

# Modelling Volatility Cycles: The MF2-GARCH Model\*

Christian Conrad<sup>†</sup> and Robert F. Engle<sup>‡</sup>

January 14, 2025

## Abstract

We propose a novel multiplicative factor multi-frequency GARCH (MF2-GARCH) model, which exploits the empirical fact that the daily standardized forecast errors of one-component GARCH models are predictable by a moving average of past standardized forecast errors. In contrast to other multiplicative component GARCH models, the MF2-GARCH features stationary returns, and long-term volatility forecasts are mean-reverting. When applied to the S&P 500, the new component model significantly outperforms the one-component GJR-GARCH, the GARCH-MIDAS-RV, and the log-HAR model in long-term out-of-sample forecasting. We illustrate the MF2-GARCH's scalability by applying the new model to more than 2,100 individual stocks in the Volatility Lab at NYU Stern.

*Keywords:* Volatility forecasting, long- and short-term volatility, mixed frequency data, volatility component models, long-term forecasting

---

\*We are grateful to the co-editor, Eric Ghysels, and two anonymous referees for their comments, which greatly improved our paper. We would like to thank Jörg Breitung, Christian Brownlees, Rob Capellini, Zeno Enders, Jean-David Fermanian, Christian Francq, Christian Gourieroux, Onno Kleen, Robinson Kruse-Becher, Enno Mammen, Anne Opschoor, Timo Teräsvirta, and Jean Michel Zakoïan as well as seminar and conference participants at CREST (January 2020), the ES World Congress (2020), the 13th Annual SoFiE Conference (2021), and the 11th ECB Conference on Forecasting Techniques (2021) for their feedback on earlier versions of the paper. Funding by the German Federal Ministry of Education and Research (BMBF) and the Baden-Württemberg Ministry of Science as part of Germany's Excellence Strategy (ExU 10.2.31) is gratefully acknowledged.

<sup>†</sup>Corresponding author. Alfred-Weber-Institute, Heidelberg University; HEiKA - Heidelberg Karlsruhe Strategic Partnership, Heidelberg University, Karlsruhe Institute of Technology; KOF Swiss Economic Institute, ZEW Mannheim. Email: christian.conrad@awi.uni-heidelberg.de.

<sup>‡</sup>Stern School of Business, New York University. Email: re21@stern.nyu.edu

# 1 Introduction

There is strong empirical evidence that the conditional variance of stock returns consists of several components. Early evidence for volatility components was provided in, for example, Ding and Granger (1996) and Engle and Lee (1999). More recent evidence can be found in Christoffersen et al. (2008), Kim and Nelson (2013), Dorion (2016), and Conrad and Kleen (2020), among others. While the GARCH models of Ding and Granger (1996) and Engle and Lee (1999) have additive volatility components, more recent GARCH-type models decompose the conditional variance into multiplicative short- and long-term components. For example, in the Spline-GARCH model of Engle and Rangel (2008) and the multiplicative time-varying GARCH (MTV-GARCH) of Amado and Teräsvirta (2013, 2017), the long-term volatility component is a deterministic function of calendar time. In contrast, in the GARCH-MIDAS of Engle et al. (2013), the long-term component depends either on a rolling window realized variance (henceforth GARCH-MIDAS-RV, see also Wang and Ghysels, 2015) or on low-frequency macroeconomic or financial variables (henceforth GARCH-MIDAS-X, see also Asgharian et al., 2013, and Conrad and Loch, 2015). Those multiplicative volatility models are based on the idea that returns follow a stationary GARCH process once divided by the long-term volatility component. However, there is no consensus yet on the most suitable approach for modeling the long-term component.

We propose a novel specification for the long-term volatility component in multiplicative GARCH models. The specification is motivated by a new empirical fact that we document for the volatility forecast errors of one-component GARCH models: While the daily standardized forecast errors are essentially unpredictable based on past *daily* standardized forecast errors, a *rolling window moving average* of the past daily standardized forecast errors does have predictive power. This is because one-component models tend to under- or over-predict volatility for *extended* time periods. While over-prediction typically happens during economic expansions, under-prediction materializes during economic recessions and other crises periods.

The new model combines a short-term GJR-GARCH (see Glosten et al., 1993) component with a long-term component specified as a multiplicative error model (MEM) for the past forecast errors of the GARCH component. That is, the long-term component exploits the predictability in the averaged standardized forecast errors of the short-term component. Intuitively, the long-term component scales the GARCH component's volatility forecast up/down if the short-term component's forecasts have under-/overestimated

volatility in the recent past, i.e., the model is learning from past forecast errors. Because the long-term component can either evolve at the same frequency as the short-term component or at a lower frequency, the new specification belongs to the class of mixed frequency data sampling (MIDAS) models pioneered by Ghysels et al. (2004). We refer to the proposed specification as **Multiplicative Factor Multi-Frequency GARCH**. As the ‘MF’ appears twice, we abbreviate the model as MF2-GARCH.

The properties of the MF2-GARCH model clearly distinguish it from previous specifications. First, while in other multiplicative models, there is typically no feedback from the short-term to the long-term component (e.g., in the Spline-GARCH), the MF2-GARCH explicitly specifies the long-term component as a function of the short-term component’s past forecast errors. Second, because the actual economic drivers of long-term volatility are unknown and may vary over time, it is challenging to correctly specify the long-term component in the GARCH-MIDAS-X in real time. Our specification avoids this problem and is based on a simple MEM equation for the long-term component. Interestingly, the MF2-GARCH can be rewritten as a GARCH-MIDAS-X with an explanatory variable “generated within the model.” Third, because the MF2-GARCH is dynamically complete, i.e., it fully specifies the dynamics of the conditional variance, it is straightforward to construct multi-step ahead volatility forecasts.

We obtain the following theoretical results for the MF2-GARCH: First, we derive the unconditional variance of the daily returns. While the unconditional variance is time-varying in the Spline-GARCH and infinite in the GARCH-MIDAS-RV of Wang and Ghysels (2015), returns are covariance stationary in the MF2-GARCH. Due to the feedback between the short- and long-term components, the unconditional variance depends not only on the model parameters but also on the fourth moment of the innovation. Second, we obtain the news impact curve (NIC). The NIC illustrates that the responsiveness to news changes with the level of volatility which is another essential feature that distinguishes the MF2-GARCH from other models in the GARCH family. Specifically, conditional volatility is more responsive to news during low volatility periods than during high volatility periods. Third, we derive expressions for multi-step ahead forecasts of conditional volatility. Our results show that the forecasts are much more flexible than forecasts from the nested GARCH model. The forecasts reflect the current stance of the conditional variance and the prevailing volatility regime. In the short term, the conditional volatility forecast will approach the forecast of the long-term component before it converges to the unconditional volatility in the long run. Forecasts from the MF2-GARCH also differ from forecasts of

standard Spline-GARCH or GARCH-MIDAS models. The forecasts of the latter models are typically assumed to converge to the current level of the long-term component. Thus, the forecasts from these models do not feature mean reversion in the long run. Forth, we discuss the quasi-maximum likelihood estimation of the MF2-GARCH and provide a Monte Carlo simulation showing that standard asymptotic results lead to valid inference. Finally, we provide an analysis of the degree of misspecification of the nested one-component GJR-GARCH and the GARCH-MIDAS-RV when the true data-generating process is an MF2-GARCH. As discussed in Patton (2020), in the presence of estimation error and depending on the employed loss function, a parsimoniously misspecified model might dominate the true but more complex model in terms of forecast performance. We show by simulations that for reasonable parameter values of the MF2-GARCH, the degree of misspecification of the GJR-GARCH and the GARCH-MIDAS-RV is so severe that the MF2-GARCH outperforms both models when evaluated by the squared error and the QLIKE loss.

Our empirical results strongly support the MF2-GARCH. First, we estimate the MF2-GARCH for the S&P 500 and 2,142 US and international equities in the Volatility Laboratory (V-Lab) at NYU Stern.<sup>1</sup> Our in-sample results show that the MF2-GARCH is clearly preferred to the nested one-component GJR-GARCH, the GARCH-MIDAS-RV and the Spline-GARCH. For the S&P 500, we show that the MF2-GARCH's estimated long-term component is closely related to news about the macroeconomy and monetary policy, particularly news about inflation and interest rates. Thus, we provide further evidence for the close link between economic conditions and long-term volatility (see, for example, Engle et al., 2013, and Conrad and Loch, 2015). We also illustrate that the MF2-GARCH's volatility forecasts, which feature cyclical behavior, are much more flexible than the forecasts of the competitor models. In contrast to the (overly) smooth long-term component of the Spline-GARCH, the MF2-GARCH's long-term component adjusts in response to short-lived periods of market turmoil. Nevertheless, compared to the one-component GJR-GARCH, the MF2-GARCH's forecasts avoid overestimating volatility after a short-lived surge in volatility due to the low persistence in its short-term component.

While most of the literature on volatility forecasting focuses on short-term (e.g., one-day ahead) prediction horizons, Christoffersen and Diebold (2000), Engle (2009b), and Ghysels et al. (2019), among others, have highlighted that in many areas of finance long-term risk forecasts are the relevant inputs. This leads to the question of how far ahead into the future

---

<sup>1</sup>In V-Lab, the MF2-GARCH is now estimated for more than 18,000 assets from different asset classes on a weekly basis. See: <https://vlab.stern.nyu.edu/docs/volatility/MF2-GARCH>

we can forecast volatility. Hence, in evaluating the out-of-sample forecast performance of the MF2-GARCH, our focus is on medium- and long-term forecast horizons of up to eight months. We test whether the MF2-GARCH, which is designed to capture volatility cycles, leads to better long-term predictions than the nested GJR-GARCH, the GARCH-MIDAS-RV, the Spline-GARCH, and Corsi and Reno's (2012) log-HAR with leverage. For the S&P 500, it turns out that the MF2-GARCH outperforms all competitor models when the forecast horizon is beyond two months. The MF2-GARCH's forecast performance is particularly strong during periods of high volatility, where it dominates the competitor models at all forecast horizons. For a cross-section of twenty equities, out-of-sample results from the V-Lab confirm that the MF2-GARCH strongly outperforms the competitor models.

The paper is organized as follows. In Section 2, we show that the volatility forecast errors of the GJR-GARCH are predictable. We introduce the MF2-GARCH and discuss its properties in Section 3. The empirical results are presented in Section 4 and Section 5 concludes. Further model details, proofs, and additional tables and figures can be found in the Supporting Information Appendix.

## 2 A New Empirical Fact of Volatility Forecast Errors

In this section, we provide evidence for a new empirical fact of volatility forecast errors: Rolling window moving averages of the standardized forecast errors of one-component GARCH models behave counter-cyclically and have predictive power for future standardized forecast errors.

We denote the log-return on day  $t$  by  $r_t$ . The conditional heteroskedasticity in daily stock returns is commonly modeled as a GARCH process. Daily stock returns are written as

$$r_t = \sqrt{\tilde{h}_t} \zeta_t, \quad (1)$$

where  $\tilde{h}_t$  denotes the conditional variance and the  $\zeta_t$  are assumed to be *i.i.d.* with mean and variance equal to zero and one, respectively. For illustration, we estimate a GJR-GARCH(1,1) specification for a long time series of daily S&P 500 log-returns covering January 1971 to June 2023.<sup>2</sup> We obtain the following result:

$$\tilde{h}_t = 0.018 + (0.022 + 0.116 \mathbf{1}_{\{r_{t-1} < 0\}}) r_{t-1}^2 + 0.905 \tilde{h}_{t-1}, \quad (2)$$

---

<sup>2</sup>The data will be introduced and discussed in more detail in Section 4.1. See also Panel A of Table 1.

where the numbers in parentheses are Bollerslev-Wooldridge robust standard errors and  $\mathbf{1}_{\{r_{t-1} < 0\}}$  equals one if  $r_{t-1} < 0$ , and zero else. As expected, the conditional variance is highly persistent and there is strong evidence for asymmetry.

Several test statistics have been proposed to check a GARCH specification's adequacy. For example, Engle and Ng (1993) and Halunga and Orme (2009) propose Lagrange Multiplier (LM) tests for the null hypothesis that a (GJR-)GARCH(1,1) is correctly specified.

An alternative approach is to check whether equation (1) is misspecified in the sense that  $\zeta_t = \sqrt{\tau_t} Z_t$ , where the  $Z_t$  are *i.i.d.* and  $\tau_t$  represents an omitted multiplicative long-term volatility component. The long-term component evolves either at the same frequency as the daily returns or at a lower (e.g., monthly or quarterly) frequency. The daily returns,  $r_t$ , can be either stationary or non-stationary. For example, in the Spline-GARCH model  $\tau_t$  evolves at the daily frequency and – because the long-term component is a deterministic function of time – the daily returns have a time-varying unconditional second moment. In either case, the scaled returns,  $r_t/\sqrt{\tau_t}$ , are assumed to follow a stationary GARCH process. LM tests for an omitted  $\tau_t$  component have been proposed in Lundbergh and Teräsvirta (2002) and Amado and Teräsvirta (2017) for daily long-term components. The LM test of Conrad and Schienle (2020) allows for explanatory variables in the long-term component and either a daily or lower frequency  $\tau_t$ .

The tests of Lundbergh and Teräsvirta (2002) and Conrad and Schienle (2020) exploit that the squared standardized errors,  $\zeta_t^2 = r_t^2/\tilde{h}_t$ , are *i.i.d.* under the null hypothesis of a constant long-term component. The Conrad and Schienle (2020) LM test checks whether  $\zeta_t^2$  is predictable by  $x_{t-1}, x_{t-2}, \dots, x_{t-K}$ , where  $x_t$  is a predictor variable that can be exogenous or “generated within the model.”<sup>3</sup> Under the null hypothesis, the LM test is  $\chi^2$  distributed with  $K$  degrees of freedom.

We propose to use the  $m$ -days rolling window average of the squared standardized errors,

$$\tilde{V}_{t-1}^{(m)} = \frac{1}{m} \sum_{j=1}^m \frac{r_{t-j}^2}{\tilde{h}_{t-j}}, \quad (3)$$

as the predictor variable. Under the null hypothesis,  $r_t^2$  is a conditionally unbiased proxy of the true but unobservable conditional variance and  $\tilde{h}_t$  is a one-step-ahead forecast for the same quantity. If the GARCH model is correctly specified, the standardized volatility forecast errors,  $r_t^2/\tilde{h}_t$ , have an expected value of one and a variance of two if  $\zeta_t$  is Gaussian.

---

<sup>3</sup>For details, see the discussion below Assumption 6 in Conrad and Schienle (2020), as well as Remark 7 in the supplementary Appendix to their paper.

Hence, we think of  $\tilde{V}_t^{(m)}$  as a measure of the *local bias* of the GARCH conditional variance. For  $m = 1$ , we obtain  $\tilde{V}_{t-1}^{(1)} = \zeta_{t-1}^2$ , which is the predictor variable in the ‘ARCH nested in GARCH’ test of Lundbergh and Teräsvirta (2002).

Figure 1 shows  $\tilde{V}_t^{(m)}$  based on the conditional variances from the GJR-GARCH in equation (2) for  $m \in \{1, 15, 25, 45\}$ . In the upper right and both lower panels, it is visible that, as expected,  $\tilde{V}_t^{(m)}$  fluctuates around the value of one. However, as the two lower panels show, there are extended periods during which the one-component GARCH model under- or overestimates volatility. That is, for  $m = 25$  and  $m = 45$  the evolution of  $\tilde{V}_{t-1}^{(m)}$  is in line with a local bias of the GJR-GARCH conditional variance. The local bias appears to be counter-cyclical: The one-component GARCH model tends to overestimate volatility during expansions and to underestimate it during recessions. For  $m = 1$ ,  $\tilde{V}_t^{(m)}$  is too noisy to reveal this bias. There are also some spikes in  $\tilde{V}_t^{(m)}$ . These spikes occur due to extraordinary events with unexpectedly high volatility. For example, the two largest spikes are due to the stock market crashes on October 19, 1987 (“Black Monday”) and October 13, 1989 (“Mini-Crash”). The spike in March 2020 is due to the emergence of the Covid-19 pandemic and the spike on September 14, 2022 due to the release of higher-than-expected inflation numbers.

In Panel B of Table 1, we formally test whether  $\tilde{V}_{t-1}^{(m)}$  has predictive power for  $\zeta_t^2$  using the Conrad and Schienle (2020) LM test with  $K = 1$ . For  $m \in \{1, 5, 15\}$ , the null hypothesis of a constant long-term component is not rejected. When the true long-term component smoothly varies over time, this is to be expected because for small  $m$ ,  $\tilde{V}_{t-1}^{(m)}$  is too noisy to have explanatory power for  $\zeta_t^2$ . In contrast, for  $m \in \{25, 35, 45, 55\}$ , we strongly reject the null hypothesis. Thus, the LM test provides evidence for an omitted long-term component and suggests that  $\tilde{V}_{t-1}^{(m)}$  is suitable for modeling the dynamics of the long-term component when  $m$  is appropriately chosen.

Importantly, the behavior of the standardized volatility forecast errors is not specific to the one-component GJR-GARCH. We also estimated EGARCH (Nelson, 1991), FIGARCH (Baillie et al., 1996), and Realized GARCH (Hansen et al., 2012) models and obtained very similar results. For example, the correlation between the  $\tilde{V}_{t-1}^{(45)}$  of the GJR-GARCH and the  $\tilde{V}_{t-1}^{(45)}$  of the EGARCH, FIGARCH, and the Realized GARCH is 0.92, 0.82, and 0.77, respectively. Figure A.1 in the Supporting Information Appendix plots  $\tilde{V}_{t-1}^{(45)}$  for all four models and confirms that there is strong co-movement. Furthermore, our findings do not only hold for the S&P 500 but also for other international stock indices. For illustration,

Figure A.2 replicates Figure 1 for the FTSE 100.<sup>4</sup>

In summary, the evidence suggests that one-component GARCH models are misspecified and that the misspecification is detectable when using suitable moving averages of past standardized forecast errors to predict the current standardized forecast error.

