Analytic moments for GJR-GARCH (1,1) processes

Analytic Moments for GJR-GARCH (1,1) Processes

Carol Alexander* Emese Lazar† Silvia Stanescu‡

March 17, 2020

Abstract

For a GJR-GARCH(1,1) specification with a generic innovation distribution we derive analytic expressions for the first four conditional moments of the forward and aggregated returns and variances. Moments for the most commonly used GARCH models are stated as special cases. We also derive the limits of these moments as the time horizon increases, establishing regularity conditions for the moments of aggregated returns to converge to normal moments. A simulation study using these analytic moments produces approximate predictive distributions which are free from the bias affecting simulations. An empirical study using almost 30 years of daily equity index, exchange rate and interest rate data applies Johnson SU and Edgeworth expansion distribution fitting to our closed-form formulae for higher moments of returns.

Keywords: Approximate predictive distributions, conditional and unconditional moments, GARCH, kurtosis, skewness, simulation

JEL Code: C53

*University of Sussex Business School. c.alexander@sussex.ac.uk
†Henley Business School. e.lazar@rdg.ac.uk
‡Cantab Capital Partners. stanescu.silvia@gmail.com

1 Introduction

Return distributions have a great variety of financial applications to risk assessment and portfolio optimization so their prediction attracts much interest in academic research. It has long been recognized that time series of asset returns are not well described by normal, independent processes (Mandelbrot, 1963; Fama, 1965). Typically, their conditional distributions are non-normal and they exhibit volatility clustering, so they are not independent. Hence, we require forecasts of the entire distribution, not only of the first two moments of returns.

The family of generalized autoregressive conditional heteroskedasticity (GARCH) models is highly successful in capturing (at least partially) the salient empirical features of both conditional and unconditional returns distributions. Following the pioneering work of Engle (1982), Bollerslev (1986) and Taylor (1986) numerous alternative specifications for GARCH processes have been proposed. In many financial markets, especially equities and commodities, the GARCH conditional variance equation captures the asymmetric response of volatility to innovations with different signs. Well-known asymmetric GARCH models include the EGARCH model of Nelson (1991), the AGARCH model of Engle (1990) and Engle and Ng (1993), the NGARCH model also proposed by Engle and Ng (1993), and the model of Glosten, Jagannathan and Runkle (1993), henceforth denoted GJR. Additionally, GARCH models with non-normal innovation distributions have been developed by Bollerslev (1987), Nelson (1991), Haas, Mittnik, Paolella (2004) and many others.

The performance of various GARCH models has been empirically assessed by numerous authors, following Andersen and Bollerslev (1998), Marcucci (2005) and many others.[1] Virtually all this literature refers to the accuracy of forward or aggregated returns distributions when a point GARCH variance forecast is used. However, only the one-step-ahead GARCH variance forecast can be made with certainty: due to the uncertainty about future returns, the forward returns variances, and variances of aggregated future returns, are stochastic. So

[1] Also, Andersen, Bollerslev and Diebold (2009) give a broad overview of volatility modelling procedures, focusing on the GARCH methodology and Bauwens, Laurent and Rombouts (2006) review some important contributions to the multivariate ARCH literature.

a point GARCH variance forecast represents only an expected value of the GARCH variance, under its distribution. Until now, the only papers to examine this conditional distribution are Ishida and Engle (2002), who derive the conditional variance of the forward conditional variance for a symmetric GARCH(1,1) model with symmetric innovations, and Christoffersen et al. (2010) who derive the second conditional moment of the two-step-ahead forward variance for eight GARCH processes with affine vs. non-affine and conditionally-normal vs. conditionally GED alternatives.

By contrast, there is considerable research on the unconditional moments of returns generated by GARCH processes.[2] However, since returns are not identically distributed, it is the conditional moments of the returns and variances, and their dynamics that are most important for financial applications. Knowledge of the dynamics of the conditional mean and variance is sufficient only when conditional distributions are normal: more generally, the dynamics of higher order conditional moments are needed. Hence, whilst most research focuses on the (first four) conditional moments of forward returns for some specific GARCH processes, several recent papers also consider the aggregated returns.

Duan et al. (1999) derive expressions for the first four conditional moments of the aggregated returns generated by the normal NGARCH model under the risk-neutral probability measure, and Duan et al. (2006) extend these results to the risk-neutral moments of aggregated returns under normal GJR and normal EGARCH processes.[3] Wong and So (2003) derive an expression for the variance of aggregated QGARCH returns and, under the additional assumption that the innovation is symmetric, expressions for the third and fourth

[2]See Engle (1982), Nemec (1985), Milhoj (1985), Bollerslev (1986), He and Terasvirta (1999a, 1999b), Karanasos (1999, 2001), He, Terasvirta and Malmsten (2002), Demos (2002), Ling and McAleer (2002a, 2002b), Karanasos and Kim (2003), Bai, Russell and Tiao (2003), Karanasos, Psaradakis and Sola (2004) and Francq and Zakoian (2010). A flexible approach that offers analytic expressions for the moments of the absolute values of the returns is given in Harvey (2013) for the Beta-t-EGARCH model and Harvey and Lange (2017) for the more general Beta-Gen-t-EGARCH model combined with conditional score dynamics.

[3]See also Nelson (1991) for related results on the moments of the EGARCH model under the physical distribution and Mazzoni (2010) who considers a related application to option pricing in the implied measure, using the Heston and Nandi (2000) model. Note that the finance literature makes a clear distinction between physical and implied distributions, the former being obtained from time series of asset returns and the latter being obtained from option prices under the assumption of complete markets and risk-neutral valuation.

order conditional moments of aggregate returns.[4] Breuer and Jandacka (2010) derive the limit of the variance and kurtosis of forward and aggregated returns for a generic symmetric GARCH(1,1) process, for which the forward and aggregated skewness are both zero.

We extend this previous research in a unified framework for which the results cited above may be derived as special cases. Assuming a GJR specification and a generic conditional distribution that accommodates both skewness and kurtosis in the innovations, we derive formulae for the first four conditional moments of forward and aggregated returns and of forward and aggregated variances. We also derive the limits of these moments as the returns horizon increases. Furthemore, we compare our analytic moments with those generated by Monte Carlo simuation, and by bootstrapping. This shows that our moments are unbiased, whereas even a very large number of simulations can be highly inaccurate, and bootstrapping methods can have large biases. Our results may be used to generate accurate forward and aggregated GARCH distributions for real time series data, without simulations. To illustrate this, an empirical study estimates a variety of GARCH models for the S&P 500 index, Euro–US dollar exchange rate and the 3-month Treasury bill rate and again demonstrates a very good fit between the approximate and simulated predictive returns distributions.[5]

The formulae for the moments are presented in Section 2 and those for the limits in Section 4. The proofs are lengthy and are presented in a separate on-line technical appendix, which also details the results for important special cases of the generic model. Section 3 compares our analytic moments with those obtained using Monte Carlo simulation and bootstrapping. Section 5 applies the analytic moments to derive approximate predictive distributions for the forward and aggregated returns, with an extensive empirical study to examine the goodness-of-fit between these distributions and those generated via simulation; Section 6 concludes.

[4]Their QGARCH model is the same as an AGARCH(p,q). Also in the context of a QGARCH process, Simonato (2013) considers the approximation of *multivariate* aggregate returns distributions based on their first four moments. Of some relation to this research is the paper by Christoffersen et al. (2008), proposing a new two-component GARCH model for European options valuation, for which the authors also derive the conditional moment generating function of the (log) price distribution.

[5]The return distributions could, alternatively, be studied via computing their bounds as in Goncalves et al. (2016).

2 Moments of Generic GJR Returns and Variances

Here we present analytic expressions for the first four conditional moments of both forward one-period and future aggregated (also called cumulative) returns and variances for the GJR model with a generic innovation distribution having zero mean, unit variance and finite higher moments. For the n^{th} conditional moment of returns to exist we need the first n moments of z to be finite. So, fourth moments exist for financial returns if $n \geq 4$, which is often but not always satisfied empirically. For instance, Deschamps (2012), Harvey and Lange (2017) and many others obtain degrees of freedom estimates greater than 4 in t-GARCH models. Others obtain an estimate greater than 4 for the maximum moment exponent,[6] as in Huisman et al. (2001) and Alexander and Lazar (2009), among others. Similarly, for the n^{th} conditional moment of variances to exist we need the first $2n$ moments of z to be finite and again, this is verified by several but not all prior empirical studies of financial returns. For example, Lux and Morales-Arias (2010), Zhu and Galbraith (2011), Krause and Paolella (2014), Theodossiou and Savva (2015) and Harvey and Lange (2017) all estimate general t-GARCH models with degrees of freedom greater than 8; and Huisman et al. (2002) derive a maximum moment exponent greater than 8. However, Harvey and Sucarrat (2014) find several financial assets with degrees of freedom estimates less than 8 in t-GARCH models. Note that our results can be applied depending on the moments of the data: e.g. if only the first 6 moments of z exist, then our results can be applied for all four moments of the returns and the first 3 moments of the variance.

We assume that the one-period log return $r_t = \log\left(\frac{P_{t+1}}{P_t}\right)$ on a financial asset with market price P_t follows a stationary process. The mathematical specification of the generic GJR model is:

$$r_t = \mu + \varepsilon_t, \quad \varepsilon_t = z_t h_t^{1/2}, \quad z_t \underset{i.i.d.}{\sim} D\left(0,1\right), \quad h_t = \omega + \alpha\, \varepsilon_{t-1}^2 + \lambda\, \varepsilon_{t-1}^2 I_{t-1}^{-} + \beta\, h_{t-1}, \qquad (1)$$

[6] For a discussion on the link between the tail index estimator and the degrees of freedom of a t distribution see Paolella (2016).

where I_t^- is an indicator function which equals 1 if $\varepsilon_t < 0$ and zero otherwise. We note that z_t and h_t are contemporaneously independent because z_t is i.i.d. and h_t depends only on the past. We assume that $D(0,1)$ is a generic conditional distribution with zero mean, unit variance, constant skewness τ_z and kurtosis κ_z, and with constant higher moments $\mu_z^{(i)} = E\left(z_t^i\right)$ for any $i > 4,\ i \in N$.[7]

The aggregated return over n consecutive time periods is $\sum_{s=1}^{n} r_{t+s}$, and the following notation is used for the conditional un-centred and centred moments of the forward and aggregated returns:

$$\tilde{\mu}_{r,s}^{(i)} = E_t\left(r_{t+s}^i\right), \quad \mu_{r,s}^{(i)} = E_t\left(\left(r_{t+s} - \tilde{\mu}_{r,s}^{(1)}\right)^i\right)$$
$$\tilde{M}_{r,n}^{(i)} = E_t\left[\left(\sum_{s=1}^{n} r_{t+s}\right)^i\right], \quad M_{r,n}^{(i)} = E_t\left(\left(\sum_{s=1}^{n}\left(r_{t+s} - \tilde{\mu}_{r,s}^{(1)}\right)\right)^i\right)$$

for $s = 1, 2, ..., n$ and $i = 1, 2, 3, 4$. Thus, the skewness and kurtosis of the forward return distributions are:

$$\tau_{r,s} = \mu_{r,s}^{(3)}\left(\mu_{r,s}^{(2)}\right)^{-3/2} \quad \text{and} \quad \kappa_{r,s} = \mu_{r,s}^{(4)}\left(\mu_{r,s}^{(2)}\right)^{-2}$$

and the skewness and kurtosis of the aggregated return distributions are:

$$\mathrm{T}_{r,n} = M_{r,n}^{(3)}\left(M_{r,n}^{(2)}\right)^{-3/2} \quad \text{and} \quad \mathrm{K}_{r,n} = M_{r,n}^{(4)}\left(M_{r,n}^{(2)}\right)^{-2}.$$

Similarly, the following notation is used for the conditional un-centred and centred moments of the forward and aggregated variances:

$$\tilde{\mu}_{h,s}^{(i)} = E_t\left(h_{t+s}^i\right), \quad \mu_{h,s}^{(i)} = E_t\left(\left(h_{t+s} - \tilde{\mu}_{h,s}^{(1)}\right)^i\right)$$
$$\tilde{M}_{h,n}^{(i)} = E_t\left[\left(\sum_{s=1}^{n} h_{t+s}\right)^i\right], \quad M_{h,n}^{(i)} = E_t\left(\left(\sum_{s=1}^{n}\left(h_{t+s} - \tilde{\mu}_{h,s}^{(1)}\right)\right)^i\right)$$

[7] To be more precise, we have:

$$\mu_z^{(i)} = E\left(z_t^i\right) = E_t\left(z_t^i\right) = E_t\left(z_t - E_t\left(z_t\right)\right)^i = E_t\left(z_t - E_t\left(z_t\right)\right)^i E_t\left(z_t^2\right)^{-i/2}$$

since un-centred, centred and standardized moments are all equal for a zero mean, unit variance distribution. Also, since the z process is i.i.d., conditional and unconditional moments of z are also identical.

for $s = 1, 2, ..., n$ and $i = 1, 2, 3, 4$. Thus, the skewness and kurtosis of the forward variance distributions are:

$$\tau_{h,s} = \mu_{h,s}^{(3)} \left(\mu_{h,s}^{(2)}\right)^{-3/2} \quad \text{and} \quad \kappa_{h,s} = \mu_{h,s}^{(4)} \left(\mu_{h,s}^{(2)}\right)^{-2}$$

and the skewness and kurtosis of the aggregated variance distributions are:

$$\mathrm{T}_{h,n} = M_{h,n}^{(3)} \left(M_{h,n}^{(2)}\right)^{-3/2} \quad \text{and} \quad \mathrm{K}_{h,n} = M_{h,n}^{(4)} \left(M_{h,n}^{(2)}\right)^{-2}.$$

We start with the un-centred moments, namely:

$$E_t\left(x_{t+s}^i\right) \quad \text{and} \quad E_t\left[\left(\sum_{s=1}^{n} x_{t+s}\right)^i\right]$$

for $x = r$ and $x = h$ in turn, $s = 1, 2, ..., n$ and $i = 1, 2, 3, 4$. Subsequently, we obtain the centred and standardized moments of the GJR process with a generic innovation distribution. The derivations rely on the observation that, although $E_t\left(h_{t+1}\right) = h_{t+1}$ (i.e. $V_t\left(h_{t+1}\right) = 0$) both $\left\{h_{t+s}\left|\Omega_t : s \in N\backslash\{0,1\}\right.\right\}$ and $\left\{\sum_{s=1}^{n} h_{t+s}\left|\Omega_t : n \in N\backslash\{0,1\}\right.\right\}$ are random. Moreover, both $\left\{r_{t+s}\left|\Omega_t : s \in N\backslash\{0\}\right.\right\}$ and $\left\{\sum_{s=1}^{n} r_{t+s}\left|\Omega_t : n \in N\backslash\{0\}\right.\right\}$ are random and have distributions that can also be approximated using moments that we derive.

