BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel—2023-40
Last updated: August 2, 2023

Generalized Autoregressive Gamma Processes

by Bruno Feunou

Financial Markets Department
Bank of Canada
bfeunou@bankofcanada.ca

DOI: https://doi.org/10.34989/swp-2023-40 | **ISSN** 1701-9397

Acknowledgements

I am indebted to Atsushi Inoue (the editor) and the associate editor from the *Journal of Business and Economic Statistics*, three anonymous referees, Virginie Traclet and Jean-Sébastien Fontaine for helpful comments that improved the article. I pay a special tribute to my dear father, Kamkui Emmanuel, who passed away on January 13, 2022: "Father, your guidance and support greatly shaped me and everything I do today, including this research agenda. Thank you for everything and rest in peace." I am grateful to Liam Lindsay for excellent research assistance. The views expressed in this paper are mine and do not necessarily reflect those of the Bank of Canada.

Abstract

We introduce generalized autoregressive gamma (GARG) processes, a class of autoregressive and moving-average processes that extends the class of existing autoregressive gamma (ARG) processes in one important dimension: each conditional moment dynamic is driven by a different and identifiable moving average of the variable of interest. The paper provides ergodicity conditions for GARG processes and derives closed-form conditional and unconditional moments. The paper also presents estimation and inference methods, illustrated by an application to European option pricing where the daily realized variance follows a GARG dynamic. Our results show that using GARG processes reduces pricing errors by substantially more than using ARG processes does.

Topics: Econometric and statistical methods; Asset pricing
JEL codes: C58, G12

Résumé

Nous présentons les processus gamma autorégressifs généralisés (GARG), une catégorie de processus autorégressifs et moyennes mobiles qui est un prolongement de la catégorie existante de processus gamma autorégressifs dans une dimension importante : la dynamique de chacun des moments conditionnels est influencée par une différente moyenne mobile identifiable de la variable d'intérêt. Nous fournissons les conditions d'ergodicité pour les processus GARG et en établissons les moments conditionnels et inconditionnels de forme fermée. Nous présentons aussi des méthodes d'estimation et d'inférence, puis les appliquons à l'évaluation d'options européennes où la variance quotidienne réalisée suit la dynamique des processus GARG. Nos résultats montrent que l'utilisation de ces processus réduit les erreurs d'évaluation de façon nettement plus importante que les processus gamma autorégressifs.

Sujets : Méthodes économétriques et statistiques; Évaluation des actifs
Codes JEL : C58, G12

1 Introduction

The finance literature has widely relied on autoregressive gamma (ARG) processes to model variations in the distribution of positive time series. In modeling the term structure of interest rates, volatility factors are traditionally designed to follow ARG dynamics (Le et al., 2010; Monfort et al., 2017). In option pricing, when modeling the key component, the conditional variance, several papers use ARG dynamics (Feunou and Tedongap, 2012; Majewski et al., 2015). ARG processes have also been used to model intraday financial market activity, in particular intertrade duration (Gourieroux and Jasiak, 2006; Gourieroux et al., 1999).

The ARG process studied in Gourieroux and Jasiak (2006) corresponds to the discretization of the Cox-Ingersoll-Ross diffusion process (Cox et al., 1985). It is a univariate positive random process whose cumulant generating function is defined for the scalar $u < 1/\varphi$ and given by:

$$\psi_t(u) \equiv \ln\left[E\left[\exp\left(ux_{t+1}\right)|I_t\right]\right] = \omega(u) + \alpha(u)x_t, \tag{1}$$

where I_t is the sigma algebra generated by $(x_s, s \leq t)$ and

$$\omega(u) = -\nu\log(1-u\varphi), \text{ and } \alpha(u) = \frac{\phi u}{1-u\varphi}, \tag{2}$$

with $\nu \geq 0$, $\varphi > 0$ and $\phi \geq 0$. It admits the following state space representation:

$$\begin{aligned} \frac{x_{t+1}}{\varphi}|U_{t+1}, I_t &\sim \gamma\left(\nu + U_{t+1}\right) \\ U_{t+1}|I_t &\sim P\left(\frac{\phi x_t}{\varphi}\right), \end{aligned}$$

where U_{t+1} is a latent process that follows a Poisson distribution denoted by $P(\cdot)$ and $\gamma(\cdot)$ is the standard gamma distribution.

Gourieroux and Jasiak (2006) provide the following generalization to any order (p, q):

$$\psi_t(u) \equiv \ln\left[E\left[\exp\left(ux_{t+1}\right)|I_t\right]\right] = \omega(u) + \alpha(u)m_t, \tag{3}$$

where

$$m_t = \sum_{j=0}^{p-1} \phi_j x_{t-j} + \sum_{k=1}^{q} \theta_k m_{t-j}, \tag{4}$$

$\phi_0 = 1$ for identification, $p \geq 1$ and $q \geq 1$.

One implication of this ARG process in definition (3) is that all the conditional cumulants (the derivatives of $\psi_t(u)$ at $u = 0$) are driven by the same factor m_t. Indeed, we have

$$\psi_t^{(n)}(0) = \omega^{(n)}(0) + \alpha^{(n)}(0) m_t, \tag{5}$$

with $f^{(n)}(u)$, the n^{th} order derivative of function $f(\cdot)$ at u.

This suggests that both the conditional expectation and the conditional variance of x_t are driven by m_t, and are, therefore, perfectly positively correlated, and that all moments are highly positively correlated. There is considerable empirical evidence contradicting the very tight restriction between the first two moments imposed by affine models, in particular when considering interest-rate and variance modeling (Cieslak and Povala, 2016). Using swap data, Collin-Dufresne et al. (2004) find that a popular and well-documented three-factor affine model implies volatility paths that are negatively correlated with the GARCH volatility estimates of weekly changes in the six-month rate. Andersen and Benzoni (2010) use intraday Treasury data to show that realized yield volatility is unrelated to principal components extracted from the cross-section, which proxy for model-implied volatility. Regarding the dynamic of the stock market, there is also evidence that the expectation of the realized variance has a distinct dynamic apart from the variance of the realized variance (Corsi et al., 2008).

Numerous contributions in the literature focus on building complex parametric and semi-parametric time series models where the first four moments have very distinct dynamics. Hansen (1994) and Jondeau and Rockinger (2003) are pioneers in that literature and have shown that in the first four moments, behavior and persistence are very different.

For instance, the skewness is strongly persistent while kurtosis is much less so. Chang et al. (2011) extract non-parametric measures of risk-neutral variance and skewness from option prices and show that the correlation between option-implied skewness and option-implied volatility for the S&P500 is -0.06 and the correlation between the average option-implied skewness and average option-implied volatility for the S&P100 components is 0.05.

This inability of affine dynamics to fit all the moments jointly implies that they cannot fit the conditional density and, hence, they generate large option pricing errors, as shown in our empirical investigations. To mitigate these shortcomings of affine models in general, and ARG in particular, we introduce GARG processes to extend ARG processes in one important dimension: each conditional moment dynamic is driven by a different moving average of the variable of interest (x_t). Moreover, GARG processes are parsimonious and, importantly, they maintain one of the key advantages of ARG processes: a closed-form multi-step ahead distribution. Hence, we are able to add much-needed complexity to the ARG dynamic while keeping its main advantage, i.e., computing derivative prices in closed-form.

The key principle of our generalization of ARG processes and affine models in general to a GARG dynamic is simple: contrary to affine dynamics where all the cumulants are driven by the same factor m_t (see equation (5)), we want each cumulant $\psi_t^{(n)}(0)$ to be driven by its own specific factor (say, $m_t^{(n)}$), that is,

$$\psi_t^{(n)}(0) = \omega_n + \alpha_n m_t^{(n)}, \tag{6}$$

where $m_t^{(n)}$ is a moving average of x_t:

$$m_t^{(n)} = x_t + \theta_n m_{t-1}^{(n)}. \tag{7}$$

One way to achieve that with a minimal number of additional parameters is to set $\theta_n = \beta\theta^n$, which is equivalent to the following recursive formulation of the conditional cumulant generating function:

$$\psi_t(u) = \omega(u) + \alpha(u)x_t + \beta\psi_{t-1}(\theta u) \qquad \text{for } t \geq 1, \tag{8}$$

where functions $\alpha(\cdot)$ and $\omega(\cdot)$ are given in equation (2). The main theoretical challenge of this paper is to build a process whose the cumulant generating function has the recursive formulation in equation (8). We successfully tackle that challenge, and our findings go beyond ARG dynamics as we extend this result to other positive-valued affine models. This includes the autoregressive gamma-zero (ARGZero) processes of Monfort et al. (2017), the integer-valued autoregressive processes (INAR) of Al-Osh and Alzaid (1987), the affine GARCH models of Heston and Nandi (2000) and Christoffersen et al. (2006).

Our approach is not the only way to enhance ARG and disentangle its moments dynamics. An alternative is to model x_t as the sum of independent univariate ARG dynamics, which we refer to as the multi-factor ARG, or MARG for short. The MARG pioneered by Le et al. (2010) is the leading framework in discrete-time affine term structure of interest rate models. Its appeal is simple and intuitive: if one is interested in disentangling the first K moments, it is enough to sum K independent ARGs. However, as we discuss in detail in section 5, the MARG loses its analytical tractability when we analyse the dynamic conditional only on the variable of interest x_t. Another approach in option pricing is the Wishart autoregressive (WAR) process studied in details in Gourieroux et al. (2009), Yu et al. (2017) and Gourieroux and Sufana (2010). Although this is the leading extension of the ARG dynamic, it suffers from the same shortcomings as the MARG.

This paper is organized as follows. Section 2 defines the generalized autoregressive gamma process (GARG), with a conditional distribution from a convolution of non-centered gamma and the noncentrality parameters written as linear functions of lagged variables. Section 3 derives the short-term and long-term dynamics of conditional moments. Section 4 discusses ergodicity conditions and derives moments of the unconditional distribution.

Section 5 compares the GARG and MARG models. Section 6 discusses issues related to identification and statistical inference. An application to option pricing is presented in section 7, where we show that the GARG dynamic dominates ARG. A final section concludes the paper. Proofs and extensions are gathered in the online appendices.

2 Specification

2.1 GARG dynamics

We consider a univariate positive random process x_t with $t \geq 1$ and generalize the ARG process to the GARG process, built through the following state space representation:

$$x_{t+1} = \bar{Z}_{t+1} + \mathbf{1}_{[t>0]} \left[\sum_{j=0}^{t-1} Z_{t+1}^{(j)} \right], \quad t \geq 0, \tag{9}$$

where $\mathbf{1}_{[\cdot]}$ is an indicator function, and for $t > 0$, $\bar{Z}_{t+1}$ and $Z_{t+1}^{(j)}$ with $j = 0, ..., t-1$ are $t+1$ conditionally (conditional on I_t) independent random variables with the following state-space representation:

$$\frac{Z_{t+1}^{(j)}}{\varphi_j} | U_{t+1}^{(j)}, I_t \sim \gamma \left(\nu_j + U_{t+1}^{(j)} \right) \tag{10}$$

$$U_{t+1}^{(j)} | I_t \sim P \left(\frac{\phi_j x_{t-j}}{\varphi_j} \right), \tag{11}$$

where

$$\nu_j = \nu \beta^j, \quad \varphi_j = \varphi \theta^j, \quad \phi_j = \phi \beta^j \theta^j. \tag{12}$$

The cumulant generating function of $\bar{Z}_{t+1}$ is $\beta^t \psi_0 \left(\theta^t u \right)$, with

$$\psi_0 \left(u \right) = \frac{\phi}{1 - \beta\theta} \frac{\theta u}{1 - \theta\varphi u} \mu - \frac{\nu}{1 - \beta\theta} \ln \left(1 - \theta\varphi u \right). \tag{13}$$

The GARG has five parameters, ν, φ, ϕ, β and θ, with the following restrictions:

$$\nu, \varphi, \phi, \beta, \theta \geq 0, \quad \beta\theta < 1. \tag{14}$$

The ARG dynamic is nested within the GARG and is obtained by setting $\beta = 0$ or $\theta = 0$.

2.2 Alternative formulation of the GARG

The parametric forms for the coefficients ν_j, φ_j and ϕ_j given in equation (12) bear no particular intuition or economic meaning. These forms are chosen to get a recursive dynamic for the cumulant generating function (8).

We can now establish the main result of this paper.

Proposition 1 *Let us assume that the positive-valued univariate process of interest x_t follows the dynamics described in equations (9), (10), and (11). Then, for a scalar u such that $1 - u\varphi_j > 0$ for all $j > 0$, the conditional cumulant generating function of x_{t+1} ($\psi_t(u)$) exists and evolves according to the following recursive dynamic:*

$$\psi_t(u) = \omega(u) + \alpha(u)x_t + \beta\psi_{t-1}(\theta u) \qquad \text{for } t \geq 1, \tag{15}$$

where functions $\alpha(\cdot)$ and $\omega(\cdot)$ are given in equation (2).

The condition for the existence of the cumulant generating function in equation (15) is $1 - u\varphi_j > 0$ for all $j > 0$. That condition is equivalent to $u < \frac{1}{\varphi} \min_{j \geq 0} \left\{ \left(\frac{1}{\theta}\right)^j \right\}$. Hence, the set of values for u depends on θ: $\begin{cases} \theta \leq 1, & u \leq \frac{1}{\varphi\theta}; \\ \theta > 1, & u \leq 0. \end{cases}$ The fact that the cumulant generating function only exists for negative arguments when $\theta > 1$ is an argument for constraining θ to be below one. Indeed, in most applications, we would build positive processes that are obtained as linear combinations of GARG processes with positive loadings. We show in section 4 that $\theta \leq 1$ is a necessary condition for weak ergodicity.

Equation (15) is an alternative formulation of the new model. In other words, the GARG dynamic has two equivalent formulations: The first one is the state-space representation given by equations (9), (10) and (11), which is useful for the simulation of the GARG dynamic. The second one is specified by means of the conditional cumulant generating function ($\psi_t(\cdot)$) and is given by equation (15). All the results and implications of this paper use the second formulation of the model.

The GARG dynamics are nested within the class of generalized affine models introduced by Feunou and Meddahi (2009). Belonging to the class of generalized affine dynamics is essential for the derivation of closed-form multi-step ahead dynamics. While Feunou and Meddahi (2009) focus on building dynamics on the cumulant generating function directly, which presents some challenging theoretical issues, the GARG dynamic is well defined by construction, and the recursion given by equation (15) is a well-defined cumulant generating function at any point in time t.

Proof of Proposition 1 Using equation (9) and the fact that all the Zs on the right-hand side are conditionally independent, we have:

$$\psi_t(u) = \ln E_t\left[\exp\left(u\bar{Z}_{t+1}\right)\right] + \sum_{j=0}^{t-1} \ln E_t\left[\exp\left(uZ_{t+1}^{(j)}\right)\right].$$

By assumption,

$$\ln E_t\left[\exp\left(u\bar{Z}_{t+1}\right)\right] = \beta^t\psi_0\left(\theta^t u\right),$$

and the state-space representation given by equations (10) and (11) implies that

$$\ln E_t\left[\exp\left(uZ_{t+1}^{(j)}\right)\right] = \beta^j\omega(\theta^j u) + \beta^j\alpha(\theta^j u)x_{t-j}.$$

Hence,

$$\begin{aligned}
\psi_t(u) &= \beta^t\psi_0\left(\theta^t u\right) + \sum_{j=0}^{t-1}\beta^j\left[\omega(\theta^j u) + \alpha(\theta^j u)x_{t-j}\right] \\
&= \omega(u) + \alpha(u)x_t + \beta^t\psi_0\left(\theta^t u\right) + \sum_{j=1}^{t-1}\beta^j\left[\omega(\theta^j u) + \alpha(\theta^j u)x_{t-j}\right] \\
&= \omega(u) + \alpha(u)x_t + \beta\left\{\beta^{t-1}\psi_0\left(\theta^{t-1}\theta u\right) + \sum_{k=0}^{t-2}\beta^k\left[\omega(\theta^k\theta u) + \alpha(\theta^k\theta u)x_{t-1-k}\right]\right\} \\
&= \omega(u) + \alpha(u)x_t + \beta\psi_{t-1}(\theta u),
\end{aligned}$$

which establishes proposition 1.

2.3 Extensions

Our model can be generalized in several directions. In sections 1.3, 1.4 and 1.5 of the Appendix we show that other positive-valued affine models such as the GARCH models of Heston and Nandi (2000) and Christoffersen et al. (2006) and the integer-valued autoregressive (INAR) of Al-Osh and Alzaid (1987) can also be generalized using similar techniques. We also discuss a multivariate and a multi-lag extension of the GARG dynamic in sections 1.1 and 1.2 of the Appendix.

Here we extend our generalization of the ARG to the autoregressive gamma-zero (ARG_0) of Monfort et al. (2017) which is essential for term structure of interest rates modelling at the zero-lower bound as it encompasses a zero-point mass, which is not possible with ARG dynamic.

Generalized autoregressive gamma-zero processes First, both the ARG_0 and the ARG can be written within a single affine framework (also known as extented-ARG processes) as follows:

$$\psi_t(u) \equiv \ln\left[E\left[\exp\left(ux_{t+1}\right)|I_t\right]\right] = \omega_0(u) + \alpha(u)x_t, \tag{16}$$

where I_t is the sigma algebra generated by $(x_s, s \leq t)$, and

$$\omega_0(u) = -\nu\log(1-u\varphi) + \frac{\xi u}{1-u\varphi} \text{ and } \alpha(u) = \frac{\phi u}{1-u\varphi}. \tag{17}$$

The ARG dynamic is obtained by setting $\xi = 0$ while the ARG_0 dynamic is obtained by setting $\nu = 0$. Using the decomposition of x_{t+1} given in equation (9), we generalize these extented-ARG processes in the following state-space representation:

$$\begin{aligned} \frac{Z_{t+1}^{(j)}}{\varphi_j}|U_{t+1}^{(j)}, I_t &\sim \gamma\left(\nu_j + U_{t+1}^{(j)}\right) \\ U_{t+1}^{(j)}|I_t &\sim P\left(\frac{\xi_j + \phi_j x_{t-j}}{\varphi_j}\right), \end{aligned}$$

where ν_j, φ_j and ϕ_j are given in equation (12) and $\xi_j \equiv \psi(\beta\theta)^j$. Using the same steps as for the generalization of the ARG dynamic, we show that the cumulant generating function of the generalized autoregressive gamma-zero processes follows a recursion similar to equation (15):

$$\psi_t(u) = \omega_0(u) + \alpha(u)x_t + \beta\psi_{t-1}(\theta u) \quad \text{for } t \geq 1. \tag{18}$$

For the remainder of this paper, we focus on the GARG dynamic given by equations (9), (10) and (11). However, all our findings apply to any of the extensions discussed in this section. To keep that degree of generality, we will not use the explicit expression of functions $\alpha(\cdot)$ and $\omega(\cdot)$ given in equation (2).

3 Conditional moments dynamic

3.1 Moments dynamic

In this section, we derive the dynamics of conditional moments, including expectation, variance, skewness and kurtosis. As equation (15) specifies the dynamic of the log-conditional moment generating function (cumulant generating function), it is convenient to derive the law of motion of conditional cumulants (derivatives of the cumulant generating function at 0). From equation (15), we have

$$\psi_t^{(n)}(0) = \omega^{(n)}(0) + \alpha^{(n)}(0)x_t + \beta\theta^n\psi_{t-1}^{(n)}(0), \tag{19}$$

in particular,

$$\begin{aligned} \psi_t'(0) &= \omega'(0) + \alpha'(0)x_t + \beta\theta\psi_{t-1}'(0) && (20) \\ \psi_t''(0) &= \omega''(0) + \alpha''(0)x_t + \beta\theta^2\psi_{t-1}''(0). && (21) \end{aligned}$$

All positive affine processes considered in this paper share one important property: all derivatives $\omega^{(n)}(0)$ and $\alpha^{(n)}(0)$ are positive for all n. We can also easily establish that x_t

is an ARMA(1,1) with the autoregressive parameter $\alpha'(0) + \beta\theta$ and the moving average parameter $\beta\theta$. Indeed, using equation 20, we have

$$\begin{aligned} x_{t+1} &= \psi'_t(0) + \underbrace{x_{t+1} - \psi'_t(0)}_{u_{t+1}} = \underbrace{\omega'(0) + \alpha'(0)\,x_t + \beta\theta\psi'_{t-1}(0)}_{=\psi'_t(0)} + u_{t+1} \\ &= \omega'(0) + \alpha'(0)\,x_t + \beta\theta\underbrace{(x_t - u_t)}_{\psi'_{t-1}(0)} + u_{t+1} = \omega'(0) + (\alpha'(0) + \beta\theta)\,x_t + u_{t+1} - \beta\theta u_t \end{aligned} \quad (22)$$

3.2 Importance of parameters β and θ

Equation (19) implies that:

$$\psi_t^{(n)}(0) = \frac{\omega^{(n)}(0)}{1 - \beta\theta^n} + \alpha^{(n)}(0)\left(\sum_{j=0}^{\infty} (\beta\theta^n)^j\, x_{t-j}\right). \quad (23)$$

Consequently, each conditional cumulant (i.e., $\psi_t^{(n)}(0)$ for a given n) is driven by its own factor ($m_t^{(n)}$):

$$m_t^{(n)} = \sum_{j=0}^{\infty} (\beta\theta^n)^j\, x_{t-j}, \quad (24)$$

which is a moving average of the variable of interest x_t. Hence, with only two additional parameters (β and θ), we are able to generate a parsimonious generalization of ARG processes that disentangles the dynamics of all the conditional moments. Further, the ability to disentangle moments dynamics stems from parameter θ. Indeed, when $\theta = 1$, all the conditional cumulants are driven by the same factor, $\sum_{j=0}^{\infty} \beta^j x_{t-j}$, and thus are perfectly positively correlated.

To confirm and complete this central point, we compute the implied correlation between two conditional cumulants at different orders n and m. If $\rho < 1$, we establish (section 2 of the Appendix) that the correlation between $\psi_t^{(n)}(0)$ and $\psi_t^{(m)}(0)$ is given by

$$Corr\left(\psi_t^{(n)}(0), \psi_t^{(m)}(0)\right) = \sqrt{\frac{1 - \left[\frac{\beta(\theta^n - \theta^m)}{1 - \beta^2\theta^{n+m}}\right]^2}{1 - \left[\frac{\beta(\theta^n - \theta^m)}{\xi}\right]^2}}, \quad (25)$$

where

$$\xi \equiv \left(1-\rho\beta\theta^{n}\right)\left(1-\rho\beta\theta^{m}\right)\vartheta+\frac{\beta\theta^{n}}{1-\rho\beta\theta^{m}}+\frac{\beta\theta^{m}}{1-\rho\beta\theta^{n}},$$

with

$$\rho \equiv \alpha'\left(0\right)+\beta\theta, \tag{26}$$

and $\vartheta=\frac{1-\rho^{2}+(\alpha'(0))^{2}}{\alpha'(0)(1-\rho\beta\theta)}$. From equation (25), it is readily apparent that $\theta=1$ implies that all the cumulants are perfectly correlated and, thus, $\theta\neq 1$ is essential to break the tight links between moments that are inherent within ARG processes.

3.3 Initial cumulant $\psi_0(u)$ and the dynamic of x_1

In practice, when computing the conditional cumulant function through recursion (15), we need a starting cumulant function $\psi_0(u)$. We set $\psi_0(u)$ to the unconditional average $E\left[\psi_t\left(u\right)\right]$. This is similar to the practice in the GARCH literature, where the initial variance is typically set to the unconditional expectation of the conditional variance process. Under the conditions $\rho<1$ and $\beta\theta^{n}<1$, we show in section 2 of the Appendix that the unconditional expectation of $\psi_t^{(n)}(0)$ is given by

$$E\left[\psi_t^{(n)}\left(0\right)\right]=\frac{\alpha^{(n)}(0)\omega'(0)+\left(1-\rho\right)\omega^{(n)}(0)}{\left(1-\rho\right)\left(1-\beta\theta^{n}\right)}. \tag{27}$$

We can thus derive the unconditional expectation of $\psi_t(u)$ using the following identity:

$$E\left[\psi_t\left(u\right)\right]=\sum_{n=1}^{\infty}\frac{u^{n}}{n!}E\left[\psi_t^{(n)}\left(0\right)\right]=\sum_{n=1}^{\infty}\frac{u^{n}}{n!}\frac{\alpha^{(n)}(0)\mu+\omega^{(n)}(0)}{1-\beta\theta^{n}}=\mu\left\{\sum_{n=1}^{\infty}\frac{u^{n}}{n!}\frac{\alpha^{(n)}(0)}{1-\beta\theta^{n}}\right\}+\left\{\sum_{n=1}^{\infty}\frac{u^{n}}{n!}\frac{\omega^{(n)}(0)}{1-\beta\theta^{n}}\right\},$$

where $\mu\equiv\frac{\omega'(0)}{1-\rho}$ is the unconditional expectation of x_t. Using functions $\omega(u)$ and $\alpha(u)$ defined in equation (2), we deduce that

$$E\left[\psi_t\left(u\right)\right]=\frac{\phi}{1-\beta\theta}\frac{\theta u}{1-\theta\varphi u}\mu-\frac{\nu}{1-\beta\theta}\ln\left(1-\theta\varphi u\right). \tag{28}$$

It is also worth stressing that $E\left[\psi_t\left(u\right)\right]$ is not the unconditional cumulant function of x_t. Later we discuss the conditions required for the unconditional distribution (and hence the unconditional cumulant function) of x_t to exist. Since $\psi_0(u)$ given in equation (28) is the cumulant generating function of x_1, it implies that x_1 has the following state-space representation:

$$\begin{aligned} \frac{x_1}{\theta\varphi}|U &\sim \gamma\left(\frac{\nu}{1-\beta\theta}+U\right) \\ U &\sim P\left(\frac{\phi}{1-\beta\theta}\frac{\mu}{\varphi}\right). \end{aligned} \tag{29}$$

3.4 Multi-horizon dynamic

Like affine models, an important characteristic of GARG processes is the existence of a closed-form forecast of any nonlinear transformation of a GARG process at any horizon. This characteristic enables financial applications such as closed-form bonds and option pricing. The multi-horizon cumulant generating function defined as

$$\psi_t(u;h) \equiv \ln\left[E_t\left[\exp\left(ux_{t+h}\right)\right]\right]$$

is computed analytically in section 3.1 of the Appendix where we establish that:

$$\begin{aligned} \psi_t(u;h) &= \sum_{j=1}^{h} \beta^{j-1}\psi_t\left(\theta^{j-1}u_j\right) + \sum_{j=2}^{h}\sum_{i=0}^{j-2}\beta^i\omega\left(\theta^i u_j\right) \text{ for } h \geq 2 \qquad (30) \\ u_h &= u, \quad u_\tau = \sum_{i=\tau+1}^{h} \beta^{i-(\tau+1)}\alpha\left(\theta^{i-(\tau+1)}u_i\right) \text{ for } 1 \leq \tau \leq h-1. \end{aligned}$$

The derivatives of $\psi_t(u;h)$ at $u=0$ give closed-form expressions of moments of the time t distribution of x_{t+h}. We show in section 3.2 of the Appendix that

$$\begin{aligned} \psi_t^{(n)}(0;h) &= \sum_{k=1}^{n} \mathcal{D}_n^{\circ h-1}\left(\bar{C}_n\right)[n,k]\left[\psi_t^{(k)}(0) - \frac{\omega^{(k)}(0)}{1-\beta\theta^k}\right] \\ &\quad + \sum_{k=1}^{n}\left[\beta\theta^k\mathcal{D}_n^{\circ h-1}\left(\bar{C}_n\right) + \left(1-\beta\theta^k\right)\left(\sum_{\tau=0}^{h-1}\mathcal{D}_n^{\circ\tau}\left(\bar{C}_n\right)\right)\right][n,k]\frac{\omega^{(k)}(0)}{1-\beta\theta^k}, \end{aligned} \tag{31}$$

where

$$\bar{C}_n \equiv \begin{bmatrix} 1 & 0_{1\times n-1} \\ 0_{n-1\times 1} & 0_{n-1\times n-1} \end{bmatrix}, \quad \mathcal{D}_n(X) \equiv \mathcal{B}_n(XA_n) + \beta X\Theta_n, \tag{32}$$

$$\Theta_n \equiv \begin{bmatrix} \theta^1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \theta^n \end{bmatrix}, \quad A_n \equiv \begin{pmatrix} \alpha^{(1)}(0) \\ \vdots \\ \alpha^{(n)}(0) \end{pmatrix}, \quad \mathcal{B}_n(z) \equiv \begin{bmatrix} B_{1,1}(z) & 0 & 0 \\ \vdots & \ddots & 0 \\ B_{n,1}(z) & \cdots & B_{n,n}(z) \end{bmatrix}, \tag{33}$$

with $B_{n,k}$ being the partial or incomplete exponential Bell polynomial, $\mathcal{D}_n^{\circ h}$ the function $\mathcal{D}_n$ compounded h times with itself and $M[n,k]$ the $(n,k)th$ entry of a generic matrix M. We provide the definition and explicit expression of $B_{n,k}$ in section 4.3 of the Appendix.

4 Ergodicity and unconditional distribution

4.1 Weak ergodicity

This section discusses weak ergodicity conditions, which are conditions under which the distribution at horizon h tends to a limiting distribution. For more on ergodicity, see Darolles et al. (2006), where the focus is on affine processes. Weak ergodicity is equivalent to the convergence of the multi-horizon cumulant generating function, which is also equivalent to the convergence of the h-step ahead n-th conditional cumulant $\psi_t^{(n)}(0;h)$ derived in equation (31) as h increases and for all n.

Let us denote $\mathcal{X}_h^{(n)} \equiv \mathcal{D}_n^{\circ h}\left(\bar{C}_n\right)$, where matrix $\bar{C}_n$ and matrix function $\mathcal{D}_n(\cdot)$ are given in equation (32). $\psi_t^{(n)}(0;h)$ converges if and only if

$$\lim_{h\to\infty}\sum_{\tau=0}^{h}\mathcal{X}_\tau^{(n)} < \infty. \tag{34}$$

A necessary condition for the convergence of the series $\sum_{\tau=0}^{h}\mathcal{X}_\tau^{(n)}$ is $\lim_{h\to\infty}\mathcal{X}_h^{(n)} = 0$, which implies that $\lim_{h\to\infty}\psi_t^{(n)}(0;h)$ is independent of t. In section 4.1 of the Appendix, we show the following result:

Proposition 2 $\psi_t^{(n)}(0;h)$ *converges as* h *increases if and only if* $\rho < 1$ *and* $\beta\theta^j < 1$ *for* $j = 1, ..., n$.

Recall that $\rho \equiv \alpha'(0) + \beta\theta = \phi + \beta\theta$, thus the following corollary:

Corollary 1 *If $\phi < 1$, $\beta\theta < 1 - \phi$, and $\theta \leq 1$, the h-step ahead conditional distribution of a GARG process converges as h increases.*

Assuming that $\phi < 1$, $\beta\theta < 1 - \phi$, and $\theta \leq 1$, let us denote $\mathcal{Y}_n \equiv \lim_{h\to\infty} \sum_{\tau=0}^{h} \mathcal{X}_{\tau}^{(n)}$. We have

$$\psi_{\infty}^{(n)} \equiv \lim_{h\to\infty} \psi_t^{(n)}(0; h) = \sum_{k=1}^{n} \mathcal{Y}_n [n, k] \, \omega^{(k)}(0),$$

and the cumulant generating function of the unconditional distribution is

$$\psi_{\infty}(u) = \sum_{n=1}^{\infty} \psi_{\infty}^{(n)} \frac{u^n}{n!} = \sum_{n=1}^{\infty} \left\{ \sum_{k=1}^{n} \mathcal{Y}_n [n, k] \, \omega^{(k)}(0) \right\} \frac{u^n}{n!}.$$

Unlike the unconditional expectation of the conditional cumulant generating function $E(\psi_t(u))$, which has been characterized analytically and shown to belong to the gamma distribution family (see equation (28)), we have not been able to compute $\psi_{\infty}(u)$ in closed-form and thus are unable to assess whether the unconditional distribution belongs to a known family of distribution. It is important to stress again that $\psi_{\infty}(u) \neq E(\psi_t(u))$. Indeed,

$$\psi_{\infty}(u) \equiv \ln(E[\exp(u x_{t+1})]) = \ln(E[\exp(\psi_t(u))]) \neq E[\psi_t(u)].$$

Finally, the weak ergodicity implies the existence of an invariant distribution whose cumulant generating function is $\psi_{\infty}(u)$, such that if the process is initialized from its invariant distribution, it is stationary.

4.2 Autocorrelation functions

Various methods exist for examining serial dependence in stationary GARG processes. In this section, we consider the first- and second-order autocorrelograms. Our goal is to show how the flexibility (throughout parameters β and θ) of GARG impacts the serial dependence of the series of interest x_t.