### 3 The MF2-GARCH Model

This section introduces the MF2-GARCH model. In the main specification, the short- and the long-term components evolve at a daily frequency. For this specification, we derive the unconditional variance of returns, the NIC and multi-step ahead forecasts. In Section 3.2, we suggest several directions in which the MF2-GARCH can be extended. In particular, we introduce a parametrization that allows for multiple frequencies, i.e., the short-term component evolves at the daily frequency, while the long-term component varies at a lower frequency. Further details on the MF2-GARCH are provided in the Supporting Information Appendix: Section A.1 discusses quasi-maximum likelihood estimation and provides a Monte Carlo simulation. The degree of misspecification of the nested one-component GJR-GARCH and the GARCH-MIDAS-RV when the true model is an MF2-GARCH is analyzed in Section A.2. A comparison of the MF2-GARCH with other component models is provided in Section A.3.

In general and as motivated in Section 2, daily log-returns are defined as  $r_t = \sigma_t Z_t = \sqrt{h_t \tau_t} Z_t$ . We denote the information set on day  $t$  by  $\mathcal{F}_t$ .  $\sigma_t^2$  denotes the conditional variance and the short- and long-term volatility components are given by  $h_t$  and  $\tau_t$ . We make the following assumption about the innovations  $Z_t$ .

**Assumption 1.** *Let  $Z_t$  be i.i.d. The density of  $Z_t$  is symmetric with  $\mathbf{E}[Z_t] = 0$  and  $\mathbf{E}[Z_t^2] = 1$ . Further,  $Z_t^2$  has a nondegenerate distribution and  $\kappa = \mathbf{E}[Z_t^4] < \infty$ .*

The assumption that the density of  $Z_t$  is symmetric is commonly made for GJR-GARCH models because it allows for a straightforward computation of the multi-step ahead conditional variance forecast (see, for example, Zivot, 2009). Also, Ling and McAleer (2002) make this assumption when deriving conditions for the stationarity and the existence of the fourth moment of the GJR-GARCH. Importantly, as shown in Alexander et al. (2021),

---

<sup>4</sup>We obtained similar results for the DAX and the Hang Seng Index (HSI). In addition, we found evidence for considerable co-movement in standardized volatility forecast errors internationally (see also Engle and Campos-Martins, 2023).

the symmetry of the density of  $Z_t$  does not preclude that the  $s$ -step ahead aggregated returns exhibit skewness. The assumption that  $\kappa = \mathbf{E}[Z_t^4] < \infty$  is a necessary condition for ensuring the finiteness of the unconditional variance of the returns and for the existence of the variance of the score of the likelihood function.

### 3.1 Daily short- and long-term components

We specify the short-term volatility component as a unit variance GJR-GARCH(1,1)

$$h_t = (1 - \phi) + (\alpha + \gamma \mathbf{1}_{\{r_{t-1} < 0\}}) \frac{r_{t-1}^2}{\tau_{t-1}} + \beta h_{t-1}, \quad (4)$$

where  $\phi = \alpha + \gamma/2 + \beta$ . Note that the driving variable in equation (4) is  $r_{t-1}^2/\tau_{t-1}$ . This distinguishes  $h_t$  from the daily conditional variance,  $\tilde{h}_{i,t}$ , in equation (2). We make the following assumption about the parameters of the short-term component:

**Assumption 2.** *The parameters of the short-term GJR-GARCH component satisfy the conditions  $\alpha > 0$ ,  $\alpha + \gamma > 0$ ,  $\beta > 0$  and  $\phi = \alpha + \gamma/2 + \beta < 1$ .*

If Assumptions 1 and 2 hold, then  $r_t/\sqrt{\tau_t} = \sqrt{h_t} Z_t$  follows a covariance stationary GJR-GARCH(1,1) with  $\mathbf{E}[h_t Z_t^2] = \mathbf{E}[h_t] = 1$ . If the GJR-GARCH(1,1) fully captures the conditional heteroskedasticity, then  $\tau_t$  is equal to a constant and the multiplicative model reduces to a one-component GJR-GARCH for the daily returns.

Following Engle (2009a), we refer to  $r_t/\sqrt{h_t}$  as “deGARCHed returns” and define  $V_t = r_t^2/h_t = \tau_t Z_t^2$  as the squared deGARCHed returns. If the GARCH component fully captures the conditional heteroskedasticity, then by Assumption 1 the  $V_t$  are *i.i.d.* However, if  $\tau_t$  is time-varying and persistent, then  $V_t$  is autocorrelated. Because  $V_t$  is a non-negative variable, we specify the long-term component as a MEM equation for the conditional expectation of  $V_t$ :

$$\tau_t = \lambda_0 + \lambda_1 V_{t-1}^{(m)} + \lambda_2 \tau_{t-1}, \quad (5)$$

where

$$V_{t-1}^{(m)} = \frac{1}{m} \sum_{j=1}^m V_{t-j} = \frac{1}{m} \sum_{j=1}^m \frac{r_{t-j}^2}{h_{t-j}}. \quad (6)$$

Recall that we can think of the squared deGARCHed returns as standardized volatility forecast errors. Hence,  $V_{t-1}^{(m)}$  is a *rolling window* measure of the local bias of the short-term component’s conditional variance over the previous  $m$  days. Note that equation (5) can be written as a MEM(1,  $m$ ) with the restriction that the  $m$  ‘ARCH’ coefficients are given

by  $\lambda_1/m$ . The MEM(1,  $m$ ) is covariance stationary if the sum of the ARCH and GARCH coefficients is less than one. By construction, this is satisfied if  $\lambda_1 + \lambda_2 < 1$ .

**Assumption 3.** *The parameters of the long-term component satisfy the conditions  $\lambda_0 > 0$ ,  $\lambda_1 > 0$ ,  $\lambda_2 > 0$  and  $\lambda_1 + \lambda_2 < 1$ .*

Under Assumptions 1 and 3, it holds that  $V_t = \tau_t Z_t^2$  is a covariance stationary MEM with  $\mathbf{E}[V_t | \mathcal{F}_{t-1}] = \tau_t$  and  $\mathbf{E}[V_t] = \lambda_0/(1 - \lambda_1 - \lambda_2)$ . We refer to the parametrization given by equations (4) and (5) as MF2-GARCH-rw- $m$ , where ‘rw- $m$ ’ stands for *rolling window* of length  $m$ .

### 3.1.1 Unconditional variance of daily returns

Next, we derive the unconditional variance of the daily returns. First, note that the unconditional mean and variance are given by  $\mathbf{E}[r_t] = 0$  and  $\mathbf{Var}[r_t] = \mathbf{E}[r_t^2] = \mathbf{E}[\tau_t h_t]$ . The following theorem provides an expression for  $\mathbf{Var}[r_t]$ .

**Theorem 1.** *Let Assumptions 1-3 be satisfied. If the MF2-GARCH-rw- $m$  process,  $(r_t)_{t \in \mathbb{Z}}$ , is covariance stationary, then*

$$\Gamma_m = \left( \lambda_1 \frac{1}{m} \phi_\kappa + \lambda_2 \phi \right) + \lambda_1 \phi_\kappa \frac{1}{m} \sum_{j=2}^m \phi^{j-1} < 1 \quad (7)$$

with  $\phi_\kappa = (\alpha + \gamma/2)\kappa + \beta$ . The unconditional variance of the daily returns is given by

$$\mathbf{Var}[r_t] = (\lambda_0 + \frac{\lambda_0}{1 - \lambda_1 - \lambda_2} (1 - \phi)(\lambda_1 + \lambda_2) + \Delta_m) / (1 - \Gamma_m), \quad (8)$$

where

$$\Delta_m = (1 - \phi) \lambda_1 \phi \frac{\lambda_0}{1 - \lambda_1 - \lambda_2} \left( \frac{m-1}{m} + \frac{1}{m} \sum_{j=2}^m \sum_{k=1}^{j-2} \phi^k \right).$$

In the proof of Theorem 1, we show that  $\Gamma_m < 1$  is satisfied if the returns are covariance stationary. For example, for  $m = 1$ , the condition reduces to  $\Gamma_1 = \lambda_1[(\alpha + \gamma/2)\kappa + \beta] + \lambda_2[\alpha + \gamma/2 + \beta] < 1$ . This illustrates that even if the conditions which ensure that  $r_t/\sqrt{\tau_t}$  and  $V_t = r_t^2/h_t$  are individually covariance stationary (i.e., Assumptions 1-3) are satisfied, the condition in equation (7) may be violated if  $\kappa$  is sufficiently large.

In general, the variance of the daily returns will depend on all the model parameters and the innovation’s fourth moment,  $\kappa$ . The fact that the unconditional variance depends on  $\kappa$  distinguishes the MF2-GARCH-rw- $m$  from standard GARCH models and is due to

the correlation between  $\tau_t$  and  $h_t$ . The unconditional variance of the returns increases in  $\kappa$  and decreases as  $m$  increases.<sup>5</sup> Theorem 1 reveals that the MF2-GARCH is fundamentally different from the Spline-GARCH and the GARCH-MIDAS-RV. In the Spline-GARCH, the unconditional variance of the returns is time-varying, and  $\text{Var}[r_t] = \infty$  in the GARCH-MIDAS-RV (see Wang and Ghysels, 2015).

When imposing the restrictions  $m = 1$  and  $\gamma = 0$ , we obtain a model that can be considered a ‘multiplicative version’ of the additive component model of Engle and Lee (1999). For  $m = 1$ , both components are symmetric and we impose the restriction  $\alpha + \beta < \lambda_1 + \lambda_2 < 1$  to ensure that  $\tau_t$  is the long-run component. In the following corollary, we derive the condition for the covariance stationarity of the daily returns when  $m = 1$  and  $\gamma = 0$  and state their unconditional variance.

**Corollary 1.** *Let Assumptions 1-3 be satisfied,  $m = 1$ , and  $\gamma = 0$ . The necessary and sufficient condition for the covariance stationary of the MF2-GARCH-rw-1 process is  $\Gamma_1 = \lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta) < 1$ . The unconditional variance of the returns is given by*

$$\text{Var}[r_t] = \frac{\lambda_0 + \frac{\lambda_0}{1-\lambda_1-\lambda_2}(1-\alpha-\beta)(\lambda_1+\lambda_2)}{1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]}. \quad (9)$$

If the short-term component is constant (i.e.  $\alpha = \beta = 0$ ), the expression in equation (9) reduces to  $\text{Var}[r_t] = \mathbf{E}[\tau_t] = \lambda_0/(1 - \lambda_1 - \lambda_2)$ . If the long-term component is constant (i.e.  $\lambda_1 = \lambda_2 = 0$ ), the expression in equation (9) reduces to  $\text{Var}[r_t] = \mathbf{E}[\lambda_0 h_t] = \lambda_0$ . In the latter case, the MF2-GARCH-rw-1 reduces to a GARCH(1,1).

As expected, empirically we find that  $V_{t-1}^{(1)} = r_{t-1}^2/h_{t-1}$  is too noisy to serve as a proxy of the local bias. In Section 4.2.1, we show that the optimal choice of  $m$  is around 63 for the S&P 500, which corresponds to a quarterly moving average.

### 3.1.2 News impact curve

Following Engle and Ng (1993), we use the NIC to illustrate how the conditional volatility is updated in response to new information. For a GJR-GARCH(1,1) with conditional variance  $\tilde{h}_t = \alpha_0 + (\alpha + \gamma \mathbf{1}_{\{r_{t-1} < 0\}})r_{t-1}^2 + \beta \tilde{h}_{t-1}$ , the NIC is defined as

$$NIC_{t+1}^{GJR} = \tilde{h}_{t+1}(r_t | \tilde{h}_t) = A_t^{GJR} + (\alpha + \gamma \mathbf{1}_{\{r_t < 0\}})r_t^2,$$

---

<sup>5</sup>For a graphical illustration, Figure A.3 in the Supporting Information Appendix plots the annualized unconditional variance as a function of  $m$  and for  $\kappa \in \{3, 5, 7\}$ . The model parameters are chosen as  $\alpha = 0.02$ ,  $\gamma = 0.10$ ,  $\beta = 0.8$ ,  $\lambda_0 = 0.01$ ,  $\lambda_1 = 0.05$ , and  $\lambda_2 = 0.94$ .

where  $A_t^{GJR} = \alpha_0 + \beta \tilde{h}_t$ . That is, the NIC is a function of today's return and  $A_t^{GJR}$  is known conditional on  $\mathcal{F}_{t-1}$ . If  $\gamma > 0$ , negative news,  $r_t < 0$ , have a stronger effect on volatility than positive news. However, the size of the effect does not depend on the current level of volatility.

The NIC of the MF2-GARCH-rw- $m$  consists of three terms:

$$NIC_{t+1}^{MF} = \sigma_{t+1}^2(r_t | \tau_t, h_t) = A_t^{MF} + B_t^{MF} + \left( \lambda_1 \beta \frac{1}{m} + \lambda_2 (\alpha + \gamma \mathbf{1}_{\{r_t < 0\}}) \right) r_t^2, \quad (10)$$

where

$$\begin{aligned} A_t^{MF} &= \lambda_0(1 - \phi) + \lambda_0 \beta h_t + \lambda_1(1 - \phi) \frac{1}{m} \sum_{j=1}^{m-1} \frac{r_{t-j}^2}{h_{t-j}} \\ &\quad + \lambda_1 \beta h_t \frac{1}{m} \sum_{j=1}^{m-1} \frac{r_{t-j}^2}{h_{t-j}} + \lambda_2(1 - \phi) \tau_t + \lambda_2 \beta \sigma_t^2 \end{aligned}$$

and

$$\begin{aligned} B_t^{MF} &= \lambda_0(\alpha + \gamma \mathbf{1}_{\{r_t < 0\}}) \frac{r_t^2}{\tau_t} + \lambda_1(1 - \phi) \frac{1}{m} \frac{r_t^2}{h_t} \\ &\quad + \lambda_1(\alpha + \gamma \mathbf{1}_{\{r_t < 0\}}) \frac{1}{m} \frac{r_t^4}{\sigma_t^2} + \lambda_1(\alpha + \gamma \mathbf{1}_{\{r_t < 0\}}) \frac{r_t^2}{\tau_t} \frac{1}{m} \sum_{j=1}^{m-1} \frac{r_{t-j}^2}{h_{t-j}}. \end{aligned}$$

The intercept,  $A_t^{MF}$ , is known conditional on  $\mathcal{F}_{t-1}$ . The last term depends on  $r_t^2$  and model parameters and, hence, is similar to the term  $(\alpha + \gamma \mathbf{1}_{\{r_t < 0\}}) r_t^2$  in the GJR-GARCH. However, the term  $B_t^{MF}$  shows that the marginal effect of  $r_t^2$  also depends on the current conditional variance,  $\sigma_t^2$ , as well as  $h_t$  and  $\tau_t$  individually.

Figure 2 shows the (standardized) NIC of the MF2-GARCH-rw- $m$  with  $m = 63$  (left panel) and  $m = 21$  (right panel). The NIC is plotted as a function of the return,  $r_t$ , and for  $\tau_t = 1$  (blue),  $\tau_t = 1.5$  (red),  $\tau_t = 0.5$  (green). The parameters are chosen as in Figure A.3. The short-term component is fixed at its unconditional expectation (i.e.,  $h_t = 1$ ). Note that parameters of the long-term component are chosen such that  $\mathbf{E}[\tau_t] = 1$ . That is, the solid black NIC represents a situation in which both the short- and long-term component are at their unconditional expectation. Due to the asymmetry term in the short-term component, the impact of negative returns is stronger than the impact of positive returns. The other two lines show that the effect of new information becomes stronger/weaker when long-term volatility is below/above its expectation. That is, the conditional volatility is more sensitive to news during a low-volatility period than during a high-volatility period. This is reasonable because a large value of  $|r_t|$  is to be expected during turbulent times but less so during tranquil times. In line with the observation that the variance of the returns decreases as  $m$  increases, the news impact is slightly weaker for  $m = 63$  than for  $m = 21$ .

### 3.1.3 Forecasting volatility

Because the MF2-GARCH-rw- $m$  model is dynamically complete, we can analytically derive volatility forecasts for any desired horizon. In the following, we assume that a researcher has observed returns up to day  $t$ . Based on the information set  $\mathcal{F}_t$ , she intends to compute a volatility forecast for day  $t+s$ . First, recall that the  $s$ -step ahead forecast of the short-term component can be computed as  $\mathbf{E}[h_{t+s}|\mathcal{F}_t] = 1 + \phi^{s-1}(h_{t+1} - 1)$ ,  $s \geq 2$  (see, for example, Zivot, 2009). The forecasts for the long-term component are slightly more involved and are presented in Appendix A.4 of the Supporting Information Appendix. Using these results,  $\mathbf{E}[\sigma_{t+s}^2|\mathcal{F}_t]$  can be computed as follows:

**Theorem 2.** *Let Assumptions 1-3 and the constraint in equation (7) be satisfied. Then, in the MF2-GARCH-rw- $m$  the forecast of the conditional variance on day  $t+s$ ,  $s \geq 1$ , can be computed as follows: First,  $\mathbf{E}[\sigma_{t+1}^2|\mathcal{F}_t] = h_{t+1}\tau_{t+1}$ . Second, for  $s = 2, \dots, m$  the forecasts can be recursively calculated as*

$$\begin{aligned} \mathbf{E}[\sigma_{t+s}^2|\mathcal{F}_t] &= (1 - \phi)\mathbf{E}[\tau_{t+s}|\mathcal{F}_t] + \lambda_0\phi\mathbf{E}[h_{t+s-1}|\mathcal{F}_t] + \left(\lambda_1\frac{1}{m}\phi_\kappa + \lambda_2\phi\right)\mathbf{E}[\sigma_{t+s-1}^2|\mathcal{F}_t] \\ &\quad + \lambda_1\phi\mathbf{E}[h_{t+s-1}|\mathcal{F}_t]\frac{1}{m}\sum_{j=s}^m\frac{r_{t+s-j}^2}{h_{t+s-j}} + (1 - \phi)\lambda_1\phi\frac{1}{m}\sum_{j=2}^{s-1}\mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t]\left(1 + \sum_{k=1}^{j-2}\phi^k\right) \\ &\quad + \lambda_1\phi_\kappa\phi\frac{1}{m}\sum_{j=2}^{s-1}\phi^{j-2}\mathbf{E}[\sigma_{t+s-j}^2|\mathcal{F}_t]. \end{aligned} \tag{11}$$

Third, for  $s > m$  the following recursion applies:

$$\begin{aligned} \mathbf{E}[\sigma_{t+s}^2|\mathcal{F}_t] &= (1 - \phi)\mathbf{E}[\tau_{t+s}|\mathcal{F}_t] + \lambda_0\phi\mathbf{E}[h_{t+s-1}|\mathcal{F}_t] + \left(\lambda_1\frac{1}{m}\phi_\kappa + \lambda_2\phi\right)\mathbf{E}[\sigma_{t+s-1}^2|\mathcal{F}_t] \\ &\quad + (1 - \phi)\lambda_1\phi\frac{1}{m}\sum_{j=2}^m\mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t]\left(1 + \sum_{k=1}^{j-2}\phi^k\right) \\ &\quad + \lambda_1\phi_\kappa\phi\frac{1}{m}\sum_{j=2}^m\phi^{j-2}\mathbf{E}[\sigma_{t+s-j}^2|\mathcal{F}_t]. \end{aligned} \tag{12}$$

We will illustrate the behavior of the volatility forecasts in Section 4.2.2.

