UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche ed Aziendali "Marco Fanno"

CHASING VOLATILITY. A PERSISTENT MULTIPLICATIVE ERROR MODEL WITH JUMPS

MASSIMILIANO CAPORIN
University of Padova

EDUARDO ROSSI
University of Pavia

PAOLO SANTUCCI DE MAGISTRIS
University of Aahrus

September 2014

"MARCO FANNO" WORKING PAPER N.186

Chasing volatility
A persistent multiplicative error model with jumps

Massimiliano Caporin [*] Eduardo Rossi [†]
Paolo Santucci de Magistris [‡]

This version: August 29, 2014

Abstract

The realized volatility of financial returns is characterized by persistence and occurrence of unpredictable large increments. To capture those features, we introduce the Multiplicative Error Model with jumps (MEM-J). When a jump component is included in the multiplicative specification, the conditional density of the realized measure is shown to be a countably infinite mixture of Gamma and K distributions. Strict stationarity conditions are derived. A Monte Carlo simulation experiment shows that maximum likelihood estimates of the model parameters are reliable even when jumps are rare events. We estimate alternative specifications of the model using a set of daily bipower measures for 7 stock indexes and 16 individual NYSE stocks. The estimates of the jump component confirm that the probability of jumps dramatically increases during the financial crises. Compared to other realized volatility models, the introduction of the jump component provides a sensible improvement in the fit, as well as for in-sample and out-of-sample volatility tail forecasts.

Keywords: Multiplicative Error Model with Jumps, Jumps in volatility, Realized measures, Volatility-at-Risk

J.E.L. codes: C22, C58, G10

[*]Dipartimento di Scienze Economiche e Aziendali "Marco Fanno", University of Padova, Italy.
[†]Dipartimento di Scienze Economiche e Aziendali, University of Pavia, Italy.
[‡]Department of Economics and Business and CREATES, University of Aarhus, Denmark.

1 Introduction

A great deal of the recent literature on volatility modeling exploits realized volatility measures as ex-post estimates of the return variation over a given horizon. The recent financial crisis has been an important test for existing volatility models. In general, models of realized measures are unable to fit the abnormal levels reached by the volatility during the financial turmoil. This seems to call for a more realistic econometric specification of such models. The inclusion of jumps in the volatility process is a step forward a more appropriate description of the volatility dynamics. Recently, the analysis of jumps in prices and volatility, and their interactions, in a continuous-time framework has shown the importance of both components in fitting the observed dynamics of prices, see e.g. Chernov et al. (2003), Duffie et al. (2000), Pan (2002), Eraker (2004), Eraker et al. (2003), Jones (2003), Broadie et al. (2007), Todorov and Tauchen (2011), Andersen et al. (2012), Bandi and Renò (2012, 2013).

In a discrete-time setting, the analysis has focused on the role that jumps in prices have in predicting the future volatility. Andersen et al. (2007) extend the HAR-RV model to include past price jumps, i.e. the HAR-RV-J model. Instead Caporin et al. (2014) explicitly model the volatility jumps in a HAR setup. This allows a direct estimation of volatility jumps which is used to analyze the economic determinants. One of the results of their analysis is that volatility jumps increase significantly the fit of the model in the right tail. It emerges that it is important to allow for the presence of jumps because this component can contribute to explain the level of the daily volatility during periods of market turmoils. One limitation of the HAR models in the log-transformed volatility series is that to obtain the forecasts distribution of the levels can be problematic.

We propose the Asymmetric HAR-MEM-J (AHAR-MEM-J) which is an extension of the multiplicative error model (MEM) by Engle (2002) and Engle and Gallo (2006). We extend the MEM approach to the modeling of the realized measures by including a latent process, labeled *jump*, that causes infrequent large moves in the volatility. The AHAR-MEM-J is a three-factor model: first, a long-run factor, modeled by the Asymmetric HAR, which replicates the long-run dependence present in volatility; second, a short-run factor, which represents the transitory component of the volatility process; and third, the jump factor, which is responsible for the presence of realizations in the right tail of volatility distribution. For an analogous interpretation, see Ghysels et al. (2004). Thanks to the availability of realized volatility measures which sterilize the effect of price jumps on volatility, we can easily focus on time series that include only volatility jumps, if they are present. Modeling the volatility by including a jump process increases the model's capability of capturing extreme movements, or tail events. Potential sources of jump innovations to volatility can be important news, data releases, or unexpected events, which might induce market participants to suddenly revise their portfolios, thus producing large variations in the volatility level. During the financial crises of 2008 the volatilities of stock markets across the world have experienced such abnormal movements.

Our approach is similar to that of Bauwens and Veredas (2004). We specify the volatility process as a combination of a continuous volatility component and a discrete compound Poisson process for the jumps; the two elements determine the level of the volatility in a multiplicative framework. It follows that the conditional density of the realized measure is a countably infinite mixture of two random variables: one distributed as a Gamma, and the second, when the number of jumps is strictly larger than zero, is distributed as a Kappa, henceforth K. The K is a product distribution, known in physics and radar applications, but never used in econometrics, to the best of our knowledge. The K is obtained as the product of two Gamma-distributed random variables. Exploiting the knowledge of the mixture density that characterizes the conditional distribution of the observed volatility measure, it is possible to obtain in closed form the conditional moments, the likelihood function and the quantiles. In order to account for the empirical evidence of jump clustering, the intensity parameter, governing the jump occurrence in the compound Poisson process, is specified in a time varying form, according to an autoregressive specification, in the spirit of Hansen (1994) and Maheu and McCurdy (2004). For what concerns the continuous volatility component, we have adopted an Asymmetric Heterogeneous Autoregressive (AHAR) specification. A common finding in the empirical literature that employs MEMs in volatility modeling is indeed that the estimates of GARCH-type specifications of the conditional mean turn out to be close to be integrated. Such an evidence highlights the necessity of having a mean model specification that takes into account the persistence observed in the realized measure series. Our empirical results confirm this finding. Indeed, the HAR specification sensibly improves the fit compared to simpler, and less persistent, specifications of the continuous volatility component. The model parameters can be estimated by maximum likelihood methods. However, given the mixture structure, different local maxima may exist, see Frühwirth-Schnatter (2006). For our model, a Monte Carlo simulation experiment shows the appropriateness of the finite-sample features of maximum likelihood estimation. In addition, the maximum likelihood estimates of the jump component seem to be reliable, even when jumps are rare events.

The empirical application is based on daily bipower volatility series of individual stocks and equity indexes. The estimation results highlight a positive probability of jumps in volatility, which is consistent with the findings of previous studies on the topic. The AHAR-MEM-J with time-varying jump intensity allows for a greater flexibility in accommodating extremely large volatilty realizations, dramatically improving the fit of the baseline MEM. By analogy to the Value-at-Risk (VaR), we introduce the *Volatility-at-Risk* (VolaR) which constitutes a natural measure of risk when designing volatility trading strategy. The evaluation of the VolaR estimation provided by alternative specifications is in favour of the MEM-J against models without jumps.

In summary, the contributions of the paper are at least three. Firstly, we generalize the baseline MEM of Engle and Gallo (2006) by including a jump term, which captures the occasional boosts of volatility, and a pseudo long-memory component which is able to account for the observed persistence. Secondly, the conditional density of the model's dependent variable is derived as well as the log-likelihood function. Finally, we provide evidence that the jumps are a relevant component of the realized measure series, thhus supporting the claim that ignoring them might lead to an under-estimation of the VolaR. This under-coverage of the right tail of the volatility density leads to an underestimation of the volatility risk especially in periods of markets turmoils, with consequences for the pricing of derivatives and volatility trading strategies.

The paper is organized as follows. Section 2 sets the notation of the baseline MEM. Section 3 describes the MEM-J and the finite mixture distribution that characterizes the conditional density of the model's dependent variable. Conditional moments are also presented. Section 4 discusses both model extensions with HAR dynamics and time-varying parameters, and the model's properties, such as conditions for covariance stationarity and maximum likelihood estimation. Section 5 illustrates the empirical results. In particular, Subsection 5.1 describes the dataset and the construction of the volatility series, while the subsection 5.2 provides a discussion of the empirical results obtained with stocks indexes and individual S&P 500 stocks under different model specifications; in subsection 5.3 results with alternative MEM specifications are presented. In Section 6 the results of the VolaR analysis are reported and discussed. Finally, Section 7 concludes. Proofs, selected derivations of relevant quantities and additional theoretical details are included in Appendix A. A Web Appendix also contains some additional results and details on the K distribution.

2 The baseline MEM

In this section we briefly present the MEM, in its simplest form, as introduced by Engle and Gallo (2006). Let RM_t be the daily realized volatility measure.[1] We assume that RM_t follows a general MEM, i.e.

$$RM_t = \mu_t \varepsilon_t \tag{1}$$

with

$$\mu_t = \omega + \alpha' X_{t-1} + \beta \mu_{t-1}$$

[1]We assume that the realized volatility measure used in the empirical analysis is corrected for microstructure noise and filtered from price jumps. The estimator adopted in the present paper is briefly described in Section 5.1.

and where X_{t-1} is a vector that contains variables included in the information set at time $t-1$. Moreover, the innovation ε_t is a random variable with scale-shape Gamma density[2]

$$\varepsilon_t | I_{t-1} \overset{iid}{\sim} \Gamma\left(\frac{1}{\nu}, \nu\right) \tag{2}$$

$$f\left(\varepsilon_t | I_{t-1}\right) = \frac{1}{\varepsilon_t}\left(\varepsilon_t \nu\right)^{\nu} \frac{1}{\Gamma\left(\nu\right)} e^{-\varepsilon_t \nu}, \quad \varepsilon_t \geq 0 \tag{3}$$

where $\frac{1}{\nu}$ is the scale and ν is the shape of the Gamma density, both driven by the common parameter ν. In this case, we have $\mathbb{E}_{t-1}\left[\varepsilon_t\right] = 1$ and $\mathbb{V}_{t-1}\left[\varepsilon_t\right] = \frac{1}{\nu}$. By the properties of the Gamma distribution (in particular the product of a Gamma-variate by a scalar assuming it is known or included in the information set) we have

$$f\left(RM_t | I_{t-1}\right) = \frac{1}{RM_t}\left(\frac{RM_t \nu}{\mu_t}\right)^{\nu} \frac{1}{\Gamma\left(\nu\right)} e^{-\frac{RM_t \nu}{\mu_t}}, \quad RM_t \geq 0.$$

If the realized measure follows a MEM, the conditional mean and variance are given as

$$\mathbb{E}\left[RM_t | I_{t-1}\right] = \mu_t = \omega + \alpha' X_{t-1} + \beta \mu_{t-1}, \tag{4}$$

and

$$\mathbb{V}\left[RM_t | I_{t-1}\right] = \mu_t^2 \nu^{-1}. \tag{5}$$

The form of μ_t is sufficiently flexible to include simple auto-regressive patterns, HAR terms, asymmetry, or predetermined variables. Examples of possible specifications for μ_t are given, among others, in Engle and Gallo (2006) and Brownlees et al. (2012). Interestingly, the term μ_t induces the conditional variance of the realized measure to be time-varying, thus making the MEM consistent with the so-called volatility-of-volatility feature, studied in Corsi et al. (2008) among others. The literature on multiplicative models of volatility includes several extensions of the baseline MEM. For example, Gallo and Otranto (2012) extend the MEM to include time-varying parameters as in the case of regime-switching MEM. The latter specification allows for changing parameters but requires to impose a priori structures on the form of the transition and on the number of underlying regimes. Alternatively, Haerdle et al. (2012) propose to adaptively estimate the MEM based on a window of varying length and thus providing updated parameter estimates at each point in time.

[2]See the Web Appendix for some further details on Gamma, and related, random variables.

3 A Multiplicative Error Model with Jumps

The baseline MEM with a Gamma distributed error term is poorly designated to account for the presence of large and abrupt movements, i.e. the jumps, that characterize the volatility dynamics. The presence and the effects of volatility jumps have been already documented in the literature either in a continuous time framework, or, in discrete time, see Caporin et al. (2014), among others. Excluding volatility jumps reduces the fit of the model to real data, thus resulting in a worsening of the forecasting performance. We therefore propose a generalization of the MEM of Engle and Gallo (2006), which we call MEM-J. The new model introduces, in a multiplicative way, an additional volatility jump term to the standard MEM of Engle and Gallo (2006). The dynamic of the MEM is also generalized by the inclusion of HAR terms following Corsi (2009), we defer the discussion of this to Section 4. Under the MEM-J specification, the realized volatility measure RM_t is decomposed into the product of three elements

$$RM_t = \mu_t Z_t \varepsilon_t \tag{6}$$

where μ_t is a function measurable with respect to the information set at time $t-1$, Z_t is the volatility jump component, and the innovation ε_t is a scale-shape Gamma, $\varepsilon_t|I_{t-1} \sim \Gamma\left(\frac{1}{\nu}, \nu\right)$. Hereafter, to simplify the interpretation of the model outcome, the Gamma density of the innovation term is expressed in the mean-shape representation, i.e. $\varepsilon_t|I_{t-1} \sim \Gamma(1, \nu)$, which is, by construction, equivalent to the scale-shape representation. Hence, a number of assumptions on Z_t and ε_t are required in order to identify and separate the two sources of shocks. The jump term, Z_t, is defined as

$$Z_t = \begin{cases} 1 & N_t = 0 \\ \sum_{j=1}^{N_t} Y_{j,t} & N_t > 0 \end{cases} \tag{7}$$

where N_t is a non-negative integer-valued random variable that represents the number of jumps occurring at time t. When $N_t = 0$, i.e. jumps are absent, the MEM-J reduces to the MEM. The random variable determining the occurrence and the number of jumps, N_t, is modeled as a Poisson with intensity λ,

$$P\left(N_t = m|I_{t-1}\right) = \frac{e^{-\lambda}\lambda^m}{m!}, \quad m = 0, 1, 2, ... \tag{8}$$

The second characterizing element of Z_t defines the size of the jumps. This is determined by the sum of independent Gamma random variables, $Y_{j,t} \sim \Gamma(1, \varsigma)$ (in mean-shape form). Note that the jump density is not dependent on time and the parameter characterizing the jump evolution is assumed to be time-invariant.

Assumption 1 *In the MEM-J*

i. ε_t *is an i.i.d. process defined on positive support with* $\mathbb{E}[\varepsilon_t] = 1$.

ii. ε_t, N_t *and the variables* $Y_{j,t}$, $\quad j = 1, 2, \ldots, N_t$, *are assumed to be independent for any* t.

By the properties of the Gamma density,[3] it follows that, if $N_t = m > 0$,

$$Z_t | N_t = m > 0, I_{t-1} \sim \Gamma\left(m, m\varsigma\right) \tag{9}$$

in mean-shape representation. It is interesting to note that the jump component has mean and variance which depend on the number of jumps, i.e. $\mathbb{E}\left[Z_t | N_t = m > 0, I_{t-1}\right] = m$ and $\mathbb{V}\left[Z_t | N_t = m > 0, I_{t-1}\right] = \frac{m}{\varsigma}$. So far, all parameters are assumed to be time invariant. In Section 4 we discuss the introduction of time varying parameters. Additional flexibility in the model parameters can potentially capture the increase in the jumps contribution to the overall variability of the volatility during market turmoils.

It follows from equation (6) that the MEM-J can be written as

$$RM_t = \mu_t \eta_t \tag{10}$$

where the innovation term $\eta_t = Z_t \varepsilon_t$ is the product of two sources of shocks, one depending on jumps. In the next paragraphs we will study the properties of the conditional density of η_t and of RM_t which clearly depend on the distributional assumptions made on Z_t and ε_t.

3.1 The conditional density of η_t

The density of η_t depends on the realization of N_t. When $N_t = 0$, we have that $\eta_t | N_t = 0, I_{t-1}$ is simply equal to $\varepsilon_t | I_{t-1}$, since $Z_t = 1$. In this case, the conditional density of η_t, in mean-shape form, coincides with that of ε_t, i.e.

$$\eta_t | N_t = 0, I_{t-1} \equiv \varepsilon_t | I_{t-1} \sim \Gamma\left(1, \nu\right). \tag{11}$$

Differently, when $N_t = m > 0$ the conditional density of η_t given Z_t is Gamma in mean-shape form

$$\eta_t | Z_t, N_t = m > 0, I_{t-1} \sim \Gamma\left(Z_t, \nu\right). \tag{12}$$

In order to derive the conditional density of η_t given N_t and I_{t-1}, a fundamental element for the construction of the model likelihood, we have to evaluate the following integral:

$$\int_0^\infty f(\eta_t | N_t = m > 0, Z_t, I_{t-1}) f(z_t | N_t = m > 0, I_{t-1}) \mathrm{d}z, \tag{13}$$

[3] See the discussion in the Web Appendix.

where both conditional densities in the integral are Gamma expressed in mean-shape form. We thus introduce the following proposition (proof in Appendix A).

Proposition 1 *Under Assumption 1, consider $\eta_t = Z_t \varepsilon_t$ where Z_t defined in (7) has the conditional density in (9) and $\varepsilon_t \sim \Gamma(1, \nu)$. Assuming that Z_t and ε_t are independent at all leads and lags, it follows that*

$$f(\eta_t | N_t = m > 0, I_{t-1}) = \frac{2}{\eta_t} \big(\eta_t \varsigma \nu\big)^{\frac{m\varsigma+\nu}{2}} \frac{1}{\Gamma(m\varsigma)\Gamma(\nu)} \mathbb{K}_{m\varsigma-\nu}\Big(2\sqrt{\eta_t \varsigma \nu}\Big), \tag{14}$$

where $\mathbb{K}_a(\cdot)$ is the modified Bessel function of the second kind. Thus the innovation term η_t, conditional on $N_t = m > 0$ and I_{t-1}, has a K distribution, see Redding (1999), denoted as

$$\eta_t | N_t = m > 0, I_{t-1} \sim K\big(m, m\varsigma, \nu\big).$$

The first two moments of η_t, conditional on $N_t = m > 0$ and I_{t-1}, are

$$\begin{aligned}
\mathbb{E}\left[\eta_t | N_t = m > 0, I_{t-1}\right] &= m \\
\mathbb{V}\left[\eta_t | N_t = m > 0, I_{t-1}\right] &= m^2 \frac{m\varsigma + \nu + 1}{m\varsigma\nu} = m^2 \left(\frac{1}{\nu} + \frac{1}{m\varsigma} + \frac{1}{m\nu\varsigma}\right).
\end{aligned}$$

The K density is governed by three parameters which have specific meanings in our case. The first parameter is the mean of the K density, and it is equal to the number of jumps, m. The second parameter depends on the shape of the jump component Z_t, while the third also depends on the shape parameter of the innovation term ε_t. Additional details on the K distribution are presented in the Web Appendix. Interestingly, the conditional variance of η_t is an increasing function of the number of jumps arrivals, m. Hence, periods with a larger number of jumps arrivals are characterized by a higher volatility-of-volatility.

The innovation term conditional on the information set, I_{t-1}, might be seen as characterized by a countably infinite mixture

$$f(\eta_t | I_{t-1}) = P\left(N_t = 0 | I_{t-1}\right) \Gamma\left(1, \nu\right) + \sum_{m=1}^{\infty} P\left(N_t = m | I_{t-1}\right) \times K\left(m, m\varsigma, \nu\right), \tag{15}$$

where

$$P\left(N_t = 0 | I_{t-1}\right) = e^{-\lambda}.$$

The mixing variable is the Poisson process N_t, which depends on the parameter λ. As λ increases, more weight is given to the K distribution, while when $\lambda = 0$ the density of η_t is $\Gamma\left(1, \nu\right)$ and the MEM-J reduces to the MEM.

3.2 The conditional density of RM_t

The conditional density of the realized measure, given $N_t = m > 0$ and I_{t-1}, follows from the distribution of the term η_t in equation (14). The following proposition reports the density and the subsequent corollary introduces the conditional moments of RM_t.

