A non-Gaussian, structure-preserving stochastic volatility and option pricing model in discrete time *

Simon Feistle[a] Matthias R. Fengler[a,b] Alexander Melnikov[a]

Version: February 6, 2024

Abstract

We provide a novel stochastic volatility model based on the autoregressive Gamma process that allows for both a structure-preserving change to the risk-neutral measure and a non-Gaussian distribution for the return innovations. The model employs the Meixner distribution, which enriches the return dynamics with conditional stochastic skewness and kurtosis. We propose a fast and accurate estimation method by combining the approximate maximum likelihood method of Bates (2006) with a numerical integration technique suitable for highly oscillatory functions. We derive a closed-form discrete-time option pricing formula. The Meixner specification is superior to the benchmark of their family, especially when calibrated to option data.

Keywords: approximate maximum likelihood, autoregressive Gamma process, discrete-time option pricing, exponentially affine models, Meixner distribution, stochastic volatility

JEL Classification: C22, G12, G13

*Corresponding author: Matthias Fengler. Email: matthias.fengler@unisg.ch.
a: School of Economics and Political Science, University of St.Gallen; *b*: Swiss Finance Institute.

1 Introduction

Discrete-time option pricing models based on the class of autoregressive Gamma (ARG) processes are an appealing alternative to those within the ARCH-GARCH family because they feature a latent stochastic volatility (SV) factor as an additional source of risk. Consequently, conditional return moments are not completely determined by past observations as is the case for ARCH-GARCH models. The use of conditionally deterministic states, particularly in periods leading up to major market movements, can result in an inaccurate representation of the true conditional risk. This is because forecasts are only updated after a significant fluctuation occurs, resulting in an underestimation of risk. Such underestimation can have detrimental consequences for derivative pricing and risk management applications.

Despite this advantage, the number of SV models based on the ARG process (ARGSV models) is relatively small.[1] Existing ARGSV models mostly depend on the Gaussian distribution (see, e.g., Gagliardini et al., 2011; Corsi et al., 2013; Majewski et al., 2015; Han et al., 2020). This often facilitates a structure-preserving change to a risk-neutral measure but also imposes limitations. The only non-Gaussian ARGSV model, the standardized inverse Gaussian (SIG) model of Feunou and Tédongap (2012), does not allow for a structure-preserving change of measure. On the other hand, a multitude of models featuring non-Gaussian error term distributions have been proposed within GARCH or GARCH-like frameworks.[2]

To the best of our knowledge, there does not exist a discrete-time closed-form SV option pricing model that preserves its structure under the risk-neutral measure and features asymmetric and leptokurtic return innovations. As a first contribution, we propose a new member of the ARGSV family called ARGSV-MXN that unifies those properties. This model draws on Meixner distributed innovations and is structure-preserving under the risk-neutral measure if combined with an exponentially linear pricing kernel. The Meixner distribution equips the model with time-varying conditional skewness and kurtosis in the returns, neither of which are completely determined by past return observations. Through a variance-in-mean parameter, the ARGSV-MXN model also accounts for the possibility of an IV smirk because it accommodates a non-zero leverage effect.

[1] Gouriéroux and Jasiak (2006) introduce ARG processes. Apart from SV models, they are also used in multivariate settings (Gouriéroux et al., 2009) and to model times between trades (see, e.g., Gouriéroux and Jasiak, 2006), affine interest rate term structures (see, e.g., Monfort and Pegoraro, 2006, Monfort et al., 2017), and credit risk (Monfort et al., 2021).

[2] E.g., Christoffersen et al. (2006) allow for inverse Gaussian, Mercuri (2008) for tempered stable, Christoffersen et al. (2010) for Poisson-normal, Mercuri and Bellini (2014) for Variance-Gamma, Fengler and Melnikov (2018) for Meixner innovations.

The Meixner distribution is a subclass of the generalized z-distributions with semi-heavy tails (Grigelionis, 1999). To date, the option pricing literature appears to employ the Meixner distribution predominantly in continuous time settings. Grigelionis (1999) and Schoutens (2002) are the first to use processes of the Meixner type (pure jump processes) instead of standard Brownian motions. Itkin (2014) allows for a Brownian diffusion along with a Meixner jump component. The only discrete-time option pricing model using the Meixner distribution that we are aware of is the GARCH model of Fengler and Melnikov (2018).

The second contribution of this paper is to develop a fast and accurate estimation strategy for ARGSV models. A potential reason for the prevalence of GARCH models in discrete-time option pricing is that their estimation is relatively straightforward, whereas estimating ARGSV models typically requires more complex techniques. As a consequence, many studies employ realized variance (RV) as a proxy for the latent variance factor (Gagliardini et al., 2011; Corsi et al., 2013; Majewski et al., 2015; and Han et al., 2020). While intuitive from an empirical perspective, using RV as a proxy for the latent variance factor may introduce additional measurement error. Additionally, RV measures do not capture overnight price variation and may miss jumps in the first few minutes of trading, depending on their specification. Lilla (2021) proposes an estimation strategy based on Kalman filtering that can address these problems by using RV in a measurement equation. GMM approaches do not rely on observational data for the latent variance factor but require a large number of analytical moment conditions (see Feunou and Tédongap (2012) for ARGSV).

For the proposed estimation approach, we tailor the approximate maximum likelihood (AML) method of Bates (2006) to estimating ARGSV models.[3] The main contribution is to combine the AML method with the numerical integration schemes of Ooura (2005). This is pivotal because ARGSV models have characteristic functions that lead to integrands that are singular at the origin, highly oscillatory, and slowly decaying towards infinity. The double exponential method of Ooura (2005) is able to cope with such integrands allowing us to perform the large number of Fourier inversions required by AML in a fast and reliable manner. A simulation study shows that our estimation strategy can recover the true parameter values for different error term distributions.

In our empirical application, we estimate the return parameters under the physical measure for the new ARGSV-MXN model and two benchmark models on S&P 500 data. We jointly recover the risk-neutral dynamics of the models and estimate the price of variance risk. Applying Fourier-based pricing, we calibrate the price of variance risk to option im-

[3] Feunou (2023) suggests the use of the methodology of Bates (2006) to estimate the class of generalized ARG models. Our implementation may facilitate numerically the estimation of this class of models as well.

plied volatilities. The risk-neutral dynamics are determined in this last step, as the pricing kernel keeps the price of variance risk as a free parameter. While the MXN model only slightly outperforms under these terms, it unfolds its true power when its risk-neutral parameters are calibrated directly to option prices. Under this setting, the fit to market implied IVs is remarkable compared to benchmark models and ARGSV models employed in the literature.

The paper is structured as follows. In Section 2, we introduce the ARGSV model and derive the risk-neutral dynamics based on an exponentially affine stochastic discount factor. Building on this result, we provide the multi-step risk-neutral dynamics required for option pricing. Section 3 provides the estimation strategy and reports the results of a simulation study. Section 4 is devoted to the applications.

2 ARGSV model class

In this section, we present the class of ARGSV models and derive its properties, when specified based on the Meixner (MXN) distribution. Combining the model with an exponentially linear pricing kernel, we derive the risk-neutral dynamics and describe how to use the model for option pricing.

2.1 Return dynamics

We use a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to model uncertainty in the evolution of the market prices, where $\mathbb{P}$ is the physical probability measure in a discrete-time economy having the time index set $\{1, \dots, T\}$, where $T < \infty$. Let Y denote the process of observable (log) returns of some asset and r^f denote the process of the risk-free rate. We specify the data generating process of the excess return as

$$Y_{t+1} - r^f_{t+1} = \gamma h_{t+1} + \varepsilon_{t+1}\,, \tag{1}$$

$$\psi^{\mathbb{P}}_{\varepsilon_{t+1}|t,h_{t+1}}(u) = A(u) h_{t+1}\,, \tag{2}$$

$$h_{t+1}|\mathcal{F}_t \sim \Gamma(\delta, \beta h_t, c)\,, \tag{3}$$

where r^f_{t+1} is assumed to be predictable with respect to the filtration $\mathcal{F}_t = \sigma(\varepsilon_t, h_t)$, which is the σ-algebra of the joint process $\{\varepsilon_t, h_t\}$, containing the information available up to and including time t. We denote by $\psi^{\mathbb{P}}_{\varepsilon_{t+1}|t,h_{t+1}}(u) \equiv \log \mathrm{E}_t[\mathrm{e}^{u\varepsilon_{t+1}}|h_{t+1}]$ the cumulant generating function of the error term ε_{t+1} given time-t information and the latent variable h_{t+1}.

Generally, we use the notation $\mathrm{E}_t[\cdot|h_{t+1}]$ for expectations conditioning on $\mathcal{F}_t$ and h_{t+1}. The distribution of ε is captured by the function $A(\cdot)$.

Stochastic variance in the returns is modeled through the latent process h that follows an ARG(1) process[4] with parameters $\delta, \beta, c > 0$ (see Appendix B for a brief and Gouriéroux and Jasiak (2006) for an in-depth discussion of the ARG process). The conditional mean and variance of h are given by $\mu_t^h \equiv \mathrm{E}_t[h_{t+1}] = c\delta + \rho h_t$ and $v_t^h \equiv \mathrm{Var}_t[h_{t+1}] = c^2\delta + 2\rho c h_t$, with $\rho = \beta c$, respectively. The $\mathcal{F}_t$-conditional variance of the return process is mainly determined by h_t. In these formulae, $\rho < 1$ is the autoregressive coefficient of the ARG(1) process, δ determines the conditional dispersion of the process, and c is a scale parameter. Importantly, the ARG process guarantees the positivity of the variance component and can be seen as the discrete-time analogue of the Feller (1951) process. Therefore, ARGSV models are closely related to the classical Heston (1993) model.

