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

Yannick Dillschneider* and Raimond Maurer
Goethe University Frankfurt

March 8, 2020
(Preliminary and incomplete)

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-type estimation procedure. Invoking new results from generalized transform analysis, we derive analytically tractable expressions for exact moments and devise a computationally efficient approximation procedure. We exemplify our methodology by deriving moment conditions that jointly incorporate stock returns as well as prices of equity and volatility options.

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

Keywords: generalized method of moments; generalized transform analysis; stochastic volatility models; option pricing

* 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: 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. 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. The latent character of certain state variables can significantly complicate estimation of model parameters. Possible consequences range from inaccurate parameter estimates to even the failure to properly identify important parameters. Moreover, 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 additional information about risk premia. By contributing accurate information about otherwise poorly identified or unidentified parameters, including derivatives prices into the estimation procedure can be expected to yield substantial statistical efficiency gains and overcome various identification issues.

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. 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 option-implied volatility measure. 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. Bardgett et al. (2019) further conclude that volatility index derivatives contain incremental information about stock return volatility that is not already spanned by equity derivatives. Moreover, Song and Xiu (2016) find evidence that standard model specifications capture risk premia reflected by equity options well, but fail to adequately reflect risk premia embedded in volatility index 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 mostly either rely on computationally intensive techniques, such as extensive simulations and large-dimensional optimizations, or impose strong and somewhat arbitrary assumptions on measurement errors of derivatives prices, or both. This paper develops a novel and unified approach to incorporate a broad class of derivatives into a GMM estimation procedure, which is both computationally fast and

---

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

2The most prominent volatility index is probably CBOE’s VIX, which is derived from the prices of S&P 500 options.
compatible with realistic measurement error models. 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. Deriving exact and approximate moment conditions using our methodology, we devise a GMM estimation procedure that incorporates moments involving polynomials of derivatives prices.

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 the results in the literature, we derive the required standard transforms of the state vector, which rely on the solution of a system of generalized Riccati equations. In practical applications, these can be solved numerically in an efficient way by using vectorization techniques.

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) 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.

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. 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

3These 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.
a limited amount of the information available in derivatives markets. Yet, most of the much finer information contained in the cross section of option prices is not 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. In a Bayesian framework, Eraker (2004) achieve this by Markov chain Monte Carlo methods, while Christoffersen et al. (2010) 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) 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. Equivalent assumptions in a maximum likelihood framework allow Aït-Sahalia and Kimmel (2007) to obtain (approximate) transition densities involving option prices. The latter approach eliminates the huge computational burden embedded in explicit inversions, but maintains the 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. 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 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)). A vast literature exists also on the use of various approximation techniques for the purpose of option pricing. 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. Finally, section 6 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 framework. 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.

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 \(X \subset \mathbb{R}^{nx}\) and is defined on the real-world filtered probability space \((\Omega, \Sigma, \mathcal{F}, \mathbb{P})\), in which the sample space \(\Omega\) is equipped with a \(\sigma\)-algebra \(\Sigma\) 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 \(X = \mathbb{R} \times \mathcal{Z}\) for \(\mathcal{Z} \subset \mathbb{R}^{nz}\), is composed of the stock price \(S_t\) and the \(nz\)-element vector \(Z_t\) containing additional state variables. In

---

7See also Kristensen and Salanié (2017) for other types of approximations and potential improvements of the resulting estimation approaches.

8Existing 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)), and saddlepoint approximation (e.g., Glasserman and Kim (2009)), among others. Relatedly, Aït-Sahalia et al. (2019) propose a closed-form expansion method for option-implied volatilities.
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

\[
\begin{align*}
    d\log S_t &= \mu^{\mathbb{M}}(Z_{t-}) dt + \sigma^{\mathbb{M}}(Z_{t-}) dW_t^\mathbb{M} + J_{S,t} dN_t, \\
    dZ_t &= \mu^{\mathbb{M}}_{Z}(Z_{t-}) dt + \sigma^{\mathbb{M}}_{Z}(Z_{t-}) dW_t^\mathbb{M} + J_{Z,t} dN_t,
\end{align*}
\]

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 \( \mu^{\mathbb{M}} = [\mu^{\mathbb{M}}_S; \mu^{\mathbb{M}}_Z] \), \( \sigma^{\mathbb{M}} = [\sigma_S; \sigma_Z] \), and \( J_{X,t} = [J_{S,t}; J_{Z,t}] \), we can write the dynamics in equation (2.1) in the general form

\[
    dX_t = \mu^{\mathbb{M}}_X(Z_{t-}) dt + \sigma^{\mathbb{M}}_X(Z_{t-}) dW_t^\mathbb{M} + J_{X,t} dN_t .
\]

Analogous to Duffie et al. (2000), we impose the following affine restrictions on the drift vector \( \mu^{\mathbb{M}}_X \), instantaneous diffusive covariance matrix \( \Omega_x = \sigma^\top_x \sigma_x \), jump intensity vector \( \Lambda^{\mathbb{M}} \), and joint distribution \( \nu^\mathbb{M} \) of jump sizes \( J_{X,t} \):

- \( \mu^\mathbb{M}_x(z) = A^\mathbb{M}_{\mu,X} + B^\mathbb{M}_{\mu,X} z \) with \( A^\mathbb{M}_{\mu,X} \in \mathbb{R}^{n x} \) and \( B^\mathbb{M}_{\mu,X} \in \mathbb{R}^{n x \times n z} \),
- \( \text{vec}[\Omega(z)] = A_{\Omega,X} + B_{\Omega,X} z \) with \( A_{\Omega,X} \in \mathbb{R}^{n_x} \) and \( B_{\Omega,X} \in \mathbb{R}^{n_x \times n_z} \),
- \( \Lambda^{\mathbb{M}}(z) = A^\mathbb{M}_\Lambda + B^\mathbb{M}_\Lambda z \) with \( A^\mathbb{M}_\Lambda \in \mathbb{R}^{n_j} \) and \( B^\mathbb{M}_\Lambda \in \mathbb{R}^{n_j \times n_z} \), 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:

- \( A^\mathbb{Q}_{\mu,S} = r - q - \frac{1}{2}A_{\Omega,S} - (\Phi^\mathbb{Q}_{\mu}([1; 0]) - \nu^\mathbb{Q})^\top A^\mathbb{Q}_\Lambda \) and
- \( B^\mathbb{Q}_{\mu,S} = -\frac{1}{2}B_{\Omega,S} - (\Phi^\mathbb{Q}_{\mu}([1; 0]) - \nu^\mathbb{Q})^\top B^\mathbb{Q}_\Lambda \)

for instantaneous diffusive stock price variance \( \Omega_S(z) = A_{\Omega,S} + B_{\Omega,S} z \), denoting by \( \nu \) the vector of ones and by \( \Phi^\mathbb{Q}_{\mu} \) the vector of marginal jump transforms of jump sizes under the law \( \nu^\mathbb{M} \). Specifically, the \( i \)-th element of \( \Phi^\mathbb{Q}_{\mu} \) for \( \omega \in \mathbb{C}^{n x} \) is given by \( \int \exp(\omega \cdot J_{i,X}) d\nu^\mathbb{M} \). Under standard integrability conditions, this assures that \( \exp((q-r)t) S_t \) is a 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.\(^9\)

### 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.

---

\(^9\)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.
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

$$
\begin{align*}
    d\log S_t &= (\theta_0^M + \theta_1^M Z_{1,t-}) dt + Z_{1,t-}^{1/2} d\tilde{W}_{1,t}^M + J_{11,X,t} dN_{1,t} \\
    dZ_{1,t} &= \kappa_1^M (\theta_1^M - Z_{1,t-}) dt + \varsigma_1 Z_{1,t-}^{1/2} d\tilde{W}_{2,t}^M + J_{21,X,t} dN_{1,t},
\end{align*}
$$

(2.2a)

(2.2b)

where $d\tilde{W}_{1,t}^M = (1 - \rho_I^2)^{1/2} dW_{1,t}^M + \rho_I dW_{2,t}^M$ with constant correlation $\rho_I$ and $\Lambda_M(z) = \lambda_0^M + \lambda_1^M z_1$. Equation (2.2) takes into account that jumps in stock prices and volatility tend to occur simultaneously within this model is the instantaneous diffusion variance 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

$$
\begin{align*}
    d\log S_t &= (\theta_0^M + \theta_1^M Z_{1,t-} + \theta_2^M Z_{2,t-}) dt + (Z_{1,t-} + Z_{2,t-})^{1/2} d\tilde{W}_{1,t}^M + J_{11,X,t} dN_{1,t} \\
    dZ_{1,t} &= \kappa_1^M (\theta_1^M - Z_{1,t-}) dt + \varsigma_1 Z_{1,t-}^{1/2} dW_{2,t}^M + J_{21,X,t} dN_{1,t} \\
    dZ_{2,t} &= \kappa_2^M (\theta_2^M - Z_{2,t-}) dt + \varsigma_2 Z_{2,t-}^{1/2} dW_{3,t}^M + J_{21,X,t} dN_{1,t},
\end{align*}
$$

(2.3a)

(2.3b)

(2.3c)

where $d\tilde{W}_{1,t}^M = (1 - \rho_I^2)^{1/2} dW_{1,t}^M + \rho_I dW_{2,t}^M$ with stochastic correlation $\rho_I = \rho_I Z_{1,t-}^{1/2}/(Z_{1,t} + Z_{2,t})^{1/2}$. Moreover, jump intensities are given by $\Lambda_M(z) = \lambda_0^M + \lambda_1^M z_1 + \lambda_2^M z_2$. 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.

2.2.2 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 variance $Z_{1,t} + Z_{2,t}$ of the stock price process. The idea 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 $\theta_1^M$ 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

$$
d\log S_t = (b_{10}^M + b_{1}^M Z_{1,t-} + b_{2}^M Z_{2,t-}) dt + Z_{1,t-}^{1/2} d\tilde{W}_{1,t}^M + J_{11,t} dN_{1,t}
$$

$$
dZ_{1,t} = \kappa_1^M (\psi_{11}^M - Z_{1,t-}) dt + \varsigma_1 Z_{1,t-}^{1/2} dW_{2,t}^M + J_{21,t} dN_{1,t}
$$

$$
dZ_{2,t} = \kappa_2^M (\psi_{12}^M - Z_{2,t-}) dt + \varsigma_2 Z_{2,t-}^{1/2} dW_{3,t}^M,
$$

where $d\tilde{W}_{1,t}^M = (1 - \rho_1^2)^{1/2} dW_{1,t}^M + \rho_1 dW_{2,t}^M$ with constant correlation $\rho_1$ and $\Lambda^M(z) = \lambda_0^M + \lambda_2^M z_1$.

### 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. (2015)). 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

$$
d\log S_t = (b_{10}^M + b_{1}^M Z_{1,t-} + b_{2}^M Z_{2,t-}) dt + Z_{1,t-}^{1/2} d\tilde{W}_{1,t}^M + J_{11,t} dN_{1,t}
$$

$$
dZ_{1,t} = \kappa_1^M (\psi_{11}^M - Z_{1,t-}) dt + \varsigma_1 Z_{1,t-}^{1/2} dW_{2,t}^M + J_{21,t} dN_{1,t}
$$

$$
dZ_{2,t} = \kappa_2^M (\psi_{12}^M - Z_{2,t-}) dt + \varsigma_2 Z_{2,t-}^{1/2} dW_{3,t}^M + J_{31,t} dN_{1,t},
$$

where $d\tilde{W}_{1,t}^M = (1 - \rho_1^2)^{1/2} dW_{1,t}^M + \rho_1 dW_{2,t}^M$ with constant correlation $\rho_1$ and $\Lambda^M(z) = \lambda_0^M + \lambda_2^M z_2$.

### 2.3 Standard transform analysis

The methodology developed in this paper relies on the tractability of certain classes of moments of the state vector. Well-established variants of 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 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 $\tau \in \mathbb{R}_+^n$, corresponding to a non-decreasing sequence of time points $\tau_i$ such that $\tau_{i+1} \geq \tau_i \geq 0$ with the convention that $\tau_0 = 0$, define the stacked vectors $\tilde{S}_{t+\tau} = [\log S_{t+\tau_1}; \cdots; \log S_{t+\tau_n}]$ and $Z_{t+\tau} = [Z_{t+\tau_1}; \cdots; Z_{t+\tau_n}]$. From these, moreover define the vector $\tilde{X}_{t+\tau} = [\tilde{S}_{t+\tau}; Z_{t+\tau}]$. Economically, the elements of $\tilde{S}_{t+\tau}$ thus correspond to log returns between consecutive time points in $\tau$. In what follows, we state expressions for certain moments of $\tilde{X}_{t+\tau}$.

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. 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 $\tilde{X}_{t+\tau}$ in the following proposition. It should be noted that under the dynamics specified by equation (2.1), the conditional exponential moments of $\tilde{X}_{t+\tau}$ 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 $\tilde{X}_{t+\tau}$ by a limiting procedure.
Proposition 2.1. Consider an argument $\omega \in \mathbb{C}^{n \times \bar{n}}$.

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

$$
\Phi^M(\omega; \tilde{\tau}, 0, Z_t) = E^M[\exp(\omega \cdot \tilde{X}_{t+\tau}) | \mathcal{F}_t] \\
= \exp(A_{\Phi}(\omega; \tilde{\tau}, 0) + B_{\Phi}(\omega; \tilde{\tau}, 0) \cdot Z_t)
$$

(2.6)

with coefficients $A_{\Phi}(\omega; \tilde{\tau}, 0) \in \mathbb{C}$ and $B_{\Phi}(\omega; \tilde{\tau}, 0) \in \mathbb{C}^{n \times \bar{n}}$ given in equation (A.13).

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

$$
\Phi^M(\omega; \tilde{\tau}, \infty) = E^M[\exp(\omega \cdot \tilde{X}_{t+\tau})] \\
= \exp(A_{\Phi}(\omega; \tilde{\tau}, \infty))
$$

(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 Faa di Bruno formula, 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. 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 $\tilde{X}_{t+\tau}$ in the following proposition. In essence, under the imposed regularity conditions, the respective moment expressions are formed by differentiation the expressions obtained in proposition 2.1 so that $\Phi^M(\alpha) = \partial_\alpha^2 \Phi^M$ holds. It is therefore not surprising that the resulting unconditional pl-linear moments of $\tilde{X}_{t+\tau}$ 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 $\tilde{X}_{t+\tau}$ by a limiting argument.

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

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

$$
\Phi^M(\alpha; \tilde{\tau}, 0, Z_t) = E^M[\exp(\omega \cdot \tilde{X}_{t+\tau}) (\tilde{X}_{t+\tau})^\alpha | \mathcal{F}_t] \\
= \Phi^M(\omega; \tilde{\tau}, 0) \sum_{Q(\alpha)} M_{k,\ell}^{Q(\alpha)} (A^{(\ell)}_{\Phi}(\omega; \tilde{\tau}, 0) + B^{(\ell)}_{\Phi}(\omega; \tilde{\tau}, 0) \cdot Z_t)^k
$$

(2.8)

with coefficients $A^{(\beta)}_{\Phi}(\omega; \tilde{\tau}, 0) \in \mathbb{C}$ and $B^{(\beta)}_{\Phi}(\omega; \tilde{\tau}, 0) \in \mathbb{C}^{n \times \bar{n}}$ for $\beta \leq \alpha$ given in equation (A.23).

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

$$
\Phi^M(\alpha; \tilde{\tau}, \infty) = E^M[\exp(\omega \cdot \tilde{X}_{t+\tau}) (\tilde{X}_{t+\tau})^\alpha] \\
= \Phi^M(\omega; \tilde{\tau}, \infty) \sum_{Q(\alpha)} M_{k,\ell}^{Q(\alpha)} (A^{(\ell)}_{\Phi}(\omega; \tilde{\tau}, \infty))^k
$$

(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), $Q(\alpha) = \bigcup_{|\beta| \leq |\alpha|} Q(\alpha, \beta)$ is a disjoint union and each $Q(\alpha, \beta)$ is a set of multi-indices $k$ and $\ell$, defined in equation (A.2). Moreover, $M_{k,\ell}^{Q(\alpha)}$ denotes the associated generalized 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 Dillschneider (2020).

