) \in \tilde{\mathcal{S}}(\mathcal{Y}^{\beta} \times \mathcal{Z}^{\tilde{n}}; \mathbb{1} \otimes q_{\beta})$;

(ii) $\mathrm{E}^{\mathbb{M}}[|\exp(\omega \cdot X_{t\oplus\tilde{\tau}})\,(X_{t\oplus\tilde{\tau}})^{\alpha}|\, q_{\beta}(Z_{t+\tilde{\tau}})^{-1}] < \infty$.

The upcoming proposition states an extension of the unconditional pl-linear moments in equation (2.9). With obvious modifications, an equivalent expression can be derived for conditional pl-linear moments, which are, however, of minor importance for the purpose of our paper.

Proposition 4.1. *Consider an argument $\omega \in \mathbb{C}^{n_X\tilde{n}}$ as well as multi-indices $\alpha \in \mathbb{N}^{n_X\tilde{n}}$ and $\beta \in \mathbb{N}^{n_V\tilde{n}}$. Let assumption 4.1 hold. Then we have*

$$\begin{aligned} \tilde{\Phi}^{\mathbb{M},[\alpha,\beta]}(\omega, 0; \tilde{\tau}, \infty) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t\oplus\tilde{\tau}})\,(X_{t\oplus\tilde{\tau}})^{\alpha}\,(V_{t+\tilde{\tau}})^{\beta}] \\ &= \langle w^{\beta}(y; K), \exp(A^{\beta}_{\Pi}(\mathfrak{b}(y); \tilde{T}))\,\Phi^{\mathbb{M},[\alpha]}(\omega + [0; B^{\beta}_{\Pi}(\mathfrak{b}(y); \tilde{T})]; \tilde{\tau}, \infty)\rangle \end{aligned} \tag{4.4}$$

with w^{β}, A^{β}_{Π}, and B^{β}_{Π} given in lemma 4.1.

The crucial result leading to these moments is the interchange of the tempered distribution and the expectation operator in equation (4.4), which can be justified using the extended Schwartz theory (cf. appendix D.1). For evaluating the integrand in equation (4.4), $\Phi^{\mathbb{M},[\alpha]}$ can be determined from equation (2.9) under the conditions of proposition 2.2. While $\tilde{\Phi}^{\mathbb{M},[\alpha,\beta]}$ thereby admits an analytically tractable expression for arbitrary multi-indices (α, β), in general, its computation is only feasible for low orders of β, since w^{β} requires $|\beta|$-dimensional numerical integration.

4.4 Approximate moments

To avoid the computational cost of the exact pl-linear moments involving derivatives prices in section 4.3 while exploiting the feasibility of low-order moments, this section proposes an effective polynomial approximation approach.

Take $\mathcal{L}^2(\mathcal{Z},\mathbb{M})$ to be the set of square-integrable functions on $\mathcal{Z}$ against the probability measure $\mathbb{M}$, i.e., comprising all f satisfying $\mathrm{E}^{\mathbb{M}}[|f(Z_t)|^2] < \infty$, where the choice of t is arbitrary due to stationarity. In order to assure that functions in $\mathcal{L}^2(\mathcal{Z},\mathbb{M})$ can be approximated by monomials of Z_t, we impose the following standard assumption, under which $\mathbb{M}$ is said to have exponential tails.

Assumption 4.2. $\mathrm{E}^{\mathbb{M}}[\exp(\epsilon\,\|Z_t\|)] < \infty$ *for some* $\epsilon > 0$.

As a consequence of assumption 4.2, the set of monomials $\{z^\gamma : \gamma \in \mathbb{N}^{n_Z}\}$ forms a basis for $\mathcal{L}^2(\mathcal{Z},\mathbb{M})$ (e.g., theorem 3.2.18 in Dunkl and Xu (2014)). Employing the well-known Gram-Schmidt procedure, the set of monomials can be transformed into an orthonormal basis $\{\phi_\gamma(z) : \gamma \in \mathbb{N}^{n_Z}\}$. By construction, we have $\phi_\gamma(z) = \sum_{\eta \preccurlyeq \gamma} b^{(\gamma)}_{\phi,\eta}\, z^\eta$ for coefficients $b^{(\gamma)}_{\phi,\eta} \in \mathbb{R}$ depending on the unconditional monomial moments of Z_t under $\mathbb{M}$ up to order $2|\gamma|$, with $\preccurlyeq$ denoting the lexicographic order. Under the conditions of proposition 2.3, these moments can be computed in closed form by equation (2.11).

In order to approximate derivatives prices, we need to assure that the price functions are contained in $\mathcal{L}^2(\mathcal{Z},\mathbb{M})$. Hence, we further impose the following assumption on $\mathcal{V}_i$ in equation (4.1).

Assumption 4.3. $(z \mapsto \mathcal{V}_i(z; K_i, \tilde{T}_i)) \in \mathcal{L}^2(\mathcal{Z},\mathbb{M})$ *for all* $1 \le i \le n_V$.

Combining assumptions 4.2 and 4.3, we construct an approximant $V_{t,(p)}$ for V_t by projecting each of its elements $V_{i,t} = \mathcal{V}_i(Z_t; K_i, \tilde{T}_i)$ onto the truncated set of basis functions $\{\phi_\gamma(z) : \gamma \in \mathbb{N}^{n_Z}, |\gamma| \le p\}$. Due to stationarity, the projection is independent of the particular choice of t. We summarize the construction of $V_{t,(p)}$ in the upcoming lemma, whose proof is standard and thus omitted.

Lemma 4.2. *Let assumptions 4.2 and 4.3 hold. Then* $V_{t,(p)}$ *is given as*

$$V_{t,(p)} = \sum_{|\eta| \le p} \tilde{c}_{V,\eta}\, \phi_\eta(Z_t) \tag{4.5}$$

with $\tilde{c}_{V,\eta} = \mathrm{E}^{\mathbb{M}}[V_t\, \phi_\eta(Z_t)] \in \mathbb{R}^{n_V}$. *Moreover,* $V_{t,(p)} \to V_t$ *elementwise in* $\mathcal{L}^2(\mathcal{Z},\mathbb{M})$ *as* $p \to \infty$.

Using lemma 4.2, we can now construct an approximant $V_{t+\tilde{\tau},(p)}$ for $V_{t+\tilde{\tau}}$, given a non-decreasing time vector $\tilde{\tau} \in \mathbb{R}^{\tilde{n}}$. By a change of basis, equation (4.5) can be expressed as

$$V_{t,(p)} = \sum_{|\eta| \le p} \tilde{b}_{V,\eta,(p)}\, (Z_t)^\eta \tag{4.6}$$

for $\tilde{b}_{V,\eta,(p)} \in \mathbb{R}^{n_V}$, depending on the expansion order p. Constructing $V_{t+\tilde{\tau}_j,(p)}$ as in equation (4.6) separately for every $1 \le j \le \tilde{n}$, we define the stacked vector $V_{t+\tilde{\tau},(p)} = [V_{t+\tilde{\tau}_1,(p)}; \ldots; V_{t+\tilde{\tau}_{\tilde{n}},(p)}]$. Padding the coefficients in equation (4.6) with zeros then yields

$$V_{t+\tilde{\tau},(p)} = \sum_{|\eta| \le p} b_{V,\eta,(p)}\, (Z_{t+\tilde{\tau}})^\eta \tag{4.7}$$

for $b_{V,\eta,(p)} \in \mathbb{R}^{n_V \tilde{n}}$. Monomials of $V_{t+\tilde{\tau},(p)}$ in equation (4.7) thus obtain as polynomials in $Z_{t+\tilde{\tau}}$, given by

$$(V_{t+\tilde{\tau},(p)})^\beta = \sum_{|\eta| \le p|\beta|} b^{(\beta)}_{V,\eta,(p)}\, (Z_{t+\tilde{\tau}})^\eta \tag{4.8}$$

for $b^{(\beta)}_{V,\eta,(p)} \in \mathbb{R}$ determined as polynomials of the coefficients $\tilde{c}_{V,\eta}$ in equation (4.5).

It is now a natural question whether the proposed approximation of monomials of $V_{t+\tilde{\tau}}$ via $V_{t+\tilde{\tau},(p)}$ in equation (4.8) yields a sensible approximation of pl-linear moments involving derivatives prices. In general, elementwise convergence in the $\mathcal{L}^2(\mathcal{Z}, \mathbb{M})$ sense does not imply convergence of the associated pl-linear moments, unless an additional regularity condition is imposed.

Assumption 4.4. $(\exp(\omega \cdot X_{t\oplus\tilde{\tau}})\,(X_{t\oplus\tilde{\tau}})^{\alpha}\,(V_{t+\tilde{\tau},(p)})^{\beta})_p$ *is uniformly integrable.*[16]

Under the additional condition in assumption 4.4, the upcoming proposition formalizes the aspired moment approximation procedure. Approximate pl-linear moments involving the derivatives prices $V_{t+\tilde{\tau}}$ can be obtained by computing exact pl-linear moments involving the approximant $V_{t+\tilde{\tau},(p)}$. The resulting sequence of moment approximants converges to the exact moments derived in proposition 4.1.

Proposition 4.2. *Consider an argument $\omega \in \mathbb{C}^{n_X \tilde{n}}$ as well as multi-indices $\alpha \in \mathbb{N}^{n_X \tilde{n}}$ and $\beta \in \mathbb{N}^{n_V \tilde{n}}$. Let assumptions 4.2 to 4.5 hold. Then we have that*

$$\begin{aligned}\tilde{\Phi}^{\mathbb{M},[\alpha,\beta]}_{(p)}(\omega, 0; \tilde{\tau}, \infty) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t\oplus\tilde{\tau}})\,(X_{t\oplus\tilde{\tau}})^{\alpha}\,(V_{t+\tilde{\tau},(p)})^{\beta}] \\ &= \sum_{|\eta| \le p|\beta|} b^{(\beta)}_{V,\eta,(p)}\,\Phi^{\mathbb{M},[\alpha+[0;\eta]]}(\omega; \tilde{\tau}, \infty)\end{aligned} \tag{4.9}$$

with $b^{(\beta)}_{V,\eta,(p)}$ given in equation (D.13) satisfies $\tilde{\Phi}^{\mathbb{M},[\alpha,\beta]}_{(p)} \to \tilde{\Phi}^{\mathbb{M},[\alpha,\beta]}$ as $p \to \infty$.

Except for the coefficients $b^{(\beta)}_{V,\eta,(p)}$, the approximate pl-linear moment $\tilde{\Phi}^{\mathbb{M},[\alpha,\beta]}_{(p)}$ in equation (4.9) does not require the evaluation of any moments involving derivatives prices. Only the pl-linear moments $\Phi^{\mathbb{M},[\alpha+[0;\eta]]}$ for $|\eta| \le p|\beta|$ need to be computed, which can be achieved at low computational cost as in equation (2.9) under the conditions of proposition 2.2, or even in closed form as in equation (2.11) under the conditions of proposition 2.3 when $\omega = 0$.

It remains to establish a practicable procedure for computing the coefficients $b^{(\beta)}_{V,\eta,(p)}$, which are polynomials of the coefficients $\tilde{c}_{V,\eta}$ in lemma 4.2, in order to evaluate the approximate moments in proposition 4.2. For this purpose, we rely on the exact moments derived in section 4.3 and, hence, impose the following regularity conditions in order to access these results.

Assumption 4.5. *For every $1 \le i \le n_V$ there exists positive $q_i \in \mathcal{C}^{\infty}(\mathcal{Z})$ satisfying the following conditions for all $\gamma \in \mathbb{N}^{n_V}$ with $|\gamma| \le p$:*

(i) $((y_i, z) \mapsto \Pi(\mathfrak{b}_i(y_i); \tilde{T}_i, z)) \in \tilde{\mathcal{S}}(\mathcal{Y}_i \times \mathcal{Z}; \mathbb{1} \otimes q_i)$*;*

(ii) $\mathrm{E}^{\mathbb{M}}[|(Z_t)^{\gamma}|\,q_i(Z_t)^{-1}] < \infty$.

The conditions in assumption 4.5 are derived from assumption 4.1 and thereby allow to exploit the exact moment expressions stated in proposition 4.1. In this respect, it is important to note that only moments involving first-order polynomials in derivatives prices need to be evaluated.

Lemma 4.3. *Fix $p \in \mathbb{N}$. Let assumptions 4.2, 4.3 and 4.5 hold. Then $\tilde{c}_{V,\eta}$ in equation (4.5) for every $|\eta| \le p$ is given by*

$$\tilde{c}_{V,\eta} = \sum_{i=1}^{n_V} e_i \sum_{\gamma \preccurlyeq \eta} b^{(\eta)}_{\phi,\gamma}\,\langle w_i(y_i; K_i), \exp(A_{\Pi}(\mathfrak{b}_i(y_i); \tilde{T}_i))\,\Phi^{\mathbb{M},[[0;\gamma]]}([0; B_{\Pi}(\mathfrak{b}_i(y_i); \tilde{T}_i)]; 0, \infty)\rangle\,, \tag{4.10}$$

where $e_i \in \mathbb{N}^{n_V}$ denotes the i-th standard unit vector.

In order to determine $V_{t,(p)}$, it is necessary according to equation (4.10) to compute moments of the form $\tilde{\Phi}^{\mathbb{M},[[0;\gamma],e_i]}$ as in equation (4.4) for each combination of i and $|\gamma| \le p$. Hence, in general, $V_{t,(p)}$ can be

[16]A sequence $(\xi_p)_p$ is called uniformly integrable whenever $\sup_p \mathrm{E}^{\mathbb{M}}[|\xi_p|\,\mathfrak{U}(|\xi_p| - K)] \to 0$ as $K \to \infty$.

computed by performing a series of one-dimensional numerical integration problems. Lemma 4.3 thereby yields a computationally feasible procedure to compute the moment approximation in proposition 4.2.

4.5 Examples

To illustrate our theoretical results and make them more easily accessible, we now briefly discuss three groups of examples with increasing complexity. For each of these groups, we without further notice suppose that the conditions of proposition 4.1 hold, so that moments involving derivatives prices can be determined according to equation (4.4).

4.5.1 Single-period, first-order moments

As a first group of examples, we illustrate polynomial moments involving the derivatives price $V_{i,t+\tilde{\tau}_1}$ for $\tilde{\tau}_1 \geq 0$ and $1 \leq i \leq n_V$. For this, consider a time vector $\tilde{\tau} = [\tilde{\tau}_1] \in \mathbb{R}_+$ and a first-order price moment with $\beta = e_i \in \mathbb{N}^{n_V}$. Further setting $\alpha = [\alpha_1] \in \mathbb{N}^{n_X}$ in equation (4.4) yields the polynomial moment

$$\begin{aligned} \tilde{\Phi}^{\mathbb{M},[\alpha,e_i]}(0,0;\tilde{\tau},\infty) &= \mathrm{E}^{\mathbb{M}}[(X_{t\oplus[\tilde{\tau}_1]})^{\alpha_1}\, V_{i,t+\tilde{\tau}_1}] \\ &= \langle w^{e_i}(y;K), \exp(A_\Pi^{e_i}(\mathfrak{b}(y);\tilde{T}))\, \Phi^{\mathbb{M},[\alpha]}([0;B_\Pi^{e_i}(\mathfrak{b}(y);\tilde{T})];\tilde{\tau},\infty)\rangle\,, \end{aligned} \tag{4.11}$$

with lemma 4.1 providing the required expressions for w^{e_i}, $A_\Pi^{e_i}$, and $B_\Pi^{e_i}$:

$$\begin{aligned} w^{e_i}(y;K) &= w_i(y_1;K_i) \\ A_\Pi^{e_i}(\mathfrak{b}(y);\tilde{T}) &= A_\Pi(\mathfrak{b}_i(y_1);\tilde{T}_i) \\ B_\Pi^{e_i}(\mathfrak{b}(y);\tilde{T}) &= B_\Pi(\mathfrak{b}_i(y_1);\tilde{T}_i)\,. \end{aligned}$$

Here, w^{e_i} corresponds to the tempered distribution w_i associated to $V_{i,t}$, while $A_\Pi^{e_i}$ and $B_\Pi^{e_i}$ are the respective coefficients of the pricing transform.

It remains to determine a tractable expression for the standard transform in equation (4.11), using the general results in section 2.3. Specifically, for argument $\omega = [\omega_1] \in \mathbb{C}^{n_X}$, the exponential transform $\Phi^{\mathbb{M}}$ takes the form

$$\Phi^{\mathbb{M}}(\omega;\tilde{\tau},\infty) = \exp(A_\Phi(\omega_1;\tilde{\tau}_1,\infty) + B_\Phi(\omega_1;\tilde{\tau}_1,\infty))$$

under the conditions of proposition 2.1, yielding a special case of equation (2.7). Essentially by taking partial derivatives with the Faà di Bruno formula (A.1), we further obtain an expression for the pl-linear transform $\Phi^{\mathbb{M},[\alpha]}$ as

$$\Phi^{\mathbb{M},[\alpha]}(\omega;\tilde{\tau},\infty) = \Phi^{\mathbb{M}}(\omega;\tilde{\tau},\infty) \sum_{\tilde{\mathcal{Q}}(\alpha)} M_{k,\ell}^{\alpha}\,(A_\Phi^{(\ell)}(\omega_1;\tilde{\tau}_1,\infty) + B_\Phi^{(\ell)}(\omega_1;\tilde{\tau}_1,\infty))^k\,,$$

which is justified by the conditions of proposition 2.2 and yields a special case of equation (2.9).

With these settings, we can now investigate the concrete implementation of the polynomial moment in equation (4.11) for the case of equity and volatility derivatives as in sections 3.3 and 3.4, respectively. To simplify notation, we introduce the auxiliary transform $\Upsilon(\mathfrak{b}([y_1])) = \Upsilon_i(\mathfrak{b}(y_1))$ given by

$$\Upsilon(\mathfrak{b}(y)) = \exp(A_\Pi^{e_i}(\mathfrak{b}(y);\tilde{T}))\, \Phi^{\mathbb{M},[\alpha]}([0;B_\Pi^{e_i}(\mathfrak{b}(y);\tilde{T})];\tau,\infty)\,, \tag{4.12}$$

so that $\tilde{\Phi}^{\mathbb{M},[\alpha,e_i]}(0,0;\tilde{\tau},\infty) = \langle w^{e_i}(y;K), \Upsilon(\mathfrak{b}(y))\rangle$. The following examples 4.1 and 4.2 exploit the integral representations derived in lemmas 3.1 and 3.3 to determine the respective single-period, first-order moments. Each such moment can be determined by performing one-dimensional numerical integration. Exploiting symmetry properties of the integrands, the computational complexity of each one-dimensional integration problem can be reduced by replacing the regularization Δ with the computationally more efficient regularization $\tilde{\Delta}$.

Example 4.1 (Stock moments). Suppose $V_{i,t}$ corresponds to an equity option. We can then take $w_i = w_{\text{stock}}^{O_i}$ and $F_i = F_{\text{stock}}^{O_i}$. With the integral representation in lemma 3.1, we obtain the polynomial moment in equation (4.11) as

$$\langle w^{e_i}(y;K), \Upsilon(\mathfrak{b}(y))\rangle = \Upsilon_{i,0} + \int_{\mathbb{R}_+} \frac{\Delta_{\tilde{y}_1}^{(1)}(F_i(\tilde{y}_1;K_i)\,\Upsilon_{i,1}(\tilde{y}_1))}{\mathrm{i}\tilde{y}_1}\,\mathrm{d}\tilde{y}_1\,,$$

where

$$\begin{aligned}\Upsilon_{i,0} &= \tfrac{1}{2}c^{O_i}\big(\Upsilon_i([1;0]) - \exp(K_i)\,\Upsilon_i([0;0])\big)\\ \Upsilon_{i,1}(\tilde{y}_1) &= \Upsilon_i([1;0] + \mathrm{i}\tilde{y}_1[1;0]) - \exp(K_i)\,\Upsilon_i(\mathrm{i}\tilde{y}_1[1;0])\end{aligned}$$

in terms of the auxiliary transform in equation (4.12). ■

Example 4.2 (VIX moments). Suppose $V_{i,t}$ corresponds to a volatility option. We can then take $w_i = w_{\text{vix}}^{O_i}$ and $F_i = F_{\text{vix}}^{O_i}$. With the integral representation in lemma 3.3, we obtain the polynomial moment in equation (4.11) as

$$\langle w^{e_i}(y), \Upsilon(\mathfrak{b}(y))\rangle = \Upsilon_{i,0} + \int_{\mathbb{R}_+} \frac{\Delta_{\tilde{y}_1}^{(3/2)}(F_i(\tilde{y}_1;K_i)\,\Upsilon_{i,1}(\tilde{y}_1))}{(\mathrm{i}\tilde{y}_1)^{3/2}}\,\mathrm{d}\tilde{y}_1\,,$$

where

$$\begin{aligned}\Upsilon_{i,0} &= \tfrac{1}{2}(1 - c^{O_i})\,K_i^{1/2}\,\Upsilon_i([0;0])\\ \Upsilon_{i,1}(\tilde{y}_1) &= \Upsilon_i(\mathrm{i}\tilde{y}_1[0;b_{\text{vix}}])\end{aligned}$$

in terms of the auxiliary transform in equation (4.12). ■

4.5.2 Single-period, second-order moments

As a second group of examples, we essentially maintain the previous setting, but now consider polynomial moments involving the product of contemporaneous derivatives prices $V_{i,t+\tilde{\tau}_1}$ and $V_{j,t+\tilde{\tau}_1}$ for $\tilde{\tau}_1 \geq 0$ and $1 \leq i,j \leq n_V$. For this, we still consider the time vector $\tilde{\tau} = [\tilde{\tau}_1] \in \mathbb{R}_+$, but now a second-order price moment with $\beta = e_{ij} = e_i + e_j \in \mathbb{N}^{n_V}$. With $\alpha = [\alpha_1] \in \mathbb{N}^{n_X}$, equation (4.4) then yields the polynomial moment

$$\begin{aligned}\tilde{\Phi}^{\mathbb{M},[\alpha,e_{ij}]}(0,0;\tilde{\tau},\infty) &= \mathrm{E}^{\mathbb{M}}[(X_{t\oplus[\tilde{\tau}_1]})^{\alpha_1}\,V_{i,t+\tilde{\tau}_1}\,V_{j,t+\tilde{\tau}_1}]\\ &= \langle w^{e_{ij}}(y;K), \exp(A_{\Pi}^{e_{ij}}(\mathfrak{b}(y);\tilde{T}))\,\Phi^{\mathbb{M},[\alpha]}([0;B_{\Pi}^{e_{ij}}(\mathfrak{b}(y);\tilde{T})];\tilde{\tau},\infty)\rangle\,,\end{aligned} \tag{4.13}$$

where from lemma 4.1

$$\begin{aligned}w^{e_{ij}}(y;K) &= w_i(y_1;K_i) \otimes w_j(y_2;K_j)\\ A_{\Pi}^{e_{ij}}(\mathfrak{b}(y);\tilde{T}) &= A_{\Pi}(\mathfrak{b}_i(y_1);\tilde{T}_i) + A_{\Pi}(\mathfrak{b}_j(y_2);\tilde{T}_j)\\ B_{\Pi}^{e_{ij}}(\mathfrak{b}(y);\tilde{T}) &= B_{\Pi}(\mathfrak{b}_i(y_1);\tilde{T}_i) + B_{\Pi}(\mathfrak{b}_j(y_2);\tilde{T}_j)\,.\end{aligned}$$

Unlike for first-order moments, $w^{e_{ij}}$ now is a tensor product of the tempered distributions w_i and w_j associated to the derivatives prices $V_{i,t}$ and $V_{j,t}$, respectively. Similarly, $A_{\Pi}^{e_{ij}}$ and $B_{\Pi}^{e_{ij}}$ can be interpreted as tensor sums of the associated coefficients of the pricing transform.

To determine a tractable expression for the standard transform in equation (4.13), we again use the general results in section 2.3. For argument $\omega = [\omega_1] \in \mathbb{C}^{n_X}$, the exponential transform $\Phi^{\mathbb{M}}$ remains

$$\Phi^{\mathbb{M}}(\omega; \tilde{\tau}, \infty) = \exp(A_\Phi(\omega_1; \tilde{\tau}_1, \infty) + B_\Phi(\omega_1; \tilde{\tau}_1, \infty)) \; ,$$

valid under the conditions of proposition 2.1 as a special case of equation (2.7). Likewise, the expression for the associated pl-linear transform $\Phi^{\mathbb{M},[\alpha]}$ is still given by

$$\Phi^{\mathbb{M},[\alpha]}(\omega; \tilde{\tau}, \infty) = \Phi^{\mathbb{M}}(\omega; \tilde{\tau}, \infty) \sum_{\tilde{\mathcal{Q}}(\alpha)} M^{\alpha}_{k,\ell} \, (A^{(\ell)}_\Phi(\omega_1; \tilde{\tau}_1, \infty) + B^{(\ell)}_\Phi(\omega_1; \tilde{\tau}_1, \infty))^k$$

under the conditions of proposition 2.2, yielding a special case of equation (2.9).

To implement the polynomial moment in equation (4.13) for the case of equity and volatility derivatives as in sections 3.3 and 3.4, respectively, we again introduce some additional notation. Specifically, we now define the auxiliary transform $\Upsilon(\mathfrak{b}([y_1; y_2])) = \Upsilon_{ij}(\mathfrak{b}(y_1), \mathfrak{b}(y_2))$ by

$$\Upsilon(\mathfrak{b}(y)) = \exp(A^{e_{ij}}_\Pi(\mathfrak{b}(y); \tilde{T})) \, \Phi^{\mathbb{M},[\alpha]}([0; B^{e_{ij}}_\Pi(\mathfrak{b}(y); \tilde{T})]; \tau, \infty) \; . \tag{4.14}$$

Consequently, we may write $\tilde{\Phi}^{\mathbb{M},[\alpha,e_{ij}]}(0, 0; \tilde{\tau}, \infty) = \langle w^{e_{ij}}(y; K), \Upsilon(\mathfrak{b}(y))\rangle$. Exploiting the integral representations in lemmas 3.1 and 3.3, the following examples 4.3 to 4.5 determine the respective single-period, second-order moments. Determining each of these moments requires performing up to two-dimensional numerical integration. Exploiting symmetries of the integrands, the computational complexity of the integration problems can be reduced by replacing the regularization Δ with the computationally more efficient regularization $\tilde{\Delta}$ along one dimension.

Example 4.3 (Stock-stock moments). Suppose both $V_{i,t}$ and $V_{j,t}$ correspond to equity options. We can then take $w_i = w^{O_i}_{\text{stock}}$ and $w_j = w^{O_j}_{\text{stock}}$ as well as $F_i = F^{O_i}_{\text{stock}}$ and $F_j = F^{O_j}_{\text{stock}}$. With the integral representation in lemma 3.1, we obtain the polynomial moment in equation (4.13) as

$$\begin{aligned}\langle w^{e_{ij}}(y; K), \Upsilon(\mathfrak{b}(y))\rangle = \Upsilon_{ij,0} &+ \int_{\mathbb{R}_+} \frac{\Delta^{(1)}_{\tilde{y}_1}(F_i(\tilde{y}_1; K_i) \, \Upsilon_{ij,1}(\tilde{y}_1))}{\mathrm{i}\tilde{y}_1} \, \mathrm{d}\tilde{y}_1 + \int_{\mathbb{R}_+} \frac{\Delta^{(1)}_{\tilde{y}_2}(F_j(\tilde{y}_2; K_j) \, \Upsilon_{ij,2}(\tilde{y}_2))}{\mathrm{i}\tilde{y}_2} \, \mathrm{d}\tilde{y}_2 \\ &+ \int_{\mathbb{R}_+ \times \mathbb{R}_+} \frac{\Delta^{(1)}_{\tilde{y}_1} \Delta^{(1)}_{\tilde{y}_2}(F_i(\tilde{y}_1; K_i) \, F_j(\tilde{y}_2; K_j) \, \Upsilon_{ij,12}(\tilde{y}_1, \tilde{y}_2))}{\mathrm{i}\tilde{y}_1 \, \mathrm{i}\tilde{y}_2} \, \mathrm{d}[\tilde{y}_1; \tilde{y}_2] \; ,\end{aligned}$$

where

$$\begin{aligned}\Upsilon_{ij,0} &= \tfrac{1}{4} c^{O_i} c^{O_j} \big(\Upsilon_{ij}([1; 0], [1; 0]) + \exp(K_i + K_j) \, \Upsilon_{ij}([0; 0], [0; 0]) \\ &\qquad - \exp(K_i) \, \Upsilon_{ij}([0; 0], [1; 0]) - \exp(K_j) \, \Upsilon_{ij}([1; 0], [0; 0])\big) \\ \Upsilon_{ij,1}(\tilde{y}_1) &= \tfrac{1}{2} c^{O_j} \big(\Upsilon_{ij}([1; 0] + \mathrm{i}\tilde{y}_1[1; 0], [1; 0]) + \exp(K_i + K_j) \, \Upsilon_{ij}(\mathrm{i}\tilde{y}_1[1; 0], [0; 0]) \\ &\qquad - \exp(K_i) \, \Upsilon_{ij}(\mathrm{i}\tilde{y}_1[1; 0], [1; 0]) - \exp(K_j) \, \Upsilon_{ij}([1; 0] + \mathrm{i}\tilde{y}_1[1; 0], [0; 0])\big) \\ \Upsilon_{ij,2}(\tilde{y}_2) &= \tfrac{1}{2} c^{O_i} \big(\Upsilon_{ij}([1; 0], [1; 0] + \mathrm{i}\tilde{y}_2[1; 0]) + \exp(K_i + K_j) \, \Upsilon_{ij}([0; 0], \mathrm{i}\tilde{y}_2[1; 0]) \\ &\qquad - \exp(K_i) \, \Upsilon_{ij}([0; 0], [1; 0] + \mathrm{i}\tilde{y}_2[1; 0]) - \exp(K_j) \, \Upsilon_{ij}([1; 0], \mathrm{i}\tilde{y}_2[1; 0])\big) \\ \Upsilon_{ij,12}(\tilde{y}_1, \tilde{y}_2) &= \Upsilon_{ij}([1; 0] + \mathrm{i}\tilde{y}_1[1; 0], [1; 0] + \mathrm{i}\tilde{y}_2[1; 0]) + \exp(K_i + K_j) \, \Upsilon_{ij}(\mathrm{i}\tilde{y}_1[1; 0], \mathrm{i}\tilde{y}_2[1; 0]) \\ &\qquad - \exp(K_i) \, \Upsilon_{ij}(\mathrm{i}\tilde{y}_1[1; 0], [1; 0] + \mathrm{i}\tilde{y}_2[1; 0]) - \exp(K_j) \, \Upsilon_{ij}([1; 0] + \mathrm{i}\tilde{y}_1[1; 0], \mathrm{i}\tilde{y}_2[1; 0])\end{aligned}$$

in terms of the auxiliary transform in equation (4.14). ■

Example 4.4 (VIX-VIX moments). Suppose both $V_{i,t}$ and $V_{j,t}$ correspond to volatility options.

We can then take $w_i = w_{\text{vix}}^{O_i}$ and $w_j = w_{\text{vix}}^{O_j}$ as well as $F_i = F_{\text{vix}}^{O_i}$ and $F_j = F_{\text{vix}}^{O_j}$. With the integral representation in lemma 3.3, we obtain the polynomial moment in equation (4.13) as

$$\begin{aligned}\langle w^{e_{ij}}(y;K), \Upsilon(\mathfrak{b}(y))\rangle = \Upsilon_{ij,0} &+ \int_{\mathbb{R}_+} \frac{\Delta^{(3/2)}_{\tilde{y}_1}(F_i(\tilde{y}_1;K_i)\,\Upsilon_{ij,1}(\tilde{y}_1))}{(\mathrm{i}\tilde{y}_1)^{3/2}}\,\mathrm{d}\tilde{y}_1 + \int_{\mathbb{R}_+} \frac{\Delta^{(3/2)}_{\tilde{y}_2}(F_j(\tilde{y}_2;K_j)\,\Upsilon_{ij,2}(\tilde{y}_2))}{(\mathrm{i}\tilde{y}_2)^{3/2}}\,\mathrm{d}\tilde{y}_2 \\ &+ \int_{\mathbb{R}_+\times\mathbb{R}_+} \frac{\Delta^{(3/2)}_{\tilde{y}_1}\Delta^{(3/2)}_{\tilde{y}_2}(F_i(\tilde{y}_1;K_i)\,F_j(\tilde{y}_2;K_j)\,\Upsilon_{ij,12}(\tilde{y}_1,\tilde{y}_2))}{(\mathrm{i}\tilde{y}_1)^{3/2}(\mathrm{i}\tilde{y}_2)^{3/2}}\,\mathrm{d}[\tilde{y}_1;\tilde{y}_2]\,,\end{aligned}$$

where

$$\begin{aligned}\Upsilon_{ij,0} &= \tfrac{1}{4}(1-c^{O_i})(1-c^{O_j})\,K_i^{1/2}K_j^{1/2}\,\Upsilon_{ij}([0;0],[0;0]) \\ \Upsilon_{ij,1}(\tilde{y}_1) &= \tfrac{1}{2}(1-c^{O_j})\,K_j^{1/2}\,\Upsilon_{ij}(\mathrm{i}\tilde{y}_1[0;b_{\text{vix}}],[0;0]) \\ \Upsilon_{ij,2}(\tilde{y}_2) &= \tfrac{1}{2}(1-c^{O_i})\,K_i^{1/2}\,\Upsilon_{ij}([0;0],\mathrm{i}\tilde{y}_2[0;b_{\text{vix}}]) \\ \Upsilon_{ij,12}(\tilde{y}_1,\tilde{y}_2) &= \Upsilon_{ij}(\mathrm{i}\tilde{y}_1[0;b_{\text{vix}}],\mathrm{i}\tilde{y}_2[0;b_{\text{vix}}])\end{aligned}$$

in terms of the auxiliary transform in equation (4.14). ∎

Example 4.5 (Stock-VIX moments). Suppose $V_{i,t}$ corresponds to an equity option, while $V_{j,t}$ corresponds to a volatility option. We can then take $w_i = w_{\text{stock}}^{O_i}$ and $w_j = w_{\text{vix}}^{O_j}$ as well as $F_i = F_{\text{stock}}^{O_i}$ and $F_j = F_{\text{vix}}^{O_j}$. With the integral representations in lemmas 3.1 and 3.3, we obtain the polynomial moment in equation (4.13) as

$$\begin{aligned}\langle w^{e_{ij}}(y;K), \Upsilon(\mathfrak{b}(y))\rangle = \Upsilon_{ij,0} &+ \int_{\mathbb{R}_+} \frac{\Delta^{(1)}_{\tilde{y}_1}(F_i(\tilde{y}_1;K_i)\,\Upsilon_{ij,1}(\tilde{y}_1))}{\mathrm{i}\tilde{y}_1}\,\mathrm{d}\tilde{y}_1 + \int_{\mathbb{R}_+} \frac{\Delta^{(3/2)}_{\tilde{y}_2}(F_j(\tilde{y}_2;K_j)\,\Upsilon_{ij,2}(\tilde{y}_2))}{(\mathrm{i}\tilde{y}_2)^{3/2}}\,\mathrm{d}\tilde{y}_2 \\ &+ \int_{\mathbb{R}_+\times\mathbb{R}_+} \frac{\Delta^{(1)}_{\tilde{y}_1}\Delta^{(3/2)}_{\tilde{y}_2}(F_i(\tilde{y}_1;K_i)\,F_j(\tilde{y}_2;K_j)\,\Upsilon_{ij,12}(\tilde{y}_1,\tilde{y}_2))}{\mathrm{i}\tilde{y}_1\,(\mathrm{i}\tilde{y}_2)^{3/2}}\,\mathrm{d}[\tilde{y}_1;\tilde{y}_2]\,,\end{aligned}$$

where

$$\begin{aligned}\Upsilon_{ij,0} &= \tfrac{1}{4}c^{O_i}(1-c^{O_j})\,K_j^{1/2}\left(\Upsilon_{ij}([1;0],[0;0]) - \exp(K_i)\,\Upsilon_{ij}([0;0],[0;0])\right) \\ \Upsilon_{ij,1}(\tilde{y}_1) &= \tfrac{1}{2}(1-c^{O_j})\,K_j^{1/2}\left(\Upsilon_{ij}([1;0]+\mathrm{i}\tilde{y}_1[1;0],[0;0]) - \exp(K_i)\,\Upsilon_{ij}(\mathrm{i}\tilde{y}_1[1;0],[0;0])\right) \\ \Upsilon_{ij,2}(\tilde{y}_2) &= \tfrac{1}{2}c^{O_i}\left(\Upsilon_{ij}([1;0],\mathrm{i}\tilde{y}_2[0;b_{\text{vix}}]) - \exp(K_i)\,\Upsilon_{ij}([0;0],\mathrm{i}\tilde{y}_2[0;b_{\text{vix}}])\right) \\ \Upsilon_{ij,12}(\tilde{y}_1,\tilde{y}_2) &= \Upsilon_{ij}([1;0]+\mathrm{i}\tilde{y}_1[1;0],\mathrm{i}\tilde{y}_2[0;b_{\text{vix}}]) - \exp(K_i)\,\Upsilon_{ij}(\mathrm{i}\tilde{y}_1[1;0],\mathrm{i}\tilde{y}_2[0;b_{\text{vix}}])\end{aligned}$$

in terms of the auxiliary transform in equation (4.14). ∎

4.5.3 Multi-period, second-order moments

As a third and final group of examples, we consider polynomial moments involving the product of asynchronous derivatives prices $V_{i,t+\tilde{\tau}_1}$ and $V_{j,t+\tilde{\tau}_2}$ for $\tilde{\tau}_2 > \tilde{\tau}_1 \geq 0$ and $1 \leq i,j \leq n_V$. Unlike in the preceding examples, we now have two distinct time points $\tilde{\tau}_1$ and $\tilde{\tau}_2$, collected in a non-decreasing time vector $\tilde{\tau} = [\tilde{\tau}_1;\tilde{\tau}_2] \in \mathbb{R}_+^2$. Further setting $\beta = e_{ij'} = e_i + e_{j'} \in \mathbb{N}^{2m}$ with $j' = n_V + j$ as well as $\alpha = [\alpha_1;\alpha_2] \in \mathbb{N}^{2n_X}$, equation (4.4) yields

$$\begin{aligned}\tilde{\Phi}^{\mathbb{M},[\alpha,e_{ij'}]}(0,0;\tilde{\tau},\infty) &= \mathrm{E}^{\mathbb{M}}[(X_{t\oplus[\tilde{\tau}_1]})^{\alpha_1}\,(X_{t+\tilde{\tau}_1\oplus[\tilde{\tau}_2-\tilde{\tau}_1]})^{\alpha_2}\,V_{i,t+\tilde{\tau}_1}\,V_{j,t+\tilde{\tau}_2}] \\ &= \langle w^{e_{ij'}}(y;K), \exp(A_\Pi^{e_{ij'}}(\mathfrak{b}(y);\tilde{T}))\,\Phi^{\mathbb{M},[\alpha]}([0;B_\Pi^{e_{ij'}}(\mathfrak{b}(y);\tilde{T})];\tilde{\tau},\infty)\rangle\,,\end{aligned} \tag{4.15}$$

where from lemma 4.1, we get

$$\begin{aligned} w^{e_{ij}'}(y;K) &= w_i(y_1;K_i) \otimes w_j(y_2;K_j) \\ A_\Pi^{e_{ij}'}(\mathfrak{b}(y);\tilde{T}) &= A_\Pi(\mathfrak{b}_i(y_1);\tilde{T}_i) + A_\Pi(\mathfrak{b}_j(y_2);\tilde{T}_j) \\ B_\Pi^{e_{ij}'}(\mathfrak{b}(y);\tilde{T}) &= [B_\Pi(\mathfrak{b}_i(y_1);\tilde{T}_i); B_\Pi(\mathfrak{b}_j(y_2);\tilde{T}_j)] \,. \end{aligned}$$

Analogous to single-period, second-order moments, $w^{e_{ij}'} = w^{e_{ij}}$ is a tensor product of the tempered distributions w_i and w_j, while $A_\Pi^{e_{ij}'}$ is a tensor sum of the respective coefficients of the pricing transform. However, $B_\Pi^{e_{ij}'}$ is not a tensor sum, but rather a block vector in $\mathbb{C}^{2n_Z}$.

