Economic Modelling 76 (2019) 135–152

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.journals.elsevier.com/economic-modelling

On the asymmetric impact of macro–variables on volatility☆

Alessandra Amendola [a], Vincenzo Candila [a,*], Giampiero M. Gallo [b]

[a] Dept. of Economics and Statistics, University of Salerno, Italy
[b] Italian Court of Audits (Corte dei conti), Italy

ARTICLE INFO

Keywords:
Volatility
Asymmetry
GARCH–MIDAS
Forecasting

ABSTRACT

We extend the GARCH–MIDAS model to take into account possible different impacts from positive and negative macroeconomic variations on financial market volatility: a Monte Carlo simulation which shows good properties of the estimator with realistic sample sizes. The empirical application is performed on the daily *S&P*500 volatility dynamics with the U.S. monthly industrial production and national activity index as additional (signed) determinants. We estimate the Relative Marginal Effect of macro variable movements on volatility at different lags. In the out-of-sample analysis, our proposed GARCH–MIDAS model not only statistically outperforms the competing specifications (GARCH, GJR-GARCH and GARCH–MIDAS models), but shows significant utility gains for a mean-variance investor under different risk aversion parameters. Attention to robustness is given by choosing different samples and estimating the model in an international context (six different stock markets).

1. Introduction

The connection between macro–variables and market volatility has been extensively studied, possibly starting with Officer (1973), in regards to the responsiveness of macro–variable volatility to stock market volatility. Volatility in financial markets has a strong interdependence with the business cycle: while an early result by Schwert (1989) (with monthly data and very simple volatility measures) documents that asset returns exhibit larger volatility during recessions, the Great Recession has highlighted the bidirectional nature of uncertainty transmission between the real and the financial sectors, and the fact that increases in volatility occur also during expansions. In the development of more sophisticated volatility analysis following the diffusion of GARCH models (Bollerslev, 1986), the idea that volatility dynamics has two main components, one related to a slow moving long–term local average (low frequency component), and one related to short–term movements has found a variety of solutions: from early specifications in additive forms (Ding and Granger (1996), Engle and Lee (1999)), the two components are present in a number of extensions. In the Spline–GARCH by Engle and Rangel (2008) short and long–run components are combined multiplicatively, with the latter being a slowly time-varying and deterministic function of time. Within the GARCH framework, other two–component models can be found, for instance, in Amado and Teräsvirta (2013), and Bauwens and Storti (2009); within the class of Multiplicative Error Models (modeling of realized volatility measures), B–Splines are used in Brownlees and Gallo (2010) and the Engle and Lee (1999) approach is included in the Asymmetric Composite MEM in Brownlees et al. (2012) and, more recently, in Cipollini and Gallo (2018).

The dependence of the low frequency component upon macro–variables can be established *ex–post* as in Engle and Rangel (2008) or can be tackled directly. Within a GARCH framework, this requires to find a solution to the problem of having daily data for returns and monthly (at best) for macro–variables. A mixed frequency approach was proposed by Engle et al. (2013) in the GARCH–MIDAS model with two multiplicative components: the short–run component is a mean–reverting GARCH(1,1), possibly enriched by other variables observed at the same frequency (daily), while the long–run component of volatility is a smoothing filter of lower frequency observations, in the spirit of the MIDAS regression (Ghysels et al. (2007), Ghysels et al. (2005)) with a filter for lagged (and possibly forward–looking) values of the macro–variables.

Several interesting contributions are present in the recent literature for GARCH–MIDAS model: Asgharian et al. (2013) applied the model together with principal components to combine information provided by different variables such as interest rates, unemployment, inflation, exchange rates and the industrial production growth rate. Facing the traditional GARCH(1,1) as a competing model, their

☆ The views expressed in the article are those of the authors and do not involve the responsibility of the Corte dei conti.

* Corresponding author.
E-mail addresses: alamendola@unisa.it (A. Amendola), vcandila@unisa.it (V. Candila), giampiero.gallo@corteconti.it (G.M. Gallo).

https://doi.org/10.1016/j.econmod.2018.07.025
Received 24 March 2018; Received in revised form 24 July 2018; Accepted 27 July 2018
Available online 11 August 2018

Table 1
Granger test.

	Lag	1	2	3	4	5	6	7	8	9	10	11	12
$IP(lag) \xrightarrow{GnC} S\&P500_{ARV}$	F	24.012	12.728	7.039	7.186	7.492	5.792	4.902	4.249	4.031	3.535	3.360	3.243
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$NAI(lag) \xrightarrow{GnC} S\&P500_{ARV}$	F	0.961	0.996	1.929	1.471	2.459	1.875	1.807	1.618	1.921	1.604	1.754	1.604
	p-value	0.328	0.371	0.126	0.212	0.035	0.087	0.088	0.122	0.052	0.109	0.066	0.095

Notes: The table shows the F-statistics as well as their p-values from the regressions "$S\&P500_{ARV}$ on lagged $S\&P500_{ARV}$ and *IP*" (upper panel) and "$S\&P500_{ARV}$ on lagged $S\&P500_{ARV}$ and *NAI*" (lower panel). The expression "X GnC Y" means that "X does not Granger cause Y". $S\&P500_{ARV}$ stands for *S&P*500 monthly aggregated realized volatility, obtained summing the daily realized volatility. The daily realized volatility is computed by taking the square root of the summation of the 5-min squared intradaily and overnight returns. The regressions cover the period January 2000–December 2016, for a total of 204 observations.

GARCH–MIDAS showed itself to have a superior forecasting ability in terms of monthly variance. In Conrad and Loch (2015), the links between macro–variables and stock market volatility are investigated in terms of a lead–lag relationship. By applying different GARCH–MIDAS models where a number of macro–variables (for instance, lagged and expected unemployment rate, real GDP and industrial production) are included in the long-run component, the authors found that most of the specifications provide better volatility forecasts with respect to the chosen benchmark. Amendola et al. (2017) consider the Effective Federal Funds Rate (an indicator of U.S. Monetary Policy) as a volatility determinant of Crude Oil future price. The out-of-sample forecasting performance of their GARCH–MIDAS generally has a better performance than two benchmarks, the lagged realized volatility (a benchmark suggested by Asgharian et al. (2015)), and the GARCH(1,1). Recently, Mo et al. (2017) use the GARCH–MIDAS model to forecast the volatility of commodity futures in the markets of China and India. The GARCH–MIDAS has also several applications in the multivariate framework, see, for example, the works by Bauwens et al. (2017), Asgharian et al. (2015), Conrad et al. (2014) and Colacito et al. (2011), among others.

While the idea that macro–variables may have different impacts on other macro–variables according to their signs is present in the literature (Segal et al. (2015), Moshiri and Banihashem (2012), Cologni and Manera (2009), among others), to the best of our knowledge, this is the first contribution investigating how, and how much, these impacts may separately affect the volatility of the variable of interest.

In this paper we want to address the specific issue of the low–frequency component of the conditional variance of financial returns reacting differently to macro–variables growth rates according to their signs (according to the motivation detailed in Section 2). We do so by proposing three possible channels for asymmetry: the first is captured by the sign of the returns in the short–run component of GARCH (Glosten et al., 1993). The second channel of asymmetry is associated with different shapes in the function weighing lagged macro–variable realizations according to whether they are positive or negative. The third asymmetric effect comes from *sign–specific* parameters in the relationship between the weighted sum of lagged macro–variables and financial volatility in the long–run component (Section 3). It is important to stress that in the standard GARCH–MIDAS specification, having a single coefficient ends up imposing the same value on the parameter measuring the effect of the growth rate of the macro–variable on the long–run component of volatility; if such a parameter is, say, negative, negative growth rates would increase volatility, while positive ones would decrease it. This corresponds to a specific testable restriction on the sign–specific parameters in our specification (their sum being equal to zero and equal weighting functions). By contrast, the more general specification allows to see not only that they may both be *volatility–enhancing* (opposite signs), but also of different magnitude, providing evidence for an *asymmetric* impact. We provide evidence of the features of our model by means of a Monte Carlo simulation aiming at assessing its behavior in relationship to the configuration of sample sizes at different frequencies (Section 4). The main illustration is performed on the *S&P*500 volatility, with the U.S. Industrial Production and National Activity Index as additional volatility determinants, over a period spanning from 1980 to 2016 (Section 5). We show that the weighting functions are different, and the parameters associated to sign–specific past industrial production have indeed opposite signs, signaling the relevance of a different type of asymmetry in volatility, next to the one induced by negative returns. Out–of–sample (with a rolling window), we show that the proposed model generally outperforms all the competing specifications in a forecasting exercise evaluated by year. In Section 6 we perform some robustness check which confirm qualitatively the importance of the asymmetric specification, but signal some (minor parameter instability). As a matter of fact, in–sample (Section 6.1), the proposed model generates macro–variable parameters with opposite signs even for different periods, with varying magnitude. What matters is that the out-of-sample superior performance of our suggestion is confirmed also for different rolling windows and loss functions (Section 6.2). Finally, the relevance of the asymmetry of macro–variables on the volatility is successfully tested also in an international setting (Section 6.3), where we take indices for four advanced (Euro Stoxx 50 for the Euro–area, DAX for Germany, CAC 40 for France, S&P/TSX for Canada) and two emerging markets, (Hang Seng for Hong Kong and BOVESPA for Brazil). Conclusions follow.

2. Some motivation for the asymmetric analysis

Empirical and theoretical evidences may be provided to illustrate that positive and negative macro–variables variations have separate impacts on the volatility of interest (in our case, *S&P*500 volatility). We consider two different macro–variables: the U.S. Industrial Production growth rate (*IP*) and (differenced) Chicago Fed National Activity Index (*NAI*). Both variables have been used as additional volatility determinants in different contributions (Amendola et al. (2017), Conrad and Loch (2015), among others). *IP* has been traditionally used as a proxy of the aggregate supply, while *NAI* is an indicator of the health of the economy (position relative to the historical trend rate of growth). Both macro–variables are observed monthly. Segal et al. (2015) provide a theoretical and empirical framework to evaluate separately the volatility associated with positive and negative shocks to macro–variables. The former are defined as "good" uncertainty and the latter as "bad" uncertainty. These uncertainties are measured by *IP* semivariances. By considering them separately, the authors find that both *IP* semivariances have a statistically significant impact on several macro–variable growth rates (consumption, GDP, earnings, market dividend, capital and R&D investment).

Another piece of evidence is given by the influence that (lagged) *IP* and *NAI* have on *S&P*500 monthly (aggregated realized) volatility ($S\&P500_{ARV}$), in the sense of Granger (1969) (Table 1). $S\&P500_{ARV}$ is obtained summing the close–to close daily realized volatility.[1] The

[1] The daily realized volatility is computed by taking the square root of the summation of the 5-min squared intradaily returns – as provided by the Oxford-Man Institute– and squared overnight returns.

Table 2
*S&P*500 aggregated monthly volatility OLS regressions.

Regressor	Benchmark	Model$_1$	Model$_2$
Panel A			
$S\&P500_{ARV}(-1)$	0.933***	0.927***	0.777***
$IP(-1)$		−0.025	
$IP^{+}(-1)$			0.059***
$IP^{-}(-1)$			−0.081***
AIC	−459.187	−466.563	−503.656
BIC	−450.551	−449.291	−477.748
Adj-R^2	0.875	0.880	0.901
MSE	0.604	0.577	0.476
Panel B			
$S\&P500_{ARV}(-1)$	0.933***	0.930***	0.869***
$NAI(-5)$		0.001	
$NAI^{+}(-5)$			0.028*
$NAI^{-}(-5)$			−0.031*
AIC	−459.187	−446.068	−450.767
BIC	−450.551	−428.796	−424.858
Adj-R^2	0.875	0.871	0.875
MSE	0.604	0.610	0.590

Notes: The table reports OLS regressions of $S\&P500_{ARV}$ on lagged $S\&P500_{ARV}$ volatility and *IP* (Panel A) and *NAI* (Panel B) with Newey-West (HAC) standard errors. The regression covers the period January 2000–December 2016, for a total of 204 observations. The MSE has been multiplied by 100. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively.

upper panel of the table refers to the relationship between lagged *IP* and $S\&P500_{ARV}$, for the period January 2000–December 2016. The null of Granger non–causality (*GnC*) of *IP* at lag l, with $l = 1, \dots, 12$ on $S\&P500_{ARV}$ is always rejected.

The lower panel of the same table investigates the relationship between lagged *NAI* and $S\&P500_{ARV}$, under the same period. In this case, the null of *GnC* is rejected at lags 5, 6, 7, 9, 11 and 12, with a significance level of 0.10.

Let us now determine whether and how much (lagged) *IP* and *NAI* realizations have an influence on current *S&P*500 volatility, for the same period under consideration. We regress $S\&P500_{ARV}$ on $S\&P500_{ARV}(-1)$. This regression represents the benchmark of the first relationship in view of volatility persistence and is reported in the first column of Table 2. If we add to the previous regression also $IP(-1)$, as shown in the second column of the same table, the additional regressor is not significant. But, splitting $IP(-1)$ between positive and negative values ($IP(-1)^{+}$ and $IP(-1)^{-}$, respectively) both *IP* coefficients become significant but with opposite signs, and the autoregressive coefficient decreases (cf. the third column of Table 2). The indication is that macro–variable variations have different impacts, according to whether they are positive or negative, but both in the direction of increasing volatility, a point we will discuss further. Moreover, the Adjusted-R^2 of this specification increases while the Mean Squared Error (MSE) decreases.

The results concerning the relationship between lagged *NAI* and $S\&P500_{ARV}$ are reported in the lower panel of Table 2. Considering only *NAI* as an additional regressor does not lead to any substantial change in the regression, but, if we consider positive and negative *NAI* values separately, the value of the MSE decreases.

A final example supporting the idea that positive and negative macro–variable variations have different impact on the volatility of the variable of interest is represented in Fig. 1. In this figure, we graphically represent the values of the volatility $S\&P500_{ARV}$ against values of lagged monthly returns and of lagged growth rates of industrial production *IP* (Fig. 1a) and *NAI* (Fig. 1b) with the idea that shades represent quartiles of volatility. The figures clearly show that higher volatility values tend to be in the lower left quadrant, while lower values are in the top right.