The following notation will be used: $\varphi = \alpha + \lambda F_0 + \beta$, with F_0 being the distribution function for $D(0,1)$ evaluated at zero; $\bar{h} = \omega(1-\varphi)^{-1}$ is the steady-state variance towards which the conditional variance mean reverts, if $\varphi \in (0,1)$;

$$\tilde{\mu}_{h,s}^{(2)} = c_1 + c_2\varphi^{s-1} + \left(h_{t+1}^2 - c_3\right)\gamma^{s-1}, \tag{2}$$

where $\gamma = \varphi^2 + (\kappa_z - 1)(\alpha + \lambda F_0)^2 + \kappa_z\lambda^2 F_0(1 - F_0)$, $c_1 = \left(\omega^2 + 2\omega\varphi\bar{h}\right)(1-\gamma)^{-1}$, $c_2 = 2\omega\varphi\left(h_{t+1} - \bar{h}\right)(\varphi - \gamma)^{-1}$ and $c_3 = c_1 + c_2$; in the following expectations f is the density function of $D(0,1)$ and $E_t\left(h_{t+s}^{3/2}\right) \simeq \frac{5}{8}\left(\tilde{\mu}_{h,s}^{(1)}\right)^{3/2} + \frac{3}{8}\tilde{\mu}_{h,s}^{(2)}\left(\tilde{\mu}_{h,s}^{(1)}\right)^{-1/2}$ with $\tilde{\mu}_{h,s}^{(2)}$ given in (2) and $\tilde{\mu}_{h,s}^{(1)}$ given in Theorem 1 below:

$$E_t\left(\varepsilon_{t+s}\varepsilon_{t+s+u}^2\right) = \varphi^{u-1}\left(\alpha\tau_z + \lambda\int_{x=-\infty}^{0} x^3 f(x)\,dx\right)E_t\left(h_{t+s}^{3/2}\right), \tag{3}$$

and $E_t\left(\varepsilon_{t+s}\varepsilon^3_{t+s+u}\right)=\tau_z\theta^{(3/2)}_{su}$, with

$$\theta^{(3/2)}_{su}=\frac{3}{4}\left(\tilde{\mu}^{(1)}_{h,s+u}\right)^{1/2}\left(\begin{array}{l} c_4\varphi^{u-1}E_t\left(h^{3/2}_{t+s}\right)+\frac{1}{2}\left(\tilde{\mu}^{(1)}_{h,s+u}\right)^{-1} \\ \left(c_5\gamma^{u-1}E_t\left(h^{5/2}_{t+s}\right)+2\omega c_4\left(\varphi(\varphi-\gamma)^{-1}\left(\varphi^{u-1}-\gamma^{u-1}\right)+\gamma^{u-1}\right)E_t\left(h^{3/2}_{t+s}\right)\right)\end{array}\right),$$

and $c_5=\alpha\left(\alpha\mu^{(5)}_z+2\beta\tau_z\right)+\lambda\left(2\alpha+\lambda\right)\int\limits_{x=-\infty}^{0}x^5f\left(x\right)dx+2\beta\lambda\int\limits_{x=-\infty}^{0}x^3f\left(x\right)dx$; and $c_4=\left(\alpha\tau_z+\lambda\int\limits_{x=-\infty}^{0}x^3f\left(x\right)dx\right)$. Next, $E_t\left(\varepsilon_{t+s}\varepsilon_{t+s+u}\varepsilon^2_{t+s+u+v}\right)=c_4\varphi^{v-1}\theta^{(3/2)}_{su}$ and

$$E_t\left(\varepsilon^2_{t+s}\varepsilon^2_{t+s+u}\right)=\bar{h}\left(1-\varphi^u\right)\tilde{\mu}^{(1)}_{h,s}+\varphi^{u-1}\kappa_z\left(\alpha+\lambda F_0+\kappa_z^{-1}\beta\right)\tilde{\mu}^{(2)}_{h,s};$$

Finally, using a second order Taylor expansion for $h^{5/2}_{t+s}$ around $E_t\left(h_{t+s}\right)$ an approximation for $E_t\left(h^{5/2}_{t+s}\right)$ is: $E_t\left(h^{5/2}_{t+s}\right)\simeq\frac{1}{8}\left(\tilde{\mu}^{(1)}_{h,s}\right)^{1/2}\left(15\tilde{\mu}^{(2)}_{h,s}-7\left(\tilde{\mu}^{(1)}_{h,s}\right)^2\right)$.

Theorem 1: Moments of Forward and Aggregated Returns

The conditional moments of forward one-period returns of model (1) are:

$$\begin{aligned}\tilde{\mu}^{(1)}_{r,s}=\mu,\quad \mu^{(2)}_{r,s}=\tilde{\mu}^{(1)}_{h,s}&=\bar{h}+\varphi^{s-1}\left(h_{t+1}-\bar{h}\right),\\ \tau_{r,s}=\tau_zE_t\left(h^{3/2}_{t+s}\right)\left(\tilde{\mu}^{(1)}_{h,s}\right)^{-3/2}&\simeq\tau_z\left(\tfrac{5}{8}+\tfrac{3}{8}\tilde{\mu}^{(2)}_{h,s}\left(\tilde{\mu}^{(1)}_{h,s}\right)^{-2}\right),\\ \kappa_{r,s}&=\kappa_z\tilde{\mu}^{(2)}_{h,s}\left(\tilde{\mu}^{(1)}_{h,s}\right)^{-2}.\end{aligned}$$

The conditional moments of the aggregated returns of model (1) are:

$$\begin{aligned}\tilde{M}^{(1)}_{r,n}=n\mu,\quad M^{(2)}_{r,n}&=n\bar{h}+\left(1-\varphi\right)^{-1}\left(1-\varphi^n\right)\left(h_{t+1}-\bar{h}\right),\\ \mathrm{T}_{r,n}&\simeq\left(\tau_z\sum_{s=1}^{n}\left(\tfrac{5}{8}\left(\tilde{\mu}^{(1)}_{h,s}\right)^{3/2}+\tfrac{3}{8}\tilde{\mu}^{(2)}_{h,s}\left(\tilde{\mu}^{(1)}_{h,s}\right)^{-1/2}\right)+3\sum_{s=1}^{n}\sum_{u=1}^{n-s}E_t\left(\varepsilon_{t+s}\varepsilon^2_{t+s+u}\right)\right)\left(M^{(2)}_{r,n}\right)^{-3/2},\\ \mathrm{K}_{r,n}&=\left(\begin{array}{l}\kappa_z\sum\limits_{s=1}^{n}\tilde{\mu}^{(2)}_{h,s}+\sum\limits_{s=1}^{n}\sum\limits_{u=1}^{n-s}\left(4E_t\left(\varepsilon_{t+s}\varepsilon^3_{t+s+u}\right)+6E_t\left(\varepsilon^2_{t+s}\varepsilon^2_{t+s+u}\right)\right)\\ +12\sum\limits_{s=1}^{n}\sum\limits_{u=1}^{n-s}\sum\limits_{v=1}^{n-s-u}E_t\left(\varepsilon_{t+s}\varepsilon_{t+s+u}\varepsilon^2_{t+s+u+v}\right)\end{array}\right)\left(M^{(2)}_{r,n}\right)^{-2}.\end{aligned}$$

The first conditional moments $\tilde{\mu}^{(1)}_{r,s}$ and $\tilde{M}^{(1)}_{r,n}$ simply state that, with a constant conditional mean equation, the time t conditional expectation of the s-step-ahead one-period return is equal to the constant conditional mean, whereas the expected return aggregated over n periods scales with time. The second moment of the forward return $\mu_{r,s}^{(2)}$ shows that the conditional expectation of the s-step-ahead variance is equal to the steady state variance $\bar{h}$, plus an exponentially decreasing correction term to account for the distance between the one-step-ahead variance h_{t+1} and the steady state variance $\bar{h}$. Because we assume that the returns are not autocorrelated, the variance of aggregated returns over n time periods $M_{r,n}^{(2)}$ is simply equal to the sum of the s-step-ahead variances for $s = 1, 2, ..., n$.

The expressions for the forward and aggregated skewness are obtained using a second order Taylor series expansion for $E_t\left(h_{t+s}^{3/2}\right)$, as detailed in the Technical Appendix A1. It is easily observed that if the innovation is symmetric ($\tau_z = 0$) then the forward returns distribution is also symmetric. By contrast, considering the expression for $E_t\left(\varepsilon_{t+s}\varepsilon_{t+s+u}^2\right)$ in (3), cumulative returns have an independent source of skewness in addition to that of the innovations τ_z, due to the asymmetric response parameter λ in the conditional variance equation. When returns are negative and extreme, the volatility increases (due to the lambda parameter). Consequently, there are more extreme returns in the lower tail than in the upper tail of the distribution where there is no corresponding increase in the volatility. As a result, aggregated returns can exhibit skewness even if the innovation is symmetric. Figure 1 illustrates the behaviour of the skewness of aggregated returns for the normal GJR model with $\omega = 0.05, \alpha = 0.075$, and $\beta = 0.9$, and for various values of the leverage parameter λ.

For $s = 1$, the forward kurtosis equals the kurtosis of the innovation process. But for $s > 1$, the forward excess kurtosis can be non-zero even when the innovation has zero excess kurtosis, due to the uncertainty in forward variance. The conditional variance of the conditional variance varies with s and, as it must be positive, the forward kurtosis will be greater than the kurtosis of the innovation, whenever $s > 1$, and will itself be time-varying. The net effect of uncertainty in variance is a greater weight in the tails of forward one-period returns. Also, the time-varying conditional variance of the conditional variance introduces dynamics in the higher moments of the forward returns.

Regarding the kurtosis of aggregated returns, with Gaussian innovations this increases

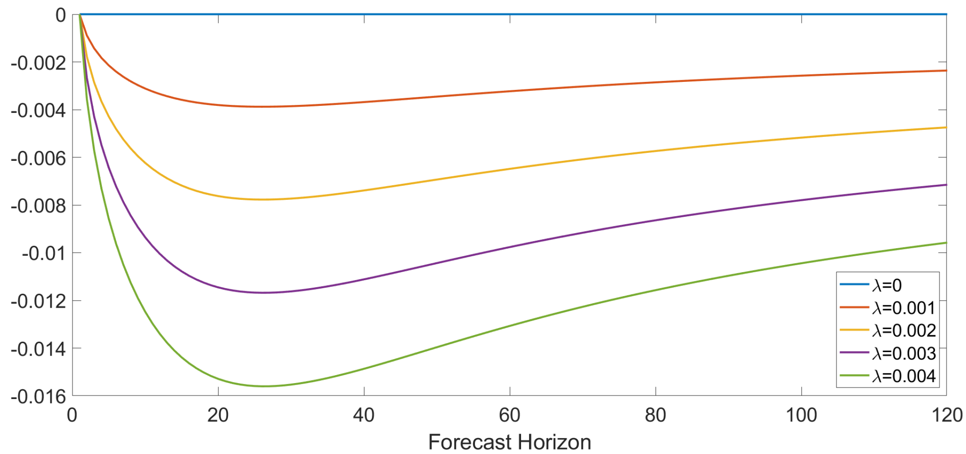

Figure 1: **Skewness of aggregated returns** The behaviour of the skewness of s-step-ahead aggregated returns, as a function of h. We assume returns have GJR conditional variances with $\omega = 0.05, \alpha = 0.075$, and $\beta = 0.9$, and we plot the skewness of aggregated returns $\mathrm{T}_{r,n}$ in Theorem 1 for various values of λ. The innovations z_t are normal.

with s initially, but starts to decrease as the central limit theorem becomes effective and eventually converges to 3. The upper part of Figure 2 illustrates the behaviour of this kurtosis for the GARCH(1,1) model with $\omega = 0.05, \alpha = 0.075$, and $\beta = 0.9$, and for various values of the initial variance σ_{t+1}^2. However, with thick-tailed innovations, there appears to be a third effect which reduces the kurtosis over the first few aggregation steps. For instance, if the innovations are drawn from a unit-variance Student t distribution with 6 degrees of freedom then, with the same parameters for the GARCH(1,1) the kurtosis of aggregated returns as a function of the forecast horizon is shown in the lower part of Figure 2. The reason for this is that for low values of s the fourth moment of aggregated returns increases very quickly (compared with the increase in the variance) but for higher values of s this increase reduces, to match the increase in the variance. The driving factor in this is the parameter φ; as its value increases towards 1 the aggregated kurtosis increases more at low values of s. As a result, it takes longer to converge towards its long-term value of 3. This effect is depicted in Figure 3.

The moments of variances require the following results (see Technical Appendix A2):

$$\tilde{\mu}_{h,s}^{(3)} = \sum_{i=0}^{s-2} c_6^i \left(\omega^3 + 3\omega^2 \varphi \tilde{\mu}_{h,s-i-1}^{(1)} + 3\omega\gamma \tilde{\mu}_{h,s-i-1}^{(2)} \right) + c_6^{s-1} h_{t+1}^3,$$

with

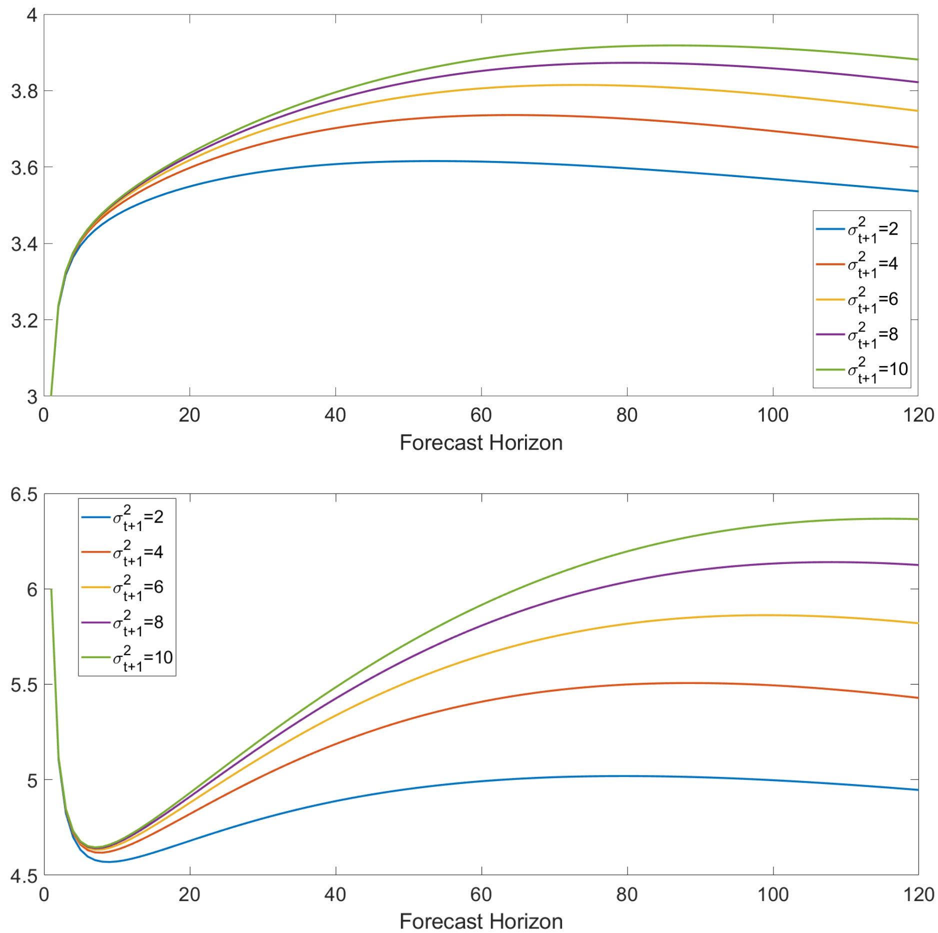

Figure 2: **Kurtosis of aggregated returns for GARCH(1,1)** as a function of forecast horizon s for different starting values of the variance. We assume returns have GARCH(1,1) conditional variances with $\omega = 0.05, \alpha = 0.075$, and $\beta = 0.9$, and we plot the kurtosis of aggregated returns $\mathrm{K}_{r,n}$ in Theorem 1 for various values of the initial variance σ^2_{t+1}. The upper graph assumes normal innovations and the lower graph considers innovations that follow a standardised Student t distribution with 6 degrees of freedom.

$$c_6 = \mu_z^{(6)}\left(\alpha^3 + 3\alpha\lambda\left(\alpha+\lambda\right)F_0 + \lambda^3 F_0\right) + 3\beta\gamma - \beta^2\left(2\beta + 3\left(\alpha + \lambda F_0\right)\right) \tag{4}$$

$$\tilde{\mu}^{(4)}_{h,s} = \sum_{j=0}^{s-2} c_7^j \left(\omega^4 + 4\omega^3\varphi\tilde{\mu}^{(1)}_{h,s-j-1} + 6\omega^2\gamma\tilde{\mu}^{(2)}_{h,s-j-1} + 4\omega c_6\tilde{\mu}^{(3)}_{h,s-j-1}\right) + c_7^{s-1} h^4_{t+1},$$

with

$$\begin{aligned} c_7 &= \mu_z^{(8)}\left(\alpha^4 + F_0\left(\lambda^4 + 4\left(\alpha^3\lambda + \alpha\lambda^3\right) + 6\alpha^2\lambda^2\right)\right) + \beta^4 \\ &+ 4\left[\mu_z^{(6)}\beta\left(\alpha^3 + F_0\left(\lambda^3 + 3\left(\alpha^2\lambda + \alpha\lambda^2\right)\right)\right) + \beta^3\left(\alpha + \lambda F_0\right)\right] + 6\kappa_z\beta^2\left(\alpha^2 + \lambda^2 F_0 + 2\alpha\lambda F_0\right). \end{aligned}$$