## 3.2 Extensions of the MF2-GARCH-rw

### 3.2.1 Modifications of the daily long-term component

**MF2-GARCH with Beta-weights:** In equation (6), we assume that  $V_{t-1}^{(m)}$  is based on the average of the last  $m$  standardized forecast errors. Instead of imposing equal weights,

we can take a weighted average of the form

$$V_{t-1}^{(m)} = \sum_{j=1}^m w_j(\omega) V_{t-j} = \sum_{j=1}^m w_j(\omega) \frac{r_{t-j}^2}{h_{t-j}}. \quad (13)$$

Following a common choice in the MIDAS literature (see Ghysels et al., 2006, and Ghysels et al., 2007), we parsimoniously model the weights  $w_j(\omega)$  according to a restricted Beta weighting scheme:  $w_j(\omega) = ((1-j/(m+1))^{\omega-1})/(\sum_{k=1}^m (1-k/(m+1))^{\omega-1})$ . By construction, the weights sum to one. In addition, we impose the constraint that the weights are non-increasing ( $\omega \geq 1$ ). For  $\omega > 1$  the weights decline from the first lag. For  $\omega = 1$  the weights are given by  $1/m$  and, hence, we obtain the model with rolling window  $V_t^{(m)}$ . We will refer to this parametrization as MF2-GARCH-bw- $m$ , where ‘bw- $m$ ’ stands for *Beta-weights* of length  $m$ .

**Realized volatility MEM:** When  $m$  is small, the average standardized forecast error of the short-term component,  $V_t^{(m)} = 1/m \sum_{j=1}^m r_{t-j}^2/h_{t-j}$ , can be a noisy proxy for the local bias. Instead, if we observe daily realized variances,  $RV_{i,t}$ , we can base  $\tau_t$  on the realized measure:  $V_t^{(m,RV)} = 1/m \sum_{j=1}^m RV_{t-j}/h_{t-j}$ . Again, we can apply the MEM specification from equation (5). However, without further assumptions, this specification is no longer dynamically complete. We leave this specification for future research.

### 3.2.2 Low-frequency long-term component

In the MF2-GARCH-rw- $m$ ,  $\tau_t$  varies at the daily frequency. Instead, we can specify the MF2-GARCH so that the long-term component varies at a lower frequency. For this specification, we introduce a notation that allows for mixed frequencies. We distinguish between a low-frequency period  $t$  and a high-frequency period  $i$ . The high-frequency period  $i$  represents days while  $t$  might represent a monthly, quarterly, or semi-annual frequency. We assume that there are  $n$  days within each period  $t$ , i.e.  $i = 1, \dots, n$ , and that we observe  $t = 1, \dots, T$  low-frequency periods. Using this notation, we denote the log-return on day  $i$  of period  $t$  by  $r_{i,t}$  (where we use the convention that  $r_{0,t} = r_{n,t-1}$ ). Similarly, we denote the information set on day  $i$  in period  $t$  by  $\mathcal{F}_{i,t}$  and define  $\mathcal{F}_t := \mathcal{F}_{n,t}$ . Note that the new notation reduces to the notation with a daily long-term component when  $n = 1$ . In the following, we assume that Assumption 1 holds for the innovations  $Z_{i,t}$ .

Using the notation for multiple frequencies, we write the short-term volatility component as

$$h_{i,t} = (1 - \phi) + (\alpha + \gamma \mathbf{1}_{\{r_{i-1,t} < 0\}}) \frac{r_{i-1,t}^2}{\tau_t} + \beta h_{i-1,t} \quad (14)$$

for  $i = 2, \dots, n$  and with  $h_{0,t} = h_{n,t-1}$ . Thus, for  $i = 1$ , we obtain  $h_{1,t} = (1 - \phi) + (\alpha + \gamma \mathbf{1}_{\{r_{n,t-1} < 0\}}) r_{n,t-1}^2 / \tau_{t-1} + \beta h_{n,t-1}$ . As before, we assume that Assumption 2 holds. We close the model by defining a MEM specification at the lower frequency. By averaging the squared deGARCHed returns within low-frequency period  $t$ , we obtain

$$V_t = \frac{1}{n} \sum_{i=1}^n \frac{r_{i,t}^2}{h_{i,t}} = \tau_t \frac{1}{n} \sum_{i=1}^n Z_{i,t}^2 = \tau_t \bar{Z}_t, \quad (15)$$

where  $\bar{Z}_t = n^{-1} \sum_{i=1}^n Z_{i,t}^2$  with  $\mathbf{E}[\bar{Z}_t] = 1$  and  $\mathbf{Var}[\bar{Z}_t] = (\kappa - 1)/n$ . Again, it follows from Assumption 1 that the  $\bar{Z}_t$  are *i.i.d.* This suggests the specification

$$\tau_t = \lambda_0 + \lambda_1 V_{t-1} + \lambda_2 \tau_{t-1}. \quad (16)$$

Note that the low-frequency component  $\tau_t$  still measures volatility in daily units. We will refer to this parametrization of the long-term component as MF2-GARCH-lf- $n$ , where ‘lf’ stands for *low-frequency* and  $n$  refers to the choice of the low-frequency period.<sup>6</sup> For example, when setting  $n = 21$  or  $n = 63$ , the long-term component varies at the monthly or quarterly frequency. If Assumptions 1 and 3 hold, then  $V_t = \tau_t \bar{Z}_t$  is a covariance stationary MEM(1,1) with  $\mathbf{E}[V_t | \mathcal{F}_{t-1}] = \tau_t$  and  $\mathbf{E}[V_t] = \lambda_0 / (1 - \lambda_1 - \lambda_2)$ . A major drawback of the low-frequency updating of the long-term component is that it introduces a discontinuity into the daily conditional variances. Thus, the daily returns are no longer covariance stationary.

## 4 Empirical Application

### 4.1 Stock market data

We use daily return data for the S&P 500 starting in January 1971 and ending in June 2023. Using the notation for  $n = 1$ , daily log returns are computed as  $r_t = 100(\log(P_t) - \log(P_{t-1}))$ , where  $P_t$  is the close price on day  $t$ . We employ realized variances based on intraday data provided by Tick Data to evaluate the forecast performance. Daily realized variances,  $RV_t$ , are defined as the sum of the squared five-minute intraday log-returns on day  $t$  plus the squared overnight log-return (see Bollerslev et al., 2018). We compute realized variances for the period January 2000 to June 2023. Table 1 shows summary statistics for the daily returns and realized variances. The annualized daily returns have a sample mean of 7.38 percent. The mean of the annualized daily realized volatility is 15.81

---

<sup>6</sup>We treat  $n$  as a fixed number that is not ‘too large.’ If  $n \rightarrow \infty$ , then  $\bar{Z}_t$  converges to one in probability and, hence, an identification issue arises due to the linear dependence of  $V_t$  and  $\tau_t$ .

percent during the period January 2010 to June 2023, which is the period that is used for the out-of-sample forecast evaluation. In addition, in Section 4.2.3 we use return data from the V-Lab for 2,142 US and international equities.

## 4.2 MF2-GARCH in action

### 4.2.1 Application to S&P 500

We first apply the MF2-GARCH model to the daily log returns of the S&P 500. We mainly focus on models with a daily long-term component and estimate the MF2-GARCH-rw- $m$  and the MF2-GARCH-bw- $m$  for the entire sample period from January 1971 to June 2023.

**Choice of  $m$ :** We first estimate both models for values of  $m$  up to 160 and determine the optimal  $m$  as the one that minimizes the BIC.<sup>7</sup> For both models, the upper left panel of Figure 3 shows the BIC as a function of  $m$ . The lowest value of the BIC materializes for  $m = 63$  for both models. For this value of  $m$ , the MF2-GARCH-rw- $m$  is clearly preferred relative to the MF2-GARCH-bw- $m$ . To investigate whether this pattern holds more generally, we reestimate both models for three subsamples: January 1971 - December 2009 (upper right panel), January 1980 - December 2009 (lower left panel), and January 1980 to Juni 2023 (lower right panel). In all three panels, the optimal choice of  $m$  is around 63 and the MF2-GARCH-rw- $m$  is the preferred model. Overall, the subsample analysis shows that the optimal choice of  $m$  is very stable for the S&P 500 and that equal weights are preferred to Beta-weights.

**Parameter estimates (sample period January 1971 to June 2023):** The first two rows of Table 2 (labeled as ‘ $\tau_t$  const.’) show the parameter estimates of the nested one-component GJR-GARCH. The parameter estimates of  $\alpha$ ,  $\gamma$  and  $\beta$  take typical values and indicate strong persistence in the GARCH component ( $\alpha + \gamma/2 + \beta = 0.982$ ). Next, the table displays the parameter estimates of the MF2-GARCH-rw- $m$  model. First, for the optimal window length, i.e.,  $m = 63$ , the estimates of the parameters in the long-term component,  $\lambda_1$  and  $\lambda_2$ , are both significant. As expected, the persistence in the long-term component (0.982) is much stronger than the persistence in the short-term component (0.924). Also, due to introducing a time-varying long-term component, the short-term GJR-component

---

<sup>7</sup>Since the choice of  $m$  does not affect the number of parameters, we could also determine the optimal value based on the likelihood function. We prefer the BIC because this allows for a meaningful comparison with the nested one-component GJR-GARCH.

is much less persistent than the one-component GJR-GARCH. The BIC clearly favors the two-component MF2-GARCH-rw-63 over the one-component model. To illustrate the consequences of choosing  $m$  too small and too large, we present parameter estimates for a monthly ( $m = 21$ ) and semi-annual ( $m = 126$ ) averaging of past forecast errors. For  $m = 21$ , the persistence in the long-term component increases to 0.995. Presumably, this is because increasing  $\lambda_2$  smoothes the long-term component and, thereby, counteracts the effect of decreasing  $m$ . For  $m = 126$ , the estimates of the long-term component's parameters are almost the same as for  $m = 63$ , but the standard errors increase considerably.

For the MF2-GARCH-bw- $m$  specification, we focus on the optimal choice of  $m = 63$ . For this value of  $m$ , the estimates of the short- and long-term component's parameters are close to those of the MF2-GARCH-rw-63 model. The estimate of  $\omega$  is equal to one. That is, the optimal Beta-weighting scheme for  $m = 63$  is flat. As suggested by the BIC comparison, this confirms that the MF2-GARCH-rw-63 is the preferred model. Thus, for the S&P 500, allowing for a flexible weighting scheme, only increases the standard errors of the remaining parameters. For the MF2-GARCH-lf- $n$  models with a low-frequency long-term component, we display parameter estimates for a quarterly ( $n = 63$ ) and a monthly ( $n = 21$ ) long-term component. When the long-term component varies at the quarterly frequency, the estimate of  $\lambda_2$  is insignificant. That is, when  $n$  is large, a MEM(0,1) is sufficient to capture the movements in the conditional expectation of  $V_t$ . In contrast, when  $\tau_t$  evolves at a monthly ( $n = 21$ ) frequency,  $\lambda_2$  becomes significant and a MEM(1,1) is preferred. In terms of the BIC, the MF2-GARCH-rw-63 model with a daily long-term component is clearly preferred to models with a low-frequency long-term component. Finally, we estimate the GARCH-MIDAS-RV model as in equation (31) of Section A.3 in the Supporting Information Appendix with quarterly ( $m = 63$ ) or monthly ( $m = 21$ ) rolling window realized variance and a Beta-weighting scheme with  $K = 252$ . In this model,  $\lambda_1$  measures the effect of  $RV_t^{rw}$  on long-term volatility. As expected, this effect is estimated to be positive. Again, the BIC favors the MF2-GARCH-rw-63 over the GARCH-MIDAS-RV models.

Figure 4 shows the annualized conditional volatility,  $\sqrt{h_{i,t}\tau_t}$ , and the long-term volatility,  $\sqrt{\tau_t}$ , as implied by the MF2-GARCH-rw-63. As expected, the behavior of  $\sqrt{\tau_t}$  mirrors that of the averaged standardized forecast errors from the one-component GJR-GARCH in Figure 1. That is, the figure confirms that there are prolonged cycles in long-term volatility.<sup>8</sup> Specifically, long-term volatility is high during the Great Recession and the Covid-19

---

<sup>8</sup>We checked whether there is still predictability in the volatility forecast errors of the MF2-GARCH-

pandemic but also after the crash on October 19, 1987. Using Baker et al.'s (2021) equity market volatility (EMV) trackers, we show in Appendix A.5 of the Supporting Information Appendix that the MF2-GARCH's long-term volatility is driven by various types of economic news (such as macroeconomic and monetary policy news), which a GARCH-MIDAS-X model with a single macroeconomic or financial explanatory variable is unlikely to capture fully.

#### 4.2.2 Implications for volatility forecasting

In this section, we illustrate the behavior of the volatility forecasts of the MF2-GARCH-rw- $m$ . In particular, we hint at the differences between the MF2-GARCH's forecasts and those of the nested GJR-GARCH. We rely on the parameter estimates for the MF2-GARCH-rw-63 and the GJR-GARCH for the S&P 500 in Table 2. The left panels of Figure 5 show the conditional volatility (solid black line) and long-component (solid gray line) of the MF2-GARCH for two subsamples. The horizontal line represents the unconditional volatility. We compute volatility forecasts from day  $t = 50$  onwards. In the upper/lower panel,  $t = 50$  corresponds to August 10, 2011, and September 14, 2016, respectively. The dashed black lines correspond to the forecasts from the MF2-GARCH. In the right panels, we plot the volatility forecasts from the GJR-GARCH as dashed black lines, whereby we impose the same unconditional variance as for the MF2-GARCH-rw-63 and the same conditional variance at the forecast origin.<sup>9</sup> We have chosen the forecast origin so that we can illustrate the differences between the forecasts from the two models. In the upper panel, the conditional volatility as well as the long-term component are below the unconditional volatility until there is a jump in volatility up to a level above 50%.<sup>10</sup> As expected, the right panel shows that the forecasts from the GJR-GARCH overestimate volatility after the surge in volatility. This can be explained by the high persistence in the one-component GARCH model. In contrast, the forecasts from the MF2-GARCH (left panel) are much more in line with the actual behavior of the conditional volatility. This is due to the lower persistence in the short-term component, which dominates the behavior of the MF2-GARCH forecast

---

rw-63. We found no evidence for autocorrelation in the daily squared standardized residuals,  $\hat{Z}_t^2$ , when averaged at lower frequencies. In addition, the empirical density of  $\hat{Z}_t$  is close to being symmetric.

<sup>9</sup>That is, the  $s$ -step ahead forecast is given by:  $\mathbf{Var}[r_t] + (\alpha^{GA} + \gamma^{GA}/2 + \beta^{GA})^{s-1}(h_{t+1}\tau_{t+1} - \mathbf{Var}[r_t])$ , where  $\alpha^{GA}$ ,  $\gamma^{GA}$  and  $\beta^{GA}$  are the parameters of the GJR-GARCH and  $\mathbf{Var}[r_t]$  is the unconditional variance of the MF2-GARCH.

<sup>10</sup>The increase in volatility was driven by the European sovereign debt crisis and a downgrade of the U.S.'s credit rating by Standard & Poor's.

in the short run. In the lower panel, volatility stayed below the unconditional volatility for an extended period. Thus, the long-term volatility component is below the unconditional volatility at the forecast origin. However, shortly before the forecast origin, the conditional volatility spikes and is still slightly above the level of the unconditional volatility at the forecast origin.<sup>11</sup> The forecasts from the GJR-GARCH directly converge to the unconditional volatility. The MF2-GARCH forecast behaves very differently. In the medium run, the forecast converges towards the forecast of the long-term component (dashed gray line). That is, the forecast decreases *below* the unconditional volatility. Only in the very long run, the MF2-GARCH forecast will converge towards the unconditional volatility. This illustrates that the MF2-GARCH forecast captures the empirical observation that there are persistent cyclical movements of the conditional volatility around the unconditional volatility. In summary, both examples illustrate that forecasts from the MF2-GARCH are more realistic than forecasts from the one-component GJR-GARCH. The MF2-GARCH forecasts reflect both the current level of volatility,  $\sqrt{h_{t+1}\tau_{t+1}}$ , as well as the prevailing volatility trend,  $\sqrt{\tau_{t+1}}$ .