3. The A-GARCH-MIDAS model

In the GARCH-MIDAS framework, the log-return for day i within a lower frequency period t, $r_{i,t}$, is such that

$$r_{i,t} - \mu_{i,t} = \sqrt{\tau_t \times g_{i,t}}\varepsilon_{i,t}, \quad \text{with} \quad i \mid t = 1, \dots, N_t \quad \text{and} \quad t = 1, \dots, T, \tag{1}$$

where $\mu_{i,t}$ represents the conditional mean and N_t the number of days for period t. A period t may be a week, a month, a quarter, etc., depending on the frequency at which the additional (stationary) macro–variable X_t (usually a growth rate) is observed. In Eq. (1), $\varepsilon_{i,t}$ represents the error term, assumed to follow an iid distribution, conditional on the information set up to day $i-1$ of period t (that is: $I_{i-1,t}$). Formally, $\varepsilon_{i,t} \mid I_{i-1,t}$ could be $\mathcal{N}(0,1)$, or a Student's t, but the choice does not alter the consistency properties of the corresponding estimator, based on quasi-maximum likelihood arguments, a standard result in the GARCH literature. The long and short-run components are expressed as τ_t and $g_{i,t}$, respectively. The low–frequency component τ_t, obtained as a one-sided MIDAS filter of the lagged variable X_{t-k}, is as follows:

$$\tau_t = \exp\left(m + \theta \sum_{k=1}^{K} \delta_k(\omega) X_{t-k}\right), \tag{2}$$

where $\delta_k(\omega)$ is a suitable function weighting the past K realizations of X_t, θ measures the impact of such a sum (labelled as $WSUM_{(t-1):(t-K)}$) on volatility around an average level m. In this respect, almost all the empirical works use the Beta function as weighting function:

$$\delta_k(\omega) = \frac{(k/K)^{\omega_1 - 1}(1 - k/K)^{\omega_2 - 1}}{\sum_{j=1}^{K}(j/K)^{\omega_1 - 1}(1 - j/K)^{\omega_2 - 1}}. \tag{3}$$

This function allows the weights to put equal, increasing or decreasing emphasis on the past realizations of X_t, provided that $\omega_n \geq 1$, with $n = 1, 2$. For instance, when $\omega_1 = \omega_2 = 1$, the equally weighting scheme is adopted. When $\omega_1 > \omega_2$, farther observations are weighted more and finally when $\omega_1 < \omega_2$ more recent observations have a larger importance. Since we are interested only in this latter case, we adopt a restricted weighting scheme, imposing $\omega_1 = 1$, so as to allow just for a monotonically decreasing weighting scheme. Obviously, $\sum_{k=1}^{K} \delta_k(\omega_2) = 1$.

In Eq. (2), the exponential transformation delivers a positive τ_t which varies according to the sign of $WSUM_{(t-1):(t-K)}$. Suppose that both $WSUM_{(t-1):(t-K)} > 0$ and $WSUM_{(t-2):(t-K-1)} > 0$ and $\theta < 0$. What matters in whether τ_t increases or decreases is the sign of the difference between the two, $\Delta WSUM_t$. If it is positive (i.e. $WSUM_{(t-1):(t-k)} > WSUM_{(t-2):(t-K-1)}$), $\Delta\tau_t < 0$ and the long–run volatility will decrease as the result of an increased accumulation of growth rates. Note that, in this specification, positive and negative values may cancel each other out (i.e. $WSUM_{(t-1):(t-K)} \approx 0$).

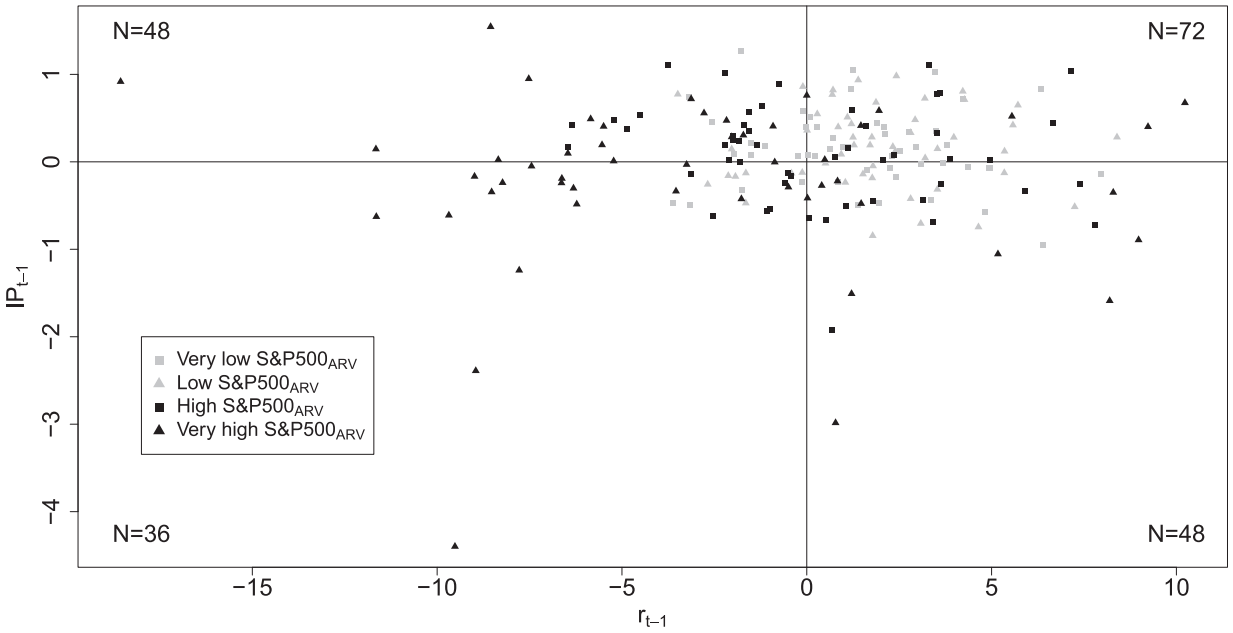
The short-run component $g_{i,t}$ follows a unit mean-reverting GARCH(1,1), with an asymmetric term related to past returns, that is:

$$g_{i,t} = (1 - \alpha - \beta - \gamma/2) + \left(\alpha + \gamma \cdot \mathbb{1}_{(r_{i-1,t} - \mu_{i-1,t} < 0)}\right) \frac{\left(r_{i-1,t} - \mu_{i-1,t}\right)^2}{\tau_t} + \beta g_{i-1,t}, \tag{4}$$

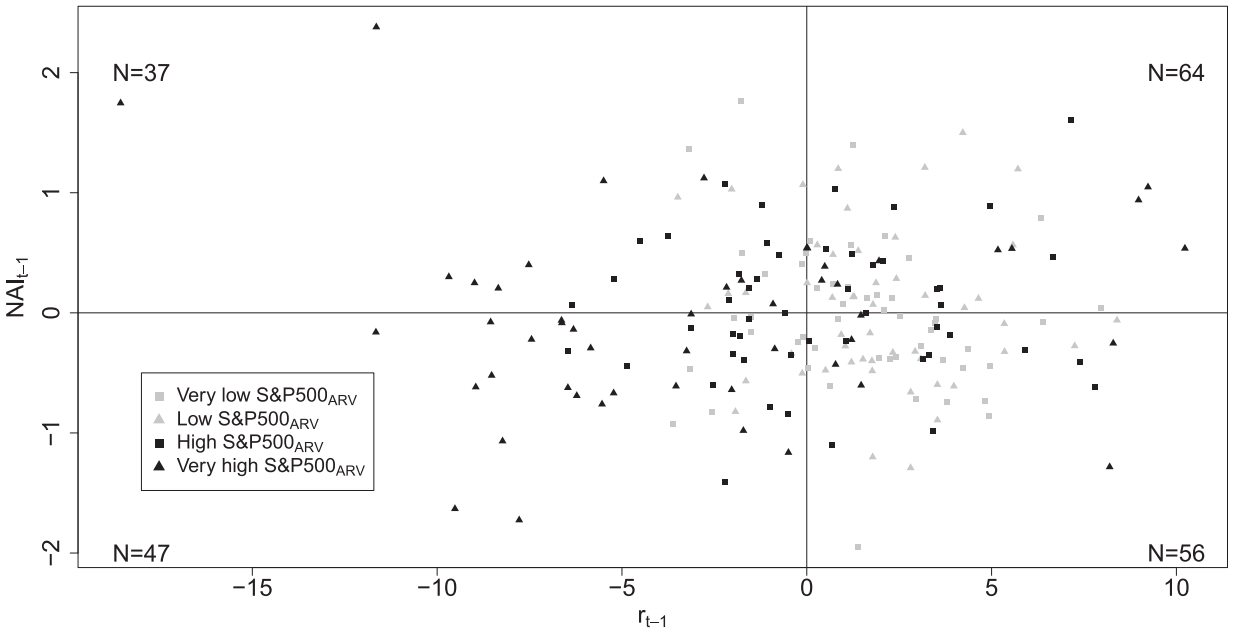
where $\mu_{i-1,t}$ represent the conditional mean of the daily log-returns, and $\mathbb{1}_{(.)}$ is an indicator function that assumes value one if the argument is true; note that when $i = 1$, $(i-1, t) = (N_{t-1}, t-1)$. The following constraints are used to assure positivity of short-run component: $\alpha \geq 0$, $\beta \geq 0$ and $\alpha + \beta + \gamma/2 < 1$. Under these conditions, the unconditional

[2] The analysis on the RME_k^{τ} when *NAI* is used is not reported for sake of brevity, but it is available upon request.

(a) *IP*



(b) *NAI*



Notes: in this figure we synthesize the behavior of $S\&P500_{ARV}$ (the $S\&P500$ monthly aggregated realized volatility, obtained summing the 5-minute squared intradaily and overnight returns) as a function of r_{t-1} (monthly aggregated daily returns, i.e. $r_t = \sum_i^{N_t} r_{i,t}$) and of the monthly changes in *IP* (top panel) and *NAI* (bottom panel). We grouped $S\&P500_{ARV}$ by levels of volatility, and we report the number N of observations within each quadrant, spanning the years 2000 - 2016.

Fig. 1. Levels of volatility according to the sign of lagged returns and macro–variables (*IP* and *NAI*).

expectation of the short-run component is $E(g_{i,t}) = 1$. If we assume that $r_{i,t}$ is zero mean as well as $E(g_{i,t} \mid I_{t-1}) = 1$ and given that at period $t-1$ the long-run component for period t is determined and equal to τ_t, the local level of the daily returns variance calculated at period t is:

$$E\left[(r_{i,t})^2 \mid I_{t-1}\right] = E\left(\tau_t \times g_{i,t} \mid I_{t-1}\right) = \tau_t E\left(g_{i,t} \mid I_{t-1}\right) = \tau_t, \tag{5}$$

i.e. it varies with t. However, if θ in Eq. (2) is zero, then the long-term component reduces to a constant $\forall t$, and we would have a constant unconditional variance of $r_{i,t}$:

$$E\left[(r_{i,t})^2 \mid I_{t-1}\right] = \exp(m) E\left(g_{i,t} \mid I_{t-1}\right) = \exp(m). \tag{6}$$

Finally, conditional on the previous day $i-1$, the variance of $r_{i,t}$, for any day within a period t, calculated at day $i-1$, is given by:

$$E\left[r_{i,t}^2 \mid I_{i-1,t}\right] = \tau_t \times g_{i,t}. \tag{7}$$

In order to introduce the presence of an asymmetric response to positive and negative macro–variable changes as a testable proposition, the A-GARCH-MIDAS defines the long-run component as:

$$\tau_t = \exp\left(m + \theta^+ \sum_{k=1}^{K} \delta_k(\omega_2^+) X_{t-k} \mathbb{1}_{(X_{t-k} \geq 0)} + \theta^- \sum_{k=1}^{K} \delta_k(\omega_2^-) X_{t-k} \mathbb{1}_{(X_{t-k} < 0)}\right), \tag{8}$$

where asymmetry shows up in two places: the *sign–specific* parameters (θ^+ and θ^-), and the weights associated with each positive and negative kth lagged value ($\delta_k(\omega_2^+)$ and $\delta_k(\omega_2^-)$, respectively). As a matter of fact, the GARCH–MIDAS is a (strongly restricted) special case of our A–GARCH–MIDAS when $\theta^+ = |\theta^-|$, and $\delta_k(\omega_2^+) = \delta_k(\omega_2^-)$, for all $k = 1, \ldots, K$.

From such an expression, following Engle et al. (2013), we can calculate the *Relative Marginal Effect* on τ_t, RME_k^τ, of a change ΔX_{t-k} in X_{t-k} (according to its sign, $\Delta X_{t-k}^+ > 0$, when $X_{t-k} > 0$; $\Delta X_{t-k}^- < 0$, when $X_{t-k} < 0$) as

$$RME_k^\tau = \begin{cases} \exp\left(\theta^+ \cdot \delta_k(\omega_2^+) \cdot \Delta X_{t-k}^+\right) - 1, & \text{if } X_{t-k} > 0 \\ \exp\left(\theta^- \cdot \delta_k(\omega_2^-) \cdot \Delta X_{t-k}^-\right) - 1, & \text{if } X_{t-k} < 0, \end{cases} \tag{9}$$

which further substantiates asymmetric contributions.