Determining a tractable expression for the standard transform in equation (4.15) entails some additional complexity due to the multi-period structure. Nevertheless, the general results in section 2.3 conveniently yield such an expression in terms of the single-period coefficients. Specifically, for argument $\omega = [\omega_1;\omega_2] \in \mathbb{C}^{2n_X}$, the exponential transform $\Phi^{\mathbb{M}}$ can be given as

$$\Phi^{\mathbb{M}}(\omega;\tilde{\tau},\infty) = \exp(A_\Phi(\omega_2;\tilde{\tau}_2-\tilde{\tau}_1,0) + A_\Phi(\tilde{\omega}_1;\tilde{\tau}_1,\infty) + B_\Phi(\tilde{\omega}_1;\tilde{\tau}_1,\infty))$$

with $\tilde{\omega}_1 = \omega_1 + [0; B_\Phi(\omega_2;\tilde{\tau}_2-\tilde{\tau}_1,0)]$, justified by the conditions of proposition 2.1 as a special case of equation (2.7). Essentially by taking partial derivatives with repeated help of the Faà di Bruno formula (A.1), the associated pl-linear transform $\Phi^{\mathbb{M},[\alpha]}$ can be written as the tensor expression

$$\begin{aligned} \Phi^{\mathbb{M},[\alpha]}(\omega;\tilde{\tau},\infty) = \Phi^{\mathbb{M}}(\omega;\tilde{\tau},\infty) \sum_{\tilde{\mathcal{Q}}([\alpha_1;\alpha_2])} M_{k,[\ell_1;\ell_2]}^{[\alpha_1;\alpha_2]} \Big(& 0^{\ell_1} A_\Phi^{(\ell_2)}(\omega_2;\tilde{\tau}_2-\tilde{\tau}_1,0) \\ & + \sum_{|\gamma_2|\le|\ell_2|} A_\Phi^{(\ell_1+[0;\gamma_2])}(\tilde{\omega}_1;\tilde{\tau}_1,\infty) \sum_{\mathcal{Q}(\ell_2,\gamma_2)} M_{k',\ell'}^{\ell_2} \, (B_\Phi^{(\ell')}(\omega_2;\tilde{\tau}_2-\tilde{\tau}_1,0))^{k'} \\ & + \sum_{|\gamma_2|\le|\ell_2|} B_\Phi^{(\ell_1+[0;\gamma_2])}(\tilde{\omega}_1;\tilde{\tau}_1,\infty) \sum_{\mathcal{Q}(\ell_2,\gamma_2)} M_{k',\ell'}^{\ell_2} \, (B_\Phi^{(\ell')}(\omega_2;\tilde{\tau}_2-\tilde{\tau}_1,0))^{k'} \Big)^k , \end{aligned}$$

which holds under the conditions of proposition 2.2 as a special case of equation (2.9).

With these expressions, we may implement the polynomial moment in equation (4.15) for the case of equity and volatility derivatives as in sections 3.3 and 3.4, respectively. As before, we simplify notation by introducing the auxiliary transform $\Upsilon(\mathfrak{b}([y_1;y_2])) = \Upsilon_{ij}(\mathfrak{b}(y_1),\mathfrak{b}(y_2))$ given by

$$\Upsilon(\mathfrak{b}(y)) = \exp(A_\Pi^{e_{ij}'}(\mathfrak{b}(y);\tilde{T}))\, \Phi^{\mathbb{M},[\alpha]}([0;B_\Pi^{e_{ij}'}(\mathfrak{b}(y);\tilde{T})];\tau,\infty) \,. \tag{4.16}$$

This allows to write $\tilde{\Phi}^{\mathbb{M},[\alpha,e_{ij}']}(0,0;\tilde{\tau},\infty) = \langle w^{e_{ij}'}(y;K), \Upsilon(\mathfrak{b}(y))\rangle$. It should be noted that in the multi-period setting, only the auxiliary transform differs from the respective single-period auxiliary transform. Not surprisingly therefore, we may determine multi-period, second-order moments from the formulas in examples 4.3 to 4.5 when using the auxiliary transform in equation (4.16) instead.

4.6 Including measurement errors

Thus far, this section has considered moments involving derivatives prices under the implicit assumption that these price are observed exactly. In practical applications, however, one usually observes derivatives prices only with measurement errors stemming from various sources (e.g., price discreteness and bid-ask spreads). For this reason, we briefly discuss the generalization of our results in the presence of measurement errors.

In general, measurement errors may exhibit complex interdependencies and dependencies with respect to the state variables. To accommodate such features, we consider an augmented state vector $Z_t = [\tilde{Z}_t;\varepsilon_t]$

and define $Z_{t+\tilde{\tau}} = [\tilde{Z}_{t+\tilde{\tau}}; \varepsilon_{t+\tilde{\tau}}]$ for notational convenience. For generalizing proposition 4.1, suppose that $\tilde{V}_t$ denotes the vector of derivatives prices measured with error ε_t and construct $\tilde{V}_{t+\tilde{\tau}}$ as usual. Analogous to equation (4.4), define the pl-linear moments

$$\tilde{\tilde{\Phi}}^{\mathbb{M},[\alpha,\beta]}(\omega, 0; \tilde{\tau}, \infty) = \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t\oplus\tilde{\tau}})\,(X_{t\oplus\tilde{\tau}})^{\alpha}\,(\tilde{V}_{t+\tilde{\tau}})^{\beta}]\,. \tag{4.17}$$

To give an expression for $\tilde{\tilde{\Phi}}^{\mathbb{M},[\alpha,\beta]}$ in equation (4.17), we consider three exemplary cases. First, for exponential measurement errors with $\tilde{V}_t = V_t\,\exp(\varepsilon_t)$, we have

$$\tilde{\tilde{\Phi}}^{\mathbb{M},[\alpha,\beta]}(\omega, 0; \tilde{\tau}, \infty) = \tilde{\Phi}^{\mathbb{M},[\alpha,\beta]}(\omega + [0; 0; \beta], 0; \tilde{\tau}, \infty)\,. \tag{4.18}$$

Second, for multiplicative measurement errors with $\tilde{V}_t = V_t\,\varepsilon_t$, we obtain

$$\tilde{\tilde{\Phi}}^{\mathbb{M},[\alpha,\beta]}(\omega, 0; \tilde{\tau}, \infty) = \tilde{\Phi}^{\mathbb{M},[\alpha+[0;0;\beta],\beta]}(\omega, 0; \tilde{\tau}, \infty)\,. \tag{4.19}$$

Third, for additive measurement errors with $\tilde{V}_t = V_t + \varepsilon_t$, the multi-binomial theorem yields

$$\tilde{\tilde{\Phi}}^{\mathbb{M},[\alpha,\beta]}(\omega, 0; \tilde{\tau}, \infty) = \sum_{\eta \le \beta} \binom{\beta}{\eta} \tilde{\Phi}^{\mathbb{M},[\alpha+[0;0;\eta],\beta-\eta]}(\omega, 0; \tilde{\tau}, \infty)\,. \tag{4.20}$$

Equations (4.18) to (4.20) give pl-linear moments involving $\tilde{V}_t$ in terms of the pl-linear moments involving V_t determined by proposition 4.1 using an augmented state space definition. Only sufficiently strong independence assumptions allow to treat measurements error moments separately. Specifically, making the assumption that $\tilde{Z}_{t+\tilde{\tau}}$ and $\varepsilon_{t+\tilde{\tau}}$ are independent, we obtain

$$\tilde{\Phi}^{\mathbb{M},[\alpha,\beta]}(\omega, 0; \tilde{\tau}, \infty) = \tilde{\Phi}^{\mathbb{M},[[\alpha_S;\alpha_Z;0],\beta]}([\omega_S; \omega_Z; 0], 0; \tilde{\tau}, \infty)\,\tilde{\Phi}^{\mathbb{M},[[0;0;\alpha_\varepsilon],0]}([0; 0; \omega_\varepsilon], 0; \tilde{\tau}, \infty)\,,$$

where the measurement errors affect only the second term.

5 Estimation methodology

In this section, we devise a GMM estimation procedure that incorporates the moments involving derivatives prices developed in section 4. After laying out the basic setup in section 5.1, we suggest two GMM estimators: an exact one in section 5.2 and an approximate one in section 5.3, using the exact and approximate moments derived in sections 4.3 and 4.4, respectively. Subsequently, section 5.4 discusses the proposed methodology and some extensions.

5.1 Basic setup

For estimating the affine stochastic volatility model (2.1), the data set comprises the stock price S_t and a panel of option prices, collected in the vector V_t. In order to obtain such a panel in practice, it is usually necessary to interpolate observed market prices. We may also easily accommodate additional observables that are polynomial in the state vector, such as the squared VIX and related static portfolios constructed from option prices. Maintaining the setup of section 4, each element of V_t is given by proposition 3.1 as in equation (4.1). Measurement errors can straightforwardly be accounted for along the lines of section 4.6. The interval between adjacent observation dates equals Δ. Define a parameter space Θ such that a parameter vector $\vartheta \in \Theta$ contains all relevant model parameters. For any $\vartheta \in \Theta$, we denote by $\mathbb{P}(\vartheta)$ and $\mathbb{Q}(\vartheta)$ the associated parameterized real-world and risk-neutral probability measures, respectively. Suppose that all data is generated by the model (2.1) under the parameter vector $\vartheta_0 \in \Theta$.

5.2 Exact GMM estimation

Exact GMM estimation relies on the exact moments involving asset prices derived in section 4.3. Specifically, we set a vector of moment conditions $f_t(\vartheta)$ of the form

$$f_t(\vartheta) = P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau}}) - \mathrm{E}^{\mathbb{P}(\vartheta)}[P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau}})] \,, \tag{5.1}$$

where $P(x,v) = \sum_{\alpha,\beta} c_{\alpha,\beta}\, x^\alpha\, v^\beta$ is a vector-valued polynomial and the model-based expected value in equation (5.1) may be written in terms of the exact extended transforms $\tilde{\Phi}^{\mathbb{P}(\vartheta),[\alpha,\beta]}$ defined in section 4.3,

$$\mathrm{E}^{\mathbb{P}(\vartheta)}[P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau}})] = \sum_{\alpha,\beta} c_{\alpha,\beta}\, \tilde{\Phi}^{\mathbb{P}(\vartheta),[\alpha,\beta]}(0,0;\tilde{\tau},\infty) \,. \tag{5.2}$$

Our main result concerning exact moments involving asset prices in proposition 4.1 for the choice $\mathbb{M} = \mathbb{P}(\vartheta)$ yields a tractable expression for each exact extended transform, given by equation (4.4). By construction, the exact moment conditions $f_t(\vartheta)$ determined by equations (5.1) and (5.2) thus satisfy $\mathrm{E}^{\mathbb{P}(\vartheta_0)}[f_t(\vartheta_0)] = 0$.

Defining the sample average $\hat{g}_T(\vartheta) = \frac{1}{T}\sum_{t=1}^T f_{t\Delta}(\vartheta)$, the GMM estimator associated to the exact moment conditions in equation (5.2) can be written as

$$\hat{\vartheta}_T(W) = \operatorname*{argmin}_{\vartheta\in\Theta} \hat{g}_T(\vartheta)^\top W\, \hat{g}_T(\vartheta) \tag{5.3}$$

for some weighting matrix W. Within the present setting, the optimal weighting matrix can be computed from the data alone. Hence, the efficient GMM estimator can be realized in a single step. Specifically, we have that

$$\Omega = \lim_{T\to\infty} T\, \mathrm{E}^{\mathbb{P}(\vartheta_0)}[\hat{g}_T(\vartheta_0)\, \hat{g}_T(\vartheta_0)^\top] = \Gamma_0 + \sum_{\ell=1}^{\infty} \Gamma_\ell + \Gamma_\ell^\top \,,$$

assuming that $\Gamma_\ell = \mathrm{E}^{\mathbb{P}(\vartheta_0)}[f_{t\Delta}(\vartheta_0)\, f_{(t+\ell)\Delta}(\vartheta_0)^\top]$ are absolutely summable, with

$$\Gamma_\ell = \mathrm{E}^{\mathbb{P}(\vartheta_0)}[P(X_{t\Delta\oplus\tilde{\tau}}, V_{t\Delta+\tilde{\tau}})\, P(X_{(t+\ell)\Delta\oplus\tilde{\tau}}, V_{(t+\ell)\Delta+\tilde{\tau}})^\top] \,.$$

The efficient GMM estimator results from equation (5.3) when constructing an estimator $\hat{\Omega}_T$ for Ω from the data (using, e.g., the Newey and West (1987) procedure) and setting $W \propto \hat{\Omega}_T^{-1}$.

Under standard regularity conditions, as stated in Dillschneider (2020) within a more general setting, the exact GMM estimator in equation (5.3) is consistent and asymptotically normal. In practical applications, however, we may find these regularity conditions being violated as the polynomial moments in equation (5.2) cannot be determined exactly. Like other pl-linear moments, these generally require numerical solutions to ODEs as well as the evaluation of multi-dimensional numerical integrals. As discussed before, the latter moreover renders the moment conditions $f_t(\vartheta)$ in equation (5.1) computationally infeasible except for low orders of β.

5.3 Approximate GMM estimation

Approximate GMM estimation uses the approximate moments involving asset prices derived in section 4.4 for some prespecified approximation order p. Analogous to equation (5.1), we specify a vector of moment conditions $f_{t,(p)}(\vartheta)$ of the form

$$f_{t,(p)}(\vartheta) = P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau}}) - \mathrm{E}^{\mathbb{P}(\vartheta)}[P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau},(p)})] \,, \tag{5.4}$$

where $P(x,v) = \sum_{\alpha,\beta} c_{\alpha,\beta}\, x^\alpha\, v^\beta$ is a vector-valued polynomial and $V_{t+\tilde{\tau},(p)}$ is constructed as in section 4.4. Hence, the model-based expected value in equation (5.4) may be written in terms of the approximate extended transforms $\tilde{\Phi}^{\mathbb{P}(\vartheta),[\alpha,\beta]}_{(p)}$ defined in section 4.4,

$$\mathrm{E}^{\mathbb{P}(\vartheta)}[P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau},(p)})] = \sum_{\alpha,\beta} c_{\alpha,\beta}\, \tilde{\Phi}^{\mathbb{P}(\vartheta),[\alpha,\beta]}_{(p)}(0,0;\tilde{\tau},\infty)\,. \tag{5.5}$$

The construction of each approximate extended transform follows the procedure laid out in section 4.4 using expansion under the measure $\mathbb{M} = \mathbb{P}(\vartheta)$. Ultimately, proposition 4.2 yields a tractable expression for each approximate extended transform, given by equation (4.9). By construction, the approximate moment conditions $f_{t,(p)}(\vartheta)$ determined by equations (5.1) and (5.2) satisfy $\mathrm{E}^{\mathbb{P}(\vartheta_0)}[f_{t,(p)}(\vartheta_0)] \to 0$ as $p \to \infty$ under the regularity conditions in proposition 4.2.

The GMM estimator associated to the approximate moment conditions in equation (5.5) can be constructed analogously to the exact GMM estimator in equation (5.3). Setting the sample average $\hat{g}_{T,(p)}(\vartheta) = \frac{1}{T}\sum_{t=1}^{T} f_{t\Delta,(p)}(\vartheta)$ and a weighting matrix W, the approximate GMM estimator is given by

$$\hat{\vartheta}_{T,(p)}(W) = \operatorname*{argmin}_{\vartheta\in\Theta}\, \hat{g}_{T,(p)}(\vartheta)^\top W\, \hat{g}_{T,(p)}(\vartheta)\,. \tag{5.6}$$

The choice of an optimal weighting matrix can follow the considerations discussed for the exact estimator.

In contrast to the exact case, the asymptotic properties of the approximate GMM estimator in equation (5.6) essentially depend on the supposed behavior of the approximation order $p(T)$ as $T \to \infty$. Dillschneider (2020) formalizes regularity conditions representing an ideal situation, in which $p(T)$ grows fast enough to preserve the asymptotic properties of the associated exact estimator. Under weaker assumptions in the framework of Armstrong and Kolesár (2019), the approximate estimator may turn out to exhibit local bias, while remaining consistent but less efficient. Especially when taking the approximation order to be fixed at $p(T) = p$, the approximate estimator is globally biased and inconsistent, a case formally treated in Hall and Inoue (2003). Additional higher-order properties of approximate GMM estimators are discussed in Kristensen and Salanié (2017).

In practical applications, additional approximation errors are incurred since the polynomial moments in equation (5.5) rely on numerical solutions to ODEs as well as numerical integration. Yet, contrary to their exact analogues, the approximate polynomial moments only require the evaluation of one-dimensional integrals. Apart from that, they depend merely on polynomial moments of the state vector, which are even available in closed form (cf. proposition 2.3). This renders approximate moment conditions and, hence, the approximate GMM estimator computationally feasible even for larger orders of β.

5.4 Discussion and further extensions

The estimation approach suggested in this section avoids computationally intensive filtering techniques for latent state variables. Instead, it is based on unconditional moments that integrate out latent state variables, but nevertheless captures state dynamics through multi-period moments of augmented state and asset price vectors. Such an approach is common in the literature in related settings (e.g., Duffie and Singleton (1993)).

The choice of an unconditional GMM estimation setup is motivated by the concrete application pursued in this paper. From a theoretical perspective, however, our methodology in section 4 to develop expressions for moments involving transform-based derivatives prices is not limited to an unconditional perspective. Hence, we may extend the estimation approach to a conditional setting.

For exact GMM estimation, we could instead of the unconditional ones in equation (5.1) consider conditional moment conditions of the form

$$f_t(\vartheta; Z_t) = P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau}}) - \mathrm{E}^{\mathbb{P}(\vartheta)}[P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau}}) \,|\, \mathcal{F}_t] \,. \tag{5.7}$$

Compared to the unconditional setting, we only switch to a conditional model-based expected value in equation (5.7), which analogous to equation (5.2) can be written as

$$\mathrm{E}^{\mathbb{P}(\vartheta)}[P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau}}) \,|\, \mathcal{F}_t] = \sum_{\alpha,\beta} c_{\alpha,\beta}\, \tilde{\Phi}^{\mathbb{P}(\vartheta),[\alpha,\beta]}(0,0;\tilde{\tau},0,Z_t) \,. \tag{5.8}$$

Unlike before, we now require conditional extended transforms $\tilde{\Phi}^{\mathbb{P}(\vartheta),[\alpha,\beta]}$ in equation (5.8). With obvious modifications in proposition 4.1 for the choice $\mathbb{M} = \mathbb{P}(\vartheta)$, we may obtain a tractable expression for each such transform by employing conditional standard transforms in equation (4.4). As usual, we use instruments of the form $h(Z_t)$ for some functions h to construct unconditional moment conditions from equation (5.7), noting that the condition $\mathrm{E}^{\mathbb{P}(\vartheta_0)}[f_t(\vartheta_0; Z_t)\, h(Z_t)] = 0$ is automatically assured.

For approximate GMM estimation, we could instead of equation (5.4) take conditional moment conditions of the form

$$f_{t,(p)}(\vartheta; Z_t) = P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau}}) - \mathrm{E}^{\mathbb{P}(\vartheta)}[P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau},(p)}) \,|\, \mathcal{F}_t] \,. \tag{5.9}$$

To avoid expansion coefficients that are highly nonlinear in Z_t and thus interfere with tractability, we maintain an unconditional construction of $V_{t+\tilde{\tau},(p)}$ as in section 4.4. However, equation (5.9) now depends on its conditional moments through

$$\mathrm{E}^{\mathbb{P}(\vartheta)}[P(X_{t\oplus\tilde{\tau}}, V_{t+\tilde{\tau},(p)}) \,|\, \mathcal{F}_t] = \sum_{\alpha,\beta} c_{\alpha,\beta}\, \tilde{\Phi}^{\mathbb{P}(\vartheta),[\alpha,\beta]}_{(p)}(0,0;\tilde{\tau},0,Z_t) \,, \tag{5.10}$$

unlike the unconditional moments in equation (5.5). The conditional extended transforms $\tilde{\Phi}^{\mathbb{P}(\vartheta),[\alpha,\beta]}_{(p)}$ in equation (5.10) obtain from equation (4.9) when employing conditional standard transforms under $\mathbb{M} = \mathbb{P}(\vartheta)$. Again, we rely on instruments of the form $h(Z_t)$ to construct unconditional moment conditions from equation (5.9). To render this approach meaningful, we should assure that $\mathrm{E}^{\mathbb{P}(\vartheta_0)}[f_{t,(p)}(\vartheta_0; Z_t)\, h(Z_t)] \to 0$ as $p \to \infty$. This requires somewhat stronger regularity conditions than those imposed in proposition 4.2, which essentially only captures the case $h = \mathbb{1}$.

A conditional GMM procedure of the described form presupposes observability of the state vector Z_t. However, it will be regularly the case that Z_t, or at least part of it, is latent. In that case, a conditional GMM procedure is not directly feasible, but only after forming state estimates $\hat{Z}_t$ with exact or approximate moment conditions formed from $f_t(\vartheta; \hat{Z}_t)$ or $f_{t,(p)}(\vartheta; \hat{Z}_t)$, respectively, using instruments $h(\hat{Z}_t)$. Yet, employing state estimates instead of actual states introduces (additional) approximation errors into the estimation procedure that need to be accounted for.

6 Numerical results

In this section, we provide some numerical results to further support our methodology for determining moments involving derivatives prices (section 4) and the resulting estimation procedure (section 5). While we provide analytically tarctable expressions for both exact and approximate moments involving derivatives prices as well as moment-based estimation procedures building thereon, the exact versions generally pose significantly higher challenges in their implementation. These are due to the requirement of multi-dimensional numerical integration in conjunction with a highly oscillatory behavior of integrands, which necessitates sophisticated integration techniques. We leave a further investigation of general exact moments for future research and here concentrate on approximate moments involving derivatives prices, which rely only on one-dimensional numerical integration. To eventually allow for a computationally attractive implementation of approximate moments, in addition to the established convergence result (cf. proposition 4.2), we need to verify that the approximation works already well for rather low approximation orders. Hereby having set the agenda for this section, we proceed as follows. Section 6.1 describes the basic setup used for our analyses. In section 6.2, we analyze the pricing errors resulting from our polynomial approximation approach introduced in section 4.4. Equipped with the intuition from this analysis, we proceed to the investigation of moment errors in section 6.3.

6.1 Basic setup

For our numerical analyses, the focus will be on two important models, each constituting a special case of the general affine specification in section 2.1. Both models are introduced in section 2.2. The first model is the Heston model in equation (2.2). In fact, we consider two variants of this model, one without jumps (SV1) and one with jumps (SV1J). The second model is the stochastic mean reversion model in equation (2.4) with jumps (SV2J). Table 1 specifies the relevant model parameters used in our analyses, which are chosen to be roughly in accordance with the empirical estimates obtained in Aït-Sahalia et al. (2020) and Bardgett et al. (2019).

[Table 1 about here.]

Within all models, we consider both equity derivatives as in section 3.3 and volatility derivatives as in section 3.4. For each type of derivatives, we take into account a number of different option specifications. In the case of equity derivatives, we take out-of-the-money put and call options, as is standard in the literature. We consider four different (constant) maturities at 1, 3, 6, and 12 months. For each maturity, we determine a set of strikes located at the $0.01, 0.05, 0.1, 0.2, \ldots, 0.8, 0.9, 0.95$ percentiles of a Gaussian distribution with mean zero and standard deviation equal to the risk-neutral standard deviation of maturity-congruent stock returns, with initial state at the unconditional mean of the latent state vector. We choose a slightly asymmetric strike set to account for the left-skewed conditional return distribution. In the case of volatility derivatives, we take only call options. We again consider four different (constant) maturities at 1, 3, 6, and 12 months. For each maturity, we determine a set of strikes located at the $0, 0.1, 0.2, \ldots, 0.8, 0.9, 0.95$ percentiles of a Gamma distribution, whose moments are matched to a mean and standard deviation equal to the respective risk-neutral moments of the squared VIX, with initial state at the unconditional mean of the latent state vector. Here, the zero percentile effectively includes the (prepaid) forward contract.

For our analyses, we require both exact and approximate option prices. For exact option prices $V_{i,t}$, we use the transform-based pricing formulas developed in section 3. Specifically, equity options rely on the formulas in corollary 3.1 and lemma 3.1, while volatility options rely on the formulas in corollary 3.2 and lemma 3.3. For approximate option prices $V_{i,t,(p)}$ at some approximation order p, we follow the polynomial approximation procedure described in section 4.4. Specifically we employ the expansion captured by lemma 4.2 with coefficients determined according to lemma 4.3 under the real-world measure $\mathbb{P}$.

In addition, we require certain polynomial moments involving option prices. Specifically, we require exact moments $\mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]$ as well as approximate moments $\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)})^N]$ at some approximation order p. In principle, for both we could follow the procedures laid out in sections 4.3 and 4.4 and captured by propositions 4.1 and 4.2, respectively. Additionally, we require joint moments of the form $\mathrm{E}^{\mathbb{P}}[P(V_{i,t}, V_{i,t,(p)})]$ for some polynomial P, which in principle also could be determined from our methodology. For the purpose of our numerical analyses here, exact moments will serve as a means to evaluate the accuracy of approximate moments. To determine exact moments and conduct out-of-methodology comparisons, we revert to standard methods, whose computational complexity is quite large, but bearable for this particular exercise as all required moments need to be determined only once. For the SV1 model, we can rely on direct integration against the exact unconditional density, which is known to be that of a Gamma distribution. For the SV1J and SV2J models, the unconditional densities are unknown and need to be approximated. We use two common techniques, Monte Carlo simulation and density approximation with the Fourier-cosine expansion method of Fang and Oosterlee (2009). For our application, the latter performs superior to the density approximation via ordinary Fourier inversion along the lines of Shephard (1991a,b), especially in multiple dimensions. To numerically determine the expected value by direct integration, we span a (sparse) grid over the state space and evaluate option prices as well as the (approximate) density at each of the grid points, before summing all values with quadrature weights obtained from a trapezoidal rule. For Monte Carlo simulation, we use a 25-year burn-in period to approximately draw 10,000 sample paths from an unconditional state distribution by an Euler-Maruyama scheme as in Lord et al. (2010) with 10 intra-day steps, before evaluating option prices at each realized state vector and performing Monte Carlo integration thereon to obtain the expected value.

6.2 Pricing errors

We begin our numerical investigation with an analysis of the pricing errors incurred by the polynomial approximation procedure described in section 4.4. For this purpose, we define a pricing error as the mean-squared difference between an exact option price $V_{i,t}$ and the corresponding approximate option price $V_{i,t,(p)}$ at approximation order p, as captured by lemma 4.2. More precisely, we numerically determine the *relative* pricing error measure $\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)} - V_{i,t})^2]^{1/2}/\mathrm{E}^{\mathbb{P}}[V_{i,t}]$, which coincides up to a monotonic transformation with the objective function that the approximation procedure seeks to minimize. In theory, this construction prescribes a monotonic improvement in the pricing error measure when increasing the approximation order. As detailed in section 6.1, each expected value is determined by integration against the exact unconditional density for the SV1 model, whereas we employ Monte Carlo simulation as well as Fourier-cosine expansion methods for the SV1J and SV2J models. While pricing errors per se are not our main concern in this paper, we expect the analysis to provide some helpful intuition about general properties of the approximation procedure itself that may be helpful for our later analysis.

We start with equity derivatives as in section 3.3. Within the SV1 model, figure 1 visually illustrates the incurred pricing errors for different approximation orders and option specifications.

[Figure 1 about here.]

As expected, the approximation quality generally improves with increasing approximation order. Overall, the approximation procedure achieves pricing errors that are quite low in relative terms. Median pricing errors across all option specifications attain an order of magnitude of roughly 10^{-3} for the highest approximation order. Beyond aggregate pricing error levels, figure 1 prompts several observations regarding structural patterns of pricing errors across option specifications. First, we observe that pricing errors tend to decrease for longer maturities. Moreover, the improvement achievable by increasing approximation orders seems to be smaller for shorter-dated options. Second, we further observe that approximation errors tend to increase somewhat for deeper out-of-the-money options, especially for calls. This effect becomes markedly more pronounced at longer maturities.

We conduct an analogous analysis within the SV1J and SV2J models, for which table 2 reports aggregate pricing errors formed by determining the median and maximum of raw pricing errors across all considered option specifications.

[Table 2 about here.]

For the SV1J model, median pricing errors achieve an order of magnitude of roughly 10^{-3} under both Monte Carlo simulation in panel (a) and density approximation in panel (b), which is comparable to the closed-form results for the SV1 model reported in panel (c). For the SV2J model, pricing errors turn out to be of larger magnitude under both employed numerical methods in panels (a) and (b), with median pricing errors achieving roughly an order of magnitude of 10^{-2}. Despite the possible interference of numerical errors, we generally find that pricing errors improve monotonically when increasing the approximation order. Beyond that, under either numerical method, the underlying raw pricing errors (not reported) for the SV1J and SV2J models display structural patterns that are similar to those encountered in figure 1 for the SV1 model. Hence, the maximum pricing errors reported in table 2 tend to be essentially driven by deep out-of-the-money (call) options.

We repeat the preceding analysis for the case of volatility derivatives as in section 3.4. Within the SV1 model, figure 2 plots the pricing errors incurred for different approximation orders and option specifications.

[Figure 2 about here.]

Again, we generally witness the expected monotonic improvement of pricing errors with increasing approximation order. Aggregate pricing errors achieve comparable levels as for equity derivatives, with median pricing errors of about 10^{-3} at the highest approximation order. However, a larger dispersion can be observed across different option specifications. In particular, figure 2 suggests a strong maturity effect on the level of pricing errors as well as on the improvement due to larger approximation orders. Generally, levels decrease and improvements increase substantially for longer maturities. Likewise, we again observe a sizable moneyness effect in the sense that pricing errors tend to increase for deeper out-of-the-money options. This leads to rather monotonic patterns in figure 2, as we consider only calls. Additional unreported results show an analogous increasing behavior for deeper out-of-the-money puts.

For the SV1J and SV2J models, table 3 reports aggregate pricing errors across the considered option specifications.

[Table 3 about here.]

Overall, we find a similar picture as for equity derivatives. For the SV1J model, Monte Carlo simulation in panel (a) and density approximation in panel (b) achieve an order of magnitude of about 10^{-3}, comparable to the closed-form results obtained for the SV1 model reported in panel (c). For the SV2J model, however, pricing errors are higher at an order of magnitude of roughly $5{\times}10^{-2}$. As evidenced by the raw pricing error measures (not reported), the structural patters observed in figure 2 for the SV1 model remain largely intact also for the SV1J and SV2J models under either of the employed numerical methods. In consequence, deep out-of-the-money options typically determine the maximum pricing errors reported in table 3.

Overall, our approximation approach achieves acceptable pricing errors for both equity and volatility derivatives, even for the relatively low approximation orders that we consider in our analysis. What of course needs to be taken into account in their assessment is that our employed pricing error measure is itself subjected to numerical errors, as expected values need to be determined by numerical methods. Nevertheless, some structural patterns of pricing errors across different option specifications stand out.