3 Transform-based derivatives pricing

The results from generalized transform analysis can be used to determine the prices of a large class 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 generalized transform analysis, whose foundations are briefly reviewed in section 3.1. 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.

3.1 Basic Schwartz distribution theory

In order to determine a broad class of general moments, some additional structure is required. Drawing on the generalized transform analysis introduced in Chen and Joslin (2012) and further developed in Dillschneider (2020), we briefly introduce the necessary terminology and notation. Further details are given in appendix B.1.

By \( S(\mathbb{R}^m) \), referred to as the Schwartz space, we denote the space of rapidly decaying smooth functions, regularly succinctly referred to as Schwartz functions. The associated continuous dual space \( S^*(\mathbb{R}^m) \), whose elements are called tempered distributions, contains all continuous linear functionals on \( S(\mathbb{R}^m) \). To denote the action of a tempered distribution \( g \in S^*(\mathbb{R}^m) \) on a Schwartz function \( f \in S(\mathbb{R}^m) \), we use the duality pairing notation \( \langle g(y), f(y) \rangle \). A sufficiently regular ordinary function \( g \) identifies with a tempered distribution via integration, \( \langle g(y), f(y) \rangle = \int_{\mathbb{R}^m} g(y) f(y) \, dy \). Using a regularization approach, this notion can be extended to a larger class of functions for which the ordinary integrals do not exist. Any Schwartz function \( f \in S(\mathbb{R}^m) \) has a Fourier transform \( \hat{f} \in \mathcal{F}(\mathbb{R}^m) \), which allows to define the Fourier transform \( \hat{g} = \mathcal{F} g \in S^*(\mathbb{R}^m) \) of the tempered distribution \( g \in S^*(\mathbb{R}^m) \) via the requirement that \( \langle \hat{g}(y), \hat{f}(y) \rangle = \langle g(y), f(y) \rangle \) holds for all \( f \in S(\mathbb{R}^m) \).

We extend the definitions above to subsets \( \mathcal{Y} \subset \mathbb{R}^m \) as follows. The space \( S(\mathcal{Y}) \) consists of all functions \( f \) such that there exists some \( \hat{f} \in S(\mathbb{R}^m) \) coinciding with \( f \) on \( \mathcal{Y} \). Likewise, the dual space \( S^*(\mathcal{Y}) \) consists of those \( g \in S^*(\mathbb{R}^m) \) whose support is contained in \( \mathcal{Y} \). Consequently, we can define \( \langle g(y), f(y) \rangle = \int_{\mathbb{R}^m} g(y) f(y) \, dy \) for \( g \in S(\mathcal{Y}) \) and \( f \in S^*(\mathcal{Y}) \), where the choice of \( \hat{f} \in S(\mathbb{R}^m) \) 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 be paid immediately at inception of the contract. Other empirically relevant features, such as futures-style margining, can be incorporated by minor modifications.\(^{10}\)

\(^{10}\)E.g., for futures-style margining, the results hold with the risk-neutral transform \( \Phi^Q \) replacing the pricing transform \( \Pi \) defined in equation (3.2) below (e.g., Cox et al. (1981)).
To cover a broad class of relevant derivatives prices, fix a non-decreasing time vector \( \tilde{T} \in \mathbb{R}^\tilde{n} \) and suppose that the derivative contract features a single payoff, which occurs \( \tilde{T}_{\tilde{n}} \) periods ahead and is determined by \( h(\tilde{X}_{t+\tilde{T}}; K) \) 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

\[
h(\tilde{x}; K) = \sum_{i=1}^{n_h} \exp(\tilde{\omega}_i \cdot \tilde{x}) \hat{g}_i(\tilde{\omega} \cdot \tilde{x}; K) \quad (3.1)
\]

for \( \tilde{\omega}_i, \tilde{\omega} \in \mathbb{R}^{n \times \tilde{n}} \) and \( (\tilde{y} \mapsto \hat{g}_i(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R}) \).

Within this setting, denote by \( D_t(\tau) = \exp(-\tau \tau) \) the \( \tau \)-period discount factor. Using \( \Phi^\mathcal{Q} \) 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 \( \tilde{X}_{t+\tilde{T}} \) by

\[
\Pi(\omega; \tilde{T}, Z_t) = E^\mathcal{Q}[D_t(\tilde{T}_{\tilde{n}}) \exp(\omega \cdot \tilde{X}_{t+\tilde{T}}) | \mathcal{F}_t] = \exp(A(\omega; \tilde{T}) + B(\omega; \tilde{T} \cdot Z_t)),
\]

where \( A(\omega; \tilde{T}) = A^\mathcal{Q}(\omega; \tilde{T}) - r \tilde{T}_{\tilde{n}} \) and \( B(\omega; \tilde{T}) = B^\mathcal{Q}(\omega; \tilde{T}) \). In order to access the results of generalized transform analysis, we impose the following assumption on \( \Pi \) in equation (3.2).

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

With the general payoff function (3.1) and the pricing transform in equation (3.2), 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 \( \mathcal{V} \).

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

\[
\mathcal{V}(K, \tilde{T}; Z_t) = E^\mathcal{Q}[D_t(\tilde{T}_{\tilde{n}}) h(\tilde{X}_{t+\tilde{T}}; K) | \mathcal{F}_t] = \langle w(y; K), \Pi(u(y); \tilde{T}, Z_t) \rangle,
\]

where \( y = [\tilde{\omega}; \tilde{y}] \) and \( u([\tilde{\omega}; \tilde{y}]) = \tilde{\omega} + i\tilde{y} \tilde{\omega} \). Moreover, \( (y \mapsto w(y; K)) \in \mathcal{S}^*(\mathcal{Y}) \) is given by the distributional tensor product\(^{11}\)

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

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 \( \Pi \) needs to be determined numerically in most cases, evaluation of the price function \( \mathcal{V} \) in equation (3.3) will, therefore, generally require solving a series of one-dimensional numerical integration problems.

\(^{11}\)The distributional tensor product \( \otimes \) is formally defined in appendix B.1.
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.\(^{12}\) Fixing the time vector $\hat{T} = [\tau]$ as well as the log moneyness strike $K$, the call and put payoff of this option are given by $h^C_{stock}(\hat{X}_{t+\hat{T}}; K)$ and $h^P_{stock}(\hat{X}_{t+\hat{T}}; K)$, respectively, where

\[
\begin{align*}
    h^C_{stock}(\tilde{x}; K) &= (\exp([1; 0] \cdot \tilde{x}) - \exp(K)) U([1; 0] \cdot \tilde{x} - K) \\
    h^P_{stock}(\tilde{x}; K) &= (\exp(K) - \exp([1; 0] \cdot \tilde{x})) U(K - [1; 0] \cdot \tilde{x}).
\end{align*}
\]

Here, $U$ denotes the Heaviside step function. Each of the payoff functions in equation (3.5) satisfies the conditions of assumption 3.1 and can straightforwardly be expressed in the form of equation (3.1).

Denote the derivatives price associated to $h^O_{stock}$ in equation (3.5) by $V^O_{stock}$ for option type $O \in \{C, P\}$. The following corollary to proposition 3.1 yields an expression for $V^O_{stock}$ as a special case of equation (3.3).

**Corollary 3.1.** Let $h^O_{stock}$ be as in equation (3.5). Moreover, let assumption 3.2 hold for $\omega_1 = [1; 0]$ and $\omega_2 = [0; 0]$ and $\omega = [1; 0]$. Then we have

\[
V^O_{stock}(K, \hat{T}; Z_t) = \left< w^O_{stock}(y; K), \Pi(u_{stock}(y); \hat{T}, Z_t) \right>,
\]

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

\[
\begin{align*}
    w^C_{stock}([\tilde{\omega}; \tilde{y}]; K) &= \frac{1}{2\pi} (\delta(\tilde{\omega} - [1; 0]) - \exp(K) \delta(\tilde{\omega})) \otimes (\pi \delta(\tilde{y}) - i \exp(-iK\tilde{y}) \tilde{y}^{-1}) \\
    w^P_{stock}([\tilde{\omega}; \tilde{y}]; K) &= \frac{1}{2\pi} (\exp(K) \delta(\tilde{\omega}) - \delta(\tilde{\omega} - [1; 0])) \otimes (\pi \delta(\tilde{y}) + i \exp(-iK\tilde{y}) \tilde{y}^{-1}).
\end{align*}
\]

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

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 $\Pi$ as in equation (3.6), but also for other transforms. Therefore, the following lemma treats a generic transform $\Upsilon$, which covers $\Pi$ as a special case. When applied to $\Pi$, the integral representation of $w^O_{stock}$ in equation (3.8) 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 $(y \mapsto \Upsilon(u_{stock}(y))) \in S'(\mathcal{Y})$ be Hermitian as a function of $\tilde{y}$. Then $w^O_{stock}$ in corollary 3.1 can be represented in integral form as

\[
\left< w^O_{stock}(y; K), \Upsilon(u_{stock}(y)) \right> = \frac{e^C_{stock}}{2} \Upsilon([1; 0]) + \int_{\mathbb{R}_+} \Upsilon(F_{stock}(\tilde{y}; K) \Upsilon([1; 0] + i\tilde{y}[1; 0])) \frac{d\tilde{y}}{\tilde{y}} - \exp(K) \left( \frac{e^P_{stock}}{2} \Upsilon([0; 0]) + \int_{\mathbb{R}_+} \Upsilon(F_{stock}(\tilde{y}; K) \Upsilon(i\tilde{y}[1; 0])) \frac{d\tilde{y}}{\tilde{y}} \right),
\]

where $e^C_{stock} = +1$, $e^P_{stock} = -1$, and $F_{stock}(\tilde{y}; K) = \frac{1}{i} \exp(-iK\tilde{y})$.

\(^{12}\)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).
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 $\tau_{\text{vix}}$ equal to 30 calendar days, we obtain the representation

$$VIX_t^2 = \frac{2 \exp(r \tau_{\text{vix}})}{\tau_{\text{vix}}} \left( \int_{-\infty}^{0} \exp(-K) \mathcal{V}^{P}_{\text{stock}}(K, \tau_{\text{vix}}; Z_t) dK + \int_{0}^{\infty} \exp(-K) \mathcal{V}^{C}_{\text{stock}}(K, \tau_{\text{vix}}; Z_t) dK \right).$$

(3.9)

This theoretical construction forms the conceptual basis for practical VIX-type indices, which aim at approximating the right-hand side of equation (3.9) using a finite number of observed option quotes. For the purpose of derivatives pricing, the thus constructed VIX is hardly tractable in the form of equation (3.9). 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 $VIX_t^2$ on the latent state vector $Z_t$ for the state dynamics in equation (2.1).

Lemma 3.2. Let $VIX_t^2$ be given as in equation (3.9). It holds that

$$VIX_t^2 = a_{\text{vix}} + b_{\text{vix}} \cdot Z_t$$

(3.10)

for coefficients $a_{\text{vix}} \in \mathbb{R}$ and $b_{\text{vix}} \in \mathbb{R}^{n_Z}$ given in equation (B.11).

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 $h^C_{\text{vix}}(\tilde{X}_{t+\tilde{T}}; K)$ and $h^P_{\text{vix}}(\tilde{X}_{t+\tilde{T}}; K)$, respectively. Using the affine expression for $VIX_t^2$ in equation (3.10), we have

$$h^C_{\text{vix}}(\tilde{x}; K) = ((a_{\text{vix}} + [0; b_{\text{vix}}] \cdot \tilde{x})^{1/2} - K^{1/2}) U((a_{\text{vix}} + [0; b_{\text{vix}}] \cdot \tilde{x}) - K)$$

(3.11a)

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

(3.11b)

As before, $U$ denotes the Heaviside step function. Each of the payoff functions in equation (3.11) satisfies the conditions of assumption 3.1 and constitutes a special case of equation (3.1).

Denote the derivatives price associated to $h^O_{\text{vix}}$ in equation (3.11) by $\mathcal{V}^O_{\text{vix}}$ for option type $O \in \{C, P\}$. The following corollary to proposition 3.1 states an expression for $\mathcal{V}^O_{\text{vix}}$ as a special case of equation (3.3).

Corollary 3.2. Let $h^O_{\text{vix}}$ be as in equation (3.11). Moreover, let assumption 3.2 hold for $\bar{\omega}_1 = [0; 0]$ and $\tilde{\omega} = [0; b_{\text{vix}}]$. Then we have

$$\mathcal{V}^O_{\text{vix}}(K, \tilde{T}; Z_t) = \langle w^O_{\text{vix}}(y; K), \Pi(u_{\text{vix}}(y); \tilde{T}, Z_t) \rangle,$$

(3.12)

For an analysis of the approximation errors incurred in the practical realization of equation (3.9), the interested reader is referred to, e.g., Jiang and Tian (2005, 2007).
where \( y = [\tilde{\omega}; \tilde{y}] \) and \( v_{\text{vix}}([\tilde{\omega}; \tilde{y}]) = \tilde{\omega} + i\tilde{y}[0; b_{\text{vix}}] \). The associated \( (y \mapsto w_{\text{vix}}^C(y; K)) \in S^*(\mathcal{Y}) \) are given by

\[
\begin{align*}
w_{\text{vix}}^C([\tilde{\omega}; \tilde{y}]; K) &= \frac{1}{4\pi} \delta(\tilde{\omega}) \otimes \exp(i a_{\text{vix}} \tilde{y}) (-\operatorname{sgn}(\tilde{y}) i)^{3/2} \Gamma(1/2, iK\tilde{y}) |\tilde{y}|^{-3/2} \tag{3.13a} \\
w_{\text{vix}}^P([\tilde{\omega}; \tilde{y}]; K) &= \frac{1}{4\pi} \delta(\tilde{\omega}) \otimes (4\pi K^{1/2} \delta(\tilde{y}) - \exp(i a_{\text{vix}} \tilde{y}) (-\operatorname{sgn}(\tilde{y}) i)^{3/2} \gamma(1/2, iK\tilde{y}) |\tilde{y}|^{-3/2}) \tag{3.13b}
\end{align*}
\]

where \( \Gamma \) and \( \gamma \) denote the upper and lower incomplete Gamma function, respectively.

The (prepaid) forward contract with payoff \( h_{\text{vix}}^C(\tilde{X}_{t+T}; 0) \) has the price price function \( V_{\text{vix}}^C(0, \tilde{T}; Z_t) \), 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 \( w_{\text{vix}}^C \) in corollary 3.2. Analogous to lemma 3.1, the upcoming lemma gives such a representation for a general transform \( \Psi \), which covers \( \Pi \) in equation (3.12) as a special case. When applied to \( \Pi \), the integral representation of \( w_{\text{vix}}^C \) in equation (3.14) 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).\(^{14}\)

**Lemma 3.3.** Let \( (y \mapsto \Psi(v_{\text{stock}}(y))) \in S(\mathcal{Y}) \) be Hermitian as a function of \( \tilde{y} \). Then \( w_{\text{vix}}^C \) in corollary 3.2 can be represented in integral form as

\[
\langle w_{\text{vix}}^C(y; K), \Psi(v_{\text{vix}}(y)) \rangle = \frac{1 - c_{\text{vix}}^C}{2} K^{1/2} T([0; 0]) + \int_{\mathbb{R}_+} \tilde{X}_y^{(2)} \frac{\mathcal{R}(F_{\text{vix}}^C(\tilde{y}; K) \Psi(i\tilde{y}[0; b_{\text{vix}}]))}{\tilde{y}^{3/2}} \, d\tilde{y}, \tag{3.14}
\]

where \( \tilde{X}_y^{(2)} f(\tilde{y}) = f(\tilde{y}) - f(+0), c_{\text{vix}}^C = +1, c_{\text{vix}}^P = -1, \) and

\[
\begin{align*}
F_{\text{vix}}^C(\tilde{y}; K) &= +\frac{1}{2\pi} \exp(i a_{\text{vix}} \tilde{y}) (-\operatorname{sgn}(\tilde{y}) i)^{3/2} \Gamma(1/2, iK\tilde{y}) \tag{3.15a} \\
F_{\text{vix}}^P(\tilde{y}; K) &= -\frac{1}{2\pi} \exp(i a_{\text{vix}} \tilde{y}) (-\operatorname{sgn}(\tilde{y}) i)^{3/2} \gamma(1/2, iK\tilde{y}) \tag{3.15b}
\end{align*}
\]

### 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. Section 4.2 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 4.3. While not the focus of our presentation, section 4.4 shows that our methodology can straightforwardly be extended to include realistic measurement errors in derivatives prices. Finally, to make our results more easily accessible, section 4.5 studies several examples. Derivations and proofs are contained in appendix C.