Proposition 2 *Consider model (10) where $\eta_t = Z_t\varepsilon_t$ with Z_t defined in equation (7) and $\varepsilon_t \sim \Gamma(1,\nu)$. Assuming that Z_t and ε_t are independent at all leads and lags, it follows that*

$$f(RM_t|N_t = m > 0, I_{t-1}; \theta) = \frac{2}{RM_t}\left(\frac{RM_t}{\mu_t}\varsigma\nu\right)^{\frac{m\varsigma+\nu}{2}} \frac{1}{\Gamma(m\varsigma)\Gamma(\nu)} \mathbb{K}_{m\varsigma-\nu}\left(2\sqrt{\frac{RM_t}{\mu_t}\varsigma\nu}\right), \tag{16}$$

where θ is the vector of parameters. Thus the realized measure RM_t, conditional on $N_t = m > 0$ and I_{t-1}, has a K distribution, denoted as

$$RM_t|N_t = m > 0, I_{t-1} \sim K\big(m\mu_t, m\varsigma, \nu\big).$$

The first two moments of RM_t, conditional on $N_t = m > 0$ and I_{t-1}, are

$$\begin{aligned}
\mathbb{E}\left[RM_t|N_t = m > 0, I_{t-1}\right] &= \mu_t m, \\
\mathbb{V}\left[RM_t|N_t = m > 0, I_{t-1}\right] &= \mu_t^2 m^2 \frac{m\varsigma + \nu + 1}{m\varsigma\nu}.
\end{aligned}$$

As a result, both the conditional mean and the variance of the RM sequence are not only time-varying and driven by μ_t as in the baseline MEM, but also dependent on the realized number of jumps, m. On the other hand, when jumps are absent, i.e. $m = 0$, the conditional density $f(RM_t|N_t = 0, I_{t-1}; \theta)$ is that of the baseline MEM. Integrating out the realized number of jumps, the density of RM_t conditional on the information set I_{t-1} is a countably infinite mixture

$$f(RM_t|I_{t-1}; \theta) = P\left(N_t = 0|I_{t-1}\right)\Gamma\left(\mu_t, \nu\right) + \sum_{m=1}^{\infty} P\left(N_t = m|I_{t-1}\right) \times K\left(m\mu_t, m\varsigma, \nu\right). \tag{17}$$

The conditional distribution of RM_t depends both on the element μ_t as well as on the jump intensity, λ. The expected value of Z_t can then be used to derive the expected value of the realized measure RM_t.[4] Integrating out the dependence on N_t, it is possible to obtain the expected value and the variance of RM_t with respect to the information set I_{t-1} only, see Section 4.2.

[4]See Appendix A.3 for details on the derivation of the moments of Z_t.

4 A persistent MEM-J with time-varying parameters

The main stylized fact that emerges from the empirical analysis of the financial returns is that their volatility is characterized by several dynamic and distributional features. High persistence, leverage effects, clusters of jumps and heteroskedastic effects in volatility are indeed relevant characteristic that must be addressed by a proper model. In this section, we show how the MEM-J presented in Section 3 can account for all these features.

4.1 Specification of μ_t

The role played by the specification of μ_t becomes clear when looking at the dynamics of the model's residuals. As volatility is an highly persistent series characterized by a slow and hyperbolic decay of the autocorrelation function, it becomes clear that a simple ARMA(1,1) specification, as implied by the baseline MEM, is not suited to describe such a rich dynamic behaviour. As a consequence, the model's residuals display significant autocorrelation. A successful and simple approach to capture the (pseudo) long-memory property of the volatility series has been proposed by Corsi (2009) with the HAR model. The HAR is long autoregressive model, subject to linear constraints, designed to capture the persistence of the logarithm of realized volatility. We therefore consider two alternative specifications for μ_t:

- Asymmetric MEM:

$$\mu_t = \omega + \alpha RM_{t-1} + \beta\mu_{t-1} + \gamma RM^-_{t-1} \tag{18}$$

  where $RM^-_t \equiv RM_t \cdot \mathrm{I}\{r_t < 0\}$. This base specification takes the presence of asymmetric responses of volatility to the sign of the returns (see Engle and Gallo, 2006), and the coefficient γ captures a stronger response to negative returns, the well known *leverage* effect. The baseline MEM can be simply obtained setting $\gamma = 0$.

- Asymmetric HAR-MEM (AHAR-MEM), which combines the distinctive elements of the Asymmetric MEM and the HAR

$$\mu_t = \omega + \beta\mu_{t-1} + \alpha_1 RM_{t-1} + \alpha_2 RM_{t-1:t-5} + \alpha_3 RM_{t-1:t-21} + \gamma RM^-_{t-1} \tag{19}$$

  where $RM_{t-1:t-5} = \frac{1}{5}\sum_{j=1}^{5} RM_{t-j}$ and $RM_{t-1:t-21} = \frac{1}{21}\sum_{j=1}^{21} RM_{t-j}$. The restriction $\alpha_2 = \alpha_3 = 0$ gives the Asymmetric MEM. Hence, it is possible to test whether the inclusion of the weekly and monthly volatility terms provide a significant improvement in modeling the volatility dynamics.

4.2 Time-varying jump intensity

The previous specification of the MEM-J is inherently limited given that the Poisson process governing the occurrence of jumps and the Gamma density characterizing the jump size are all driven by time invariant parameters. To increase the model flexibility we introduce time variation in the parameter of the Poisson process. Instead, we maintain a time invariant jump size since preliminary evaluations of the proposed model show that letting the parameter to vary across time does not improve upon time-invariant specifications, but it increases the computational burden associated with the model estimation. Nevertheless, if needed (and supported by the data), even the jump size can be time-varying. We first specify the dynamic evolution of the parameter λ_t, i.e. the jump intensity, for which we suggest the Auto Regressive Jump Intensity (ARJI) specification of Chan and Maheu (2002). The innovation in the jump intensity dynamic are derived from the jump probability as follows:

$$\lambda_t = \phi_1 + \phi_2 \lambda_{t-1} + \phi_3 \xi_{t-1} \tag{20}$$

where

$$\xi_t = \mathbb{E}\left[N_t | I_t\right] - \lambda_t = \sum_{m=0}^{\infty} mP\left(N_t = m | I_t\right) - \lambda_t. \tag{21}$$

The restrictions $\phi_1 > 0$ and $\phi_2 > \phi_3 > 0$ are sufficient to guarantee the positiveness of λ_t as in Chan and Maheu (2002). Note that the innovation term depends on the conditional probabilities of observing m jumps given the information set at time t, and those are determined following the hypothesis of having a Poisson process governing the jumps number, see (8). However, as the conditioning set is different, those probabilities must be appropriately evaluated. We will discuss this issue in Section 4.4 when dealing with the model estimation. From a distributional point of view, letting the mixing parameter λ to be dynamic implies that the conditional density of RM_t in (16) has a time-varying weight associated with the K density. This provides an extremely flexible specification of the density of RM_t, which can be exploited to infer a precise probability of occurrence of tail events, see Section 6.[5]

Proposition 3 *Consider model* (10) *where* $\eta_t = Z_t \varepsilon_t$ *with* Z_t *defined in equation* (7) *and* $\varepsilon_t \sim \Gamma(1, \nu)$. *Assuming that* Z_t *and* ε_t *are independent at all leads and lags, it follows that the first two moments of* RM_t *and* η_t *conditional on* I_{t-1} *are*

$$\mathbb{E}\left[\eta_t | I_{t-1}\right] = \left(e^{-\lambda_t} + \lambda_t\right), \tag{22}$$

$$\mathbb{E}\left[RM_t | I_{t-1}\right] = \mu_t \left(e^{-\lambda_t} + \lambda_t\right), \tag{23}$$

[5] Creal et al. (2013) derives the Generalized Autoregressive Score (GAS) representation for both the time-varying intensity Poisson process and the dynamic mixtures of models. We believe that an extension of the MEM-J model within the GAS framework is a natural advancement but this is left to future investigation.

$$\mathbb{V}\left[\eta_t|I_{t-1}\right] = \left[\frac{\lambda_t}{\varsigma} + \lambda_t^2\right](1+\nu^{-1}) + (e^{-\lambda_t}+\lambda_t)\left[1+\nu^{-1}-e^{-\lambda_t}-\lambda_t\right] \tag{24}$$

$$\mathbb{V}\left[RM_t|I_{t-1}\right] = \mu_t^2\left\{\left[\frac{\lambda_t}{\varsigma} + e^{-\lambda_t} + (\lambda_t+\lambda_t^2)\right](1+\nu^{-1}) - (e^{-\lambda_t}+\lambda_t)^2\right\}. \tag{25}$$

The conditional expected value and variance of RM_t depend on the time-varying mean component as well as on the time-varying jump intensity. We stress that the conditional variance of RM_t is time-varying thanks to the evolution in time of both μ_t and λ_t, thus allowing the MEM-J, similarly to the MEM, to capture the volatility-of-volatility effect studied in Corsi et al. (2008) among others. The above results highlight how the conditional moments of RM_t depend on the marginal moments of the jump term. In fact, the availability of the first and second moments of the jump component potentially allows for the construction of confidence intervals around the impact of the expected jump. For example, compared to the baseline MEM, the conditional expectation of RM_t is inflated by a time-varying factor $\left(e^{-\lambda_t}+\lambda_t\right)$, which is never smaller than one by construction and acts as a state-dependent boosting factor, whose introduction is expected to find strong support in the data.

4.3 Stationarity

We first provide the stationarity conditions for the simple case of time-invariant jump intensity with μ_t specified as

$$\mu_t = \omega + \sum_{i=1}^{q}\alpha_i RM_{t-i} + \sum_{i=1}^{p}\beta_i\mu_{t-i}. \tag{26}$$

Given the multiplicative structure of the MEM-J, the conditions for strictly stationarity of RM_t can be studied writing (26) in vector form (Markov representation):

$$z_t = b_t + A_t z_{t-1}. \tag{27}$$

Theorem 1 *For the MEM-J process in* (6) *with Assumption 1 where η_t is a sequence i.i.d. random variables with intensity parameter λ, there exist a unique strictly stationary solution (which is also weakly stationary) if*

$$(e^{-\lambda}+\lambda)\sum_{j=1}^{q}\alpha_j + \sum_{i=1}^{p}\beta_i < 1. \tag{28}$$

Proof in Appendix A.5.
We also remark that if the MEM-J has a second-order stationary solution and if $\omega > 0$ then condition (28) holds. We note that the introduction of the jump term in the MEM-J is expressed in a different form with respect to the baseline MEM of Engle and Gallo

(2006). We have a multiplicative term, depending on λ, that impacts only on the *ARCH*-related coefficients, those capturing the impact of innovations on the mean-evolution. This is a consequence of the fact that the jump term, Z_t, is a constituent of the model innovations, η_t, and it is not persistent by construction. Interestingly, the larger the coefficient λ, the larger is the inflating factor, and the smaller is the stationarity region given the parameters α_i and β_i.

We now move to the mode complex case of the time-varying jump intensity. In that case, we provide a sufficient condition for the stationarity of the MEM-J. Consider the MEM-J process in (6) with density, conditional to η_t, defined in (16), where η_t is a sequence of random variables with time-varying intensity parameter λ_t defined as in (20), and $A_t = A \odot E_t$.

Theorem 2 *Given the MEM-J in* (6) *with Assumption 1 and the processes for* λ_t *and* μ_t *specified as in* (20) *and* (26)*, respectively, a sufficient condition for the existence of a strictly stationary solution is*

$$\rho(A) < \exp(-E[\log [(p+q)(\eta_t + (p+q) - 1)]])$$

where $\rho(A)$ *is the spectral radius of* A *(i.e. the greatest modulus of its eigenvalues).*

This second result is less intuitive than in the constant intensity case. Nevertheless, we stress the sufficient condition depends both on the number of dynamic parameters affecting μ_t and on the expectation of the innovation η_t that combined the jump term and the error term ϵ_t.

4.4 Maximum likelihood estimation

The AHAR-MEM-J can be estimated by maximum likelihood. Under the maintained assumption that $N_t|I_{t-1} \sim Poisson(\lambda_t)$ with N_t and ε_t independent processes, the conditional density of RM_t, $f\left(RM_t|I_{t-1}\right)$, can be computed in closed form as in equation (17).[6] Model parameters are estimated by maximizing the sample log-likelihood $\ell(\theta) = \sum_{t=1}^{T} \log f\left(RM_t|I_{t-1};\theta\right)$, where $f\left(RM_t|I_{t-1}\right)$ is defined in equation (17) and $\theta \in \Theta$ is the vector of parameters for the AHAR-MEM-J with time-varying λ, i.e. $\theta = [\omega, \beta, \alpha_1, \alpha_2, \alpha_3, \gamma_1, \gamma_2, \gamma_3, \phi_1, \phi_2, \phi_3, \varsigma, \nu]'$. The log-likelihood function $\ell(\theta)$ is the log-trasform of a mixture density. In general, the mixture likelihood function can be unbounded, that is the function is characterized by the presence of singularities. Thus the ML estimator as global maximizer of the mixture likelihood function does not exist. Nevertheless, statistical theory outlined in Kiefer (1978) guarantees that a particular local maximizer of

[6] The conditional density of RM_t involves an infinite sum of densities that depends on the number of jumps. Therefore, a truncation on the maximum number of jumps $\bar{m} < \infty$ is required in practical applications. See Section 4.5 for a discussion of the choice of $\bar{m}$.

the mixture likelihood function is consistent, efficient, and asymptotically normal if the mixture is not overfitting. Several local maximizers may exist for a given sample, and a major difficulty with the ML approach is to identify if the correct one has been found, see Frühwirth-Schnatter (2006).

In the case of estimation of MixNormal-GARCH models, Ausín and Galeano (2007) and Bauwens et al. (2007) have devised a Bayesian estimation procedures to avoid such degenerated states. Alternatively, Broda et al. (2013) have proposed an augmented likelihood function. We don't adopt any of these computational devices since the Monte Carlo results, reported below, show that this problem is not a major concern in our case.

The information used to evaluate the likelihood function involves the computation of further quantities that drive the dynamics of the jump intensity, λ_t. From the Bayes rule, it follows that the filtered jumps probabilities are equal to

$$P\left(N_t = m|I_t\right) = \frac{P\left(RM_t|N_t = m, I_{t-1}\right) \times P\left(N_t = m|I_{t-1}\right)}{P\left(RM_t|I_{t-1}\right)}, \quad j = 0, 1, 2, \dots \tag{29}$$

which are then used to recover the jump intensity innovations in equation (21).

4.5 Monte Carlo Simulations

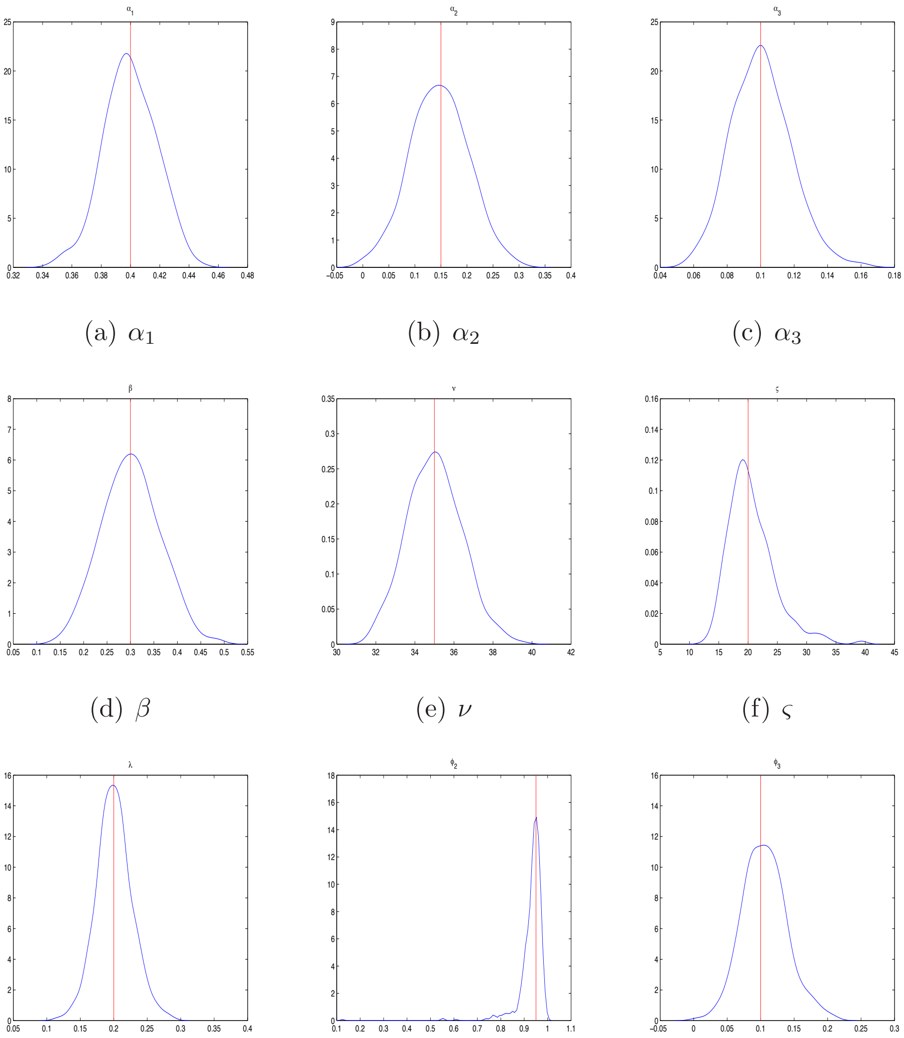
We run a set of Monte Carlo simulations to explore the performance of the ML estimation of MEM-J in finite samples. We simulate three different specifications with the same μ_t as in (19) with no asymmetric effect: HAR-MEM (model in (1)), HAR-MEM-J with constant λ (model in (6)) and HAR-MEM-J with time-varying λ_t (model in (6) with λ_t specified as in (20)). The algorithm to simulate pseudo-random variates from a K density is illustrated in Appendix B. The parameters used in simulation and the results are reported in Table 1. The simulated sample size is set equal to 3000. Due to the computational burden in estimating the MEM-J the Monte Carlo replications are 500. We investigate the effects that the over-specification of the jump component can have on the maximum likelihood estimates. This can be a typical situation which arises when we have to specify nonlinear models with latent components. We estimate over-specified models (upper and middle panel of Table 1), i.e. the HAR-MEM-J with time-varying jump intensity, when the data have been generated with either λ equal to zero (i.e. $\phi_1 = \phi_2 = \phi_3 = 0$) or with a constant λ. The infinite sum of densities required to compute the likelihood, see (15), is truncated at $\bar{m} = 10$. For values of λ that are rarely larger than 1, the probability of observing more than 10 jumps is of order 10^{-8}. When the jumps are totally absent, the unconditional mean of $\lambda_t = \frac{\phi_1}{1-\phi_2}$ is estimated almost equal to zero, meaning that there is a very limited mixing effect in the conditional density of RM_t due to the estimated jump term. This means that the estimated model is very close to the HAR-MEM which is the DGP. Indeed, the parameters governing μ_t are estimated correctly and with small

RMSE. It should also be noted that the parameter ς is not defined under the DGP, but is estimated when fitting the HAR-MEM-J on the data as it determines the shape of the K distribution. When jumps are absent, i.e. $\lambda = 0$, the parameter ς is not identified, this is reflected in a very high RMSE.

When the jumps are present, but λ is constant, the impact of the over-specification is fairly limited, see the middle panel of Table 1, and the distribution of all parameters is centered on the true values. The estimate of the parameter ϕ_3 is close to zero with a quite large RMSE. This evidence also confirms that the HAR-MEM-J with time varying jump intensity is identified also when λ is constant and the estimated model is overspecified. This means that the sources of variation in $\mathbb{E}\left[RM_t|I_{t-1}\right]$, i.e. μ_t and λ_t, are separately identified. In fact, if we look at the estimates of the HAR parameters, they seem unaffected by this misspecification. In the third case considered, which is the estimate of the correctly specified model, the maximum likelihood estimates have a very small finite sample bias and the RMSE's of $\phi_1/(1-\phi_2)$, ϕ_2 and ϕ_3 have the same order of magnitude of the other parameters. In Figure 1, the kernel density estimates of the Monte Carlo distribution are displayed. The plots show that the finite sample distributions for all parameters are centred on the true values. Furthermore, the Monte Carlo estimates based on the mixture density are well behaved and we don't find any evidence of the presence of multiple local maxima.