To gain some insights on the behavior of τ_t in this model, where compensations are unlikely, let us define the weighted sum of positive and negative lagged X_t between $t-K$ and $t-1$ as $WSUM_{(t-1):(t-K)}^+ = \sum_{k=1}^K \delta_k(\omega_2^+) X_{t-k} \mathbb{1}_{(X_{t-k} \geq 0)}$ and $WSUM_{(t-1):(t-K)}^- = \sum_{k=1}^K \delta_k(\omega_2^-) X_{t-k} \mathbb{1}_{(X_{t-k} < 0)}$, respectively. Assuming still $\theta^+ > 0$ and $\theta^- < 0$, in Eq. (8), m now represents a lower bound for τ_t, attainable only when both $WSUM_{(t-1):(t-K)}^+ = 0$ and $WSUM_{(t-1):(t-K)}^- = 0$. The long-run volatility τ_t will differ from τ_{t-1} based on the difference between the weighted sums for the periods $(t-1) : (t-K)$ and the ones for $(t-2) : (t-K-1)$, i.e. on

$$\Delta WSUM_t^+ = WSUM_{(t-1):(t-K)}^+ - WSUM_{(t-2):(t-K-1)}^+$$

$$\Delta WSUM_t^- = WSUM_{(t-1):(t-K)}^- - WSUM_{(t-2):(t-K-1)}^-.$$

To simplify the discussion with an example, take the case where $\Delta WSUM_t^+ = 0$ and $\Delta WSUM_t^- < 0$: this happens when $\delta_1(\omega_2^-) X_{t-1} < \delta_K(\omega_2^-) X_{t-K-1} < 0$, i.e. when the incoming (weighted) most recent term is "worse" than the corresponding outgoing term: as a consequence, τ_t will be greater than τ_{t-1}. Other configurations could be evaluated, but it is clear that there are cases when τ_t will be lower than τ_{t-1}: in the previous example, when, other things being equal, $\delta_K(\omega_2^-) X_{t-K-1} < \delta_1(\omega_2^-) X_{t-1} < 0$ and $\Delta WSUM_t^- > 0$. What is to be remarked is that the presence of a most recent positive or negative value of X_{t-1} does not determine *per se* an increase or a decrease in τ_t.

Statistical inference is carried out as with the GARCH-MIDAS. Its asymptotic properties are the same as in Wang and Ghysels (2015). For both models, the estimation of the parameters is carried out by maximizing the following log-likelihood:

$$\ln \mathscr{L} = -\frac{1}{2} \sum_{t=1}^{T} \left[\sum_{i=1}^{N_t} \left[\log(2\pi) + \log(g_{i,t}\tau_t) + \frac{(r_{i,t} - \mu_{i,t})^2}{g_{i,t}\tau_t}\right]\right]. \tag{10}$$

Also in view of the stylized facts just examined, both the traditional and the A-GARCH-MIDAS models can include an asymmetric term also in the short-run equation, in the spirit of the Glosten-Jagannathan-Runkle (GJR) GARCH model (Glosten et al., 1993), capturing the customary sign–specific reactions to lagged returns. Just to keep acronyms simple, in what follows we will refer to GJR-G-M instead of GJR-GARCH-MIDAS and GJR-A-G-M instead of GJR-A-GARCH-MIDAS.

3.1. GJR-G-M and GJR-A-G-M news impact curves

A convenient way to summarize the overall impact of sign–specific values in returns and the macro–variable, comes from adapting the news impact curves (NICs) of Engle and Ng (1993) to the GJR-G-M and GJR − A − G − M. In the original definition, the NIC is a function of a range of potential returns which describes graphically how the current conditional variance σ_i^2 reacts to (past) shocks, for a given intercept when past conditional variance is kept constant (at its unconditional value). In the GARCH-MIDAS framework, we can consider that this past conditional variance can move slowly, changing along the low frequency dimension, and the NIC in this context should rather be also a function of a range of (lagged) values of macro–variables. To calculate the NICs for GJR-G-M and GJR-A-G-M, we consider a generic day i within the period t, with conditional variance equal to:

$$\sigma_{i,t}^2 = \tau_t \times g_{i,t}, \tag{11}$$

with τ_t specified as in (2) or in (8), and $\sigma_{i,t}^2$ is conditional on information of the previous day. Assuming for simplicity a zero conditional mean of the daily log-returns (as in (1)), we can mirror the original derivation of the NICs, and replace lagged conditional variances with suitable values. Note that, rewriting (4) and substituting $g_{i-1,t}$ with its unit expectation,

$$g_{i,t} = (1 - \alpha - \beta - \gamma/2) + \left(\alpha + \gamma \cdot \mathbb{1}_{(r_{i-1,t} < 0)}\right) \frac{(r_{i-1,t})^2}{\tau_t} + \beta g_{i-1,t} \tag{12}$$

$$= (1 - \alpha - \beta - \gamma/2) + \left(\alpha + \gamma \cdot \mathbb{1}_{(r_{i-1,t} < 0)}\right) \frac{(r_{i-1,t})^2}{\tau_t} + \beta;$$

$$= (1 - \alpha - \gamma/2) + \left(\alpha + \gamma \cdot \mathbb{1}_{(r_{i-1,t} < 0)}\right) \frac{(r_{i-1,t})^2}{\tau_t}. \tag{13}$$

Replacing (13) in (11), our modified NIC is:

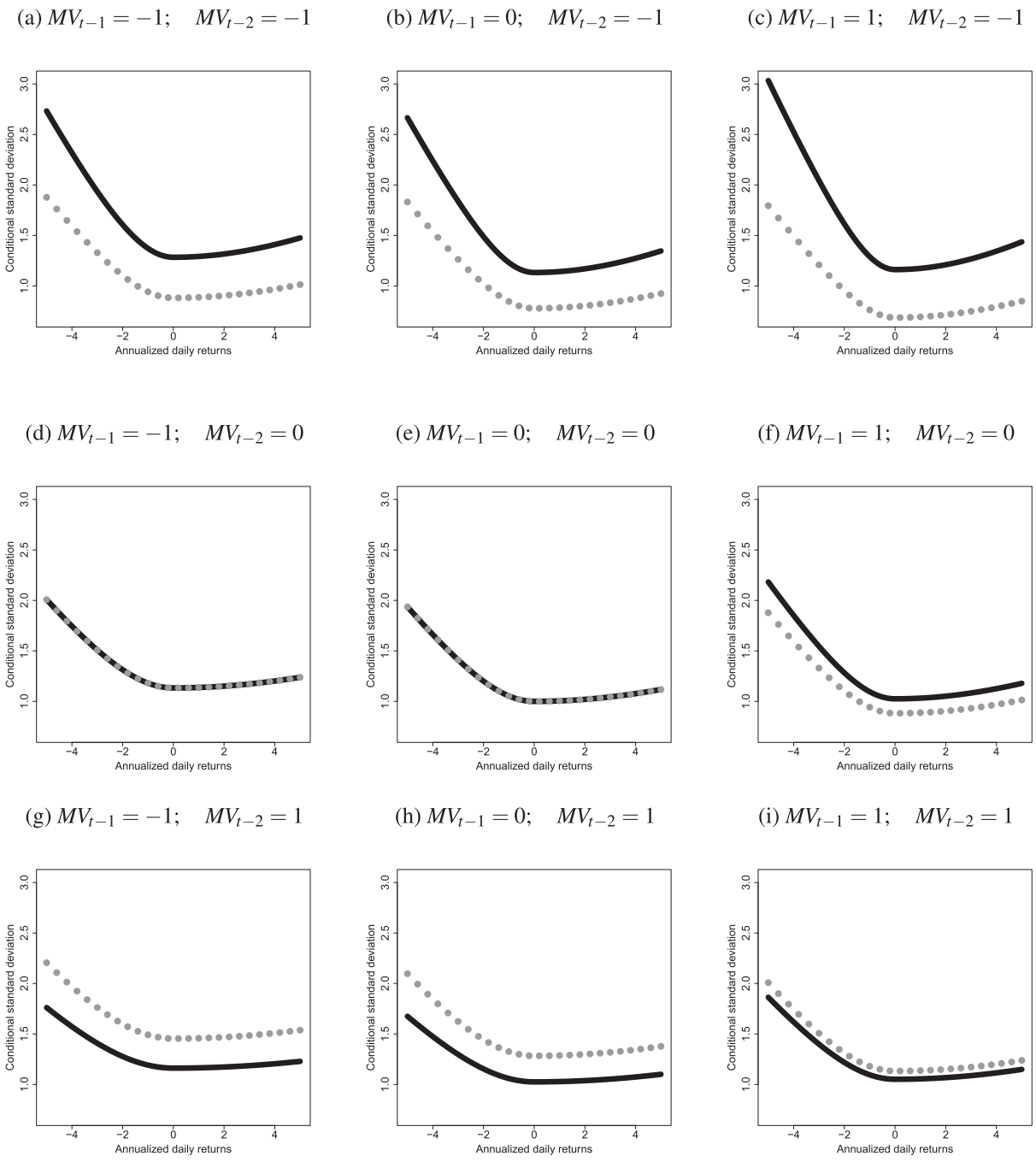
$$\sigma_{i,t}^2 = (1 - \alpha - \gamma/2)\,\tau_t + \left(\alpha + \gamma \cdot \mathbb{1}_{(r_{i-1,t} < 0)}\right) (r_{i-1,t})^2. \tag{14}$$

In (14), a theoretical conditional variance depends on past macro–variable realizations (influencing τ_t) and past daily log-returns. To convert it into an illustrative graphical representation, for both GJR-G-M and GJR-A-G-M models, let us assume that only $K = 2$ past macro–variable realizations may influence the volatility. Moreover, to facilitate the comparison between the two NICs, we use the same coefficients for the short-run equation, that is: $\alpha = 0.01$, $\beta = 0.85$, $\gamma = 0.1$. Finally, for both models we set the coefficient m such that, when past daily log-returns and macro–variable values are equal to zero, the conditional variance is equal to one. Hence, from:

$$(1 - \alpha - \gamma/2)\,\tau_t = 1, \tag{15}$$

we get $m = -\log(1 - \alpha - \gamma/2)$. Keeping in mind that $\omega_1 = 1$, the weight ω_2 can be derived allowing for an equal–weight scheme for the two past macro–variable realizations, such that for both models $\omega_2 = \omega_2^+ = \omega_2^- = 1$. As regards the coefficients influencing past macro–variable realizations, we set $\theta = \theta^- = -0.5$ and $\theta^+ = 0.1$. This setup entails that positive past macro–variable variations let the volatility increase, but generally less so than negative ones.

Under this design, several NICs for GJR-G-M and GJR-A-G-M are plotted in Fig. 2. We notice first that all the NICs have the usual asymmetric pattern, due to the inclusion of the γ coefficient in the short-run equation. Second, when the macro–variable realizations are zero at time $t-1$ and $t-2$, the two NICs are identical (Fig. 2e). More interestingly, looking at the case reported in Fig. 2a where the macro–variable decreases both at time $t-1$ and $t-2$, the NIC for the GJR-A-G-M is much larger than that for the GJR-G-M. Finally, (cf. Fig. 2i) when we have two positive increases in the macro variable we observe a lower impact on volatility than in the case of negative values. Clearly, without the extra asymmetry, the potential different impact of sign–specific macro–variable variations would not emerge.



Notes: The figure shows some typical extended NIC configurations for GJR-G-M (dotted grey lines) and GJR-A-G-M (solid black lines) models, for different lagged macro–variable values (see text for definitions).

Fig. 2. GJR-G-M and GJR-A-G-M extended News Impact Curves.

Table 3
Summary statistics of simulated daily returns.

	$T = 1500$	$T = 2250$	$T = 3000$
Mean	−0.000	−0.000	−0.001
Std. Dev.	0.683	0.735	0.760
Min	−3.817	−4.472	−4.918
Max	3.815	4.509	4.971
Skewness	−0.001	−0.007	0.002
Kurtosis	4.604	4.917	5.284

Notes: Overall average of some summary statistics across replications for three different values of T.

4. A Monte Carlo experiment

Let us focus on the finite sample performance of the GJR-A-G-M by conducting some Monte Carlo experiments. A simulation study can help in understanding how it performs when estimating the parameters (assuming them as known), in particular as a function of the number of observations. We replicate over $R = 1000$ simulations the series $r_{i,t}$, whose conditional variance is driven by an A-G-M, where, for simplicity, the asymmetric term in the short-run equation is not included:

$$r_{i,t} = \sqrt{\tau_t \times g_{i,t}}\varepsilon_{i,t}, \quad \text{with } i = 1, \dots, N_t \quad \text{and} \quad \varepsilon_{i,t} \sim T_\nu. \tag{16}$$

In (16),

- $\varepsilon_{i,t}$ follows a Student's t distribution, with ν degrees of freedom;
- $N_t = N = 30$, $\forall t$, with $t = 1, \dots, T_{MV}$;
- the total number of observations is $T = N \cdot T_{MV}$;
- the long-run component is given in Eq. (8), with $m = m_0$, $\theta^+ = \theta_0^+$ and $\theta^- = \theta_0^-$;
- the short-run component is given in Eq. (4), with $\alpha = \alpha_0$ and $\beta = \beta_0$.
- X_t is such that $X_t = \Delta Y_t$, with $Y_t \sim RW$, that is Y_t follows a Random Walk process.