For estimation and option pricing, it is useful to characterize the excess return and latent variance component by their joint conditional cumulant generating function under $\mathbb{P}$. Following Gagliardini et al. (2011), we have

$$\begin{aligned}\psi^{\mathbb{P}}_{Y_{t+1}-r^f_{t+1},h_{t+1}|t}(u,v) &\equiv \log \mathrm{E}_t[\mathrm{e}^{u(Y_{t+1}-r^f_{t+1})+vh_{t+1}}] \\ &= \psi^{\mathbb{P}}_{\varepsilon_{t+1},h_{t+1}|t}(u, v+u\gamma) \\ &= \psi^{\mathbb{P}}_{h_{t+1}|t}\left(u\gamma + A(u) + v\right) \\ &= B\left(u\gamma + A(u) + v\right) + C\left(u\gamma + A(u) + v\right) h_t\,, \qquad (4)\end{aligned}$$

where $B(u) = -\delta \log(1 - cu)$ and $C(u) = \frac{cu}{1-cu}\beta$. The functions $B(\cdot)$ and $C(\cdot)$ originate from the conditional cumulant generating function of the ARG(1) process presented in Appendix B.

If $\varepsilon_{t+1}|\mathcal{F}_t, h_{t+1} \sim \mathcal{N}(0, h_{t+1})$, we have $A(u) = \frac{1}{2}u^2$, which is most widely used (e.g., Corsi et al., 2013; Gagliardini et al., 2011; Han et al., 2020). Alternatively, Feunou and Tédongap (2012) consider the standardized inverse Gaussian (SIG) distribution by assuming $\varepsilon_{t+1}|\mathcal{F}_t, h_{t+1} \sim SIG\left(0, h_{t+1}, \eta h_{t+1}^{-\frac{1}{2}}\right)$; see Table 1 for $A(\cdot)$ in this specification. The additional parameter η determines the conditional skewness of the error term distribution. However, using an exponentially affine SDF, this specification is not structure-preserving under the risk-neutral measure (Feunou and Tédongap, 2012).

We now suggest a novel specification based on the MXN distribution that adds additional flexibility in modeling conditional fat-tailedness and asymmetry of financial data compared to Gaussian error terms:

[4]The extension to the ARG(p), $p > 1$ process is straightforward.

Proposition 2.1. *Let ε_{t+1} be the translated MXN random variate*

$$\varepsilon_{t+1}|\mathcal{F}_t, h_{t+1} \sim MXN\left(\breve{a}, b, -\eta_1 h_{t+1}, d h_{t+1}\right) , \tag{5}$$

where $\breve{a} = \sqrt{\frac{2\cos^2\left(\frac{b}{2}\right)}{d}}$, $\eta_1 = \frac{d \tan\left(\frac{b}{2}\right)}{\sqrt{\frac{d}{2\cos^2(b/2)}}}$, $d > 0$, $b \in (-\pi, \pi)$.

Then

$$A(u) = 2d \log\left\{ \frac{\cos\left(\frac{b}{2}\right)}{\cos\left[\frac{1}{2}(b + \breve{a}u)\right]} \right\} - \eta_1 u\,, \tag{6}$$

and

$$\mathrm{E}_t[\varepsilon_{t+1}|h_{t+1}] = 0\,, \tag{7}$$

$$\mathrm{Var}_t[\varepsilon_{t+1}|h_{t+1}] = h_{t+1}\,. \tag{8}$$

Proof. The results (6), (7), and (8) follow directly from the general results on the MXN distribution presented in Appendix A. □

Here, $MXN(\cdot,\cdot,\cdot,\cdot)$ denotes the four-parameter Meixner distribution (see Appendix A). Similarly to the SIG model of Feunou and Tédongap (2012), the MXN model exhibits conditional skewness and conditional kurtosis. They are given by:

$$\mathrm{Skew}_t[\varepsilon_{t+1}|h_{t+1}] = \sqrt{\frac{2}{dh_{t+1}}}\sin\left(\frac{b}{2}\right)\,, \tag{9}$$

$$\mathrm{Kurt}_t[\varepsilon_{t+1}|h_{t+1}] = 3 + \frac{3 - 2\cos^2\left(\frac{b}{2}\right)}{dh_{t+1}}\,. \tag{10}$$

The conditional skewness (conditional on $\mathcal{F}_t$ and h_t)is mainly determined by b: For $b \approx 0$, the error terms are symmetric, with positive skewness for $b > 0$ and negative skewness for $b < 0$. The conditional skewness of the returns $\mathrm{Skew}_t(Y_{t+1})$ is derived in Appendix C and depends both on the sign of b and the sign of γ. If $\gamma < 0$ and $b < 0$, the conditional skewness is negative. In general, the conditional skewness may or may not be negative, and even may change sign as time evolves. This is a similar interplay as Feunou and Tédongap (2012) observe for the conditionally SIG shocks. The conditional kurtosis of the MXN error terms is mainly determined by d. Larger parameter values drive the kurtosis to 3, while smaller values cause additional excess kurtosis. Both conditional skewness and kurtosis increase for small values of h_{t+1}. This property provides the process with a jump-type behavior. The models of Christoffersen et al. (2006, Eq. (5e)) and Feunou and Tédongap (2012) exhibit similar properties.

Table 1: Alternative distributions of the error term

Gaussian	$A(u) = \frac{1}{2}u^2$
Standardized inverse Gaussian (SIG)	$A(u) = -\frac{3u}{\eta} + \frac{9}{\eta^2}\left(1 - \sqrt{1 - \frac{2}{3}u\eta}\right)$
Meixner (MXN)	$A(u) = 2d \log\left\{\frac{\cos\left(\frac{b}{2}\right)}{\cos\left[\frac{1}{2}(b + \breve{a}u)\right]}\right\} - \eta_1 u$

Parameters of the error term distributions: $\eta \in (-1, 1)$ is the shape parameter of the SIG error term of Feunou and Tédongap (2012). $b \in [-\pi, \pi]$ and $d > 0$ are the shape parameters of the MXN error term; $\eta_1 = d \tan\left(\frac{b}{2}\right)\left(\frac{d}{2\cos^2(b/2)}\right)^{-\frac{1}{2}}$ and $\breve{a} = \left(\frac{2\cos^2\left(\frac{b}{2}\right)}{d}\right)^{\frac{1}{2}}$ are normalizing constants.

Through the variance-in-mean parameter γ, the ARGSV model allows for instantaneous correlation between the returns and their variance commonly referred to as leverage effect for the case of a negative sign. To examine this co-movement of the returns and their conditional variances, we derive in Appendix D that:

$$\mathrm{Cov}_t\left(Y_{t+1}, v^Y_{t+1}\right) = \gamma\rho\left(1 + 2\gamma^2 c\right) v^h_t \,. \tag{11}$$

As in Feunou and Tédongap (2012) the sign of the leverage effect is driven by the sign of γ. We note that the models do not fall under the conditionally normal ARGSV class with leverage effect defined in Han et al. (2020). Even with a Gaussian error term, the returns Y_{t+1} of the ARGSV model are not Gaussian conditional on $\mathcal{F}_t$ only. The option prices implied by our models allow for a skewed volatility smile, considered equivalent to the existence of a leverage effect by Han et al. (2020).

2.2 Risk-neutral dynamics

It is often of particular interest to study models that retain their structure under the risk-neutral measure, as this facilitates their use, e.g., in option pricing. For a Gaussian error term in (1), Gagliardini et al. (2011) have shown, using an exponentially affine SDF, that under the risk-neutral probability measure $\mathbb{Q}$, the model has the same structure as under $\mathbb{P}$, featuring adjusted ARG parameters and $\gamma = -\frac{1}{2}$. For the Meixner case, we also retain the structure of the return model under $\mathbb{Q}$, without fixing γ to a specific value.
We assume the four-dimensional, exponentially affine SDF:

$$\mathfrak{m}_{t,t+1} = \mathrm{e}^{-r^f_{t+1}} \exp(\theta_{0,t} + \theta_{1,t}(Y_{t+1} - r^f_{t+1}) + \theta_{2,t}h_t + \theta_{3,t}h_{t+1}) \,, \tag{12}$$

where the coefficients in $\theta_t = (\theta_{0,t}, \theta_{1,t}, \theta_{2,t}, \theta_{3,t})^\top$ are $\mathcal{F}_t$-measurable. The rationale of such an SDF from a general equilibrium perspective is given, e.g., in Bansal and Yaron (2004).

