Stochastic volatility with heterogeneous time scales

To appear in *Quantitative Finance*, Vol. 00, No. 00, Month 20XX, 1–21

Stochastic volatility with heterogeneous time scales

D. DELPINI*†‡ and G. BORMETTI§

† Department of Economics and Business, University of Sassari,
via Muroni 25, 07100 Sassari, Italy
‡ IMT Institue for Advanced Studies, P.za San Ponziano 6, 55100 Lucca, Italy
§ Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy

(*v2.1 released February 2014*)

The heterogeneity of market agents is recognized as a possible driver mechanism for the persistence of financial volatility. In this work we focus on the multiplicity of investment strategies' horizons, we embed this concept in a continuous-time stochastic volatility framework and prove that a parsimonious, two-factor model captures effectively the long memory of volatility as measured from real data. Since estimating parameters in a continuous-time model is challenging, we introduce a robust methodology based on the Generalized Method of Moments. In addition to the volatility clustering, the estimated model also captures other relevant stylized facts, emerging as a minimal but realistic and complete framework for modelling financial time series.

Keywords: Stochastic Processes; Financial Markets; Volatility Modelling; Generalized Method of Moments

JEL Classification: C32, C13, C16

1. Introduction

In (Mandelbrot 1963) Benoît Mandelbrot refers to the volatility clustering as "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes" (see also Mandelbrot 1997). Since then this effect has remained one of the most intriguing properties exhibited by financial time series. In the early Nineties the long memory property, the other face of the volatility clustering, is independently investigated by (Dacorogna *et al.* 1993) and (Ding *et al.* 1993). In the former work, after amending price changes from the heteroscedasticity due to seasonal effects, the authors find a persistent, positive autocorrelation of absolute stock returns which declines hyperbolically with the time lag. In the latter, through the analysis of the daily closing prices of Standard&Poor 500 index for the time span January 3, 1928 - August 30, 1991, Ding and collaborators study the power correlation of absolute returns $|r_t|^d$ for positive d, finding a strong persistence especially for d close to one.

The volatility clustering is usually measured by the persistent correlation between squared returns, or the logarithm of absolute returns, and shows that large variations are more likely to be followed by large than small ones (Dacorogna *et al.* 1993, Guillaume *et al.* 1997, Cont *et al.* 1997, Liu *et al.* 1997, Muzy *et al.* 2000). The slow decay of the volatility can be ascribed to various mechanisms. Formally, persistence might be due to the presence of a genuine long-memory data generating process, as in (Comte and Renault 1998), or to the aggregation of a finite or infinite number of short-memory processes (Granger 1980, Biró and Jakovác 2005, Ding *et al.* 1993, Bollerslev 1986, Barndorff-Nielsen

*Corresponding author. Email: ddelpini@uniss.it

and Shephard 2001, Calvet and Fisher A. J. 2004, McAleer and Medeiros 2008, Banerjee and Urga 2005). But it could also be the result of structural breaks and regime switches (Diebold and Inoue 2001, Gourieroux and Jasiak 2001, Granger and Hyung 2004, Granger and Teräsvirta 1999, Mikosch and Stărică 2004). For a wider survey of contributions on this topic we refer the interested reader to (Bouchaud 2001, Cont 2001, Andersen *et al.* 2009).

When we focus on the behaviour of market agents, Agent Based Models provide a first explanatory framework, where macroscopic evidences are explained in terms of microscopic interactions among market participants. As clarified in the seminal papers (Lux and Marchesi 1999, 2000) the alternation of the economic agents between chartist and fundamentalist regime can be identified as the source of the observed volatility clustering, an empirical signature of persistence. The same mechanism leading to the previous regime switching is further investigated in (Alfi *et al.* 2009a,b), where the authors identify the minimal assumptions required for an agent based model to capture the empirical stylized facts. In a different approach (Müller *et al.* 1994) persistence is induced by the coexistence of agents differing in their perceptions of the market, risk profiles, institutional constraints, degree of information, prior beliefs, and other characteristics such as geographical locations. In (Müller *et al.* 1993) the role of heterogeneous time horizons for the investment strategies is specifically addressed. In (Corsi 2009) the daily, weekly and monthly time scales are isolated as the relevant ones, while the first direct evidence of these three scales as well as an attempt to capture them with an ARCH model is provided in (Lynch and Zumbach 2003). As a major achievement of the latter works we see that a small subset of time scales succeeds in capturing the long run behaviour of the squared return correlation. Interestingly, those horizons reflect typical time scales of the human activity, which noticeably follow a pseudo-geometric progression (Bouchaud 2001). Generalizing the concept of a finite mixture of time scales to a continuum of agents, an attractive intuition is that the integrated effect of exponential heterogeneous strategies may lead to persistence. On a formal basis, this amounts to expressing the correlation function as

$$C(\tau) = \int_0^{1/\tau_{\min}} \exp\left(-\tau/\tau_{\mathrm{agent}}\right) p(1/\tau_{\mathrm{agent}}) \mathrm{d}(1/\tau_{\mathrm{agent}}) \,,$$

which at the leading order for $\tau \to +\infty$ is determined by the behaviour of the density $p(1/\tau_{\mathrm{agent}})$ around the origin. Indeed, by virtue of Watson's Lemma, we obtain $C(\tau) \sim 1/\tau^{1+\alpha}$ provided that $p(1/\tau_{\mathrm{agent}}) \sim \tau_{\mathrm{agent}}^{-\alpha}$ with $\alpha > -1$.

While the absence of serial of correlation is verified empirically with good approximation, the volatility clustering itself provides evidence that financial returns can not just follow a random walk process. The latter would imply independent, identically distributed price increments and any non linear function of the returns would exhibit zero autocorrelation, a property that simply does not hold in practice. Another empirical evidence of this violation is the *leverage effect*, the negative correlation between past returns and future instantaneous volatility. It measures the tendency of market volatility to increase after a price downfall (Christie 1982, Campbell and Hentschel 1992, Glosten *et al.* 1993, Bouchaud *et al.* 2001, Bouchaud and Potters 2003, Perelló *et al.* 2004, Bollerslev *et al.* 2006).

As far as the distributional properties of the volatility proxies are concerned, (Miccichè *et al.* 2002) identifies the inverse gamma distribution as an effective approximation for both the low and high volatility regimes. The simplest model reproducing this distribution as a result of a volatility feedback effect corresponds to the ARCH-like equation reading, in the continuous time limit, as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = -\kappa(\sigma - \sigma_\infty) + \eta\ \sigma\ \zeta(t)\,,$$

where $\zeta(t)$ is the Brownian motion and κ, σ_∞, η are positive constants. For this specific case, the stationary distribution of the volatility has the form of an inverse gamma

$$\frac{\lambda^\nu}{\Gamma(\nu)} \frac{\mathrm{e}^{-\lambda/\sigma}}{\sigma^{1+\nu}} ,$$

with $\nu = 1 + 2\kappa/\eta^2$ and $\lambda = 2\kappa\sigma_\infty/\eta^2$.

In the following section we propose an approach inspired by the latter evidence, as well as by the idea of a mixture of heterogeneous investment horizons, in a spirit similar to the Heston multi-factor model (Bates 2000). Specifically, we introduce a continuous-time stochastic volatility model where the volatility is driven by unobserved stochastic factors with inverse gamma stationary distribution. We limit the number of volatility factors to two, allowing for a clear interpretation in terms of a long-run and a short-run components. Despite this simplification, the model is able to account for the persistent behaviour of the volatility, the return-volatility correlation and the fat tails of return distributions. It achieves a remarkable degree of realism, and still it stays analytically tractable to a large extent. Differently from a multi-factor Heston model with the same number of factors, in our framework we can compute exactly the non-linear correlation functions with the most relevant financial implications. As a direct consequence and a further advantage, the model proves to be easy to estimate from the data series. In particular, exploiting the analytical information, we are able to develop a novel, efficient approach to the estimation of parameters in continuous-time stochastic volatility models which is rooted in the Generalized Method of Moments proposed in (Hansen 1982).

The remainder of the paper is organized as follows. In Section 3 we derive the analytical expressions for the leverage and volatility autocorrelation, while in Section 4 we detail a calibration procedure which is inspired by the Generalized Method of Moments. We conclude in Section 5, and we postpone the analytical derivations in the Appendices at the end of the paper.

2. The model

A quite general expression for the asset price at time t, reminiscent of the Geometric Brownian motion paradigm, is given by

$$S_t = S_0 \, \exp\left(\mu\, t + X_t\right) ,$$

where X_t is the stochastic centred log-return and μ a constant drift coefficient. The time evolution of X_t can be modelled in terms of the stochastic differential equation (SDE)

$$\mathrm{d}X_t = \sigma_t \, \mathrm{d}W_t^X , \tag{1}$$

where σ_t is the instantaneous volatility of the price and $\mathrm{d}W_t^X$ the increment of a standard Wiener process. Since X_0 is equal to zero, we also have $\mathrm{E}\left[X_t\right] = 0$ and $\mathrm{E}\left[\ln S_t - \ln S_0\right] = \mu t$ for any t. A common choice accounting for the stochastic behaviour of the volatility, as measured by suitable proxies, is $\sigma_t = \sigma(Y_t)$ as a function of an unobserved driving process Y_t. General financial considerations regarding the mean-reverting behaviour of the volatility process lead to a second SDE of the form

$$\mathrm{d}Y_t = -\kappa_Y (Y_t - y_\infty)\, \mathrm{d}t + \sqrt{\Sigma(Y_t)}\, \mathrm{d}W_t^Y , \tag{2}$$

with $\kappa_Y = 1/\tau_Y > 0$, and $y_\infty > 0$. In (Delpini and Bormetti 2011) $\Sigma(Y_t)$ is equal to $\sigma_Y^2 Y_t^2$ with $\sigma_Y > 0$, from which it follows that σ_t is proportional to Y_t. This choice leads to an inverse gamma stationary distribution with shape and scale parameters $\nu = 1 + 2\kappa_Y/\sigma_Y^2$ and $\lambda = 2\kappa_Y y_\infty/\sigma_Y^2$, respectively. In light of the considerations presented in the Introduction, our assumption is motivated by the will of recovering the most effective statistical description of the volatility distribution. Different choices for Σ have been suggested in the literature and among the most popular ones it is worth mentioning the Heston (Heston 1993) and Stein-Stein (Stein and Stein 1991) models. For a complete overview of continuous time models as well as widely employed discrete time approaches like ARCH, GARCH and their generalizations, we refer to the handbook of financial time series (Andersen *et al.* 2009).