#### 4.2.3 Evidence from the V-Lab

We complement the previous findings with evidence from the V-Lab. Overall, we consider 2,142 US and international stocks. The estimation period for all stocks ends on April 5, 2024. The beginning of the estimation period is stock-specific and depends on data availability. For each stock, we select the optimal MF2-GARCH-rw- $m$  specification by choosing the value of  $m$  that minimizes the BIC. Table 3 presents summary statistics for comparing the MF2-GARCH parameter estimates with those for the nested GJR-GARCH.<sup>12</sup> In addition, we report results for the Spline-GARCH and the Zero Slope Spline-GARCH (ZS-Spline-GARCH). For those models, we only report the estimates of the parameters in the short-term component. For all models, Table 3 shows the median parameter estimates of  $\alpha$ ,  $\gamma$  and  $\beta$ . In addition, the table displays the median persistence in the short- ( $\phi$ ) and long-term ( $\lambda_1 + \lambda_2$ ) component and the median value of  $m$ . The table shows that the estimate of  $\beta$  decreases substantially when estimating an MF2-GARCH instead of a GJR-GARCH. Conversely, the median estimates of  $\alpha$  and  $\gamma$  slightly increase. Overall, the persistence in the short-term component is strongly reduced when introducing the long-term component. Naturally, the persistence in the smooth long-term component of the MF2-GARCH is con-

---

<sup>11</sup>The spike in volatility was associated with fears that the Federal Reserve might raise interest rates.

<sup>12</sup>We only include stocks for which the MF2-GARCH parameter estimates satisfy Assumptions 2 and 3.

siderably higher than in the short-term component. Interestingly, for the Spline-GARCH and the ZS-Spline-GARCH, the reduction of the persistence in the short-term component is less pronounced. Presumably, this is because the long-term components of these models are too smooth. The smooth long-term components capture low-frequency variations (i.e., at the business cycle) in long-term volatility but not increases in volatility due to isolated events such as the crash on October 19, 1987. Across the 2,142 stocks, the median number of knots that are selected for both the Spline-GARCH and the ZS-Spline-GARCH is eight (this number is not reported in the table). Thus, the Spline-GARCH models require the estimation of considerably more parameters than the MF2-GARCH. The last column of Table 3 reports the percentage of stocks for which the BIC of the MF2-GARCH-rw is smaller than the BIC of the models in the respective rows. The MF2-GARCH is preferred over the GJR-GARCH for 95% of the stocks. In addition, it is preferred over the two Spline-GARCH models for over 60% of the stocks. In summary, our findings from the V-Lab confirm the results for the S&P 500 from Table 2 for a large number of stocks.

### 4.3 Out-of-sample forecasting

Finally, we evaluate the MF2-GARCH’s out-of-sample forecast performance. Most previous literature focuses on short forecast horizons, such as one-day ahead. However, as discussed in, for example, Christoffersen and Diebold (2000), Engle (2009b), Ederington and Guan (2010), positions are often held for more extended periods and, hence, long-term risk predictions should also be evaluated. Indeed, the failure of risk management during the financial crisis can be partly attributed to focusing on short-term risks, while neglecting long-term risks (Engle, 2009b). Thus, our particular focus is on longer-term volatility predictions. The MF2-GARCH is designed to capture slow movements in the long-run volatility, and we intend to test whether the model helps to improve longer-term volatility forecasts. In addition, our results will provide new evidence for answering the question of how far ahead we can forecast volatility (see also Ghysels et al., 2019).

#### 4.3.1 S&P 500

We focus on the MF2-GARCH specification with a daily long-term component and rolling window weighting scheme. As competitor models, we consider the nested GJR-GARCH, the GARCH-MIDAS-RV and the log-HAR model with leverage. While the MF2-GARCH, the GJR-GARCH, and the GARCH-MIDAS-RV model are based on daily returns, the log-HAR with leverage makes use of realized variances based on intraday data. Within the

respective model classes, the GJR-GARCH (see, for instance, Hansen and Lunde, 2005, and Brownlees et al., 2012) and the log-HAR (see, for example, Bollerslev et al., 2018, and Conrad and Kleen, 2020) are well known for their excellent short-term forecast performance.

The out-of-sample (OOS) period is January 2010 to June 2023. On each day of the OOS period, we compute forecasts of the variance on the next day, over the next week, as well as the  $s$ -months ahead *forward variance* with  $s \in \{1, \dots, 8\}$ . For computing forward variances, we assume that each month has 21 trading days. For example, the MF2-GARCH forecast of 2-months forward variance is  $\hat{\sigma}_{t+22,t+42|t}^2 = \sum_{j=1}^{21} \hat{\sigma}_{t+21+j|t}^2$ .

We reestimate all models on a rolling window of observations and update parameters on a monthly basis. For the MF2-GARCH, we select the optimal  $m$  based on the BIC and estimate  $\kappa$  based on the respective in-sample observations. The GARCH-MIDAS-RV models are based on a rolling window  $RV_t^{rw}$  with  $m \in \{21, 63\}$  and  $K = 252$ . When forecasting the  $s$ -months ahead *forward variance*, we specify the log-HAR model with leverage as

$$\begin{aligned} \log(RV_{t+(s-1)\cdot21+1,t+s\cdot21}) &= b_0 + b_1 \log RV_t + b_2 \log \left( \frac{RV_{t-4,t}}{5} \right) + b_3 \log \left( \frac{RV_{t-20,t}}{21} \right) \\ &\quad + b_4 r_t \mathbf{1}_{\{r_t < 0\}} + b_5 \frac{r_{t-4,t}}{5} \mathbf{1}_{\{r_{t-4,t} < 0\}} + b_6 \frac{r_{t-20,t}}{21} \mathbf{1}_{\{r_{t-20,t} < 0\}} + \zeta_t^{(s)}, \end{aligned} \tag{17}$$

where  $RV_{t+1,t+k} = \sum_{j=1}^k RV_{t+j}$ ,  $r_{t+1,t+k} = \sum_{j=1}^k r_{t+j}$  and  $\mathbf{1}_{\{r_{t+1,t+k} < 0\}}$  is an indicator function which takes the value one if  $r_{t+1,t+k} < 0$ , and zero otherwise. For daily and weekly forecast horizons, the dependent variable is  $\log(RV_{t+1})$  and  $\log(RV_{t+1,t+5})$ , respectively. The log-HAR is estimated by OLS and forecasts are based on the assumption that the  $\zeta_t^{(s)}$  are normally distributed.<sup>13</sup>

We evaluate the forecasts based on the SE and the QLIKE loss, which are both robust loss functions (see Patton, 2011). For the forecast evaluation, we proxy the true latent volatility by the corresponding realized variances. The realization of the  $s$ -months forward variance is  $RV_{t+(s-1)\cdot21+1,t+s\cdot21} = \sum_{j=1}^{21} RV_{t+(s-1)\cdot21+j}$ . To increase the readability of the forecast losses, we compare them to the losses of a simple benchmark. As benchmark, we consider the ‘historical volatility forecast’ which we define as the appropriately scaled

---

<sup>13</sup>We also considered log-HAR models with (the log of) quarterly and semi-annual averages of realized variances as additional explanatory variables. However, those specifications did not lead to an improved forecast performance relative to the baseline model. In addition, we considered the log-HAR without leverage and the pure HAR model. Again, both specifications did not lead to improvements in forecast performance.

mean of the daily realized variances during the ten years prior to the forecast origin. For each model and forecast horizon, we report the model performance as the ratio of the root mean square error (RMSE) of the respective model and the RMSE of historical volatility. Similarly, we present the respective QLIKE ratio. In addition, we compute a model confidence set (MCS) as proposed in Hansen et al. (2011). In Appendix A.6 of the Supporting Information Appendix, we describe in detail how the MCS is computed.

We first evaluate the forecasts during the January 2010 to December 2019 period, which ends before the Covid-19 pandemic. Panels A and B of Table 4 present results for the RMSE and the QLIKE, respectively. For each forecast horizon, the bold numbers indicate the model that achieves the lowest RMSE or QLIKE ratio. Blue shaded cells indicate which models are included in the MCS. In Panel A, the MF2-GARCH achieves the lowest RMSE for all forecast horizons. The entry of 0.685 at the one month horizon implies that the RMSE of the MF2-GARCH is 31.5% lower than the RMSE of the historical volatility forecast. Even at the longest forecast horizon of eight months, the RMSE is almost 20% lower. The MF2-GARCH is included in the MCS as the single model for all forecast horizons longer than one week. In Panel B, the log-HAR is the preferred model for forecast horizons up to two months.<sup>14</sup> Beyond that horizon, the MF2-GARCH produces the best forecasts. Overall, the MF2-GARCH is successful in forecasting volatility far into the future. We believe that the MF2-GARCH’s outperformance at long forecast horizons is due to the specification of the long-term component, which is designed to capture the underlying volatility cycle. In the medium term, the MF2-GARCH’s forecasts are strongly influenced by the dynamic structure of the long-term component but eventually converge to the unconditional variance. On the other hand, the GARCH-MIDAS-RV’s long-term component is a proxy of the level of volatility in the recent past. Imposing that the forecast converges to this level does not pay off.

The full OOS period ends in June 2023 and covers the Covid-19 pandemic. The beginning of the Covid-19 pandemic in March 2020 is characterized by days with extremely high realized variances. For example, on March 12 and 16, 2020, realized variances took values corresponding to an annualized daily volatility of 122% and 160%. This creates a short period of instability. In particular, the monthly realized forward variances are strongly shifted upwards once an extreme realized variance enters the measure and stay at elevated

---

<sup>14</sup>As discussed in Patton (2020), the ranking of models implied by the MSE and QLIKE can differ due to model misspecification or parameter estimation error. See also Section A.2 of the Supporting Information Appendix.

levels until the extreme realized variance drops out of the 21-days window. Although this period of extreme volatility is unpredictable for all models, it has a strong influence on the Diebold and Mariano (1995) tests underlying the MCS. As shown in Iacone et al. (2024), Diebold and Mariano (1995) tests become uninformative due to low power when short-lived periods of instability are included in the OOS period.<sup>15</sup> Following van Dijk and Franses (2003), we avoid this problem by considering a weighted average loss differential, where we attach zero weight to unpredictable volatility outliers. We define those outliers as a daily realized variance that is above the 99th percentile of the OOS daily realized variances' empirical distribution. Because almost all of these outliers materialize at the beginning of the Covid-19 pandemic, our approach is essentially the same as omitting this short period of instability from the evaluation sample.

Table 5 shows the results for the full OOS period. The results are similar to those in Table 4. Independently of the loss function, the MF2-GARCH's forecasts dominate for forecast horizons of three or more months. At shorter horizons, the log-HAR performs well, in particular according to the QLIKE loss. At the shortest forecast horizon of one day, the log-HAR clearly benefits from the information in the realized volatility. Overall, our previous results are confirmed. Figure A.5 in the Supporting Information Appendix visualizes the forecast performance of the MF2-GARCH, the log-HAR, and the benchmark model for the  $s$ -months forward variances. The figure shows the cumulated losses (left panels: SE, right panels: QLIKE) for forecast horizons of one month (upper panels), four months (middle panels) and eight months (lower panels). In line with the results from Table 5, the MF2-GARCH outperforms the log-HAR in all panels but the upper right one. Most importantly, the ranking of the two models is relatively stable across the entire OOS period.

To shed more light on the potentially time-varying relative forecast performance of the different models, we present a forecast evaluation that is conditional on being in a high volatility regime on the day a forecast is made. We classify a day as belonging to the high volatility regime if the realized variance of that day is above the 70th percentile of the OOS empirical distribution of realized variances. As Table A.4 in the Supporting Information Appendix shows, the evidence in favor of the MF2-GARCH becomes much stronger when focusing on forecasts that were made during high volatility regimes.

---

<sup>15</sup>In line with their result, we find that the MCS includes essentially all models when we do not control for this period of instability. In this setting, the assumption that the loss differences are stationary, which underlies the MCS procedure (see Hansen et al., 2011, Assumption 2), is likely to be violated.

Finally, we compare the relative forecast performance of the MF2-GARCH and the log-HAR by applying the conditional predictive ability test of Giacomini and White (2006). We use the realized variance,  $RV_t$ , that materializes on the day the forecasts are made as a predictor for the loss difference. That is, for each forecast horizon, we run a regression of the loss difference on a constant and the realized variance. The results (not reported) show that a high realized variance today predicts that the MF2-GARCH will outperform the log-HAR at essentially all forecast horizons. To illustrate this effect, Figure A.6 in the Supporting Information Appendix shows the estimated conditional mean of the regression and the corresponding 95% confidence interval as a function of  $RV_t$ .

#### 4.3.2 V-Lab

We also evaluated the forecast performance of the MF2-GARCH in the V-Lab. Here, we focus on the GJR-GARCH, the Spline-GARCH and the Zero-Slope Spline-GARCH as competitor models.<sup>16</sup> The forecast performance of the different models is evaluated for twenty (randomly selected) US stocks: Boeing (BA), Berkshire Hills Bancorp (BHLB), BlackRock Inc (BLK), Cardinal Health Inc (CAH), Edison International (EIX), Farmer Brothers Co (FARM), Guess (GES), HEICO Corp (HEI), Harley-Davidson Inc (HOG), Houston American Energy (HUSA), Marriott International Inc (MAR), Moody's Corp (MCO), MetLife Inc (MET), Microsoft Corp (MSFT), Novavax (NVAX), Oil-Dri Corp of America (ODC), Pfizer (PFE), Ralph Lauren (RL), Union Pacific (UNP), Whirlpool Corp (WHR). Because there is no intraday data available in V-Lab, squared daily returns are used to compute realized variances. The out-of-sample period is January 2010 to June 2023. For each forecast horizon and for each stock, based on the RMSE (respectively the QLIKE) the four models are ranked from rank 1 (best) to rank 4 (worst). For each forecast horizon, Table 6 presents the average rank across the twenty stocks. According to both loss functions, the MF2-GARCH performs best for all forecast horizons. Thus, Table 6 provides further evidence for the superior forecast performance of the MF2-GARCH within the class of GARCH models.

## 5 Conclusions

We suggest a multiplicative factor multi-frequency component GARCH model. The new model is motivated by the observation that a rolling window average of past standardized

---

<sup>16</sup>V-Lab does not produce forecasts for the HAR model or the GARCH-MIDAS.

forecast errors of one-component GARCH models has predictive power for current and future standardized forecast errors. This predictability is due to counter-cyclical movements in financial volatility, which simple GARCH models do not adequately capture. In contrast, the MF2-GARCH explicitly models these volatility cycles. We show that the MF2-GARCH is straightforward to estimate, features stationary returns and a NIC that is more responsive to news during low volatility periods than during high volatility periods. In addition, multi-step ahead volatility forecasts can be easily computed. Overall, the properties of the MF2-GARCH clearly distinguish it from other component models. Finally, our empirical results show that the new specification outperforms the nested GJR-GARCH, the Spline-GARCH, the GARCH-MIDAS-RV and the log-HAR in out-of-sample forecast performance.

In general, the MF2-GARCH model will benefit applications that require long-term forecasts of financial volatility, such as long-run value-at-risk predictions or measurement of systemic risk. For example, Conrad et al. (2024) show that the MF2-GARCH allows to compute ‘volatility news’ separately for the short-term and the long-term components. Assuming a positive relation between expected returns and the conditional variance of returns in a GARCH-in-Mean type model, unexpected returns can be decomposed into cash flow and discount rate news. In this framework, discount rate news is mainly driven by news to the MF2-GARCH’s long-term component. This is because only news to the long-term component is persistent enough to generate sizable variation in discount rates. From this, it follows that the MF2-GARCH’s long-term volatility component is a strong predictor for the strength of the instantaneous response of the stock market to surprises in macroeconomic announcements (see Conrad et al., 2024).

It will also be interesting to employ the long-term component in applications that require low-frequency estimates of volatility, for example, when analyzing the link between financial volatility and financial crises (see, for instance, Danielsson et al., 2018).

## References

- [1] Alexander, C., Lazar, E., Stanescu, S., 2021. Analytic moments for GJR-GARCH (1,1) processes. *International Journal of Forecasting*, 37, 105-124.
- [2] Amado, C., Teräsvirta, T., 2013. Modelling volatility by variance decomposition. *Journal of Econometrics*, 175, 142-153.

- [3] Amado, C., Teräsvirta, T., 2017. Specification and testing of multiplicative time-varying GARCH Models with applications. *Econometric Reviews*, 36, 421-446.
- [4] Asgharian, H., Hou, A. J., Javed, F., 2013. The importance of the macroeconomic variables in forecasting stock return variance: A GARCH-MIDAS approach. *Journal of Forecasting*, 32, 600-612.
- [5] Baillie, R.T., Bollerslev, T., Mikkelsen, H.O., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74, 3-30.
- [6] Baker, S.R., Bloom, N., Davis, S.J., Kost, K., 2021. Policy news and stock market volatility. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.3363862>
- [7] Bollerslev, T., Hood, B., Huss, J., Pedersen, L.H., 2018. Risk everywhere: Modeling and managing volatility. *Review of Financial Studies*, 31, 2730-2773.
- [8] Brownlees, C., Engle, R.F., Bryan, K., 2012. A practical guide to volatility forecasting through calm and storm. *Journal of Risk*, 14, 3-22.
- [9] Christoffersen, P.F., Diebold, F.X., 2000. How relevant is volatility forecasting for financial risk management? *The Review of Economics and Statistics*, 82, 12-22.
- [10] Christoffersen, P.F., Jacobs, K., Ornthanalai, C., Wang, Y. 2008. Option valuation with long-run and short-run volatility components. *Journal of Financial Economics*, 90, 272-297.
- [11] Conrad, C., Loch, K., 2015. Anticipating long-term stock market volatility. *Journal of Applied Econometrics*, 30, 1090-1114.
- [12] Conrad, C., Kleen, O., 2020. Two are better than one: Volatility forecasting using multiplicative component GARCH-MIDAS models. *Journal of Applied Econometrics*, 35, 19-45.
- [13] Conrad, C., Schienle, M., 2020. Testing for an omitted multiplicative long-term component in GARCH models. *Journal of Business & Economic Statistics*, 38, 229-242.
- [14] Conrad, C., Schoelkopf, J.T., Tushteva, N., 2024. Long-term volatility shapes the stock market's sensitivity to news. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.4632733>

- [15] Corsi, F., Reno, R., 2012. Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics*, 30, 368-380.
- [16] Danielsson, J., Valenzuela, M., Zer, I., 2018. Learning from history: Volatility and financial crises. *Review of Financial Studies*, 31, 2774-2805.
- [17] Diebold, F. X., Mariano, R.S., 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13, 253-263.
- [18] Ding, Z., Granger, C., 1996. Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics*, 73, 185-215.
- [19] Dorion, C., 2016. Option valuation with macro-finance variables. *Journal of Financial and Quantitative Analysis*, 51, 1359–1389.
- [20] Ederington, L.H., Guan, W., 2010. Longer-term time-series volatility forecasts. *Journal of Financial and Quantitative Analysis*, 45, 1055-1076.
- [21] Engle, R.F., 2009a. *Anticipating correlations: A new paradigm for risk management*. Princeton University Press.
- [22] Engle, R.F., 2009b. The risk that risk will change. *Journal of Investment Management*, 7, 1-5.
- [23] Engle, R.F., Ng, V.K., 1993. Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48, 1749-1778.
- [24] Engle, R., Lee, G., 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, 475-497.
- [25] Engle, R.F., Rangel, J.G., 2008. The Spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies*, 21, 1187-1222.
- [26] Engle, R.F., Ghysels, E., Sohn, B., 2013. Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics*, 95, 776-797.
- [27] Engle, R.F., Campos-Martins, S., 2023. What are the events that shake our world? Measuring and hedging global COVOL. *Journal of Financial Economics*, 147, 221-242.