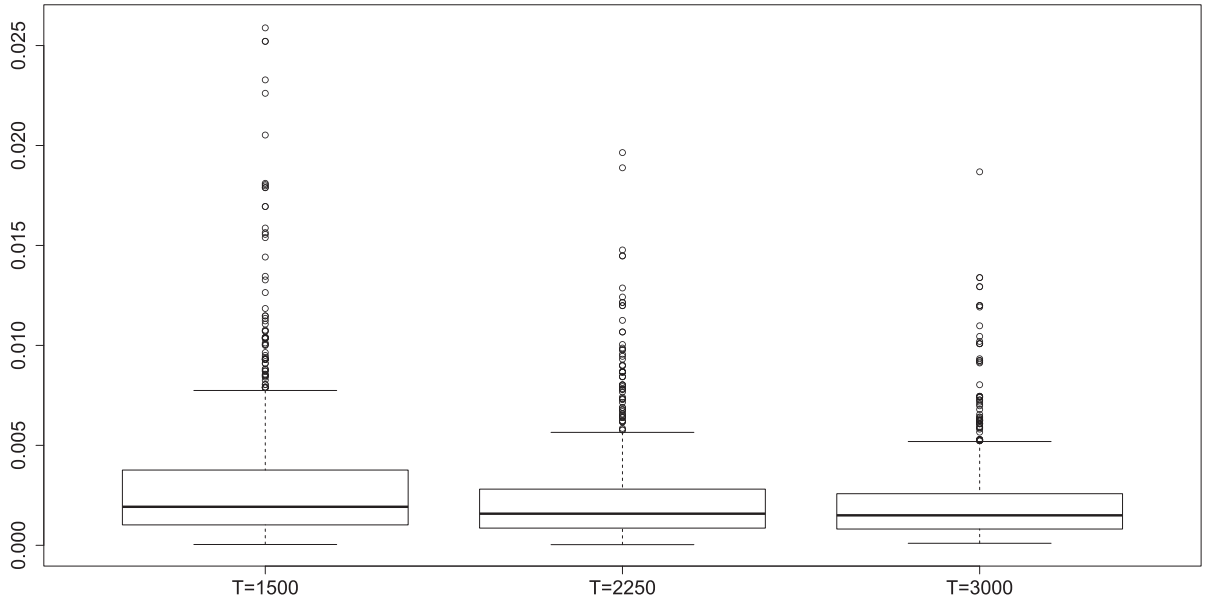
In particular, we consider three different lengths for X_t: $T_{MV} = \{50, 100, 150\}$. This allows us to verify the performance of the A-G-M for $T = \{1500, 2250, 3000\}$. The degrees of freedom, fixed for all the simulations, is $\nu = 25$. To speed up the estimation, we adopt the equally weighting scheme, that is: $\omega_2^+ = \omega_2^- = 1$. In doing so, the parametric space under the Monte Carlo simulation is reduced to $\Phi^{MC} = \{\alpha, \beta, m, \theta^+, \theta^-\}$. For both the simulation and the estimation procedures, we set $K = 12$, meaning that the MIDAS regression of the long-run equation includes 12 lagged low–frequency variable realizations.

Summary statistics of the simulated daily returns, for all the Monte Carlo configurations, are shown in Table 3: on average, the simulated daily returns have those characteristics noted as "stylized facts", with the desired positive excess of kurtosis; the null hypothesis of the ARCH-

Table 4
Monte Carlo estimates.

	ϕ_0	$T_{MV} = 50, T = 1500$			$T_{MV} = 75, T = 2250$			$T_{MV} = 100, T = 3000$		
		$E(\widehat{\phi})$	95% CI		$E(\widehat{\phi})$	95% CI		$E(\widehat{\phi})$	95% CI	
α	0.05	0.054	0.053	0.056	0.056	0.055	0.057	0.056	0.055	0.057
β	0.80	0.725	0.715	0.735	0.749	0.742	0.757	0.760	0.754	0.766
m	−5.00	−4.878	−4.900	−4.852	−4.879	−4.897	−4.863	−4.906	−4.919	−4.892
θ^+	4.00	3.983	3.947	4.019	3.975	3.951	4.002	4.020	3.999	4.038
θ^-	−6.00	−5.972	−6.009	−5.933	−5.994	−6.020	−5.970	−6.023	−6.040	−6.002

Notes: The table presents the overall average of the Monte Carlo estimates as well as the 95% confidence intervals (CI), for three different configurations of T.



Notes: Box–plots of MSE between the real volatility and that estimated by means of the A-G-M, as the number of observation increases.

Fig. 3. The distribution of the MSE across simulations.

Table 5
Summary statistics.

	Obs	Min	Max	Mean	Std. Dev.	Skew.	Kurt.
*S&P*500 daily returns	9333	−22.900	10.957	0.032	1.119	−1.151	26.493
IP	444	−4.398	2.026	0.148	0.674	−1.208	6.170
IP^+	285	0.004	2.026	0.515	0.386	1.283	2.094
IP^-	159	−4.398	−0.003	−0.509	0.571	−3.326	2.094
NAI	444	−2.403	3.690	0.002	0.765	0.467	2.042
NAI^+	217	0.002	3.690	0.588	0.559	2.083	5.982
NAI^-	227	−2.403	−0.003	−0.559	0.453	−1.426	2.485

Notes: The table presents the number of observations (Obs), the minimum (Min) and the maximum (Max) observation value, the mean, the standard deviation (Std. Dev.), the skewness (Skew.), the (excess) kurtosis (Kurt.). *IP* represents the U.S. Industrial Production growth rate (i.e. $100 \cdot [\log(IP_t) - \log(IP_{t-1})]$). *NAI* represents the first difference of the Chicago Fed (seasonally adjusted) National Activity Index.

LM test is rejected (results not reported for the sake of space), giving us the desired conditional heteroskedasticity.

The Monte Carlo estimates of the A-G-M are summarized in Table 4, where we report the true values of the parameters in the first column as a reference. In general, the bias between the estimated and the true parameters decreases as the sample period increases; as expected, the width of the 95% confidence intervals decreases with larger T's; the average estimates of the parameters of interest, θ^+ and θ^-, when $T = 3000$, are very accurate.

The distance between the estimated and the true volatility is to be investigated under the MSE loss function (LF): we report the box–plots of the MSE between the true and estimated volatility, when T increases, in Fig. 3. Even if at a slow rate, the MSE decreases as long as T increases.

Overall, the performance of the A-G-M estimation when the data generating process is known are fully satisfactory. The estimated parameters are unbiased and the MSE decreases as the sample size increases.

5. Empirical application

In the empirical application, we estimate a number of GARCH models on the *S&P*500 index, adding, in turn, *IP* and *NAI* as additional long–term volatility determinants. In particular, the set of models con-

Table 6
In-sample estimation.

	GARCH	GJR	GJR-G-M$_{IP}$	GJR-A-G-M$_{IP}$	GJR-G-M$_{NAI}$	GJR-A-G-M$_{NAI}$
α	0.090***	0.021***	0.020***	0.015***	0.02***	0.015***
	(0.017)	(0.005)	(0.006)	(0.006)	(0.005)	(0.006)
β	0.896***	0.893***	0.892***	0.889***	0.894***	0.893***
	(0.021)	(0.014)	(0.017)	(0.017)	(0.015)	(0.016)
γ		0.135***	0.129***	0.134***	0.128***	0.135***
		(0.020)	(0.026)	(0.024)	(0.025)	(0.024)
m			5.446***	4.612***	5.569***	4.768***
			(0.126)	(0.229)	(0.226)	(0.249)
θ			0.196*		−0.126	
			(0.118)		(0.17)	
ω_2			1.001		1.024	
			(1.074)		(5.446)	
θ^+				0.664***		0.164
				(0.149)		(0.163)
ω_2^+				1.001***		1.023
				(0.335)		(0.822)
θ^-				−0.142**		−0.572***
				(0.058)		(0.155)
ω_2^-				4.219***		1.001***
				(0.381)		(0.352)
AIC	72,570.682	72,355.033	72,344.037	**72,321.455**	72,356.87	72,332.940
BIC	72,584.964	**72,376.456**	72,386.885	72,378.585	72,399.72	72,390.070
QLIKE	−3.777	−3.782	−3.783	**−3.787**	−3.783	−3.786
MSE	0.181	0.166	0.156	**0.148**	0.157	0.152
LB(5)$_{sq}$	0.686	0.843	0.876	0.868	0.927	0.910
LB(10)$_{sq}$	0.890	0.961	0.974	0.963	0.983	0.975
LB(20)$_{sq}$	0.883	0.970	0.988	0.978	0.988	0.982
LM(5)$_{sq}$	0.998	1.000	1.000	1.000	1.000	1.000
LM(10)$_{sq}$	1.000	1.000	1.000	1.000	1.000	1.000
LM(20)$_{sq}$	1.000	1.000	1.000	1.000	1.000	1.000

Notes: The table contains the estimated parameters for GARCH and GJR (with variance targeting), GJR-G-M and GJR-A-G-M models. The dependent variable is the *S&P*500 volatility. The number of lagged realizations entering the long-run equation is $K = 24$. Newey-West (HAC) standard errors are in parentheses. The sample covers the period from January 2, 1980 to December 31, 2016, for a total of 9333 daily observations. *IP* and *NAI* are the additional volatility determinants, entering the GJR-G-M and GJR-A-G-M models. Both macro–variables are observed monthly, for a total of 444 observations. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. QLIKE and MSE map the distance between the volatility proxy and estimated volatility, for the period January 2, 2000 to December 31, 2016. The volatility proxy consists of the realized volatility, obtained by taking the square root of the summation of the 5-min squared intradaily returns and overnight returns. The MSE has been multiplied by 10,000. LB(l)$_{sq}$ represents the p-values of the Ljung-Box test at l lag, applied on squared standardized residuals. LM(l)$_{sq}$ represents the p-values of the Engle's Lagrange Multiplier test for autoregressive conditional heteroskedasticity at l lag, applied on the squared standardized residuals. The best performance is indicated in bold.

sists of the GARCH, GJR-GARCH, GJR-G-M and our model, the GJR-A-G-M, all in their own (1,1) specification. With reference to the latter two models, the labels GJR-G-M$_{MV}$ and GJR-A-G-M$_{MV}$ indicate that a macro–variable *MV* has been considered as an additional volatility determinant.

We present two sets of results: we first focus on the in-sample estimation, where the sample period covers the interval 1980–2016, for a total of 9333 daily and 444 monthly observations. Second, we compare the out-of-sample performance of the GJR-A-G-M (with *IP* as macro–variable entering the long-run equation) to the other specifications. Under both the in-sample and out-of-sample evaluations, the number of lagged realizations entering the long-run equation is $K = 24$. The results reported here are quite robust to changes in K. Some summary statistics of the variables considered are presented in Table 5. As largely documented in literature, daily returns and *IP* are left-skewed. Moreover, all the variables present a high level of kurtosis.

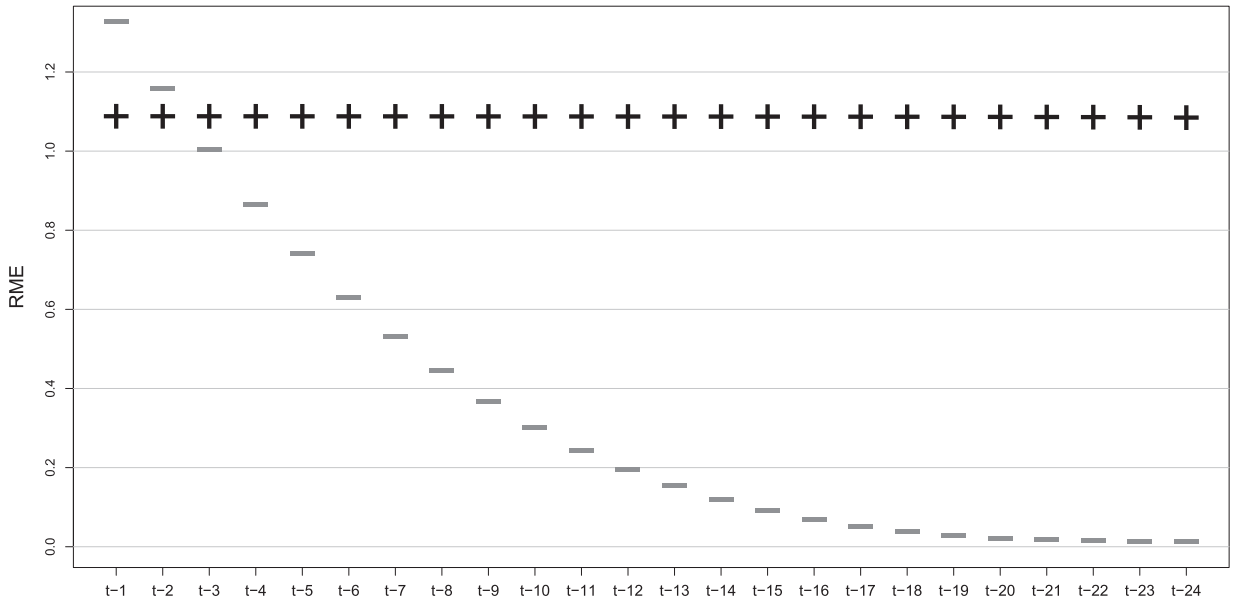
5.1. In-sample evaluation

The results of the in-sample estimation are given in Table 6. To ensure the comparability of the models, we estimated also the GARCH(1,1) and GJR specifications using variance targeting (Francq et al., 2011; Engle and Mezrich, 1996).

Unsurprisingly, given well established empirical regularities, all the GARCH and GJR estimated coefficients are statistically different from zero and with the expected signs for the short–run asymmetric effects. Also the short-run coefficients of all the GARCH-MIDAS specifications, irrespective of the macro–variable considered, are highly significant. The sign of the θ coefficient for the GJR-G-M$_{IP}$ model is positive, meaning that if from one month to the next the sum of changes in *IP* increases, the long-run volatility will increase. Apart from increases in the short–run volatility induced by negative returns, in reference to our proposed model GJR-A-G-M including *IP* as additional volatility determinant, we have a positive, respectively, negative sign for $\widehat{\theta}^{+}$ and $\widehat{\theta}^{-}$. The net effect on long–run volatility from one month to the next will depend on the two sources of asymmetry discussed above. We calculate the Relative Marginal Effect (as in Eq. (9)) resulting from a one-standard-deviation change in *IP* at lag one, and compare it to the corresponding RME_1^{τ} for the GJR-G-M$_{IP}$ model. For the latter specification, $\widehat{\theta} = 0.196$, $\delta_1(\widehat{\omega}_2) = 0.044$, and $\widehat{\sigma}_{IP} = 0.674$ giving $RME_1^{\tau} = 0.58\%$, which just changes sign in the case of a one standard deviation decrease. With our GJR-A-G-M model, instead, we have two separate θ's, i.e. $\widehat{\theta}^{+} = 0.664$ and $\widehat{\theta}^{-} = -0.142$ and two separate weights at lag one, namely $\delta_1(\widehat{\omega}_2^{+}) = 0.044$ and $\delta_1(\widehat{\omega}_2^{-}) = 0.168$, for positive and negative *IP* changes, respectively. This time, the marginal effects are calculated on sign–specific semi-standard deviations (see Table 5): a one standard deviation increase in IP^{+} at lag one (that is, 0.386), leads to a 1.12% relative marginal effect on the long-run volatility. On the other hand, assuming that IP^{-} decreases by one semi−standard deviation at lag one (that is, by 0.571) leads to a 1.37% RME_1^{τ}. Extending the calculations to all $k = 1, \dots, K = 24$, we get the positive and negative portions of the RME_k^{τ} depicted in Fig. 4. The substantial constancy of the former, and the sharp decay of the latter, point to an interesting interpretation of the contribution that the cumulated positive sum $WSUM_t^{+}$ and its negative counterpart $WSUM_t^{-}$ give to the low–frequency component.