4.2.1 Autocorrelation of the level

Since x_t is an ARMA(1,1) (see discussions at the end of section 3.1) and if $\rho < 1$ and $\beta\theta^2 < 1$ (which is equivalent to the covariance-stationarity of x_t), we have

$$Corr\left(x_t, x_{t+h}\right) = \rho^{h-1} Corr(x_t, x_{t+1}) \text{ if } h \geq 1,$$

with

$$Corr\left(x_t, x_{t+1}\right) = \alpha'(0)\left[\frac{1-(\beta\theta)^2 - \alpha'(0)\beta\theta}{1-(\beta\theta)^2 - 2\alpha'(0)\beta\theta}\right].$$

4.2.2 Autocorrelation of the squared values of the process

The second-order autocorrelogram represents the serial dependence in squared values of the process. Let us denote $z_t \equiv \left(x_t, \varepsilon_t^2, x_t^2\right)^\top$, where $\varepsilon_t \equiv x_t - \psi'_{t-1}(0)$. z_t is a VARMA(1,1) since its conditional expectation

$$M_{z,t-1} = \left(\psi'_{t-1}(0), \psi''_{t-1}(0), \psi''_{t-1}(0) + \psi'_{t-1}(0)^2\right)^\top$$

is a VAR(1). Indeed, we have

$$E_{t-1}\left[M_{z,t}\right] = K_z + \Phi_z M_{z,t-1},$$

where

$$K_z \equiv \begin{pmatrix} \omega'(0) \\ \omega''(0) \\ \omega''(0) + \omega'(0)^2 \end{pmatrix}, \quad \Phi_z \equiv \begin{bmatrix} \rho & 0 & 0 \\ \alpha''(0) & \beta\theta^2 & 0 \\ \varphi_1 & \varphi_2 & \rho^2 \end{bmatrix}$$

$$\varphi_1 = \alpha''(0) + 2\omega'(0)\rho, \quad \varphi_2 = \beta\theta^2 + \alpha'(0)^2 - \rho^2.$$

Hence, using the conditions $\rho < 1$ and $\beta\theta^n < 1$ for $n \leq 4$, we have

$$cov(z_{t+h}, z_t) = \Phi_z^{h-1} cov(z_{t+1}, z_t).$$

We provide closed-form expressions for $cov(z_{t+1}, z_t)$ and $Var\left[z_t\right]$ in section 6.2 of the Appendix.

4.3 How β and θ impact unconditional moments

We consider different values for β and θ and fix $\nu = 0.039$, $\varphi = 0.017$, and $\phi = 0.252$. These values are the results of the estimation on realized variance data (see section 7.1). Figure 1 plots unconditional moments as functions of β and θ and reveals interesting insights. First, unsurprisingly low values for β imply a model very close to the original ARG. When β is low, the model is broadly similar for different values of θ. Similarly, when θ is low the model is broadly similar for different values of β. This is not surprising since θ is only identified if β is sufficiently different from zero and vice versa. While the volatility increases with β, the skewness and kurtosis decrease with β. The same findings apply when looking at variation across θ. The skewness and kurtosis decrease with θ, but the volatility increases with θ. In conclusion, adding $\theta < 1$ improves the ability of the new model to fit highly skewed and fat-tailed time series.

Turning to the correlation between cumulants given by equation (25) and plotted in the last row of Figure 1, we observe that the correlations between cumulants decrease with β, with a perfect correlation for low values of β. For very high values of β, the correlation between cumulants decreases, with θ reaching values as low as 0.8. There is a U-shaped pattern as a function of θ for the medium value of β. In conclusion, adding parameters β and θ breaks the tight link between cumulants that is embedded in ARG dynamics and affine models in general. Finally, in the ARG dynamic the autocorrelogram of the level is very similar to that of the squared values of the process, as shown by the first plot in the second row of Figure 1. We can see how increasing θ and thus departing progressively from the affine structure enables the disentanglement of these autocorrelograms.

5 Comparison with multi-factor affine models

Our discussion so far has focused on a simple extension of a single-factor ARG process (and affine processes in general). Our proposal is similar to the generalization of the AR dynamic to ARMA or ARCH to GARCH. Our approach is not the only way to enhance ARG and disentangle its moments dynamics. An alternative is to model x_t as the sum of independent univariate ARG dynamics, which we refer to as the multi-factor ARG, or MARG for short. The MARG pioneered by Le et al. (2010) is the leading framework in the discrete-time affine term structure of interest rate models. Its appeal is simple and intuitive: if one is interested in disentangling the first K moments, it is enough to sum K independent ARGs. However, as we discuss in detail in this section, the MARG loses its analytical tractability when we analyse the dynamic conditional only on the variable of interest x_t. In option pricing, the Wishart autoregressive (WAR) process studied in detail in Gourieroux et al. (2009), Yu et al. (2017) and Gourieroux and Sufana (2010) is the leading extension of the ARG dynamic. The WAR has the same shortcomings as the MARG.

The MARG resembles our model specification given in equation (9), where the multiple latent factors are assumed to be independent both conditionally and unconditionally. Formally, the MARG is given by

$$x_{t+1} = \sum_{k=1}^{K} a_k x_{t+1,k}, \tag{35}$$

where $a_k > 0$; $k = 1, \cdots, K$ and $x_{t+1,k}$; $k = 1, \cdots, K$ are independent ARG processes. A MARG may seem less constrained than our GARG model given in equation (9). First, each latent factor $x_{t+1,k}$ has its own set of parameters while the parameters for the factors ($Z_{t+1}^{(j)}$) in the GARG model given in equation (9) are all related, as shown in equation (12).

Second, while the $x_{t+1,k}$, $k = 1, \cdots, K$ in equation (35) are independent, their counterparts (the $Z_{t+1}^{(j)}$) in equation (9) are *conditionally* independent but are dependent unconditionally. In fact, the conditional distribution of $Z_{t+1}^{(j)}$ depends only on x_{t-j} and not on $Z_t^{(j)}$.

In other words, although in the GARG dynamic, the $Z_{t+1}^{(j)}$ are latent, their distributions depend only on the observed process of interest x_t. This property enables us to compute analytically the distribution of x_{t+1} conditional on its own past, without resorting to any filtering procedure. This is not the case of the MARG, which requires a filtering procedure.

In fact, it is fair to compare the two models when we derive the distribution of x_{t+1} conditional on its own past implied by the MARG dynamic.

5.1 Cumulant generating of x_{t+1} conditional on $(x_s, s \leq t)$.

To avoid cumbersome mathematical derivations, we focus on the case K=2. We can also, without loss of generality, assume that $a_k = 1$, as the a_k are not separately identifiable for the parameters of the latent process $x_{t+1,k}$. In fact, $a_k x_{t+1,k}$ is also an ARG process. Hence, we have

$$\begin{aligned} x_t &= x_{1t} + x_{2t} \\ E_t\left[\exp\left(u x_{jt+1}\right)\right] &= \exp\left(\omega_j\left(u\right) + \alpha_j\left(u\right) x_{jt}\right) \end{aligned}$$

$$\omega_j(u) = -\nu_j \log(1 - u\varphi_j), \text{ and } \alpha_j(u) = \frac{\phi_j u}{1 - u\varphi_j}.$$

Similar to equation (1), $\psi_t(u)$ denotes the one-step ahead cumulant generating function of x_{t+1} conditional on its own past. Formally, we have $\psi_t(u) \equiv \ln\left[E\left[\exp\left(u x_{t+1}\right)|I_t\right]\right]$, where I_t is the sigma algebra generated by $(x_s, s \leq t)$. We show the following result in section 8 of the Appendix:

Proposition 3 *The dynamic of $\psi_t(u)$ implied by the MARG dynamic is given by*

$$\psi_{t+1}\left(u\right) = \omega_1\left(u\right) + \omega_2\left(u\right) + \alpha_2\left(u\right) x_{t+1} + \ln\left[\frac{\int \exp\left(\tilde{\omega}\left(iy,u\right) + \tilde{\alpha}\left(iy,u\right) x_t + \psi_t\left(\tilde{\theta}\left(iy,u\right)\right) - iyx_{t+1}\right) dy}{\int \exp\left(\psi_t\left(iy\right) - iyx_{t+1}\right) dy}\right] \tag{36}$$

where

$$\tilde{\theta}\left(y,u\right) \equiv \alpha_0^{-1}\left(\alpha_1\left(y + \alpha_0\left(u\right)\right) - \alpha_2\left(y\right)\right)$$

$$\tilde{\omega}\left(y,u\right) \equiv \omega_2\left(y\right) - \omega_2\left(\tilde{\theta}\left(y,u\right)\right) + \omega_1\left(y + \alpha_0\left(u\right)\right) - \omega_1\left(\tilde{\theta}\left(y,u\right)\right)$$

$$\tilde{\alpha}(y,u) \equiv \alpha_2(y) - \alpha_2\left(\tilde{\theta}(y,u)\right), \quad \alpha_0(u) \equiv \alpha_1(u) - \alpha_2(u).$$

The two dynamics to compare are given by equation (36) for the MARG model and equation (8) for the GARG model. While both models express the conditional cumulant generating function recursively, the recursion for the GARG is simple as it has a closed-form expression while the one for the MARG is non-linear and requires computing tedious integrals.

5.2 Mean and variance of x_{t+1} conditional on $(x_s, s \leq t)$

We push the comparison one step further by evaluating the first two moments implied by equation (36). This is done by taking the first two derivatives of equation (36) with respect to u. To ease the mathematical derivations, we focus on the case $\varphi_1 = \varphi_2 = \varphi$, which implies that both the MARG and the GARG have the same number of parameters. $\varphi_1 = \varphi_2 = \varphi$ implies that the MARG dynamic collapses to the ARG dynamic if and only if $\phi_1 = \phi_2$.

The following proposition, proven in sections 8.1 and 8.2 of the Appendix, gives the dynamic of the first two moments:

Proposition 4 *The dynamic of $\psi_t'(0)$ implied by the MARG dynamic is given by*

$$\psi_{t+1}'(0) = (\nu_1+\nu_2)\varphi + \frac{(\phi_1+\phi_2)}{2}x_{t+1} + \frac{(\phi_1+\phi_2)}{2}\frac{\int\left(\psi_t'(iy) - \tilde{\psi}_t'(iy)\right)e^{\psi_t(iy)-iyx_{t+1}}dy}{\int e^{\psi_t(iy)-iyx_{t+1}}dy}. \tag{37}$$

The dynamic of $\psi_t''(0)$ implied by the MARG dynamic is given by

$$\begin{aligned}\psi_{t+1}''(0) \;=\;& (\nu_1+\nu_2)\varphi^2 + 2\phi_2\varphi x_{t+1} - \left(\psi_{t+1}'(0) - (\nu_1+\nu_2)\varphi - \phi_2 x_{t+1}\right)^2 \\ &+\frac{\int\left[\phi_1\bar{\psi}_t''(iy) + \left(\bar{\psi}_t'(iy)\right)^2 + 2\varphi\bar{\psi}_t'(iy) - 2\varphi\phi_2\frac{\bar{\psi}_t'(iy)}{1-iy\varphi} + \frac{\phi_2\varphi^2(\phi_1-\phi_2)\nu_1}{(1-iy\varphi)^2}\right]e^{\psi_t(iy)-iyx_{t+1}}dy}{\int e^{\psi_t(iy)-iyx_{t+1}}dy}\end{aligned} \tag{38}$$

where

$$\tilde{\psi}_t(u) = -\nu\ln(1-u\varphi) + \frac{\phi u x_t}{1-u\varphi},\ \ \nu \equiv \frac{2(\nu_2\phi_1+\nu_1\phi_2)}{\phi_1+\phi_2},\ \ \phi \equiv \frac{2\phi_1\phi_2}{\phi_1+\phi_2},$$

and

$$\bar{\psi}_t(u) \equiv \phi_1\psi_t(u) - \frac{(\phi_1+\phi_2)}{2}\tilde{\psi}_t(u).$$

Equations (37) and (38) contain integrals that complicate our analysis. To solve these integrals and gather more intuition, we resort to approximations by assuming that

$$\psi_t(iy) \approx \left(\psi_t'(0) - \tilde{\psi}_t'(0)\right) iy + \left(\psi_t''(0) - \tilde{\psi}_t''(0)\right) \frac{(iy)^2}{2} + \tilde{\psi}_t(iy). \tag{39}$$

Combining equations (37) and (38) with the approximation given in equation (39) leads to the following corollary:

Corollary 2 *The dynamics of $\psi_t'(0)$ and $\psi_t''(0)$ implied by the MARG dynamic are given by*

$$\psi_{t+1}'(0) = (\nu_1+\nu_2)\varphi + \frac{(\phi_1+\phi_2)}{2}\left[\left(2\psi_t'(0) - \tilde{\psi}_t'(0)\right) + \left(\frac{2\psi_t''(0) - \tilde{\psi}_t''(0)}{\psi_t''(0)}\right)(x_{t+1} - \psi_t'(0))\right] \tag{40}$$

and

$$\psi_{t+1}''(0) = (\phi_1\phi_2 - 1)\varphi^2(\nu_1+\nu_2) + 2\varphi\left(\psi_{t+1}'(0) - \phi_1\phi_2\psi_t'(0)\right) + \left(\frac{\phi_1+\phi_2}{2}\right)^2\left(\frac{2\psi_t''(0) - \tilde{\psi}_t''(0)}{\psi_t''(0)}\right)\tilde{\psi}_t''(0). \tag{41}$$

We verify numerically that these approximations are accurate. In fact, we show that the same dynamic is obtained when Kalman filters methods are used (see Monfort et al., 2017). First, the two moments dynamics are interrelated since to compute the time $t+1$ expectation ($\psi_{t+1}'(0)$), we need past variance ($\psi_t''(0)$) and vice versa. This contrasts with the GARG dynamic where each cumulant dynamic is computed independently. This connection between moments dynamics is at the heart of our problematic. Second, the dynamic here is non-linear, which complicates the temporal aggregation. This is in contrast to the linear dynamic of the GARG. It is important to stress that the model is affine when we condition on the unknown unobserved component but becomes non-affine if we condition on the observed variable x_t. This paper is about the dynamic of x_t conditional on the past of x_t. However, we nonetheless investigate the MARG empirically in section 7.

6 Statistical Inference

The univariate GARG has five parameters, $\phi, \varphi, \nu, \beta, \theta$; the goal of this section is to discuss their estimation and statistical inference. First, the five parameters are well identified. In section 6.1 of the Appendix, we discuss an identification approach that consists of expressing $\phi, \varphi, \nu, \beta$ and θ as functions of quantities that can be directly estimated: the unconditional mean, variance, skewness and the first two autocorrelations. We have also run simulation exercises to establish heuristically the identification of the five parameters. The results are displayed in Tables 1 and 2 of the Appendix. In this section we review standard estimation methods such as the pseudo-maximum likelihood and the maximum likelihood and show that they can be used for the estimation of GARG processes. This implies that standard statistical inferences related to these methods remain valid in the context of the GARG dynamic.

6.1 Maximum likelihood estimators

We exploit the Fourier inversion formula to compute the conditional density as follows:

$$f_t\left(x_{t+1}\right) = \frac{1}{\pi}\int_0^{\infty} Re\left[e^{-iux_{t+1}+\psi_t(iu)}\right] du, \tag{42}$$

where i stands for the imaginary unit. This enables the estimation of the GARG's parameters through the maximum likelihood (ML) procedure and the use of standard inference to compute standard errors. Because a numerical inversion is involved, some practical challenges could arise.

6.2 Pseudo-maximum likelihood estimators

Since conditional moments are available in closed-form, the GARG can be estimated using pseudo-maximum likelihood. The order-2 pseudo-maximum likelihood estimators are the solutions of:

$$(\hat{\phi}, \hat{\nu}, \hat{\varphi}, \hat{\beta}, \hat{\theta})' = \arg \max_{\phi,\nu,\varphi,\beta,\theta} \sum_{t=1}^{T} \left\{ -\frac{1}{2} \log\left(\psi''_{t-1}(0)\right) - \frac{1}{2} \frac{\left(x_t - \psi'_{t-1}(0)\right)^2}{\psi''_{t-1}(0)} \right\}, \tag{43}$$

where $\psi''_{t-1}(0)$ and $\psi'_{t-1}(0)$ are computed recursively using equation (19).

Because GARG processes have a closed-form conditional characteristic function, one alternative to the ML-based method is the empirical characteristic function (ECF) estimation method. Finally, some applications in the stochastic volatility and term structure of interest rates literatures require latent factors. In sections 6.2, 6.3 and 6.4 of the Appendix, we discuss the ECF estimator, the generalized method of moments and the generalized method of moments.

7 Empirical analysis: Option pricing model

7.1 Fitting the historical joint dynamic of S&P 500 returns and realized variances

In this section we denote the day t stock price and return by S_t and R_t, with $R_t \equiv \ln(S_t/S_{t-1})$. We design an option pricing model where returns and realized variances (RV_t) are modeled jointly in line with the literature (Majewski et al., 2015; Christoffersen et al., 2014). Our measure of realized variances, RV_t on day t, is the sum of the squared 5-min log-returns observed within day t. To highlight the usefulness of the GARG process, we assume that the realized variance follows a GARG dynamic (instead of the ARG dynamic), that is,

$$R_{t+1} = \ln(S_{t+1}/S_t) = r + \left(\lambda - \frac{1}{2}\right) RV_{t+1} + \sqrt{RV_{t+1}}\varepsilon_{t+1} \tag{44}$$

$$\varepsilon_{t+1} \sim i.i.d N(0,1)$$

$$RV_{t+1} \sim GARG(\phi, \varphi, \nu, \beta, \theta), \tag{45}$$

where r is the risk-free rate (calibrated to the sample average of the 3-month Treasury Bill rate) and λ is interpreted as the price of risk as it indicates the variation in the equity risk-premium per unit of variation in the realized variance. Equation (44) is well motivated empirically by several studies in the literature (Andersen et al., 2001; Andersen et al., 2007), where it is shown that time t return conditional on time t realized variance follows a Gaussian distribution. The remaining challenge is to model the conditional distribution of the realized variance. Many studies have relied on the ARG process; we will instead assume a GARG dynamic, as in equation (45).

7.1.1 Benchmark models

Our benchmark models are variants of the ARG(p,q) model defined in equations (3) and (4), that is,

$$\psi_t(u) \equiv \ln \left[E \left[\exp \left(uRV_{t+1} \right) | I_t \right] \right] = \omega(u) + \alpha(u) m_t, \tag{46}$$

where

1. ARG0: $m_t = RV_t$
2. ARG1: $m_t = RV_t + \theta_1 m_{t-1}$
3. ARG2: $m_t = RV_t + \theta_1 m_{t-1} + \theta_2 m_{t-2}$.
4. MARG: We also add the MARG model discussed in section 5, formally

$$\begin{aligned} RV_t &= x_{1,t} + x_{2,t} \\ x_{j,t} &\sim ARG(\nu_j, \varphi, \phi_j), \text{ with } j = 1, 2. \end{aligned}$$

Note that both the ARG2 and the MARG models have the same number of parameters (5) as the GARG. Before giving details on option pricing, we evaluate the relative performance of GARG processes in fitting salient facts of the dynamic of the observed realized variance series.

7.1.2 Data and empirical results

The empirical investigation begins by obtaining daily historical realized variances for the S&P 500 index from oxford-man.ox.ac.uk. The data cover the period from January 01, 2000, to December 31, 2017. Table 1 contains the maximum likelihood estimation for the benchmark ARG processes and the GARG process on daily historical realized variances. The likelihood and BIC figures indicate that the GARG is the best performing model. Unsurprisingly, the likelihood ratio tests favor all the alternative specifications against the basic ARG model (ARG0). To better gauge the ability of the different models to fit the data, we report the observed sample mean, variance, skewness and kurtosis and compare them with each model's implied moments. Contrary to ARG models, the GARG is able to match these unconditional moments.

To further shed light on these results, we plot sample autocorrelations and cross-correlations in Figure 1 of the Appendix. The top left panel displays the realized variance's autocorrelation function across models. The other panels display the cross-correlations between the level and the squared, $corr(RV_t, RV_{t+h}^2)$, the cross-correlation between the squared and the level, $corr(RV_t^2, RV_{t+h})$, and the autocorrelation of the squared $corr(RV_t^2, RV_{t+h}^2)$. We can see that none of the ARG models are able to capture the long memory inherent in the observed variance dynamic. In contrast, the GARG dynamic better matches the sample autocorrelograms.

To diagnose the different models, we extract the conditional mean $E_{t-1}[x_t]$ and the conditional variance $Var_{t-1}[x_t]$, then compute the standardized residuals as

$$z_t \equiv \frac{x_t - E_{t-1}[x_t]}{\sqrt{Var_{t-1}[x_t]}}.$$

In principle, the better a model is at fitting the conditional mean (the first conditional cumulant), the smaller the autocorrelation in z_t. In the same vein, the better a model is at fitting the conditional variance (the second conditional cumulant), the smaller the autocorrelation in z_t^2. Figure 2 displays the autocorrelograms of the level and square of the standardized residuals z_t along with 95% confidence bounds and depicts interesting insights. Only the GARG model is able to extract the first two moments dynamics with great accuracy as its autocorrelograms lie mostly within the confidence bounds. The basic ARG model, ARG0, is unable to fit both moments, especially the conditional mean. The other two versions of the ARG processes, ARG1 and ARG2, are able to fit the second moment with the same accuracy as the GARG dynamic but at the cost of the first moment fitting. This highlights the central point of our proposal: there is a tension between fitting the first two moments that are inherent in the ARG dynamic, which we are able to overcome with the GARG process. The MARG model provides a clear improvement over ARG dynamics regarding the first moment, but it is still outperformed by the GARG dynamic on both dimensions.

The physical properties of the realized variance dynamic we have investigated above are likely to have important implications for the models' ability to fit a large panel of options. This is the task we now turn to.

7.2 Option pricing

7.2.1 Risk-neutral estimation

Similar to the ARG processes, the GARG processes are built to enable closed-form option prices. In this exercise, we assume that the joint dynamic given by equations (44) and (45) is under the risk-neutral probability measure. This implies that r is the risk-free rate and $\lambda = 0$. We provide full details on option pricing under the GARG and MARG dynamics in sections 7.2 and 8.2 of the Appendix. We estimate the different models by optimizing their fit on option data. This analysis aims at exploring the ability of each specification to properly match the risk-neutral distribution embedded in option contracts. We start by presenting the key features of the option data panel used in our empirical analysis and then study the performance of the various models relying on the implied volatility root-mean-squared-error.

7.2.2 Option data

We use European-style options written on the S&P 500 index. The observations span the period January 10, 1996, to August 28, 2013. This data set is available through OptionMetrics, which supplies data for the U.S. option markets. In line with the literature, we only include out-of-the-money (OTM) options with maturities ranging from 15 to 180 days. This selection procedure is intended to guarantee that the contracts considered are liquid. We also filter out options that violate basic no-arbitrage criteria. For each maturity quoted on Wednesdays, we select only the six most liquid strike prices, which amounts to a data set of 21,283 option contracts. To ease calculation and interpretation, OTM put prices are converted into corresponding in-the-money call values, by exploiting the call-put parity relationship. We provide a detailed description of option data in section 7.3 of the Appendix.

7.2.3 Fitting options

We explore the performance of the different models by relying on the implied volatility root-mean-squared error (IVRMSE) (see Appendix 7.3 for details). Table 2 contains the results of the option-based estimation. Clearly, our option-fitting strategy yields accurate parameter estimates, as evidenced by fairly small standard errors and sizeable model likelihoods. Because we are fitting the model only on options, the estimates correspond to risk-neutral parameters. The proposed GARG model clearly outperforms the alternative ARG specifications, as it delivers the highest likelihood value, the lowest BIC and the smallest global IVRMSE. Specifically, the GARG model offers about 10% and 20% improvement over the benchmark ARG0 model in terms of log-likelihood and IVRMSE, respectively. While the MARG model outperforms the ARG dynamics, it has a marginally lower IVRMSE than the GARG. At the bottom of the table, we report P-values of the Diebold-Mariano Test (see Diebold and Mariano, 2002) to assess whether differences in pricing errors across models are statistically significant. We test ARG1 against ARG0, ARG2 against ARG1 and finally, GARG against ARG2. The results indicate P-values of 0 except for the test of ARG2 against ARG1, implying that ARG0 errors are statistically significantly higher than ARG1, while GARG errors are statistically lower than ARG2. There is no statistical difference between ARG1 and ARG2 option pricing errors.

7.2.4 Model fit by moneyness, maturity and VIX levels

We now scrutinize the overall performance results reported in the bottom panel of Table 2. To this end, we report the IVRMSE by moneyness, maturity and VIX levels in Table 3. We see that all models offer a satisfactory performance (low IVRMSEs) in matching at-the-money options contracts. By contrast, fitting deep OTM call and put options seems more challenging. Interestingly, the ability of the various specifications to match the observed option-implied volatility appears consistent across the term structure of the options, as the IVRMSEs are of comparable magnitude. Moreover, the performance of these models tends to deteriorate nearly monotonically as a function of the VIX level. This observation suggests that the ability of the models to generate realistic option prices weakens in highly volatile times. Nevertheless, the GARG model dominates the other models along the moneyness, maturity, and VIX level dimensions. However, we would have expected the fit to be significantly better for long-maturities since the GARG process is able to generate slowly decaying autocorrelations. The fit in this longer maturity dimension would certainly improve using a higher order GARG dynamic, which we introduce in section 1.2 of the Appendix.

8 Conclusion

This study introduces the generalized autoregressive gamma (GARG) dynamic. The GARG is a parsimonious generalization of ARG dynamics able to overcome tight links between conditional moments that are implicit within ARG processes. The GARG dynamic enables each moment to be driven by its specific moving average of the variable of interest. Besides, the new process maintains the practical advantage of ARG dynamics and affine models in general: it has a closed-form multi-step ahead moment generating function. Empirically, we show that the GARG dynamic dominates ARG in fitting the historical dynamic of realized variance, and most importantly in describing the behavior of a large panel of option prices. Our generalization so far has focused on models with finite moments at all orders, which clearly restricts applications to a certain type of financial data. Generalization to processes with infinite moments is an exciting topic for future research.

References

Al-Osh, M. A., and Alzaid, A. A. (1987), "First-order integer-valued autoregressive (inar(1)) process," *Journal of Time Series Analysis,* 8(3), 261–275.

Andersen, Torben G., and Benzoni, Luca (2010), "Do bonds span volatility risk in the U.S. Treasury market? A specification test for affine term structure models," *The Journal of Finance,* 65(2), 603–653.

Andersen, Torben G., Bollerslev, Tim, and Dobrev, Dobrislav (2007), "No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications," *Journal of Econometrics,* 138(1), 125–180.

Andersen, Torben G., Bollerslev, Tim, Diebold, Francis X., and Ebens, Heiko (2001), "The distribution of realized stock return volatility," *Journal of Financial Economics,* 61(1), 43–76.

Chang, Bo-Young, Christoffersen, Peter, Jacobs, Kris, and Vainberg, Gregory (2011), "Option-implied measures of equity risk," *Review of Finance,* 16(2), 385–428.

Christoffersen, Peter, Feunou, Bruno, Jacobs, Kris, and Meddahi, Nour (2014), "The economic value of realized volatility: Using high-frequency returns for option valuation," *Journal of Financial and Quantitative Analysis,* 49(3), 663–697.

Christoffersen, Peter, Heston, Steve, and Jacobs, Kris (2006), "Option valuation with conditional skewness," *Journal of Econometrics,* 131(1), 253–284.

Cieslak, Anna, and Povala, Pavol (2016), "Information in the term structure of yield curve volatility," *The Journal of Finance,* 71(3), 1393–1436.

Collin-Dufresne, Pierre, Jones, Christopher S., and Goldstein, Robert S. (2004), "Can interest rate volatility be extracted from the cross section of bond yields? An investigation of unspanned stochastic volatility," Working Paper 10756, National Bureau of Economic Research, September.

Corsi, Fulvio, Mittnik, Stefan, Pigorsch, Christian, and Pigorsch, Uta (2008), "The volatility of realized volatility," *Econometric Reviews,* 27(1-3), 46–78.

Cox, John C., Ingersoll, Jonathan E., and Ross, Stephen A. (1985), "A theory of the term structure of interest rates," *Econometrica,* 53(2), 385–407.

Darolles, Serge, Gourieroux, Christian, and Jasiak, Joann (2006), "Structural laplace transform and compound autoregressive models," *Journal of Time Series Analysis,* 27(4), 477–503.

Diebold, Francis X., and Mariano, Robert S. (2002), "Comparing predictive accuracy," *Journal of Business & Economic Statistics,* 20(1), 134–144.

Feunou, B., and Meddahi, N. (2009), "Generalized affine models," Working Paper, Available at SSRN.

Feunou, Bruno, and Tedongap, Romeo (2012), "A stochastic volatility model with conditional skewness," *Journal of Business & Economic Statistics,* 30(4), 576–591.

Gourieroux, Christian, and Jasiak, Joann (2006), "Autoregressive gamma processes," *Journal of Forecasting,* 25(2), 129–152.

Gourieroux, Christian, and Sufana, Razvan (2010), "Derivative pricing with Wishart multivariate stochastic volatility," *Journal of Business & Economic Statistics,* 28(3), 438–451.

Gourieroux, Christian, Jasiak, Joann, and Sufana, R. (2009), "The Wishart autoregressive process of multivariate stochastic volatility," *Journal of Econometrics,* 150(2), 167–181. Recent Development in Financial Econometrics.

Gourieroux, Christian, Jasiak, Joanna, and Fol, Gaelle Le (1999), "Intra-day market activity," *Journal of Financial Markets,* 2(3), 193–226.

Hansen, Bruce E. (1994), "Autoregressive conditional density estimation," *International Economic Review,* 35(3), 705–730.

Heston, Steven L., and Nandi, Saikat (2000), "A closed-form GARCH option valuation model," *The Review of Financial Studies,* 13(3), 585–625.

Jondeau, Eric, and Rockinger, Michael (2003), "Conditional volatility, skewness, and kurtosis: Existence, persistence, and comovements," *Journal of Economic Dynamics and Control,* 27(10), 1699–1737.

Le, Anh, Singleton, Kenneth J., and Dai, Qiang (2010), "Discrete-time affineQ term structure models with generalized market prices of risk," *The Review of Financial Studies,* 23(5), 2184–2227.

Majewski, Adam A., Bormetti, Giacomo, and Corsi, Fulvio (2015), "Smile from the past: A general option pricing framework with multiple volatility and leverage components," *Journal of Econometrics,* 187(2), 521–531.

Monfort, Alain, Pegoraro, Fulvio, Renne, Jean-Paul, and Roussellet, Guillaume (2017), "Staying at zero with affine processes: An application to term structure modelling," *Journal of Econometrics,* 201(2), 348–366.

Yu, Philip L. H., Li, W. K., and Ng, F. C. (2017), "The generalized conditional autoregressive Wishart model for multivariate realized volatility," *Journal of Business & Economic Statistics,* 35(4), 513–527.

Figure 1: Unconditional moments

These figures represent unconditional moments as functions of β and θ. We set $\nu = 0.039$, $\varphi = 0.017$, $\phi = 0.252$.