- [28] Giacomini, R., White, H., 2006. Tests of conditional predictive ability. *Econometrica*, 74, 1545-1578.
- [29] Ghysels, E., Santa-Clara, P., Valkanov, R., 2004. The MIDAS touch: Mixed data sampling regression models. Working Paper, UNC and UCLA.
- [30] Ghysels, E., Santa-Clara, P., Valkanov, R., 2006. Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131, 59-95.
- [31] Ghysels, E., Sinko, A., Valkanov, R., 2007. MIDAS regressions: Further results and new directions. *Econometric Reviews*, 26, 53-90.
- [32] Ghysels, E., Plazzi, A., Valkanov, R., Serrano, A.R., Dossani, A., 2019. Direct versus iterated multi-period volatility forecasts. *Annual Review of Financial Economics*, 11, 173-195.
- [33] Glosten, L. R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of nominal excess return on stocks. *Journal of Finance*, 48, 1779-1801.
- [34] Halunga, A.G., Orme, C. D., 2009. First-order asymptotic theory for parametric misspecification tests of GARCH models. *Econometric Theory*, 25, 364-410.
- [35] Hansen, P.R., Lunde., A., 2005. A forecast comparison of volatility models: Does anything beat a GARCH(1,1). *Journal of Applied Econometrics*, 20, 873-889.
- [36] Hansen, P.R., Lunde, A., Nason, J. M., 2011. The model confidence set. *Econometrica*, 79, 453–497.
- [37] Hansen, P.R., Huang, Z., Shek, H.H., 2012. Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27, 877-906.
- [38] Iacone, F., Rossini, L., Viselli, A., 2024. Comparing predictive ability in presence of instability over a very short time. Papers 2405.11954, arXiv.org. Available at: <https://arxiv.org/abs/2405.11954>
- [39] Kim, Y., Nelson, C.R., 2013. Pricing stock market volatility: Does it matter whether the volatility is related to the business cycle? *Journal of Financial Econometrics*, 12, 307-328.

- [40] Ling, S., McAleer, M., 2002. Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics*, 106, 109-117.
- [41] Lundbergh, S., Teräsvirta, T., 2002. Evaluating GARCH models. *Journal of Econometrics*, 110, 417-435.
- [42] Nelson D.B., 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59, 347–370.
- [43] Patton, A.J., 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160, 246-256.
- [44] Patton, A.J., 2020. Comparing possibly misspecified forecasts. *Journal of Business & Economic Statistics*, 38, 796-809.
- [45] van Dijk, D., Franses, P.H., 2003. Selecting a nonlinear time series model using weighted tests of equal forecast accuracy. *Oxford Bulletin of Economics and Statistics*, 65, 727-744.
- [46] Wang, F., Ghysels, E., 2015. Econometric analysis of volatility component models. *Econometric Theory*, 31, 362-393.
- [47] Zivot, E., 2009. Practical issues in the analysis of univariate GARCH models. In: Mikosch, T., Kreiß, J.P., Davis, R., Andersen, T. (eds) *Handbook of Financial Time Series*. Springer, Berlin, Heidelberg.

# Figures

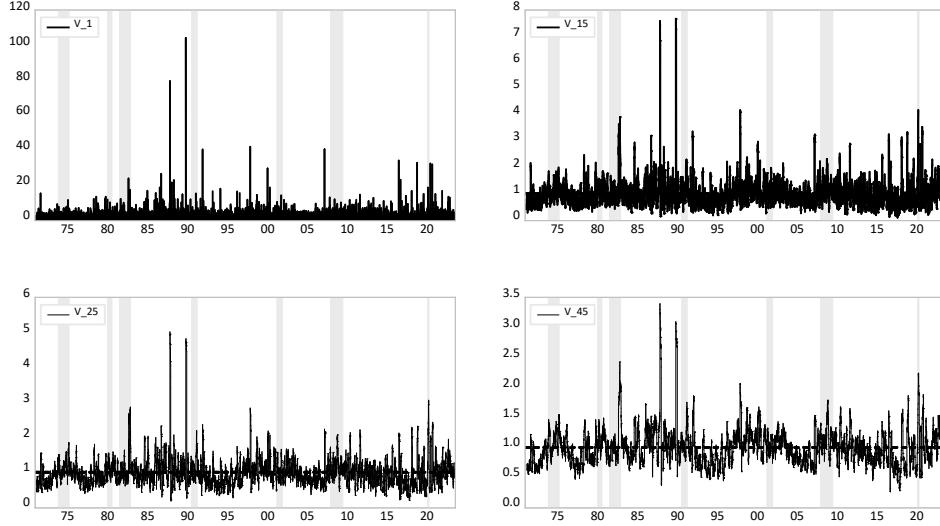


Figure 1: A GJR-GARCH(1,1) is estimated for daily S&P 500 return data for the January 1971 to June 2023 period. The figure shows  $\tilde{V}_t^{(m)}$  for  $m = 1$  (upper left panel),  $m = 15$  (upper right panel),  $m = 25$  (lower left panel), and  $m = 45$  (lower right panel). Grey shaded areas represent NBER recession periods.

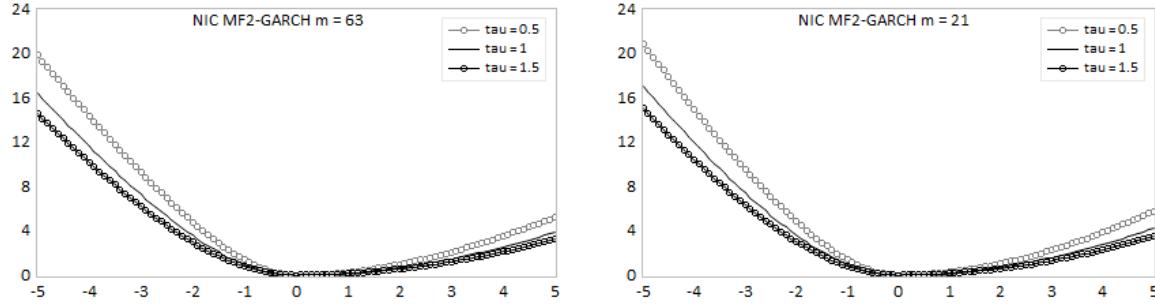


Figure 2: The figure shows the NIC for an MF2-GARCH-rw- $m$  model with  $m = 63$  (left panel) and  $m = 21$  (right panel) and parameters as in Figure A.3. We fix  $h_t = 1$  and assume that the short-term component correctly predicts volatility on days  $t-1$  to  $t-m-1$ , i.e. we set  $r_{t-j}^2/h_{t-j} = 1$  for  $j = 1, \dots, m-1$ . The NICs are plotted as a function of the return,  $r_t$ , and for  $\tau_t \in \{0.5, 1, 1.5\}$ . The NICs are standardized such that news impact is zero for  $r_t = 0$  and presented as annualized volatilities, i.e. we plot  $\sqrt{252(\sigma_{t+1}^2(r_t|\tau_t, h_t = 1) - \sigma_{t+1}^2(r_t = 0|\tau_t, h_t = 1))}$ .

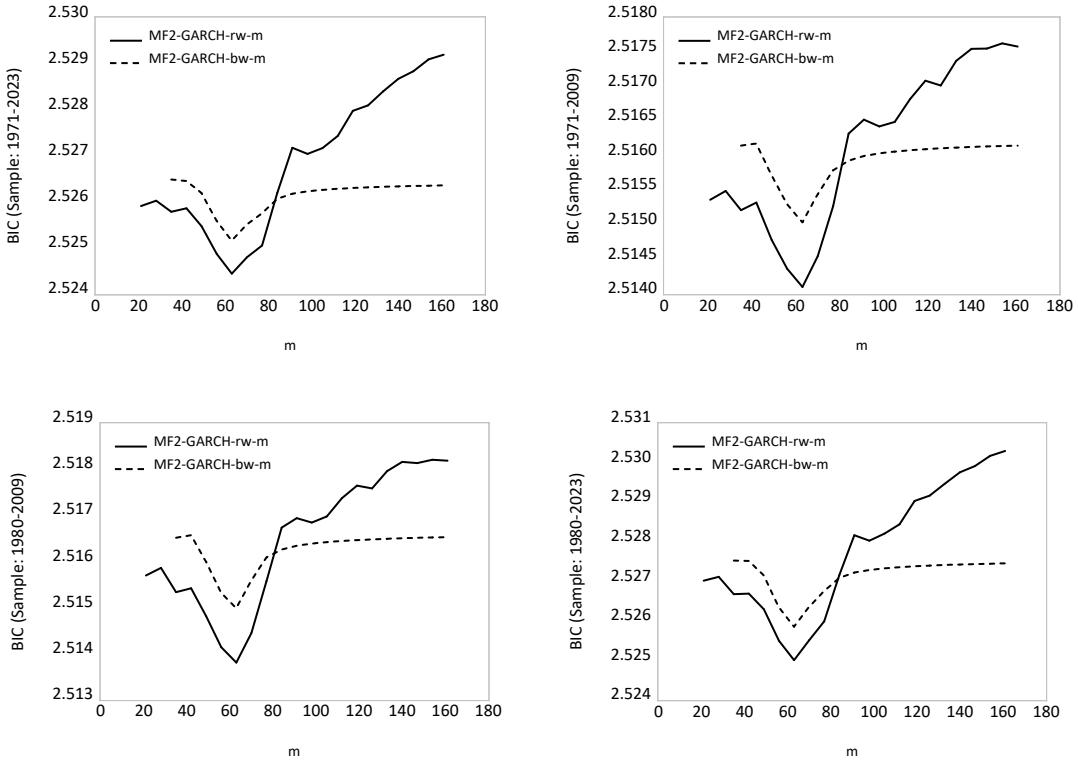


Figure 3: MF2-GARCH-rw- $m$  and MF2-GARCH-bw- $m$  models are estimated using daily S&P 500 return data for the sample periods January 1971 to June 2023 (upper left), January 1971 to December 2009 (upper right), January 1980 to December 2009 (lower left), and January 1980 to June 2023 (lower right). The figure plots the BIC as a function of  $m$ .

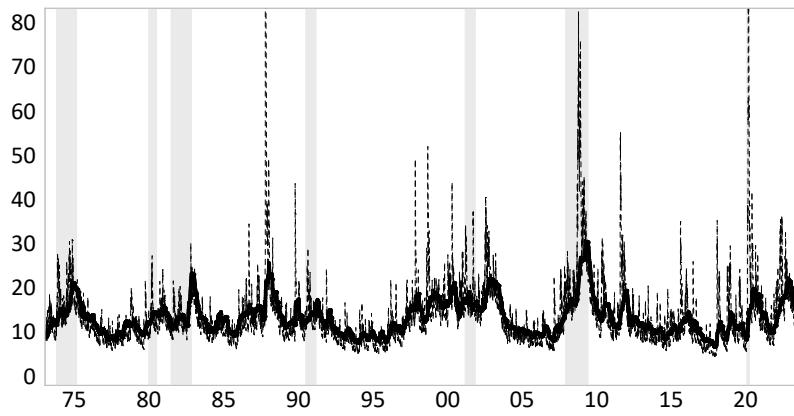


Figure 4: The figure shows the estimated conditional volatility,  $\sqrt{h_t \tau_t}$ , (dashed line) and long-term volatility,  $\sqrt{\tau_t}$ , (solid line) from the MF2-GARCH-rw-63 model for the daily S&P 500 returns presented in Table 2. All quantities are annualized. Grey shaded areas represent NBER recession periods.

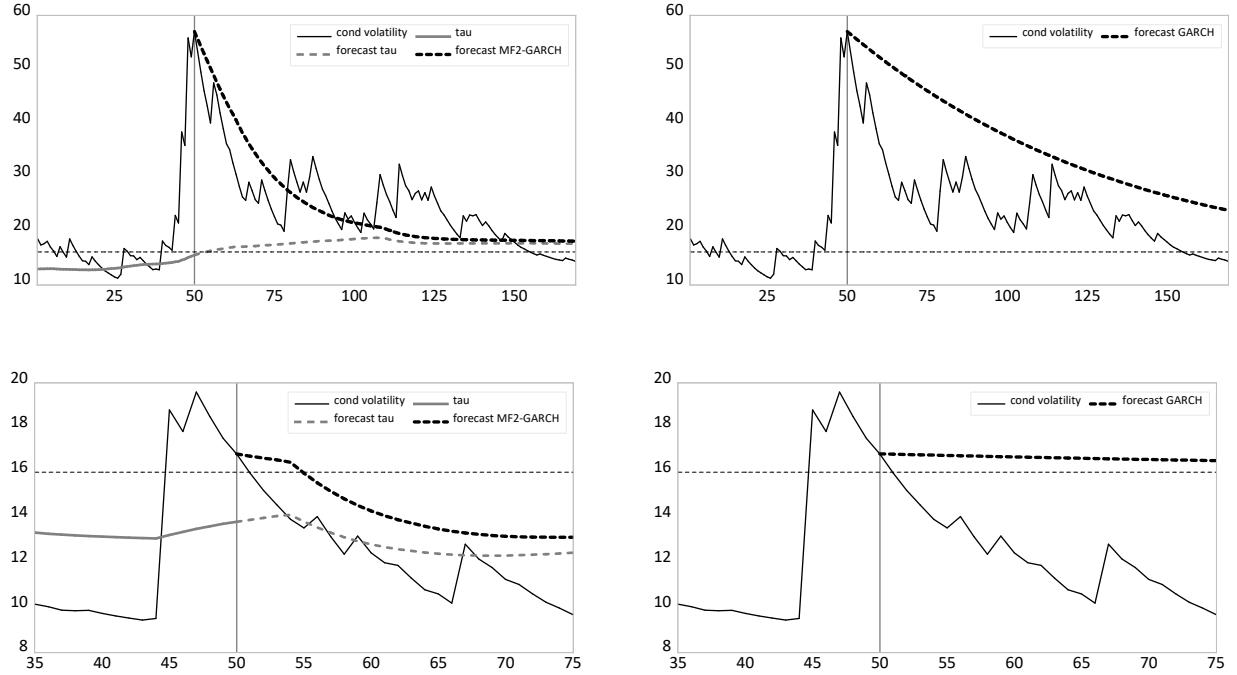


Figure 5: The left panels show the conditional volatility (solid black line) from an MF2-GARCH-rw-63 estimated for S&P 500 returns (see Table 2). From day  $t = 50$  (indicated by the black vertical line) onwards, we compute volatility forecasts (dashed black line). In the upper/lower panels  $t = 50$  corresponds to August 10, 2011 and September 14, 2016, respectively. The plots also show the long-term components (gray line) and the forecast of long-term volatility (dashed gray line). The dashed lines in the right panels are the volatility forecast from a GJR-GARCH(1,1). We impose the same conditional volatility at the forecast origin and the same unconditional volatility. All quantities are annualized.

# Tables

Table 1: Summary statistics S&P 500 and LM test.

<b>Panel A:</b> Summary Statistics							
	mean	sd	skewness	kurtosis	min	max	AC(1)
$r_t$	0.03	1.09	-1.00	27.11	-22.93	10.71	-0.02
$RV_t$	0.99	2.94	18.01	486.49	0.02	101.29	0.63

<b>Panel B:</b> LM test: Explanatory variable $\tilde{V}_t^{(m)}$							
$m$	1	5	15	25	35	45	55
$p$ -value	0.930	0.960	0.450	0.030	0.001	0.001	0.010

**Notes:** Panel A shows summary statistics for the daily returns,  $r_t$ , and the daily realized variances,  $RV_t$ , of the S&P 500. The columns present the mean, the standard deviation (sd), skewness, kurtosis, the minimum (min) and maximum (max) as well as the first-order autocorrelation coefficient (AC(1)). Daily returns for the S&P 500 cover the period January 1971 to June 2023. Realized variances are for the period January 2010 to June 2023. Panel B shows the results of the Conrad and Schienle (2020) LM test for an omitted long-term component under the null hypothesis of a one-component GJR-GARCH. We set  $K = 1$ . The table shows the  $p$ -values of the test for different choices of  $m$ .

Table 2: Estimation results MF2-GARCH models: S&amp;P 500.