In particular, we find that pricing errors are generally decreasing for longer maturities. With the spacing of strikes, derived from approximate conditional distributions, we try to ensure that options are in fact somewhat comparable across maturities. After controlling for differences in conditional distributions, the observed maturity effects on pricing errors appear to be driven by the shapes of the option price functions for different maturities. This is particularly apparent for volatility derivatives. At a hypothetical zero maturity, each option price as a function of latent states has a hockey stick shape, whose accurate approximation will require polynomials of relatively high order. As the maturity increases, the hockey stick shape will be smoothed out more and more, so that accurate approximation becomes feasible even with lower-order polynomials. In contrast, for equity options at the hypothetical zero maturity, each out-of-the-money option price as a function of latent states would be constant (equal to zero) and, therefore, easy to approximate by polynomials. For shorter maturities, out-of-the-money options bear value mostly for extremely large values of the latent state vector in the tails of the state distribution, while having close to zero value otherwise. This shape will again be smoothed out for larger maturities and become easier to approximate by low-order polynomials. What additionally explains part of the maturity effect of pricing errors for equity and volatility derivatives is the maturity dependence of the denominator in the *relative* pricing error measure (i.e., expected option prices), which generally decreases for shorter maturities and scales up the pricing error measure more aggressively.

In addition, we observe a moneyness effect in the sense that deeper out-of-the-money (call) options incur larger pricing errors. Their price function again is effectively supported in the tails of the state distribution, while being close to zero for typical values of the latent state vector. Moreover, what also contributes to the observed asymmetry is an asymmetric price effect due to the model-implied option smirk, whereby the denominator for call options much more strongly scales up the pricing error measure compared to the respective put options. In conjunction with our scaling of the strike range, this asymmetry becomes even more pronounced for the longer maturities that we consider.

Finally, the finding that pricing error measures generally increase for the two-dimensional SV2J model compared to the one-dimensional SV1 and SV1J models can be explained, at least to some degree, by the numerical errors incurred in the numerical benchmark methods. When using a direct integration approach, such errors may be driven by the use of sparse grids due to the infeasibility of dense full grids achieving a comparable resolution as attained in the one-dimensional cases. Similar effects are at work for Monte Carlo simulation, in the sense that a given sample size will accomplish a lower resolution of the state space in higher dimensions (whereas the convergence *rate* is known to be independent of the dimensionality).

6.3 Moment errors

While providing intuition about our approximation procedure, pricing errors themselves are not the primary concern in this paper. More importantly, we proceed to study moment errors that arise when replacing the exact moments in section 4.3 by the approximate moments in section 4.4. Specifically, we analyze *relative* moment errors of the form $|\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)})^N] - \mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]|/\mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]$ for some integer moment order N, where $V_{i,t}$ is an exact option price and $V_{i,t,(p)}$ is the corresponding approximate option price at some approximation order p. Unlike for pricing errors in section 6.2, this moment error measure does not directly imply monotonicity in the approximation order. It should be noted, however, that approximate first-order moments are in fact exact at any approximation order, for which reason we restrict our attention to moments of orders two to four. To determine $\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)})^N]$, we employ our approximation methodology developed in section 4.4, specifically the form established in proposition 4.2. For $\mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]$, we compute expected values as detailed in section 6.1, using integration against the exact unconditional density for the SV1 model, while relying on Monte Carlo simulation and Fourier-cosine expansion for the SV1J and SV2J models. This effectively yields a comparison of our methodology for constructing approximate moments involving option prices against some common (but computationally expensive) approximation procedures.

We again start the analysis with equity derivatives as in section 3.3. For the SV1 model, figure 3 plots moment errors for different moment orders, approximation orders, and option specifications.

[Figure 3 about here.]

Overall, moment errors settle on small levels for all option specifications considered, even for the relatively low approximation orders. Aggregate moment errors achieve a median order of magnitude of almost 10^{-6} for $N = 2$, 10^{-5} for $N = 3$, and around 10^{-4} for $N = 4$ at the highest approximation order. This reflects the general pattern that moment errors tend to increase with higher moment orders. Although monotonicity is not automatically assured, we generally observe that the approximation quality improves with increasing approximation order. We witness some further structural patterns of moment errors across different option specifications, consistent with those observed for pricing errors. First, moment errors tend to be larger in magnitude for longer maturities, at which also the effect of increasing approximation orders is slightly more pronounced. Second, we find that moment errors generally increase for deeper out-of-the-money options, which is particularly the case for calls at longer maturities.

For the SV1J and SV2J models, table 4 summarizes aggregate moment errors across the considered options specifications.

[Table 4 about here.]

For the SV1J model, moment errors computed from the density approximation method in panel (b) achieves levels comparable to those for the SV1 model reported in panel (c). Interestingly, moment errors obtained by Monte Carlo simulation in panel (a) are eventually unable to reach similar accuracy levels at the highest approximation order. For the SV2J model, we observe a similar discrepancy of Monte Carlo simulation in panel (a) compared to density approximation in panel (b). Relative to the SV1J model, each of the numerical methods yields moment errors that are of a larger order of magnitude, consistent with what has been diagnosed for pricing errors. As for the SV1 model in figure 3, the raw moment error measures (not reported) within the SV1J and SV2J models tend to increase with higher moment orders and display analogous structural patterns under each numerical method. In particular, this implies that the maximum moment errors reported in table 4 are mainly driven by deep out-of-the-money (call) options.

We continue the analysis with volatility derivatives as in section 3.4. For the SV1 model, figure 4 shows moment errors for different moment orders, approximation orders, and option specifications.

[Figure 4 about here.]

Aggregate moment errors for volatility derivatives are slightly higher than for equity derivatives, with a median order of magnitude of around 10^{-5} for $N = 2$, $5{\times}10^{-5}$ for $N = 3$, and 10^{-4} for $N = 4$ at the highest approximation order. Again, we find that moment errors tend to increase for higher moment orders. Generally, moment errors also tend to decrease more or less monotonically with increasing approximation order. Moreover, we observe some structural patters of moment errors across different option specifications, consistent with those encountered for pricing errors. In this respect, moment errors tend to be substantially lower for longer maturities, where also the effect of increasing approximation orders is more pronounced. With respect to the strike dimension, we find that moment errors are generally higher for deeper out-of-the-money options, which leads to a rather monotonic picture in figure 4 as only calls are considered.

For the SV1J and SV2J models, table 5 reports aggregate moment errors across the considered option specifications.

[Table 5 about here.]

Moment errors for volatility derivatives tend to be slightly higher than for equity derivatives. For the SV1J model, the density approximation method in panel (b) yields similar magnitudes compared to the closed-form results for the SV1 model reported in panel (c). Similar to the case of equity derivatives, Monte Carlo simulation in panel (a) is eventually unable to achieve the same accuracy levels as the direct integration methods. For the SV2J model, we observe a similar discrepancy of Monte Carlo simulation in panel (a) and density approximation in panel (b), while each generally yields larger moment errors than the respective method within the SV1J model. The raw moment error measures (not reported) within the SV1J and SV2J models imply that their general structural patterns continue to holds as for the SV1 model in figure 4 under each of the employed numerical methods. This especially implies that the maximum moment errors reported in table 5 are essentially determined by deep out-of-the-money options.

Overall, despite considering only relatively low approximation orders in the analysis, the moment error measures attain rather low magnitudes, even lower than for pricing errors. For some exceptional cases, especially short-dated volatility derivatives, larger approximations orders may be necessary in order to achieve a sufficient accuracy. Nevertheless, we observe similar structural patterns for moment errors as previously for pricing errors. In particular, moment errors exhibit a maturity effect in the sense that they decrease for longer maturities as well as a moneyness effect by which they increase for deeper out-of-the-money (call) options. The rationalizations discussed for the analogous structural patterns of pricing errors largely carry over to the case of moment errors. By considering higher-order monomials of option prices, some effects will even be aggravated, which may thus explain that moment errors tend to increase with moment orders.

Specific methodological challenges are posed by an even stronger relevance of the tails of the state distribution. In particular, this will be the case for deeper out-of-the-money options under higher moment orders. Numerical methodologies hereby need to accurately capture the tails of the distribution, which is a challenging task unless a closed-form distribution is known. For Monte Carlo simulation and the Fourier-cosine expansion, this may result in an even further increased computational burden to assure an appropriate accuracy, especially in multiple dimensions. Within our approximation procedure, deeper out-of-the-money options tend to go along with higher-frequency oscillations of integrands, which can effectively be dealt with using a sufficiently dense grid for one-dimensional numerical integration, thereby incurring some additional computational cost. An alternative approach could resort to expressions from complex Fourier theory and choose an appropriate regularization parameter without necessarily increasing the computational cost.

What is interesting beyond the structural patterns of moment errors is the apparent inability of Monte Carlo simulation to attain a similar approximation quality compared to the other methods considered, a discrepancy that we did not observe in our pricing error analysis. After some additional investigation, we attribute this mainly to some biases of the employed Euler–Maruyama scheme, which seem quite persistent even when increasing the number of sample paths or the number of intra-day steps.[17] Such biases eventually pollute the analysis of moment errors once they dominate the error measure.

7 Conclusion

In this paper, we develop a novel and unified methodology to incorporate observed derivatives prices into a GMM estimation procedure. To achieve this, we obtain a general pricing formula, covering a broad class of derivatives, using the generalized transform analysis introduced in Chen and Joslin (2012)

[17] A relatively low accuracy of Monte Carlo simulation with the Euler–Maruyama scheme is in line with results in Broadie and Kaya (2006) in the context of option pricing. In addition, Kloeden and Neuenkirch (2013) provide a theoretical overview of potential convergence issues of Euler–Maruyama schemes.

and further developed in Dillschneider (2020). Building on this general pricing formula, we then obtain exact and approximate expressions for moments involving polynomials of derivatives prices. While exact moments are analytically tractable, due to the requirement of multi-dimensional numerical integration, they fail to be computationally feasible except for low orders. In contrast, approximate moments require only one-dimensional numerical integration, making their implementation both analytically tractable and computationally attractive. We theoretically verify convergence of the approximate moments to their exact counterparts under standard regularity conditions. Numerical results within state-of-the-art stochastic volatility models further support our proposed methodology.

While we present our results within an affine jump diffusion framework for stochastic volatility models, the scope of our methodology extends far beyond this specific case. Affine jump diffusions allow to devise convenient procedures for determining various standard transforms of the state vector, which we provide by extending the procedure in Duffie et al. (2000). As our main results apart from that rely on generalized transform analysis, our methodology is applicable also to other model classes covered by this versatile theory (see Chen and Joslin (2012) and Dillschneider (2020) for various examples) and can even be extended further. We also discuss possible alternative formulations to arrive at similar results, building on so-called complex Fourier theory (see also Dillschneider (2020) for further details). Likewise, our approach applies beyond stochastic volatility models for equity indices. Potential further topics amenable to our methodology include interest rates, credit risk, dividends, and exchange rates. Certain pragmatic approximations may even make our approach applicable to American options.

Our methodology proposed in this paper is subject to several limitations and admits further extensions, which are left for future research. While exact expressions for moments involving derivatives prices are derived, their implementation — except for low orders — is beyond the scope of this paper. The challenge here is the curse of dimensionality in conjunction with highly oscillatory integrands, which makes sophisticated integration approaches necessary. For GMM estimation, option data is assumed to be available in regular panel form. However, observed data is typically not in this form, which requires an additional interpolation step prior to estimation that introduces measurement errors. Finally, within the present setup, it could be interesting to additionally incorporate high-frequency data on stock returns, analogous to Bollerslev and Zhou (2002) and Garcia et al. (2011).

Appendix

A Supplement to standard transform analysis

This appendix contains supplementary details of standard transform analysis. In section A.1, we state an extension of the classical Faà di Bruno formula, which yields an expression for derivatives of a composite function. The subsequent sections discuss details of the standard transform analysis for affine jump diffusions, particularly derivations of exponential moments (section A.2), pl-linear moments (section A.3), and polynomial moments (section A.4).

A.1 Faà di Bruno formula

Constantine and Savits (1996) generalize the classical Faà di Bruno formula to allow for partial derivatives of a composite function. For future reference, we state their main result in the following proposition for the case of complex differentiable functions.

Proposition A.1. *Let $f \circ g : \mathbb{C}^n \to \mathbb{C}^p$ with $x \mapsto f(g(x))$, where $f : \mathbb{C}^m \to \mathbb{C}^p$ with $y \mapsto f(y)$ and $g : \mathbb{C}^n \to \mathbb{C}^m$ with $x \mapsto g(x)$. For $\alpha \in \mathbb{N}^n$, we have that*

$$\partial_x^\alpha f(g(x)) = \sum_{|\beta| \leq |\alpha|} \partial_y^\beta f(g(x)) \sum_{\mathcal{Q}(\alpha,\beta)} M_{k,\ell}^\alpha \left(\partial_x^\ell g(x)\right)^k , \tag{A.1}$$

where for multi-indices $\alpha \in \mathbb{N}^n$ and $\beta \in \mathbb{N}^m$, the set $\mathcal{Q}(\alpha, \beta)$, consisting of ordered multi-multi-indices $k = [k_1, \dots, k_s] \in \mathbb{N}^{m \times s}$ and $\ell = [\ell_1, \dots, \ell_s] \in \mathbb{N}^{n \times s}$ for $s \leq |\beta|$, is defined by

$$\mathcal{Q}(\alpha, \beta) = \bigcup_{s=1}^{|\beta|} \left\{ (k, \ell) \in \mathbb{N}^{m \times s} \times \mathbb{N}^{n \times s} : |k_i| > 0,\, \ell_1 \succ \cdots \succ \ell_s \succ 0,\, \sum_{i=1}^{s} k_i = \beta,\, \sum_{i=1}^{s} |k_i| \ell_i = \alpha \right\} , \tag{A.2}$$

where $\succ$ denotes the (strict) lexicographic order. For elements of $\mathcal{Q}(\alpha, \beta)$, we further define the multi-multinomial coefficient

$$M_{k,\ell}^\alpha = \frac{(\alpha!)}{\prod_{i=1}^r (k_i!)(\ell_i!)^{|k_i|}} \tag{A.3}$$

and the tensor expression

$$(\partial_x^\ell g(x))^k = \prod_{i=1}^{r} (\partial_x^{\ell_i} g(x))^{k_i} . \tag{A.4}$$

Proof. See Constantine and Savits (1996). □

A.2 Exponential moments

In this section, we derive the exponential moments in proposition 2.1. We proceed in two steps. First, we state single-period exponential moments based on the standard transform analysis of Duffie et al. (2000). Second, we use these results to iteratively determine multi-period exponential moments.

To formulate the required regularity conditions, we consider the characteristic χ of the state process, formally defined by $\chi = (A_{\mu,X}, B_{\mu,X}, A_{\Omega,X}, B_{\Omega,X}, A_\lambda, B_\lambda, \nu)$, containing all affine coefficients and the jump size distribution driving the affine state dynamics in equation (2.1). To simplify the notation, we suppress the dependence of the elements of χ on $\mathbb{M}$.

A.2.1 Single-period exponential moments

We closely follow the exposition in Dillschneider (2020). Fixing $\omega = [\omega_S; \omega_Z] \in \mathbb{C}^{n_X}$ and $t, T \in \mathbb{R}_+$, we define the complex-valued process $(\mathit{\Psi}_\tau)_{0 \le \tau \le T}$ by

$$\mathit{\Psi}_\tau = \exp(A_\Psi(\omega; T-\tau) + [\omega_S; B_\Psi(\omega; T-\tau)] \cdot X_{t\oplus[\tau]}) \,. \tag{A.5}$$

Here, the complex-valued coefficient functions A_Ψ and B_Ψ solve the system of ODEs (A.8) of generalized Riccati type. Applying Ito's lemma to equation (A.5), $\mathit{\Psi}_\tau$ satisfies

$$\mathrm{d}\mathit{\Psi}_\tau = \mu_{\mathit{\Psi},\tau-}\,\mathrm{d}\tau + \sigma_{\mathit{\Psi},\tau-}\,\mathrm{d}W_{t+\tau} + J_{\mathit{\Psi},\tau}\,\mathrm{d}N_{t+\tau} \,, \tag{A.6}$$

where $\mu_{\mathit{\Psi},\tau}$ and $\sigma_{\mathit{\Psi},\tau}$ can be given as functions of $X_{t\oplus[\tau]}$, while $J_{\mathit{\Psi},\tau}$ depends exponentially on $X_{t\oplus[\tau]-}$ and $J_{X,t+\tau}$.

If the characteristic χ is well-behaved in the sense of assumption A.1, $(\mathit{\Psi}_\tau)_{0\le\tau\le T}$ is a well-defined martingale. The martingale property yields a well-known result due to Duffie et al. (2000), by which conditional exponential moments of the joint state vector are exponentially affine. The particular form presented in proposition A.2 holds for the state dynamics in equation (2.1).

If the characteristic χ is moreover well-behaved in the sense of assumption A.2, unconditional exponential moments can be obtained as a limit of the conditional exponential moments in proposition A.2 as in Dillschneider (2020). In essence, proposition A.3 yields the ergodicity result that $\Psi^{\mathbb{M}}(\omega; \infty) = \lim_{\tau\to\infty} \Psi^{\mathbb{M}}(\omega; \tau, z)$ irrespective of z.

Assumption A.1. *The characteristic χ is well-behaved at $(\omega, T) \in \mathbb{C}^{n_X} \times \mathbb{R}_+$ for $\omega = [\omega_S; \omega_Z]$ by satisfying the following conditions:*

(i) $\mathit{\Phi}_\nu(\tilde{\omega})$ exists for all $\tilde{\omega}$ in an open set $\mathcal{O}$ containing $\bigcup_{0\le\tau\le T}\{[\omega_S; B_\Psi(\omega;\tau)]\}$;

(ii) the system of ODEs (A.8) is solved uniquely on $[0, T]$;

(iii) for every $t \in \mathbb{R}_+$, the process $(\mathit{\Psi}_\tau)_{0\le\tau\le T}$ with dynamics in equation (A.6) satisfies:

- $\mathrm{E}^{\mathbb{M}}[|\mathit{\Psi}_0|] < \infty$,
- $\mathrm{E}^{\mathbb{M}}[(\int_0^T \Omega_{\mathit{\Psi},\tau-}\,\mathrm{d}\tau)^{1/2}] < \infty$ *with* $\Omega_{\mathit{\Psi},\tau} = \sigma_{\mathit{\Psi},\tau}\,\sigma_{\mathit{\Psi},\tau}^\top$,
- $\mathrm{E}^{\mathbb{M}}[\int_0^T |\tilde{J}_{\mathit{\Psi},\tau}\,\Lambda_{\mathit{\Psi},\tau}|\,\mathrm{d}\tau] < \infty$ *with* $\tilde{J}_{\mathit{\Psi},\tau} = \int J_{\mathit{\Psi},\tau}\,\mathrm{d}\nu$ *and* $\Lambda_{\mathit{\Psi},\tau} = \lambda(Z_{t+\tau-})$.

Proposition A.2. *Let χ be well-behaved at $(\omega, T) \in \mathbb{C}^{n_X} \times \mathbb{R}_+$ in the sense of assumption A.1 for $\omega = [\omega_S; \omega_Z]$. Then for all $t \in \mathbb{R}_+$ and $0 \le \tau \le T$, we have*

$$\begin{aligned}\Psi^{\mathbb{M}}(\omega; \tau, Z_t) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t\oplus[\tau]})\,|\,\mathcal{F}_t] \\ &= \exp(A_\Psi(\omega;\tau) + B_\Psi(\omega;\tau)\cdot Z_t)\end{aligned} \tag{A.7}$$

with coefficients $A_\Psi(\omega;\tau) \in \mathbb{C}$ and $B_\Psi(\omega;\tau) \in \mathbb{C}^{n_Z}$ determined by the system of ODEs[18]

$$\partial_\tau A_\Psi = A_{\mu,X}^\top[\omega_S; B_\Psi] + \tfrac{1}{2}A_{\Omega,X}^\top([\omega_S; B_\Psi]\otimes[\omega_S; B_\Psi]) + A_\lambda^\top(\mathit{\Phi}_\nu([\omega_S; B_\Psi]) - \iota) \tag{A.8a}$$

$$\partial_\tau B_\Psi = B_{\mu,X}^\top[\omega_S; B_\Psi] + \tfrac{1}{2}B_{\Omega,X}^\top([\omega_S; B_\Psi]\otimes[\omega_S; B_\Psi]) + B_\lambda^\top(\mathit{\Phi}_\nu([\omega_S; B_\Psi]) - \iota) \tag{A.8b}$$

subject to the initial conditions $A_\Psi(\omega; 0) = 0$ and $B_\Psi(\omega; 0) = \omega_Z$.

Proof. See Dillschneider (2020). □

[18] Here, $\otimes$ denotes the ordinary Kronecker product.

Assumption A.2. *The characteristic χ is well-behaved at $\omega \in \mathbb{C}^{nx}$ for $\omega = [0; \omega_Z]$ by satisfying the following conditions:*

(i) χ *is well-behaved at* $([0; \omega_Z], T)$ *in the sense of assumption A.1 for all* $T \geq 0$*;*

(ii) $\Psi^{\mathbb{M}}([0; \tilde{\omega}_Z]; \infty)$ *exists at* $\tilde{\omega}_Z = B_\Psi([0; \omega_Z]; \tau)$ *for all* $\tau \geq 0$*;*

(iii) $\tilde{\omega}_Z \mapsto \Psi^{\mathbb{M}}([0; \tilde{\omega}_Z]; \infty)$ *is continuous at* $\tilde{\omega}_Z = 0$*;*

(iv) $[0; \omega_Z] \in \mathcal{R}_\Psi$*, where* $\mathcal{R}_\Psi$ *denotes the stability region of the system of ODEs (A.8) containing all* $\tilde{\omega} = [0; \tilde{\omega}_Z] \in \mathbb{C}^{nx}$ *such that:*

- $A_\Psi(\tilde{\omega}; \infty) = \lim_{\tau \to \infty} A_\Psi(\tilde{\omega}; \tau)$ *exists and is finite,*
- $B_\Psi(\tilde{\omega}; \infty) = \lim_{\tau \to \infty} B_\Psi(\tilde{\omega}; \tau)$ *equals zero.*

Proposition A.3. *Let χ be well-behaved at $\omega \in \mathbb{C}^{nx}$ in the sense of assumption A.2 for $\omega = [0; \omega_Z]$. Then for all $t \in \mathbb{R}_+$, we have*

$$\begin{aligned} \Psi^{\mathbb{M}}(\omega; \infty) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_t)] \\ &= \exp(A_\Psi(\omega; \infty)) \end{aligned} \tag{A.9}$$

with $A_\Psi(\omega; \infty) = \lim_{\tau \to \infty} A_\Psi(\omega; \tau)$ *as in proposition A.2.*

Proof. See Dillschneider (2020). □

A.2.2 Multi-period exponential moments

With additional regularity conditions, the single-period moments in propositions A.2 and A.3 allow to iteratively determine multi-period exponential moments. Essentially, the conditions in assumptions A.3 and A.4 assure that the law of iterated expectations can be applied, which yields the conditional and unconditional exponential moment expression in propositions A.4 and A.5, respectively. Both are related via the limiting procedure $\Phi^{\mathbb{M}}(\omega; \tilde{\tau}, \infty) = \lim_{\tau \to \infty} \Phi^{\mathbb{M}}(\omega; \tilde{\tau}, \tau, z)$ irrespective of z.

Assumption A.3. *The characteristic χ is well-behaved at $(\omega, \tilde{\tau}, \tau) \in \mathbb{C}^{nx\tilde{n}} \times \mathbb{R}_+^{\tilde{n}} \times \mathbb{R}_+$ for $\omega = [\omega_S; \omega_Z]$ by satisfying the following conditions:*

(i) χ *is well-behaved at* $(\omega_i + [0; B_{\Phi,(i)}(\omega; \tilde{\tau})], \Delta_i)$ *in the sense of assumption A.1 for all* $1 \leq i \leq \tilde{n}$*;*

(ii) χ *is well-behaved at* $([0; B_{\Phi,(0)}(\omega; \tilde{\tau})], \tau)$ *in the sense of assumption A.1.*

Proposition A.4. *Let χ be well-behaved at $(\omega, \tilde{\tau}, \tau) \in \mathbb{C}^{nx\tilde{n}} \times \mathbb{R}_+^{\tilde{n}} \times \mathbb{R}_+$ in the sense of assumption A.3 for $\omega = [\omega_S; \omega_Z]$. Then for all $t \in \mathbb{R}_+$, we have*

$$\begin{aligned} \Phi^{\mathbb{M}}(\omega; \tilde{\tau}, \tau, Z_t) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t+\tau\oplus\tilde{\tau}}) \,|\, \mathcal{F}_t] \\ &= \exp(A_\Phi(\omega; \tilde{\tau}, \tau) + B_\Phi(\omega; \tilde{\tau}, \tau) \cdot Z_t) \end{aligned} \tag{A.10}$$

with coefficients $A_\Phi(\omega; \tilde{\tau}, \tau) \in \mathbb{C}$ *and* $B_\Phi(\omega; \tilde{\tau}, \tau) \in \mathbb{C}^{nz}$ *given by*

$$A_\Phi(\omega; \tilde{\tau}, \tau) = A_{\Phi,(0)}(\omega; \tilde{\tau}) + A_\Psi([0; B_{\Phi,(0)}(\omega; \tilde{\tau})]; \tau) \tag{A.11a}$$

$$B_\Phi(\omega; \tilde{\tau}, \tau) = B_\Psi([0; B_{\Phi,(0)}(\omega; \tilde{\tau})]; \tau)\ , \tag{A.11b}$$

depending on A_Ψ and A_Ψ as in proposition A.2. Defining $\Delta_i = \tilde{\tau}_i - \tilde{\tau}_{i-1}$ and $\omega_i = [\omega_{i,S}; \omega_{i,Z}]$, the auxiliary coefficients $A_{\Phi,(i)}(\omega; \tilde{\tau}) \in \mathbb{C}$ and $B_{\Phi,(i)}(\omega; \tilde{\tau}) \in \mathbb{C}^{nz}$ are determined by the backward recursion

$$A_{\Phi,(i-1)}(\omega; \tilde{\tau}) = A_{\Phi,(i)}(\omega; \tilde{\tau}) + A_\Psi(\omega_i + [0; B_{\Phi,(i)}(\omega; \tilde{\tau})]; \Delta_i) \tag{A.12a}$$

$$B_{\Phi,(i-1)}(\omega; \tilde{\tau}) = B_\Psi(\omega_i + [0; B_{\Phi,(i)}(\omega; \tilde{\tau})]; \Delta_i) \tag{A.12b}$$

for $i = \tilde{n}, \ldots, 1$ subject to the initial conditions $A_{\Phi,(\tilde{n})}(\omega; \tilde{\tau}) = 0$ and $B_{\Phi,(\tilde{n})}(\omega; \tilde{\tau}) = 0$, depending on A_Ψ and B_Ψ as in proposition A.2.

Proof. Under the imposed assumptions, we may repeatedly invoke the law of iterated expectations in conjunction with proposition A.2 to obtain the result. □

Assumption A.4. *The characteristic χ is well-behaved at $(\omega, \tilde{\tau}) \in \mathbb{C}^{nx\tilde{n}} \times \mathbb{R}_+^{\tilde{n}}$ for $\omega = [\omega_S; \omega_Z]$ by satisfying the following conditions:*

(i) χ is well-behaved at $(\omega_i + [0; B_{\Phi,(i)}(\omega; \tilde{\tau})], \Delta_i)$ in the sense of assumption A.1 for all $1 \leq i \leq \tilde{n}$;

(ii) χ is well-behaved at $[0; B_{\Phi,(0)}(\omega; \tilde{\tau})]$ in the sense of assumption A.2.

Proposition A.5. *Let χ be well-behaved at $(\omega, \tilde{\tau}) \in \mathbb{C}^{nx\tilde{n}} \times \mathbb{R}_+^{\tilde{n}}$ in the sense of assumption A.4 for $\omega = [\omega_S; \omega_Z]$. Then for all $t \in \mathbb{R}_+$, we have*

$$\begin{aligned} \Phi^{\mathbb{M}}(\omega; \tilde{\tau}, \infty) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t \oplus \tilde{\tau}})] \\ &= \exp(A_\Phi(\omega; \tilde{\tau}, \infty)) \end{aligned} \tag{A.13}$$

with coefficients $A_\Phi(\omega; \tilde{\tau}, \infty) \in \mathbb{C}$ and $B_\Phi(\omega; \tilde{\tau}, \infty) \in \mathbb{C}^{nz}$ given by

$$A_\Phi(\omega; \tilde{\tau}, \infty) = A_{\Phi,(0)}(\omega; \tilde{\tau}) + A_\Psi([0; B_{\Phi,(0)}(\omega; \tilde{\tau})]; \infty) \tag{A.14a}$$

$$B_\Phi(\omega; \tilde{\tau}, \infty) = B_\Psi([0; B_{\Phi,(0)}(\omega; \tilde{\tau})]; \infty) = 0\,, \tag{A.14b}$$

depending on A_Ψ and A_Ψ as in proposition A.3. Defining $\Delta_i = \tilde{\tau}_i - \tilde{\tau}_{i-1}$ and $\omega_i = [\omega_{i,S}; \omega_{i,Z}]$, the auxiliary coefficients $A_{\Phi,(i)}(\omega; \tilde{\tau}) \in \mathbb{C}$ and $B_{\Phi,(i)}(\omega; \tilde{\tau}) \in \mathbb{C}^{nz}$ are determined by the backward recursion (A.12), depending on A_Ψ and B_Ψ as in proposition A.2.

Proof. Under the imposed assumptions, proposition A.3 implies that $\Phi^{\mathbb{M}}(\omega; \tilde{\tau}, \infty)$ can be determined as the limit when $\tau \to \infty$ in equation (A.10), leading to equation (A.13). Letting $\tau \to \infty$ in equation (A.11) thereby yields the associated coefficients in equation (A.14), where $A_{\Phi,(i)}$ and $B_{\Phi,(i)}$ are given by the recursion in equation (A.12). □

A.3 Pl-linear moments

In this section, we derive the pl-linear moments in proposition 2.2. Again, we proceed in two steps. First, we state single-period pl-linear moments based on the transform analysis discussed in Dillschneider (2020), extending that of Duffie et al. (2000) used for exponential moments. Second, we use these results to iteratively determine multi-period pl-linear moments. As before, to formulate the required regularity conditions, we define the characteristic $\chi = (A_{\mu,X}, B_{\mu,X}, A_{\Omega,X}, B_{\Omega,X}, A_\lambda, B_\lambda, \nu)$, suppressing the dependence of the elements of χ on $\mathbb{M}$.

A.3.1 Single-period pl-linear moments

We closely follow the exposition in Dillschneider (2020). Fixing $\omega = [\omega_S; \omega_Z] \in \mathbb{C}^{n_X}$, $\alpha = [\alpha_S; \alpha_Z] \in \mathbb{N}^{n_X}$, and $t, T \in \mathbb{R}_+$, we define the complex-valued process $(\Psi_\tau^{(\alpha)})_{0 \le \tau \le T}$ by

$$\Psi_\tau^{(\alpha)} = \Psi_\tau \sum_{\tilde{\mathcal{Q}}(\alpha)} M_{k,\ell}^{\alpha} \, (A_\Psi^{(\ell)}(\omega; T - \tau) + [\omega_S^{(\ell)}; B_\Psi^{(\ell)}(\omega; T - \tau)] \cdot X_{t \oplus [\tau]})^k \,, \tag{A.15}$$

in terms of the complex-valued derivatives $A_\Psi^{(\beta)} = \partial_\omega^\beta A_\Psi$ and $B_\Psi^{(\beta)} = \partial_\omega^\beta B_\Psi$ as well as $\omega_S^{(\beta)} = \partial_\omega^\beta \omega_S$ and $\omega_Z^{(\beta)} = \partial_\omega^\beta \omega_Z$ for $\beta \in \mathbb{N}^{n_X}$. Under the regularity conditions formalized below, the coefficient functions $A_\Psi^{(\beta)}$ and $B_\Psi^{(\beta)}$ for $\beta \le \alpha$ solve the joint system of ODEs (A.18) of generalized Riccati type. Applying Ito's lemma to equation (A.15), $\Psi_t^{(\alpha)}$ satisfies

$$\mathrm{d}\Psi_\tau^{(\alpha)} = \mu_{\Psi,\tau-}^{(\alpha)} \, \mathrm{d}\tau + \sigma_{\Psi,\tau-}^{(\alpha)} \, \mathrm{d}W_{t+\tau} + J_{\Psi,\tau}^{(\alpha)} \, \mathrm{d}N_{t+\tau} \,, \tag{A.16}$$

where $\mu_{\Psi,\tau}^{(\alpha)}$ and $\sigma_{\Psi,\tau}^{(\alpha)}$ can be given as functions of $X_{t \oplus [\tau]}$, while $J_{\Psi,\tau}^{(\alpha)}$ depends pl-linearly (up to order $|\alpha|$) on $X_{t \oplus [\tau]-}$ and $J_{X,t+\tau}$.

If the characteristic χ is well-behaved in the sense of assumption A.5, $(\Psi_\tau^{(\alpha)})_{0 \le \tau \le T}$ is a well-defined martingale. These conditions generalize those in assumption A.1 imposed for exponential moments. As a consequence, the martingale property yields pl-linear moments of the joint state vector as in Dillschneider (2020). For the specific dynamics in equation (2.1), these are obtained by formally differentiating equation (A.7), i.e., $\Psi^{\mathbb{M},[\alpha]}(\omega; \tau, z) = \Psi^{\mathbb{M},(\alpha)}(\omega; \tau, z) = \partial_\omega^\alpha \Psi^{\mathbb{M}}(\omega; \tau, z)$.

If the characteristic χ is moreover well-behaved in the sense of assumption A.6, unconditional pl-linear moments can be obtained as a limit of the conditional pl-linear moments in proposition A.6 as in Dillschneider (2020). These conditions generalize those in assumption A.2 for exponential moments. Specifically, it holds that $\Psi^{\mathbb{M},[\alpha]}(\omega; \infty) = \lim_{\tau \to \infty} \Psi^{\mathbb{M},[\alpha]}(\omega; \tau, z)$ irrespective of z. Likewise, we have from proposition A.3 that $\Psi^{\mathbb{M},[\alpha]}(\omega; \infty) = \Psi^{\mathbb{M},(\alpha)}(\omega; \infty) = \partial_\omega^\alpha \Psi^{\mathbb{M}}(\omega; \infty)$.

Assumption A.5. *The characteristic χ is well-behaved at $(\omega, \alpha, T) \in \mathbb{C}^{n_X} \times \mathbb{N}^{n_X} \times \mathbb{R}_+$ for $\omega = [\omega_S; \omega_Z]$ and $\alpha = [\alpha_S; \alpha_Z]$ by satisfying the following conditions:*

(i) $\Phi_\nu^{[\beta]}(\tilde{\omega}) = \Phi_\nu^{(\beta)}(\tilde{\omega})$ exists in an open set $\mathcal{O}$ containing $\bigcup_{0 \le \tau \le T} \{[\omega_S; B_\Psi(\omega; \tau)]\}$ for all $|\beta| \le |\alpha|$;

(ii) the system of ODEs (A.18) is solved uniquely on $[0, T]$ for all $\beta \le \alpha$;

(iii) for every $t \in \mathbb{R}_+$, the process $(\Psi_\tau^{(\alpha)})_{0 \le \tau \le T}$ with dynamics in equation (A.16) satisfies:

- $\mathrm{E}^{\mathbb{M}}[|\Psi_0^{(\alpha)}|] < \infty$,
- $\mathrm{E}^{\mathbb{M}}[(\int_0^T \Omega_{\Psi,\tau-}^{(\alpha)} \, \mathrm{d}\tau)^{1/2}] < \infty$ *with* $\Omega_{\Psi,\tau}^{(\alpha)} = \sigma_{\Psi,\tau}^{(\alpha)} \, \sigma_{\Psi,\tau}^{(\alpha)\top}$,
- $\mathrm{E}^{\mathbb{M}}[\int_0^T |\tilde{J}_{\Psi,\tau}^{(\alpha)} \, \Lambda_{\Psi,\tau}| \, \mathrm{d}\tau] < \infty$ *with* $\tilde{J}_{\Psi,\tau}^{(\alpha)} = \int J_{\Psi,\tau}^{(\alpha)} \, \mathrm{d}\nu$ *and* $\Lambda_{\Psi,\tau} = \lambda(Z_{t+\tau-})$.