Proposition 2.2. *The SDF in (12) satisfies the one-step ahead pricing relations for the risk-free and the risky asset if*

$$\theta^*_{1,t} = -\frac{2}{\breve{a}} \arctan\left(\frac{\zeta - \cos\left(\frac{\breve{a}}{2}\right)}{\sin\left(\frac{\breve{a}}{2}\right)}\right) - \frac{b}{\breve{a}}\,, \tag{13}$$

$$\theta^*_{0,t} = -B\Big(\theta^*_{3,t} + A(\theta^*_{1,t}) + \theta^*_{1,t}\gamma\Big)\,, \tag{14}$$

$$\theta^*_{2,t} = -C\Big(\theta^*_{3,t} + A(\theta^*_{1,t}) + \theta^*_{1,t}\gamma\Big)\,, \tag{15}$$

and $\theta^*_{3,t}$ is a free parameter, where $\zeta = \exp(\frac{\gamma-\eta_1}{2d})$.

Proof. See Appendix E. □

The proposition shows that time variation in θ_t only depends on $\theta^*_{3,t}$ and r^f_{t+1}. For the ease of exposition, we now assume that $\theta^*_{3,t} = \theta^*_3$ and $r^f_{t+1} = r^f$, and drop the time index on θ.

Proposition 2.3. *The conditional one-step ahead risk-neutral cumulant generating function $\psi^{\mathbb{Q}}_{Y_{t+1}-r^f_{t+1}, h_{t+1}|t}$ is given by*

$$\psi^{\mathbb{Q}}_{Y_{t+1}-r^f_{t+1}, h_{t+1}|t}(u,v) = B^*\big(A^*(u) + u\gamma^* + v\big) + C^*\big(A^*(u) + u\gamma^* + v\big)h_t \tag{16}$$

where

$$A^*(u) = 2d^* \log \frac{\cos(b^*/2)}{\cos[\frac{1}{2}(b^* + \breve{a}^* u)]} - \eta^*_1 u \tag{17}$$

$$B^*(u) = -\delta^* \log(1 - c^* u) \tag{18}$$

$$C^*(u) = \frac{c^* u}{1 - c^* u}\beta^* \tag{19}$$

and

$$\delta^* = \delta, \quad c^* = \frac{c}{1 - cz(\theta^*_1, \theta^*_3)}, \quad \beta^* = \frac{\beta}{1 - cz(\theta^*_1, \theta^*_3)}, \tag{20}$$
$$b^* = b + \breve{a}\theta^*_1, \quad d^* = d, \quad \breve{a}^* = \breve{a}, \quad \eta^*_1 = \eta_1, \quad \gamma^* = \gamma\,,$$
$$z(u,v) = A(u) + \gamma u + v\,.$$

Proof. See Appendix E. □

The explicit formulae (20) provide the mapping between the parameters of the model under $\mathbb{P}$ and $\mathbb{Q}$. Because the conditional cumulant generating function in (16) is of the same form as (4), we can conclude that, under $\mathbb{Q}$, the joint process $\{Y_t, h_t\}$ has the same model structure as under $\mathbb{P}$, i.e., the variance process is also an ARG(1) process, yet with parameters δ^*, β^*, c^*, and the innovation ε_t follows, conditional on h_t, the MXN distribution $MXN(\breve{a}^*, b^*, -\eta_1^* h_t, h_t d^*)$. Thus, under $\mathbb{Q}$, we obtain a different skewness of the MXN distribution. In contrast, in a model where ε_t is driven by mean zero Gaussian noise with variance h_t, the process under $\mathbb{Q}$ may be interpreted as a process with a specific variance-in-mean parameter $\gamma = -\frac{1}{2}$ (Gagliardini et al., 2011). In the Meixner case, the parameter γ is preserved in the dynamics under the risk-neutral measure and the location of the driving MXN noise is altered. For both models, however, the scale and persistence of the variance factor h may change under the risk-neutral measure.

2.3 Option pricing

To price options, we first derive the ϑ-step ahead conditional cumulant generating function of the cumulative return process.

Corollary 2.4. *Let $Y_t(\vartheta) = \sum_{i=1}^{\vartheta} Y_{t+i} - \vartheta r^f$. The risk-neutral ϑ-step ahead conditional cumulant generating function of $Y_t(\vartheta)$ is given by*

$$\psi^{\mathbb{Q}}_{Y_t(\vartheta)|t}(u) = \log \mathrm{E}_t^{\mathbb{Q}} \exp\left(u \sum_{i=1}^{\vartheta} Y_{t+i}\right) \tag{21}$$

$$= \mathcal{B}^*_{\vartheta}(u) + \mathcal{C}^*_{\vartheta}(u) h_t \tag{22}$$

where the functions $\mathcal{B}^*_{\vartheta}(u)$ and $\mathcal{C}^*_{\vartheta}(u)$ follow the recursions

$$\mathcal{B}^*_{\vartheta}(u) = \mathcal{B}^*_{\vartheta-1}(u) + B^*(z(u, \mathcal{C}^*_{\vartheta-1}(u))) \tag{23}$$

with $\mathcal{B}^*_1(u) = B^*(z(u,0))$ and

$$\mathcal{C}^*_{\vartheta}(u) = C^*(z(u, \mathcal{C}^*_{\vartheta-1}(u))) \tag{24}$$

with $\mathcal{C}^*_1(u) = C^*(z(u,0))$. Recall that $z(u,v) = A(u) + \gamma u + v$

Proof. Follows from ideas in (Darolles et al., 2006, Appendix B). □

Using the multi-step ahead conditional cumulant generating function and applying risk-neutral valuation principles, we price options following Carr and Madan (1999). We choose their method on the grounds of the well-known efficiency advantages. See Appendix G for more details.

3 Estimation of ARGSV models

3.1 Estimation strategy

ARGSV models require sophisticated estimation strategies because the variance factor is latent. Therefore, an approach followed frequently is to replace the latent factor by observational data, usually realized variance (see, e.g., Corsi et al., 2013; Han et al., 2020). This facilitates the estimation enormously because the estimation problem becomes standard. However, the obtained parameter estimates are likely to be confounded by measurement error introduced by noisy RV estimates.

Our estimation strategy follows Bayesian ideas and relies on return data only.[5] More specifically, we tailor the approximate maximum likelihood (AML) method proposed by Bates (2006) to estimating the ARGSV. The method operates in the transform space of characteristic functions (CF) and recursively updates the joint CF of the latent variable and the observed data, conditional on past data. The Bayesian updating of the CF can be seen as an extension to updating conditional Gaussian densities through Kalman filtering.

By implementing the method for the ARGSV, we can estimate all model parameters using only daily return data, thereby bypassing the necessity for RV data. As opposed to GMM approaches, we obtain a filtered time series of the latent variance factor. We expect that our estimation approach could broaden use case of ARGSV models.

3.2 Implementation

The ARGSV models are of the semi-affine form as described by Bates (2006), which refers to the fact that their conditional CFs are exponentially linear in the latent variable but not necessarily in the observed one. This can be immediately observed from (4).

[5] Alternatively, Feunou and Tédongap (2012) suggest estimating by means of a GMM procedure relying on up to 171 analytical moment conditions.

We outline more details of the estimation strategy in Appendix F. As a specific issue, the Fourier inversions of the characteristic function of h and the derivatives of the joint characteristic function of Y and h, see (40) to (42), present a particular challenge to using this method for ARGSV models. They have a pole at $\delta = 0$ and, particularly for small values of δ, are highly oscillatory and converge very slowly towards infinity. Even though it appears that in finance applications δ is not necessarily close to zero, it is critical to be able to reliably evaluate the likelihood for such values when employing an iterative gradient-based maximization algorithm. Standard numerical integration methods such as the Newton-Cotes formulae and Gaussian quadrature methods may not provide the desired accuracy.

As a solution, we propose to take advantage of the double-exponential quadrature schemes suggested by Ooura (2005). They are specifically targeted at performing accurate Fourier inversions of functions that converge very slowly at infinity and/or possess a singularity. Besides the double-exponential dampening of the integrand, another advantage of the method is the automatic selection of the node points to consider specific oscillatory behavior of the integrand at hand. Details of this selection procedure and the double-exponential dampening formula adapted to our application are deferred to the Appendix F. In our implementation, we always make use of the analytical formulae of all the CFs and their derivatives, which are available upon request.

The above amendments allow us to use the AML method in the following likelihood maximization task. For a sample of observed returns $\{Y_s\}_{s=0}^{T}$, we maximize the approximate likelihood over the parameter vector ξ

$$\mathcal{L}(\xi|\{Y_s\}_{s=0}^{T}) = \log p(\xi; Y_0) + \sum_{t=1}^{T} \log p(\xi; Y_t|\{Y_s\}_{s=0}^{t-1}), \tag{25}$$

where $p(\cdot)$ is the conditional density of the returns, and ξ contains the ARG parameters δ, c, and β as well as γ. For the SIG and MXN models, it contains further parameters, including $\xi = (\delta, c, \beta, \gamma, \eta)^\top$ for the SIG and $\xi = (\delta, c, \beta, \gamma, b, d)^\top$ for the MXN model.