#### 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 (3.3),

\(^{14}\)Related pricing formulas are also derived for the case of derivatives on quadratic variation (e.g., Broadie and Jain (2008), Carr and Lee (2009), Frix and Gatheral (2005), and Sepp (2008a)).
Let assumption 4.1 hold. Then we have

\[ V_{i,t} = V_i(K_i, \tilde{T}_i; Z_t) = \langle w_i(y_i; K_i), \Pi(u_i(y_i); \tilde{T}_i, Z_t) \rangle \]  

(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 \( \Pi \) in equation (3.2). For the expression in equation (4.1) to be well-defined, we suppose that assumption 3.2 holds accordingly, so that \( (y_i \mapsto \Pi(u_i(y_i); \tilde{T}_i, z)) \in \mathcal{S}(\mathcal{Y}_i) \) for any \( z \in \mathcal{Z} \). As before, for a non-decreasing time vector \( \tau \in \mathbb{R}^{\tilde{n}} \), we moreover construct the stacked vector \( V_{t+\tau} = [V_{t+\tau_1}; \ldots; V_{t+\tau_n}] \).

### 4.2 Exact moments

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

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

\[ (V_{t+\tau})^\beta = \langle w^\beta(y; K), \Pi^\beta(u(y); \tilde{T}, Z_{t+\tau}) \rangle, \]

(4.2)

where the tempered distribution \( (y \mapsto w^\beta(y; K)) \in \mathcal{S}^*(\mathcal{Y}) \) is given in equation (C.5). Moreover, the associated pricing transform \( (y \mapsto \Pi^\beta(u(y); \tilde{T}, z)) \in \mathcal{S}(\mathcal{Y}) \) for any \( z \in \mathcal{Z} \) has the form

\[ \Pi^\beta(u(y); \tilde{T}, Z_{t+\tau}) = \exp(A_{11}^\beta(u(y); \tilde{T}) + B_{11}^\beta(u(y); \tilde{T}) \cdot Z_{t+\tau}), \]

(4.3)

where \( A_{11}^\beta \) and \( B_{11}^\beta \) are given in equation (C.7).

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

Using lemma 4.1, we are now in place to determine joint moments of state variables \( \tilde{X}_{t+\tau} \) and derivatives prices \( V_{t+\tau} \). 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 notion of an "extended" Schwartz space \( \tilde{\mathcal{S}} \), which is formally introduced in appendix C.1. Essentially, we require that the transform is Schwartz after appropriate regularization.

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

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

(ii) \( E^M[|\exp(\omega \cdot \tilde{X}_{t+\tau})^\alpha| q_\beta(Z_{t+\tau})^{-1}] < \infty \).

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 \( \omega \in \mathcal{C}^{n \times \tilde{n}} \) as well as multi-indices \( \alpha \in \mathbb{N}^{n \times \tilde{n}} \) and \( \beta \in \mathbb{N}^{\nu \times \tilde{n}} \). Let assumption 4.1 hold. Then we have

\[ \Phi^{M, (\alpha, \beta)}(\omega, 0; \tilde{\tau}, \infty) = E^M[\exp(\omega \cdot \tilde{X}_{t+\tau})^\alpha (V_{t+\tau})^\beta] \]

(4.4)

\[ = \langle w^\beta(y; K), \exp(A_{11}^\beta(u(y); \tilde{T})) \Phi^{M, (\alpha)}(\omega + [0; B_{11}^\beta(u(y); \tilde{T})]; \tilde{\tau}, \infty) \rangle \]
with \( w^3, A_\Pi^3, \) and \( B_\Pi^3 \) 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 theory in appendix C.1. For evaluating the integrand in equation (4.4), \( \Phi^{M, (\alpha)} \) can be determined from equation (2.9) under the conditions of proposition 2.2. While \( \hat{\Phi}^{M, (\alpha, \beta)} \) thereby admits an analytically tractable expression for arbitrary multi-indices \((\alpha, \beta)\), in general, its computation is only feasible for low orders of \( \beta \), since \( w^3 \) requires \(|\beta|\)-dimensional numerical integration.

### 4.3 Approximate moments

To avoid the computational cost of the exact pl-linear moments involving derivatives prices in section 4.2 while exploiting the feasibility of low-order moments, this section proposes an effective polynomial approximation approach.

Take \( L^2(\mathcal{Z}, \mathbb{M}) \) to be the set of square integrable functions on \( \mathcal{Z} \) against the probability measure \( \mathbb{M} \), i.e., comprising all \( f \) satisfying \( E^\mathbb{M}[|f(Z_t)|^2] < \infty \), where the choice of \( t \) is arbitrary due to stationarity. In order to assure that functions in \( L^2(\mathcal{Z}, \mathbb{M}) \) 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.** \( E^\mathbb{M} [\exp(\epsilon \|Z_t\|)] < \infty \) for some \( \epsilon > 0 \).

As a consequence of assumption 4.2, the set of monomials \( \{z^\gamma : \gamma \in \mathbb{N}^n\} \) forms a basis for \( L^2(\mathcal{Z}, \mathbb{M}) \) (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 \( \{\phi_\gamma(z) : \gamma \in \mathbb{N}^n\} \). By construction, we have \( \phi_\gamma(z) = \sum_{\gamma\leq\gamma'} b_{\phi,\eta}^\langle \gamma \rangle z^\gamma \) for coefficients \( b_{\phi,\eta}^\gamma \in \mathbb{R} \) depending on the unconditional monomial moments of \( Z_t \) under \( \mathbb{M} \) up to order \( 2|\gamma| \), with \( \prec \) denoting the lexicographic order.

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

**Assumption 4.3.** \( (z \mapsto \mathcal{V}_t(K_i, \hat{T}_i; z)) \in L^2(\mathcal{Z}, \mathbb{M}) \) for all \( 1 \leq i \leq n \).

Combining assumptions 4.2 and 4.3, we construct an approximant \( V_{t,(p)} \) for \( \mathcal{V}_t \) by projecting each of its elements \( V_{t,i}(K_i, \hat{T}_i; Z_t) \) onto the truncated set of basis functions \( \{\phi_\gamma(z) : \gamma \in \mathbb{N}^n, |\gamma| \leq p\} \). 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

\[
V_{t,(p)} = \sum_{|\gamma| \leq p} \tilde{c}_{V,\eta} \phi_\gamma(Z_t) \tag{4.5}
\]

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

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

\[
V_{t,(p)} = \sum_{|\gamma| \leq p} \tilde{b}_{V,\eta,(p)} (Z_t)^\gamma \tag{4.6}
\]

for \( \tilde{b}_{V,\eta,(p)} \in \mathbb{R}^{n \nu} \), depending on the expansion order \( p \). Constructing \( V_{t+\bar{\tau},(p)} \) as in equation (4.6) separately for every \( 1 \leq j \leq \bar{n} \), we define the stacked vector \( V_{t+\bar{\tau}}(p) = [V_{t+\bar{\tau},1,(p)}; \ldots; V_{t+\bar{\tau},\bar{n},(p)}] \). Padding
the coefficients in equation (4.6) with zeros then yields
\[ V_{t+\bar{t},(p)} = \sum_{|\eta| \leq p} b_{\bar{\eta},(p)} (Z_{t+\bar{t}})^\eta \] (4.7)
for \( b_{\bar{\eta},(p)} \in \mathbb{R}^{n \times \bar{n}} \). Monomials of \( V_{t+\bar{t},(p)} \) in equation (4.7) thus obtain as polynomials in \( Z_{t+\bar{t}} \), given by
\[ (V_{t+\bar{t},(p)})^\beta = \sum_{|\eta| \leq p|\beta|} b_{\eta,(p)} (Z_{t+\bar{t}})^\eta \] (4.8)
for \( b_{\eta,(p)} \in \mathbb{R} \) determined as polynomials of the coefficients \( \tilde{c}_{\bar{\eta}} \) in equation (4.5).

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

**Assumption 4.4.** \( (\exp(\omega \cdot \tilde{X}_{t+\bar{t}}) (\tilde{X}_{t+\bar{t}})^\alpha (V_{t+\bar{t},(p)})^\beta)_p \) is uniformly integrable.\(^{15}\)

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 \( V_{t+\bar{t}} \) can be obtained by computing exact pl-linear moments involving the approximant \( V_{t+\bar{t},(p)} \). The resulting sequence of moment approximants converges to the exact moments derived in proposition 4.1.

**Proposition 4.2.** Consider an argument \( \omega \in \mathbb{C}^{n \times \bar{n}} \) as well as multi-indices \( \alpha \in \mathbb{N}^{n \times \bar{n}} \) and \( \beta \in \mathbb{N}^{n \times \bar{n}} \). Let assumptions 4.2 to 4.5 hold. Then we have that
\[
\tilde{\Phi}_{m,(\alpha,\beta)}^{(p)}(\omega,0;\bar{t},\infty) = \mathbb{E}^\mathbb{M}[\exp(\omega \cdot \tilde{X}_{t+\bar{t}}) (\tilde{X}_{t+\bar{t}})^\alpha (V_{t+\bar{t},(p)})^\beta] = \sum_{|\eta| \leq p|\beta|} b_{\eta,(p)} \Phi_{m,(\alpha+|\eta|)}^{(p)}(\omega;\bar{t},\infty) \] (4.9)
with \( b_{\eta,(p)} \) given in equation (C.11) satisfies \( \tilde{\Phi}_{m,(\alpha,\beta)}^{(p)} \to \tilde{\Phi}_{m,(\alpha,\beta)} \) as \( p \to \infty \).

Except for the coefficients \( b_{\eta,(p)} \), the approximate pl-linear moment \( \tilde{\Phi}_{m,(\alpha,\beta)}^{(p)} \) in equation (4.9) does not require the evaluation of any moments involving derivatives prices. Only the pl-linear moments \( \Phi_{m,(\alpha+|\eta|)}^{(p)} \) for \( |\eta| \leq p|\beta| \) need to be computed, which can be achieved at low computational cost as in equation (2.9) under the conditions of proposition 2.2.

It remains to establish a practicable procedure for computing the coefficients \( b_{\eta,(p)} \), which are polynomials of the coefficients \( \tilde{c}_{\bar{\eta}} \) 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 4.2 and, hence, impose the following regularity conditions in order to access these results.

**Assumption 4.5.** For every \( 1 \leq i \leq n_V \) there exists positive \( q_i \in \mathbb{C}^\infty(Z) \) satisfying the following conditions for all \( \gamma \in \mathbb{N}^{n_V} \) with \( |\gamma| \leq p \):
(i) \( ((y_i,z) \mapsto \Pi(u_i(y_i);\bar{T}_i,z)) \in \mathcal{S}(Y_i \times Z; \mathbb{I} \otimes q_i) \);
(ii) \( \mathbb{E}^\mathbb{M}[|Z_i|^1 q_i(Z_i)^{-1}] < \infty \).

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

\(^{15}\)A sequence \((\xi_p)_p\) is called uniformly integrable whenever \( \sup_p \mathbb{E}^\mathbb{M}[|\xi_p| U(|\xi_p| - K)] \to 0 \) as \( K \to \infty \).
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 \( |\eta| \leq p \) is given by

\[
\tilde{c}_{V,\eta} = \sum_{i=1}^{n_V} c_i \sum_{\gamma \in \eta} b^{(\gamma)}_\phi \exp(A_{\eta}(y_i; K_i, \exp(A_{\Pi}(u_i(y_i; T_i)))) \Phi^{M, (0, \gamma)}([0; B_{\eta}(u_i(y_i; T_i)]; 0, \infty)),
\]

(4.10)

where \( c_i \in \mathbb{N}^{n_V} \) 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 \( \tilde{\Phi}^{M, (0, \gamma)} \) as in equation (4.4) for each combination of \( i \) and \( |\gamma| \leq p \). Hence, in general, \( V_{t,(p)} \) can be 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.

4.4 Including measurement errors

Thus far, sections 4.2 and 4.3 have 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. 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 \( \bar{z}_t = [\bar{z}_t; \varepsilon_t] \) and define \( Z_{t+\tau} = [\bar{z}_{t+\tau}; \varepsilon_{t+\tau}] \) for notational convenience. For generalizing proposition 4.1, suppose that \( \tilde{V}_t \) the vector of derivatives prices measured with error \( \varepsilon_t \) and construct \( \tilde{V}_{t+\tau} \) as usual. Analogous to equation (4.4), define the pl-linear moments

\[
\tilde{\Phi}^{M, (\alpha, \beta)}(\omega, 0; \tau, \infty) = E^M[\exp(\omega \cdot \tilde{X}_{t+\tau}) (\tilde{X}_{t+\tau})^\alpha (\tilde{V}_{t+\tau})^\beta].
\]

(4.11)

To give an expression for \( \tilde{\Phi}^{M, (\alpha, \beta)} \) in equation (4.11), we consider three cases. First, for exponential measurement errors with \( \tilde{V}_t = V_t \exp(\varepsilon_t) \), we have

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

(4.12)

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

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

(4.13)

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

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

(4.14)

Equations (4.12) to (4.14) 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 measurement errors moment moments separately. Specifically, making the assumption that \( \bar{z}_{t+\tau} \) and \( \varepsilon_{t+\tau} \) are independent, we obtain

\[
\tilde{\Phi}^{M, (\alpha, \beta)}(\omega, 0; \tau, \infty) = \tilde{\Phi}^{M, (\alpha + [0, 0, 0], \beta)}([\omega_S; \omega_Z; 0], 0; \tau, \infty) \tilde{\Phi}^{M, (0, [0, 0, 0], 0)}(0; \varepsilon_0; 0; \tau, \infty),
\]

18
where the measurement errors affect only the second term.

4.5 Examples

To illustrate our theoretical results and make them more easily accessible, we now briefly discuss three examples with increasing complexity. For each of these examples, 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 Example 1

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

\[
\Phi^{M,(\alpha,\beta)}(0,0;\tilde{\tau},\infty) = \mathbb{E}^{M}([\tilde{X}_{t+|\tau_i|}]^\alpha V_{i,t+\tau_i})
= \langle w^\beta(y;K), \exp(A^\beta_H(u(y);\tilde{T})) \Phi^{M,(\alpha)}([0;B^\beta_H(u(y);\tilde{T}]|;\tilde{\tau},\infty) \rangle,
\]

with lemma 4.1 providing the required expressions for \( w^\beta, A^\beta_H, \) and \( B^\beta_H \).

Here, \( w^\beta \) corresponds to the tempered distribution \( w_i \) associated to \( V_{i,t} \), while \( A^\beta_H \) and \( B^\beta_H \) are the respective coefficients of the pricing transform. Moreover, for argument \( \omega = [\omega_1] \in \mathbb{C}^{nx} \), \( \Phi^{M,(\alpha)} \) results from taking partial derivatives of

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

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

4.5.2 Example 2

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

\[
\Phi^{M,(\alpha,\beta)}(0,0;\tilde{\tau},\infty) = \mathbb{E}^{M}([\tilde{X}_{t+|\tau_i|}]^\alpha V_{i,t+\tau_i} V_{j,t+\tau_j})
= \langle w^\beta(y;K), \exp(A^\beta_H(u(y);\tilde{T})) \Phi^{M,(\alpha)}([0;B^\beta_H(u(y);\tilde{T}]|;\tilde{\tau},\infty) \rangle,
\]

where from lemma 4.1

\[
\begin{align*}
w^\beta(y;K) &= w_i(y_1;K_i) \otimes w_j(y_2;K_j) \\
A^\beta_H(u(y);\tilde{T}) &= A_H(u_i(y_1);\tilde{T}_i) + A_H(u_j(y_2);\tilde{T}_j) \\
B^\beta_H(u(y);\tilde{T}) &= B_H(u_i(y_1);\tilde{T}_i) + B_H(u_j(y_2);\tilde{T}_j).
\end{align*}
\]
Unlike in example 1, \( w^\beta \) now is a tensor product of the tempered distributions \( w_i \) and \( w_j \) associated to the derivatives prices \( V_{i,\tau} \) and \( V_{j,\tau} \), respectively. Similarly, \( A_{H}^{\beta} \) and \( B_{H}^{\beta} \) can be interpreted as tensor sums of the associated coefficients of the pricing transform. Finally, the expression for \( \Phi_{\alpha,\beta}^{\kappa,\lambda} \) is identical to the one in example 1.

