# Book of Short Papers SIS 2020

Editors: Alessio Pollice, Nicola Salvati and Francesco Schirripa Spagnolo

PUBLISHED BY PEARSON

WWW.PEARSON.COM

*ISBN 9788891910776*

# Contents

## Specialized sessions

# Solicited Sessions

# Contributed papers and Posters

# Double Asymmetric GARCH–MIDAS model - new insights and results

## *Modello Double Asymmetric GARCH–MIDAS: nuovi approfondimenti e risultati*

Alessandra Amendola, Vincenzo Candila, Giampiero M. Gallo

**Abstract** The recently proposed Double Asymmetric GARCH-MIDAS (DAGM) model aims at separating the positive and negative macro variable variations within the long-run term and adds an asymmetric effect in the short-run component. In this work, the intent is to further extend the model in two main directions. A realized measure is included as a daily lagged variable in the short-run component (the so-called "–X" term) and a multi-step-ahead forecasting procedure is implemented for the class of GARCH–MIDAS (GM) models with the additional "–X" term. The extended DAGM-X model, which nests the DAGM and GM, is extensively evaluated under alternative configurations concerning the S&P 500 Index.

**Abstract** *Il presente lavoro illustra una estensione del modello Double Asymmetric GARCH–MIDAS (DAGM), recentemente proposto. Nella modellizzazione, oltre agli effetti asimmetrici nelle componenti di lungo e di breve periodo, è stata introdotta una misura di volatilità realizzata giornaliera come variabile addizionale per la componente di breve periodo (la cosiddetta parte "–X"). Inoltre, è stata sviluppata una procedura per le previsioni multi-step-ahead, valida per tutti i modelli GARCH–MIDAS (GM), anche con un termine aggiuntivo "–X". La performance del DAGM–X, che generalizza il modello DAGM e il modello GM, è stata valutata in riferimento all'indice S&P 500.*

**Key words:** Volatility, Asymmetry, GARCH–MIDAS, Forecasting.

---

Alessandra Amendola
Department of Economics and Statistics, University of Salerno, Italy, e-mail: alamendola@unisa.it

Vincenzo Candila
MEMOTEF Depart., Sapienza University of Rome, Italy, e-mail: vincenzo.candila@uniroma1.it

Giampiero M. Gallo
Italian Court of Audits (Corte dei conti – disclaimer) and NYU in Florence, Italy, e-mail: giampiero.gallo@nyu.edu

## 1 Introduction

The connection between the volatility of financial assets and the macroeconomic variables (MVs) has a long history. During the last decade, the GARCH–MIDAS model has been successfully used to include such a MVs into the so-called long-run volatility component, which varies with the same frequency as the MVs and around which the daily short-run component fluctuates as well. The Double Asymmetric GARCH–MIDAS (DAGM) model, recently proposed in Amendola et al. (2019), separates the positive and negative MV variations within the long-run term and adds an asymmetric effect in the short-run part. In this paper, we push one step further the DAGM model, including a realized measure as a daily lagged variable in the short-run component (the so-called "–X" term). Moreover, we introduce a multi-step-ahead forecasting procedure suitable for the class of GARCH–MIDAS (GM) models (Engle et al., 2013) with the "–X" term. The proposed DAGM–X model, which nests the standard DAGM and GM, is extensively evaluated under several alternative configurations, focusing on the S&P 500 Index. The daily S&P 500 5-minute realized volatility is included as a daily lagged variable in the short-run component while the rate of change of the monthly US Industrial Production, Housing Starts, and New Orders are used as the driver of market volatility. The idea is that a succession of negative values (over a somewhat long past, and suitably weighted) transmits a different type of increasing impulse to volatility than its positive counterpart. The empirical results give evidence that the DAGM–X model can outperform the standard DAGM, GM, and GM–X models, independently of the forecasting horizon considered. Furthermore, the proposed model is very often superior to the well known HAR model of Corsi (2009), mainly if longer forecasting horizons are taken into account.