Proposition A.6. *Let χ be well-behaved at $(\omega, \alpha, T) \in \mathbb{C}^{n_X} \times \mathbb{N}^{n_X} \times \mathbb{R}_+$ in the sense of assumption A.5 for $\omega = [\omega_S; \omega_Z]$ and $\alpha = [\alpha_S; \alpha_Z]$. Then for all $t \in \mathbb{R}_+$, and $0 \le \tau \le T$, we have*

$$\begin{aligned} \Psi^{\mathbb{M},[\alpha]}(\omega; \tau, Z_t) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t \oplus [\tau]}) \, (X_{t \oplus [\tau]})^\alpha \,|\, \mathcal{F}_t] \\ &= \Psi^{\mathbb{M}}(\omega; \tau, Z_t) \sum_{\tilde{\mathcal{Q}}(\alpha)} M_{k,\ell}^{\alpha} \, (A_\Psi^{(\ell)}(\omega; \tau) + B_\Psi^{(\ell)}(\omega; \tau) \cdot Z_t)^k \end{aligned} \tag{A.17}$$

with coefficients $A_\Psi^{(\beta)}(\omega; \tau) = \partial_\omega^\beta A_\Psi(\omega; \tau) \in \mathbb{C}$ and $B_\Psi^{(\beta)}(\omega; \tau) = \partial_\omega^\beta B_\Psi(\omega; \tau) \in \mathbb{C}^{n_Z}$ for $\beta \le \alpha$ jointly *determined by the system of ODEs*

$$
\begin{aligned}
\partial_\tau A_\Psi^{(\beta)} &= A_{\mu,X}^\top[\omega_S^{(\beta)};B_\Psi^{(\beta)}] + \tfrac{1}{2}A_{\Omega,X}^\top \sum_{\eta\leq\beta}\binom{\beta}{\eta}([\omega_S^{(\eta)};B_\Psi^{(\eta)}]\otimes[\omega_S^{(\beta-\eta)};B_\Psi^{(\beta-\eta)}]) \\
&\quad + A_\lambda^\top \sum_{|\eta|\leq|\beta|}\Phi_\nu^{[\eta]}([\omega_S;B_\Psi])\sum_{\mathcal{Q}(\beta,\eta)} M_{k,\ell}^\beta([\omega_S^{(\ell)};B_\Psi^{(\ell)}])^k - \delta_0(\beta)\,A_\lambda^\top\iota \qquad \text{(A.18a)} \\
\partial_\tau B_\Psi^{(\beta)} &= B_{\mu,X}^\top[\omega_S^{(\beta)};B_\Psi^{(\beta)}] + \tfrac{1}{2}B_{\Omega,X}^\top \sum_{\eta\leq\beta}\binom{\beta}{\eta}([\omega_S^{(\eta)};B_\Psi^{(\eta)}]\otimes[\omega_S^{(\beta-\eta)};B_\Psi^{(\beta-\eta)}]) \\
&\quad + B_\lambda^\top \sum_{|\eta|\leq|\beta|}\Phi_\nu^{[\eta]}([\omega_S;B_\Psi])\sum_{\mathcal{Q}(\beta,\eta)} M_{k,\ell}^\beta([\omega_S^{(\ell)};B_\Psi^{(\ell)}])^k - \delta_0(\beta)\,B_\lambda^\top\iota \qquad \text{(A.18b)}
\end{aligned}
$$

subject to the initial conditions $A_\Psi^{(\beta)}(\omega;0)=0$ *and* $B_\Psi^{(\beta)}(\omega;0)=\omega_Z^{(\beta)}$. *Here, we set the Dirac indicator* $\delta_0(\beta)=1$ *if* $\beta=0$ *and* $\delta_0(\beta)=0$ *otherwise. Moreover,* $\Phi_\nu^{[\beta]}=\Phi_\nu^{(\beta)}=\partial_\omega^\beta\Phi_\nu$ *determine the pl-linear moments of jump sizes.*

Proof. See Dillschneider (2020). □

Assumption A.6. *The characteristic* χ *is well-behaved at* $(\omega,\alpha)\in\mathbb{C}^{n_X}\times\mathbb{N}^{n_X}$ *for* $\omega=[0;\omega_Z]$ *and* $\alpha=[0;\alpha_Z]$ *by satisfying the following conditions:*

(i) χ *is well-behaved at* $([0;\omega_Z],[0;\alpha_Z],T)$ *in the sense of assumption A.5 for all* $T\geq 0$;

(ii) $\Psi^{\mathbb{M},[\beta]}([0;\tilde\omega_Z];\infty)$ *exists at* $\tilde\omega_Z=B_\Psi([0;\omega_Z];\tau)$ *for all* $\tau\geq 0$ *and all* $\beta=[0;\beta_Z]$ *with* $|\beta|\leq|\alpha|$;

(iii) $\tilde\omega_Z\mapsto\Psi^{\mathbb{M},[\beta]}([0;\tilde\omega_Z];\infty)$ *is continuous at* $\tilde\omega_Z=0$ *for all* $\beta=[0;\beta_Z]$ *with* $|\beta|\leq|\alpha|$;

(iv) $[0;\omega_Z]\in\mathcal{R}_\Psi^{(\alpha)}$, *where* $\mathcal{R}_\Psi^{(\alpha)}$ *denotes the stability region of the system of ODEs (A.18) containing all* $\tilde\omega=[0;\tilde\omega_Z]\in\mathbb{C}^{n_X}$ *such that for all* $\beta\leq\alpha$:

- $A_\Psi^{(\beta)}(\tilde\omega;\infty)=\lim_{\tau\to\infty}A_\Psi^{(\beta)}(\tilde\omega;\tau)$ *exists and is finite,*
- $B_\Psi^{(\beta)}(\tilde\omega;\infty)=\lim_{\tau\to\infty}B_\Psi^{(\beta)}(\tilde\omega;\tau)$ *equals zero.*

Proposition A.7. *Let* χ *be well-behaved at* $(\omega,\alpha)\in\mathbb{C}^{n_X}\times\mathbb{N}^{n_X}$ *in the sense of assumption A.6 for* $\omega=[0;\omega_Z]$ *and* $\alpha=[0;\alpha_Z]$. *Then for all* $t\in\mathbb{R}_+$, *we have*

$$
\begin{aligned}
\Psi^{\mathbb{M},[\alpha]}(\omega;\infty) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega\cdot X_t)\,(X_t)^\alpha] \\
&= \exp(A_\Psi(\omega;\infty))\sum_{\tilde{\mathcal{Q}}(\alpha)} M_{k,\ell}^\alpha\,(A_\Psi^{(\ell)}(\omega;\infty))^k
\end{aligned}
\qquad \text{(A.19)}
$$

with $A_\Psi^{(\beta)}(\omega;\infty)=\lim_{\tau\to\infty}A_\Psi^{(\beta)}(\omega;\tau)$ *for* $\beta\leq\alpha$ *as in proposition A.6.*

Proof. See Dillschneider (2020). □

A.3.2 Multi-period pl-linear moments

As for exponential moments, additional regularity conditions allow to use the single-period moments in propositions A.6 and A.7 to iteratively determine multi-period pl-linear moments. The conditions in assumptions A.7 and A.8 assure the applicability of the law of iterated expectations in this case. Analogous to propositions A.4 and A.5, we can state conditional and unconditional pl-linear moment expressions in propositions A.8 and A.9, respectively. The results justify that $\Phi^{\mathbb{M},[\alpha]}(\omega;\tilde\tau,\sim)=\Phi^{\mathbb{M},(\alpha)}(\omega;\tilde\tau,\sim)=\partial_\omega^\alpha\Phi^{\mathbb{M}}(\omega;\tilde\tau,\sim)$ as well as $\Phi^{\mathbb{M},[\alpha]}(\omega;\tilde\tau,\infty)=\lim_{\tau\to\infty}\Phi^{\mathbb{M},[\alpha]}(\omega;\tilde\tau,\tau,z)$ irrespective of z.

Assumption A.7. *The characteristic χ is well-behaved at $(\omega, \alpha, \tilde{\tau}, \tau) \in \mathbb{C}^{nx\tilde{n}} \times \mathbb{N}^{nx\tilde{n}} \times \mathbb{R}_+^{\tilde{n}} \times \mathbb{R}_+$ for $\omega = [\omega_S; \omega_Z]$ and $\alpha = [\alpha_S; \alpha_Z]$ by satisfying the following conditions:*

(i) χ is well-behaved at $(\omega_i + [0; B_{\Phi,(i)}(\omega; \tilde{\tau})], \beta_i, \Delta_i)$ in the sense of assumption A.5 for all $\beta_i = [\alpha_{i,S}; \beta_{i,Z}]$ with $|\beta_{i,Z}| \leq \sum_{j=i}^{\tilde{n}} |\alpha_j|$ and all $1 \leq i \leq \tilde{n}$;

(ii) χ is well-behaved at $([0; B_{\Phi,(0)}(\omega; \tilde{\tau})], \beta_0, \tau)$ in the sense of assumption A.5 for all $\beta_0 = [0; \beta_{0,Z}]$ with $|\beta_{0,Z}| \leq \sum_{j=1}^{\tilde{n}} |\alpha_j|$.

Proposition A.8. *Let χ be well-behaved at $(\omega, \alpha, \tilde{\tau}, \tau) \in \mathbb{C}^{nx\tilde{n}} \times \mathbb{N}^{nx\tilde{n}} \times \mathbb{R}_+^{\tilde{n}} \times \mathbb{R}_+$ in the sense of assumption A.7 for $\omega = [\omega_S; \omega_Z]$ and $\alpha = [\alpha_S; \alpha_Z]$. Then for all $t \in \mathbb{R}_+$, we have*

$$\begin{aligned} \Phi^{\mathbb{M},[\alpha]}(\omega; \tilde{\tau}, \tau, Z_t) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t+\tau\oplus\tilde{\tau}}) \, (X_{t+\tau\oplus\tilde{\tau}})^{\alpha} \,|\, \mathcal{F}_t] \\ &= \Phi^{\mathbb{M}}(\omega; \tilde{\tau}, \tau, Z_t) \sum_{\tilde{\mathcal{Q}}(\alpha)} M_{k,\ell}^{\alpha} \, (A_{\Phi}^{(\ell)}(\omega; \tilde{\tau}, \tau) + B_{\Phi}^{(\ell)}(\omega; \tilde{\tau}, \tau) \cdot Z_t)^k \end{aligned} \tag{A.20}$$

with coefficients $A_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, \tau) = \partial_{\omega}^{\beta} A_{\Phi}(\omega; \tilde{\tau}, \tau) \in \mathbb{C}$ and $B_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, \tau) = \partial_{\omega}^{\beta} B_{\Phi}(\omega; \tilde{\tau}, \tau) \in \mathbb{C}^{nz}$ for $\beta \leq \alpha$ given by

$$A_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, \tau) = A_{\Phi,(0)}^{(\beta)}(\omega; \tilde{\tau}) + \sum_{|\eta| \leq |\beta|} A_{\Psi}^{([0;\eta])}([0; B_{\Phi,(0)}(\omega; \tilde{\tau})]; \tau) \sum_{\mathcal{Q}(\beta,\eta)} M_{k,\ell}^{\beta} \, (B_{\Phi,(0)}^{(\ell)}(\omega; \tilde{\tau}))^k \tag{A.21a}$$

$$B_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, \tau) = \sum_{|\eta| \leq |\beta|} B_{\Psi}^{([0;\eta])}([0; B_{\Phi,(0)}(\omega; \tilde{\tau})]; \tau) \sum_{\mathcal{Q}(\beta,\eta)} M_{k,\ell}^{\beta} \, (B_{\Phi,(0)}^{(\ell)}(\omega; \tilde{\tau}))^k \,, \tag{A.21b}$$

depending on $A_{\Psi}^{(\beta)}$ and $B_{\Psi}^{(\beta)}$ as in proposition A.6. Defining $\Delta_i = \tilde{\tau}_i - \tilde{\tau}_{i-1}$ and $\omega_i^{(\beta)} = \partial_{\omega}^{\beta} \omega_i$ for $\omega_i = [\omega_{i,S}; \omega_{i,Z}]$, the auxiliary coefficients $A_{\Phi,(i)}^{(\beta)}(\omega; \tilde{\tau}) = \partial_{\omega}^{\beta} A_{\Phi,(i)}(\omega; \tilde{\tau}) \in \mathbb{C}$ and $B_{\Phi,(i)}^{(\beta)}(\omega; \tilde{\tau}) = \partial_{\omega}^{\beta} B_{\Phi,(i)}(\omega; \tilde{\tau}) \in \mathbb{C}^{nz}$ for $\beta \leq \alpha$ are determined by the backward recursion

$$\begin{aligned} A_{\Phi,(i-1)}^{(\beta)}(\omega; \tilde{\tau}) = {} & A_{\Phi,(i)}^{(\beta)}(\omega; \tilde{\tau}) \\ & + \sum_{|\eta| \leq |\beta|} A_{\Psi}^{(\eta)}(\omega_i + [0; B_{\Phi,(i)}(\omega; \tilde{\tau})]; \Delta_i) \sum_{\mathcal{Q}(\beta,\eta)} M_{k,\ell}^{\beta} \, (\omega_i^{(\ell)} + [0; B_{\Phi,(i)}^{(\ell)}(\omega; \tilde{\tau})])^k \end{aligned} \tag{A.22a}$$

$$B_{\Phi,(i-1)}^{(\beta)}(\omega; \tilde{\tau}) = \sum_{|\eta| \leq |\beta|} B_{\Psi}^{(\eta)}(\omega_i + [0; B_{\Phi,(i)}(\omega; \tilde{\tau})]; \Delta_i) \sum_{\mathcal{Q}(\beta,\eta)} M_{k,\ell}^{\beta} \, (\omega_i^{(\ell)} + [0; B_{\Phi,(i)}^{(\ell)}(\omega; \tilde{\tau})])^k \tag{A.22b}$$

for $i = \tilde{n}, \ldots, 1$ subject to the initial conditions $A_{\Phi,(\tilde{n})}^{(\beta)}(\omega; \tilde{\tau}) = 0$ and $B_{\Phi,(\tilde{n})}^{(\beta)}(\omega; \tilde{\tau}) = 0$, depending on $A_{\Psi}^{(\beta)}$ and $B_{\Psi}^{(\beta)}$ as in proposition A.6.

Proof. Under the imposed assumptions, repeatedly invoking the law of iterated expectations and proposition A.6 yields $\Phi^{\mathbb{M},[\beta]}(\omega; \tilde{\tau}, \tau, z) = \Phi^{\mathbb{M},(\beta)}(\omega; \tilde{\tau}, \tau, z)$. Proposition A.4 and the Faà di Bruno formula (A.1) thus yield the required results. Specifically, equations (A.20) to (A.22) obtain from equations (A.10) to (A.12), respectively, by differentiation. ☐

Assumption A.8. *The characteristic χ is well-behaved at $(\omega, \alpha, \tilde{\tau}) \in \mathbb{C}^{nx\tilde{n}} \times \mathbb{N}^{nx\tilde{n}} \times \mathbb{R}_+^{\tilde{n}}$ for $\omega = [\omega_S; \omega_Z]$ and $\alpha = [\alpha_S; \alpha_Z]$ by satisfying the following conditions:*

(i) χ is well-behaved at $(\omega_i + [0; B_{\Phi,(i)}(\omega; \tilde{\tau})], \beta_i, \Delta_i)$ in the sense of assumption A.5 for all $\beta_i = [\alpha_{i,S}; \beta_{i,Z}]$ with $|\beta_{i,Z}| \leq \sum_{j=i}^{\tilde{n}} |\alpha_j|$ and all $1 \leq i \leq \tilde{n}$;

(ii) χ is well-behaved at $([0; B_{\Phi,(0)}(\omega; \tilde{\tau})], \beta_0)$ in the sense of assumption A.6 for all $\beta_0 = [0; \beta_{0,Z}]$ with $|\beta_{0,Z}| \leq \sum_{j=1}^{\tilde{n}} |\alpha_j|$.

Proposition A.9. *Let χ be well-behaved at $(\omega, \alpha, \tilde{\tau}) \in \mathbb{C}^{n_X \tilde{n}} \times \mathbb{N}^{n_X \tilde{n}} \times \mathbb{R}_+^{\tilde{n}}$ in the sense of assumption A.8 for $\omega = [\omega_S; \omega_Z]$ and $\alpha = [\alpha_S; \alpha_Z]$. Then for all $t \in \mathbb{R}_+$, we have*

$$\begin{aligned}\Phi^{\mathbb{M},[\alpha]}(\omega; \tilde{\tau}, \infty) &= \mathrm{E}^{\mathbb{M}}[\exp(\omega \cdot X_{t \oplus \tilde{\tau}})\,(X_{t\oplus\tilde{\tau}})^{\alpha}] \\ &= \Phi^{\mathbb{M}}(\omega; \tilde{\tau}, \infty) \sum_{\tilde{\mathcal{Q}}(\alpha)} M_{k,\ell}^{\alpha}\,(A_{\Phi}^{(\ell)}(\omega; \tilde{\tau}, \infty))^k\end{aligned} \tag{A.23}$$

with coefficients $A_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, \infty) = \partial_{\omega}^{\beta} A_{\Phi}(\omega; \tilde{\tau}, \infty) \in \mathbb{C}$ and $B_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, \infty) = \partial_{\omega}^{\beta} B_{\Phi}(\omega; \tilde{\tau}, \infty) \in \mathbb{C}^{n_Z}$ for $\beta \le \alpha$ given by

$$A_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, \infty) = A_{\Phi,(0)}^{(\beta)}(\omega; \tilde{\tau}) + \sum_{|\eta| \le |\beta|} A_{\Psi}^{([0;\eta])}([0; B_{\Phi,(0)}(\omega; \tilde{\tau})]; \infty) \sum_{\mathcal{Q}(\beta,\eta)} M_{k,\ell}^{\beta}\,(B_{\Phi,(0)}^{(\ell)}(\omega; \tilde{\tau}))^k \tag{A.24a}$$

$$B_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, \infty) = \sum_{|\eta| \le |\beta|} B_{\Psi}^{([0;\eta])}([0; B_{\Phi,(0)}(\omega; \tilde{\tau})]; \infty) \sum_{\mathcal{Q}(\beta,\eta)} M_{k,\ell}^{\beta}\,(B_{\Phi,(0)}^{(\ell)}(\omega; \tilde{\tau}))^k = 0\,, \tag{A.24b}$$

depending on $A_{\Psi}^{(\beta)}$ and $B_{\Psi}^{(\beta)}$ as in proposition A.7. Defining $\Delta_i = \tilde{\tau}_i - \tilde{\tau}_{i-1}$ and $\omega_i^{(\beta)} = \partial_{\omega}^{\beta} \omega_i$ for $\omega_i = [\omega_{i,S}; \omega_{i,Z}]$, the auxiliary coefficients $A_{\Phi,(i)}^{(\beta)}(\omega; \tilde{\tau}) = \partial_{\omega}^{\beta} A_{\Phi,(i)}(\omega; \tilde{\tau}) \in \mathbb{C}$ and $B_{\Phi,(i)}^{(\beta)}(\omega; \tilde{\tau}) = \partial_{\omega}^{\beta} B_{\Phi,(i)}(\omega; \tilde{\tau}) \in \mathbb{C}^{n_Z}$ for $\beta \le \alpha$ are determined by the backward recursion (A.22), depending on $A_{\Psi}^{(\beta)}$ and $B_{\Psi}^{(\beta)}$ as in proposition A.6.

Proof. Under the imposed assumptions, proposition A.7 implies that $\Phi^{\mathbb{M},[\alpha]}(\omega; \tilde{\tau}, \infty)$ can be determined as the limit when $\tau \to \infty$ in equation (A.20), leading to equation (A.23). Letting $\tau \to \infty$ in equation (A.21) thereby yields the associated coefficients in equation (A.24), where $A_{\Phi,(i)}^{(\beta)}$ and $B_{\Phi,(i)}^{(\beta)}$ are given by the recursion in equation (A.22). □

A.4 Polynomial moments

In this section, we derive the polynomial moments in proposition 2.3. As such, polynomial moments may be computed as a special case of the pl-linear moments in section A.3, generally relying on numerical solutions to the ODEs (A.18). This section discusses an alternative approach to arrive at closed-form expressions for polynomial moments, for which we proceed in two steps. First, we state single-period polynomial moments based on the result of Dillschneider (2020). Second, we use these results to iteratively determine multi-period polynomial moments. As before, we formulate the required regularity conditions in terms of the characteristic $\chi = (A_{\mu,X}, B_{\mu,X}, A_{\Omega,X}, B_{\Omega,X}, A_{\lambda}, B_{\lambda}, \nu)$, suppressing the dependence on $\mathbb{M}$.

A.4.1 Single-period polynomial moments

We closely follow the exposition in Dillschneider (2020). Considering $\alpha \in \mathbb{N}^{n_X}$ with $|\alpha| \le p$, Ito's lemma yields that each monomial $(X_t)^{\alpha}$ satisfies

$$\mathrm{d}(X_t)^{\alpha} = \mu_{X,t-}^{(\alpha)}\,\mathrm{d}t + \sigma_{X,t-}^{(\alpha)}\,\mathrm{d}W_t + J_{X,t}^{(\alpha)}\,\mathrm{d}N_t\,.$$

Here, $\mu_{X,t}^{(\alpha)}$ and $\sigma_{X,t}^{(\alpha)}$ can be given as functions of X_t, whereas $J_{X,t}^{(\alpha)}$ depends polynomially (up to order $|\alpha|$) on X_{t-} and $J_{X,t}$. Collecting terms, we can write

$$\mu_{X,t-}^{(\alpha)} + \tilde{J}_{X,t}^{(\alpha)}\,\lambda(Z_{t-}) = \sum_{|\beta| \le |\alpha|} b_{X,\beta}^{(\alpha)}\,(X_{t-})^{\beta}$$

for coefficients $b_{X,\beta}^{(\alpha)} \in \mathbb{R}$, where $\tilde{J}_{X,t}^{(\alpha)} = \int J_{X,t}^{(\alpha)}\,\mathrm{d}\nu$.

Fixing $t, T \in \mathbb{R}_+$, now define complex-valued processes $(\mathit{\Psi}_\tau^{(\alpha)})_{0\le\tau\le T}$ by

$$\mathit{\Psi}_\tau^{(\alpha)} = \sum_{|\beta|\le|\alpha|} b_{\Psi,\beta}^{(\alpha)}(T-\tau)\,(X_{t\oplus[\tau]})^\beta\,, \tag{A.25}$$

where the coefficient functions $b_{\Psi,\beta}^{(\alpha)}(\tau) \in \mathbb{R}$ in terms of $b_{X,\beta}^{(\eta)}$ for each $|\alpha| \le p$ and $|\beta| \le |\eta| \le |\alpha|$ are determined by the ODEs (A.28). Applying Ito's lemma to equation (A.25), $\mathit{\Psi}_\tau^{(\alpha)}$ satisfies

$$\mathrm{d}\mathit{\Psi}_\tau^{(\alpha)} = \mu_{\Psi,\tau-}^{(\alpha)}\,\mathrm{d}\tau + \sigma_{\Psi,\tau-}^{(\alpha)}\,\mathrm{d}W_{t+\tau} + J_{\Psi,\tau}^{(\alpha)}\,\mathrm{d}N_{t+\tau}\,, \tag{A.26}$$

where $\mu_{\Psi,\tau}^{(\alpha)}$ and $\sigma_{\Psi,\tau}^{(\alpha)}$ can be given as functions of $X_{t\oplus[\tau]}$, while $J_{\Psi,\tau}^{(\alpha)}$ depends polynomially (up to order $|\alpha|$) on $X_{t\oplus[\tau]-}$ and $J_{X,t+\tau}$.

If the characteristic χ is well-behaved in the sense of assumption A.9, $(\mathit{\Psi}_\tau^{(\alpha)})_{0\le\tau\le T}$ is a well-defined martingale for each $|\alpha| \le p$ and all $t \in \mathbb{R}_+$. The martingale property then yields polynomial moments of the joint state vector as in Dillschneider (2020), providing an analogue to proposition A.6. Unlike before, the conditional polynomial moments in assumption A.9 can be given in closed form, as the matrix exponential and its integral in equation (A.29) allow for closed-form expressions.

If the characteristic χ is moreover well-behaved in the sense of assumption A.10, we further obtain an analogue to proposition A.7 as in Dillschneider (2020). Consistent with the pl-linear case, unconditional polynomial moments can be derived from conditional polynomial moments by a limiting procedure such that $\Psi^{\mathbb{M},[\alpha]}(0;\infty) = \lim_{\tau\to\infty} \Psi^{\mathbb{M},[\alpha]}(0;\tau,z)$ irrespective of z. Unlike before, the unconditional polynomial moments in assumption A.10 can be given in closed form using equation (A.31).

Assumption A.9. *The characteristic χ is well-behaved at $(p,T) \in \mathbb{N}\times\mathbb{R}_+$ by satisfying the following conditions for all $\alpha \in \mathbb{N}^n$ with $\alpha = [\alpha_S;\alpha_Z]$ such that $|\alpha| \le p$:*

(i) $\mathit{\Phi}_\nu^{[\alpha]}(\tilde\omega) = \mathit{\Phi}_\nu^{(\alpha)}(\tilde\omega)$ *exist in an open set $\mathcal{O}$ containing the origin;*

(ii) the system of ODEs (A.28) is solved uniquely on $[0,T]$;

(iii) for every $t \in \mathbb{R}_+$, the process $(\mathit{\Psi}_\tau^{(\alpha)})_{0\le\tau\le T}$ with dynamics in equation (A.26) satisfies:

- $\mathrm{E}^{\mathbb{M}}[|\mathit{\Psi}_0^{(\alpha)}|] < \infty$,
- $\mathrm{E}^{\mathbb{M}}[(\int_0^T \Omega_{\Psi,\tau-}^{(\alpha)}\,\mathrm{d}\tau)^{1/2}] < \infty$ *with* $\Omega_{\Psi,\tau}^{(\alpha)} = \sigma_{\Psi,\tau}^{(\alpha)}\,\sigma_{\Psi,\tau}^{(\alpha)\top}$,
- $\mathrm{E}^{\mathbb{M}}[\int_0^T |\tilde J_{\Psi,\tau}^{(\alpha)}\,\Lambda_{\Psi,\tau}|\,\mathrm{d}\tau] < \infty$ *with* $\tilde J_{\Psi,\tau}^{(\alpha)} = \int J_{\Psi,\tau}^{(\alpha)}\,\mathrm{d}\nu$ *and* $\Lambda_{\Psi,\tau} = \lambda(Z_{t+\tau-})$.

Proposition A.10. *Let χ be well-behaved at $(p,T) \in \mathbb{N}\times\mathbb{R}_+$ in the sense of assumption A.9 for $\alpha \in \mathbb{N}^n$ with $\alpha = [\alpha_S;\alpha_Z]$ and $|\alpha| \le p$. Then for all $t \in \mathbb{R}_+$ and $0 \le \tau \le T$, we have*

$$\begin{aligned}\Psi^{\mathbb{M},[\alpha]}(0;\tau,Z_t) &= \mathrm{E}^{\mathbb{M}}[(X_{t\oplus[\tau]})^\alpha\,|\,\mathcal{F}_t] \\ &= \sum_{|\beta|\le|\alpha|} b_{\Psi,[0;\beta]}^{(\alpha)}(\tau)\,(Z_t)^\beta\end{aligned} \tag{A.27}$$

with coefficients $b_{\Psi,\beta}^{(\alpha)}(\tau) \in \mathbb{R}$ for $|\beta| \le |\alpha| \le p$ jointly determined by the system of ODEs

$$\partial_\tau b_{\Psi,\beta}^{(\alpha)}(\tau) = \sum_{|\beta|\le|\eta|\le|\alpha|} b_{\Psi,\eta}^{(\alpha)}(\tau)\,b_{X,\beta}^{(\eta)} \tag{A.28}$$

subject to the initial condition $b_{\Psi,\beta}^{(\alpha)}(0) = \delta_0(\alpha-\beta)$.

Collecting all coefficients in vectors and matrices such that $[\tilde{A}_X(\tau)]_\alpha = b^{(\alpha)}_{\Psi,0}(\tau)$ *and* $[\tilde{B}_X(\tau)]_{\alpha,\beta} = b^{(\alpha)}_{\Psi,\beta}(\tau)$ *as well as* $[\tilde{C}_X]_\alpha = b^{(\alpha)}_{X,0}$ *and* $[\tilde{D}_X]_{\alpha,\beta} = b^{(\alpha)}_{X,\beta}$*, we have the closed-form solutions*

$$\tilde{A}_X(\tau) = \int_0^\tau \exp((\tau - t)\tilde{D}_X)\,\tilde{C}_X\,\mathrm{d}t \tag{A.29a}$$

$$\tilde{B}_X(\tau) = \exp(\tau\tilde{D}_X)\,. \tag{A.29b}$$

Proof. See Dillschneider (2020). □

Assumption A.10. *The characteristic* χ *is well-behaved at* $p \in \mathbb{N}$ *by satisfying the following conditions for all* $\alpha \in \mathbb{N}^n$ *with* $\alpha = [0; \alpha_Z]$ *such that* $|\alpha| \leq p$*:*

(i) χ *is well-behaved at* T *in the sense of assumption A.9 for all* $T \geq 0$*;*

(ii) $\Psi^{\mathbb{M},[\alpha]}(0;\infty)$ *exists;*

(iii) the system of ODEs (A.28) is stable such that:

- $b^{(\alpha)}_{\Psi,\beta}(\infty) = \lim_{\tau\to\infty} b^{(\alpha)}_{\Psi,\beta}(\tau)$ *exists and is finite for* $\beta = [0;\beta_Z] = 0$*,*
- $b^{(\alpha)}_{\Psi,\beta}(\infty) = \lim_{\tau\to\infty} b^{(\alpha)}_{\Psi,\beta}(\tau)$ *equals zero for* $\beta = [0;\beta_Z] \neq 0$*.*

Proposition A.11. *Let* χ *be well-behaved at* $(\omega, p) \in \mathbb{C}^{n_X} \times \mathbb{N}$ *in the sense of assumption A.10 for* $\alpha \in \mathbb{N}^n$ *with* $\alpha = [0; \alpha_Z]$ *and* $|\alpha| \leq p$*. Then for all* $t \in \mathbb{R}_+$*, we have*

$$\begin{aligned}\Psi^{\mathbb{M},[\alpha]}(0;\infty) &= \mathrm{E}^{\mathbb{M}}[(X_t)^\alpha] \\ &= b^{(\alpha)}_{\Psi,0}(\infty)\end{aligned} \tag{A.30}$$

with $b^{(\alpha)}_{\Psi,0}(\infty) = \lim_{\tau\to\infty} b^{(\alpha)}_{\Psi,0}(\tau)$ *as in proposition A.10.*

Collecting all coefficients in vectors and matrices such that $[\tilde{A}_X(\infty)]_\alpha = b^{(\alpha)}_{\Psi,0}(\infty)$ *and* $[\tilde{B}_X(\infty)]_{\alpha,\beta} = b^{(\alpha)}_{\Psi,\beta}(\infty)$ *as well as* $[\tilde{C}_Z]_\alpha = b^{([0;\alpha])}_{X,0}$ *and* $[\tilde{D}_Z]_{\alpha,\beta} = b^{([0;\alpha])}_{X,[0;\beta]}$*, we have the closed-form solutions*

$$\begin{aligned}\tilde{A}_Z(\infty) &= -\,\tilde{D}_Z^{-1}\,\tilde{C}_Z \\ \tilde{B}_Z(\infty) &= 0\,.\end{aligned} \tag{A.31}$$

Proof. See Dillschneider (2020). □

A.4.2 Multi-period polynomial moments

Under appropriate regularity conditions, the single-period moments in propositions A.10 and A.11 allow to iteratively determine multi-period polynomial moments. The conditions in assumptions A.11 and A.12 assure that the law of iterated expectations can be applied in this case. Analogous to propositions A.8 and A.9, this yields the conditional and unconditional polynomial moment expressions in propositions A.12 and A.13, respectively. The results preserve the relations $\Phi^{\mathbb{M},[\alpha]}(0;\tilde{\tau},\sim) = \Phi^{\mathbb{M},(\alpha)}(0;\tilde{\tau},\sim) = \partial^\alpha_\omega \Phi^{\mathbb{M}}(0;\tilde{\tau},\sim)$ and $\Phi^{\mathbb{M},[\alpha]}(0;\tilde{\tau},\infty) = \lim_{\tau\to\infty}\Phi^{\mathbb{M},[\alpha]}(0;\tilde{\tau},\tau,z)$ irrespective of z.

Assumption A.11. *The characteristic* χ *is well-behaved at* $(\alpha,\tilde{\tau},\tau) \in \mathbb{N}^{n_X\tilde{n}} \times \mathbb{R}^{\tilde{n}}_+ \times \mathbb{R}_+$ *for* $\alpha = [\alpha_S;\alpha_Z]$ *by satisfying the following conditions:*

(i) χ *is well-behaved at* $(\sum_{j=i}^{\tilde{n}}|\alpha_j|, \Delta_i)$ *in the sense of assumption A.9 for all* $1 \leq i \leq \tilde{n}$*;*

(ii) χ *is well-behaved at* $(\sum_{j=1}^{\tilde{n}}|\alpha_j|, \tau)$ *in the sense of assumption A.9.*

Proposition A.12. *Let χ be well-behaved at $(\alpha, \tilde{\tau}, \tau) \in \mathbb{N}^{nx\tilde{n}} \times \mathbb{R}_+^{\tilde{n}} \times \mathbb{R}_+$ in the sense of assumption A.11 for $\alpha = [\alpha_S; \alpha_Z]$. Then for all $t \in \mathbb{R}_+$, we have*

$$\begin{aligned}\Phi^{\mathbb{M},[\alpha]}(0; \tilde{\tau}, \tau, Z_t) &= \mathrm{E}^{\mathbb{M}}[(X_{t+\tau\oplus\tilde{\tau}})^{\alpha} \,|\, \mathcal{F}_t] \\ &= \sum_{|\beta|\leq|\alpha|} b_{\Phi,\beta}^{(\alpha)}(\tilde{\tau}, \tau)\, (Z_t)^{\beta}\end{aligned} \tag{A.32}$$

with coefficients $b_{\Phi,\beta}^{(\alpha)}(\tilde{\tau}, \tau) \in \mathbb{R}$ given by

$$b_{\Phi,\beta}^{(\alpha)}(\tilde{\tau}, \tau) = \sum_{\eta} b_{\Phi,\eta,(0)}^{(\alpha)}(\tilde{\tau})\, b_{\Psi,[0;\beta]}^{([0;\eta])}(\tau)\, . \tag{A.33}$$

Defining $\Delta_i = \tilde{\tau}_i - \tilde{\tau}_{i-1}$, the auxiliary coefficients $b_{\Phi,\eta,(i)}^{(\alpha)}(\tilde{\tau}) \in \mathbb{R}$ are determined by the backward recursion

$$b_{\Phi,\beta,(i-1)}^{(\alpha)}(\tilde{\tau}) = \sum_{\eta} b_{\Phi,\eta,(i)}^{(\alpha)}(\tilde{\tau})\, b_{\Psi,[0;\beta]}^{(\alpha_i+[0;\eta])}(\Delta_i) \tag{A.34}$$

for $i = \tilde{n}, \ldots, 1$ subject to the initial condition $b_{\Phi,\beta,(\tilde{n})}^{(\alpha)}(\tilde{\tau}) = 0$. In equation (A.34), the coefficients $b_{\Psi,\beta}^{(\eta+\alpha_i)}(\Delta_i)$ are taken from $A_X(\Delta_i)$ and $B_X(\Delta_i)$ in equation (A.29).