4.5.3 Example 3

As a third and final example, we consider polynomial moments involving the product of (potentially) asynchronous derivatives prices \( V_{i,\tau} \) and \( V_{j,\tau} \) for \( \tilde{\tau}_2 \geq \tilde{\tau}_1 \geq 0 \) and \( 1 \leq i, j \leq n_{V} \). Unlike in the preceding examples, we now have two (potentially) 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 \( \beta = \epsilon_i + \epsilon_{n_{V} + j} \in \mathbb{N}^{2n} \) as well as \( \alpha = [\alpha_1; \alpha_2] \in \mathbb{R}^{2n} \), equation (4.4) yields

\[
\tilde{\Phi}_{\alpha,\beta}^{\kappa,\lambda}(0; \tilde{\tau}, \infty) = \mathbb{E}^{\kappa}([\tilde{X}_t+\tau_1])^{\alpha_1} (\tilde{X}_t+\tau_1+\tau_2) V_{i,\tau+\tau_1} V_{j,\tau+\tau_2} = (w^\beta(y; K), \exp(A_{H}^{\beta}(y; \tilde{T})) \Phi_{\alpha,\beta}^{\kappa,\lambda}([0; B_{H}^{\beta}(y; \tilde{T})]; \tilde{\tau}, \infty)),
\]

where from lemma 4.1, we get

\[
w^\beta(y; K) = w_i(y_1; K_1) \otimes w_j(y_2; K_2) \\
A_{H}^{\beta}(y; \tilde{T}) = A_{H}(y_1; \tilde{T}_1) + A_{H}(y_2; \tilde{T}_2) \\
B_{H}^{\beta}(y; \tilde{T}) = [B_{H}(y_1; \tilde{T}_1); B_{H}(y_2; \tilde{T}_2)].
\]

Analogous to example 2, \( w^\beta \) is a tensor product of the tempered distributions \( w_i \) and \( w_j \), while \( A_{H}^{\beta} \) is a tensor sum of the respective coefficients of the pricing transform. However, \( B_{H}^{\beta} \) is not a tensor sum, but rather a block vector in \( \mathbb{C}^{2n} \). Moreover, for argument \( \omega = [\omega_1; \omega_2] \in \mathbb{C}^{2n} \), \( \Phi_{\alpha,\beta}^{\kappa,\lambda} \) results from taking partial derivatives of

\[
\Phi_{\alpha,\beta}^{\kappa,\lambda}(\omega; \tilde{\tau}, \infty) = \exp(A_{\Phi}(\omega_2; \tau_2 - \tau_1, 0) + A_{\Phi}(\omega_1 + [0; B_{\Phi}(\omega_2; \tau_2 - \tau_1, 0)]; \tau_1, \infty)) \\
+ B_{\Phi}(\omega_1 + [0; B_{\Phi}(\omega_2; \tau_2 - \tau_1, 0)]; \tau_1, \infty))
\]

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

With these specifications in example 3, we can reproduce example 2 by choosing \( \tilde{\tau}_2 = \tilde{\tau}_1 \) as well as \( \alpha_2 = 0 \), thereby returning effectively to the case \( \tilde{\tau} = [\tilde{\tau}_1] \).

5 Estimation methodology

In this section, we devise a GMM estimation procedure incorporating moments involving derivatives prices derived 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 4.2 and 4.3, respectively. Appendix D contains additional technical details and proofs.

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 \( V_t \). Maintaining the setup of section 4, each element of \( V_t \) is given by proposition 3.1 as in equation (4.1). In order to obtain such a panel in practice, it is usually necessary to interpolate observed market prices. The interval between adjacent observation dates equals \( \Delta \). Define a parameter space \( \Theta \) such that a parameter vector \( \vartheta \in \Theta \) contains all relevant model
parameters. For any \( \vartheta \in \Theta \), we denote by \( P(\vartheta) \) and \( Q(\vartheta) \) 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

To estimate the model using an exact GMM approach, we define a vector of moment conditions \( f_t(\vartheta) \), depending on the model parameters \( \vartheta \) as well as on the data. For a weighting matrix \( W \), the GMM estimator can be written as

\[
\hat{\vartheta}_T(W) = \text{argmin}_{\vartheta \in \Theta} \left[ \frac{1}{T} \sum_{t=1}^{T} f_t(\vartheta) \right]^T W \left[ \frac{1}{T} \sum_{t=1}^{T} f_t(\vartheta) \right].
\]  

(5.1)

Specifically, for our purposes, we consider in equation (5.1) exact moment conditions \( f_t(\vartheta) \) of the form

\[
f_t(\vartheta) = P(\tilde{X}_{t+\tau}, V_{t+\tau}) - E^{\vartheta_0}[P(\tilde{X}_{t+\tau}, V_{t+\tau})],
\]  

(5.2)

where \( P(\tilde{x}, v) = \sum_{\alpha,\beta} c_{\alpha,\beta} \tilde{x}^\alpha v^\beta \) is a vector-valued polynomial. By construction, the moment conditions in equation (5.2) satisfy \( E^{\vartheta_0}[f_t(\vartheta_0)] = 0 \). Under the conditions of proposition 4.1, linearity of the expectation operator yields

\[
E^{\vartheta_0}[P(\tilde{X}_{t+\tau}, V_{t+\tau})] = \sum_{\alpha,\beta} c_{\alpha,\beta} \tilde{\Phi}^{\vartheta_0}(\alpha,\beta)(0,0; \tau, \infty)
\]  

(5.3)

with each \( \tilde{\Phi}^{\vartheta_0}(\alpha,\beta) \) given by equation (4.4). Like other pl-linear moments, polynomial moments as in equation (5.3) generally require the evaluation of \( |\beta| \)-dimensional numerical integrals, which usually renders the moment conditions \( f_t(\vartheta) \) in equation (5.2) computationally infeasible except for low orders of \( \beta \).

With moment conditions in the form of equation (5.2), the efficient GMM estimator can be realized in a single step. Setting the sample average \( \hat{g}_T(\vartheta) = \frac{1}{T} \sum_{t=1}^{T} f_t(\vartheta) \), we generally have

\[
\Gamma = \lim_{T \to \infty} T E^{\vartheta_0}[\hat{g}_T(\vartheta_0) \hat{g}_T(\vartheta_0)^T] = \Gamma_0 + \sum_{t=1}^{\infty} \Gamma_t + \Gamma_t^T,
\]

assuming that \( \Gamma_t = E^{\vartheta_0}[f_t(\vartheta_0) f_{(t+\ell)}(\vartheta_0)^T] \) are absolutely summable. For the particular form of moment conditions considered here, we specifically obtain

\[
\Gamma_t = E^{\vartheta_0}[P(\tilde{X}_{t+\Delta+\tau}, V_{\Delta+\tau}) P(\tilde{X}_{(t+\ell)+\Delta+\tau}, V_{(t+\ell)+\Delta+\tau})^T].
\]

Hence, we can construct an estimator \( \hat{\Omega}_T \) for \( \Omega \) from the data alone; e.g., using the estimator proposed in Newey and West (1987). Efficient GMM estimation results when setting \( W \propto \hat{\Omega}_T^{-1} \).

For later comparison, we summarize the asymptotic properties of the exact estimator \( \hat{\vartheta}_T(W) \) in equation (5.1) in the upcoming proposition. To guarantee consistency and asymptotic normality of the estimator, we rely on standard regularity conditions.

**Proposition 5.1.** Let assumptions D.1 and D.2 hold. Then the estimator \( \hat{\vartheta}_T(W) \) in equation (5.1) is consistent and asymptotically normally distributed:

(i) \( \hat{\vartheta}_T(W) \to^p \vartheta_0 \);
(ii) \( \sqrt{T}(\hat{\theta}_T(W_T) - \theta_0) \xrightarrow{d} N(0, (G^T W G)^{-1} G^T W \Omega W G (G^T W G)^{-1}) \).

### 5.3 Approximate GMM estimation

To estimate the model using an approximate GMM approach, we consider a vector of approximate moment conditions \( f_{\ell,(p)}(\vartheta) \), again depending on the model parameters \( \vartheta \) and the data, but also on the approximation order \( p \). As an alternative to the exact GMM estimator in equation (5.1), define the approximate GMM estimator as a function of the weighting matrix \( W \) by

\[
\hat{\vartheta}_{T,(p)}(W) = \arg\min_{\vartheta \in \Theta} \left[ \frac{1}{T} \sum_{t=1}^{T} f_{t,\Delta,(p)}(\vartheta) \right]^T W \left[ \frac{1}{T} \sum_{t=1}^{T} f_{t,\Delta,(p)}(\vartheta) \right].
\]  
(5.4)

Analogous to equation (5.2), we now specify the approximate moment conditions \( f_{\ell,(p)}(\vartheta) \) in equation (5.4) by

\[
f_{\ell,(p)}(\vartheta) = P(\tilde{X}_{t+\tau}, V_{t+\tau}) - E^{\vartheta}[P(\tilde{X}_{t+\tau}, V_{t+\tau})],
\]  
(5.5)

where \( V_{t+\tau,(p)} \) is constructed as in section 4.3 and \( P(\tilde{x}, v) = \sum_{\alpha,\beta} c_{\alpha,\beta} \tilde{x}^\alpha v^\beta \) is a vector-valued polynomial. Therefore, generally, \( E^{\vartheta}(\vartheta_0)[f_{\ell,(p)}(\vartheta_0)] \neq 0 \) for the moment conditions in equation (5.5). Under the conditions of proposition 4.2, linearity of the expectation operator yields

\[
E^{\vartheta}[P(\tilde{X}_{t+\tau}, V_{t+\tau})] = \sum_{\alpha,\beta} c_{\alpha,\beta} \tilde{\Phi}^{(\vartheta),(\alpha,\beta)}(0,0; \tau, \infty)
\]  
(5.6)

with each \( \tilde{\Phi}^{(\vartheta),(\alpha,\beta)} \) determined by equation (4.9) using expansion under \( \mathbb{P}(\vartheta) \). In contrast to the exact polynomial moments in equation (5.3), computation of approximate polynomial moments in equation (5.6) only requires one-dimensional numerical integration. Apart from that, they depend merely on polynomial moments of the state vector, which are even available in closed form. This renders approximate moment conditions and, hence, the approximate GMM estimator computationally feasible even for larger orders of \( \beta \).

While standard asymptotic theory holds for the exact GMM estimator in equation (5.1) under the usual assumptions, the properties of the approximate GMM estimator crucially depend on the behavior of the approximation order \( p(T) \) as \( T \to \infty \). Under additional regularity conditions formalized in assumptions D.3 and D.4, the following proposition captures a situation in which \( p(T) \to \infty \) as \( T \to \infty \) at a rate such that the approximation error of the moment conditions vanishes fast enough. In that case, the approximate estimator \( \hat{\theta}_{T,(p)}(W) \) in equation (5.4) is asymptotically equivalent in terms of the properties in proposition 5.1 to the corresponding exact estimator.

**Proposition 5.2.** Let assumptions D.1 to D.4 hold. Then the estimator \( \hat{\theta}_{T,(p(T))}(W_T) \) in equation (5.4) is consistent and asymptotically normally distributed:

(i) \( \hat{\theta}_{T,(p(T))}(W_T) \Rightarrow \theta_0 \);

(ii) \( \sqrt{T}(\hat{\theta}_{T,(p(T))}(W_T) - \theta_0) \xrightarrow{d} N(0, (G^T W G)^{-1} G^T W \Omega W G (G^T W G)^{-1}) \).

The assumptions underlying proposition 5.2 represent an ideal situation. Several further cases resting on weaker requirements can be distinguished. When \( p(T) \) increases such that the approximation error of the moments decreases only at square-root rate, a local bias in the sense of Armstrong and Kolesár (2019) results. Under otherwise intact regularity conditions, the resulting approximate estimator will be consistent, but less efficient. For fixed expansion order \( p(T) = p \), the approximate estimator is globally biased and inconsistent, a case formally treated in Hall and Inoue (2003). Further in-depth analysis of
approximate GMM estimators and proposed improvements are available in Kristensen and Salanié (2017).

6 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 prices, using the generalized transform analysis introduced in Chen and Joslin (2012) 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 feasible. We verify convergence of the approximate moments to their exact counterparts under standard regularity conditions.

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. As our results rely on generalized transform analysis, our results hold also for other model classes covered by this theory (see Chen and Joslin (2012) for various examples) and can even be extended further. 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 (potentially 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, which might introduce additional 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 Transform analysis for affine jump diffusions

This appendix contains supplementary details of standard transform analysis. In appendix A.1, we state an extension of the classical Faa 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 (appendix A.2) and pl-linear moments (appendix A.3).

A.1 Faa di Bruno formula

In order to state the Faa di Bruno formula, we need to introduce some further notation. For equidimensional multi-indices \( \alpha, \beta \), we define by the lexicographic order. Hence, we have \( \alpha \succ \beta \) provided that one of the following conditions is satisfied:

(i) \( |\alpha| > |\beta| \) or

(ii) \( |\alpha| = |\beta|, \alpha_i > \beta_i \) for some \( i \) and \( \alpha_j = \beta_j \) for all \( j < i \).

Moreover, we set \( \alpha \succeq \beta \) if either \( \alpha \succ \beta \) or \( \alpha = \beta \).

Constantine and Savits (1996) generalize the classical Faa di Bruno formula to allow for partial derivatives of a composite function. Using the notation just established, their version of the Faa di Bruno formula reads as in the following proposition.

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

\[
\partial^\alpha_x f(g(x)) = \sum_{|\beta| \leq |\alpha|} \partial^\beta_y f(g(x)) \sum_{Q(\alpha, \beta)} M^{\alpha}_{k, \ell} (\partial^\ell_x g(x))^k,
\]

where for multi-indices \( \alpha \in \mathbb{N}^n \) and \( \beta \in \mathbb{N}^m \), the set \( Q(\alpha, \beta) \), consisting of ordered multi-indices \( k_i \in \mathbb{N}^m \) and \( \ell_i \in \mathbb{N}^n \), is defined by

\[
Q(\alpha, \beta) = \bigcup_{r=1}^{|\beta|} \left\{ (k_1, \ldots, k_r; \ell_1, \ldots, \ell_r) : |k_i| > 0, \ell_1 \succ \cdots \succ \ell_r \succ 0, \sum_{i=1}^r k_i = \beta, \sum_{i=1}^r |k_i| \ell_i = \alpha \right\}. \tag{A.2}
\]

Moreover, for elements \( (k_1, \ldots, k_r; \ell_1, \ldots, \ell_r) \) of \( Q(\alpha, \beta) \), we define the generalized multinomial coefficient

\[
M^{\alpha}_{k, \ell} = \frac{(\alpha!)}{\prod_{i=1}^r (k_i!)(\ell_i!)^{k_i}}, \tag{A.3}
\]

and the tensor expression

\[
(\partial^\ell_x g(x))^k = \prod_{i=1}^r (\partial^{\ell_i}_x g(x))^{k_i}. \tag{A.4}
\]


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 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 define the characteristic $\chi = (A_{\mu,X}, B_{\mu,X}, A_{\Omega,X}, B_{\Omega,X}, A_{\lambda}, B_{\lambda}, \Phi_{\nu})$, containing all affine coefficients and the jump transform $\Phi_{\nu}$ driving the affine state dynamics in equation (2.1). To simplify the notation, we suppress the dependence of the elements of $\chi$ on $M$.