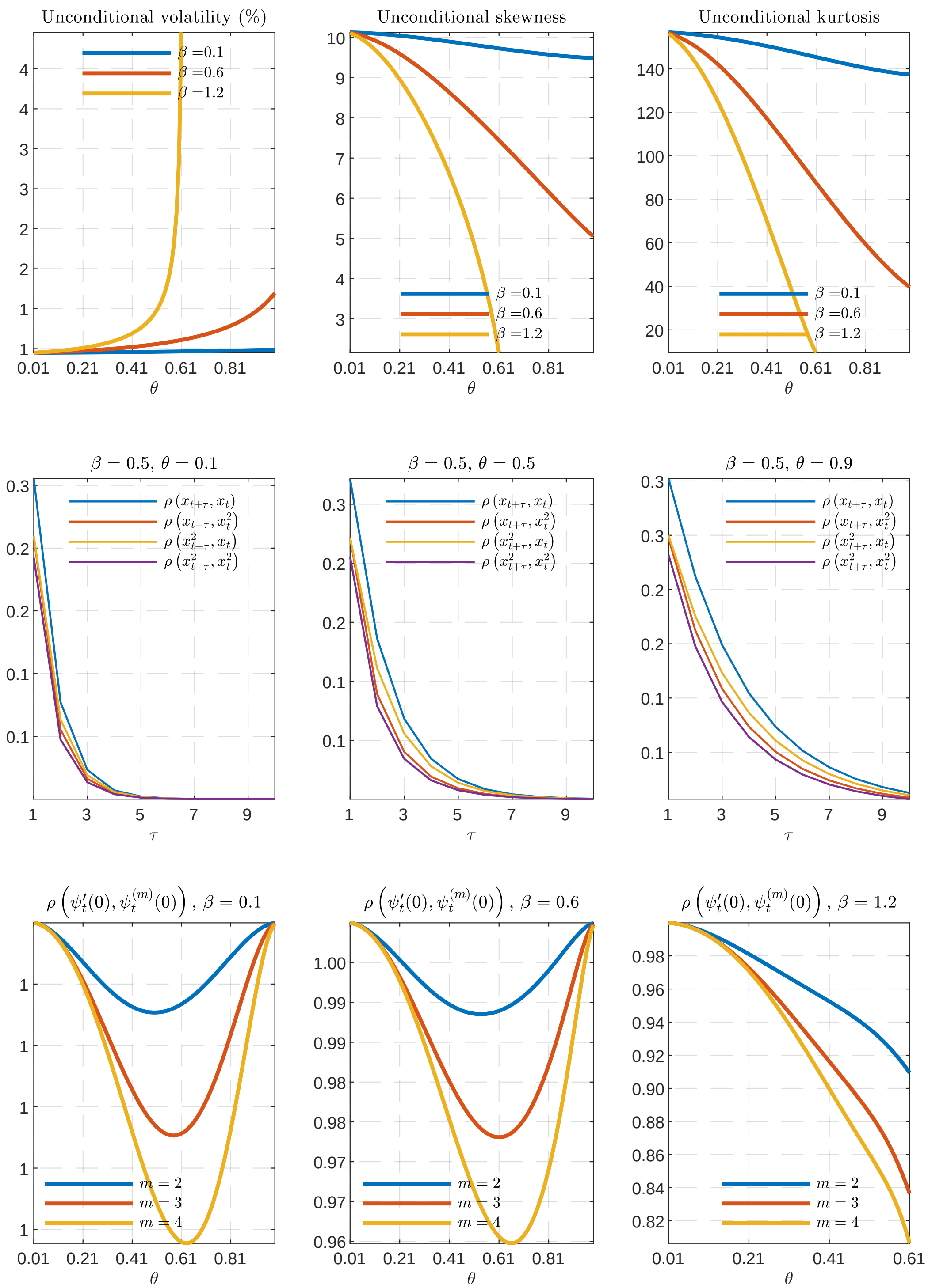
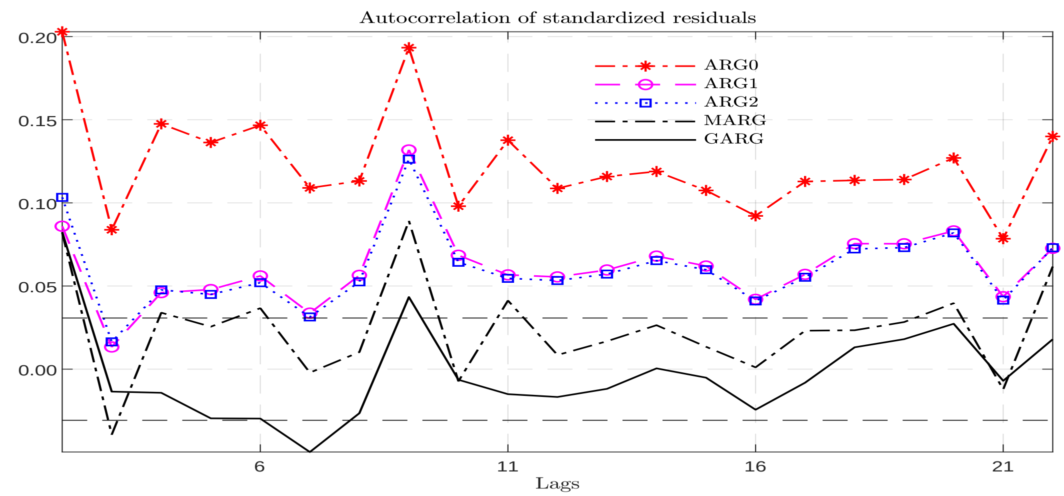


Figure 2: Autocorrelograms of residuals

These figures plot the autocorrelation of the level and the squared values of the standardized residual z_t. For each model, we extract the implied conditional mean $E_{t-1}[x_t]$ and the implied conditional variance $Var_{t-1}[x_t]$, then compute the standardized residuals as $z_t \equiv (x_t - E_{t-1}[x_t]) / \sqrt{Var_{t-1}[x_t]}$. The horizontal dashed lines represent the upper and lower confidence bounds. The sample begins from January 01, 2000, and ends December 31, 2017.



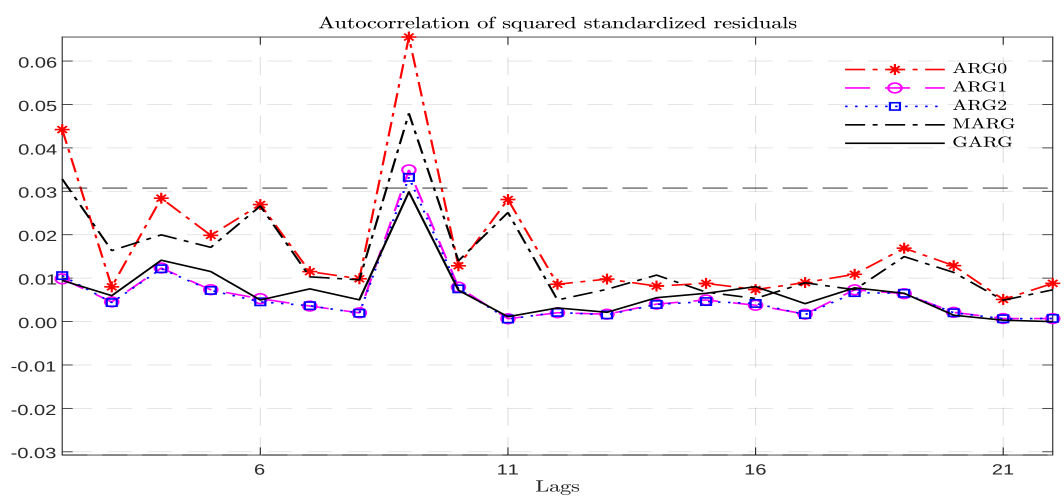


Table 1: Estimation using Historical Realized Variance

This table shows maximum likelihood estimation results for four different models. We used daily historical realized variances for the S&P 500 index from January 01, 2000, to December 31, 2017. We report the estimated parameters (Est) with their corresponding standard errors (SE).

		ARG Models						MARG Model		GARG Model	
		ARG0		ARG1		ARG2		MARG		GARG	
Parameters		Est	SE	Est	SE	Est	SE	Est	SE	Est	SE
ϕ		0.711	5.14E-03	0.362	6.53E-03	0.371	7.54E-03	0.954	1.53E-02	0.252	4.18E-04
φ		8.19E-03	5.93E-05	6.71E-03	4.79E-05	6.69E-03	4.83E-05	0.011	3.36E-05	0.017	6.97E-05
ν		1.017	0.034	0.978	0.042	0.975	0.042	0.013	2.06E-02	0.039	1.62E-04
θ_1		0		0.531	8.29E-03	0.440	0.027				
θ_2		0		0		0.081	0.021				
β										1.171	5.91E-03
θ										0.619	3.12E-03
ϕ_2								0.914	6.00E-03		
ν_2								0.198	1.88E-02		
Model Properties	Obs										
Avg	16.98	16.98		17.00		17.00		16.98		16.98	
Vol	18.27	16.91		15.44		15.37		18.27		18.27	
Skew	2.74	1.96		1.49		1.47		2.05		2.19	
Kurt	12.07	8.69		6.20		6.13		10.04		9.80	
AC(1)	0.67	0.71		0.57		0.57		0.93		0.67	
Log Likelihoods		13320		13577		13580		13597		14025	
BIC		-5.94		-6.05		-6.05		-6.07		-6.25	
LR P-Value, H0: ARG0				0.00		0.00		0.00		0.00	

Table 2: Estimation using Options

This table shows estimation results for six different models. We used Wednesday closing out-of-the-money (OTM) call and put contracts from OptionMetrics for the period beginning January 10, 1996, and ending August 28, 2013. We report the estimated parameters (Est) along with their corresponding standard errors (SE). The second-to-last row shows the implied volatility root-mean-squared errors (IVRMSEs). For comparison, the second-to-last row reports the IVRMSE ratio of each specification to the benchmark ARG0 model.

	ARG Models						MARG Model		GARG Model	
	ARG0		ARG1		ARG2		MARG		GARG	
Parameters	Est	SE	Est	SE	Est	SE	Est	SE	Est	SE
ϕ	0.938	7.65E-06	0.016	8.17E-05	0.016	8.28E-05	0.962	1.43E-04	0.020	5.15E-05
φ	2.90E-05	1.74E-07	9.50E-04	3.72E-05	9.49E-04	1.07E-05	1.88E-05	8.82E-07	1.75E-04	1.23E-05
ν	0.219	1.51E-05	0.032	1.20E-03	0.032	6.20E-05	0.145	9.43E-03	4.84E-03	3.07E-06
θ_1	0		0.974	1.41E-04	0.963	2.12E-04				
θ_2			0		0.011	7.18E-05				
β									1.079	1.13E-06
θ									0.897	4.66E-07
ϕ_2							0.547	1.17E-02		
ν_2							0.697	4.77E-02		
Model Properties										
Log Likelihoods	31554		37271		37273		37845		38236	
BIC	-2.96		-3.50		-3.50		-3.55		-3.59	
LR P-Value, H0: ARG0			0.00		0.00		0.00		0.00	
Avg. Model IV	20.54		20.74		20.74		20.78		20.80	
Variance Persistence										
	0.938		0.9905		0.9904		0.998		0.9801	
Option Errors										
IVRMSE	5.493		4.199		4.199		3.945		3.862	
Ratio to ARG0	1.000		0.764		0.764		0.720		0.703	
DM test P-Value			0.00		0.156		0.00		0.00	

Table 3: IVRMSE Option Error by Moneyness, Maturity

Panel A reports IVRMSE for contracts sorted by moneyness. Panel B reports IVRMSE for contracts sorted by days to maturity (DTM). The IVRMSE is expressed in percentage.

	OTM Call			OTM Put		
	Delta $<$ 0.3	0.3 $\leq$ Delta $<$ 0.4	0.4 $\leq$ Delta $<$ 0.5	0.5 $\leq$ Delta $<$ 0.6	0.6 $\leq$ Delta $<$ 0.7	Delta $\geq$ 0.7
Panel A: IVRMSE by Moneyness						
ARG0	6.284	4.721	4.739	5.465	5.737	5.318
ARG1	5.183	2.985	3.046	3.335	4.018	4.414
ARG2	5.184	2.985	3.045	3.335	4.017	4.413
MARG	4.032	3.078	2.924	3.676	4.202	5.727
GARG	3.985	2.757	2.910	3.034	3.375	4.490
	DTM $<$ 30	30 $\leq$ DTM $<$ 60	60 $\leq$ DTM $<$ 90	90 $\leq$ DTM $<$ 120	120 $\leq$ DTM $<$ 150	DTM $\geq$ 150
Panel B: IVRMSE by Maturity						
ARG0	5.845	5.205	5.525	5.673	5.532	5.525
ARG1	4.344	4.270	4.076	3.937	4.265	4.356
ARG2	4.344	4.270	4.075	3.937	4.264	4.355
MARG	4.208	4.094	3.943	3.528	3.886	3.724
GARG	3.736	3.953	3.710	3.677	4.293	3.947

Appendix for "Generalized Autoregressive Gamma Processes"

1 Extensions of the GARG

1.1 The multivariate case

In order to build the multivariate generalized autoregressive gamma process, we need two components, which are a standard version of the multivariate gamma distribution and a multivariate Poisson distribution. For the first component, we follow the approach in Carpenter and Diawara (2007). $Z = (Z_1, ..., Z_n)$ follows a standard multivariate gamma distribution of parameters $(k_0, k_1, ..., k_n)$ (with $0 \leq k_0 \leq \min_{1 \leq n} k_i$) denoted by $M\gamma(k_0, k_1, ..., k_n)$ and $Z_i = Y_0 + Y_i$, where $Y_0, Y_1, ..., Y_n$ are n independent random variables that follow univariate standard gamma distribution with respective parameters k_0,$k_1 - k_0$,...,$k_n - k_0$.

For the second component we assume that $U = (U_1, ..., U_n)$ follows a standard multivariate distribution of parameters $(\lambda_1, ..., \lambda_n)$ (with $0 \leq \min_{1 \leq n} \lambda_i$) denoted by $MP(\lambda_1, ..., \lambda_n)$ and $U_1, ..., U_n$, are n independent random variable univariate random variables that follow a standard univariate Poisson distribution with respective parameters $\lambda_1, ..., \lambda_n$

The multivariate GARG process is built through the following state space representation:

$$x_{t+1} = \bar{Z}_{t+1} + \mathbf{1}_{[t>0]} \left[\sum_{j=0}^{t-1} Z_{t+1}^{(j)} \right], \tag{1}$$

where $\bar{Z}_{t+1}$, $Z_{t+1}^{(j)}$ for $j = 0, ..., t-1$ are $t+1$ conditionally (conditional on I_t) independent random variables, $\bar{Z}_{t+1} \sim \beta^t \psi_0 (\theta^t u)$, and

$$\begin{aligned} \frac{Z_{t+1}^{(j)}}{\varphi^{(j)}} = \left(\frac{Z_{1,t+1}^{(j)}}{\varphi_1^{(j)}}, ..., \frac{Z_{n,t+1}^{(j)}}{\varphi_n^{(j)}} \right)' |U_{t+1}^{(j)}, I_t &\sim M\gamma \left(\nu_0^{(j)} + U_{0,t+1}^{(j)}, \nu_1^{(j)} + V_{1,t+1}^{(j)}, ..., \nu_{(n)}^{j} + V_{n,t+1}^{(j)} \right) \\ V_{i,t+1}^{(j)} &= U_{0,t+1}^{(j)} + U_{i,t+1}^{(j)}, i = 1, ..., n \\ U_{t+1}^{(j)} = \left(U_{0,t+1}^{(j)}, U_{1,t+1}^{(j)}, ..., U_{n,t+1}^{(j)} \right) |I_t &\sim MP \left(\phi_0^{(j)'} x_{t-j}, \frac{\phi_1^{(j)'} x_{t-j}}{\varphi_1^{(j)}}, ..., \frac{\phi_n^{(j)'} x_{t-j}}{\varphi_n^{(j)}} \right), \end{aligned}$$

where

$$\nu_i^{(j)} = \nu_j \beta^j, \quad \varphi_i^{(j)} = \varphi_i \theta_i^j, \quad \phi_i^{(j)} = \beta^j \theta_i^j \phi_i.$$

We show in the next section that

$$\psi_t(u) = \omega(u) + \alpha(u) x_t + \beta \psi_{t-1}(\theta u), \tag{2}$$

where

$$\omega(u) = -\nu_0 \ln \left(1 - \sum_{i=1}^{n} \varphi_i u_i \right) - \sum_{i=1}^{n} (\nu_i - \nu_0) \ln (1 - \varphi_i u_i),$$

and

$$\alpha(u) = \frac{\sum_{i=1}^{n} \varphi_i u_i}{1 - \sum_{i=1}^{n} \varphi_i u_i} \phi_0 + \sum_{i=1}^{n} \frac{u_i}{1 - u_i \varphi_i} \phi_i,$$

$\theta = diag(\theta_1, ..., \theta_n)$ and $0 \leq \nu_0 \leq \min (\nu_i)_{i=1}^n$.

Le et al. (2010) obtain a version of the multivariate autoregressive gamma process by imposing $\nu_0 = 0$, $\phi_0 = 0$ and $\beta = 0$ or $\theta = 0$. It is important to mention that in Le et al.'s setting, conditionally on information known at time t, the components of the vector x_{t+1} are independent. This is different from our set-up, where $\nu_0 \neq 0$ or $\phi_0 \neq 0$.

Our multivariate version does not boil down to a multivariate process where each entry is conditionally independent given the past. However, Granger causality can be obtained with the following Poisson intensities:

$$U_{t+1}^{(j)} \sim P \left(\frac{\Phi_j X_{t-j}}{\varphi_j} \right),$$

where the ratio is taken element-by-element, Φ_j is a matrix and X_t is a vector.

Cumulant generating function of multivariate GARG processes. The derivation of the cumulant generating function of multivariate GARG processes starts with equation 1:

$$x_{t+1} = \bar{Z}_{t+1} + \sum_{j=0}^{t-1} Z_{t+1}^{(j)},$$

with $\bar{Z}_{t+1}$, $Z_{t+1}^{(j)}$ for $j = 0, ..., t-1$, $t+1$ conditionally (conditional on I_t) independent random variables, $\bar{Z}_{t+1} \sim \beta^t \psi_0 \left(\theta^t u\right)$. Then

$$\begin{aligned} \psi_t(u) &= \ln E_t\left[\exp\left(u' x_{t+1}\right)\right] = \ln E_t\left[\exp\left(u'\left(\bar{Z}_{t+1} + \sum_{j=0}^{t-1} Z_{t+1}^{(j)}\right)\right)\right] \\ &= \beta^t \psi_0\left(\theta^t u\right) + \sum_{j=0}^{t-1} \ln E_t\left[\exp\left(u' Z_{t+1}^{(j)}\right)\right], \end{aligned}$$

since $\bar{Z}_{t+1}$, $Z_{t+1}^{(j)}$ for $j = 0, ..., t-1$ are conditionally independent.

$$\begin{aligned} E_t\left[\exp\left(u' Z_{t+1}^{(j)}\right)\right] &= E_t\left[\exp\left(\sum_{i=1}^{n} \varphi_i^{(j)} u_i \frac{Z_{i,t+1}^{(j)}}{\varphi_i^{(j)}}\right)\right] = E_t\left[\exp\left(\sum_{i=1}^{n} \varphi_i^{(j)} u_i \frac{Z_{i,t+1}^{(j)}}{\varphi_i^{(j)}}\right) \middle| U_{t+1}^{(j)}\right] \\ &= E_t\left[\exp\left(\begin{array}{c} -\left(\nu_0^{(j)} + U_{0,t+1}^{(j)}\right) \ln\left(1 - \sum_{i=1}^{n} \varphi_i^{(j)} u_i\right) \\ +\left(\nu_0^{(j)} + U_{0,t+1}^{(j)}\right) \sum_{i=1}^{n} \ln\left(1 - \varphi_i^{(j)} u_i\right) \\ -\sum_{i=1}^{n}\left(\nu_i^{(j)} + U_{0,t+1}^{(j)} + U_{i,t+1}^{(j)}\right) \ln\left(1 - \varphi_i^{(j)} u_i\right) \end{array}\right)\right], \end{aligned}$$

where

$$\nu_i^{(j)} = \nu_j \beta^j, \quad \varphi_i^{(j)} = \varphi_i \theta_i^j, \quad \phi_i^{(j)} = \beta^j \theta_i^j \phi_i. \tag{3}$$

Thus

$$\begin{aligned} E_t\left[\exp\left(u' Z_{t+1}^{(j)}\right)\right] &= \exp\left(-\nu_0^{(j)} \ln\left(1 - \sum_{i=1}^{n} \varphi_i^{(j)} u_i\right) + \nu_0^{(j)} \sum_{i=1}^{n} \ln\left(1 - \varphi_i^{(j)} u_i\right) - \sum_{i=1}^{n} \nu_i^{(j)} \ln\left(1 - \varphi_i^{(j)} u_i\right)\right) \\ &\quad \times E_t\left[\exp\left(\begin{array}{c} -U_{0,t+1}^{(j)} \ln\left(1 - \sum_{i=1}^{n} \varphi_i^{(j)} u_i\right) + U_{0,t+1}^{(j)} \sum_{i=1}^{n} \ln\left(1 - \varphi_i^{(j)} u_i\right) \\ -\sum_{i=1}^{n}\left(U_{0,t+1}^{(j)} + U_{i,t+1}^{(j)}\right) \ln\left(1 - \varphi_i^{(j)} u_i\right) \end{array}\right)\right], \end{aligned}$$

which implies that

$$\begin{aligned} E_t\left[\exp\left(u' Z_{t+1}^{(j)}\right)\right] &= \exp\left(\omega^{(j)}(u)\right) E_t\left[\exp\left(-U_{0,t+1}^{(j)} \ln\left(1 - \sum_{i=1}^{n} \varphi_i^{(j)} u_i\right) - \sum_{i=1}^{n} U_{i,t+1}^{(j)} \ln\left(1 - \varphi_i^{(j)} u_i\right)\right)\right] \\ &= \exp\left(\omega^{(j)}(u) + \left(\frac{1}{1 - \sum_{i=1}^{n} \varphi_i^{(j)} u_i} - 1\right) \phi_0^{(j)\prime} x_{t-j} + \sum_{i=1}^{n}\left(\frac{1}{1 - \varphi_i^{(j)} u_i} - 1\right) \frac{\phi_i^{(j)\prime} x_{t-j}}{\varphi_i}\right) \\ &= \exp\left(\omega^{(j)}(u) + \alpha^{(j)}(u)' x_{t-j}\right), \end{aligned}$$

where

$$\begin{aligned} \omega^{(j)}(u) &= -\nu_0^{(j)} \ln\left(1 - \sum_{i=1}^{n} \varphi_i^{(j)} u_i\right) - \sum_{i=1}^{n}\left(\nu_i^{(j)} - \nu_0^{(j)}\right) \ln\left(1 - \varphi_i^{(j)} u_i\right) \\ \alpha^{(j)}(u) &= \frac{\sum_{i=1}^{n} \varphi_i^{(j)} u_i}{1 - \sum_{i=1}^{n} \varphi_i^{(j)} u_i} \phi_0^{(j)} + \sum_{i=1}^{n} \frac{u_i}{1 - u_i \varphi_i^{(j)}} \phi_i^{(j)}. \end{aligned}$$

Replacing $\nu_i^{(j)}$, $\varphi_i^{(j)}$ and $\phi_i^{(j)}$ by their values given in equation (3), one obtains

$$\omega^{(j)}(u) = \beta^j \omega(\theta^j u), \;\; \alpha^{(j)}(u) = \beta^j \alpha(\theta^j u),$$

with

$$
\begin{aligned}
\omega(u) &= -\nu_0 \ln\left(1-\sum_{i=1}^{n}\varphi_i u_i\right) - \sum_{i=1}^{n}(\nu_i-\nu_0)\ln(1-\varphi_i u_i) \\
\alpha(u) &= \frac{\sum_{i=1}^{n}\varphi_i u_i}{1-\sum_{i=1}^{n}\varphi_i u_i}\phi_0 + \sum_{i=1}^{n}\frac{u_i}{1-u_i\varphi_i}\phi_i.
\end{aligned}
$$

Recall that

$$
\begin{aligned}
\psi_t(u) &= \beta^t\psi_0\left(\theta^t u\right) + \sum_{j=0}^{t-1}\ln E_t\left[\exp\left(u' Z_{t+1}^{(j)}\right)\right] \\
&= \beta^t\psi_0\left(\theta^t u\right) + \sum_{j=0}^{t-1}\beta^j\left[\omega(\theta^j u) + \alpha(\theta^j u)' x_{t-j}\right].
\end{aligned}
$$

Hence,

$$
\psi_t(u) = \omega(u) + \alpha(u)x_t + \beta\psi_{t-1}(\theta u).
$$

1.2 Multilags version

This paper focuses on the GARG of order $(1,1)$, but a generalization to any order (p,q) is defined as follows:

$$
\psi_t(u) = \omega(u) + \sum_{j=1}^{p}\alpha_j(u)x_{t+1-j} + \sum_{i=1}^{q}\beta_i\psi_{t-i}(\theta_i u). \tag{4}
$$

We establish that the recursion given by equation (4) is a well-defined cumulant generating function dynamic (see Feunou and Meddahi, 2009 for details on these issues). Equation (4) implies that x_t is an $ARMA(p,q)$.

1.3 Extension of the Heston and Nandi model

The Heston and Nandi (2000) model is arguably the most popular discrete-time option pricing model. It is an affine-GARCH model where the dynamic of the conditional variance is given by

$$
x_{t+1} = w + bx_t + a\left(\varepsilon_{t+1} - c\sqrt{x_t}\right)^2, \tag{5}
$$

where

$$
\varepsilon_{t+1} \sim i.i.dN(0,1).
$$

Heston and Nandi (2000) show that the log-conditional moment generating function x_{t+1} is affine in x_t:

$$
\psi_t(u) = \ln\left\{E_t\left[\exp\left(ux_{t+1}\right)\right]\right\} = \omega_{hn}(u) + \alpha_{hn}(u)x_t,
$$

where

$$
\begin{aligned}
\omega_{hn}(u) &= uw - \frac{1}{2}\ln(1-2ua) \\
\alpha_{hn}(u) &= ub + \frac{uac^2}{1-2ua}.
\end{aligned}
$$

Using the exact same steps following the generalisation of the ARG process, we are able to generalize the Heston and Nandi model by decomposing x_{t+1} as in equation (1) and by specifying the following state-space for $Z_{t+1}^{(j)}$:

$$
Z_{t+1}^{(j)} = w_j + b_j x_{t-j} + \bar{Z}_{t+1}^{(j)},
$$

where

$$\left.\frac{\bar{Z}_{t+1}^{(j)}}{\varphi_j}\right| U_{t+1}^{(j)}, I_t \quad \sim \quad \gamma\left(\nu_j + U_{t+1}^{(j)}\right)$$

$$\left. U_{t+1}^{(j)}\right| I_t \quad \sim \quad P\left(\frac{\phi_j x_{t-j}}{\varphi_j}\right)$$

and

$$w_j = (\beta\theta)^j w, \quad b_j = (\beta\theta)^j b, \quad \nu_j = \frac{\beta^j}{2}, \quad \phi_j = ac^2 (\beta\theta)^j, \quad \varphi_j = 2\theta^j a.$$

We show that the cumulant generating function of the new dynamic is recursive:

$$\psi_t(u) = \omega_{hn}(u) + \alpha_{hn}(u)x_t + \beta\psi_{t-1}(\theta u) \qquad \text{for } t \geq 1. \tag{6}$$

1.4 Extension of the IG GARCH process

The IG GARCH process of Christoffersen et al. (2006) is also a dynamic for the conditional variance x_t that is affine but non-Gaussian, and builds on inverse-Gaussian innovations instead. The process is specified as follows:

$$x_{t+1} = w + bx_t + cy_{t+1} + a\frac{x_t^2}{y_{t+1}},$$

where

$$y_{t+1}|x_t \sim IG\left(\frac{x_t}{\eta^2}\right).$$

The IG GARCH is affine. Indeed, Christoffersen et al. (2006) establish that

$$E_t\left[\exp\left(ux_{t+1}\right)\right] = \exp\left(\omega_{ig}(u) + \alpha_{ig}(u)\, x_t\right),$$

where

$$\omega_{ig}(u) = uw - \frac{1}{2}\ln\left(1 - 2ua\eta^4\right)$$

$$\alpha_{ig}(u) = ub + \frac{1}{\eta^2}\left(1 - \sqrt{\left(1 - 2ua\eta^4\right)\left(1 - 2uc\right)}\right).$$

In the same vein as the ARG and the Heston-Nandi GARCH, we are able to generalize the IG GARCH by decomposing x_{t+1} as in equation (1) and by specifying the following state-space for $Z_{t+1}^{(j)}$:

$$Z_{t+1}^{(j)} = w_j + b_j x_{t-j} + c_j y_{t+1}^{(j)} + a_j \frac{x_{t-j}^2}{y_{t+1}^{(j)}} + \varepsilon_{t+1}^{(j)},$$

where

$$y_{t+1}^{(j)}|x_t \sim IG\left(\frac{x_{t-j}}{\eta_j^2}\right), \quad \frac{\varepsilon_{t+1}^{(j)}}{\varphi_j} \sim \gamma\left(\beta^j - 1\right)$$

and

$$w_j \equiv \beta^j\theta^j w, \quad b_j \equiv \beta^j\theta^j b, \quad c_j \equiv \theta^j c, \quad \eta_j^2 \equiv \frac{\eta^2}{\beta^j}, \quad a_j \equiv \theta^j\beta^{2j} a, \quad \varphi_j \equiv 2a_j\eta_j^4.$$

We show that the cumulant generating function of the new dynamic is recursive:

$$\psi_t(u) = \omega_{ig}(u) + \alpha_{ig}(u)x_t + \beta\psi_{t-1}(\theta u) \qquad \text{for } t \geq 1. \tag{7}$$

1.5 Generalized INAR processes

In risk analysis, the variable of interest x_t is often integer-valued and measures the number of claims in period t. The processes in this class are called integer-valued autoregressive (INAR) and have been explored in the time series and insurance literature (see Darolles et al., 2006 for a detailed discussion). The INAR is an affine model with a linear conditional cumulant generating function similar to those of ARG dynamics:

$$\psi_t(u) \equiv \ln\left[E\left[\exp\left(ux_{t+1}\right)|I_t\right]\right] = \omega_{inar}(u) + \alpha_{inar}(u)x_t, \tag{8}$$

where I_t is the sigma algebra generated by $(x_s, s \leq t)$,

$$\omega_{inar}(u) = \lambda\left(\exp(u) - 1\right), \text{ and } \alpha_{inar}(u) = \ln\left(\rho\exp(u) + 1 - \rho\right), \tag{9}$$

where $0 < \rho < 1$ and $\lambda > 1$. Using the decomposition of x_{t+1} given in equation (1) we generalize this INAR process in the following state-space:

$$\begin{aligned} Z_{t+1}^{(j)} &= \theta^j\left(\sum_{i=1}^{\beta^j x_{t-j}} y_{i,t+1}^{(j)} + \varepsilon_{t+1}^{(j)}\right), \\ y_{i,t+1}^{(j)} &= \begin{cases} 1 \\ 0 \end{cases} ;\Pr\left[y_{i,t+1}^{(j)} = 1\right] = \rho, \\ \varepsilon_{t+1}^{(j)} &\sim P\left(\lambda\beta^j\right), \end{aligned}$$

where β and θ are positive integers. Through straightforward derivations, we establish that the generalized INAR processes follow a recursion similar to equation (2):

$$\psi_t\left(u\right) = \omega_{inar}(u) + \alpha_{inar}(u)x_t + \beta\psi_{t-1}\left(\theta u\right) \quad \text{for } t \geq 1. \tag{10}$$

For the remainder of this appendix, we focus solely on the univariate case with the following expression of functions $\alpha(u)$ and $\omega(u)$:

$$\omega(u) = -\nu\log(1 - u\varphi) \text{ and } \alpha(u) = \frac{\phi u}{1 - u\varphi}, \tag{11}$$

with $\nu \geq 0$ $\varphi > 0$ and $\phi \geq 0$.

2 Computing $E\left[\psi_t\left(u\right)\right]$ and $Corr\left(\psi_t^{(n)}(0), \psi_t^{(m)}(0)\right)$

Since $\varepsilon_t \equiv x_t - \psi_{t-1}'\left(0\right)$ is a martingale difference, we have

$$\psi_t'\left(0\right) = \omega'(0) + \left(\alpha'(0) + \beta\theta\right)\psi_{t-1}'\left(0\right) + \alpha'(0)\varepsilon_t. \tag{12}$$

Hence, the conditional expectation, $\psi_t'\left(0\right)$, is a $AR(1)$, implying that x_t is a ARMA(1,1).

Using the same rationale, we have

$$\begin{pmatrix} \psi_t'\left(0\right) \\ \psi_t^{(n)}\left(0\right) \end{pmatrix} = \begin{pmatrix} \omega'(0) \\ \omega^{(n)}(0) \end{pmatrix} + \begin{bmatrix} \alpha'(0) + \beta\theta & 0 \\ \alpha^{(n)}(0) & \beta\theta^n \end{bmatrix}\begin{pmatrix} \psi_{t-1}'\left(0\right) \\ \psi_{t-1}^{(n)}\left(0\right) \end{pmatrix} + \begin{pmatrix} \alpha'(0) \\ \alpha^{(n)}(0) \end{pmatrix}\varepsilon_t, \tag{13}$$

which implies that $\begin{pmatrix} \psi_t'\left(0\right) \\ \psi_t^{(n)}\left(0\right) \end{pmatrix}$ is a VAR(1), in particular $\begin{pmatrix} x_t \\ \varepsilon_t^2 \end{pmatrix}$ is a VARMA(1,1).[1] It follows that the n^{th} order cumulant $\psi_t^{(n)}\left(0\right)$ is mean-stationary if and only if $\rho < 1$ and $\beta\theta^n < 1$, where

$$\rho \equiv \alpha'\left(0\right) + \beta\theta. \tag{14}$$

[1] Indeed, $\begin{pmatrix} \psi_t'\left(0\right) \\ \psi_t^{(2)}\left(0\right) \end{pmatrix} = E_t\left[\begin{pmatrix} x_{t+1} \\ \varepsilon_{t+1}^2 \end{pmatrix}\right]$, and since $\begin{pmatrix} \psi_t'\left(0\right) \\ \psi_t^{(2)}\left(0\right) \end{pmatrix}$ is a VAR(1), it implies that $\begin{pmatrix} x_t \\ \varepsilon_t^2 \end{pmatrix}$ is a VARMA(1,1).