$\tau_t$	$\alpha$	$\gamma$	$\beta$	$\lambda_0$	$\lambda_1$	$\lambda_2$	$\omega$	LLF	BIC
GJR-GARCH									
constant	0.023*** (0.006)	0.115*** (0.021)	0.901*** (0.016)	1.015*** (0.136)				-16753.53	2.53433
MF2-GARCH-rw- $m$									
$m = 63$	0.003 (0.007)	0.162*** (0.021)	0.840*** (0.018)	0.018* (0.009)	0.112** (0.054)	0.870*** (0.062)		-16678.61	<b>2.52445</b>
$m = 21$	0.002 (0.007)	0.174*** (0.021)	0.814*** (0.023)	0.005*** (0.002)	0.034*** (0.011)	0.961*** (0.012)		-16688.31	2.52591
$m = 126$	0.012* (0.007)	0.145*** (0.022)	0.864*** (0.018)	0.019 (0.016)	0.101 (0.094)	0.879*** (0.109)		-16702.77	2.52809
MF2-GARCH-bw- $m$									
$m = 63$	0.003 (0.018)	0.162*** (0.055)	0.840*** (0.103)	0.018 (0.045)	0.112 (0.247)	0.870*** (0.294)	1.000 (3.505)	-16678.61	2.52516
$m = 126$	0.001 (0.008)	0.169*** (0.022)	0.824*** (0.027)	0.008** (0.003)	0.058*** (0.021)	0.933*** (0.023)	4.900*** (1.416)	-16686.25	2.52632
MF2-GARCH-lf- $n$									
$n = 63$	0.016*** (0.006)	0.145*** (0.021)	0.864*** (0.016)	0.153*** (0.045)	0.745*** (0.130)	0.098 (0.137)		-16699.87	2.52766
$n = 21$	0.008 (0.007)	0.169*** (0.023)	0.825*** (0.026)	0.062*** (0.020)	0.390*** (0.101)	0.547*** (0.108)		-16702.62	2.52807
GARCH-MIDAS-RV									
$m = 63$	0.005 (0.006)	0.177*** (0.021)	0.808*** (0.018)	0.313*** (0.044)	0.628*** (0.057)		73.671*** (7.047)	-16692.25	2.52651
$m = 21$	0.010 (0.006)	0.180*** (0.023)	0.813*** (0.023)	0.323*** (0.055)	0.608*** (0.072)		4.933*** (1.738)	-16694.05	2.52678

**Notes:** The table reports estimation results for GJR-GARCH, MF2-GARCH-rw- $m$ , MF2-GARCH-bw- $m$ , MF2-GARCH-lf- $n$ , and GARCH-MIDAS-RV (with  $K = 252$ ) models. For the GARCH-MIDAS-RV,  $\lambda_0$  corresponds to  $c$ . All models are estimated using daily return data for the period January 1971 to June 2023. The numbers in parentheses are Bollerslev–Wooldridge robust standard errors. \*\*\*, \*\* and \* indicate significance at the 1%, 5% and 10% level. LLF is the value of the maximized log-likelihood function. BIC is the Bayesian information criterion.

Table 3: Parameter estimates for 2,142 stocks in V-Lab: Summary statistics.

	$\alpha$	$\gamma$	$\beta$	$\phi$	$\lambda_1 + \lambda_2$	$m$	BIC(MF2)<BIC( $\bullet$ )
GJR-GARCH	0.079	0.032	0.883	0.978	-	-	0.959
MF2-GARCH-rw	0.111	0.043	0.726	0.861	0.993	51.000	-
Spline-GARCH	0.118	-	0.800	0.920	-	-	0.632
ZS-Spline-GARCH	0.117	-	0.804	0.923	-	-	0.658

**Notes:** The table presents median parameter estimates for 2,142 US and international stocks. The GJR-GARCH, MF2-GARCH-rw- $m$ , Spline-GARCH and ZS-Spline-GARCH models were estimated in V-Lab. The last column, BIC(MF2)<BIC( $\bullet$ ), reports the percentage of equities for which the BIC of the MF2-GARCH is smaller than the BIC of the model in the respective row. The estimation period ended on April 5, 2024.

Table 4: Out-of-sample forecasting: S&amp;P 500 – pre Covid-19.

Forecast Horizon	day	week	1m	2m	3m	4m	5m	6m	7m	8m
<b>Panel A: relative RMSE</b>										
Historical Forecast	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
GJR-GARCH	0.788	0.763	0.798	0.864	0.901	0.910	0.899	0.883	0.874	0.866
MF2-GARCH-rw	<b>0.744</b>	<b>0.699</b>	<b>0.685</b>	<b>0.790</b>	<b>0.834</b>	<b>0.846</b>	<b>0.832</b>	<b>0.820</b>	<b>0.815</b>	<b>0.811</b>
log-HAR lev.	0.766	0.739	0.725	0.829	0.900	0.926	0.913	0.910	0.899	0.888
GARCH-MIDAS $m = 21$	0.770	0.738	0.715	0.861	0.924	0.958	0.958	0.950	0.938	0.927
GARCH-MIDAS $m = 63$	0.791	0.770	0.746	0.886	0.954	0.992	0.997	0.988	0.978	0.974
<b>Panel B: relative QLIKE</b>										
Historical Forecast	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
GJR-GARCH	0.647	0.640	0.723	<b>0.847</b>	<b>0.901</b>	0.922	0.920	0.918	0.919	0.918
MF2-GARCH-rw	0.615	0.613	0.708	<b>0.827</b>	<b>0.884</b>	<b>0.894</b>	<b>0.875</b>	<b>0.866</b>	<b>0.861</b>	<b>0.859</b>
log-HAR lev.	<b>0.544</b>	<b>0.552</b>	<b>0.677</b>	<b>0.822</b>	0.889	0.924	0.903	0.903	0.902	0.893
GARCH-MIDAS $m = 21$	0.636	0.641	0.734	0.857	0.909	0.918	0.907	0.907	0.899	0.885
ARCH-MIDAS $m = 63$	0.638	0.645	0.745	0.876	0.918	0.926	0.910	0.907	0.910	0.908

**Notes:** The out-of-sample period is January 2010 to December 2019. Using a rolling window scheme, all models are re-estimated on a monthly basis (i.e., every 21 days). For both GARCH-MIDAS-RV models, we set  $K = 252$ . The first in-sample period ends in December 2009. The forecast horizons are one day, one week, and from 1 month (1m) to 8 months (8m). Numbers reported are the RMSE (Panel A) or QLIKE (Panel B) losses for each model and forecast horizon relative to the respective RMSE (Panel A) or QLIKE (Panel B) of historical volatility. Bold numbers indicate the model with the lowest RMSE or QLIKE. Blue shaded cells indicate that the respective model is included in the 85% model confidence set.

Table 5: Out-of-sample forecasting: S&amp;P 500 – full OOS period.

Forecast Horizon	day	week	1m	2m	3m	4m	5m	6m	7m	8m
<b>Panel A: relative RMSE</b>										
Historical Forecast	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
GJR-GARCH	0.990	0.837	0.869	1.618	0.863	0.987	0.912	0.834	0.814	0.798
MF2-GARCH-rw	0.781	<b>0.641</b>	<b>0.602</b>	0.766	<b>0.735</b>	<b>0.757</b>	<b>0.772</b>	<b>0.780</b>	<b>0.781</b>	<b>0.786</b>
log-HAR lev.	<b>0.776</b>	0.662	0.628	<b>0.765</b>	0.818	0.841	0.859	0.856	0.848	0.843
GARCH-MIDAS $m = 21$	1.197	0.891	0.938	1.101	1.215	1.293	1.296	1.306	1.333	1.368
GARCH-MIDAS $m = 63$	1.167	0.916	0.903	1.339	1.329	1.411	1.440	1.437	1.492	1.495
<b>Panel B: relative QLIKE</b>										
Historical Forecast	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
GJR-GARCH	0.676	0.652	0.702	0.829	0.873	<b>0.892</b>	0.898	0.894	0.891	0.889
MF2-GARCH-rw	0.646	0.613	0.664	<b>0.774</b>	<b>0.832</b>	<b>0.853</b>	<b>0.857</b>	<b>0.857</b>	<b>0.856</b>	<b>0.861</b>
log-HAR lev.	<b>0.572</b>	<b>0.549</b>	<b>0.625</b>	<b>0.770</b>	0.840	0.874	0.885	0.888	0.890	0.886
GARCH-MIDAS $m = 21$	0.669	0.650	0.707	0.821	0.878	0.898	0.909	0.911	0.920	0.927
GARCH-MIDAS $m = 63$	0.669	0.653	0.713	0.833	0.879	0.899	0.905	0.906	0.918	0.932

**Notes:** The out-of-sample period is January 2010 to June 2023. See notes of Table 4.

Table 6: Out-of-sample forecasting in V-Lab: Twenty US stocks.

	day	week	1m	2m	3m	4m	5m	6m	7m	8m
<b>Panel A: Ranking according to RMSE</b>										
GJR-GARCH	2.45	2.45	2.70	3.05	2.80	3.10	2.90	2.80	2.75	2.70
MF2-GARCH-rw	<b>1.90</b>	<b>1.80</b>	<b>1.85</b>	<b>1.60</b>	<b>1.70</b>	<b>1.60</b>	<b>1.60</b>	<b>1.55</b>	<b>1.50</b>	<b>1.50</b>
Spline-GARCH	3.35	3.55	3.25	3.35	3.50	3.35	3.45	3.50	3.40	3.50
ZS-Spline-GARCH	2.30	2.20	2.20	2.00	2.00	1.95	2.05	2.15	2.35	2.30
<b>Panel B: Ranking according to QLIKE</b>										
GJR-GARCH	2.20	2.10	1.75	<b>1.90</b>	2.20	2.45	2.50	2.55	2.65	2.70
MF2-GARCH-rw	<b>1.40</b>	<b>1.35</b>	<b>1.70</b>	<b>1.90</b>	<b>1.70</b>	<b>1.65</b>	<b>1.70</b>	<b>1.70</b>	<b>1.75</b>	<b>1.80</b>
Spline-GARCH	3.95	4.00	3.85	3.60	3.60	3.40	3.30	3.25	3.15	3.20
ZS-Spline-GARCH	2.45	2.55	2.70	2.60	2.50	2.50	2.50	2.45	2.45	2.30

**Notes:** The out-of-sample period is January 2010 to June 2023. All models are re-estimated on a daily basis. The first in-sample period ends in December 2009. For each forecast horizon, the numbers represent the average ranking of the respective models across the twenty US stocks according to the RMSE (Panel A) and the QLIKE (Panel B). Bold numbers indicate the model with the lowest ranking.

Online Appendix for

## **Modelling Volatility Cycles: The MF2-GARCH Model**

Christian Conrad (Heidelberg University)  
Robert F. Engle (New York University)

January 14, 2025

This appendix provides further details and supporting evidence for “Modelling Volatility Cycles: The MF2-GARCH Model.”

# A Further Details and Results

## A.1 Quasi-Maximum Likelihood Estimation

Wang and Ghysels (2015) provide the asymptotic theory of the quasi-maximum likelihood estimator (QMLE) for the GARCH-MIDAS-RV model. Francq et al. (2023) extend the results of Wang and Ghysels (2015) to a case without moment assumption on the observed process. Following the arguments in those papers, we sketch the corresponding theory for the MF2-GARCH in Section A.1.1. In addition, we provide a Monte-Carlo simulation to evaluate the finite sample performance of the QMLE in Section A.1.2.

### A.1.1 Discussion of estimator

Let  $\boldsymbol{\theta}_{GA} = (\alpha, \beta, \gamma)'$  and  $\boldsymbol{\theta}_{MEM} = (\lambda_0, \lambda_1, \lambda_2)'$  denote the vectors of parameters in the short-term and long-term component. Further, let  $\boldsymbol{\theta} = (\boldsymbol{\theta}'_{GA}, \boldsymbol{\theta}'_{MEM})'$ . The Gaussian quasi-loglikelihood function (omitting the constant) takes the same form as for the GARCH-MIDAS-RV model of Wang and Ghysels (2015) and can be written as  $L(\boldsymbol{\theta}|r_T, r_{T-1}, \dots) = \sum_{t=1}^T l_t$  with

$$l_t = -\frac{1}{2} \left[ \ln(h_t) + \ln(\tau_t) + \frac{r_t^2}{h_t \tau_t} \right], \quad (18)$$

where  $h_t = h_t(\boldsymbol{\theta})$  and  $\tau_t = \tau_t(\boldsymbol{\theta})$  are defined as in equations (4) and (5). The first and second derivative of the log-likelihood are denoted by

$$\mathbf{s}_t(\boldsymbol{\theta}) = \frac{\partial l_t}{\partial \boldsymbol{\theta}} \quad \text{and} \quad \mathbf{d}_t(\boldsymbol{\theta}) = \frac{\partial^2 l_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} . \quad (19)$$

The first derivative can be written as

$$\mathbf{s}_t(\boldsymbol{\theta}) = \frac{1}{2} \left[ \frac{r_t^2}{h_t \tau_t} - 1 \right] \mathbf{y}_t(\boldsymbol{\theta}) \quad (20)$$

with

$$\mathbf{y}_t(\boldsymbol{\theta}) = \left( \frac{1}{h_t} \frac{\partial h_t}{\partial \boldsymbol{\theta}} + \frac{1}{\tau_t} \frac{\partial \tau_t}{\partial \boldsymbol{\theta}} \right) . \quad (21)$$

Because in the MF2-GARCH, not only does the short-term component depend on the long-term component but also vice versa, the first derivative is more complicated than in the GARCH-MIDAS-RV. For example, consider the derivative with respect to  $\alpha$ :

$$\frac{\partial h_t}{\partial \alpha} = -1 + \frac{r_{t-1}^2}{\tau_{t-1}} \left( 1 - \frac{\alpha}{\tau_{t-1}} \frac{\partial \tau_{t-1}}{\partial \alpha} \right) + \beta \frac{\partial h_{t-1}}{\partial \alpha} \quad (22)$$

and

$$\frac{\partial \tau_t}{\partial \alpha} = -\frac{\lambda_1}{m} \sum_{j=1}^m \frac{r_{t-j}^2}{h_{t-j}^2} \frac{\partial h_{t-j}}{\partial \alpha} + \lambda_1 \frac{\partial \tau_{t-1}}{\partial \alpha} . \quad (23)$$

That is, the derivative of  $h_t$  with respect to  $\alpha$  depends on the derivative of  $\tau_{t-1}$  with respect to  $\alpha$  and vice versa. In contrast, in the GARCH-MIDAS-RV the derivative of  $\tau_t$  with respect to the short-term component's parameters is zero. Similarly, the derivative with respect to  $\lambda_1$  is given by

$$\frac{\partial h_t}{\partial \lambda_1} = -\alpha \frac{r_{t-1}^2}{\tau_{t-1}^2} \frac{\partial \tau_{t-1}}{\partial \lambda_1} + \beta \frac{\partial h_{t-1}}{\partial \lambda_1} \quad (24)$$

and

$$\frac{\partial \tau_t}{\partial \lambda_1} = \frac{1}{m} \sum_{j=1}^m \frac{r_{t-j}^2}{h_{t-j}} - \frac{\lambda_1}{m} \sum_{j=1}^m \frac{r_{t-j}^2}{h_{t-j}^2} \frac{\partial h_{t-j}}{\partial \lambda_1} + \lambda_1 \frac{\partial \tau_{t-1}}{\partial \lambda_1}. \quad (25)$$

Again, equation (25) is more complicated than in the GARCH-MIDAS-RV. In the GARCH-MIDAS-RV,  $\tau_t$  depends only on parameters and observed data.

The second derivative can be written as

$$\mathbf{d}_t(\boldsymbol{\theta}) = \frac{\partial \mathbf{s}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} = -\frac{1}{2} \frac{r_t^2}{h_t \tau_t} \mathbf{y}_t(\boldsymbol{\theta}) \mathbf{y}_t(\boldsymbol{\theta})' + \frac{1}{2} \left[ \frac{r_t^2}{h_t \tau_t} - 1 \right] \frac{\partial \mathbf{y}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}. \quad (26)$$

Because we only observe  $r_t$  for  $t = 1, \dots, T$  with  $T \gg m$ ,  $L(\boldsymbol{\theta}|r_T, r_{T-1}, \dots, r_0, r_{-1}, \dots)$  is unobserved. We denote by  $\tilde{L}(\boldsymbol{\theta}|r_T, r_{T-1}, \dots, r_1) = \sum_{t=1}^T \tilde{l}_t$  the observable version that depends on initial values for the short- and long-term component. Using similar notation as before, we write the first and second derivative of the observable likelihood as  $\tilde{\mathbf{s}}_t(\boldsymbol{\theta})$  and  $\tilde{\mathbf{d}}_t(\boldsymbol{\theta})$ . For the quasi-maximum likelihood estimator,  $\hat{\boldsymbol{\theta}}$ , it holds that  $\sum_t \tilde{\mathbf{s}}_t(\hat{\boldsymbol{\theta}}) = \mathbf{0}$ .

