Journal of Econometrics 235 (2023) 470–483

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

GARCH density and functional forecasts☆

Karim M. Abadir [a], Alessandra Luati [b], Paolo Paruolo [c,*]

[a] Business School, Imperial College London, UK
[b] Department of Statistical Sciences "Paolo Fortunati", University of Bologna, Italy
[c] European Commission, Joint Research Centre (JRC), Ispra, Italy

A R T I C L E   I N F O

Article history:
Received 2 March 2021
Received in revised form 1 April 2022
Accepted 16 April 2022
Available online 21 June 2022

JEL classification:
C22
C53
C58
G17
D81

Keywords:
GJR GARCH(1, 1)
Prediction density
Functionals
Value at Risk
Expected Shortfall

A B S T R A C T

This paper derives the analytic form of the multi-step ahead prediction density of a Gaussian GARCH(1,1) process with a possibly asymmetric news impact curve in the GJR class. These results can be applied when single-period returns are modeled as a GJR Gaussian GARCH(1,1) and interest lies in single-period returns at some future forecast horizon. The Gaussian density has been used in applications as an approximation to this as yet unknown prediction density; the analytic form derived here shows that this prediction density, while symmetric, can be far from Gaussian. This explicit form can be used to compute exact tail probabilities and functionals, such as the Value at Risk and the Expected Shortfall, to quantify expected future required risk capital for single-period returns. Finally, the paper shows how estimation uncertainty can be mapped onto uncertainty regions for any functional of this prediction distribution.

1. Introduction

Since their introduction in Engle (1982) and Bollerslev (1986), Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) processes have been widely employed in Financial Econometrics, see e.g. Bollerslev et al. (2010). In the GARCH original formulation, the conditional distribution of innovations was assumed to be Gaussian.

Even with Gaussian innovations, GARCH processes were shown to generate volatility-clustering, and, when stationary, excess kurtosis in the stationary distribution, see e.g. Bollerslev et al. (1992). Other distributions are commonly used for the innovations, starting with the Student t proposed in Bollerslev (1987); these distributions are usually symmetric with fatter tails than the Gaussian. Despite its importance, the exact form of the h-step ahead prediction probability density function (p.d.f.) for GARCH processes is however still unknown, see Andersen et al. (2006), page 811.

The present paper derives the analytical form of the multi-step-ahead prediction density of a Gaussian GJR GARCH(1,1) process with asymmetric news-impact-curve, see Glosten et al. (1993), which nests the Gaussian GARCH(1,1). The derived

☆ The authors acknowledge useful comments from Enrique Sentana, Robert Engle, Eric Renault, Leopoldo Catania, Barbara Rossi, Dimitris Politis, Ngai Chan, Christian Brownlees, Nour Meddahi, Torben Andersen (the Journal Editor), the Associate Editor, anonymous referees and from participants to seminars in University of Verona, Bocconi University of Milan, CREST Paris, the Seventh Italian Congress of Econometrics and Empirical Economics 2017, EC[2] 2017 in Amsterdam, Barcellona GSE Summer Forum 2017 and the NBER-NSF Time Series Conference 2018. The views set out in this paper are those of the authors and do not necessarily reflect the ones of the institutions of affiliation.
* Correspondence to: European Commission, Joint Research Centre (JRC), Via Enrico Fermi 2749, TP 723, 21027, Ispra (VA), Italy.
E-mail addresses: k.m.abadir@imperial.ac.uk (K.M. Abadir), alessandra.luati@unibo.it (A. Luati), paolo.paruolo@ec.europa.eu (P. Paruolo).

https://doi.org/10.1016/j.jeconom.2022.04.010

formulae are valid for stationary as well as non-stationary GJR GARCH(1,1) processes. They also cover the case when the intercept of the conditional variance equation equals zero, as in the RiskMetrics Value at Risk (VaR) calculations.

The form of the multi-step ahead GJR GARCH(1,1) prediction distribution is proved to be symmetric for any symmetric distribution of the innovations. The closed form expression for Gaussian innovations shows that the prediction distribution can be very far from Gaussian, calling into question the use of the Gaussian to approximate it, as in Lau and McSharry (2010, page 1322). It is found, however, that for selected special cases such as the RiskMetrics one, the prediction distribution is very close to the Gaussian.

The exact closed form of the prediction density can be used to compute exact tail probabilities and risk functionals for single-period returns h-steps in the future, such as the Expected Shortfall. The paper also shows how to construct uncertainty intervals for these functionals to reflect estimation uncertainty on the GARCH parameters.

The rest of the paper is organized as follows. Section 2 states the prediction problem and derives generic results on the prediction p.d.f. for symmetric distributions of innovations. Section 3 states the main results; proofs of this section are collected in the Appendix. Section 4 presents graphs of the prediction density and discusses estimation-uncertainty regions for prediction functionals. Section 5 concludes.

The Supplementary Material contains a MATLAB implementation of the formulae for the Gaussian GJR GARCH(1,1) prediction p.d.f. This paper follows the notational conventions in Abadir and Magnus (2002).

2. Problem definition

This section introduces notation for the reference GJR GARCH(1,1) process, it defines the prediction problem and discusses the symmetry of the prediction density for symmetric distributions of the innovations, which is the most common case in practice. Consider the asymmetric GJR GARCH(1,1)

$$x_t = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = \omega + \alpha_{t-1} x_{t-1}^2 + \beta \sigma_{t-1}^2, \qquad \alpha_t := \alpha + \lambda 1_{x_t<0} = \alpha + \frac{\lambda}{2}(1-\varsigma_t) \tag{2.1}$$

where $\alpha > 0$, $\beta > 0$, $\omega \geq 0$, $\lambda \geq 0$ are the parameters, collected in the vector $\boldsymbol{\theta} = (\omega, \alpha, \beta, \lambda)'$, $1_{x_t<0} = \frac{1}{2}(1-\varsigma_t)$ is the indicator function for the event $x_t < 0$, and $\varsigma_t := \operatorname{sgn}(\varepsilon_t) = \operatorname{sgn}(x_t)$ is the sign of ε_t or of x_t; these signs are the same because $\sigma_t > 0$ almost surely. Process (2.1) contains several special cases: the GARCH(1,1) case is obtained for $\lambda = 0$ and the RiskMetrics method for computing the VaR corresponds to $\lambda = 0$, $\omega = 0$, $\alpha = 0.06$, $\beta = 0.94$, see Tsay (2010) section 7.2. The parameter vector $\boldsymbol{\theta}$ is assumed to be known; this assumption is relaxed in Section 4.

The sequence $\{\varepsilon_t\}$ is assumed to be independent and identically distributed (i.i.d), with mean zero, variance 1, and with Gaussian probability density function $f_\varepsilon(v) := g(v^2) := (2\pi)^{-\frac{1}{2}} \exp(-v^2/2)$. Hereafter $f_x(u)$ (respectively $F_x(u)$) denotes the p.d.f. (respectively the cumulative distribution function - c.d.f.) of random variable x evaluated at u.

The process is observed up to and including time $t = 0$, and the problem is to derive the p.d.f. of x_h for $h = 1, 2, 3, \ldots$, conditional on x_0 and σ_0, $\sigma_0 > 0$. In other words, the object of interest is the density $f_{x_h|x_0,\sigma_0}(\cdot|\cdot,\cdot)$, which for simplicity is abbreviated as $f_{x_h}(\cdot)$.

In financial applications, the process x_t is taken to represent single-period log returns on an asset bought at $t-1$ and sold at t, with $x_t = \log p_t - \log p_{t-1}$, where p_t is the asset price. The prediction density for x_t derived in this paper hence corresponds to a single-period return between $t-1$ and t, which for $t > 1$ implies buying an asset in the future and selling it one period after.

This is different from a buy-and-hold strategy of buying an asset at time 0, say, and selling it at time $t = 2, 3, \ldots$, which correspond to multi-period returns of the type $\sum_{j=1}^t x_j = \log p_t - \log p_0$. Multi-period returns are not covered by the results in the present paper. Andersen et al. (2006, pp 811–812) discuss approximations to the prediction of multi-period returns, and the possible use of the Central Limit Theorem on the sum $\sum_{j=1}^t x_j$ as proposed in Diebold (1988).

The map from x_{t-1} to σ_t^2, evaluated replacing σ_{t-1}^2 by the unconditional variance (assuming it exists), is known as the News Impact Curve, see Engle and Ng (1993). As it is well known, when $\lambda > 0$ the News Impact Curve is asymmetric, with a steeper slope for $x_{t-1} < 0$. Despite the possible presence of an asymmetric News Impact Curve, the prediction p.d.f. of x_t is symmetric as long as the p.d.f. of ε_t is symmetric – a characteristic shared by most innovation densities used in practice – as stated in the following proposition.

Proposition 2.1 (*Symmetry*)**.** *Let the p.d.f. of ε_t be symmetric with form $f_\varepsilon(v) = g(v^2)$ for some Riemann-integrable function $g : \mathbb{R}_{0,+} \to \mathbb{R}_{0,+}$; then the prediction density $f_{x_t}(u)$ of x_t is symmetric, i.e. $f_{x_t}(u) = f_{x_t}(-u)$.*

Proof. Because σ_t and ε_t are independent, the conditional density of x_t given σ_t is $f_{x_t|\sigma_t}(u|r) = \frac{1}{r} f_{\varepsilon_t}(\frac{u}{r})$ by the usual transformation theorem. By assumption $f_{\varepsilon_t}(\frac{u}{r}) = g(\frac{u^2}{r^2})$, and hence $f_{x_t|\sigma_t}(u|r) = \frac{1}{r} g(\frac{u^2}{r^2})$. This implies

$$f_{x_t}(u) = \int_0^\infty f_{x_t,\sigma_t}(u,r)\mathrm{d}r = \int_0^\infty f_{x_t|\sigma_t}(u|r) f_{\sigma_t}(r)\mathrm{d}r = \int_0^\infty \frac{1}{r} g\left(\frac{u^2}{r^2}\right) f_{\sigma_t}(r)\mathrm{d}r = f_{x_t}(-u). \quad \square \tag{2.2}$$

This result is in contrast with the expectation that the prediction density of x_t could have asymmetries induced by the asymmetric News Impact Curve alone, see Andersen et al. (2006) page 811. Proposition 2.1 applies as a special case

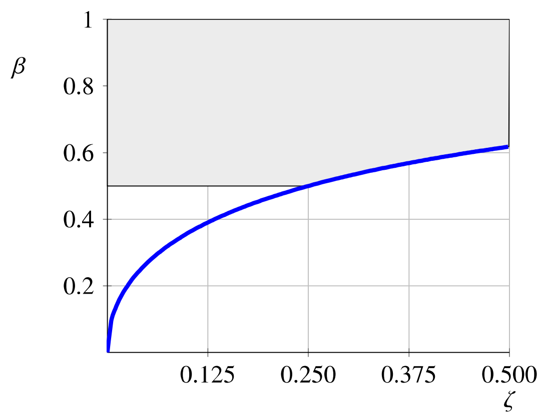

Fig. 1. Function $\underline{\beta}(\zeta)$ of $\zeta := \omega/(2\sigma_1^2)$. Blue line: $\underline{\beta}(\zeta) := -\zeta + \sqrt{\zeta^2 + 2\zeta}$. Shaded area: region $\beta \geq \max(\frac{1}{2}, \underline{\beta}(\zeta))$, see Assumption 2.2.*a* and *b*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to the case of Gaussian errors ε_t considered in this paper, and implies that the prediction density derived below will be symmetric.