In that case, the unconditional mean is given by

$$E\left[\psi_t^{(n)}(0)\right] = \frac{\alpha^{(n)}(0)\omega'(0) + (1-\rho)\,\omega^{(n)}(0)}{(1-\rho)\,(1-\beta\theta^n)}. \tag{15}$$

We can thus derive the unconditional expectation of $\psi_t(u)$ using the following identity:

$$\begin{aligned} E\left[\psi_t(u)\right] &= \sum_{n=1}^{\infty} \frac{u^n}{n!} E\left[\psi_t^{(n)}(0)\right] = \sum_{n=1}^{\infty} \frac{u^n}{n!} \frac{\alpha^{(n)}(0)\mu + \omega^{(n)}(0)}{1-\beta\theta^n} \\ &= \mu \left\{ \sum_{n=1}^{\infty} \frac{u^n}{n!} \frac{\alpha^{(n)}(0)}{1-\beta\theta^n} \right\} + \left\{ \sum_{n=1}^{\infty} \frac{u^n}{n!} \frac{\omega^{(n)}(0)}{1-\beta\theta^n} \right\}, \end{aligned}$$

where $\mu \equiv \frac{\omega'(0)}{1-\rho}$ is the unconditional expectation of x_t. Using functions $\omega(u)$ and $\alpha(u)$ defined in equation (11), we deduce that

$$E\left[\psi_t(u)\right] = \frac{\phi}{1-\beta\theta} \frac{\theta u}{1-\theta\varphi u}\mu - \frac{\nu}{1-\beta\theta} \ln\left(1-\theta\varphi u\right). \tag{16}$$

It is also worth stressing that $E\left[\psi_t(u)\right]$ is not the unconditional cumulant function of x_t. Later we discuss the required conditions for the unconditional distribution (and hence the unconditional cumulant function) of x_t to exist. Since $\psi_0(u)$ given in equation (16) is the cumulant generating function of x_1, it implies that x_1 has the following state-space representation:

$$\begin{aligned} \frac{x_1}{\theta\varphi}|U &\sim \gamma\left(\frac{\nu}{1-\beta\theta} + U\right) \\ U &\sim P\left(\frac{\phi}{1-\beta\theta}\frac{\mu}{\varphi}\right). \end{aligned} \tag{17}$$

Equation (13) implies that the vector $\begin{pmatrix} \psi_t'(0) \\ \psi_t^{(n)}(0) \end{pmatrix}$ is covariance-stationary if and only if $\rho < 1$ and $\beta\theta^n < 1$. This result can easily be generalized to the vector $\begin{pmatrix} \psi_t^{(n)}(0) \\ \psi_t^{(m)}(0) \end{pmatrix}$ for two positive integers n and m. $\begin{pmatrix} \psi_t^{(n)}(0) \\ \psi_t^{(m)}(0) \end{pmatrix}$ is covariance-stationary if and only if $\rho < 1$, $\beta\theta^n < 1$ and $\beta\theta^m < 1$. In that case, the unconditional covariance is given by

$$Cov\left(\psi_t^{(n)}(0), \psi_t^{(m)}(0)\right) = E\left[\psi_t^{(2)}(0)\right]\left(\vartheta + \frac{\beta\theta^n}{1-\rho\beta\theta^n} + \frac{\beta\theta^m}{1-\rho\beta\theta^m}\right)\frac{\alpha'(0)\,(1-\rho\beta\theta)\,\alpha^{(n)}(0)\alpha^{(m)}(0)}{(1-\rho^2)\,(1-\beta^2\theta^{n+m})},$$

which implies that the correlation is

$$Corr\left(\psi_t^{(n)}(0), \psi_t^{(m)}(0)\right) = \sqrt{\frac{1-\left[\frac{\beta(\theta^n-\theta^m)}{1-\beta^2\theta^{n+m}}\right]^2}{1-\left[\frac{\beta(\theta^n-\theta^m)}{\xi}\right]^2}}, \tag{18}$$

where

$$\xi \equiv (1-\rho\beta\theta^n)\,(1-\rho\beta\theta^m)\,\vartheta + \frac{\beta\theta^n}{1-\rho\beta\theta^m} + \frac{\beta\theta^m}{1-\rho\beta\theta^n},$$

and $\vartheta = \frac{1-\rho^2+\left(\alpha'(0)\right)^2}{\alpha'(0)(1-\rho\beta\theta)}$. From equation (18), it is readily apparent that $\theta = 1$ implies that all the cumulants are perfectly correlated, and thus $\theta \neq 1$ is essential to break the tight link between moments that are inherent within ARG processes.

3 Multi-step ahead cumulant generating function and cumulants

3.1 Multi-step ahead cumulant generating function

We denote the h-step ahead cumulant generating function by $\psi_t(u;h)$, that is,

$$\psi_t(u;h) \equiv \ln\left[E_t\left[\exp\left(ux_{t+h}\right)\right]\right]$$

$$\begin{aligned}
E_t\left[\exp\left(ux_{t+h+1}\right)\right] &= E_t\left[E_{t+h}\left[\exp\left(ux_{t+h+1}\right)\right]\right] \\
&= E_t\left[\exp\left(\psi_{t+h}(u)\right)\right] \\
&= E_t\left[\exp\left(\beta^h\psi_t\left(\theta^h u\right) + \sum_{i=0}^{h-1}\beta^i\left(\omega\left(\theta^i u\right) + \alpha\left(\theta^i u\right)x_{t+h-i}\right)\right)\right] \\
&= \exp\left(\beta^h\psi_t\left(\theta^h u\right) + \sum_{i=0}^{h-1}\beta^i\omega\left(\theta^i u\right)\right)E_t\left[\exp\left(\sum_{i=0}^{h-1}\beta^i\alpha\left(\theta^i u\right)x_{t+h-i}\right)\right]
\end{aligned}$$

$$\Psi_t\left(u^{(h)};h\right) \equiv \ln\left[E_t\left[\exp\left(\sum_{j=1}^{h}u_j^{(h)}x_{t+j}\right)\right]\right]$$

$$\begin{aligned}
& E_t\left[\exp\left(\sum_{j=1}^{h+1}u_j^{(h+1)}x_{t+j}\right)\right] \\
= \; & E_t\left[E_{t+h}\left[\exp\left(\sum_{j=1}^{h+1}u_jx_{t+j}\right)\right]\right] \\
= \; & E_t\left[\exp\left(\psi_{t+h}\left(u_{h+1}^{(h+1)}\right) + \sum_{j=1}^{h}u_j^{(h+1)}x_{t+j}\right)\right] \\
= \; & E_t\left[\exp\left(\begin{array}{c}\beta^h\psi_t\left(\theta^h u_{h+1}^{(h+1)}\right) + \sum_{i=0}^{h-1}\beta^i\left(\omega\left(\theta^i u_{h+1}^{(h+1)}\right) + \alpha\left(\theta^i u_{h+1}^{(h+1)}\right)x_{t+h-i}\right) \\ + \sum_{j=1}^{h}u_j^{(h+1)}x_{t+j}\end{array}\right)\right] \\
= \; & \exp\left(\beta^h\psi_t\left(\theta^h u_{h+1}^{(h+1)}\right) + \sum_{i=0}^{h-1}\beta^i\omega\left(\theta^i u_{h+1}^{(h+1)}\right)\right)E_t\left[\exp\left(\sum_{j=1}^{h}\left(u_j^{(h+1)} + \beta^{h-j}\alpha\left(\theta^{h-j}u_{h+1}^{(h+1)}\right)\right)x_{t+j}\right)\right] \\
= \; & \exp\left(\Psi_t\left(u^{(h)};h\right) + \beta^h\psi_t\left(\theta^h u_{h+1}^{(h+1)}\right) + \sum_{i=0}^{h-1}\beta^i\omega\left(\theta^i u_{h+1}^{(h+1)}\right)\right)
\end{aligned}$$

$$\begin{aligned}
\Psi_t\left(u^{(h+1)};h+1\right) &= \Psi_t\left(u^{(h)};h\right) + \beta^h\psi_t\left(\theta^h u_{h+1}^{(h+1)}\right) + \sum_{i=0}^{h-1}\beta^i\omega\left(\theta^i u_{h+1}^{(h+1)}\right) \\
u_j^{(h)} &= u_j^{(h+1)} + \beta^{h-j}\alpha\left(\theta^{h-j}u_{h+1}^{(h+1)}\right),\ j = 1,...,h,
\end{aligned}$$

with

$$\begin{aligned}
\Psi_t\left(u^{(h+1)};h+1\right) &= \psi_t\left(u^{(1)}\right) + \sum_{j=2}^{h+1}\left[\beta^{j-1}\psi_t\left(\theta^{j-1}u_j^{(j)}\right) + \sum_{i=0}^{j-2}\beta^i\omega\left(\theta^i u_j^{(j)}\right)\right] \\
&= \sum_{j=1}^{h+1}\beta^{j-1}\psi_t\left(\theta^{j-1}u_j^{(j)}\right) + \sum_{j=2}^{h+1}\sum_{i=0}^{j-2}\beta^i\omega\left(\theta^i u_j^{(j)}\right)
\end{aligned}$$

and

$$\begin{aligned}\psi_t(u;h) &= \sum_{j=1}^{h}\beta^{j-1}\psi_t\left(\theta^{j-1}u_j^{(j)}\right)+\sum_{j=2}^{h}\sum_{i=0}^{j-2}\beta^{i}\omega\left(\theta^{i}u_j^{(j)}\right) \text{ for } h\geq 2\\ u_j^{(h)} &= 0 \text{ for } j\leq h,\ u_h^{(h)}=u\\ u_j^{(\tau)} &= u_j^{(\tau+1)}+\beta^{\tau-j}\alpha\left(\theta^{\tau-j}u_{\tau+1}^{(\tau+1)}\right),\ \text{for } j=1,...,\tau \text{ and } 1\leq\tau\leq h-1\end{aligned}$$

$$u_j^{(h-1)}=\beta^{h-1-j}\alpha\left(\theta^{h-1-j}u\right) \text{ for } j=1,...,h-1,$$

in particular,

$$u_{h-1}^{(h-1)}=\alpha(u)$$

$$u_j^{(h-2)}=\beta^{h-1-j}\alpha\left(\theta^{h-1-j}u\right)+\beta^{h-2-j}\alpha\left(\theta^{h-2-j}u_{h-1}^{(h-1)}\right),\quad \text{for } j=1,...,h-2,$$

and

$$u_{h-2}^{(h-2)}=\beta\alpha(\theta u)+\alpha\left(u_{h-1}^{(h-1)}\right)$$

$$\begin{aligned}u_j^{(h-3)} &= u_j^{(h-2)}+\beta^{h-3-j}\alpha\left(\theta^{h-3-j}u_{h-2}^{(h-2)}\right),\ \text{for } j=1,...,h-3\\ &= \beta^{h-1-j}\alpha\left(\theta^{h-1-j}u\right)+\beta^{h-2-j}\alpha\left(\theta^{h-2-j}\alpha(u)\right)+\beta^{h-3-j}\alpha\left(\theta^{h-3-j}u_{h-2}^{(h-2)}\right)\end{aligned}$$

and

$$u_{h-3}^{(h-3)}=\beta^2\alpha\left(\theta^2u\right)+\beta\alpha(\theta\alpha(u))+\alpha\left(u_{h-2}^{(h-2)}\right)$$

$$\begin{aligned}u_j^{(\tau)} &= u_j^{(\tau+1)}+\beta^{\tau-j}\alpha\left(\theta^{\tau-j}u_{\tau+1}^{(\tau+1)}\right)\\ u_j^{(\tau-1)} &= u_j^{(\tau)}+\beta^{\tau-1-j}\alpha\left(\theta^{\tau-1-j}u_{\tau}^{(\tau)}\right)\\ &\vdots\\ u_j^{(\tau-k)} &= u_j^{(\tau-k+1)}+\beta^{\tau-k-j}\alpha\left(\theta^{\tau-k-j}u_{\tau-k+1}^{(\tau-k+1)}\right).\end{aligned}$$

Hence,

$$u_j^{(\tau-k)}=\sum_{i=0}^{k}\beta^{\tau-i-j}\alpha\left(\theta^{\tau-i-j}u_{\tau-i+1}^{(\tau-i+1)}\right),$$

in particular

$$\begin{aligned}u_j^{(h-k)} &= \sum_{i=0}^{k}\beta^{h-i-j}\alpha\left(\theta^{h-i-j}u_{h-i+1}^{(h-i+1)}\right),\quad j=1,...,h-k\\ &= \sum_{s=1}^{k+1}\beta^{h-k+s-(j+1)}\alpha\left(\theta^{h-k+s-(j+1)}u_{h-k+s}^{(h-k+s)}\right),\quad j=1,...,h-k,\quad 0<k<h\end{aligned}$$

or

$$\begin{aligned}u_j^{(\tau)} &= \sum_{s=1}^{h-\tau+1}\beta^{\tau+s-(j+1)}\alpha\left(\theta^{\tau+s-(j+1)}u_{\tau+s}^{(\tau+s)}\right),\quad j=1,...,\tau,\quad 0<\tau<h\\ u_j^{(\tau)} &= \sum_{i=\tau+1}^{h}\beta^{i-(j+1)}\alpha\left(\theta^{i-(j+1)}u_i^{(i)}\right),\quad j=1,...,\tau,\quad 0<\tau<h,\end{aligned}$$

and

$$u_\tau^{(\tau)}=\sum_{i=\tau+1}^{h}\beta^{i-(\tau+1)}\alpha\left(\theta^{i-(\tau+1)}u_i^{(i)}\right).\quad 0<\tau<h$$

3.2 Multi-step ahead conditional cumulant

$$\begin{aligned}
\psi_t(u;h) &= \sum_{j=1}^{h}\beta^{j-1}\psi_t\left(\theta^{j-1}u_j\right)+\sum_{j=2}^{h}\sum_{i=0}^{j-2}\beta^{i}\omega\left(\theta^{i}u_j\right) \text{ for } h\geq 2\\
u_h &= u\\
u_\tau &= \sum_{i=\tau+1}^{h}\beta^{i-(\tau+1)}\alpha\left(\theta^{i-(\tau+1)}u_i\right) \text{ for } 1\leq\tau\leq h-1.
\end{aligned}$$

Let us denote

$$f_\tau(u)=cu_\tau$$

and

$$g_\tau(u)=\psi\circ f_\tau(u)=\psi\left(f_\tau(u)\right).$$

Then we have

$$\begin{aligned}
\frac{d}{du^n}g_\tau(0) &= \sum_{k=1}^{n}\psi^{(k)}(0)\,B_{n,k}\left(f_\tau'(0),f_\tau''(0),\cdots,f_\tau^{(n-k+1)}(0)\right)\\
&= \sum_{k=1}^{n}\psi^{(k)}(0)\,c^kB_{n,k}\left(u_\tau'(0),u_\tau''(0),\cdots,u_\tau^{(n-k+1)}(0)\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
\psi_t^{(n)}(0;h) &= \sum_{j=1}^{h}\beta^{j-1}\sum_{k=1}^{n}\psi_t^{(k)}(0)\left(\theta^{j-1}\right)^kB_{n,k}\left(u_j'(0),u_j''(0),\cdots,u_j^{(n-k+1)}(0)\right)\\
&\quad+\sum_{j=2}^{h}\sum_{i=0}^{j-2}\beta^{i}\sum_{k=1}^{n}\omega^{(k)}(0)\left(\theta^{i}\right)^kB_{n,k}\left(u_j'(0),u_j''(0),\cdots,u_j^{(n-k+1)}(0)\right)
\end{aligned}$$

$$\begin{aligned}
\psi_t^{(n)}(0;h) &= \sum_{k=1}^{n}\left[\sum_{j=1}^{h}\left(\beta\theta^k\right)^{j-1}B_{n,k}\left(u_j'(0),u_j''(0),\cdots,u_j^{(n-k+1)}(0)\right)\right]\psi_t^{(k)}(0)\\
&\quad+\sum_{k=1}^{n}\left[\sum_{j=1}^{h}\frac{1-\left(\beta\theta^k\right)^{j-1}}{1-\beta\theta^k}B_{n,k}\left(u_j'(0),u_j''(0),\cdots,u_j^{(n-k+1)}(0)\right)\right]\omega^{(k)}(0)
\end{aligned}$$

$$\begin{aligned}
\psi_t^{(n)}(0;h) &= \sum_{k=1}^{n}\left[\sum_{j=1}^{h}\left(\beta\theta^k\right)^{j-1}B_{n,k}\left(u_j'(0),u_j''(0),\cdots,u_j^{(n-k+1)}(0)\right)\right]\left[\psi_t^{(k)}(0)-\frac{\omega^{(k)}(0)}{1-\beta\theta^k}\right]\\
&\quad+\sum_{k=1}^{n}\left[\sum_{j=1}^{h}B_{n,k}\left(u_j'(0),u_j''(0),\cdots,u_j^{(n-k+1)}(0)\right)\right]\frac{\omega^{(k)}(0)}{1-\beta\theta^k}
\end{aligned}$$

$$\begin{aligned}
u_h^{(n)}(0) &= 1_{[n=1]}+u1_{[n=0]}\\
u_\tau^{(n)}(0) &= \sum_{k=1}^{n}\left[\sum_{i=\tau+1}^{h}\left(\beta\theta^k\right)^{i-(\tau+1)}B_{n,k}\left(u_i'(0),u_i''(0),\cdots,u_i^{(n-k+1)}(0)\right)\right]\alpha^{(k)}(0) \text{ for } 1\leq\tau\leq h-1
\end{aligned}$$

for $0 \leq \tau \leq h-1$. Let us denote

$$\begin{aligned}
a_{n,k}(\tau) &\equiv \sum_{j=\tau}^{h} \left(\beta\theta^k\right)^{j-1} B_{n,k}\left(u_j'(0), u_j''(0), \cdots, u_j^{(n-k+1)}(0)\right) \\
b_{n,k}(\tau) &\equiv \sum_{j=\tau}^{h} B_{n,k}\left(u_j'(0), u_j''(0), \cdots, u_j^{(n-k+1)}(0)\right) \\
c_{n,k}(\tau) &\equiv \left(\beta\theta^k\right)^{1-\tau} a_{n,k}(\tau) = \sum_{j=\tau}^{h} \left(\beta\theta^k\right)^{j-\tau} B_{n,k}\left(u_j'(0), u_j''(0), \cdots, u_j^{(n-k+1)}(0)\right).
\end{aligned}$$

We have

$$\begin{aligned}
u_\tau^{(n)}(0) &= \sum_{k=1}^{n} \left[\sum_{i=\tau+1}^{h} \left(\beta\theta^k\right)^{i-(\tau+1)} B_{n,k}\left(u_i'(0), u_i''(0), \cdots, u_i^{(n-k+1)}(0)\right)\right] \alpha^{(k)}(0) \text{ for } 1 \leq \tau \leq h-1 \\
&= \sum_{k=1}^{n} c_{n,k}(\tau+1)\, \alpha^{(k)}(0) \text{ for } 1 \leq \tau \leq h-1.
\end{aligned}$$

For the sake of conciseness, we will denote

$$\begin{aligned}
U_\tau^{(n)}(0) &\equiv \begin{pmatrix} u_\tau^{(1)}(0) \\ \vdots \\ u_\tau^{(n)}(0) \end{pmatrix} \\
B_{n,k}\left(U_j^{(n-k+1)}(0)\right) &\equiv B_{n,k}\left(u_j'(0), u_j''(0), \cdots, u_j^{(n-k+1)}(0)\right).
\end{aligned}$$

We have

$$\begin{aligned}
b_{n,k}(\tau) &= \sum_{j=\tau}^{h} B_{n,k}\left(U_j^{(n-k+1)}(0)\right) = B_{n,k}\left(U_\tau^{(n-k+1)}(0)\right) + b_{n,k}(\tau+1) \\
c_{n,k}(\tau) &= \sum_{j=\tau}^{h} \left(\beta\theta^k\right)^{j-\tau} B_{n,k}\left(U_j^{(n-k+1)}(0)\right) = B_{n,k}\left(U_\tau^{(n-k+1)}(0)\right) + \sum_{j=\tau+1}^{h} \left(\beta\theta^k\right)^{j-\tau} B_{n,k}\left(U_j^{(n-k+1)}(0)\right) \\
&= B_{n,k}\left(U_\tau^{(n-k+1)}(0)\right) + \beta\theta^k \sum_{j=\tau+1}^{h} \left(\beta\theta^k\right)^{j-\tau-1} \left(U_j^{(n-k+1)}(0)\right) \\
&= B_{n,k}\left(U_\tau^{(n-k+1)}(0)\right) + \beta\theta^k c_{n,k}(\tau+1).
\end{aligned}$$

This implies that

$$B_{n,k}\left(U_\tau^{(n-k+1)}(0)\right) = b_{n,k}(\tau) - b_{n,k}(\tau+1) = c_{n,k}(\tau) - \beta\theta^k c_{n,k}(\tau+1),$$

and for $1 \leq \tau \leq h-1$, we have

$$u_\tau^{(n)}(0) = \sum_{k=1}^{n} c_{n,k}(\tau+1)\, \alpha^{(k)}(0),$$

which implies that

$$\begin{aligned}
u_\tau^{(1)}(0) &= c_{1,1}(\tau+1)\, \alpha^{(1)}(0) \\
u_\tau^{(2)}(0) &= c_{2,1}(\tau+1)\, \alpha^{(1)}(0) + c_{2,2}(\tau+1)\, \alpha^{(2)}(0) \\
&\vdots \\
u_\tau^{(n)}(0) &= c_{n,1}(\tau+1)\, \alpha^{(1)}(0) + c_{n,2}(\tau+1)\, \alpha^{(2)}(0) + \cdots + c_{n,n}(\tau+1)\, \alpha^{(n)}(0).
\end{aligned}$$

Hence,

$$\begin{pmatrix} u_\tau^{(1)}(0) \\ \vdots \\ u_\tau^{(n)}(0) \end{pmatrix} = \begin{bmatrix} c_{1,1}(\tau+1) & 0 & 0 \\ \vdots & \ddots & 0 \\ c_{n,1}(\tau+1) & \cdots & c_{n,n}(\tau+1) \end{bmatrix} \begin{pmatrix} \alpha^{(1)}(0) \\ \vdots \\ \alpha^{(n)}(0) \end{pmatrix}$$

and

$$U_\tau^{(n)}(0) = C_n(\tau+1)\,A_n,$$

with

$$C_n(\tau) \equiv \begin{bmatrix} c_{1,1}(\tau) & 0 & 0 \\ \vdots & \ddots & 0 \\ c_{n,1}(\tau) & \cdots & c_{n,n}(\tau) \end{bmatrix}, \quad , A_n \equiv \begin{pmatrix} \alpha^{(1)}(0) \\ \vdots \\ \alpha^{(n)}(0) \end{pmatrix}$$

$$c_{n,k}(\tau) = B_{n,k}\left(U_\tau^{(n-k+1)}(0)\right) + \beta\theta^k c_{n,k}(\tau+1).$$

This implies that

$$\begin{aligned} C_n(\tau) &= \begin{bmatrix} B_{1,1}\left(U_\tau^{(1)}(0)\right) & 0 & 0 \\ \vdots & \ddots & 0 \\ B_{n,1}\left(U_\tau^{(n)}(0)\right) & \cdots & B_{n,n}\left(U_\tau^{(1)}(0)\right) \end{bmatrix} + \beta \begin{bmatrix} \theta^1 c_{1,1}(\tau+1) & 0 & 0 \\ \vdots & \ddots & 0 \\ \theta^1 c_{n,1}(\tau+1) & \cdots & \theta^n c_{n,n}(\tau+1) \end{bmatrix} \\ &= \begin{bmatrix} B_{1,1}\left(U_\tau^{(1)}(0)\right) & 0 & 0 \\ \vdots & \ddots & 0 \\ B_{n,1}\left(U_\tau^{(n)}(0)\right) & \cdots & B_{n,n}\left(U_\tau^{(1)}(0)\right) \end{bmatrix} + \beta \begin{bmatrix} c_{1,1}(\tau+1) & 0 & 0 \\ \vdots & \ddots & 0 \\ c_{n,1}(\tau+1) & \cdots & c_{n,n}(\tau+1) \end{bmatrix} \begin{bmatrix} \theta^1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \theta^n \end{bmatrix} \end{aligned}$$

and

$$C_n(\tau) = B_n\left(U_\tau^{(n)}(0)\right) + \beta C_n(\tau+1)\,\Theta_n,$$

where

$$\mathcal{B}_n\left(U_\tau^{(n)}(0)\right) \equiv \begin{bmatrix} B_{1,1}\left(U_\tau^{(1)}(0)\right) & 0 & 0 \\ \vdots & \ddots & 0 \\ B_{n,1}\left(U_\tau^{(n)}(0)\right) & \cdots & B_{n,n}\left(U_\tau^{(1)}(0)\right) \end{bmatrix}, \quad , \Theta_n \equiv \begin{bmatrix} \theta^1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \theta^n \end{bmatrix}.$$

Hence,

$$\begin{aligned} C_n(\tau) &= \mathcal{B}_n\left(U_\tau^{(n)}(0)\right) + \beta C_n(\tau+1)\,\Theta_n \\ &= \mathcal{B}_n\left(C_n(\tau+1)\,A_n\right) + \beta C_n(\tau+1)\,\Theta_n \\ &\equiv \mathcal{D}_n\left(C_n(\tau+1)\right) \end{aligned}$$

with the following terminal condition

$$c_{n,k}(h) = 1_{[n=k=1]}.$$

Thus,

$$C_n(h) = \begin{bmatrix} 1 & 0_{1\times n-1} \\ 0_{n-1\times 1} & 0_{n-1\times n-1} \end{bmatrix} \equiv \bar{C}_n$$

and

$$C_n(\tau) = \mathcal{D}_n^{\circ^{h-\tau}}\left(\bar{C}_n\right)$$

and

$$C_n(1) = \mathcal{D}_n^{\circ^{h-1}}\left(\bar{C}_n\right)$$

$$\psi_t^{(n)}(0;h) = \sum_{k=1}^{n} c_{n,k}(1)\left[\psi_t^{(k)}(0) - \frac{\omega^{(k)}(0)}{1-\beta\theta^k}\right] + \sum_{k=1}^{n} b_{n,k}(1)\frac{\omega^{(k)}(0)}{1-\beta\theta^k},$$

with

$$\begin{aligned} b_{n,k}(1) &= c_{n,k}(h) + \sum_{\tau=1}^{h-1}\left(c_{n,k}(\tau) - \beta\theta^k c_{n,k}(\tau+1)\right) \\ &= c_{n,k}(h) + \sum_{\tau=1}^{h-1} c_{n,k}(\tau) - \beta\theta^k \sum_{\tau=1}^{h-1} c_{n,k}(\tau+1) \\ &= c_{n,k}(h) + \sum_{\tau=1}^{h-1} c_{n,k}(\tau) - \beta\theta^k \sum_{\tau=2}^{h} c_{n,k}(\tau) \\ &= c_{n,k}(h) + c_{n,k}(1) - \beta\theta^k c_{n,k}(h) + \left(1-\beta\theta^k\right)\sum_{\tau=2}^{h-1} c_{n,k}(\tau) \\ &= \beta\theta^k c_{n,k}(1) + \left(1-\beta\theta^k\right)\sum_{\tau=1}^{h} c_{n,k}(\tau) \end{aligned}$$

$$\begin{aligned} c_{n,k}(1) &= \mathcal{D}_n^{\circ^{h-1}}\left(\bar{C}_n\right)[n,k] \\ \sum_{\tau=1}^{h} c_{n,k}(\tau) &= \sum_{\tau=1}^{h} C_n(\tau)[n,k] = \left(\sum_{\tau=1}^{h} \mathcal{D}_n^{\circ^{h-\tau}}\left(\bar{C}_n\right)\right)[n,k] \\ &= \left(\sum_{\tau=0}^{h-1} \mathcal{D}_n^{\circ^{\tau}}\left(\bar{C}_n\right)\right)[n,k]. \end{aligned}$$

4 Weak ergodicity

4.1 Proof of Proposition 2

$\mathcal{X}_\tau^{(n)}$ can be rewritten as follows:

$$\mathcal{X}_\tau^{(n)} = \begin{bmatrix} \mathcal{X}_\tau^{(n-1)} & 0_{n-1\times 1} \\ \mathcal{X}_\tau^{(n)}[n,1:n-1] & \mathcal{X}_\tau^{(n)}[n,n] \end{bmatrix}, \tag{19}$$

with

$$\mathcal{X}_\tau^{(n)}[n,n] = \left(\alpha^{(1)}(0)\right)^n \frac{\left(\beta\theta^n\right)^\tau - \left(\rho^n\right)^\tau}{\beta\theta^n - \rho^n} \tag{20}$$

and

$$\mathcal{X}_\tau^{(n)}[n,1:n-1]' = B_{n,1:n-1}\left(\mathcal{X}_{\tau-1}^{(n)} A_n\right)' + \beta\Theta_{n-1}\mathcal{X}_{\tau-1}^{(n)}[n,1:n-1]'. \tag{21}$$

Equation (21) can be rewritten as

$$\mathcal{X}_\tau^{(n)}[n,1:n-1]' = \begin{pmatrix} \alpha^{(n)}(0)\,\mathcal{X}_{\tau-1}^{(n)}[n,n] \\ B_{n,2:n-1}\left(\mathcal{X}_{\tau-1}^{(n-1)} A_{n-1}\right)' \end{pmatrix} + \left(e_1^{(n-1)} A_{n-1}' + \beta\Theta_{n-1}\right)\mathcal{X}_{\tau-1}^{(n)}[n,1:n-1]', \tag{22}$$

where $e_1^{(n-1)}$ is the $(n-1)\times 1$ vector taking 1 at the first entry and 0 everywhere else. To establish the proposition, we proceed by induction. Equation (20) implies that $\lim_{h\to\infty}\sum_{\tau=0}^{h}\mathcal{X}_\tau^{(n)}[n,n] < \infty$ if and only if $\beta\theta^n < 1$ and $\rho < 1$. Furthermore, assume that $\lim_{h\to\infty}\sum_{\tau=0}^{h}\mathcal{X}_\tau^{(n-1)} < \infty$ and $\lim_{h\to\infty}\sum_{\tau=0}^{h}\mathcal{X}_\tau^{(n)} < \infty$ if the following conditions are met:

1. **Condition (1):** All the eigenvalues of $e_1^{(n-1)}A'_{n-1}+\beta\Theta_{n-1}$ have modulus strictly less than 1.

2. **Condition (2):** $\lim_{h\to\infty}\sum_{\tau=0}^{h}B_{n,2:n-1}\left(\mathcal{X}_{\tau}^{(n-1)}A_{n-1}\right)<\infty$.

Condition (1) is equivalent to $\rho<1$ and $\beta\theta^{j}<1$ for $j=2,...,n-1$. Let us tackle condition (2) by considering any integer $2\leq k\leq n-1$ and using the following inequality:

$$\begin{aligned}B_{n,k}\left(\mathcal{X}_{\tau}^{(n-1)}A_{n-1}\right) &\leq S(n,k)\left\{\|\mathcal{X}_{\tau}^{(n-1)}A_{n-1}\|_{\infty}\right\}^{k}\\ &\leq S(n,k)\left[\|\mathcal{X}_{\tau}^{(n-1)}\|_{\infty}\right]^{k}\left(\|A_{n-1}\|_{\infty}\right)^{k},\end{aligned}\tag{23}$$

where $S(n,k)$ denotes the sequence of Stirling numbers of the second kind[2] and $\|\cdot\|_{\infty}$ is a matrix norm that is simply the maximum absolute row sum of the matrix. Taking the sum ($\sum_{\tau=0}^{h}$) on both sides of equation (23), we have

$$\sum_{\tau=0}^{h}B_{n,k}\left(\mathcal{X}_{\tau}^{(n-1)}A_{n-1}\right)\leq S(n,k)\left(\|A_{n-1}\|_{\infty}\right)^{k}\sum_{\tau=0}^{h}\left\{\|\mathcal{X}_{\tau}^{(n-1)}\|_{\infty}\right\}^{k}.$$

Furthermore, using d'Alembert's convergence ratio,[3] $\lim_{h\to\infty}\sum_{\tau=0}^{h}\mathcal{X}_{\tau}^{(n-1)}<\infty$ is equivalent to $\lim_{\tau\to\infty}\frac{\|\mathcal{X}_{\tau+1}^{(n-1)}\|_{\infty}}{|\mathcal{X}_{\tau}^{(n-1)}\|_{\infty}}<1$ and thus $\lim_{\tau\to\infty}\frac{\left\{\|\mathcal{X}_{\tau+1}^{(n-1)}\|_{\infty}\right\}^{k}}{\left\{|\mathcal{X}_{\tau}^{(n-1)}\|_{\infty}\right\}^{k}}<1$, and $\lim_{h\to\infty}\sum_{\tau=0}^{h}\left\{\|\mathcal{X}_{\tau}^{(n-1)}\|_{\infty}\right\}^{k}<\infty$, implying that condition (2) is met. Since we have already established the convergence of $\psi_t^{(n)}(0;h)$ for $n=1,2,3$, proposition 2 has been proven.