Proof. Under the imposed assumptions, we repeatedly invoke the law of iterated expectations and proposition A.10. Collecting terms in equation (A.27) thus yields the expressions in equations (A.32) to (A.34). □

Assumption A.12. *The characteristic χ is well-behaved at $(\alpha, \tilde{\tau}) \in \mathbb{N}^{nx\tilde{n}} \times \mathbb{R}_+^{\tilde{n}}$ for $\alpha = [\alpha_S; \alpha_Z]$ by satisfying the following conditions:*

(i) χ is well-behaved at $(\sum_{j=i}^{\tilde{n}} |\alpha_j|, \Delta_i)$ in the sense of assumption A.9 for all $1 \leq i \leq \tilde{n}$;

(ii) χ is well-behaved at $\sum_{j=1}^{\tilde{n}} |\alpha_j|$ in the sense of assumption A.10.

Proposition A.13. *Let χ be well-behaved at $(\alpha, \tilde{\tau}) \in \mathbb{N}^{nx\tilde{n}} \times \mathbb{R}_+^{\tilde{n}}$ in the sense of assumption A.12 for $\alpha = [\alpha_S; \alpha_Z]$. Then for all $t \in \mathbb{R}_+$, we have*

$$\begin{aligned}\Phi^{\mathbb{M},[\alpha]}(0; \tilde{\tau}, \infty) &= \mathrm{E}^{\mathbb{M}}[(X_{t+\tau\oplus\tilde{\tau}})^{\alpha}] \\ &= \sum_{|\beta|\leq|\alpha|} b_{\Phi,\beta}^{(\alpha)}(\tilde{\tau}, \infty)\, (Z_t)^{\beta}\end{aligned} \tag{A.35}$$

with coefficients $b_{\Phi,\beta}^{(\alpha)}(\tilde{\tau}, \infty) \in \mathbb{R}$ given by

$$b_{\Phi,\beta}^{(\alpha)}(\tilde{\tau}, \infty) = \sum_{\eta} b_{\Phi,\eta,(0)}^{(\alpha)}(\tilde{\tau})\, b_{\Psi,[0;\beta]}^{([0;\eta])}(\infty)\, , \tag{A.36}$$

depending on $b_{\Psi,\beta}^{(\eta)}$ as in proposition A.11. Defining $\Delta_i = \tilde{\tau}_i - \tilde{\tau}_{i-1}$, the auxiliary coefficients $b_{\Phi,\eta,(i)}^{(\alpha)}(\tilde{\tau}) \in \mathbb{R}$ are determined by the backward recursion (A.34), depending on $b_{\Psi,\beta}^{(\eta)}$ as in proposition A.10.

Proof. Under the imposed assumptions, proposition A.11 implies that $\Phi^{\mathbb{M},[\alpha]}(0; \tilde{\tau}, \infty)$ can be determined as the limit when $\tau \to \infty$ in equation (A.32), leading to equation (A.35). Letting $\tau \to \infty$ in equation (A.33) thereby yields the associated coefficients in equation (A.36), where $b_{\Phi,\eta,(i)}^{(\alpha)}(\tilde{\tau})$ are given by the recursion in equation (A.34). □

B Supplement to transform-based derivatives pricing

This appendix contains the proofs for the results in section 3.

B.1 Auxiliary results

To derive the transform-based pricing formulas for the equity and volatility derivatives in sections 3.3 and 3.4, respectively, we establish a general result. Specifically, lemma B.1 considers a general payoff function $\mathfrak{g}(y; a, b) = y^a\, \mathfrak{U}(y - b)$, where $\mathfrak{U}$ is the (Heaviside) unit step function, and determines its distributional Fourier transform $\hat{\mathfrak{g}}(y; a, b)$. Expressions are stated in terms of a tempered distribution $(\mathrm{i}y)^{-(1+a)}$, which will be characterized by its integral representation in lemma B.2 below. Certain special cases of these general results in lemmas B.1 and B.2 will be relevant for equity options ($a = 0$) and volatility options ($a = 0$ and $a = 1/2$).

Lemma B.1. *Let $\mathfrak{g}(y; a, b) = y^a\, \mathfrak{U}(y - b)$ with $a \geq 0$ and $b \in \mathbb{R}$ such that $b \geq 0$ if $a \notin \mathbb{N}$. Then the associated distributional Fourier transform $\hat{\mathfrak{g}}(y; a, b)$ is given by*

$$\begin{aligned} \hat{\mathfrak{g}}(y; a, b) &= \frac{\Gamma(1 + a, \mathrm{i}by)}{\Gamma(1 + a)}\, \hat{\mathfrak{g}}(y; a, 0) \\ &= \frac{a\, \Gamma(a, \mathrm{i}by)}{\Gamma(1 + a)}\, \hat{\mathfrak{g}}(y; a, 0) + b^a\, \hat{\mathfrak{g}}(y; 0, b) \end{aligned} \tag{B.1}$$

with $\hat{\mathfrak{g}}(y; 0, b) = \Gamma(1, \mathrm{i}by)\, \hat{\mathfrak{g}}(y; 0, 0)$ and

$$\hat{\mathfrak{g}}(y; a, 0) = \Gamma(1 + a)\, (0 + \mathrm{i}y)^{-(1+a)} = \begin{cases} \Gamma(1 + a)\, (\mathrm{i}y)^{-(1+a)} & , a \notin \mathbb{N} \\ \mathrm{i}^a \pi\, \delta^{(a)}(y) + a!\, (\mathrm{i}y)^{-(1+a)} & , a \in \mathbb{N}\, . \end{cases} \tag{B.2}$$

Here, $\Gamma(a, z) = \int_z^\infty \xi^{a-1} \exp(-\xi)\, \mathrm{d}\xi$ denotes the upper incomplete Gamma function and $(0 + \mathrm{i}y)^{-(1+a)}$ as well as $(\mathrm{i}y)^{-(1+a)}$ denote tempered distributions that are further characterized in lemma B.2 below.

Proof. Following Gel'fand and Shilov (1964), define $\mathfrak{g}_\epsilon(y) = \exp(-\epsilon y)\, \mathfrak{g}(y)$ for $\epsilon > 0$. By direct integration in either of the cases considered in the lemma, it holds by the properties of the incomplete Gamma function that

$$\begin{aligned} \hat{\mathfrak{g}}_\epsilon(y; a, b) &= \int_b^\infty \tilde{y}^a\, \exp(-\mathrm{i}\tilde{y}y - \epsilon\tilde{y})\, \mathrm{d}\tilde{y} \\ &= \int_{b\epsilon + \mathrm{i}by}^\infty \tilde{y}^a\, \exp(-\tilde{y})\, \mathrm{d}\tilde{y}\, \, (\epsilon + \mathrm{i}y)^{-(1+a)} \\ &= \Gamma(1 + a, b\epsilon + \mathrm{i}by)\, (\epsilon + \mathrm{i}y)^{-(1+a)} \\ &= a\, \Gamma(a, b\epsilon + \mathrm{i}by)\, (\epsilon + \mathrm{i}y)^{-(1+a)} + b^a\, \Gamma(1, b\epsilon + \mathrm{i}by)\, (\epsilon + \mathrm{i}y)^{-1}\, . \end{aligned} \tag{B.3}$$

Taking the limit in equation (B.3) such that $\hat{\mathfrak{g}}(y; a, b) = \lim_{\epsilon \downarrow 0} \hat{\mathfrak{g}}_\epsilon(y; a, b)$ in the sense of distributions, we have

$$\begin{aligned} \hat{\mathfrak{g}}(y; a, b) &= \Gamma(1 + a, \mathrm{i}by)\, (0 + \mathrm{i}y)^{-(1+a)} \\ &= a\, \Gamma(a, \mathrm{i}by)\, (0 + \mathrm{i}y)^{-(1+a)} + b^a\, \Gamma(1, \mathrm{i}by)\, (0 + \mathrm{i}y)^{-1} \end{aligned} \tag{B.4}$$

in terms of the tempered distribution $(0 + \mathrm{i}y)^{-(1+a)} = \lim_{\epsilon \downarrow 0} (\epsilon + \mathrm{i}y)^{-(1+a)}$, as established in lemma B.2. Distinguishing the two cases in equation (B.5) for $a \notin \mathbb{N}$ and $a \in \mathbb{N}$ yields equations (B.1) and (B.2) and completes the proof. □

Lemma B.2. *For $a \geq 1$, the tempered distribution $(0+\mathrm{i}y)^{-a} = \lim_{\epsilon\downarrow 0}(\epsilon+\mathrm{i}y)^{-a} \in \mathcal{S}^*(\mathbb{R})$ is the limit as $\epsilon \downarrow 0$ of the regular tempered distribution $(\epsilon+\mathrm{i}y)^{-a} \in \mathcal{S}^*(\mathbb{R})$, acting as $\langle(\epsilon+\mathrm{i}y)^{-a}, f(y)\rangle = \int_{\mathbb{R}}(\epsilon+\mathrm{i}y)^{-a} f(y)\,\mathrm{d}y$. Specifically, it holds that*

$$(0+\mathrm{i}y)^{-a} = \begin{cases} (\mathrm{i}y)^{-a} & , a \notin \mathbb{N} \\ \dfrac{\mathrm{i}^{a-1}\pi}{(a-1)!}\,\delta^{(a-1)}(y) + (\mathrm{i}y)^{-a} & , a \in \mathbb{N} \end{cases} \tag{B.5}$$

with integral representation

$$\lim_{\epsilon\downarrow 0}\int_{-\infty}^{+\infty}\frac{f(y)}{(\epsilon+\mathrm{i}y)^a}\,\mathrm{d}y = \begin{cases} \displaystyle\int_0^\infty \frac{\Delta_y^{(a)} f(y)}{(\mathrm{i}y)^a}\,\mathrm{d}y & , a \notin \mathbb{N} \\ \dfrac{(-\mathrm{i})^{a-1}\pi}{(a-1)!}\, f^{(a-1)}(0) + \displaystyle\int_0^\infty \frac{\Delta_y^{(a)} f(y)}{(\mathrm{i}y)^a}\,\mathrm{d}y & , a \in \mathbb{N}\,. \end{cases} \tag{B.6}$$

Here, the tempered distribution $(\mathrm{i}y)^{-a} \in \mathcal{S}^*(\mathbb{R})$ can be represented in integral form as

$$\langle(\mathrm{i}y)^{-a}, f(y)\rangle = \int_0^\infty \frac{\Delta_y^{(a)} f(y)}{(\mathrm{i}y)^a}\,\mathrm{d}y \tag{B.7}$$

with regularization (Taylor residual)

$$\Delta_y^{(a)} f(y) = f(y) + (-1)^a f(-y) - \sum_{k=0}^{\lfloor a\rfloor -1}\frac{1+(-1)^{a+k}}{k!}\, f^{(k)}(0)\, y^k\,. \tag{B.8}$$

Proof. We start by defining the Taylor residual

$$\tilde{f}_a(y) = f(y-\mathrm{i}\epsilon) - \sum_{k=0}^{\lfloor a-\rfloor -1}\frac{1}{k!}\, f^{(k)}(0)\, y^k\,, \tag{B.9}$$

where $\lfloor a-\rfloor = \lim_{\epsilon'\downarrow 0}\lfloor a-\epsilon'\rfloor$ denotes the strict floor (i.e., the largest integer smaller than a). Moreover, note that $\int_{-\infty}^{+\infty}(\epsilon+\mathrm{i}y)^{-a}\,\mathrm{d}y = 0$ for $\epsilon > 0$ and $a > 1$. Hence, using the binomial formula, it also holds that $\int_{-\infty}^{+\infty}(\epsilon+\mathrm{i}y)^{-a}y^k\,\mathrm{d}y = 0$ for $\epsilon > 0$ and $a > k+1$. Adding and subtracting the Taylor polynomial in equation (B.9) thus yields the identity

$$\int_{-\infty}^{+\infty}\frac{f(y)}{(\epsilon+\mathrm{i}y)^a}\,\mathrm{d}y = \int_{-\infty}^{+\infty}\frac{\tilde{f}_a(y)}{(\epsilon+\mathrm{i}y)^a}\,\mathrm{d}y\,. \tag{B.10}$$

Taking into account that $f = \tilde{f}_a$ for $a = 1$, equation (B.10) is valid for all $\epsilon > 0$ and $a \geq 1$. Splitting the integration domain on the right-hand-side of equation (B.10), we may write

$$\begin{aligned}\int_{-\infty}^{+\infty}\frac{f(y)}{(\epsilon+\mathrm{i}y)^a}\,\mathrm{d}y &= \int_{-\rho}^{+\rho}\frac{\tilde{f}_a(y)}{(\epsilon+\mathrm{i}y)^a}\,\mathrm{d}y + \int_{-\infty}^{-\rho}\frac{\tilde{f}_a(y)}{(\epsilon+\mathrm{i}y)^a}\,\mathrm{d}y + \int_{+\rho}^{+\infty}\frac{\tilde{f}_a(y)}{(\epsilon+\mathrm{i}y)^a}\,\mathrm{d}y \\ &= \int_{-\rho}^{+\rho}\frac{\tilde{f}_a(y)}{(\epsilon+\mathrm{i}y)^a}\,\mathrm{d}y + \int_{+\rho}^{+\infty}\frac{\tilde{f}_a(y)+(-1)^a\tilde{f}_a(-y)}{(\epsilon+\mathrm{i}y)^a}\,\mathrm{d}y\end{aligned} \tag{B.11}$$

for any $\rho > 0$, where the last equality follows by a change of variable and combining two integrals. Since terms of order $a-1$ cancel when $a \in \mathbb{N}$, we have $\Delta_y^{(a)} f(y) = \tilde{f}_a(y) + (-1)^a\tilde{f}_a(-y)$ for all $a \geq 1$.

For the remainder of the proof, we distinguish the cases $a \notin \mathbb{N}$ and $a \in \mathbb{N}$. In each of the cases, our main effort will be dedicated to determining the limit of the first integral in the terminal expression of equation (B.11), whereas the limit of the second integral may be determined rather straightforwardly.

Beginning with the easier case $a \notin \mathbb{N}$, define $Q_a(\rho, \epsilon)$ by

$$Q_a(\rho, \epsilon) = \int_{-\rho}^{+\rho} \frac{\tilde{f}_a(y)}{(\epsilon + \mathrm{i}y)^a} \,\mathrm{d}y = \int_{-\rho}^{+\rho} \frac{f^{(\lfloor a \rfloor)}(\xi(y))\, y^{\lfloor a \rfloor}}{(\epsilon + \mathrm{i}y)^a} \,\mathrm{d}y \,, \tag{B.12}$$

where the second equality holds for some mean value $\xi(y)$ between 0 and y by Taylor's theorem. Therefore, we obtain the bounds

$$|Q_a(\rho, \epsilon)| \le \int_{-\rho}^{+\rho} \left| \frac{f^{(\lfloor a \rfloor)}(\xi(y))\, y^{\lfloor a \rfloor}}{(\epsilon + \mathrm{i}y)^a} \right| \mathrm{d}y \le \frac{2\rho^{1-(a-\lfloor a \rfloor)}}{1-(a-\lfloor a \rfloor)} \sup_{|y|\le\rho} |f^{(\lfloor a \rfloor)}(y)|$$

since derivatives of f are bounded as a Schwartz function. It immediately follows for Q_a in equation (B.12) that $Q_a(\rho, \epsilon) \to 0$ as $\rho, \epsilon \downarrow 0$. Taking the limit in equation (B.11) thus yields

$$\lim_{\rho,\epsilon\downarrow 0} \int_{-\infty}^{+\infty} \frac{f(y)}{(\epsilon + \mathrm{i}y)^a} \,\mathrm{d}y = \int_{0}^{+\infty} \frac{\tilde{f}_a(y) + (-1)^a \tilde{f}_a(-y)}{(\mathrm{i}y)^a} \,\mathrm{d}y \,, \tag{B.13}$$

where the right-hand-side integral converges as the Taylor residual around zero is of order $\lfloor a \rfloor$. In fact, we may even write the right-hand-side of equation (B.13) as a sum of two individually convergent integrals. With the definitions in equations (B.7) and (B.8), equation (B.13) yields the non-integer case in equations (B.5) and (B.6).

Turning to the case $a \in \mathbb{N}$, define the integral

$$J_a(\rho, \epsilon) = \int_{-\rho}^{+\rho} \frac{y^{a-1}}{(\epsilon + \mathrm{i}y)^a} \,\mathrm{d}y = -2\mathrm{i}^{a+1} \Im B(\mathrm{i}\rho/\epsilon, a, 1-a) \tag{B.14}$$

in terms of the incomplete Beta function B. It holds that $J_a(\rho, \epsilon) \to (-\mathrm{i})^{a-1}\pi$ as $\rho, \epsilon \to 0$ with $\rho/\epsilon \to \infty$. For integer a, we now define $Q_a(\rho, \epsilon)$ by

$$\begin{aligned} Q_a(\rho, \epsilon) &= \int_{-\rho}^{+\rho} \frac{\tilde{f}_a(y)}{(\epsilon + \mathrm{i}y)^a} \,\mathrm{d}y - \frac{(-\mathrm{i})^{a-1}\pi}{(a-1)!} f^{(a-1)}(0) \\ &= \int_{-\rho}^{+\rho} \frac{\tilde{f}_a(y) - \frac{1}{(a-1)!} f^{(a-1)}(0)\, y^{a-1}}{(\epsilon + \mathrm{i}y)^a} \,\mathrm{d}y + \frac{J_a(\rho, \epsilon) - (-\mathrm{i})^{a-1}\pi}{(a-1)!} f^{(a-1)}(0) \\ &= \int_{-\rho}^{+\rho} \frac{f^{(a)}(\xi(y))\, y^a}{(\epsilon + \mathrm{i}y)^a} \,\mathrm{d}y + \frac{J_a(\rho, \epsilon) - (-\mathrm{i})^{a-1}\pi}{(a-1)!} f^{(a-1)}(0) \,, \end{aligned} \tag{B.15}$$

where the second equality uses the definition of J_a in equation (B.14) and the third equality invokes Taylor's theorem for some mean value $\xi(y)$ between 0 and y. We thus obtain the bounds

$$\begin{aligned} |Q_a(\rho, \epsilon)| &\le \int_{-\rho}^{+\rho} \left| \frac{f^{(a)}(\xi(y))\, y^a}{(\epsilon + \mathrm{i}y)^a} \right| \mathrm{d}y + \frac{|J_a(\rho, \epsilon) - (-\mathrm{i})^{a-1}\pi|}{(a-1)!} |f^{(a-1)}(0)| \\ &\le 2\rho \sup_{|y|\le\rho} |f^{(a)}(y)| + \frac{|J_a(\rho, \epsilon) - (-\mathrm{i})^{a-1}\pi|}{(a-1)!} |f^{(a-1)}(0)| \end{aligned}$$

due to the boundedness of derivatives of f as a Schwartz function. Accounting for the limiting behavior of J_a, Q_a in equation (B.15) satisfies $Q_a(\rho, \epsilon) \to 0$ as $\rho, \epsilon \downarrow 0$ with $\rho/\epsilon \to \infty$. Taking appropriate limits in equation (B.11) thus yields

$$\lim_{\rho,\epsilon\downarrow 0} \int_{-\infty}^{+\infty} \frac{f(y)}{(\epsilon + \mathrm{i}y)^a} \,\mathrm{d}y = \frac{\mathrm{i}^{a-1}\pi}{(a-1)!} \langle \delta^{(a-1)}(y), f(y) \rangle + \int_{0}^{+\infty} \frac{\tilde{f}_a(y) + (-1)^a \tilde{f}_a(-y)}{(\mathrm{i}y)^a} \,\mathrm{d}y \,. \tag{B.16}$$

Here, it should be noted that after combining the integrals, the Taylor terms of order $a-1$ cancel, so that the residual around zero is in fact of order a. Hence, the integral on the right-hand-side of equation (B.16) converges in the sense of Cauchy principal value integral. Using the definition of the Dirac delta distribution and distributional derivatives together with the definitions in equations (B.7) and (B.8), equation (B.16) justifies the integer case in equations (B.5) and (B.6). $\square$

B.2 General derivatives

From the generalized transform analysis of Chen and Joslin (2012), we have the following result.

Proposition B.1. *Let $g \in \mathcal{S}^*(\mathbb{R})$ and $(y \mapsto \Pi(\omega + \mathrm{i}y\hat{\omega}; \tilde{T}, z)) \in \mathcal{S}(\mathbb{R})$ for all $z \in \mathcal{Z}$. Then*

$$\begin{aligned}\Pi_g(\omega, \hat{\omega}; \tilde{T}, Z_t) &= \mathrm{E}^{\mathbb{Q}}[D_t(\tilde{T}_{\tilde{m}}) \exp(\omega \cdot X_{t\oplus\tilde{T}})\, g(\hat{\omega} \cdot X_{t\oplus\tilde{T}}) \,|\, \mathcal{F}_t] \\ &= \frac{1}{2\pi} \langle \hat{g}(y), \Pi(\omega + \mathrm{i}y\hat{\omega}; \tilde{T}, Z_t)\rangle\end{aligned}$$

in terms of the distributional Fourier transform $\hat{g} \in \mathcal{S}^*(\mathbb{R})$.

Proof. See Chen and Joslin (2012) or Dillschneider (2020). $\square$

Proof of proposition 3.1. Use the definition of the pricing transform Π in equation (3.3). By the imposed assumptions, proposition B.1 yields

$$\mathcal{V}(Z_t; K, \tilde{T}) = \frac{1}{2\pi} \sum_{i=1}^{n_h} \langle \hat{g}_i(\tilde{y}; K), \Pi(\mathfrak{b}([\bar{\omega}_i; \tilde{y}]); \tilde{T}, Z_t)\rangle \,. \tag{B.17}$$

Since $\mathcal{Y} \subset \mathbb{R}^{n_X+1}$, there exists some $\tilde{\Pi}$ with $(y \mapsto \tilde{\Pi}(\mathfrak{b}(y); \tilde{T}, z)) \in \mathcal{S}(\mathbb{R}^{n_X+1})$ that coincides with Π on $\mathcal{Y}$. As the support of each g_i is contained in $\mathcal{Y}$, we can rewrite equation (B.17) as

$$\begin{aligned}\mathcal{V}(Z_t; K, \tilde{T}) &= \frac{1}{2\pi} \sum_{i=1}^{n_h} \langle \hat{g}_i(\tilde{y}; K), \tilde{\Pi}(\mathfrak{b}([\bar{\omega}_i; \tilde{y}]); \tilde{T}, Z_t)\rangle \\ &= \frac{1}{2\pi} \sum_{i=1}^{n_h} \langle \delta(\tilde{\omega} - \bar{\omega}_i) \otimes \hat{g}_i(\tilde{y}; K), \tilde{\Pi}(\mathfrak{b}([\tilde{\omega}; \tilde{y}]); \tilde{T}, Z_t)\rangle \\ &= \langle w(y; K), \tilde{\Pi}(\mathfrak{b}(y); \tilde{T}, Z_t)\rangle \,.\end{aligned} \tag{B.18}$$

By construction, the support of $(y \mapsto w(y; K)) \in \mathcal{S}^*(\mathbb{R}^{n_X+1})$ is contained in $\mathcal{Y}$, so that equation (B.18) thus yields equation (3.4). $\square$

B.3 Equity derivatives

Proof of corollary 3.1. For the call payoff in equation (3.6a), define $(\tilde{y} \mapsto g^C_{\mathrm{stock}}(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ by $g^C_{\mathrm{stock}}(\tilde{y}; K) = \mathfrak{g}(\tilde{y}; 0, K)$ for $\mathfrak{g}$ as in lemma B.1. Combining equations (B.1) and (B.2), the associated distributional Fourier transform $(\tilde{y} \mapsto \hat{g}^C_{\mathrm{stock}}(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ is given by

$$\begin{aligned}\hat{g}^C_{\mathrm{stock}}(\tilde{y}; K) &= \hat{\mathfrak{g}}(\tilde{y}; 0, K) \\ &= \pi\, \delta(\tilde{y}) + \exp(-\mathrm{i}K\tilde{y})\, (\mathrm{i}\tilde{y})^{-1} \,,\end{aligned}$$

in terms of Dirac delta distribution $\delta(\tilde{y})$ and the tempered distribution $(\mathrm{i}\tilde{y})^{-1}$ for $\Gamma(1, \mathrm{i}Ky) = \exp(-\mathrm{i}Ky)$. Using $\bar{\omega}_1 = [1; 0]$, $\bar{\omega}_2 = [0; 0]$, $\hat{\omega} = [1; 0]$, $g_1(\tilde{y}; K) = g^C_{\mathrm{stock}}(\tilde{y}; K)$, and $g_2(\tilde{y}; K) = -\exp(K)\, g^C_{\mathrm{stock}}(\tilde{y}; K)$, exploiting the linearity of the Fourier transform, then yields equations (3.7) and (3.8a) as a special case of proposition 3.1.

Proceeding analogously for the put payoff in equation (3.6b), define $(\tilde{y} \mapsto g^P_{\text{stock}}(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ by $g^P_{\text{stock}}(\tilde{y}; K) = -\mathfrak{g}(-\tilde{y}; 0, -K)$ for $\mathfrak{g}$ as in lemma B.1. Due to the scaling property[19] of the Fourier transform, equations (B.1) and (B.2) yield the associated distributional Fourier transform $(\tilde{y} \mapsto \hat{g}^P_{\text{stock}}(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ as

$$\begin{aligned} \hat{g}^P_{\text{stock}}(\tilde{y}; K) &= - \hat{\mathfrak{g}}(-\tilde{y}; 0, -K) \\ &= - \pi\, \delta(\tilde{y}) + \exp(-\mathrm{i}K\tilde{y})\, (\mathrm{i}\tilde{y})^{-1} , \end{aligned}$$

again in terms of the tempered distributions $\delta(\tilde{y})$ and $(\mathrm{i}\tilde{y})^{-1}$. Using $\bar{\omega}_1 = [1; 0]$, $\bar{\omega}_2 = [0; 0]$, $\hat{\omega} = [1; 0]$, $g_1(\tilde{y}; K) = g^P_{\text{stock}}(\tilde{y}; K)$, and $g_2(\tilde{y}; K) = -\exp(K)\, g^P_{\text{stock}}(\tilde{y}; K)$, again exploiting the linearity of the Fourier transform, then yields equations (3.7) and (3.8b) as a special case of proposition 3.1. □

Proof of lemma 3.1. As in lemma B.2, we define the tempered distribution $(\mathrm{i}\tilde{y})^{-1} \in \mathcal{S}^*(\mathbb{R})$ for every Schwartz function $f \in \mathcal{S}(\mathbb{R})$ by the convergent integral

$$\langle (\mathrm{i}\tilde{y})^{-1}, f(\tilde{y}) \rangle = \int_{\mathbb{R}_+} \frac{\Delta^{(1)}_{\tilde{y}} f(\tilde{y})}{\mathrm{i}\tilde{y}}\, \mathrm{d}\tilde{y} , \tag{B.19}$$

using the regularization $\Delta^{(1)}_{\tilde{y}} f(\tilde{y}) = f(\tilde{y}) - f(-\tilde{y})$. To avoid redundancies, write $(y \mapsto w^O_{\text{stock}}(y; K)) \in \mathcal{S}^*(\mathcal{Y})$ in equation (3.8) compactly as

$$w^O_{\text{stock}}([\tilde{\omega}; \tilde{y}]; K) = (\delta(\tilde{\omega} - [1; 0]) - \exp(K)\, \delta(\tilde{\omega})) \otimes \tilde{g}^O_{\text{stock}}(\tilde{y}; K) . \tag{B.20}$$

Define the tempered distribution $(\tilde{y} \mapsto \tilde{g}^O_{\text{stock}}(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ in equation (B.20) by

$$\tilde{g}^O_{\text{stock}}(\tilde{y}; K) = \frac{c^O}{2}\, \delta(\tilde{y}) + F_{\text{stock}}(\tilde{y}; K)\, (\mathrm{i}\tilde{y})^{-1} ,$$

with equation (B.19) implying the integral representation

$$\langle \tilde{g}^O_{\text{stock}}(\tilde{y}; K), \Upsilon(\mathfrak{b}_{\text{stock}}([\tilde{\omega}; \tilde{y}])) \rangle = \frac{c^O}{2} \Upsilon(\tilde{\omega}) + \int_{\mathbb{R}_+} \frac{\Delta^{(1)}_{\tilde{y}} (F_{\text{stock}}(\tilde{y}; K)\, \Upsilon(\tilde{\omega} + \mathrm{i}\tilde{y}[1; 0]))}{\mathrm{i}\tilde{y}}\, \mathrm{d}\tilde{y} .$$

An application of the definition of the distributional tensor product in equation (3.1) to $\hat{w}^O_{\text{stock}}$ in terms of $\tilde{g}^O_{\text{stock}}$ in equation (B.20) then yields equation (3.9), as required. □

B.4 Volatility derivatives

Proof of lemma 3.2. Define by $[\log S]^*_t$ the jump-adjusted quadratic variation of the stock price, having dynamics

$$\mathrm{d}[\log S]^*_t = \Omega_S(Z_{t-})\, \mathrm{d}t + 2\, (\exp(J_{S,t}) - \iota^\top - J_{S,t})\, \mathrm{d}N_t .$$

Here, $\Omega_S(z) = A_{\Omega,S} + B_{\Omega,S} z$ is the instantaneous diffusive variance of the stock price and $\exp(J_{S,t})$ denotes the elementwise exponential of the stock jump sizes $J_{S,t}$. Under constant interest rates and dividend yields, the jump-adjusted quadratic variations of the stock price and of its forward price coincide. Utilizing the reasoning of Carr and Wu (2009), we therefore get the identity

$$VIX^2_t = \frac{1}{\tau_{\text{vix}}} \mathrm{E}^{\mathbb{Q}}[[\log S]^*_{t+\tau_{\text{vix}}} - [\log S]^*_t \,|\, \mathcal{F}_t] \tag{B.21}$$

[19]By the scaling property, $g(\tilde{y}) = f(a\tilde{y})$ has Fourier transform $\hat{g}(\tilde{y}) = \hat{f}(\tilde{y}/a)/|a|$.

for VIX_t^2 defined in equation (3.10).

It remains to establish an expression for the expected value on the right-hand side of equation (B.21). For this, we obtain

$$\mathrm{E}^{\mathbb{Q}}[[\log S]^*_{t+\tau_{\text{vix}}} - [\log S]^*_t \,|\, \mathcal{F}_t] = \int_0^{\tau_{\text{vix}}} C_{[\log S]^*} + D_{[\log S]^*} \mathrm{E}^{\mathbb{Q}}[Z_{t+\tau} \,|\, \mathcal{F}_t] \, \mathrm{d}\tau \,, \tag{B.22}$$

where the coefficients $C_{[\log S]^*} \in \mathbb{R}$ and $D_{[\log S]^*} \in \mathbb{R}^{1\times nz}$ are given by

$$C_{[\log S]^*} = A_{\Omega,S} + 2\,\mathrm{E}^{\mathbb{Q}}[\exp(J_{S,t}) - \iota^\top - J_{S,t}]\, A_\lambda^{\mathbb{Q}} \tag{B.23a}$$

$$D_{[\log S]^*} = B_{\Omega,S} + 2\,\mathrm{E}^{\mathbb{Q}}[\exp(J_{S,t}) - \iota^\top - J_{S,t}]\, B_\lambda^{\mathbb{Q}} \,. \tag{B.23b}$$

We moreover have

$$\mathrm{E}^{\mathbb{Q}}[Z_{t+\tau} \,|\, \mathcal{F}_t] = (\exp(\tau D_Z) - I)\, D_Z^{-1}\, C_Z + \exp(\tau D_Z)\, Z_t \tag{B.24}$$

with $C_Z \in \mathbb{R}^{nz}$ and $D_Z \in \mathbb{R}^{nz\times nz}$ determined as

$$C_Z = A_{\mu,Z}^{\mathbb{Q}} + \mathrm{E}^{\mathbb{Q}}[J_{Z,t}]\, A_\lambda^{\mathbb{Q}} \tag{B.25a}$$

$$D_Z = B_{\mu,Z}^{\mathbb{Q}} + \mathrm{E}^{\mathbb{Q}}[J_{Z,t}]\, B_\lambda^{\mathbb{Q}} \,. \tag{B.25b}$$

Substituting equation (B.24) into equation (B.22) and performing the integration yields the coefficients $a_{\text{vix}} \in \mathbb{R}$ and $b_{\text{vix}} \in \mathbb{R}^{nz}$ in equation (3.11) by

$$a_{\text{vix}} = C_{[\log S]^*} + \frac{1}{\tau_{\text{vix}}} D_{[\log S]^*}\, (\exp(\tau_{\text{vix}} D_Z) - I - \tau_{\text{vix}} D_Z) D_Z^{-2} C_Z \tag{B.26a}$$

$$b_{\text{vix}} = \frac{1}{\tau_{\text{vix}}} (D_{[\log S]^*}\, (\exp(\tau_{\text{vix}} D_Z) - I)\, D_Z^{-1})^\top \,, \tag{B.26b}$$

depending on the coefficients $C_{[\log S]^*}$ and $D_{[\log S]^*}$ in equation (B.23) as well as C_Z and D_Z in equation (B.25). □

Proof of corollary 3.2. For the call payoff in equation (3.12a), define $(\tilde{y} \mapsto g_{\text{vix}}^C(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ by $g_{\text{vix}}^C(\tilde{y}; K) = \mathfrak{g}(\tilde{y}; 1/2, K) - K^{1/2}\mathfrak{g}(\tilde{y}; 0, K)$ for $\mathfrak{g}$ as in lemma B.1. Combining equations (B.1) and (B.2), we obtain the associated distributional Fourier transform $(\tilde{y} \mapsto \hat{g}_{\text{vix}}^C(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ as

$$\begin{aligned} \hat{g}_{\text{vix}}^C(\tilde{y}; K) &= \hat{\mathfrak{g}}(\tilde{y}; 1/2, K) - K^{1/2}\hat{\mathfrak{g}}(\tilde{y}; 0, K) \\ &= \frac{\frac{1}{2}\Gamma(1/2, \mathrm{i}K\tilde{y})}{\Gamma(3/2)}\, \hat{\mathfrak{g}}(\tilde{y}; 1/2, 0) \\ &= \tfrac{1}{2}\Gamma(1/2, \mathrm{i}K\tilde{y})\, (\mathrm{i}\tilde{y})^{-3/2} \,, \end{aligned}$$

in terms of the tempered distribution $(\mathrm{i}\tilde{y})^{-3/2}$, where Γ denotes the upper incomplete Gamma function. Setting $\bar{\omega}_1 = [0; 0]$, $\hat{\omega} = [0; b_{\text{vix}}]$, and $g_1(\tilde{y}; K) = g_{\text{vix}}^C(\tilde{y}; K)$, exploiting the shift property[20] of the Fourier transform, this yields equations (3.13) and (3.14a) as a special case of proposition 3.1.

Proceeding analogously for the put payoff in equation (3.12b), define $(\tilde{y} \mapsto g_{\text{vix}}^P(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ by $g_{\text{vix}}^P(\tilde{y}; K) = K^{1/2}\big(\mathfrak{g}(\tilde{y}; 0, 0) - \mathfrak{g}(\tilde{y}; 0, K)\big) - \big(\mathfrak{g}(\tilde{y}; 1/2, 0) - \mathfrak{g}(\tilde{y}; 1/2, K)\big)$ for $\mathfrak{g}$ as in lemma B.1. Due to the scaling property of the Fourier transform, equations (B.1) and (B.2) yield the associated distributional

[20] By the shift property, $g(\tilde{y}) = f(a + \tilde{y})$ has Fourier transform $\hat{g}(\tilde{y}) = \exp(\mathrm{i}a\tilde{y})\, \hat{f}(\tilde{y})$.