In the specific case of Gaussian errors ε_t, the prediction density $f_{x_t}(u)$ is a normal variance mixture, see Barndorff-Nielsen et al. (1982) for a definition, thanks to the argument used in the proof of Proposition 2.1. These distributions are also called mixed-normal distributions, and they include the Student t and many other distributions in Econometrics, such as the ones in Abadir and Paruolo (1997).

When $\lambda > 0$, the signs $\boldsymbol{\varsigma} := (\varsigma_1, \ldots, \varsigma_{h-1})'$ of the innovations at times $1, \ldots, h-1$ are relevant for the conditional variance equation in (2.1), where each ς_i takes values $\{-1, 1\}$.[1] The set of all possible signs is indicated by $\mathcal{S}$, with 2^{h-1} elements, corresponding to all $h-1$-vectors $\boldsymbol{s} = (s_1, \ldots, s_{h-1})'$ with entries equal to ± 1.

The prediction p.d.f. is derived solving explicitly an integral similar to the one in (2.2). The integrand is expanded using binomial expansions which are convergent under appropriate conditions on the β parameter, which needs to be not too small relative to ω and σ_1^2, as specified in Assumption 2.2.[2] For instance $\beta > 0.61803$ is sufficient for Assumption 2.2 to hold, which is usually satisfied in practice, see e.g. Bampinas et al. (2018).

In order to state Assumption 2.2, define $\zeta := \omega/2\sigma_1^2$; note that ζ is known at time 0 for given parameter values $\boldsymbol{\theta}$, and that one has $0 < \zeta < \frac{1}{2}$ because $\sigma_1^2 > \omega$. Finally define the function $\underline{\beta}(\zeta) := -\zeta + \sqrt{\zeta^2 + 2\zeta}$, which is increasing in ζ and bounded by $\lim_{\zeta \to \frac{1}{2}} \underline{\beta}(\zeta) = (-1 + \sqrt{5})/2 \approxeq 0.61803$.

Assumption 2.2 (*Assumption on* β)**.** If $\omega > 0$ then assume that:

a. for $h = 3$, one has $\beta \geq \underline{\beta}(\zeta)$;
b. for $h > 3$, one has $\beta \geq \max(\frac{1}{2}, \underline{\beta}(\zeta))$.

If $\omega = 0$ no assumption is made on β.

Note that when $\omega = 0$, as in the RiskMetrics case, no restriction is imposed on β; the same applies for $h = 2$ step ahead also when $\omega > 0$. Assumption 2.2.*a* and *b* are graphed in Fig. 1.

A final piece of notation is introduced, which involves $\Psi(a; c; z)$, the confluent hypergeometric function of the second kind, also known as Tricomi function or Kummer U function, see Abadir (1999) and Gradshteyn and Ryzhik (2007, section 9.21). Its integral representation is

$$\Psi(a; c; z) = \frac{1}{\Gamma(a)} \int_{\mathbb{R}_+} \exp(-zt)\, t^{a-1} (1+t)^{c-a-1} \, \mathrm{d}t, \tag{2.3}$$

where $\Gamma(a) = \int_{\mathbb{R}_+} \exp(-t) t^{a-1} \mathrm{d}t$ is the Gamma function and here $\mathrm{Re}(z) > 0$, $\mathrm{Re}(a) > 0$. Eq. (2.3) has a structure similar to the Gamma function; computationally, $\Psi(a; c; z)$ is pre-programmed in various software packages, like the exp function.

The Ψ function enters the $A_h(r, \boldsymbol{\theta}, \boldsymbol{\varsigma})$ quantities defined below, which depend on a scalar r, the vector of parameters $\boldsymbol{\theta}$ and on the vector of signs $\boldsymbol{\varsigma}$. For conciseness $A_h(r, \boldsymbol{\theta}, \boldsymbol{\varsigma})$ is simply denoted as $A_{h,\boldsymbol{s}}(r)$ where $\boldsymbol{s}$ is the value taken by the

[1] Note that 0 is not included in the set of possible values for ς_i because $\Pr(\varsigma_i = 0) = 0$ whenever ε_i is a continuous random variable.

[2] The condition is only sufficient but it may not be necessary; further research will show if this condition can be improved.

vector of signs ς, and equals

$$A_{h,\boldsymbol{s}}(r) := \left(\omega+\beta\sigma_1^2\right)^{r+\frac{1}{2}} \left(\beta\sigma_1^2\right)^{-\frac{1}{2}} \pi^{\frac{h-1}{2}} \sum_{k_1,\ldots,k_{h-2}} \binom{r}{k_1}\binom{r-k_1}{k_2}\cdots\binom{r-K_{h-3}}{k_{h-2}} \left(\frac{\omega}{\omega+\beta\sigma_1^2}\right)^{K_{h-2}} \beta^{(h-2)r-\sum_{i=1}^{h-2}K_i} \cdot \prod_{t=1}^{h-2} \Psi\left(\frac{1}{2}; r-K_t+\frac{3}{2}; \frac{\beta}{2\alpha_{h-t}}\right) \Psi\left(\frac{1}{2}; r-K_{h-2}+\frac{3}{2}; \frac{\omega+\beta\sigma_1^2}{2\alpha_1\sigma_1^2}\right), \tag{2.4}$$

where $K_0 := 0$, $K_i := \sum_{t=1}^{i} k_t$. The multiple sum is defined for $h \geq 3$ as $\sum_{k_1,\ldots,k_{h-2}} := \sum_{k_1=0}^{r}\sum_{k_2=0}^{r-k_1}\cdots\sum_{k_{h-2}=0}^{r-K_{h-3}}$, where the individual sums extend to ∞ if $r \notin \mathbb{N}_0 := \mathbb{N}\cup 0$. Note that $A_{h,\boldsymbol{s}}(r)$ depends on $\boldsymbol{s}$ via α_{h-t}, $t = 1,\ldots,h-1$. For $h = 2$ the sum $\sum_{k_1,\ldots,k_{h-2}}$ and the product $\prod_{t=1}^{h-2}$ are empty and (2.4) reduces to

$$A_{2,s_1}(r) = \left(\omega+\beta\sigma_1^2\right)^{r+\frac{1}{2}} \left(\beta\sigma_1^2\right)^{-\frac{1}{2}} \pi^{\frac{1}{2}} \Psi\left(\frac{1}{2}; r+\frac{3}{2}; \frac{\omega+\beta\sigma_1^2}{2\alpha_1\sigma_1^2}\right),$$

which depends on s_1 via α_1.

3. Main results

This section presents explicit closed formulae for the prediction p.d.f. $f_{x_h}(u)$ in Theorem 3.1 and in Theorem 3.3.

Theorem 3.1 (*GARCH(1,1) Prediction Density*)**.** *Assume that ε_t are i.i.d.* N(0, 1) *and let Assumption* 2.2 *hold; then one has, for $h \geq 2$, and $\omega \geq 0$*

$$f_{x_h}(u) = (2\pi)^{-\frac{h}{2}} \sum_{j=0}^{\infty} \frac{1}{j!}\left(\frac{-u^2}{2}\right)^j c_j, \qquad -\infty < u < \infty \tag{3.1}$$

where $c_j := 2^{-h+1}\sum_{\boldsymbol{s}\in\mathcal{S}} c_{j,\boldsymbol{s}}$, $c_{j,\boldsymbol{s}} := \gamma_{h,\boldsymbol{s}}^{\frac{1}{2}} A_{h,\boldsymbol{s}}(-j-\frac{1}{2})$, $A_{h,\boldsymbol{s}}(\cdot)$ is defined in (2.4) *and $\gamma_{h,\boldsymbol{s}} := \beta^{h-1}/\prod_{t=1}^{h-1}\alpha_t$. The conditional density $f_{x_h|\varsigma}(u|\boldsymbol{s})$ has the same form as $f_{x_h}(u)$ in* (3.1) *with $c_{j,\boldsymbol{s}}$ in place of c_j.*

Proofs of this section are placed in the Appendix. Note that $c_{j,\boldsymbol{s}}$, like $\gamma_{h,\boldsymbol{s}}$ and $A_{h,\boldsymbol{s}}(\cdot)$, depends on the value $\boldsymbol{s}$ of the vector of signs ς via α_t, $t = 1,\ldots,h-1$. Observe also that $c_{j,\boldsymbol{s}}$ does not depend on the evaluation point u, and hence the $c_{j,\boldsymbol{s}}$ coefficients can be computed only once for the whole density. Immediate consequences of Theorem 3.1 are collected in the following corollary.

Corollary 3.2 (*C.d.f. and Moments*)**.** *The prediction c.d.f. for $h \geq 2$ is given by*

$$F_{x_h}(u) = \frac{1}{2} + (2\pi)^{-\frac{h}{2}} \sum_{j=0}^{\infty} \frac{u}{j!\,(2j+1)}\left(-\frac{u^2}{2}\right)^j c_j \qquad -\infty < u < \infty \tag{3.2}$$

where c_j is as in Theorem 3.1*; the odd moments for x_h are* 0 *and the even moments are equal to*

$$\mathrm{E}(x_h^{2m}) = 2^{m-\frac{3}{2}(h-1)}\pi^{-\frac{h}{2}}\Gamma\left(m+\frac{1}{2}\right)\sum_{\boldsymbol{s}\in\mathcal{S}}\gamma_{h,\boldsymbol{s}}^{\frac{1}{2}} A_{h,\boldsymbol{s}}(m) \qquad m = 1,2,\ldots \tag{3.3}$$

Recall that $\gamma_{h,\boldsymbol{s}} := \beta^{h-1}/\prod_{t=1}^{h-1}\alpha_t$ and $A_{h,\boldsymbol{s}}(m)$ depend on $\boldsymbol{s}$ via α_t, $t = 1,\ldots,h-1$, see (2.4). The moments of a GARCH(1,1) were first obtained in Baillie and Bollerslev (1992), see their equations (34) and (35). Note that the $A_{h,\boldsymbol{s}}(m)$ coefficients that appear in the moments are made of finite sums extending to m and do not fall in the logarithmic case, unlike in Theorem 3.1.[3] In fact, $m-k \in \{0,1,\ldots,m\}$ implies that

$$\Psi\left(\frac{1}{2}; \frac{3}{2}+m-k; \xi\right) = \left(\frac{1}{2}\right)_{m-k} \xi^{-\frac{1}{2}-m+k} \sum_{j=0}^{m-k} \frac{\binom{m-k}{j}}{\binom{-\frac{1}{2}+m-k}{j}} \frac{\xi^j}{j!}$$

is a finite sum, where $(a)_j := \prod_{i=0}^{j-1}(a+i)$ denotes Pochhammer's symbol, see e.g. Abadir (1999) eq. (2).

Formula (3.1) in Theorem 3.1 is alternating in sign, and converges absolutely; the absolute value of its terms tends to first increase and then decrease. For large values of u^2 this behavior may cause the accumulation of precision errors, resulting in numerical failure to produce an accurate value of the sum. This observation led to an alternative expression for the p.d.f., which is presented in the next theorem.

[3] The logarithmic case is when the second argument in Tricomi's Ψ function is in $\mathbb{Z}$ (i.e. is an integer), see Abadir (1999) and reference therein.