4.2 Other details

We can express $\mathcal{B}_n(z)$ recursively as follows:

$$\mathcal{B}_n(z)=\begin{bmatrix}\mathcal{B}_{n-1}(z) & 0_{n-1\times1}\\ B_{n,1:n-1}(z) & z(1)^{n}\end{bmatrix}.$$

We also have

$$\begin{aligned}\bar{C}_nA_n &= \begin{pmatrix}\alpha^{(1)}(0)\\0\\\vdots\\0\end{pmatrix},\\ \bar{C}_n\Theta_n &= \begin{bmatrix}\bar{C}_{n-1} & 0_{n-1\times1}\\0_{1\times n-1} & 0\end{bmatrix}\begin{bmatrix}\Theta_{n-1} & 0_{n-1\times1}\\0_{1\times n-1} & \theta^{n}\end{bmatrix}=\begin{bmatrix}\bar{C}_{n-1}\Theta_{n-1} & 0_{n-1\times1}\\0_{1\times n-1} & 0\end{bmatrix}.\end{aligned}$$

Hence,

$$\mathcal{D}_n\left(\bar{C}_n\right)=\begin{bmatrix}\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right) & 0_{n-1\times1}\\ B_{n,1:n-1}\left(\bar{C}_nA_n\right) & \left(\alpha^{(1)}(0)\right)^{n}\end{bmatrix}$$

$$\begin{aligned}\mathcal{D}_n\left(\bar{C}_n\right)A_n &= \begin{bmatrix}\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right) & 0_{n-1\times1}\\ B_{n,1:n-1}\left(\bar{C}_nA_n\right) & \left(\alpha^{(1)}(0)\right)^{n}\end{bmatrix}\begin{pmatrix}A_{n-1}\\ \alpha^{(n)}(0)\end{pmatrix}\\ &= \begin{pmatrix}\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)A_{n-1}\\ B_{n,1:n-1}\left(\bar{C}_nA_n\right)A_{n-1}+\left(\alpha^{(1)}(0)\right)^{n}\alpha^{(n)}(0)\end{pmatrix},\end{aligned}$$

$$\begin{aligned}\mathcal{D}_n\left(\bar{C}_n\right)\Theta_n &= \begin{bmatrix}\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right) & 0_{n-1\times1}\\ B_{n,1:n-1}\left(\bar{C}_nA_n\right) & \left(\alpha^{(1)}(0)\right)^{n}\end{bmatrix}\begin{bmatrix}\Theta_{n-1} & 0_{n-1\times1}\\0_{1\times n-1} & \theta^{n}\end{bmatrix}\\ ] &= \begin{bmatrix}\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\Theta_{n-1} & 0_{n-1\times1}\\ B_{n,1:n-1}\left(\bar{C}_nA_n\right)\Theta_{n-1} & \left(\theta\alpha^{(1)}(0)\right)^{n}\end{bmatrix}\end{aligned}$$

[2] We provide the definition and explicit expression of $S(n,k)$ in Section 4.3.
[3] See "Convergence Tests," §1.3.3 in Zwillinger (2018).

$$
\begin{aligned}
&\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)\\
=\;& \mathcal{B}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)A_n\right)+\beta\mathcal{D}_n\left(\bar{C}_n\right)\Theta_n\\
=\;& \left[\begin{array}{cc} \mathcal{B}_{n-1}\left(\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)A_{n-1}\right) & 0_{n-1\times 1}\\ B_{n,1:n-1}\left(\mathcal{D}_n\left(\bar{C}_n\right)A_n\right) & \rho^n\left(\alpha^{(1)}(0)\right)^n\end{array}\right]+\left[\begin{array}{cc}\beta\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\Theta_{n-1} & 0_{n-1\times 1}\\ \beta B_{n,1:n-1}\left(\bar{C}_nA_n\right)\Theta_{n-1} & \beta\left(\theta\alpha^{(1)}(0)\right)^n\end{array}\right]\\
=\;& \left[\begin{array}{cc} \mathcal{B}_{n-1}\left(\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)A_{n-1}\right)+\beta\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\Theta_{n-1} & 0_{n-1\times 1}\\ B_{n,1:n-1}\left(\mathcal{D}_n\left(\bar{C}_n\right)A_n\right)+\beta B_{n,1:n-1}\left(\bar{C}_nA_n\right)\Theta_{n-1} & \rho^n\left(\alpha^{(1)}(0)\right)^n+\beta\left(\theta\alpha^{(1)}(0)\right)^n\end{array}\right]\\
=\;& \left[\begin{array}{cc} \mathcal{D}_{n-1}\left(\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\right) & 0_{n-1\times 1}\\ B_{n,1:n-1}\left(\mathcal{D}_n\left(\bar{C}_n\right)A_n\right)+\beta B_{n,1:n-1}\left(\bar{C}_nA_n\right)\Theta_{n-1} & \left(\alpha^{(1)}(0)\right)^n\left(\rho^n+\beta\theta^n\right)\end{array}\right]\\
=\;& \left[\begin{array}{cc} \mathcal{D}_{n-1}\left(\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\right) & 0_{n-1\times 1}\\ B_{n,1:n-1}\left(\mathcal{D}_n\left(\bar{C}_n\right)A_n\right)+\beta\mathcal{D}_n\left(\bar{C}_n\right)[n,1:n-1]\,\Theta_{n-1} & \left(\alpha^{(1)}(0)\right)^n\left(\rho^n+\beta\theta^n\right)\end{array}\right].
\end{aligned}
$$

The first element of $\mathcal{D}_n\left(\bar{C}_n\right)A_n$ is $\left(\alpha^{(1)}(0)\right)^2+\beta\theta\alpha^{(1)}(0)=\alpha^{(1)}(0)\,\rho$.

We also have

$$
\mathcal{D}_n\left(\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)\right)=\mathcal{B}_n\left(\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)A_n\right)+\beta\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)\Theta_n
$$

$$
\begin{aligned}
&\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)A_n\\
=\;& \left[\begin{array}{cc} \mathcal{D}_{n-1}\left(\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\right) & 0_{n-1\times 1}\\ B_{n,1:n-1}\left(\mathcal{D}_n\left(\bar{C}_n\right)A_n\right)+\beta B_{n,1:n-1}\left(\bar{C}_nA_n\right)\Theta_{n-1} & \left(\alpha^{(1)}(0)\right)^n\left(\rho^n+\beta\theta^n\right)\end{array}\right]\left(\begin{array}{c} A_{n-1}\\ \alpha^{(n)}(0)\end{array}\right)\\
=\;& \left(\begin{array}{c} \mathcal{D}_{n-1}\left(\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\right)A_{n-1}\\ B_{n,1:n-1}\left(\mathcal{D}_n\left(\bar{C}_n\right)A_n\right)A_{n-1}+\beta B_{n,1:n-1}\left(\bar{C}_nA_n\right)\Theta_{n-1}A_{n-1}+\left(\alpha^{(1)}(0)\right)^n\left(\rho^n+\beta\theta^n\right)\alpha^{(n)}(0)\end{array}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)\Theta_n\\
=\;& \left[\begin{array}{cc} \mathcal{D}_{n-1}\left(\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\right) & 0_{n-1\times 1}\\ B_{n,1:n-1}\left(\mathcal{D}_n\left(\bar{C}_n\right)A_n\right)+\beta B_{n,1:n-1}\left(\bar{C}_nA_n\right)\Theta_{n-1} & \left(\alpha^{(1)}(0)\right)^n\left(\rho^n+\beta\theta^n\right)\end{array}\right]\left[\begin{array}{cc}\Theta_{n-1} & 0_{n-1\times 1}\\ 0_{1\times n-1} & \theta^n\end{array}\right]\\
=\;& \left[\begin{array}{cc} \mathcal{D}_{n-1}\left(\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\right) & 0_{n-1\times 1}\\ B_{n,1:n-1}\left(\mathcal{D}_n\left(\bar{C}_n\right)A_n\right)\Theta_{n-1}+\beta B_{n,1:n-1}\left(\bar{C}_nA_n\right)\left(\Theta_{n-1}\right)^2 & \left(\theta\alpha^{(1)}(0)\right)^n\left(\rho^n+\beta\theta^n\right)\end{array}\right]
\end{aligned}
$$

$$
\mathcal{B}_n(z)=\left[\begin{array}{cc}\mathcal{B}_{n-1}(z) & 0_{n-1\times 1}\\ B_{n,1:n-1}(z) & z(1)^n\end{array}\right]
$$

$$
\begin{aligned}
&\mathcal{D}_n\left(\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)\right)\\
=\;& \left[\begin{array}{cc} \mathcal{B}_{n-1}\left(\mathcal{D}_{n-1}\left(\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\right)A_{n-1}\right) & 0_{n-1\times 1}\\ B_{n,1:n-1}\left(\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)A_n\right) & \left(\alpha^{(1)}(0)\,\rho^2\right)^n\end{array}\right]\\
&+\left[\begin{array}{cc} \beta\mathcal{D}_{n-1}\left(\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\right) & 0_{n-1\times 1}\\ \beta B_{n,1:n-1}\left(\mathcal{D}_n\left(\bar{C}_n\right)A_n\right)\Theta_{n-1}+\beta^2B_{n,1:n-1}\left(\bar{C}_nA_n\right)\left(\Theta_{n-1}\right)^2 & \beta\left(\theta\alpha^{(1)}(0)\right)^n\left(\rho^n+\beta\theta^n\right)\end{array}\right]\\
=\;& \left[\begin{array}{cc} \mathcal{D}_{n-1}\left(\mathcal{D}_{n-1}\left(\mathcal{D}_{n-1}\left(\bar{C}_{n-1}\right)\right)\right) & 0_{n-1\times 1}\\ B_{n,1:n-1}\left(\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)A_n\right)+\beta\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)[n,1:n-1]\Theta_{n-1} & \alpha^{(1)}(0)^n\left(\left(\rho^n\right)^2+\beta\theta^n\rho^n+\left(\beta\theta^n\right)^2\right)\end{array}\right].
\end{aligned}
$$

The first element of $\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)A_n$ is $\alpha^{(1)}(0)\,\rho^2$:

$$
\mathcal{D}_n\left(\mathcal{D}_n\left(\bar{C}_n\right)\right)[n,1:n-1]=B_{n,1:n-1}\left(\mathcal{D}_n\left(\bar{C}_n\right)A_n\right)+\beta B_{n,1:n-1}\left(\bar{C}_nA_n\right)\Theta_{n-1}.
$$

Hence,

$$
\mathcal{D}_n^{\circ^{\tau}}\left(\bar{C}_n\right)[n,1:n-1]=B_{n,1:n-1}\left(\mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)A_n\right)+\beta\mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)[n,1:n-1]\,\Theta_{n-1}
$$

$$\mathcal{D}_n^{\circ^4}\left(\bar{C}_n\right) = \mathcal{B}_n\left(\mathcal{D}_n^{\circ^3}\left(\bar{C}_n\right)A_n\right) + \beta\mathcal{D}_n^{\circ^3}\left(\bar{C}_n\right)\Theta_n$$

$$\mathcal{D}_n^{\circ^3}\left(\bar{C}_n\right)A_n = \begin{pmatrix} \mathcal{D}_{n-1}^{\circ^3}\left(\bar{C}_{n-1}\right)A_{n-1} \\ \mathcal{D}_n^{\circ^3}\left(\bar{C}_n\right)[n,1:n-1]A_{n-1} + \alpha^{(1)}(0)^n\left((\rho^n)^2 + \beta\theta^n\rho^n + (\beta\theta^n)^2\right)\alpha^{(n)}(0) \end{pmatrix}$$

$$\mathcal{D}_n^{\circ^3}\left(\bar{C}_n\right)\Theta_n = \begin{bmatrix} \mathcal{D}_{n-1}^{\circ^3}\left(\bar{C}_{n-1}\right)\Theta_{n-1} & 0_{n-1\times 1} \\ \mathcal{D}_n^{\circ^3}\left(\bar{C}_n\right)[n,1:n-1]\Theta_{n-1} & \alpha^{(1)}(0)^n\left((\rho^n)^2 + \beta\theta^n\rho^n + (\beta\theta^n)^2\right)\theta^n \end{bmatrix}$$

$$\beta\mathcal{D}_n^{\circ^3}\left(\bar{C}_n\right)\Theta_n = \begin{bmatrix} \beta\mathcal{D}_{n-1}^{\circ^3}\left(\bar{C}_{n-1}\right)\Theta_{n-1} & 0_{n-1\times 1} \\ \beta\mathcal{D}_n^{\circ^3}\left(\bar{C}_n\right)[n,1:n-1]\Theta_{n-1} & \alpha^{(1)}(0)^n\left((\rho^n)^2\beta\theta^n + (\beta\theta^n)^2\rho^n + (\beta\theta^n)^3\right) \end{bmatrix}$$

$$\mathcal{B}_n(z) = \begin{bmatrix} \mathcal{B}_{n-1}(z) & 0_{n-1\times 1} \\ B_{n,1:n-1}(z) & z(1)^n \end{bmatrix}$$

$$\begin{aligned} &\mathcal{D}_n^{\circ^4}\left(\bar{C}_n\right) \\ = &\begin{bmatrix} \mathcal{B}_{n-1}\left(\mathcal{D}_{n-1}^{\circ^3}\left(\bar{C}_{n-1}\right)A_{n-1}\right) & 0_{n-1\times 1} \\ B_{n,1:n-1}\left(\mathcal{D}_n^{\circ^3}\left(\bar{C}_n\right)A_n\right) & \left(\alpha^{(1)}(0)\rho^3\right)^n \end{bmatrix} \\ &+ \begin{bmatrix} \beta\mathcal{D}_{n-1}^{\circ^3}\left(\bar{C}_{n-1}\right)\Theta_{n-1} & 0_{n-1\times 1} \\ \beta\mathcal{D}_n^{\circ^3}\left(\bar{C}_n\right)[n,1:n-1]\Theta_{n-1} & \alpha^{(1)}(0)^n\left((\rho^n)^2\beta\theta^n + (\beta\theta^n)^2\rho^n + (\beta\theta^n)^3\right) \end{bmatrix} \\ = &\begin{bmatrix} \mathcal{D}_{n-1}^{\circ^4}\left(\bar{C}_{n-1}\right) & 0_{n-1\times 1} \\ \mathcal{D}_n^{\circ^4}\left(\bar{C}_n\right)[n,1:n-1] & \alpha^{(1)}(0)^n\left((\rho^n)^3 + (\rho^n)^2\beta\theta^n + (\beta\theta^n)^2\rho^n + (\beta\theta^n)^3\right) \end{bmatrix}. \end{aligned}$$

We can then write

$$\mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right) = \begin{bmatrix} \mathcal{D}_{n-1}^{\circ^\tau}\left(\bar{C}_{n-1}\right) & 0_{n-1\times 1} \\ \mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right)[n,1:n-1] & \mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right)[n,n] \end{bmatrix}$$

, with

$$\begin{aligned} \mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right)[n,1:n-1] &= B_{n,1:n-1}\left(\mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)A_n\right) + \beta\mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)[n,1:n-1]\Theta_{n-1} \\ \mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right)[n,n] &= (\beta\theta^n)\mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right)[n,n] + \left(\alpha^{(1)}(0)\rho^{\tau-1}\right)^n. \end{aligned}$$

$\mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right)[n,n]$ converges if $\beta\theta^n < 1$ and $\rho < 1$. In fact, we can compute $\mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right)[n,n]$ in closed-form. Indeed, we have

$$\mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right)[n,n] = \left(\alpha^{(1)}(0)\right)^n\frac{(\beta\theta^n)^\tau - (\rho^n)^\tau}{\beta\theta^n - \rho^n}$$

$$\mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right)[n,1:n-1]' = B_{n,1:n-1}\left(\mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)A_n\right)' + \beta\Theta_{n-1}\mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)[n,1:n-1]'$$

$$\begin{aligned} \mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right)[n,1] &= \mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)[n,1:n-1]A_{n-1} + \mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)[n,n]\alpha^{(n)}(0) \\ &\quad + \beta\begin{bmatrix} \theta & 0 & \cdots & 0 \end{bmatrix}\mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)[n,1:n-1]'. \end{aligned}$$

Hence,

$$\begin{aligned}\mathcal{D}_n^{\circ^\tau}\left(\bar{C}_n\right)[n,1:n-1]' &= \begin{pmatrix}\mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)[n,1:n-1]\,A_{n-1}+\mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)[n,n]\,\alpha^{(n)}(0)\\ B_{n,2:n-1}\left(\mathcal{D}_{n-1}^{\circ^{\tau-1}}\left(\bar{C}_{n-1}\right)A_{n-1}\right)'\end{pmatrix}\\ &\quad+\beta\Theta_{n-1}\mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)[n,1:n-1]'\\ &= \left(\alpha^{(n)}(0)\left(\alpha^{(1)}(0)\right)^n\frac{\left(\beta\theta^n\right)^{\tau-1}-\left(\rho^n\right)^{\tau-1}}{\beta\theta^n-\rho^n}\right)e_1^{(n-1)}\\ &\quad+\begin{pmatrix}0\\ B_{n,2:n-1}\left(\mathcal{D}_{n-1}^{\circ^{\tau-1}}\left(\bar{C}_{n-1}\right)A_{n-1}\right)'\end{pmatrix}\\ &\quad+\left(e_1^{(n-1)}A_{n-1}'+\beta\Theta_{n-1}\right)\mathcal{D}_n^{\circ^{\tau-1}}\left(\bar{C}_n\right)[n,1:n-1]'\end{aligned}$$

$$\mathcal{X}_\tau^{(n)}[n,1:n-1]'=\begin{pmatrix}\alpha^{(n)}(0)\,\mathcal{X}_{\tau-1}^{(n)}[n,n]\\ B_{n,2:n-1}\left(\mathcal{X}_{\tau-1}^{(n-1)}A_{n-1}\right)'\end{pmatrix}+\left(e_1^{(n-1)}A_{n-1}'+\beta\Theta_{n-1}\right)\mathcal{X}_{\tau-1}^{(n)}[n,1:n-1]'.$$

4.3 Exponential Bell polynomials and Stirling numbers of the second kind

The partial or incomplete exponential Bell polynomials are a triangular array of polynomials given by

$$B_{n,k}(x_1,x_2,\ldots,x_{n-k+1})=\sum\frac{n!}{j_1!j_2!\cdots j_{n-k+1}!}\left(\frac{x_1}{1!}\right)^{j_1}\left(\frac{x_2}{2!}\right)^{j_2}\cdots\left(\frac{x_{n-k+1}}{(n-k+1)!}\right)^{j_{n-k+1}},$$

where the sum is taken over all sequences, j_1, j_2, j_3, ..., j_{n-k+1}, of non-negative integers such that these two conditions are satisfied:

$$\begin{aligned}j_1+j_2+\cdots+j_{n-k+1} &= k\\ j_1+2j_2+3j_3+\cdots+(n-k+1)j_{n-k+1} &= n.\end{aligned}$$

The Stirling numbers of the second kind, written $S(n,k)$, count the number of ways to partition a set of n labelled objects into k nonempty unlabelled subsets. They can be calculated using the following explicit formula:

$$S(n,k)=\frac{1}{k!}\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n.$$

5 Unconditional moments

5.1 Unconditional skewness

The multi-step ahead third cumulant is given by

$$\mu_{3,t}\left(x_{t+h}\right)=E_t\left[\psi_{t+h-1}^{(3)}(0)\right]+\mu_{3,t}\left(\psi_{t+h-1}'(0)\right)+3cov_t\left(\psi_{t+h-1}'(0),\psi_{t+h-1}''(0)\right),$$

with

$$\begin{aligned}E_t\left[\psi_{t+h}^{(3)}(0)\right] &= E\left[\psi_t^{(3)}(0)\right]+\alpha^{(3)}(0)\frac{\left(\beta\theta^3\right)^h-\rho^h}{\beta\theta^3-\rho}\left(\psi_t'(0)-E\left[\psi_t'(0)\right]\right)\\ &\quad+\left(\beta\theta^3\right)^h\left(\psi_t^{(3)}(0)-E\left[\psi_t^{(3)}(0)\right]\right),\end{aligned}$$

$$
\begin{aligned}
& cov_t\left(\psi'_{t+h}(0), \psi''_{t+h}(0)\right) \\
= \;& \frac{\alpha'(0)\alpha''(0)E\left[\psi''_t(0)\right]}{(1-\rho^2)(\rho-\beta\theta^2)}\left[\frac{\beta\theta(1-\theta)\left(1-\rho^2\right)}{1-\rho\beta\theta^2}\left(1-\left(\rho\beta\theta^2\right)^h\right)+\alpha'(0)\left(1-\rho^{2h}\right)\right] \\
& +\frac{\alpha'(0)\alpha''(0)^2\left(\psi'_t(0)-E\left[\psi'_t(0)\right]\right)}{(\beta\theta^2-\rho)(1-\rho)}\left[\begin{array}{c}\frac{(\theta-\rho)\left(1-\left(\beta\theta^2\right)^h\right)}{(\rho^2-\beta\theta^2)\theta}+\frac{(1-\theta)\left(1-\left(\rho\beta\theta^2\right)^h\right)}{\rho\theta(1-\beta\theta^2)} \\ +\frac{(1-\beta\theta)\left(1-\rho^h\right)}{\rho(1-\beta\theta^2)}+\frac{\alpha'(0)\left(1-\rho^{2h}\right)}{\rho(\beta\theta^2-\rho^2)}\end{array}\right] \\
& +\frac{\alpha'(0)\alpha''(0)\left(\psi''_t(0)-E\left[\psi''_t(0)\right]\right)}{\beta\theta^2-\rho^2}\left[\frac{\theta-\rho}{\theta(1-\rho)}\left(\left(\beta\theta^2\right)^h-\left(\beta\theta^2\rho\right)^h\right)-\frac{\alpha'(0)}{\rho-\beta\theta^2}\left(\left(\rho^2\right)^h-\left(\beta\theta^2\rho\right)^h\right)\right],
\end{aligned}
$$

and

$$
\begin{aligned}
& \mu_{3,t}\left(\psi'_{t+h}(0)\right) \\
= \;& E\left[\psi^{(3)}_t(0)\right]\alpha'(0)^3\frac{1-\left(\rho^3\right)^h}{1-\rho^3}+\left(\psi'_t(0)-E\left[\psi'_t(0)\right]\right)\frac{\alpha^{(3)}(0)\alpha'(0)^3}{\beta\theta^3-\rho}\left(\frac{\left(\beta\theta^3\right)^h-\left(\rho^3\right)^h}{\beta\theta^3-\rho^3}-\frac{\rho^h-\left(\rho^3\right)^h}{(1-\rho^2)\rho}\right) \\
& +\left(\psi^{(3)}_t(0)-E\left[\psi^{(3)}_t(0)\right]\right)\alpha'(0)^3\frac{\left(\beta\theta^3\right)^h-\left(\rho^3\right)^h}{\beta\theta^3-\rho^3} \\
& +3\rho\alpha'(0)^2E\left[\psi''_t(0)\right]\alpha'(0)\alpha''(0)\left\{\begin{array}{c}\frac{1-\rho\beta\theta}{(1-\rho^2)(1-\rho\beta\theta^2)}\frac{1-\left(\rho^3\right)^h}{1-\rho^3}-\frac{\alpha'(0)}{(1-\rho^2)(\rho-\beta\theta^2)}\frac{\left(\rho^2\right)^h-\left(\rho^3\right)^h}{(1-\rho)\rho^2} \\ +\frac{\beta\theta(\theta-1)}{\rho-\beta\theta^2}\frac{\left(\beta\theta^2\rho\right)^h-\left(\rho^3\right)^h}{(\beta\theta^2-\rho^2)\rho}\end{array}\right\} \\
& +3\rho\alpha'(0)^2\alpha'(0)\alpha''(0)^2\frac{\left(\psi'_t(0)-E\left[\psi'_t(0)\right]\right)}{(\beta\theta^2-\rho)(1-\rho)}\left[\begin{array}{c}\frac{(\theta-\rho)}{\theta(\beta\theta^2-\rho^2)}\frac{\left(\beta\theta^2\right)^h-\left(\rho^3\right)^h}{\beta\theta^2-\rho^3}+\frac{(\theta-1)}{\theta\rho^2(1-\beta\theta^2)}\frac{\left(\beta\theta^2\rho\right)^h-\left(\rho^3\right)^h}{(\beta\theta^2-\rho^2)} \\ -\frac{\alpha'(0)}{\rho^3(\beta\theta^2-\rho^2)}\frac{\left(\rho^2\right)^h-\left(\rho^3\right)^h}{1-\rho}-\frac{1-\beta\theta}{(1-\beta\theta^2)\rho^2}\frac{\rho^h-\left(\rho^3\right)^h}{1-\rho^2}\end{array}\right] \\
& +3\rho\alpha'(0)^2\left(\psi''_t(0)-E\left[\psi''_t(0)\right]\right)\frac{\alpha'(0)\alpha''(0)}{\beta\theta^2-\rho^2}\left[\begin{array}{c}\frac{\theta-\rho}{(1-\rho)\theta}\frac{\left(\beta\theta^2\right)^h-\left(\rho^3\right)^h}{\beta\theta^2-\rho^3}+\frac{\theta-1}{\rho(1-\rho)\theta}\frac{\left(\beta\theta^2\rho\right)^h-\left(\rho^3\right)^h}{\rho-\beta\theta^2} \\ -\frac{\alpha'(0)}{\rho^2(\rho-\beta\theta^2)}\frac{\left(\rho^2\right)^h-\left(\rho^3\right)^h}{1-\rho}\end{array}\right].
\end{aligned}
$$

$\psi^{(3)}_t(0;h)$ converge if and only if $|\rho|<1$, $\left|\beta\theta^2\right|<1$ and $\left|\beta\theta^3\right|<1$. In that case, $\psi^{(3)}_t(0;h)$ converge to

$$
\mu_3(x_t)\equiv\left(1+\frac{\alpha'(0)^3}{1-\rho^3}\right)E\left[\psi^{(3)}_t(0)\right]+3\left(1+\frac{\rho\alpha'(0)^2}{1-\rho^3}\right)\frac{1-\rho\beta\theta}{(1-\rho^2)(1-\rho\beta\theta^2)}\alpha'(0)\alpha''(0)E\left[\psi''_t(0)\right].
$$

Hence, the unconditional skewness is

$$
\begin{aligned}
Skew\left[x_t\right] \;=\; & \frac{\mu_3(x_t)}{Var\left[x_t\right]^{3/2}}=\frac{\left(1+\frac{\alpha'(0)^3}{1-\rho^3}\right)}{\left(1+\frac{\alpha'(0)^2}{1-\rho^2}\right)^{3/2}}\frac{E\left[\psi^{(3)}_t(0)\right]}{E\left[\psi''_t(0)\right]^{3/2}} \\
& +3\frac{(1-\rho\beta\theta)\left(1+\frac{\rho\alpha'(0)^2}{1-\rho^3}\right)\alpha'(0)\alpha''(0)}{(1-\rho^2)(1-\rho\beta\theta^2)\left(1+\frac{\alpha'(0)^2}{1-\rho^2}\right)^{3/2}}E\left[\psi''_t(0)\right]^{-1/2}.
\end{aligned}
$$

5.2 Analytical expression for $cov(z_{t+1}, z_t)$

The goal here is to compute

$$
cov(z_{t+1},z_t)=\left[\begin{array}{ccc}cov(x_{t+1},x_t) & cov(x_{t+1},\varepsilon^2_t) & cov(x_{t+1},x^2_t) \\ cov(\varepsilon^2_{t+1},x_t) & cov(\varepsilon^2_{t+1},\varepsilon^2_t) & cov(\varepsilon^2_{t+1},x^2_t) \\ cov(x^2_{t+1},x_t) & cov(x^2_{t+1},\varepsilon^2_t) & cov(x^2_{t+1},x^2_t)\end{array}\right].
$$

5.2.1 Expression of $cov\left(x_{t+1},x_t\right)$

$$cov\left(x_{t+1},x_t\right)=\alpha'(0)\left[\frac{1-(\beta\theta)^2-\alpha'(0)\beta\theta}{1-(\beta\theta)^2-2\alpha'(0)\beta\theta}\right]Var\left[x_t\right],$$

with

$$\begin{aligned}
Var\left[x_t\right] &= \left(\frac{\alpha'(0)^2}{1-\left(\alpha'(0)+\beta\theta\right)^2}+1\right)E\left[\psi_t''\left(0\right)\right]\\
E\left[\psi_t''\left(0\right)\right] &= \frac{\omega''(0)+\alpha''(0)E\left[x_t\right]}{1-\beta\theta^2}\\
E\left[x_t\right] &= \frac{\omega'(0)}{1-\left(\alpha'(0)+\beta\theta\right)}.
\end{aligned}$$

5.2.2 Expression of $cov(x_{t+1},\varepsilon_t^2)$

$$\begin{aligned}
cov(x_{t+1},\varepsilon_t^2) &= cov\left(\psi_t'\left(0\right),\varepsilon_t^2\right)=cov\left(\omega'(0)+\rho\psi_{t-1}'\left(0\right)+\alpha'(0)\varepsilon_t,\varepsilon_t^2\right)\\
&= \rho cov\left(\psi_{t-1}'\left(0\right),\varepsilon_t^2\right)+\alpha'(0)cov\left(\varepsilon_t,\varepsilon_t^2\right)\\
&= \rho cov\left(\psi_{t-1}'\left(0\right),\psi_{t-1}''\left(0\right)\right)+\alpha'(0)E\left[\varepsilon_t^3\right].
\end{aligned}$$

We have

$$E\left[\varepsilon_t^3\right]=E\left[\psi_t^{(3)}\left(0\right)\right]=\frac{\omega^{(3)}(0)+\alpha^{(3)}(0)E\left[x_t\right]}{1-\beta\theta^3}$$

$$\begin{aligned}
& cov\left(\psi_t'\left(0\right),\psi_t''\left(0\right)\right)\\
= & cov\left(\omega'(0)+\rho\psi_{t-1}'\left(0\right)+\alpha'(0)\varepsilon_t,\omega''(0)+\beta\theta^2\psi_{t-1}''\left(0\right)+\alpha''(0)\psi_{t-1}'\left(0\right)+\alpha''(0)\varepsilon_t\right)\\
= & cov\left(\rho\psi_{t-1}'\left(0\right)+\alpha'(0)\varepsilon_t,\beta\theta^2\psi_{t-1}''\left(0\right)+\alpha''(0)\psi_{t-1}'\left(0\right)+\alpha''(0)\varepsilon_t\right)\\
= & \rho\beta\theta^2cov\left(\psi_{t-1}'\left(0\right),\psi_{t-1}''\left(0\right)\right)+\rho\alpha''(0)cov\left(\psi_{t-1}'\left(0\right),\psi_{t-1}'\left(0\right)\right)+\rho\alpha''(0)cov\left(\psi_{t-1}'\left(0\right),\varepsilon_t\right)\\
& +\alpha'(0)\beta\theta^2cov\left(\varepsilon_t,\psi_{t-1}''\left(0\right)\right)+\alpha'(0)\alpha''(0)cov\left(\varepsilon_t,\psi_{t-1}'\left(0\right)\right)+\alpha'(0)\alpha''(0)cov\left(\varepsilon_t,\varepsilon_t\right)\\
= & \rho\beta\theta^2cov\left(\psi_t'\left(0\right),\psi_t''\left(0\right)\right)+\rho\alpha''(0)var\left[\psi_t'\left(0\right)\right]+\alpha'(0)\alpha''(0)E\left[\psi_t''\left(0\right)\right].
\end{aligned}$$

Hence,

$$cov\left(\psi_t'\left(0\right),\psi_t''\left(0\right)\right)=\frac{\rho\alpha''(0)var\left[\psi_t'\left(0\right)\right]+\alpha'(0)\alpha''(0)E\left[\psi_t''\left(0\right)\right]}{1-\rho\beta\theta^2}$$

with

$$Var\left[\psi_t'\left(0\right)\right]=\frac{\alpha'(0)^2E\left[\psi_t''\left(0\right)\right]}{1-\rho^2}.$$