Deriving the complete asymptotic theory for the QMLE of the MF2-GARCH is beyond the scope of our paper. However, we expect that the assumptions stated in Wang and Ghysels (2015) and Francq et al. (2023) can be appropriately adapted to the MF2-GARCH. For example, note that our Assumption 2 ensures that  $\sqrt{h_t} Z_t$  is strictly stationary (see Assumption 2.2 in Whang and Ghysels, 2015) and Assumption 3 ensures that  $\sqrt{\tau_t} Z_t$  is strictly stationary. Under Assumption 1,  $\kappa < \infty$  and  $Z_t^2$  has a non-degenerate distribution as in Assumption 4.3 in Wang and Ghysels (2015). Let  $\boldsymbol{\theta}_0$  denote the vector of true parameters. We assume that  $\boldsymbol{\theta}_0$  lies in the interior of an appropriately defined compact parameter space. In particular, we preclude the case that the parameters in  $h_t$  and  $\tau_t$  are zero (see below). Finally, because  $\frac{r_t^2}{h_t(\boldsymbol{\theta}_0)\tau_t(\boldsymbol{\theta}_0)} = Z_t^2$ , it follows that  $\mathbf{E}[\mathbf{s}_t(\boldsymbol{\theta}_0)|\mathcal{F}_{t-1}] = \mathbf{0}$ , i.e.  $\mathbf{s}_t(\boldsymbol{\theta}_0)$  follows a martingale difference sequence. The asymptotic distribution of the QMLE is then expected to be given by

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{D}^{-1} \boldsymbol{\Omega} \mathbf{D}^{-1}), \quad (27)$$

where  $\boldsymbol{\Omega} = \boldsymbol{\Omega}(\boldsymbol{\theta}_0) = \mathbf{E}[\mathbf{s}_t(\boldsymbol{\theta}_0)\mathbf{s}_t(\boldsymbol{\theta}_0)']$  and  $\mathbf{D} = \mathbf{D}(\boldsymbol{\theta}_0) = -\mathbf{E}[\mathbf{d}_t(\boldsymbol{\theta}_0)]$ . These expressions simplify to  $\boldsymbol{\Omega} = \frac{1}{4}(\kappa - 1)\mathbf{E}[\mathbf{y}_t(\boldsymbol{\theta}_0)\mathbf{y}_t(\boldsymbol{\theta}_0)']$  and  $\mathbf{D} = \frac{1}{2}\mathbf{E}[\mathbf{y}_t(\boldsymbol{\theta}_0)\mathbf{y}_t(\boldsymbol{\theta}_0)']$ . Thus, as in Wang

and Ghysels (2015), we obtain

$$\mathbf{D}^{-1}\boldsymbol{\Omega}\mathbf{D}^{-1} = (\kappa - 1) (\mathbf{E}[\mathbf{y}_t(\boldsymbol{\theta}_0)\mathbf{y}_t(\boldsymbol{\theta}_0)'])^{-1} = (\kappa - 1) \left( \mathbf{E} \left[ \frac{1}{h_t^2\tau_t^2} \frac{\partial\sigma_t^2}{\partial\boldsymbol{\theta}} \frac{\partial\sigma_t^2}{\partial\boldsymbol{\theta}'}(\boldsymbol{\theta}_0) \right] \right)^{-1}. \quad (28)$$

From the results of Francq and Zakoian (2007), we know that the asymptotic distribution of the QMLE for the parameters of a GARCH model is no longer Gaussian when some of the coefficients are null. As discussed in Francq and Zakoian (2009), this implies that Wald tests of the nullity of the GARCH coefficients require modifications of the standard critical regions. Extending their results to testing the nullity of the coefficients in the long-term component of the MF2-GARCH is left for future research. In any case, we suggest estimating a one-component GJR-GARCH in the first step and then applying the Conrad and Schienle (2020) test with  $\tilde{V}_t^{(m)}$ , as defined in equation (3), as the explanatory variable. If the null hypothesis is rejected, the MF2-GARCH can be estimated.

### A.1.2 Simulation: Parameter estimates and standard errors

We set the true parameters to  $\alpha_0 = 0.02$ ,  $\gamma_0 = 0.10$ ,  $\beta_0 = 0.8$ ,  $\lambda_{00} = 0.02$ ,  $\lambda_{01} = 0.05$ , and  $\lambda_{02} = 0.94$ . Overall, we perform  $R = 1,000$  Monte Carlo replications. Table A.1 presents the mean parameter estimates across the 1,000 simulations. To estimate the model parameters, we rely on the MaxSQPF algorithm, as implemented in OxMetrics. In addition, the table shows the root mean squared error (RMSE) of the estimates of each parameter as well as the asymptotic standard error (ASE). The ASE is the square root of the mean (across the  $R$  replications) of the asymptotic variance of each parameter. The asymptotic variance is computed based on numerical derivatives. For example, for parameter  $j$ , the ASE is defined as  $\sqrt{1/R \sum_{i=1}^R \widehat{\text{Var}}(j,j)}$ , where  $\widehat{\text{Var}}(j,j)$  is the estimate of  $(j,j)$ -th element of the asymptotic covariance matrix in equation (27). Panel A presents simulation results for  $Z_t$  being standard normal and Panel B for  $Z_t$  following a student  $t$ -distribution with five degrees of freedom. The latter is standardized to have a variance of one. Panel A shows results for  $m = 21$  (i.e., a monthly averaging) and sample sizes of  $T = 5,000$  and  $T = 10,000$ , which roughly correspond to 20 and 40 years of daily observations. In both cases, the bias of all the parameter estimates is essentially zero. The ASE is close to the RMSE, indicating that the QMLE procedure leads to valid inference. As expected, the RMSE and the ASE decrease with increasing sample size. When setting  $m = 63$  (i.e., a quarterly averaging), the parameter estimates remain unbiased. However, there is a considerable increase in the RMSE and ASE of the parameters in the long-term component, which are nevertheless still close to each other. This increase can be

explained by the fact that the long-term component now depends on the sum of 63 lagged forecast errors of the short-term component, which complicates the optimization and the computation of the derivatives. As before, RMSE and the ASE decrease with increasing sample size. In Panel B, the parameter estimates are still unbiased in all specifications. Nevertheless, replacing the normal distribution with the student  $t$ -distribution increases the RMSE and ASE for all parameters (not only the parameters of the long-term component). The reason is that for five degrees of freedom, the fourth moment of  $Z_t$  is  $\kappa = 6$ , which inflates the asymptotic variance.

## A.2 Simulation: Misspecified models

In this section, we evaluate the consequences of estimating a one-component GJR-GARCH or a GARCH-MIDAS-RV model when the true data generating process follows an MF2-GARCH-rw- $m$ .<sup>17</sup> The true parameters of the MF2-GARCH are chosen as in Section A.1.2. Again, we employ  $R = 1000$  Monte Carlo replications. We present the mean parameter estimates as well as the squared error (SE) and the QLIKE loss of the one-day ahead conditional variance predictions. For example, for each replication, the SE and QLIKE losses of model  $j \in \{\text{GJR-GARCH}, \text{GARCH-MIDAS-RV}, \text{MF2-GARCH}\}$  are computed as

$$SE_j = \frac{1}{T} \sum_{t=1}^T (\sigma_{t+1}^2 - \hat{\sigma}_{t+1|t,j}^2)^2 \quad (29)$$

and

$$QLIKE_j = \frac{1}{T} \sum_{t=1}^T \left( \frac{\sigma_{t+1}^2}{\hat{\sigma}_{t+1|t,j}^2} + \ln(\sigma_{t+1}^2 / \hat{\sigma}_{t+1|t,j}^2) - 1 \right), \quad (30)$$

where  $\sigma_{t+1}^2$  is the true conditional variance and  $\hat{\sigma}_{t+1|t,j}^2$  is the conditional variance forecast from model  $j$ . We average the losses across the  $R$  replications and present the average loss relative to the average loss of the estimated MF2-GARCH model. The focus on the relative losses is motivated by Patton (2020). In the presence of estimation error, the forecasts of the misspecified one-component GJR-GARCH or the GARCH-MIDAS-RV model might dominate the forecasts of the correctly specified MF2-GARCH. Whether this happens depends on the ‘degree’ of misspecification, estimation error, and the employed loss function (see Patton, 2020). In particular, a “parsimoniously misspecified model” might be preferred to a “correctly specified model that is subject to estimation error” (Patton,

---

<sup>17</sup>The GARCH-MIDAS-RV is formally introduced in Section A.3. For the simulation, we set  $K = 252$  and employ a quarterly realized variance based on daily return data.

2020, p.804). Only in the ideal setting of comparing models that are correctly specified for the respective loss function, no estimation error and nested information sets, the SE and QLIKE loss will asymptotically yield the same model ranking.

The Panel A of Table A.2 shows simulation results for  $m = 21$  and Panel B for  $m = 63$ . In both panels, we assume that  $Z_t$  is standard normal and  $T = 5,000$ . First, the GJR-GARCH estimate of  $\beta_0$  is way above the true value. Conversely,  $\gamma_0$  is severely underestimated. As expected, the estimated parameters reflect the so-called ‘IGARCH’ effect: The misspecified one-component GJR-GARCH tends to overestimate volatility persistence. Although there is some upward bias in the estimate of  $\beta_0$ , the parameter estimates of the GARCH-MIDAS-RV’s short-term component are much closer to the true parameters values. The columns SE and QLIKE reveal whether there is a trade-off between bias in the estimated parameters and estimation error. For example, while the one-component GJR-GARCH parameter estimates are biased, in principle, the model might outperform the MF2-GARCH in terms of forecast accuracy due to less estimation error. However, the SE-column shows that this is not the case. The average squared error of the GJR-GARCH is 6.5 times the average squared error of the MF2-GARCH. For the GARCH-MIDAS-RV the ratio is 3.4. We obtain similar ratios for the QLIKE.<sup>18</sup> Panel B shows that the biases in the estimated parameters of the GJR-GARCH and the GARCH-MIDAS-RV tend to decrease when  $m = 63$ . Note that for  $m = 63$ , the MF2-GARCH’s long-term component is smoother, making the QML estimates of the parameters in the MF2-GARCH’s long-term component less precise (see Section A.1.2). Although a smoother long-term component decreases the variance of the daily returns, the deviations of the conditional variance from the long-term component are more persistent. This makes the forecasts from the GJR-GARCH ‘more misspecified’ and, hence, the relative SE and QLIKE increase compared to Panel A. In contrast, the GARCH-MIDAS-RV can handle a smoothly varying long-term component; hence, we observe the opposite effect. The relative SE and QLIKE of the GARCH-MIDAS-RV decrease. Overall, the simulation results show that for reasonable parameter values of the MF2-GARCH the degree of misspecification of the one-component GJR-GARCH and the GARCH-MIDAS-RV is so strong that both loss functions will be able to identify the true model. However, there is a caveat. In real-world forecasting applications, the true conditional variance is not observable and needs to be replaced by a proxy, e.g., the realized variance. As discussed in Patton (2011), the SE and the QLIKE

---

<sup>18</sup>We also computed how often the RMSE (QLIKE) ratios are below one across the  $R$  replications, i.e., how often the MF2-GARCH wins the ranking. This happens in 99% of the simulations.

are robust loss functions in the sense that they preserve the expected loss ranking of models when the latent conditional variance is replaced by a conditionally unbiased proxy.

### A.3 Relation to other components models

Amado et al. (2019) provide a survey on multiplicative volatility models defined as  $r_t/\sqrt{\tau_t} = \sqrt{h_t}Z_t$ . Those models assume that  $h_t$  follows some GARCH-type process but differ concerning the specification of  $\tau_t$ . As note before, in the Spline-GARCH of Engle and Rangel (2008) and the MTV-GARCH of Amado and Teräsvirta (2013, 2017), the long-term component is deterministic. The MF2-GARCH is closely related to the GARCH-MIDAS of Engle et al. (2013). In the latter model,  $\tau_t$  depends on lagged values either of the realized variance or one or several exogenous explanatory variables. The GARCH-MIDAS-RV (see also Wang and Ghysels, 2015) assumes a daily long-term component that depends on a rolling window realized variance and can be written as

$$\tau_t^{GM} = c + \lambda_1 \sum_{j=1}^K w_j(\omega) RV_{t-j}^{rw}, \quad (31)$$

where  $RV_{t-1}^{rw} = \sum_{s=1}^m r_{t-s}^2$  and  $w_j(\omega)$  is a weighting scheme (as in Section 3.2). If we rewrite the long-term component of the MF2-GARCH-rw- $m$  as a  $\text{MEM}(\infty)$  with  $\tau_t = \lambda_0/(1 - \lambda_2) + \lambda_1 \sum_{j=1}^{\infty} \lambda_2^{j-1} V_{t-j}^{(m)}$ , it becomes clear that the MF2-GARCH corresponds to a GARCH-MIDAS-X with an exponential weighting scheme and explanatory variable  $V_t^{(m)}$ .

The difference in how the long-term component is constructed has important implications for how the two models update volatility in response to news. We illustrate this by looking at the NIC of both models in a simplified setting. We assume that the MF2-GARCH and the GARCH-MIDAS-RV have the same short-term component  $h_t$  with parameters  $\alpha = 0.02$ ,  $\gamma = 0.1$  and  $\beta = 0.8$ . For the long-term components, we assume the following specifications:

$$\tau_{t+1} = \lambda_0 + \lambda_1 \frac{\varepsilon_t^2}{h_t} \quad (32)$$

$$\tau_{t+1}^{GM} = c + \lambda_1 \varepsilon_t^2. \quad (33)$$

For the MF2-GARCH, with equation (32) we assume that  $\lambda_2 = 0$  and  $m = 1$ . Equation (33) is the long-term component of a GARCH-MIDAS-RV with  $K = 1$  and  $m = 1$ . Obviously, in the MF2-GARCH, the response of the long-term component to news is the stronger the lower  $h_t$ . In contrast, in the GARCH-MIDAS the response of the long-term component does not depend on the short-term component. Figure A.4 shows the news impact functions for

both models. We set  $\lambda_0 = c = 0.3$ ,  $\lambda_1 = 0.7$ , fix  $\tau_t = \tau_t^{GM} = 1$  and plot the NIC for  $h_t \in \{0.5, 1, 1.5\}$ . When  $h_t = 1$ , the NIC of both models is the same. Similarly as in Figure 2, when short-term volatility decreases (from  $h_t = 1$  to  $h_t = 0.5$ ), the MF2-GARCH becomes much more responsive to news, while it becomes considerably less responsive to news when short-term volatility increases (from  $h_t = 1$  to  $h_t = 1.5$ ). For the GARCH-MIDAS-RV, we see the opposite effect: It responds slightly less to news when short-term volatility decreases and slightly more when short-term volatility increases. Thus, both models adjust their responsiveness in opposite directions, whereby the MF2-GARCH adjusts more strongly.<sup>19</sup>

While multi-step ahead volatility forecasts can be easily computed for the MF2-GARCH, this is less straightforward for the GARCH-MIDAS-RV. Most often, multi-step ahead forecasts of the short-term component are computed and then multiplied by the one-step ahead forecast of the long-term component (see, for example, Conrad and Kleen, 2020). However, with an increasing forecast horizon, simply fixing the long-term component at the one-step ahead forecast becomes increasingly disadvantageous because no mean reversion is allowed (see Section 4.3).

Further related models are the Realized EGARCH-MIDAS of Borup and Jakobsen (2019) and the realized-measure-based component model of Noureldin (2022). In contrast to the MF2-GARCH, those specifications combine daily returns with high-frequency volatility measures. While all the previous models are of the GARCH-type, there are also pure component MEM models. For example, see the Composite-MEM for intraday volumes by Brownlees et al. (2011) or the Component-MEM for realized volatilities by Amendola et al. (2020).

#### A.4 Forecasting the $\tau_t$ of the MF2-GARCH-rw- $m$

First,  $\mathbf{E}[\tau_{t+1}|\mathcal{F}_t] = \tau_{t+1}$ . Second, for  $s = 2, \dots, m$  the forecasts can be recursively calculated as

$$\begin{aligned}\mathbf{E}[\tau_{t+s}|\mathcal{F}_t] &= \lambda_0 + (\lambda_1 \frac{1}{m} + \lambda_2) \mathbf{E}[\tau_{t+s-1}|\mathcal{F}_t] + \lambda_1 \frac{1}{m} \sum_{j=2}^{s-1} \mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t] \\ &\quad + \lambda_1 \frac{1}{m} \sum_{j=s}^m \frac{r_{t+s-j}^2}{h_{t+s-j}}.\end{aligned}\tag{34}$$

---

<sup>19</sup>Note that the Spline-GARCH's long-term component does not depend on news or the current level of volatility. Thus, Spline-GARCH's news impact function simply equals the news impact function of the short-term component scaled with the long-term component.

Third, for  $s > m$  the following recursion applies:

$$\mathbf{E}[\tau_{t+s}|\mathcal{F}_t] = \lambda_0 + (\lambda_1 \frac{1}{m} + \lambda_2) \mathbf{E}[\tau_{t+s-1}|\mathcal{F}_t] + \lambda_1 \frac{1}{m} \sum_{j=2}^m \mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t]. \quad (35)$$

## A.5 What does the long-term component capture?

Baker et al. (2021) relate stock market volatility (as measured by the VIX or realized volatility) to news about the economy and policy. For this, they construct equity market volatility (EMV) trackers based on a counting of newspaper articles on stock market volatility. From the content of the newspaper articles, they infer which types of news are the most important drivers of volatility. Thereby, they decompose the ‘Overall’ EMV tracker into subcomponents. To better understand the economic determinants of long-term volatility, we relate the long-term component of the MF2-GARCH-rw-63 to the EMV trackers. The EMV trackers can be obtained from 1985 onwards from the FRED database maintained by the Federal Reserve Bank of St. Louis. The highest frequency at which the EMV trackers are available is monthly. We regress the difference between the long-term component at the end of the current and the end of the previous month (using the mixed frequency notation:  $\tau_{n,t} - \tau_{n,t-1}$ ) on selected EMV trackers for the current month. As shown in Column (1) of Table A.3, we find a strong positive association between the monthly change in the long-term component and the ‘Overall’ tracker. When focusing on sub-trackers, ‘Macroeconomic News & Outlook’ and ‘Monetary Policy’ have a significant effect on long-term volatility (see Column (2)). Finally, the sub-trackers ‘Inflation’ and ‘Interest Rates’, which are components of ‘Macroeconomic News & Outlook’, appear to be particularly relevant. Overall, the evidence suggests that the long-term volatility of the MF2-GARCH is driven by various types of economic news, which a GARCH-MIDAS-X model with a single macroeconomic or financial explanatory variable is unlikely to capture fully.