A.2.1 Single-period exponential moments

For given $\omega = [\omega_S; \omega_Z] \in \mathbb{C}^{nx}$ and $t, T \in \mathbb{R}_+$, define the complex-valued process $(\Psi_t)_{0 \leq t \leq T}$ by

$$
\Psi_t = \exp(A_{\Psi}(\omega; T - t) + [\omega_S; B_{\Psi}(\omega; T - t)] \cdot \tilde{X}_{t+\tau}) ,
$$

(A.5)

where the complex-valued coefficient functions $A_{\Psi}$ and $B_{\Psi}$ solve the system of ODEs (A.8) of generalized Riccati type. By Ito’s lemma applied to equation (A.5), $\Psi_t$ follows dynamics of the form

$$
d\Psi_t = \mu_{\Psi,t} - \sigma_{\Psi,t} \, d\tau + \sigma_{\Psi,t} \, dW_{t+\tau} + J_{\Psi,t} \, dN_{t+\tau} .
$$

(A.6)

If the characteristic $\chi$ is well-behaved in the sense of assumption A.1, the above construction assures that the process $(\Psi_t)_{0 \leq t \leq T}$ 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).

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

(i) $\Phi_{\nu}(\tilde{\omega})$ exists for all $\tilde{\omega}$ in an open set $O$ containing $\bigcup_{0 \leq t \leq T} \{[\omega_S; B_{\Psi}(\omega; t)]\}$;

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

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

- $\mathbb{E}^M[|\Psi_0|] < \infty$,
- $\mathbb{E}^M[\int_0^T \Omega_{\Psi,t} - d\tau]^{1/2} < \infty$ with $\Omega_{\Psi,t} = \sigma_{\Psi,t} \sigma_{\Psi,t}^\top$,
- $\mathbb{E}^M[\int_0^T |\mathbb{E}[J_{\Psi,t} \mid F_{t+\tau} - ] \Lambda(Z_{t+\tau})| \, d\tau] < \infty$.

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

$$
\Psi^M(\omega; \tau, Z_t) = \mathbb{E}^M[\exp(\omega \cdot \tilde{X}_{t+\tau}) \mid F_t] = \exp(A_{\Psi}(\omega; \tau) + B_{\Psi}(\omega; \tau) \cdot Z_t)
$$

(A.7)

with coefficients $A_{\Psi}(\omega; \tau) \in \mathbb{C}$ and $B_{\Psi}(\omega; \tau) \in \mathbb{C}^{nx}$ determined by the system of ODEs\(^{16}\)

$$
\partial_t A_{\Psi} = A_{\mu,X}^{\top} [\omega_S; B_{\Psi}] + \frac{1}{2} A_{\Omega,X}^{\top} ( [\omega_S; B_{\Psi}] \otimes [\omega_S; B_{\Psi}] ) + A_{\lambda} (\Phi_{\nu}( [\omega_S; B_{\Psi}] ) - t) \quad \text{(A.8a)}
$$

$$
\partial_t B_{\Psi} = B_{\mu,X}^{\top} [\omega_S; B_{\Psi}] + \frac{1}{2} B_{\Omega,X}^{\top} ( [\omega_S; B_{\Psi}] \otimes [\omega_S; B_{\Psi}] ) + B_{\lambda} (\Phi_{\nu}( [\omega_S; B_{\Psi}] ) - t) \quad \text{(A.8b)}
$$

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

Proof. See Duffie et al. (2000).

\(^{16}\)Here, $\otimes$ denotes the ordinary Kronecker product.

25
In addition to the stationarity assumption, the regularity conditions specified by assumption A.2 allow to obtain unconditional exponential moments as a limit of the conditional exponential moments in proposition A.2. In essence, proposition A.3 yields the ergodicity result that \( \Psi_M(\omega; \infty) = \lim_{\tau \to \infty} \Psi_M(\omega; \tau, z) \) irrespective of \( z \).

**Assumption A.2.** The characteristic \( \chi \) is well-behaved at \( \omega \in \mathbb{C}^{n \times} \) for \( \omega = [0; \omega] \) by satisfying the following conditions:

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

(ii) \( \Psi^M([0; \omega]; \infty) \) exists at \( \omega = B_\Phi([0; \omega]; \tau) \) for all \( \tau \geq 0 \);

(iii) \( \omega \to \Psi^M([0; \omega]; \infty) \) is continuous at \( \omega = 0 \);

(iv) \( [0; \omega] \in \mathcal{R}_\Phi \), where \( \mathcal{R}_\Phi \) denotes the stability region of the system of ODEs (A.8) containing all \( \omega = [0; \omega] \in \mathbb{C}^{n \times} \) such that:

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

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

\[
\Psi^M(\omega; \infty) = E^M[\exp(\omega \cdot X_t)]
\]

\[
= \exp(A_\Phi(\omega; \infty))
\]

(A.9)

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

**Proof.** See Dillschneider (2020). \( \square \)

**A.2.2 Multi-period exponential moments**

With additional regularity conditions, the single-period moments in proposition A.2 allow to iteratively determine multi-period exponential moments. Essentially, the conditions in assumption A.3 assure that the law of iterated expectations can be applied. We can then state an auxiliary result in proposition A.4 that will be used for proving both statements in proposition 2.1.

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

(i) \( \chi \) is well-behaved at \( (\omega_1 + [0; B_\Phi(1)(\omega; \hat{\tau})], \Delta_i) \) in the sense of assumption A.1 for all \( 1 \leq i \leq \hat{n} \);

(ii) \( \chi \) is well-behaved at \( (0; B_\Phi(0)(\omega; \hat{\tau})) \) in the sense of assumption A.1.

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

\[
\Phi^M(\omega; \hat{\tau}, \tau, Z_t) = E^M[\exp(\omega \cdot \tilde{X}_{t+\tau}) | F_t]
\]

\[
= \exp(A_\Phi(\omega; \hat{\tau}, \tau) + B_\Phi(\omega; \hat{\tau}, \tau) \cdot Z_t)
\]

(A.10)

with coefficients \( A_\Phi(\omega; \hat{\tau}, \tau) \in \mathbb{C} \) and \( B_\Phi(\omega; \hat{\tau}, \tau) \in \mathbb{C}^{n \times} \) given by

\[
A_\Phi(\omega; \hat{\tau}, \tau) = A_\Phi(0)(\omega; \hat{\tau}) + A_\Phi([0; B_\Phi(0)(\omega; \hat{\tau})]; \tau)
\]

(A.11a)

\[
B_\Phi(\omega; \hat{\tau}, \tau) = B_\Phi([0; B_\Phi(0)(\omega; \hat{\tau})]; \tau).
\]

(A.11b)
Defining $\Delta_i = \tilde{\tau}_i - \tilde{\tau}_{i-1}$ and $\omega_i = [\omega_i; \omega_i; Z]$, the auxiliary coefficients $A_{\Phi,(i)}(\omega; \tilde{\tau}) \in \mathbb{C}$ and $B_{\Phi,(i)}(\omega; \tilde{\tau}) \in \mathbb{C}^{n \times 2}$ are determined by the backward recursion

\begin{align}
A_{\Phi,(i-1)}(\omega; \tilde{\tau}) &= A_{\Phi,(i)}(\omega; \tilde{\tau}) + A_{\Phi}(\omega_i + [0; B_{\Phi,(i)}(\omega; \tilde{\tau})]; \Delta_i) \\
B_{\Phi,(i-1)}(\omega; \tilde{\tau}) &= B_{\Phi}(\omega_i + [0; B_{\Phi,(i)}(\omega; \tilde{\tau})]; \Delta_i)
\end{align}

for $i = \hat{n}, \ldots, 1$ subject to the initial conditions $A_{\Phi,(\hat{n})}(\omega; \tilde{\tau}) = 0$ and $B_{\Phi,(\hat{n})}(\omega; \tilde{\tau}) = 0$, depending on $A_{\Phi}$ and $B_{\Phi}$ 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.

Under the imposed assumptions, the conditional moment expression in proposition 2.1(i) results as a special case of proposition A.4 just established.

**Proof of proposition 2.1(i).** From proposition A.4, we obtain the exponentially affine form in equation (2.6) for the special case $\tau = 0$. Moreover, the coefficients in equation (A.11) simplify to

\begin{align}
A_{\Phi}(\omega; \tilde{\tau}, 0) &= A_{\Phi,(0)}(\omega; \tilde{\tau}) \\
B_{\Phi}(\omega; \tilde{\tau}, 0) &= B_{\Phi,(0)}(\omega; \tilde{\tau})
\end{align}

with $A_{\Phi,(i)}$ and $B_{\Phi,(i)}$ determined from the recursion in equation (A.12).

Additional assumptions are required to show proposition 2.1(ii). Under the conditions in assumption A.4, proposition A.4 in conjunction with proposition A.3 yields the unconditional moment expression in proposition 2.1(ii). Essentially, we then have $\Phi^{M}(\omega; \tilde{\tau}, \infty) = \lim_{\tau \to \infty} \Phi^{M}(\omega; \tilde{\tau}, \tau, z)$ irrespective of $z$.

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

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

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

**Proof of proposition 2.1(ii).** Under the imposed assumptions, proposition A.3 implies that $\Phi^{M}(\omega; \tilde{\tau}, \infty)$ can be determined as the limit when $\tau \to \infty$ in equation (A.10), leading to equation (2.7). Letting $\tau \to \infty$ in equation (A.11) thereby yields the associated coefficients

\begin{align}
A_{\Phi}(\omega; \tilde{\tau}, \infty) &= A_{\Phi,(0)}(\omega; \tilde{\tau}) + A_{\Phi}([0; B_{\Phi,(0)}(\omega; \tilde{\tau})]; \infty) \\
B_{\Phi}(\omega; \tilde{\tau}, \infty) &= 0,
\end{align}

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 $\chi = (A_{\mu,X}, B_{\mu,X}, A_{\Omega,X}, B_{\Omega,X}, A_{\Lambda}, B_{\Lambda}, \Phi_{\nu})$, suppressing the dependence of the elements of $\chi$ on $M$. 

27
A.3.1 Single-period pl-linear moments

For given $\omega = [\omega_S; \omega_Z] \in \mathbb{C}^{n_x}$, $\alpha = [\alpha_S; \alpha_Z] \in \mathbb{N}^{n_x}$, and $t, T \in \mathbb{R}_+$, define the complex-valued process $(\Psi^{(\alpha)}_t)_{0 \leq \tau \leq T}$ by

$$
\Psi^{(\alpha)}_\tau = \Psi_0 \sum_{Q(\alpha)} M^\alpha_\tau (A^{(\alpha)}_\psi (\omega; T - \tau) + [\omega^{(\alpha)}_S; B^{(\alpha)}_\psi (\omega; T - \tau)] \cdot X_{t+|\tau|})^k,
$$

(A.15)

setting $A^{(\beta)}_\psi = \partial^\beta_A \Phi$ and $B^{(\beta)}_\psi = \partial^\beta_B \Phi$ as well as $\omega^{(\beta)}_S = \partial^\beta_S \omega_S$ and $\omega^{(\beta)}_Z = \partial^\beta_Z \omega_Z$ for $\beta \in \mathbb{N}^{n_x}$. By Ito’s lemma applied to equation (A.15), $\Psi^{(\alpha)}_t$ follows dynamics of the form

$$
d\Psi^{(\alpha)}_\tau = \mu^{(\alpha)}_\Psi,\tau^- d\tau + \sigma^{(\alpha)}_\Psi,\tau^- dW_{t+\tau} + J^{(\alpha)}_\Psi dN_{t+\tau}.
$$

(A.16)

In essence, $\Psi^{(\alpha)}_t$ in equation (A.15) is constructed from $\Psi_T$ by applying the Faà di Bruno formula (A.1) to the right-hand side of equation (A.5). Under the regularity conditions in assumption A.5, by formal differentiation of equation (A.8), we have that the coefficient functions $A^{(\beta)}_\psi$ and $B^{(\beta)}_\psi$ for $\beta \leq \alpha$ solve the joint system of ODEs (A.18) of generalized Riccati type.

With this construction, well-behavedness of the characteristic $\chi$ in the sense of assumption A.5 assures that the process $(\Psi^{(\alpha)}_t)_{0 \leq \tau \leq T}$ 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 Dillenschneider (2020). For the specific dynamics in equation (2.1), these are obtained by formally differentiating equation (A.7), i.e., $\Psi^{M,(\alpha)}(\omega; \tau, z) = \partial^\alpha \Psi^M(\omega; \tau, z)$.

**Assumption A.5.** The characteristic $\chi$ 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) $\Phi^{(\beta)}_\psi (\omega)$ exists in an open set $O$ containing $\bigcup_{0 \leq \tau \leq T} \{[\omega_S; B^\beta(\omega; \tau)]\}$ for all $|\beta| \leq |\alpha|$;

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

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

- $E^M[|\Psi^{(\alpha)}_0|] < \infty$,
- $E^M[\int_0^T \Omega^{(\alpha)}_\Psi,\tau^- (d\tau)^{1/2}] < \infty$ with $\Omega^{(\alpha)}_\Psi,\tau^- = \sigma^{(\alpha)}_\Psi,\tau^- \sigma^{(\alpha)}_\Psi,\tau^T$,
- $E^M[\int_0^T |M^{(\alpha)}_\Psi,\tau^- [F_{t+\tau^-}] \Lambda(Z_{t+\tau^-})| d\tau] < \infty$.

**Proposition A.5.** Let $\chi$ 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 $0 \leq \tau \leq T$, we have

$$
\Psi^{M,(\alpha)}(\omega; \tau, Z_t) = E^M[\exp(\omega \cdot \tilde{X}_{t+|\tau|}) (\tilde{X}_{t+|\tau|})^\alpha | F_t] = \Psi^{M}(\omega; \tau, Z_t) \sum_{Q(\alpha)} M^\alpha_\tau (A^{(\alpha)}_\psi (\omega; \tau) + B^{(\alpha)}_\psi (\omega; \tau) \cdot Z_t)^k
$$

(A.17)

with coefficients $A^{(\beta)}_\psi (\omega; \tau) = \partial^{\beta}_\omega A^\beta(\omega; \tau) \in \mathbb{C}$ and $B^{(\beta)}_\psi (\omega; \tau) = \partial^{\beta}_\omega B^\beta(\omega; \tau) \in \mathbb{C}^{n_x}$ for $\beta \leq \alpha$ jointly.
determined by the system of ODEs
\[
\partial_t A_\psi^{(\beta)} = A_{\mu,X}^T [\omega_S^{(\beta)}; B_\psi^{(\beta)}] + \frac{1}{2} A_{\Omega,X}^T \sum_{\eta \leq \beta} \left( \sum_{\eta \leq \beta} M_{\beta,\beta}[\omega_S^{(\beta)}; B_\psi^{(\beta)}] \otimes [\omega_S^{(\beta-\eta)}; B_\psi^{(\beta-\eta)}] \right)
\]
\[
+ A_{\Lambda}^T \sum_{|\eta| \leq |\beta|} \Phi^{(\beta)}_{\nu} ([\omega_S; B_\psi]) \sum_{Q(\beta,\eta)} M_{\beta,\beta}^{\nu} ([\omega_S^{(\beta)}; B_\psi^{(\beta)}])^k - \delta_0(\beta) A_{\Lambda}^T \tag{A.18a}
\]
\[
\partial_t B_\psi^{(\beta)} = B_{\mu,X}^T [\omega_S^{(\beta)}; B_\psi^{(\beta)}] + \frac{1}{2} B_{\Omega,X}^T \sum_{\eta \leq \beta} \left( \sum_{\eta \leq \beta} M_{\beta,\beta}[\omega_S^{(\beta)}; B_\psi^{(\beta)}] \otimes [\omega_S^{(\beta-\eta)}; B_\psi^{(\beta-\eta)}] \right)
\]
\[
+ B_{\Lambda}^T \sum_{|\eta| \leq |\beta|} \Phi^{(\beta)}_{\nu} ([\omega_S; B_\psi]) \sum_{Q(\beta,\eta)} M_{\beta,\beta}^{\nu} ([\omega_S^{(\beta)}; B_\psi^{(\beta)}])^k - \delta_0(\beta) B_{\Lambda}^T \tag{A.18b}
\]
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, the derivatives \( \Phi^{(\beta)}_{\nu} = \partial_\nu \Phi_\nu \) determine the pl-linear moments of jump sizes.