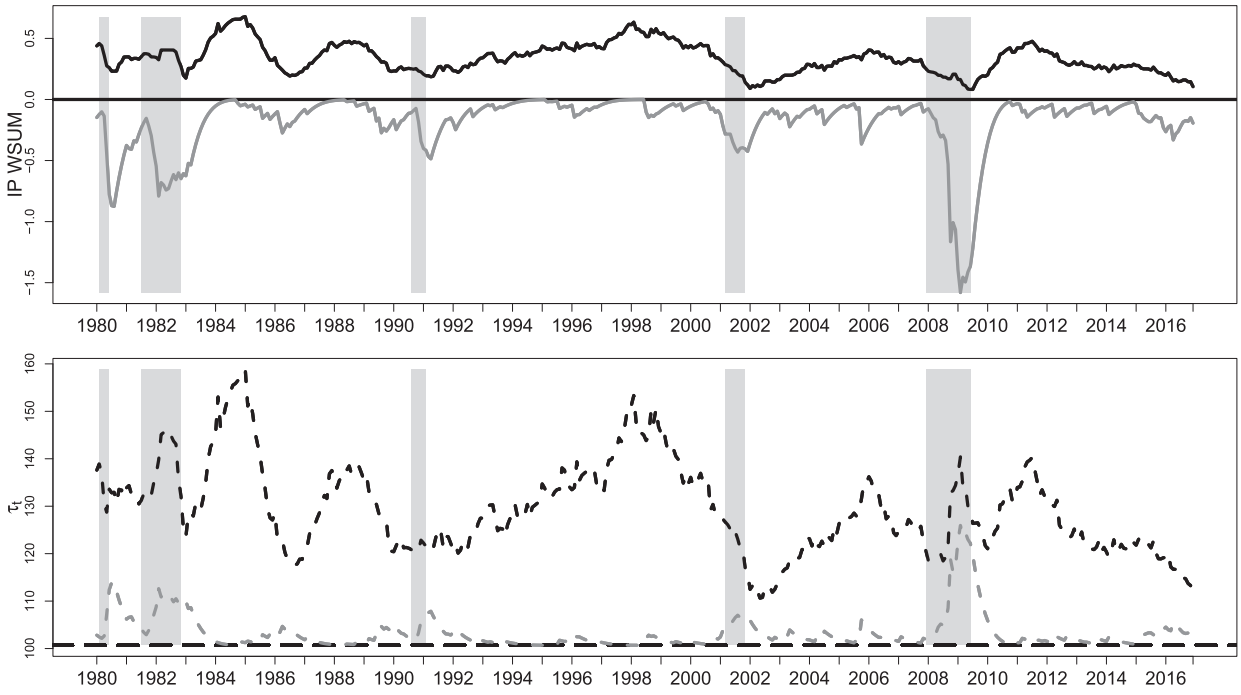
This behavior is reported in Fig. 5 where we have built a graph for $WSUM^{+}_{(t-1):(t-K)}$ and $WSUM^{-}_{(t-1):(t-K)}$ in the top panel, superimposing grey areas corresponding to the NBER recession dates in the US: the former is a very smooth series, in view of the fact that it is *de facto* a (one-sided) moving average of past positive values of growth rates of *IP*; the second shows, by and large, values close to zero, with sharp negative peaks in correspondence to the main recessions. The latest peak, corresponding to the Great Recession, is particularly deep into negative territory.

The bottom panel of Fig. 5 puts together the various portions, reconstructing the estimated long–run component $\widehat{\tau}_t$: we have in the lowest part the horizontal line corresponding to $\exp(\widehat{m})$; above it, we have reported the line corresponding to $\exp(\widehat{m} + \widehat{\theta}^{-}\widehat{WSUM}^{-}_{(t-1):(t-K)})$, which, as expected, provides the main contributions to τ_t during recessions only, and, finally the whole sequence of $\widehat{\tau}_t$, including the contribution



Notes: The figure plots the relative marginal effect. Black "plus"and grey "minus" represent the relative marginal effect for one standard deviation increase and decrease, respectively in the growth rate of *IP*.

Fig. 4. Relative marginal effect RME_k^{τ}, split into positive (+) and negative (−) contributions as a function of $k = 1, \dots, 24$.



Notes: In the top panel, black and grey lines represent the positive ($\widehat{WSUM}^{+}_{(t-1):(t-K)}$) and negative ($\widehat{WSUM}^{-}_{(t-1):(t-K)}$) weighted functions, respectively; the bottom panel reconstructs how $\hat{\tau}_t$ (top line) is built, starting from the horizontal line corresponding to $\exp(\hat{m})$ (lowest line); the grey dashed line corresponds to $\exp(\hat{m} + \hat{\theta}^{-}\widehat{WSUM}^{-}_{(t-1):(t-K)})$ and the black dashed line corresponds to $\exp(\hat{m} + \hat{\theta}^{+}\widehat{WSUM}^{+}_{(t-1):(t-K)} + \hat{\theta}^{-}\widehat{WSUM}^{-}_{(t-1):(t-K)})$. Shaded areas represent NBER recession periods.

Fig. 5. $WSUM^{+}_{(t-1):(t-K)}$ and $WSUM^{-}_{(t-1):(t-K)}$ for *IP* (top panel) and the corresponding decomposition of the long–run component (bottom panel).

of $\hat{\theta}^{+}\widehat{WSUM}^{+}_{(t-1):(t-K)}$. As noted elsewhere, the increase and decrease of $\hat{\tau}_t$ occur even as a consequence of the contribution of the positive portion (as a reflection of the *good* uncertainty, in the language of Segal et al. (2015)). In the occasion of the Great Recession such a contribution sharply decreases, given the prolonged sequence of negative *IP* results.

As per the other macro–variable employed, *NAI*, the coefficient $\hat{\theta}$ appears to be not significant, even though we had evidence of it Granger causing the *S&P*500 volatility. By contrast, allowing separate effects, under GJR-A-G-M$_{NAI}$, we have a significant effect on the long–run volatility, although from only $WSUM^{-}_{(t-1):(t-K)}$.[2] From an economic point of view, this means that only when the U.S. economy is expanding below its historical trend rate of growth, the stock market (long-run) volatility increases, a feature that it is not automatically embedded in the *IP* itself. The conclusion we reach is thus that any macroeconomic variable may have a relevant asymmetric effect on the low–frequency component of volatility of a market index, and that our proposed model is able to estimate this effect in a forecasting context.

In terms of information criteria, the Akaike Information Criterion (AIC) signals the GJR-A-G-M$_{IP}$ as the best model while the Bayesian Information Criterion (BIC) considers the GJR specification as the best one. However, as pointed out by Raftery (1995), the BIC differences between the best model (GJR) and the second best (GJR-A-G-M$_{IP}$) is just above 2, which is the upper limit of the class of BIC differences indicating *weak* evidence of a better fit in one model over another.

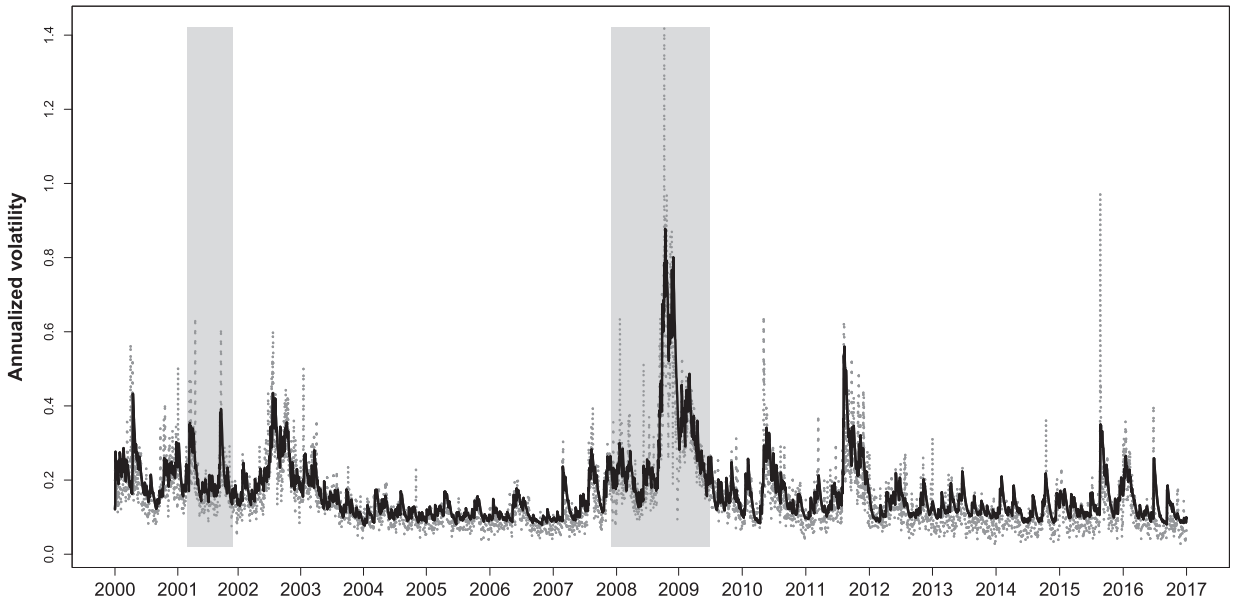
When the focus is on the average distance between the volatility proxy (a 5-min daily realized volatility augmented by squared overnight returns) and the estimated volatility, both LFs adopted signal the GJR-A-G-M$_{IP}$ as the best model. The LFs used are the QLIKE and MSE, both robust to the presence of some noise in the volatility proxy (Patton, 2011). Thus, for the rest of this section, the GJR-A-G-M$_{IP}$ model will be considered as benchmark.

The last six rows in Table 6 report some residual diagnostics. In particular, after having obtained the standardized residuals as $(r_{i,t} - \mu_{i,t})/\hat{\sigma}_{i,t}$, where $\hat{\sigma}_{i,t}$ represents the estimated conditional standard deviation, the Ljung-Box (LB) test and the Engle ARCH Lagrange Multiplier (LM) tests are applied. The LB test verifies the absence of autocorrelation up to a given lag, while the LM test checks that no other ARCH effects are present. Both the LB and LM tests, applied on squared standardized residuals, give evidence that all the models fit the data well, independently of the lag used in the tests.

The patterns of the GJR-A-G-M$_{IP}$ estimated conditional volatility and the *S&P*500 realized volatility are highlighted in Fig. 6. Globally, the estimated volatility of the GJR-A-G-M$_{IP}$ follows quite closely the *S&P*500 realized volatility.

5.2. Out-of-sample evaluation

This paragraph is devoted to verify if the chosen benchmark helps in better explaining the S&P500 volatility with respect to the competing models, in the out-of-sample perspective.



Notes: The figure plots the GJR-A-G-M$_{IP}$ volatility (solid black line) and the *S&P*500 realized volatility (dashed grey line). Shaded areas represent NBER recession periods.

Fig. 6. Realized and GJR-A-G-M$_{IP}$ volatilities.

In reference to the out-of-sample evaluation, we use the same rolling scheme of Amendola et al. (2017) and Asgharian et al. (2013). We obtain out-of-sample volatility forecasts for the period 2004–2016, by considering a rolling estimation window of 12 years. Therefore, the rolling forecasting scheme adopted is as follows:

1. Estimate model parameters using the period from January 1992 to December 2003.
2. Keep the estimated parameters fixed from January 2004 to December 2004. Obtain the out-of-sample volatility forecasts by considering the estimated parameters, observed daily return, *IP* and *NAI* values.
3. Move forward the rolling window by 1 year.
4. Repeat the procedure until the end of the series.

The previous scheme is adopted for all the competing models. The volatility proxy we use consists of the daily realized volatility, $S\&P500_{RV}$ (*RV*), obtained by taking the square root of the sum of squared 5-min intradaily and overnight returns.

The general performance of the proposed model against the competing specifications is syntesized in Table 7. In particular, according with Narayan and Bannigidadmath (2015), this latter table reports some absolute and relative goodness-of-fit indeces of our proposed specification with respect to the competing models, for the global out-of-sample period (2004–2016). In addition to QLIKE and MSE LFs, we also calculate the values of the Mean-Absolute-Error (MAE) and the Mincer-Zarnowitz R^2 (RMZ) from the regression:

$$RV_{i,t} = c + d\hat{\sigma}_{i,t} + \eta_{i,t}. \tag{17}$$

The relative goodness-of-fit indices are:

Table 7
Out-of-sample forecast evaluation.

	vs GARCH	vs GJR	vs GJR-G-M$_{IP}$	vs GJR-G-M$_{NAI}$	vs GJR-A-G-M$_{NAI}$
QLIKE	−3.588	−3.600	−3.594	−3.573	−3.580
MSE	0.019	0.018	0.016	0.017	0.016
MAE	3.009	2.898	2.815	2.897	2.887
R_{QLIKE}	1.020	1.017	1.018	1.025	1.022
R_{MSE}	1.322	1.225	1.095	1.149	1.096
R_{MAE}	1.153	1.110	1.078	1.110	1.105
R_{RMZ}	1.098	1.001	1.013	1.008	1.008

Notes: The table presents the forecast performance results for the GJR-A-G-M$_{IP}$ specification against the competing models, for the out-of-sample period 2004–2016, with a rolling window of 12 years. *MSE* and *MAE* value have been multiplied by 1000. R_{QLIKE}, R_{MSE} and R_{MAE} are the relative *QLIKE*, *MSE* and *MAE*. In particular, R_{QLIKE} is the ratio between the *QLIKE* of the considered benchmark, that is the GJR-A-G-M$_{IP}$ model, and the model in column. R_{MSE} and R_{MAE} are the ratio between the *MSE* and the *MAE* of the model in column and the GJR-A-G-M$_{IP}$ model. All the LFs consider as volatility proxy the realized volatility, obtained by taking the square root of the summation of the 5-min squared intradaily returns and overnight returns. R_{QLIKE}, R_{MSE} and R_{MAE} values larger than one indicate that the model in column under-performs the GJR-A-G-M$_{IP}$ specification. R_{RMZ} provides the relative Mincer-Zarnowitz $R,^2$, given by $R^2_{\text{GJR}-\text{A}-\text{G}-\text{M}_{IP}}/R^2_M$, where $R^2_{(\cdot)}$ is the R^2 of the Mincer-Zarnowitz regression (Eq. (17)). R_{RMZ} values greater than one indicate that the model in column has a smaller R^2 than that of the GJR-A-G-M$_{IP}$ specification.