Theorem 3.3 (*Alternative Formulae for the GARCH(1,1) Prediction Density*)**.** *Under the same assumptions of Theorem 3.1 one has for $h \geq 2$,*

$$f_{x_h}(u) = (2\pi)^{-\frac{h}{2}} \sum_{j=0}^{\infty} \frac{\mu_j(u^2)}{j!} c_j^\star, \qquad -\infty < u < \infty \tag{3.4}$$

where $c_j^\star := 2^{-1} \sum_{\boldsymbol{s} \in \mathcal{S}} c_{j,\boldsymbol{s}}^\star$; $\mu_j(\cdot)$ and $c_{j,\boldsymbol{s}}^\star$ depend on h and ω in the following way:

- *for $h = 2$ and $\omega \geq 0$, $\mu_j(\cdot)$ is defined as $\mu_j(w) := \mathrm{e}^{-\rho w}(\rho w)^j$, $\rho := 1/(2(\omega + \beta\sigma_1^2))$ and*

$$c_{j,s}^\star := \pi^{\frac{1}{2}} (\sigma_1^2 \alpha_1)^{-\frac{1}{2}} \left(\frac{1}{2}\right)_j \Psi\left(j + \frac{1}{2}; 1; \frac{\omega + \beta\sigma_1^2}{2\alpha_1\sigma_1^2}\right); \tag{3.5}$$

- *for $h \geq 3$ and $\omega > 0$, $\mu_j(\cdot)$ is defined as $\mu_j(w) := \mathrm{e}^{-w/(2\omega)}(w/(2\omega))^j$, and*

$$c_{j,s}^\star := \pi^{\frac{h-1}{2}} \left(\prod_{t=1}^{h-1} \alpha_t^{-\frac{1}{2}}\right) (\sigma_1^2)^{-\frac{1}{2}} \sum_{k_1=0}^{\infty} \binom{-j-\frac{1}{2}}{k_1} \sum_{k_2=0}^{\infty} \cdots \sum_{k_{h-2}=0}^{\infty} \beta^{-\sum_{i=1}^{h-2} K_i} \left(\frac{\omega}{\omega + \beta\sigma_1^2}\right)^{K_{h-2}} \prod_{i=2}^{h-2} \binom{-\frac{1}{2} - K_{i-1}}{k_i} \cdot \prod_{t=1}^{h-2} \Psi\left(\frac{1}{2}; 1 - K_t; z_{h-t}\right) \Psi\left(\frac{1}{2}; 1 - K_{h-2}; z_1\right), \tag{3.6}$$

with $K_0 := 0$, $K_t := \sum_{i=1}^{t} k_i$, $z_1 := (\omega + \beta\sigma_1^2)/(2\alpha_1\sigma_1^2)$, $z_i := \beta/(2\alpha_i)$, $i = 2, \ldots, h-1$;

- *for $h \geq 2$ and $\omega = 0$, $\mu_j(\cdot)$ is defined as $\mu_j(w) := (-w/(2\sigma_1^2\beta^{h-1}))^j$ and*

$$c_{j,s}^\star := \pi^{\frac{h-1}{2}} \left(\sigma_1^2 \prod_{t=1}^{h-1} \alpha_t\right)^{-\frac{1}{2}} \prod_{t=1}^{h-1} \Psi\left(\frac{1}{2}; 1 - j; \frac{\beta}{2\alpha_t}\right). \tag{3.7}$$

The conditional density $f_{x_h|\mathcal{S}}(u|\boldsymbol{s})$ has the same form as $f_{x_h}(u)$ in (3.4) with $c_{j,\boldsymbol{s}}^\star$ in place of $c_j^\star$.

The prediction p.d.f. formulae in Eqs. (3.4) and (3.7) for $\omega = 0$ hold for all values of the parameters, see Assumption 2.2. Unlike the sums in Theorem 3.1, the expressions in (3.4), (3.5), (3.6) involve non-negative terms, containing a factor e^{-au^2} with $a > 0$. When $u = 0$, e^{-au^2} equals 1, while for large $|u|$, e^{-au^2} becomes very small. Eqs. (3.1) and (3.4) offer alternative ways to compute the same p.d.f.

4. Implementation

The prediction p.d.f.s are plotted in Figs. 2 to 4 for various values of the parameters $\boldsymbol{\theta}$, using formula (3.1).[4] In order to compare these densities with the N(0,1) p.d.f., the variate x_t is standardized as $\widetilde{x}_h := x_h/\sigma_{h|0}$ where $\sigma_{h|0}^2$ indicates the variance of x_t conditional on information available at time 0.[5]

Fig. 2 shows the standardized prediction densities in case $h = 2$ for $\lambda = 0$ and different pairs of (α, β) values. It shows that the deviations from the Gaussian can be substantial; the prediction densities are more similar to a Gaussian when the ratio β/α is large.

Fig. 3 shows the GJR GARCH(1,1) case with $\lambda > 0$. It is seen that $\lambda > 0$ increases the kurtosis, but has no effect on symmetry, see Proposition 2.1. Finally, Fig. 4 shows that in the RiskMetrics case the prediction p.d.f. is very close to the Gaussian.

The explicit form of the prediction density allows one to compute the VaR $Q_{h,p}$ defined as the (negative) of the p quantile of the prediction distribution, i.e. $p = \Pr(x_h < -Q_{h,p})$ or the Expected Shortfall $ES_{h,p}$, defined as $ES_{h,p} := -\mathrm{E}(x_h|x_h < -Q_{h,p})$, see e.g. Francq and Zakoian (2015). These functionals depend on the parameter vector $\boldsymbol{\theta} = (\omega, \alpha, \beta, \lambda)'$.

Suppose now that the GJR GARCH(1,1) in Eq. (2.1) has been estimated on a sample of data $\{x_t\}_{t=-T+1}^{0}$ by Quasi Maximum Likelihood (QML). Let $\widehat{\boldsymbol{\theta}}$ be the corresponding QML Estimator and $\boldsymbol{\theta}_0$ the (pseudo)-true value of the parameters. Under appropriate regularity conditions, see Lee and Hansen (1994), Jensen and Rahbek (2004) and Arvanitis and Louka (2017) and references therein, one has results of the type $T^{\frac{1}{2}}\boldsymbol{R}'(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{w} \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Omega_R})$, where $\xrightarrow{w}$ indicates convergence in distribution as $T \to \infty$, and $\boldsymbol{R}$ indicates a full-column-rank matrix with r columns.

The associated asymptotic confidence region has form

$$A_\eta = \{\boldsymbol{R}'\boldsymbol{\theta} : (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\boldsymbol{R}\boldsymbol{\Omega}_{\boldsymbol{R}}^{-1}\boldsymbol{R}'(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq c_\eta\} \tag{4.1}$$

---

4 Computations were performed in Mathematica. A Matlab implementation is provided in the Supplementary Material.

5 One finds $\sigma_{1|0}^2 = \sigma_1^2 = \omega + \alpha_0 x_0^2 + \beta\sigma_0^2$ and $\sigma_{t|0}^2 = \omega + \sigma_{t-1|0}^2(\alpha^\circ + \beta)$ for $t = 2, 3, \ldots$ where $\alpha^\circ := \mathrm{E}_{t-1}(\alpha_t) = \alpha + \lambda/2$, and use is made of $\Pr(x_{t-1} < 0) = \frac{1}{2}$.

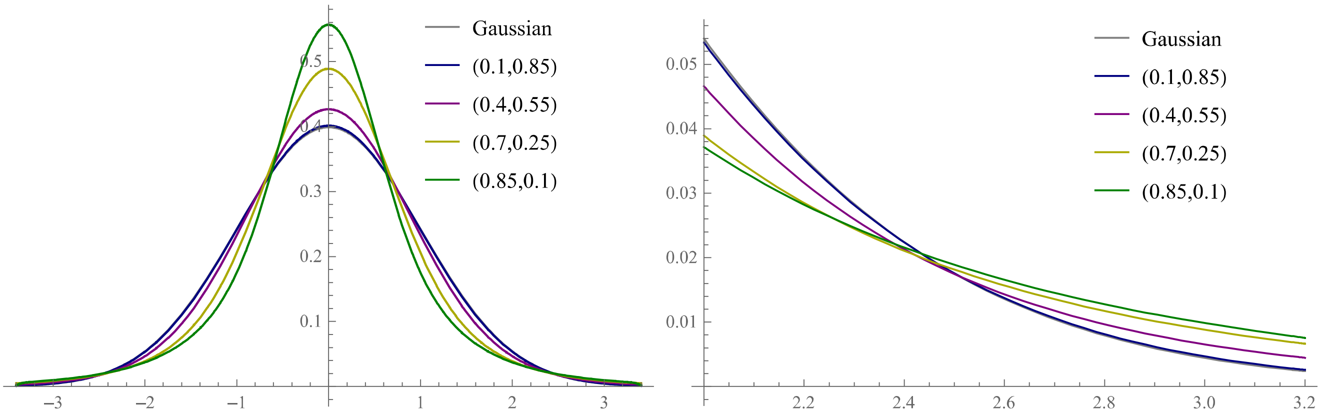

Fig. 2. Prediction density $f_{x_2}(u)$ for standardized x_2, $\omega = 0.1, \sigma_0^2 = 1; x_0^2 = 1$ varying values of (α, β). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

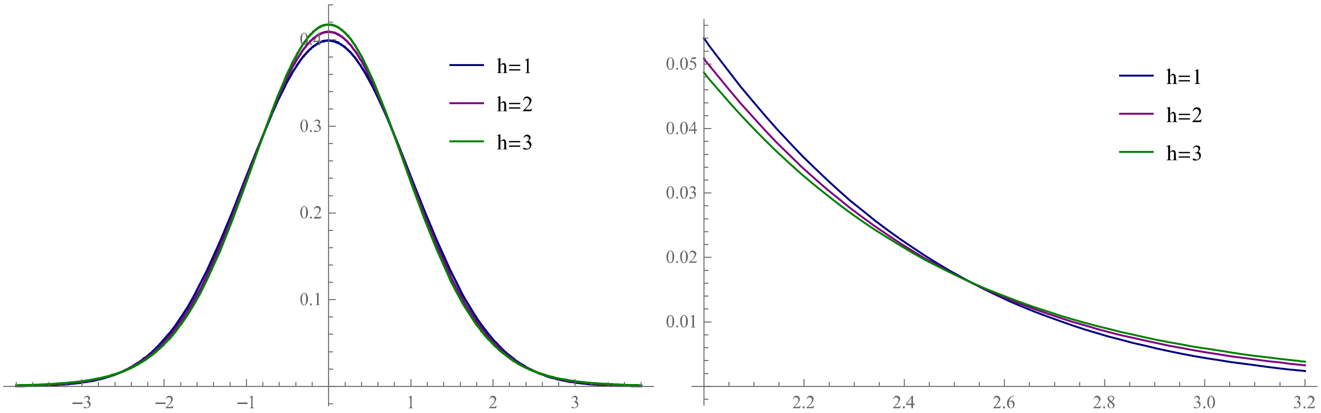

Fig. 3. Prediction density for standardized x_h, $h = 1, 2, 3$, in blue, purple and green respectively, $\omega = 0.25, \alpha = 0.1, \beta = 0.7, \sigma_0^2 = 1; x_0^2 = 1, \lambda = 0.2$ ($h = 1$ is standard Gaussian). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

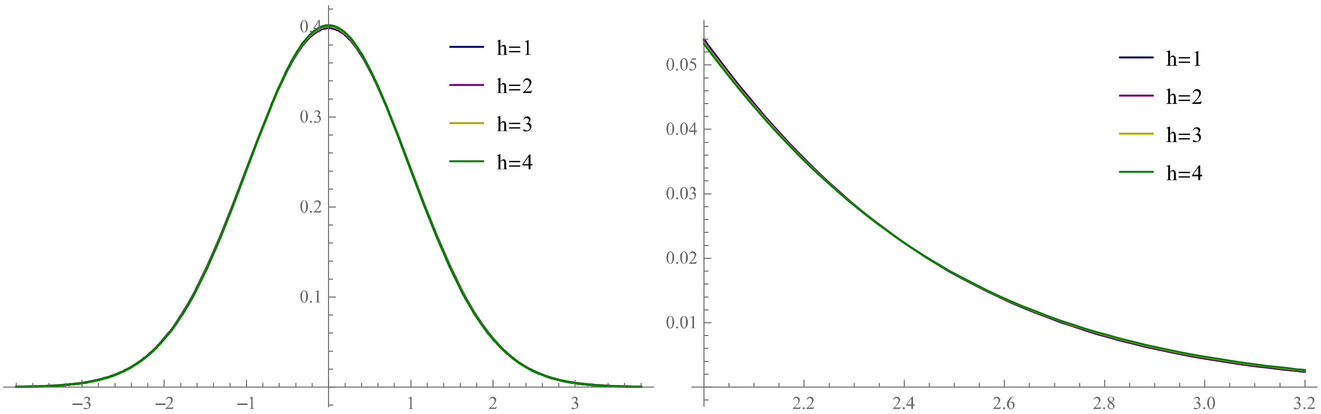

Fig. 4. RiskMetrics prediction p.d.f. $f_{x_h}(u)$ for standardized x_h, $\omega = 0, \sigma_0^2 = 1; x_0^2 = 1$ and $\alpha = 0.06$, $\beta = 0.94$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where c_η is the $1 - \eta$ quantile of a $\chi^2(r)$ distribution. The region A_η has the property that $\Pr(\boldsymbol{R}'\boldsymbol{\theta}_0 \in A_\eta) \to 1 - \eta$. Note that in (4.1) $\boldsymbol{\Omega}_{\boldsymbol{R}}$ can be replaced by a consistent estimator.