3.3 Estimation on simulated data

In this section, we examine the accuracy of our estimation strategy in a simulation study. For each of the models considered under a given parameter constellation, we simulate 250 independent sample paths of length $T = 5,000$ according to (1). We also present the estimates of the ARG parameters using an alternative parametrization which is easier to interpret. It invokes the mean, $\mu = \frac{c\delta}{1-\rho}$, the unconditional variance, $\omega = \frac{c^2\delta}{(1-\rho)^2}$, and the

persistence, $\rho = c\beta$, of the process h.
Panel 1 of Table 2 displays a set of estimation results on simulated data for Gaussian error terms. We tested our estimation method on a wide range of parameter values for all parameters and the results are similar. Our AML implementation coupled with the double exponential numerical integration of Ooura (2005) estimates all parameters of the process accurately on average, with an average bias much smaller than the respective standard deviation of the parameter estimates across all independent sample paths (empirical standard errors). The estimation of the model is unfeasible with standard numerical integration methods. The average asymptotic standard errors are close to the empirical ones for all parameters, except for the parameters c and β, for which the asymptotic standard errors tend to underestimate the empirical ones. We conjecture that this is likely due to the particular role of those two parameters that jointly determine the autoregressive coefficient of the process because, once considered in the alternative representation, the asymptotic and empirical standard errors are close.
Panel 2 of Table 2 exhibits the results for SIG error terms. We find that they are similar to the ones obtained with the Gaussian error term as far as the ARG parameters and γ are concerned. Additionally, η is estimated accurately and its standard error estimates are reliable.
Panel 3 of Table 2 presents results for the MXN model. Again, we find that all parameters are accurately estimated, which is best seen when studying the mean (μ), variance (ω), and persistence (ρ) of the process. The estimates of the MXN parameters b and d are precise as well. The average asymptotic standard errors are close to the empirical ones for all parameters.
Two remarks on the estimation of the MXN model are in order. The first is related to the simulation of the sample paths. The parameters of the ARG process together with d have to be chosen in a way such that the product $h_t d$, which determines the kurtosis of the MXN variate for the next time step, never becomes too small or excessively large. Otherwise, the rejection method described in Grigoletto and Provasi (2008) becomes inefficient when simulating MXN variates. Thus, we can not analyze the estimation method across all permitted parameter values. However, this is not a serious limitation of the estimator, as it is related to simulation of sample paths only.
The second remark is refers to a general property of the MXN distribution. The larger d, the harder it becomes to estimate. This is to be expected, considering that the MXN distribution approaches the Gaussian distribution as d approaches infinity, causing d to become unidentifiable. In such cases, the optimization may fail. In conclusion, our estimation approach enables us to effectively estimate different ARGSV models.

Table 2: Results of simulation study

Panel 1: ARGSV-GAU

	δ	$c \times 10^4$	$\beta \times 10^{-4}$	γ	ρ	$\mu \times 10^4$	$\sqrt{\omega} \times 10^4$
True Values	2.00	3.00	0.317	-5.000	0.950	120.0	84.9
Avg. Est.	2.09	3.01	0.321	-4.982	0.948	120.2	83.5
Avg. Bias	0.09	0.01	0.005	0.018	-0.002	0.2	-1.4
Std. Dev.	0.19	0.43	0.047	0.166	0.008	7.8	6.5
RMSE	0.20	0.43	0.047	0.167	0.008	7.8	6.7
Avg. as. SE	0.17	0.20	0.021	0.174	0.006	7.8	6.5

Panel 2: ARGSV-SIG

	δ	$c \times 10^4$	$\beta \times 10^{-4}$	γ	η	ρ	$\mu \times 10^4$	$\sqrt{\omega} \times 10^4$
True Values	2.00	3.00	0.317	-5.000	-0.050	0.950	120.0	84.9
Avg. Est.	2.05	3.04	0.319	-5.016	-0.049	0.948	119.9	83.9
Avg. Bias	0.05	0.04	0.003	-0.016	0.001	-0.002	-0.1	-1.0
Std. Dev.	0.17	0.45	0.050	0.148	0.004	0.008	8.1	6.8
RMSE	0.18	0.45	0.050	0.149	0.005	0.008	8.1	6.9
Avg. as. SE	0.17	0.20	0.022	0.152	0.005	0.006	7.9	6.6

Panel 3: ARGSV-MXN

	δ	$c \times 10^4$	$\beta \times 10^{-4}$	γ	b	d	ρ	$\mu \times 10^4$	$\sqrt{\omega} \times 10^4$
True Values	2.00	3.00	0.317	-5.000	-0.250	100.00	0.950	120.0	84.9
Avg. Est.	2.06	3.06	0.317	-5.012	-0.246	104.36	0.947	119.8	83.7
Avg. Bias	0.06	0.06	0.0003	-0.012	0.004	4.36	-0.003	-0.2	-1.1
Std. Dev.	0.19	0.46	0.053	0.157	0.092	19.98	0.008	8.1	6.9
RMSE	0.20	0.47	0.053	0.157	0.092	20.45	0.009	8.1	7.0
Avg. as. SE	0.20	0.53	0.058	0.197	0.098	20.57	0.010	9.9	8.2

Estimation results on 250 independent sample paths with length $T = 5,000$. The latent ARG process h is generated from a Poisson mixture of gamma distributions as described in Gouriéroux and Jasiak (2006) using the true values of the parameters δ, c, and β. We generate SIG random variables by standardizing inverse Gaussian ones using the true value of the parameter η, and MXN random variables with the rejection method described in Grigoletto and Provasi (2008) using the true values of the parameters b and d. Starting values for the optimization of each path are chosen randomly and uniformly as follows: $\delta_0 \in (0.3, 4)$, $c_0 \times 10^4 \in (0.4, 4)$, $\rho_0 \in (0.8, 0.965)$, $\gamma_0 \in (-10, 10)$, $\eta_0 \in (-0.2, 0.2)$, $b_0 \in (-\frac{\pi}{2}, \frac{\pi}{2})$, $d_0 \in (100, 1000)$. The constraints $\delta, c, \beta, d > 0$ and $b \in [-\pi, \pi]$ are imposed through transformations. The asymptotic standard errors are computed from the inverse of the Hessian at the final parameter estimates. Standard errors of the back-transformed parameters as well as of the persistence $\rho = c\beta$, the unconditional mean $\mu = \frac{c\delta}{1-\rho}$, and the unconditional variance $\omega = \frac{c^2\delta}{(1-\rho)^2}$ of the ARG process are obtained with the delta method. The estimation results are shown for 250 paths for which the optimization converged.

4 Empirical application

In this section, we estimate the ARGSV models on return data. Starting from these estimates, we obtain the risk-neutral dynamics of the structure-preserving models. We conclude by providing a comparison between the MXN and the SIG model by directly calibrating their critical parameters to option IVs.

4.1 Estimation of return dynamics

Panel A of Table 3 presents the parameter estimates and standard errors for δ, c, β, and γ, as well as those of b and d (for the MXN model) and of η (for the SIG model). They are estimated on daily S&P 500 log returns from January 3, 2000 to December 31, 2020 (N =5,284). To obtain excess returns, we calculate log risk-free rates using daily risk-free rates from Kenneth R. French's data library. The daily excess returns on the S&P 500 index have a mean of 1.16×10^{-4}, a variance of 1.58×10^{-4} (annualized volatility of 19.85%), a skewness of -0.39, and a kurtosis of 13.93.

As seen from Table 3, the estimate for d in the MXN model is relatively large. The MXN parameter b is estimated with the expected negative sign. When evaluating the conditional skewness in (9) at $h_t = \mu$, i.e., its unconditional expectation, we find a value of about -0.15. Hence, despite the large estimate for d, the MXN model partly captures the skewness of the return data. Regarding the SIG model, η is estimated with the expected negative sign and in a range similar to what Feunou and Tédongap (2012) report – see their Table 4, p. 58.

Both models show a very high persistence (ρ) in the latent variance factor. The unconditional mean μ and the unconditional variance ω of the latent factor are estimated in a similar range for both models, and μ is close to the actual daily variance of the return data. Because the latent factor is the main driver of the unconditional volatility of the returns, both models reflect well the annualized volatility of the data.

4.2 Estimation of risk-neutral dynamics

We use S&P 500 option data from OptionMetrics for the first 21 Wednesdays of the year 2021. For each of these trading days, we eliminate from the dataset any option that was not traded on that day. Further, we only consider the sufficiently liquid out-of-the-money put and call options that are less than 10% out of the money. We also limit the data to options with more than 28 and less than 100 days to maturity. This leaves us with a sample of $N = 17,654$ option contracts (599 to 1,140 observations per trading day).

Panels B to D of Table 3 display the average values of the results for each of the trading days. We derive the risk-neutral parameters according to (20). Moreover, we calibrate the free SDF parameter θ_3^* and the initial value of the latent variance factor h_0[6] to the option implied volatilities (IVs) by minimizing the IVRMSE given by

$$\text{IVRMSE} = \sqrt{\frac{1}{N}\sum_{i}^{N}\left(\text{IV}_i^{Market} - \text{IV}_i^{Model}\right)^2}. \tag{26}$$

We obtain market IVs and those of our models by inverting the Black-Scholes formula.

The unconditional mean of the variance factor is higher under the risk-neutral measure for both models ($\mu^* > \mu$). This difference is larger for the Gaussian model since it has to compensate for the negative drift introduced by imposing $\gamma^* = -\frac{1}{2}$, as required by the no-arbitrage conditions. On the other hand, the MXN model has the additional advantage of allowing us to retain γ under the risk-neutral measure. For both models, the persistence of the variance component is slightly higher under the risk-neutral measure ($\rho^* > \rho$).