	ω	α_1	α_2	α_3	β	ν	ς	$\frac{\phi_1}{1-\phi_2}$	ϕ_2	ϕ_3
DGP: $\lambda = 0$										
	0.001	**0.4**	**0.15**	**0.1**	**0.3**	**20**	**0**	**0**	**0**	**0**
Mean	0.001	0.399	0.148	0.097	0.302	20.106	38.010	0.009	0.599	0.254
RMSE	0.000	0.021	0.069	0.023	0.076	0.577	40.972	0.033	0.636	0.334
DGP: $\lambda = 0.25$										
	0.001	**0.4**	**0.15**	**0.1**	**0.3**	**35**	**20**	**0.25**	**0**	**0**
Mean	0.001	0.400	0.152	0.099	0.296	34.957	20.625	0.250	0.479	0.019
RMSE	0.000	0.017	0.050	0.017	0.056	1.646	3.710	0.018	0.611	0.033
DGP: $\lambda_t > 0$										
	0.001	**0.4**	**0.15**	**0.1**	**0.3**	**35**	**20**	**0.2**	**0.95**	**0.1**
Mean	0.001	0.400	0.147	0.100	0.301	35.011	20.620	0.201	0.931	0.106
RMSE	0.000	0.018	0.056	0.018	0.062	1.394	4.051	0.027	0.060	0.034

Table 1: Monte Carlo results. The true parameter values used in simulation are in bold. Sample mean and Root mean squared error (RMSE) of maximum likelihood estimates of simulated HAR-MEM's models.



(a) α_1 (b) α_2 (c) α_3

(d) β (e) ν (f) ς

(g) ϕ_1 (h) ϕ_2 (i) ϕ_3

Figure 1: Kernel densities of the Monte Carlo estimates of the MEM-J-λ_t, where λ_t varies according to (20).

5 Empirical Results

5.1 Database and estimation of volatility

Our purpose is to model realized measures to estimate the probability and the size of the volatility jumps once that *price jumps* have been disentangled from the volatility dynamics. Indeed, when price jumps are present, the total price variation, or *quadratic variation*, is equal to the sum of integrated volatility plus the squared price jumps. The quadratic variation can be estimated by the sum of the intraday squared returns, $r^2_{t_j}$, i.e.

the realized volatility (or realized variance), RV,

$$RV_t = \sum_{j=1}^{M} (p_{t_j} - p_{t_{j-1}})^2 = \sum_{j=1}^{M} r_{t_j}^2 \quad t = 1, ..., T \tag{30}$$

where $r_{t_j} \equiv p_{t_j} - p_{t_{j-1}}$ is the intraday return and M is the number of intraday observations. The realized volatility converges to the integrated variance plus the squared jump component. Barndorff-Nielsen and Shephard (2004), have shown that RV allows for a direct nonparametric decomposition of the total price variation into its two separate components: a continuous part, called *Bipower Variation* (BPV), and a discontinuous one, the squared price jumps. Disentangling the price jumps is important because, as it has been noted by Huang and Tauchen (2005), their relative contribution to the total price variability is about 7%. The BPV is defined as

$$BPV_t = \frac{\pi}{2} \sum_{j=2}^{M} |r_{t_j}||r_{t_{j-1}}| \quad t = 1, ..., T \tag{31}$$

and converges to the integrated variance as M diverges. As M increases, the bipower variation will converge to integrated continuous volatility, which is likely to be affected by volatility jumps, as shown, for instance by Caporin et al. (2014). Hence, the following empirical analysis will be based on the estimation of alternative MEM specifications on the BPV series. It should be noted that, in finite samples, any volatility estimator, including the BPV, is contaminated by a measurement bias, but it is assumed that the measurement error can be considered negligible for the purposes of this paper.

The empirical analysis is based on two data sets. The first one includes the bipower variation of seven stock indexes: S&P500, FTSE 100, DAX, DJIA, NASDAQ 100, CAC 40, Bovespa, sampled from January 3, 2000 through January 31, 2013, as made available by the Oxford-Man Institute's Realised Library. The second data set employed consists of the intradaily returns of 16 large cap equities quoted on the New York market: Boeing, Bank of America, City Group, Caterpillar, Federal Express, Honeywell, Hewlett-Packard, IBM, JP Morgan, Kraft, Pepsi, Procter & Gamble, AT&T, Time Warner, Texas Instruments, and Wells Fargo. Prices are sampled at one minute frequency, from January 2, 2003 to June 30, 2012, and they are provided by TickData. The bipower variation is estimated from the 1-minute prices. In both datasets, the realized measure is expressed as daily volatility, i.e. the square root of the bipower variation, $RM_t = \sqrt{BPV_t}$.

Table A.1 in the Web Appendix reports descriptive statistics of $\sqrt{BPV_t}$ for the seven indexes. We observe well-known stylized facts such as high kurtosis and asymmetry, due to the long upper tail characterizing the empirical density of bipower variation time series, and the presence of a strong serial correlation as suggested by the very high values of the auto-correlations at the selected lags 1, 5 and 22.

Several alternative multiplicative specifications are considered for characterizing the dynamic behaviour of the square root of the bipower variation of the stock indexes and individual S&P 500 stocks. We estimate the simple AMEM and the AHAR-MEM in order to evaluate if there is empirical support for a HAR specification in μ_t. For what concerns the multiplicative models with jumps, we consider the following cases:

- Constant jump intensity: from equation (20) we set $\lambda_t = \phi_1$ and $\phi_2 = \phi_3 = 0$;
- Time-varying jump intensity: with λ_t evolving as in equation (20).

To compare the alternative models we consider two different approaches. Firstly, we pursuit a full-sample evaluation approach, where the MEM and MEM-J specifications are compared with respect to their fit on the empirical data and a series of statistical tests for restrictions on the parameters are performed. Secondly, we evaluate model abilities in capturing the behaviour of the upper tail of the realized measure both in-sample and out-of-sample. This is not only crucial for risk-management purposes, but also consistent with the expected ability of the MEM-J in explaining sudden increases in the volatility that cause observations of the realized measures located in the upper tail.

5.2 Full-sample model comparison

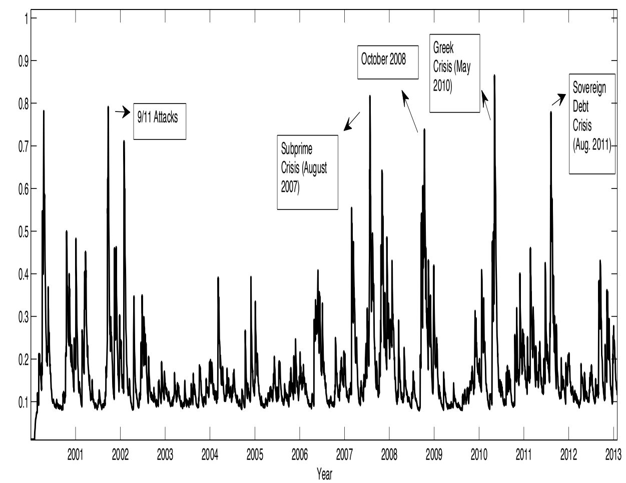
The different specifications adopted for the estimation of the BPV dynamic behavior are first compared in terms of their ability in capturing the dynamics of the series. To this end, we analyze the dynamic properties of the residuals

$$e_t^* = \frac{RM_t}{\mathbb{E}\left[RM_t|I_{t-1}\right]}. \tag{32}$$

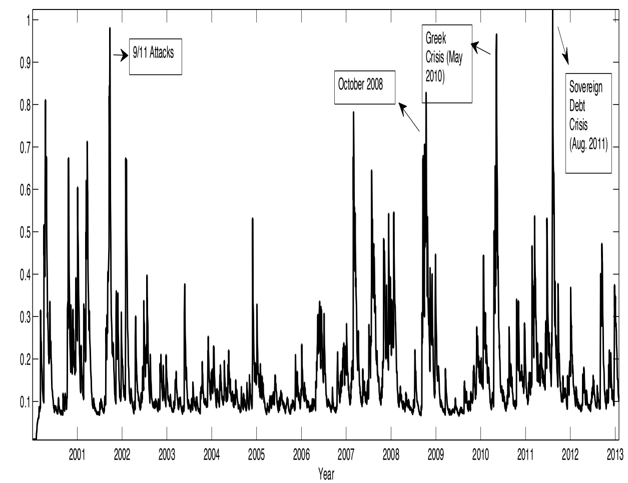
The residuals obtained from different model specifications might be compared in terms of density, moments, as well as with respect to the presence of serial correlation in the first and second order moments.[7] Tables from 2 to 5 report the parameter estimates and the residual statistics. The standardized residuals, $\widehat{\epsilon}_t$, are obtained by the transformation $\widehat{\epsilon}_t^* = F_N^{-1}\left[F_\Gamma(\widehat{\epsilon}_t)\right]$, for $t = 1, \ldots, n$, where $F_N()$ and $F_\Gamma()$ are the cumulative density functions of the standard normal and Gamma distributions, respectively. Since the standard diagnostic statistics and graphs are designed for residuals that are assumed normal distributed and since the model disturbances are assumed to come from a Gamma density, this transformation for the residuals is justified. From Tables 2 and 3, which report estimation results for the Asymmetric MEM and the Asymmetric HAR-MEM respectively, it

[7] As an alternative, model residuals might be computed by standardization of RM_t with respect to its expected value, involving the impact of μ_t and (when present) Z_t. In this case, innovations are defined as Pearson's residuals

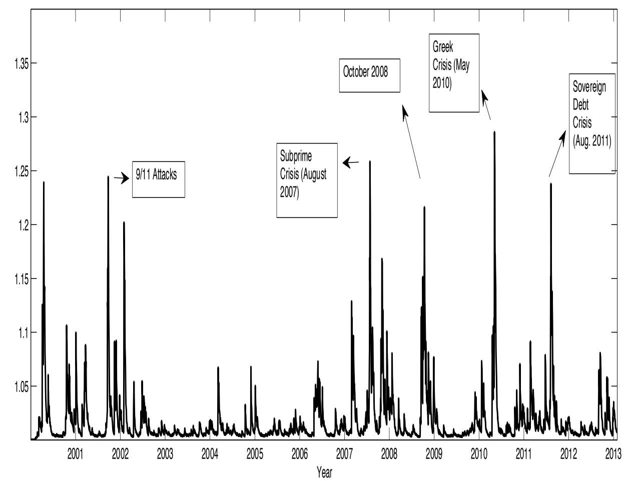
$$\hat{\varepsilon}_t = \frac{RM_t - \mathbb{E}\left[RM_t|I_{t-1}\right]}{\mathbb{V}\left[RM_t|I_{t-1}\right]^{1/2}}. \tag{33}$$



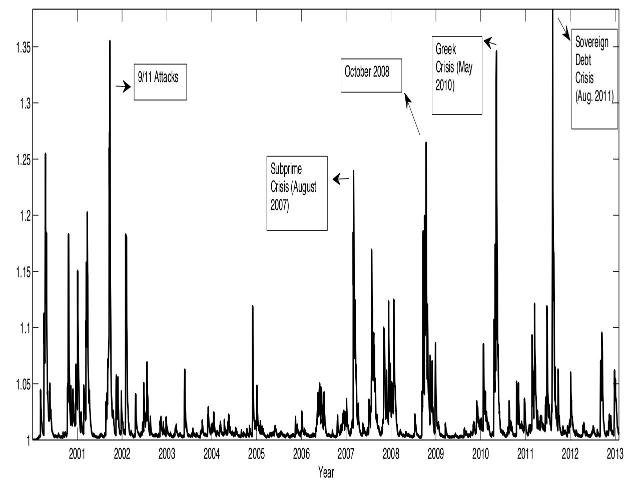
(a) λ_t of S&P500



(b) λ_t of CAC



(c) $E(Z_t|I_{t-1})$ of S&P500



(d) $E(Z_t|I_{t-1})$ of CAC

Figure 2: Jump intensity and expected jump component from AHAR-MEM-J-(λ_t) model for S&P500 and CAC 40.

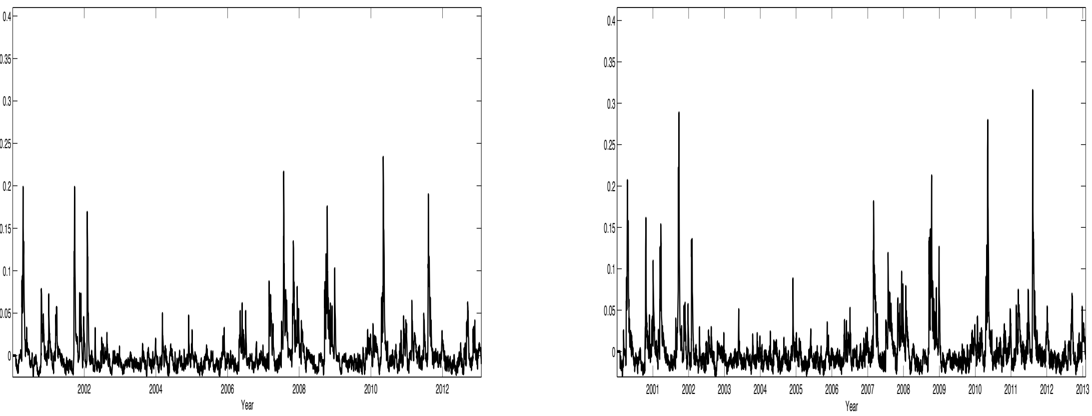
	ω	α	β	γ	ν	$LogLL$	Q_1	Q_{10}	Q_{22}
SP500	0.0003^a	0.3278^a	0.5944^a	0.0895^a	15.9492^a	15673	0.004	0.000	0.000
FTSE 100	0.0002^a	0.3120^a	0.6393^a	0.0556^a	16.6881^a	16104	0.001	0.000	0.000
DAX	0.0003^a	0.3247^a	0.6112^a	0.0703^a	17.4627^a	15202	0.000	0.000	0.000
DJIA	0.0003^a	0.3234^a	0.6037^a	0.0814^a	15.6513^a	15742	0.003	0.000	0.000
NSDQ	0.0003^a	0.3404^a	0.5915^a	0.0759^a	16.1353^a	15199	0.001	0.000	0.000
CAC	0.0002^a	0.3051^a	0.6327^a	0.0782^a	17.8968^a	15503	0.000	0.000	0.000
BOVESPA	0.0010^a	0.3561^a	0.5402^a	0.0562^a	15.1477^a	13829	0.007	0.000	0.000
BA	0.0005^a	0.3618^a	0.5859^a	0.0326^a	22.4478^a	10559	0.0387	0.0003	0.0002
BAC	0.0003^a	0.5227^a	0.4376^a	0.0479^a	19.8641^a	10198	0.0832	0.0000	0.0005
C	0.0002^a	0.5214^a	0.4535^a	0.0285^a	24.6862^a	10049	0.0370	0.0000	0.0005
CAT	0.0007^a	0.4162^a	0.5206^a	0.0403^a	22.8666^a	10340	0.0046	0.0000	0.0020
FDX	0.0005^a	0.3995^a	0.5435^a	0.0355^a	21.1708^a	10516	0.0227	0.0000	0.0000
HON	0.0008^a	0.4063^a	0.5172^a	0.0506^a	23.0828^a	10502	0.0516	0.0022	0.0033
HPQ	0.0007^a	0.4000^a	0.5302^a	0.0452^a	20.6334^a	10335	0.0390	0.0000	0.0004
IBM	0.0005^a	0.4420^a	0.4968^a	0.0371^a	25.3562^a	11303	0.0968	0.0071	0.0260
JPM	0.0004^a	0.4864^a	0.4731^a	0.0366^a	22.4146^a	10323	0.0039	0.0000	0.0000
KFT	0.0005^a	0.3565^a	0.5880^a	0.0106^a	16.5042^a	10845	0.0030	0.0000	0.0003
PEP	0.0003^a	0.3389^a	0.6153^a	0.0310^a	22.4093^a	11373	0.3856	0.0229	0.0003
PG	0.0004^a	0.4130^a	0.5288^a	0.0279^a	23.2476^a	11544	0.0953	0.0005	0.0003
T	0.0003^a	0.3980^a	0.5617^a	0.0253^a	25.2941^a	10981	0.0234	0.0000	0.0001
TWX	0.0005^a	0.4032^a	0.5499^a	0.0287^a	28.8549^a	10672	0.0698	0.0024	0.0044
TXN	0.0006^a	0.3546^a	0.5984^a	0.0270^a	25.2586^a	10120	0.0527	0.0018	0.0006
WFC	0.0002^a	0.4590^a	0.5184^a	0.0158	18.9402^a	10255	0.4878	0.0000	0.0002

Table 2: Estimates of the baseline AMEM (see (18)). The upper part of the table reports the results for several stock indexes while the lower part refers to 16 NYSE stocks. a, b and c stand for significance at 1%, 5% and 10% respectively. Q_1, Q_{10} and Q_{22} are the p-values of the Ljung-Box test for absence of autocorrelation in the residuals, where the latter are computed as $\hat{\epsilon}_t = \frac{RM_t}{E[RM_t|I_{t-1}]}$.

is evident that the Asymmetric MEM is unable to capture the persistence that is present in the bipower variation time series. On the contrary, in the case of the Asymmetric HAR-MEM the Ljung-Box statistics does not reject the null of no residual autocorrelation in 4 out of the 7 stock indexes considered, and only when we focus on lags up to the 22-nd. Looking at the individual stocks, at the 5% confidence level we have only 4 out of the 16 equities with some evidences of residual serial correlation, and only over 22 lags. The number of stocks with serial correlation in the residuals of the AHAR-MEM decreases to 1 if we take the 1% confidence level. If we compare the estimated parameters of the stocks to those of the indexes, we note the following: stocks are characterized by a somewhat higher impact of previous day bipower variation levels. Differently, last week and last month average bipower variations, have a more heterogeneous impact, with some cases of reduced significance. Heterogeneity in the single assets might be expected. Finally, we observe how the innovation parameter, ν, is sensibly higher than that of the indexes. This reflects the differences in the density, where the volatility of volatility of the indexes is higher than that of the individual stocks. This may be due to the different sample

	ω	α_1	α_2	α_3	β	γ	ν	$LogLL$	Q_1	Q_{10}	Q_{22}	LR_{α_2,α_3}
SP500	0.0003^a	0.3081^a	0.1670^a	0.0980^a	0.3297^a	0.1129^a	16.2920^a	15708	0.9042	0.4782	0.3882	70.29^a
FTSE 100	0.0002^a	0.3218^a	0.2458^a	0.1246^a	0.2396^a	0.0684^a	16.9945^a	16135	0.9106	0.6457	0.0119	60.46^a
DAX	0.0004^a	0.3331^a	0.1294^a	0.1103^a	0.3506^a	0.0817^a	17.7801^a	15232	0.2201	0.1974	0.0000	60.40^a
DJIA	0.0004^a	0.3120^a	0.1962^a	0.0928^a	0.3052^a	0.1013^a	15.9279^a	15771	0.9133	0.4444	0.5854	57.96^a
NSDQ	0.0004^a	0.3205^a	0.1844^a	0.1397^a	0.2633^a	0.1038^a	16.6296^a	15249	0.7336	0.1068	0.0944	99.88^a
CAC	0.0003^a	0.3216^a	0.1508^a	0.0991^a	0.3541^a	0.0877^a	18.1808^a	15530	0.3166	0.0889	0.0000	53.05^a
B0VESPA	0.0011^a	0.3495^a	0.1634^a	0.1552^a	0.2089^a	0.0724^a	15.5308^a	13870	0.7588	0.6229	0.2247	80.81^a
BA	0.0006^a	0.3840^a	0.0735	0.1183^a	0.3669^a	0.0338^a	22.8003^a	10578	0.7975	0.7728	0.0449	37.44^a
BAC	0.0003^a	0.5316^a	0.0457	0.1297^a	0.2551^a	0.0442^a	20.3631^a	10228	0.7118	0.6958	0.6892	59.75^a
C	0.0002^a	0.5230^a	0.0941^c	0.1013^a	0.2548^a	0.0281^a	25.1744^a	10072	0.6971	0.7133	0.2519	47.02^a
CAT	0.0009^a	0.4127^a	0.2341^a	0.1009^a	0.1706^b	0.0475^a	23.2921^a	10362	0.8485	0.6575	0.8894	44.30^a
FDX	0.0006^a	0.4181^a	0.0834^c	0.1635^a	0.2770^a	0.0339^a	21.7095^a	10546	0.7435	0.1739	0.4699	60.45^a
HON	0.0008^a	0.4075^a	0.1281^c	0.0744^a	0.3052^a	0.0544^a	23.3126^a	10514	0.7512	0.8410	0.1363	23.80^a
HPQ	0.0009^a	0.4007^a	0.1822^a	0.1222^a	0.2076^a	0.0577^a	21.1230^a	10363	0.7678	0.4510	0.4235	56.44^a
IBM	0.0005^a	0.4445^a	0.1275^b	0.0629^a	0.2973^a	0.0401^a	25.5941^a	11314	0.9893	0.9638	0.5213	22.40^a
JPM	0.0005^a	0.4918^a	0.2013^a	0.1077^a	0.1527^b	0.0349^a	22.9204^a	10349	0.6510	0.8022	0.3092	53.64^a
KFT	0.0009^a	0.3813^a	0.2703^a	0.1829^a	0.0772	0.0025	16.9128^a	10875	0.6344	0.5964	0.9161	59.10^a
PEP	0.0004^a	0.3507^a	0.0655	0.0678^b	0.4645^a	0.0319^a	22.5384^a	11380	0.6252	0.3843	0.0052	13.83^a
PG	0.0006^a	0.4099^a	0.2367^a	0.0916^a	0.1841^a	0.0354^a	23.6511^a	11565	0.4232	0.9140	0.0508	41.37^a
T	0.0004^a	0.4151^a	0.1219^b	0.1062^a	0.3077^a	0.0298^a	25.6652^a	10999	0.8620	0.6676	0.0327	34.94^a
TWX	0.0006^a	0.4109^a	0.1384^b	0.1129^a	0.2815^a	0.0327^a	29.3522^a	10692	0.7001	0.9871	0.4933	40.95^a
TXN	0.0008^a	0.3661^a	0.1848^b	0.1288^a	0.2606^a	0.0334^a	25.6894^a	10140	0.4013	0.8918	0.1558	40.59^a
XOM	0.0004^a	0.4773^a	0.2018^a	0.1539^a	0.1374	0.0146	19.4596^a	10287	0.0834	0.4063	0.3548	65.20^a

Table 3: Estimates of the Asymmetric HAR-MEM (see (19)). The upper part of the table reports the results for several stock indexes while the lower part refers to 16 NYSE stocks. a, b and c stand for significance at 1%, 5% and 10% respectively. Q_1, Q_{10} and Q_{22} are the p-values of the Ljung-Box test for absence of autocorrelation in the residuals, where the latter are computed as $\hat{\epsilon}_t = \frac{RM_t}{E[RM_t|I_{t-1}]}$. LR_{α_2,α_3} is the likelihood ratio test for the nullity of α_2 and α_3.