Following the spirit of the Introduction, in this paper we extend the model given by (1) and (2) allowing the instantaneous volatility to depend on multiple stochastic unobserved factors. Before stating explicitly our model's equations, it is worth noting that such a generalization is inspired by the multi-factor stochastic volatility model introduced by Bates in (Bates 2000) as a possible description of the S&P 500 futures price, and later revisited in (Christoffersen *et al.* 2009) as a model for the dynamics of the volatility smirk in the option pricing context. In the generalized form used in (Corsi and Renò 2012), the multi-factor model with jumps reads

$$\begin{aligned} \mathrm{d}X_t &= \sum_{i=1}^{N} \sqrt{Y_t^i}\,\mathrm{d}W_t^i + \mathrm{d}J_t^X \\ \mathrm{d}Y_t^i &= -\kappa_i \left(Y_t^i - y_\infty^i\right) \mathrm{d}t + \eta_i \sqrt{Y_t^i}\,\mathrm{d}W_t^{i+N} + \mathrm{d}J_t^i \end{aligned} \tag{3}$$

where $i = 1, \ldots, N$, $(W_t^1, \ldots, W_t^{2N})$ is a multivariate possibly correlated Brownian motion and $\{J^X, J^1, \ldots, J^N\}$ is a multivariate, possibly correlated, Poisson process with constant intensities. In principle, each factor may be linked to the sensitivity of the economic agents to different investment horizons, and in light of this heterogeneity the modelling could reflect N volatility components. The model that we consider in this work is a special case of the dynamics (3): We take $N = 2$, discard the jump contribution and assume perfect correlation between the Wiener processes of the two factors, $\mathrm{corr}(\mathrm{dW}_\mathrm{t}^1, \mathrm{dW}_\mathrm{t}^2) = \mathrm{d}t$. However, at variance with equation (3) and as a major contribution of our paper, we consider inverse gamma driving factors, each one being described by the same mean-reverting dynamics provided in (2) with Σ proportional to the squared process. In conclusion, the model we propose reduces to

$$\begin{aligned} \mathrm{d}X_t &= Y_t\,\mathrm{d}W_t^X + Z_t\,\mathrm{d}W_t^X \\ \mathrm{d}Y_t &= -\kappa_Y\left(Y_t - y_\infty\right)\mathrm{d}t + \sigma_Y\,Y_t\,\mathrm{d}W_t^Y \\ \mathrm{d}Z_t &= -\kappa_Z\left(Z_t - z_\infty\right)\mathrm{d}t + \sigma_Z\,Z_t\,\mathrm{d}W_t^Z\,, \end{aligned} \tag{4}$$

where we impose the initial time conditions $X_{t=0} = X_0 = 0$, $Y_{t=t_0} = y_0 > 0$ and $Z_{t=t_0} = z_0 > 0$, with $\kappa_Y = 1/\tau_Y > 0$, and $\kappa_Z = 1/\tau_Z > 0$. We also indicate $\nu_Y = 1 + 2\kappa_Y/\sigma_Y^2$ and $\nu_Z = 1 + 2\kappa_Z/\sigma_Z^2$ the tail exponents of the inverse gamma stationary distributions of Y_t and Z_t. The correlation structure among the three Brownian motions is described by the following matrix

$$\begin{pmatrix} 1 & \rho_{XY} & \rho_{XZ} \\ \rho_{XY} & 1 & \rho_{YZ} \\ \rho_{XZ} & \rho_{YZ} & 1 \end{pmatrix}.$$

It is worth noting that the volatility factors and the return process originate from different initial times. More specifically, as the processes Y_t and Z_t are unobserved factors and we are mainly concerned with their dynamics at the stationary state, we assume they started at $t_0 < 0$ in the past and we recover the stationary limit by letting $t_0 \to -\infty$. On the other hand, X_t represents the observed (detrended) logarithmic increment of the price for a fixed time lag and, therefore, it seems natural to take the spot time $t = 0$ as the initial time of the return process.

Some considerations are due regarding our choice of the factors specification. In (Delpini and Bormetti 2011) the single factor Y_t corresponds (up to a constant) to the instantaneous volatility itself. As such, Y_t has a clear interpretation and we choose its dynamics specifically with the intent to accommodate the distributional properties of the volatility observed in the reality. Here, in the spirit of the factor model (3), the evolution of log-returns is given in terms of two additive factors; it follows that Y_t and Z_t can not be interpreted, separately, as the return volatility (or possibly the variance as in ARCH/GARCH models), but as the underlying unobserved factors. The choice of full correlation between W_t^1 and W_t^2 allows us to introduce formally $\sigma_t = Y_t + Z_t$ motivating the inverse gamma dynamics which drives the factors. As the volatility is now a sum of factors, we will not end up exactly with an asymptotic inverse gamma law for σ_t. Nevertheless, we expect the tail asymptotic to be preserved under suitable assumptions (see discussion at the end of this section) and here we give priority to capturing the observed, long range memory of the squared return correlation.

As far as the leverage is concerned, we know that a negative ρ_{XY} suffices to accommodate the observed short range scaling of the return-volatility correlation. In Appendix B and in the numerical section we set ρ_{XZ} equal to zero to prevent Z_t from impacting the leverage. Nonetheless, in what follows we derive the relation between the factors behaviour and the moments of X_t under the general case of non trivial correlations between the Brownian motions. As we show in Appendix A, the structure of the model (4) allows to compute the moments of the probability density function (PDF) of X_t at all times recursively. After cumbersome calculations, and by exploiting Itô's Lemma to compute the cross correlations between the two volatility factors, it can be verified that the moments of X can be expressed always as a superposition of exponential functions of $(t - t_0)$

$$\mathrm{E}\left[X_t^n\right] = \sum_{i,j=0;\ i+j\leq n}^{n} H_{i,j}^{(n)}(t; y_0, z_0)\, \mathrm{e}^{F_{i,j}(t-t_0)}\,, \tag{5}$$

where the constants read $F_{m,n} = F_m^Y + F_n^Z + mn\rho_{YZ}\sqrt{\sigma_Y^2\sigma_Z^2}$, with $F_m^Y = -\kappa_Y m + m(m-1)\sigma_Y^2/2$, and $F_n^Z = -\kappa_Z n + n(n-1)\sigma_Z^2/2$. Due to the linearity of the ODEs (A2), the coefficients $H_{i,j}^{(n)}$ are polynomially-weighted combinations of exponentials in t. They also depend parametrically on the parameters of the dynamics of the volatility factors, especially the starting values y_0, z_0. We recall that t is the time horizon associated to the price log-return X_t, while t_0 is the initial time of the volatility process.

In the following, we report the explicit expressions of the coefficients $H_{i,j}^{(n)}(t; y_0, z_0)$ for the case $n = 2$ (the constants $k_{i,j}^{(m,n)}$, which depend on the initial conditions y_0, z_0, are defined recursively in Appendix A)

$$\begin{aligned}
H_{0,0}^{(2)} &= \left[k_{0,0}^{(2,0)} + 2k_{0,0}^{(1,1)} + k_{0,0}^{(0,2)}\right] t\,,\\
H_{1,0}^{(2)} &= \left[k_{1,0}^{(2,0)} + 2k_{1,0}^{(1,1)}\right] \frac{1-\mathrm{e}^{-F_{1,0}t}}{F_{1,0}}\,,\\
H_{0,1}^{(2)} &= \left[k_{0,1}^{(2,0)} + 2k_{0,1}^{(1,1)}\right] \frac{1-\mathrm{e}^{-F_{0,1}t}}{F_{0,1}}\,,\\
H_{2,0}^{(2)} &= k_{2,0}^{(2,0)} \frac{1-\mathrm{e}^{-F_{2,0}t}}{F_{2,0}}\,,\\
H_{1,1}^{(2)} &= 2k_{1,1}^{(1,1)} \frac{1-\mathrm{e}^{-F_{1,1}t}}{F_{1,1}}\,,\\
H_{0,2}^{(2)} &= k_{0,2}^{(0,2)} \frac{1-\mathrm{e}^{-F_{0,2}t}}{F_{0,2}}\,.
\end{aligned}$$

Since t is finite, the coefficients $H_{i,j}^{(n)}$ are finite quantities themselves, and all the relevant information about the behaviour of $\mathrm{E}\left[X_t^n\right]$ in the stationary limit of Y and Z is retained by the t_0-exponentials in Eq. (5). Since $F_{0,0} = 0$, $\mathrm{E}\left[X_t^n\right]$ is finite in the stationary limit $t_0 \to -\infty$ provided that all the $F_{i,j}$ for $i, j = 0, \ldots, n$ with $i + j \leq n$ are negative, otherwise it diverges indicating the emergence of fat tails in the PDF $p_t(x)$ of X_t. In the latter case, the tail behaviour would be compatible with an hyperbolic scaling with a tail exponent smaller than the order of the lowest diverging moment.

In (Delpini and Bormetti 2011) the hyperbolic scaling of $p_t(x)$ is induced by the power-law tail of the asymptotic (inverse gamma) distribution of the volatility, and a simple relation exists between the tail exponent of the latter and the order of the first diverging moment of $p_t(x)$. In the present case, the asymptotic distribution of σ_t is that of the sum of the two factors Y and Z, both inverse gamma distributed with tail indices ν_Y and ν_Z respectively. In the limit of Y independent of Z[1], the distribution of the sum behaves as a power-law with tail index $\nu_\sigma = \min\{\nu_Y, \nu_Z\} = \nu$, see e.g. (Bouchaud and Potters 2003, Wilke *et al.* 1998). Therefore, the same mechanism triggering the divergence of the return moments applies here asymptotically and the PDF $p_t(x)$ manifests a decay compatible with a power-law scaling with tail index determined by the value of ν_σ.

The theoretical PDF is also compliant with the more basic properties of the returns. In particular, the stationary limit of (5), and the expression of $H^{(2)}_{0,0}$, show that the variance is linear in the return time lag t. Furthermore, the explicit expressions of the functions $H^{(3)}_{0,0}$ and $H^{(4)}_{0,0}$, not reported here for the sake of parsimony, would also reveal that the skewness and kurtosis vanish in the limit of large t, according to the observed Gaussian-like shape of the distribution for large time horizons.

3. Non linear dependence

In this section we discuss the main properties of the return correlation structure predicted by model (4).

Our model inherits the property of absence of serial correlation, which is supported empirically, from the class of stochastic volatility models. More importantly, it deals with non linear correlation functions explicitly. The expressions of the volatility autocorrelation and the return-volatility correlation functions, whose derivation we postpone to Appendices A - C, provide a valuable analytical characterization of model (4). This information can be exploited for the estimation of the model from empirical data, but it is also crucial to grasp the relevant information about the process time scaling, as we discuss in the rest of this section.