For our second set of estimation results on S&P 500 option data, we use the same option data as described above. For the SIG model, we do not perform an analytic change of measure but assume that the model follows the structure laid out in Section 2.1 under the risk-neutral measure. This allows us to include this model which does not provide a structure-preserving change of measure, as a benchmark model for the MXN model for this set of results. The calibrated risk-neutral parameters for the MXN model are obtained performing the analytic change of measure, but calibrating the physical-parameters along with θ_3^* directly to option data, guaranteeing put-call parity for this model. For both models, we fix δ^* to the respective values estimated on the return data to control the size of the optimization problem.

Table 4 shows the results of this calibration. We find a larger negative skewness parameter of the MXN model (b) compared to the econometric approach. At the same time, we estimate a substantially smaller value for the kurtosis parameter of the MXN model (d). This implies a large excess kurtosis. We also find the expected negative sign for the skewness parameter of the SIG model (η).

The ability of the MXN model to fit the implied IV surface is better than that of the SIG model. For the SIG model, we obtain a fit that is generally in line with the findings of

[6] We calibrate h_0 instead of using filtered values, which we obtain from the estimation using the return data. This allows us to avoid any influence on the comparison of the models arising from (arbitrary) differences in the accuracy of the filtered variance paths.

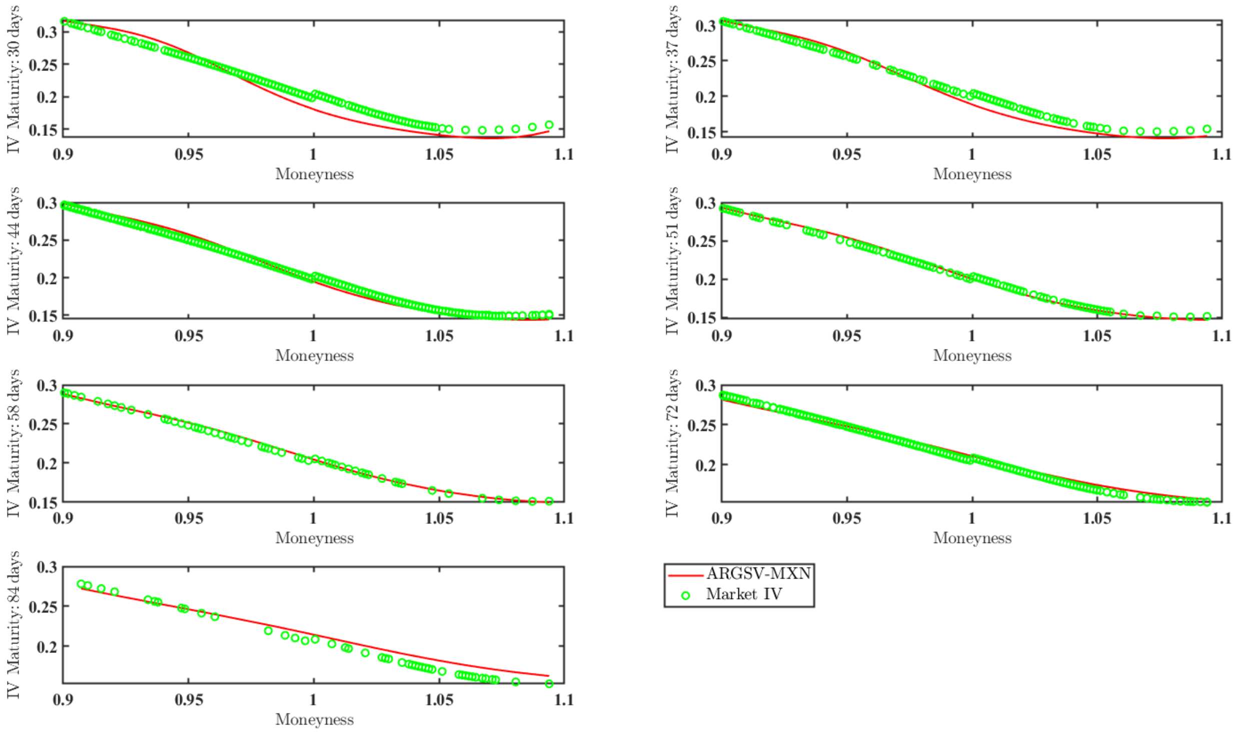

Figure 1: This figure displays the fit of the ARGSV-MXN model to option IV when the risk-neutral paramters are calibrated.

Feunou and Tédongap (2012) for a similar model. The MXN model is able to produce a very competitive fit within the ARGSV family.

Figure 1 illustrates the ARGSV-MXN model fit to option IVs. To guarantee put-call parity, we calibrate the physical parameters and θ_3^* instead of the risk-neutral parameters to option IVs. As is nicely visible, the model allows us to fit option IVs across a range of maturities and moneyness levels. For the out-of-the-money calls, the model neatly captures the stylized fact of a 'smiley' IV of short-dated maturities.

Table 3: Estimation results for ARGSV model

	ARGSV-MXN	ARGSV-GAU	ARGSV-SIG
Panel A: Parameters estimated on return data			
δ	1.57	1.36	1.40
	(0.14)	(0.10)	(0.10)
$c \times 10^6$	1.7685	2.510	2.418
	(0.08)	(0.12)	(0.12)
$\beta \times 10^{-6}$	0.53	0.39	0.40
	(0.03)	(0.02)	(0.02)
γ	0.87	0.86	0.82
	(1.18)	(1.18)	(1.16)
b	-0.49		
	(0.11)		
d	37,732		
	(11,767)		
$\eta \times 10^3$			-1.22
			(0.35)
ρ	0.980	0.974	0.976
	(0.003)	(0.003)	(0.003)
$\mu \times 10^4$	1.41	1.32	1.44
	(0.17)	(0.16)	(0.16)
$\sqrt{\omega} \times 10^4$	1.12	1.14	1.21
	(0.15)	(0.15)	(0.15)
$\sqrt{\text{Var}[Y_t]}$ (%, ann.)	18.94	18.19	18.96
Panel B: Risk-neutral parameters			
$\delta^* = \delta$			
$c^* \times 10^6$	1.7688	2.511	
$\beta^* \times 10^{-6}$	0.54	0.39	
b^*	-0.49		
γ^*	0.87	-0.50	
ρ^*	0.981	0.975	
$\mu^* \times 10^4$	1.55	1.71	
$\sqrt{\omega^*} \times 10^4$	1.14	1.19	
$\sqrt{\text{Var}^*[Y_t]}$ (%, ann.)	19.67	22.76	
Panel C: Parameters calibrated to option IV			
θ_3^*	68	205	
$h_0 \times 10^4$	0.78	0.75	
Panel D: Other SDF parameters			
$\theta_0^* \times 10^4$	1.94	7.03	
θ_1^*	-1.37	-1.63	
θ_2^*	-71	-203	
VRP (%, ann.)	-0.92	-2.41	
$IVRMSE$ (%)	4.22	4.26	

VRP is calculated as the difference between the variance of Y under the physical and risk-neutral measure. $IVRMSE$ is the root mean squared error of the Black-Scholes implied volatilities of the model prices compared to thosse of the the market prices. Parameters θ_3, and h_0 are calibrated to minimize the $IVRMSE$.

Table 4: Calibration results for ARGSV models

	ARGSV-MXN	ARGSV-SIG
Panel A: Parameters estimated on return data		
δ^*	1.57	1.40
Panel B: Parameters calibrated to option IV		
$c^* \times 10^6$	1.76	12.2
ρ^*	0.977	0.892
b^*	-2.031	
d^*	47.99	
η^*		-0.14
γ^*	0.18	-0.47
$IVRMSE$ (%)	0.776	1.142

The value of δ^* is fixed based on the estimation results obtained on return data. The other parameters are calibrated to minimize the IVRMSE. For the MXN model, physical paramters are calibrated together with θ_3^* to guarantee put-call parity. The prices for the SIG model do not guarantee put-call parity.

5 Conclusion

Building on the Meixner distribution, we introduce a novel ARGSV model featuring a non-Gaussian error term distribution while maintaining its structure under a risk-neutral measure. We derive the risk-neutral dynamics of the model, its multi-step-ahead risk-neutral distribution, and a closed-form option pricing formula that can be evaluated using Fourier inversions. By capitalizing on a latent risk source, stochastic volatility models of this kind evade the conditionally deterministic states inherent in GARCH and GARCH-like models, positioning them as compelling competitors.

The ARGSV family is more difficult to estimate than GARCH models. Therefore, the literature often relies on proxy variables, such as realized variance data, for their estimation. We present an estimation strategy that avoids the use of proxy variables. Looking forward, we anticipate that our approach may facilitate the estimation of related models, e.g., the generalized ARG process introduced in Feunou (2023) or the models considered in Han et al. (2020).

Using S&P 500 (option) data, we simultaneously obtain the risk-neutral parameters and an estimate of the price of variance risk. We confirm previous findings on the price of variance risk and show that the Meixner model provides a very good fit of the IV surface when all Meixner parameters are calibrated.

References

Bansal, R. and Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* **59**: 1481–1509.

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

Bertholon, H., Monfort, A. and Pegoraro, F. (2008). Econometric asset pricing modelling, *Journal of Financial Econometrics* **6**(4): 407–458.