Consider next a (multivariate) functional of interest $\boldsymbol{\varphi}$, such as the VaR Q or the Expected Shortfall ES or both, which depends on $\boldsymbol{\theta}$, $\boldsymbol{\varphi} = \boldsymbol{\varphi}(\boldsymbol{\theta})$. Define also the set of values B_η taken by the $\boldsymbol{\varphi}$ map for any value of $\boldsymbol{\theta}$ in A_η, i.e.

$$B_\eta = \boldsymbol{\varphi}(A_\eta) = \{\boldsymbol{\varphi}(\boldsymbol{\theta}), \boldsymbol{\theta} \in A_\eta\}. \tag{4.2}$$

Then the following proposition shows that B_η is an uncertainty region for $\boldsymbol{\varphi}$ with at least asymptotic coverage equal to $1-\eta$, similarly to Fanelli and Paruolo (2010) Proposition 1.

Proposition 4.1 (*Uncertainty Region*). *B_η is an uncertainty region for $\boldsymbol{\varphi}$ with at least asymptotic coverage equal to $1-\eta$, i.e.* $\Pr(\boldsymbol{\varphi}(\boldsymbol{\theta}_0) \in B_\eta) \to \nu \geq 1-\eta$.

Proof. Simply note that the counter-images satisfy $\boldsymbol{\varphi}^{-1}(B_\eta) \subset A_\eta$. □

In practice, one needs to compute the set $\boldsymbol{\varphi}(A_\eta)$. Assuming for simplicity that $\boldsymbol{\varphi}(A_\eta)$ is univariate, indicated here as $\varphi(A_\eta)$, an uncertainty region would be the interval (φ_1, φ_2) where $\varphi_1 = \inf_{\boldsymbol{\theta}\in A_\eta}\{\varphi(\boldsymbol{\theta})\}$ and $\varphi_2 = \sup_{\boldsymbol{\theta}\in A_\eta}\{\varphi(\boldsymbol{\theta})\}$. One way to approximate the interval (φ_1, φ_2) is to calculate the extremes of $\varphi(\theta)$ for a grid of points θ in A_η. A grid of points in A_η can for instance be constructed as $\boldsymbol{R}'\boldsymbol{\theta} = (c_\eta \boldsymbol{\Omega}_{\boldsymbol{R}})^{\frac{1}{2}}\boldsymbol{u} + \boldsymbol{R}'\widehat{\boldsymbol{\theta}}$ starting from uniformly distributed points $\boldsymbol{u}$ in the unit ball of dimension r, generated e.g. as in Harman and Lacko (2010).

5. Conclusions

This paper derives the analytical form of the prediction density of a Gaussian GJR GARCH(1,1) process. This can be used to evaluate the probability of tail events or of functionals that may be of interest for risk valuations for single-period returns at any time horizon in the future. The analytical form of the prediction density can be very far from Gaussian. Estimation uncertainty for the estimation of the GJR GARCH parameters can be translated in uncertainty regions for risk functionals for single-period returns at any future time horizon.

The present results can be extended to other innovation distribution of current use; these extensions are left to future research.

Appendix. Proofs of the main results in Section 3

This Appendix reports proofs of Theorem 3.1, Corollary 3.2 and Theorem 3.3, which are based on several preliminary lemmas. The following compact notation is used for $h-1$-dimensional vectors: $\boldsymbol{z} := (x_1^2, \ldots, x_{h-1}^2)'$, $\boldsymbol{\varsigma} := (\varsigma_1, \ldots, \varsigma_{h-1})'$. Where no ambiguity arises, parentheses are dropped in product notation for conciseness.

Proposition A.1 (*Densities*). *For symmetric $f_\varepsilon(v) = g(v^2)$, the symmetric p.d.f. $f_{x_h}(\cdot)$ is related to the p.d.f. of x_h^2 in the following way*

$$f_{x_h}(u) = f_{x_h^2}(u^2)\,|u|\,. \tag{A.1}$$

Moreover, $\Pr(\varsigma_h = -1) = \Pr(\varsigma_h = 1) = \frac{1}{2}$ *and one has*

$$f_{\boldsymbol{z},x_h^2|\boldsymbol{\varsigma}}(\boldsymbol{w}, w_h|\boldsymbol{s}) = \prod_{t=1}^{h} \left(w_t\sigma_t^2\right)^{-\frac{1}{2}} g\left(\frac{w_t}{\sigma_t^2}\right) \tag{A.2}$$

where σ_t^2 depends on w_{t-j} (the value of x_{t-j}^2) and s_{t-j} (the sign of x_{t-j}) for $j = 1, \ldots, t-1$ via (2.1).

Proof. Consider the transformation theorem to obtain x_h^2; from standard results (Mood et al., 1974, page 201, Example 19) or (Abadir et al., 2018, Exercise 7.32) one has

$$f_{x_h^2}(w) = \left(\frac{1}{2}\frac{1}{\sqrt{w}}f_{x_h}(-\sqrt{w}) + \frac{1}{2}\frac{1}{\sqrt{w}}f_{x_h}(\sqrt{w})\right)1_{w>0}. \tag{A.3}$$

By Proposition 2.1 one has $f_{x_h}(-\sqrt{w}) = f_{x_h}(\sqrt{w})$, and (A.3) simplifies into $f_{x_h^2}(w) = w^{-\frac{1}{2}} f_{x_h}(\sqrt{w})1_{w>0}$, or, letting u indicate $w^{\frac{1}{2}}$, and solving for $f_{x_h}(u)$, one finds $f_{x_h}(u) = |u| f_{z_h}(u^2)$, which proves (A.1).

One has by assumption that $f_\varepsilon(\epsilon) := g(\epsilon^2) := (2\pi)^{-\frac{1}{2}}\exp(-\epsilon^2/2)$. Hence, applying the transformation theorem to $z_t = x_t^2$ using the conditional distribution $f_{x_t|x_1,\ldots,x_{t-1}}(u|u_1, \ldots u_{t-1}) = (2\pi\sigma_t^2)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\frac{u^2}{\sigma_t^2}\right) = (\sigma_t^2)^{-\frac{1}{2}} g\left(\frac{u^2}{\sigma_t^2}\right)$ one finds

$$f_{z_t|x_1,\ldots,x_{t-1}}(w|u_1, \ldots u_{t-1}) = (w)^{-\frac{1}{2}} f_{x_t|x_1,\ldots,x_{t-1}}(w^{\frac{1}{2}}|u_1, \ldots u_{t-1})1_{w>0} = (w\sigma_t^2)^{-\frac{1}{2}} g\left(\frac{w}{\sigma_t^2}\right),$$

from which Eq. (A.2) follows. □

Densities are first computed conditionally on $\boldsymbol{\varsigma}$ and later they are marginalized with respect to it. The basic building block is given by the expression in (A.2). This density can be marginalized with respect to $\boldsymbol{z}$ as follows

$$f_{x_h^2|\boldsymbol{\varsigma}}(w_h|\boldsymbol{s}) = \int_{\mathbb{R}_+^{h-1}} f_{\boldsymbol{z},x_h^2|\boldsymbol{\varsigma}}(\boldsymbol{w}, w_h|\boldsymbol{s})\mathrm{d}\boldsymbol{w}. \tag{A.4}$$

Finally, $f_{x_h^2|\varsigma}(w_h|\mathbf{s})$ can be marginalized with respect to the signs ς using the mutual independence of the signs ς_{t-j} and the fact that $\Pr(\varsigma_h = -1) = \Pr(\varsigma_h = 1) = \frac{1}{2}$ for all t, due to the symmetry of g. One hence finds

$$f_{x_h^2}(w) = \sum_{\mathbf{s}\in\mathcal{S}} f_{x_h^2|\varsigma}(w|\mathbf{s})\Pr(\mathbf{s}) = 2^{-h+1}\sum_{\mathbf{s}\in\mathcal{S}} f_{x_h^2|\varsigma}(w|\mathbf{s}) \tag{A.5}$$

where the sum $\sum_{\mathbf{s}\in\mathcal{S}}$ is over $s_j \in \{-1, 1\}$, for $j = 1, \dots, h-1$. The prediction density $f_{x_h^2}(w)$ is found by combining (A.5), (A.4), (A.2), (A.1).

The next proposition reports a recursion for the volatility process, similar in spirit to eq. (6) in Nelson (1990), that turns out to be useful when solving the integral in (A.4). Let $\boldsymbol{y} := (y_1, \dots, y_{h-1})'$ and $\boldsymbol{v} := (v_1, \dots, v_{h-1})'$, where v_i indicates a value of the random variable y_i defined below.

Proposition A.2 (*Volatility and Transformations*). *The volatility process can be written as*

$$\sigma_h^2 = \omega + (1 + y_{h-1})\beta\sigma_{h-1}^2 \qquad y_h := \frac{\alpha_h}{\beta}\varepsilon_h^2. \tag{A.6}$$

For $h \geq 2$, σ_h^2 has the following recursive expression in terms of $y_1, \dots, y_{h-1}$:

$$\sigma_h^2 = \omega + (1 + y_{h-1})\left\{\omega\beta + (1 + y_{h-2})\left(\dots\left(\omega\beta^{h-2} + (1 + y_1)\,\beta^{h-1}\sigma_1^2\right)\right)\right\} \tag{A.7}$$

with $\sigma_1^2 = \omega + \beta\sigma_0^2 + \alpha_0 x_0^2$, which is measurable with respect to the information set at time 0*. Moreover, one has*

$$f_{x_h^2|\varsigma}(w|\mathbf{s}) = \left(\frac{\gamma_{h,\mathbf{s}}}{w}\right)^{\frac{1}{2}} \int_{\mathbb{R}_+^{h-1}} \prod_{t=1}^{h-1}\left(v_t^{-\frac{1}{2}} g\left(\frac{\beta}{\alpha_t}v_t\right)\right)\sigma_h^{-1} g\left(\frac{w}{\sigma_h^2}\right)\mathrm{d}v_t, \tag{A.8}$$

where $\gamma_{h,\mathbf{s}} := \beta^{h-1}/(\prod_{t=1}^{h-1}\alpha_t)$ and the right-hand side of (A.8) depend on values of the vector of signs $\mathbf{s}$ via $\alpha_t = \alpha + \frac{1}{2}\lambda(1 - s_t)$ with $t = 1, \dots, h-1$.