## 2 DAGM–X model

Letting $r_{i,t}$ represent the log–return, that is, the first difference of the log–closing prices for day $i$ of the period (week or month) $t$ (with $i = 1, \cdots, N_t$, where $N_t$ is the number of days for period $t$), our GM framework defines:

$$r_{i,t} = \sqrt{\tau_t \times g_{i,t}}\varepsilon_{i,t}. \tag{1}$$

In this expression, $\varepsilon_{i,t}$ is the innovation term, $g_{i,t}$ follows a unit–mean reverting GJR-GARCH(1,1) process (short–run component), and $\tau_t$ provides the slow–moving local level of volatility (long–run component).

The short–run component is then given by:

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

where $\mathbb{1}_{(.)}$ is an indicator function.

The long–run component is defined as:

$$\tau_t = exp\left(m + \theta^+ \sum_{k=1}^{K} \delta_k(\omega)^+ MV_{t-k} \mathbb{1}_{(MV_{t-k} \geq 0)} + \theta^- \sum_{k=1}^{K} \delta_k(\omega)^- MV_{t-k} \mathbb{1}_{(MV_{t-k} < 0)}\right), \tag{3}$$

where $m$ plays the role of an intercept, $\theta^+$ and $\theta^-$ represent the asymmetric responses to the one–sided filter, and $\delta_k(\omega)^+$ and $\delta_k(\omega)^-$ are suitable functions weighing the past $K$ realizations of the additional stationary predetermined variable labelled $MV_t$ as the *MIDAS variable*. Throughout this work, the Beta function will be used as weighting function of all the GM models, that is:

$$\delta_k(\omega)^+ = \frac{(k/K)^{\omega_1^+ - 1}(1-k/K)^{\omega_2^+ - 1}}{\sum_{j=1}^{K} (j/K)^{\omega_1^+ - 1}(1-j/K)^{\omega_2^+ - 1}}, \tag{4}$$

$$\delta_k(\omega)^- = \frac{(k/K)^{\omega_1^- - 1}(1-k/K)^{\omega_2^- - 1}}{\sum_{j=1}^{K} (j/K)^{\omega_1^- - 1}(1-j/K)^{\omega_2^- - 1}}, \tag{5}$$

with the restriction $\omega_1^+ = \omega_1^- = 1$, which gives a higher weight to the most recent observations (monotonically decreasing weighting scheme). This is in line with what suggested in Ghysels and Qian (2019). Note that the Beta functions assure that $\sum_{k=1}^{K} \delta_k(\omega_2^+) = 1$ and $\sum_{k=1}^{K} \delta_k(\omega_2^-) = 1$.

The extension we suggest is to extend the short-run equation to some additional volatility determinants, observed at the same frequency of $r_{i,t}$. This allows us to move out of the classical GM framework with just MVs. In particular, the short-run component will change to:

$$g_{i,t} = (1-\alpha-\beta-\gamma/2) + \left(\alpha + \gamma \cdot \mathbb{1}_{(r_{i-1,t}<0)}\right)\frac{(r_{i-1,t})^2}{\tau_t} + \beta g_{i-1,t} + \tilde{z} \cdot X_{i-1,t}, \tag{6}$$

where the variable $X_{i-1,t}$ is observed at the same frequency as $r_{i,t}$.[1] We define the model with the short-run component in (6) as DAGM–X, which nests the DAGM if $z = 0$. In order to estimate the DAGM–X (and, hence, also the GM–X), the following assumptions are made:

**Assumption 1** *The innovation* $\varepsilon_{i,t}$ *in (1) is iid, with* $E\left[\varepsilon_{i,t}\right] = 0$ *and* $E\left[\varepsilon_{i,t}^2\right] = 1$.