3.1. *Leverage effect*

The leverage, a measure of the correlation between returns and volatility, is usually defined as

$$\mathcal{L}(\tau; t) = \frac{\mathrm{E}\left[\mathrm{d}X_t\, \mathrm{d}X^2_{t+\tau}\right]}{\mathrm{E}\left[\mathrm{d}X^2_t\right]^2} \ .$$

Empirically and for arbitrary t, $\mathcal{L}(\tau; t)$ has been found to be negative and exponentially decaying for positive τ and approximately zero otherwise: A correlation exists between past returns and the volatility in the future and not *vice versa.*

For our case, a finite time, explicit expression is derived in Appendix B. Equation (B2) reveals that the leverage function is characterized by the superposition of three exponential functions, with

[1] In fact, we make this assumption when estimating the model from the empirical data in Section 4.

different characteristic times $\tau_{\mathcal{L}}$, $\tau_{\sim\mathcal{L}}$, $\tau_{<}$, which are ordered according to the following hierarchy

$$\begin{aligned}
\tau_{\mathcal{L}} &= \frac{2}{2-\tau_Y\sigma_Y^2}\tau_Y = \frac{\nu_Y-1}{\nu_Y-2}\tau_Y \, ; \\
\tau_{\sim\mathcal{L}} &= \left(\frac{1}{\tau_Y} - \frac{\sigma_Y^2}{2} + \frac{1}{\tau_Z}\right)^{-1} = \frac{\tau_Z}{\tau_Z+\tau_{\mathcal{L}}}\tau_{\mathcal{L}} < \tau_{\mathcal{L}} \, ; \\
\tau_{<} &= \left[2\left(\frac{1}{\tau_Y} - \frac{\sigma_Y^2}{2}\right) - \frac{\sigma_Y^2}{2}\right]^{-1} = \frac{2}{4-\tau_{\mathcal{L}}\sigma_Y^2}\tau_{\mathcal{L}} = \frac{\nu_Y-1}{2\nu_Y-3}\tau_{\mathcal{L}} \, .
\end{aligned}$$

If $\nu_Y \to 3^+$, $\tau_{\mathcal{L}}$ converges to $2\tau_Y$, while for $\nu_Y \to +\infty$ we have that $\tau_{\mathcal{L}}$ goes to τ_Y. The time scale $\tau_{\sim\mathcal{L}}$ is strictly smaller than $\tau_{\mathcal{L}}$; however, we are implicitly assuming that the characteristic time of the Z process accounts for the volatility persistence, that is $\tau_Z \gg \tau_{\mathcal{L}}$, implying that $\tau_{\sim\mathcal{L}}$ is expected to be only slightly smaller than the leverage scale. Ultimately, if $\nu_Y \to 3^+$, $\tau_{<}$ converges to $2\tau_{\mathcal{L}}/3$, while under the Gaussian limit we have that $\tau_{<}$ converges to $\tau_{\mathcal{L}}/2$. In fact, the three leverage scales are constrained in a narrow range, which empirically has been found to be of order ten days for indexes and slightly larger for single stocks (Bouchaud and Potters 2003).

3.2. *Autocorrelation function of squared increments*

The volatility clustering is commonly measured by the quantity $\mathcal{A}'(\tau;t) = \mathrm{E}\left[\mathrm{d}X_t^2\,\mathrm{d}X_{t+\tau}^2\right]$ and the volatility autocorrelation can be estimated in terms of the following standardized quantity

$$\mathcal{A}(\tau;t) = \frac{\mathcal{A}'(\tau;t) - \mathrm{E}\left[dX_t^2\right]\mathrm{E}\left[dX_{t+\tau}^2\right]}{\sqrt{\mathrm{Var}[dX_t^2]\,\mathrm{Var}[dX_{t+\tau}^2]}} \, . \tag{6}$$

To complete our analytical characterization of the two-factor model (4), in Appendix C we derive the explicit expression of (6). Specifically the expression (C7) features five different exponential scales $\tau_{\mathcal{A}}^{(i=1,\ldots,5)}$. Similarly to the leverage, the characteristic times are organized in a hierarchy as follows

$$\begin{aligned}
\tau_{\mathcal{A}}^{(1)} &= -\frac{1}{F_2^Z} = \frac{\tau_{\mathcal{L}}}{\tau_Y}\tau_Z = \tau_{>Z} > \tau_Z \, , \\
\tau_{\mathcal{A}}^{(2)} &= -\frac{1}{F_1^Z} = \tau_Z \, , \\
\tau_{\mathcal{A}}^{(3)} &= -\frac{1}{F_1^Y} = \tau_Y < \tau_{\mathcal{L}} \, , \\
\tau_{\mathcal{A}}^{(4)} &= -\frac{1}{(F_1^Y + F_1^Z)} = \frac{\tau_Z}{\tau_Z+\tau_Y}\tau_Y < \tau_Y \, , \\
\tau_{\mathcal{A}}^{(5)} &= -\frac{1}{F_2^Y} = \frac{\tau_{\mathcal{L}}}{2} \, .
\end{aligned}$$

For ν_Y varying in $(4, +\infty)$, τ_Y is inferiorly bounded by $2\tau_{\mathcal{L}}/3$, while the upper bound is given by $\tau_{\mathcal{L}}$. Therefore $\tau_{>Z}$ ranges between τ_Z and $3\tau_Z/2$, and we can conclude that the previous five scales indeed cluster into two groups, a long-range and a short-range one. The first group corresponds to $\{\tau_Z, \tau_{>Z}\}$, whose typical scale is given by τ_Z, while the second one contains the three remaining scales, superiorly bounded by $\tau_{\mathcal{L}}$ and of order τ_Y. These relations enlighten the reason why the inclusion of the second factor process Z_t significantly improves the performance of the model. While

Y_t is entirely responsible for the emergence of the leverage and influences the short-run behaviour of σ_t, Z_t provides the auxiliary degree of freedom required to capture the persistence of squared returns. Since Z_t is independent from W_t^X ($\rho_{XZ} = 0$), it does not affect the volatility response to large negative returns but it is exclusively responsible for the volatility behaviour in the long-run.

4. Calibration via Generalized Method of Moments

In this section, we propose and discuss an application of the Generalized Method of Moments to the estimation of the model's parameters from an empirical time series of price increments.

Twelve free parameters, τ_Y, y_∞, y_0, σ_Y^2, τ_Z, z_∞, z_0, σ_Z^2, ρ_{XY}, ρ_{XZ}, ρ_{YZ}, and μ characterize the stochastic model (4). The estimation of parameters in a continuous-time stochastic volatility model is a challenging task primarily due to the latency of the volatility state variable. This issue has inspired many scholars, and there exists a specialised literature on computationally intensive methods mimicking likelihood-based inference. In general, these belong to the class of non linear filtering methods, and among possible approaches we mention Kalman filters, Particle filters, and Monte Carlo Markov Chain approaches. For more techniques and further discussion we refer the reader to the handbook (Andersen *et al.* 2009). Here we take the opportunity to quote the interesting proposal discussed in (Javaheri 2005) and rooted on the spectral approach to nonlinear filtering. A relatively simpler approach to estimation, which does not rely on any ad hoc approximation of the density of returns, is based on the computable moments of the model. For continuous-time stochastic volatility models, it is generally very hard to derive closed form solutions for the return moments, but this is not the case for the model under consideration in this paper. For this reason, we follow a methodology inspired by the Generalized Method of Moments (GMM).

Here we briefly review the general theory of the GMM and refer the reader to (Hamilton 1994), Chapter 14, where an introduction is provided based on Hansen's formulation of the estimation problem (Hansen 1982).

Given T observations $\{\boldsymbol{W}_t\}$ for $t = 1, \ldots, T$, each one being an h dimensional vector, and a vector $\theta \in \mathrm{R}^{\mathbf{k}}$ of unknown parameters (from a parameter space $\boldsymbol{\Theta}$), in order to apply GMM there should be a function $\mathbf{h}(\theta, \boldsymbol{W}_\mathbf{t}) : \mathrm{R}^{\mathbf{k}} \times \mathrm{R}^{\mathbf{h}} \to \mathrm{R}^{\mathbf{r}}$ characterized by the property that

$$\mathrm{E}\left[\mathbf{h}(\theta, \boldsymbol{W}_\mathbf{t})\right] = \mathbf{0}\,. \tag{7}$$

These r equalities are usually described as orthogonality conditions. The basic idea of GMM is to replace these conditions with sample averages and to solve the following optimization problem

$$\hat{\theta} = \underset{\theta\in\boldsymbol{\Theta}}{\operatorname{argmin}} \left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{h}(\theta, \boldsymbol{W}_\mathbf{t})\right)^{\mathrm{t}} \hat{\boldsymbol{\Omega}}_T^{-1} \left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{h}(\theta, \boldsymbol{W}_\mathbf{t})\right) ,$$

where $\hat{\boldsymbol{\Omega}}_T$ is a positive-definite weighting matrix depending on the available data set and on the value of θ itself. The practical procedure is the one which follows. An initial estimate $\hat{\theta}^{(0)}$ is obtained by minimizing the previous quantity with an arbitrary choice of $\hat{\boldsymbol{\Omega}}_T$, e.g. $\hat{\boldsymbol{\Omega}}_T = \mathrm{I}_{r\times r}$. Supposing that $\mathbf{h}(\theta, \boldsymbol{W}_\mathbf{t})$ is serially uncorrelated, the estimate $\hat{\theta}^{(0)}$ is then used in

$$\hat{\boldsymbol{\Omega}}_T = \frac{1}{T}\sum_{t=1}^{T}\mathbf{h}(\hat{\theta}^{(0)}, \boldsymbol{W}_t)\ \mathbf{h}^{\mathrm{t}}(\hat{\theta}^{(0)}, \boldsymbol{W}_t)$$

to arrive to a new GMM estimate $\hat{\theta}^{(1)}$. This process can be iterated until an arbitrary stopping criterion is invoked [1]. Let $\bar{\theta}$ be the true value of θ, and define the $k \times r$ matrix

$$\boldsymbol{G}(\theta, \boldsymbol{W}_{\mathbf{t}}) = \frac{\mathbf{1}}{\mathbf{T}} \frac{\partial}{\partial \theta} \sum_{\mathbf{t=1}}^{\mathbf{T}} \mathbf{h}^{\mathrm{t}}(\theta, \boldsymbol{W}_{\mathbf{t}}) \ ;$$

then the theory behind the GMM states that $\hat{\theta}^{(1)}$ is approximately distributed as Normal$(\bar{\theta}, \hat{\boldsymbol{V}}_T/T)$ with $\hat{\boldsymbol{V}}_T/T$ given by

$$\hat{\boldsymbol{V}}_T = \left\{ \boldsymbol{G}(\hat{\theta}^1, \boldsymbol{W}_t) \ \hat{\boldsymbol{\Omega}}_T^{-1} \boldsymbol{G}^t(\hat{\theta}^1, \boldsymbol{W}_t) \right\}^{-1} \ .$$

For the case under consideration in this paper, we need to estimate the following vector of parameters

$$\theta^{\mathrm{t}} = \left(\mu, \tau_{\mathbf{Y}}, \mathbf{y}_{\infty}, \mathbf{y_0}, \sigma_{\mathbf{Y}}^{\mathbf{2}}, \tau_{\mathbf{Z}}, \mathbf{z}_{\infty}, \mathbf{z_0}, \sigma_{\mathbf{Z}}^{\mathbf{2}}, \rho_{\mathbf{XY}}, \rho_{\mathbf{XZ}}, \rho_{\mathbf{YZ}} \right) \ .$$