Proof. Consider $f_{\boldsymbol{z},x_h^2|\varsigma}(\boldsymbol{w}, w_h|\mathbf{s})$ from (A.2), and consider the transformation of from $\boldsymbol{z}$ to $\boldsymbol{y}$. Observe that the domain of integration remains $\mathbb{R}_+^{h-1}$, that the inverse transformation is $z_t = \beta\sigma_t^2 y_t/\alpha_t$, with Jacobian $\gamma_{h,\mathbf{s}}\prod_{t=1}^{h-1}\sigma_t^2$, $\gamma_{h,\mathbf{s}} := \beta^{h-1}/(\prod_{t=1}^{h-1}\alpha_t)$. Hence one finds

$$f_{\boldsymbol{y},x_h^2|\varsigma}(\boldsymbol{v}, w|\mathbf{s}) = \gamma_{h,\mathbf{s}}^{\frac{1}{2}} \prod_{t=1}^{h-1}\left(v_t^{-\frac{1}{2}} g\left(\frac{\beta}{\alpha_t}v_t\right)\right)\left(w\sigma_h^2\right)^{-\frac{1}{2}} g\left(\frac{w}{\sigma_h^2}\right)$$

from which (A.8) follows. □

Lemma A.3 (*Limits of Ψ*). *$\Psi(a; c; z) \to 0$ for $c \to -\infty$ for real and positive a and z; $\Psi(a; c; z) \to 0$ for $a \to \infty$ for any c and real and positive z.*

Proof. The proof uses the Lebesgue dominated convergence theorem, see e.g. Theorem 10.27 in Apostol (1974). Consider the integral representation (2.3) of $\Psi(a; c; z)$ for real and positive a and z. Note that for negative c and $t \geq 0$ one has

$$k_n(t) := \frac{1}{\Gamma(a)}\mathrm{e}^{-zt}t^{a-1}(1 + t)^{c-a-1} \leq \frac{1}{\Gamma(a)}\mathrm{e}^{-zt}t^{a-1} =: r(t),$$

where $k_n(t), r(t) > 0$, $\int_{\mathbb{R}_+} k_n(t)\mathrm{d}t = \Psi(a; c; z)$ and $\int_{\mathbb{R}_+} r(t)\mathrm{d}t = \frac{1}{\Gamma(a)}\int_{\mathbb{R}_+}\mathrm{e}^{-zt}t^{a-1}\mathrm{d}t = \frac{1}{\Gamma(a)}z^{-a}\Gamma(a) = z^{-a}$. The notation $k_n(t)$ is chosen here to indicate that a sequence of values a_n or c_n will be constructed. The above expression shows that $k_n(t)$ is dominated by the function $r(t)$, which is Lebesgue-integrable on $\mathbb{R}_+$.

Next observe that for any $t > 0$, and for $c_n \to -\infty$, one has $k_n(t) \to 0$. Hence $k_n(t)$ converges to the zero function $k(t) := 0$ on the whole $\mathbb{R}_+$, except for the point $t = 0$. By the dominated convergence theorem, $\lim_{c_n\to-\infty}\Psi(a; c_n; z) = \lim_{c_n\to-\infty}\int_{\mathbb{R}_+} k_n(t)\mathrm{d}t = \int_{\mathbb{R}_+} k(t)\mathrm{d}t = 0$. This proves that $\Psi(a; c; z) \to 0$ for $c \to -\infty$ for real and positive a and z.

Let now $a_n \to \infty$ for any c and real and positive z. Observe that for any $t \geq 0$, $\frac{t}{1+t} < 1$ and $\Gamma(a_n) \to \infty$, and hence

$$k_n(t) := \frac{1}{\Gamma(a_n)}\mathrm{e}^{-zt}t^{a_n-1}(1 + t)^{c-a_n-1} = \frac{1}{\Gamma(a_n)}\mathrm{e}^{-zt}(1 + t)^{c-2}\left(\frac{t}{1 + t}\right)^{a_n-1} \to 0.$$

Hence $k_n(t)$ converges to the zero function $k(t) := 0$ on the whole $\mathbb{R}_+$. By the dominated convergence theorem, $\lim_{a_n\to\infty}\Psi(a_n; c; z) = \lim_{a_n\to\infty}\int_{\mathbb{R}_+} k_n(t)\mathrm{d}t = \int_{\mathbb{R}_+} k(t)\mathrm{d}t = 0$. This proves that $\Psi(a; c; z) \to 0$ for $a \to \infty$ for real and positive z and any c. □

Lemma A.4 (*Conditions on* β). *Assumption 2.2 ensures that for any* $j \geq 2$

$$\omega\left(1-\sum_{i=1}^{j-1}\beta^i\right) \leq \beta^j\sigma_1^2, \tag{A.9}$$

which implies that in (A.6) one has

$$\omega \leq (1+\upsilon_{h-1})\,\beta\sigma_{h-1}^2. \tag{A.10}$$

Proof. For $j = 2$ the inequality (A.9) reads $\beta^2\sigma_1^2 + \omega\beta - \omega \geq 0$. Solving the quadratic on the left-hand side for β one finds two roots, $\beta_1 = (-\omega - \sqrt{\omega^2 + 4\omega\sigma_1^2})/(2\sigma_1^2) < 0$ and $\underline{\beta} = (-\omega + \sqrt{\omega^2 + 4\omega\sigma_1^2})/(2\sigma_1^2) > 0$, so that the quadratic is non-negative for $\beta \leq \beta_1$ or for $\beta > \underline{\beta}$. Note that $\underline{\beta} = \underline{\beta}(\zeta) = -\zeta + (\zeta^2 + 2\zeta)^{\frac{1}{2}}$ where $\zeta = \omega/(2\sigma_1^2)$. Because $\beta_1 < 0$ is not possible, this holds only when $\beta \geq \underline{\beta}(\zeta)$. This proves that (A.9) is valid for $j = 2$ for $\beta \geq \underline{\beta}(\zeta)$ and a fortiori also for $\beta \geq \max\{\frac{1}{2}, \underline{\beta}(\zeta)\}$.

An induction approach is used for $j > 2$. Assume that (A.9) is valid for some $j = j_0 \geq 2$ and $\beta \geq \max\{\frac{1}{2}, \underline{\beta}(\zeta)\}$; it can then be shown that (A.9) is valid also replacing j with $j+1$. To see this, take (A.9) for $j = j_0$ and multiply by β. One finds

$$\omega\left(\beta-\sum_{i=1}^{j_0-1}\beta^{i+1}\right) \leq \beta^{j_0+1}\sigma_1^2.$$

Because $\beta \geq \frac{1}{2}$, one has $\omega(1-\beta) \leq \omega\beta$, so that,

$$\omega\left(1-\beta-\sum_{i=1}^{j_0-1}\beta^{i+1}\right) \leq \omega\left(\beta-\sum_{i=1}^{j_0-1}\beta^{i+1}\right) \leq \beta^{j_0+1}\sigma_1^2.$$

Rearranging $1-\beta-\sum_{i=1}^{j_0-1}\beta^{i+1}$ as $1-\sum_{i=1}^{j_0}\beta^i$, one finds that (A.9) holds also for $j = j_0+1$. The induction step hence proves that (A.9) holds for any j if $\beta \geq \max\{\frac{1}{2}, \underline{\beta}(\zeta)\}$.

To show (A.10), observe that the minimum value for $\beta\sigma_{h-1}^2$ corresponds to $\upsilon_{h-2} = \cdots = \upsilon_1 = 0$ and equals $\omega\sum_{i=1}^{h-2}\beta^i + \beta^{h-1}\sigma_1^2$, which is greater than or equal to ω by (A.9), and hence $\omega \leq \beta\sigma_{h-1}^2 \leq (1+\upsilon_{h-1})\,\beta\sigma_{h-1}^2$. □

Lemma A.5 (*Binomial Expansion*). *Let* $\omega > 0$*; under Assumption 2.2 the following expansion holds for any* $m \in \mathbb{R}$ *and for* $q = 2, 3, \ldots$*:*

$$\left(\frac{\sigma_q^2}{\omega}\right)^m = \beta^{(q-2)m}\sum_{n_1=0}^{m}\sum_{n_2=0}^{m-n_1}\cdots\sum_{n_{q-2}=0}^{m-N_{q-3}}\beta^{-\sum_{i=1}^{q-2}N_i}\prod_{i=1}^{q-2}\binom{m-N_{i-1}}{n_i}(1+\upsilon_{q-i})^{m-N_i}\left(1+(1+\upsilon_1)\frac{\beta\sigma_1^2}{\omega}\right)^{m-N_{q-2}}, \tag{A.11}$$

where $N_0 := 0$, $N_t := \sum_{i=1}^t n_i$ *and the sums* $\sum_{n_1=0}^m\sum_{n_2=0}^{m-n_1}\cdots\sum_{n_{h-2}=0}^{m-N_{h-3}}$ *extend to* ∞ *if* $m \notin \mathbb{N}_0$. *Eq. (A.11) is implied by the following recursions, defined in terms of* $\varphi_h := \sigma_h^2/\omega$

$$\varphi_{q+1}^m = \sum_{n=0}^m\binom{m}{n}(1+\upsilon_q)^{m-n}\beta^{m-n}\varphi_q^{m-n}, \qquad q = 1, 2, \ldots. \tag{A.12}$$

If $\omega = 0$, *then one has directly*

$$\sigma_q^2 = \beta^{q-1}\sigma_1^2\prod_{i=1}^{q-1}(1+\upsilon_i). \tag{A.13}$$

Proof. Define $\varphi_h := \sigma_h^2/\omega$, and observe that from (A.6) one has

$$\varphi_{h-t+1} = 1 + (1+\upsilon_{h-t})\beta\varphi_{h-t} \qquad t = 1, \ldots, h-2.$$

Under Assumption 2.2, Eq. (A.10) in Lemma A.4 implies that one can employ binomial expansions of φ_{h-t+1}^m using decreasing powers of $(1+\upsilon_{h-t})\beta\varphi_{h-t}$. Hence

$$\varphi_{h-t+1}^m = \sum_{n=0}^m\binom{m}{n}(1+\upsilon_{h-t})^{m-n}\beta^{m-n}\varphi_{h-t}^{m-n} \qquad t = 1, \ldots, h-2.$$

This proves (A.12); note that if m is not a non-negative integer, the sum extends to ∞. Next reiterate the recursive expansion for $t = 1, 2, \ldots, h-2$ until the expression contains φ_2; this gives (A.11). Eq. (A.13) is obtained by recursive substitutions in (A.6) with $\omega = 0$. □

Lemma A.6 (*Integral*). *One has for $a, b, n, z > 0$*

$$\int_{\mathbb{R}_+} \exp(-zv)(a+bv)^n v^{-\frac{1}{2}} \mathrm{d}v = a^{n+\frac{1}{2}} b^{-\frac{1}{2}} \pi^{\frac{1}{2}} \Psi\left(\frac{1}{2}; n+\frac{3}{2}; \frac{a}{b}z\right). \tag{A.14}$$

