SYRTO

SYstemic Risk TOmography
*Signals, Measurements, Transmission Channels,
and Policy Interventions*

Chasing volatility
A persistent multiplicative error model with jumps

Massimiliano Caporin, Eduardo Rossi, Paolo Santucci de Magistris

Chasing volatility
A persistent multiplicative error model with jumps

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

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,

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

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 $Y_{j,t}$, $j = 1, 2, \ldots, N_t$, are assumed to be independent for any t.

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

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

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. $\mathbb{E}[Z_t | N_t = m > 0, I_{t-1}] = m$ and $\mathbb{V}[Z_t | N_t = m > 0, I_{t-1}] = \frac{m}{\varsigma}$. 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.

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

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

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

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*

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

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

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

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

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

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

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

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 $\Gamma(1,\nu)$ 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

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

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

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

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

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

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.

10

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

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

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* $\varepsilon_t \sim \Gamma\left(1, \nu\right)$. *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.

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

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

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

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

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):

$$z_t = b_t + A_t z_{t-1}. \qquad (27)$$

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*

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

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*

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

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, $f(RM_t|I_{t-1})$, can be computed in closed form as in equation (17).[6] Model parameters are estimated by maximizing the sample log-likelihood $\ell(\theta) = \sum_{t=1}^{T} \log f(RM_t|I_{t-1}; \theta)$, where $f(RM_t|I_{t-1})$ 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

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

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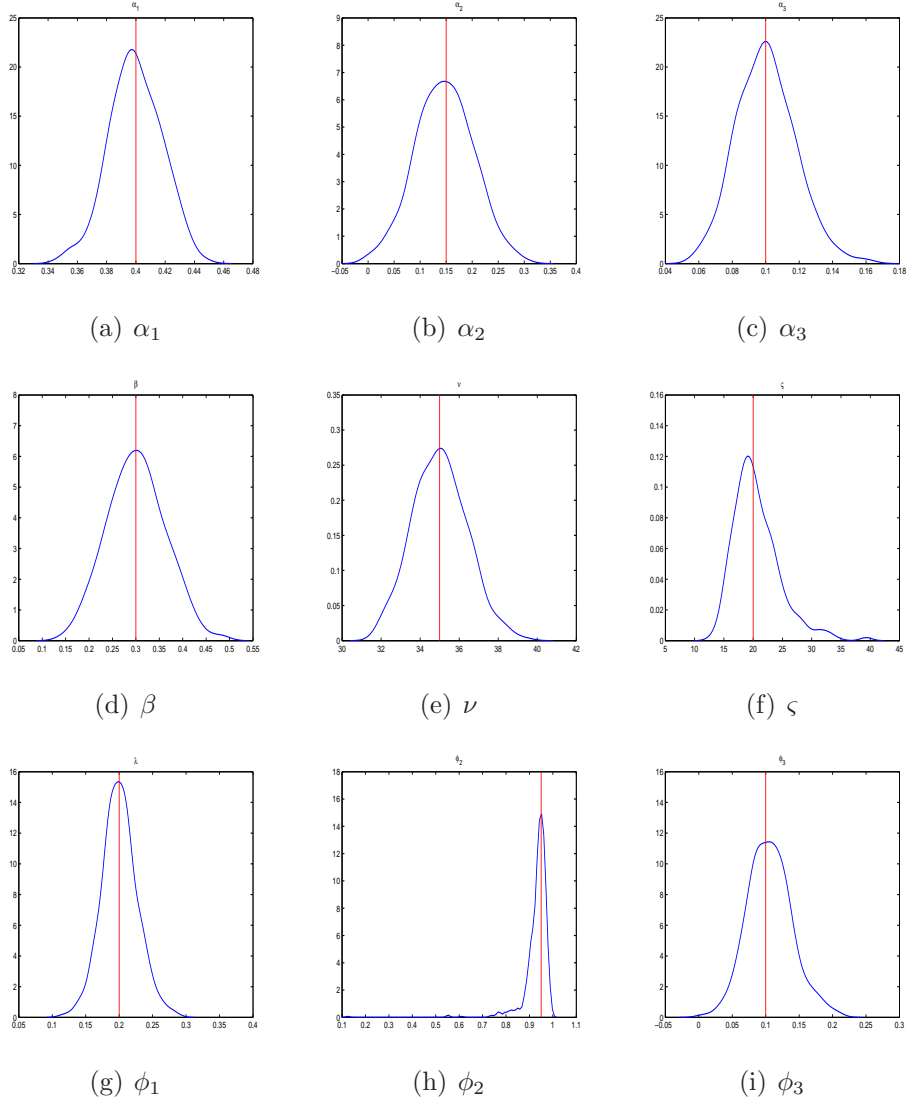
Figure 1: Kernel densities of the Monte Carlo estimates of the MEM-J-λ_t, where λ_t varies according to (20).