5.2.3 Expression of $cov(x_{t+1},x_t^2)$

$$\begin{aligned}
& cov(x_{t+1},x_t^2)\\
= & cov(\psi_t'\left(0\right),x_t^2)=cov(\omega'(0)+\rho\psi_{t-1}'\left(0\right)+\alpha'(0)\varepsilon_t,x_t^2)\\
= & \rho cov(\psi_{t-1}'\left(0\right),x_t^2)+\alpha'(0)cov(\varepsilon_t,x_t^2)\\
= & \rho cov(\psi_{t-1}'\left(0\right),\psi_{t-1}''\left(0\right)+\psi_{t-1}'\left(0\right)^2)+\alpha'(0)cov(\varepsilon_t,x_t^2)\\
= & \rho cov(\psi_{t-1}'\left(0\right),\psi_{t-1}''\left(0\right))+\rho cov(\psi_{t-1}'\left(0\right),\psi_{t-1}'\left(0\right)^2)+\alpha'(0)cov(\varepsilon_t,x_t^2)\\
= & \rho cov(\psi_t'\left(0\right),\psi_t''\left(0\right))+\rho cov(\psi_t'\left(0\right),\psi_t'\left(0\right)^2)+\alpha'(0)cov(\varepsilon_t,x_t^2)
\end{aligned}$$

$$
\begin{aligned}
& cov(\psi_t'(0), \psi_t'(0)^2) \\
= \ & cov(\rho\psi_{t-1}'(0), 2\omega'(0)\rho\psi_{t-1}'(0) + \alpha'(0)^2\psi_{t-1}''(0) + \rho^2\psi_{t-1}'(0)^2) + E\left[cov_{t-1}(\psi_t'(0), \psi_t'(0)^2)\right] \\
= \ & 2\omega'(0)\rho^2 cov(\psi_{t-1}'(0), \psi_{t-1}'(0)) + \rho\alpha'(0)^2 cov(\psi_{t-1}'(0), \psi_{t-1}''(0)) + \rho^3 cov(\psi_{t-1}'(0), \psi_{t-1}'(0)^2) \\
& + E\left[cov_{t-1}(\alpha'(0)\varepsilon_t, \alpha'(0)^2\varepsilon_t^2 + 2\alpha'(0)\left(\omega'(0) + \rho\psi_{t-1}'(0)\right)\varepsilon_t)\right] \\
= \ & \rho^3 cov(\psi_t'(0), \psi_t'(0)^2) + 2\omega'(0)\rho^2 Var\left[\psi_t'(0)\right] + \rho\alpha'(0)^2 cov\left(\psi_t'(0), \psi_t''(0)\right) \\
& + \alpha'(0)^3 E\left[\varepsilon_t^3\right] + 2\alpha'(0)^2 E\left[\left(\omega'(0) + \rho\psi_{t-1}'(0)\right)\psi_{t-1}''(0)\right] \\
= \ & \rho^3 cov(\psi_t'(0), \psi_t'(0)^2) + 2\omega'(0)\rho^2 Var\left[\psi_t'(0)\right] + \rho\alpha'(0)^2 cov\left(\psi_t'(0), \psi_t''(0)\right) \\
& + \alpha'(0)^3 E\left[\varepsilon_t^3\right] + 2\alpha'(0)^2\omega'(0) E\left[\psi_t''(0)\right] + 2\alpha'(0)^2\rho E\left[\psi_t'(0)\psi_t''(0)\right] \\
= \ & \rho^3 cov(\psi_t'(0), \psi_t'(0)^2) + \alpha'(0)^3 E\left[\varepsilon_t^3\right] + 2\alpha'(0)^2\left(\omega'(0) + \rho E\left[x_t\right]\right) E\left[\psi_t''(0)\right] \\
& + 2\omega'(0)\rho^2 Var\left[\psi_t'(0)\right] + 3\rho\alpha'(0)^2 cov\left(\psi_t'(0), \psi_t''(0)\right) \\
= \ & \rho^3 cov(\psi_t'(0), \psi_t'(0)^2) + \alpha'(0)^3 E\left[\varepsilon_t^3\right] + 2\alpha'(0)^2 E\left[x_t\right] E\left[\psi_t''(0)\right] \\
& + 2\omega'(0)\rho^2 Var\left[\psi_t'(0)\right] + 3\rho\alpha'(0)^2 cov\left(\psi_t'(0), \psi_t''(0)\right).
\end{aligned}
$$

Hence,

$$
cov(\psi_t'(0), \psi_t'(0)^2) = \frac{\begin{array}{c}\alpha'(0)^3 E\left[\varepsilon_t^3\right] + 2\alpha'(0)^2 E\left[x_t\right] E\left[\psi_t''(0)\right] \\ + 2\omega'(0)\rho^2 Var\left[\psi_t'(0)\right] + 3\rho\alpha'(0)^2 cov\left(\psi_t'(0), \psi_t''(0)\right)\end{array}}{1-\rho^3}
$$

$$
\begin{aligned}
cov(\varepsilon_t, x_t^2) &= E\left[cov_{t-1}(\varepsilon_t, x_t^2)\right] = E\left[cov_{t-1}(\varepsilon_t, \varepsilon_t^2 - 2\psi_{t-1}'(0)\varepsilon_t)\right] \\
&= E\left[cov_{t-1}(\varepsilon_t, \varepsilon_t^2)\right] - 2E\left[\psi_{t-1}'(0) cov_{t-1}(\varepsilon_t, \varepsilon_t)\right] \\
&= E\left[\varepsilon_t^3\right] - 2E\left[\psi_{t-1}'(0)\psi_{t-1}''(0)\right] \\
&= E\left[\varepsilon_t^3\right] - 2cov(\psi_t'(0), \psi_t''(0)) - 2E\left[x_t\right] E\left[\psi_t''(0)\right].
\end{aligned}
$$

5.2.4 Expression of $cov(\varepsilon_{t+1}^2, x_t)$

$$
\begin{aligned}
cov(\varepsilon_{t+1}^2, x_t) &= cov(\psi_t''(0), x_t) = cov(\omega''(0) + \alpha''(0)x_t + \beta\theta^2\psi_{t-1}''(0), x_t) \\
&= \alpha''(0) cov(x_t, x_t) + \beta\theta^2 cov(\psi_{t-1}''(0), x_t) \\
&= \alpha''(0) var(x_t) + \beta\theta^2 cov(\psi_{t-1}''(0), \psi_{t-1}'(0)) \\
&= \alpha''(0) var(x_t) + \beta\theta^2 cov(\psi_t''(0), \psi_t'(0))
\end{aligned}
$$

5.2.5 Expression of $cov(x_{t+1}^2, x_t)$

$$
\begin{aligned}
& cov(x_{t+1}^2, x_t) \\
= \ & cov(\psi_t''(0) + \psi_t'(0)^2, x_t) = cov(\psi_t''(0), x_t) + cov(\psi_t'(0)^2, x_t) \\
= \ & cov(\varepsilon_{t+1}^2, x_t) + cov(2\omega'(0)\rho\psi_{t-1}'(0) + \alpha'(0)^2\psi_{t-1}''(0) + \rho^2\psi_{t-1}'(0)^2, \psi_{t-1}'(0)) \\
& + E\left[cov_{t-1}(\varepsilon_t, \alpha'(0)^2\varepsilon_t^2 + 2\alpha'(0)\left(\omega'(0) + \rho\psi_{t-1}'(0)\right)\varepsilon_t)\right] \\
= \ & cov(\varepsilon_{t+1}^2, x_t) + 2\omega'(0)\rho var(\psi_t'(0)) + \alpha'(0)^2 cov(\psi_t''(0), \psi_t'(0)) + \rho^2 cov(\psi_t'(0)^2, \psi_t'(0)) \\
& + \alpha'(0)^2 E\left[\varepsilon_t^3\right] + E\left[2\alpha'(0)\left(\omega'(0) + \rho\psi_{t-1}'(0)\right)\psi_{t-1}''(0)\right] \\
= \ & cov(\varepsilon_{t+1}^2, x_t) + 2\omega'(0)\rho var(\psi_t'(0)) + \alpha'(0)^2 cov(\psi_t''(0), \psi_t'(0)) + \rho^2 cov(\psi_t'(0)^2, \psi_t'(0)) \\
& + \alpha'(0)^2 E\left[\varepsilon_t^3\right] + 2\alpha'(0)\omega'(0) E\left[\psi_t''(0)\right] + 2\alpha'(0)\rho E\left[\psi_t'(0)\psi_t''(0)\right] \\
= \ & cov(\varepsilon_{t+1}^2, x_t) + 2\omega'(0)\rho var(\psi_t'(0)) + \left(\alpha'(0)^2 + 2\alpha'(0)\rho\right) cov(\psi_t''(0), \psi_t'(0)) \\
& + \rho^2 cov(\psi_t'(0)^2, \psi_t'(0)) + \alpha'(0)^2 E\left[\varepsilon_t^3\right] + 2\alpha'(0) E\left[x_t\right] E\left[\psi_t''(0)\right]
\end{aligned}
$$

5.2.6 Expression of $cov(\varepsilon_{t+1}^2, \varepsilon_t^2)$

$$\begin{aligned}
& cov(\varepsilon_{t+1}^2, \varepsilon_t^2) \\
= \; & cov(\psi_t''(0), \varepsilon_t^2) = cov(\omega''(0) + \beta\theta^2\psi_{t-1}''(0) + \alpha''(0)\psi_{t-1}'(0) + \alpha''(0)\varepsilon_t, \varepsilon_t^2) \\
= \; & \beta\theta^2 cov(\psi_{t-1}''(0), \varepsilon_t^2) + \alpha''(0)cov(\psi_{t-1}'(0), \varepsilon_t^2) + \alpha''(0)cov(\varepsilon_t, \varepsilon_t^2) \\
= \; & \beta\theta^2 cov(\psi_t''(0), \psi_t''(0)) + \alpha''(0)cov(\psi_t'(0), \psi_t''(0)) + \alpha''(0)E(\varepsilon_t^3) \\
= \; & \beta\theta^2 Var[\psi_t''(0)] + \alpha''(0)cov(\psi_t'(0), \psi_t''(0)) + \alpha''(0)E(\varepsilon_t^3)
\end{aligned}$$

$$\begin{aligned}
Var[\psi_t''(0)] &= Var[\beta\theta^2\psi_{t-1}''(0) + \alpha''(0)\psi_{t-1}'(0)] + \alpha''(0)^2 E[\psi_t''(0)] \\
&= \left(\beta\theta^2\right)^2 Var[\psi_{t-1}''(0)] + \alpha''(0)^2 Var[\psi_{t-1}'(0)] \\
&\quad +2\beta\theta^2\alpha''(0)cov(\psi_t'(0), \psi_t''(0)) + \alpha''(0)^2 E[\psi_t''(0)].
\end{aligned}$$

Hence,

$$Var[\psi_t''(0)] = \frac{\alpha''(0)^2 Var[\psi_t'(0)] + 2\beta\theta^2\alpha''(0)cov(\psi_t'(0), \psi_t''(0)) + \alpha''(0)^2 E[\psi_t''(0)]}{1 - \left(\beta\theta^2\right)^2}.$$

5.2.7 Expression of $cov(\varepsilon_{t+1}^2, x_t^2)$

$$\begin{aligned}
cov(\varepsilon_{t+1}^2, x_t^2) &= cov(\psi_t''(0), x_t^2) = cov(\beta\theta^2\psi_{t-1}''(0) + \alpha''(0)x_t, x_t^2) \\
&= \beta\theta^2 cov(\psi_{t-1}''(0), x_t^2) + \alpha''(0)cov(x_t, x_t^2) \\
&= \beta\theta^2 cov(\psi_{t-1}''(0), \psi_{t-1}''(0) + \psi_{t-1}'(0)^2) + \alpha''(0)cov(x_t, x_t^2) \\
&= \beta\theta^2 Var[\psi_t''(0)] + \beta\theta^2 cov(\psi_t'(0), \psi_t'(0)^2) + \alpha''(0)cov(x_t, x_t^2)
\end{aligned}$$

$$\begin{aligned}
& cov(x_t, x_t^2) \\
= \; & cov(\psi_{t-1}'(0), \psi_{t-1}''(0) + \psi_{t-1}'(0)^2) + E\left[cov_{t-1}\left(x_t, x_t^2\right)\right] \\
= \; & cov\left(\psi_t'(0), \psi_t''(0)\right) + cov\left(\psi_t'(0), \psi_t'(0)^2\right) + E\left[cov_{t-1}\left(\varepsilon_t, \varepsilon_t^2 - 2\varepsilon_t\psi_{t-1}'(0)\right)\right] \\
= \; & cov\left(\psi_t'(0), \psi_t''(0)\right) + cov\left(\psi_t'(0), \psi_t'(0)^2\right) + E(\varepsilon_t^3) - 2E\left[\psi_t'(0)\,\psi_t''(0)\right] \\
= \; & cov\left(\psi_t'(0), \psi_t'(0)^2\right) + E(\varepsilon_t^3) - 2E\left[\psi_t'(0)\right]E\left[\psi_t''(0)\right] - cov\left(\psi_t'(0), \psi_t''(0)\right)
\end{aligned}$$

5.2.8 Expression of $cov(x_{t+1}^2, \varepsilon_t^2)$

$$\begin{aligned}
& cov(x_{t+1}^2, \varepsilon_t^2) \\
= \; & cov(\psi_t''(0) + \psi_t'(0)^2, \varepsilon_t^2) = cov(\psi_t''(0), \varepsilon_t^2) + cov(\psi_t'(0)^2, \varepsilon_t^2) \\
= \; & cov(\varepsilon_{t+1}^2, \varepsilon_t^2) + cov(2\omega'(0)\rho\psi_{t-1}'(0) + \alpha'(0)^2\psi_{t-1}''(0) + \rho^2\psi_{t-1}'(0)^2, \psi_{t-1}''(0)) \\
& +E\left[cov_{t-1}(\alpha'(0)^2\varepsilon_t^2 + 2\alpha'(0)\left(\omega'(0) + \rho\psi_{t-1}'(0)\right)\varepsilon_t, \varepsilon_t^2)\right] \\
= \; & cov(\varepsilon_{t+1}^2, \varepsilon_t^2) + 2\omega'(0)\rho cov(\psi_t'(0), \psi_t''(0)) + \alpha'(0)^2 Var[\psi_t''(0)] + \rho^2 cov(\psi_t'(0)^2, \psi_t''(0))
\end{aligned}$$

5.2.9 Expression of $cov(x_{t+1}^2, x_t^2)$

$$
\begin{aligned}
& cov(x_{t+1}^2, x_t^2) \\
= \;& cov(\psi_t''(0) + \psi_t'(0)^2, x_t^2) = cov(\psi_t''(0), x_t^2) + cov(\psi_t'(0)^2, x_t^2) = cov(\varepsilon_{t+1}^2, x_t^2) + cov(\psi_t'(0)^2, x_t^2) \\
= \;& cov(\varepsilon_{t+1}^2, x_t^2) + cov(2\omega'(0)\rho\psi_{t-1}'(0) + \alpha'(0)^2\psi_{t-1}''(0) + \rho^2\psi_{t-1}'(0)^2, \psi_{t-1}''(0) + \psi_{t-1}'(0)^2) \\
& + E[cov_{t-1}\left(\alpha'(0)^2\varepsilon_t^2 + 2\alpha'(0)\left(\omega'(0) + \rho\psi_{t-1}'(0)\right)\varepsilon_t, \varepsilon_t^2 + 2\psi_{t-1}'(0)\varepsilon_t\right)] \\
= \;& cov(\varepsilon_{t+1}^2, x_t^2) + cov(2\omega'(0)\rho\psi_{t-1}'(0), \psi_{t-1}''(0) + \psi_{t-1}'(0)^2) + cov(\alpha'(0)^2\psi_{t-1}''(0), \psi_{t-1}''(0) + \psi_{t-1}'(0)^2) \\
& + cov(\rho^2\psi_{t-1}'(0)^2, \psi_{t-1}''(0) + \psi_{t-1}'(0)^2) + \alpha'(0)^2 E[cov_{t-1}\left(\varepsilon_t^2, \varepsilon_t^2 + 2\psi_{t-1}'(0)\varepsilon_t\right)] \\
& + 2\alpha'(0)E[\left(\omega'(0) + \rho\psi_{t-1}'(0)\right)cov_{t-1}\left(\varepsilon_t, \varepsilon_t^2 + 2\psi_{t-1}'(0)\varepsilon_t\right)] \\
= \;& cov(\varepsilon_{t+1}^2, x_t^2) + 2\omega'(0)\rho cov(\psi_t'(0), \psi_t''(0)) + 2\omega'(0)\rho cov(\psi_t'(0), \psi_t'(0)^2) \\
& + \alpha'(0)^2 Var[\psi_t''(0)] + \alpha'(0)^2 cov(\psi_t''(0), \psi_t'(0)^2) \\
& + \rho^2 cov(\psi_t'(0)^2, \psi_t''(0)) + \rho^2 Var\left[\psi_t'(0)^2\right] + \alpha'(0)^2 E[Var_{t-1}\left(\varepsilon_t^2\right)] + 2\alpha'(0)^2 E[\psi_t'(0)\psi_t^{(3)}(0)] \\
& + 2\alpha'(0)\omega'(0)E[\psi_t^{(3)}(0)] + 2\alpha'(0)\rho E[\psi_t'(0)\psi_t^{(3)}(0)] \\
& + 4\alpha'(0)\omega'(0)E[\psi_t'(0)\psi_t''(0)] + 4\alpha'(0)\rho E[\psi_t'(0)^2\psi_t''(0)]
\end{aligned}
$$

$$
\begin{aligned}
cov(\psi_t''(0), \psi_t'(0)^2) &= ? \\
Var\left[\psi_t'(0)^2\right] &= ? \\
E[Var_{t-1}\left(\varepsilon_t^2\right)] &= ? \\
E[\psi_t'(0)\psi_t^{(3)}(0)] &= ?
\end{aligned}
$$

$$
\begin{aligned}
& E[Var_{t-1}\left(\varepsilon_t^2\right)] \\
= \;& E[E_{t-1}\left(\varepsilon_t^4\right) - \psi_{t-1}''(0)^2] \\
= \;& E[\psi_{t-1}^{(4)}(0) + 3\psi_{t-1}''(0)^2 - \psi_{t-1}''(0)^2] \\
= \;& E[\psi_{t-1}^{(4)}(0) + 2\psi_{t-1}''(0)^2] \\
= \;& E\left[\psi_t^{(4)}(0)\right] + 2E[\psi_t''(0)^2]
\end{aligned}
$$

with

$$E[\psi_t^{(4)}(0)] = \frac{\omega^{(4)}(0) + \alpha^{(4)}(0)E[x_t]}{1 - \beta\theta^4}$$

$$Var\left[\psi_t'(0)^2\right] = Var\left[E_{t-1}\left[\psi_t'(0)^2\right]\right] + E\left[Var_{t-1}\left[\psi_t'(0)^2\right]\right]$$

$$
\begin{aligned}
& Var\left[E_{t-1}\left[\psi_t'(0)^2\right]\right] \\
= \;& Var\left[2\omega'(0)\rho\psi_{t-1}'(0) + \alpha'(0)^2\psi_{t-1}''(0) + \rho^2\psi_{t-1}'(0)^2\right] \\
= \;& 4\omega'(0)^2\rho^2 Var[\psi_t'(0)] + \alpha'(0)^4 Var[\psi_t''(0)] + \rho^4 Var\left[\psi_t'(0)^2\right] \\
& + 4\omega'(0)\rho\alpha'(0)^2 cov\left(\psi_t'(0), \psi_t''(0)\right) + 4\omega'(0)\rho^3 cov\left(\psi_t'(0), \psi_t'(0)^2\right) \\
& + 2\alpha'(0)^2\rho^2 cov(\psi_t''(0), \psi_t'(0)^2)
\end{aligned}
$$

$$
\begin{aligned}
& E\left[Var_{t-1}\left[\psi_t'(0)^2\right]\right] \\
= \; & E\left[Var_{t-1}\left[\alpha'(0)^2\varepsilon_t^2 + 2\alpha'(0)\left(\omega'(0)+\rho\psi_{t-1}'(0)\right)\varepsilon_t\right]\right] \\
= \; & E\left[\begin{array}{c} \alpha'(0)^4 Var_{t-1}\left[\varepsilon_t^2\right] + 4\alpha'(0)^2\left(\omega'(0)+\rho\psi_{t-1}'(0)\right)^2\psi_{t-1}''(0) \\ +4\alpha'(0)^3\left(\omega'(0)+\rho\psi_{t-1}'(0)\right)\psi_{t-1}^{(3)}(0) \end{array}\right] \\
= \; & \alpha'(0)^4 E\left[Var_{t-1}\left[\varepsilon_t^2\right]\right] + 4\alpha'(0)^2\omega'(0)^2 E\left[\psi_{t-1}''(0)\right] \\
& +4\alpha'(0)^2\rho^2 E\left[\psi_{t-1}'(0)^2\psi_{t-1}''(0)\right] + 8\alpha'(0)^2\rho\omega'(0)E\left[\psi_{t-1}'(0)\psi_{t-1}''(0)\right] \\
& +4\alpha'(0)^3\omega'(0)E\left[\psi_{t-1}^{(3)}(0)\right] + 4\alpha'(0)^3\rho E\left[\psi_{t-1}'(0)\psi_{t-1}^{(3)}(0)\right] \\
= \; & \alpha'(0)^4 E\left[\psi_t^{(4)}(0)\right] + 2\alpha'(0)^4 E[\psi_t''(0)^2] + 4\alpha'(0)^2\omega'(0)^2 E\left[\psi_t''(0)\right] \\
& +4\alpha'(0)^2\rho^2 E\left[\psi_t'(0)^2\psi_t''(0)\right] + 8\alpha'(0)^2\rho\omega'(0)E\left[\psi_t'(0)\psi_t''(0)\right] \\
& +4\alpha'(0)^3\omega'(0)E\left[\psi_t^{(3)}(0)\right] + 4\alpha'(0)^3\rho E\left[\psi_t'(0)\psi_t^{(3)}(0)\right]
\end{aligned}
$$

$$
\begin{aligned}
& \left(1-\rho^4\right)Var\left[\psi_t'(0)^2\right] \\
= \; & 4\omega'(0)^2\rho^2 Var[\psi_t'(0)] + \alpha'(0)^4 Var[\psi_t''(0)] + 12\alpha'(0)^2\rho\omega'(0)cov\left(\psi_t'(0),\psi_t''(0)\right) \\
& +\alpha'(0)^4 E\left[\psi_t^{(4)}(0)\right] + 2\alpha'(0)^4 E[\psi_t''(0)^2] + 4\omega'(0)\rho^3 cov\left(\psi_t'(0),\psi_t'(0)^2\right) \\
& +4\alpha'(0)^2\left(\omega'(0)(1+\rho)E\left[\psi_t'(0)\right] + \rho^2 E\left[\psi_t'(0)^2\right]\right)E\left[\psi_t''(0)\right] + 4\alpha'(0)^3\omega'(0)E\left[\psi_t^{(3)}(0)\right] \\
& +4\alpha'(0)^3\rho E\left[\psi_t'(0)\psi_t^{(3)}(0)\right] + 6\alpha'(0)^2\rho^2 cov(\psi_t''(0),\psi_t'(0)^2)
\end{aligned}
$$

$$
\begin{aligned}
& cov(\psi_t''(0),\psi_t'(0)^2) \\
= \; & cov(\alpha''(0)\psi_{t-1}'(0) + \beta\theta^2\psi_{t-1}''(0), 2\omega'(0)\rho\psi_{t-1}'(0) + \alpha'(0)^2\psi_{t-1}''(0) + \rho^2\psi_{t-1}'(0)^2) \\
& +E\left[cov_{t-1}(\alpha''(0)\varepsilon_t, \alpha'(0)^2\varepsilon_t^2 + 2\alpha'(0)\left(\omega'(0)+\rho\psi_{t-1}'(0)\right)\varepsilon_t)\right] \\
= \; & 2\omega'(0)\rho\alpha''(0)Var\left[\psi_t'(0)\right] + \left(\alpha'(0)^2\alpha''(0) + 2\omega'(0)\rho\beta\theta^2\right)cov(\psi_t'(0),\psi_t''(0)) \\
& +\alpha''(0)\rho^2 cov(\psi_t'(0),\psi_t'(0)^2) + \alpha'(0)^2\beta\theta^2 Var[\psi_t''(0)] + \rho^2\beta\theta^2 cov(\psi_t''(0),\psi_t'(0)^2) \\
& +\alpha'(0)^2\alpha''(0)E\left[\psi_t^{(3)}(0)\right] + 2\alpha'(0)\alpha''(0)E\left[\psi_t'(0)\right]E\left[\psi_t''(0)\right] + 2\rho\alpha'(0)\alpha''(0)cov(\psi_t'(0),\psi_t''(0)).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
& cov(\psi_t''(0),\psi_t'(0)^2) \\
= \; & \frac{\begin{array}{c} 2\omega'(0)\rho\alpha''(0)Var\left[\psi_t'(0)\right] + \left(\alpha'(0)^2\alpha''(0) + 2\omega'(0)\rho\beta\theta^2 + 2\rho\alpha'(0)\alpha''(0)\right)cov(\psi_t'(0),\psi_t''(0)) \\ +\alpha''(0)\rho^2 cov(\psi_t'(0),\psi_t'(0)^2) + \alpha'(0)^2\beta\theta^2 Var[\psi_t''(0)] \\ +\alpha'(0)^2\alpha''(0)E\left[\psi_t^{(3)}(0)\right] + 2\alpha'(0)\alpha''(0)E\left[\psi_t'(0)\right]E\left[\psi_t''(0)\right] \end{array}}{1-\rho^2\beta\theta^2}
\end{aligned}
$$

$$
\begin{aligned}
& cov\left(\psi_t'(0),\psi_t^{(3)}(0)\right) \\
= \; & cov\left(\rho\psi_{t-1}'(0), \alpha^{(3)}(0)\psi_{t-1}'(0) + \beta\theta^3\psi_{t-1}^{(3)}(0)\right) + E\left[cov_{t-1}(\alpha'(0)\varepsilon_t, \alpha^{(3)}(0)\varepsilon_t)\right] \\
= \; & \rho\alpha^{(3)}(0)Var\left[\psi_t'(0)\right] + \rho\beta\theta^3 Cov\left(\psi_t'(0),\psi_t^{(3)}(0)\right) + \alpha'(0)\alpha^{(3)}(0)E\left[\psi_t''(0)\right].
\end{aligned}
$$

Hence,

$$
cov\left(\psi_t'(0),\psi_t^{(3)}(0)\right) = \frac{\rho\alpha^{(3)}(0)Var\left[\psi_t'(0)\right] + \alpha'(0)\alpha^{(3)}(0)E\left[\psi_t''(0)\right]}{1-\rho\beta\theta^3}
$$

5.2.10 Expressions of $Var\left[\varepsilon_t^2\right]$ and $Var\left[x_t^2\right]$

We have

$$\begin{aligned} Var\left[\varepsilon_t^2\right] &= E[Var_{t-1}\left(\varepsilon_t^2\right)] + Var\left[\psi_t''\left(0\right)\right] = E\left[\psi_t^{(4)}\left(0\right)\right] + 2E[\psi_t''\left(0\right)^2] + Var\left[\psi_t''\left(0\right)\right] \\ &= E\left[\psi_t^{(4)}\left(0\right)\right] + 2E[\psi_t''\left(0\right)]^2 + 3Var\left[\psi_t''\left(0\right)\right] \end{aligned}$$

$$\begin{aligned} Var\left[x_t^2\right] &= E[Var_{t-1}\left(x_t^2\right)] + Var\left[\psi_t''\left(0\right) + \psi_t'\left(0\right)^2\right] \\ &= E[Var_{t-1}\left(\varepsilon_t^2 + 2\psi_{t-1}'\left(0\right)\varepsilon_t\right)] + Var\left[\psi_t''\left(0\right)\right] + Var\left[\psi_t'\left(0\right)^2\right] + 2cov\left(\psi_t''\left(0\right), \psi_t'\left(0\right)^2\right) \\ &= E[Var_{t-1}\left(\varepsilon_t^2\right)] + 4E[\psi_{t-1}'\left(0\right)^2 Var_{t-1}\left(\varepsilon_t\right)] + 4E\left[\psi_{t-1}'\left(0\right)cov_{t-1}\left(\varepsilon_t^2, \varepsilon_t\right)\right] \\ &\quad + Var\left[\psi_t''\left(0\right)\right] + Var\left[\psi_t'\left(0\right)^2\right] + 2cov\left(\psi_t''\left(0\right), \psi_t'\left(0\right)^2\right) \\ &= E[Var_{t-1}\left(\varepsilon_t^2\right)] + 4E[\psi_t'\left(0\right)^2\psi_t''\left(0\right)] + 4E\left[\psi_t'\left(0\right)\psi_t^{(3)}\left(0\right)\right] \\ &\quad + Var\left[\psi_t''\left(0\right)\right] + Var\left[\psi_t'\left(0\right)^2\right] + 2cov\left(\psi_t''\left(0\right), \psi_t'\left(0\right)^2\right). \end{aligned}$$

6 Parameters identification and estimation

6.1 Identification

Our approach for identification consists of expressing unknown parameters as functions of quantities, which can be directly estimated. Using the expressions for the unconditional expectation and variance, we have

$$\begin{aligned} \mu \equiv E\left[x_t\right] &= \frac{\nu\varphi}{1-\rho} \\ \sigma^2 \equiv Var\left[x_t\right] &= \nu\varphi^2\left(\frac{1-\rho^2+\phi^2}{1-\rho^2}\right)\frac{1-\rho+2\phi}{\left(1-\theta(\rho-\phi)\right)\left(1-\rho\right)}. \end{aligned}$$

We deduce parameters φ and ν as follows:

$$\begin{aligned} \varphi &= \frac{\sigma^2}{\mu}\frac{\left(1-\theta(\rho-\phi)\right)}{1-\rho+2\phi}\frac{1-\rho^2}{1-\rho^2+\phi^2} \\ \nu &= \frac{\left(1-\rho\right)\mu}{\varphi}. \end{aligned}$$

In addition, using analytical expressions of autocorrelation functions, we have

$$\begin{aligned} \rho_1 &\equiv Corr\left(x_t, x_{t+1}\right) = \phi\frac{1-\left(\rho-\phi\right)^2-\left(\rho-\phi\right)\phi}{1-\left(\rho-\phi\right)^2-2\left(\rho-\phi\right)\phi} \\ \rho_k &\equiv Corr\left(x_t, x_{t+k}\right) = \rho^{k-1}\rho_1. \end{aligned}$$

It is apparent that:

$$\rho = \frac{\rho_2}{\rho_1}.$$

This enables us to establish that ϕ is the solution of the following second-order equation:

$$\left(\rho_1-\rho\right)\phi^2 - \left(1-\rho^2\right)\phi + \rho_1\left(1-\rho^2\right) = 0,$$

which always has a solution $\phi < 1-\rho$. Given that we can solve for ϕ and ρ, it implies that we also have solved for $\beta\theta$. Indeed $\beta\theta = \rho - \phi$. To identify β and θ, we use the skewness. Let us denote the unconditional skewness by μ_3, and

$$\begin{aligned}
\bar{\mu}_3 &\equiv \frac{\mu\left(1-\rho+2\phi\right)^2\left(1-\rho^2+\phi^2\right)^2}{2\sigma\left(1-\rho^2\right)^2}\mu_3 \\
\eta &\equiv \left(1-\rho+3\phi\right)\left(1+\frac{\phi^3}{1-\rho^3}\right) \\
\chi &\equiv 3\left(1+\frac{\rho\phi^2}{1-\rho^3}\right)\frac{\left(1-\rho\left(\rho-\phi\right)\right)\left(1-\rho+2\phi\right)\phi^2}{\left(1-\rho^2\right)}.
\end{aligned}$$

Using the analytical derivations of the unconditional skewness, we show that θ is the solution for the following third-order polynomial equation:

$$d\theta^3 + a\theta^2 + b\theta + c = 0,$$

where

$$\begin{aligned}
c &\equiv \eta + \chi - \bar{\mu}_3, \\
b &\equiv \left(\bar{\mu}_3\rho - \chi - \eta\left(\rho+2\right)\right)\left(\rho-\phi\right) \\
a &\equiv \left(\bar{\mu}_3 - \chi + \eta\left(1+2\rho\right)\left(\rho-\phi\right)\right)\left(\rho-\phi\right) \\
d &\equiv \left(\chi - \bar{\mu}_3\rho - \eta\rho(\rho-\phi)\right)\left(\rho-\phi\right)^2,
\end{aligned}$$

which enables us to solve for θ and deduce β as $\beta = \frac{(\rho-\phi)}{\theta}$. In conclusion, all the 5 parameters of the GARG dynamics are recovered from the unconditional mean, variance, skewness and the first two autocorrelations. Thus, the structural parameters are identified.

6.2 Empirical characteristic function estimators

Because GARG processes have closed-form conditional characteristic functions, one alternative to the ML-based method is the empirical characteristic function (ECF) estimation method (see Knight and Yu, 2002 for details on the ECF' theory). Let us denote $\lambda_0 = (\phi, \nu, \varphi, \beta, \theta)$, the unknown parameter to estimate. We can write the GARG dynamic as follows:

$$E\left[\exp\left(ux_{t+1}\right) - \exp\left(\psi_t(u;\lambda_0)\right)|X_t\right], \tag{24}$$

where $X_t = (x_1, ..., x_t)'$ and

$$\psi_t\left(u\right) = \beta^t\psi_0\left(\theta^t u\right) + \sum_{j=0}^{t-1}\left[\beta^j\omega(\theta^j u) + \beta^j\alpha(\theta^j u)x_{t-j}\right]. \tag{25}$$

This implies that for any weighting function, $W(X_t, v)$ (often termed instruments in the GMM literature), $\forall u, v$, we have:

$$E\left[\left(\exp\left(ux_{t+1}\right) - \exp\left(\psi_t(u;\lambda_0)\right)\right)W(X_t, v)\right]. \tag{26}$$

This leads to continuum of moments restrictions. Hence, we can estimate the model by applying the GMM to the continuum of moments restrictions of Carrasco and Florens (2000). The ML efficiency is achieved by choosing the Carrasco and Florens (2002) weighting function, i.e., $W(X_t, v) = \exp(X_t'v)$. However, there are also difficulties with the implementation of the ECF approach. The biggest challenge is the need to use a large set of moment conditions, leading to the singularity of the covariance matrix. Carrasco et al. (2007) have proposed solutions to these difficulties in the framework of GMM with a continuum of moment conditions.