Expressions for $\tilde{\mu}^{(i,j,k)}_{h,suv}$, with $i,\ j,\ k \in \{0,1,2\}$ are also required but since most are rather lengthy they are only stated in the Technical Appendix A2, with the proof of the following:

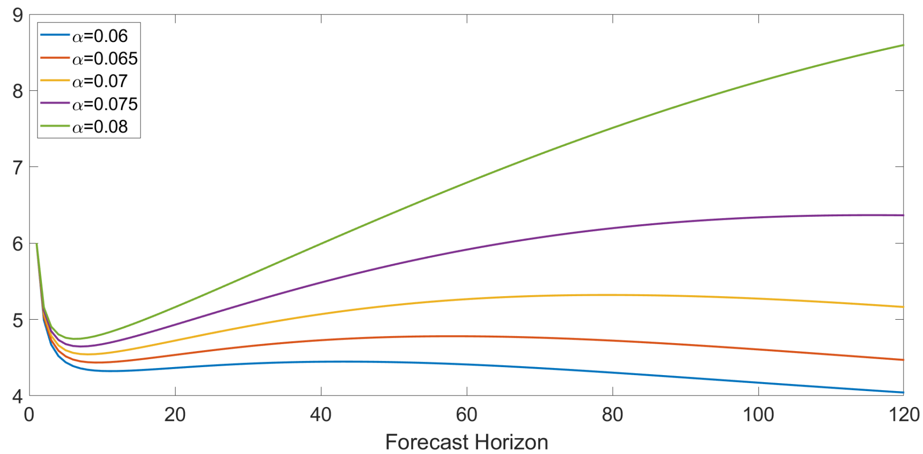

Figure 3: **Kurtosis of aggregated returns for different values of α.** We assume returns have Student t GARCH(1,1) conditional variances with $\mu = 0, \lambda = 0, \beta = 0.9$ where α takes values $0.06, 0.065, 0.07, 0.075$ and 0.08 and the corresponding values of φ are: $0.96, 0.965, 0.97, 0.975$ and 0.98. The initial variance $\sigma^2_{t+1} = 10$.

Theorem 2: Moments of Forward and Aggregated Variances

The conditional moments of forward one-period variances of model (1) are:

$$\tilde{\mu}^{(1)}_{h,s} = \bar{h} + \varphi^{s-1}\left(h_{t+1} - \bar{h}\right), \quad \mu^{(2)}_{h,s} = \tilde{\mu}^{(2)}_{h,s} - \left(\tilde{\mu}^{(1)}_{h,s}\right)^2,$$

$$\tau_{h,s} = \left[\tilde{\mu}^{(3)}_{h,s} - 3\tilde{\mu}^{(2)}_{h,s}\tilde{\mu}^{(1)}_{h,s} + 2\left(\tilde{\mu}^{(1)}_{h,s}\right)^3\right]\left(\tilde{\mu}^{(2)}_{h,s} - \tilde{\mu}^{(1)}_{h,s}\right)^{-3/2},$$

$$\kappa_{h,s} = \left(\tilde{\mu}^{(4)}_{h,s} - 4\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(3)}_{h,s} + 6\left(\tilde{\mu}^{(1)}_{h,s}\right)^2\tilde{\mu}^{(2)}_{h,s} - 3\left(\tilde{\mu}^{(1)}_{h,s}\right)^4\right)\left(\tilde{\mu}^{(2)}_{h,s} - \left(\tilde{\mu}^{(1)}_{h,s}\right)^2\right)^{-2}.$$

The conditional moments of the aggregated future variances of model (1) are:

$$\tilde{M}^{(1)}_{h,n} = n\bar{h} + \left(h_{t+1} - \bar{h}\right)(1-\varphi)^{-1}(1-\varphi^n),$$

$$M^{(2)}_{h,n} = \sum_{s=1}^{n}\left(\tilde{\mu}^{(2)}_{h,s} - \left(\tilde{\mu}^{(1)}_{h,s}\right)^2\right) + 2\sum_{s=1}^{n}\sum_{u=1}^{n-s}\left(\tilde{\mu}^{(1,1)}_{h,su} - \tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u}\right),$$

$$\mathrm{T}_{h,n} = M^{(3)}_{h,n}\left(M^{(2)}_{h,n}\right)^{-3/2},$$

$$M^{(3)}_{h,n} = \sum_{s=1}^{n}\left(\tilde{\mu}^{(3)}_{h,s} - 3\tilde{\mu}^{(2)}_{h,s}\tilde{\mu}^{(1)}_{h,s} + 2\left(\tilde{\mu}^{(1)}_{h,s}\right)^3\right) + 3\sum_{s=1}^{n}\sum_{u=1}^{n-s}A_{h,s,u} + 6\sum_{s=1}^{n}\sum_{u=1}^{n-s}\sum_{v=1}^{n-s-u}B_{h,s,u,v},$$

$$A_{h,s,u} = \tilde{\mu}^{(2,1)}_{h,su} + \tilde{\mu}^{(1,2)}_{h,su} + 2\left(\tilde{\mu}^{(1)}_{h,s} + \tilde{\mu}^{(1)}_{h,s+u}\right)\left(\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u} - \tilde{\mu}^{(1,1)}_{h,su}\right) - \tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(2)}_{h,s+u} - \tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(2)}_{h,s},$$

$$B_{h,s,u,v} = \tilde{\mu}^{(1,1,1)}_{h,suv} - \tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1,1)}_{h,(s+u)v} - \tilde{\mu}^{(1)}_{h,(s+u)}\tilde{\mu}^{(1,1)}_{h,s(u+v)} - \tilde{\mu}^{(1)}_{h,(s+u+v)}\tilde{\mu}^{(1,1)}_{h,su} + 2\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,(s+u)}\tilde{\mu}^{(1)}_{h,(s+u+v)},$$

$$
\mathrm{K}_{h,n} = M^{(4)}_{h,n}\left(M^{(2)}_{h,n}\right)^{-2}
$$

$$
M^{(4)}_{h,n} = \sum_{s=1}^{n}\mu^{(4)}_{h,s} + \sum_{s=1}^{n}\sum_{u=1}^{n-s}C_{h,s,u} + 12\sum_{s=1}^{n}\sum_{u=1}^{n-s}\sum_{v=1}^{n-s-u}D_{h,s,u,v} + 24\sum_{s=1}^{n}\sum_{u=1}^{n-s}\sum_{v=1}^{n-s-u}\sum_{w=1}^{n-s-u-v}E_{h,s,u,v,w},
$$

$$
\begin{aligned}
C_{h,s,u} = {} & 4\left(\begin{array}{l}
\tilde{\mu}^{(3,1)}_{h,su} + \tilde{\mu}^{(1,3)}_{h,su} - 3\left(\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(2,1)}_{h,su} + \tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1,2)}_{h,su}\right) - \left(\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(3)}_{h,s} + \tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(3)}_{h,s+u}\right) \\
+3\tilde{\mu}^{(1)}_{h,s}\left(\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1,1)}_{h,su} + \tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(2)}_{h,s} - \left(\tilde{\mu}^{(1)}_{h,s}\right)^2\tilde{\mu}^{(1)}_{h,s+u}\right) \\
+3\tilde{\mu}^{(1)}_{h,s+u}\left(\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1,1)}_{h,su} + \tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(2)}_{h,s+u} - \tilde{\mu}^{(1)}_{h,s}\left(\tilde{\mu}^{(1)}_{h,s+u}\right)^2\right)
\end{array}\right) \\
& +6\left(\begin{array}{l}
\tilde{\mu}^{(2,2)}_{h,su} - 2\left(\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1,2)}_{h,su} + \tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(2,1)}_{h,su}\right) + \left(\tilde{\mu}^{(1)}_{h,s}\right)^2\tilde{\mu}^{(2)}_{h,s+u} \\
+\left(\tilde{\mu}^{(1)}_{h,s+u}\right)^2\tilde{\mu}^{(2)}_{h,s} + 4\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1,1)}_{h,su} - 3\left(\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u}\right)^2
\end{array}\right),
\end{aligned}
$$

$$
\begin{aligned}
D_{h,s,u,v} = {} & \tilde{\mu}^{(2,1,1)}_{h,suv} - 2\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1,1,1)}_{h,suv} - \tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(2,1)}_{h,s(u+v)} - \tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(2,1)}_{h,su} + \left(\tilde{\mu}^{(1)}_{h,s}\right)^2\tilde{\mu}^{(1,1)}_{h,(s+u)v} \\
& +2\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1,1)}_{h,s(u+v)} + 2\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(1,1)}_{h,su} + \tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(2)}_{h,s} \\
& -3\left(\tilde{\mu}^{(1)}_{h,s}\right)^2\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1)}_{h,s+u+v} + \tilde{\mu}^{(1,2,1)}_{h,suv} - 2\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1,1,1)}_{h,suv} - \tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(2,1)}_{h,(s+u)v} - \tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(1,2)}_{h,su} \\
& +\left(\tilde{\mu}^{(1)}_{h,s+u}\right)^2\tilde{\mu}^{(1,1)}_{h,s(u+v)} + 2\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1,1)}_{h,(s+u)v} + 2\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(1,1)}_{h,su} \\
& +\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(2)}_{h,s+u} - 3\tilde{\mu}^{(1)}_{h,s}\left(\tilde{\mu}^{(1)}_{h,s+u}\right)^2\tilde{\mu}^{(1)}_{h,s+u+v} + \tilde{\mu}^{(1,1,2)}_{h,suv} - 2\tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(1,1,1)}_{h,suv} \\
& -\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1,2)}_{h,(s+u)v} - \tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1,2)}_{h,s(u+v)} + \left(\tilde{\mu}^{(1)}_{h,s+u+v}\right)^2\tilde{\mu}^{(1,1)}_{h,su} + 2\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(1,1)}_{h,(s+u)v} \\
& +2\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(1,1)}_{h,s(u+v)} + \tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(2)}_{h,s+u+v} - 3\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u}\left(\tilde{\mu}^{(1)}_{h,s+u+v}\right)^2,
\end{aligned}
$$

$$
\begin{aligned}
E_{h,s,u,v,w} = {} & \tilde{\mu}^{(1,1,1,1)}_{h,suvw} - \tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(1,1,1)}_{h,su(v+w)} - \tilde{\mu}^{(1)}_{h,s+u+v+w}\tilde{\mu}^{(1,1,1)}_{h,suv} + \tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(1)}_{h,s+u+v+w}\tilde{\mu}^{(1,1)}_{h,su} - \tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1,1,1)}_{h,(s+u)vw} \\
& +\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(1,1)}_{h,(s+u)(v+w)} + \tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u+v+w}\tilde{\mu}^{(1,1)}_{h,(s+u)v} - \tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1,1,1)}_{h,(s+u)vw} + \tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(1,1)}_{h,s(u+w+v)} \\
& +\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1)}_{h,s+u+v+w}\tilde{\mu}^{(1,1)}_{h,s(u+v)} + \tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1,1)}_{h,(s+u+v)w} - 3\tilde{\mu}^{(1)}_{h,s}\tilde{\mu}^{(1)}_{h,s+u}\tilde{\mu}^{(1)}_{h,s+u+v}\tilde{\mu}^{(1)}_{h,s+u+v+w}.
\end{aligned}
$$

Ishida and Engle (2002) and others argue that the conditional variance of the conditional variance grows faster than linearly with the current variance. Our formula for $\mu^{(2)}_{h,s}$ shows that the conditional variance of the forward conditional variance is a quadratic function of the current variance h_{t+1}, in model (1). Hence, the uncertainty around the point variance forecast increases much more than linearly when variance levels are high, much reducing the reliability of the point forecast. This highlights the importance of our analytic formulae for the higher moments of GARCH variances.[8]

[8]Intuitively, we expect distributions of forward variances to be positively skewed, since jumps in variance

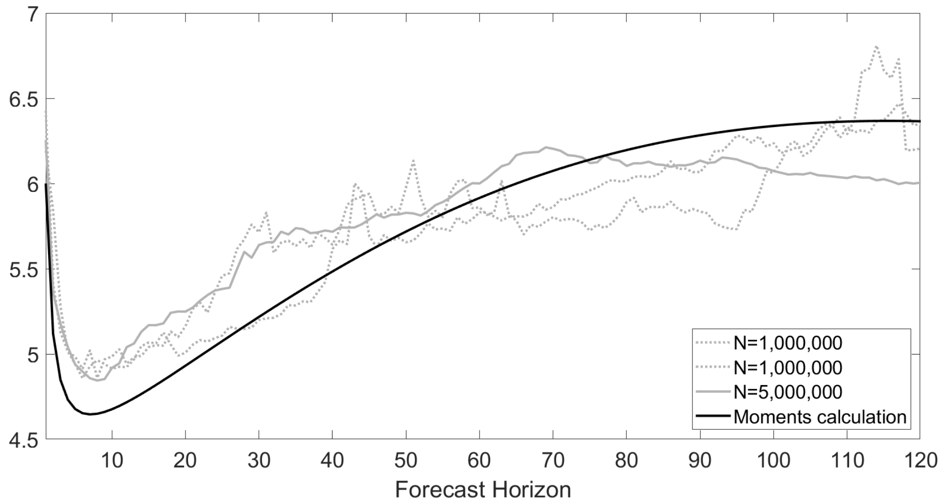

Figure 4: **Simulated kurtosis of aggregated returns.** The figure compares our theoretical values with simulated values of the kurtosis of aggregated returns for the Student t-GARCH(1,1) model. The black line is our theoretical value, the grey line is based on 5 million simulations for 120 steps ahead of forward returns, and the dotted grey lines are two other simulated values for kurtosis based on 1 million simulations. The parameter values are: $\mu = 0$, $\omega = 0.05$, $\alpha = 0.075$, $\beta = 0.9$, df $= 6$, and $\sigma^2_{t+1} = 10$.

3 Simulation and Bootstrapping Results

Simulations offer simplicity, and with modern computers they are also fast, but they can be inaccurate. To support this observation Figure 4 compares our theoretical formulae for the kurtosis of aggregated returns 120 days ahead with their simulated values.[9] We generate this figure by taking $N \times 120$ simulations with standardised Student t innovations. We conclude that simulations can easily over- or under-estimate kurtosis, even with $N = 5$ million runs.

To examine this bias further, we undertook a bootstrapping exercise. For this, we assume a standardised symmetric t-GARCH(1,1) process with 6 degrees of freedom and the following base set of GARCH parameters: $\mu = 0, \omega = 0.05, \alpha = 0.075; \beta = 0.9$. Our aim is to estimate the distribution of the skewness and kurtosis of the aggregated returns using a bootstrap approach, and to compare this with our closed-form expressions for the skewness and kurtosis, up to 120 days ahead. We construct two distributions, one via bootstrapping and one using

are usually positive rather than negative. In an empirical implementation of the moments formulae derived in this section we find that the skewness of forward variance is indeed positive, for all horizons and all three samples considered; we also find that the excess kurtosis of variance is always positive. These empirical results are excluded from this paper for reasons of space, but can be obtained from the authors on request.

[9]We choose 120 days because it is the longest risk horizon recommended by Basel III banking regulations.

the closed-form formulae, for the skewness and kurtosis of aggregated returns. Assuming an initial variance of $\sigma_{t+1}^2 = 10$, we simulate a time series of 5,000 t-GARCH returns based on the above parameters. Then we re-estimate the parameters of the t-GARCH(1,1) model.[10] We repeat this simulation and parameter re-estimation exercise 1,000 times and each time we also save the standardised residuals, for use in the bootstrap.