Fourier transform $(\tilde{y} \mapsto \hat{g}^P_{\mathrm{vix}}(\tilde{y};K)) \in \mathcal{S}^*(\mathbb{R})$ as

$$\begin{aligned}
\hat{g}^P_{\mathrm{vix}}(\tilde{y};K) &= K^{1/2}\big(\hat{\mathfrak{g}}(\tilde{y};0,0) - \hat{\mathfrak{g}}(\tilde{y};0,K)\big) - \big(\hat{\mathfrak{g}}(\tilde{y};1/2,0) - \hat{\mathfrak{g}}(\tilde{y};1/2,K)\big) \\
&= K^{1/2}\hat{\mathfrak{g}}(\tilde{y};0,0) - \frac{\frac{1}{2}\gamma(1/2,\mathrm{i}Ky)}{\Gamma(3/2)}\,\hat{\mathfrak{g}}(\tilde{y};1/2,0) \\
&= 2\pi\, K^{1/2}\delta(\tilde{y}) - \tfrac{1}{2}\gamma(1/2,\mathrm{i}Ky)\,(\mathrm{i}y)^{-3/2}
\end{aligned}$$

in terms of the Dirac delta distributiona and again the tempered distribution $(\mathrm{i}\tilde{y})^{-3/2}$, where γ denotes the lower incomplete Gamma function. The last equality follows by realizing that $\hat{\mathfrak{g}}(\tilde{y};0,0)$ is the distributional Fourier transform of $\mathfrak{g}(\tilde{y};0,0) = \mathfrak{U}(\tilde{y})$, which is redundant when positivity is assured. Setting $\bar{\omega}_1 = [0;0]$, $\hat{\omega} = [0;b_{\mathrm{vix}}]$, and $g_1(\tilde{y};K) = g^P_{\mathrm{vix}}(\tilde{y};K)$, again exploiting the shift property of the Fourier transform, this yields equations (3.13) and (3.14b) as a special case of proposition 3.1. □

Proof of lemma 3.3. As in lemma B.2, we define the tempered distribution $(\mathrm{i}\tilde{y})^{-3/2} \in \mathcal{S}^*(\mathbb{R})$ for every Schwartz function $f \in \mathcal{S}(\mathbb{R})$ by the convergent integral

$$\langle (\mathrm{i}\tilde{y})^{-3/2}, f(\tilde{y})\rangle = \int_{\mathbb{R}_+} \frac{\Delta^{(3/2)}_{\tilde{y}} f(\tilde{y})}{(\mathrm{i}\tilde{y})^{3/2}}\,\mathrm{d}\tilde{y}\,, \tag{B.27}$$

using the regularization $\Delta^{(3/2)}_{\tilde{y}} f(\tilde{y}) = f(\tilde{y}) - \mathrm{i}\, f(-\tilde{y}) - (1-\mathrm{i})\, f(0)$. For $(y \mapsto w^O_{\mathrm{vix}}(y;K)) \in \mathcal{S}^*(\mathcal{Y})$ in equation (3.14), we have the compact expression

$$w^O_{\mathrm{vix}}([\tilde{\omega};\tilde{y}];K) = \delta(\tilde{\omega}) \otimes \tilde{g}^O_{\mathrm{vix}}(\tilde{y};K)\,. \tag{B.28}$$

The tempered distribution $(\tilde{y} \mapsto \tilde{g}^O_{\mathrm{vix}}(\tilde{y};K)) \in \mathcal{S}^*(\mathbb{R})$ in equation (B.28) is defined as

$$\tilde{g}^O_{\mathrm{vix}}(\tilde{y};K) = \frac{1-c^O}{2}\,K^{1/2}\,\delta(\tilde{y}) + F^O_{\mathrm{vix}}(\tilde{y};K)\,(\mathrm{i}\tilde{y})^{-3/2}\,,$$

by equation (B.27) having the integral representation

$$\langle \tilde{g}^O_{\mathrm{vix}}(\tilde{y};K), \Upsilon(\mathfrak{b}_{\mathrm{vix}}([\tilde{\omega};\tilde{y}]))\rangle = \frac{1-c^O}{2} K^{1/2}\,\Upsilon(\tilde{\omega}) + \int_{\mathbb{R}_+} \frac{\Delta^{(3/2)}_{\tilde{y}}\big(F^O_{\mathrm{vix}}(\tilde{y};K)\,\Upsilon(\tilde{\omega} + \mathrm{i}\tilde{y}[0;b_{\mathrm{vix}}])\big)}{(\mathrm{i}\tilde{y})^{3/2}}\,\mathrm{d}\tilde{y}\,.$$

An application of the definition of the distributional tensor product in equation (3.1) to $\hat{w}^O_{\mathrm{vix}}$ in terms of $\tilde{g}^O_{\mathrm{vix}}$ in equation (B.28) then yields equation (3.15), as required. □

C Extensions of transform-based derivatives pricing

In this appendix, we develop an extension of the transform-based pricing approach of section 3. Instead of relying on the general formulation there using Schwartz distribution theory, we here rely on a functional generalization of Fourier transformation, henceforth referred to as *complex* Fourier transformation. This approach allows us to treat derivatives with certain payoff functions, namely those that grow exponentially (at infinity). For these, the in this appendix yields prices in terms of ordinary integrals, avoiding the integral regularization approach necessary in section 3 (see also appendix B for further details). In particular, we obtain alternative (but equivalent) formulas for equity and volatility derivatives. To maintain the connection to our other results, we still integrate a Schwartz-based formulation, which allows to interpret our previous results as a limiting case. In principle, it may even be possible to drop the Schwartz requirement altogether — at the expense of more tedious regularity conditions. For details, the interested reader is referred to Dillschneider (2020).

C.1 Complex Fourier theory

Classical Fourier theory focuses primarily on the class of square-integrable functions. Following the classical exposition in Titchmarsh (1975), we use a simple approach to extend Fourier transformation to certain functions outside this class. Specifically, we will be concerned with functions g exhibiting exponential growth such that $\exp(-\epsilon|y|)\, g(y)$ is bounded for some $\epsilon = \epsilon^* > 0$ and, hence, square-integrable for each $\epsilon > \epsilon^*$. Without loss of generality, we constrain our attention to one-sided exponential growth in the sense that the regularized function $g_\epsilon(y) = \exp(-\epsilon y)\, g(y)$ is square-integrable. Two-sided exponential growth may by accounted for by appropriately splitting the support of g; e.g., using $g(y) = g(y)\,\mathfrak{U}(y) + g(y)\,\mathfrak{U}(-y)$ and applying the reasoning the both $g(y)\,\mathfrak{U}(y)$ and $g(y)\,\mathfrak{U}(-y)$ separately.

For such square-integrable g_ϵ, the Fourier transform exists in the ordinary sense as a square-integrable function $\hat{g}_\epsilon$. By construction of g_ϵ and definition of Fourier transformation, $\hat{g}_\epsilon$ satisfies the relation

$$\hat{g}_\epsilon(y) = \int_{\mathbb{R}} \exp(-\mathrm{i}y\tilde{y})\, g_\epsilon(\tilde{y})\, \mathrm{d}\tilde{y} = \int_{\mathbb{R}} \exp(-\mathrm{i}(y - \mathrm{i}\epsilon)\tilde{y})\, g(\tilde{y})\, \mathrm{d}\tilde{y} = \hat{g}(y - \mathrm{i}\epsilon)\,.$$

This last equality defines the complex Fourier transform $\hat{g}(y - \mathrm{i}\epsilon)$. It usually exists in a certain interval of ϵ; when it exists for $\epsilon = 0$, the complex Fourier transform of course coincides with the ordinary Fourier transform.

C.2 Auxiliary results

In order to prepare the pricing of equity and volatility derivatives in the present setting, we first derive some auxiliary results. We revisit the function $\mathfrak{g}$ encountered in lemma B.1 and derive its complex Fourier transform $\hat{\mathfrak{g}}(y - \mathrm{i}\epsilon; a, b) = \hat{\mathfrak{g}}_\epsilon(y; a, b)$. In essence, lemma B.1 exploits the expressions in lemma C.1 and uses the limit $\hat{\mathfrak{g}}(y; a, b) = \lim_{\epsilon\downarrow 0} \hat{\mathfrak{g}}_\epsilon(y; a, b)$ in the sense of distributions.

Lemma C.1. *Let $\mathfrak{g}_\epsilon(y; a, b) = \exp(-\epsilon y)\, \mathfrak{g}(y; a, b) = \exp(-\epsilon y)\, y^a\, \mathfrak{U}(y - b)$ for $\epsilon > 0$ with $a \geq 0$ and $b \in \mathbb{R}$ such that $b \geq 0$ if $a \notin \mathbb{N}$. Then the associated distributional Fourier transform $\hat{\mathfrak{g}}(y; a, b)$ is given by*

$$\begin{aligned}\hat{\mathfrak{g}}_\epsilon(y; a, b) &= \frac{\Gamma(1 + a, \epsilon b + \mathrm{i}by)}{\Gamma(1 + a)}\, \hat{\mathfrak{g}}_\epsilon(y; a, 0) \\ &= \frac{a\,\Gamma(a, \epsilon b + \mathrm{i}by)}{\Gamma(1 + a)}\, \hat{\mathfrak{g}}_\epsilon(y; a, 0) + b^a\, \hat{\mathfrak{g}}_\epsilon(y; 0, b)\end{aligned} \tag{C.1}$$

with $\hat{\mathfrak{g}}_\epsilon(y; 0, b) = \Gamma(1, \epsilon b + \mathrm{i}by)\, \hat{\mathfrak{g}}_\epsilon(y; 0, 0)$ and

$$\hat{\mathfrak{g}}_\epsilon(y; a, 0) = \Gamma(1 + a)\, (\epsilon + \mathrm{i}y)^{-(1+a)}\,. \tag{C.2}$$

Here, $\Gamma(a, z) = \int_z^\infty \xi^{a-1} \exp(-\xi)\, \mathrm{d}\xi$ denotes the upper incomplete Gamma function and $(\epsilon + \mathrm{i}y)^{-(1+a)}$ denotes the (regular) tempered distribution acting as $\langle(\epsilon + \mathrm{i}y)^{-(1+a)}, f(y)\rangle = \int_{\mathbb{R}}(\epsilon + \mathrm{i}y)^{-(1+a)} f(y)\, \mathrm{d}y$.

Proof. Equations (C.1) and (C.2) immediately follow by the calculations performed in equation (B.3) for proving lemma B.1. □

Lemma C.2 conveys that the tempered distribution $\hat{\mathfrak{g}}_\epsilon$ is asymptotically equivalent to $\hat{\mathfrak{g}}$, through the convergence of $(\epsilon + \mathrm{i}y)^{-a}$ to $(0 + \mathrm{i}y)^{-a}$. In addition, it can be established that $\hat{\mathfrak{g}}_\epsilon$ is in fact equivalent to $\hat{\mathfrak{g}}$ once the function f can be extended in a certain way to a complex domain. Such extensions will naturally arise when dealing with complex Fourier transforms. The following lemma formally states these equivalence results in terms of the tempered distributions $(\epsilon + \mathrm{i}y)^{-a}$ and $(0 + \mathrm{i}y)^{-a}$.

Lemma C.2. *Let $(y, \epsilon) \mapsto f(y - \mathrm{i}\epsilon)$ exist and be complex differentiable in a neighborhood of $\mathbb{R} \times [0, \epsilon^*]$ for some $\epsilon^* > 0$, such that $(y \mapsto f(y - \mathrm{i}\epsilon)) \in \mathcal{S}(\mathbb{R})$ for $\epsilon \in [0, \epsilon^*]$. For $a \geq 1$, the regular tempered distribution $(\epsilon + \mathrm{i}y)^{-a} \in \mathcal{S}^*(\mathbb{R})$ acting on $f(y - \mathrm{i}\epsilon)$ is equivalent to the tempered distribution $(0 + \mathrm{i}y)^{-a} \in \mathcal{S}^*(\mathbb{R})$ acting on $f(y)$. Specifically, it holds that*

$$\langle (\epsilon + \mathrm{i}y)^{-a}, f(y - \mathrm{i}\epsilon) \rangle = \begin{cases} \langle (\mathrm{i}y)^{-a}, f(y) \rangle & , a \notin \mathbb{N} \\ \mathrm{i}^{a-1}\pi \, \langle \delta^{(a-1)}(y), f(y) \rangle + \langle (\mathrm{i}y)^{-a}, f(y) \rangle & , a \in \mathbb{N} \end{cases} \tag{C.3}$$

with integral representation

$$\int_{-\infty}^{+\infty} \frac{f(y - \mathrm{i}\epsilon)}{(\epsilon + \mathrm{i}y)^a} \, \mathrm{d}y = \begin{cases} \displaystyle\int_0^{+\infty} \frac{\Delta_y^{(a)} f(y)}{(\mathrm{i}y)^a} \, \mathrm{d}y & , a \notin \mathbb{N} \\ \displaystyle(-\mathrm{i})^{a-1}\pi \, f^{(a-1)}(0) + \int_0^{+\infty} \frac{\Delta_y^{(a)} f(y)}{(\mathrm{i}y)^a} \, \mathrm{d}y & , a \in \mathbb{N}\,. \end{cases} \tag{C.4}$$

Proof. Extending the definition before, we now use the (complex) Taylor residual

$$\tilde{f}_a(y - \mathrm{i}\epsilon) = f(y - \mathrm{i}\epsilon) - \sum_{k=0}^{\lfloor a- \rfloor - 1} \frac{(-\mathrm{i})^k}{k!} \, f^{(k)}(0) \, (\epsilon + \mathrm{i}y)^k \,, \tag{C.5}$$

denoting by $\lfloor a- \rfloor = \lim_{\epsilon' \downarrow 0} \lfloor a - \epsilon' \rfloor$ the strict floor. As noted previously, we have $\int_{-\infty}^{+\infty} (\epsilon + \mathrm{i}y)^{-a} \, \mathrm{d}y = 0$ for $\epsilon > 0$ and $a > 1$. Adding and subtracting the Taylor polynomial in equation (C.5) thus yields

$$\int_{-\infty}^{+\infty} \frac{f(y - \mathrm{i}\epsilon)}{(\epsilon + \mathrm{i}y)^a} \, \mathrm{d}y = \int_{-\infty}^{+\infty} \frac{\tilde{f}_a(y - \mathrm{i}\epsilon)}{(\epsilon + \mathrm{i}y)^a} \, \mathrm{d}y \,. \tag{C.6}$$

Further noting that $f = \tilde{f}_a$ for $a = 1$, the identity displayed in equation (C.6) in fact holds for $\epsilon > 0$ and $a \geq 1$. As a consequence of the Cauchy integral theorem (e.g., theorem 10.35 in Rudin (1987)), we find that the right-hand-side of equation (C.6) may be written as

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{f(y - \mathrm{i}\epsilon)}{(\epsilon + \mathrm{i}y)^a} \, \mathrm{d}y &= \oint_{\mathcal{C}(\rho)} \frac{\tilde{f}_a(y)}{(\mathrm{i}y)^a} \, \mathrm{d}y + \int_{-\infty}^{-\rho} \frac{\tilde{f}_a(y)}{(\mathrm{i}y)^a} \, \mathrm{d}y + \int_{+\rho}^{+\infty} \frac{\tilde{f}_a(y)}{(\mathrm{i}y)^a} \, \mathrm{d}y \\ &= \oint_{\mathcal{C}(\rho)} \frac{\tilde{f}_a(y)}{(\mathrm{i}y)^a} \, \mathrm{d}y + \int_{+\rho}^{+\infty} \frac{\tilde{f}_a(y) + (-1)^a \tilde{f}_a(-y)}{(\mathrm{i}y)^a} \, \mathrm{d}y \,, \end{aligned} \tag{C.7}$$

where $\mathcal{C}(\rho)$ is a positively oriented (open) semicircle from $-\rho$ to $+\rho$, so that the pole at the origin lies in the exterior of the (infinite) closed curve along which the path integral is computed. Letting $\rho \downarrow 0$ in equation (C.7), we obtain

$$\int_{-\infty}^{+\infty} \frac{f(y - \mathrm{i}\epsilon)}{(\epsilon + \mathrm{i}y)^a} \, \mathrm{d}y = \mathrm{i}\pi \, \mathrm{Res}((\mathrm{i}y)^{-a} \tilde{f}_a(y), 0) + \int_0^{+\infty} \frac{\tilde{f}_a(y) + (-1)^a \tilde{f}_a(-y)}{(\mathrm{i}y)^a} \, \mathrm{d}y \,, \tag{C.8}$$

where $\mathrm{Res}(f, z)$ denotes the residue of the function f at the point z. Intuitively, the integral along the semicircle $\mathcal{C}(\rho)$ corresponds to half of the integral of the full circle, which obtains directly by virtue of the Cauchy residue theorem (e.g., theorem 10.42 in Rudin (1987)). Moreover, following the argument in the proof of lemma B.2, the integral on the right-hand-side of equation (C.8) is convergent in the sense of a Cauchy principal value integral.

It remains to determine the residue term in equation (C.8), for which we distinguish the cases $a \notin \mathbb{N}$ and $a \in \mathbb{N}$. For the easier case $a \notin \mathbb{N}$, the residue term vanishes and we obtain

$$\int_{-\infty}^{+\infty} \frac{f(y - \mathrm{i}\epsilon)}{(\epsilon + \mathrm{i}y)^a} \, \mathrm{d}y = \int_0^{+\infty} \frac{\tilde{f}_a(y) + (-1)^a \tilde{f}_a(-y)}{(\mathrm{i}y)^a} \, \mathrm{d}y \, . \tag{C.9}$$

Evidently, equation (C.9) corresponds to the non-integer cases in equations (C.3) and (C.4).

For $a \in \mathbb{N}$, we instead have that $\mathrm{Res}((\mathrm{i}y)^{-a} \tilde{f}_a(y), 0) = (-\mathrm{i})^a f^{(a-1)}(0)$, so that

$$\int_{-\infty}^{+\infty} \frac{f(y - \mathrm{i}\epsilon)}{(\epsilon + \mathrm{i}y)^a} \, \mathrm{d}y = (-\mathrm{i})^{a-1} \pi \, f^{(a-1)}(0) + \int_0^{+\infty} \frac{\tilde{f}_a(y) + (-1)^a \tilde{f}_a(-y)}{(\mathrm{i}y)^a} \, \mathrm{d}y \, . \tag{C.10}$$

By definition of the Dirac delta distribution and distributional derivatives, equation (C.10) corresponds to integer cases in equations (C.3) and (C.4). □

C.3 General derivatives

To study the pricing of general derivatives using complex Fourier theory, we start with an analogue of proposition B.1. For this, we impose that the pricing transform Π is a Schwartz function after regularization (shifting) and, hence, square-integrable. Instead of requiring that g represents a tempered distribution, we now impose that g is square-integrable after regularization (scaling). In that case, the regularized g_ϵ is locally integrable and, hence, represents a regular tempered distribution. Moreover, the ordinary Fourier transform $\hat{g}_\epsilon$ is also square-integrable and represents a regular tempered distribution. Under these conditions, a generalized transform relation can be derived using well-established results from ordinary Fourier theory, as in Dillschneider (2020). Formulating our result for one-side regularization is not restrictive in this context, since within two-side regularization proposition C.1 may straightforwardly be invoked twice after appropriately splitting the support of g.

Proposition C.1. *Let* $(y \mapsto g_\epsilon(y) = \exp(-\epsilon y)\, g(y)) \in \mathcal{S}^*(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$ *and* $(y \mapsto \Pi(\omega + \epsilon\hat{\omega} + \mathrm{i}y\hat{\omega}; \tilde{T}, z)) \in \mathcal{S}(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$ *for some* $\epsilon \in \mathbb{R}$ *and all* $z \in \mathcal{Z}$*. Then*

$$\begin{aligned}
\Pi_g(\omega, \hat{\omega}; \tilde{T}, Z_t) &= \mathrm{E}^{\mathbb{Q}}[D_t(\tilde{T}_{\tilde{m}}) \, \exp(\omega \cdot X_{t \oplus \tilde{T}}) \, g(\hat{\omega} \cdot X_{t \oplus \tilde{T}}) \,|\, \mathcal{F}_t] \\
&= \frac{1}{2\pi} \langle \hat{g}_\epsilon(y), \Pi(\omega + \epsilon\hat{\omega} + \mathrm{i}y\hat{\omega}; \tilde{T}, Z_t) \rangle \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \hat{g}_\epsilon(y) \, \Pi(\omega + \epsilon\hat{\omega} + \mathrm{i}y\hat{\omega}; \tilde{T}, Z_t) \, \mathrm{d}y
\end{aligned}$$

in terms of the distributional Fourier transform $\hat{g}_\epsilon \in \mathcal{S}^*(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$.

Proof. By construction, we have the identity $\Pi_g(\omega, \hat{\omega}; \tilde{T}, Z_t) = \Pi_{g_\epsilon}(\omega + \epsilon\hat{\omega}, \hat{\omega}; \tilde{T}, Z_t)$. Under the imposed conditions, invoking the results in Dillschneider (2020) as a consequence of Parseval's formula (e.g., p. 189 in Rudin (1991)), it follows that

$$\Pi_{g_\epsilon}(\omega + \epsilon\hat{\omega}, \hat{\omega}; \tilde{T}, z) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{g}_\epsilon(y) \, \Pi(\omega + \epsilon\hat{\omega} + \mathrm{i}y\hat{\omega}; \tilde{T}, z) \, \mathrm{d}y \, .$$

As $\hat{g}_\epsilon$ represents a regular tempered distribution, rewriting yields the stated expressions. □

With proposition C.1 at hand, we may revisit the pricing of general derivatives as in section 3.2. Staying within the framework introduced there, we appropriately modify the imposed assumptions in order to apply the complex Fourier theory. Analogous to assumptions 3.1 and 3.2, we impose the following conditions on the payoff function and pricing transform in assumption 3.1 and equation (3.3), respectively.

Assumption C.1. *The payoff function h satisfies equation (3.2) for $\bar{\omega}_i, \hat{\omega} \in \mathbb{R}^{n \times \tilde{m}}$ and $(\tilde{y} \mapsto g_{i,\varepsilon_i}(\tilde{y}; K) = \exp(-\varepsilon_i \tilde{y})\, g_i(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$ with $\varepsilon_i \in \mathbb{R}$.*

Assumption C.2. *$(y \mapsto \Pi(\mathfrak{b}(y); \tilde{T}, z)) \in \mathcal{S}(\mathcal{Y}_\varepsilon)$ for $\mathfrak{b}([\omega; \tilde{y}]) = \omega + \mathrm{i}\tilde{y}\hat{\omega}$, $\mathcal{Y}_\varepsilon = \bigcup_{i=1}^{n_h} \{\bar{\omega}_i + \varepsilon_i \hat{\omega}\} \times \mathbb{R}$, and all $z \in \mathcal{Z}$.*

The modifications in assumptions C.1 and C.2 allow the application of the generalized transform result in proposition C.1. Analogous to proposition 3.1, we obtain a transform-based pricing formula for general derivatives. Unlike before, we can now routinely express derivatives prices in terms of Dirac delta distributions and regular tempered distributions $\hat{g}_{i,\varepsilon_i}$. The last equality in equation (C.11) uses the integral notation for Dirac delta distribution to highlight the integral form.

Proposition C.2. *Let assumptions C.1 and C.2 hold. Then we have*

$$\begin{aligned}\mathcal{V}(Z_t; K, \tilde{T}) &= \mathrm{E}^{\mathbb{Q}}[D_t(\tilde{T}_{\tilde{m}})\, h(X_{t \oplus \tilde{T}}; K) \,|\, \mathcal{F}_t] \\ &= \langle w_\varepsilon(y; K), \Pi(\mathfrak{b}(y); \tilde{T}, Z_t)\rangle \\ &= \int_{\mathcal{Y}_\varepsilon} w_\varepsilon(y; K)\, \Pi(\mathfrak{b}(y); \tilde{T}, Z_t)\, \mathrm{d}y\,,\end{aligned} \tag{C.11}$$

where $y = [\tilde{\omega}; \tilde{y}]$ and $\mathfrak{b}([\tilde{\omega}; \tilde{y}]) = \tilde{\omega} + \mathrm{i}\tilde{y}\hat{\omega}$. Moreover, $(y \mapsto w_\varepsilon(y; K)) \in \mathcal{S}^*(\mathcal{Y}_\varepsilon)$ is given by the distributional tensor product

$$w_\varepsilon([\tilde{\omega}; \tilde{y}]; K) = \frac{1}{2\pi} \sum_{i=1}^{n_h} \delta(\tilde{\omega} - \bar{\omega}_i - \varepsilon_i \hat{\omega}) \otimes \hat{g}_{i,\varepsilon_i}(\tilde{y}; K)\,, \tag{C.12}$$

in terms of the distributional Fourier transforms $(\tilde{y} \mapsto \hat{g}_{i,\varepsilon_i}(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$.

Proof. Using the definition of the pricing transform Π in equation (3.3), the imposed assumptions together with proposition C.1 yield

$$\mathcal{V}(Z_t; K, \tilde{T}) = \frac{1}{2\pi} \sum_{i=1}^{n_h} \langle \hat{g}_{i,\varepsilon_i}(\tilde{y}; K), \Pi(\mathfrak{b}([\bar{\omega}_i + \varepsilon_i \hat{\omega}; \tilde{y}]); \tilde{T}, Z_t)\rangle\,. \tag{C.13}$$

Following the reasoning in the proof of proposition 3.1 with obvious modifications and admitting the integral notation for the Dirac delta distribution, equation (C.13) thus justifies equations (C.11) and (C.12). □

C.4 Equity derivatives

Consider the pricing of equity derivatives in the setting of section 3.3. As an analogue of corollary 3.1, the following corollary to proposition C.2 yields an expression for $\mathcal{V}^O_{\text{stock}}$ using complex Fourier theory.

Corollary C.1. *Let h^O_{stock} be as in equation (3.6). Moreover, let assumption C.2 hold for $\bar{\omega}_1 = [1; 0]$, $\bar{\omega}_2 = [0; 0]$, $\hat{\omega} = [1; 0]$, $\varepsilon_1 = \varepsilon_2 = c^O \epsilon$ with $\epsilon > 0$. Then we have*

$$\mathcal{V}^O_{\text{stock}}(Z_t; K, \tilde{T}) = \langle w^O_{\text{stock},\epsilon}(y; K), \Pi(\mathfrak{b}_{\text{stock}}(y); \tilde{T}, Z_t)\rangle\,, \tag{C.14}$$

where $y = [\tilde{\omega}; \tilde{y}]$ and $\mathfrak{b}_{\text{stock}}([\tilde{\omega}; \tilde{y}]) = \tilde{\omega} + \mathrm{i}\tilde{y}[1; 0]$. The associated $(y \mapsto w^O_{\text{stock},\epsilon}(y; K)) \in \mathcal{S}^*(\mathcal{Y}_\varepsilon)$ are given by

$$w^C_{\text{stock},\epsilon}([\tilde{\omega}; \tilde{y}]; K) = (\delta(\tilde{\omega} - (1+\epsilon)[1; 0]) - \exp(K)\, \delta(\tilde{\omega} - \epsilon[1; 0])) \otimes F^C_{\text{stock},\epsilon}(\tilde{y}; K)\, (\epsilon + \mathrm{i}\tilde{y})^{-1} \tag{C.15a}$$

$$w^P_{\text{stock},\epsilon}([\tilde{\omega}; \tilde{y}]; K) = (\exp(K)\, \delta(\tilde{\omega} + \epsilon[1; 0]) - \delta(\tilde{\omega} - (1-\epsilon)[1; 0])) \otimes F^P_{\text{stock},\epsilon}(\tilde{y}; K)\, (\epsilon + \mathrm{i}\tilde{y})^{-1} \tag{C.15b}$$

with $F^O_{\text{stock},\epsilon}(\tilde{y};K) = F^O_{\text{stock}}(\tilde{y} - \mathrm{i}\epsilon;K)$.

Proof. We proceed analogous to the proof of corollary 3.1. For the call payoff in equation (3.6a), now define $(\tilde{y} \mapsto g^C_{\text{stock},\epsilon}(\tilde{y};K)) \in \mathcal{S}^*(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$ by $g^C_{\text{stock},\epsilon}(\tilde{y};K) = \mathfrak{g}_\epsilon(\tilde{y};0,K)$ for $\mathfrak{g}_\epsilon$ as in lemma C.1. With equations (C.1) and (C.2), we then obtain the associated Fourier transform $(\tilde{y} \mapsto \hat{g}^C_{\text{stock},\epsilon}(\tilde{y};K)) \in \mathcal{S}^*(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$ as

$$\begin{aligned}\hat{g}^C_{\text{stock},\epsilon}(\tilde{y};K) &= \hat{\mathfrak{g}}_\epsilon(\tilde{y};0,K) \\ &= \exp(-\epsilon K - \mathrm{i}K\tilde{y})\,(\epsilon + \mathrm{i}\tilde{y})^{-1}\,,\end{aligned}$$

in terms of the (regular) tempered distribution $(\epsilon + \mathrm{i}\tilde{y})^{-1}$. Continuing the argument as before, we arrive at equations (C.14) and (C.15a) as a special case of proposition C.2.

For the put payoff in equation (3.6b), define $(\tilde{y} \mapsto g^P_{\text{stock},\epsilon}(\tilde{y};K)) \in \mathcal{S}^*(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$ by $g^P_{\text{stock},\epsilon}(\tilde{y};K) = -\mathfrak{g}_{-\epsilon}(-\tilde{y};0,-K)$ for $\mathfrak{g}_\epsilon$ as in lemma C.1. Invoking the scaling property of the Fourier transform, we obtain $(\tilde{y} \mapsto \hat{g}^P_{\text{stock},\epsilon}(\tilde{y};K)) \in \mathcal{S}^*(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$ as

$$\begin{aligned}\hat{g}^P_{\text{stock},\epsilon}(\tilde{y};K) &= -\,\hat{\mathfrak{g}}_{-\epsilon}(-\tilde{y};0,-K) \\ &= \exp(-\epsilon K - \mathrm{i}K\tilde{y})\,(\epsilon + \mathrm{i}\tilde{y})^{-1}\,,\end{aligned}$$

again in terms of the (regular) tempered distribution $(\epsilon + \mathrm{i}\tilde{y})^{-1}$. Following the rest of the argument thus yields equations (C.14) and (C.15b) as a special case of proposition C.2. □

For practical implementation, we may now easily obtain an integral representation of $w^O_{\text{stock},\epsilon}$ in corollary C.1. Intuitively, the corresponding representation in lemma 3.1 corresponds to the limiting case of lemma C.3 when letting $\epsilon \downarrow 0$. Whenever $\Upsilon(\mathfrak{b}_{\text{stock}}([\tilde{\omega};\tilde{y}]))$ is Hermitian as a function of $\tilde{y}$, a computationally more efficient representation may be obtained. Noting that both integrands in equation (C.16) are products of Hermitian functions and, hence, themselves Hermitian, the integrands may be replaced by twice their real parts.

Lemma C.3. *Let* $(y \mapsto \Upsilon(\mathfrak{b}_{\text{stock}}(y))) \in \mathcal{S}(\mathcal{Y}_\varepsilon)$. *Then* $w^O_{\text{stock},\epsilon}$ *in corollary C.1 can be represented in integral form as*

$$\begin{aligned}\langle w^O_{\text{stock},\epsilon}(y;K), \Upsilon(\mathfrak{b}_{\text{stock}}(y))\rangle &= \int_{\mathbb{R}} \frac{F^O_{\text{stock},\epsilon}(\tilde{y};K)\,\Upsilon((1+c^O\epsilon)[1;0] + \mathrm{i}\tilde{y}[1;0])}{\epsilon + \mathrm{i}\tilde{y}}\,\mathrm{d}\tilde{y} \\ &\quad - \exp(K) \int_{\mathbb{R}} \frac{F^O_{\text{stock},\epsilon}(\tilde{y};K)\,\Upsilon(c^O\epsilon[1;0] + \mathrm{i}\tilde{y}[1;0])}{\epsilon + \mathrm{i}\tilde{y}}\,\mathrm{d}\tilde{y}\,,\end{aligned} \tag{C.16}$$

with option indicators $c^C = +1$ *and* $c^P = -1$.

Proof. We proceed along the lines of the proof of lemma 3.1. The regular tempered distribution $(\epsilon + \mathrm{i}\tilde{y})^{-1}$ in equation (C.15) can be represented by a convergent integral, such that for every Schwartz function $f \in \mathcal{S}(\mathbb{R})$, we have

$$\langle(\epsilon + \mathrm{i}\tilde{y})^{-1}, f(\tilde{y})\rangle = \int_{\mathbb{R}} \frac{f(\tilde{y})}{\epsilon + \mathrm{i}\tilde{y}}\,\mathrm{d}\tilde{y}\,. \tag{C.17}$$

We may write $(y \mapsto w^O_{\text{stock},\epsilon}(y;K)) \in \mathcal{S}^*(\mathcal{Y}_\varepsilon)$ in equation (C.15) compactly as

$$w^O_{\text{stock},\epsilon}([\tilde{\omega};\tilde{y}];K) = (\delta(\tilde{\omega} - (1+c^O\epsilon)[1;0]) - \exp(K)\,\delta(\tilde{\omega} - c^O\epsilon)) \otimes \tilde{g}^O_{\text{stock},\epsilon}(\tilde{y};K) \tag{C.18}$$

with $(\tilde{y} \mapsto \tilde{g}^O_{\text{stock},\epsilon}(\tilde{y};K)) \in \mathcal{S}^*(\mathbb{R})$ in equation (C.18) given by $\tilde{g}^O_{\text{stock},\epsilon}(\tilde{y};K) = F^O_{\text{stock},\epsilon}(\tilde{y};K)\,(\epsilon + \mathrm{i}\tilde{y})^{-1}$, whose integral representation in equation (C.17) implies

$$\langle \tilde{g}^O_{\text{stock},\epsilon}(\tilde{y};K), \Upsilon(\mathfrak{b}_{\text{stock}}([\tilde{\omega};\tilde{y}]))\rangle = \int_{\mathbb{R}} \frac{F_{\text{stock},\epsilon}(\tilde{y};K)\,\Upsilon(\tilde{\omega}+\mathrm{i}\tilde{y}[1;0])}{\epsilon+\mathrm{i}\tilde{y}}\,\mathrm{d}\tilde{y}\,.$$

An application of the definition of the distributional tensor product in equation (3.1) to $\hat{w}^O_{\text{stock},\epsilon}$ in terms of $\tilde{g}^O_{\text{stock},\epsilon}$ in equation (C.18) then yields equation (C.16), as required. □

C.5 Volatility derivatives

Consider the pricing of volatility derivatives in the setting of section 3.4. As an analogue of corollary 3.2, the following corollary to proposition C.2 yields an expression for $\mathcal{V}^O_{\text{vix}}$ using complex Fourier theory.