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

Christoffersen, P., Elkamhi, R., Feunou, B. and Jacobs, K. (2010). Option valuation with conditional heteroskedasticity and nonnormality, *Review of Financial Studies* **23**(5): 2139–2183.

Christoffersen, P., Heston, S. and Jacobs, K. (2006). Option valuation with conditional skewness, *Journal of Econometrics* **131**: 253–284.

Corsi, F., Fusari, N. and La Vecchia, D. (2013). Realizing smiles: Options pricing with realized volatility, *Journal of Financial Economics* **107**: 284–304.

Darolles, S., Gouriéroux, C. and Jasiak, J. (2006). Structural Laplace transform and compound autoregressive model, *Journal of Time Series Analysis* **27**(4): 477–503.

Feller, W. (1951). Two singular diffusion problems, *Annals of Mathematics* **54**(1): 173–182.

Fengler, M. and Melnikov, A. (2018). GARCH option pricing models with Meixner innovations, *Review of Derivatives Research* **21**(3): 277–305.

Feunou, B. (2023). Generalized autoregressive positive-valued processes, *Journal of Business & Economic Statistics* pp. 1–15.

Feunou, B. and Tédongap, R. (2012). A stochastic volatility model with conditional skewness, *Journal of Business & Economic Statistics* **30**(4): 576–591.

Gagliardini, P., Gouriéroux, C. and Renault, E. (2011). Efficient derivative pricing by the extended method of moments, *Econometrica* **79**(4): 1181–1232.

Gouriéroux, C. and Jasiak, J. (2006). Autoregressive Gamma processes, *Journal of Forecasting* **25**: 129–152.

Gouriéroux, C., Jasiak, J. and Sufana, R. (2009). The Wishart autoregressive processes of multivariate stochastic volatility, *Journal of Econometrics* **150**: 167–181.

Gouriéroux, C. and Monfort, A. (2007). Econometric specification of stochastic discount factor models, *Journal of Econometrics* **136**: 509–530.

Grigelionis, B. (1999). Processes of Meixner type, *Lithuanian Mathematical Journal* **39**(1): 33–41.

Grigelionis, B. (2001). Generalized z-distributions and related stochastic processes, *Lithuanian Mathematical Journal* **41**(3): 239–251.

Grigoletto, M. and Provasi, C. (2008). Simulation and estimation of the Meixner distribution, *Communications in Statistics - Simulation and Computation* **38**(1): 58–77.

Han, H., Khrapov, S. and Renault, E. (2020). The leverage effect puzzle revisited: Identification in discrete time, *Journal of Econometrics* **217**(2): 230–258. Nonlinear Financial Econometrics.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* **6**: 327–343.

Itkin, A. (2014). Splitting and matrix exponential approach for jump-diffusion models with Inverse Normal Gaussian, Hyperbolic and Meixner jumps, *Algorithmic Fianance* **3:3-4**: 233–250.

Lilla, F. (2021). Volatility Bursts: A Discrete-Time Option Model with Multiple Volatility Components*, *Journal of Financial Econometrics* **21**(3): 678–713.

Majewski, A. A., Bormetti, G. and Corsi, F. (2015). Smile from the past: A general option pricing framework with multiple volatility and leverage components, *Journal of Econometrics* **187**(2): 521–531. Econometric Analysis of Financial Derivatives.

Mercuri, L. (2008). Option pricing in a GARCH model with tempered stable innovations, *Finance Research Letters* **5**(3): 172–182.

Mercuri, L. and Bellini, F. (2014). Option pricing in a dynamic variance-gamma model, *Technical report*.

Monfort, A. and Pegoraro, F. (2006). Switching VARMA Term Structure Models, *Journal of Financial Econometrics* **5**(1): 105–153.

Monfort, A., Pegoraro, F., Renne, J.-P. and Roussellet, G. (2017). Staying at zero with affine processes: An application to term structure modelling, *Journal of Econometrics* **201**(2): 348–366. THEORETICAL AND FINANCIAL ECONOMETRICS: ESSAYS IN HONOR OF C. GOURIEROUX.

Monfort, A., Pegoraro, F., Renne, J.-P. and Roussellet, G. (2021). Affine modeling of credit risk, pricing of credit events, and contagion, *Management Science* **67**(6): 3674–3693.

Ooura, T. (2005). A double exponential formula for the fourier transforms, *Publications of The Research Institute for Mathematical Sciences* **41**: 971–977.

Schoutens, W. (2002). The Meixner process: theory and applications in finance, *Report 2002–004*, K.U. Leuven.

Schoutens, W. and Teugels, J. L. (1998). Lévy processes, polynomials and martingales, *Communications in Statistics. Stochastic Models* **14**(1–2): 335–349.

A The Meixner distribution

A random variable X has the Meixner distribution $MXN(a,b,m,d)$ if its probability density function is given by:

$$p_X(x) = \frac{\left(2\cos\left(\frac{b}{2}\right)\right)^{2d}}{2a\pi\Gamma(2d)} \exp\left(b\,\frac{x-m}{a}\right) \left|\Gamma\left(d + \mathrm{i}\frac{x-m}{a}\right)\right|^2 , \tag{27}$$

where $\Gamma(\cdot)$ is the gamma function, $\mathrm{i} = \sqrt{-1}$, $a > 0$, $-\pi < b < \pi$, $m \in \mathbb{R}$, $d > 0$ and $x \in \mathbb{R}$ – see Schoutens and Teugels (1998) and Grigelionis (1999). In (27), the parameter d influences the peakedness and the parameter b the skewness of the Meixner distribution. The parameters a and m define scale and location, respectively. Moments of all orders exist. In particular,

$$\begin{aligned} \mathrm{E}[X] &= m + ad\tan\left(\frac{b}{2}\right) = \mu_{MXN} , \\ \mathrm{Var}[X] &= \frac{a^2 d}{2}\,\frac{1}{\cos^2(b/2)} = \sigma^2_{MXN} , \\ \mathrm{Skew}[X] &= \sqrt{\frac{2}{d}}\,\sin\left(\frac{b}{2}\right) , \\ \mathrm{Kurt}[X] &= 3 + \frac{3 - 2\cos^2(b/2)}{d} . \end{aligned} \tag{28}$$

Thus the distribution is symmetric for $b = 0$, skewed to the left (right) for $b < 0$ ($b > 0$). The Meixner distribution always has excess kurtosis.

The MGF of the Meixner distribution $MXN(a,b,m,d)$ exists (Grigelionis, 2001). It is given by

$$\mathrm{E}[\mathrm{e}^{uX}] = \mathrm{e}^{mu} \left(\frac{\cos\left(\frac{b}{2}\right)}{\cos\left(\frac{b+au}{2}\right)}\right)^{2d} , \quad u \in \left(\frac{-\pi - b}{a}, \frac{\pi - b}{a}\right) . \tag{29}$$

Suppose that $X \sim MXN(a,b,m,d)$ and $Y = AX + B$ with $A \in \mathbb{R}_{>0}$ and $B \in \mathbb{R}$. Because Y has the density $p_Y(x) = \frac{1}{A} p_X\left(\frac{x-B}{A}\right)$, we observe that

$$Y \sim MXN(Aa, b, Am + B, d) . \tag{30}$$

Hence, the Meixner distribution is closed under affine transformations. This allows one to define a zero mean, unit variance Meixner distribution as

$$\frac{X - \mu_{MXN}}{\sigma_{MXN}} \sim MXN\left(\tilde{a},\, b\,, \tilde{m},\, d\right) . \tag{31}$$

where the parameters a and m cancel out as they can be expressed as functions of b and d:

$$\tilde{a} = \sqrt{\frac{2\cos^2\left(\frac{b}{2}\right)}{d}}; \tag{32}$$

$$\tilde{m} = -\frac{d\tan\left(\frac{b}{2}\right)}{\sqrt{\frac{d}{2\cos^2(b/2)}}}. \tag{33}$$

B The autoregressive Gamma process

The process h is an autoregressive gamma process of order one , ARG(1), if and only if the conditional distribution of h_t given h_{t-1} is non-centered gamma $\bar{\gamma}(\delta, \beta h_{t-1}, c)$, $\delta, c, \beta > 0$. The ARG(1) is weakly stationary if $\rho = \beta c < 1$ (Gouriéroux and Jasiak, 2006). The conditional cumulant generating function of the ARG(1) process is given by:

$$\psi_{h_{t+1}|h_t} = B(u) + C(u)h_t\,, \tag{34}$$

where $B(u) = -\delta\log(1-cu)$ and $C(u) = \frac{cu}{1-cu}\beta$.