Proof. Set $m := b/a$ and $t := mv$; note that, using the integral representation of Ψ in (2.3), one has

$$\int_{\mathbb{R}_+} (a+bv)^n \exp(-zv) v^{-\frac{1}{2}} \mathrm{d}v = a^n \int_{\mathbb{R}_+} (1+t)^n \exp\left(-\frac{z}{m}t\right)\left(m^{-1}t\right)^{-\frac{1}{2}} m^{-1}\mathrm{d}t$$
$$= a^n m^{-\frac{1}{2}} \int_{\mathbb{R}_+} (1+t)^n \exp\left(-\frac{z}{m}t\right) t^{-\frac{1}{2}}\mathrm{d}t = a^{n+\frac{1}{2}} b^{-\frac{1}{2}} \pi^{\frac{1}{2}} \Psi\left(\frac{1}{2}; n+\frac{3}{2}; \frac{az}{b}\right). \quad \square$$

Lemma A.7 (*Coefficients $A_{h,\mathbf{s}}(\cdot)$*). *Let*

$$A_{h,\mathbf{s}}(r) := \int_{\mathbb{R}_+^{h-1}} \exp\left(-\frac{1}{2}\sum_{t=1}^{h-1}\frac{\beta}{\alpha_t}v_t\right)(\sigma_h^2)^r \prod_{t=1}^{h-1}\frac{\mathrm{d}v_t}{\sqrt{v_t}}; \tag{A.15}$$

then

$$A_{2,s_1}(r) = (\omega+\beta\sigma_1^2)^{r+\frac{1}{2}}\left(\beta\sigma_1^2\right)^{-\frac{1}{2}}\pi^{\frac{1}{2}}\Psi\left(\frac{1}{2}; r+\frac{3}{2}; \frac{\omega+\beta\sigma_1^2}{2\alpha_1\sigma_1^2}\right), \tag{A.16}$$

and, if $\omega > 0$, under Assumption 2.2 one has for $h \geq 3$

$$A_{h,\mathbf{s}}(r) = \beta^{(h-2)r}\left(\beta\sigma_1^2\right)^{-\frac{1}{2}}\pi^{\frac{h-1}{2}}\sum_{k_1=0}^{r}\sum_{k_2=0}^{r-k_1}\cdots\sum_{k_{h-2}=0}^{r-K_{h-3}}\beta^{-\sum_{i=1}^{h-2}K_i}\prod_{t=1}^{h-2}\binom{r-K_{t-1}}{k_t}\Psi\left(\frac{1}{2}; r-K_t+\frac{3}{2}; \frac{\beta}{2\alpha_{h-t}}\right)\cdot$$
$$\cdot\,\omega^{K_{h-2}}(\omega+\beta\sigma_1^2)^{r-K_{h-2}+\frac{1}{2}}\Psi\left(\frac{1}{2}; r-K_{h-2}+\frac{3}{2}; \frac{\omega+\beta\sigma_1^2}{2\alpha_1\sigma_1^2}\right), \tag{A.17}$$

where $K_0 := 0$, $K_i := \sum_{t=1}^{i} k_i$, and the sums extend to ∞ if $r \notin \mathbb{N}_0$; (A.17) reduces to (A.16) for $h = 2$.

If $\omega = 0$ then

$$A_{h,\mathbf{s}}(r) = \beta^{(h-1)r}\pi^{\frac{h-1}{2}}(\sigma_1^2)^r\prod_{t=1}^{h-1}\Psi\left(\frac{1}{2}; r+\frac{3}{2}; \frac{\beta}{2\alpha_t}\right), \tag{A.18}$$

which equals expression (A.17) for $\omega = 0$.

Proof. Set $h = 2$ in (A.15) and use (A.14) with $a = \omega+\beta\sigma_1^2$, $b = \beta\sigma_1^2$ to show that

$$\int_{\mathbb{R}_+}\exp\left(-\frac{\beta}{2\alpha_1}v_1\right)\left(\omega+(1+v_1)\beta\sigma_1^2\right)^r v_1^{-\frac{1}{2}}\mathrm{d}v_1 = (\omega+\beta\sigma_1^2)^{r+\frac{1}{2}}(\beta\sigma_1^2)^{-\frac{1}{2}}\pi^{\frac{1}{2}}\Psi\left(\frac{1}{2}; r+\frac{3}{2}; \frac{\omega+\beta\sigma_1^2}{2\alpha_1}\right);$$

this proves Eq. (A.16).

Next consider the case $h \geq 3$. If $\omega > 0$, under Assumption 2.2 Eq. (A.9) holds for $2 \leq j \leq h$, and one can use expansion (A.11) in (A.15). One finds

$$A_{h,\mathbf{s}}(r) = \omega^r\beta^{(h-2)r}\sum_{k_1=0}^{r}\sum_{k_2=0}^{r-k_1}\cdots\sum_{k_{h-2}=0}^{r-K_{h-3}}\beta^{-\sum_{i=1}^{h-2}K_i}\prod_{t=1}^{h-2}\binom{r-K_{t-1}}{k_t}\int_{\mathbb{R}_+^{h-2}}(1+v_{h-t})^{r-K_t}\exp\left(-\frac{\beta v_{h-t}}{2\alpha_{h-t}}\right)v_{h-t}^{-\frac{1}{2}}\mathrm{d}v_{h-i}\cdot$$
$$\cdot\int_{\mathbb{R}_+}\left(1+(1+v_1)\frac{\beta\sigma_1^2}{\omega}\right)^{r-K_{h-2}}\exp\left(-\frac{\beta v_1}{2\alpha_1}\right)v_1^{-\frac{1}{2}}\mathrm{d}v_1$$
$$= \omega^r\beta^{(h-2)r}\left(\beta\sigma_1^2\right)^{-\frac{1}{2}}\pi^{\frac{h-2}{2}}\sum_{k_1=0}^{r}\sum_{k_2=0}^{r-k_1}\cdots\sum_{k_{h-2}=0}^{r-K_{h-3}}\beta^{-\sum_{i=1}^{h-2}K_i}\prod_{t=1}^{h-2}\binom{r-K_{t-1}}{k_t}\Psi\left(\frac{1}{2}; r-K_t+\frac{3}{2}; \frac{\beta}{2\alpha_{h-t}}\right)\cdot$$
$$\cdot\,\omega^{K_{h-2}-r}(\omega+\beta\sigma_1^2)^{r-K_{h-2}+\frac{1}{2}}(\beta\sigma_1^2)^{-\frac{1}{2}}\pi^{\frac{1}{2}}\Psi\left(\frac{1}{2}; r-K_{h-2}+\frac{3}{2}; \frac{\omega+\beta\sigma_1^2}{2\alpha_1\sigma_1^2}\right),$$

where the second equality follows by repeated applications of (A.14). Simplifying, one finds (A.17). If instead $\omega = 0$, using (A.13) and (A.14) one finds

$$A_{h,\mathbf{s}}(r) = \beta^{(h-1)r}(\sigma_1^2)^r\prod_{t=1}^{h-1}\int_{\mathbb{R}_+}(1+v_t)^r\exp\left(-\frac{\beta}{2\alpha_t}v_t\right)v_t^{-\frac{1}{2}}\mathrm{d}v_t = \beta^{(h-1)r}\pi^{\frac{h-1}{2}}(\sigma_1^2)^r\prod_{t=1}^{h-1}\Psi\left(\frac{1}{2}; r+\frac{3}{2}; \frac{\beta}{2\alpha_t}\right).$$

This proves (A.18). □

Proof of Theorem 3.1. The integral to be solved is

$$f_{z_h|\varsigma}(w_h|\mathbf{s}) = \frac{\sqrt{\gamma_{h,\mathbf{s}}/w_h}}{(2\pi)^{\frac{h}{2}}} \int_{\mathbb{R}_+^{h-1}} \exp\left(-\frac{1}{2}\left(\sum_{t=1}^{h-1} \frac{\beta}{\alpha_t} v_t + \frac{w_h}{\sigma_h^2}\right)\right)(\sigma_h^2)^{-\frac{1}{2}} \prod_{t=1}^{h-1} \frac{\mathrm{d}v_t}{\sqrt{v_t}}. \tag{A.19}$$

Expand $\exp(-w_h/(2\sigma_h^2))$ as $\sum_{j=0}^{\infty}(-w_h/2)^j(\sigma_h^2)^{-j}/j!$ and note that

$$\begin{aligned} f_{z_h|\varsigma}(w_h|\mathbf{s}) &= \frac{\gamma_{h,\mathbf{s}}^{\frac{1}{2}} w_h^{-\frac{1}{2}}}{(2\pi)^{\frac{h}{2}}} \sum_{j=0}^{\infty} \frac{1}{j!}\left(\frac{-w_h}{2}\right)^j \int_{\mathbb{R}_+^{h-1}} \exp\left(-\frac{1}{2}\sum_{t=1}^{h-1}\frac{\beta}{\alpha_t}v_t\right)(\sigma_h^2)^{-j-\frac{1}{2}} \prod_{t=1}^{h-1}\frac{\mathrm{d}v_t}{\sqrt{v_t}}. \\ &= \frac{\gamma_{h,\mathbf{s}}^{\frac{1}{2}} w_h^{-\frac{1}{2}}}{(2\pi)^{\frac{h}{2}}} \sum_{j=0}^{\infty} \frac{1}{j!}\left(\frac{-w_h}{2}\right)^j A_{h,\mathbf{s}}\left(-j-\frac{1}{2}\right) \end{aligned}$$

where $A_{h,\mathbf{s}}(\cdot)$ is defined in (A.15). Substituting from (A.17) one finds the expression for $c_{j,\mathbf{s}}$, which is valid both for $\omega > 0$ as well as for $\omega = 0$, see (A.18). When marginalizing with respect to $\boldsymbol{\varsigma}$, all elements in $\mathcal{S}$ are equally likely, and one deduces $c_j := 2^{-h+1}\sum_{\mathbf{s}\in\mathcal{S}} c_{j,\mathbf{s}}$. Note that if all $c_{j,\mathbf{s}}$ do not vary with $\mathbf{s}$, one has $c_j = c_{j,\mathbf{s}}$.

In order to show that $f_{x_h}(u)$ is absolutely summable for finite u, consider for instance the case $h = 2$; the p.d.f is the average of

$$f_{x_2|\varsigma_1}(u|s_1) = \frac{1}{2}(\pi\alpha_1\sigma_1^2)^{-\frac{1}{2}} \sum_{j=0}^{\infty}\frac{1}{j!}\left(-\rho u^2\right)^j \Psi\left(\frac{1}{2}; 1-j; z\right),$$

where

$$\rho := \frac{1}{2\left(\omega+\beta\sigma_1^2\right)}, \qquad z = \frac{\omega+\beta\sigma_1^2}{2\alpha_1\sigma_1^2} > 0.$$

Because $\Psi\left(\frac{1}{2}; 1-j; z\right)$ is finite and non-negative, monotonically decreasing towards 0 by Lemma A.3 for increasing j, one has $\sup_{j\in\mathbb{N}} \Psi\left(\frac{1}{2}; 1-j; z\right) = M < \infty$, so that

$$\sum_{j=0}^{\infty}\left|\frac{1}{j!}\left(-\rho u^2\right)^j \Psi\left(\frac{1}{2}; 1-j; \xi\right)\right| \le \sum_{j=0}^{\infty}\frac{1}{j!}(\rho u^2)^j \Psi\left(\frac{1}{2}; 1-j; \xi\right) \le M\sum_{j=0}^{\infty}\frac{1}{j!}(\rho u^2)^j = M\exp\left(\rho u^2\right)$$

where $\exp\left(\rho u^2\right)$ is finite for any finite u. One hence concludes that the series is absolutely convergent for any finite u. The case for $h > 2$ is similar. □

Proof of Corollary 3.2. The c.d.f.s are found by integrating termwise the p.d.f. from 0 to $u_h > 0$, where termwise integration is guaranteed by Theorem 10.26 in Apostol (1974). This delivers (3.2) for positive u. The symmetry of $f_{x_h}(\cdot)$ implies $F_{x_h}(0) = \frac{1}{2}$ and $F_{x_h}(u) = 1 - F_{x_h}(-u)$. Hence for $0 > u = -a$, say, with $a > 0$, one has

$$\begin{aligned} F_{x_h}(u) = 1 - F_{x_h}(a) &= \frac{1}{2} - (2\pi)^{-\frac{h}{2}} \sum_{j=0}^{\infty}\frac{1}{j!\,(2j+1)} a^{2j+1}(-2(\omega+\beta\sigma_1^2))^{-j} c_j \\ &= \frac{1}{2} + (2\pi)^{-\frac{h}{2}} \sum_{j=0}^{\infty}\frac{1}{j!\,(2j+1)} (-a)^{2j+1}(-2(\omega+\beta\sigma_1^2))^{-j} c_j \end{aligned}$$

which proves that the expressions in (3.2) is valid also for negative u.