Corollary C.2. *Let h^O_{vix} be as in equation (3.12). Moreover, let assumption C.2 hold for $\bar{\omega}_1 = [0;0]$, $\hat{\omega} = [0;b_{\text{vix}}]$, $\varepsilon_1 = \epsilon$ with $\epsilon > 0$. Then we have*

$$\mathcal{V}^O_{\text{vix}}(Z_t;K,\tilde{T}) = \langle w^O_{\text{vix},\epsilon}(y;K), \Pi(\mathfrak{b}_{\text{vix}}(y);\tilde{T},Z_t)\rangle\,, \tag{C.19}$$

where $y=[\tilde{\omega};\tilde{y}]$ and $\mathfrak{b}_{\text{vix}}([\tilde{\omega};\tilde{y}]) = \tilde{\omega}+\mathrm{i}\tilde{y}[0;b_{\text{vix}}]$. The associated $(y\mapsto w^O_{\text{vix},\epsilon}(y;K))\in\mathcal{S}^*(\mathcal{Y}_\varepsilon)$ are given by

$$w^C_{\text{vix},\epsilon}([\tilde{\omega};\tilde{y}];K) = \delta(\tilde{\omega}-\epsilon[0;b_{\text{vix}}])\otimes F^C_{\text{vix},\epsilon}(\tilde{y};K)\,(\epsilon+\mathrm{i}\tilde{y})^{-3/2} \tag{C.20a}$$

$$w^P_{\text{vix},\epsilon}([\tilde{\omega};\tilde{y}];K) = \delta(\tilde{\omega})\otimes K^{1/2}\,\delta(\tilde{y}) + \delta(\tilde{\omega}-\epsilon[0;b_{\text{vix}}])\otimes F^P_{\text{vix},\epsilon}(\tilde{y};K)\,(\epsilon+\mathrm{i}\tilde{y})^{-3/2} \tag{C.20b}$$

with $F^O_{\text{vix},\epsilon}(\tilde{y};K) = F^O_{\text{vix}}(\tilde{y}-\mathrm{i}\epsilon;K)$.

Proof. We proceed analogous to the proof of corollary 3.2. For the call payoff in equation (3.12a), define $(\tilde{y}\mapsto g^C_{\text{vix},\epsilon}(\tilde{y};K))\in\mathcal{S}^*(\mathbb{R})\cap\mathcal{L}^2(\mathbb{R})$ by $g^C_{\text{vix},\epsilon}(\tilde{y};K) = \mathfrak{g}_\epsilon(\tilde{y};1/2,K) - K^{1/2}\mathfrak{g}_\epsilon(\tilde{y};0,K)$ for $\mathfrak{g}_\epsilon$ as in lemma C.1. With equations (C.1) and (C.2), we obtain the associated Fourier transform $(\tilde{y}\mapsto \hat{g}^C_{\text{vix},\epsilon}(\tilde{y};K))\in\mathcal{S}^*(\mathbb{R})\cap\mathcal{L}^2(\mathbb{R})$ as

$$\begin{aligned}\hat{g}^C_{\text{vix},\epsilon}(\tilde{y};K) &= \hat{\mathfrak{g}}_\epsilon(\tilde{y};1/2,K) - K^{1/2}\hat{\mathfrak{g}}_\epsilon(\tilde{y};0,K)\\ &= \frac{\frac{1}{2}\Gamma(1/2,\epsilon K+\mathrm{i}K\tilde{y})}{\Gamma(3/2)}\,\hat{\mathfrak{g}}_\epsilon(\tilde{y};1/2,0)\\ &= \tfrac{1}{2}\Gamma(1/2,\epsilon K+\mathrm{i}K\tilde{y})\,(\epsilon+\mathrm{i}\tilde{y})^{-3/2}\,,\end{aligned}$$

in terms of the (regular) tempered distribution $(\epsilon+\mathrm{i}\tilde{y})^{-3/2}$. Keeping the setting as before and using the shift property of the Fourier transform yields equations (C.19) and (C.20a) as a special case of proposition C.2.

For the put payoff in equation (3.12b), define $(\tilde{y}\mapsto g^P_{\text{vix},\epsilon}(\tilde{y};K))\in\mathcal{S}^*(\mathbb{R})\cap\mathcal{L}^2(\mathbb{R})$ by $g^P_{\text{vix},\epsilon}(\tilde{y};K) = K^{1/2}\big(\mathfrak{g}_\epsilon(\tilde{y};0,0)-\mathfrak{g}_\epsilon(\tilde{y};0,K)\big) - \big(\mathfrak{g}_\epsilon(\tilde{y};1/2,0)-\mathfrak{g}_\epsilon(\tilde{y};1/2,K)\big)$ for $\mathfrak{g}_\epsilon$ as in lemma C.1. The scaling property of the Fourier transform together with equations (C.1) and (C.2) yields the associated Fourier transform $(\tilde{y}\mapsto \hat{g}^P_{\text{vix},\epsilon}(\tilde{y};K))\in\mathcal{S}^*(\mathbb{R})\cap\mathcal{L}^2(\mathbb{R})$ as

$$\begin{aligned}\hat{g}^P_{\text{vix},\epsilon}(\tilde{y};K) &= K^{1/2}\big(\hat{\mathfrak{g}}_\epsilon(\tilde{y};0,0)-\hat{\mathfrak{g}}_\epsilon(\tilde{y};0,K)\big) - \big(\hat{\mathfrak{g}}_\epsilon(\tilde{y};1/2,0)-\hat{\mathfrak{g}}_\epsilon(\tilde{y};1/2,K)\big)\\ &= K^{1/2}\hat{\mathfrak{g}}_\epsilon(\tilde{y};0,0) - \frac{\frac{1}{2}\gamma(1/2,\epsilon K+\mathrm{i}Ky)}{\Gamma(3/2)}\,\hat{\mathfrak{g}}_\epsilon(\tilde{y};1/2,0)\\ &= 2\pi\,K^{1/2}\,\delta(\tilde{y}-\mathrm{i}\epsilon) - \tfrac{1}{2}\gamma(1/2,\epsilon K+\mathrm{i}Ky)\,(\epsilon+\mathrm{i}y)^{-3/2}\,,\end{aligned}$$

in terms of the (complex) Dirac delta distribution, acting as $\langle \delta(\tilde{y} - \mathrm{i}\epsilon), f(\tilde{y}) \rangle = f(\mathrm{i}\epsilon)$, and the (regular) tempered distribution $(\epsilon + \mathrm{i}y)^{-3/2}$. The last equality follows by realizing that $\hat{\mathfrak{g}}_\epsilon(\tilde{y}; 0, 0) = (\epsilon + \mathrm{i}\tilde{y})^{-1}$ is the complex Fourier transform of $\mathfrak{g}(\tilde{y}; 0, 0) = \mathfrak{U}(\tilde{y})$, which is redundant when positivity is assured. Invoking the shift property of the Fourier transform thus yields equations (C.19) and (C.20b) as a special case of proposition C.2. $\square$

For practical implementation, $w^O_{\mathrm{vix},\epsilon}$ in corollary C.2 straightforwardly implies an integral representation. Intuitively, the corresponding representation in lemma 3.3 may be interpreted as a limiting case of lemma C.4 when letting $\epsilon \downarrow 0$. From a computational perspective, a more efficient representation can be devised for the special case of $\Upsilon(\mathfrak{b}_{\mathrm{vix}}([\tilde{\omega}; \tilde{y}]))$ being Hermitian as a function of $\tilde{y}$. In that case, the integrand in equation (C.21) is a product of Hermitian functions of $\tilde{y}$ and, hence, itself Hermitian. Therefore, the integrand may be replaced by twice its real part.

Lemma C.4. *Let $(y \mapsto \Upsilon(\mathfrak{b}_{\mathrm{vix}}(y))) \in \mathcal{S}(\mathcal{Y}_\varepsilon)$. Then $w^O_{\mathrm{vix},\epsilon}$ in corollary C.2 can be represented in integral form as*

$$\langle w^O_{\mathrm{vix},\epsilon}(y; K), \Upsilon(\mathfrak{b}_{\mathrm{vix}}(y)) \rangle = \frac{1 - c^O}{2} K^{1/2}\, \Upsilon([0; 0]) + \int_{\mathbb{R}} \frac{F^O_{\mathrm{vix},\epsilon}(\tilde{y}; K)\, \Upsilon(\epsilon[0; b_{\mathrm{vix}}] + \mathrm{i}\tilde{y}[0; b_{\mathrm{vix}}])}{(\epsilon + \mathrm{i}\tilde{y})^{3/2}}\, \mathrm{d}\tilde{y}\,, \quad \text{(C.21)}$$

with option indicators $c^C = +1$ and $c^P = -1$.

Proof. We proceed as in the proof of lemma 3.3. The regular tempered distribution $(\epsilon + \mathrm{i}\tilde{y})^{-3/2} \in \mathcal{S}^*(\mathbb{R})$ for every Schwartz function $f \in \mathcal{S}(\mathbb{R})$ can be represented by the convergent integral

$$\langle (\epsilon + \mathrm{i}\tilde{y})^{-3/2}, f(\tilde{y}) \rangle = \int_{\mathbb{R}} \frac{f(\tilde{y})}{(\epsilon + \mathrm{i}\tilde{y})^{3/2}}\, \mathrm{d}\tilde{y}\,. \quad \text{(C.22)}$$

Moreover, we may compactly express $(y \mapsto w^O_{\mathrm{vix},\epsilon}(y; K)) \in \mathcal{S}^*(\mathcal{Y}_\varepsilon)$ in equation (C.20) as

$$w^O_{\mathrm{vix},\epsilon}([\tilde{\omega}; \tilde{y}]; K) = \frac{1 - c^O}{2} K^{1/2}\, \delta(\tilde{\omega}) \otimes \delta(\tilde{y}) + \delta(\tilde{\omega} - \epsilon[0; b_{\mathrm{vix}}]) \otimes \tilde{\tilde{g}}^O_{\mathrm{vix},\epsilon}(\tilde{y}; K) \quad \text{(C.23)}$$

with $(\tilde{y} \mapsto \tilde{\tilde{g}}^O_{\mathrm{vix},\epsilon}(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ in equation (C.23) defined as $\tilde{\tilde{g}}^O_{\mathrm{vix},\epsilon}(\tilde{y}; K) = F^O_{\mathrm{vix},\epsilon}(\tilde{y}; K)\, (\epsilon + \mathrm{i}\tilde{y})^{-3/2}$, whose integral representation in equation (C.22) implies

$$\langle \tilde{\tilde{g}}^O_{\mathrm{vix},\epsilon}(\tilde{y}; K), \Upsilon(\mathfrak{b}_{\mathrm{vix}}([\tilde{\omega}; \tilde{y}])) \rangle = \int_{\mathbb{R}} \frac{F^O_{\mathrm{vix},\epsilon}(\tilde{y}; K)\, \Upsilon(\tilde{\omega} + \mathrm{i}\tilde{y}[0; b_{\mathrm{vix}}])}{(\epsilon + \mathrm{i}\tilde{y})^{3/2}}\, \mathrm{d}\tilde{y}\,.$$

An application of the definition of the distributional tensor product in equation (3.1) to $\hat{w}^O_{\mathrm{vix},\epsilon}$ in terms of the Dirac delta distribution and $\tilde{\tilde{g}}^O_{\mathrm{vix},\epsilon}$ in equation (C.23) then yields equation (C.21), as required. $\square$

D Supplement to moments involving derivatives prices

This appendix contains the proofs for the results in section 4.

D.1 Auxiliary results

Using the definition of extended Schwartz spaces in section 4.2, consider a generic transform Υ satisfying $((y, z) \mapsto \Upsilon(y; z)) \in \tilde{\mathcal{S}}(\tilde{\mathcal{Y}} \times \mathcal{Z}^{\tilde{n}}; \mathbb{1} \otimes v)$ for $\tilde{\mathcal{Y}} \subset \mathbb{R}^m$ and some positive weighting function $v \in \mathcal{C}^\infty(\mathcal{Z}^{\tilde{n}})$. Given this construction, we define $\psi(z) = \langle \tilde{g}(y), \Upsilon(y; z) \rangle$ for $\tilde{g} \in \mathcal{S}^*(\tilde{\mathcal{Y}}) = \tilde{\mathcal{S}}^*(\tilde{\mathcal{Y}}; \mathbb{1})$. We aim at determining moments involving $\psi(Z_{t+\tilde{\tau}})$ and functions of the augmented state vector $X_{t\oplus\tilde{\tau}}$. Using the results in

Dillschneider (2020), we have the following Fubini-type result that allows to interchange the order of the tempered distribution and the expectation operator.

Proposition D.1. *Let* $\tilde{g} \in \tilde{\mathcal{S}}^*(\tilde{\mathcal{Y}}; \mathbb{1})$ *and* $((y, z) \mapsto \Upsilon(y; z)) \in \tilde{\mathcal{S}}(\tilde{\mathcal{Y}} \times \mathcal{Z}^{\tilde{n}}; \mathbb{1} \otimes v)$ *with*

$$\mathrm{E}^{\mathbb{M}}[|h(X_{t\oplus\tilde{\tau}})|\, v(Z_{t+\tilde{\tau}})^{-1}] < \infty\,. \tag{D.1}$$

Then $\psi(z) = \langle \tilde{g}(y), \Upsilon(y; z)\rangle$ *satisfies*

$$\begin{aligned} \mathrm{E}^{\mathbb{M}}[f(X_{t\oplus\tilde{\tau}})\, \psi(Z_{t+\tilde{\tau}})] &= \mathrm{E}^{\mathbb{M}}[h(X_{t\oplus\tilde{\tau}})\, \langle \tilde{g}(y), \Upsilon(y; Z_{t+\tilde{\tau}})\rangle] \\ &= \langle \tilde{g}(y), \mathrm{E}^{\mathbb{M}}[h(X_{t\oplus\tilde{\tau}})\, \Upsilon(y; Z_{t+\tilde{\tau}})]\rangle\,. \end{aligned} \tag{D.2}$$

Proof. The first equality in equation (D.2) holds by definition of ψ. For the second equality, denote the (linear) expectation functional $\mathcal{E}_h$ by $\langle \mathcal{E}_h(z), f(z)\rangle = \mathrm{E}^{\mathbb{M}}[h(X_{t\oplus\tilde{\tau}})\, f(Z_{t+\tilde{\tau}})]$. Condition (D.1) assures that $\mathcal{E}_h \in \tilde{\mathcal{S}}^*(\mathcal{Z}^{\tilde{n}}; v)$. It holds by construction of extended Schwartz spaces and tempered distributions that $\tilde{g} \otimes \mathcal{E}_h \in \tilde{\mathcal{S}}^*(\tilde{\mathcal{Y}} \times \mathcal{Z}^{\tilde{n}}; \mathbb{1} \otimes v)$. Here, the action of the tensor product $\langle \tilde{g}(y) \otimes \mathcal{E}_h(z), \Upsilon(y; z)\rangle$ equals $\langle \mathcal{E}_h(z), \langle \tilde{g}(y), \Upsilon(y; z)\rangle\rangle = \langle \tilde{g}(y), \langle \mathcal{E}_h(z), \Upsilon(y; z)\rangle\rangle$, analogous to equation (3.1). This yields the second equality in equation (D.2). □

D.2 Exact moments

Proof of lemma 4.1. The proof is split into two steps. First, we derive the relevant expressions for each $(V_{t+\tilde{\tau}_j})^{\beta_j}$. By construction of the vector V_t, from equation (4.1), we have

$$V_{i,t+\tilde{\tau}_j} = \mathcal{V}_i(Z_{t+\tilde{\tau}_j}; K_i, \tilde{T}_i) = \langle w_i(y_{i,j}; K_i), \Pi(\mathfrak{b}_i(y_{i,j}); \tilde{T}_i, Z_{t+\tilde{\tau}_j})\rangle$$

for $(y_{i,j} \mapsto w_i(y_{i,j}; K_i)) \in \mathcal{S}^*(\mathcal{Y}_i)$ and $(y_{i,j} \mapsto \Pi(\mathfrak{b}_i(y_{i,j}); \tilde{T}_i, Z_{t+\tilde{\tau}_j})) \in \mathcal{S}(\mathcal{Y}_i)$. Given multi-indices $\beta_j \in \mathbb{N}^{n_V}$, define associated index vectors $q(\beta_j) \in \mathbb{N}^{|\beta_j|}$ with multiplicities according to β_j such that $(V_{t+\tilde{\tau}_j})^{\beta_j} = \prod_{i=1}^{|\beta_j|} V_{q_i(\beta_j),t+\tilde{\tau}_j}$. By definition of tensor products of tempered distributions in equation (3.1), we thus have

$$(V_{t+\tilde{\tau}_j})^{\beta_j} = \langle w^{\beta_j}(y_j; K), \Pi^{\beta_j}(\mathfrak{b}(y_j); \tilde{T}, Z_{t+\tilde{\tau}_j})\rangle \tag{D.3}$$

for $(y_j \mapsto w^{\beta_j}(y_j; K)) \in \mathcal{S}^*(\mathcal{Y}^{\beta_j})$ and $(y_j \mapsto \Pi^{\beta_j}(\mathfrak{b}(y_j); \tilde{T}, Z_{t+\tilde{\tau}_j})) \in \mathcal{S}(\mathcal{Y}_i)$ on the Cartesian product space $\mathcal{Y}^{\beta_j} = \prod_{i=1}^{|\beta_j|} \mathcal{Y}_{q_i(\beta_j)}$. Each tempered distributions w^{β_j} in equation (D.3) is given by the distributional tensor product

$$w^{\beta_j}(y_j; K) = \bigotimes_{i=1}^{|\beta_j|} w_{q_i(\beta_j)}(y_{i,j}; K_{q_i(\beta_j)}) \tag{D.4}$$

and each Π^{β_j} in equation (D.3) by

$$\begin{aligned} \Pi^{\beta_j}(\mathfrak{b}(y_j); \tilde{T}, Z_{t+\tilde{\tau}_j}) &= \prod_{i=1}^{|\beta_j|} \Pi_{q_i(\beta_j)}(\mathfrak{b}_{q_i(\beta_j)}(y_{i,j}); \tilde{T}_{q_i(\beta_j)}, Z_{t+\tilde{\tau}_j}) \\ &= \exp(A_\Pi^{\beta_j}(\mathfrak{b}(y_j); \tilde{T}) + B_\Pi^{\beta_j}(\mathfrak{b}(y_j); \tilde{T}) \cdot Z_{t+\tilde{\tau}_j})\,. \end{aligned} \tag{D.5}$$

Moreover, the coefficients $A_\Pi^{\beta_j}$ and $B_\Pi^{\beta_j}$ in equation (D.5) are determined as

$$A_\Pi^{\beta_j}(\mathfrak{b}(y_j);\tilde{T}) = \sum_{i=1}^{|\beta_j|} A_\Pi(\mathfrak{b}_{q_i(\beta_j)}(y_{i,j});\tilde{T}_{q_i(\beta_j)}) \tag{D.6a}$$

$$B_\Pi^{\beta_j}(\mathfrak{b}(y_j);\tilde{T}) = \sum_{i=1}^{|\beta_j|} B_\Pi(\mathfrak{b}_{q_i(\beta_j)}(y_{i,j});\tilde{T}_{q_i(\beta_j)}) \,. \tag{D.6b}$$

This completes the first step.

Second, we combine the expressions relating to $(V_{t+\tilde{\tau}_j})^{\beta_j}$ to obtain analogous expressions relating to $(V_{t+\tilde{\tau}})^{\beta} = \prod_{j=1}^{\tilde{n}}(V_{t+\tilde{\tau}_j})^{\beta_j}$. Using w^{β_j} from equation (D.4) and again invoking the definition of the distributional tensor product in equation (3.1) yields equation (4.2) with w^β defined by the distributional tensor product

$$w^\beta(y;K) = \bigotimes_{j=1}^{\tilde{n}} w^{\beta_j}(y_j;K) \,, \tag{D.7}$$

where $(y \mapsto w^\beta(y;K)) \in \mathcal{S}^*(\mathcal{Y}^\beta)$ on the Cartesian product space $\mathcal{Y}^\beta = \prod_{j=1}^{\tilde{n}} \mathcal{Y}^{\beta_j}$. Moreover, Π^β in equation (4.2) is given by

$$\Pi^\beta(\mathfrak{b}(y);\tilde{T},Z_{t+\tilde{\tau}}) = \prod_{j=1}^{\tilde{n}} \Pi^{\beta_j}(\mathfrak{b}(y_j);\tilde{T},Z_{t+\tilde{\tau}_j}) \tag{D.8}$$

with $(y \mapsto \Pi^\beta(\mathfrak{b}(y);\tilde{T},Z_{t+\tilde{\tau}})) \in \mathcal{S}(\mathcal{Y}^\beta)$. Since each of the Π^{β_j} in equation (D.5) is exponentially affine in $Z_{t+\tilde{\tau}_j}$, it follows that Π^β in equation (D.8) is exponentially affine in $Z_{t+\tilde{\tau}}$, yielding equation (4.3) with coefficients A_Π^β and B_Π^β given as

$$A_\Pi^\beta(\mathfrak{b}(y);\tilde{T}) = \sum_{j=1}^{\tilde{n}} A_\Pi^{\beta_j}(\mathfrak{b}(y_j);\tilde{T}) \tag{D.9a}$$

$$B_\Pi^\beta(\mathfrak{b}(y);\tilde{T}) = [B_\Pi^{\beta_1}(\mathfrak{b}(y_1);\tilde{T});\ldots;B_\Pi^{\beta_{\tilde{n}}}(\mathfrak{b}(y_{\tilde{n}});\tilde{T})] \tag{D.9b}$$

in terms of the coefficients $A_\Pi^{\beta_j}$ and $B_\Pi^{\beta_j}$ in equation (D.6). This concludes the second step and, hence, the proof. $\square$

Proof of proposition 4.1. To ease notation, we introduce

$$F_{t+\tilde{\tau}}^{(\alpha)}(\omega) = \exp(\omega \cdot X_{t\oplus\tilde{\tau}})\,(X_{t\oplus\tilde{\tau}})^\alpha \,, \tag{D.10}$$

satisfying $\Phi^{\mathbb{M},[\alpha]}(\omega;\tilde{\tau},\infty) = \mathrm{E}^{\mathbb{M}}[F_{t+\tilde{\tau}}^{(\alpha)}(\omega)]$. Further using $(V_{t+\tilde{\tau}})^\beta$ from equation (4.2), we obtain

$$\begin{aligned}
\tilde{\Phi}^{\mathbb{M},[\alpha,\beta]}(\omega,0;\tilde{\tau},\infty) &= \mathrm{E}^{\mathbb{M}}[F_{t+\tilde{\tau}}^{(\alpha)}(\omega)\,(V_{t+\tilde{\tau}})^\beta] \\
&= \mathrm{E}^{\mathbb{M}}[\langle w^\beta(y;K), \Pi^\beta(\mathfrak{b}(y);\tilde{T},Z_{t+\tilde{\tau}})\,F_{t+\tilde{\tau}}^{(\alpha)}(\omega)\rangle] \\
&= \langle w^\beta(y;K), \mathrm{E}^{\mathbb{M}}[\Pi^\beta(\mathfrak{b}(y);\tilde{T},Z_{t+\tilde{\tau}})\,F_{t+\tilde{\tau}}^{(\alpha)}(\omega)]\rangle \,,
\end{aligned} \tag{D.11}$$

the last equality following from proposition D.1.

Finally, using the exponentially affine form of Π^β in equation (4.3) and the definition of $F_{t+\tilde{\tau}}^{(\alpha)}$ in equation (D.10), we arrive at

$$\begin{aligned}
\mathrm{E}^{\mathbb{M}}[\Pi^\beta(\mathfrak{b}(y);\tilde{T},Z_{t+\tilde{\tau}})\,F_{t+\tilde{\tau}}^{(\alpha)}(\omega)] &= \exp(A_\Pi^\beta(\mathfrak{b}(y);\tilde{T}))\,\mathrm{E}^{\mathbb{M}}[F_{t+\tilde{\tau}}^{(\alpha)}(\omega + [0;B_\Pi^\beta(\mathfrak{b}(y);\tilde{T})])] \\
&= \exp(A_\Pi^\beta(\mathfrak{b}(y);\tilde{T}))\,\Phi^{\mathbb{M},[\alpha]}(\omega + [0;B_\Pi^\beta(\mathfrak{b}(y);\tilde{T})];\tilde{\tau},\infty)
\end{aligned} \tag{D.12}$$

with $\Phi^{\mathbb{M},[\alpha]}$ as in equation (2.9). Substituting equation (D.12) into equation (D.11) yields equation (4.4), as required. $\square$

D.3 Approximate moments

Proof of proposition 4.2. From lemma 4.2 and basis representation in terms of monomials, equation (4.6) holds with coefficients

$$\tilde{b}_{V,\eta,(p)} = \sum_{|\gamma|\leq p} b^{(\gamma)}_{\phi,\eta}\, \tilde{c}_{V,\gamma} \,.$$

Moreover, using $V_{t+\tilde{\tau},(p)} = \sum_{j=1}^{\tilde{n}} e_j \otimes V_{t+\tilde{\tau}_j,(p)}$ with $e_j \in \mathbb{N}^{\tilde{n}}$ and $V_{t+\tilde{\tau}_j,(p)}$ in the form of equation (4.6), we conclude that equation (4.7) is valid with all non-zero coefficients being of the form

$$b_{V,e_j\otimes\eta} = e_j \otimes \tilde{b}_{V,\eta} \,.$$

By Taylor expansion, using the Faà di Bruno formula (A.1) and $b_{V,\eta,(p)}$ from equation (4.7), it moreover follows that equation (4.8) holds with coefficients

$$b^{(\beta)}_{V,\eta,(p)} = \frac{1}{\eta!} \sum_{\substack{|\rho|\leq|\eta| \\ \rho\leq\beta}} \frac{\beta!}{(\beta-\rho)!} b^{\beta-\rho}_{V,0,(p)} \sum_{\mathcal{Q}(\eta,\rho)} M^{\rho}_{k,\ell} [b_{V,\ell,(p)}(\ell!)]^k \,. \tag{D.13}$$

Equation (4.9) is then a straightforward consequence of the form of $(V_{t+\tilde{\tau},(p)})^\beta$ in equation (4.8) and the definition of pl-linear moments in equation (2.9).

For the convergence result, we invoke the Vitali convergence theorem (e.g., p. 187 in Folland (1999)). Defining $\mathcal{L}^1(\mathcal{Z},\mathbb{M})$ to be the space of integrable functions on $\mathcal{Z}$ against the probability measure $\mathbb{M}$, we thereby have that the following are equivalent: (i) $\mathcal{L}^1(\mathcal{Z},\mathbb{M})$ convergence and (ii) convergence in the measure $\mathbb{M}$ and uniform integrability. By $\mathcal{L}^2(\mathcal{Z},\mathbb{M})$ convergence of $((V_{t+\tilde{\tau},(p)})^\beta)_p$ and the imposed uniform integrability, the convergence result in proposition 4.2 follows. $\square$

Proof of lemma 4.3. To obtain equation (4.10), note that

$$\begin{aligned} \tilde{c}_{V,\alpha} &= \sum_{i=1}^{n_V} e_i\, \mathrm{E}^{\mathbb{M}}[V_{i,t}\, \phi_\alpha(Z_t)] \\ &= \sum_{i=1}^{n_V} e_i \sum_{\gamma\preccurlyeq\alpha} b^{(\alpha)}_{\phi,\gamma}\, \mathrm{E}^{\mathbb{M}}[V_{i,t}\,(Z_t)^\gamma] \\ &= \sum_{i=1}^{n_V} e_i \sum_{\gamma\preccurlyeq\alpha} b^{(\alpha)}_{\phi,\gamma}\, \tilde{\Phi}^{\mathbb{M},[[0;\gamma],e_i]}(0,0;0,\infty) \,. \end{aligned} \tag{D.14}$$

By the imposed assumptions, according to proposition 4.1, each $\tilde{\Phi}^{\mathbb{M},[[0;\gamma],e_i]}$ in equation (D.14) can be computed as a special case of equation (4.4). Using that $w^{e_i}(y;K) = w_i(y_i;K_i)$, $A^{e_i}_\Pi(\mathfrak{b}(y);\tilde{T}) = A_\Pi(\mathfrak{b}_i(y_i);\tilde{T}_i)$, and $B^{e_i}_\Pi(\mathfrak{b}(y);\tilde{T}) = B_\Pi(\mathfrak{b}_i(y_i);\tilde{T}_i)$ by lemma 4.1 thus yields the result stated in equation (4.10). $\square$

References

Ackerer, D. and D. Filipović (2019). "Option Pricing with Orthogonal Polynomial Expansions". *Mathematical Finance*, forthcoming.

Aït-Sahalia, Y. (2002). "Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-form Approximation Approach". *Econometrica* 70 (1), pp. 223–262.

— (2008). "Closed-form likelihood expansions for multivariate diffusions". *Annals of Statistics* 36 (2), pp. 906–937.

Aït-Sahalia, Y., D. Amengual, and E. Manresa (2015a). "Market-based estimation of stochastic volatility models". *Journal of Econometrics* 187 (2), pp. 418–435.

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

Aït-Sahalia, Y., M. Karaman, and L. Mancini (2020). "The term structure of equity and variance risk premia". *Journal of Econometrics*, forthcoming.

Aït-Sahalia, Y. and R. L. Kimmel (2007). "Maximum likelihood estimation of stochastic volatility models". *Journal of Financial Economics* 83 (2), pp. 413–452.

Aït-Sahalia, Y., C. Li, and C. X. Li (2021a). "Closed-form implied volatility surfaces for stochastic volatility models with jumps". *Journal of Econometrics*, forthcoming.

— (2021b). "Implied Stochastic Volatility Models". *Review of Financial Studies* 34 (1), pp. 394–450.

Andersen, T. G., L. Benzoni, and J. Lund (2002). "An Empirical Investigation of Continuous-Time Equity Return Models". *Journal of Finance* 57 (3), pp. 1239–1284.

Andersen, T. G., N. Fusari, and V. Todorov (2015). "Parametric Inference and Dynamic State Recovery From Option Panels". *Econometrica* 83 (3), pp. 1081–1145.

Andersen, T. G., N. Fusari, V. Todorov, and R. T. Varneskov (2018). "Inference for Option Panels in Pure-Jump Settings". Working paper.

— (2020). "Spatial Dependence in Option Observation Errors". *Econometric Theory*, forthcoming.

Armstrong, T. B. and M. Kolesár (2019). "Sensitivity Analysis using Approximate Moment Condition Models". Working paper.

Bakshi, G., C. Cao, and Z. Chen (1997). "Empirical Performance of Alternative Option Pricing Models". *Journal of Finance* 52 (5), pp. 2003–2049.

Bakshi, G. and N. Ju (2005). "A Refinement to Aït-Sahalia's (2002) "Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach"". *Journal of Business* 78 (5), pp. 2037–2052.

Bakshi, G., N. Ju, and H. Ou-Yang (2006). "Estimation of continuous-time models with an application to equity volatility dynamics". *Journal of Financial Economics* 82 (1), pp. 227–249.

Bakshi, G. and D. B. Madan (2000). "Spanning and derivative-security valuation". *Journal of Financial Economics* 55 (2), pp. 205–238.

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

Barletta, A. and E. Nicolato (2017). "Orthogonal expansions for VIX options under affine jump diffusions". *Quantitative Finance* 18 (6), pp. 951–967.

Barras, L. and A. Malkhozov (2016). "Does variance risk have two prices? Evidence from the equity and option markets". *Journal of Financial Economics* 121 (1), pp. 79–92.

Bates, D. S. (2000). "Post-'87 crash fears in the S&P 500 futures option market". *Journal of Econometrics* 94 (1–2), pp. 181–238.

Bates, D. S. (2006). “Maximum Likelihood Estimation of Latent Affine Processes”. *Review of Financial Studies* 19 (3), pp. 909–965.

Bollerslev, T., M. Gibson, and H. Zhou (2011). “Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities”. *Journal of Econometrics* 160 (1), pp. 235–245.

Bollerslev, T. and H. Zhou (2002). “Estimating stochastic volatility diffusion using conditional moments of integrated volatility”. *Journal of Econometrics* 109 (1), pp. 33–65.

Boswijk, H. P., R. J. A. Laeven, and A. Lalu (2016). “Asset Returns with Self-Exciting Jumps: Option Pricing and Estimation with a Continuum of Moments”. Working paper.

Branger, N., A. Hansis, and C. Schlag (2010). “Expected Option Returns and the Structure of Jump Risk Premia”. Working paper.

Branger, N., A. Kraftschik, and C. Völkert (2016). “The Fine Structure of Variance: Pricing VIX Derivatives in Consistent and Log-VIX Models”. Working paper.

Broadie, M. and A. Jain (2008). “The Effect of Jumps and Discrete Sampling on Volatility and Variance Swaps”. *International Journal of Theoretical and Applied Finance* 11 (8), pp. 761–797.

Broadie, M. and Ö. Kaya (2006). “Exact Simulation of Stochastic Volatility and Other Affine Jump Diffusion Processes”. *Operations Research* 54 (2), pp. 217–231.

Carr, P. and R. Lee (2009). “Robust Replication of Volatility Derivatives”. Working paper.

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

— (2001). “Optimal positioning in derivative securities”. *Quantitative Finance* 1 (1), pp. 19–37.

— (2014). “Joint modeling of VIX and SPX options at a single and common maturity with risk management applications”. *IIE Transactions* 46 (11), pp. 1125–1131.

Carr, P. and L. Wu (2009). “Variance Risk Premiums”. *Review of Financial Studies* 22 (3), pp. 1311–1341.

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

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

Chen, H., W. W. Dou, and L. Kogan (2019). “Measuring ”dark matter” in asset pricing models”. Working paper.

Chen, H. and S. Joslin (2012). “Generalized Transform Analysis of Affine Processes and Applications in Finance”. *Review of Financial Studies* 25 (7), pp. 2225–2256.

Cheridito, P., D. Filipović, and R. L. Kimmel (2007). “Market price of risk specifications for affine models: Theory and evidence”. *Journal of Financial Economics* 83 (1), pp. 123–170.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen (2003). “Alternative models for stock price dynamics”. *Journal of Econometrics* 116 (1–2), pp. 225–257.

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

Christoffersen, P., K. Jacobs, and K. Mimouni (2010). “Volatility Dynamics for the S&P500: Evidence from Realized Volatility, Daily Returns, and Option Prices”. *Review of Financial Studies* 23 (8), pp. 3141–3189.

Constantine, G. M. and T. H. Savits (1996). “A Multivariate Faà Di Bruno Formula With Applications”. *Transactions of the American Mathematical Society* 348 (2), pp. 503–520.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1981). “The relation between forward prices and futures prices”. *Journal of Financial Economics* 9 (4), pp. 321–346.

Cuchiero, C., M. Keller-Ressel, and J. Teichmann (2012). “Polynomial processes and their applications to mathematical finance”. *Finance and Stochastics* 16 (4), pp. 711–740.

Davydov, D. and V. Linetsky (2003). “Pricing Options on Scalar Diffusions: An Eigenfunction Expansion Approach”. *Operations Research* 51 (2), pp. 185–209.

Dillschneider, Y. (2020). “Generalized Transform Analysis for Asset Pricing and Parameter Estimation”. Working paper.

Dotsis, G., D. Psychoyios, and G. Skiadopoulos (2007). “An empirical comparison of continuous-time models of implied volatility indices”. *Journal of Banking & Finance* 31 (12), pp. 3584–3603.

Duan, J.-C. and C.-Y. Yeh (2010). “Jump and volatility risk premiums implied by VIX”. *Journal of Economic Dynamics and Control* 34 (11), pp. 2232–2244.

Duarte, J., C. S. Jones, and J. L. Wang (2020). “Very Noisy Option Prices and Inference Regarding Option Returns”. Working paper.

Duffie, D., D. Filipović, and W. Schachermayer (2003). “Affine Processes and Applications in Finance”. *Annals of Applied Probability* 13 (3), pp. 984–1053.

Duffie, D. and R. Kan (1996). “A Yield-factor Model of Interest Rates”. *Mathematical Finance* 6 (4), pp. 379–406.

Duffie, D., J. Pan, and K. Singleton (2000). “Transform Analysis and Asset Pricing for Affine Jump-Diffusions”. *Econometrica* 68 (6), pp. 1343–1376.

Duffie, D. and K. J. Singleton (1993). “Simulated Moments Estimation of Markov Models of Asset Prices”. *Econometrica* 61 (4), pp. 929–952.

Dunkl, C. F. and Y. Xu (2014). *Orthogonal Polynomials of Several Variables.* 2nd ed. Cambridge University Press.

Durham, G. B. (2006). “Monte Carlo methods for estimating, smoothing, and filtering one- and two-factor stochastic volatility models”. *Journal of Econometrics* 133 (1), pp. 273–305.