6.3 Generalized method of moments

One of the advantages of the discrete-time affine models is that we can compute unconditional moments of any component of the process of interest. This point have been studied in detail in Feunou and Tédongap (2012) and is an important result for estimation purposes because even when there are unobserved components in the process of interest, we can still compute the moments of observed components and use them to implement a GMM estimation routine. It turns out that we keep this advantage in the GARG. In the case of an observable variable of interest, there is no need to compute the unconditional moments. We can use the conditional moments equations derived in the paper.

The following moment conditions have been used by Bollerslev and Zhou (2002) to estimate one-factor and two-factors stochastic volatility models by means of conditional moments of realized variance:

$$
\begin{gathered}
E\left[x_{t+1}-\mu_{1,0}\right]=0\ ,\ \ E\left[x_{t+1}^2-\mu_{2,0}\right]=0,\\
E\left[\left(x_{t+1}-\mu_{1,t}\right)x_t\right]=0\ ,\ \ E\left[\left(x_{t+1}^2-\mu_{2,t}\right)x_t\right]=0,\\
E\left[\left(x_{t+1}-\mu_{1,t}\right)x_t^2\right]=0\ ,\ \ E\left[\left(x_{t+1}^2-\mu_{2,t}\right)x_t^2\right]=0,\\
E\left[\left(x_{t+1}-\mu_{1,t}\right)x_{t-1}\right]=0\ ,\ \ E\left[\left(x_{t+1}^2-\mu_{2,t}\right)x_{t-1}\right]=0,\\
E\left[\left(x_{t+1}-\mu_{1,t}\right)x_{t-1}^2\right]=0\ ,\ \ E\left[\left(x_{t+1}^2-\mu_{2,t}\right)x_{t-1}^2\right]=0,
\end{gathered}
$$

where $\mu_{1,0}=E(x_{t+1})$, $\mu_{2,0}=E(x_{t+1}^2)$, $\mu_{1,t}=E_t\left[x_{t+1}\right]=\psi_t^{\top}(0)$ and $\mu_{2,t}=E_t\left[x_{t+1}^2\right]=\psi_t^{\top\top}(0)+\left(\psi_t^{\top}(0)\right)^2$. We simulate the GARG model with $\varphi=2.784E{-}05$, $\nu=0.1394$, $\phi=0.1125$, $\beta=0.9227$ and $\theta=0.9066$. For different sample sizes (T) and number of replications (N), we estimate the GARG model and report in Table 1 the statistics (mean, median and root mean square errors (RMSE)) across different replication sizes. The GMM does well when we consider the longest sample size (1000) and the biggest number of replications (4000).

6.4 Filtration-based maximum likelihood estimators

Some applications in the literature on stochastic volatility and term structures of interest rates require latent factors. Let y_t denote a vector of variables observed at discrete dates indexed by t. Let x_t represent a latent state variable affecting the dynamics of y_t. We assume that x_t follows a GARG dynamic while the moment generating function of y_t conditional on x_t is an exponentially affine function of the latent variables x_t, that is,

$$
\begin{aligned}
E\left[\exp\left(v'y_t\right)|x_t\right] &= \exp\left(c_y(v)+d_y(v)x_t\right) &(27)\\
x_t &\sim GARG(\phi,\nu,\varphi,\beta,\theta). &(28)
\end{aligned}
$$

To estimate this latent version of the model, we adapt the direct filtration-based maximum likelihood methodology of Bates (2006). Let $Y_t\equiv\{y_1,...,y_t\}$ denote the data observed by the econometrician up to date t and $G_{t|t}(u)\equiv E\left[e^{ux_t}|Y_t\right]$ the moment generating function of the latent variable x_t at time t conditional upon observing Y_t. The filtered moment generating function $G_{t|t}(u)$ can be recursively updated as follows:

- **Step 1**: Given past filtered values $x_{1|1},\cdots,x_{t-1|t-1}$ and $G_{t|t}(u)$, the joint characteristic function of the next period's (y_{t+1},x_{t+1}) conditional on data observed through date t can be evaluated by iterated expectations

$$
\begin{aligned}
F\left(v,u|Y_t\right) &\equiv E\left[\exp\left(v'y_{t+1}+ux_{t+1}\right)|Y_t\right]=E\left[\exp\left(c_y\left(v\right)+\left(u+d_y\left(v\right)\right)x_{t+1}\right)|Y_t\right]\\
&= E\left[\exp\left(c_y\left(v\right)+\psi_t\left(u+d_y\left(v\right)\right)\right)|Y_t\right]\\
&= E\left[\exp\left(c_y\left(v\right)+\omega\left(u+d_y\left(v\right)\right)+\beta\psi_{t-1|t-1}\left(\theta\left(u+d_y\left(v\right)\right)\right)+\alpha\left(u+d_y\left(v\right)\right)x_t\right)|Y_t\right]\\
&= \exp\left(c_y\left(v\right)+\omega\left(u+d_y\left(v\right)\right)+\beta\psi_{t-1|t-1}\left(\theta\left(u+d_y\left(v\right)\right)\right)\right)G_{t|t}\left(\alpha\left(u+d_y\left(v\right)\right)\right),
\end{aligned}
$$

  where $\psi_{t-1|t-1}\left(\cdot\right)$ is computed recursively as in equation (2) using the filtered values of x_t: $x_{t|t}=E[x_t|Y_t]=G'_{t|t}(0)$.

- **Step 2**: The density function of the next period's datum y_{t+1} conditional upon data observed through date t can be evaluated by the Fourier inversion of its conditional characteristic function:

$$
p\left(y_{t+1}|Y_t\right)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}F\left(iv,0|Y_t\right)e^{-ivy_{t+1}}dv.
$$

- **Step 3**: The conditional characteristic function of the next period's x_{t+1} is

$$G_{t+1|t+1}(u) = \frac{\frac{1}{2\pi}\int_{-\infty}^{+\infty} F\left(iv, u|Y_t\right) e^{-ivy_{t+1}} dv}{p\left(y_{t+1}|Y_t\right)}.$$

- **Step 4**: Repeat steps 1–3 for subsequent values of t. Given the underlying parameters, the log likelihood function for the maximum likelihood estimation is $\ln L(Y_t) = \sum_{t=1}^{T} \ln p\left(y_t|Y_{t-1}\right).$

7 Empirical Results

7.1 Results from the physical dynamic estimation

To further shed light on these results, we plot sample autocorrelations and cross-correlations in Figure 2. As expected, for the ARG processes these correlations are almost identical to the autocorrelation. This is clearly at odds with the observed facts (last row of Figure 2). The GARG is able to disentangle those dynamics.

7.2 Option pricing under the GARG dynamic

The risk-neutral dynamic is given by

$$\begin{aligned} R_{t+1} &= \ln\left(S_{t+1}/S_t\right) = r - \frac{1}{2}RV_{t+1} + \sqrt{RV_{t+1}}\varepsilon_{t+1} && (29) \\ \varepsilon_{t+1} &\sim i.i.d N(0,1) \\ RV_{t+1} &\sim GARG\left(\phi, \varphi, \nu, \beta, \theta\right), && (30) \end{aligned}$$

where r is the risk-free rate. Let us denote the conditional cumulant generating function of the aggregate returns between $t+1$ and $t+\tau$ by $\psi_{t,\tau}^{R}(u)$. This means that

$$\psi_{t,\tau}^{R}(u) \equiv \ln\left(E_t\left[\exp\left(u\sum_{j=1}^{\tau} R_{t+j}\right)\right]\right).$$

The time t call option values with strike X and maturity τ, denoted by $C_t^{Mod}(X,\tau)$, is the discounted risk-neutral expectation of the terminal payoff $\max(S_{t+\tau} - X, 0)$,

$$C_t^{Mod}(X,\tau) \equiv E_t\left[\exp(-r\tau)\max(S_{t+\tau} - X, 0)\right],$$

and computed via standard Fourier inversion techniques:

$$\begin{aligned} C_t^{Mod}(X,\tau) &= S_t P_1(t,\tau) - \exp(-r\tau) X P_2(t,\tau), \text{ where} && (31) \\ P_1(t,\tau) &= \frac{1}{2} + \int_0^{\infty} Re\left(\frac{\exp\left(\psi_{t,\tau}^{R}(1+iu) - r\tau - iu\ln\left(\frac{X}{S_t}\right)\right)}{iu\pi}\right) du \\ P_2(t,\tau) &= \frac{1}{2} + \int_0^{\infty} Re\left(\frac{\exp\left(\psi_{t,\tau}^{R}(iu) - iu\ln\left(\frac{X}{S_t}\right)\right)}{iu\pi}\right) du, \end{aligned}$$

and i stands for the imaginary unit. Let us denote the conditional cumulant generating function of the aggregate variance between $t+1$ and $t+\tau$ by $\psi_{t,\tau}^{RV}(u)$. This means that

$$\psi_{t,\tau}^{RV}(u) \equiv \ln\left(E_t\left[\exp\left(u\sum_{j=1}^{\tau} RV_{t+j}\right)\right]\right),$$

and straight iterative expectation manipulations imply that

$$\psi_{t,\tau}^{R}(u) = ur\tau + \psi_{t,\tau}^{RV}\left(\frac{1}{2}(u^2 - u)\right). \quad (32)$$

Following the exact same steps used in deriving the multi-horizon dynamic, we establish that

$$\begin{aligned}
\psi_{t,h}^{RV}(u) &= b_h(u) + a_h(u)\, RV_t + \left(\sum_{j=1}^{h} \beta^j \psi_{t-1}\left(\theta^j v_j\right)\right) \\
v_h &= u, \quad v_\tau = u + \sum_{j=\tau+1}^{h} \beta^{j-(\tau+1)} \alpha\left(\theta^{j-(\tau+1)} v_j\right), \quad for \;\; \tau < h, \\
b_h(u) &= \sum_{j=0}^{h-1}\left(\sum_{i=0}^{j} \beta^i \omega\left(\theta^i v_{j+1}\right)\right), \quad a_h(u) = \sum_{j=0}^{h-1} \beta^j \alpha\left(\theta^j v_{j+1}\right).
\end{aligned}$$

We discuss in detail how to price options under the MARG dynamic in Section 8.2 of this appendix.

7.3 IVRMSE and risk-neutral model estimation

The IVRMSE synthesizes the discrepancy between model-based and market-based implied volatilities. To compute the IVRMSE, we invert the model-based option price C_j^{Mod} of each contract j using the Black-Scholes formula (BS). Thus, the model-based implied volatility can be formally extracted according to

$$IV_j^{Mod} = BS^{-1}\left(C_j^{Mod}\right).$$

Applying a similar procedure to the set of observed option contracts $\{C_j^{Mkt}\}$ yields, market-based implied volatilities are defined as

$$IV_j^{Mkt} = BS^{-1}\left(C_j^{Mkt}\right).$$

Accordingly, the implied volatility error is computed as

$$e_j = IV_j^{Mkt} - IV_j^{Mod}.$$

It follows that the IVRMSE is given by

$$IVRMSE \equiv \sqrt{\frac{1}{N}\sum_{j=1}^{N} e_j^2},$$

where N denotes the options sample size. Finally, risk-neutral parameters are estimated by maximizing the Gaussian-implied volatility error likelihood:

$$\ln L^O = -\frac{1}{2}\sum_{j=1}^{N}\left\{\ln\left(IVRMSE^2\right) + e_j^2/IVRMSE^2\right\}. \tag{33}$$

7.4 Exploring options data

We use European-style options written on the S&P 500 index. The observations span the period January 10, 1996, to August 28, 2013.[4] In line with the literature, we only include out-of-the-money (OTM) options with maturities ranging from 15 to 180 days. This selection procedure is intended to guarantee that the contracts we use are sufficiently liquid. We also filter out options that violate basic no-arbitrage criteria. For each maturity quoted on Wednesdays, we select only the six most liquid strike prices, which amounts to a data set of 21,283 option contracts. To ease calculation and interpretation, OTM put prices are converted into corresponding in-the-money call values by exploiting the call-put parity relationship.

Table 3 provides a concise description of the options data. To highlight the main characteristics of S&P 500 index options, we sort the data by moneyness, maturity, and market volatility index (VIX) level. Panel A of Table 3 groups the data by six moneyness buckets and shows the number of contracts, the average option price, the average Black and Scholes (1973) implied

[4]Data are available through OptionMetrics, which supplies data for the U.S. option markets.

volatility, and the average bid-ask spread in dollars. Our measure of moneyness is based on the Black-Scholes delta computed as

$$\Phi\left(\frac{\ln\left(S_t/X\right)+r_f M+1/2\left(IV^{Mkt}\right)^2 M/365}{IV^{Mkt}\sqrt{M/365}}\right),$$

where $\Phi(*)$ stands for the normal cumulative distribution function (CDF), X is the strike price, r_f is the non-annualized daily risk-free rate, M is the time-to-maturity expressed in days, and IV^{Mkt} denotes the annualized implied Black-Scholes volatility computed at the market price of the option. A few empirical regularities emerge at this point. We observe that deep OTM puts, which include the largest number of contracts with deltas exceeding 0.7, are the most expensive. This echoes the well-documented volatility smirk pattern in index options across moneyness.

Panel B of Table 3 sorts the data by maturity expressed in calendar days. Even though the term structure of volatility is nearly flat on average during the sample period, we notice that options with longer maturities are relatively more expensive. Panel C of Table 3 categorizes the data by the VIX level. It is immediately obvious that a large portion of the selected option contracts (75%) are quoted on days with VIX levels ranging between 15% and 35%. Overall, a typical "median" contract features a delta above 0.6 and a time-to-expiry between 30 and 90 days, and is quoted on "normal" days when the VIX lies within the $[15-25]\,\%$ interval.

8 Multi-factor affine model

x_t follows a multi-factor affine model, or MARG, if

$$x_t = x_{1t} + x_{2t},$$

where x_{1t} and x_{2t} are two independent ARG dynamics.

$$\begin{aligned} E_t\left[\exp\left(u x_{1t+1}\right)\right] &= \exp\left(\omega_1(u) + \alpha_1(u)\, x_{1t}\right) \\ E_t\left[\exp\left(u x_{2t+1}\right)\right] &= \exp\left(\omega_2(u) + \alpha_2(u)\, x_{2t}\right) \end{aligned}$$

$$\omega_j(u) = -\nu_j \log(1-u\varphi_j), \text{ and } \alpha_j(u) = \frac{\phi_j u}{1-u\varphi_j},$$

$$\begin{aligned} E_t\left[\exp\left(u x_{t+1}\right)\right] &= \exp\left(\omega_1(u)+\omega_2(u)+\alpha_2(u)\, x_{2t}+\alpha_1(u)\, x_{1t}\right) \\ &= \exp\left(\omega_1(u)+\omega_2(u)+\alpha_2(u)\left(x_t-x_{1t}\right)+\alpha_1(u)\, x_{1t}\right) \\ &= \exp\left(\omega_1(u)+\omega_2(u)+\alpha_2(u)\, x_t+\left(\alpha_1(u)-\alpha_2(u)\right) x_{1t}\right) \end{aligned}$$

$$\begin{aligned} E_t\left[\exp\left(u x_{t+1}+v x_{1t+1}\right)\right] &= E_t\left[\exp\left(\left(u+v\right) x_{1t+1}+u x_{2t+1}\right)\right] \\ &= \exp\left(\omega_2(u)+\alpha_2(u)\, x_{2t}+\omega_1(u+v)+\alpha_1(u+v)\, x_{1t}\right) \\ &= \exp\left(\omega_2(u)+\omega_1(u+v)+\alpha_2(u)\, x_t+\left(\alpha_1(u+v)-\alpha_2(u)\right) x_{1t}\right) \end{aligned}$$

$$\begin{aligned} \exp\left(\psi_t(u)\right) &\equiv E\left[\exp\left(u x_{t+1}\right)|x_s, s\le t\right] \\ \exp\left(g_{t|t}(u)\right) &\equiv E\left[\exp\left(u x_{1t}\right)|x_s, s\le t\right] \end{aligned}$$

$$\begin{aligned} \exp\left(\psi_t(u)\right) &= E\left[\exp\left(u x_{t+1}\right)|x_s, s\le t\right] = E\left[E_t\left[\exp\left(u x_{t+1}\right)\right]|x_s, s\le t\right] \\ &= E\left[\exp\left(\omega_1(u)+\omega_2(u)+\alpha_2(u)\, x_t+\left(\alpha_1(u)-\alpha_2(u)\right) x_{1t}\right)|x_s, s\le t\right] \\ &= \exp\left(\omega_1(u)+\omega_2(u)+\alpha_2(u)\, x_t\right) E\left[\exp\left(\left(\alpha_1(u)-\alpha_2(u)\right) x_{1t}\right)|x_s, s\le t\right] \\ &= \exp\left(\omega_1(u)+\omega_2(u)+\alpha_2(u)\, x_t+g_{t|t}\left(\alpha_1(u)-\alpha_2(u)\right)\right). \end{aligned}$$

Hence,

$$\psi_t(u) = \omega_1(u) + \omega_2(u) + \alpha_2(u) x_t + g_{t|t}(\alpha_1(u) - \alpha_2(u)),$$

which is equivalent to

$$g_{t|t}(\alpha_1(u) - \alpha_2(u)) = \psi_t(u) - \omega_1(u) - \omega_2(u) - \alpha_2(u) x_t.$$

Let us denote

$$\alpha_0(u) \equiv \alpha_1(u) - \alpha_2(u).$$

Hence,

$$v = \alpha_0(u) \Longleftrightarrow u = \alpha_0^{-1}(v)$$

and

$$g_{t|t}(v) = \psi_t\left(\alpha_0^{-1}(v)\right) - \omega_1\left(\alpha_0^{-1}(v)\right) - \omega_2\left(\alpha_0^{-1}(v)\right) - \alpha_2\left(\alpha_0^{-1}(v)\right) x_t$$

$$\begin{aligned}
& E\left[\exp\left(u x_{t+1} + v x_{1t+1}\right) | x_s, s \leq t\right] \\
= \; & E\left[E_t\left[\exp\left(u x_{t+1} + v x_{1t+1}\right)\right] | x_s, s \leq t\right] \\
= \; & E\left[\exp\left(\omega_2(u) + \omega_1(u+v) + \alpha_2(u) x_t + \left(\alpha_1(u+v) - \alpha_2(u)\right) x_{1t}\right) | x_s, s \leq t\right] \\
= \; & \exp\left(\omega_2(u) + \omega_1(u+v) + \alpha_2(u) x_t\right) E\left[\exp\left(\left(\alpha_1(u+v) - \alpha_2(u)\right) x_{1t}\right) | x_s, s \leq t\right] \\
= \; & \exp\left(\omega_2(u) + \omega_1(u+v) + \alpha_2(u) x_t + g_{t|t}\left(\alpha_1(u+v) - \alpha_2(u)\right)\right)
\end{aligned}$$

$$\begin{aligned}
& \exp\left(g_{t+1|t+1}(v)\right) \\
\equiv \; & E\left[\exp\left(v x_{1t+1}\right) | x_s, s \leq t+1\right] \\
= \; & \frac{\frac{1}{2\pi} \int E\left[\exp\left(iu x_{t+1} + v x_{1t+1}\right) | x_s, s \leq t\right] \exp\left(-iu x_{t+1}\right) du}{p_t(x_{t+1})} \\
= \; & \frac{\frac{1}{2\pi} \int \exp\left(\omega_2(iu) + \omega_1(iu+v) + \alpha_2(iu) x_t + g_{t|t}\left(\alpha_1(iu+v) - \alpha_2(iu)\right) - iu x_{t+1}\right) du}{p_t(x_{t+1})}
\end{aligned}$$

where

$$p_t(x_{t+1}) = \frac{1}{2\pi} \int \exp\left(\psi_t(iu) - iu x_{t+1}\right) du.$$

Hence,

$$\begin{aligned}
g_{t+1|t+1}(v) = \; & \ln\left[\int \exp\left(\omega_2(iu) + \omega_1(iu+v) + \alpha_2(iu) x_t + g_{t|t}\left(\alpha_1(iu+v) - \alpha_2(iu)\right) - iu x_{t+1}\right) du\right] \\
& - \ln\left[\int \exp\left(\psi_t(iu) - iu x_{t+1}\right) du\right]
\end{aligned}$$

$$\begin{aligned}
\psi_{t+1}(u) = \; & \omega_1(u) + \omega_2(u) + \alpha_2(u) x_{t+1} + g_{t+1|t+1}\left(\alpha_1(u) - \alpha_2(u)\right) \\
= \; & \omega_1(u) + \omega_2(u) + \alpha_2(u) x_{t+1} + g_{t+1|t+1}\left(\alpha_0(u)\right) \\
= \; & \omega_1(u) + \omega_2(u) + \alpha_2(u) x_{t+1} \\
& + \ln\left[\int \exp\left(\omega_2(iy) + \omega_1(iy + \alpha_0(u)) + \alpha_2(iy) x_t + g_{t|t}\left(\alpha_1(iy + \alpha_0(u)) - \alpha_2(iy)\right) - iy x_{t+1}\right) dy\right] \\
& - \ln\left[\int \exp\left(\psi_t(iy) - iy x_{t+1}\right) dy\right]
\end{aligned}$$

$$\begin{aligned}
\psi_{t+1}(u) = \; & \omega_1(u) + \omega_2(u) + \alpha_2(u) x_{t+1} - \ln\left[\int \exp\left(\psi_t(iy) - iy x_{t+1}\right) dy\right] \\
& + \ln\left[\int \exp\begin{pmatrix} \omega_2(iy) - \omega_2\left(\alpha_0^{-1}\left(\alpha_1(iy + \alpha_0(u)) - \alpha_2(iy)\right)\right) + \omega_1(iy + \alpha_0(u)) \\ -\omega_1\left(\alpha_0^{-1}\left(\alpha_1(iy + \alpha_0(u)) - \alpha_2(iy)\right)\right) \\ +\alpha_2(iy) x_t - \alpha_2\left(\alpha_0^{-1}\left(\alpha_1(iy + \alpha_0(u)) - \alpha_2(iy)\right)\right) x_t \\ +\psi_t\left(\alpha_0^{-1}\left(\alpha_1(iy + \alpha_0(u)) - \alpha_2(iy)\right)\right) - iy x_{t+1} \end{pmatrix} dy\right]
\end{aligned}$$

$$\psi_{t+1}(u) = \omega_1(u) + \omega_2(u) + \alpha_2(u)x_{t+1} - \ln\left[\int \exp\left(\psi_t(iy) - iyx_{t+1}\right)dy\right]$$
$$+\ln\left[\int \exp\left(\tilde{\omega}(iy,u) + \tilde{\alpha}(iy,u)x_t + \psi_t\left(\tilde{\theta}(iy,u)\right) - iyx_{t+1}\right)dy\right],$$

where

$$\tilde{\theta}(y,u) \equiv \alpha_0^{-1}\left(\alpha_1(y+\alpha_0(u)) - \alpha_2(y)\right)$$
$$\tilde{\omega}(y,u) \equiv \omega_2(y) - \omega_2\left(\tilde{\theta}(y,u)\right) + \omega_1(y+\alpha_0(u)) - \omega_1\left(\tilde{\theta}(y,u)\right)$$
$$\tilde{\alpha}(y,u) \equiv \alpha_2(y) - \alpha_2\left(\tilde{\theta}(y,u)\right).$$

In the sequel, we denote

$$q_t(y,u;x_{t+1}) \equiv \exp\left(\tilde{\omega}(iy,u) + \tilde{\alpha}(iy,u)x_t + \psi_t\left(\tilde{\theta}(iy,u)\right) - iyx_{t+1}\right).$$

8.1 Derivatives

8.1.1 First order

$$\frac{\partial\tilde{\theta}(y,u)}{\partial u} = \frac{\alpha_1'(y+\alpha_0(u))\,\alpha_0'(u)}{\alpha_0'\left(\alpha_0^{-1}\left(\alpha_1(y+\alpha_0(u)) - \alpha_2(y)\right)\right)} \iff \frac{\partial\tilde{\theta}(y,0)}{\partial u} = \frac{\alpha_1'(y)\,\alpha_0'(0)}{\alpha_0'(y)}$$

$$\frac{\partial\tilde{\omega}(y,u)}{\partial u} = -\omega_2'\left(\tilde{\theta}(y,u)\right)\frac{\partial\tilde{\theta}(y,u)}{\partial u} + \omega_1'(y+\alpha_0(u))\,\alpha_0'(u) - \omega_1'\left(\tilde{\theta}(y,u)\right)\frac{\partial\tilde{\theta}(y,u)}{\partial u}$$
$$\frac{\partial\tilde{\omega}(y,0)}{\partial u} = -\left(\omega_2'(y) + \omega_1'(y)\right)\frac{\partial\tilde{\theta}(y,0)}{\partial u} + \omega_1'(y)\,\alpha_0'(0)$$

$$\frac{\partial\tilde{\alpha}(y,u)}{\partial u} \equiv -\alpha_2'\left(\tilde{\theta}(y,u)\right)\frac{\partial\tilde{\theta}(y,u)}{\partial u} \iff \frac{\partial\tilde{\alpha}(y,0)}{\partial u} \equiv -\alpha_2'(y)\frac{\partial\tilde{\theta}(y,0)}{\partial u}$$

$$\tilde{\theta}(y,0) \equiv \alpha_0^{-1}\left(\alpha_1(y) - \alpha_2(y)\right) = y$$
$$\tilde{\omega}(y,0) \equiv \omega_2(y) - \omega_2(y) + \omega_1(y) - \omega_1(y) = 0$$
$$\tilde{\alpha}(y,0) \equiv \alpha_2(y) - \alpha_2(y) = 0$$

$$\frac{\partial\tilde{\theta}(y,0)}{\partial u} = \frac{\alpha_1'(y)\,\alpha_0'(0)}{\alpha_0'(y)};\ \frac{\partial\tilde{\omega}(y,0)}{\partial u} = -\left(\omega_2'(y) + \omega_1'(y)\right)\frac{\partial\tilde{\theta}(y,0)}{\partial u} + \omega_1'(y)\,\alpha_0'(0);\ \frac{\partial\tilde{\alpha}(y,0)}{\partial u} \equiv -\alpha_2'(y)\frac{\partial\tilde{\theta}(y,0)}{\partial u}$$

$$\psi_{t+1}'(u)$$
$$= \omega_1'(u) + \omega_2'(u) + \alpha_2'(u)x_{t+1}$$
$$+\frac{\int\left(\frac{\partial\tilde{\omega}(iy,u)}{\partial u} + \frac{\partial\tilde{\alpha}(iy,u)}{\partial u}x_t + \psi_t'\left(\tilde{\theta}(iy,u)\right)\frac{\partial\tilde{\theta}(iy,u)}{\partial u}\right)q_t(y,u;x_{t+1})\,dy}{\int q_t(y,u;x_{t+1})\,dy}$$

$$\psi_{t+1}'(0) = \omega_1'(0) + \omega_2'(0) + \alpha_2'(0)x_{t+1} + \frac{\int\left(\frac{\partial\tilde{\omega}(iy,0)}{\partial u} + \frac{\partial\tilde{\alpha}(iy,0)}{\partial u}x_t + \psi_t'(iy)\frac{\partial\tilde{\theta}(iy,0)}{\partial u}\right)e^{\psi_t(iy) - iyx_{t+1}}dy}{\int e^{\psi_t(iy) - iyx_{t+1}}dy}$$

Let us denote

$$\nu \equiv \frac{2\left(\nu_2\phi_1 + \nu_1\phi_2\right)}{\phi_1 + \phi_2};\ \phi \equiv \frac{2\phi_1\phi_2}{\phi_1 + \phi_2}$$

$$\begin{aligned}
\psi'_{t+1}(0) &= (\nu_1+\nu_2)\varphi + (\phi_1+\phi_2)x_{t+1} - \frac{\frac{1}{\pi}\int_0^\infty Re\left[\left((\nu_2\phi_1+\nu_1\phi_2)\varphi + \frac{\phi_1\phi_2 x_t}{1-iy\varphi}\right)\frac{e^{\psi_t(iy)-iyx_{t+1}}}{1-iy\varphi}\right]dy}{\frac{1}{\pi}\int_0^\infty Re\left[e^{\psi_t(iy)-iyx_{t+1}}\right]dy} \\
&= (\nu_1+\nu_2)\varphi + (\phi_1+\phi_2)x_{t+1} - \frac{\frac{1}{\pi}\int_0^\infty Re\left[\left(\frac{(\phi_1+\phi_2)}{2}\nu\varphi + \frac{(\phi_1+\phi_2)}{2}\frac{\phi x_t}{1-iy\varphi}\right)\frac{e^{\psi_t(iy)-iyx_{t+1}}}{1-iy\varphi}\right]dy}{\frac{1}{\pi}\int_0^\infty Re\left[e^{\psi_t(iy)-iyx_{t+1}}\right]dy} \\
&= (\nu_1+\nu_2)\varphi + (\phi_1+\phi_2)x_{t+1} - \frac{(\phi_1+\phi_2)}{2}\frac{\frac{1}{\pi}\int_0^\infty Re\left[\left(\nu\varphi + \frac{\phi x_t}{1-iy\varphi}\right)\frac{e^{\psi_t(iy)-iyx_{t+1}}}{1-iy\varphi}\right]dy}{\frac{1}{\pi}\int_0^\infty Re\left[e^{\psi_t(iy)-iyx_{t+1}}\right]dy} \\
&= (\nu_1+\nu_2)\varphi + \frac{(\phi_1+\phi_2)}{2}x_{t+1} + \frac{(\phi_1+\phi_2)}{2}\frac{\frac{1}{\pi}\int_0^\infty Re\left[\left(x_{t+1} - \frac{1}{1-iy\varphi}\left(\nu\varphi + \frac{\phi x_t}{1-iy\varphi}\right)\right)e^{\psi_t(iy)-iyx_{t+1}}\right]dy}{\frac{1}{\pi}\int_0^\infty Re\left[e^{\psi_t(iy)-iyx_{t+1}}\right]dy}
\end{aligned}$$

$$\tilde{\psi}'_t(u) = \frac{1}{1-u\varphi}\left(\nu\varphi + \frac{\phi x_t}{1-u\varphi}\right) \Longleftrightarrow \tilde{\psi}'_t(u) = \frac{\nu\varphi}{1-u\varphi} + \frac{\phi x_t}{(1-u\varphi)^2} \Longleftrightarrow \tilde{\psi}_t(u) = -\nu\ln(1-u\varphi) + \frac{\phi u x_t}{1-u\varphi}$$

$$\tilde{\psi}'_t(u) = \frac{\nu\varphi}{1-u\varphi} + \frac{\phi x_t}{(1-u\varphi)^2}; \tilde{\psi}''_t(u) = \frac{\nu\varphi^2}{(1-u\varphi)^2} + 2\frac{\varphi\phi x_t}{(1-u\varphi)^3}$$

$$\tilde{\psi}''_t(0) = \nu\varphi^2 + 2\varphi\phi x_t$$

$$\left(x_{t+1} - \tilde{\psi}'_t(u)\right)e^{\psi_t(u)-ux_{t+1}} = -\left(\tilde{\psi}'_t(u) - x_{t+1}\right)e^{\tilde{\psi}_t(u)-ux_{t+1}}e^{\psi_t(u)-\tilde{\psi}_t(u)}$$

$$\int\left(x_{t+1} - \tilde{\psi}'_t(u)\right)e^{\psi_t(u)-ux_{t+1}}du = -e^{\psi_t(u)-ux_{t+1}} + \int\left(\psi'_t(u) - \tilde{\psi}'_t(u)\right)e^{\psi_t(u)-ux_{t+1}}du$$

$$\int\left(x_{t+1} - \tilde{\psi}'_t(iy)\right)e^{\psi_t(iy)-iyx_{t+1}}dy = ie^{\psi_t(iy)-iyx_{t+1}} + \int\left(\psi'_t(iy) - \tilde{\psi}'_t(iy)\right)e^{\psi_t(iy)-iyx_{t+1}}dy$$

$$\begin{aligned}
\psi'_{t+1}(0) &= (\nu_1+\nu_2)\varphi + \frac{(\phi_1+\phi_2)}{2}x_{t+1} + \frac{(\phi_1+\phi_2)}{2}\frac{\frac{1}{\pi}\int_0^\infty Re\left[\left(x_{t+1} - \frac{1}{1-iy\varphi}\left(\nu\varphi + \frac{\phi x_t}{1-iy\varphi}\right)\right)e^{\psi_t(iy)-iyx_{t+1}}\right]dy}{\frac{1}{\pi}\int_0^\infty Re\left[e^{\psi_t(iy)-iyx_{t+1}}\right]dy} \\
&= (\nu_1+\nu_2)\varphi + \frac{(\phi_1+\phi_2)}{2}x_{t+1} + \frac{(\phi_1+\phi_2)}{2}\frac{\frac{1}{\pi}\int_0^\infty Re\left[\left(x_{t+1} - \tilde{\psi}'_t(iy)\right)e^{\psi_t(iy)-iyx_{t+1}}\right]dy}{\frac{1}{\pi}\int_0^\infty Re\left[e^{\psi_t(iy)-iyx_{t+1}}\right]dy} \\
\psi'_{t+1}(0) &= (\nu_1+\nu_2)\varphi + \frac{(\phi_1+\phi_2)}{2}x_{t+1} + \frac{(\phi_1+\phi_2)}{2}\frac{\frac{1}{\pi}\int_0^\infty Re\left[\left(\psi'_t(iy) - \tilde{\psi}'_t(iy)\right)e^{\psi_t(iy)-iyx_{t+1}}\right]dy}{\frac{1}{\pi}\int_0^\infty Re\left[e^{\psi_t(iy)-iyx_{t+1}}\right]dy}
\end{aligned}$$