The orthogonality conditions can be obtained computing the lowest order moments of returns, the leverage correlation, and the squared return autocorrelation. More precisely, we require the following expectations to be zero

$$\mathrm{E} \begin{bmatrix} \Delta X_t \\ |\Delta X_t| - \sqrt{\frac{2\Delta t}{\pi}} \sum_{l=0}^{1} \binom{1}{l} \sum_{i=0}^{1-l} \sum_{j=0}^{l} k_{i,j}^{(1-l,l)} e^{F_{i,j}(t-t_0)} \\ (\Delta X_t)^2 - \Delta t \sum_{l=0}^{2} \binom{2}{l} \sum_{i=0}^{2-l} \sum_{j=0}^{l} k_{i,j}^{(2-l,l)} e^{F_{i,j}(t-t_0)} \\ |\Delta X_t|^3 - \sqrt{\frac{8\Delta t^3}{\pi}} \sum_{l=0}^{3} \binom{3}{l} \sum_{i=0}^{3-l} \sum_{j=0}^{l} k_{i,j}^{(3-l,l)} e^{F_{i,j}(t-t_0)} \\ \Delta X_t \Delta X_{t+\Delta t}^2 - \Delta t^2 \left(C_{2,0} + 2C_{1,1} + C_{0,2} \right)^2 \mathcal{L}(L'\Delta t; t) \\ \vdots \\ \Delta X_t \Delta X_{t+L\Delta t}^2 - \Delta t^2 \left(C_{2,0} + 2C_{1,1} + C_{0,2} \right)^2 \mathcal{L}(L''\Delta t; t) \\ \Delta X_t^2 \Delta X_{t+K'\Delta t}^2 - \mathcal{A}'(K'\Delta t; t) \\ \vdots \\ \Delta X_t^2 \Delta X_{t+K''\Delta t}^2 - \mathcal{A}'(K''\Delta t; t) \end{bmatrix} = \mathbf{0} \, ,$$

where $\Delta X_t = \ln S_{t+\Delta t} - \ln S_t - \mu\Delta t$, $\Delta t = 1/250$ yr and the coefficients $C_{m,n}$ (depending on $t - t_0$) are defined recursively by Eq. (A3) in Appendix A. In the previous equation, the conditions referring to the return-volatility and the squared return correlations depend on the four positive integers $L' < L''$, $K' < K''$. The choice of these values can be made based on prior analysis of the time scales of the correlation functions, and we detail them later. Thus, the dimension r of the vector $\mathbf{h}(\theta, \boldsymbol{W}_{\mathbf{t}})$ reduces to $4 + L'' - L' + 1 + K'' - K' + 1$. From an econometric viewpoint the problem of the parameter estimation is now cast into a sound statistical framework. By means of GMM we can obtain an estimate of central values and associated statistical uncertainty for all the unknowns of the problem. However, the quantity to be optimized is highly non linear, the optimization procedure of the twelve dimensional problem is problematic *per se*, and finding a solution under unsupervised

[1]When the process $\mathbf{h}(\theta, \boldsymbol{W}_{\mathbf{t}})$ for $t = 1, \ldots, T$ is serially correlated, the Newey-West estimate for $\hat{\boldsymbol{\Omega}}_T$ can be used, please refer to equation 14.1.19 in (Hamilton 1994) for further details.

search can be extremely demanding. For this reason and in order to reduce the number of parameters to be estimated, we invoke some reasonable assumptions concerning the volatility process.

The starting point of our heuristic is the role played by the parameter t_0. Even though it could be treated as an unknown parameter, its role is peculiar and different from the others. It mainly determines the regime of the factors processes and we assume $t_0 \to -\infty$. Said differently, we assume that the data we are observing reflect stationary realizations of Y_t and Z_t. Under this regime, both mean-reverting processes do not depend on the initial time values y_0 and z_0, and we identify y_∞ with y_0, and z_∞ with z_0. Since both Y_t and Z_t are unobserved processes reflecting the presence in the market of investment strategies with heterogeneous time horizons, it is plausible to assume that the Brownian motions driving those processes are uncorrelated. Even though this assumption can be relaxed, fixing $\rho_{YZ} = 0$ the problem greatly simplifies, all $F_{m,n}$ reduce to $F^Y_m + F^Z_n$, and all terms $C_{m,n}$ split into $C_{m,0} \times C_{0,n}$. Finally, since we expect that a non-zero negative ρ_{XY} should suffices in reproducing the leverage effect, and in order to avoid that the extra factor Z_t perturbates the leverage structure, we fix ρ_{XZ} equal to zero.

From the considerations after Eq. (5) in Section 2, we know the tail exponent of the distribution of the volatility factors is responsible for the divergence of the moments of X_t. Even though Y_t and Z_t were characterized by two different tail exponents, the order of the first divergent moment of X_t would be determined by the lowest of them. Without loss of generality, we then assume that the stationary distributions of Y_t and Z_t have the same shape parameter $\nu = \nu_Y = \nu_Z$. Now the reduced vector of parameters reads $\theta^{\mathrm{t}} = (\mu, \mathbf{y}_\infty, \mathbf{z}_\infty, \tau_{\mathbf{Y}}, \tau_{\mathbf{Z}}, \rho_{\mathbf{XY}}, \nu)$, while the orthogonality relations simplify. For instance, the first four relations reduce to

$$\begin{aligned}
&\mathrm{E}\left[\Delta X_t\right] = 0\,, \\
&\mathrm{E}\left[|\Delta X_t| - \sqrt{\frac{2\Delta t}{\pi}}(y_\infty + z_\infty)\right] = 0\,, \\
&\mathrm{E}\left[(\Delta X_t)^2 - (y_\infty + z_\infty)^2 \Delta t + \frac{y_\infty^2 + z_\infty^2}{\nu - 2}\Delta t\right] = 0\,, \\
&\mathrm{E}\left[|\Delta X|^3 - \sqrt{\frac{8\Delta t^3}{\pi}}\frac{(\nu-1)^2}{(\nu-3)(\nu-2)}(y_\infty^3 + z_\infty^3) - 3\sqrt{\frac{8\Delta t^3}{\pi}}\frac{\nu-1}{\nu-2}(y_\infty + z_\infty)y_\infty z_\infty\right] = 0\,.
\end{aligned} \tag{8}$$

The numerator of the leverage for positive τ becomes

$$\begin{aligned}
&\rho_{XY}\sqrt{\frac{8}{\tau_Y(\nu-1)}}\mathrm{e}^{-\frac{\tau}{\tau_{\mathcal{L}}}} \times \left\{ C^{\mathrm{st}}_{0,1}\left[C^{\mathrm{st}}_{2,0} - \frac{\nu-1}{\nu-3}y_\infty C^{\mathrm{st}}_{1,0}\right]\mathrm{e}^{-\left(\frac{\nu-2}{\nu-1}\right)\frac{\tau}{\tau_{\mathcal{L}}}} \right. \\
&\left. + \left[C^{\mathrm{st}}_{2,0}\left(C^{\mathrm{st}}_{0,1} - z_\infty\right) + C^{\mathrm{st}}_{1,0}\left(C^{\mathrm{st}}_{0,2} - C^{\mathrm{st}}_{0,1}z_\infty\right)\right]\mathrm{e}^{-\frac{\tau}{\tau_Z}} + \left[\left(\frac{\nu-1}{\nu-3}y_\infty + z_\infty\right)\left(C^{\mathrm{st}}_{2,0} + C^{\mathrm{st}}_{1,0}C^{\mathrm{st}}_{0,1}\right)\right]\right\}\,,
\end{aligned}$$

where the superscript $^{\mathrm{st}}$ stands for the stationary regime corresponding to $t_0 \to -\infty$. Even though the leverage introduces a superposition of three exponential functions, we have seen at the end of the Section 3.1 that the characteristic exponents have the same magnitude and are all dominated by $\tau_{\mathcal{L}}$. For this reason, and recalling that the typical decay time for the leverage is smaller than one hundred days, we set $L' = 1$, and $L'' = 250$ (roughly one trading year) which greatly enhances the numerical convergence. We first solve the optimization problem for the 7 parameters using the first $4 + L'' - L' + 1 = 254$ orthogonal relations (that is, in this step we leave out the orthogonality conditions involving the volatility autocorrelation). In this way, we get a first estimate of the leverage characteristic scale $\hat{\tau}_{\mathcal{L}} \approx 25$ days. Then we perform the final optimization on the entire set of $254 + K'' - K' + 1$ orthogonal relations. In order to include the relationships involving the volatility persistence we should fix $K' = 1$ and K'' equal to the maximum lag compatible with the length

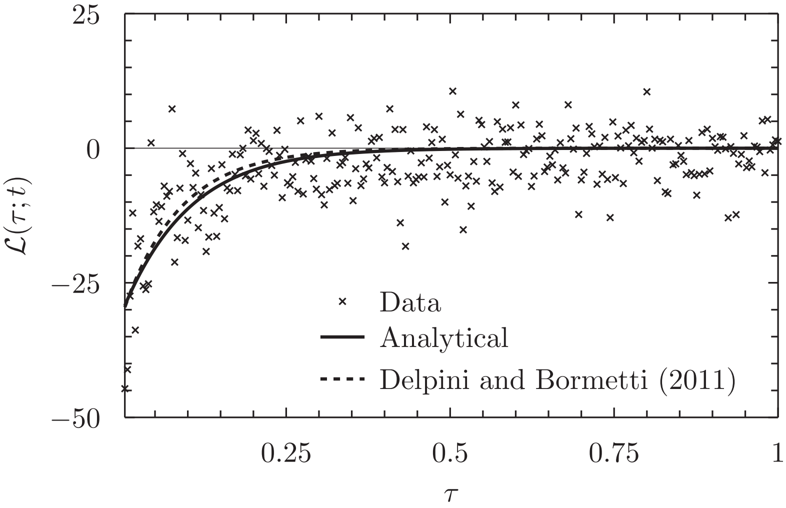

Figure 1. Analytical description of the empirical leverage correlation with values of parameters estimated by GMM with $K' = \lfloor 2\hat{\tau}_{\mathcal{L}} \rfloor$.

Table 1. Estimated values of the parameters ($K' = \lfloor 2\hat{\tau}_{\mathcal{L}} \rfloor$) from daily returns of the S&P500 index 1970-2010.