(a) $S\&P500$ (b) CAC

Figure 3: Relative expected jump contribution. The figures display the ratio $\frac{\mathbb{E}_J[RM_t|I_{t-1}]-\mathbb{E}_0[RM_t|I_{t-1}]}{\mathbb{E}_0[RM_t|I_{t-1}]}$, where $\mathbb{E}_J[RM_t|I_{t-1}]$ is the conditional expectation under AHAR-MEM-J-λ_t, and $\mathbb{E}_0[RM_t|I_{t-1}]$ is the conditional expectation under AHAR-MEM.

periods under exam. Given these results, in the rest of the paper we will consider only the AHAR specification.

As it emerges from the theoretical analysis in Section 3, the MEM-J specification is well suited to provide a high degree of flexibility for the conditional moments of RM_t. The comparison of AHAR-MEM and AHAR-MEM-J specifications starts off from the simple evaluation of parameters significance. In fact, obtaining significant coefficients for the jump intensity, either in the constant or dynamic specification, might be seen as a first evidence that jumps in volatility are a significant component of the stock indexes bipower variation. With regard to this aspect, Tables 3 and 4 allow for a first comparison. By looking at Table 4, we observe that for all series the parameters associated with the jumps, ς and λ, are all statistically significant, thus supporting the potential relevance of the jump component Z_t. In addition, the parameters driving the evolution of μ_t are close to those of the baseline AHAR-MEM specification in Table 3, with the exception of the innovation term, whose scale parameter ν is characterized by a sensible increase. As expected, this confirms that accounting for the presence of a jump component in the bipower variation affects the estimate of the parameters of the innovation term distribution. Since the nuisance parameter ς is not defined in the AHAR-MEM, it is not possible to evaluate the significance of the jump term by a standard LR test, see the discussion in Hansen (1996).[8] However, from a comparison of the values in Tables 3 and 4, it clearly emerges that the log-likelihood functions of the AHAR-MEM-J are much larger than those of the AHAR-MEM, as their difference is often larger than 100.

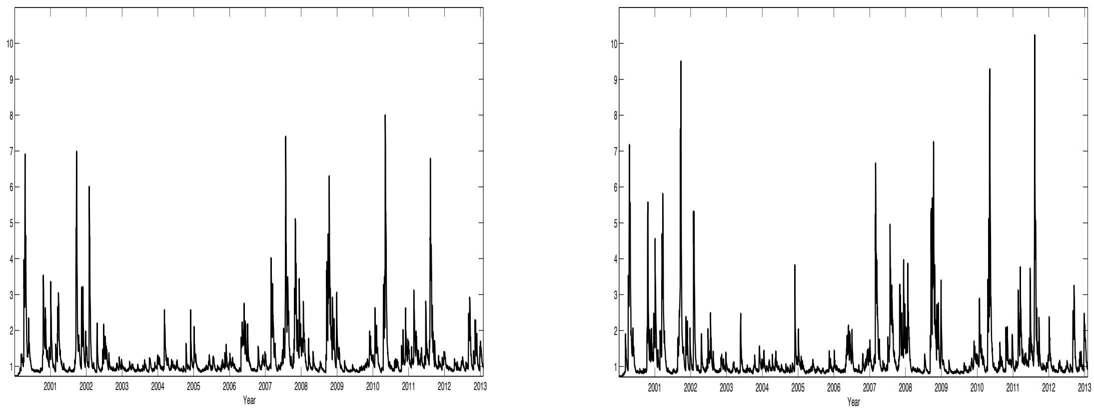
[8] In this case simulation based approaches can be used to recover likelihood ratio test critical values following Hansen (1996). We don't pursue that strategy due to the computational burden implied in the estimation of the MEM-J.

	ω	α_1	α_2	α_3	β	γ	ν	ς	ϕ_1	$LogLL$	Q_1	Q_{10}	Q_{22}
SP500	0.0003^a	0.3022^a	0.1675^a	0.1129^a	0.3129^a	0.1086^a	22.1149^a	17.9948^a	0.1791^a	15789	0.7856	0.4161	0.3145
FTSE 100	0.0002^a	0.2943^a	0.2179^a	0.1604^a	0.2595^a	0.0591^a	26.2087^a	10.4352^a	0.1498^a	16313	0.1217	0.6439	0.0016
DAX	0.0003^a	0.3037^a	0.1427^a	0.1317^a	0.3508^a	0.0738^a	25.1605^a	13.4345^a	0.1643^a	15310	0.0040	0.0420	0.0000
DJIA	0.0003^a	0.3001^a	0.1844^a	0.1060^a	0.3097^a	0.0961^a	21.6582^a	15.9597^a	0.1719^a	15852	0.4610	0.3420	0.4822
NSDQ	0.0003^a	0.3070^a	0.1947^a	0.1502^a	0.2568^a	0.1019^a	22.5996^a	10.9389^a	0.1371^a	15323	0.4407	0.1010	0.0563
CAC40	0.0003^a	0.2983^a	0.1551^a	0.1171^a	0.3556^a	0.0782^a	27.2816^a	11.6089^a	0.1678^a	15623	0.0130	0.0475	0.0000
BOVESPA	0.0009^a	0.3216^a	0.1287^b	0.1673^a	0.2559^a	0.0704^a	22.1820^a	20.2391^a	0.2233^b	13946	0.0860	0.2011	0.0660
BA	0.0006^a	0.3841^a	0.0831	0.1486^a	0.3134^a	0.0251^a	35.5468^a	20.2166^a	0.1970^a	10670	0.6809	0.9054	0.0349
BAC	0.0003^a	0.4990^a	0.1000^c	0.1368^a	0.2114^a	0.0326^a	30.7986^a	31.7204^a	0.1909^a	10406	0.2802	0.8761	0.7291
C	0.0003^a	0.4749^a	0.1188^b	0.1075^a	0.2585^a	0.0201^a	40.2461^a	16.9422^a	0.1743^a	10176	0.0187	0.0695	0.0834
CAT	0.0008^a	0.4080^a	0.2081^a	0.1387^a	0.1603^a	0.0417^a	32.8795^a	25.9484^a	0.1751^a	10445	0.6661	0.5964	0.7196
FDX	0.0005^a	0.3871^a	0.1284^a	0.2018^a	0.2180^a	0.0349^a	31.7331^a	33.7029^a	0.2088^a	10648	0.8995	0.2012	0.3635
HON	0.0007^a	0.3840^a	0.0838	0.1116^a	0.3364^a	0.0498^a	31.9527^a	29.6532^a	0.1734^a	10590	0.1232	0.4163	0.0273
HPQ	0.0008^a	0.4055^a	0.1367^a	0.1519^a	0.2146^a	0.0432^a	33.4704^a	17.5766^a	0.1750^a	10506	0.9266	0.6977	0.3116
IBM	0.0006^a	0.4173^a	0.1476^b	0.0866^a	0.2700^a	0.0314^a	35.7822^a	20.8649^a	0.1508^a	11385	0.2993	0.7714	0.3067
JPM	0.0005^a	0.4543^a	0.1842^a	0.1355^a	0.1648^a	0.0320^a	31.7622^a	34.5964^a	0.1889^a	10434	0.0524	0.2157	0.1139
KFT	0.0011	0.3201^a	0.2528^a	0.2212^a	0.0697^a	0.0002	28.5538^a	35.8665^a	0.2717^a	11061	0.1221	0.3387	0.6590
PEP	0.0004^a	0.3140^a	0.0953	0.1163^a	0.4086^a	0.0293^a	33.3037^a	27.9358^a	0.1862^a	11503	0.5895	0.3141	0.0028
PG	0.0006^a	0.3817^a	0.2081^a	0.1253^a	0.1891^a	0.0319^a	36.3629^a	16.8767^a	0.1639^a	11677	0.6694	0.8454	0.0130
T	0.0005^a	0.3915^a	0.1344^a	0.1321^a	0.2758^a	0.0255^a	36.8422^a	21.6597^a	0.1536^a	11094	0.5551	0.6199	0.0105
TWX	0.0006^a	0.3923^a	0.1262^b	0.1486^a	0.2651^a	0.0280^a	44.0405^a	27.6775^a	0.1605^a	10827	0.6941	0.8620	0.3519
TXN	0.0009^a	0.3758^a	0.1710^b	0.1384^a	0.2450^a	0.0291^a	32.8834^a	27.8139^a	0.1362^a	10194	0.2234	0.7120	0.1184
WFC	0.0004^a	0.4490^a	0.1759^b	0.1586^a	0.1598^c	0.0178^b	33.9598^a	20.8867^a	0.2171^a	10485	0.3035	0.7847	0.5023

Table 4: Estimates of the Asymmetric HAR-MEM-J. The upper part of the table reports the results for several stock indexes while the lower part refers to 16 NYSE stocks. a, b and c stand for significance at 1%, 5% and 10% respectively. Q_1, Q_{10} and Q_{22} are the p-values of the Ljung-Box test for absence of autocorrelation in the residuals, where the latter are computed as $\hat{\epsilon}_t = \frac{RM_t}{E[RM_t|I_{t-1}]}$.

	ω	α_1	α_2	α_3	β	γ	ν	ς	$\frac{\phi_1}{1-\phi_2}$	ϕ_2	ϕ_3	$LogLL$	Q_1	Q_{10}	Q_{22}	LR_{ϕ_2,ϕ_3}
SP500	0.0003^a	0.2913^a	0.1694^a	0.1182^a	0.3162^a	0.1061^a	22.9479^a	13.9452^a	0.1757^a	0.9408^a	0.1930^a	15808	0.3521	0.2656	0.2197	38.14
FTSE 100	0.0002^a	0.2748^a	0.2189^a	0.1648^a	0.2705^a	0.0564^a	26.5095^a	9.7599^a	0.1477^a	0.6411^c	0.3950^b	16335	0.0099	0.2004	0.0003	45.66
DAX	0.0002^a	0.2886^a	0.1415^b	0.1356^a	0.3653^a	0.0699^a	26.2173^a	12.1181^a	0.1760^a	0.9193^a	0.2146^a	15330	0.0007	0.0021	0.0000	40.44
DJIA	0.0003^a	0.2868^a	0.1688^a	0.1127^a	0.3307^a	0.0946^a	22.4206^a	12.9378^a	0.1761^a	0.9489^a	0.2248^a	15875	0.1050	0.1101	0.2587	47.00
NSDQ	0.0003^a	0.3003^a	0.1888^a	0.1587^a	0.2597^a	0.0973^a	23.6406^a	10.2317^a	0.1585^a	0.9634^a	0.1665^a	15342	0.1816	0.0700	0.0283	38.09
CAC	0.0002^a	0.2774^a	0.1365^b	0.1152^a	0.3978^a	0.0770^a	27.5657^a	10.8710^a	0.1630^a	0.9012^a	0.2313^a	15652	0.0001	0.0015	0.0000	56.81
BOVESPA	0.0009^a	0.3127^a	0.1226^b	0.1705^a	0.2687^a	0.0700^a	22.6130^a	17.4224^a	0.2147^a	0.3461^a	0.3455^a	13960	0.0232	0.0766	0.0298	27.41
BA	0.0006^a	0.3664^a	0.0893	0.1424^a	0.3322^a	0.0245^a	35.8425^a	19.0657^a	0.1904^a	0.8863^a	0.1048^b	10683	0.7242	0.9225	0.0431	26.30
BAC	0.0004^a	0.4785^a	0.0937^c	0.1371^a	0.2334^a	0.0346^a	33.6804^a	19.4741^a	0.1945^a	0.9814^a	0.0986^a	10427	0.0762	0.6096	0.5985	40.52
C	0.0003^a	0.4483^a	0.1164^b	0.1069^a	0.2876^a	0.0218^a	41.6322^a	13.6892^a	0.1541^a	0.9911^a	0.1105^a	10230	0.0004	0.0028	0.0112	106.29
CAT	0.0008^a	0.3901^a	0.2122^a	0.1395^a	0.1718^a	0.0419^a	33.8332^a	21.4409^a	0.1696^a	0.5204	0.2761^b	10457	0.2384	0.4199	0.6357	23.97
FDX	0.0005^a	0.3742^a	0.1433^a	0.1946^a	0.2224^a	0.0348^a	33.6948^a	22.4715^a	0.2008^a	0.8536^a	0.1597^a	10660	0.5496	0.1314	0.3446	25.09
HON	0.0007^a	0.3665^a	0.0947	0.1054^a	0.3505^a	0.0493^a	33.8111^a	19.3674^a	0.1631^a	0.9065^a	0.2174^a	10614	0.0257	0.2464	0.0204	48.14
HPQ	0.0008^a	0.3892^a	0.1621^a	0.1445^a	0.2137^a	0.0437^a	33.9732^a	17.4021^a	0.1697^a	0.8364^a	0.1549^a	10537	0.7540	0.6728	0.3541	63.04
IBM	0.0006^a	0.3947^a	0.1263^c	0.0791^a	0.3220^a	0.0298^a	36.6859^a	17.7869^a	0.1490^a	0.9819^a	0.1292^c	11416	0.0218	0.2105	0.1295	62.64
JPM	0.0005^a	0.4541^a	0.1832^a	0.1361^a	0.1824^a	0.0322^a	33.1899^a	26.7587^a	0.1918^a	0.9878^a	0.0548	10450	0.0524	0.2157	0.1139	31.39
KFT	0.0011^b	0.3065^a	0.2625^a	0.2090^a	0.0846	0.0012	30.0529^a	27.2279^a	0.2694^a	0.9673^a	0.0791^a	11079	0.0457	0.2383	0.6070	35.68
PEP	0.0003^a	0.2846^a	0.0275	0.0765	0.5545^a	0.0341^a	34.5631^a	21.7868^a	0.1710^a	0.8060^a	0.2754^a	11535	0.0208	0.0455	0.0007	64.31
PG	0.0007^a	0.3705^a	0.2044^a	0.1263^a	0.1975^a	0.0298^a	36.6526^a	15.1394^a	0.1541^a	0.9483^a	0.1342^a	11696	0.3166	0.5629	0.0057	37.96
T	0.0005^a	0.3682^a	0.1356^a	0.1256^a	0.3060^a	0.0246^a	37.1412^a	16.5691^a	0.1539^a	0.9723^a	0.1493^a	11126	0.0960	0.2222	0.0051	63.91
TWX	0.0006^a	0.3757^a	0.1253^c	0.1431^a	0.2888^a	0.0261^a	45.4493^a	23.1275^a	0.1539^a	0.7662^a	0.2464^a	10846	0.2432	0.6658	0.3053	37.17
TXN	0.0009^a	0.3625^a	0.1785^b	0.1390^a	0.2463^a	0.0296^a	34.3740^a	19.1028^a	0.1306^a	0.9704^a	0.1152^a	10208	0.4644	0.8732	0.1554	26.78
WFC	0.0004^a	0.4281^a	0.1665^b	0.1531^a	0.1926^a	0.0201^b	35.7612^a	16.2251^a	0.2086^a	0.9266^a	0.1493^a	10505	0.7835	0.9313	0.5639	40.06

Table 5: Estimates of the Asymmetric HAR-MEM-J with time-varying λ, see (20). The upper part of the table reports the results for several stock indexes while the lower part refers to 16 NYSE stocks. a, b and c stand for significance at 1%, 5% and 10% respectively. Q_1, Q_{10} and Q_{22} are the p-values of the Ljung-Box test for absence of autocorrelation in the residuals, where the latter are computed as $\hat{\epsilon}_t = \frac{RM_t}{E[RM_t|I_{t-1}]}$. LR_{ϕ_2,ϕ_3} is the likelihood ratio test for the nullity of ϕ_2 and ϕ_3.



(a) $S\&P500$ (b) CAC

Figure 4: Volatility-of-volatility ratio. The figures display the ratio $\frac{\mathbb{V}_J[RM_t|I_{t-1}]}{\mathbb{V}_0[RM_t|I_{t-1}]}$, where $\mathbb{V}_J[RM_t|I_{t-1}]$ is the conditional variance under AHAR-MEM-J-(λ_t), and $\mathbb{V}_0[RM_t|I_{t-1}]$ is the conditional variance under AHAR-MEM.

When the jump intensity parameter is assumed to be time-varying, see Table 5, we observe changes in parameters associated with the innovation term ε_t and with the jump component Z_t, while the parameters in μ_t are not much affected. The estimates of ν and ς are generally higher for the individual stocks than those of the indexes. The estimated unconditional mean of λ_t is between 0.15 and 0.20 for most stocks, and there are not relevant differences between stock indexes and individual stocks. Interestingly, most markets and stocks, among those considered, display estimates of ϕ_2 larger than 0.9. Two notable exceptions are FTSE-100, CAT and BOVESPA. For the latter, the sensitivity to the news arrival, measured by the parameter ϕ_3 is close to persistence parameter, i.e. ϕ_2, this might suggest that the time-varying jump intensity specification is not needed. However, the LR test for the joint nullity of ϕ_2 and ϕ_3 takes very large values. Even though in this case we don't have any asymptotic theory for the LR test, we believe that the observed values of the test statistic can reasonably lead to the rejection of the null hypothesis in all cases considered. Introducing dynamics in the jump intensity is therefore important to provide the necessary degree of flexibility in the characterization of the conditional moments of RM_t.