Using the closed-form expressions derived in Theorem 1, we calculate the skewness and kurtosis of the returns aggregated over n days for $n = 1, ...120$, for each set of re-estimated t-GARCH(1,1) paramaters. For the symmetric t-GARCH(1,1) process the skewness $\mathrm{T}_{r,n}$ is always zero, but the kurtosis $\mathrm{K}_{r,n}$ is typically non-zero. Hence, based on 1,000 sets of re-estimated GARCH parameters, we have generated a simulated distribution for the kurtosis of aggregated returns. We shall illustrate this distribution using its quantiles in Figure 5. In the bootstrap approach, for each set of re-estimated GARCH parameters, we:

(a) Simulate t-GARCH returns, 120 days ahead, using the bootstrapped returns from the saved residuals and aggregate the returns over n days ahead, for $n = 1, ...120$;

(b) Repeat step (a) 1,000 times;

(c) Calculate the skewness and kurtosis of the aggregated returns, for $n = 1, ..., 120$.

This yields an empirical distribution of 1,000 values for the skewness and kurtosis of aggregated returns, one value for each set of re-estimated parameters. In Figure 5 we plot the means and the $[10\%, 90\%]$ and $[25\%, 75\%]$ confidence intervals of the skewness and kurtosis of aggregated returns, computed using the closed-form formulae (only for the kurtosis) and via bootstrapping (for both skewness and kurtosis), and compare them with the values given by the closed-form formulae using the base set of parameters.

We draw the following conclusions: (i) the bootstrapped skewness is unbiased; (ii) the bootstrapped kurtosis presents a downwards bias; (iii) the 'true' value of the kurtosis lies outside the inter-quartile range for the bootstrapped kurtosis, in most cases (the effect is

[10]The parameters were re-estimated using the R statistical package for t-GARCH. We tried using fewer returns but there was considerable variation in the re-estimated paramaters. There was some variation in the re-estimated parameters even with a sample of size 5,000 but, for instance, with only 1,000 simulated returns the re-estimated α parameter varied between 0.05 and 0.1 while the β varied between 0.95 and 0.85 – most of the time. In fact, a few values lay outside this range. Results available on request.

more visible at longer horizons); and (iv) the kurtosis estimate computed using the closed-form formulae is unbiased.

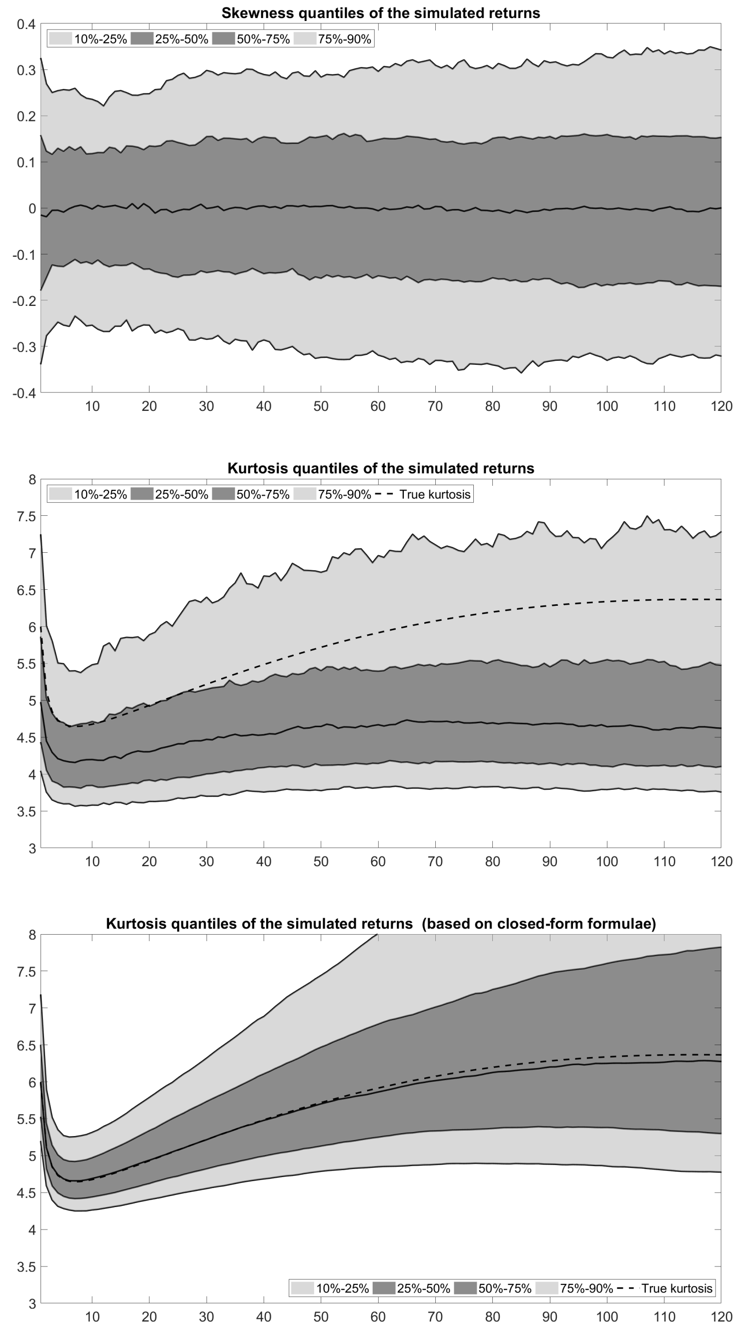

Figure 5: **Skewness and kurtosis of aggregated returns – bootstrapping exercise.** The figure compares our theoretical values with bootstrapped values of the skewness and kurtosis of aggregated returns for the Student t-GARCH(1,1) model. The dotted line is our theoretical value, the black line is the mean based on 1,000 bootstraps for n aggregated returns, with $n = 1, ..., 120$ along the horizontal axis. The inter-quartile range and the $[10\%, 90\%]$ confidence intervals are depicted for each distribution. The base parameter values for the bootstrap are: $\mu = 0, \omega = 0.05, \alpha = 0.075, \beta = 0.9$, df $= 6$, and we set $\sigma_{t+1}^2 = 10$.

4 Limits of the Moments

The limits of the forward and aggregated mean are trivial and immediate, but the convergence behaviour of the conditional moments as the time horizon increases are more complex:

Theorem 3: Limits of Moments of Forward and Aggregated Returns

Suppose $\varphi \in (0,1)$ and $\varphi \neq \gamma$ for model (1). Then:[11]

$$\lim_{s\to\infty} \mu_{r,s}^{(2)} = \bar{h}, \tag{5}$$

$$\lim_{s\to\infty} \tau_{r,s} = \begin{cases} \tau_z \left(\frac{5}{8} + \frac{3}{8}\left(\omega^2 + 2\omega\varphi\bar{h}\right)(1-\gamma)^{-1}\left(\bar{h}\right)^{-2}\right) & \text{if } \gamma \in (0,1), \\ \operatorname{sgn}(\tau_z)\,\infty & \text{if } \gamma \in [1,\infty), \end{cases} \tag{6}$$

$$\lim_{s\to\infty} \kappa_{r,s} = \begin{cases} \kappa_z \left(\omega^2 + 2\omega\varphi\bar{h}\right)(1-\gamma)^{-1}\left(\bar{h}\right)^{-2} & \text{if } \gamma \in (0,1), \\ \infty & \text{if } \gamma \in [1,\infty), \end{cases} \tag{7}$$

$$\lim_{n\to\infty} \frac{M_{r,n}^{(2)}}{n} = \bar{h}, \tag{8}$$

$$\lim_{n\to\infty} \mathrm{T}_{r,n} = \begin{cases} 0 & \text{if } \gamma \in (0,1), \\ \operatorname{sgn}\left(\tau_z\left(\alpha + \frac{\gamma-\varphi}{3}\right) + \lambda \int\limits_{x=-\infty}^{0} x^3 f(x)\,dx\right)\infty & \text{if } \gamma \in [1,\infty), \end{cases} \tag{9}$$

$$\lim_{n\to\infty} \mathrm{K}_{r,n} = \begin{cases} 3 & \text{if } \gamma \in (0,1), \\ 3 + \frac{\kappa_z}{2}\left(1-\varphi^2\right)\left(1 + 6\left(\alpha + \lambda F_0 + \kappa_z^{-1}\beta\right)(1-\varphi)^{-1}\right) + \operatorname{sgn}\left(|\lambda| + |\tau_z|\right)\infty & \text{if } \gamma = 1, \\ \infty & \text{if } \gamma \in (1,\infty). \end{cases} \tag{10}$$

Hence, under suitable parameter conditions, the conditional moments of forward one-period returns converge to finite limits that are the unconditional counterparts of the respective conditional moments, and these parameter conditions are the necessary and sufficient condi-

[11] $\operatorname{sgn}(x) = \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0 \end{cases}$ and we use the convention that $\operatorname{sgn}(0)\infty = 0$.

tions for the existence of the corresponding unconditional moments. Indeed, $\varphi \in (0,1)$ is a necessary and sufficient condition for the existence of the unconditional variance, as can be shown using Theorem 2.2 (and Example 2.1) in Ling and McAleer (2002b), and for $\varphi \in (0,1)$ a steady-state level of variance exists, i.e. $\exists\, h_0$ such that $E(h_t) = h_0$, for any $t \in N$. It is easy to show that, when it exists, the unconditional variance h_0 is given by: $h_0 = \frac{\omega}{1-\varphi} = \bar{h}$.[12] Thus we obtain that for $\varphi \in (0,1)$, $\lim_{s\to\infty} \mu_{r,s}^{(2)} = \lim_{s\to\infty} E_t(h_{t+s}) = \bar{h} = E(h_{t+s})$.

Again, using Theorem 2.2 from Ling and McAleer (2002b), a necessary and sufficient condition for the existence of the fourth unconditional moment is $\gamma \in (0,1)$. This is the condition required in Theorem 3 for the fourth conditional moment to converge to a finite limit. It is easy to show that, when it exists, the fourth unconditional moment is given by:

$$E\left(\varepsilon_t^4\right) = \kappa_z E\left(h_t^2\right) = \kappa_z \frac{\omega^2 + 2\omega\varphi\bar{h}}{1-\gamma} = \kappa_z c_1$$

and, as a result, the unconditional kurtosis is given by the same expression as in (7) above, for $\gamma \in (0,1)$. A special case of this result is the unconditional kurtosis for a GARCH(1,1) process with symmetric innovations, derived by Ishida and Engle (2002).

The unconditional skewness is given by:

$$\frac{E\left(\varepsilon_t^3\right)}{\left[E\left(\varepsilon_t^2\right)\right]^{3/2}} = \frac{E\left(z_t^3 h_t^{3/2}\right)}{\left[E\left(h_t\right)\right]^{3/2}} = \tau_z \frac{E\left(h_t^{3/2}\right)}{\left[E\left(h_t\right)\right]^{3/2}}.$$

Since $E\left(h_t^{3/2}\right)$ cannot be computed analytically in this framework, we use a second order Taylor series expansion to approximate it. Thus, $E\left(h_t^{3/2}\right) \simeq \frac{5}{8}\left(E\left(h_t\right)\right)^{3/2} + \frac{3}{8} E\left(h_t^2\right)\left(E\left(h_t\right)\right)^{-1/2}$ and a resulting approximation of the unconditional skewness is the same as the expression given in (6) for $\gamma \in (0,1)$.

For the aggregated returns, if gamma is below 1, the limit of the skewness is zero, and the limit of the excess kurtosis is zero as well, which illustrates the results of the central limit theorem that the limiting distribution is normal. The same conditions are needed for the convergence of s-step ahead skewness and kurtosis, as for the convergence of skewness

[12] Applying the expectation operator on both sides of the equation for the GJR conditional variance and using that the indicator I_t^- and the even powers of the contemporaneous innovations ε_t^{2k}, where $k \in N$, are independent, we get: $E(h_t) = \omega + \alpha\, E\left(\varepsilon_{t-1}^2\right) + \lambda E\left(\varepsilon_{t-1}^2\right) F_0 + \beta\, E(h_t)$. Using that $E_{t-2}\left(\varepsilon_{t-1}^2\right) = h_{t-1}$ and the tower law of expectations, we can write: $E(h_t) = \omega + (\alpha + \lambda F_0 + \beta)\; E(h_t)$, which yields $E(h_t) = \frac{\omega}{1-\varphi} = \bar{h}$.

and kurtosis of aggregated returns. This is in line with the results of Francq and Zakoian (2010), Appendix A.3.3 with respect to the limit of stationary martingale differences. The conditional moments of the aggregated returns converge to the corresponding moments of a normal distribution, provided that certain parameter conditions are met. Outside of the regularity conditions, the conditional skewness of aggregated returns diverges to $\pm\infty$ and the conditional kurtosis of aggregated returns diverges to $+\infty$. This is similar to a result of Diebold (1988) who shows that the unconditional distribution of the aggregated returns for a conditionally normal AR-ARCH (m, p) process also converges to a normal distribution, under suitable parameter conditions.

Interestingly, identical convergence conditions apply for the moments of both forward and aggregated returns. Whenever the moments of forward returns converge to the unconditional moments, the aggregated moments converge to the corresponding moments of a normal distribution. Moreover, for a special case of the generic framework, namely for the normal GARCH(1,1) model with $\gamma = 1$, the limit of the kurtosis of forward returns is infinite whilst the kurtosis of aggregated returns converges to a constant value different from 3. In fact, this additional convergence case for $\gamma = 1$ is not specific to the normal GARCH(1,1): it applies to any GARCH(1,1) model with symmetric innovations. This result, for the symmetric special case, is in agreement with Breuer and Jandacka (2010) even though our proof is different from theirs.

Theorem 4: Limits of Moments of Forward and Aggregated Variances

Suppose $\varphi \in (0, 1)$ and $\gamma \neq \varphi$ (as above); additionally $c_6 \neq \gamma$ and $c_6 \neq \varphi$. Then we have:

a) The limit of the conditional variance of the forward conditional variance of model (1) is:

$$\lim_{s\to\infty} \mu_{h,s}^{(2)} = \begin{cases} \left(\left(\omega^2 + 2\omega\varphi\bar{h}\right)(1-\gamma)^{-1} - \bar{h}^2\right) & \text{if } \gamma \in (0,1)\,, \\ \infty & \text{if } \gamma \in [1,\infty)\,. \end{cases} \tag{11}$$

b) The limit of the conditional variance of the aggregated conditional variance (per unit of time) of model (1) is:

$$\lim_{n\to\infty}\frac{M_{h,n}^{(2)}}{n}=\begin{cases}\left(\left(\omega^2+2\omega\varphi\bar{h}\right)(1-\gamma)^{-1}-\bar{h}^2\right)\left(1+2\varphi(1-\varphi)^{-1}\right) & \text{if } \gamma\in(0,1)\,,\\ \infty & \text{if } \gamma\in[1,\infty)\,.\end{cases} \tag{12}$$

c) The limit of the conditional skewness of the forward conditional variance of model (1) is:

$$\lim_{s\to\infty}\tau_{h,s}=\begin{cases}M_1 & \text{if } c_6\in(0,1)\,,\\ \infty & \text{if } c_6\in[1,\infty)\,,\end{cases} \tag{13}$$

where

$$M_1=\frac{\omega\left(\omega^2+3\omega\varphi\bar{h}+3\gamma c_1\right)(1-c_6)^{-1}-3\bar{h}c_1+2\bar{h}^3}{\left(c_1-\bar{h}^2\right)^{3/2}}$$

d) For $\gamma\in(0,1)$ the conditional skewness of the aggregated conditional variance of model (1) has limit:[13]

$$\lim_{n\to\infty}\mathrm{T}_{h,n}=\begin{cases}0 & \text{if } c_6\in(0,1)\,,\\ \infty & \text{if } c_6\in(1,\infty)\,,\\ \operatorname{sgn}(N)\infty & \text{if } c_6=1,\end{cases} \tag{14}$$

where

$$\begin{aligned}N=&\frac{\omega\left(\omega^2+3\omega\varphi\bar{h}+3\gamma c_1\right)}{2}+3\frac{\bar{h}}{2}\left(c_1+\varphi\left(\omega^2+3\omega\varphi\bar{h}+3\gamma c_1\right)+\omega^2(1-\gamma)^{-1}+2\omega\varphi\bar{h}\right)\\&+3\gamma(1-\gamma)^{-1}\frac{\left(\omega^2+3\omega\varphi\bar{h}+3\gamma c_1\right)}{2}\left(\omega+2\varphi\bar{h}\right)\\&+3(1-\varphi)^{-1}\bar{h}\left[c_1\left(1+\varphi\right)-2\varphi\bar{h}^2+\bar{h}\left(\left(h_{t+1}-\bar{h}\right)-\omega\right)-\varphi\left(2c_1-\bar{h}^2\right)\right].\end{aligned}$$

The returns process has no autocorrelation so the variance of aggregated returns is just the sum of the forward one-period variances. However, the variance process is autocorrelated. As a result the variances of the two processes have different limiting behaviour. The limit of the variance of aggregated returns per unit of time is equal to the limit of forward variance (i.e. $\lim_{n\to\infty}\frac{\mathrm{M}_{r,n}^{(2)}}{n}=\lim_{s\to\infty}\mu_{r,s}^{(2)}$), but the same does not hold for the variance of variance. Indeed, $\lim_{n\to\infty}\frac{\mathrm{M}_{h,n}^{(2)}}{n}>\lim_{s\to\infty}\mu_{h,s}^{(2)}$.