	$\hat{\theta}^{(1)}$	$\hat{\boldsymbol{\sigma}}_T$	$\hat{\boldsymbol{\rho}}_T$						
μ	2.1×10^{-4}	6×10^{-5}	1.00	−0.01	0.02	−0.28	−0.01	−0.01	−0.01
y_∞	0.095	0.004	−0.01	1.00	−0.97	−0.04	−0.14	0.00	0.99
z_∞	0.052	0.004	0.02	−0.97	1.00	0.03	0.25	0.00	−0.94
τ_Y	0.07 yr	0.01 yr	−0.28	−0.04	0.03	1.00	0.05	0.01	−0.05
τ_Z	0.40 yr	0.02 yr	−0.01	−0.14	0.25	0.05	1.00	0.00	−0.12
ρ_{XY}	−0.77	0.09	−0.01	0.00	0.00	0.01	0.00	1.00	0.00
ν	4.15	0.01	−0.01	0.99	−0.94	−0.05	−0.12	0.00	1.00

of the data series. However, from numerical experiments we found that the fit improves when K' increases and there exists a trade-off between the ability of the model to capture the short-run and the long-run behaviour of the volatility autocorrelation. As a consequence of the hierarchy of time scales detailed in Section 3.2, the leverage determines the scaling of $\mathcal{A}$ for short time lags. It follows that when $K' = 1$, we fit the behaviour of the whole curve, but this produces a sizeable distortion in the central region of the curve. Therefore, since we aim at reproducing the long-run persistence of the volatility, we fix $K' = \lfloor 2\hat{\tau}_{\mathcal{L}} \rfloor = 50$ [1]. As far as the value of K'' is concerned, since we know empirically that the correlation vanishes after a few hundred days, we fix $K'' = L'' = 250$. In Figure 2 we show the result of the fit for $\mathcal{A}$ with both values of K'.

In this setup, by an optimization with 7 unknown parameters and 454 orthogonal relations, we obtain the GMM estimates $\hat{\theta}^{(1)}$ and $\hat{\boldsymbol{V}}_T/T$. Then, in order to compute a consistent estimate of σ_Y^2 and the associated confidence level, we draw a random sample from Normal$(\hat{\theta}^{(1)}, \hat{\boldsymbol{V}}_T/T)$ and obtain a statistics of σ_Y^2 through the relation $2/(\tau_Y(\nu-1))$. We proceed in an analogous way for σ_Z^2, and for $\tau_{\mathcal{L}} = \tau_Y(\nu-1)/(\nu-2)$. The time series on which we perform the analysis consists of a data set from the Standard & Poor's 500 index daily returns from 1970 to 2010. This allows to evaluate the ability of the extended model to capture the persistence of the volatility, not only in absolute terms but also in comparison with the previous estimate from a simpler model. In Tab. 1 we report the central values $\hat{\theta}^{(1)}$, the standard errors $\hat{\boldsymbol{\sigma}}_T = \sqrt{\mathrm{diag}(\hat{\mathbf{V}}_T/T)}$, and the correlation structure $\hat{\boldsymbol{\rho}}_T = \hat{\mathbf{V}}_T/(T\hat{\boldsymbol{\sigma}}_T\hat{\boldsymbol{\sigma}}_T^{\mathrm{t}})$ for all the parameters. As far as the other relevant parameters of the model are concerned, we have $\sigma_Y^2 = 9.9 \pm 1.9$, $\sigma_Z^2 = 1.58 \pm 0.07$, and $\tau_{\mathcal{L}} = 0.10 \pm 0.02$ yr. The new values confirm

[1] In the following we adopt standard mathematical notation $\lfloor\cdot\rfloor$ for the integer part function.

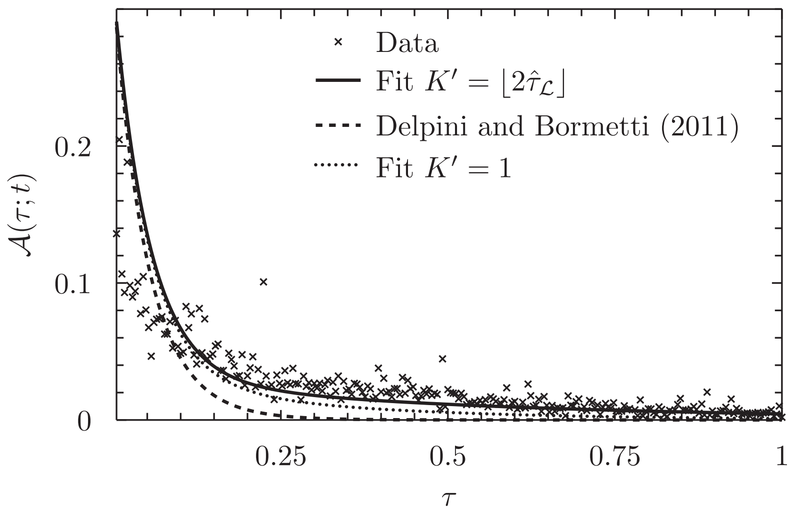

Figure 2. Empirical volatility autocorrelation function of the daily returns of the S&P500 index 1970-2010 (data points), and analytical descriptions: Bold and dotted lines, new expressions with GMM estimates for different values of K'; dashed line, formula and values of parameters as in (Delpini and Bormetti 2011).

the goodness of the estimate provided in (Delpini and Bormetti 2011), in particular the value of ρ_{XY} is strictly negative and the level of the tail exponent ν predicts the divergence of moments higher than the fourth one. It is worth to comment explicitly the relationship between the shortest and longest time scales which characterize our process. Indeed, while the shortest time scale corresponds to the typical relaxation time of Y_t, which is equal to 0.07 ± 0.01 yr and is therefore dominated by the leverage time scale 0.10 ± 0.02 yr (to be compared with the old estimate $\tau_{\mathcal{L}} = 0.09$ yr), the new time scale τ_Z for the process Z_t is larger by a factor of six and fully determines the long-run volatility behaviour.

Since the number of orthogonality conditions is much higher than the number of parameters to be estimated, in order to assess the reliability of our results we perform a standard J-test (Sargan 1958, Hansen 1982) for over-identifying restrictions. We obtain a value for the J-statistics $J = 159.18$, to be compared with the critical value 497.29 for a χ^2_{447} at 95% significance level, and, then, we conclude that the null hypothesis that the moment conditions match the data well can not be rejected.

In Figures 1 and 2 we plot the leverage function and the normalized autocorrelation of squared returns. The exponential decay of the leverage is described correctly by the analytical formula, and no relevant differences are noticeable with respect to the description obtained via the model introduced in (Delpini and Bormetti 2011). Different considerations apply to the persistence of the volatility as predicted by the extended model. The "slow" volatility factor Z_t introduces a longer time scale and allows to capture the long range memory of the autocorrelation function. This is clear from the comparison between the dashed line for the old model, and the bold one corresponding to model (4). Our results demonstrate the ability of a multi-factor approach to stochastic volatility to effectively describe several phenomena. We conclude that a simple two factor model with inverse gamma distributed factors is able to capture the emergence of multiple time scales for the volatility autocorrelation, as well as the exponential decay of the return-volatility correlation. The measured value for the tail parameter ν is coherent with the internal consistency of the model requiring ν to be greater than four (in order for Eq. (C7) to converge in the stationary limit). In particular, $\nu = 4.15$ predicts an hyperbolic decay of the daily return distribution, which captures correctly the non Gaussian probability of extreme events in the real data.

5. Conclusions

In this work, the model for the statistical description of financial stylized facts proposed in (Delpini and Bormetti 2011) is amended from the unrealistically fast decay of the volatility autocorrelation. This is achieved introducing an extra stochastic factor that drives the volatility. The intuition behind this generalization traces back to the early empirical analysis of the FX market in (Müller *et al.* 1994) and the model in (Müller *et al.* 1993), where the role played by heterogeneous investors is strongly emphasized. Evidences from these papers are rooted in the econometric analysis of publicly available financial time series, but a convincing micro-founded model is still lacking. Access to electronic order book data and to agents' identifiers would allow to estimate the individual components of this heterogeneity. An even approximate estimation of the distribution of typical investment horizons from this information would provide a valuable trader-based foundation.

With respect to previous approaches and analyses of continuous-time stochastic volatility models, we believe that the estimation procedure proposed here represents a further improvement and fulfils the desirable requirements of statistical soundness. At the same time, it also allows to focus on those facts which are established as relevant for the description of financial data. In particular, we pursue an heuristic approach to optimization that reduces the dimensionality of the parameters space and retains only the ingredients needed to capture the aforementioned empirical evidences.

In our case, the emergence of power-law tails in the return distribution complies with past empirical analysis (Mantegna and Stanley 2000), but also poses serious limitations to the usage of the model in the context of option pricing. This is certainly true for vanilla instruments, whose payoff grows exponentially with the log-price. On the other hand, our framework could deliver, in principle, better estimates of the role played by rare events for market risk evaluation.

On the whole, we believe our model provides valuable insight about the volatility clustering phenomenon. It also achieves a remarkable degree of realism, higher than previous attempts in continuous-time stochastic volatility modelling, yet allowing for the derivations of relevant analytic relations, i.e. the moments of the return probability distribution and non linear correlation functions with profound financial implications, whose knowledge greatly enhances the estimation methodology inspired by the Generalized Method of Moments.

Acknowledgements

We warmly thank Fulvio Corsi for many inspiring discussions and the two anonymous referees for numerous comments and suggestions.

Funding

We both acknowledge support of the Scuola Normale Superiore Grant 'Giovani Ricercatori 2011'.

References

Alfi, V., Cristelli, M., Pietronero, L. and Zaccaria, A., Minimal agent based model for financial markets I: Origin and self-organization of stylized facts. *Eur. Phys. J. B*, 2009a, **67**, 385–397.

Alfi, V., Cristelli, M., Pietronero, L. and Zaccaria, A., Minimal agent based model for financial markets II: Statistical properties of the linear and multiplicative dynamics. *Eur. Phys. J. B*, 2009b, **67**, 399–417.

Andersen, T.G., Davis, R.A., Kreiss, J.P. and Mikosch (eds.), T., *Handbook of Financial Time Series*, 2009 (Springer-Verlag: Berlin).

Banerjee, A. and Urga, G., Modelling structural breaks, long memory and stock market volatility: an overview. *J. Econom.*, 2005, **129**, 1–34.

Barndorff-Nielsen, O.E. and Shephard, N., Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. B*, 2001, **63**, 167–241.

Bates, D., Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, 2000, **94**, 181–238.

Biró, T.S. and Jakovác, A., Power-Law Tails from Multiplicative Noise. *Phys. Rev. Lett.*, 2005, **94**, 132302.

Bollerslev, T., Litvinova, J. and Tauchen, G., Leverage and Volatility Feedback Effects in High-Frequency Data. *J. Financ. Econom.*, 2006, **4**, 353–384.

Bollerslev, T., Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 1986, **31**, 307 – 327.

Bouchaud, J.P., Power laws in economics and finance: some ideas from physics. *Quant. Finance*, 2001, **1**, 105–112.

Bouchaud, J.P., Matacz, A. and Potters, M., Leverage Effect in Financial Markets: The Retarded Volatility Model. *Phys. Rev. Lett.*, 2001, **87**, 228701.

Bouchaud, J.P. and Potters, M., *Theory of Financial Risk and Derivative Pricing: from Statistical Physics to Risk Management*, 2003, Cambridge University Press.

Calvet, L.E. and Fisher A. J., How to Forecast Long-Run Volatility: Regime Switching and the Estimation of Multifractal Processes. *J. Financ. Econom.*, 2004, **2**, 49–83.

Campbell, J.Y. and Hentschel, L., No news is good news. *J. financ. econ.*, 1992, **31**, 281–318.