**Assumption 2** *The short-run parameters are subject to:* $\alpha > 0$; $\beta \geq 0$; $\alpha + \beta + \gamma/2 < 1$;

**Assumption 3** $\tilde{z} \geq 0$; $X_{i,t} \geq 0$, $\forall i$ *and* $t$; $X_{i,t}$ *is stationary and ergodic.*

[1] Note that, because of this extension,

$$E(g_{i,t}) = 1 + \frac{\tilde{z}}{(1-\alpha-\beta-\gamma/2)}\bar{X} \equiv 1 + z\bar{X}.$$

Assumptions 1–2 are standard in the GARCH literature. Together with Assumption 3 (Han and Kristensen, 2014), the positiveness of $g_{i,t}$, $\forall i$ and $t$ is guaranteed. The parameter space of the DAGM–X- model[2] is $\Theta = \left\{\alpha, \beta, \gamma, \tilde{z}, m, \theta^{+}, \omega_2^{+}, \theta^{-}, \omega_2^{-}\right\}$. The maximum likelihood (ML) estimates for $\Theta$ are obtained once a distribution is chosen for the innovation term in (1). If $\varepsilon_{i,t}$, conditional on the information set up to day $i-1$ of period $t$, denoted by $\Phi_{i-1,t}$, is assumed to be standard normally distributed, then the following log-likelihood can be maximized:

$$\mathscr{L}(\Theta) = -\frac{1}{2}\sum_{t=t_s}^{T}\sum_{i=1}^{N_t}\left[\log(2\pi) + \log(g_{i,t}\tau_t) + \frac{r_{i,t}^2}{g_{i,t}\tau_t}\right], \tag{7}$$

where $t_s$ is taken sufficiently large in order to assure that different models, with different choices of $K$, will have the same information set. This is crucial when several models, with alternative $K$ lags, are evaluated in terms of information criteria.

If it is straightforward to derive the one-step-ahead volatility, the multi-step-ahead predictions require some preliminary assumptions, mainly if the model at hand includes the "–X" part. Firstly, let $\sigma_{i,t} = \sqrt{\tau_t \times g_{i,t}}$ be the standard deviation in a GM framework, conditionally to $\Phi_{i-1,t}$. Moreover, let $\Phi_{N_T,T}$ denote the information set available the last day ($N_T$) of the last period ($T$). Let $\tilde{\sigma}_{f,T+1}$ be the conditional standard deviation forecast for the day $f$ of the period $T+1$, conditional on the information set $\Phi_{N_T,T}$, that is: $\tilde{\sigma}_{f,T+1} = \left(\sqrt{g_{f,T+1} \times \tau_{T+1}}|\Phi_{N_T,T}\right)$. The main difference between $\tilde{\sigma}_{f,T+1}$ and $\sigma_{i,t}$ is that the former is based on a fixed information set, while the latter on an information set which updates daily. The one-step-ahead predictions depend on $\tau_t$, which in turn is based on the MIDAS variable observed up to period $t-1$. Therefore, in the multi-step-ahead context, $\tau_{T+1}$ is a function of historical values of the MIDAS variable observed up to period $T$, which are included in $\Phi_{N_T,T}$ and is kept constant for all days in $T+1$. The multi-step-ahead predictions require an estimate for the short-run component $g_{f,T+1}$: extending, e.g. Conrad and Kleen (2019), we have:

$$E(g_{f,T+1}|\Phi_{N_T,T}) = 1 + z\cdot\bar{X} + (\alpha+\beta+\gamma/2)^{f-1}\left(g_{1,T+1} - 1 - z\cdot\bar{X}\right), \tag{8}$$

where $\bar{X}$ is the average of the X variable calculated over the period up to $\Phi_{N_T,T}$. Note that, as usual, the short-run prediction depends just on $g_{1,T+1}$ (a function of the information set $\Phi_{N_T,T}$) and reduces to the customary $1 + (\alpha+\beta+\gamma/2)^{f-1}\left(g_{1,T+1} - 1\right)$ when the short-run unconditional expectation is equal to one ($z = 0$). In either case, as $f$ increases, the short-run prediction will approach its unconditional mean, and that will combine with $\tau_{T+1}$ to produce the overall forecast.