— (2013). “Risk-neutral Modeling with Affine and Nonaffine Models”. *Journal of Financial Econometrics* 11 (4), pp. 650–681.

Egloff, D., M. Leippold, and L. Wu (2010). “The Term Structure of Variance Swap Rates and Optimal Variance Swap Investments”. *Journal of Financial and Quantitative Analysis* 45 (5), pp. 1279–1310.

Eraker, B. (2001). “MCMC Analysis of Diffusion Models with Application to Finance”. *Journal of Business & Economic Statistics* 19 (2), pp. 177–191.

— (2004). “Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices”. *Journal of Finance* 59 (3), pp. 1367–1403.

Eraker, B., M. Johannes, and N. Polson (2003). “The Impact of Jumps in Volatility and Returns”. *Journal of Finance* 58 (3), pp. 1269–1300.

Fang, F. and C. W. Oosterlee (2009). “A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions”. *SIAM Journal on Scientific Computing* 31 (2), pp. 826–848.

Feunou, B. and C. Okou (2018). “Risk-Neutral Moment-Based Estimation of Affine Option Pricing Models”. *Journal of Applied Econometrics* 33 (7), pp. 1007–1025.

Filipović, D. and E. Mayerhofer (2009). “Affine Diffusion Processes: Theory and Applications”. Working paper.

Filipović, D., E. Mayerhofer, and P. Schneider (2013). “Density approximations for multivariate affine jump-diffusion processes”. *Journal of Econometrics* 176 (2), pp. 93–111.

Folland, G. B. (1999). *Real Analysis: Modern Techniques and Their Applications.* 2nd ed. Wiley.

Fouque, J.-P. and Y. F. Saporito (2018). “Heston stochastic vol-of-vol model for joint calibration of VIX and S&P 500 options”. *Quantitative Finance* 18 (6), pp. 1003–1016.

Friz, P. and J. Gatheral (2005). “Valuation of volatility derivatives as an inverse problem”. *Quantitative Finance* 5 (6), pp. 531–542.

Fulop, A. and J. Li (2019). “Bayesian estimation of dynamic asset pricing models with informative observations”. *Journal of Econometrics* 209 (1), pp. 114–138.

Gagliardini, P., C. Gourieroux, and E. Renault (2011). “Efficient Derivative Pricing by the Extended Method of Moments”. *Econometrica* 79 (4), pp. 1181–1232.

Gallant, A. R. and G. Tauchen (1996). “Which Moments to Match?” *Econometric Theory* 12 (4), pp. 657–681.

Garcia, R., M.-A. Lewis, S. Pastorello, and É. Renault (2011). “Estimation of objective and risk-neutral distributions based on moments of integrated volatility”. *Journal of Econometrics* 160 (1), pp. 22–32.

Gel’fand, I. M. and G. E. Shilov (1964). *Generalized Functions I: Properties and Operations*. Academic Press.

Glasserman, P. and K.-K. Kim (2009). “Saddlepoint approximations for affine jump-diffusion models”. *Journal of Economic Dynamics and Control* 33 (1), pp. 15–36.

Hall, A. R. and A. Inoue (2003). “The large sample behaviour of the generalized method of moments estimator in misspecified models”. *Journal of Econometrics* 114 (2), pp. 361–394.

Harvey, A. C. and N. Shephard (1996). “Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns”. *Journal of Business & Economic Statistics* 14 (4), pp. 429–434.

Heston, S. L. (1993). “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”. *Review of Financial Studies* 6 (2), pp. 327–343.

Huang, J.-z. and L. Wu (2004). “Specification Analysis of Option Pricing Models Based on Time-Changed Lévy Processes”. *Journal of Finance* 59 (3), pp. 1405–1439.

Jacquier, E., N. G. Polson, and P. E. Rossi (1994). “Bayesian Analysis of Stochastic Volatility Models”. *Journal of Business & Economic Statistics* 12 (4), pp. 371–389.

Jarrow, R. and A. Rudd (1982). “Approximate option valuation for arbitrary stochastic processes”. *Journal of Financial Economics* 10 (3), pp. 347–369.

Jiang, G. J. and R. C. A. Oomen (2007). “Estimating Latent Variables and Jump Diffusion Models Using High-Frequency Data”. *Journal of Financial Econometrics* 5 (1), pp. 1–30.

Jiang, G. J. and Y. S. Tian (2005). “The Model-Free Implied Volatility and Its Information Content”. *Review of Financial Studies* 18 (4), pp. 1305–1342.

— (2007). “Extracting Model-Free Volatility from Option Prices”. *Journal of Derivatives* 14 (3), pp. 35–60.

Jones, C. S. (2003). “The dynamics of stochastic volatility: evidence from underlying and options markets”. *Journal of Econometrics* 116 (1–2), pp. 181–224.

Kloeden, P. and A. Neuenkirch (2013). “Convergence of numerical methods for stochastic differential equations in mathematical finance”. *Recent Developments in Computational Finance*. World Scientific, pp. 49–80.

Kokholm, T. and M. Stisen (2015). “Joint pricing of VIX and SPX options with stochastic volatility and jump models”. *Journal of Risk Finance* 16 (1), pp. 27–48.

Kou, S. G. and H. Wang (2004). “Option Pricing under a Double Exponential Jump Diffusion Model”. *Management Science* 50 (9), pp. 1178–1192.

Kristensen, D. and A. Mele (2011). “Adding and subtracting Black-Scholes: A new approach to approximating derivative prices in continuous-time models”. *Journal of Financial Economics* 102 (2), pp. 390–415.

Kristensen, D. and B. Salanié (2017). “Higher-order properties of approximate estimators”. *Journal of Econometrics* 198 (2), pp. 189–208.

Lewis, A. L. (1998). “Applications of Eigenfunction Expansions in Continuous-Time Finance”. *Mathematical Finance* 8 (4), pp. 349–383.

Lian, G.-H. and S.-P. Zhu (2013). “Pricing VIX options with stochastic volatility and random jumps”. *Decisions in Economics and Finance* 36 (1), pp. 71–88.

Linetsky, V. (2004). “The Spectral Decomposition of the Option Value”. *International Journal of Theoretical and Applied Finance* 7 (3), pp. 337–384.

— (2007). “Spectral Methods in Derivatives Pricing”. *Handbooks in Operations Research and Management Science: Financial Engineering.* Ed. by J. R. Birge and V. Linetsky. Vol. 15. Elsevier, pp. 223–299.

Lord, R., R. Koekkoek, and D. Van Dijk (2010). “A comparison of biased simulation schemes for stochastic volatility models”. *Quantitative Finance* 10 (2), pp. 177–194.

Madan, D. B. and F. Milne (1994). “Contingent claims valued and hedged by pricing and investing in a basis”. *Mathematical Finance* 4 (3), pp. 223–245.

Mencía, J. and E. Sentana (2013). “Valuation of VIX derivatives”. *Journal of Financial Economics* 108 (2), pp. 367–391.

Merton, R. C. (1973). “Theory of Rational Option Pricing”. *Bell Journal of Economics and Management Science* 4 (1), pp. 141–183.

Newey, W. K. and K. D. West (1987). “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”. *Econometrica* 55 (3), pp. 703–708.

Pacati, C., G. Pompa, and R. Renò (2018). “Smiling twice: The Heston++ model”. *Journal of Banking & Finance* 96, pp. 185–206.

Pan, J. (2000). “Integrated Time-Series Analysis of Spot and Option Prices”. Working paper.

— (2002). “The jump-risk premia implicit in options: evidence from an integrated time-series study”. *Journal of Financial Economics* 63 (1), pp. 3–50.

Papanicolaou, A. and R. Sircar (2014). “A regime-switching Heston model for VIX and S&P 500 implied volatilities”. *Quantitative Finance* 14 (10), pp. 1811–1827.

Rudin, W. (1987). *Real and Complex Analysis.* 3rd ed. McGraw-Hill.

— (1991). *Functional Analysis.* 2nd ed. McGraw-Hill.

Ruiz, E. (1994). “Quasi-maximum likelihood estimation of stochastic volatility models”. *Journal of Econometrics* 63 (1), pp. 289–306.

Sandmann, G. and S. J. Koopman (1998). “Estimation of stochastic volatility models via Monte Carlo maximum likelihood”. *Journal of Econometrics* 87 (2), pp. 271–301.

Sepp, A. (2008a). “Pricing Options on Realized Variance in the Heston model with Jumps in Returns and Volatility”. *Journal of Computational Finance* 11 (4), pp. 33–70.

— (2008b). “VIX Option Pricing in a Jump-Diffusion Model”. *Risk*, pp. 84–89.

Shephard, N. G. (1991a). “From Characteristic Function to Distribution Function: A Simple Framework for the Theory”. *Econometric Theory* 7 (4), pp. 519–529.

— (1991b). “Numerical integration rules for multivariate inversions”. *Journal of Statistical Computation and Simulation* 39 (1–2), pp. 37–46.

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

Song, Z. and D. Xiu (2016). “A tale of two option markets: Pricing kernels and volatility risk”. *Journal of Econometrics* 190 (1), pp. 176–196.

Stanton, R. (1997). “A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk”. *Journal of Finance* 52 (5), pp. 1973–2002.

Titchmarsh, E. C. (1975). *Introduction to the Theory of Fourier Integrals.* Oxford University Press.

Wachter, J. A. (2013). “Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?” *Journal of Finance* 68 (3), pp. 987–1035.

Wu, L. (2011). “Variance dynamics: Joint evidence from options and high-frequency returns”. *Journal of Econometrics* 160 (1), pp. 280–287.

Xiu, D. (2014). “Hermite polynomial based expansion of European option prices”. *Journal of Econometrics* 179 (2), pp. 158–177.

Yu, J. (2007). “Closed-form likelihood approximation and estimation of jump-diffusions with an application to the realignment risk of the Chinese Yuan”. *Journal of Econometrics* 141 (2), pp. 1245–1280.

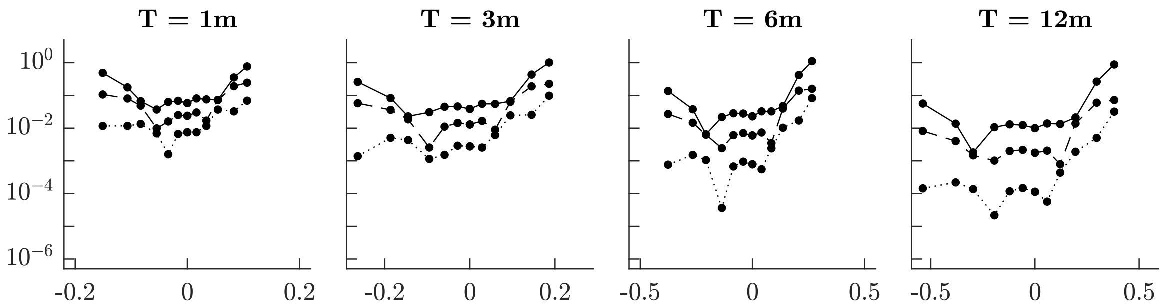

Figure 1: Pricing errors for equity derivatives, SV1 model

This figure shows *relative* pricing errors for equity derivatives within the SV1 model of the polynomial price approximation procedure in lemma 4.2 for different approximation orders p. Each plot shows relative pricing errors $\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)} - V_{i,t})^2]^{1/2}/\mathrm{E}^{\mathbb{P}}[V_{i,t}]$ (y-axis, in log scale) plotted against the option strike K (x-axis) for the given maturity T (in months). Each line style corresponds to a different approximation order: —— $p = 1$, - - - $p = 2$, ······ $p = 4$. $V_{i,t}$ denotes the exact option price for the respective option specification; $V_{i,t,(p)}$ denotes its polynomial approximation according to lemmas 4.2 and 4.3. Equity derivatives are defined as in section 3.3. Further details are provided in section 6.1.

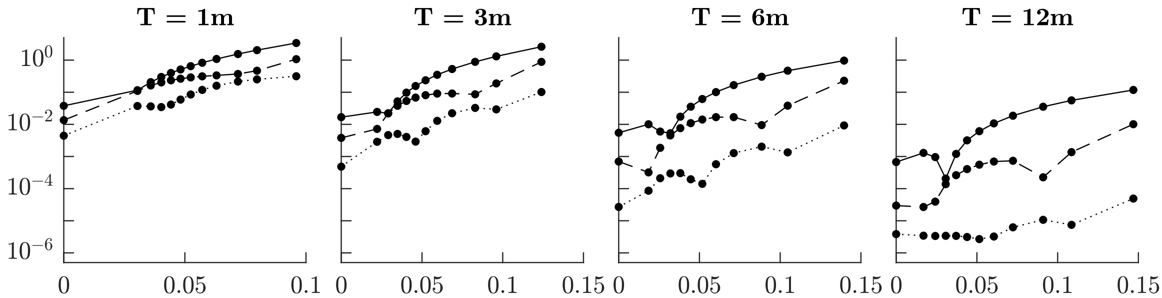

Figure 2: Pricing errors for volatility derivatives, SV1 model

This figure shows *relative* pricing errors for volatility derivatives within the SV1 model of the polynomial price approximation procedure in lemma 4.2 for different approximation orders p. Each plot shows relative pricing errors $\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)} - V_{i,t})^2]^{1/2}/\mathrm{E}^{\mathbb{P}}[V_{i,t}]$ (y-axis, in log scale) plotted against the option strike K (x-axis) for the given maturity T (in months). Each line style corresponds to a different approximation order: —— $p = 1$, - - - $p = 2$, ······ $p = 4$. $V_{i,t}$ denotes the exact option price for the respective option specification; $V_{i,t,(p)}$ denotes its polynomial approximation according to lemmas 4.2 and 4.3. Volatility derivatives are defined as in section 3.4. Further details are provided in section 6.1.

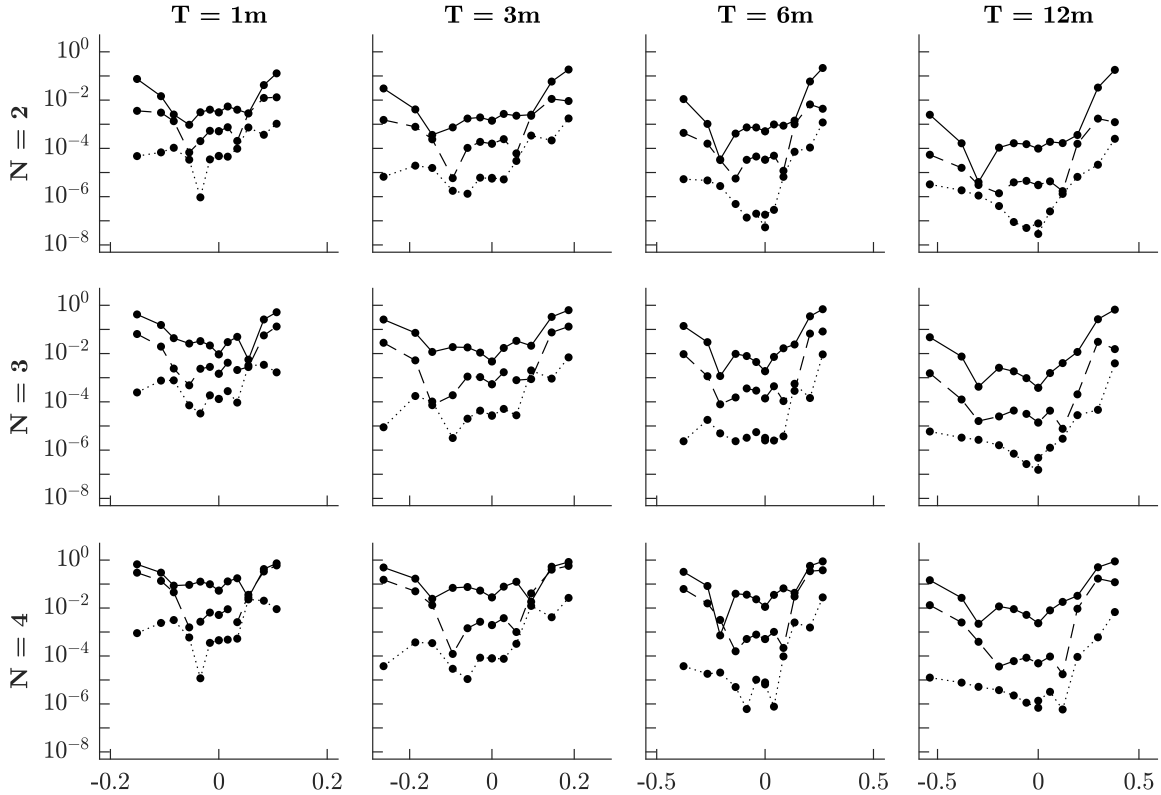

Figure 3: Moment errors for equity derivatives, SV1 model

This figure shows *relative* moment errors for equity derivatives within the SV1 model of the polynomial moment approximation procedure in proposition 4.2 for different moment orders N and approximation orders p. Each plot shows relative moment errors $|\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)})^N] - \mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]|/\mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]$ (y-axis, in log scale) plotted against the option strike K (x-axis) for the given maturity T (in months) and moment order N. Each line style corresponds to a different approximation order: —— $p = 1$, - - - $p = 2$, ⋯⋯ $p = 4$. $V_{i,t}$ denotes the exact option price for the respective option specification; $V_{i,t,(p)}$ denotes its polynomial approximation according to lemmas 4.2 and 4.3. Equity derivatives are defined as in section 3.3. Further details are provided in section 6.1.

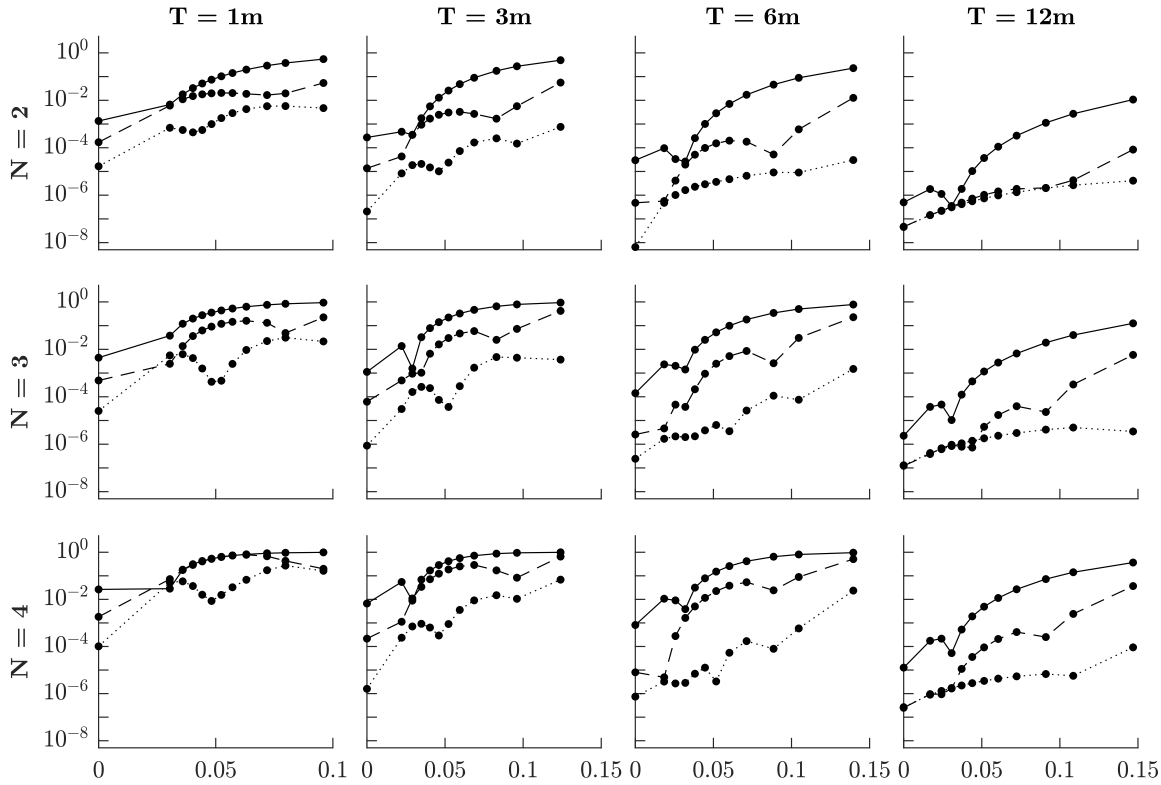

Figure 4: Moment errors for volatility derivatives, SV1 model

This figure shows *relative* moment errors for volatility derivatives within the SV1 model of the polynomial moment approximation procedure in proposition 4.2 for different moment orders N and approximation orders p. Each plot shows relative moment errors $|\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)})^N] - \mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]|/\mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]$ (y-axis, in log scale) plotted against the option strike K (x-axis) for the given maturity T (in months) and moment order N. Each line style corresponds to a different approximation order: —— $p = 1$, - - - $p = 2$, ⋯⋯ $p = 4$. $V_{i,t}$ denotes the exact option price for the respective option specification; $V_{i,t,(p)}$ denotes its polynomial approximation according to lemmas 4.2 and 4.3. Volatility derivatives are defined as in section 3.4. Further details are provided in section 6.1.

	$\mathbb{P}$				$\mathbb{Q}$		
	SV1	SV1J	SV2J		SV1	SV1J	SV2J
$b_0^{\mathbb{P}}$	0.048	0.048	0.048	$b_0^{\mathbb{Q}}$	0.038	0.038	0.038
$b_1^{\mathbb{P}}$	1.190	1.190	1.190	$b_1^{\mathbb{Q}}$	−0.310	−0.310	−0.310
$\kappa_1^{\mathbb{P}}$	2.5	2.5	2.5	$\kappa_1^{\mathbb{Q}}$	2	2	2
$\theta_1^{\mathbb{P}}$	0.04	0.04		$\theta_1^{\mathbb{Q}}$	0.05	0.05	
ς_1	0.3	0.3	0.3				
$\kappa_2^{\mathbb{P}}$			0.625	$\kappa_2^{\mathbb{Q}}$			0.5
$\theta_2^{\mathbb{P}}$			0.04	$\theta_2^{\mathbb{Q}}$			0.05
ς_2			0.2				
ρ_1	−0.8	−0.8	−0.8				
$\lambda_0^{\mathbb{P}}$	0	2	2	$\lambda_0^{\mathbb{Q}}$	0	2	2
$\lambda_1^{\mathbb{P}}$	0	10	10	$\lambda_1^{\mathbb{Q}}$	0	10	10
$\mu_{0,J}^{\mathbb{P}}$		−0.01	−0.01	$\mu_{0,J}^{\mathbb{Q}}$		−0.02	−0.02
$\sigma_{0,J}^{\mathbb{P}}$		0.04	0.04	$\sigma_{0,J}^{\mathbb{Q}}$		0.04	0.04
$\mu_{1,J}^{\mathbb{P}}$		0.005	0.005	$\mu_{1,J}^{\mathbb{Q}}$		0.01	0.01

Table 1: Model parameters

This table specifies the parameters used for the numerical analyses in section 6. The SV1 and SV1J models are based on the dynamics (2.2). The SV2J model is based on the dynamics (2.4). Jump size distributions are specified such that under $\mathbb{M}$, $J_{11,X,t}$ follows a Gaussian distribution with mean $\mu_{0,J}^{\mathbb{M}}$ and standard deviation $\sigma_{0,J}^{\mathbb{M}}$, while $J_{21,X,t}$ follows an exponential distribution with mean $\mu_{1,J}^{\mathbb{M}}$.

	SV1J			SV2J		
	$p=1$	$p=2$	$p=4$	$p=1$	$p=2$	$p=4$
(a) Monte Carlo simulation						
	4.06×10^{-2}	1.16×10^{-2}	1.72×10^{-3}	1.16×10^{-1}	5.55×10^{-2}	1.71×10^{-2}
	(9.34×10^{-1})	(2.00×10^{-1})	(7.04×10^{-2})	(1.43×10^{0})	(3.37×10^{-1})	(1.62×10^{-1})
(b) Density approximation						
	4.01×10^{-2}	1.14×10^{-2}	1.62×10^{-3}	1.14×10^{-1}	5.48×10^{-2}	1.92×10^{-2}
	(9.15×10^{-1})	(1.93×10^{-1})	(7.11×10^{-2})	(1.42×10^{0})	(3.93×10^{-1})	(1.69×10^{-1})
(c) Closed-form benchmark						
	4.66×10^{-2}	1.45×10^{-2}	2.66×10^{-3}			
	(1.12×10^{0})	(2.45×10^{-1})	(9.80×10^{-2})			

Table 2: Pricing errors for equity derivatives, SV1J and SV2J models

This table reports aggregate *relative* pricing errors for equity derivatives within the SV1J and SV2J models of the polynomial price approximation procedure in lemma 4.2 for different approximation orders p. Panels (a) and (b) report aggregate pricing errors determined by Monte Carlo simulation and density approximation, respectively. Panel (c) reports aggregate pricing errors within the SV1 model. Each entry reports the median (maximum in parenthesis) of relative pricing errors $\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)} - V_{i,t})^2]^{1/2}/\mathrm{E}^{\mathbb{P}}[V_{i,t}]$ taken over option strikes K and maturities T. $V_{i,t}$ denotes the exact option price for the respective option specification; $V_{i,t,(p)}$ denotes its polynomial approximation according to lemmas 4.2 and 4.3. Equity derivatives are defined as in section 3.3. Further details are provided in section 6.1.

	SV1J			SV2J		
	$p = 1$	$p = 2$	$p = 4$	$p = 1$	$p = 2$	$p = 4$
(a) Monte Carlo simulation						
	1.22×10^{-1}	1.54×10^{-2}	1.19×10^{-3}	5.56×10^{-1}	2.41×10^{-1}	4.72×10^{-2}
	(4.20×10^{0})	(1.61×10^{0})	(4.22×10^{-1})	(4.19×10^{0})	(1.34×10^{0})	(5.95×10^{-1})
(b) Density approximation						
	1.18×10^{-1}	1.51×10^{-2}	9.78×10^{-4}	5.39×10^{-1}	2.44×10^{-1}	5.01×10^{-2}
	(4.16×10^{0})	(1.65×10^{0})	(3.93×10^{-1})	(4.28×10^{0})	(1.34×10^{0})	(8.11×10^{-1})
(c) Closed-form benchmark						
	9.91×10^{-2}	1.68×10^{-2}	2.45×10^{-3}			
	(3.41×10^{0})	(1.07×10^{0})	(3.15×10^{-1})			

Table 3: Pricing errors for volatility derivatives, SV1J and SV2J models

This table reports aggregate *relative* pricing errors for volatility derivatives within the SV1J and SV2J models of the polynomial price approximation procedure in lemma 4.2 for different approximation orders p. Panels (a) and (b) report aggregate pricing errors determined by Monte Carlo simulation and density approximation, respectively. Panel (c) reports aggregate pricing errors within the SV1 model. Each entry reports the median (maximum in parenthesis) of relative pricing errors $\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)} - V_{i,t})^2]^{1/2}/\mathrm{E}^{\mathbb{P}}[V_{i,t}]$ taken over option strikes K and maturities T. $V_{i,t}$ denotes the exact option price for the respective option specification; $V_{i,t,(p)}$ denotes its polynomial approximation according to lemmas 4.2 and 4.3. Volatility derivatives are defined as in section 3.4. Further details are provided in section 6.1.

	SV1J			SV2J		
	$p=1$	$p=2$	$p=4$	$p=1$	$p=2$	$p=4$
(a) Monte Carlo simulation						
$N=2$	1.33×10^{-3}	1.92×10^{-3}	1.94×10^{-3}	1.51×10^{-2}	1.80×10^{-2}	1.71×10^{-2}
	(2.05×10^{-1})	(3.86×10^{-2})	(3.70×10^{-2})	(2.50×10^{-1})	(2.84×10^{-2})	(2.30×10^{-2})
$N=3$	1.29×10^{-2}	3.18×10^{-3}	3.06×10^{-3}	7.78×10^{-2}	2.26×10^{-2}	1.63×10^{-2}
	(6.69×10^{-1})	(7.13×10^{-2})	(7.26×10^{-2})	(7.03×10^{-1})	(2.19×10^{-1})	(4.15×10^{-2})
$N=4$	4.54×10^{-2}	7.37×10^{-3}	1.13×10^{-2}	2.11×10^{-1}	4.05×10^{-2}	1.94×10^{-2}
	(8.77×10^{-1})	(3.72×10^{-1})	(1.05×10^{-1})	(8.77×10^{-1})	(1.11×10^{0})	(2.86×10^{-1})
(b) Density approximation						
$N=2$	1.28×10^{-3}	1.06×10^{-4}	1.88×10^{-6}	8.14×10^{-3}	1.83×10^{-3}	3.70×10^{-4}
	(1.76×10^{-1})	(9.98×10^{-3})	(1.21×10^{-3})	(2.41×10^{-1})	(2.38×10^{-2})	(3.71×10^{-3})
$N=3$	1.29×10^{-2}	5.99×10^{-4}	1.43×10^{-5}	6.27×10^{-2}	5.42×10^{-3}	1.01×10^{-3}
	(6.41×10^{-1})	(1.15×10^{-1})	(7.24×10^{-3})	(7.04×10^{-1})	(1.95×10^{-1})	(1.53×10^{-2})
$N=4$	4.59×10^{-2}	3.18×10^{-3}	5.10×10^{-5}	2.00×10^{-1}	2.77×10^{-2}	4.90×10^{-3}
	(8.62×10^{-1})	(5.10×10^{-1})	(2.43×10^{-2})	(8.83×10^{-1})	(8.95×10^{-1})	(2.78×10^{-1})
(c) Closed-form benchmark						
$N=2$	1.82×10^{-3}	1.71×10^{-4}	6.36×10^{-6}			
	(2.14×10^{-1})	(1.30×10^{-2})	(1.74×10^{-3})			
$N=3$	1.75×10^{-2}	9.67×10^{-4}	2.79×10^{-5}			
	(6.88×10^{-1})	(1.32×10^{-1})	(9.22×10^{-3})			
$N=4$	5.97×10^{-2}	3.49×10^{-3}	8.30×10^{-5}			
	(8.83×10^{-1})	(5.99×10^{-1})	(2.78×10^{-2})			

Table 4: Moment errors for equity derivatives, SV1J and SV2J models

This table reports aggregate *relative* moment errors for equity derivatives within the SV1J and SV2J models of the polynomial moment approximation procedure in proposition 4.2 for different moment orders N and approximation orders p. Panels (a) and (b) report aggregate moment errors determined by Monte Carlo simulation and density approximation, respectively. Panel (c) reports aggregate moment errors within the SV1 model. Each entry reports the median (maximum in parenthesis) of relative moment errors $|\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)})^N] - \mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]|/\mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]$ taken over option strikes K and maturities T. $V_{i,t}$ denotes the exact option price for the respective option specification; $V_{i,t,(p)}$ denotes its polynomial approximation according to lemmas 4.2 and 4.3. Equity derivatives are defined as in section 3.3. Further details are provided in section 6.1.

	SV1J			SV2J		
	$p = 1$	$p = 2$	$p = 4$	$p = 1$	$p = 2$	$p = 4$
(a) Monte Carlo simulation						
$N = 2$	$1.56{\times}10^{-2}$	$5.69{\times}10^{-3}$	$4.82{\times}10^{-3}$	$7.15{\times}10^{-2}$	$2.15{\times}10^{-2}$	$1.71{\times}10^{-2}$
	$(6.59{\times}10^{-1})$	$(1.85{\times}10^{-1})$	$(1.02{\times}10^{-1})$	$(5.95{\times}10^{-1})$	$(1.24{\times}10^{-1})$	$(8.78{\times}10^{-2})$
$N = 3$	$1.31{\times}10^{-1}$	$9.90{\times}10^{-3}$	$1.93{\times}10^{-2}$	$3.66{\times}10^{-1}$	$5.07{\times}10^{-2}$	$1.59{\times}10^{-2}$
	$(9.65{\times}10^{-1})$	$(5.85{\times}10^{-1})$	$(1.33{\times}10^{-1})$	$(9.54{\times}10^{-1})$	$(4.23{\times}10^{-1})$	$(1.23{\times}10^{-1})$
$N = 4$	$2.56{\times}10^{-1}$	$3.28{\times}10^{-2}$	$3.05{\times}10^{-2}$	$5.37{\times}10^{-1}$	$3.00{\times}10^{-1}$	$4.99{\times}10^{-2}$
	$(9.96{\times}10^{-1})$	$(7.83{\times}10^{-1})$	$(1.39{\times}10^{-1})$	$(9.94{\times}10^{-1})$	$(1.39{\times}10^{0})$	$(2.07{\times}10^{0})$
(b) Density approximation						
$N = 2$	$1.02{\times}10^{-2}$	$1.56{\times}10^{-4}$	$6.02{\times}10^{-6}$	$7.36{\times}10^{-2}$	$1.27{\times}10^{-2}$	$9.48{\times}10^{-4}$
	$(6.22{\times}10^{-1})$	$(9.77{\times}10^{-2})$	$(6.71{\times}10^{-3})$	$(5.88{\times}10^{-1})$	$(5.82{\times}10^{-2})$	$(1.90{\times}10^{-2})$
$N = 3$	$1.12{\times}10^{-1}$	$6.15{\times}10^{-3}$	$7.20{\times}10^{-5}$	$3.58{\times}10^{-1}$	$6.03{\times}10^{-2}$	$2.92{\times}10^{-3}$
	$(9.60{\times}10^{-1})$	$(5.26{\times}10^{-1})$	$(2.93{\times}10^{-2})$	$(9.52{\times}10^{-1})$	$(3.41{\times}10^{-1})$	$(3.22{\times}10^{-2})$
$N = 4$	$2.23{\times}10^{-1}$	$3.26{\times}10^{-2}$	$4.30{\times}10^{-4}$	$5.27{\times}10^{-1}$	$3.75{\times}10^{-1}$	$3.01{\times}10^{-2}$
	$(9.96{\times}10^{-1})$	$(7.73{\times}10^{-1})$	$(2.27{\times}10^{-1})$	$(9.95{\times}10^{-1})$	$(1.30{\times}10^{0})$	$(1.61{\times}10^{0})$
(c) Closed-form benchmark						
$N = 2$	$6.10{\times}10^{-3}$	$1.90{\times}10^{-4}$	$9.07{\times}10^{-6}$			
	$(5.46{\times}10^{-1})$	$(5.60{\times}10^{-2})$	$(5.73{\times}10^{-3})$			
$N = 3$	$6.53{\times}10^{-2}$	$2.54{\times}10^{-3}$	$3.39{\times}10^{-5}$			
	$(9.35{\times}10^{-1})$	$(4.18{\times}10^{-1})$	$(3.11{\times}10^{-2})$			
$N = 4$	$1.60{\times}10^{-1}$	$3.53{\times}10^{-2}$	$2.02{\times}10^{-4}$			
	$(9.93{\times}10^{-1})$	$(7.81{\times}10^{-1})$	$(2.71{\times}10^{-1})$			

Table 5: Moment errors for volatility derivatives, SV1J and SV2J models

This table reports aggregate *relative* moment errors for volatility derivatives within the SV1J and SV2J models of the polynomial moment approximation procedure in proposition 4.2 for different moment orders N and approximation orders p. Panels (a) and (b) report aggregate moment errors determined by Monte Carlo simulation and density approximation, respectively. Panel (c) reports aggregate moment errors within the SV1 model. Each entry reports the median (maximum in parenthesis) of relative moment errors $|\mathrm{E}^{\mathbb{P}}[(V_{i,t,(p)})^N] - \mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]|/\mathrm{E}^{\mathbb{P}}[(V_{i,t})^N]$ taken over option strikes K and maturities T. $V_{i,t}$ denotes the exact option price for the respective option specification; $V_{i,t,(p)}$ denotes its polynomial approximation according to lemmas 4.2 and 4.3. Volatility derivatives are defined as in section 3.4. Further details are provided in section 6.1.