Christie, A., The stochastic behavior of common stock variances Value, leverage and interest rate effects. *J. financ. econ.*, 1982, **10**, 407–432.

Christoffersen, P., Heston, S. and Jacobs, K., The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science*, 2009, **55**, 1914–1932.

Comte, F. and Renault, E., Long memory in continuous-time stochastic volatility models. *Math. Financ.*, 1998, **8**, 291–323.

Cont, R., Empirical properties of asset returns: stylized facts and statistical issues. *Quant. Finance*, 2001, **1**, 223–236.

Cont, R., Potters, M. and Bouchaud, J.P., Scaling in stock market data: stable laws and beyond. In *Scale Invariance and Beyond*, edited by B. Dubrulle, F. Graner and D. Sornette, 1997, Les Ulis: EDP Sciences.

Corsi, F., A simple approximate long memory model of realized volatility. *J. Financ. Economet.*, 2009, **7**, 174–196.

Corsi, F. and Renò, R., Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *J. Bus. Econ. Stat.*, 2012, **3**, 368–380.

Dacorogna, M.M., Müller, U.A., Nagler, R.J., Olsen, R.B. and Pictet, O.V., A Geographical Model for the Daily and Weekly Seasonal Volatility in the Foreign Exchange Market. *J. Int. Money Finance*, 1993, **12**, 413–438.

Delpini, D. and Bormetti, G., Minimal model of financial stylized facts. *Phys. Rev. E*, 2011, **83**, 041111.

Diebold, F.X. and Inoue, A., Long memory and regime switching. *J. Econom.*, 2001, **105**, 131–159.

Ding, Z., Granger, C.W.J. and Engle, R.F., A long memory property of stock market returns and a new model. *J. Empirical Finance*, 1993, **1**, 83–106.

Glosten, L.R., Jagannathan, R. and Runkle, D.E., On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *J. Finance*, 1993, **48**, 1779–1801.

Gourieroux, C. and Jasiak, J., Memory and infrequent breaks. *Econ. Lett.*, 2001, **70**, 29–41.

Granger, C.W.J. and Hyung, N., Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. *J. Empir. Financ.*, 2004, **11**, 399–421.

Granger, C.W.J. and Teräsvirta, T., A simple nonlinear time series model with misleading linear properties. *Econ. Lett.*, 1999, **62**, 161–165.

Granger, C., Long memory relationships and the aggregation of dynamic models. *J. Econom.*, 1980, **14**, 227–238.

Guillaume, D.M., Dacorogna, M.M., Davé, R.D., Müller, U.A., Olsen, R.B. and Pictet, O.V., From the Bird's Eye to the Microscope: a Survey of New Stylized facts of the Intra-daily Foreign Exchange Markets. *Finance and Stochastics*, 1997, **1**, 95–129.

Hamilton, J.D., *Time Series Analysis*, 1994, Princeton University Press.

Hansen, L.P., Large Sample Properties of Generalized Methods of Moments Estimators. *Econometrica*, 1982, **50**, 1029–1054.

Heston, S.L., A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.*, 1993, **6**, 327–343.

Javaheri, A., Volatility Estimation via Chaos Expansions. *Wilmott Magazine*, 2005, **May/June**, 78–81.

Liu, Y., Cizeau, P., Meyer, M., Peng, C.K. and Stanley, H.E., Correlations in economic time series. *Physica A*, 1997, **245**, 437–440.

Lux, T. and Marchesi, M., Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 1999, **397**, 498–500.

Lux, T. and Marchesi, M., Volatility clustering in financial markets: A micro-simulation of interacting agents. *Int. J. Theoretical Appl. Finance*, 2000, **3**, 675–702.

Lynch, P.E. and Zumbach, G.O., Market heterogeneities and causal structure of volatility. *Quant. Finance*, 2003, **3**, 320–331.

Mandelbrot, B.B., The Variation of Certain Speculative Prices. *J. Bus.*, 1963, **36**, 394–419.

Mandelbrot, B.B., *Fractals and Scaling in Finance*, 1997, Springer-Verlag.

Mantegna, R.N. and Stanley, H.E., *An Introduction to Econophysics: Correlations and Complexity in Finance*, 2000 (Cambridge University Press: Cambridge).

McAleer, M. and Medeiros, M.C., A multiple regime smooth transition Heterogeneous Autoregressive model for long memory and asymmetries. *J. Econom.*, 2008, **147**, 104–119.

Miccichè, S., Bonanno, G., Lillo, F. and Mantegna, R.N., Volatility in financial markets: stochastic models and empirical results. *Physica A*, 2002, **314**, 756–761.

Mikosch, T. and Stărică, C., Nonstationarities in Financial Time Series, the Long-Range Dependence, and the IGARCH Effects. *Rev. Econ. Stat.*, 2004, **86**, 378–390.

Müller, U.A., Dacorogna, M.M., Davé, R.D., Olsen, R.B., Pictet, O.V. and von Weizsäcker, J.E., Volatilities of different time resolutions - analyzing the dynamics of market components. *J. Empirical Finance*, 1993, **4**, 213–239.

Müller, U.A., Dacorogna, M.M., Davé, R.D., Olsen, R.B., Pictet, O.V. and Ward, J.R., Fractals and intrinsic time - a challenge to econometricians. In *Proceedings of the B. Lüthje*, edited by E. Zinsprognose, Vol. 4, pp. 213–239, 1994 (Verband öffentlicher Banken, Bonn).

Muzy, J.F., Delour, J. and Bacry, E., Modelling fluctuations of financial time series: from cascade process to stochastic volatility model. *Eur. Phys. J. B*, 2000, **17**, 537–548.

Novikov, E.A., Functionals and the random-force method in turbulence theory. *Sov. Phys. JETP-USSR*, 1965, **20**, 1290–1294.

Perelló, J. and Masoliver, J., Random diffusion and leverage effect in financial markets. *Phys. Rev. E*, 2003, **67**, 037102.

Perelló, J., Masoliver, J. and Bouchaud, J.P., Multiple time scales in volatility and leverage correlations: a stochastic volatility model. *Appl. Math. Finance*, 2004, **11**, 27–50.

Sargan, J., The Estimation of Economic Relationships using Instrumental Variables. *Econometrica*, 1958, **26**, 393–415.

Stein, E.M. and Stein, J.C., Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *Rev. Financ. Stud.*, 1991, **4**, 727–52.

Wilke, C., Altmeyer, S. and Martinetz, T., Large-scale evolution and extinction in a hierarchically structured environment. In *Proceedings of the Artificial Life VI, Proceedings of the Sixth International Conference on Artificial Life*, edited by C. Adami, R.K. Belew, H. Kitano and T.C. E., pp. 266–272, 1998.

Appendix A: Moments computation

Given model (4), we first report the definitions of some recurring quantities entering the analytical manipulations presented in this Appendix

$$\begin{aligned} A^Y_m &= m\kappa_Y y_\infty \,, \\ A^Z_n &= n\kappa_Z z_\infty \,, \\ F^Y_m &= -\kappa_Y m + m(m-1)\sigma^2_Y/2 \,, \\ F^Z_n &= -\kappa_Z n + n(n-1)\sigma^2_Z/2 \,, \\ F_{m,n} &= F^Y_m + F^Z_n + mn\rho_{YZ}\sqrt{\sigma^2_Y\sigma^2_Z} \,. \end{aligned}$$

Starting from Eq. (1), application of Itô's Lemma to the function X^l_t readily provides

$$\mathrm{E}\left[X^l_t\right] = \frac{l}{2}(l-1)\int_0^t \mathrm{E}\left[X^{l-2}_s\left(Y_s+Z_s\right)^2\right]\mathrm{d}s \,, \tag{A1}$$

and the same Lemma proves that the correlation functions between integer powers of X_t, Y_t, and Z_t satisfy the following differential equation

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}\left[X^l_tY^m_tZ^n_t\right] &= F_{m,n}\,\mathrm{E}\left[X^l_tY^m_tZ^n_t\right] + A^Y_m\,\mathrm{E}\left[X^l_tY^{m-1}_tZ^n_t\right] + A^Z_n\,\mathrm{E}\left[X^l_tY^m_tZ^{n-1}_t\right] \\ &+ \frac{l}{2}(l-1)\,\mathrm{E}\left[X^{l-2}_tY^m_tZ^n_t\left(Y_t+Z_t\right)^2\right] \\ &+ l\left(m\,\rho_{XY}\sigma_Y + n\,\rho_{XZ}\sigma_Z\right)\,\mathrm{E}\left[X^{l-1}_tY^m_tZ^n_t\left(Y_t+Z_t\right)\right] \,. \end{aligned} \tag{A2}$$

Previous equations correspond to a system of nested linear ordinary differential equations (ODE), which can be solved recursively starting from the lowest order of l, m, and n, and whose solution involves integration of the two point correlations $C_{m,n}(t;t_0) = \mathrm{E}\left[Y^m_tZ^n_t\right]$ [1]. From application of Itô's Lemma we obtain

$$\begin{aligned} \mathrm{d}\left(Y^m Z^n\right) &= \left[F^Y_mY^mZ^n + F^Z_nY^mZ^n\right]\mathrm{d}t + \left[A^Y_mY^{m-1}Z^n + A^Z_nY^mZ^{n-1}\right]\mathrm{d}t \\ &+ \rho_{YZ}mn\sqrt{\sigma^2_Y\sigma^2_Z}Y^mZ^n\,\mathrm{d}t + m\sqrt{\sigma^2_Y}Y^mZ^n\,\mathrm{d}W^Y_t + n\sqrt{\sigma^2_Z}Y^mZ^n\mathrm{d}W^Z_t \,; \end{aligned}$$

then, taking expectation and differentiating *w.r.t* time we derive the following ODE

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}\left[Y^m_tZ^n_t\right] &= A^Y_m\mathrm{E}\left[Y^{m-1}_tZ^n_t\right] + A^Z_n\mathrm{E}\left[Y^m_tZ^{n-1}_t\right] \\ &+ \left(F^Y_m + F^Z_n + \rho_{YZ}mn\sqrt{\sigma^2_Y\sigma^2_Z}\right)\mathrm{E}\left[Y^m_tZ^n_t\right] \,. \end{aligned}$$

For instance, for the case $m = n = 1$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}\left[Y_tZ_t\right] = A^Y_1\mathrm{E}\left[Z_t\right] + A^Z_1\mathrm{E}\left[Y_t\right] + \left(F^Y_1 + F^Z_1 + \rho_{YZ}\sqrt{\sigma^2_Y\sigma^2_Z}\right)\mathrm{E}\left[Y_tZ_t\right] \,,$$

[1] In the following we drop the dependence on t and t_0.