C Third and fourth conditional central moments of Y

The conditional third central moment of the returns is given by

$$\mathrm{E}_t[(Y_{t+1} - \mu_t^Y)^3] = \sqrt{\frac{2}{d}}\sin(b/2)\,\mu_t^h + 3\gamma v_t^h + \gamma^3\mu_{3,t}^h\,, \tag{35}$$

where $\mu_{3,t}^h \equiv \mathrm{E}_t[(h_{t+1} - \mu_t^h)^3] = 2c^3\delta + 6c^2\rho h_t$. The conditional fourth central moment of the returns is given by

$$\begin{aligned}\mathrm{E}_t[(Y_{t+1} - \mu_t^Y)^4] = {} & \frac{3 - 2\cos^2(b/2)}{d}\mu_t^h + 3(\mu_t^h)^2 + 12\gamma^2(\mu_t^h)^3 + (3 + 4\gamma^2\sqrt{2/d}\sin(b/2))v_t^h \\ & + 6\gamma^2 v_t^h\mu_t^h + 6\gamma^2\mu_{3,t}^h + \gamma^4\mu_{4,t}^h\,. \end{aligned} \tag{36}$$

D Conditional covariance of returns and their conditional variance

$$
\begin{aligned}
\mathrm{Cov}_t\left(Y_{t+1}, v^Y_{t+1}\right) &= \mathrm{Cov}_t\left(\gamma h_{t+1} + \varepsilon_{t+1}, \mu^h_{t+1} + \gamma^2 v^h_{t+1}\right) \\
&= \mathrm{Cov}_t\left(\gamma h_{t+1} + \varepsilon_{t+1}, \rho h_{t+1} + 2\gamma^2 \rho c h_{t+1}\right) \\
&= \rho\,\mathrm{Cov}_t\left(\varepsilon_{t+1}, h_{t+1}\right) + 2\gamma^2\rho c\,\mathrm{Cov}_t\left(\varepsilon_{t+1}, h_{t+1}\right) + \gamma\rho v^h_t + 2\gamma^3\rho c v^h_t \\
&= \gamma\rho\left(1 + 2\gamma^2 c\right) v^h_t\,.
\end{aligned}
$$

E Proofs

Proof of Propsition 2.2. The joint risk-neutral dynamics of log-returns and the variance process are derived following Gouriéroux and Monfort (2007) and Bertholon et al. (2008). We observe from the no arbitrage conditions

$$
\begin{aligned}
\mathrm{E}^{\mathbb{P}}_t[\mathfrak{m}_{t,t+1}\mathrm{e}^{r^f_{t+1}}] &= 1 \\
&= \mathrm{e}^{\theta_{0,t}+\theta_{2,t}h_t+B\left((\theta_{1,t}\gamma+A(\theta_{1,t})+\theta_{3,t}\right)+C\left((\theta_{1,t}\gamma+A(\theta_{1,t})+\theta_{3,t}\right)h_t}\,,
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{E}^{\mathbb{P}}_t[\mathfrak{m}_{t,t+1}\mathrm{e}^{Y_{t+1}}] &= 1\,, \\
&= \mathrm{e}^{\theta_{0,t}+\theta_{2,t}h_t+B\left((\theta_{1,t}+1)\gamma+A(\theta_{1,t}+1)+\theta_{3,t}\right)+C\left((\theta_{1,t}+1)\gamma+A(\theta_{1,t}+1)+\theta_{3,t}\right)h_t}
\end{aligned}
$$

that

$$
\begin{aligned}
\theta^*_{0,t} &= -B\left(\theta^*_{3,t} + A(\theta^*_{1,t}) + \theta^*_{1,t}\gamma\right), \\
\theta^*_{0,t} &= -B\left(\theta^*_{3,t} + A(\theta^*_{1,t}+1) + (\theta^*_{1,t}+1)\gamma\right), \\
\theta^*_{2,t} &= -C\left(\theta^*_{3,t} + A(\theta^*_{1,t}) + \theta^*_{1,t}\gamma\right), \\
\theta^*_{2,t} &= -C\left(\theta^*_{3,t} + A(\theta^*_{1,t}+1) + (\theta^*_{1,t}+1)\gamma\right),
\end{aligned}
$$

and thus

$$
A(\theta^*_{1,t}) = A(\theta^*_{1,t}+1) + \gamma\,.
$$

Solving for $\theta^*_{1,t}$ where we use $A(u)$ implied by the Meixner model, we get:

$$
\begin{aligned}
2d\log\left\{\frac{\cos\left(\frac{b}{2}\right)}{\cos\left(\frac{1}{2}(b+\breve{a}\theta^*_{1,t})\right)}\right\}-\eta_1\theta^*_{1,t} &= \\
2d\log\left\{\frac{\cos\left(\frac{b}{2}\right)}{\cos\left(\frac{1}{2}(b+\breve{a}(\theta^*_{1,t}+1))\right)}\right\} & -\eta_1(\theta^*_{1,t}+1)+\gamma \\
\frac{\cos\left(\frac{1}{2}(b+\breve{a}(\theta^*_{1,t}+1))\right)}{\cos\left(\frac{1}{2}(b+\breve{a}\theta^*_{1,t})\right)} &= \mathrm{e}^{\frac{\gamma-\eta_1}{2d}} \\
\cos\left(\frac{b+\breve{a}\theta^*_{1,t}}{2}\right)\cos\left(\frac{\breve{a}}{2}\right)-\sin\left(\frac{b+\breve{a}\theta^*_{1,t}}{2}\right)\sin\left(\frac{\breve{a}}{2}\right) &= \zeta\cos\left(\frac{b+\breve{a}\theta^*_{1,t}}{2}\right) \\
\frac{\sin\left(\frac{b+\breve{a}\theta^*_{1,t}}{2}\right)}{\cos\left(\frac{b+\breve{a}\theta^*_{1,t}}{2}\right)} &= -\frac{\zeta-\cos\left(\frac{\breve{a}}{2}\right)}{\sin\left(\frac{\breve{a}}{2}\right)} \\
\frac{b+\breve{a}\theta^*_{1,t}}{2} &= -\arctan\left(\frac{\zeta-\cos\left(\frac{\breve{a}}{2}\right)}{\sin\left(\frac{\breve{a}}{2}\right)}\right) \\
\theta^*_{1,t} &= -\frac{2}{\breve{a}}\arctan\left(\frac{\zeta-\cos\left(\frac{\breve{a}}{2}\right)}{\sin\left(\frac{\breve{a}}{2}\right)}\right)-\frac{b}{\breve{a}},
\end{aligned}
$$

where $\zeta=\exp(\frac{\gamma-\eta_1}{2d})$. □

Proof of Propsition 2.3. The conditional CGF under the risk-neutral measure $\mathbb{Q}$ is

$$
\begin{aligned}
\psi^{\mathbb{Q}}_{Y_{t+1},h_{t+1}|t}(u,v) &\equiv \log \mathrm{E}^{\mathbb{Q}}_t[\exp(uY_{t+1}+vh_{t+1})] \\
&= \log\frac{\mathrm{E}^{\mathbb{P}}_t[\mathfrak{m}_{t,t+1}\exp(uY_{t+1}+vh_{t+1})]}{\mathrm{E}^{\mathbb{P}}_t[\mathfrak{m}_{t,t+1}]} \\
&= ur^f+\psi^{\mathbb{P}}_{\varepsilon_{t+1},h_{t+1}|t}(\theta^*_1+u,\theta^*_1\gamma_1+\theta^*_3+u\gamma_1+v) \\
&\quad -\psi^{\mathbb{P}}_{\varepsilon_{t+1},h_{t+1}|t}(\theta^*_1,\theta^*_1\gamma_1+\theta^*_3) \\
&= ur^f+\psi^{\mathbb{P}}_{h_{t+1}|t}\big(A(\theta^*_1+u)+\theta^*_1\gamma_1+\theta^*_3+u\gamma_1+v\big) \\
&\quad -\psi^{\mathbb{P}}_{h_{t+1}|t}\big(A(\theta^*_1)+\theta^*_1\gamma_1+\theta^*_3\big) \\
&= ur^f+\psi^{\mathbb{P}}_{h_{t+1}|t}\big(z(\theta^*_1+u,\theta^*_3+v)\big)-\psi^{\mathbb{P}}_{h_{t+1}|t}\big(z(\theta^*_1,\theta^*_3)\big) \\
&= ur^f+B^*\big(A^*(u)+u\gamma_1+v\big)+C^*\big(A^*(u)+u\gamma_1+v\big)h_t
\end{aligned}
$$

□

F Details of estimation strategy

Throughout, π is the circle constant and i the imaginary unit defined as $\mathrm{i} = \sqrt{-1}$. From (4), we obtain the joint conditional CF

$$\phi_{Y_{t+1}-r^f_{t+1},\, h_{t+1}|t}(\mathrm{i}u, \mathrm{i}v) \equiv \exp \psi_{Y_{t+1}-r^f_{t+1},\, h_{t+1}}(\mathrm{i}u, \mathrm{i}v)$$
$$= \mathrm{e}^{B(\mathrm{i}u,\mathrm{i}v)+C(\mathrm{i}u,\mathrm{i}v)h_t} . \quad (37)$$