Figure 2 reports two examples of the fitted time-varying jump intensity, λ_t, and of the expected jump component $\mathbb{E}[Z_t|I_{t-1}]$. Both the jump intensity and the expected jump (which is a non-linear function of the jump intensity) increase during the recent crises: the end of technology market bubble in 2001-2002, the subprime crisis in 2007-2008 and the European sovereign crisis in 2010. Notably, the most recent crisis seem to be more relevant in France compared to the others, a somewhat expected result. Figure 3 plots the impact of a change in the model structure, moving from an AHAR-MEM without jumps to a specification including jumps. The figure shows the relative difference between the expected realized measure, i.e. $\mathbb{E}\left[RM_t|I_{t-1}\right]$, of the two models (expected bipower variation of the model with jumps minus the expected value from the model without jumps). We observe that the introduction of jumps leads to an increase in expected values in particular during market turmoil. This can be interpreted as a further evidence in favour of a model including a combination of continuous mean evolution and discontinuous (jumps) elements. The same evidence arises from a visual inspection of Figure 4 which reports the ratio between the conditional variance obtained from the model with jumps and the model without jumps. We note that the ratio is generally centred on 1, but during financial turmoil the conditional variance generated by jumps is by far larger than that generated without jumps. As a consequence, the model with jumps is expected to provide a better description of tail volatility events, as it is able to generate large and sudden increases in the conditional volatility-of-volatility levels.

5.3 Alternative specifications for the distribution of ϵ_t

An alternative specification of the AHAR-MEM can be based on a flexible distribution for ϵ_t that is able to generate fat-tails. To this scope, we choose the generalized Gamma distribution of Stacy (1962), leading to the AHAR-MEM-GG specification

$$\begin{aligned}
RM_t &= \mu_t \epsilon_t \\
\mu_t &= \omega + \beta\mu_{t-1} + \alpha_1 RM_{t-1} + \alpha_2 RM_{t-1:t-5} + \alpha_3 RM_{t-1:t-21} + \gamma RM^-_{t-1} \\
\epsilon^S_t &\sim GenGamma(1, \xi, \nu)
\end{aligned}$$

where the density of the generalized Gamma is

$$f_{GG}(\epsilon_t, \xi, \nu) = \frac{(\xi/\delta^\nu)\epsilon_t^{\xi-1} e^{(-\epsilon_t/\delta)^\xi}}{\Gamma(\nu/\xi)}$$

with $\delta = \frac{\Gamma(\nu/\xi)}{\Gamma((\nu+1)/\xi)}$ to ensure that $E(\epsilon_t) = 1$. Lunde (1999) and Andres and Harvey (2012) have already successfully adopted the generalized Gamma distribution in the ACD-MEM framework. The parameter estimates of the AHAR-MEM-GG model are reported in Table A.2 in the Web Appendix. As expected, the parameters governing the conditional mean, μ_t, are not affected by the choice of the distribution of ϵ_t and are very close to those obtained under the Gamma specification in Table 3 and the Ljung-Box tests do not signal significant autocorrelation in the residuals. However, the estimates of ξ and ν suggest a strong deviation from the Gamma distribution, which is obtained under the assumption that $\xi = 1$. The overall fitting of the AHAR-MEM-GG model, as measured by the value of the log-likelihood function, improves considerably compared to the AHAR-MEM. In particular, the estimates of ξ are far from 1 and generally close to 0.05, and the LR test for $\xi = 1$ strongly rejects the null hypothesis that ϵ_t is generated by a Gamma distribution.

According to Corsi et al. (2008), realized volatility sequences show significant evidences of heteroskedasticity, the so-called volatility-of-volatility effect. In order to accommodate this stylized fact in the MEM, beyond what is already accounted for by the time variation in the conditional mean μ_t,[9] we suggest a time-varying specification for a transformation of the shape parameter of the innovation term, i.e. $\bar{\nu} = \frac{1}{\nu}$. The reparameterized innovation shape follows the Auto Regressive Shape (ARS) specification

$$\bar{\nu}_t = \theta_0 + \theta_1 \bar{\nu}_{t-1} + \theta_2 \zeta_{t-1} \tag{34}$$

and the innovation ζ_t is given as

$$\zeta_t = \left(\frac{RM_t}{\mathbb{E}\left[RM_t | I_{t-1}\right]} \right)^2 - 1 - \bar{\nu}_t \tag{35}$$

Therefore, $\bar{\nu}_t$ has an autoregressive behaviour coupled with an innovation term linked to the squared value of the estimated innovation ε_t. The form of ζ_t comes from the properties of the Gamma density. Recall that $\varepsilon_t | I_{t-1} \sim \Gamma\left(1, \nu_t\right) = \Gamma\left(1, \frac{1}{\bar{\nu}_t}\right)$ we have that $\mathbb{E}\left[\varepsilon_t | I_{t-1}\right] = 1$ and $\mathbb{V}\left[\varepsilon_t | I_{t-1}\right] = \bar{\nu}_t$. In addition, $\mathbb{E}\left[\varepsilon_t^2 | I_{t-1}\right] = 1 + \bar{\nu}_t$. The model with time varying shape is named AHAR-MEM-ν_t. The estimates of this model are reported in Table 6. The point estimates of the parameters governing the dynamics of μ_t are very close to those obtained with the AHAR-MEM specification with constant ν reported in Table 3. Interestingly, the parameters governing the dynamics of $\bar{\nu}_t$ are significant in most cases, and the LR test for the joint nullity of θ_1 and θ_2 rejects the null in most cases. This evidence suggests that a restrictive assumption on the distribution of the innovation term, i.e. $\varepsilon_t \sim \Gamma(1, \nu)$ with constant ν, is rejected by the data. Indeed, despite the AHAR-MEM with constant ν generates time-varying conditional variance through μ_t^2, see equation (5), the results suggests that an extra degree of flexibility is required to model the conditional second moment.

Finally, we introduce the mixture AHAR-MEM specification (M-AHAR-MEM), extending the mixture model of Lanne (2006) to allow for HAR dynamics in each volatility

[9] The conditional mean μ_t already provides some form of heteroskedasticity, as we have already noted. However, this might not be sufficiently flexible to capture the volatility-of-volatility effect.

	ω	α_1	α_2	α_3	β	γ	θ_0	θ_1	θ_2	$LogLL$	Q_1	Q_{10}	Q_{22}	LR_{θ_1,θ_2}
SP500	0.0003^a	0.3041^a	0.1672^a	0.1060^a	0.3363^a	0.1096^a	0.0038^a	0.9378^a	0.0050^a	15734	0.9344	0.4857	0.5390	50.05
FTSE 100	0.0002	0.3161^a	0.2464^a	0.1201^a	0.2552^a	0.0646^a	0.0128^a	0.7787^a	0.0104	16176	0.7626	0.0060	0.0020	81.84
DAX	0.0003^a	0.3250^a	0.1275^a	0.1179^a	0.3672^a	0.0752^a	0.0103^a	0.8160^a	0.0079^a	15256	0.5220	0.0032	0.0000	47.54
DJIA	0.0002^a	0.3077^a	0.1895^a	0.1017^a	0.3260^a	0.0988^a	0.0032^a	0.9489^a	0.0058^a	15804	0.9329	0.5755	0.6027	67.61
NSDQ	0.0001^a	0.3222^a	0.1866^a	0.1635^a	0.2714^a	0.0974^a	0.0013^a	0.9779^a	0.0046^a	15276	0.7925	0.5182	0.0711	53.52
CAC40	0.0002^a	0.3109^a	0.1342^a	0.1035^a	0.3856^a	0.0852^a	0.0062^a	0.8858^a	0.0055^a	15554	0.5590	0.0001	0.0000	49.57
BOVESPA	0.0010^a	0.3448^a	0.1515^a	0.1658^a	0.2183^a	0.0726^a	0.0494^a	0.2305^a	0.0110^a	13881	0.9376	0.1709	0.1980	23.65
BA	0.0006^a	0.3850^a	0.0789^a	0.1168^a	0.3624^a	0.0302^a	0.0048^a	0.8892^a	0.0041^a	10588	0.8739	0.0621	0.0456	20.24
BAC	0.0003^a	0.5189^a	0.0531^a	0.0986^a	0.2926^a	0.0416^a	0.0119^a	0.7562^a	0.0100^a	10245	0.7037	0.4143	0.5928	35.58
C	0.0001	0.4803^a	0.1130^a	0.0670^a	0.3255^a	0.0340^a	0.0004^a	0.9887^a	0.0069^a	10185	0.7980	0.0598	0.0713	225.68
CAT	0.0008^a	0.3989^a	0.2480^a	0.1094^a	0.1683^a	0.0469^a	0.0075^a	0.8226^a	0.0072^a	10384	0.9899	0.8039	0.8961	42.91
FDX	0.0004	0.4218^a	0.0831^a	0.1627^a	0.2870^a	0.0314	0.0080	0.8249^a	0.0068	10559	0.7456	0.4557	0.3413	24.91
HON	0.0006^a	0.4125^a	0.1303^a	0.0853^a	0.3064^a	0.0508^a	0.0043^a	0.8985^a	0.0079^a	10548	0.9684	0.7746	0.0664	68.11
HPQ	0.0004^a	0.3912^a	0.1783^a	0.1528^a	0.2290^a	0.0496^a	0.0019^a	0.9586^a	0.0051^a	10396	0.7268	0.9478	0.5159	64.74
IBM	0.0001^a	0.4302^a	0.0953^a	0.0803^a	0.3682^a	0.0403^a	0.0007^a	0.9811^a	0.0047^a	11352	0.7672	0.5905	0.3508	76.10
JPM	0.0004^a	0.4735^a	0.1944^a	0.0976^a	0.1894^a	0.0345^a	0.0058^a	0.8664^a	0.0082^a	10378	0.9654	0.2145	0.2509	56.75
KFT	0.0009^a	0.4063^a	0.2399^a	0.1231^a	0.1485^a	0.0069^a	0.0102^a	0.8259^a	0.0175^a	10922	0.1772	0.0982	0.2973	94.44
PEP	0.0003^a	0.3402^a	0.0870^a	0.0656^a	0.4607^a	0.0340^a	0.0059^a	0.8622^a	0.0111^a	11441	0.9868	0.0011	0.0090	121.66
PG	0.0005^a	0.3946^a	0.2505^a	0.0942^a	0.1935^a	0.0335^a	0.0035^a	0.9141^a	0.0085^a	11630	0.4945	0.9344	0.3630	130.38
T	0.0001	0.3933^a	0.0980^a	0.1141^a	0.3714^a	0.0248^a	0.0006^a	0.9832^a	0.0045^a	11055	0.6114	0.4681	0.0133	112.14
TWX	0.0006^a	0.4050^a	0.1598^a	0.1139^a	0.2721^a	0.0243^a	0.0038^a	0.8858^a	0.0083^a	10733	0.7376	0.6394	0.6280	81.21
TXN	0.0006^a	0.3719^a	0.1885^a	0.1576^a	0.2354^a	0.0338^a	0.0012^a	0.9688^a	0.0041^b	10166	0.6269	0.7724	0.1390	51.45
WFC	0.0004^a	0.4719^a	0.1988^a	0.1233^a	0.1700^a	0.0263^a	0.0079^a	0.8437^a	0.0099^b	10330	0.1354	0.3657	0.4412	84.11

Table 6: Estimates of the Asymmetric HAR-MEM with time varying shape, ν_t. The upper part of the table reports the results for several stock indexes while the lower part refers to 16 NYSE stocks. a, b and c stand for significance at 1%, 5% and 10% respectively. Q_1, Q_{10} and Q_{22} are the p-values of the Ljung-Box test for absence of autocorrelation in the residuals, where the latter are computed as $\hat{\epsilon}_t = \frac{RM_t}{E[RM_t|I_{t-1}]}$. LR_{θ_1,θ_2} is the likelihood ratio test for the nullity of θ_1 and θ_2.

component,

$$
\begin{aligned}
RM_t &= \pi\mu_t^L\epsilon_t^L + (1-\pi)\mu_t^S\epsilon_t^S \\
\mu_t^L &= \omega^L + \beta^L\mu_{t-1}^L + \alpha_1^L RM_{t-1} + \alpha_2^L RM_{t-1:t-5} + \alpha_3^L RM_{t-1:t-21} + \gamma^L RM_{t-1}^- \\
\mu_t^S &= \omega^S + \beta^S\mu_{t-1}^S + \alpha_1^S RM_{t-1} + \alpha_2^S RM_{t-1:t-5} + \alpha_3^S RM_{t-1:t-21} + \gamma^S RM_{t-1}^- \\
\epsilon_t^L &\sim \Gamma(1, \frac{1}{\nu^L}) \\
\epsilon_t^S &\sim \Gamma(1, \frac{1}{\nu^S})
\end{aligned}
$$

Lanne (2006) reports strong evidence in favour of the mixture MEM in capturing not only the long-memory dynamics of the realized volatility, but also its heavy tail marginal distribution as generated by the mixture of the two Gamma densities governing ϵ_t^L and ϵ_t^S. The estimates of the model parameters, reported in Table A.3 in the Web Appendix, confirm the results in Lanne (2006) supporting the evidence of two volatility factors with distinct dynamic features. The dynamics of μ_t^L are characterized by high persistence and low variance of the error term; indeed α_2^L and α_3^L are significant and ν^L lies generally above 25. Conversely, μ_t^S is the rapidly moving volatility factor which is characterized by higher innovation variance, with ν^S below 10 in many cases, and non persistent dynamics, i.e. α_2^S and α_3^S are not different from zero in most cases. The estimates of the mixture parameter π are between 0.69 and 0.92, meaning that more weight is given to the persistent volatility factor.

In the next section, we will evaluate the ability of the MEM specifications considered in this study to correctly predict the probability of tail events.

6 Volatility-at-Risk

One of the advantages of the MEM-J with respect to multiplicative specifications without jumps is given by the ability of the model in capturing the evolution of the variance upper tail. In general, the volatility distribution is heavily skewed to the right with many more extreme volatility realizations than we would expect if the distribution was normal. Volatility jumps are responsible of rapid boosts in the volatility which, if not correctly modeled, would imply model's residuals with an upper tail fatter than that hypothesized for the errors.

By analogy to the VaR introduced for quantifying extreme risk on returns, we can define the VolaR, i.e. the risk of extreme high volatility as

$$\Pr\{RM_t > v(\alpha)|I_{t-1}\} = \alpha$$

where $\Pr\{\cdot|I_{t-1}\}$ denotes the conditional distribution at date t of the one-step-ahead volatility, whereas $v(\alpha)$ defines the volatility level that may occur with probability α. The VolaR might be of interest for investors trading volatility, see Zhang et al. (2010), and Euan (2013) and therein cited references. In fact, the knowledge of the probability that volatility will exceed a given threshold is useful both in designing volatility trading strategies based on options (allowing for an optimal calibration of the option maturity as well as the option strike) and for strategies based on volatility indices or exchange traded volatility products (having an impact on the choice of the investment direction as well as on the size of the position). In addition, the evaluation of volatility risk might be of interest for options traders and market makers to define optimal prices and order execution, and, finally, to portfolio managers willing to determine the need and the amount of a volatility hedge.

	AMEM	AHAR	AHAR-ν_t	M-AHAR	AHAR-GG	HAR-V-J	AHAR-J	AHAR-J-λ_t
SP500	0.0000	0.0000	0.0000	0.0080	0.0000	0.5181	0.6749	0.3456
FTSE 100	0.0000	0.0000	0.0000	0.4882	0.0000	0.7394	0.3289	0.4896
DAX	0.0000	0.0000	0.0000	0.0241	0.0001	0.6660	0.1415	0.1337
DJIA	0.0000	0.0000	0.0000	0.0010	0.0000	0.4189	0.4334	0.5880
NSDQ	0.0000	0.0000	0.0000	0.6016	0.0000	0.6651	0.5216	0.5999
CAC	0.0000	0.0000	0.0000	0.0013	0.0004	0.0984	0.0677	0.3594
BOVESPA	0.0000	0.0000	0.0000	0.5952	0.0519	0.8380	0.0055	0.0145
BA	0.0000	0.0000	0.0000	0.0229	0.0000	0.5486	0.0423	0.1483
BAC	0.0000	0.0000	0.0000	0.0268	0.0000	0.0008	0.1095	0.2011
C	0.0000	0.0000	0.0000	0.0045	0.0000	0.0217	0.1608	0.6656
CAT	0.0000	0.0000	0.0000	0.0921	0.0000	0.5137	0.2131	0.2528
FDX	0.0000	0.0000	0.0000	0.8566	0.0000	0.5309	0.0746	0.0815
HON	0.0000	0.0000	0.0000	0.0013	0.0000	0.0369	0.2604	0.2923
HPQ	0.0000	0.0000	0.0000	0.5133	0.0000	0.0621	0.0601	0.1515
IBM	0.0000	0.0000	0.0000	0.0038	0.0000	0.5957	0.3150	0.0909
JPM	0.0000	0.0000	0.0000	0.0856	0.0000	0.0059	0.1312	0.1218
KFT	0.0000	0.0000	0.0000	0.0112	0.0000	0.6487	0.0000	0.2042
PEP	0.0000	0.0000	0.0000	0.0000	0.0000	0.6652	0.1350	0.9984
PG	0.0000	0.0000	0.0000	0.0123	0.0000	0.9482	0.0267	0.0290
T	0.0000	0.0000	0.0000	0.0078	0.0000	0.5549	0.5883	0.7226
TWX	0.0000	0.0000	0.0000	0.0023	0.0000	0.0487	0.4656	0.3480
TXN	0.0000	0.0000	0.0000	0.0001	0.0000	0.5385	0.1078	0.2782
WFC	0.0000	0.0000	0.0000	0.4959	0.0000	0.0128	0.1071	0.0930

Table 7: P-values of the Berkowitz (2001) test for the in-sample VolaR at 1% level, corresponding to a value of 2.3263.

In order to evaluate the estimation of the VolaR (i.e. the right tail coverage) obtained with models with and without jumps, we consider the method introduced by Berkowitz (2001), which allows testing the adherence of the hypothesised density with the realization of the modeled variable. The test has been proposed for density forecast evaluation, but can also be applied in-sample. The test of Berkowitz (2001) makes use of the Rosenblatt (1952) transformation, and is built within a likelihood framework. Moreover, the test is flexible and can be applied to the fit of the entire density as well as over specific segments

	AMEM	AHAR	AHAR-ν_t	M-AHAR	AHAR-GG	HAR-V-J	AHAR-J	AHAR-J-λ_t
S&P500	0.0000	0.0000	0.0000	0.0035	0.0002	0.3503	0.0048	0.5133
FTSE 100	0.5577	0.1526	0.0000	0.0615	0.5078	0.0000	0.3551	0.4917
DAX	0.0010	0.0002	0.0000	0.0436	0.4529	0.1246	0.0171	0.1212
DJIA	0.0000	0.0000	0.0000	0.0009	0.0001	0.0260	0.0358	0.5111
NSDQ	0.0000	0.0000	0.0000	0.0025	0.0000	0.0010	0.0293	0.9532
CAC	0.0000	0.0000	0.0000	0.0696	0.1039	0.2384	0.0711	0.4825
BOVESPA	0.0279	0.0117	0.0000	0.3168	0.3036	0.8389	0.4567	0.2591
BA	0.0000	0.0000	0.0000	0.0477	0.0002	0.3503	0.2482	0.4850
BAC	0.0000	0.0000	0.0000	0.9360	0.0000	0.0010	0.0185	0.0768
C	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0043	0.1416
CAT	0.0000	0.0000	0.0000	0.0231	0.0003	0.0631	0.0262	0.3891
FDX	0.0000	0.0000	0.0000	0.0362	0.0000	0.1076	0.1889	0.6195
HON	0.0000	0.0000	0.0000	0.0002	0.0006	0.6167	0.6935	0.1595
HPQ	0.0000	0.0000	0.0000	0.0416	0.0000	0.1246	0.0142	0.3614
IBM	0.0000	0.0000	0.0000	0.0000	0.0000	0.8587	0.7541	0.7128
JPM	0.0000	0.0000	0.0000	0.0000	0.0000	0.0260	0.0453	0.1329
KFT	0.0000	0.0000	0.0000	0.3501	0.0000	0.0007	0.7700	0.6569
PEP	0.0000	0.0000	0.0000	0.0356	0.0001	0.0010	0.5746	0.5788
PG	0.0000	0.0000	0.0000	0.0119	0.0000	0.0000	0.3249	0.9246
T	0.0000	0.0000	0.0000	0.0140	0.0000	0.2384	0.7679	0.1859
TWX	0.0000	0.0000	0.0000	0.1001	0.0000	0.0409	0.0117	0.2544
TXN	0.0000	0.0000	0.0000	0.0031	0.0069	0.8389	0.0000	0.0055
WFC	0.0000	0.0000	0.0000	0.0000	0.0000	0.5799	0.0121	0.0002

Table 8: P-values of the Berkowitz (2001) test for the out-of-sample VolaR at 1% level, corresponding to a value of 2.3263.

of the density support. For our purposes, we apply the test over the upper $q\%$ tail of the RM_t density. In details, given the density of the RM_t, we compute the conditional CDF of RM_t

$$y_t = F\left(RM_t|I_{t-1}\right) = \int_0^{RM_t} f\left(x|I_{t-1}\right)\mathrm{d}x,$$

where $F(RM_t|I_{t-1})$ for the MEM-J is given by the mixture of Gamma and K conditional CDFs (see Appendix B for the K CDF). Then, under correct model specification, the empirical CDF values should be distributed according to the standard uniform, i.e. $y_t \sim U(0,1)$. Berkowitz (2001) considers the further transformation

$$z_t = \Phi^{-1}\left(y_t\right)$$

where $\Phi(\cdot)$ is the standard normal CDF, and z_t would be, under correct model specification, distributed as a standardized normal. To test the tail coverage of alternative models, we choose a VolaR level of 1% (i.e. VolaR = 2.3263), and calculate a new truncated variable

$$z_t^* = \begin{cases} \text{VolaR} & \text{if} \quad z_t \leq \text{VolaR} \\ z_t & \text{if} \quad z_t > \text{VolaR} \end{cases} \tag{36}$$

A tail coverage test can be derived using the LR principle. Under the null, the mean and the variance of z_t are 0 and 1, respectively, while under the alternative they are unrestricted. Under the null of correct tail coverage the test statistic is distributed as $\chi^2(2)$. See Berkowitz (2001) for further details on this test.