where the mean values read

$$\mathrm{E}\left[Y_t\right] = -\frac{A_1^Y}{F_1^Y} + e^{F_1^Y(t-t_0)}\left[y_0 + \frac{A_1^Y}{F_1^Y}\right],$$
$$\mathrm{E}\left[Z_t\right] = -\frac{A_1^Z}{F_1^Z} + e^{F_1^Z(t-t_0)}\left[z_0 + \frac{A_1^Z}{F_1^Z}\right],$$

and $t_0 \leq 0$ is the starting time of the factors processes. More generally, by iterative solution it can be verified that $C_{m,n}$ admits the expansion

$$C_{m,n} = \mathrm{E}\left[Y_t^m Z_t^n\right] = \sum_{i=0}^{m}\sum_{j=0}^{n} k_{i,j}^{(m,n)} e^{F_{i,j}(t-t_0)}, \tag{A3}$$

where the coefficients depend on the initial conditions y_0, z_0 and satisfy the following recursive relations

$$\begin{aligned}
k_{i<m,j<n}^{(m,n)} &= -\frac{A_m^Y k_{i,j}^{(m-1,n)} + A_n^Z k_{i,j}^{(m,n-1)}}{F_{m,n} - F_{i,j}}, \\
k_{i<m,n}^{(m,n)} &= -\frac{A_m^Y k_{i,n}^{(m-1,n)}}{F_{m,n} - F_{i,n}}, \\
k_{m,j<n}^{(m,n)} &= -\frac{A_n^Z k_{m,j}^{(m,n-1)}}{F_{m,n} - F_{m,j}}, \\
k_{m,n}^{(m,n)} &= \mathrm{E}\left[Y_{t_0}^m Z_{t_0}^n\right] + A_m^Y \sum_{i=0}^{m-1}\sum_{j=0}^{n} \frac{k_{i,j}^{(m-1,n)}}{F_{m,n} - F_{i,j}} + A_n^Z \sum_{i=0}^{m}\sum_{j=0}^{n-1} \frac{k_{i,j}^{(m,n-1)}}{F_{m,n} - F_{i,j}}.
\end{aligned} \tag{A4}$$

The moments $\mu_m^Y(t) = \mathrm{E}\left[Y_t^m\right]$ and $\mu_n^Z(t) = \mathrm{E}\left[Z_t^n\right]$ are specific cases of the expansion (A3), where coefficients are given by the column vector $(k_{i,0}^{(m,0)})_{i\leq m}$ and the row vector $(k_{0,j}^{(0,n)})_{j\leq n}$, while in general the set of coefficients $k_{i,j}^{(m,n)}$ can be cast in a $(m+1)\times(n+1)$ real matrix. For instance, for the case $C_{2,1} = \mathrm{E}\left[Y_t^2 Z_t\right]$ we obtain

$$\begin{aligned}
k_{0,0}^{(2,1)} &= -\frac{1}{F_{2,1}}\left[A_2^Y k_{0,0}^{(1,1)} + A_1^Z k_{0,0}^{(2,0)}\right], \\
k_{0,1}^{(2,1)} &= -\frac{A_2^Y k_{0,1}^{(1,1)}}{F_{2,1} - F_1^Z}, \\
k_{1,0}^{(2,1)} &= -\frac{1}{F_{2,1} - F_1^Y}\left[A_2^Y k_{1,0}^{(1,1)} + A_1^Z k_{1,0}^{(2,0)}\right], \\
k_{1,1}^{(2,1)} &= -\frac{A_2^Y k_{1,1}^{(1,1)}}{F_{2,1} - F_{1,1}}, \\
k_{2,0}^{(2,1)} &= -\frac{A_1^Z k_{2,0}^{(2,0)}}{F_{2,1} - F_2}, \\
k_{2,1}^{(2,1)} &= \mathrm{E}\left[Y_{t_0}^2 Z_{t_0}\right] - \left[k_{0,0}^{(2,1)} + k_{0,1}^{(2,1)} + k_{1,0}^{(2,1)} + k_{1,1}^{(2,1)} + k_{2,0}^{(2,1)}\right].
\end{aligned}$$

Ultimately, given the expansion (A3) and by inspection of (A1), we recognize that the moments of X can be cast in the form of Eq. (5) in the main text.

Appendix B: Computation of the return-volatility correlation

The following results are derived under the assumption that ρ_{XZ} is equal to zero. The numerator of the return-volatility correlation $\mathcal{L}(\tau;t) = \mathrm{E}\left[\mathrm{d}X_t\,\mathrm{d}X^2_{t+\tau}\right]/\mathrm{E}\left[\mathrm{d}X_t^2\right]^2$ can be cast in the form

$$\mathrm{E}\left[\mathrm{d}X_t\,\mathrm{d}X^2_{t+\tau}\right] = \mathrm{E}\left[\sigma_t\sigma^2_{t+\tau}\,\zeta^X_t\right]\,\mathrm{d}t^2\,.$$

where $\sigma_t = Y_t + Z_t$. Adopting the same convention of (Perelló *et al.* 2004), we formally [1] express the Wiener increment as $\mathrm{d}W^X_t = \zeta^X_t\,\mathrm{d}t$, where ζ^X_t is a Gaussian noise with zero mean and variance $\mathrm{E}\left[(\zeta^X_t)^2\right] = 1/\mathrm{d}t$. Novikov's theorem (Novikov 1965, Perelló and Masoliver 2003) allows to compute the expectation involving ζ^X_t, giving us

$$\frac{\mathrm{E}\left[\mathrm{d}X_t\,\mathrm{d}X^2_{t+\tau}\right]}{\mathrm{d}t^2} = 2\rho_{XY}\sqrt{\sigma^2_Y}H(\tau)\mathrm{e}^{-\kappa_Y\tau}\;\mathrm{E}\left[Y_t\;\sigma_t\sigma_{t+\tau}\;\exp\left(\sqrt{\sigma^2_Y}\Delta_t W^Y(\tau)\right)\right]\,, \tag{B1}$$

where we define $\Delta_t W(\tau) \doteq \int_t^{t+\tau} dW_s$. We refer the interested reader to Section IV in (Delpini and Bormetti 2011) for further details regarding the derivation of the previous equation. The right hand side of (B1) can be split into four pieces proportional to the expectations

$$\begin{aligned}
f_{YYY}(\tau,t) &\doteq \mathrm{E}\left[Y^2_t\,Y_{t+\tau}\exp\left(\sqrt{\sigma^2_Y}\,\Delta_t W^Y(\tau)\right)\right]\,,\\
f_{YYZ}(\tau,t) &\doteq \mathrm{E}\left[Y^2_t\,Z_{t+\tau}\exp\left(\sqrt{\sigma^2_Y}\,\Delta_t W^Y(\tau)\right)\right]\,,\\
f_{YZY}(\tau,t) &\doteq \mathrm{E}\left[Y_t\,Z_t\,Y_{t+\tau}\exp\left(\sqrt{\sigma^2_Y}\,\Delta_t W^Y(\tau)\right)\right]\,,\\
f_{YZZ}(\tau,t) &\doteq \mathrm{E}\left[Y_t\,Z_t\,Z_{t+\tau}\exp\left(\sqrt{\sigma^2_Y}\,\Delta_t W^Y(\tau)\right)\right]\,.
\end{aligned}$$

Then, it is possible to show that these quantities satisfy the relations

$$\begin{aligned}
f_{YYY}(\tau,t) - (\sigma^2_Y-\kappa_Y)\int_0^\tau f_{YYY}(\tau',t)\;\mathrm{e}^{\frac{\sigma^2_Y}{2}(\tau-\tau')}\mathrm{d}\tau' &= \mathrm{e}^{\frac{\sigma^2_Y}{2}\tau}\left[C_{3,0}+\kappa_Y y_\infty\tau C_{2,0}\right]\,,\\
f_{YYZ}(\tau,t) + \kappa_Z\int_0^\tau f_{YYZ}(\tau',t)\;\mathrm{e}^{\frac{\sigma^2_Y}{2}(\tau-\tau')}\mathrm{d}\tau' &= \mathrm{e}^{\frac{\sigma^2_Y}{2}\tau}\left[C_{2,0}C_{0,1}+\kappa_Z z_\infty\tau C_{2,0}\right]\,,\\
f_{YZY}(\tau,t) - (\sigma^2_Y-\kappa_Y)\int_0^\tau f_{YZY}(\tau',t)\;\mathrm{e}^{\frac{\sigma^2_Y}{2}(\tau-\tau')}\mathrm{d}\tau' &= \mathrm{e}^{\frac{\sigma^2_Y}{2}\tau}\left[C_{2,0}C_{0,1}+\kappa_Y y_\infty\tau C_{1,0}C_{0,1}\right]\,,\\
f_{YZZ}(\tau,t) + \kappa_Z\int_0^\tau f_{YZZ}(\tau',t)\;\mathrm{e}^{\frac{\sigma^2_Y}{2}(\tau-\tau')}\mathrm{d}\tau' &= \mathrm{e}^{\frac{\sigma^2_Y}{2}\tau}\left[C_{1,0}C_{0,2}+\kappa_Z z_\infty\tau C_{1,0}C_{0,1}\right]\,.
\end{aligned}$$

The above relations correspond to a set of Volterra integro-differential equations of the second kind. Their solutions are known in closed-form, and after plugging them in Eq. (B1), the final expression

[1] This convention is somewhat unusual, but admissible as the distributional assumption on ζ^X_t guarantees that dW_t is still Gaussian with zero mean and variance equal to $\mathrm{d}t$. This choice proves to be convenient to develop the following calculations.

of the leverage correlation reads

$$
\begin{aligned}
\mathcal{L}(\tau;t) = {} & \frac{2\rho_{XY}\sqrt{\sigma_Y^2}H(\tau)\,\mathrm{e}^{\left(\frac{\sigma_Y^2}{2}-\kappa_Y\right)\tau}}{\left(C_{2,0}+2C_{1,0}C_{0,1}+C_{0,2}\right)^2} \times \left\{ \left[\left(\frac{\kappa_Y y_\infty}{\sigma_Y^2-\kappa_Y} - z_\infty\right)\left(C_{2,0}+C_{1,0}C_{0,1}\right)\right] \right. \\
& + \left[C_{2,0}C_{0,1}+C_{1,2}-z_\infty\left(C_{2,0}+C_{1,0}C_{0,1}\right)\right]\mathrm{e}^{-\kappa_Z\tau} \\
& \left. - \left[C_{3,0}+C_{2,0}C_{0,1}+\frac{\kappa_Y y_\infty}{\sigma_Y^2-\kappa_Y}\left(C_{2,0}+C_{1,0}C_{0,1}\right)\right]\mathrm{e}^{\left(\sigma_Y^2-\kappa_Y\right)\tau} \right\}. \qquad \text{(B2)}
\end{aligned}
$$

A meaningful comparison of the previous expression with real data requires to take the stationary limit $t_0 \to -\infty$, replacing the functions $C_{m,n}$ with the asymptotic values $C_{m,n}^{\mathrm{st}}$.