The CF of the latent variable h_t at time t, conditional upon observing $\{Y_k\}_{k=0}^{\ell}$ is $\phi_{h_t|\{Y_k\}_{k=0}^{\ell}}(\mathrm{i}v) \equiv \mathrm{E}\left[\mathrm{e}^{\mathrm{i}vh_t}|\{Y_k\}_{k=0}^{\ell}\right]$. The recursive updating procedure requires an approximation of the function $\phi_{h_{t+1}|\{Y_k\}_{k=0}^{t+1}}(\cdot)$. Because the unconditional distribution of the ARG process is gamma (Gouriéroux and Jasiak, 2006), we use the gamma distribution for the approximation

$$\hat{\psi}_{h_{t+1}|\{Y_k\}_{k=0}^{t+1}}(\mathrm{i}v) \equiv \log \hat{\phi}_{h_{t+1}|\{Y_k\}_{k=0}^{t+1}}(\mathrm{i}v)$$
$$= -\nu_{t+1} \log(1 - \mathrm{i}v\kappa_{t+1}) , \quad (38)$$

with the parameters ν (shape) and κ (scale).[7] At $t = 0$ we initialize the parameters for $\hat{\psi}_{h_t}$ with the parameters of the unconditional distribution of the ARG process at the chosen starting values. We define $\phi_{Y_{t+1},\, h_{t+1}|\{Y_k\}_{k=0}^{\ell}} \equiv \mathrm{E}\left[\mathrm{e}^{\mathrm{i}uY_{t+1}+\mathrm{i}vh_{t+1}}|\{Y_k\}_{k=0}^{\ell}\right]$. Bates (2006) shows that given the time-t CF of the latent variable, the time-t prior joint CF of the next period's (Y_{t+1}, h_{t+1}) conditional on the data observed until time t can be written as

$$\phi_{Y_{t+1},\, h_{t+1}|\{Y_k\}_{k=0}^{t}} = \mathrm{e}^{B(\mathrm{i}u,\mathrm{i}v)} \psi_{h_t|\{Y_k\}_{k=0}^{t}} \left[C(\mathrm{i}u, \mathrm{i}v)\right] . \quad (39)$$

Now, the density function of next period's Y_{t+1} conditional on the data observed through time t can be evaluated by Fourier inversion of its conditional CF

$$p(Y_{t+1}|\{Y_k\}_{k=0}^{t}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{Y_{t+1},h_{t+1}|\{Y_k\}_{k=0}^{t}}(\mathrm{i}u, 0)\mathrm{e}^{-\mathrm{i}uY_{t+1}}\,\mathrm{d}u . \quad (40)$$

We use that $\phi_{h_{t+1}|\{Y_k\}_{k=0}^{t+1}}(\mathrm{i}v) = \frac{1}{2\pi p(Y_{t+1}|\{Y_k\}_{k=0}^{t})} \int_{-\infty}^{\infty} \phi_{Y_{t+1},h_{t+1}|\{Y_k\}_{k=0}^{t}}(\mathrm{i}u, \mathrm{i}v)\mathrm{e}^{-\mathrm{i}uY_{t+1}}\,\mathrm{d}u$ to obtain the filtered estimate $\hat{h}_{t+1|\{Y_k\}_{k=0}^{t+1}}$ for h_{t+1} and of its variance from the derivatives of

[7] The probability density function corresponding to this parameterization of the gamma distribution is $f(x) = \frac{1}{\Gamma(\nu)\kappa^{\nu}} x^{\nu-1}\mathrm{e}^{-\frac{x}{\kappa}}$.

the MGF $\phi_{h_{t+1}|\{Y_k\}_{k=0}^{t+1}}(v)$ evaluated at 0

$$\begin{aligned}\hat{h}_{t+1|\{Y_k\}_{k=0}^{t+1}} &= \phi'_{h_{t+1}|\{Y_k\}_{k=0}^{t+1}}(0) \\ &= \frac{1}{2\pi p(Y_{t+1}|\{Y_k\}_{k=0}^{t})} \int_{-\infty}^{\infty} \phi'_{Y_{t+1},h_{t+1}|\{Y_k\}_{k=0}^{t}}(\mathrm{i}u, 0)\mathrm{e}^{-\mathrm{i}uY_{t+1}}\,\mathrm{d}u \end{aligned} \tag{41}$$

$$\begin{aligned}\mathrm{Var}(h_{t+1}|\{Y_k\}_{k=0}^{t+1}) &=\phi''_{h_{t+1}|\{Y_k\}_{k=0}^{t+1}}(0) - \hat{h}^2_{t+1|\{Y_k\}_{k=0}^{t+1}} \\ &= \frac{1}{2\pi p(Y_{t+1}|\{Y_k\}_{k=0}^{t})} \int_{-\infty}^{\infty} \phi''_{Y_{t+1},h_{t+1}|\{Y_k\}_{k=0}^{t}}(\mathrm{i}u, 0)\mathrm{e}^{-\mathrm{i}uY_{t+1}}\,\mathrm{d}u \\ &\quad - \hat{h}^2_{t+1|\{Y_k\}_{k=0}^{t+1}}\,. \end{aligned} \tag{42}$$

We denote by $\phi'_{Y_{t+1},h_{t+1}|\{Y_k\}_{k=0}^{t}}(\mathrm{i}u, 0)$ and $\phi''_{Y_{t+1},h_{t+1}|\{Y_k\}_{k=0}^{t}}(\mathrm{i}u, 0)$ the first and second-order derivatives of $\phi_{Y_{t+1},h_{t+1}|\{Y_k\}_{k=0}^{t}}(\mathrm{i}u, \mathrm{i}v)$ with respect to v evaluated at $v = 0$, respectively.

Finally, the parameters of the CF of the gamma distribution for the approximation of $\phi_{h_{t+1}|\{Y_k\}_{k=0}^{t+1}}$ are updated by

$$\kappa_{t+1} = \frac{\mathrm{Var}\left(h_{t+1}|\{Y_k\}_{k=0}^{t+1}\right)}{\hat{h}_{t+1|\{Y_k\}_{k=0}^{t+1}}}\,, \tag{43}$$

$$\nu_{t+1} = \frac{\hat{h}^2_{t+1|\{Y_k\}_{k=0}^{t+1}}}{\mathrm{Var}\left(h_{t+1}|\{Y_k\}_{k=0}^{t+1}\right)}\,, \tag{44}$$

following the method of moments.

The double-exponential approximation formula for the integral in (40) according to Ooura (2005) is

$$\begin{aligned}\tilde{p}_h^{(N)}(Y_{t+1}) =& \frac{1}{\pi}\,\mathrm{Re}\Bigg\{ \frac{2\pi\mathrm{i}}{\tilde{Y}} \sum_{n=-N_-}^{N_+} \phi_{Y_{t+1},h_{t+1}|t}\left(\frac{\pi}{\tilde{Y}h} g(nh), 0\right) \sin\left(\frac{\pi}{2h}\tilde{g}(nh)\right) \\ & g'(nh)\exp\left(\frac{\pi\mathrm{i}(-Y_{t+1})}{\tilde{Y}h} g(nh) - \frac{\pi\mathrm{i}}{2h}\tilde{g}(nh)\right)\Bigg\} \end{aligned} \tag{45}$$

Let $\mathrm{Re}(z)$ denote the real part of a complex number z. We use that the integrand for negative u is the complex conjugate of the one with positive u values. Thus the integral can be evaluated by taking the real part of the integrand over $[0, \infty)$ and doubling the result. $\tilde{Y}$ is chosen such that $-Y_{t+1} \in (0, 2\tilde{Y})$, and

$$g(x) = \frac{x}{1 - \exp(-2x - \alpha(1 - \mathrm{e}^{-x}) - \beta(\mathrm{e}^{x} - 1))}\,, \tag{46}$$

$$\tilde{g}(x) = g(x) - x\,, \tag{47}$$

where $\alpha = \beta \left(\frac{1+\log\left(1+\frac{\pi}{\tilde{Y}h}\right)}{4\tilde{Y}h} \right)^{-\frac{1}{2}}$, and $\beta = 0.25$. For positive values of Y_{t+1}, we multiply Y_{t+1} by (-1) and use the complex conjugate of ϕ.

Further parameters of the approximation formula are N_-, N_+, and h. We use $N_- = 94$, $N_+ = 69$, and $h = 0.075$. These parameter values are, among other combinations, suggested by Ooura (2005). Our estimation results are very robust to the choice of these parameter combinations. In general, when the mesh size h is reduced, accuracy is expected to improve but larger values of N_- and N_+ are required. The approximation of the integrals (41) and (42) works equivalently.

G Carr Madan representation

Applying risk-neutral valuation principles, we find the price of a European style call with strike K and expiry date $T = t + \vartheta$ as

$$P_c(K,T) = e^{x_t}\Pi_1 - e^{-r^f\vartheta} K\Pi_2 \tag{48}$$

where $x_t = \log S_t$ and Π_j, for $j = 1,2$, is the probability that $x_T > \log K$ obtained under different probability measures.

We use the representation of Carr and Madan (1999) to price and define the modified call price:

$$c(k) = e^{\alpha k} P_c(K,T) \tag{49}$$

which includes the dampening factor α. The Fourier transform of $c(k)$ can be written as

$$\hat{c}(u) = \frac{e^{-r^f\vartheta} \exp(\psi^{\mathbb{Q}}_{Y_t(\vartheta)|t}(u - (\alpha+1)i) S_T^{iu+\alpha-1}}{\alpha^2 + \alpha - u^2 + iu(2\alpha+1)}. \tag{50}$$

The call price is then given by the following transform of the modified call price

$$P_c(K,T) = \frac{e^{-\alpha k}}{\pi} \int_0^\infty \mathrm{Re}\left[e^{-iuk}\hat{c}(u) \right] du. \tag{51}$$

For the single numerical integration required for pricing according to (51), we again employ the double exponential formula of Ooura (2005).