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, $r_{t_j}^2$, 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

$$\hat{\varepsilon}_t = \frac{RM_t - \mathbb{E}\left[RM_t | I_{t-1}\right]}{\mathbb{V}\left[RM_t | I_{t-1}\right]^{1/2}}. \tag{33}$$
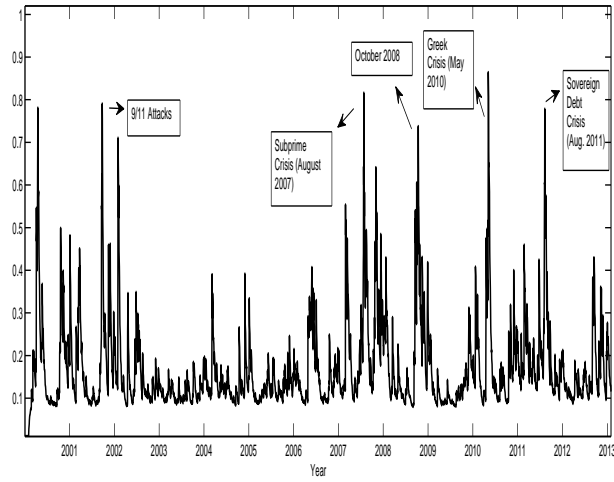
(a) λ_t of S&P500
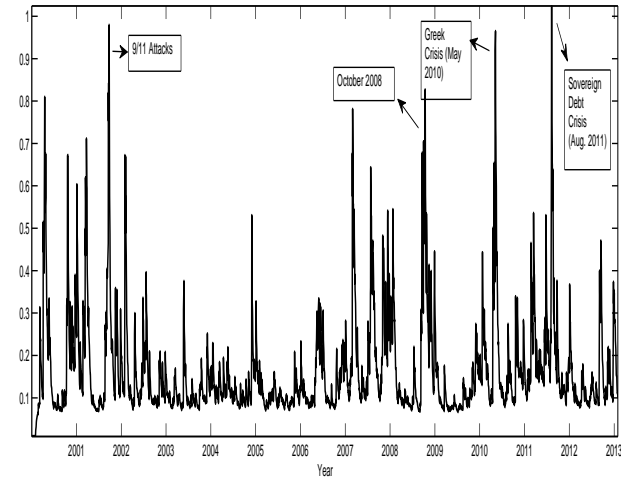
(b) λ_t of CAC
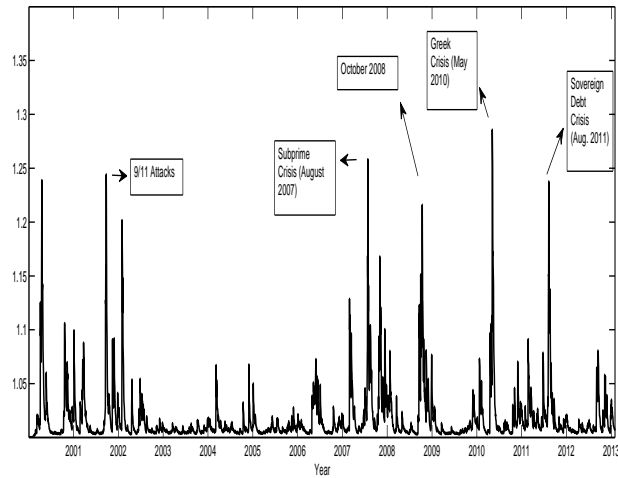
(c) $E(Z_t|I_{t-1})$ of S&P500
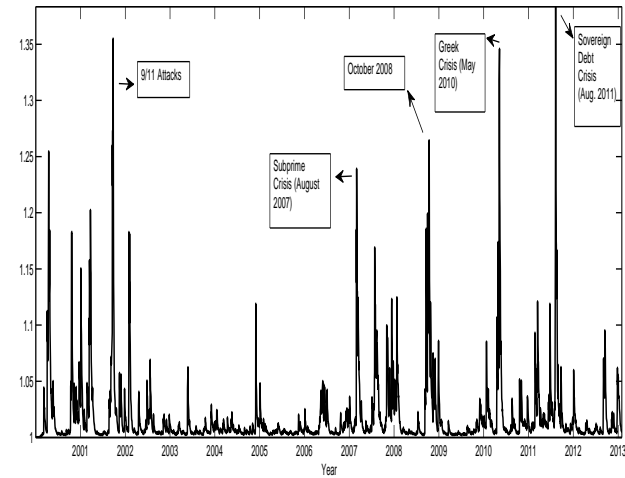
(d) $E(Z_t|I_{t-1})$ of CAC