**Proof.** See Dillschneider (2020). \( \square \)

For unconditional pl-linear moments, we proceed as for exponential moments. In addition to the stationarity assumption, we require the regularity conditions in assumption A.6 to obtain unconditional pl-linear moments, generalizing assumption A.2 for exponential moments. Analogous to the case of exponential moments, assumption A.6 allows to obtain unconditional pl-linear moments as a limit of the pl-linear moments, generalizing assumption A.2 for exponential moments. Likewise, we have from proposition A.3 that \( \Psi^{M,(\alpha)}(\omega; \tau, z) \) is well-behaved at \( \omega \), \( \alpha \in [0; \omega_Z] \) and \( \alpha = [0; \alpha_Z] \) by satisfying the following conditions:

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

(ii) \( \Psi^{M,(\beta)}([0; \omega_Z]; \infty) \) 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) \( \tilde{\omega}_Z \mapsto \Psi^{M,(\beta)}([0; \omega_Z]; \infty) \) 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 \( \tilde{\omega} = [0; \omega_Z] \in \mathbb{C}^{n_x} \) 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.6.** Let \( \chi \) 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

\[
\Psi^{M,(\alpha)}(\omega; \infty) = E^M [\exp(\omega \cdot X_t) (X_t)^{\alpha}] = \exp(A_\psi(\omega; \infty)) \sum_{Q^{(\alpha)}} M_{\beta,\beta}^{a} (A_\psi^{(\beta)}(\omega; \infty))^k \tag{A.19}
\]

with \( A_\psi^{(\beta)}(\omega; \infty) = \lim_{\tau \to \infty} A_\psi^{(\beta)}(\omega; \tau) \) for \( \beta \leq \alpha \) as in proposition A.5.

Proof. See Dillschneider (2020). \( \square \)
A.3.2 Multi-period pl-linear moments

As for exponential moments, additional regularity conditions allow to use the single-period moments in proposition A.5 to iteratively determine multi-period pl-linear moments. The conditions in assumption A.7 assure the applicability of the law of iterated expectations in this case. Analogous to proposition A.4, we can now state an auxiliary result that will be used to prove both statements in proposition 2.2. The result effectively states that $\Phi^{M,(\alpha)}(\omega; \bar{\tau}, \tau, z) = \partial_\alpha^\beta \Phi^M(\omega; \bar{\tau}, \tau, z)$.

**Assumption A.7.** The characteristic $\chi$ is well-behaved at $(\omega, \alpha, \tau, \tau) \in \mathbb{C}^{n x \hat{n}} \times \mathbb{R}^n_x \times \mathbb{R}^n_\tau \times \mathbb{R}_+$ for $\omega = [\omega^S; \omega^Z]$ and $\alpha = [\alpha^S; \alpha^Z]$ by satisfying the following conditions:

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

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

**Proposition A.7.** Let $\chi$ be well-behaved at $(\omega, \alpha, \tau, \tau) \in \mathbb{C}^{n x \hat{n}} \times \mathbb{R}^n_x \times \mathbb{R}^n_\tau \times \mathbb{R}_+$ 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

$$
\Phi^{M,(\alpha)}(\omega; \bar{\tau}, \tau, Z_t) = E^M[\exp(\omega \cdot \bar{X}_{t+\tau+\tau})](\bar{X}_{t+\tau+\tau})^\alpha |F_t]
= \Phi^M(\omega; \bar{\tau}, \tau, Z_t) \sum_{Q(\alpha)} M^\alpha_{k,\ell} (A^{(t)}_\Phi(\omega; \bar{\tau}, \tau) + B^{(t)}_\Phi(\omega; \bar{\tau}, \tau) \cdot Z_t)^k
$$

(A.20)

with coefficients $A^{(t)}_\Phi(\omega; \bar{\tau}, \tau) = \partial_\alpha^\beta A^{(t)}_\Phi(\omega; \bar{\tau}) \in \mathbb{C}$ and $B^{(t)}_\Phi(\omega; \bar{\tau}, \tau) = \partial_\alpha^\beta B^{(t)}_\Phi(\omega; \bar{\tau}, \tau) \in \mathbb{C}^{n z}$ for $\beta \leq \alpha$ given by

$$
A^{(t)}_\Phi(\omega; \bar{\tau}) = A^{(t)}_{\Phi,(0)}(\omega; \bar{\tau}) + \sum_{|\eta| \leq |\beta|} A^{(t)}_{\Phi,(\eta)}([0; B_{\Phi,(0)}(\omega; \bar{\tau})]; \tau) \sum_{Q(\beta, \eta)} M^\beta_{k,\ell} ([0; B_{\Phi,(0)}(\omega; \bar{\tau})])^k
$$

(A.21a)

and

$$
B^{(t)}_\Phi(\omega; \bar{\tau}, \tau) = \sum_{|\eta| \leq |\beta|} B^{(t)}_{\Phi,(\eta)}([0; B_{\Phi,(0)}(\omega; \bar{\tau})]; \tau) \sum_{Q(\beta, \eta)} M^\beta_{k,\ell} ([0; B_{\Phi,(0)}(\omega; \bar{\tau})])^k
$$

(A.21b)

Defining $\Delta_i = \bar{\tau}_i - \bar{\tau}_{i-1}$ and $\omega_{(i)} = \partial_\alpha^\beta \omega_i$ for $\omega_i = [\omega^S; \omega^i; \omega^Z]$, the auxiliary coefficients $A^{(t)}_{\Phi,(i)}(\omega; \bar{\tau}) = \partial_\alpha^\beta A^{(t)}_{\Phi,(i)}(\omega; \bar{\tau}) \in \mathbb{C}$ and $B^{(t)}_{\Phi,(i)}(\omega; \bar{\tau}) = \partial_\alpha^\beta B^{(t)}_{\Phi,(i)}(\omega; \bar{\tau}) \in \mathbb{C}^{n z}$ for $\beta \leq \alpha$ are determined by the backward recursion

$$
A^{(t)}_{\Phi,(i-1)}(\omega; \bar{\tau}) = A^{(t)}_{\Phi,(i)}(\omega; \bar{\tau}) + \sum_{|\eta| \leq |\beta|} A^{(t)}_{\Phi,(\eta)}([0; B_{\Phi,(0)}(\omega; \bar{\tau})]; \Delta_i) \sum_{Q(\beta, \eta)} M^\beta_{k,\ell} ([0; B_{\Phi,(0)}(\omega; \bar{\tau})])^k
$$

(A.22a)

and

$$
B^{(t)}_{\Phi,(i-1)}(\omega; \bar{\tau}) = \sum_{|\eta| \leq |\beta|} B^{(t)}_{\Phi,(\eta)}([0; B_{\Phi,(0)}(\omega; \bar{\tau})]; \Delta_i) \sum_{Q(\beta, \eta)} M^\beta_{k,\ell} ([0; B_{\Phi,(0)}(\omega; \bar{\tau})])^k
$$

(A.22b)

for $i = \bar{n}, \ldots, 1$ subject to the initial conditions $A^{(t)}_{\Phi,(0)}(\omega; \bar{\tau}) = 0$ and $B^{(t)}_{\Phi,(0)}(\omega; \bar{\tau}) = 0$, depending on $A^{(t)}_{\Phi}$ and $B^{(t)}_{\Phi}$ as in proposition A.5.

**Proof.** Repeatedly invoking the law of iterated expectations and proposition A.5, we obtain that $\Phi^{M,(\beta)}(\omega; \bar{\tau}, \tau, z) = \partial_\alpha^\beta \Phi^M(\omega; \bar{\tau}, \tau, z)$. Proposition A.4 and the Faa 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.

Under the imposed assumptions, the conditional moment expression in proposition 2.2(i) results as a
special case of proposition A.7 just established.

**Proof of proposition 2.2(i).** From proposition A.7, we obtain the exponentially affine form in equation (2.8) for the special case $\tau = 0$. Moreover, the coefficients in equation (A.21) simplify to

$$A_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, 0) = A_{\Phi, (0)}^{(\beta)}(\omega; \tilde{\tau})$$  \hspace{1cm} (A.23a)

$$B_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, 0) = B_{\Phi, (0)}^{(\beta)}(\omega; \tilde{\tau})$$  \hspace{1cm} (A.23b)

with $A_{\Phi, (i)}^{(\beta)}$ and $B_{\Phi, (i)}^{(\beta)}$ determined from the recursion in equation (A.22).

Additional assumptions are required to show proposition 2.2(ii). Under the conditions in assumption A.8, proposition A.7 in conjunction with proposition A.6 yields the unconditional moment expression in proposition 2.2(ii). Essentially, we have $\Phi^{M, (\alpha)}(\omega; \tilde{\tau}, \infty) = \lim_{\tau \to \infty} \Phi^{M, (\alpha)}(\omega; \tilde{\tau}, \tau, z)$ irrespective of $z$.

**Assumption A.8.** The characteristic $\chi$ is well-behaved at $(\omega, \alpha, \tilde{\tau}) \in C^\alpha \times R^m$ for all $\alpha = [\alpha_S; \alpha_Z]$ by satisfying the following conditions:

(i) $\chi$ 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 $|\beta_i, Z| \leq \sum_{j=1}^{\tilde{n}} |\alpha_j|$ and all $1 \leq i \leq \tilde{n}$;

(ii) $\chi$ 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 $|\beta_0, Z| \leq \sum_{j=1}^{\tilde{n}} |\alpha_j|$.

**Proof of proposition 2.2(ii).** Under the imposed assumptions, proposition A.6 implies that $\Phi^{M, (\alpha)}(\omega; \tilde{\tau}, \infty)$ can be determined as the limit when $\tau \to \infty$ in equation (A.20), leading to equation (2.9). Letting $\tau \to \infty$ in equation (A.21) thereby yields the associated coefficients

$$A_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, \infty) = A_{\Phi, (0)}^{(\beta)}(\omega; \tilde{\tau}) + \sum_{|\eta| \leq |\beta|} A_{\Phi}^{(\eta)}([0; B_{\Phi}(\omega; \tilde{\tau})]; \infty) \sum_{\mathcal{Q}(\beta, \eta)} M_{\mathcal{h}, \mathcal{Q}}^{(\beta)}([0; B_{\Phi, (0)}^{(\beta)}(\omega; \tilde{\tau})])^k$$  \hspace{1cm} (A.24a)

$$B_{\Phi}^{(\beta)}(\omega; \tilde{\tau}, \infty) = 0$$,  \hspace{1cm} (A.24b)

where $A_{\Phi, (i)}^{(\beta)}$ and $B_{\Phi, (i)}^{(\beta)}$ are given by the recursion in equation (A.22).

**B**  

**B.1 Basic Schwartz distribution theory**

**Schwartz space.** Denote by $S(R^m)$ the complex-valued Schwartz space of rapidly decaying functions (e.g., Grubb (2009)). The space $S(R^m)$ will usually be referred to simply as *the* Schwartz space and its elements will regularly be called Schwartz functions. These can be formally characterized as smooth functions (i.e., elements of $C^\infty(R^m)$) with the additional property that all derivatives decay faster at infinity than any polynomial. The latter property requires that each semi-norm $\|\cdot\|_{\alpha, \beta}$ defined through

$$\|f\|_{\alpha, \beta} = \sup_{y \in R^m} |y^\alpha \partial^\beta_y f(y)|$$

is finite for any multi-indices $\alpha, \beta \in N^m$. Convergence in the Schwartz space may be characterized by separate convergence with respect to each of the semi-norms.
Tempered distributions. The space of tempered distributions $\mathcal{S}^*(\mathbb{R}^m)$ is defined as the continuous dual space of $\mathcal{S}(\mathbb{R}^m)$, which is the space of all continuous linear functionals on the test function space $\mathcal{S}(\mathbb{R}^m)$. We use the duality pairing notation $\langle g(y), f(y) \rangle$ to denote the action of a tempered distribution $g \in \mathcal{S}^*(\mathbb{R}^m)$ on a Schwartz function $f \in \mathcal{S}(\mathbb{R}^m)$. A tempered distribution $g$ is called regular if it coincides with a locally integrable\(^{17}\) function such that

$$\langle g(y), f(y) \rangle = \int_{\mathbb{R}^m} g(y) f(y) \, dy$$

for all Schwartz functions $f$; otherwise $g$ is called singular. Standard examples of regular tempered distributions are those generated by $L^p$ functions as well as by continuous functions with at most polynomial growth. Probably the most prominent example of a singular tempered distribution is the Dirac delta functional $\delta$, defined by the assignment $\langle \delta(y), f(y) \rangle = f(0)$.

Several standard operations may be defined for tempered distributions. Addition and scalar multiplication are defined in natural ways. Moreover, for a slowly increasing\(^{18}\) function $h \in C^\infty(\mathbb{R}^m)$ and $g \in \mathcal{S}^*(\mathbb{R}^m)$, we define the product $h g \in \mathcal{S}^*(\mathbb{R}^m)$ via $\langle h(y) g(y), f(y) \rangle = \langle g(y), h(y) f(y) \rangle$ for every Schwartz function $f$.

Fourier transforms. Fourier transforms may be defined for Schwartz functions as well as for tempered distributions. For any Schwartz function $f \in \mathcal{S}(\mathbb{R}^m)$, the associated Fourier transform $\hat{f} = \mathcal{F}f \in \mathcal{S}(\mathbb{R}^m)$ exists in the ordinary sense, defined by

$$\hat{f}(y) = \mathcal{F}f(y) = \int_{\mathbb{R}^m} f(\tilde{y}) \exp(-iy \cdot \tilde{y}) \, d\tilde{y}.$$  

To define the Fourier transform $\hat{g} = \mathcal{F}g \in \mathcal{S}^*(\mathbb{R}^m)$ of a tempered distribution $g \in \mathcal{S}^*(\mathbb{R}^m)$, one sets $\langle \hat{g}(y), f(y) \rangle = \langle g(y), \hat{f}(y) \rangle$ for every Schwartz function $f \in \mathcal{S}(\mathbb{R}^m)$.

Tensor products. 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 $\mathcal{S}(\mathbb{R}^m \times \mathbb{R}^n)$. By construction, it follows from $f \in \mathcal{S}(\mathbb{R}^m \times \mathbb{R}^n)$ that both $(z \mapsto f(y, z)) \in \mathcal{S}(\mathbb{R}^n)$ and $(y \mapsto f(y, z)) \in \mathcal{S}(\mathbb{R}^m)$ for fixed $y$ and $z$, respectively. Expressions such as $\langle g(y), f(y, z) \rangle$ and $\langle h(z), f(y, z) \rangle$ for $g \in \mathcal{S}^*(\mathbb{R}^m)$ and $h \in \mathcal{S}^*(\mathbb{R}^n)$ are therefore well-defined. Moreover, it can be shown that the resulting functions are again Schwartz functions in $\mathcal{S}(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^m)$, respectively. We may thus proceed to compute $\langle h(z), \langle g(y), f(y, z) \rangle \rangle$ and $\langle g(y), \langle h(z), f(y, z) \rangle \rangle$. In fact, we may consistently define the distributional tensor product $(y, z) \mapsto g(y) \otimes h(z)) \in \mathcal{S}^*(\mathbb{R}^m \times \mathbb{R}^n)$ with $\langle g(y) \otimes h(z), f(y, z) \rangle$ determined by the equivalent expressions

$$\langle g(y), \langle h(z), f(y, z) \rangle \rangle = \langle h(z), \langle g(y), f(y, z) \rangle \rangle.$$

As such, this equality may be considered a distributional analogue to the classical Fubini integral theorem.

B.2 General derivatives

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

---

\(^{17}\)A measurable function is called locally integrable over $\mathbb{R}^m$ if it is integrable over every compact subset of $\mathbb{R}^m$.