The Berkowitz (2001) test emphasizes the ability of a model specification to well account for the probability of tail events. In Table 7 the p-values of the Berkowitz (2001) test based on the in-sample estimates are reported for alternative MEM specifications considered so far. It clearly emerges that the AHAR-MEM with jumps outperforms the corresponding specification without jumps. All the MEM specifications without jumps and Gamma distribution for ε_t strongly reject the null hypothesis of correct specification of the upper quantiles. It should be noted that also the AHAR-MEM-ν_t model is poorly designed to capture tail events. Letting the conditional variance of ε_t to be time varying is not sufficient for a proper characterization of the extremes of the conditional distribution of RM_t, thus suggesting a distinct role of the jumps from pure heteroskedastic effects in ε_t. Interestingly, the M-AHAR-MEM of Lanne (2006) provides some evidence of correct specification of the upper tail, as the Berkowitz test cannot reject the null hypothesis in 3 over 7 cases. Conversely, the AHAR-MEM-GG model fails to give the correct probability mass on the right tail, in other words, despite the generalized Gamma distribution provides a good fitting for the entire distribution, it fails to properly account for the probability of tail events.

It is noteworthy that the introduction of jumps, with constant and time-varying λ, is sufficient for a good characterization of the right tail. In only one case, BOVESPA, the presence of jumps leads to minor improvements. We still have a rejection of the null with constant jump intensity, while with time-varying jump intensity the null cannot be rejected, but only at the 5% significance level. The introduction of the time-varying jump intensity sensibly improves the performances for the CAC40 index. A possible explanation for the good performance of the model with jumps in giving the correct probability mass to the right tail can be derived from a visual inspection of Figure 5. It clearly emerges that the model with jumps is able to generate large and sudden spikes in the conditional variance of $\sqrt{BPV_t}$, as generated by the jump component Z_t, while the model with time-varying ν_t, can only generate smooth trajectories, and hence it is not able to assign enough probability to extreme volatility events. Interestingly, also the HAR-V-J model of Caporin et al. (2014), which is an HAR specification with time-varying jump intensity on the log transformed $\sqrt{BPV_t}$ series, provides a good fitting of the tails of the volatility distribution.[10] In this case, the null hypothesis cannot be rejected at 5% significance for all the indexes. This confirms once more the importance to account for the probability of

[10] Due to space constraints, we do not report a complete discussion of the HAR-V-J model, which can be found in Caporin et al. (2014). It is however important to note, that, since the HAR-V-J model is linear in the logarithms of realized volatility, this implies a multiplicative structure for the latter, similar to that obtained under the AHAR-MEM-J.

large volatility increments.

Table 8 reports the p-values of the Berkowitz (2001) test based on the out-of-sample forecasts, for a total of 1,000 observations (the holdout sample starts on February 2, 2009). When AMEM and AHAR are used, the null hypothesis is always rejected with the exception of FTSE. Better results are obtained when the Generalized Gamma is adopted. In 4 out of 7 cases, the null hypothesis cannot be rejected. A similar performance is also achieved with the M-AHAR-MEM and with the AHAR-MEM-J with constant λ. Indeed, by including the jumps (with constant intensity) the null is not rejected for DAX, DJIA, NASDAQ and CAC40, even if the associated p-values are below 10%. A slightly better performance is achieved with the HAR-V-J model as the null hypothesis cannot be rejected in four out of seven cases. A striking improvement is instead associated with the full model with persistence and time-varying jump intensity. For the AHAR-MEM-J-λ_t the lowest p-value is associated with the DAX index and equals 12%. The out-of-sample results confirm the adequacy of our model in capturing the presence of volatility jumps and support the importance of jumps in VolaR estimation.

The same tail analysis is repeated also on the individual stocks. In this case, the out-of-sample period starts on July 15, 2008. The introduction of volatility jumps produces clear improvements. The baseline AMEM as well as its generalizations with persistence and time-varying shape always reject the null hypothesis of correct tail coverage both in-sample and out-of-sample. The picture slightly improves when the M-AHAR-MEM specification is adopted, as the null cannot be rejected at 5% significance in 4 cases for the in-sample analysis and in 3 cases in the out-of-sample case. Conversely, the AHAR-MEM specification with the Generalized Gamma distribution always rejects the null hypothesis at all significance levels. Differently, when the possibility of volatility jumps is included in the model, we observe a small number of rejections. We have 3 out 16 in-sample and 8 out of 16 out-of-sample rejections for the model with constant jump intensity. Similarly, we have 5 out of 16 in-sample and 7 out of 16 out-of-sample rejections when the HAR-V-J specification is adopted. These numbers decrease to 1 and 2 for the in-sample and out-of-sample cases, respectively, when the jump intensity is dynamic. The results with single stocks further support our model, highlighting the potential benefits of accounting for volatility jumps in modeling volatility series. It is also important to stress the fact that alternative model specifications, allowing for fat-tails, fail in most cases in correctly estimating the VolaR. This evidence stresses the importance of volatility jumps in characterizing the extremes of the volatility series.

7 Concluding remarks

We have introduced a new model, the AHAR-MEM-J for realized measures that generalizes the MEM of Engle and Gallo (2006) by introducing persistence (through HAR

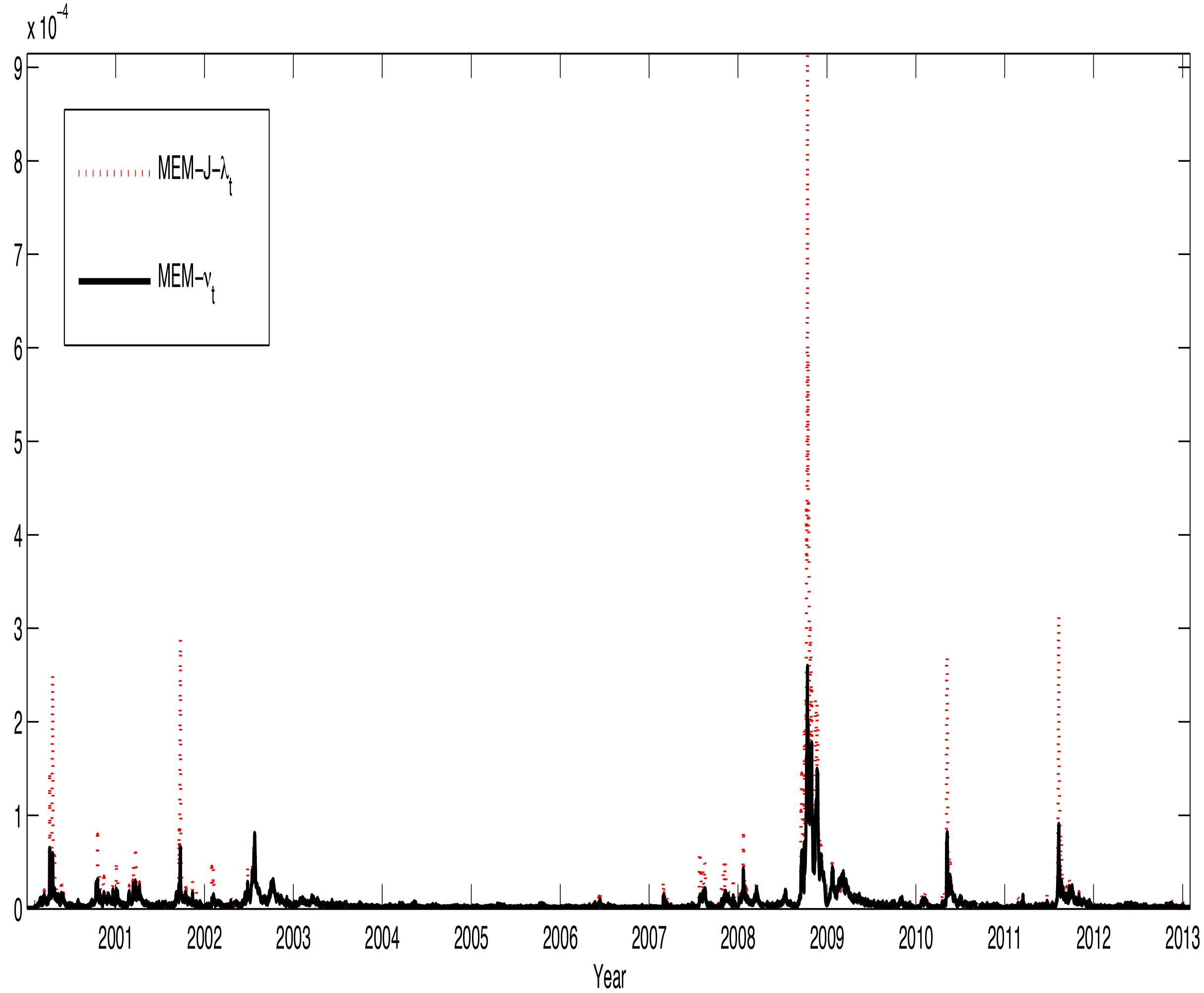


Figure 5: Conditional variance of RM_t of $S\&P500$ obtained with the AHAR-MEM-J-λ_t, dashed red line, and with the AHAR-MEM-ν_t, straight black line.

terms, see Corsi, 2009) and multiplicative volatility jumps. By specifying the volatility process as a combination of a continuous volatility component and a discrete compound Poisson process for the jumps, the conditional density of the realized measure is a countably infinite mixture of two random variables: one distributed as a Gamma, and the second, when the number of jumps is strictly larger than zero, is K distributed. We add further flexibility considering both time-varying jump intensity and time-varying specification for a transformation of the shape parameter of the innovation term. This flexible parametrization of the dynamics of the realized measure allow to capture the extreme or abnormal movements in the volatility level. We discuss also the model estimation in finite samples, by means of a small Monte Carlo simulation, and the effects of misspecification. The empirical application shows that the AHAR-MEM-J(λ_t) captures the extreme moves registered in the last years in the volatilities of individual stocks and equity indexes. We also provide statistical evidence that, for the sample period analyzed, the model correctly predicts the probability of occurrence of abnormal volatility levels, i.e. of jumps. We compare alternative models by means of a new measure called the volatility-at-risk, i.e. the risk of extreme high volatility. The empirical analysis put in evidence how models that cannot generate sudden and large movements in the realized variance, i.e. without jumps, fail in fitting the extreme right tail of the distribution. Moreover, the recent empirical evidence on the contemporaneous correlation between jumps in price and volatility would suggest an extension of our set up to include the presence of price jumps. This is left for future research. Finally, the potential application of this model is not limited to the study of volatility but it can be employed in the analysis of any positive time series that features persistence and sudden large variations, e.g. trading volume and durations.

Acknowledgements. We wish to thank Eric Ghysels, Christian Gourieroux, Andrew Harvey, Roberto Renò, Esther Ruiz, and Jean-Michel Zakoïan, for stimulating discussions and comments. We also thanks the participants to the seminar at CREST, the SoFiE-INET conference on Skewness, Heavy Tails, Market Crashes, and Dynamics (Cambridge, 2014), the XV Workshop in Quantitative Finance (Florence, 2014), the 13th OxMetrics User Conference (Aarhus 2013), the IAAE 2014 Conference (London 2014), the International Symposium on Forecasting (Rotterdam, 2014), for their valuable suggestions. All errors and omissions are our own.
Massimiliano Caporin acknowledges financial support from: the European Union, Seventh Framework Program FP7/2007-2013 under grant agreement SYRTO-SSH-2012-320270; the MIUR PRIN project MISURA - Multivariate Statistical Models for Risk Assessment; the Institut Europlace de Finance under the research program Systemic Risks. Eduardo Rossi acknowledges financial support from the MIUR PRIN project MISURA - Multivariate Statistical Models for Risk Assessment. Paolo Santucci de Magistris acknowledges financial support from the Danish National Research Foundation (DNRF78).

References

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007). Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility. *The Review of Economics and Statistics*, 89(4):701–720.

Andersen, T. G., Bondarenko, O., and Gonzalez-Perez, M. T. (2012). Uncovering novel features of equity-index return dynamics via corridor implied volatility. Technical report.

Andres, P. and Harvey, A. (2012). The dyanamic location/scale model: with applications to intra-day financial data. Technical report, Faculty of Economics, University of Cambridge.

Ausín, M.-C. and Galeano, P. (2007). Bayesian estimation of the gaussian mixture {GARCH} model. *Computational Statistics & Data Analysis*, 51(5):2636 – 2652.

Bandi, F. and Renò, R. (2012). Time-varying leverage effects. *Journal of Econometrics*, 169:94–113.

Bandi, F. and Renò, R. (2013). Price and volatility co-jumps. Technical report, University of Siena.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2:1–37.

Bauwens, L., Hafner, C., and Rombouts, J. (2007). Multivariate mixed normal conditional heteroskedasticity. *Computational Statistics & Data Analysis*, 51(7):3551 – 3566.

Bauwens, L. and Veredas, D. (2004). The stochastic conditional duration model: a latent factor model for the analysis of financial durations. *Journal of Econometrics*, 119:381–412.

Berkowitz, J. (2001). The accuracy of density forecasts in risk management. *Journal of Business and Economic Statistics*, 19:465–474.

Broadie, M., Chernov, M., and Johannes, M. (2007). Model specification and risk premia: Evidence from futures options. *The Journal of Finance*, 62:1453–1490.

Broda, S. A., Haas, M., Krause, J., Paolella, M. S., and Steude, S. C. (2013). Stable mixture GARCH models. *Journal of Econometrics*, 172(2):292 – 306.

Brownlees, C. T., Cipollini, F., and Gallo, G. M. (2012). Multiplicative error models. In Bauwens, L., Hafner, C., and Laurent, S., editors, *Handbook of Volatility Models and Their Applications*, pages 223–247. John Wiley & Sons, Inc.

Caporin, M., Rossi, E., and de Magistris, P. S. (2014). Volatility jumps and their economic determinants. *CREATES RP2014-25, Forthcoming on Journal of Financial Econometrics.*

Chan, W. H. and Maheu, J. M. (2002). Conditional jump dynamics in stock market returns. *Journal of Business and Economic Statistics*, 20(3):377–389.

Chernov, M., Gallant, R., Ghysels, E., and Tauchen, G. (2003). Alternative models for stock price dynamics. *Journal of Econometrics*, 116:225–257.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7:174–196.

Corsi, F., Mittnik, S., Pigorsch, C., and Pigorsch, U. (2008). The volatility of realized volatility. *Econometric Reviews*, 27:46–78.

Creal, D., Koopman, S. J., and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777–795.

Duffie, D., Pan, J., and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68:1343–1376.

Engle, R. F. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics*, 17:425–446.

Engle, R. F. and Gallo, G. M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, 131(12):3–27.

Eraker, B. (2004). Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *Journal of Finance*, 59(2):1367–1403.

Eraker, B., Johannes, M. S., and Polson, N. G. (2003). The impact of jumps in returns and volatility. *Journal of Finance*, 53(3):1269–1300.

Euan, S. (2013). *Volatility Trading*. Wiley Trading series. John Wiley & Sons, Inc., Hoboken (NJ), second edition.

Franq, C. and Zakoian, J.-M. (2010). *GARCH Models*. John Wiley & Sons, Inc., Chichester, UK.

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Spriger, New York.

Gallo, G. M. and Otranto, E. (2012). Realized volatility and change of regimes. Technical report, Dipartimento di statistica "G. Parenti",Universita' di Firenze.

Ghysels, E., Gouriéroux, C., and Jasiak, J. (2004). Stochastic volatility duration models. *Journal of Econometrics*, 119(2):413 – 433.

Haerdle, W., Hautsch, N., and Mihoci, A. (2012). Local adaptive multiplicative error models for high-frequency forecasts. Technical Report 31, Humboldt -Universitat zu Berlin.

Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review*, 35:705–730.

Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*, 64(2):413–30.

Huang, X. and Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, 3:456–499.

Jones, C. (2003). Nonlinear mean reversion in the short-term interest rate. *Review of Financial Studies*, 16:765–791.

Kiefer, N. (1978). Discrete parameter variation: Efficient estimation of a switching regression model. *Econometrica*, 46:427–434.

Lanne, M. (2006). A Mixture Multiplicative Error Model for Realized Volatility. *Journal of Financial Econometrics*, 4(4):594–616.

Lunde, A. (1999). A generalized gamma autoregressive conditional duration model. Technical report, Department of Economics, Politics and Public Administration, Aalborg University.

Maheu, J. M. and McCurdy, T. H. (2004). News arrival, jump dynamics, and volatility components for individual stock returns. *Journal of Finance*, LIX:755–793.

Pan, J. (2002). The jump-risk premia implicit in options: evidence from an integrated time-series study. *Journal of Financial Economics*, 63(1):3 – 50.

Redding, N. J. (1999). Estimating the parameters of the k distribution in the intensity domain. Technical report, DSTO Electronics and Surveillance Laboratory, South Australia.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The Annals of Mathematical Statistics*, 23:470–472.

Stacy, E. (1962). A generalization of the gamma distribution. *The Annals of Mathematical Statistics*, 33:1187–1192.

Todorov, V. and Tauchen, G. (2011). Volatility jumps. *Journal of Business and Economic Statistics*, 29:356–371.

Zhang, J. E., Shu, J., and Brenner, M. (2010). The new market for volatility trading. *Journal of Futures Markets*, 30(9):809–833.