Appendix C: Computation of the squared return correlation

We derive the following results under the assumption that both correlation coefficients ρ_{XZ} and ρ_{YZ} are equal to zero. Resorting to the same parametrization of the Wiener variation adopted in Appendix B we have

$$
\begin{aligned}
\mathcal{A}'(\tau;t) &= \mathrm{d}t^2\mathrm{E}\left[\sigma_t^2\sigma_{t+\tau}^2\,\mathrm{d}W_t^X\,\zeta_t^X\right] \\
&= \mathrm{d}t^2\mathrm{E}\left[\sigma_t^2\sigma_{t+\tau}^2\right] + \mathcal{O}(\mathrm{d}t^3)\,. \qquad \text{(C1)}
\end{aligned}
$$

In order to compute the autocorrelation function of squared returns, the τ-lagged correlations

$$
f_t^{(m,n,p,q)}(\tau) = \mathrm{E}\left[Y_t^m Z_t^n Y_{t+\tau}^p Z_{t+\tau}^q\right]
$$

have to be evaluated. The relevant cases correspond to $p, q \leq 2$, and below we detail the corresponding exact results, all of which are obtained replacing the process $Y_{t+\tau}^p Z_{t+\tau}^q$ with its integral representation from time t to time $t+\tau$.

Computation of $f_t^{(m,n,1,0)}(\tau) = \mathrm{E}\left[Y_t^m Z_t^n Y_{t+\tau}\right]$.

It is readily verified that $f_t^{(m,n,1,0)}(\tau)$ is solution of a linear ODE, giving

$$
f_t^{(m,n,1,0)}(\tau) = -\frac{A_1^Y}{F_1^Y}C_{m,n} + e^{F_1^Y\tau}\left[C_{m+1,n}+\frac{A_1^Y}{F_1^Y}C_{m,n}\right]. \qquad \text{(C2)}
$$

Computation of $f_t^{(m,n,0,1)}(\tau) = \mathrm{E}\left[Y_t^m Z_t^n Z_{t+\tau}\right]$.

In much the same way we have

$$
f_t^{(m,n,0,1)}(\tau) = -\frac{A_1^Z}{F_1^Z}C_{m,n} + e^{F_1^Z\tau}\left[C_{m,n+1}+\frac{A_1^Z}{F_1^Z}C_{m,n}\right].
$$

Computation of $\boldsymbol{f_t^{(m,n,2,0)}(\tau)} = \mathrm{E}\left[\boldsymbol{Y_t^m Z_t^n Y_{t+\tau}^2}\right]$.

After replacement of $Y_{t+\tau}^2$, we can write

$$f_t^{(m,n,2,0)}(\tau) = f_t^{(m,n,2,0)}(0) + F_2^Y \int_0^\tau f_t^{(m,n,2,0)}(\tau')\mathrm{d}\tau' + A_2^Y \int_0^\tau f_t^{(m,n,1,0)}(\tau')\mathrm{d}\tau' \,;$$

further, we can replace the solution (C2) for $f_t^{(m,n,1,0)}(\tau')$ in the second integral, leading to straightforward integrations of exponential functions of τ. Finally, we are left with

$$\begin{aligned} f_t^{(m,n,2,0)}(\tau) = &\left[\frac{A_2^Y A_1^Y}{F_2^Y F_1^Y} C_{m,n}\right] + e^{F_1^Y \tau}\left[-\frac{A_2^Y}{F_2^Y - F_1^Y}\left(C_{m+1,n} + \frac{A_1^Y}{F_1^Y} C_{m,n}\right)\right] \\ &+ e^{F_2^Y \tau}\left[C_{m+2,n} + \frac{A_2^Y}{F_2^Y - F_1^Y}\left(C_{m+1,n} + \frac{A_1^Y}{F_2^Y} C_{m,n}\right)\right] . \end{aligned} \tag{C3}$$

Computation of $\boldsymbol{f_t^{(m,n,0,2)}(\tau)} = \mathrm{E}\left[\boldsymbol{Y_t^m Z_t^n Z_{t+\tau}^2}\right]$.

As before, after replacement of the parameters for the dynamics of the $Z_{t+\tau}^2$ process, we get to

$$\begin{aligned} f_t^{(m,n,0,2)}(\tau) = &\left[\frac{A_2^Z A_1^Z}{F_2^Z F_1^Z} C_{m,n}\right] \quad + e^{F_1^Z \tau}\left[-\frac{A_2^Z}{F_2^Z - F_1^Z}\left(C_{m,n+1} + \frac{A_1^Z}{F_1^Z} C_{m,n}\right)\right] \\ &+ e^{F_2^Z \tau}\left[C_{m,n+2} + \frac{A_2^Z}{F_2^Z - F_1^Z}\left(C_{m,n+1} + \frac{A_1^Z}{F_2^Z} C_{m,n}\right)\right] . \end{aligned} \tag{C4}$$

Computation of $\boldsymbol{f_t^{(m,n,1,1)}(\tau)} = \mathrm{E}\left[\boldsymbol{Y_t^m Z_t^n Y_{t+\tau} Z_{t+\tau}}\right]$.

The evolution of the joint process $Y_t Z_t$ is given by

$$\mathrm{d}(Y_t Z_t) = \left(F_{1,1} Y_t Z_t + A_1^Z Y_t + A_1^Y Z_t\right)\,\mathrm{d}t + \sqrt{\sigma_Y^2}\, Y_t Z_t\,\mathrm{d}W_t^Y + \sqrt{\sigma_Z^2}\, Y_t Z_t \mathrm{d}W_t^Z \,,$$

and substitution inside the expectation gives

$$\begin{aligned} f_t^{(m,n,1,1)}(\tau) = &\left[\frac{A_1^Y A_1^Z}{F_{1,1}}\left(\frac{1}{F_1^Y} + \frac{1}{F_1^Z}\right)\right] C_{m,n} - \frac{A_1^Z}{F_{1,1} - F_1^Y}\, e^{F_1^Y \tau}\left[C_{m+1,n} + \frac{A_1^Y}{F_1^Y} C_{m,n}\right] \\ &- \frac{A_1^Y}{F_{1,1} - F_1^Z}\, e^{F_1^Z \tau}\left[C_{m,n+1} + \frac{A_1^Z}{F_1^Z} C_{m,n}\right] \\ &+ e^{F_{1,1}\tau}\left[C_{m+1,n+1} + \frac{A_1^Z}{F_{1,1} - F_1^Y} C_{m+1,n} + \frac{A_1^Y}{F_{1,1} - F_1^Z} C_{m,n+1}\right. \\ &\left. - A_1^Y A_1^Z\left(\frac{2F_{1,1} - F_1^Y - F_1^Z}{F_{1,1}(F_{1,1} - F_1^Y)(F_{1,1} - F_1^Z)}\right) C_{m,n}\right] . \end{aligned} \tag{C5}$$

As expected from the structure of model (4), and as confirmed by all previous examples, it is clear that the functions $f_t^{(m,n,p,q)}(\tau)$ admit a general expansion reading

$$f_t^{(m,n,p,q)}(\tau) = \sum_{i=1}^{p}\sum_{j=1}^{q} h_{i,j}^{(m,n,p,q)}(t)\, e^{F_{i,j}\tau}\,,$$

where the terms $h_{i,j}^{(m,n,p,q)}(t)$ can be computed exactly.

Final expression for $\mathcal{A}(\tau;t)$

Coming back to Eq. (C1) we have

$$\begin{aligned}\frac{\mathcal{A}'(\tau;t)}{\mathrm{d}t^2} &= f_t^{(2,0,2,0)}(\tau) + f_t^{(0,2,2,0)}(\tau) + 2f_t^{(0,2,2,0)}(\tau) + f_t^{(2,0,0,2)}(\tau)\\ &+ f_t^{(0,2,0,2)}(\tau) + 2f_t^{(0,2,0,2)}(\tau) + 2\left[f_t^{(2,0,1,1)}(\tau) + f_t^{(0,2,1,1)}(\tau) + 2f_t^{(0,2,1,1)}(\tau)\right]. \end{aligned} \tag{C6}$$

By means of Eq.s (C3)-(C5), and after defining the following auxiliary functions

$$\begin{aligned} T_1 &= C_{2,0} + C_{0,2} + 2C_{1,1}\,, \quad T_2 = C_{3,0} + C_{1,2} + 2C_{2,1}\,,\\ T_2^* &= C_{0,3} + C_{2,1} + 2C_{1,2}\,, \quad T_3 = C_{4,0} + C_{2,2} + 2C_{3,1}\,,\\ T_3^* &= C_{0,4} + C_{2,2} + 2C_{1,3}\,, \quad T_4 = C_{3,1} + C_{1,3} + 2C_{2,2}\,, \end{aligned}$$

we can write the final, somewhat convoluted, expression

$$\begin{aligned}\frac{\mathcal{A}'(\tau;t)}{\mathrm{d}t^2} &= \left[\frac{A_2^Y A_1^Y}{F_2^Y F_1^Y} + \frac{A_2^Z A_1^Z}{F_2^Z F_1^Z} + 2\frac{A_1^Y A_1^Z}{F_{1,1}}\left(\frac{1}{F_1^Y} + \frac{1}{F_1^Z}\right)\right] T_1\\ &- e^{F_1^Y\tau}\left(T_2 + \frac{A_1^Y}{F_1^Y}T_1\right)\left[\frac{A_2^Y}{F_2^Y - F_1^Y} + 2\frac{A_1^Z}{F_{1,1} - F_1^Y}\right]\\ &- e^{F_1^Z\tau}\left(T_2^* + \frac{A_1^Z}{F_1^Z}T_1\right)\left[\frac{A_2^Z}{F_2^Z - F_1^Z} + 2\frac{A_1^Y}{F_{1,1} - F_1^Z}\right]\\ &+ e^{F_2^Y\tau}\left[T_3 + \frac{A_2^Y}{F_2^Y - F_1^Y}\left(T_2 + \frac{A_1^Y}{F_2^Y}T_1\right)\right]\\ &+ e^{F_2^Z\tau}\left[T_3^* + \frac{A_2^Z}{F_2^Z - F_1^Z}\left(T_2^* + \frac{A_1^Z}{F_2^Z}T_1\right)\right]\\ &+ 2e^{F_{1,1}\tau}\left[T_4 + \frac{A_1^Y}{F_{1,1} - F_1^Z}T_2^* + \frac{A_1^Z}{F_{1,1} - F_1^Y}T_2\right.\\ &\left. + \frac{2F_{1,1} - F_1^Y - F_1^Z}{F_{1,1}(F_{1,1} - F_1^Y)(F_{1,1} - F_1^Z)}\right]. \end{aligned} \tag{C7}$$

Ultimately, evaluation of the volatility autocorrelation $\mathcal{A}(\tau;t)$ given by (6) requires to compute $\mathrm{Var}[\mathrm{d}X_t^2] = \mathrm{E}\left[\mathrm{d}X_t^4\right] - \mathrm{E}\left[\mathrm{d}X_t^2\right]^2$ which is given by

$$\begin{aligned} &3\left(C_{4,0} + 4C_{3,1} + 6C_{2,2} + 4C_{1,3} + C_{0,4}\right)\mathrm{d}t^2\\ &- \left(C_{2,0} + 2C_{1,1} + C_{0,2}\right)^2\mathrm{d}t^2\,. \end{aligned}$$