\(^{18}\)A smooth function is called slowly increasing (at infinity) if each of its derivatives is bounded by some polynomial.
Proposition B.1. Let \( g \in S^*(\mathbb{R}) \) and \( (y \mapsto \Pi(\omega + iy\hat{\omega}; \hat{T}, z)) \in S(\mathbb{R}) \) for all \( z \in \mathbb{Z} \). Then

\[
\Pi_g(\omega, \hat{\omega}; \hat{T}, Z_t) = E^Q[D_t(\hat{T}_n) \exp(\omega \cdot \hat{X}_{t+T}) g(\hat{\omega} \cdot \hat{X}_{t+T}) | \mathcal{F}_t] = \frac{1}{2\pi} \langle \hat{g}(y), \Pi(\omega + iy\hat{\omega}; \hat{T}, Z_t) \rangle
\]

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

Proof. See Chen and Joslin (2012).

Proof of proposition 3.1. Use the definition of pricing transform \( \Pi \) in equation (3.2). By the imposed assumptions, proposition B.1 yields

\[
\mathcal{V}(K, \hat{T}; Z_t) = \frac{1}{2\pi} \sum_{i=1}^{n_h} \langle \hat{g}(y_i; K), \Pi(u([\omega_i; \bar{y}]); \hat{T}, Z_t) \rangle. \tag{B.2}
\]

Since \( \mathcal{V} \subset \mathbb{R}^{n \times 1} \), there exists some \( \hat{\Pi} \) with \( (y \mapsto \hat{\Pi}(u(y); \hat{T}, z)) \in S(\mathbb{R}^{n \times 1}) \) that coincides with \( \Pi \) on \( \mathcal{V} \). As the support of each \( g_i \) is contained in \( \mathcal{V} \), we can rewrite equation (B.2) as

\[
\mathcal{V}(K, \hat{T}; Z_t) = \frac{1}{2\pi} \sum_{i=1}^{n_h} \langle \delta(\omega - \bar{\omega}_i) \hat{g}(y_i; K), \hat{\Pi}(u([\bar{\omega}_i; \bar{y}])); \hat{T}, Z_t) \rangle \tag{B.3}
\]

By construction, the support of \( (y \mapsto w(y; K)) \in S^*(\mathbb{R}^{n \times 1}) \) is contained in \( \mathcal{V} \), so that equation (B.3) thus yields equation (3.3).

B.3 Equity derivatives

Proof of corollary 3.1. For the call payoff in equation (3.5a), define \( (\bar{y} \mapsto g^C_{stock}(\bar{y}; K)) \in S^*(\mathbb{R}) \) by \( g^C_{stock}(\bar{y}; K) = U(\bar{y} - K) \). The associated distributional Fourier transform \( (\bar{y} \mapsto \hat{g}^C_{stock}(\bar{y}; K)) \in S^*(\mathbb{R}) \) is given by

\[
\hat{g}^C_{stock}(\bar{y}; K) = \frac{1}{2} \delta(\bar{y}) - \frac{1}{2\pi} i \exp(-iK \bar{y}) \bar{y}^{-1},
\]

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

Proceeding analogously for the put payoff in equation (3.5b), define \( (\bar{y} \mapsto g^P_{stock}(\bar{y}; K)) \in S^*(\mathbb{R}) \) by \( g^P_{stock}(\bar{y}; K) = U(K - \bar{y}) \) with associated distributional Fourier transform \( (\bar{y} \mapsto \hat{g}^P_{stock}(\bar{y}; K)) \in S^*(\mathbb{R}) \) given by

\[
\hat{g}^P_{stock}(\bar{y}; K) = -\frac{1}{2} \delta(\bar{y}) - \frac{1}{2\pi} i \exp(-iK \bar{y}) \bar{y}^{-1},
\]

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

Proof of lemma 3.1. Following Gelfand and Shilov (1964), we define the tempered distribution
\( \hat{y}^{-1} \in S^*(\mathbb{R}) \) for every Schwartz function \( f \in S(\mathbb{R}) \) by the convergent integral

\[
\langle \hat{y}^{-1}, f(\hat{y}) \rangle = \int_{\mathbb{R}^+} \frac{\Delta_y^{(1)} f(\hat{y})}{\hat{y}} \, d\hat{y},
\]

using the regularization \( \Delta_y^{(1)} f(\hat{y}) = f(\hat{y}) - f(-\hat{y}) \). It is straightforward to verify that this formulation is equivalent to the one in terms of a Cauchy principal value integral. To avoid redundancies, write \((y \mapsto w^O_{stock}(y; K)) \in S^*(\mathbb{Y}) \) in equation (3.7) compactly as

\[
w^O_{stock}([\hat{\omega}; \hat{y}]; K) = \frac{1}{2\pi} \left( \delta([\hat{\omega} - [1; 0]) - \exp(K) \delta([\hat{\omega}]) \right) \otimes \hat{g}^O_{stock}(\hat{y}; K)
\]

and define the tempered distribution \( \hat{g}^O_{stock}(\hat{y}; K) \in S^*(\mathbb{R}) \) in equation (B.4) by

\[
\hat{g}^O_{stock}(\hat{y}; K) = c^O_{stock} \pi \delta(\hat{y}) - i \exp(-iK\hat{y}) \hat{y}^{-1}.
\]

If \( Y(u(y)) \) is Hermitian as function of \( \hat{y} \), then so is \( F_{\text{stock}}(\hat{y}; K) \) \( Y(u(y)) \) for \( F_{\text{stock}}(\hat{y}; K) = \frac{1}{2} \exp(-iK\hat{y}) \). Since for Hermitian \( f \), we have that \( \Delta_y^{(1)} f(\hat{y}) = 2\mathfrak{R}f(\hat{y}) \), the tempered distribution \( \hat{g}^O_{stock} \) in equation (B.5) has the integral representation

\[
\frac{1}{2\pi} \left( \hat{g}^O_{stock}(\hat{y}; K), Y(u_{\text{stock}}([\hat{\omega}; \hat{y}])) \right) = \frac{c^O_{stock}}{2} \mathfrak{R}(\hat{\omega}) - \int_{\mathbb{R}^+} \Delta_y^{(1)} \left( \frac{1}{\hat{y}} \mathfrak{I}(F_{\text{stock}}(\hat{y}; K) \mathfrak{R}(\hat{\omega} + i\hat{y}[1; 0])) \right) \, d\hat{y}
\]

An application of the definition of the distributional tensor product in equation (B.1) to \( w^O_{stock} \) in terms of \( \hat{g}^O_{stock} \) in equation (B.4) then yields equation (3.8), as required. \( \square \)

### 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

\[
d[\log S]^*_t = \Omega_S(Z_{t^-}) \, dt + 2 \left( \exp(J_{S,t}) - i^T - J_{S,t} \right) dN_t.
\]

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

\[
\text{VIX}^2_t = \frac{1}{\tau_{\text{vix}}} \mathbb{E}^Q[[i \log S]^*_t + \tau_{\text{vix}} - [\log S]^*_t] \big| F_t
\]

(B.6)

for \( \text{VIX}^2_t \) defined in equation (3.9).

It remains to establish an expression for the expected value on the right-hand side of equation (B.6). For this, we obtain

\[
\mathbb{E}^Q[[i \log S]^*_t + \tau_{\text{vix}} - [\log S]^*_t] \big| F_t = \int_0^{\tau_{\text{vix}}} C_{[\log S]^*} + D_{[\log S]^*} \mathbb{E}^Q[|Z_{t+\tau} - F_t|] \, d\tau,
\]

(B.7)

34
where the coefficients $C_{\log S}\ast \in \mathbb{R}$ and $D_{\log S}\ast \in \mathbb{R}^{1\times n_z}$ are given by

$$
\begin{align*}
C_{\log S}\ast &= A_{\log S} + 2E^{(\ast)}[\exp(J_{S,t}) - \tau - J_{S,t}]A_{\lambda}^Q \\
D_{\log S}\ast &= B_{\log S} + 2E^{(\ast)}[\exp(J_{S,t}) - \tau - J_{S,t}]B_{\lambda}^Q .
\end{align*}
$$

(B.8a) (B.8b)

We moreover have

$$
E^{(\ast)}[Z_{t+\tau} | F_t] = [\exp(\tau D_Z) - I]D_Z^{-1}C_Z + \exp(\tau D_Z)Z_t
$$

(B.9)

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

$$
\begin{align*}
C_Z &= A_{\log \mu, Z} + E^{(\ast)}[\gamma_t]A_{\lambda}^Q \\
D_Z &= B_{\log \mu, Z} + E^{(\ast)}[\gamma_t]B_{\lambda}^Q .
\end{align*}
$$

(B.10a) (B.10b)

Substituting equation (B.9) into equation (B.7) and performing the integration yields the coefficients $a_{\text{vix}} \in \mathbb{R}$ and $b_{\text{vix}} \in \mathbb{R}^{n_z}$ in equation (3.10) by

$$
\begin{align*}
a_{\text{vix}} &= C_{\log S}\ast + \frac{1}{\tau_{\text{vix}}}D_{\log S}\ast \left(\exp(\tau_{\text{vix}} D_Z) - I - \tau_{\text{vix}} D_Z\right)D_Z^{-2}C_Z \\
b_{\text{vix}} &= \frac{1}{\tau_{\text{vix}}} (D_{\log S}\ast \left(\exp(\tau_{\text{vix}} D_Z) - I\right) D_Z^{-1})^\top,
\end{align*}
$$

(B.11a) (B.11b)

depending on the coefficients $C_{\log S}\ast$ and $D_{\log S}\ast$ in equation (B.8) as well as $C_Z$ and $D_Z$ in equation (B.10).

\[ \square \]

**Proof of corollary 3.2.** We follow the derivations in Dillschneider (2020). For the call payoff in equation (3.11a), define $(\tilde{y} \mapsto g_{\text{vix}}^C(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ by $g_{\text{vix}}^C(\tilde{y}; K) = (\tilde{y}^{1/2} - K^{1/2}) U(\tilde{y} - K)$. We obtain the associated distributional Fourier transform $(\tilde{y} \mapsto \hat{g}_{\text{vix}}^C(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ as

$$
\hat{g}_{\text{vix}}^C(\tilde{y}; K) = \frac{1}{2} (-\text{sgn}(\tilde{y}))^{1/2} \Gamma(1/2, iK\tilde{y}) |\tilde{y}|^{-3/2},
$$

in terms of the tempered distribution $|\tilde{y}|^{-3/2}$, where $\Gamma$ denotes the upper incomplete Gamma function. Setting $\tilde{\omega}_1 = [0; 0]$, $\tilde{\omega} = [0; b_{\text{vix}}]$, and $g_1(\tilde{y}; K) = g_{\text{vix}}^C(\tilde{y}; K)$, exploiting the shift property\(^{19}\) of the Fourier transform, this yields equation (3.13a) as a special case of proposition 3.1.

Proceeding analogously for the put payoff in equation (3.11b), define $(\tilde{y} \mapsto g_{\text{vix}}^P(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ by $g_{\text{vix}}^P(\tilde{y}; K) = (K^{1/2} - \tilde{y}^{1/2}) U(K - \tilde{y})$, with associated distributional Fourier transform $(\tilde{y} \mapsto \hat{g}_{\text{vix}}^P(\tilde{y}; K)) \in \mathcal{S}^*(\mathbb{R})$ given by

$$
\hat{g}_{\text{vix}}^P(\tilde{y}; K) = 2\pi K^{1/2} \delta(\tilde{y}) - \frac{1}{2} (-\text{sgn}(-1))^{1/2} \gamma(1/2, iK\tilde{y}) |\tilde{y}|^{-3/2},
$$

again in terms of the tempered distribution $|\tilde{y}|^{-3/2}$, where $\gamma$ denotes the lower incomplete Gamma function. Setting $\tilde{\omega}_1 = [0; 0]$, $\tilde{\omega} = [0; b_{\text{vix}}]$, and $g_1(\tilde{y}; K) = g_{\text{vix}}^P(\tilde{y}; K)$, again exploiting the shift property of the Fourier transform, this yields equation (3.13b) as a special case of proposition 3.1.

\[ \square \]

**Proof of lemma 3.3.** Following Gel’fand and Shilov (1964), we define the tempered distribution $|\tilde{y}|^{-3/2} \in \mathcal{S}^*(\mathbb{R})$ for every Schwartz function $f \in \mathcal{S}(\mathbb{R})$ by the convergent integral

$$
\langle |\tilde{y}|^{-3/2}, f(\tilde{y}) \rangle = \int_{\mathbb{R}_+} \frac{\Delta_{\tilde{y}}^{(2)} f(\tilde{y})}{\tilde{y}^{3/2}} \, d\tilde{y},
$$

using the regularization $\Delta_{\tilde{y}}^{(2)} f(\tilde{y}) = f(\tilde{y}) + f(-\tilde{y}) - f(+0) - f(-0)$, where $f(\pm 0) = \lim_{\epsilon \downarrow 0} f(\pm \epsilon)$. For

\(^{19}\)By the shift property, $g(\tilde{y}) = f(a + \tilde{y})$ has Fourier transform $\hat{g}(\tilde{y}) = \exp(ia\tilde{y}) \hat{f}(\tilde{y})$. For
Then, as in Dillschneider (2020), we can extend the notion of a Schwartz space to include also functions that are \( q \) whenever there exists a positive weighting function \( \beta \) for \( (y \mapsto w_{\text{vix}}^V(y; K)) \in \mathcal{S}^*(\mathcal{Y}) \) in equation (3.13), we obtain

\[
\langle w_{\text{vix}}^V(y; K), \Upsilon(u_{\text{vix}}(y)) \rangle = \frac{1 - c_{\text{vix}}^V}{2} K^{1/2} \Upsilon([0; 0]) \quad \text{and} \quad \mathcal{Y} \in \mathcal{Y} \times \mathcal{Z}^n.
\]

Hence, \( \Delta_{\tilde{p}}^{(2)} f(\tilde{y}) = \Delta_{\tilde{p}}^{(2)} (2\Re f(\tilde{y})) \) when defining \( \Delta_{\tilde{p}}^{(2)} f(\tilde{y}) = f(\tilde{y}) - f(+0) \). Using \((y \mapsto w_{\text{vix}}^V(y; K)) \in \mathcal{S}^*(\mathcal{Y}) \) in equation (3.13), we obtain

\[
\langle w_{\text{vix}}^V(y; K), \Upsilon(u_{\text{vix}}(y)) \rangle = \frac{1 - c_{\text{vix}}^V}{2} K^{1/2} \Upsilon([0; 0]) + \int_{\mathbb{R}^+} \Delta_{\tilde{p}}^{(2)} \Re (F_{\text{vix}}^V(\tilde{y}; K) \Upsilon(i\tilde{y}[0; b_{\text{vix}}])) \, d\tilde{y}.
\]

and, thus, equation (3.14).

\[
\Box
\]

C Moments involving derivatives prices

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

C.1 Expectations of generalized transforms

As in Dillschneider (2020), we can extend the notion of a Schwartz space to include also functions that are Schwartz only after appropriate regularization. To be specific, consider a smooth function \((y, z) \mapsto \Upsilon(y; z)\) on \( \mathcal{Y} \times \mathcal{Z}_n \). We define an extended Schwartz space \( \hat{S} \) such that \((y, z) \mapsto \Upsilon(y; z)) \in \mathcal{S}(\mathcal{Y} \times \mathcal{Z}_n; \mathbb{1} \otimes q) \) whenever there exists a positive weighting function \( q \) on \( \mathcal{Z}_n \) so that \((y, z) \mapsto q(z) \Upsilon(y; z)) \in \mathcal{S}(\mathcal{Y} \times \mathcal{Z}_n).

Given this construction, we define

\[
\psi(z) = \langle \tilde{g}(y), \Upsilon(y; z) \rangle
\]

for \( \tilde{g} \in \mathcal{S}^*(\mathcal{Y}) \) and aim at determining moments involving \( \psi(Z_{t+\hat{r}}) \) and functions of the state vector. From Dillschneider (2020), we have the following Fubini-type result that allows to interchange the order of the tempered distribution and the expectation operator.