8.1.2 Second order

$$\begin{aligned}
& \psi'_{t+1}(u) \\
= \; & \omega'_1(u) + \omega'_2(u) + \alpha'_2(u)x_{t+1} \\
& + \frac{\int\left(\frac{\partial\tilde{\omega}(iy,u)}{\partial u} + \frac{\partial\tilde{\alpha}(iy,u)}{\partial u}x_t + \psi'_t\left(\tilde{\theta}(iy,u)\right)\frac{\partial\tilde{\theta}(iy,u)}{\partial u}\right)q_t(y,u;x_{t+1})\,dy}{\int q_t(y,u;x_{t+1})\,dy}
\end{aligned}$$

$$
\begin{aligned}
&\psi_{t+1}''(u)\\
=\;&\omega_1''(u)+\omega_2''(u)+\alpha_2''(u)\,x_{t+1}+\\
&\frac{\int\left(\frac{\partial^2\tilde{\omega}(iy,u)}{\partial u^2}+\frac{\partial^2\tilde{\alpha}(iy,u)}{\partial u^2}x_t+\psi_t''\left(\tilde{\theta}(iy,u)\right)\left(\frac{\partial\tilde{\theta}(iy,u)}{\partial u}\right)^2+\psi_t'\left(\tilde{\theta}(iy,u)\right)\frac{\partial^2\tilde{\theta}(iy,u)}{\partial u^2}\right)q_t(y,u;x_{t+1})\,dy}{\int q_t(y,u;x_{t+1})\,dy}\\
&+\frac{\int\left(\frac{\partial\tilde{\omega}(iy,u)}{\partial u}+\frac{\partial\tilde{\alpha}(iy,u)}{\partial u}x_t+\psi_t'\left(\tilde{\theta}(iy,u)\right)\frac{\partial\tilde{\theta}(iy,u)}{\partial u}\right)^2q_t(y,u;x_{t+1})\,dy}{\int q_t(y,u;x_{t+1})\,dy}\\
&-\frac{\left[\int\left(\frac{\partial\tilde{\omega}(iy,u)}{\partial u}+\frac{\partial\tilde{\alpha}(iy,u)}{\partial u}x_t+\psi_t'\left(\tilde{\theta}(iy,u)\right)\frac{\partial\tilde{\theta}(iy,u)}{\partial u}\right)q_t(y,u;x_{t+1})\,dy\right]^2}{\left[\int q_t(y,u;x_{t+1})\,dy\right]^2}
\end{aligned}
$$

$$
\begin{aligned}
&\psi_{t+1}''(u)\\
=\;&\omega_1''(u)+\omega_2''(u)+\alpha_2''(u)\,x_{t+1}-\left(\psi_{t+1}'(u)-\omega_1'(u)-\omega_2'(u)-\alpha_2'(u)\,x_{t+1}\right)^2\\
&+\frac{\int\left(\frac{\partial\tilde{\omega}(iy,u)}{\partial u}+\frac{\partial\tilde{\alpha}(iy,u)}{\partial u}x_t+\psi_t'\left(\tilde{\theta}(iy,u)\right)\frac{\partial\tilde{\theta}(iy,u)}{\partial u}\right)^2q_t(y,u;x_{t+1})\,dy}{\int q_t(y,u;x_{t+1})\,dy}\\
&+\frac{\int\left(\frac{\partial^2\tilde{\omega}(iy,u)}{\partial u^2}+\frac{\partial^2\tilde{\alpha}(iy,u)}{\partial u^2}x_t+\psi_t''\left(\tilde{\theta}(iy,u)\right)\left(\frac{\partial\tilde{\theta}(iy,u)}{\partial u}\right)^2+\psi_t'\left(\tilde{\theta}(iy,u)\right)\frac{\partial^2\tilde{\theta}(iy,u)}{\partial u^2}\right)q_t(y,u;x_{t+1})\,dy}{\int q_t(y,u;x_{t+1})\,dy}
\end{aligned}
$$

$$
\begin{aligned}
&\psi_{t+1}''(0)\\
=\;&\omega_1''(0)+\omega_2''(0)+\alpha_2''(0)\,x_{t+1}-\left(\psi_{t+1}'(0)-\omega_1'(0)-\omega_2'(0)-\alpha_2'(0)\,x_{t+1}\right)^2\\
&+\frac{\int\left(\frac{\partial\tilde{\omega}(iy,0)}{\partial u}+\frac{\partial\tilde{\alpha}(iy,0)}{\partial u}x_t+\psi_t'(iy)\frac{\partial\tilde{\theta}(iy,0)}{\partial u}\right)^2e^{\psi_t(iy)-iyx_{t+1}}dy}{\int e^{\psi_t(iy)-iyx_{t+1}}dy}\\
&+\frac{\int\left(\frac{\partial^2\tilde{\omega}(iy,0)}{\partial u^2}+\frac{\partial^2\tilde{\alpha}(iy,0)}{\partial u^2}x_t+\psi_t''(iy)\left(\frac{\partial\tilde{\theta}(iy,0)}{\partial u}\right)^2+\psi_t'(iy)\frac{\partial^2\tilde{\theta}(iy,0)}{\partial u^2}\right)e^{\psi_t(iy)-iyx_{t+1}}dy}{\int e^{\psi_t(iy)-iyx_{t+1}}dy}
\end{aligned}
$$

$$
\begin{aligned}
&\psi_{t+1}''(0)\\
=\;&\nu_1\varphi^2+\nu_2\varphi^2+2\phi_2\varphi x_{t+1}-\left(\psi_{t+1}'(0)-\nu_1\varphi-\nu_2\varphi-\phi_2x_{t+1}\right)^2\\
&+2\varphi\frac{\int\left(\phi_1\psi_t'(iy)-\frac{(\phi_1+\phi_2)}{2}\tilde{\psi}_t'(iy)\right)e^{\psi_t(iy)-iyx_{t+1}}dy}{\int e^{\psi_t(iy)-iyx_{t+1}}dy}\\
&+\frac{\int\left(\phi_1\psi_t'(iy)-\frac{\phi_1+\phi_2}{2}\tilde{\psi}_t'(iy)\right)^2e^{\psi_t(iy)-iyx_{t+1}}dy}{\int e^{\psi_t(iy)-iyx_{t+1}}dy}+\phi_1\frac{\int\left(\phi_1\psi_t''(iy)-\frac{\phi_1+\phi_2}{2}\tilde{\psi}_t''(iy)\right)e^{\psi_t(iy)-iyx_{t+1}}dy}{\int e^{\psi_t(iy)-iyx_{t+1}}dy}\\
&+\phi_2\varphi^2(\phi_1-\phi_2)\,\nu_1\frac{\int\frac{e^{\psi_t(iy)-iyx_{t+1}}}{(1-iy\varphi)^2}dy}{\int e^{\psi_t(iy)-iyx_{t+1}}dy}-2\varphi\phi_2\frac{\int\left(\phi_1\psi_t'(iy)-\frac{\phi_1+\phi_2}{2}\tilde{\psi}_t'(iy)\right)\frac{e^{\psi_t(iy)-iyx_{t+1}}}{1-iy\varphi}dy}{\int e^{\psi_t(iy)-iyx_{t+1}}dy}.
\end{aligned}
$$

Hence, we denote

$$
\bar{\psi}_t(u)\equiv\phi_1\psi_t(u)-\frac{(\phi_1+\phi_2)}{2}\tilde{\psi}_t(u).
$$

We have

$$
\begin{aligned}
\psi_{t+1}''(0)\;=\;&\nu_1\varphi^2+\nu_2\varphi^2+2\phi_2\varphi x_{t+1}-\left(\psi_{t+1}'(0)-\nu_1\varphi-\nu_2\varphi-\phi_2x_{t+1}\right)^2\\
&+\frac{\int\left[\phi_1\bar{\psi}_t''(iy)+\left(\bar{\psi}_t'(iy)\right)^2+2\varphi\bar{\psi}_t'(iy)-2\varphi\phi_2\frac{\bar{\psi}_t'(iy)}{1-iy\varphi}+\frac{\phi_2\varphi^2(\phi_1-\phi_2)\nu_1}{(1-iy\varphi)^2}\right]e^{\psi_t(iy)-iyx_{t+1}}dy}{\int e^{\psi_t(iy)-iyx_{t+1}}dy}.
\end{aligned}
$$

8.2 Option pricing

Regarding option pricing under a MARG dynamic:

$$RV_t = x_{t,1} + x_{t,2}$$

where $x_{t,1}$ and $x_{t,2}$ follow independent ARG dynamics. To price European options under a MARG dynamic, we first compute the option price conditional on both the observed realized variance and the latent factor. In this case, $\psi_{t,h}^{RV}(u)$ is an affine function of RV_t and the latent state $x_{1,t}$:

$$C_t^{Mod}(X,\tau) = S_t P_1(t,\tau) - \exp(-r\tau) X P_2(t,\tau),$$

where

$$P_1(t,\tau) = \frac{1}{2} + \int_0^\infty Re\left(\frac{\exp\left(iur\tau + \psi_{t,\tau}^{RV}\left(\bar{f}(iu)\right) - iu\ln\left(\frac{X}{S_t}\right)\right)}{iu\pi}\right)du$$

$$P_2(t,\tau) = \frac{1}{2} + \int_0^\infty Re\left(\frac{\exp\left(iur\tau + \psi_{t,\tau}^{RV}\left(\bar{f}(-iu)\right) - iu\ln\left(\frac{X}{S_t}\right)\right)}{iu\pi}\right)du$$

$$\bar{f}(u) \equiv \frac{1}{2}(u^2 + u)$$

and

$$\psi_{t,\tau}^{RV}(u) \equiv \ln\left(E\left[\exp\left(u\sum_{j=1}^{\tau} RV_{t+j}\right)\right]|x_{1s}, x_{2s}, s \le t\right).$$

We have

$$\psi_{t,\tau}^{RV}(u) = \psi_{t,\tau}^{(1)}(u) + \psi_{t,\tau}^{(2)}(u),$$

where

$$\psi_{t,\tau}^{(j)}(u) = a_j(u;\tau)\, x_{t,j} + b_j(u;\tau),$$

$$a_j(u;\tau) = \frac{\phi_j\left(u + a_j(u;\tau-1)\right)}{1 - \left(u + a_j(u;\tau-1)\right)\varphi_j}; b_j(u;\tau) = b_j(u;\tau-1) - \nu_j \log(1 - \left(u + a_j(u;\tau-1)\right)\varphi_j),$$

with the initial values $a_j(u;0) = 0$, $b_j(u;0) = 0$. Hence,

$$P_1(t,\tau) = \frac{1}{2} + \int_0^\infty Re\left(\frac{1}{iu\pi}\exp\left(\begin{array}{c} iur\tau - iu\ln\left(\frac{X}{S_t}\right) + b_1\left(\bar{f}(iu);\tau\right) + a_{1j}\left(\bar{f}(iu);\tau\right)x_{t,1} \\ +b_2\left(\bar{f}(iu);\tau\right) + a_{2j}\left(\bar{f}(iu);\tau\right)x_{t,2}\end{array}\right)\right)du$$

$$P_2(t,\tau) = \frac{1}{2} + \int_0^\infty Re\left(\frac{1}{iu\pi}\exp\left(\begin{array}{c} iur\tau - iu\ln\left(\frac{X}{S_t}\right) + b_1\left(\bar{f}(-iu);\tau\right) + a_{1j}\left(\bar{f}(-iu);\tau\right)x_{t,1} \\ +b_2\left(\bar{f}(-iu);\tau\right) + a_{2j}\left(\bar{f}(-iu);\tau\right)x_{t,2}\end{array}\right)\right)du.$$

Substituting $x_{2,t} = RV_t - x_{1,t}$, we have

$$P_1(t,\tau) = \frac{1}{2} + \int_0^\infty Re\left(\frac{1}{iu\pi}\exp\left(\begin{array}{c} iur\tau - iu\ln\left(\frac{X}{S_t}\right) + b_2\left(\bar{f}(iu);\tau\right) + a_{2j}\left(\bar{f}(iu);\tau\right)RV_t \\ +b_1\left(\bar{f}(iu);\tau\right) + \left(a_{1j}\left(\bar{f}(iu);\tau\right) - a_{2j}\left(\bar{f}(iu);\tau\right)\right)x_{t,1}\end{array}\right)\right)du$$

$$P_2(t,\tau) = \frac{1}{2} + \int_0^\infty Re\left(\frac{1}{iu\pi}\exp\left(\begin{array}{c} iur\tau - iu\ln\left(\frac{X}{S_t}\right) + b_2\left(\bar{f}(-iu);\tau\right) + a_{2j}\left(\bar{f}(-iu);\tau\right)RV_t \\ +b_1\left(\bar{f}(-iu);\tau\right) + \left(a_{1j}\left(\bar{f}(-iu);\tau\right) - a_{2j}\left(\bar{f}(-iu);\tau\right)\right)x_{t,1}\end{array}\right)\right)du.$$

Following Bates (2006), the model price (denoted by $\hat{C}_t^{Mod}(X,\tau)$) with only the present and past realized variances in the information set is the expectation of the model price, which includes $x_{1,t}$ in the information set, that is,

$$\hat{C}_t^{Mod}(X,\tau) \equiv E\left[C_t^{Mod}(X,\tau)|RV_s, s \le t\right] = S_t E\left[P_1(t,\tau)|RV_t\right] - \exp(-r\tau) X E\left[P_2(t,\tau)|RV_t\right]. \tag{34}$$

We have

$$E\left[P_1(t,\tau)\,|RV_s, s\le t\right] = \frac{1}{2}+\int_0^\infty Re\left(\frac{1}{iu\pi}\exp\left(\begin{array}{c} iur\tau - iu\ln\left(\frac{X}{S_t}\right) + b_2\left(\bar{f}\left(iu\right);\tau\right) + a_{2j}\left(\bar{f}\left(iu\right);\tau\right)RV_t \\ +b_1\left(\bar{f}\left(iu\right);\tau\right) + g_{t|t}\left(a_{1j}\left(\bar{f}\left(iu\right);\tau\right) - a_{2j}\left(\bar{f}\left(iu\right);\tau\right)\right) \end{array}\right)\right)du$$

$$E\left[P_2(t,\tau)\,|RV_s, s\le t\right] = \frac{1}{2}+\int_0^\infty Re\left(\frac{1}{iu\pi}\exp\left(\begin{array}{c} iur\tau - iu\ln\left(\frac{X}{S_t}\right) + b_2\left(\bar{f}\left(-iu\right);\tau\right) + a_{2j}\left(\bar{f}\left(-iu\right);\tau\right)RV_t \\ +b_1\left(\bar{f}\left(-iu\right);\tau\right) + g_{t|t}\left(a_{1j}\left(\bar{f}\left(-iu\right);\tau\right) - a_{2j}\left(\bar{f}\left(-iu\right);\tau\right)\right) \end{array}\right)\right)du,$$

where $g_{t|t}\left(u\right) \equiv \ln\left[E\left[\exp\left(ux_{1t}\right)|RV_s, s\le t\right]\right]$. Following Bates (2006), we approximate $g_{t|t}\left(u\right)$ using the cumulant generating function of the gamma distribution, that is,

$$g_{t|t}\left(u\right) \approx -\left(\frac{x_{1t|t}^2}{\sigma_{1t|t}^2}\right)\ln\left(1-\frac{\sigma_{1t|t}^2}{x_{1t|t}}u\right),$$

where

$$x_{1t|t} = \frac{\psi_t'\left(0\right) - \left(\nu_1+\nu_2\right)\varphi - \phi_2 RV_t}{\phi_1-\phi_2}, \quad \sigma_{1t|t}^2 = \frac{\psi_t''\left(0\right) + \left(\nu_1+\nu_2\right)\varphi^2 - 2\varphi\psi_t'\left(0\right)}{\left(\phi_1-\phi_2\right)^2} \tag{35}$$

with

$$x_{1t|t} \equiv E\left[x_{1t}|RV_s, s\le t\right], \quad \sigma_{1t|t}^2 \equiv Var\left[x_{1t}|RV_s, s\le t\right], \tag{36}$$

and $\psi_t'\left(0\right)$ and $\psi_t''\left(0\right)$ evolve recursively as given in corollary 2 of the main paper.

Table 1: Monte Carlo Exercise for the GMM

Mean, Median and RMSE of parameters estimates across N samples (of T observations). Parameters used in the simulation of the GARG model are φ =2.78E-05, ν =0.139, ϕ =0.112, β =0.922 and θ =0.906

Par	Mean	Median	RMSE	Mean	Median	RMSE	Mean	Median	RMSE
	N=1000, T=250			N=1000, T=500			N=1000, T=1000		
φ	2.58E-05	3.66E-05	1.68E-05	2.61E-05	3.66E-05	1.66E-05	2.64E-05	3.66E-05	1.64E-05
ν	0.077	0.109	0.079	0.077	0.109	0.078	0.078	0.109	0.077
ϕ	0.104	0.115	0.018	0.104	0.115	0.018	0.105	0.115	0.018
β	0.695	0.946	0.437	0.702	0.946	0.430	0.710	0.946	0.423
θ	1.318	0.879	0.780	1.307	0.879	0.770	1.293	0.879	0.755
	N=2000, T=250			N=2000, T=500			N=2000, T=1000		
φ	2.80E-05	3.67E-05	1.55E-05	2.90E-05	3.66E-05	1.49E-05	2.97E-05	3.66E-05	1.44E-05
ν	0.083	0.109	0.072	0.086	0.109	0.069	0.088	0.109	0.066
ϕ	0.106	0.115	0.017	0.107	0.115	0.016	0.108	0.115	0.015
β	0.744	0.946	0.395	0.766	0.946	0.372	0.784	0.946	0.353
θ	1.248	0.879	0.733	1.208	0.878	0.693	1.175	0.878	0.657
	N=4000, T=250			N=4000, T=500			N=4000, T=1000		
φ	3.20E-05	3.66E-05	1.28E-05	3.24E-05	3.66E-05	1.25E-05	3.27E-05	3.66E-05	1.24E-05
ν	0.095	0.109	0.056	0.096	0.109	0.054	0.097	0.109	0.053
ϕ	0.110	0.115	0.012	0.111	0.115	0.012	0.111	0.115	0.011
β	0.837	0.946	0.291	0.847	0.946	0.278	0.852	0.946	0.270
θ	1.084	0.879	0.557	1.066	0.878	0.533	1.056	0.879	0.517

Table 2: Monte Carlo Exercise for the QMLE

Mean, Median and RMSE of parameters estimates across N samples (of T observations). Parameters used in the simulation of the GARG model are φ =2.78E-05, ν =0.139, ϕ =0.112, β =0.922 and θ =0.906

Par	Mean	Median	RMSE	Mean	Median	RMSE	Mean	Median	RMSE
	N=1000, T=250			N=1000, T=500			N=1000, T=1000		
φ	6.00E-05	5.13E-05	5.37E-05	5.97E-05	5.20E-05	5.19E-05	6.13E-05	5.31E-05	5.43E-05
ν	0.142	0.051	0.453	0.154	0.051	0.415	0.148	0.050	0.427
ϕ	0.110	0.109	0.029	0.108	0.109	0.027	0.108	0.106	0.028
β	1.028	0.959	0.356	1.048	0.960	0.415	1.066	0.964	0.488
θ	0.879	0.909	0.189	0.875	0.908	0.204	0.870	0.909	0.206
	N=2000, T=250			N=2000, T=500			N=2000, T=1000		
φ	5.45E-05	5.02E-05	4.20E-05	5.46E-05	4.99E-05	4.12E-05	5.56E-05	5.15E-05	4.23E-05
ν	0.216	0.057	0.963	0.220	0.055	0.915	0.179	0.052	0.711
ϕ	0.109	0.109	0.023	0.107	0.107	0.023	0.107	0.107	0.024
β	0.998	0.937	0.270	0.993	0.941	0.243	1.001	0.947	0.255
θ	0.891	0.918	0.164	0.893	0.918	0.158	0.888	0.915	0.161
	N=4000, T=250			N=4000, T=500			N=4000, T=1000		
φ	5.30E-05	5.07E-05	3.82E-05	5.26E-05	5.17E-05	3.57E-05	5.30E-05	5.31E-05	3.57E-05
ν	0.244	0.048	0.942	0.373	0.049	3.038	0.275	0.050	2.347
ϕ	0.108	0.109	0.022	0.108	0.107	0.021	0.107	0.106	0.021
β	0.969	0.940	0.168	0.963	0.941	0.146	0.966	0.964	0.147
θ	0.901	0.914	0.126	0.905	0.913	0.121	0.903	0.909	0.122

Table 3: S&P 500 Index Option Data

This table presents the characteristics of S&P 500 index option data by moneyness, maturity, and VIX level. We use Wednesday's closing out-of-the-money (OTM) call and put contracts from OptionMetrics for the period starting from January 10, 1996, to August 28, 2013. The moneyness is measured by the Black-Scholes delta. DTM denotes the number of calendar days to maturity. The average price is reported in dollars, and the average implied volatility is expressed in percentages.

	OTM Call			OTM Put			
	Delta < 0.3	0.3 ≤ Delta < 0.4	0.4 ≤ Delta < 0.5	0.5 ≤ Delta < 0.6	0.6 ≤ Delta < 0.7	Delta ≥ 0.7	All
Panel A: By Moneyness							
Number of contracts	3,788	1,391	1,781	2,846	2,746	8,731	21,283
Average price	7.85	20.94	32.28	45.30	65.93	132.41	74.35
Average implied volatility	16.72	18.40	19.31	20.40	21.71	25.09	21.62
Average bid-ask spread	1.046	1.674	1.955	2.018	1.834	1.228	1.470
	DTM < 30	30 ≤ DTM < 60	60 ≤ DTM < 90	90 ≤ DTM < 120	120 ≤ DTM < 150	DTM ≥ 150	All
Panel B: By Maturity							
Number of contracts	2,725	6,480	5,053	2,869	1,974	2,182	21,283
Average price	41.26	61.01	76.44	92.30	97.88	105.59	74.35
Average implied volatility	20.21	21.28	21.73	22.94	22.08	21.95	21.62
Average bid-ask spread	0.820	1.231	1.578	1.872	1.800	1.910	1.470
	VIX < 15	15 ≤ VIX < 20	20 ≤ VIX < 25	25 ≤ VIX < 30	30 ≤ VIX < 35	VIX ≥ 35	All
Panel C: By VIX Level							
Number of contracts	3,962	6,133	5,996	2,456	1,240	1,496	21,283
Average price	57.95	66.90	80.75	85.77	85.33	94.86	74.35
Average implied volatility	13.61	18.04	22.45	26.24	30.22	39.42	21.62
Average bid-ask spread	1.055	1.301	1.446	1.704	1.811	2.683	1.470

Figure 1: Autocorrelation functions across models

This figure presents the autocorrelation of the level $corr(x_t, x_{t+h})$ (top left), the cross-correlation between the level and the squared $corr(x_t, x_{t+h}^2)$ (top right), the cross-correlation between the squared and the level $corr(x_t^2, x_{t+h})$ (bottom left), and the autocorrelation of the squared $corr(x_t^2, x_{t+h}^2)$ (bottom right), of the daily realized variance. The sample begins January 01, 2000, and ends December 31, 2017.

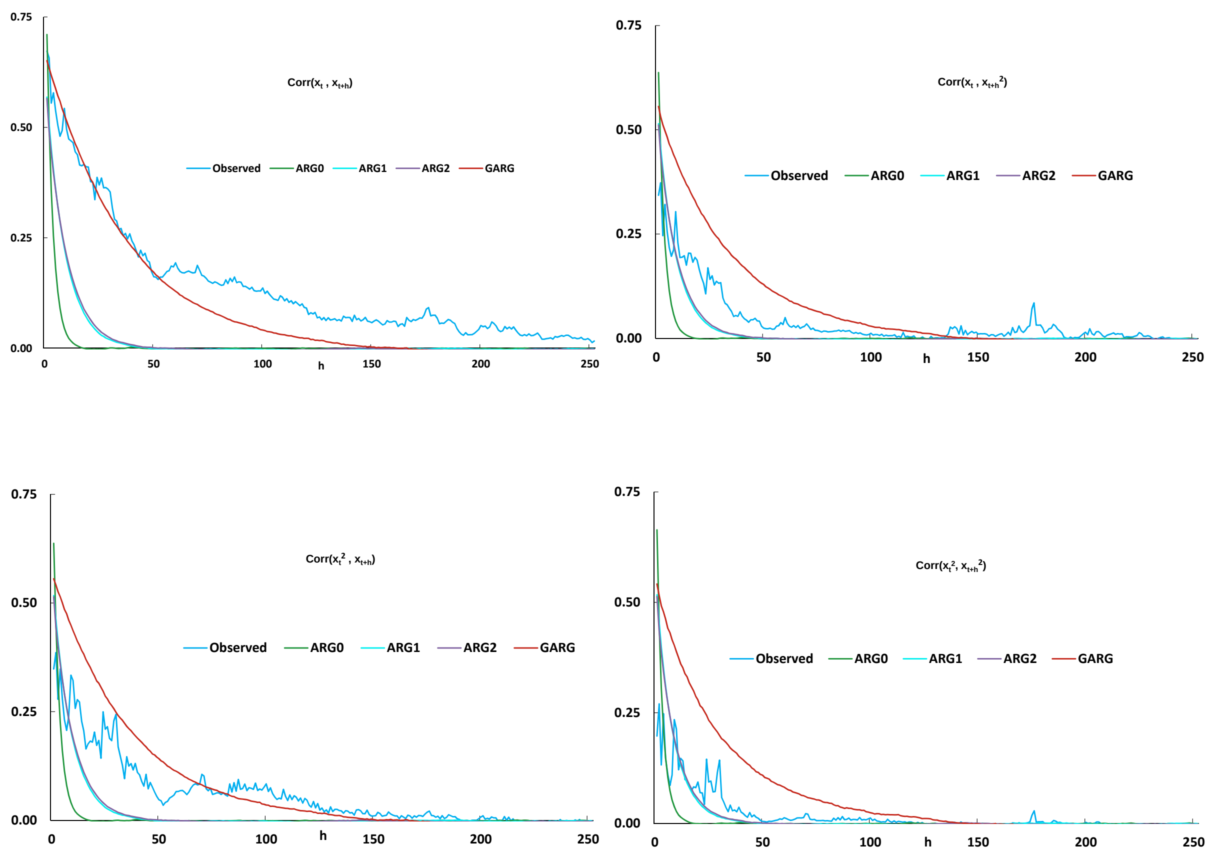
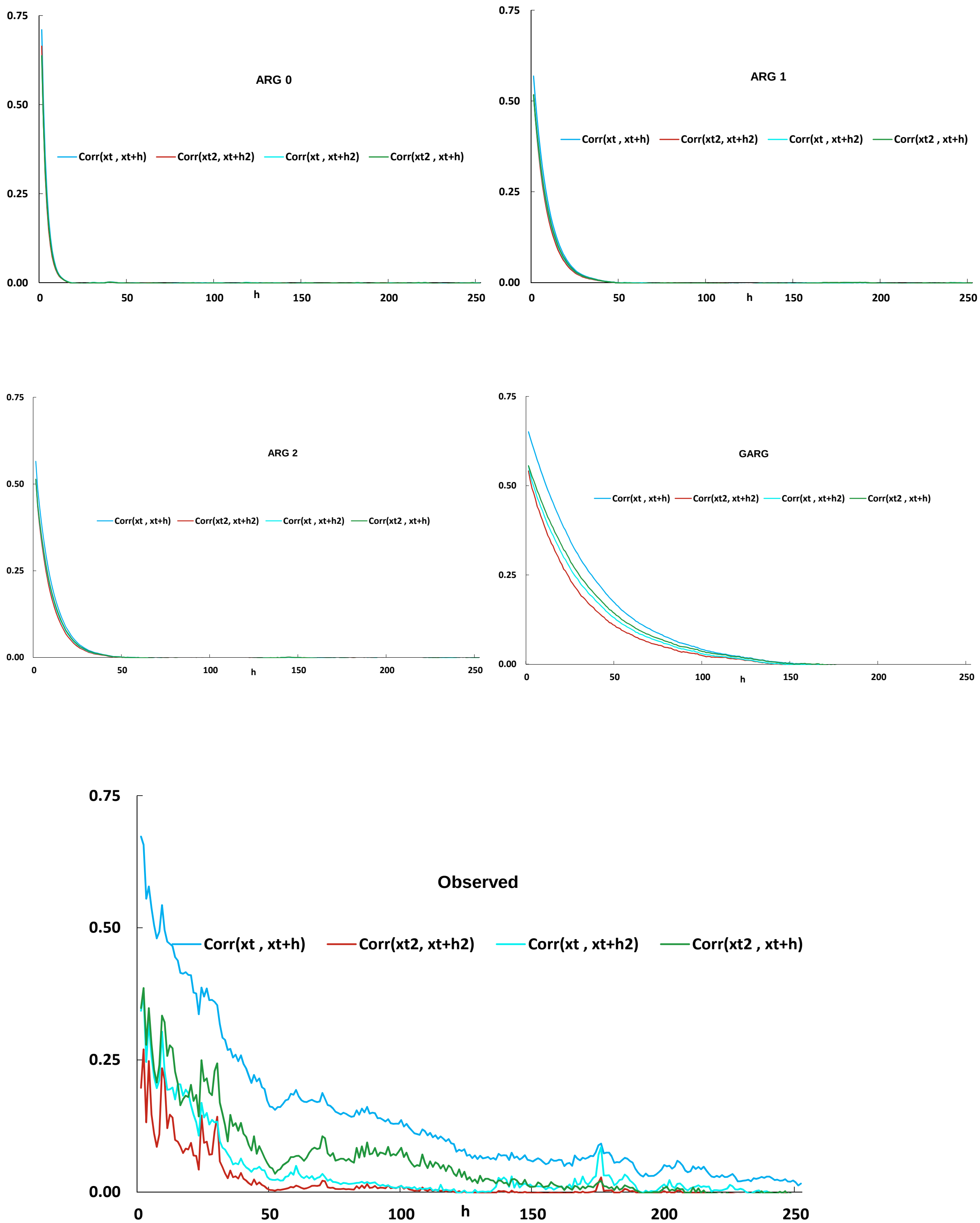


Figure 2: Autocorrelation functions

For a given model, we plot together $corr(x_t, x_{t+h})$, $corr(x_t, x^2_{t+h})$, $corr(x^2_t, x_{t+h})$, and $corr(x^2_t, x^2_{t+h})$. The last row displays the sample autocorrelogram. The sample begins January 01, 2000, and ends December 31, 2017.



References

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

Black, Fischer, and Scholes, Myron (1973), "The pricing of options and corporate liabilities," *Journal of Political Economy,* 81(3), 637–654.

Bollerslev, Tim, and Zhou, Hao (2002), "Estimating stochastic volatility diffusion using conditional moments of integrated volatility," *Journal of Econometrics,* 109(1), 33–65.

Carpenter, Mark, and Diawara, Norou (2007), "A multivariate gamma distribution and its characterizations," *American Journal of Mathematical and Management Sciences,* 27(3-4), 499–507.

Carrasco, Marine, and Florens, Jean-Pierre (2000), "Generalization of GMM to a continuum of moment conditions," *Econometric Theory,* 16(6), 797–834.

Carrasco, Marine, and Florens, Jean-Pierre (2002), "Efficient gmm estimation using the empirical characteristic function," Technical Report.

Carrasco, Marine, Chernov, Mikhail, Florens, Jean-Pierre, and Ghysels, Eric (2007), "Efficient estimation of general dynamic models with a continuum of moment conditions," *Journal of Econometrics,* 140(2), 529–573.

Christoffersen, Peter, Heston, Steve, and Jacobs, Kris (2006), "Option valuation with conditional skewness," *Journal of Econometrics,* 131(1), 253–284.

Darolles, Serge, Gourieroux, Christian, and Jasiak, Joann (2006), "Structural laplace transform and compound autoregressive models," *Journal of Time Series Analysis,* 27(4), 477–503.

Feunou, B., and Meddahi, N. (2009), "Generalized affine models," Working Paper, Available at SSRN.

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

Heston, Steven L., and Nandi, Saikat (2000), "A closed-form GARCH option valuation model," *The Review of Financial Studies,* 13(3), 585–625.

Knight, John L., and Yu, Jun (2002), "Empirical characteristic function in time series estimation," *Econometric Theory,* 18(3), 691–721.

Le, Anh, Singleton, Kenneth J., and Dai, Qiang (2010), "Discrete-time affineQ term structure models with generalized market prices of risk," *The Review of Financial Studies,* 23(5), 2184–2227.

Zwillinger, Dan (2018) *"CRC standard mathematical tables and formulas," in Advances in Applied Mathematics,* CRC Press.