[13] Since proofs become increasingly lengthy we only state the limit of the conditional skewness of the aggregated conditional variance in the case that $\gamma\in(0,1)$, and for $\gamma\geq1$ we present the principles of the derivation in the Technical Appendix A4.

Of course, the returns and variance processes are not independent, and certain aspects of this dependence are reflected in a similar behaviour in their limiting distributions. In particular, recall that the forward and aggregated returns had identical regularity conditions and that whenever a moment of forward returns converges to a finite limit, the corresponding moment of aggregated returns converges to the normal value. Theorem 4 yields a similar result for forward and aggregated variances: the skewness of the aggregated variance converges to zero when the skewness of forward variance converges to a finite value. Figure 6 shows the convergence areas for the moments of the returns and variance.

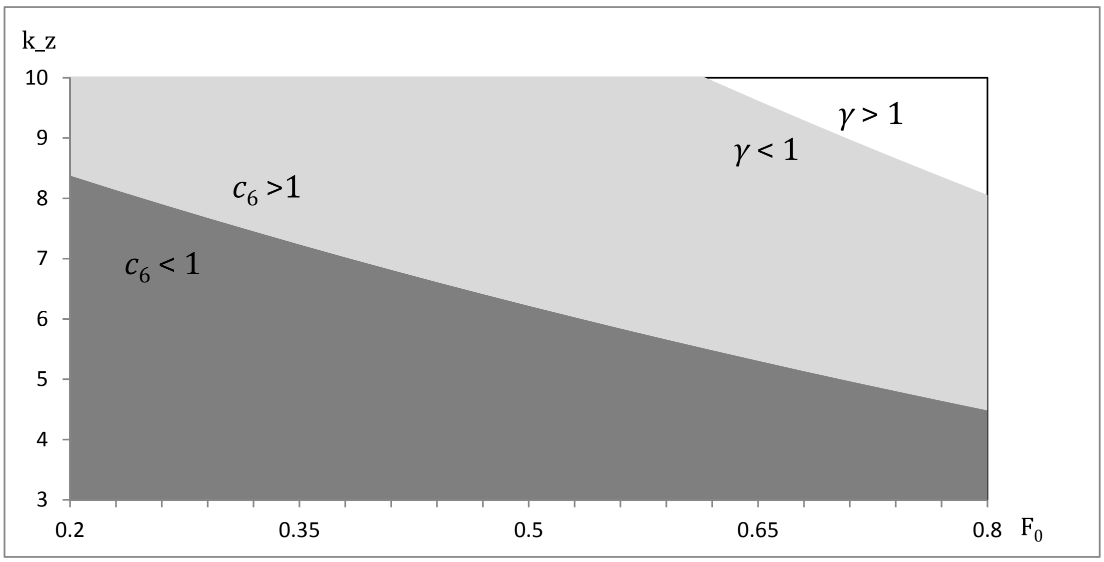

Figure 6: **Convergence of moments** The convergence areas for the moments of the returns and variances are illustrated, as a function of F_0 and κ_z. The parameter values used are $\omega = 0.000075, \alpha = 0.03887, \lambda = 0.007846$ and $\beta = 0.9476$.

How useful are these results in practice? In the previous section we have compared our theoretical results with simulated and bootstrapped values for these moments and we have demonstrated that both simulated and bootstrapped moments can be highly inaccurate. Another point to note is that our results yield parameter conditions for the existence of finite values for the skewness and kurtosis of the aggregated returns. In the next section we shall apply some standard models to daily log returns on some standard risk factors for equities, foreign exchange rates and interest rates. We used the daily log returns on the S&P 500 index, the daily log returns on the EUR/USD exchange rate, and the daily basis point changes in the 3-month US Treasury bill rate, from 1990 to 2018. The GARCH(1,1) and GJR normal and Student t GARCH models were estimated using these data, with estimations being based on an in-sample period of about 10 years of daily returns. All models were rolled forward daily, then used for forecasting s-step ahead forward and aggregated returns and variances.

We were rather surprised by our findings, which were derived from both the EVIEWS package and the MatLab Financial Econometrics (MFE) tool box provided by Kevin Sheppard, each being standard tools for academic research on GARCH models. For instance, based on the out-of-sample period 1 Jan 2018 to 30 March 2018 2018, most of these models returned a value for γ which is greater than 1, implying that both the (unconditional) skewness and kurtosis of these returns are infinite, and that these moments of aggregated returns will diverge as the risk horizon increases.[14]

5 An Application: Approximate Predictive Distributions

Density forecasting is a prime application of our moment formulae.[15] Whilst one must know all the moments of a random variable to determine its distribution,[16] it can be approximated based on the first few moments alone. Based on their relative merits and drawbacks, and given the frequency of their use in similar applications as well as the feasibility of obtaining

[14]Earlier estimations did not have this problem, except for the interest rate data – more detailed results are available on request. The only exceptions were the EUR/USD exchange rate models with normal innovations, where few of the estimations returned a value for γ which implied divergence skewness and kurtosis.

[15]Another possible application in finance is the computation of Value-at-Risk and Expected Shortfall of portfolios, for regulatory purposes. This is a natural extension of the moments formulae, which can to be used jointly with a statistical approximation method such as: the Johnson SU distribution, illustrated in Simonato (2010), or the Gram-Charlier expansions (for example the Edgeworth expansion discussed below). In this second case the formulae are lengthy (available from the authors on request). Also, a working formula for the Edgeworth Expected Shortfall is published in Boudt, Peterson and Croux (2009), which can be modified by calculating VaR via an inversion of the Edgeworth expansion (as opposed to a Cornish-Fisher approximation). Alternatively, the Cornish-Fisher ES formulae of Christoffersen (2012) can be used. These are all time-efficient approximations, the efficiency of which needs further investigation.

[16]To be precise, a distribution is uniquely determined by its moments only if the Carleman condition holds, i.e. only if $\sum_{n=1}^{\infty} \alpha_{2n}^{-0.5n} \to \infty$, where $\{\alpha_k\}$ is the moment sequence (see Serfling, 1980, p.46). In the following we assume that this condition is met.

an approximate distribution function in closed form, we have selected two approximation methods: the Edgeworth expansion and the Johnson SU distribution.

5.1 Distribution Approximations Methods

The first paper to use the Edgeworth expansion to approximate the multi-step-ahead distribution in GARCH models is Baillie and Bollerslev (1992).[17] The Edgeworth expansion can approximate a density of interest around a base density, usually the standard normal density.[18] It belongs to the class of Gram-Charlier expansions, being a rearrangement of a Gram-Charlier A series.[19] However, Gram-Charlier A series and Edgeworth series are only equivalent asymptotically, when an infinite number of terms enter the expansions. In empirical applications using finite order approximations they can differ significantly, and the Edgeworth version is preferred because it is an asymptotic expansion.[20] Nevertheless, the Edgeworth expansion may have monotonicity and convergence problems, i.e. the distribution function is not guaranteed to be monotonic and the error of approximation does not necessarily improve when we increase the order of the expansion.[21]

The first four terms of the Edgeworth expansion are:

$$f_x\left(x\right) \simeq f_x^E\left(x\right) = \varphi\left(x\right) - \tfrac{\tau_x}{6}\varphi^{(3)}\left(x\right) + \tfrac{(\kappa_x-3)}{24}\varphi^{(4)}\left(x\right) + \tfrac{\tau_x^2}{72}\varphi^{(6)}\left(x\right),$$

where $f_x^E\left(x\right)$ is the second-order Edgeworth approximation of f_x, so moments (cumulants) of order higher than four (kurtosis) are ignored, φ is the standard normal density and $\varphi^{(k)}$

[17]They constructed multi-step-ahead prediction error intervals in ARMA-GARCH models and, assuming a GARCH(1,1) error term with symmetric innovation density, explicitly computed the second and fourth conditional moments of the s-step-ahead forecast error at forecast origin t. Then they use the Cornish-Fisher expansion to approximate the quantiles of the multi-step-ahead prediction error distribution. Similarly, Alexander et al. (2013) use analytic higher-order moments to compute the multi-step-ahead Value-at-Risk from GARCH models using the Cornish-Fisher expansion.

[18]For the general theory and expansion see Edgeworth (1905), Wallace (1958) and Bhattacharya and Ghosh (1978).

[19]See Chebyshev (1860), Chebyshev (1890), Gram (1883), Charlier (1905) and Charlier (1906).

[20]An asymptotic expansion is defined in Wallace (1958) as one where the error of approximation approaches zero as one of its parameters, e.g. the sample size for approximations of the sampling distribution of a random sample of size T, approaches infinity. Furthermore, Wallace (1958) calls the Edgeworth expansion a 'formal' asymptotic expansion.

[21]See Jasche (2002) and Wallace (1958).

is its k^{th} derivative, and τ_x and κ_x denote the skewness and kurtosis of f_x. For our purposes f_x will be the density of the normalised forward returns.

A random variable x is said to follow a Johnson SU distribution (Johnson, 1949) if:

$$z = \gamma + \delta \sinh^{-1}\left(\tfrac{x-\xi}{\lambda}\right),$$

where z is a standard normal variable. The four parameters γ, δ, ξ and λ may be estimated using the moment-matching algorithm described in Tuenter (2001). Although flexible, the main disadvantage of this approach is that a Johnson SU distribution is not guaranteed to exist for any set of mean, variance, skewness and (positive) excess kurtosis.

5.2 Evaluation Methods

To assess how well these approximate distributions serve their purpose we should investigate whether they provide an adequate representation of the conditional distributions of forward returns. But these distributions are not observable, even ex-post, so we shall use simulated distributions as proxies. The null hypothesis is H_0: $F_m = F_s$, where F_m is the cumulative distribution function for the approximate distribution constructed using the first four moments and a specific approximation method, and F_s is the distribution function for the simulated forward returns based on the GARCH process. The simulated distribution F_s is given by the step-function of the sample. Thus, $F_s(x_i) = T^{-1}i$, where i is the number of returns less than or equal to x_i, with x_i being an increasingly ordered sample for $i \in \{1, ..., T\}$, and T is the number of simulations.

The Kolmogorov-Smirnov (KS) and Anderson Darling (AD) tests are standard hypothesis tests where the null is the equality of two distributions.[22] The KS test, proposed by Kolmogorov (1933) and Smirnov (1939), is based on the maximum difference D between an empirical and a hypothetical cumulative distribution. The test statistic is KS $= \sqrt{T}D$, where $D = \max_x |F_m(x) - F_s(x)|$. For practical implementations, we use the following version:[23]

[22]For alternatives and extensions of these tests see Diebold et al. (1999), Clements and Smith (2000 and 2002) and Corradi and Swanson (2006a and 2006b).

[23]See Anderson and Darling (1952) and Pearson and Hartley (1972).

$$\text{KS} = \sqrt{T} \max_{1\leq i\leq T} \left\{ \max \left[F_m\left(x_i\right) - \tfrac{i-1}{T}, \tfrac{i}{T} - F_m\left(x_i\right) \right] \right\}.$$

When comparing alternative models, the one with the lowest KS value is deemed the most accurate for predicting the distribution in question.

The framework proposed by Anderson and Darling (1952) is more flexible, allowing for different weighting of the observations. They propose two distance measures, which are actually generalisations of the KS and Cramer von Mises (CVM) statistics.[24] The respective test statistics are given by:

$$\text{AD}_1 = T^{1/2} \max_x \left| F_m\left(x\right) - F_s\left(x\right) \right| \left(\psi(F_m\left(x\right))^{1/2} \right), \tag{15}$$

$$\text{AD}_2 = T \int_x \left[\left(F_m\left(x\right) - F_s\left(x\right)\right)^2 \psi\left(F_m\left(x\right)\right) \right], \tag{16}$$

where ψ is a weighting function. Following convention we refer to the Anderson-Darling (AD) test as (16) with a weighting function $\psi\left(x\right) = \left(x\left(1-x\right)\right)^{-1}$.[25]

Conducting these tests in our setting requires the simulation of critical values. The statistics only have standard distributions if the distribution under the null hypothesis is entirely pre-specified, but in our case the F_m distribution relies on estimated parameter values so the theoretical critical values are no longer applicable.

5.3 Data and Methodology

The performance of our proposed distribution forecasts is tested using daily observations on an equity index (S&P 500), a foreign exchange rate (Euro/dollar) and an interest rate (3-month Treasury bill). These series represent three major market risk types and within each class they represent the most important risk factors in terms of volumes of exposures. Data were obtained from Datastream and each comprise almost 30 years of daily data from

[24] The KS and Cramer-von Mises tests are obtained when $\psi(t) = 1$ in (15) and (16) respectively.

[25] For practical implementations and for an increasingly ordered sample, we use the following versions of the CVM and AD test statistics (see Anderson and Darling, 1952 and Pearson and Hartley, 1972):

$$\text{CVM} = \sum_{i=1}^{T} \left[F_m\left(x_i\right) - \tfrac{2i-1}{2T} \right]^2 + \tfrac{1}{12T}, \quad \text{AD} = -\sum_{i=1}^{T} \tfrac{2i-1}{T} \left[\ln\left(F_m\left(x_i\right)\right) + \ln\left(1 - F_m\left(x_{T+1-i}\right)\right) \right] - T.$$

1st January 1990 to 30th March 2018.[26] Figure 7 plots the daily log returns for the equity and exchange rate data and the daily changes in the interest rate.[27]

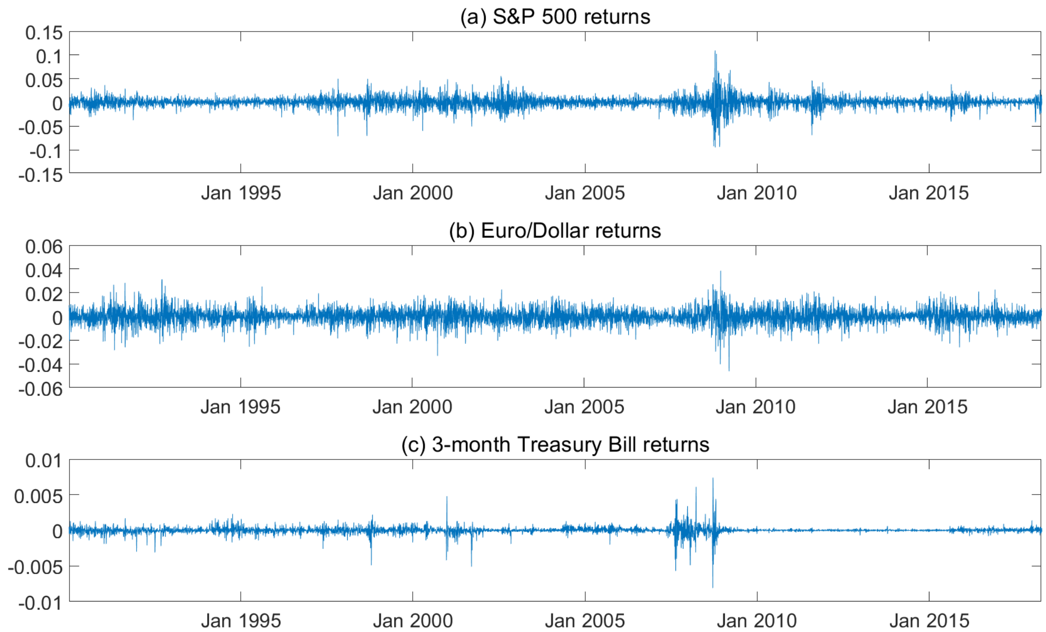

Figure 7: **Returns** The equity and exchange rate daily (log) returns are computed as the first differences of the logarithm of the S&P 500 index values and Euro/dollar exchange rates, respectively. The interest rate returns are computed as first differences in interest rate values.