The moments are derived as follows. From (A.19) one sees that

$$\mathrm{E}(x_h^{2m}|\varsigma) = \frac{\sqrt{\gamma_{h,\mathbf{s}}}}{(2\pi)^{\frac{h}{2}}} \int_{\mathbb{R}_+^h} w_h^{m-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{w_h}{\sigma_h^2}\right)\mathrm{d}w_h \prod_{t=1}^{h-1}\exp\left(-\frac{1}{2}\left(\frac{\beta}{\alpha_t}v_t\right)\right)\sigma_h^{-1} v_t^{-\frac{1}{2}}\mathrm{d}v_t.$$

Recall that $\int_{\mathbb{R}_+}\exp\left(-\frac{w}{2\sigma_h^2}\right) w^{m-\frac{1}{2}}\mathrm{d}w = \left(2\sigma_h^2\right)^{m+\frac{1}{2}}\Gamma\left(m+\frac{1}{2}\right)$ so that

$$\begin{aligned} \mathrm{E}(x_h^{2m}|\varsigma) &= 2^{m-\frac{h-1}{2}}\pi^{-\frac{h}{2}}\gamma_{h,\mathbf{s}}^{\frac{1}{2}}\Gamma\left(m+\frac{1}{2}\right)\int_{\mathbb{R}_+^{h-1}}\prod_{t=1}^{h-1}\exp\left(-\frac{1}{2}\left(\frac{\beta}{\alpha_t}v_t\right)\right)\left(\sigma_h^2\right)^m v_t^{-\frac{1}{2}}\mathrm{d}v_t \\ &= 2^{m-\frac{h-1}{2}}\pi^{-\frac{h}{2}}\Gamma\left(m+\frac{1}{2}\right)\gamma_{h,\mathbf{s}}^{\frac{1}{2}} A_{h,\mathbf{s}}(m) \end{aligned}$$

where $A_{h,\mathbf{s}}(\cdot)$ is defined in (A.15), which equals (A.17) (or (2.4)). Hence

$$\mathrm{E}(x_h^{2m}) = 2^{m-\frac{3}{2}(h-1)}\pi^{-\frac{h}{2}}\Gamma\left(m+\frac{1}{2}\right)\sum_{\mathbf{s}\in\mathcal{S}}\gamma_{h,\mathbf{s}}^{\frac{1}{2}}A_{h,\mathbf{s}}(m) \quad \square$$

Proof of Theorem 3.3. Consider first the case $h = 2$ and

$$f_{z_2|\varsigma_1}(w|s_1) = \frac{\gamma_{2,s_1}^{\frac{1}{2}} w^{-\frac{1}{2}}}{2\pi}\int_{\mathbb{R}_+}\exp\left(-\frac{1}{2}\left(\frac{\beta}{\alpha_1}v_1 + \frac{w}{\sigma_2^2}\right)\right)(\sigma_2^2)^{-\frac{1}{2}}\frac{\mathrm{d}v_1}{\sqrt{v_1}}. \tag{A.20}$$

Observe that σ_2^2 can be written as $b(1+\psi)$ with $b := \omega + \beta\sigma_1^2$ and $\psi := v_1/q$ where $q = (\omega+\beta\sigma_1^2)/\beta\sigma_1^2$, $\sigma_2^2 = \omega+(1+v_1)\beta\sigma_1^2 = \omega+\beta\sigma_1^2+\beta\sigma_1^2 v_1 = b(1+\psi)$. Next note that $w_2/(2\sigma_2^2) = \rho w_2/(1+\psi)$ with $\rho = 1/(2(\omega+\beta\sigma_1^2))$ and

$$\exp\left(-w_2/(2\sigma_2^2)\right) = \exp\left(-\rho w_2/(1+\psi)\right) = \exp\left(-\rho w_2\right)\exp\left(\rho w_2\psi/(1+\psi)\right),$$

where the last term can be expanded as $\exp\left(\rho w_2\psi/(1+\psi)\right) = \sum_{j=0}^{\infty}\frac{\rho^j w_2^j}{j!}\psi^j(1+\psi)^{-j}$.

Observe that $\gamma_{2,s_1} = \beta/\alpha_1$, $q\mathrm{d}\psi = \mathrm{d}v_1$, and set $z := (\omega+\beta\sigma_1^2)/(2\alpha_1\sigma_1^2) = \beta q/(2\alpha_1)$. Substituting these expressions in (A.20) one finds

$$\begin{aligned}
f_{z_2|\varsigma_1}(w|s_1) &= (2\pi)^{-1}\gamma_{2,s_1}^{\frac{1}{2}} w^{-\frac{1}{2}}\mathrm{e}^{-\rho w}\sum_{j=0}^{\infty}\frac{\rho^j w^j}{j!}\int_0^\infty \exp\left(-\frac{\beta q}{2\alpha_1}\psi\right)\psi^j(1+\psi)^{-j}(b(1+\psi))^{-\frac{1}{2}}(q\psi)^{-\frac{1}{2}}q\mathrm{d}\psi \\
&= (2\pi)^{-1}\left(\frac{\beta}{\alpha_1}\right)^{\frac{1}{2}} w^{-\frac{1}{2}}\left(\frac{q}{b}\right)^{\frac{1}{2}}\mathrm{e}^{-\rho w}\sum_{j=0}^{\infty}\frac{\rho^j w^j}{j!}\int_0^\infty \exp\left(-z\psi\right)(1+\psi)^{-\frac{1}{2}-j}(\psi)^{j-\frac{1}{2}}\mathrm{d}\psi \\
&= (2\pi)^{-1}\beta^{\frac{1}{2}}w^{-\frac{1}{2}}\left(\alpha_1\beta\sigma_1^2\right)^{-\frac{1}{2}}\mathrm{e}^{-\rho w}\sum_{j=0}^{\infty}\frac{\rho^j w^j}{j!}\Gamma\left(j+\frac{1}{2}\right)\Psi\left(j+\frac{1}{2};1;z\right)
\end{aligned}$$

by the integral representation of Ψ, see Eq. (2.3). By eq. (2) in Abadir (1999), one has $\Gamma\left(j+\frac{1}{2}\right) = \sqrt{\pi}\left(\frac{1}{2}\right)_j$. Substituting back and rearranging, one finds (3.4) and (3.5).

For $h \geq 3$, consider first the case $\omega > 0$ and the expression

$$f_{z_h|\varsigma}(w|\mathbf{s}) = \xi_{h,\mathbf{s}}\int_{\mathbb{R}_+^{h-1}}\exp\left(-\frac{w}{2\sigma_h^2}\right)\exp\left(-\frac{1}{2}\left(\sum_{t=1}^{h-1}\frac{\beta}{\alpha_t}v_t\right)\right)(\sigma_h^2)^{-\frac{1}{2}}\prod_{t=1}^{h-1}\frac{\mathrm{d}v_t}{\sqrt{v_t}}, \tag{A.21}$$

where $\xi_{h,\mathbf{s}} := \gamma_{h,\mathbf{s}}^{\frac{1}{2}}w^{-\frac{1}{2}}(2\pi)^{-\frac{h}{2}} = \beta^{\frac{h-1}{2}}\prod_{i=1}^{h-1}\alpha_i^{-\frac{1}{2}}w_h^{-\frac{1}{2}}(2\pi)^{-\frac{h}{2}}$. Let $\varphi_h := \sigma_h^2/\omega = 1+\psi$, $\psi := (1+v_{h-1})\beta\varphi_{h-1}$, $r := 1/(2\omega)$ and write

$$\exp\left(-\frac{w}{2\sigma_h^2}\right) = \mathrm{e}^{-rw}\exp\left(rw\frac{\psi}{1+\psi}\right) = \mathrm{e}^{-rw}\sum_{j=0}^{\infty}\frac{r^j w^j}{j!}\psi^j(1+\psi)^{-j}$$

Hence, substituting this in (A.21) and using $\sigma_h^2 = \omega\varphi_h$ one gets

$$f_{z_h|\varsigma}(w|\mathbf{s}) = \xi_{h,\mathbf{s}}\omega^{-\frac{1}{2}}\sum_{j=0}^{\infty}\mathrm{e}^{-rw}\frac{r^j w^j}{j!}\int_{\mathbb{R}_+^{h-1}}\varphi_h^{-j-\frac{1}{2}}(1+v_{h-1})^j\beta^j\varphi_{h-1}^j\exp\left(-\frac{1}{2}\left(\sum_{t=1}^{h-1}\frac{\beta}{\alpha_t}v_t\right)\right)\prod_{t=1}^{h-1}\frac{\mathrm{d}v_t}{\sqrt{v_t}} \tag{A.22}$$

Expanding $\varphi_h^{-j-\frac{1}{2}}$ as in Eq. (A.12) in Lemma A.5 one has $\varphi_h^{-j-\frac{1}{2}} = \sum_{n=0}^{\infty}\binom{-j-\frac{1}{2}}{n}(1+v_{h-t})^{-j-\frac{1}{2}-n}\beta^{-j-\frac{1}{2}-n}\varphi_{h-1}^{-j-\frac{1}{2}-n}$ and substituting in (A.22), one finds

$$\begin{aligned}
f_{z_h|\varsigma}(w|\mathbf{s}) &= \xi_{h,\mathbf{s}}\omega^{-\frac{1}{2}}\sum_{j=0}^{\infty}\mathrm{e}^{-rw}\frac{r^j w^j}{j!}\sum_{n=0}^{\infty}\binom{-j-\frac{1}{2}}{n}\cdot \\
&\cdot\int_{\mathbb{R}_+^{h-1}}(1+v_{h-1})^{-\frac{1}{2}-n}\beta^{-\frac{1}{2}-n}\varphi_{h-1}^{-\frac{1}{2}-n}\exp\left(-\frac{1}{2}\left(\sum_{t=1}^{h-1}\frac{\beta}{\alpha_t}v_t\right)\right)\prod_{t=1}^{h-1}\frac{\mathrm{d}v_t}{\sqrt{v_t}}
\end{aligned} \tag{A.23}$$

Next expand $\varphi_{h-1}^{-\frac{1}{2}-n}$ as in Eq. (A.11) in Lemma A.5 with $q = h-1$ and substitute in (A.23); one has:

$$f_{z_h|\varsigma}(w|\mathbf{s}) = \xi_{h,\mathbf{s}}\omega^{-\frac{1}{2}}\sum_{j=0}^{\infty}\mathrm{e}^{-rw}\frac{r^j w^j}{j!}\sum_{n=0}^{\infty}\beta^{(-\frac{1}{2}-n)(h-2)}\binom{-j-\frac{1}{2}}{n}\int_{\mathbb{R}_+^{h-1}}(1+v_{h-1})^{-\frac{1}{2}-n}\sum_{n_1=0}^{\infty}\cdots\sum_{n_{h-3}=0}^{\infty}\beta^{-\sum_{i=1}^{h-3}N_i}.$$

$$\cdot \prod_{t=1}^{h-3} \binom{-\frac{1}{2} - n - N_{t-1}}{n_t} (1 + v_{h-1-t})^{-\frac{1}{2}-n-N_t} \left(1 + (1 + v_1)\frac{\beta\sigma_1^2}{\omega}\right)^{-\frac{1}{2}-n-N_{h-3}} \exp\left(-\frac{1}{2}\left(\sum_{t=1}^{h-1} \frac{\beta}{\alpha_t} v_t\right)\right) \prod_{t=1}^{h-1} \frac{\mathrm{d}v_t}{\sqrt{v_t}}. \tag{A.24}$$