A Proofs and results

A.1 Proof of Proposition 1

Given $\varepsilon_t|I_{t-1} \sim \Gamma(1,\nu)$ and $Z_t|N_t = m > 0, I_{t-1} \sim \Gamma(m, m\varsigma)$, integrating out z_t we have the conditional density of η_t

$$f(\eta_t|N_t = m > 0, I_{t-1}) = \frac{2}{\eta_t}\big(\eta_t\varsigma\nu\big)^{\frac{m\varsigma+\nu}{2}}\frac{1}{\Gamma(m\varsigma)\Gamma(\nu)}\mathbb{K}_{m\varsigma-\nu}\Big(2\sqrt{\eta_t\varsigma\nu}\Big).$$

which is the K density, see Redding (1999). $\mathbb{K}_a(\cdot)$ is the modified Bessel function of the second kind. The moments of η_t, conditional on $N_t = m > 0$ and I_{t-1}, are derived from the moments of the K density, which are given by

$$\mathbb{E}\left[y^s\right] = \frac{\mu^s\Gamma\left(\nu_1+s\right)\Gamma\left(\nu_2+s\right)}{\nu_1^s\nu_2^s\Gamma\left(\nu_1\right)\Gamma\left(\nu_2\right)}.$$

See Appendix B for further details on K distribution. ■

A.2 Proof of Proposition 2

From equation (14), the conditional density of RM_t is derived as

$$\begin{aligned}
f(RM_t|N_t = m > 0, I_{t-1}) &= f\Big(\frac{RM_t}{\mu_t}|N_t = m > 0, I_{t-1}\Big)\Big|\frac{1}{\mu_t}\Big| \\
&= \frac{2\mu_t}{RM_t}\Big(\frac{RM_t}{\mu_t}\varsigma\nu\Big)^{\frac{m\varsigma+\nu}{2}u}\frac{1}{\Gamma(m\varsigma)\Gamma(\nu)}\mathbb{K}_{m\varsigma-\nu}\left(2\sqrt{\frac{RM_t}{\mu_t}\varsigma\nu}\right)\Big|\frac{1}{\mu_t}\Big| \\
&= \frac{2}{RM_t}\Big(\frac{RM_t}{\mu_t}\varsigma\nu\Big)^{\frac{m\varsigma+\nu}{2}}\frac{1}{\Gamma(m\varsigma)\Gamma(\nu)}\mathbb{K}_{m\varsigma-\nu}\left(2\sqrt{\frac{RM_t}{\mu_t}\varsigma\nu}\right)
\end{aligned}$$

Similarly to the case of η_t, the moments of RM_t conditional on $N_t = m > 0$ and I_{t-1} are derived from the moments function of the K density. ■

A.3 Moments of Z_t

Lemma A.1 *Given the MEM-J in* (6) *with Assumption 1 and the processes for λ_t specified as in* (20)*, the conditional moments of Z_t are*

$$\mathbb{E}\big[Z_t|I_{t-1}\big] = e^{-\lambda_t} + \lambda_t \tag{37}$$

$$\mathbb{V}\big[Z_t|I_{t-1}\big] = \frac{\lambda_t}{\varsigma} + e^{-\lambda_t} + \big(\lambda_t + \lambda_t^2\big) - \big(e^{-\lambda_t} + \lambda_t\big)^2. \tag{38}$$

The filtered expected jumps are

$$\mathbb{E}\left[Z_t|I_t\right] = \sum_{m=0}^{\infty} P\left(N_t = m|I_t\right) \times \mathbb{E}\left[Z_t|N_t = m, I_{t-1}\right] \tag{39}$$

Proof To prove this result, we start by computing the expectation (and for completeness the variance) of the jump term Z_t. We distinguish two cases depending on the presence of jumps. When we have no jumps, the mean and variance of Z_t are

$$\begin{aligned}
\mathbb{E}\left[Z_t|N_t=0,I_{t-1}\right] &= 1\\
\mathbb{V}\left[Z_t|N_t=0,I_{t-1}\right] &= 0
\end{aligned}$$

Differently, when $N_t>0$ we have

$$\begin{aligned}
\mathbb{E}\left[Z_t|N_t=m>0,I_{t-1}\right] &= m\\
\mathbb{V}\left[Z_t|N_t=m>0,I_{t-1}\right] &= \frac{m}{\varsigma}.
\end{aligned}$$

Integrating out the dependence on N_t, we obtain

$$\begin{aligned}
\mathbb{E}\left[Z_t|I_{t-1}\right] &= \sum_{m=0}^{\infty} P\left(N_t=m|I_{t-1}\right)\times\mathbb{E}\left[Z_t|N_t=m,I_{t-1}\right]\\
&= P\left(N_t=0|I_{t-1}\right)\times 1+\sum_{m=1}^{\infty} P\left(N_t=m|I_{t-1}\right)\times m\\
&= e^{-\lambda_t}+\lambda_t \qquad (40)
\end{aligned}$$

as, for the Poisson process governing the jumps number we have $P\left(N_t=0|I_{t-1}\right)=e^{-\lambda_t}$ and $\mathbb{E}\left[N_t|I_{t-1}\right]=\sum_{m=0}^{\infty}P\left(N_t=m|I_{t-1}\right)\times m=\lambda_t=\sum_{m=1}^{\infty}P\left(N_t=m|I_{t-1}\right)\times m$ since for $m=0$ we do not have a contribution to the expected value.

For the variance we have

$$\mathbb{V}\left[Z_t|I_{t-1}\right]=\mathbb{E}\left[\mathbb{V}\left[Z_t|N_t,I_{t-1}\right]|I_{t-1}\right]+\mathbb{V}\left[\mathbb{E}\left[Z_t|N_t,I_{t-1}\right]|I_{t-1}\right].$$

Separately evaluating the two components, we obtain first

$$\begin{aligned}
\mathbb{E}\left[\mathbb{V}\left[Z_t|N_t,I_{t-1}\right]|I_{t-1}\right] &= \sum_{m=0}^{\infty} P\left(N_t=m|I_{t-1}\right)\times\mathbb{V}\left[Z_t|N_t=m,I_{t-1}\right]\\
&= P\left(N_t=0|I_{t-1}\right)\times 0+\sum_{m=1}^{\infty} P\left(N_t=m|I_{t-1}\right)\times\frac{m}{\varsigma}\\
&= \frac{\lambda_t}{\varsigma}. \qquad (41)
\end{aligned}$$

For the second element we can write

$$\mathbb{V}\left[\mathbb{E}\left[Z_t|N_t,I_{t-1}\right]|I_{t-1}\right]=\mathbb{E}\left[\mathbb{E}\left[Z_t|N_t,I_{t-1}\right]^2|I_{t-1}\right]-\left(\mathbb{E}\left[\mathbb{E}\left[Z_t|N_t,I_{t-1}\right]|I_{t-1}\right]\right)^2,$$

where the first term is given as

$$\begin{aligned}
\mathbb{E}\left[\mathbb{E}\left[Z_t|N_t,I_{t-1}\right]^2|I_{t-1}\right] &= \sum_{m=0}^{\infty} P\left(N_t=m|I_{t-1}\right)\times\mathbb{E}\left[Z_t|N_t=m,I_{t-1}\right]^2\\
&= P\left(N_t=0|I_{t-1}\right)\times 1+\sum_{m=1}^{\infty} P\left(N_t=m|I_{t-1}\right)\times m^2\\
&= e^{-\lambda_t}+\left(\lambda_t+\lambda_t^2\right),
\end{aligned}$$

from the second order moment of a Poisson and using the fact that the contribution of the zero-jump component to the second order moment is equal to zero. By the law of iterated expectations,

$$\mathbb{E}\left[\mathbb{E}\left[Z_t|N_t, I_{t-1}\right]|I_{t-1}\right] = \mathbb{E}\left[Z_t|I_{t-1}\right] = e^{-\lambda_t} + \lambda_t.$$

■

The filtered expected jumps are given by

$$\begin{aligned}
\mathbb{E}\left[Z_t|I_t\right] &= \sum_{m=0}^{\infty} P\left(N_t = m|I_t\right) \times \mathbb{E}\left[Z_t|N_t = m, I_t\right] \\
&= \sum_{m=0}^{\infty} P\left(N_t = m|I_t\right) \times \mathbb{E}\left[Z_t|N_t = m, I_{t-1}\right] \qquad (42)
\end{aligned}$$

where $P\left(N_t = m|I_t\right)$ is derived in equation (29).

A.4 Proof of Proposition 3

The expected value of RM_t conditional on I_{t-1} is obtained, noting that μ_t is measurable on I_{t-1}, so that

$$\begin{aligned}
\mathbb{E}\left[RM_t|I_{t-1}\right] &= \mu_t \mathbb{E}\left[\eta_t|I_{t-1}\right] \\
&= \mu_t \sum_{m=0}^{\infty} \mathbb{E}\left[\eta_t|N_t = m, I_{t-1}\right] \times P\left(N_t = m|I_{t-1}\right) \\
&= \mu_t \mathbb{E}\left[Z_t|I_{t-1}\right].
\end{aligned}$$

It follows that

$$\mathbb{E}\left[RM_t|I_{t-1}\right] = \mu_t \left(e^{-\lambda_t} + \lambda_t\right), \qquad (43)$$

where the conditional expected value of Z_t is derived in (40). The conditional variance of RM_t is obtained as

$$\begin{aligned}
\mathbb{V}\left[RM_t|I_{t-1}\right] &= \mathbb{E}\left[RM_t^2|I_{t-1}\right] - \mathbb{E}\left[RM_t|I_{t-1}\right]^2 \\
&= \mu_t^2 \left\{\left[\frac{\lambda_t}{\varsigma} + e^{-\lambda_t} + (\lambda_t + \lambda_t^2)\right](1 + \nu^{-1}) - (e^{-\lambda_t} + \lambda_t)^2\right\}. \qquad (44)
\end{aligned}$$

■

A.5 Proof of Theorem 1

Let $y_t \equiv RM_t$, the vector form of the process in (26) is

$$z_t = b_t + A_t z_{t-1}$$

where

$$z_t = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-q+1} \\ \mu_t \\ \vdots \\ \mu_{t-p+1} \end{bmatrix}, \quad b_t = \begin{bmatrix} \omega\eta_t \\ \vdots \\ 0 \\ \omega \\ \vdots \\ 0 \end{bmatrix}$$

and

$$A_t = \begin{bmatrix} \alpha_1\eta_t & \dots & \alpha_{q-1}\eta_t & \alpha_q\eta_t & \beta_1\eta_t & \dots & \beta_{p-1}\eta_t & \beta_p\eta_t \\ 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & 1 & 0 & 0 & \dots & 0 & 0 \\ \alpha_1 & \dots & \alpha_{q-1} & \alpha_q & \beta_1 & \dots & \beta_{p-1} & \beta_p \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

A_t is a $(p+q) \times (p+q)$ matrix with positive and independent coefficients. Further, $A = \mathbb{E}[A_t]$ and $b = \mathbb{E}[b_t]$ do not depend on t. Because the proof follows closely that of Theorem 2.5 in Franq and Zakoian (2010) we give only a sketch. Given the condition in (28), we can construct a stationarity solution. For $t, k \in \mathbb{Z}$, we define $\mathbb{R}^d$-valued vectors as follows:

$$Z_k(t) = \begin{cases} 0 & \text{if } k < 0 \\ b_t + A_t Z_{k-1}(t-1) & \text{if } k \geq 0 \end{cases}$$

With a multiplicative norm, i.e. $\|A\| = \sum |a_{ij}|$, we have, for any random matrix A with positive coefficients, $\mathbb{E}\|A\| = \mathbb{E}\sum_{i,j}|a_{i,j}| = \|\mathbb{E}[A]\|$. For $k > 0$

$$\mathbb{E}\|Z_k(t) - Z_{k-1}(t-1)\| = \|\mathbb{E}[A_t A_{t-1} \dots A_{t-k+} b_{t-k}]\|$$

because the matrix $A_t A_{t-1} \dots A_{t-k+} b_{t-k}$ is positive. All the terms of the product are independent (because the process $\{\eta_t\}$ is i.i.d. and every term is function of a variable η_{t-j}). Provided that $A \equiv \mathbb{E}[A_t]$ and $b = \mathbb{E}[b_t]$ do not depend on t, it follows that

$$\mathbb{E}\|Z_k(t) - Z_{k-1}(t-1)\| = \|A^k b\| = \iota' A^k b$$

where $\iota = (1, \dots, 1)'$, because all the elements in the vector $A^k b$ are positive. The condition (28) implies that the eignevalues of A are strictly less than one. The characteristic polynomial can be expressed as

$$\det(\lambda I_{p+q} - A) = \lambda^{p+q}\Big(1 - (e^{-\lambda} + \lambda)\sum_{j=1}^{q}\alpha_j\lambda^{-j} - \sum_{i=1}^{p}\beta_i\lambda^{-i}\Big)$$

When $|\lambda| \geq 1$, using the inequality $|a-b| \geq |a| - |b|$ we have

$$|\det(\lambda I_{p+q} - A)| \geq \left|\Big(1 - (e^{-\lambda} + \lambda)\sum_{j=1}^{q}\alpha_j\lambda^{-j} - \sum_{i=1}^{p}\beta_i\lambda^{-i}\Big)\right| \geq 1 - (e^{-\lambda} + \lambda)\sum_{j=1}^{q}\alpha_j - \sum_{i=1}^{p}\beta_i > 0$$

it follows that the spectral radius of A is less than one, i.e. $\rho(A) < 1$. This implies that $A^k \to 0$ at exponential rate as $k \to \infty$. For any fixed t, $Z_k(t)$ converges almost surely as $k \to \infty$. Let z_t denote the limit of $\{Z_k(t)\}_{k\in\mathbb{Z}}$. For fixed k, the process $\{Z_k(t)\}_{t\in\mathbb{Z}}$ is strictly stationary. It follows that the limit process $z(t)$ is strictly stationary and it is a solution of (27).

■

A.6 Proof of Theorem 2

Differently from Theorem 1, now A_t is a $(p+q)\times(p+q)$ matrix with positive coefficients but not independent, since η_t is correlated through λ_t. The matrix A_t can be written as the Hadamard product of two matrices, i.e.

$$A_t = A \odot E_t$$

with

$$A = \begin{bmatrix} \alpha_1 & \dots & \alpha_{q-1} & \alpha_q & \beta_1 & \dots & \beta_{p-1} & \beta_p \\ 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & 1 & 0 & 0 & \dots & 0 & 0 \\ \alpha_1 & \dots & \alpha_{q-1} & \alpha_q & \beta_1 & \dots & \beta_{p-1} & \beta_p \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

and a $(p+q)\times(p+q)$ matrix

$$E_t = \begin{bmatrix} \eta_t & \eta_t & \dots & \eta_t \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix}.$$

The conditional mean of η_t is

$$E[\eta_t|I_{t-1}] = e^{-\lambda_t} + \lambda_t.$$

whereas, the unconditional mean satisfies

$$E[\eta_t] = E[e^{-\lambda_t}] + E[\lambda_t] = E[e^{-\lambda_t}] + \frac{\phi_1}{1-\phi_2}.$$

In the last equality, since $0 < E[e^{-\lambda_t}] < 1$, it follows that $E[e^{-\lambda_t}] + \frac{\phi_1}{1-\phi_2} < 1 + \frac{\phi_1}{1-\phi_2}$.

The sequence $\{A_t, t \in \mathbb{Z}\}$ is ergodic and strictly stationary. With a multiplicative norm, i.e. $\|A\| = \sum |a_{ij}|$, $\log\|A_t\| \le \log\|A\| + \log\|E_t\|$, with $\log\|E_t\| = \log[(p+q)(\eta_t + (p+q) - 1)]$. Therefore $\log^+\|A_t\| \le \log\|A\| + \log^+\|E_t\|$, where $\log^+(x) = \max(\log(x), 0)$. Given that the Lyapunov exponent γ is equal to, see Franq and Zakoian (2010, Theorem 2.3),

$$\gamma = \lim_{t\to\infty} a.s. \frac{1}{t} \log\|A_t A_{t-1} \dots A_1\| \tag{45}$$

and

$$\log\left(\|A_t A_{t-1}\dots A_1\|\right) \leq \log\|A^t\| + \sum_{i=1}^{t}\log\|E_i\|$$

Since, $\lim_{t\to\infty}\frac{1}{t}\log\|A^t\| = \log\{\rho(A)\}$, $\gamma < 0$ if and only if

$$\rho(A) < \exp\left(-E[\log\|E_t\|]\right). \tag{46}$$

Now, we turn to the proof of the existence of a stationary and ergodic solution if the condition in (46) is satisfied, i.e. $\gamma < 0$. Since the random variable η_t has finite variance, the components of the matrix A_t are integrable. Hence,

$$\mathbb{E}[\log^+\|A_t\|] \leq \mathbb{E}\|A_t\| < \infty.$$

With $\gamma < 0$ it follows from (45) that

$$\tilde{z}_t(N) = b_t + \sum_{n=0}^{N} A_t A_{t-1}\dots A_{t-n} b_{t-n-1}$$

converges a.s. when N goes to infinity, to some limit $\tilde{z}_t$. Using the multiplicative norm

$$\|\tilde{z}_t(N)\| \leq \|b_t\| + \sum_{n=0}^{\infty}\|A_t A_{t-1}\dots A_{t-n}\|\|b_{t-n-1}\|$$

and

$$\begin{aligned}\|A_t A_{t-1}\dots A_{t-n}\|^{1/n}\|b_{t-n-1}\|^{1/n} &= \exp\left[\frac{1}{n}\|A_t A_{t-1}\dots A_{t-n}\| + \frac{1}{n}\|b_{t-n-1}\|\right]\\ &\overset{a.s.}{\to} \exp(\gamma) < 1.\end{aligned}$$

To show that $n^{-1}\log\|b_{t-n-1}\| \to 0$ we have used the result that for a sequence X_n of identically distributed random variables admitting an expectation holds that $X_n/n \overset{a.s.}{\to} 0$ when $n \to \infty$. In our case this can be applied because $\mathbb{E}|\log\|b_{t-n-1}\|| < \infty$, see Franq and Zakoian (2010, Proof of Theorem 2.4, p.31). Let $\tilde{z}_{q+1,t}$ denote the $(q+1)$-th element of $\tilde{z}_t$. Setting $y_t = \tilde{z}_{q+1,t}\eta_t$, we define a solution of model (6). This solution is nonanticipative because y_t can be expressed as a measurable function of $\eta_t, \eta_{t-1}, \dots$. By the ergodicity of η_t this solution is also strictly stationary and ergodic.

■

B Web Appendix

Notes on Gamma-distributed random variables

The density of the two-parameter Gamma random variable $X \sim \Gamma(\delta, \nu)$ in *scale-shape* is

$$f_X(x) = \frac{1}{x}\left(\frac{x}{\delta}\right)^{\nu} \frac{1}{\Gamma(\nu)} e^{-\frac{x}{\delta}}, \quad x \geq 0 \tag{47}$$

where ν is the shape parameter and δ is the scale parameter. In this case we have $\mathbb{E}[x] = \nu\delta$ and $\mathbb{V}[x] = \nu\delta^2$. Special parameter combination: if we impose $\delta = \frac{1}{\nu}$, we have a Gamma with unit mean and variance $\frac{1}{\nu}$. A Gamma in *mean-shape* representation $X \sim \Gamma(\mu, \nu)$, is such that $\mathbb{E}[x] = \mu$ and $\mathbb{V}[x] = \left(\frac{\mu}{\nu}\right)^2 \nu = \frac{\mu^2}{\nu}$.

Transformations of Gamma random variables: If the scale-shape Gamma distribution is multiplied by a scalar we have

$$\tilde{X} = \alpha X \sim \Gamma(\alpha\delta, \nu)$$

In that case, mean and variance become $\mathbb{E}[\tilde{x}] = \nu\alpha\delta$ and $\mathbb{V}[\tilde{x}] = \nu\alpha^2\delta^2$. For the mean-shape Gamma we have

$$\tilde{X} = \alpha X \sim \Gamma(\alpha\mu, \nu)$$

and $\mathbb{E}[\tilde{x}] = \alpha\mu$ and $\mathbb{V}[\tilde{x}] = \left(\frac{\alpha\mu}{\nu}\right)^2 \nu = \frac{\alpha^2\mu^2}{\nu}$.

Sum of Gamma random variables: The Gamma distribution is closed under aggregation in a specific case: the sum of Gamma distributions must have the same scale parameter. Under scale-shape Gamma, this translates into

$$X_i \sim \Gamma(\delta, \nu_i)$$

$$Y = \sum_{i=1}^{k} X_i \sim \Gamma\left(\delta, \sum_{i=1}^{k} \nu_i\right)$$

$$\mathbb{E}[Y] = \delta \sum_{i=1}^{k} \nu_i,$$

$$\mathbb{V}[Y] = \delta^2 \sum_{i=1}^{k} \nu_i.$$

It follows directly from this properties of the Gamma distribution that

$$Z_t | N_t = m > 0, I_{t-1} = \sum_{j=1}^{m} Y_{j,t} \sim \Gamma(m, m\varsigma)$$

given that $Y_{j,t} \sim \Gamma(1, \varsigma)$.