	S&P 500	EUR/USD	3M Bill
Mean	0.00028	-7.9E-06	-8.3E-06
Maximum	0.11	0.0384	0.0074
Minimum	-0.0947	-0.0462	-0.0081
Volatility	0.176	0.097	0.0074
Skewness	-0.262***	-0.084***	-0.888***
Excess kurtosis	8.88***	2.48***	50.4***

Table 1: **Summary statistics.** Sample statistics for the equity and exchange rate daily log returns, and for the daily basis point changes in interest rates from 1990 to 2018. Asterisks denote significance at 5% (*), 1% (**) and 0.1%(***). The standard error of the sample mean is equal to the sample standard deviation, divided by sample size, while the standard errors are approximately $(6/T)^{1/2}$ and $(24/T)^{1/2}$ for the sample skewness and excess kurtosis, respectively, where T is the sample size. We used 252 risk days per year to annualize the standard deviation into volatility.

Table 1 presents the sample statistics of the empirical unconditional daily returns distribution over the entire sample. In accordance with stylized facts the mean of every series is not

[26]The Euro was only introduced in 1999, so the ECU/dollar exchange rate is used before this date.

[27]First differences in fixed maturity interest rates are the equivalent of log returns on corresponding bonds.

statistically different from zero and the unconditional volatility is highest for the equity and lowest for the interest rates. Skewness is negative and low (in absolute value) but significant for all three series, so extreme negative returns are more likely than extreme positive returns of the same magnitude, while excess kurtosis is always positive and highly significant, so the unconditional distributions of the series have a greater probability mass in the tails than the normal distribution with the same variance.

For each time series we repeatedly estimate four different GARCH models, namely the baseline GARCH(1,1) and the asymmetric GJR, each with normal and Student t error distributions. For each model we estimate the parameters using ten years (approximately 2500) of observations on each time series. The estimation window is then rolled daily and parameters are re-estimated until the entire dataset is exhausted. The resulting time series of estimated model parameters are subsequently used to compute corresponding time series for the higher moments of forward and aggregated returns, based on the formulae derived in Section 2.[28] This way, we construct time series of conditional moments for the forward and aggregated returns and variances, for any time horizon s, from 3rd January 2000 to 30th March 2018.

Next, we apply the two approximation methods summarised in Section 5.1 to derive distributions for the s-day forward and aggregated returns for each of the four GARCH models, yielding eight approximate distributions to evaluate and compare with distributions based on 10,000 simulations. Then we test the accuracy of our approximations using the KS distance, the CVM and AD test statistics described in Section 5.2. To capture any differences between market regimes the tests are performed for 150 days from a low volatility period (i.e. January to August 2006), 150 days from a high volatility period (i.e. August 2008 to March 2009) and the current period from 2nd January to 30th March 2018. In Tables 2 and 4 these periods are labelled 'low vol', 'high vol' and 'current' respectively. The label 'total' refers to results obtained for all three periods considered together.

[28] Recall that only the normal and Student t GJR lead to non-zero skewness forecasts for the forward and aggregated returns, but the skewness of the forward and aggregated variances is non-zero even for the symmetric models. All four models yield non-zero, positive excess kurtosis for all time series.

Data	Normal GARCH				Normal GJR				t-GARCH				t-GJR			
S&P 500	total	low vol	high vol	current	total	low vol	high vol	current	total	low vol	high vol	current	total	low vol	high vol	current
$\hat{\mu} \times 10^3$	0.228	0.296	0.074	0.440	-0.105	-0.015	-0.281	0.109	0.335	0.359	0.200	0.608	0.088	0.137	-0.079	0.378
$\hat{\omega} \times 10^5$	0.129	0.126	0.096	0.220	0.162	0.188	0.108	0.230	0.090	0.106	0.058	0.126	0.126	0.152	0.078	0.180
$\hat{\alpha}$	0.082	0.075	0.068	0.135	-0.011	-0.005	-0.016	-0.012	0.078	0.062	0.068	0.138	-0.014	-0.009	-0.019	-0.012
$\hat{\lambda}$	-	-	-	-	0.146	0.137	0.125	0.219	-	-	-	-	0.148	0.126	0.129	0.251
$\hat{\beta}$	0.910	0.917	0.927	0.850	0.924	0.921	0.945	0.878	0.920	0.930	0.931	0.866	0.928	0.932	0.947	0.872
degrees of freedom	-	-	-	-	-	-	-	-	8.485	9.120	9.267	5.000	10.291	11.164	11.376	5.475
long term volatility	34.3%	12.2%	50.4%	18.3%	36.3%	12.6%	53.2%	20.4%	35.1%	12.1%	51.7%	19.0%	37.0%	12.5%	54.4%	21.3%
Data	Normal GARCH				Normal GJR				t-GARCH				t-GJR			
Euro/Dollar	total	low vol	high vol	current	total	low vol	high vol	current	total	low vol	high vol	current	total	low vol	high vol	current
$\hat{\mu} \times 10^3$	-0.090	-0.014	-0.184	-0.050	-0.088	-0.010	-0.162	-0.101	-0.036	0.069	-0.136	-0.048	-0.036	0.070	-0.129	-0.072
$\hat{\omega} \times 10^5$	0.021	0.027	0.020	0.011	0.020	0.027	0.020	0.006	0.020	0.025	0.020	0.009	0.019	0.025	0.020	0.004
$\hat{\alpha}$	0.028	0.027	0.029	0.032	0.028	0.029	0.033	0.012	0.030	0.031	0.028	0.035	0.029	0.032	0.031	0.015
$\hat{\lambda}$	-	-	-	-	-0.001	-0.004	-0.010	0.025	-	-	-	-	0.001	-0.002	-0.006	0.025
$\hat{\beta}$	0.966	0.966	0.967	0.966	0.968	0.966	0.967	0.974	0.965	0.963	0.967	0.964	0.966	0.963	0.968	0.973
degrees of freedom	-	-	-	-	-	-	-	-	9.069	8.966	10.036	6.942	9.124	8.961	10.120	7.076
long term volatility	12.4%	8.7%	16.6%	7.1%	12.5%	8.7%	16.8%	6.5%	12.4%	8.7%	16.5%	7.1%	12.4%	8.7%	16.6%	6.5%
Data	Normal GARCH				Normal GJR				t-GARCH				t-GJR			
3M Treasury Bill	total	low vol	high vol	current	total	low vol	high vol	current	total	low vol	high vol	current	total	low vol	high vol	current
$\hat{\mu} \times 10^3$	4.167	5.014	4.615	0.983	2.993	3.775	3.074	0.869	2.637	2.978	2.944	1.042	2.423	2.852	2.580	0.982
$\hat{\omega} \times 10^5$	4.372	6.131	4.191	0.493	4.747	6.293	4.931	0.493	6.161	8.902	5.677	0.613	6.350	9.024	6.007	0.621
$\hat{\alpha}$	0.271	0.260	0.313	0.191	0.161	0.163	0.156	0.170	0.181	0.168	0.199	0.167	0.145	0.148	0.139	0.154
$\hat{\lambda}$	-	-	-	-	0.187	0.176	0.257	0.040	-	-	-	-	0.077	0.047	0.126	0.033
$\hat{\beta}$	0.763	0.760	0.743	0.817	0.770	0.765	0.756	0.818	0.803	0.799	0.800	0.818	0.799	0.796	0.795	0.815
degrees of freedom	-	-	-	-	-	-	-	-	5.472	5.452	5.277	6.000	5.630	5.549	5.560	6.000
long term volatility	156.6	52.0	235.2	49.9	159.4	50.3	240.1	48.7	142.7	52.3	213.0	47.0	146.1	51.6	218.7	46.0

Table 2: **GARCH model parameter estimates** We report an average of the time series of parameter estimates from the four different GARCH models for each of the three datasets. Each model is re-estimated daily using the last ten years of daily log returns. Averages reported are labelled 'low vol', 'high vol' and 'current' and these refer to the sub-periods: January to August 2006, August 2008 to March 2009 and observations from 2018, respectively. The label 'total' refers to the average over all three sub-periods considered together.

5.4 Empirical Results

First consider the parameter estimates obtained for each model and each dataset. Parameters are re-estimated daily, so to depict all the estimates would require 12 pages, each displaying between 4 and 6 time series over 17 years. Therefore, because we lack space here, Table 2 simply summarises the model parameter estimates as averages taken over the different sub-samples described above.[29] There is nothing particularly remarkable about these parameter estimates; they largely conform to those found by other researchers in the field, see Harvey and Lange (2017), Herwartz (2017) and many other empirical studies cited above.

Turning now to the accuracy tests for our distribution approximations, we proceed as follows: (i) Fix a model and its parameter estimates and the value of s for the horizon of interest;[30] (ii) Use these model parameters to generate exact values for the first four moments of s-day forward and aggregated returns, based on our analytic moment formulae; (iii) Fit two types of distribution approximation to these moments, i.e. (a) Johnson SU and (b) the Edgeworth expansion, as described in Section 4.1; (iv) use the same value of s and the same model parameters to generate 10,000 simulations of s-day forward and aggregate returns, whence we obtain an empirical distribution; and finally (v) apply the evaluation methods for equality of the two distributions described in 4.2.

This very comprehensive and extensive exercise is impossible to report in detail, so the only time horizon reported here is $s = 5$ days. The results we report here are qualitatively similar to our results for $s = 10$ and 20 trading days, which are available on request. Tables 3 and 4 summarize the results of the distribution tests for each of the eight approximate distributions considered.[31] The AD test gave results very similar to the CVM test, hence we only report the results for the KS and CVM tests. We report the mean values and the

[29]The conditional mean μ is assumed constant here, but note our earlier remarks about filtering through an autoregressive model if this is necessary to remove autocorrelation in daily returns.

[30]For instance, take the normal GJR model estimated using ten years of daily log returns on the S&P 500 ending on 31 December 2008, and examine the moments of s-day forward and aggregated returns distributions for $s = 10$.

[31]For the interest rate sample, fitting the Johnson SU distribution using the moments of the aggregated returns estimated for the Student t GJR was problematic and hence, for this sample, we do not report results for the Johnson SU Student t GJR in Table 3.

associated standard deviations of the test statistics and also the percentage of times when the computed test statistic was higher than the asymptotic 5% critical value. Since we perform the tests at the 5% significance level we expect a 5% rejection rate.

The 5% critical values are 0.0136 for the KS distance and 0.461 for the CVM statistic.[32] Although the asymptotic critical values do not apply exactly in our case, the model producing the lowest values is still the best among the alternatives.[33] We now discuss the results in greater detail for the equity, exchange rate and interest rate distributions in turn.

(a) S&P 500 Index

The Johnson SU approximation appears to be a more suitable approximation than the Edgeworth expansion for the forward returns, especially when applied in combination with the moments produced by the Student t GARCH(1,1) and GJR models. For the normal GARCH(1,1) and GJR models the average values of both the KS and CVM test statistics are still greater, but only marginally, for the Edgeworth approximations, compared with their Johnson SU counterparts. Indeed, if the Edgeworth expansion is used, then the models with normal innovations fit the simulated distributions significantly better than their Student t counterparts. This is not the case for the Johnson SU distribution, where all four GARCH specifications provide very close fits to the simulated distributions (although again the normal models, and especially the normal GARCH(1,1), do give slightly better fits than the Student t models). Overall, the Johnson SU normal GARCH(1,1) model produces the lowest average values of both the KS distance (0.0087) and CVM test statistic (0.1687). For the aggregated returns, the results improve even further, with the Johnson SU methodology still proving

[32]These are asymptotic results for a test where the distribution being tested for is continuous, fully known and generic (no particular family of distributions assumed). Stephens (1970) derives modified statistics for the finite sample case; however, with a sample size of 10,000 these modifications are not actually needed and the asymptotic results would apply, if the hypothetical distribution were fully specified. However, in our case this distribution is based on estimated results and we would need to simulate the correct critical values if we were to properly carry out the tests. Still, we report the percentage of times the test statistics are greater than the asymptotic critical values, so that we can infer, approximately, if the test results are at least in the vicinity of these asymptotic critical values. We also note that the results have to be interpreted with care since it is likely that the appropriate (simulated) critical values for this testing exercise are lower than the asymptotic critical values reported above (see Massey, 1951).

[33]What we mean by "best among alternatives" here means "closest to the (respective) simulated distribution". However, one has to interpret the results with care since the simulated distribution is obviously not the same for all alternative approximate distributions.

superior overall to the Edgeworth expansion (when combined with the Student t models), but to a lesser extent than in the case of the forward returns.

(b) Euro/Dollar Exchange Rate

From the results in Table 2 there is not much to choose between the two alternative approximation methods when the innovations are normal. Indeed, the results for both distribution tests are virtually identical (and good) for the normal GARCH(1,1) and GJR, using either the Johnson SU or Edgeworth expansion. For the Johnson SU approach, the fit is closest when the innovations are normal but the fit is almost as good based on the Student t models. However, when the Edgeworth expansion is employed, the results obtained with the Student t models are significantly worse than when the innovations are normal. For the distributions of aggregated Euro/dollar returns, all models produce similar and good results.

(c) 3-month Treasury Bill Rate

As in (a), the Johnson SU should be preferred to the Edgeworth expansion. The Johnson SU approximation yields the lowest average KS distance as well as the lowest value for the CVM test statistic for the normal GARCH(1,1) model, the normal GJR model being second best. The fits deteriorate when innovations have a Student t distribution.

To summarize, the predictive distributions of forward and aggregated returns on major risk factors may be well approximated using the analytic expressions for the first four conditional moments that we have derived in this paper. The best distribution approximation method overall is the Johnson SU and all four GARCH models that we have tested have predictive forward and aggregated returns distributions that can be well approximated, the easiest being that generated by the normal GARCH(1,1).