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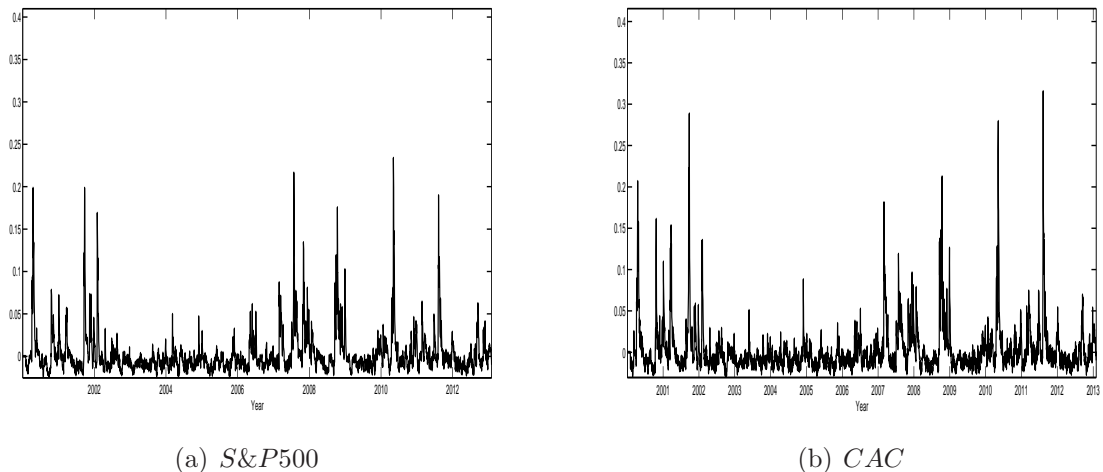
(a) $S\&P500$ (b) CAC

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.

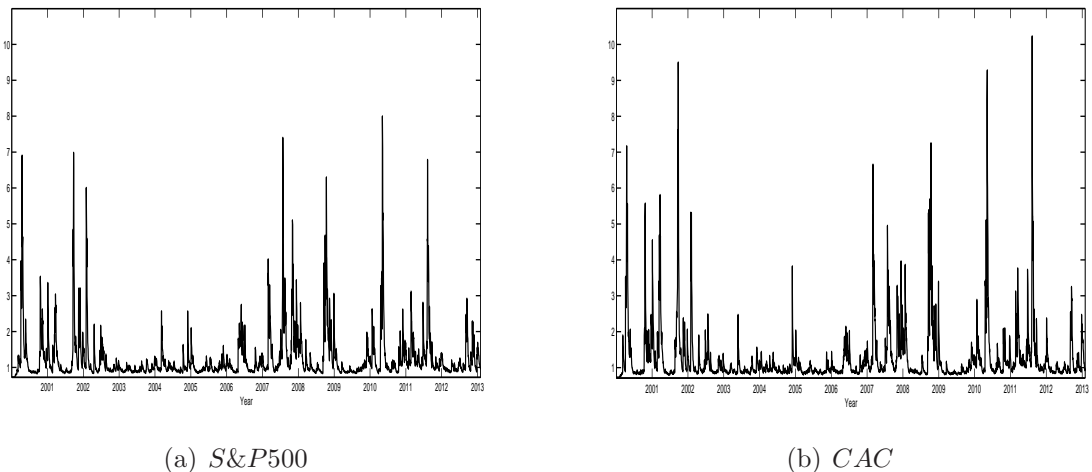(a) $S\&P500$ (b) CAC

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

25