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

\[
E^M[f(\tilde{X}_{t+\hat{r}})] \cdot q(Z_{t+\hat{r}})^{-1} < \infty.
\]

Then

\[
E^M[f(\tilde{X}_{t+\hat{r}}) \psi(Z_{t+\hat{r}})] = E^M[f(\tilde{X}_{t+\hat{r}}) \psi(y, \Upsilon(y; Z_{t+\hat{r}}}]]
= \langle \tilde{g}(y), E^M[f(\tilde{X}_{t+\hat{r}}) \Upsilon(y; Z_{t+\hat{r}})] \rangle.
\]

**Proof.** The first equality holds by definition of \( \psi \), the second equality follows from the results of Dillschneider (2020).

\[
\Box
\]

C.2 Exact moments

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

\[
V_{t+\tau_j} = V_t(K_t, \tilde{T}_t; Z_{t+\tau_j}) = (w_t(y_{i,j}; \mathbb{t}); \Pi(u_{i,j}; \tilde{T}_t, Z_{t+\tau_j}))
\]

for \((y_{i,j} \mapsto w_t(y_{i,j}; K_t)) \in \mathcal{S}^*(\mathcal{Y}_i) \) and \((y_{i,j} \mapsto \Pi(u_{i,j}; \tilde{T}_t, Z_{t+\tau_j})) \in \mathcal{S}(\mathcal{Y}_i) \). Given multi-indices \( \beta_j \in \mathbb{N}^{\nu_j} \), define associated index vectors \( q(\beta_j) \in \mathbb{N}^{5|\beta_j|} \) with multiplicities according to \( \beta_j \) such that...
\[(V_{t+\tilde{t}})^{\beta_j} = \prod_{i=1}^{[\beta_j]} V_{i,\beta_j}^{(t+\tilde{t})} . \] By definition of tensor products of tempered distributions in equation (B.1), we thus have
\[(V_{t+\tilde{t}})^{\beta_j} = \langle w^{\beta_j}(y_j; K), \Pi^{\beta_j}(u(y_j); \tilde{T}, Z_{t+\tilde{t}}) \rangle \] (C.1)
for \((y_j \mapsto w^{\beta_j}(y_j; K)) \in S^*(\mathcal{Y}^{\beta_j}) \) and \((y_j \mapsto \Pi^{\beta_j}(u(y_j); \tilde{T}, Z_{t+\tilde{t}})) \in S(\mathcal{Y}) \) on the Cartesian product space \(\mathcal{Y}^{\beta_j} = \prod_{i=1}^{[\beta_j]} \mathcal{Y}_{i,\beta_j} \). Each tempered distributions \(w^{\beta_j} \) in equation (C.1) is given by the distributional tensor product
\[w^{\beta_j}(y_j; K) = \bigotimes_{i=1}^{[\beta_j]} w_{i,\beta_j}(y_{i,j}; K_{i,\beta_j}) \] (C.2)
and each \(\Pi^{\beta_j} \) in equation (C.1) by
\[\Pi^{\beta_j}(u(y_j); \tilde{T}, Z_{t+\tilde{t}}) = \prod_{i=1}^{[\beta_j]} \Pi_{i,\beta_j}(u_{i,\beta_j}(y_{i,j}); \tilde{T}_{i,\beta_j}, Z_{t+\tilde{t}}) \] (C.3)

Moreover, the coefficients \(A^{\beta_j}_I \) and \(B^{\beta_j}_I \) in equation (C.3) are determined as
\[A^{\beta_j}_I(u(y_j); \tilde{T}) = \sum_{i=1}^{[\beta_j]} A_{i,\beta_j}(u_{i,\beta_j}(y_{i,j}); \tilde{T}_{i,\beta_j}) \] (C.4a)
\[B^{\beta_j}_I(u(y_j); \tilde{T}) = \sum_{i=1}^{[\beta_j]} B_{i,\beta_j}(u_{i,\beta_j}(y_{i,j}); \tilde{T}_{i,\beta_j}) \] (C.4b)
This completes the first step.

Second, we combine the expressions relating to \((V_{t+\tilde{t}})^{\beta_j}\) to obtain analogous expressions relating to \((V_{t+\tilde{t}})^{\beta} = \prod_{j=1}^{\tilde{n}} (V_{t+\tilde{t}})^{\beta_j}\). Using \(w^{\beta_j}\) from equation (C.2) and again invoking the definition of the distributional tensor product in equation (B.1) yields equation (4.2) with \(w^{\beta}\) defined by the distributional tensor product
\[w^{\beta}(y; K) = \bigotimes_{j=1}^{\tilde{n}} w^{\beta_j}(y_j; K) \] (C.5)
where \((y \mapsto w^{\beta}(y; K)) \in S^*(\mathcal{Y}^{\beta}) \) on the Cartesian product space \(\mathcal{Y}^{\beta} = \prod_{j=1}^{\tilde{n}} \mathcal{Y}^{\beta_j} \). Moreover, \(\Pi^{\beta}\) in equation (4.2) is given by
\[\Pi^{\beta}(u(y); \tilde{T}, Z_{t+\tilde{t}}) = \prod_{j=1}^{\tilde{n}} \Pi^{\beta_j}(u(y_j); \tilde{T}, Z_{t+\tilde{t}}) \] (C.6)
with \((y \mapsto \Pi^{\beta}(u(y); \tilde{T}, Z_{t+\tilde{t}})) \in S(\mathcal{Y}^{\beta}) \). Since each of the \(\Pi^{\beta_j}\) in equation (C.3) is exponentially affine in \(Z_{t+\tilde{t}}\), it follows that \(\Pi^{\beta}\) in equation (C.6) is exponentially affine in \(Z_{t+\tilde{t}}\), yielding equation (4.3) with coefficients \(A^{\beta}_I \) and \(B^{\beta}_I \) given as
\[A^{\beta}_I(u(y); \tilde{T}) = \sum_{j=1}^{\tilde{n}} A^{\beta_j}_I(u(y_j); \tilde{T}) \] (C.7a)
\[B^{\beta}_I(u(y); \tilde{T}) = [B^{\beta_1}_I(u(y_1); \tilde{T}); \ldots; B^{\beta_{\tilde{n}}}_I(u(y_{\tilde{n}}); \tilde{T})] \] (C.7b)
in terms of the coefficients \(A^{\beta_j}_I \) and \(B^{\beta_j}_I \) in equation (C.4). This concludes the second step and, hence, the proof. \(\square\)
Proof of proposition 4.1. To ease notation, we introduce

$$F_{t+\tilde{t}}^{(\alpha)}(\omega) = \exp(\omega \cdot \tilde{X}_{t+\tilde{t}}) (\tilde{X}_{t+\tilde{t}})^{\alpha},$$  \hspace{1cm} (C.8)

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

$$\Phi^{M,(\alpha,\beta)}(\omega, 0; \tilde{\tau}, \infty) = E^M[F_{t+\tilde{t}}^{(\alpha)}(\omega) (V_{t+\tilde{t}})^{\beta}] = E^M[(u^{\beta}(y; K), V^\beta(u(y); \tilde{T}, Z_{t+\tilde{t}}) F_{t+\tilde{t}}^{(\alpha)}(\omega))]$$ \hspace{1cm} (C.9)

the last equality following from proposition C.1.

Finally, using the exponentially affine form of $V^\beta$ in equation (4.3) and the definition of $F_{t+\tilde{t}}^{(\alpha)}$ in equation (C.8), we arrive at

$$E^M[V^\beta(u(y); \tilde{T}, Z_{t+\tilde{t}}) F_{t+\tilde{t}}^{(\alpha)}(\omega)] = \exp(A^{(\beta)}_\Pi(u(y); \tilde{T})) E^M[F_{t+\tilde{t}}^{(\alpha)}(\omega + [0; B^{(\beta)}_\Pi(u(y); \tilde{T})]))]$$ \hspace{1cm} (C.10)

with $\Phi^{M,(\alpha)}$ as in equation (2.9). Substituting equation (C.10) into equation (C.9) yields equation (4.4), as required. \hfill \Box

C.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

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

Moreover, using $V_{t+\tilde{t},(\rho)} = \sum_{j=1}^n e_j \otimes V_{t+\tilde{t},j,(\rho)}$ with $e_j \in \mathbb{N}^n$ and $V_{t+\tilde{t},j,(\rho)}$ 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 Faa di Bruno formula (A.1) and $b_{V,\gamma,(\rho)}$ from equation (4.7), it moreover follows that equation (4.8) holds with coefficients

$$b^{(\beta)}_{V,\gamma,(\rho)} = \frac{1}{\eta!} \sum_{|\rho| \leq |\gamma|} \frac{\beta!}{(\beta - \rho)!} b^{\beta-\rho}_{V,0,(\rho)} \sum_{\phi(\rho,\eta)} M^{\rho}_{\phi,\ell}(b_{V,\ell,(\eta)}(t))^{k}.$$ \hspace{1cm} (C.11)

Equation (4.9) is then a straightforward consequence of the form of $(V_{t+\tilde{t},(\rho)}^{\beta})$ 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 $L^1(\mathcal{Z},M)$ to be the space of integrable functions on $\mathcal{Z}$ against the probability measure $M$, we thereby have that the following are equivalent: (i) $L^1(\mathcal{Z},M)$ convergence and (ii) convergence in the measure $M$ and uniform integrability. By $L^2(\mathcal{Z},M)$ convergence of $((V_{t+\tilde{t},(\rho)}^{\beta})_p$ and the imposed uniform integrability, the convergence result in proposition 4.2 follows. \hfill \Box
Proof of lemma 4.3. To obtain equation (4.10), note that
\[
\hat{c}_{V,\alpha} = \sum_{i=1}^{n_V} e_i \mathbb{E}^{M}[V_{i,t} \phi_{\alpha}(Z_t)]
= \sum_{i=1}^{n_V} e_i \sum_{\gamma \leq \alpha} b^{(\alpha)}_{\phi,\gamma} \mathbb{E}^{M}[V_{i,t} (Z_t)^\gamma]
= \sum_{i=1}^{n_V} e_i \sum_{\gamma \leq \alpha} b^{(\alpha)}_{\phi,\gamma} \tilde{\Phi}^{M,([0;\gamma];e_i)}(0;0;0,\infty).
\] (C.12)

By the imposed assumptions, according to proposition 4.1, each \(\tilde{\Phi}^{M,([0;\gamma];e_i)}\) in equation (C.12) can be computed as a special case of equation (4.4). Using that \(w_{T_i}^B(y;K) = w_{T_i}^B(y_i;K_i)\), \(A_{\Pi}^T(u(y);\tilde{T}) = A_{\Pi}(u_i(y_i);\tilde{T}_i)\), and \(B_{\Pi}(u(y);\tilde{T}) = B_{\Pi}(u_i(y_i);\tilde{T}_i)\) by lemma 4.1 thus yields the result stated in equation (4.10). □

D Estimation methodology

This section contains the technical details and proofs for the results in section 5.

D.1 Exact GMM estimation

In order to formulate the required regularity conditions, we introduce some further notation. For given \(\hat{W}_T\), the exact GMM estimator \(\hat{\vartheta}_T(\hat{W}_T)\) in equation (5.1) is the minimizer of the objective function
\[
\hat{Q}_T(\vartheta) = \hat{g}_T(\vartheta)^T \hat{W}_T \hat{g}_T(\vartheta),
\] (D.1)
where \(\hat{g}_T(\vartheta) = \frac{1}{T} \sum_{t=1}^{T} f_{i,t}(\vartheta)\). Define the associated Jacobian by \(\hat{G}_T(\vartheta) = \nabla_{\vartheta} \hat{g}_T(\vartheta)\).

As \(T \to \infty\), \(\hat{Q}_T(\vartheta)\) in equation (D.1) converges in probability to the limit function
\[
Q(\vartheta) = g(\vartheta)^T W g(\vartheta),
\] (D.2)
where \(\hat{W}_T \overset{p}{\rightarrow} W\) and, by the law of large numbers, \(\hat{g}_T(\vartheta) \overset{p}{\rightarrow} g(\vartheta) = \mathbb{E}_{\vartheta_0}[f_{i}(\vartheta)]\). The associated Jacobian is defined by \(G(\vartheta) = \nabla_{\vartheta} g(\vartheta)\).

Following Newey and McFadden (1994), we can now introduce the following standard regularity conditions in assumptions D.1 and D.2 to assure consistency and asymptotic normality of the exact GMM estimator in equation (5.1).

Assumption D.1. The following conditions hold to assure consistency:

(i) \(\hat{W}_T \overset{p}{\rightarrow} W\) with \(W\) positive semidefinite;

(ii) \(W g(\vartheta) = 0\) if and only if \(\vartheta = \vartheta_0\);

(iii) \(\Theta\) is compact;

(iv) \(\sup_{\vartheta \in \Theta} \|g(\vartheta)\| < \infty\);

(v) \(\sup_{\vartheta \in \Theta} \|\hat{g}_T(\vartheta) - g(\vartheta)\| \overset{p}{\rightarrow} 0\).

Assumption D.2. The following conditions hold to assure asymptotic normality:

(i) \(\vartheta_0\) is in the interior of \(\Theta\);

(ii) \(\hat{g}_T(\vartheta)\) is continuously differentiable in a neighborhood \(O\) of \(\vartheta_0\);
Assumption D.3. Analogous to the notation introduced in appendix D.1, the approximate GMM estimator $\hat{\theta}_{T,(p)}$ where $\hat{\theta}$

Proof of proposition 5.2. For the first part of proposition 5.2, use $\hat{Q}_{T,(p)}(\theta)$ and $\hat{Q}_{T,(p)}(\theta)$, representing

The following conditions hold to assure consistency:

(i) $\sup_{\theta \in \Theta} \| \hat{g}_{T,(p)(T)}(\theta) - \tilde{g}_{T}(\theta) \| \overset{p}{\rightarrow} 0$.

Assumption D.4. The following conditions hold to assure asymptotic normality:

(i) $g_{T,(p)(T)}(\theta)$ is continuously differentiable in a neighborhood $\Theta$ of $\nu_0$;

(ii) $\sup_{\theta \in \Theta} \| \hat{G}_{T,(p)(T)}(\theta) - \hat{G}_{T}(\theta) \| \overset{p}{\rightarrow} 0$;

(iii) $\sqrt{T}(\hat{g}_{T,(p)(T)}(\nu_0) - \tilde{g}_{T}(\nu_0)) \overset{p}{\rightarrow} 0$.

Proof of proposition 5.2. For the second part of proposition 5.2, assume D.4(i) and the mean-value theorem yield

\[
\hat{g}_{T,(p)}(\theta) = \hat{g}_{T,(p)}(\theta_0) + \hat{G}_{T,(p)}(\nu_0) (\hat{\theta} - \theta_0)
\]
writing $\hat{\vartheta} = \hat{\vartheta}_{T,(p)}$ for any expansion order $p$ and for some mean value $\bar{\vartheta}$ between $\vartheta_0$ and $\hat{\vartheta}$. From the first-order condition for the estimator $\hat{\vartheta}$, we have $\hat{G}_{T,(p)}(\hat{\vartheta})^\top \hat{W}_T \hat{g}_{T,(p)}(\hat{\vartheta}) = 0$ and, therefore, equation (D.6) for large enough $T$ and $p$ gives

$$\sqrt{T}(\hat{\vartheta} - \vartheta_0) = -(\hat{G}_{T,(p)}(\hat{\vartheta})^\top \hat{W}_T \hat{G}_{T,(p)}(\hat{\vartheta}))^{-1} \hat{G}_{T,(p)}(\hat{\vartheta})^\top \hat{W}_T \sqrt{T} \hat{g}_{T,(p)}(\vartheta_0). \quad (D.7)$$

As a consequence of assumption D.4(ii), we have $\hat{G}_{T,(p(T))}(\hat{\vartheta}) \overset{p}{\rightarrow} G$ and $\hat{G}_{T,(p(T))}(\bar{\vartheta}) \overset{p}{\rightarrow} G$ and, hence,

$$(\hat{G}_{T,(p(T))}(\hat{\vartheta})^\top \hat{W}_T \hat{G}_{T,(p(T))}(\hat{\vartheta}))^{-1} \hat{G}_{T,(p(T))}(\hat{\vartheta})^\top \hat{W}_T \overset{p}{\rightarrow} (G^\top W G)^{-1} G^\top W$$

Proposition 5.2(ii) then follows from equation (D.7) and assumption D.4(iii) by an application of the Slutsky theorem. \[\square\]
References