Table 8
GW test-statistics.

	vs GARCH	vs GJR	vs GJR-G-M_{IP}	vs GJR-G-M_{NAI}	vs GJR-A-G-M_{NAI}
2004	$+^{***}$	$+^{***}$	$+$	$+^{***}$	$+^{*}$
2005	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2006	$-^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2007	$+^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2008	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2009	$+^{***}$	$+^{***}$	$+$	$+^{**}$	$+^{***}$
2010	$+^{***}$	$+^{***}$	$+$	$+^{***}$	$-^{***}$
2011	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2012	$+^{***}$	$+^{**}$	$+^{**}$	$+^{*}$	$+$
2013	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2014	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2015	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2016	$-^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2004–2016	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$

Notes: The table presents the signs of the two-tailed Giacomini–White (GW) test, whose null hypothesis is of equal predictive accuracy between models in columns and the GJR-A-G-M_{IP}, under the QLIKE LF. In particular, a "+" sign indicates that the GJR-A-G-M_{IP} outperforms the model in column, a "-" sign indicates the opposite. The volatility proxy consists of the daily realized volatility, computed by taking the square root of the summation of the 5-min squared intradaily returns and overnight returns. The rolling window is 12 years. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively.

$$R_{QLIKE} = \frac{QLIKE_{\text{GJR}-\text{A}-\text{G}-\text{M}_{IP}}}{QLIKE_M}, \tag{18}$$

$$R_{MSE} = \frac{MSE_M}{MSE_{\text{GJR}-\text{A}-\text{G}-\text{M}_{IP}}}, \tag{19}$$

$$R_{MAE} = \frac{MAE_M}{MAE_{\text{GJR}-\text{A}-\text{G}-\text{M}_{IP}}}, \tag{20}$$

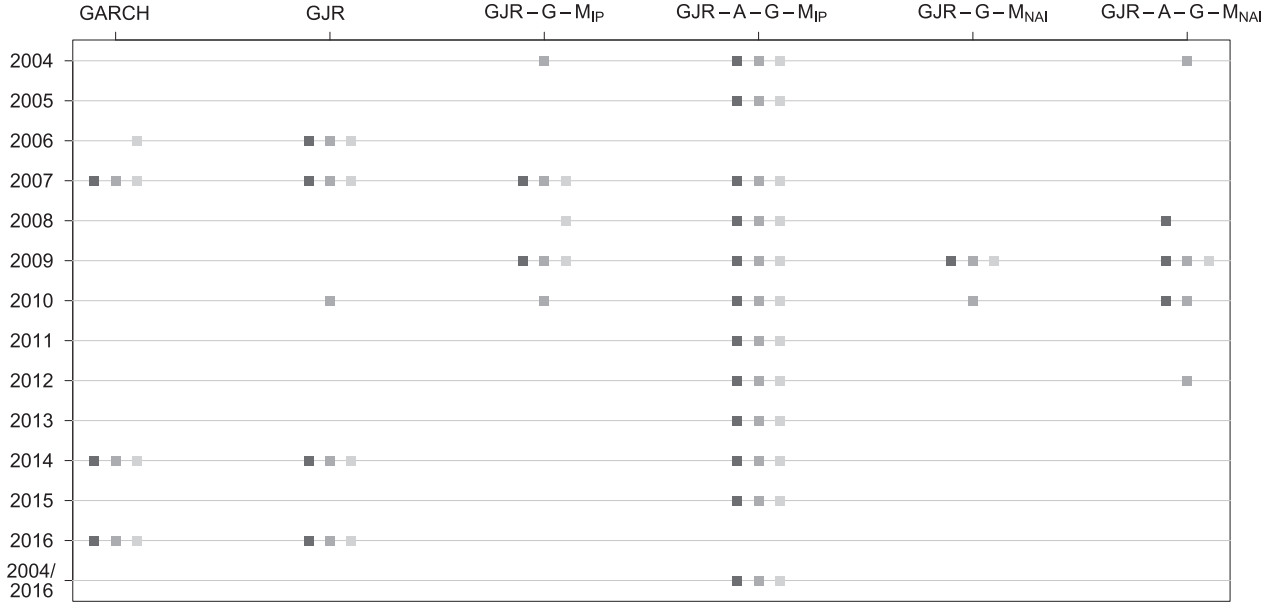
and

$$R_{RMZ} = \frac{RMZ_{\text{GJR}-\text{A}-\text{G}-\text{M}_{IP}}}{RMZ_M}, \tag{21}$$

where LF_M is the value of the LF for model M. The relative indices are built in a way that values larger than one signify that model M under-performs the GJR-A-G-M_{IP} specification according to that criterion. Table 7 clearly shows that the proposed specification always outperforms all the competing models, for the period out-of-sample period 2004–2016, both in terms of smaller LFs values and larger RMZ.

In order to further investigate the forecasting ability of the chosen specification, we use the Giacomini-White test (GW, Giacomini and White (2006)) and the Model Confidence Set (MCS) procedure (Hansen et al., 2011), year by year. The null of the two-tails GW test is of equal predictive ability between one of the competing model and the model that appears to be superior from an in-sample point of view (GJR-A-G-M_{IP}), that is:

$$H_0 : E(LF(\text{competing model}) - LF(\text{GJR} - \text{A} - \text{G} - \text{M}_{IP})) = 0.$$



Notes: The figure plots the composition of the Model Confidence Set (MCS). For different loss functions, dark (MSE), medium-dark (QLIKE) and light (MAE) shades of grey indicate that a given model is included in the MCS, at a significance level of $\alpha = 0.1$.

Fig. 7. MCS composition.

If the loss differential is positive, the proposed model has a superior forecasting ability. Otherwise, if the loss differential is negative, the competing model outperforms the GJR-A-G-M$_{IP}$. As regards the GW test, the LF employed in the analysis is the QLIKE, which has more power to reject inferior models, as suggested by Liu et al. (2015) and references therein. As done above, the LFs used in the MCS are the QLIKE, MSE and MAE.

Table 8 shows the results of the forecasts evaluation in terms of GW test. Almost always does the proposed model statistically outperform the competitors. For the whole out-of-sample period, the GJR-A-G-M$_{IP}$ has a superior forecasting ability than the other models. Focusing on the pairwise comparison between the proposed model and the GJR-G-M$_{MV}$, the latter never outperforms the GJR-A-G-M$_{IP}$, irrespective of the macro–variable *MV* used.

With reference to the MCS (Fig. 7), the GJR-A-G-M$_{IP}$ specification always belongs to the set of superior models, except in 2006 and 2016, regardless of the LFs adopted. More importantly, for the whole out-of-sample period, the proposed specification, with *IP* as additional volatility determinant, is the only model entering the MCS, again independently of the LFs adopted.

A common concern is whether a superior forecasting ability of the proposed model is also accompanied by some economic value (e.g. the information value of oil price for the volatility of a large panel of stocks, Narayan and Sharma (2014)). In this context, in line with He et al. (2015a) and He et al. (2015b), we evaluate the utility gain of the GJR-A-G-M$_{IP}$ over each competitor, assuming a mean-variance framework. In particular, considering a quadratic utility function, the utility of a trading strategy at day i, for a given model, is:

$$u_{i,t} = \mu_{i,t} - \frac{1}{2}\gamma_{RA}\hat{\sigma}^2_{i,t}, \tag{22}$$

where γ_{RA} is the coefficient of risk aversion. The utility gain of using the GJR-A-G-M$_{IP}$ model is given by the difference between the averages of Eq. (22) for the proposed specification and each of the alternative model; when expressed in annualized percentage terms and for three different coefficient of risk aversion, they are reported in Table 9. The GJR-A-G-M$_{IP}$ model brings more economic value than the GJR specification, independently of the coefficient of risk aversion adopted. When γ_{RA} increases, the utility gains guaranteed by the proposed model increase (and, in general, significantly so).

Table 9
Out-of-sample comparison of utility gain between GJR-A-G-M$_{IP}$ and the competing models.

	vs GARCH	vs GJR	vs GJR-G-M$_{IP}$	vs GJR-G-M$_{NAI}$	vs GJR-A-G-M$_{NAI}$
$\gamma_{RA} = 1$					
2004	1.337***	1.503***	0.656***	1.327***	1.071***
2005	0.234	0.556***	0.196	0.557***	0.951***
2006	−0.498**	−0.505**	0.351	0.236	1.409***
2007	−1.284**	−0.208	−0.021	0.31	1.405**
2008	2.832	8.754	0.332	4.033	0.661
2009	5.684***	4.696**	0.224	1.064	−0.572
2010	2.502***	2.403***	0.593	1.581**	0.305
2011	0.92	2.608	0.701	0.596	2.522
2012	0.969***	1.013***	1.08***	0.948***	0.755**
2013	0.23	0.225	0.725***	0.662***	0.904***
2014	−0.442	0.099	0.533*	0.5*	0.476
2015	−0.211	0.523	0.883	0.921	0.61
2016	−1.027**	−0.638	0.178	0.188	0.447
2004–2016	0.862	1.606**	0.495	0.989	0.842
$\gamma_{RA} = 3$					
2004	4.01***	4.51***	1.967***	3.981***	3.213***
2005	0.703*	1.668***	0.588	1.67***	2.852***
2006	−1.494***	−1.514***	1.053*	0.709	4.226***
2007	−3.853***	−0.624	−0.064	0.929	4.216**
2008	8.496	26.261	0.995	12.099	1.982
2009	17.053***	14.088***	0.673	3.191	−1.715
2010	7.505***	7.208***	1.778	4.742**	0.916
2011	2.761	7.825	2.104	1.789	7.565
2012	2.907***	3.04***	3.239***	2.845***	2.266***
2013	0.691	0.675	2.174***	1.98***	2.711***
2014	−1.327*	0.298	1.598**	1.501*	1.428*
2015	−0.634	1.569	2.65	2.764	1.83
2016	−3.082**	−1.913	0.535	0.565	1.341
2004–2016	2.586	4.819***	1.484	2.966*	2.525
$\gamma_{RA} = 5$					
2004	6.683***	7.516***	3.278***	6.636***	5.356***
2005	1.172*	2.78***	0.98	2.783***	4.754***
2006	−2.489***	−2.524***	1.755**	1.181	7.043***
2007	−6.422***	−1.039	−0.107	1.548	7.027***
2008	14.16	43.769*	1.658	20.164	3.304
2009	28.422***	23.48***	1.121	5.318	−2.859
2010	12.508***	12.014***	2.964	7.903**	1.526
2011	4.601	13.042	3.507	2.981	12.608
2012	4.845***	5.066***	5.398***	4.741***	3.776***
2013	1.151	1.125	3.623***	3.311***	4.518***
2014	−2.212*	0.496	2.663**	2.502**	2.381*
2015	−1.057	2.615	4.416	4.607	3.051
2016	−5.137**	−3.189	0.892	0.941	2.235
2004–2016	4.31*	8.031***	2.473	4.943*	4.208*

Notes: The table reports the utility gain of using the GJR-A-G-M$_{IP}$ model with respect to the model in column. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. See text for definitions.

Table 10
Robustness of GJR-A-G-M_{IP}: in-sample estimation.

	Period 1	Period 2	Period 3	Period 4	Period 5	Period 6	Period 7
Period	1980–1989	1990–2000	2000–2009	2010–2016	1984–Jun 2007	1985–2004	1986–2013
N	2528	2528	2515	1762	5927	5047	7060
α	0.017	0.001	0.001	0.001	0.001	0.002	0.001
	(0.027)	(0.016)	(0.017)	(0.106)	(0.004)	(0.007)	(0.007)
β	0.630***	0.834***	0.921***	0.748***	0.933***	0.917***	0.912***
	(0.210)	(0.032)	(0.020)	(0.094)	(0.011)	(0.024)	(0.013)
γ	0.240	0.204***	0.119***	0.313***	0.088***	0.107***	0.133***
	(0.187)	(0.049)	(0.022)	(0.072)	(0.013)	(0.027)	(0.018)
m	2.356***	1.726***	2.558***	2.700	4.057***	4.123***	4.485***
	(0.907)	(0.397)	(0.405)	(1.274)	(0.215)	(0.260)	(0.251)
θ^+	1.144*	2.685***	1.541***	1.548	0.939***	0.926***	0.706***
	(0.593)	(0.336)	(0.289)	(1.242)	(0.166)	(0.198)	(0.200)
ω_2^+	1.027**	2.038***	2.623***	1.305***	1.001***	1.001***	1.001***
	(0.468)	(0.247)	(0.559)	(0.154)	(0.227)	(0.221)	(0.265)
θ^-	−1.345***	−1.465***	−3.266***	−0.484*	−0.249**	−0.232*	−0.167**
	(0.336)	(0.355)	(0.423)	(0.269)	(0.124)	(0.134)	(0.081)
ω_2^-	1.001*	1.001***	1.001***	1.028***	7.628***	7.907***	3.506
	(0.542)	(0.284)	(0.334)	(0.148)	(0.157)	(0.250)	(2.414)
Diagnostics							
$LB(5)_{sq}$	0.959	0.160	0.029	0.734	0.752	0.972	0.084
$LB(10)_{sq}$	0.977	0.524	0.058	0.970	0.865	0.983	0.371
$LB(20)_{sq}$	0.998	0.728	0.351	0.997	0.967	0.994	0.623
$LM(5)_{sq}$	1.000	0.981	0.997	0.999	0.998	1.000	1.000
$LM(5)_{sq}$	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$LM(5)_{sq}$	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Notes: The table contains the estimated parameters for the GJR-A-G-M_{IP} model, over different periods. The dependent variable is the *S&P*500 volatility. The number of lagged realizations entering the long-run equation is $K = 48$ for period 1–4 and $K = 24$ elsewhere. Newey-West (HAC) standard errors are in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. $LB(l)_{sq}$ represents the p-values of the Ljung-Box test at l lag, applied on squared standardized residuals. $LM(l)_{sq}$ represents the p-values of the Engle's Lagrange Multiplier test for autoregressive conditional heteroskedasticity at l lag, applied on the squared standardized residuals.