	Normal GARCH(1,1)				Student t GARCH(1,1)				Normal GJR				Student t GJR			
	total	low vol	high vol	current	total	low vol	high vol	current	total	low vol	high vol	current	total	low vol	high vol	current
S&P 500								Johnson SU								
KS-average	0.0087	0.0088	0.0086	0.0087	0.0123	0.0091	0.0091	0.0281	0.0089	0.0088	0.0087	0.0095	0.0118	0.0090	0.0089	0.0255
KS-stdev	0.0027	0.0027	0.0026	0.0030	0.0077	0.0028	0.0026	0.0040	0.0027	0.0027	0.0025	0.0030	0.0071	0.0028	0.0025	0.0056
KS-rejections@5%	6.09%	6.00%	4.67%	9.84%	22.99%	8.67%	6.00%	100.00%	6.65%	6.67%	5.33%	9.84%	22.16%	8.00%	4.67%	100.00%
CVM-average	0.1687	0.1765	0.1577	0.1768	0.5000	0.1935	0.1809	1.1257	0.1421	0.1803	0.1593	0.0867	0.5397	0.1896	0.1688	2.3126
CVM-stdev	0.1565	0.1628	0.1298	0.1971	0.9540	0.1664	0.1346	1.4510	0.1595	0.1635	0.1284	0.1690	0.9216	0.1652	0.1296	1.0680
CVM-rejections@5%	6.93%	8.00%	4.67%	9.84%	21.88%	8.00%	4.00%	100.00%	7.20%	8.67%	4.67%	9.84%	21.88%	8.00%	4.00%	100.00%
								Edgeworth								
KS-average	0.0088	0.0088	0.0086	0.0091	0.0544	0.0163	0.0168	0.2409	0.0093	0.0089	0.0088	0.0113	0.0453	0.0142	0.0139	0.1992
KS-stdev	0.0027	0.0027	0.0026	0.0031	0.0860	0.0029	0.0031	0.0431	0.0029	0.0027	0.0025	0.0034	0.0723	0.0028	0.0030	0.0486
KS-rejections@5%	6.09%	6.00%	4.00%	11.48%	85.87%	80.67%	85.33%	100.00%	10.25%	8.00%	5.33%	27.87%	60.66%	54.67%	50.67%	100.00%
CVM-average	0.1727	0.1768	0.1579	0.1990	36.6829	0.7698	0.8121	108.4669	0.1599	0.1846	0.1614	0.1337	31.6104	0.5516	0.5059	184.4710
CVM-stdev	0.1589	0.1629	0.1296	0.2062	97.1201	0.2887	0.2912	143.6736	0.1797	0.1647	0.1284	0.2289	76.5282	0.2440	0.2278	80.9503
CVM-rejections@5%	7.20%	8.00%	4.67%	11.48%	90.86%	86.67%	91.33%	100.00%	10.25%	8.67%	4.67%	27.87%	62.33%	59.33%	50.00%	100.00%
Euro/dollar								Johnson SU								
KS-average	0.0086	0.0088	0.0086	0.0084	0.0091	0.0090	0.0088	0.0099	0.0086	0.0088	0.0086	0.0084	0.0091	0.0090	0.0088	0.0099
KS-stdev	0.0027	0.0027	0.0026	0.0028	0.0027	0.0028	0.0025	0.0030	0.0027	0.0027	0.0026	0.0028	0.0027	0.0027	0.0025	0.0030
KS-rejections@5%	4.93%	4.67%	4.67%	6.15%	7.67%	8.00%	4.00%	15.38%	4.93%	4.67%	4.67%	6.15%	7.67%	8.00%	4.00%	15.38%
CVM-average	0.1661	0.1754	0.1571	0.1655	0.1869	0.1883	0.1636	0.2372	0.1661	0.1754	0.1571	0.1654	0.1880	0.1886	0.1646	0.2408
CVM-stdev	0.1549	0.1618	0.1302	0.1887	0.1628	0.1647	0.1290	0.2123	0.1549	0.1618	0.1303	0.1886	0.1638	0.1648	0.1297	0.2147
CVM-rejections@5%	6.03%	7.33%	4.67%	6.15%	8.49%	9.33%	4.67%	15.38%	6.03%	7.33%	4.67%	6.15%	6.85%	7.33%	3.33%	13.85%
								Edgeworth								
KS-average	0.0086	0.0088	0.0086	0.0084	0.0161	0.0151	0.0122	0.0275	0.0086	0.0088	0.0086	0.0084	0.0159	0.0150	0.0122	0.0262
KS-stdev	0.0027	0.0027	0.0026	0.0028	0.0062	0.0029	0.0029	0.0028	0.0027	0.0027	0.0026	0.0028	0.0058	0.0030	0.0029	0.0029
KS-rejections@5%	4.93%	4.67%	4.67%	6.15%	57.53%	67.33%	29.33%	100.00%	4.93%	4.67%	4.67%	6.15%	57.26%	66.67%	29.33%	100.00%
CVM-average	0.1661	0.1754	0.1571	0.1655	0.8909	0.6435	0.3761	2.6497	0.1661	0.1754	0.1571	0.1654	0.8380	0.6440	0.3735	2.3576
CVM-stdev	0.1549	0.1618	0.1302	0.1887	0.8838	0.2735	0.1876	0.5287	0.1549	0.1618	0.1303	0.1886	0.7798	0.2754	0.1878	0.5101
CVM-rejections@5%	6.03%	7.33%	4.67%	6.15%	63.29%	78.67%	32.00%	100.00%	6.03%	7.33%	4.67%	6.15%	58.36%	74.67%	24.00%	100.00%
3M Bill								Johnson SU								
KS-average	0.0101	0.0103	0.0104	0.0091	0.1614	0.1315	0.2498	0.0176	0.0115	0.0113	0.0127	0.0093	0.0229	0.0212	0.0263	0.0188
KS-stdev	0.0030	0.0031	0.0029	0.0026	0.1071	0.0238	0.0996	0.0031	0.0033	0.0032	0.0030	0.0031	0.0051	0.0041	0.0046	0.0031
KS-rejections@5%	12.74%	16.67%	12.00%	4.92%	99.17%	100.00%	100.00%	95.08%	27.70%	26.00%	34.67%	14.75%	98.89%	98.67%	100.00%	96.72%
CVM-average	0.2340	0.2478	0.2360	0.1950	165.5821	85.8680	312.3000	0.9257	0.3175	0.3038	0.3786	0.2007	1.7315	1.4011	2.3533	1.0150
CVM-stdev	0.1687	0.1865	0.1548	0.1511	187.7128	33.3990	211.8000	0.3475	0.2070	0.2006	0.1921	0.2048	0.9261	0.6283	0.9624	0.3218
CVM-rejections@5%	10.25%	13.33%	8.00%	8.20%	99.72%	100.00%	100.00%	98.36%	21.05%	19.33%	27.33%	9.84%	98.61%	98.00%	100.00%	96.72%
								Edgeworth								
KS-average	0.0235	0.0187	0.0325	0.0129	0.2526	0.2097	0.3491	0.1209	0.0276	0.0217	0.0395	0.0128	0.1892	0.1496	0.2540	0.1270
KS-stdev	0.0090	0.0033	0.0054	0.0028	0.1032	0.0247	0.0824	0.0060	0.0117	0.0035	0.0068	0.0031	0.0642	0.0179	0.0471	0.0080
KS-rejections@5%	88.09%	95.33%	100.00%	40.98%	100.00%	100.00%	100.00%	100.00%	88.37%	99.33%	100.00%	32.79%	100.00%	100.00%	100.00%	100.00%
CVM-average	2.1521	1.0831	3.9245	0.4223	268.2238	181.8000	438.1000	63.0850	3.2588	1.5392	6.1278	0.4327	175.9010	99.3500	295.4000	70.1893
CVM-stdev	1.7811	0.4047	1.3952	0.2045	184.2778	40.5100	163.1000	5.9685	2.8813	0.5157	2.2873	0.2612	123.1241	25.0800	105.4000	8.4955
CVM-rejections@5%	88.37%	98.00%	100.00%	36.07%	100.00%	100.00%	100.00%	100.00%	89.47%	100.00%	100.00%	37.70%	100.00%	100.00%	100.00%	100.00%

Table 3: **Distribution tests for the approximate distributions of 5-day forward returns**
We report the average KS distance (KS-average) and CVM test statistic (CVM-average), with associated standard deviations (KS-stdev and CVM-stdev, respectively) and the percentage of cases where the test statistics are greater than the asymptotic 5% CVs(KS-rejections@5% and CVM-rejections@5%, respectively) for the 5-day forward returns for the S&P 500, Euro/dollar and 3-month Treasury Bills, respectively. Labels 'low vol', 'high vol' and 'current' refer to the sub-periods: January to August 2006, August 2008 to March 2009 and observations from 2018, respectively. The label 'total' refers to all three sub-periods, considered together.

	Normal GARCH(1,1)				Student t GARCH(1,1)				Normal GJR				Student t GJR			
	total	low vol	high vol	current	total	low vol	high vol	current	total	low vol	high vol	current	total	low vol	high vol	current
S&P 500								Johnson SU								
KS-average	0.0086	0.0085	0.0086	0.0086	0.0102	0.0086	0.0090	0.0170	0.0096	0.0095	0.0097	0.0097	0.0131	0.0108	0.0114	0.0232
KS-stdev	0.0025	0.0025	0.0024	0.0029	0.0040	0.0025	0.0026	0.0030	0.0027	0.0026	0.0027	0.0029	0.0058	0.0028	0.0029	0.0058
KS-rejections@5%	5.26%	4.67%	4.67%	8.20%	20.22%	6.00%	5.33%	91.80%	7.48%	8.00%	6.67%	8.20%	32.96%	18.00%	21.33%	98.36%
CVM-average	0.1600	0.1554	0.1608	0.1694	0.2719	0.1592	0.1740	0.7897	0.2028	0.1950	0.2051	0.2165	0.5288	0.2626	0.3007	1.7443
CVM-stdev	0.1396	0.1351	0.1396	0.1517	0.2931	0.1397	0.1382	0.3018	0.1537	0.1483	0.1552	0.1641	0.6858	0.1729	0.1850	0.9238
CVM-rejections@5%	4.99%	4.00%	5.33%	6.56%	18.28%	4.67%	4.00%	86.89%	7.48%	8.00%	6.00%	9.84%	27.42%	12.00%	15.33%	95.08%
								Edgeworth								
KS-average	0.0088	0.0086	0.0087	0.0097	0.0199	0.0102	0.0111	0.0656	0.0109	0.0108	0.0105	0.0125	0.0237	0.0142	0.0145	0.0695
KS-stdev	0.0026	0.0025	0.0025	0.0029	0.0210	0.0028	0.0028	0.0071	0.0030	0.0027	0.0029	0.0033	0.0220	0.0029	0.0032	0.0170
KS-rejections@5%	6.65%	6.00%	4.67%	13.11%	28.81%	11.33%	17.33%	100.00%	14.40%	11.33%	12.67%	26.23%	66.48%	60.00%	59.33%	100.00%
CVM-average	0.1729	0.1595	0.1643	0.2270	3.2707	0.2354	0.2904	18.0636	0.2893	0.2687	0.2594	0.4135	3.8444	0.5348	0.5863	19.9947
CVM-stdev	0.1463	0.1372	0.1389	0.1739	6.8926	0.1659	0.1692	4.1463	0.2072	0.1784	0.1886	0.2671	8.2122	0.2550	0.2897	9.2289
CVM-rejections@5%	6.09%	4.67%	4.00%	14.75%	27.42%	9.33%	16.00%	100.00%	13.30%	11.33%	8.00%	31.15%	65.65%	54.00%	63.33%	100.00%
Euro/dollar								Johnson SU								
KS-average	0.0085	0.0084	0.0086	0.0086	0.0087	0.0085	0.0087	0.0089	0.0085	0.0084	0.0086	0.0085	0.0087	0.0085	0.0088	0.0091
KS-stdev	0.0025	0.0025	0.0024	0.0027	0.0025	0.0025	0.0025	0.0028	0.0025	0.0025	0.0024	0.0027	0.0025	0.0025	0.0025	0.0028
KS-rejections@5%	5.48%	4.67%	4.67%	9.23%	4.66%	3.33%	4.00%	9.23%	4.93%	4.00%	4.67%	7.69%	4.93%	4.00%	4.00%	9.23%
CVM-average	0.1600	0.1560	0.1617	0.1655	0.1651	0.1591	0.1645	0.1804	0.1602	0.1561	0.1613	0.1673	0.1676	0.1590	0.1654	0.1927
CVM-stdev	0.1418	0.1381	0.1411	0.1533	0.1442	0.1391	0.1404	0.1646	0.1421	0.1382	0.1402	0.1566	0.1468	0.1398	0.1392	0.1765
CVM-rejections@5%	4.66%	4.67%	4.67%	4.62%	4.66%	4.00%	4.67%	6.15%	4.66%	4.67%	4.67%	4.62%	4.38%	3.33%	4.00%	7.69%
								Edgeworth								
KS-average	0.0085	0.0084	0.0086	0.0086	0.0093	0.0089	0.0090	0.0110	0.0085	0.0084	0.0086	0.0086	0.0093	0.0089	0.0091	0.0107
KS-stdev	0.0025	0.0025	0.0024	0.0027	0.0027	0.0025	0.0025	0.0030	0.0025	0.0025	0.0024	0.0027	0.0027	0.0026	0.0025	0.0029
KS-rejections@5%	5.48%	4.67%	4.67%	9.23%	7.67%	6.67%	4.00%	18.46%	5.21%	4.67%	4.67%	7.69%	7.95%	6.67%	4.67%	18.46%
CVM-average	0.1601	0.1559	0.1617	0.1657	0.1932	0.1722	0.1730	0.2886	0.1603	0.1560	0.1612	0.1683	0.1934	0.1720	0.1756	0.2838
CVM-stdev	0.1418	0.1382	0.1408	0.1536	0.1603	0.1448	0.1407	0.1999	0.1422	0.1384	0.1395	0.1583	0.1622	0.1458	0.1396	0.2115
CVM-rejections@5%	4.66%	4.67%	4.67%	4.62%	6.58%	4.00%	4.67%	16.92%	4.66%	4.67%	4.67%	4.62%	6.85%	4.00%	4.00%	20.00%
3M Bill								Johnson SU								
KS-average	0.0085	0.0084	0.0087	0.0083	0.0306	0.0144	0.0158	0.1069	0.0105	0.0097	0.0120	0.0086	-	-	-	-
KS-stdev	0.0024	0.0024	0.0024	0.0023	0.0359	0.0042	0.0031	0.0232	0.0031	0.0027	0.0029	0.0027	-	-	-	-
KS-rejections@5%	4.16%	4.00%	4.67%	3.28%	73.96%	59.33%	78.00%	100.00%	14.96%	7.33%	26.67%	4.92%	-	-	-	-
CVM-average	0.1559	0.1520	0.1645	0.1441	8.5074	0.6000	0.6860	47.1848	0.2570	0.2100	0.3398	0.1689	-	-	-	-
CVM-stdev	0.1277	0.1233	0.1349	0.1208	18.8022	0.3659	0.2650	17.0440	0.1831	0.1468	0.1973	0.1431	-	-	-	-
CVM-rejections@5%	4.43%	4.00%	4.67%	4.92%	75.90%	60.67%	81.33%	100.00%	15.79%	8.67%	26.00%	8.20%	-	-	-	-
								Edgeworth								
KS-average	0.0208	0.0179	0.0269	0.0128	0.0740	0.0511	0.0747	0.1286	0.0297	0.0245	0.0417	0.0132	0.0524	0.0468	0.0627	0.0407
KS-stdev	0.0065	0.0030	0.0041	0.0031	0.0314	0.0069	0.0094	0.0352	0.0119	0.0034	0.0065	0.0029	0.0116	0.0065	0.0089	0.0036
KS-rejections@5%	87.26%	95.33%	100.00%	36.07%	100.00%	100.00%	100.00%	100.00%	90.03%	100.00%	100.00%	40.98%	100.00%	100.00%	100.00%	100.00%
CVM-average	1.4534	0.9387	2.3913	0.4129	25.4565	10.5500	23.4370	67.0785	3.4691	1.9313	6.2325	0.4553	11.0503	8.5839	15.4580	6.2775
CVM-stdev	0.9611	0.3213	0.7176	0.2115	23.1595	2.8720	5.6867	28.1144	2.6880	0.5429	1.8308	0.2374	5.0089	2.5548	4.3200	1.0276
CVM-rejections@5%	85.60%	93.33%	100.00%	31.15%	100.00%	100.00%	100.00%	100.00%	91.14%	100.00%	100.00%	47.54%	100.00%	100.00%	100.00%	100.00%

Table 4: **Distribution tests for the approximate distributions of 5-day aggregated returns**
We report the average KS distance (KS-average) and CVM test statistic (CVM-average), with associated standard deviations (KS-stdev and CVM-stdev, respectively) and the percentage of cases where the test statistics are greater than the asymptotic 5% CVs(KS-rejections@5% and CVM-rejections@5%, respectively) for the 5-day forward returns for the S&P 500, Euro/dollar and 3-month Treasury Bills, respectively. Labels 'low vol', 'high vol' and 'current' refer to the sub-periods: January to August 2006, August 2008 to March 2009 and observations from 2018, respectively. The label 'total' refers to all three sub-periods, considered together.