Relabel $k_1 := n$, $k_i := n_{i-1}$ for $i = 2, \ldots, h-2$ and note that $K_i = n + N_{i-1}$; hence the previous expression can be simplified as follows

$$f_{z_h|\varsigma}(w|\mathbf{s}) = \xi_{h,\mathbf{s}} \omega^{-\frac{1}{2}} \sum_{j=0}^{\infty} \mathrm{e}^{-rw} \frac{r^j w^j}{j!} \sum_{k_1=0}^{\infty} \binom{-j - \frac{1}{2}}{k_1} \sum_{k_2=0}^{\infty} \cdots \sum_{k_{h-2}=0}^{\infty} \beta^{-\frac{1}{2}(h-2) - \sum_{i=1}^{h-2} K_i} \prod_{i=2}^{h-2} \binom{-\frac{1}{2} - K_{i-1}}{k_i} \cdot$$
$$\cdot \int_{\mathbb{R}_+^{h-1}} \prod_{t=1}^{h-2} (1 + v_{h-t})^{-\frac{1}{2}-K_t} \left(1 + (1 + v_1)\frac{\beta\sigma_1^2}{\omega}\right)^{-\frac{1}{2}-K_{h-2}} \exp\left(-\frac{1}{2}\left(\sum_{t=1}^{h-1} \frac{\beta}{\alpha_t} v_t\right)\right) \prod_{t=1}^{h-1} \frac{\mathrm{d}v_t}{\sqrt{v_t}}. \tag{A.25}$$

By the results in Lemma A.6 the integral in the last line in (A.25) equals

$$\omega^{-K_{h-2}-\frac{1}{2}} (\omega + \beta\sigma_1^2)^{-K_{h-2}} (\beta\sigma_1^2)^{-\frac{1}{2}} \pi^{\frac{h-1}{2}} \prod_{t=1}^{h-2} \Psi\left(\frac{1}{2}; 1 - K_t; z_{h-t}\right) \Psi\left(\frac{1}{2}; 1 - K_{h-2}; z_1\right)$$

where z_i, $i = 1, \ldots, h-1$ are defined in (3.6). Hence, recalling that $\xi_{h,\mathbf{s}} = \beta^{\frac{h-1}{2}} \prod_{i=1}^{h-1} \alpha_i^{-\frac{1}{2}} w_h^{-\frac{1}{2}} (2\pi)^{-\frac{h}{2}}$, and rearranging one finds $f_{z_h|\varsigma}(w|\mathbf{s}) = (2\pi)^{-\frac{h}{2}} w^{-\frac{1}{2}} \sum_{j=0}^{\infty} \frac{\mu_j(w)}{j!} c_j^\star$ with $\mu_j(w) := \mathrm{e}^{-rw} r^j w^j$ and

$$c_j^\star := \pi^{\frac{h-1}{2}} \prod_{i=1}^{h-1} \alpha_i^{-\frac{1}{2}} (\sigma_1^2)^{-\frac{1}{2}} \sum_{k_1=0}^{\infty} \binom{-j - \frac{1}{2}}{k_1} \sum_{k_2=0}^{\infty} \cdots \sum_{k_{h-2}=0}^{\infty} \beta^{-\sum_{i=1}^{h-2} K_i} \left(\frac{\omega}{\omega + \beta\sigma_1^2}\right)^{K_{h-2}} \prod_{i=2}^{h-2} \binom{-\frac{1}{2} - K_{i-1}}{k_i} \cdot$$
$$\cdot \prod_{t=1}^{h-2} \Psi\left(\frac{1}{2}; 1 - K_t; z_{h-t}\right) \Psi\left(\frac{1}{2}; 1 - K_{h-2}; z_1\right),$$

which gives (3.6). Note that the second parameter $1 - K_t$ (respectively $1 - j$) of the Ψ functions in (3.6) (respectively (3.7)) converges to $-\infty$ for $k_i \to \infty$, so that $\prod_{t=1}^{h-2} \Psi\left(\frac{1}{2}; 1 - K_t; z_{h-t}\right) \Psi\left(\frac{1}{2}; 1 - K_{h-2}; z_1\right)$ converges to 0 for $k_i \to \infty$ by Lemma A.3.

Finally consider the case $\omega = 0$ for $h \geq 2$, write $\exp\left(-\frac{w_h}{2\sigma_h^2}\right) = \sum_{j=0}^{\infty} \left(-\frac{w_h}{2}\right)^j \frac{1}{j!} \left(\sigma_h^2\right)^{-j}$ so that

$$f_{z_h|\varsigma}(w|\mathbf{s}) = \xi_{h,s} \sum_{j=0}^{\infty} \left(-\frac{w}{2}\right)^j \frac{1}{j!} \int_{\mathbb{R}_+^{h-1}} \exp\left(-\frac{1}{2}\left(\sum_{t=1}^{h-1} \frac{\beta}{\alpha_t} v_t\right)\right) (\sigma_h^2)^{-j-\frac{1}{2}} \prod_{t=1}^{h-1} \frac{\mathrm{d}v_t}{\sqrt{v_t}}.$$

Note that for $\omega = 0$, by (A.13) $(\sigma_h^2)^m = \prod_{t=1}^{h-1} (1 + v_{h-t})^m \beta^{(h-1)m} (\sigma_1^2)^m$, which implies that

$$f_{z_h|\varsigma}(w|\mathbf{s}) = \xi_{h,\mathbf{s}} (\sigma_1^2)^{-\frac{1}{2}} \sum_{j=0}^{\infty} \left(-\frac{w}{2\sigma_1^2}\right)^j \frac{\beta^{(-j-\frac{1}{2})(h-1)}}{j!} \prod_{t=1}^{h-1} \int_{\mathbb{R}_+^{h-1}} \exp\left(-\frac{\beta}{2\alpha_t} v_t\right) (1 + v_t)^{-j-\frac{1}{2}} v_t^{-\frac{1}{2}} \mathrm{d}v_t$$
$$= (2\pi)^{-\frac{h}{2}} \prod_{i=1}^{h-1} \alpha_i^{-\frac{1}{2}} w^{-\frac{1}{2}} (\sigma_1^2)^{-\frac{1}{2}} \sum_{j=0}^{\infty} \left(-\frac{w}{2\sigma_1^2}\right)^j \frac{\beta^{-j(h-1)}}{j!} \pi^{\frac{h-1}{2}} \prod_{t=1}^{h-1} \Psi\left(\frac{1}{2}; 1 - j; \frac{\beta}{2\alpha_t}\right),$$

where $\xi_{h,\mathbf{s}} = (2\pi)^{-\frac{h}{2}} \beta^{\frac{h-1}{2}} \prod_{i=1}^{h-1} \alpha_i^{-\frac{1}{2}} w^{-\frac{1}{2}}$. This completes the proof. □

Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2022.04.010.

References

Abadir, K.M., 1999. An introduction to hypergeometric functions for economists. Econometric Rev. 18 (3), 287–330.
Abadir, K.M., Heijmans, R.D.H., Magnus, J.R., 2018. Statistics. Econometric Exercises, Cambridge University Press.
Abadir, K.M., Magnus, J., 2002. Notation in econometrics: a proposal for a standard. Econom. J. 5 (1), 76–90.
Abadir, K.M., Paruolo, P., 1997. Two mixed normal densities from cointegration analysis. Econometrica 65 (3), 671–680.
Andersen, T., Bollerslev, T., Christoffersen, P.F., Diebold, F.X., 2006. Volatility and correlation forecasting. In: Elliott, G., Granger, C.W., Timmermann, A. (Eds.), Handbook of Economic Forecasting, Volume 1. Elsevier, New York.
Apostol, T.M., 1974. Mathematical Analysis, second ed. Addison-Wesley.

Arvanitis, S., Louka, A., 2017. Martingale Transforms with Mixed Stable Limits and the QMLE for Conditionally Heteroskedastic Models. Technical report, SSRN.
Baillie, R.T., Bollerslev, T., 1992. Prediction in dynamic models with time-dependent conditional variances. J. Econometrics 52 (1–2), 91–113.
Bampinas, G., Ladopoulos, K., Panagiotidis, T., 2018. A note on the estimated GARCH coefficients from the S & P1500 universe. Appl. Econ. 50 (34–35), 3647–3653.
Barndorff-Nielsen, O., Kent, J., Sørensen, M., 1982. Normal variance-mean mixtures and z-distributions. Internat. Statist. Rev. 50, 145–159.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econometrics 31 (3), 307–327.
Bollerslev, T., 1987. A conditionally heteroskedastic time series model for speculative prices and rates of return. Rev. Econ. Stat. 69 (3), 542–574.
Bollerslev, T., Chou, R., Kroner, K., 1992. ARCH modeling in finance: a review of the theory and empirical evidence. J. Econometrics 52, 5–59.
Bollerslev, T., Russell, J.R., Watson, M.W., 2010. Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle. Oxford University Press, Oxford.
Diebold, F.X., 1988. Empirical Modeling of Exchange Rate Dynamics / Francis X. Diebold. In: Lecture Notes in Economics and Mathematical Systems, vol. 303, Springer-Verlag, Berlin.
Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kindom inflation. Econometrica 11 (4), 122–150.
Engle, R.F., Ng, V.K., 1993. Measuring and testing the impact of news on volatility. J. Finance 48 (5), 1749–1778.
Fanelli, L., Paruolo, P., 2010. Speed of adjustment in cointegrated systems. J. Econometrics 158, 130–141.
Francq, C., Zakoian, J.-M., 2015. Risk-parameter estimation in volatility models. J. Econometrics 184 (1), 158–173.
Glosten, L., Jagannanthan, R., Runkle, D., 1993. On the relation between expected value and the volatility of the nominal excess return on stocks. J. Finance 48 (5), 1779–1802.
Gradshteyn, I., Ryzhik, I., 2007. Book of Tables of Integrals, Series, and Products, seventh ed. Academic Press, New York.
Harman, R., Lacko, V., 2010. On decompositional algorithms for uniform sampling from n-spheres and n-balls. J. Multivariate Anal. 101 (10), 2297–2304.
Jensen, S.T., Rahbek, A., 2004. Asymptotic inference for nonstationary GARCH. Econom. Theory 20, 1203–1226.
Lau, A., McSharry, P., 2010. Approaches for multi-step density forecasts with application to aggregated wind power. Ann. Appl. Stat. 4 (3), 1311–1341.
Lee, S.-W., Hansen, B.E., 1994. Asymptotic theory for the Garch(1, 1) quasi-maximum likelihood estimator. Econom. Theory 10, 29–52.
Mood, A.M., Graybill, F.A., Boes, D.C., 1974. Introduction to the Theory of Statistics, third ed. Mc Graw-Hill, New York.
Nelson, D., 1990. Stationarity and persistence in the GARCH(1, 1) model. Econom. Theory 6, 318–334.
Tsay, R.S., 2010. Analysis of Financial Time Series, third ed. Wiley, New York.