Table 11
Robustness of out-of-sample forecast evaluation.

	vs GARCH	vs GJR	vs GJR-G-M_{IP}	vs GJR-G-M_{NAI}	vs GJR-A-G-M_{NAI}
Rolling period: 10 years					
QLIKE	−3.572	−3.587	−3.565	−3.562	−3.595
MSE	0.020	0.018	0.017	0.017	0.016
MAE	3.080	2.972	2.943	2.956	2.872
R_{QLIKE}	1.012	1.008	1.014	1.015	1.006
R_{MSE}	1.276	1.181	1.114	1.095	1.067
R_{MAE}	1.110	1.072	1.061	1.066	1.035
R_{RMZ}	1.099	0.999	1.011	1.002	1.006
Rolling period: 14 years					
QLIKE	−3.588	−3.603	−3.614	−3.595	−3.605
MSE	0.019	0.017	0.015	0.016	0.016
MAE	3.000	2.870	2.737	2.805	2.810
R_{QLIKE}	1.014	1.009	1.007	1.012	1.009
R_{MSE}	1.304	1.182	1.034	1.086	1.079
R_{MAE}	1.122	1.074	1.024	1.049	1.051
R_{RMZ}	1.098	0.998	1.003	1.001	1.008

Notes: The table presents the forecast performance results for the GJR-A-G-M_{IP} specification against the competing models, for the out-of-sample period 2004–2016, with a rolling window of 12 years. *MSE* and *MAE* value have been multiplied by 1000. R_{QLIKE}, R_{MSE} and R_{MAE} are the relative *QLIKE*, *MSE* and *MAE*. In particular, R_{QLIKE} is the ratio between the *QLIKE* of the considered benchmark, that is the GJR-A-G-M_{IP} model, and the model in column. R_{MSE} and R_{MAE} are the ratio between the *MSE* and the *MAE* of the model in column and the GJR-A-G-M_{IP} model. All the LFs consider as volatility proxy the realized volatility, obtained by taking the square root of the summation of the 5-min squared intradaily returns and overnight returns. R_{QLIKE}, R_{MSE} and R_{MAE} values larger than one indicate that the model in column under-performs the GJR-A-G-M_{IP} specification. R_{RMZ} provides the relative Mincer-Zarnowitz R,[2], given by $R^2_{\text{GJR-A-G-M}_{IP}}/R^2_M$, where $R^2_{(\cdot)}$ is the R^2 of the Mincer-Zarnowitz regression (Eq. (17)). R_{RMZ} values greater than one indicate that the model in column has a smaller R^2 than that of the GJR-A-G-M_{IP} specification.

Overall, both the in-sample and out-of-sample analyses show that the GJR-A-G-M$_{MV}$, with *IP* as an additional volatility determinant, is able to produce superior volatility forecasts, at least for the periods and data under consideration. Moreover, these forecasts yield also very often economically significant utility gains over the other models.

6. Robustness checks

For our model as a benchmark, in this section we evaluate if the results of the in-sample estimates are robust to different periods and if the out-of-sample performances, expressed in terms of GW test, are confirmed for small changes of the rolling window length.

6.1. In-sample robustness

The original in-sample period (1980–2016) is divided in four, non–overlapping, subsamples. Moreover, we also include results for three additional periods: 1984–06/2007, 1985–2004 and 1986–2013. The first period corresponds to the Great Moderation period (Giannone et al. (2008) and Stock and Watson (2002), among others). The interval 1985–2004 is one of the sub-sample used in Engle et al. (2013) and the period 1986–2013 has been used in Asgharian et al. (2015).

All the GJR-A-G-M$_{IP}$ estimates for these periods are in Table 10. Overall, the results of the whole sample period are largely confirmed. Interestingly, the parameters of interest, $\widehat{\theta}^{+}$ and $\widehat{\theta}^{-}$, do not change their signs but only their magnitudes. Also of interest is the fact that, while $\widehat{\omega}_2^{+}$ is approximately constant (confirming the behavior of the positive

Table 12
GW test-statistics for alternative rolling windows and LFs.

RW	vs GARCH		vs GJR		vs GJR-G-M$_{IP}$		vs GJR-G-M$_{NAI}$		vs GJR-A-G-M$_{NAI}$	
	10	14	10	14	10	14	10	14	10	14
QLIKE										
2004	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	–	$+^{***}$	$+^{**}$	$-^{**}$	$+^{***}$
2005	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$
2006	$-^{***}$	$-^{**}$	$-^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$
2007	$+^{***}$	$-^{***}$	$-^{***}$	$-^{***}$	$-^{***}$	$+^{**}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2008	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2009	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{*}$	$+^{***}$	$+^{*}$	+	$+^{*}$
2010	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	–	$+^{**}$	$+^{***}$	$+^{***}$	$+^{***}$
2011	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{*}$	+	+	$+^{**}$	$+^{*}$	$+^{***}$
2012	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	–	$-^{***}$
2013	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$-^{***}$
2014	$-^{***}$	$-^{***}$	$-^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2015	$-^{***}$	$-^{***}$	$-^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2016	$-^{***}$	$+^{***}$	$-^{***}$	$-^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$-^{***}$
2004–2016	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
MSE										
2004	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	+	$+^{***}$	$+^{***}$	–	$+^{***}$
2005	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$-^{***}$	$+^{***}$
2006	$-^{***}$	–	$-^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$
2007	$+^{*}$	+	$+^{***}$	$-^{***}$	$+^{***}$	$+^{**}$	$+^{***}$	$+^{**}$	$+^{***}$	$-^{***}$
2008	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2009	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	+	$+^{***}$
2010	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	–	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2011	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$-^{*}$	$+^{***}$
2012	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	–	$-^{***}$
2013	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$-^{***}$
2014	$-^{***}$	$-^{***}$	$-^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2015	–	+	$-^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$+^{**}$	$+^{***}$	$+^{***}$	$+^{***}$
2016	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$+^{*}$	$+^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$-^{***}$
2004–2016	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
MAE										
2004	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{**}$	$+^{***}$	$+^{***}$	$-^{***}$	$+^{***}$
2005	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$-^{***}$	$+^{***}$
2006	$-^{***}$	$-^{**}$	$-^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$
2007	$-^{**}$	$-^{**}$	$-^{***}$	$-^{***}$	$-^{***}$	–	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2008	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2009	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	+	$+^{***}$
2010	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{**}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2011	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$-^{**}$	$+^{***}$
2012	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$-^{**}$	$-^{***}$
2013	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$-^{***}$
2014	$-^{***}$	$-^{***}$	$-^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2015	$-^{***}$	$-^{***}$	$-^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
2016	$-^{***}$	$+^{**}$	$-^{***}$	$-^{***}$	$-^{***}$	$+^{***}$	$-^{***}$	$+^{***}$	$+^{***}$	$-^{***}$
2004–2016	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$

Notes: The table presents the signs of the two-tailed Giacomini–White (GW) test, under a null hypothesis of equal predictive accuracy between models in columns and the GJR-A-G-M$_{IP}$, under the QLIKE, MSE and MAE LFs and two different rolling windows (RWs). In particular, a "+" sign indicates that the GJR-A-G-M$_{IP}$ outperforms the model in column, a "-" sign indicates the opposite. The volatility proxy consists of the daily realized volatility, computed by taking the square root of the sum of the 5-min squared intradaily returns and the overnight return. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively.

Table 13
Additional Granger tests.

Area/Country	Test	Lag	1	2	3	4	5	6	7	8	9	10	11	12
Europe	$IP(lag) \xrightarrow{GnC} STOXX50E_{ARV}$	F	16.998	8.061	4.464	4.499	4.748	3.882	3.306	2.988	2.706	2.340	2.102	2.365
		p-value	0.000	0.000	0.005	0.002	0.000	0.001	0.002	0.004	0.006	0.013	0.023	0.008
Germany	$IP(lag) \xrightarrow{GnC} GDAXI_{ARV}$	F	14.358	6.663	3.872	3.930	4.443	3.575	3.043	2.769	2.506	2.152	1.968	2.133
		p-value	0.000	0.002	0.010	0.004	0.001	0.002	0.005	0.007	0.010	0.023	0.034	0.017
France	$IP(lag) \xrightarrow{GnC} FCHI_{ARV}$	F	17.231	8.049	4.562	4.797	4.859	3.925	3.358	2.970	2.687	2.369	2.128	2.402
		p-value	0.000	0.000	0.004	0.001	0.000	0.001	0.002	0.004	0.006	0.012	0.021	0.007
Canada	$IP(lag) \xrightarrow{GnC} GSPTSE_{ARV}$	F	43.449	20.546	14.342	12.416	12.428	10.541	8.777	7.472	7.531	5.979	6.007	5.600
		p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Hong Kong	$IP(lag) \xrightarrow{GnC} HSI_{ARV}$	F	12.643	9.075	7.001	6.227	6.135	5.138	4.342	3.643	3.269	2.950	2.811	2.644
		p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.002	0.003
Brazil	$IP(lag) \xrightarrow{GnC} BVSP_{ARV}$	F	36.401	18.352	12.519	13.048	10.384	8.862	7.264	6.460	6.280	5.396	4.831	4.466
		p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: The table shows the F-statistics and the relative p-values of the Granger tests for the area/country in the first column. The expression "X GnC Y" means that "X does not Granger cause Y". $STOXX50E_{ARV}$, $GDAXI_{ARV}$, $FCHI_{ARV}$, $GSPTSE_{ARV}$, HSI_{ARV} and $BV\ SP_{ARV}$ stand for Euro Stoxx 50, DAX, CAC 40, S&P/TSX Composite Index, Hang Seng Index and BVSP BOVESPA Index monthly aggregated realized volatility, respectively. The monthly aggregated realized volatility has been obtained summing the daily realized volatility. The daily realized volatility is computed by taking the square root of the summation of the 5-min squared intradaily and overnight returns. The regressions cover the period January 2000–December 2016, for a total of 204 observations.

Table 14
Advanced and emerging markets.

Area/Country	Europe	Germany	France	Canada	Honk Kong	Brazil
Dep. Variab.	Euro Stoxx 50	DAX	CAC 40	S&P/TSX	HANG SENG	BVSP BOVESPA
Period	1987–2016	1988–2016	Mar 1990–2016	May 1980–2016	1987–2016	1997–2016
N	7667	7326	6801	9335	7413	4950
Estimates						
α	0.007	0.007	0.001	0.044^{***}	0.031^{***}	0.006
	(0.007)	(0.007)	(0.007)	(0.010)	(0.006)	(0.007)
β	0.905^{***}	0.901^{***}	0.901^{***}	0.889^{***}	0.905^{***}	0.913^{***}
	(0.009)	(0.009)	(0.009)	(0.020)	(0.010)	(0.015)
γ	0.129^{***}	0.131^{***}	0.144^{***}	0.091^{***}	0.082^{***}	0.098^{***}
	(0.015)	(0.016)	(0.015)	(0.025)	(0.014)	(0.021)
m	4.145^{***}	4.634^{***}	4.610^{***}	4.982^{***}	5.269^{***}	7.067^{***}
	(0.367)	(0.287)	(0.209)	(0.188)	(0.242)	(0.184)
θ^+	1.346^{***}	1.070^{***}	1.120^{***}	0.109	0.773^{***}	−0.449
	(0.265)	(0.222)	(0.143)	(0.074)	(0.190)	(0.130)
ω_2^+	1.126^{***}	1.001^{***}	1.139^{***}	3.040^{***}	1.614^{***}	1.468
	(0.282)	(0.342)	(0.267)	(0.234)	(0.030)	(0.299)
θ^-	-0.318^{***}	-0.202^{***}	-0.256^{***}	-0.107^{***}	-0.113^{***}	0.191^{***}
	(0.121)	(0.061)	(0.086)	(0.041)	(0.038)	(0.121)
ω_2^-	3.837^{***}	7.887^{***}	4.983^{***}	18.263^{*}	19.716^{***}	1.001^{***}
	(0.394)	(0.342)	(0.224)	(9.582)	(0.732)	(1.151)
Diagnostics						
R_{QLIKE}	0.945	0.950	0.946	0.988	0.956	0.985
R_{MSE}	0.760	0.840	0.823	0.925	0.873	0.868
R_{MAE}	0.924	0.902	0.938	0.967	0.978	0.919
$RMZ_{GJR-A-G-M}$	0.639	0.659	0.621	0.581	0.504	0.422
RMZ_{GARCH}	0.588	0.602	0.555	0.559	0.486	0.393
$LB(5)_{sq}$	0.852	0.900	0.393	0.724	0.537	0.030
$LB(10)_{sq}$	0.987	0.996	0.704	0.841	0.689	0.025
$LB(20)_{sq}$	0.998	1.000	0.654	0.609	0.948	0.008
$LM(5)_{sq}$	1.000	1.000	0.996	1.000	1.000	0.828
$LM(10)_{sq}$	1.000	1.000	1.000	1.000	1.000	0.930
$LM(20)_{sq}$	1.000	1.000	1.000	1.000	1.000	0.969

Notes: The table reports the results of the GJR-A-G-M model in advanced and emerging countries (name of indices in the first row). N represents the total number of daily observations. The additional volatility determinant is *IP*. Newey-West (HAC) standard errors are in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. In panel "Diagnostics", R_{QLIKE}, R_{MSE} and R_{MAE} are the relative *QLIKE*, *MSE* and *MAE*. In particular, R_{QLIKE} is the ratio between the *QLIKE* of the considered benchmark, that is the GARCH model, and the GJR-A-G-M model. R_{MSE} and R_{MAE} are the ratio between the *MSE* and the *MAE* of GJR-A-G-M model and the GARCH model. As a volatility proxy, all the LFs consider the realized volatility, obtained by taking the square root of the summation of the 5-min squared intradaily returns and overnight returns, for the period 2000–2016. $RMZ_{GJR-A-G-M}$ and RMZ_{GARCH} provide the Mincer-Zarnowitz R,[2], obtained from the regression (17). $LB(l)_{sq}$ represents the p-values of the Ljung-Box test at l lag, applied on squared standardized residuals. $LM(l)_{sq}$ represents the p-values of the Engle's Lagrange Multiplier test for autoregressive conditional heteroskedasticity at l lag, applied on the squared standardized residuals.

portion of RME_k^{τ} provided in Fig. 4), $\widehat{\omega}_2^-$ varies considerably, showing, at times, a value close to 1, which amounts to a larger weight given to all negative past values of X_{t-k} in the negative portion of the RME_k^{τ}.