[2] Recall the DAGM–X nests the DAGM and the GM. Therefore, the following discussion can be naturally applied to all the other models, once that the parameter space has been reduced.

## 3 Empirical application

The empirical application focuses on the volatility of S&P 500 Index, while the MVs are the US Industrial Production ($IP$), Housing Starts ($HS$), and New Orders ($NewOr$). Data on S&P 500 Index and its realized volatility (calculated aggregating at 5-minutes the squared intradaily returns) have been collected from the realised library of the Oxford-Man Institute. Data on the MVs come from the Federal Reserve (St.Louis) Economic Data archive. We use the returns as daily close–to–close log–differences of the S&P 500 Index, on a sample period between 2 January 2002 and 31 December 2019 (4500 daily observations). All the MVs have been considered in terms of month-to-month growth rate. The S&P 500 realized volatility, labelled as $RVol_{i,t}$, entering the "–X" part of the short-run component includes the overnight volatility. Some summary statistics (minimum, maximum, mean, standard deviation, skewness and kurtosis) for all variables considered are in Table 1. The last column of Table 1 reports the estimated coefficient of the predictive regression as recently proposed by Conrad and Schienle (2018), to which we refer for the details on the procedure. Overall, these variables appear to be good predictors of the long-run component.

**Table 1** Summary statistics

| | Obs. | Min. | Max. | Mean | SD | Skew. | Kurt. | $\pi_1$ |
|---|---|---|---|---|---|---|---|---|
| *Daily data* | | | | | | | | |
| $r_{i,t}$ | 4500 | -9.688 | 10.642 | 0.023 | 1.158 | -0.231 | 9.523 | |
| $RVol_{i,t}$ | 4500 | 0.128 | 8.958 | 0.835 | 0.605 | 3.392 | 20.363 | |
| *Monthly data* | | | | | | | | |
| $\Delta IP$ | 215 | -4.337 | 1.517 | 0.088 | 0.678 | -1.984 | 9.569 | $-0.125^{***}$ |
| $\Delta HS$ | 215 | -18.681 | 24.647 | 0.326 | 8.431 | 0.176 | -0.278 | $-0.01^{***}$ |
| $\Delta NewOr$ | 215 | -9.680 | 10.363 | 0.250 | 2.250 | -0.427 | 3.617 | $-0.042^{***}$ |

**Notes**: The table presents the summary statistics for the close-to-close S&P 500 log-returns ($r_{i,t}$), S&P 500 realized volatility with overnight returns ($RVol_{i,t}$), US Industrial Production, Housing Starts and New Orders growth rates ($\Delta IP$, $\Delta HS$ and $\Delta NewOr$, respectively). Sample period: 2 February 2002 - 31 December 2019. Percentage scale. The table reports the number of observations (Obs.), the minimum (Min.) and maximum (Max.), the mean, standard deviation (SD), Skewness (Skew.) and Kurtosis (Kurt.). The symbol $\Delta$ denotes the first difference. $\pi_1$ represents the estimated coefficient of the predictive regression as proposed by Conrad and Schienle (2018). $^{*}$, $^{**}$ and $^{***}$ represent the significance at levels $10\%, 5\%, 1\%$, respectively, associated to HAC robust standard errors, for the null of $\pi_1 = 0$.

In what follows, we use six different specifications for the two models at hand, the DAGM(–X) and GM(–X), labelled as $M_1, \cdots, M_6$. Even–numbered models include the $RVol_{i,t}$ in the "–X" part. Moreover, we also use the HAR model of Corsi (2009). The volatility proxy used is the 5-minute realized volatility. The results of the comparison of all the models, carried out through the Model Confidence Set (MCS, Hansen et al. (2011)), are reported in Table 2 .