6 Conclusions

We have derived analytical expressions for the moments of forward and aggregated returns and variances for an established asymmetric GARCH specification, namely the GJR model, with a generic innovations distribution. Special cases include the normal and Student t GARCH(1,1) and GJR models. We found that the distribution of forward returns is skewed only if the distribution of innovations is skewed, but the distribution of aggregated returns is skewed even if the innovation distribution is symmetric. The other source of skewness in this case is the asymmetric response of variance to positive and negative shocks (i.e. $\lambda \neq 0$).

There are two sources of kurtosis in forward returns: the degree of leptokurtosis of the innovation distribution and the uncertainty in forward variance. Since the one-step-ahead variance can be made with certainty in a GARCH setting, the kurtosis coefficient of the one-step-ahead returns distributions is equal to that of the innovation distribution. However, whenever we forecast s-steps ahead (with $s > 1$) using a GARCH(1,1) or GJR model, the s-step-ahead returns distribution for $s > 1$ will have a higher kurtosis than the one-step-ahead returns distribution, due to the positivity of the conditional variance of the conditional variance, which increases the probability mass in the tails of the forward one-period returns distribution. Also, the time-variability of the conditional variance of the conditional variance introduces dynamics in the higher moments of the forward returns.

Provided the unconditional moments exist (i) the conditional moments of forward returns converge to the corresponding unconditional moments as the time horizon increases, and (ii) the conditional moments of aggregated returns converge to the corresponding moments of a normal distribution. Otherwise, the moments of both the forward returns as well as the aggregated returns generally diverge to (plus or minus) infinity.

A bootstrapping exercise uses a t-GARCH(1,1) base model to simulate an empirical distribution for the skewness and kurtosis of the aggregated returns, using our analytic formulae and via bootstrapping. We compare this with the values given by our formulae applied to the base parameters of the GARCH model. This way, we conclude: (a) the bootstrapped skewness is unbiased; (b) the bootstrapped kurtosis presents a downwards bias; the 'true' value of the kurtosis is outside the interquartile range, increasingly so for aggregated returns over longer horizons; and (c) the kurtosis estimate computed using our formulae is unbiased.

An empirical application computed higher moments of both forward and aggregated returns of the S&P 500 index, the Euro/dollar exchange rate and the 3-month US Treasury bill rate, using our analytic expressions for the first four conditional moments based on four different GARCH processes. Subsequently, we approximated predictive distributions for forward and aggregated returns using these higher moment forecasts and the Johnson SU distribution or the Edgeworth expansion. Using established statistical tests, we evaluated the accuracy of these approximations, relative to the corresponding simulated GARCH returns distributions. The results of these tests are in general very good for the vast majority of the approximate distributions. Hence, our moment expressions may have useful applications to financial problems which, until now, have required GARCH returns distributions to be simulated.

References

Alexander, C. & E. Lazar (2009) Modelling Regime-specific Stock Price Volatility. *Oxford Bulletin of Economics and Statistics* 71, 761-797.

Alexander, C., E. Lazar & S. Stanescu (2013) Forecasting VaR using Analytic Higher Moments for GARCH Processes. *International Review of Financial Analysis* 30, 36-45

Andersen, T.G. & T. Bollerslev (1998) Answering the Critics: Yes ARCH models Do Provide Good Volatility Forecasts. *International Economic Review* 39, 885-905.

Andersen, T.G., T. Bollerslev & F.X. Diebold (2009) Parametric and Non-parametric Volatility Measurement. In L.P. Hansen & Y. Ait-Sahalia (eds.), *Handbook of Financial Econometrics*. North-Holland.

Anderson, T.W. & D.A. Darling (1952) Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *The Annals of Mathematical Statistics* 23, 193-212.

Bai, X., J.R. Russell, & G.C. Tiao (2003) Kurtosis of GARCH and Stochastic Volatility Models with Non-Normal Innovations. *Journal of Econometrics* 114, 349-360.

Baillie, $ T. Bollerslev (1982) Prediction in Dynamic Models with Time-dependent Conditional Variances *Journal of Econometrics* 52, 91-113.

Bauwens, L., S. Laurent & J.V.K. Rombouts (2006) Multivariate GARCH Models: a Survey. *Journal of Applied Econometrics* 21, 79-109.

Bhattacharya, R.N. & J.K. Ghosh (1978) On the Validity of the Formal Edgeworth Expansion. *The Annals of Statistics* 6, 435-451.

Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31, 307-327.

Bollerslev, T. (1987) A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics* 69, 542-547.

Boudt, K., Peterson, B. and Croux, C. (2009) Estimation and decomposition of downside risk for portfolios with non-normal returns. *Journal of Risk* 11, 189–200.

Breuer, T. & M. Jandacka (2010) Temporal Aggregation of GARCH Models: Conditional Kurtosis and Optimal Frequency. *Working Paper*, available from http://ssrn.com.

Charlier, C.V.L. (1905) Uber das Fehlergesetz. *Arkiv for Matematik, Astronomi Och Fysic*, 8, 1-9.

Charlier, C.V.L. (1906) Uber die Darstellung Willkurlicher Funktionen. *Arkiv for Matematik, Astronomi Och Fysic*, 20, 1-35.

Chebyshev, P.L. (1860) Sur le Developpment des Fonctions a une Seule Variable. *Bulletin de la Classe Physique-Mathematique de l'Academie Imperiale des Sciences St. Petersbourg* 3, 193-202.

Chebyshev, P.L. (1890) Sur Deux Theorems Relatifs aux Probabilites. *Acta Mathematica* 14, 305-315.

Christoffersen, P.F. (2012) *Elements of Financial Risk Management.* Academic Press.

Christoffersen, P.F., C. Dorion, K. Jacobs & Y. Wang (2010) Volatility Components: Affine Restrictions and Non-normal Innovations. *Journal of Business and Economic Statistics* 28, 483-502.

Christoffersen, P.F., K. Jacobs, C. Ornthanalai & Y. Wang (2008) Option Valuation with Long-run and Short-run Volatility Components. *Journal of Financial Economics* 90, 272-297.

Clements, M.P, and Smith, J. (2000) Evaluating the Forecast Densities of Linear and Nonlinear Models: Applications to Output Growth and Unemployment. *Journal of Forecasting* 19, 255-276.

Clements, M.P, and Smith, J. (2002) Evaluating Multivariate Forecast Densities: a Comparison of Two Approaches. *International Journal of Forecasting* 18, 397-407.

Corradi, V. and Swanson, N.R. (2006a) Bootstrap Conditional Distribution Tests in the Presence of Dynamic Misspecification. *Journal of Econometrics* 133, 779-806.

Corradi, V. and Swanson, N.R. (2006b) Predictive Density Evaluation. In G. Elliott, C.W.J. Granger and A. Timmermann (Eds), *Handbook of Economic Forecasting*, Volume 1. North Holland Elsevier, Amsterdam.

Demos, A. (2002) Moments and Dynamic Structure of a Time-Varying-Parameter Stochastic Volatility in Mean Model. *The Econometrics Journal* 5, 345-357.

Diebold, F.X. (1988) *Empirical Modelling of Exchange Rates*. Springer.

Diebold, F.X., J. Hahn & A.S. Tay (1999) Evaluating Density Forecasts with Applications to Finance and Management: High Frequency Returns on Foreign Exchange. *Review of Economics and Statistis* 81, 661-673.

Duan, J.C., G. Gauthier, C. Sasseville & J.G. Simonato (2006) Approximating the GJR-GARCH and EGARCH Option Pricing Models Analytically. *Journal of Computational Finance* 9, 41-69.

Duan, J.C., G. Gauthier & J.G. Simonato (1999) An Analytical Approximation for the GARCH Option Pricing Model. *Journal of Computational Finance* 2, 76-116.

Edgeworth, F.Y. (1905) The Law of Error. *Transactions of the Cambridge Philosophical Society* 20, 33-66 and 113 -141.

Engle, R.F. (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50, 987-1007.

Engle, R.F. (1990) Discussion: Stock Market Volatility and the Crash of '87. *Review of Financial Studies* 3, 103-106.

Engle, R.F. & V.K. Ng (1993) Measuring and Testing the Impact of News on Volatility. *Journal of Finance* 48, 1749-1778.

Fama, E. (1965) The Behaviour of Stock Market Prices. *Journal of Business* 38, 34 -105.

Francq, C. & J.-M. Zakoian (2010) *GARCH Models. Structure, Statistical Inference and Financial Applications*. Wiley.

Glosten, L.R., R. Jagannathan & D.E. Runkle (1993) On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance* 48, 1779-1801.

Goncalves, E., J. Leite & N. Mendes-Lopes (2016) On the Distribution Estimation of Power Threshold GARCH Processes. *Journal of Time Series Analysis* 37, 579-602.

Gram, J.P. (1883) Uber die Entwickelung Reeler Funktionen in Reihen mittelst der Methode der Kleinsten Quadrate. *Creile* 94, 41-73.

Haas, M., S. Mittnik, & M.S. Paolella (2004) Mixed Normal Conditional Heteroskedasticity. *Journal of Financial Econometrics* 2, 493-530.

Harvey, A. (2013) *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Cambridge University Press: Econometric Society Monograph.

Harvey, A. & R.-J. Lange (2017) Volatility Modelling with a Generalized t Distribution. *Journal of Time Series Analysis* 38, 175-190.

Harvey, A. & G. Sucarrat (2014) EGARCH Models with Fat Tails, Skewness and Leverage. *Computational Statistics and Data Analysis* 76, 320-328.

He, C. & T. Terasvirta (1999a) Properties of Moments of a Family of GARCH Processes. *Journal of Econometrics* 92, 173-192.

He, C. & T. Terasvirta (1999b) Fourth Moment Structure of the GARCH(p,q) process. *Econometric Theory* 15, 824-846.

He, C., T. Terasvirta & H. Malmsten (2002) Fourth Moments Structure of a Family of First-Order Exponential GARCH Models. *Econometric Theory* 18, 868-885.

Herwartz, H. (2017) Stock Return Prediction under GARCH - An Empirical Assessment. *International Journal of Forecasting* 33, 569-580.

Heston, S.L. & S. Nandi (2000) A Closed-Form GARCH Option Pricing Model. *Review of Financial Studies* 13, 281-300.

Huisman, R., K.G. Koedijk, C.J.M. Kool, & F. Palm (2001) Tail-index Estimator in Small sample. *Journal of Business and Economic Statistics* 19, 208–216.

Huisman, R., K.G. Koedijk, C.J.M. Kool, & F. Palm (2002) The Tail-fatness of FX Returns Reconsidered, *De Economist* 150, 299–312.

Ishida, I. & R. F. Engle (2002) Modeling Variance of Variance: The Square-Root, the Affine, and the CEV GARCH Models. *Working Paper*, available from http://www.stern.nyu.edu/rengle/.

Jaschke, S. (2002) The Cornish-Fisher-Expansion in the Context of Delta-Gamma-Normal Approximations. *Journal of Risk* 4, 33-52.

Johnson, N.L. (1949) Systems of Frequency Curves Generated by Methods of Translation. *Biometrica* 36, 149-76.

Karanasos, M. (1999) The Second Moment and the Autocovariance Function of the Squared errors of the GARCH Model. *Journal of Econometrics* 9, 63-76.

Karanasos, M. (2001) Prediction in ARMA Models with GARCH-in-Mean Effects. *Journal of Time Series Analysis* 22, 555-78.

Karanasos, M. & J. Kim (2003) Moments of ARMA-EGARCH model. *The Econometrics Journal* 6, 146-166.

Karanasos, M., Z. Psaradakis & M. Sola (2004) On the Autocorrelation Properties of Long-Memory GARCH Processes. *Journal of Time Series Analysis* 25, 265-281.

Kolmogorov, A. (1933) Sulla Determinazione Empirica di una Legge di Distribuzione. *Giornale dell'Instituto Italiano degh Attuari* 4, 1-11.

Krause, J. & M.S. Paolella (2014) A Fast, Accurate Method for Value-at-Risk and Expected Shortfall. *Econometrics* 2, 98-122.

Ling, S. & M. McAleer (2002a) Necessary and Sufficient Moment Conditions for the GARCH(r,s) and the Asymmetric Power GARCH(r,s). *Econometric Theory* 18, 722-729.

Ling, S. & M. McAleer (2002b) Stationarity and the Existence of Moments of a Family of GARCH Processes. *Journal of Econometrics* 106, 109-117.

Lux, T. & L. Morales-Arias (2010) Forecasting Volatility under Fractality, Regime-switching, Long Memory and Student-t Innovations. *Computational Statistics and Data Analysis* 54, 2676-2692.

Mandelbrot, B. (1963) The Variations of Certain Speculative Prices. *Journal of Business* 36, 394-419.

Marcucci, J. (2005) Forecasting Stock Market Volatility with Regime-Switching GARCH Models. *Studies in Nonlinear Dynamics & Econometrics* 9, 1-53.

Massey, F.J. Jr. (1951) The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association* 46, 68-78.

Mazzoni, T. (2010) Fast analytic option valuation with GARCH. *Journal of Derivatives* 18, 18-40.

Milhoj, A. (1985) The Moment Structure of ARCH Processes. *Scandinavian Journal of Statistics* 12, 281-292.

Nelson, D.B. (1991) Conditional Heteroskedasticity in Asset Returns: a New Approach. *Econometrica* 59, 347-370.

Nemec, A.F.L. (1985) Conditionally Heteroskedastic Autoregressions. *Technical Report no. 43*. Department of Statistics, University of Washington.

Paolella, M.S. (2016) Stable-GARCH Models for Financial Returns: Fast Estimation and Tests for Stability. *Econometrics* 4, 25

Pearson, E.S. & H.O. Hartley (1972) *Biometrika Tables for Statisticians*, vol.2. Cambridge University Press.

Serfling, R.J. (1980) *Approximation Theorems of Mathematical Statistics*. Wiley.

Simonato, J.G. (2010) The performance of Johnson distributions for Value at risk and expected shortfall computation, *Journal of Derivatives* 19, 7-24.

Simonato, J.G. (2013) Approximating the Multivariate Distribution of Time-Aggregated Stock Returns Under GARCH. *Journal of Risk* 16, 25-49.

Smirnov, N. (1939) Sur les Ecarts de la Courbe de Distribution Empirique. *Matematicheskii Sbornik* 48, 3-26.

Stephens, M.A. (1970) Use of the Kolmogorov-Smirnov, Cramer-von Mises and Related Statistics without Extensive Tables. *Journal of the Royal Statistical Society B* 32, 115-122.

Taylor, S. (1986) *Modeling Financial Time Series*. Wiley.

Theodossiou P. & C.S. Savva (2015) Skewness and the relation between risk and return. *Management Science* 62, 1598–1609.

Tuenter, H.J.H. (2001) An Algorithm to Determine the Parameters of SU-Curves in the Johnson System of Probability Distributions by Moment Matching. *Journal of Statistical Computation and Simulation* 70, 325-347.

Wallace, D.L. (1958) Asymptotic Approximations to Distributions. *The Annals of Mathematical Statistics* 29, 635-654.

Wong, C.M. & M.K.P. So (2003) On Conditional Moments of GARCH Models with Applications to Multiple Period Value at Risk Estimation. *Statistica Sinica* 13, 1015-1044.

Zhu, D. & J.W. Galbraith (2011) Modeling and Forecasting Expected Shortfall with the Generalized Asymmetric Student-t and Asymmetric Exponential Power Distributions. *Journal of Empirical Finance* 18, 765-778.