6.2. Out-of-sample robustness

The general performances of the GJR-A-G-M_{IP} against the competing models, under rolling windows of length 10 and 14 years, are exposed in Table 11. Independently of the size of the rolling window adopted, the performances of the previous section concerning the GJR-A-G-M_{IP} are largely confirmed. It results that the benchmark specification has almost always smaller LF values and higher RMZ.

The GW tests for rolling windows of length 10 and 14 years are reported in Table 12, under the QLIKE, MSE and MAE LFs. Again, the specification to beat is the GJR-A-G-M_{IP}.

With a smaller rolling window but irrespective of the LF adopted, the forecasting performances of the GJR-A-G-M_{IP} get slightly worse. For instance, in 2005, for a rolling window of 10 years, the GJR-A-G-M_{IP} model has always an inferior forecasting ability relative to all the other models, except the GJR-G-M_{NAI}. When the rolling window increases to 14 years, under the same LFs, GJR-A-G-M_{IP} outperforms all the competing models.

Overall, the results obtained in the previous section are robust to small changes in the rolling window and LF employed in the GW test. It is important to note that smaller rolling windows make volatility forecasts perform badly, for both MIDAS models. In fact, for rolling windows smaller than a sufficient number of years, the macro–variable observations entering these models are too few for log-likelihood maximization to converge.

6.3. Extension to an international context robustness

We can further verify whether *IP* can be considered as an additional volatility determinant also in other (both advanced and emerging) area/countries, and, consequently, whether the proposed model helps in explaining the long–term volatility component. This time, the dependent variable of interest is given by the indices of four developed and two developing area/countries: Euro Stoxx 50 for the Euro–area, DAX for Germany, CAC 40 for France, S&P/TSX for Canada, HANG SENG for Hong Kong and BVSP BOVESPA for Brazil.

First of all, we evaluate if all the resulting aggregated monthly realized volatilities are Granger caused by (lagged) *IP*, as done in Table 1. The results of the new Granger tests are in Table 13. Independently of the considered lag, all the previous aggregated monthly realized volatilities are Granger caused by the *IP* growth rate.

Then, we estimate the GJR-A-G-M_{IP} for all the considered markets, as reported in Table 14. Due to the different availability of the daily returns for each index, the sample periods change across the markets. Overall, we obtain very stable and good results. Synthesizing, the proposed model has always smaller QLIKE, MSE and MAE values relatively to the chosen benchmark, the GARCH(1,1), independently of the area/country under investigation. Moreover, the proposed specification has always larger RMZ. Finally, the ARCH test robustly signals the absence of conditional heteroskedasticity in the squared standardized residuals. Thus, we can conclude that the proposed model is able to detect asymmetries in macro-variables affecting stock market volatilities also in an international setting.

7. Conclusions

Inspired by the literature on the influence of macro–variables on volatility, this paper focuses on a research question about whether and by how much their positive and negative changes affect volatility of returns. To this end, we propose an asymmetric GARCH-MIDAS model that explicitly allows us to take into account positive and negative macroeconomic variations. The proposed specification is synthetically labelled GJR-A-G-M, because it also includes an asymmetric term for negative returns in the short–run component, as in the GJR model. First of all, we empirically verify that two macro–variables - volatility relationships hold if positive and negative changes of the economic variables are considered separately. The macro–variables used in this work are U.S. Industrial Production *IP* and the National Activity Index *NAI*, both Granger regressors for the stock market volatility, proxied by *S&P*500 realized volatility. Interestingly, lagged *IP* and *NAI* values are no statistically significant regressors for the stock market volatility. But, if considered separately, positive and negative lagged *IP* and *NAI* values are statistically significant determinants of *S&P*500 volatility. Thus, in order to take advantage of this aspect, we have modified the long-run equation of the GARCH-MIDAS model. In this new specification, a coefficient linked only to positive variations of the additional exogenous variable and another coefficient depending only on negative variations are considered. The GJR-A-G-M model was evaluated under an extensive Monte Carlo experiment. With the increase in the number of observations, the bias between the true and estimated parameters decreases. In addition, the MSE between the true and estimated volatility also decreases, though at a small rate.

We test our proposed model considering both *IP* and *NAI* as additional determinants for (daily) *S&P*500 volatility, against the standard GARCH-MIDAS, GARCH and GJR models. In the in–sample period 1980–2016, the GJR-A-G-M model confirms the intuition that positive and negative lagged *IP* changes significantly contribute to the long–run component of the model. This specification assures the best in-sample fit, in terms of AIC, QLIKE and MSE LFs, as well as R^2 in the Mincer–Zarnowitz regressions, making the GJR-A-G-M with *IP* the model to beat, in the out-of-sample analysis. Under a rolling forecasting scheme, the GJR-A-G-M_{IP} volatility forecasts are always superior (and mostly significantly so) to those of the competing models, when the Giacomini–White test is used. When the MCS procedure is employed, under the QLIKE LF, the set of superior models consists only of the GJR-A-G-M_{IP}. Moreover, the GJR-A-G-M_{IP} brings about superior utility gains than the competing models. The asymmetric impact of *IP* changes is confirmed also for different sub-samples. The out-of-sample GJR-A-G-M_{IP} performances are robust to moderate changes of the rolling estimation width and the loss functions employed in the evaluation. Finally, the different impact of macro–variables is also tested in an international setting, both for advanced and emerging markets. In particular, six additional indeces are investigated: Euro Stoxx 50 for the Euro–area, DAX for Germany, CAC 40 for France, S&P/TSX for Canada, Hang Seng for Hong Kong and BVSP BOVESPA for Brazil, respectively. Provided that all the previous indeces are Granger explained by *IP*, we obtain very stable and good results, confirming the asymmetric impact of the macro–variables on the long-run volatility. Synthesizing, also in the considered international setting, the proposed model has always smaller QLIKE, MSE and MAE values relatively to the chosen benchmark, the GARCH(1,1), independently of the area/country under investigation.

The question of the informational content of the variable driving the long–run multiplicative component can be further analyzed similarly to what Narayan and Sharma (2014) proposed to investigate how oil prices affect volatility. In our context, a large panel of stock prices can be analyzed to evaluate the additional contribution to volatility of oil price movements which is observable even at high frequency. We envisage that these monthly and daily movements can be used in our Asymmetric GARCH–MIDAS model, as drivers of the long– and the short–term components of volatility to evaluate where the MIDAS contribution fits in terms of a profit–loss appraisal.

Acknowledgements

We would like to thank the Editor and three anonymous referees for their constructive comments and suggestions which greatly helped

us in refining our work. We also thank Fabrizio Cipollini as well as participants in seminars at CFE 2017, London, for useful comments.

References

Amado, C., Teräsvirta, T., 2013. Modelling volatility by variance decomposition. J. Econom. 175 (2), 142–153.

Amendola, A., Candila, V., Scognamillo, A., 2017. On the influence of US monetary policy on crude oil price volatility. Empir. Econ. 52 (1), 155–178.

Asgharian, H., Christiansen, C., Hou, A.J., 2015. Macro-finance determinants of the long-run stock–bond correlation: the DCC-MIDAS specification. J. Financ. Econom. 14, 617–642.

Asgharian, H., Hou, A.J., Javed, F., 2013. The importance of the macroeconomic variables in forecasting stock return variance: a GARCH-MIDAS approach. J. Forecast. 32 (7), 600–612.

Bauwens, L., Braione, M., Storti, G., 2017. A dynamic component model for forecasting high-dimensional realized covariance matrices. Economet. Stat. 1, 40–61.

Bauwens, L., Storti, G., 2009. A component GARCH model with time varying weights. Stud. Nonlinear Dynam. Econom. 13 (2), 1–33.

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

Brownlees, C.T., Cipollini, F., Gallo, G.M., 2012. Multiplicative error models. In: Bauwens, L., Hafner, C., Laurent, S. (Eds.), Volatility Models and Their Applications. Wiley, pp. 223–247.

Brownlees, C.T., Gallo, G.M., 2010. Comparison of volatility measures: a risk management perspective. J. Financ. Econom. 8, 29–56.

Cipollini, F., Gallo, G.M., 2018. Modeling Euro STOXX 50 Volatility with Common and Market-specific Components. Technical Report WP18-26, Rimini Center for Economic Activity.

Colacito, R., Engle, R.F., Ghysels, E., 2011. A component model for dynamic correlations. J. Econom. 164 (1), 45–59.

Cologni, A., Manera, M., 2009. The asymmetric effects of oil shocks on output growth: a Markov–Switching analysis for the G-7 countries. Econ. Modell. 26 (1), 1–29.

Conrad, C., Loch, K., 2015. Anticipating long-term stock market volatility. J. Appl. Econom. 30 (7), 1090–1114.

Conrad, C., Loch, K., Rittler, D., 2014. On the macroeconomic determinants of long-term volatilities and correlations in U.S. stock and crude oil markets. J. Empir. Finance 29, 26–40.

Ding, Z., Granger, C.W., 1996. Modeling volatility persistence of speculative returns: a new approach. J. Econom. 73 (1), 185–215.

Engle, R.F., Ghysels, E., Sohn, B., 2013. Stock market volatility and macroeconomic fundamentals. Rev. Econ. Stat. 95 (3), 776–797.

Engle, R.F., Lee, G.J., 1999. A permanent and transitory component model of stock return volatility. In: Engle, R.F., White, H. (Eds.), Cointegration, Causality, and Forecasting: a Festschrift in Honor of Clive W. J. Granger. Oxford University Press, Oxford, pp. 475–497.

Engle, R.F., Mezrich, J., 1996. GARCH for groups. Risk 9, 36–40.

Engle, R.F., Ng, V., 1993. Measuring and testing the impact of news on volatility. J. Finance 48, 1749–1778.

Engle, R.F., Rangel, J.G., 2008. The spline-GARCH model for low frequency volatility and its global macroeconomic causes. Rev. Financ. Stud. 21, 1187–1222.

Francq, C., Horvath, L., Zakoïan, J.-M., 2011. Merits and drawbacks of variance targeting in GARCH models. J. Financ. Econom. 9, 619–656.

Ghysels, E., Santa-Clara, P., Valkanov, R., 2005. There is a risk-return trade-off after all. J. Financ. Econ. 76 (3), 509–548.

Ghysels, E., Sinko, A., Valkanov, R., 2007. MIDAS regressions: further results and new directions. Econom. Rev. 26 (1), 53–90.

Giacomini, R., White, H., 2006. Tests of conditional predictive ability. Econometrica 74 (6), 1545–1578.

Giannone, D., Lenza, M., Reichlin, L., 2008. Explaining the great moderation: it is not the shocks. J. Eur. Econ. Assoc. 6 (2–3), 621–633.

Glosten, L.R., Jagannanthan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. J. Finance 48 (5), 1779–1801.

Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37 (3), 424–438.

Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. Econometrica 79 (2), 453–497.

He, Z.L., Zhu, J., Zhu, X., 2015a. Dynamic factors and asset pricing: international and further US evidence. Pac. Basin Finance J. 32, 21–39.

He, Z.L., Zhu, J., Zhu, X., 2015b. Multi-factor volatility and stock returns. J. Bank. Finance 61, S132–S149.

Liu, L.Y., Patton, A.J., Sheppard, K., 2015. Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. J. Econom. 187 (1), 293–311.

Mo, D., Gupta, R., Li, B., Singh, T., 2017. The Macroeconomic Determinants of Commodity Futures Volatility: Evidence from Chinese and Indian Markets. Economic Modelling, Forthcoming.

Moshiri, S., Banihashem, A., 2012. Asymmetric Effects of Oil Price Shocks on Economic Growth of Oil-exporting Countries. Technical Report 2006763. SSRN.

Narayan, P.K., Bannigidadmath, D., 2015. Are Indian stock returns predictable? J. Bank. Finance 58, 506–531.

Narayan, P.K., Sharma, S.S., 2014. Firm return volatility and economic gains: the role of oil prices. Econ. Modell. 38, 142–151.

Officer, R.R., 1973. The variability of the market factor of the New York Stock Exchange. J. Bus. 46 (3), 434–453.

Patton, A., 2011. Volatility forecast comparison using imperfect volatility proxies. J. Econom. 160 (1), 246–256.

Raftery, A.E., 1995. Bayesian model selection in social research. Socio. Meth. 111–163.

Schwert, G.W., 1989. Why does stock market volatility change over time? J. Finance 44 (5), 1115–1153.

Segal, G., Shaliastovich, I., Yaron, A., 2015. Good and bad uncertainty: macroeconomic and financial market implications. J. Financ. Econ. 117 (2), 369–397.

Stock, J.H., Watson, M.W., 2002. Has the business cycle changed and why? NBER Macroecon. Annu. 17, 159–218.

Wang, F., Ghysels, E., 2015. Econometric analysis of volatility component models. Econom. Theor. 31 (2), 362–393.