**Table 2** MSE losses and MCS composition for the out-of-sample period

| | | MIDAS VAR. | –*X spec.* | 1-day | 5-days | 10-day | 1-month | 2-months | 3-months |
|---|---|---|---|---|---|---|---|---|---|
| DAGM(–X) | $M_1$ | $\Delta IP$ | | 0.122 | 0.114 | 0.147 | 0.197 | 0.237 | 0.253 |
| | $M_2$ | $\Delta IP$ | $RVol_{i-1,t}$ | 0.106 | 0.104 | 0.133 | 0.165 | 0.187 | 0.204 |
| | $M_3$ | $\Delta HS$ | | 0.114 | 0.107 | 0.14 | 0.189 | 0.229 | 0.246 |
| | $M_4$ | $\Delta HS$ | $RVol_{i-1,t}$ | 0.097 | 0.093 | 0.122 | 0.152 | 0.166 | 0.177 |
| | $M_5$ | $\Delta NewOr$ | | 0.142 | 0.128 | 0.163 | 0.221 | 0.276 | 0.316 |
| | $M_6$ | $\Delta NewOr$ | $RVol_{i-1,t}$ | 0.12 | 0.116 | 0.147 | 0.191 | 0.24 | 0.25 |
| GM(–X) | $M_1$ | $\Delta IP$ | | 0.125 | 0.114 | 0.148 | 0.2 | 0.252 | 0.278 |
| | $M_2$ | $\Delta IP$ | $RVol_{i-1,t}$ | 0.104 | 0.099 | 0.127 | 0.159 | 0.184 | 0.194 |
| | $M_3$ | $\Delta HS$ | | 0.117 | 0.108 | 0.141 | 0.188 | 0.224 | 0.242 |
| | $M_4$ | $\Delta HS$ | $RVol_{i-1,t}$ | 0.103 | 0.098 | 0.126 | 0.162 | 0.188 | 0.189 |
| | $M_5$ | $\Delta NewOr$ | | 0.135 | 0.122 | 0.156 | 0.212 | 0.268 | 0.304 |
| | $M_6$ | $\Delta NewOr$ | $RVol_{i-1,t}$ | 0.111 | 0.104 | 0.131 | 0.161 | 0.182 | 0.194 |
| | HAR | | | 0.097 | 0.113 | 0.127 | 0.163 | 0.207 | 0.199 |

**Notes**: The table reports the average MSE losses for the full out-of-sample period, starting on 2 January 2014 and ending on 31 December 2019 (1506 daily obs.), according to the 1-day to 3-months volatility forecasts. Shades of gray denote inclusion in the MCS at significance level $\alpha = 0.25$.

We find that the extended models with the realized volatility as the "–X" term always belong to the MCS and are often able to outperform the HAR model, independently of the forecasting horizon considered.

## References


Amendola, A., V. Candila, and G. M. Gallo (2019). On the asymmetric impact of macro–variables on volatility. *Economic Modelling 76*, 135–152.

Conrad, C. and O. Kleen (2019). Two are better than one: Volatility forecasting using multiplicative component GARCH models. *Jour. of Applied Econometrics*.

Conrad, C. and M. Schienle (2018). Testing for an omitted multiplicative long-term component in GARCH models. *Jour. of Business & Economic Statistics*, 1–14.

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

Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics 95*(3), 776–797.

Ghysels, E. and H. Qian (2019). Estimating MIDAS regressions via OLS with polynomial parameter profiling. *Econometrics and Statistics 9*, 1–16.

Han, H. and D. Kristensen (2014). Asymptotic theory for the QMLE in GARCH-X models with stationary and nonstationary covariates. *Journal of business & economic statistics 32*(3), 416–429.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The Model Confidence Set. *Econometrica 79*(2), 453–497.