The K distribution

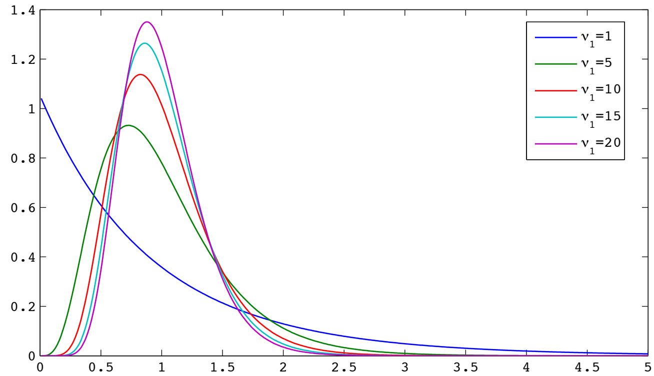
If $X \sim \Gamma(\mu, \nu_1)$, in mean-shape form, and $Y|X \sim \Gamma(X, \nu_2)$ then we have $Y \sim K(\mu, \nu_1, \nu_2)$ such that

$$f(y) = \frac{2}{y}\left(y\frac{\nu_1\nu_2}{\mu}\right)^{\frac{\nu_1+\nu_2}{2}} \frac{1}{\Gamma(\nu_1)\Gamma(\nu_2)}\mathbb{K}_{\nu_1-\nu_2}\left(2\sqrt{y\frac{\nu_1\nu_2}{\mu}}\right), \quad y \geq 0$$

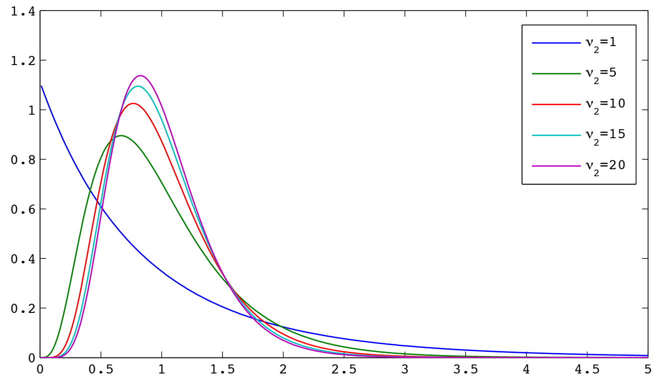
where $\mathbb{K}_a(\cdot)$ is the modified Bessel function of the second kind. Moments of the K density are given by

$$\mathbb{E}[y^s] = \frac{\mu^s\Gamma(\nu_1+s)\Gamma(\nu_2+s)}{\nu_1^s\nu_2^s\Gamma(\nu_1)\Gamma(\nu_2)}$$

Therefore we have $\mathbb{E}[Y] = \mu$ and $\mathbb{V}[Y] = \mu^2\left(\frac{\nu_1+\nu_2+1}{\nu_1\nu_2}\right)$. Furthermore, with respect to the scaling we have that μ is a scale parameter and therefore if $Y \sim K(\mu, \nu_1, \nu_2)$ then $\alpha Y \sim K(\alpha\mu, \nu_1, \nu_2)$, $\mathbb{E}[\alpha Y] = \alpha\mu$ and $\mathbb{V}[\alpha Y] = \alpha^2\mu^2\left(\frac{\nu_1+\nu_2+1}{\nu_1\nu_2}\right)$. Different K densities, corresponding to different values of ν_1 and ν_2, are plotted in Figure 6. Increasing both parameters reduces the variance, as it is apparent in both plots.



(a) K density calculated with $\nu_1 = \{1, 5, 10, 15, 20\}$ and $\nu_2 = 10$.



(b) K density calculated with $\nu_2 = \{1, 5, 10, 15, 20\}$ and $\nu_1 = 10$.

Figure 6: K density computed for different values of ν_1 (upper panel) and ν_2 (lower panel).

Further, as noted by Redding (1999, p.3), the product of two independent Gamma random variables, $Z \sim \Gamma(1, \nu_2)$ and $X \sim \Gamma(\mu, \nu_1)$, is

$$Y = Z \cdot X \sim K(\mu, \nu_1, \nu_2)$$

with density given by the following integral

$$f_Y(y) = \int_{-\infty}^{\infty} \frac{1}{x} f_Z\left(\frac{y}{x}\right) f_X(x) \, \mathrm{d}\, x.$$

Pseudo random numbers with K distribution can be generated from $Z \sim \Gamma(1, L)$ and $X \sim \Gamma(\mu, \nu)$, since $Y = Z \cdot X$ is distributed as $K(\mu, \nu, L)$.

Setting $\theta_1 = \nu_1, \theta_2 = \nu_2$ or viceversa, the cumulative distribution function (CDF) of a K-distributed random variable can be written as

$$F(y; \mu, \theta_1, \theta_2) = \frac{2^{2-\theta_1-\theta_2}}{\Gamma(\theta_1)\Gamma(\theta_1)} \int_0^{2\sqrt{\theta_1\theta_2 y/\mu}} t^{\theta_1+\theta_2-1} \mathbb{K}_{\theta_1-\theta_2}(t) \mathrm{d}t, y \geq 0. \tag{48}$$

The hypothesis of $\theta_2 \in \mathbb{N}$ instead of $\theta_2 \in \mathbb{R}_+$ is required in order to obtain the previous expression in closed form. Writing $\zeta = \theta_1 - \theta_2$, $k = 2\nu_2 - 1$ and $z = 2\sqrt{\theta_1\theta_2 y/\mu}$

$$F(y; \mu, \theta_1, \theta_2) = 1 + \frac{2^{2-\theta_1-\theta_2}}{\Gamma(\theta_1)\Gamma(\theta_2)} g(z, \zeta, k) \tag{49}$$

where

$$g(y, \zeta, k) = \begin{cases} -z^{\zeta+1}\mathbb{K}_{(\zeta+1)}(z) & k = 1 \\ (k-1)(2\zeta + k - 1)g(y, \zeta, k-2) - z^{\zeta+k}\mathbb{K}_{\zeta+1}(z) - (k-1)z^{(\zeta+k-1)}\mathbb{K}_{\zeta}(z) & \text{elsewhere} \end{cases}$$

The number of required recursions to compute a single value $F_Y(y; \mu, \theta_1, \theta_2)$ is θ_2. The best parametrization, in terms of computational speed, is $\theta_1 = \max\{\nu_1, \nu_2\}$ and $\theta_1 = \min\{\nu_1, \nu_2\}$.

Tables

	Mean	St.Dev.	Max	Min	Kurtosis	Skewness	ρ_1	ρ_5	ρ_{22}	T
S&P500	0.009	0.006	0.078	0.002	21.410	3.255	0.840	0.721	0.540	3261
FTSE100	0.008	0.005	0.072	0.001	21.079	2.957	0.805	0.697	0.531	3280
DAX	0.011	0.007	0.071	0.002	11.827	2.272	0.852	0.728	0.570	3312
DJIA	0.009	0.006	0.080	0.002	24.480	3.480	0.830	0.718	0.530	3263
NASDAQ	0.010	0.007	0.075	0.002	13.385	2.457	0.826	0.704	0.567	3266
CAC 40	0.011	0.006	0.066	0.002	12.991	2.324	0.832	0.714	0.529	3329
BOVESPA	0.013	0.006	0.079	0.003	20.660	3.138	0.755	0.610	0.400	3189
BA	0.014	0.007	0.094	0.004	18.777	3.039	0.828	0.721	0.571	2391
BAC	0.020	0.020	0.262	0.003	22.633	3.565	0.898	0.804	0.695	2391
C	0.023	0.022	0.293	0.004	26.171	3.772	0.904	0.788	0.650	2391
CAT	0.016	0.009	0.127	0.005	22.004	3.389	0.862	0.762	0.578	2391
FDX	0.014	0.007	0.087	0.004	13.938	2.588	0.834	0.737	0.611	2391
HON	0.015	0.008	0.128	0.004	32.393	3.860	0.822	0.711	0.506	2391
HPQ	0.015	0.007	0.097	0.005	22.426	3.256	0.778	0.647	0.455	2391
IBM	0.011	0.007	0.101	0.004	33.984	4.400	0.851	0.762	0.554	2391
JPM	0.018	0.014	0.160	0.004	18.676	3.244	0.905	0.788	0.651	2391
KFT	0.011	0.005	0.078	0.004	21.710	3.224	0.726	0.588	0.411	2391
PEP	0.010	0.006	0.120	0.004	74.624	5.693	0.735	0.645	0.499	2391
PG	0.010	0.005	0.107	0.003	66.383	5.520	0.768	0.656	0.480	2391
T	0.013	0.008	0.127	0.004	32.884	3.982	0.833	0.747	0.577	2391
TWX	0.016	0.008	0.114	0.005	20.938	3.331	0.861	0.766	0.628	2391
TXN	0.018	0.007	0.094	0.006	13.308	2.252	0.757	0.670	0.496	2391
WFC	0.018	0.017	0.154	0.004	15.731	3.124	0.909	0.827	0.700	2391

Table A.1: Descriptive statistics of bipower variation for 7 stock indexes and 16 NYSE stocks. ρ_1, ρ_5 and ρ_{22} are the autocorrelations at 1, 5 and 22 lags. T is the sample size.

	ω	α_1	α_2	α_3	β	γ	ξ	ν	$LogLL$	Q_1	Q_{10}	Q_{22}	$LR_{\theta=1}$
SP500	0.0004^a	0.3351^a	0.1820^a	0.1223^a	0.3099^a	0.1177^a	0.0508^a	278.5^a	15965	0.8523	0.4188	0.3381	513.3^a
FTSE 100	0.0003^a	0.3235^a	0.2429^a	0.1471^a	0.2688^a	0.0703^a	0.0558^a	269.7^a	16396	0.2639	0.6873	0.0071	523.5^a
DAX	0.0003^a	0.3410^a	0.1315	0.1311^a	0.3663^a	0.0852^a	0.0536^a	289.2^a	15457	0.0275	0.0949	0.0000	450.3^a
DJIA	0.0004^a	0.3360^a	0.2047^a	0.1187^a	0.2965^a	0.1055^a	0.0575^a	239.1^a	16032	0.6716	0.3770	0.5112	522.8^a
NSDQ	0.0004	0.3439^b	0.2043	0.1682	0.2442	0.1082^a	0.0749^a	191.2^a	15483	0.5960	0.1164	0.0678	468.4^a
CAC	0.0003^a	0.3305^a	0.1493^a	0.1116^a	0.3761^a	0.0915^a	0.0775^a	205.6^a	15749	0.0544	0.0513	0.0000	439.4^a
BOVESPA	0.0011	0.3679^a	0.1726	0.1750^b	0.2196	0.0764^a	0.0406^a	328.9^a	14135	0.3317	0.4748	0.1699	529.9^a
BA	0.0006^a	0.4066^a	0.0895	0.1377^a	0.3368^a	0.0310^a	0.0582^a	364.2^a	10732	0.7159	0.9053	0.0494	345.5^a
BAC	0.0003^a	0.5445^a	0.0838	0.1488^a	0.2244^a	0.0394^a	0.0332^a	593.2^a	10449	0.4125	0.9090	0.7371	502.8^a
C	0.0003^a	0.5237^a	0.1056^c	0.1166^a	0.2570^a	0.0263^a	0.0665^a	353.3^a	10207	0.1525	0.3626	0.1606	315.9^a
CAT	0.0009^a	0.4233^a	0.2379^a	0.1278^a	0.1690^b	0.0471^a	0.0607^a	356.7^a	10511	0.4747	0.6159	0.8139	340.7^a
FDX	0.0005^a	0.4335^a	0.0946^c	0.1951^a	0.2610^a	0.0332^a	0.0500^a	402.4^a	10709	0.9584	0.2441	0.4279	386.0^a
HON	0.0009^a	0.4162^a	0.1131	0.0960^a	0.3208^a	0.0560^a	0.0482^a	451.0^a	10664	0.3389	0.7469	0.1004	325.5^a
HPQ	0.0010^a	0.4253^a	0.1801^a	0.1438^a	0.1983^a	0.0545^a	0.0487^a	411.1^a	10555	0.8556	0.6755	0.4401	440.6^a
IBM	0.0006^a	0.4503^a	0.1410^b	0.0785^a	0.2866^a	0.0406^a	0.0706^a	339.5^a	11448	0.5940	0.9716	0.4653	291.2^a
JPM	0.0005^a	0.5003^a	0.2069^a	0.1283^a	0.1547^b	0.0355^a	0.0633^a	338.8^a	10509	0.2365	0.6009	0.2658	372.1^a
KFT	0.0011^a	0.3763^a	0.2875^a	0.2132^a	0.08	0.0004	0.0442^a	353.4^a	11107	0.4789	0.6887	0.9127	522.8^a
PEP	0.0004^a	0.3444^a	0.0704	0.0923^a	0.4630^a	0.0300^a	0.0487^a	436.8^a	11551	0.6325	0.4771	0.0054	356.6^a
PG	0.0007^a	0.4082^a	0.2353^a	0.1156^a	0.1882^a	0.0357^a	0.0513^a	431.7^a	11717	0.8590	0.9577	0.0296	345.6^a
T	0.0005^a	0.4287^a	0.1198^b	0.1222^a	0.3004^a	0.0295^a	0.0526^a	461.5^a	11142	0.8922	0.8618	0.0225	321.2^a
TWX	0.0007^a	0.4147^a	0.1409^b	0.1359^a	0.2734^a	0.0321^a	0.0530^a	534.7^a	10831	0.8800	0.9901	0.4976	318.4^a
TXN	0.0009^a	0.3786^a	0.1688^c	0.1328^a	0.2843^a	0.0336^a	0.0756^a	316.3^a	10267	0.6477	0.8915	0.1741	293.9^a
WFC	0.0004^a	0.4784^a	0.1939^b	0.1711^a	0.1695^c	0.0157	0.0453^a	407.8^a	10503	0.4608	0.8614	0.5197	496.9^a

Table A.2: Estimates of the Asymmetric HAR-MEM with Generalized Gamma distribution. The upper part of the table reports the results for several stock indexes while the lower part refers to 16 NYSE stocks. a, b and c stand for significance at 1%, 5% and 10% respectively. Q_1, Q_{10} and Q_{22} are the p-values of the Ljung-Box test for absence of autocorrelation in the residuals, where the latter are computed as $\hat{\epsilon}_t = \frac{RM_t}{E[RM_t|I_{t-1}]}$.

	ω^L	α_1^L	α_2^L	α_3^L	β^L	γ^L	ν^L	ω^S	α_1^S	α_2^S	α_3^S	β^S	γ^S	ν^S	π
SP500	0.0003^a	0.2829^a	0.1770^a	0.1456^a	0.2822^a	0.0996^a	23.4970^a	0.0006^a	0.4463^a	0.0928^a	0.0001	0.4113^a	0.1686^a	8.3977^a	0.8361^a
FTSE 100	0.0002	0.2649^a	0.1999^a	0.1681^a	0.3006^a	0.0544	24.9151^a	0.0010^b	0.6570^a	0.0011	0.0001	0.2589^a	0.2082^a	5.2148^a	0.9175^a
DAX	0.0000	0.1920^a	0.0272^a	0.0817^a	0.6517^a	0.0600^a	31.0223^a	0.0012^a	0.6165^a	0.0381	0.0001	0.2299^a	0.1137^a	11.3824^a	0.6872^a
DJIA	0.0003^a	0.2600^a	0.1810^a	0.1387^a	0.3130^a	0.0877^a	23.6464^a	0.0006	0.5105^a	0.1590^a	0.0001	0.2908^a	0.1523^a	8.5515^a	0.8114^a
NSDQ	0.0002^b	0.2708^a	0.2099^a	0.1749^a	0.2575^a	0.0945^a	24.6208^a	0.0012^a	0.5549^a	0.0448^a	0.0218^c	0.2538^a	0.1476^a	7.6523^a	0.8234^a
CAC	0.0001^b	0.2094^a	0.001	0.0899^a	0.6377^a	0.0654^a	34.3688^a	0.0006	0.5039^a	0.2475^a	0.0001	0.1867^a	0.1282^a	10.3025^a	0.6855^a
B0VESPA	0.0007^a	0.2936^a	0.0539^a	0.1804^a	0.3546^a	0.0604^a	24.1404^a	0.0022^a	0.5223^a	0.3358^a	0.0001	0.0685^a	0.1057^a	8.9983^a	0.7861^a
BA	0.0006^a	0.3523^a	0.0894	0.1801^a	0.2932	0.0187	39.8421^a	0.0005^a	0.4671	0.001	0.0001^a	0.5247^a	0.0942^a	11.5086^b	0.8044^a
BAC	0.0004^a	0.4836^a	0.1093^a	0.1362^a	0.2126^a	0.0308^a	29.4516^a	0.0000^a	0.8779^a	0.0010^a	0.0427^a	0.2653^a	0.2608^a	6.7612^a	0.9459^a
C	0.0003^a	0.4159^a	0.1358^a	0.1097	0.2875^a	0.0169^a	44.8703^a	0.0000^a	0.7817^b	0.0010^a	0.0986^a	0.1767^a	0.0713	11.3158^a	0.7967^a
CAT	0.0007^a	0.3520^a	0.1846^a	0.1661^a	0.2024^a	0.0376^a	35.7334^a	0.0016^a	0.6042^a	0.3488^a	0.0001	0.0580^a	0.1004^a	11.9857^a	0.8426^a
FDX	0.0004^a	0.3534^a	0.1488^a	0.2099^a	0.2178^a	0.0362^a	31.9676^a	0.0020^a	0.8082^a	0.0010^a	0.0001	0.2302^a	0.0170^a	10.1717^a	0.8951^a
HON	0.0004^a	0.2453^a	0.001	0.0994^a	0.5845^a	0.0416^a	43.6346^a	0.0012^a	0.7667^a	0.0653^a	0.0001	0.1477^a	0.0782^a	15.6933^a	0.6987^a
HPQ	0.0008^a	0.3864^a	0.1635^a	0.1525^a	0.2092^a	0.0395^a	31.6481^a	0.0041^a	0.6172^a	0.1140^a	0.0003^a	0.1003^a	0.3114^a	6.1879^a	0.9291^a
IBM	0.0006^a	0.3486^a	0.1661^a	0.0982^a	0.2935^a	0.0248^a	38.6967^a	0.0004^b	0.8022^a	0.0010^a	0.0001	0.2254^a	0.1332^a	13.0029^a	0.8459^a
JPM	0.0007^a	0.3457^a	0.2228^a	0.1837^a	0.1396^a	0.0336^a	39.8798^a	0.0004	0.8427^a	0.0278^a	0.0001	0.2156^a	0.0374^a	15.4291^a	0.7537^a
KFT	0.0011^a	0.2904^a	0.2667^a	0.2184^a	0.0743^a	0.0001^a	27.6085^a	0.0004^a	1.0000^a	0.1615^a	0.0001^a	0.1408^a	0.0449^a	8.7561^a	0.9064^a
PEP	0.0003^a	0.2441^a	0.001	0.0690^a	0.6298^a	0.0317^a	34.8980^a	0.0005^b	0.7892^a	0.0010^a	0.0001	0.2831^a	0.0890^a	10.8392^a	0.8809^a
PG	0.0008^a	0.3534^a	0.1920^a	0.1720^a	0.1500^a	0.0292^a	43.2445^a	0.0001^a	0.5091^a	0.2847^a	0.0001^a	0.2654^a	0.0631^a	11.1629^a	0.7847^a
T	0.0006^a	0.3307^a	0.1191	0.1437^a	0.3294^a	0.0148^a	41.4088^a	0.0001^a	0.7872^a	0.0570^a	0.0004^a	0.2213^a	0.0957^a	12.6061^a	0.8329^a
TWX	0.0006^a	0.3510^a	0.1430^a	0.1561^a	0.2746^a	0.0270^a	43.7099^a	0.0013^b	1.0000^a	0.1462^a	0.0001	0.0010^a	0.0312^a	10.9094^a	0.9180^a
TXN	0.0006^a	0.2910^a	0.0010^a	0.1133^a	0.5289^a	0.0279^a	38.1906^a	0.0000	0.7348^a	0.3632^a	0.0001	0.0010^b	0.0723^a	15.7432^a	0.7943^a
WFC	0.0004^a	0.4632^a	0.1751^a	0.1654^a	0.1299^a	0.0196^a	31.2518^a	0.0000	0.1070^a	1.0000^a	0.2410^a	0.0232^a	0.0001	6.2174^a	0.9231^a

Table A.3: Estimates of the Mixture-Asymmetric HAR-MEM of Lanne. The upper part of the table reports the results for several stock indexes while the lower part refers to 16 NYSE stocks. a, b and c stand for significance at 1%, 5% and 10% respectively. Due to space constraints, the p-values of the Ljung-Box test for absence of autocorrelation in the residuals are not reported.