## A.6 Details Model Confidence Set

Let  $\mathcal{M}$  denote the set of all competing models and  $d_{ij,t}(s)$  the sample loss difference between models  $i$  and  $j$  when forecasting volatility in period  $s$  (one day, one week,  $s$ -months ahead) at day  $t$ . The average sample loss difference between models  $i$  and  $j$  for forecast horizon  $s$  is denoted by  $\bar{d}_{ij}(s)$ . The MCS test statistic is given by  $T_{\mathcal{M}}(s) = \max_{i,j \in \mathcal{M}} |t_{i,j}(s)|$ , where the  $t_{ij}(s) = \bar{d}_{ij}(s)/\text{Var}(\bar{d}_{ij}(s))$  are pairwise  $t$ -statistics for testing equal forecast performance. The null hypothesis of the MCS test statistic is that all models in  $\mathcal{M}$  have equal forecast

performance. Under the alternative, one model has an expected loss that is greater than the loss of all other models. If the null gets rejected, the corresponding model is eliminated from  $\mathcal{M}$ . The test is repeated until the null can't be rejected. We refer to  $\mathcal{M}_{MCS}$  as the final set of surviving models. Following Hansen et al. (2011), we approximate the asymptotic distribution of the test statistic by block-bootstrapping and choose a confidence level of 85%. The block-length is horizon-specific and depends on the persistence of  $d_{i,t}(s) = n_{\mathcal{M}}^{-1} \sum_{j \in \mathcal{M}} d_{ij,t}(s)$ ,  $i = 1, \dots, n_{\mathcal{M}}$ . We first fit an AR model to each  $d_{i,t}(s)$  and determine the optimal lag order by the BIC. Then, the block-length is chosen as the maximum lag order selected by the BIC.

## B Proofs

**Proof of Theorem 1.** First, recall that Assumptions 1-3 ensure that  $\mathbf{E}[h_t] = 1$  and  $\mathbf{E}[\tau_t] = \lambda_0/(1 - \lambda_1 - \lambda_2)$ . We can write  $\mathbf{E}[h_t \tau_t]$  as

$$\begin{aligned} \mathbf{E}[h_t \tau_t] &= \mathbf{E} \left[ ((1 - \phi) + ((\alpha + \gamma \mathbf{1}_{\{Z_{t-1} < 0\}}) Z_{t-1}^2 + \beta) h_{t-1}) \right. \\ &\quad \times \left. \left( \lambda_0 + \lambda_1 \frac{1}{m} \sum_{j=1}^m Z_{t-j}^2 \tau_{t-j} + \lambda_2 \tau_{t-1} \right) \right] \\ &= (1 - \phi) (\lambda_0 + \lambda_1 \mathbf{E}[\tau_t] + \lambda_2 \mathbf{E}[\tau_t]) + \lambda_0 \phi \mathbf{E}[h_{t-1}] \\ &\quad + \lambda_1 \frac{1}{m} \phi \kappa \mathbf{E}[h_{t-1} \tau_{t-1}] + \lambda_2 \phi \mathbf{E}[h_{t-1} \tau_{t-1}] + \lambda_1 \phi \frac{1}{m} \sum_{j=2}^m \mathbf{E}[h_{t-1} Z_{t-j}^2 \tau_{t-j}]. \quad (36) \end{aligned}$$

Note that we can write

$$\begin{aligned} h_{t-1} &= (1 - \phi) \left( 1 + \sum_{k=1}^{j-2} \prod_{s=1}^k ((\alpha + \gamma \mathbf{1}_{\{Z_{t-1-s} < 0\}}) Z_{t-1-s}^2 + \beta) \right) \\ &\quad + h_{t-j} \prod_{k=1}^{j-1} ((\alpha + \gamma \mathbf{1}_{\{Z_{t-1-k} < 0\}}) Z_{t-1-k}^2 + \beta). \quad (37) \end{aligned}$$

Using equation (37), we can express the expectation  $\mathbf{E}[h_{t-1} Z_{t-j}^2 \tau_{t-j}]$  as

$$\mathbf{E}[h_{t-1} Z_{t-j}^2 \tau_{t-j}] = (1 - \phi) \mathbf{E}[\tau_t] \left( 1 + \sum_{k=1}^{j-2} \phi^k \right) + \phi^{j-2} \phi \kappa \mathbf{E}[h_{t-j} \tau_{t-j}]. \quad (38)$$

Plugging this into equation (36) leads to

$$\begin{aligned} \mathbf{E}[h_t \tau_t] &= (1 - \phi) \frac{\lambda_0}{1 - \lambda_1 - \lambda_2} + \lambda_0 \phi + (1 - \phi) \lambda_1 \phi \mathbf{E}[\tau_t] \left( \frac{m-1}{m} + \frac{1}{m} \sum_{j=2}^m \sum_{k=1}^{j-2} \phi^k \right) \\ &\quad + \left( \lambda_1 \frac{1}{m} \phi \kappa + \lambda_2 \phi \right) \mathbf{E}[h_{t-1} \tau_{t-1}] + \lambda_1 \phi \kappa \phi \frac{1}{m} \sum_{j=2}^m \phi^{j-2} \mathbf{E}[h_{t-j} \tau_{t-j}]. \quad (39) \end{aligned}$$

If  $r_t$  is covariance stationary, then

$$\left( \lambda_1 \frac{1}{m} \phi_\kappa + \lambda_2 \phi \right) - \lambda_1 \phi_\kappa \frac{1}{m} \sum_{j=2}^m \phi^{j-1} < 1. \quad (40)$$

It then follows that  $\mathbf{Var}[r_t]$  is given by equation (8). ■

**Proof of Corollary 1.** Let  $m = 1$  and  $\gamma = 0$ . If  $r_t$  is covariance stationary, we have

$$\begin{aligned} \mathbf{E}[\tau_t h_t] &= \mathbf{E} \left[ \left( \lambda_0 + \lambda_1 \frac{r_{t-1}^2}{h_{t-1}} + \lambda_2 \tau_{t-1} \right) \left( (1 - \alpha - \beta) + \alpha \frac{r_{t-1}^2}{\tau_{t-1}} + \beta h_{t-1} \right) \right] \\ &= \lambda_0(1 - \alpha - \beta) + \lambda_0 \alpha \mathbf{E}\left[\frac{r_{t-1}^2}{\tau_{t-1}}\right] + \lambda_0 \beta \mathbf{E}[h_{t-1}] \\ &\quad + \lambda_1(1 - \alpha - \beta) \mathbf{E}\left[\frac{r_{t-1}^2}{h_{t-1}}\right] + \alpha \lambda_1 \mathbf{E}\left[\frac{r_{t-1}^4}{h_{t-1} \tau_{t-1}}\right] + \lambda_1 \beta \mathbf{E}[r_{t-1}^2] \\ &\quad + \lambda_2(1 - \alpha - \beta) \mathbf{E}[\tau_{t-1}] + \alpha \lambda_2 \mathbf{E}[r_{t-1}^2] + \beta \lambda_2 \mathbf{E}[\tau_{t-1} h_{t-1}] \\ &= \lambda_0 + (1 - \alpha - \beta)(\lambda_1 + \lambda_2) \mathbf{E}[\tau_t] + [\lambda_1(\alpha \kappa + \beta) + \lambda_2(\alpha + \beta)] \mathbf{E}[\tau_t h_t], \end{aligned} \quad (41)$$

where we use that  $\mathbf{E}[\frac{r_{t-1}^2}{\tau_{t-1}}] = \mathbf{E}[h_{t-1} Z_{t-1}^2] = \mathbf{E}[h_t] = 1$ . Further,  $\mathbf{E}[\frac{r_{t-1}^2}{h_{t-1}}] = \mathbf{E}[\tau_{t-1} Z_{t-1}^2] = \mathbf{E}[\tau_{t-1}] = \mathbf{E}[\tau_t]$ . Finally,  $\mathbf{E}[\frac{r_{t-1}^4}{h_{t-1} \tau_{t-1}}] = \mathbf{E}[h_{t-1} \tau_{t-1} Z_{t-1}^4] = \kappa \mathbf{E}[h_t \tau_t]$ , and  $\mathbf{E}[r_{t-1}^2] = \mathbf{E}[h_t \tau_t]$ . If  $\lambda_1(\alpha \kappa + \beta) + \lambda_2(\alpha + \beta) < 1$ , we obtain equation (9).

Next, we show that  $\mathbf{E}[h_t \tau_t]$  has a finite variance. Note that the restrictions  $\alpha + \beta < 1$  and  $\lambda_1 + \lambda_2 < 1$  imply that the conditions for strict stationarity of  $h_t$  and  $\tau_t$  are satisfied. The limiting processes are given by

$$\bar{h}_t = (1 - \alpha - \beta) \left( 1 + \sum_{i=1}^{\infty} \prod_{j=1}^i (\alpha Z_{t-j}^2 + \beta) \right) \quad (42)$$

and

$$\bar{\tau}_t = \lambda_0 \left( 1 + \sum_{k=1}^{\infty} \prod_{j=1}^k (\lambda_1 Z_{t-j}^2 + \lambda_2) \right). \quad (43)$$

Hence, it is sufficient to show that the variance of  $r_t = \sqrt{\bar{h}_t \bar{\tau}_t} Z_t$  is finite.

Note that  $\mathbf{E}[\alpha Z_{t-j}^2 + \beta] = \alpha + \beta$  and  $\mathbf{E}[\lambda_1 Z_{t-k}^2 + \lambda_2] = \lambda_1 + \lambda_2$ . Moreover,  $\mathbf{E}[(\alpha Z_{t-j}^2 + \beta)(\lambda_1 Z_{t-k}^2 + \lambda_2)] = (\alpha + \beta)(\lambda_1 + \lambda_2)$  for  $j \neq k$  but  $\mathbf{E}[(\alpha Z_{t-j}^2 + \beta)(\lambda_1 Z_{t-j}^2 + \lambda_2)] = \lambda_1(\alpha \kappa + \beta) + \lambda_2(\alpha + \beta)$ . Next, define

$$\psi_i = \prod_{j=1}^i (\alpha Z_{t-j}^2 + \beta) \quad \text{and} \quad \delta_k = \prod_{j=1}^k (\lambda_1 Z_{t-j}^2 + \lambda_2). \quad (44)$$

It follows that

$$\mathbf{E}[\psi_i \delta_k] = [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]^k (\alpha + \beta)^{i-k} \text{ for } k \leq i \quad (45)$$

and

$$\mathbf{E}[\psi_i \delta_k] = [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]^i (\lambda_1 + \lambda_2)^{k-i} \text{ for } i < k. \quad (46)$$

Thus, we can write

$$\mathbf{E}[h_t \tau_t] = \lambda_0(1 - \alpha - \beta) \mathbf{E} \left[ \left( \sum_{i=0}^{\infty} \psi_i \right) \left( \sum_{k=0}^{\infty} \delta_k \right) \right], \quad (47)$$

with  $\psi_0 = 1$  and  $\delta_0 = 1$ . The last term can be treated as follows

$$\begin{aligned} \mathbf{E} \left( \sum_{i=0}^{\infty} \psi_i \right) \left( \sum_{k=0}^{\infty} \delta_k \right) &= \sum_{k=0}^{\infty} \sum_{i=k}^{\infty} \mathbf{E}[\psi_i \delta_k] + \sum_{i=0}^{\infty} \sum_{k=i+1}^{\infty} \mathbf{E}[\psi_i \delta_k] \\ &= \sum_{k=0}^{\infty} [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]^k \sum_{i=0}^{\infty} (\alpha + \beta)^i \\ &\quad + \sum_{i=0}^{\infty} [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]^i (\lambda_1 + \lambda_2) \sum_{k=0}^{\infty} (\lambda_1 + \lambda_2)^k \\ &= \frac{1}{[1 - (\alpha + \beta)][1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]]} \\ &\quad + \frac{(\lambda_1 + \lambda_2)}{[1 - (\lambda_1 + \lambda_2)][1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]]}. \end{aligned} \quad (48)$$

In summary, we get

$$\begin{aligned} \mathbf{E}[h_t \tau_t] &= \frac{\lambda_0}{(1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)])} \\ &\quad + \frac{\lambda_0(1 - \alpha - \beta)(\lambda_1 + \lambda_2)}{[1 - (\lambda_1 + \lambda_2)][1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]]} \\ &= \frac{\lambda_0 + (1 - \alpha - \beta)(\lambda_1 + \lambda_2) \frac{\lambda_0}{[1 - (\lambda_1 + \lambda_2)]}}{1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]}. \end{aligned} \quad (49)$$

■

**Proof of Theorem 2.** First, recall that  $\mathbf{E}[h_{t+1} \tau_{t+1} | \mathcal{F}_t] = h_{t+1} \tau_{t+1}$ . For  $2 \leq s \leq m$ , we obtain the forecasts as

$$\begin{aligned} \mathbf{E}[h_{t+s} \tau_{t+s} | \mathcal{F}_t] &= \mathbf{E}[(1 - \phi) + ((\alpha + \gamma \mathbf{1}_{\{Z_{t+s-1} < 0\}}) Z_{t+s-1}^2 + \beta) h_{t+s-1}] \\ &\quad \times \left( \lambda_0 + \lambda_1 \frac{1}{m} \sum_{j=1}^{s-1} Z_{t+s-j}^2 \tau_{t+s-j} + \lambda_1 \frac{1}{m} \sum_{j=s}^m \frac{r_{t+s-j}^2}{h_{t+s-j}} + \lambda_2 \tau_{t+s-1} \right) \Big| \mathcal{F}_t \\ &= (1 - \phi) \mathbf{E}[\tau_{t+s} | \mathcal{F}_t] + \lambda_0 \phi \mathbf{E}[h_{t+s-1} | \mathcal{F}_t] + \left( \lambda_1 \frac{1}{m} \phi \kappa + \lambda_2 \phi \right) \mathbf{E}[h_{t+s-1} \tau_{t+s-1} | \mathcal{F}_t] \\ &\quad + \lambda_1 \phi \mathbf{E}[h_{t+s-1} | \mathcal{F}_t] \frac{1}{m} \sum_{j=s}^m \frac{r_{t+s-j}^2}{h_{t+s-j}} + \lambda_1 \phi \frac{1}{m} \sum_{j=2}^{s-1} \mathbf{E}[h_{t+s-1} Z_{t+s-j}^2 \tau_{t+s-j} | \mathcal{F}_t]. \end{aligned} \quad (50)$$

A conditional version of equation (38) is given by

$$\begin{aligned}\mathbf{E}[h_{t+s-1}Z_{t+s-j}^2\tau_{t+s-j}|\mathcal{F}_t] &= (1-\phi)\mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t]\left(1+\sum_{k=1}^{j-2}\phi^k\right) \\ &\quad +\phi^{j-2}\phi_\kappa\mathbf{E}[h_{t+s-j}\tau_{t+s-j}|\mathcal{F}_t].\end{aligned}$$

Plugging this into equation (50) leads to

$$\begin{aligned}\mathbf{E}[h_{t+s}\tau_{t+s}|\mathcal{F}_t] &= (1-\phi)\mathbf{E}[\tau_{t+s}|\mathcal{F}_t]+\lambda_0\phi\mathbf{E}[h_{t+s-1}|\mathcal{F}_t] \\ &\quad +\left(\lambda_1\frac{1}{m}\phi_\kappa+\lambda_2\phi\right)\mathbf{E}[h_{t+s-1}\tau_{t+s-1}|\mathcal{F}_t] \\ &\quad +\lambda_1\phi\mathbf{E}[h_{t+s-1}|\mathcal{F}_t]\frac{1}{m}\sum_{j=s}^m\frac{r_{t+s-j}^2}{h_{t+s-j}} \\ &\quad +(1-\phi)\lambda_1\phi\frac{1}{m}\sum_{j=2}^{s-1}\mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t]\left(1+\sum_{k=1}^{j-2}\phi^k\right) \\ &\quad +\lambda_1\phi_\kappa\phi\frac{1}{m}\sum_{j=2}^{s-1}\phi^{j-2}\mathbf{E}[h_{t+s-j}\tau_{t+s-j}|\mathcal{F}_t].\end{aligned}\tag{51}$$

This allows to recursively calculate  $\mathbf{E}[h_{t+s}\tau_{t+s}|\mathcal{F}_t]$ , for  $s = 1, \dots, m$ .

Similarly, for  $s > m$  the conditional expectations are given by

$$\begin{aligned}\mathbf{E}[h_{t+s}\tau_{t+s}|\mathcal{F}_t] &= \mathbf{E}[(1-\phi)+((\alpha+\gamma\mathbf{1}_{\{Z_{t-s-1}<0\}})Z_{t+s-1}^2+\beta)h_{t+s-1}) \\ &\quad \times\left(\lambda_0+\lambda_1\frac{1}{m}\sum_{j=1}^mZ_{t+s-j}^2\tau_{t+s-j}+\lambda_2\tau_{t+s-1}\right)|\mathcal{F}_t] \\ &= (1-\phi)\mathbf{E}[\tau_{t+s}|\mathcal{F}_t]+\lambda_0\phi\mathbf{E}[h_{t+s-1}|\mathcal{F}_t] \\ &\quad +\left(\lambda_1\frac{1}{m}\phi_\kappa+\lambda_2\phi\right)\mathbf{E}[h_{t+s-1}\tau_{t+s-1}|\mathcal{F}_t] \\ &\quad +\lambda_1\phi\frac{1}{m}\sum_{j=2}^m\mathbf{E}[h_{t+s-1}Z_{t+s-j}^2\tau_{t+s-j}|\mathcal{F}_t].\end{aligned}\tag{52}$$

Again, using equation (38) the forecasts can be recursively calculated as

$$\begin{aligned}\mathbf{E}[h_{t+s}\tau_{t+s}|\mathcal{F}_t] &= (1-\phi)\mathbf{E}[\tau_{t+s}|\mathcal{F}_t]+\lambda_0\phi\mathbf{E}[h_{t+s-1}|\mathcal{F}_t] \\ &\quad +\left(\lambda_1\frac{1}{m}\phi_\kappa+\lambda_2\phi\right)\mathbf{E}[h_{t+s-1}\tau_{t+s-1}|\mathcal{F}_t] \\ &\quad +(1-\phi)\lambda_1\phi\frac{1}{m}\sum_{j=2}^m\mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t]\left(1+\sum_{k=1}^{j-2}\phi^k\right) \\ &\quad +\lambda_1\phi_\kappa\phi\frac{1}{m}\sum_{j=2}^m\phi^{j-2}\mathbf{E}[h_{t+s-j}\tau_{t+s-j}|\mathcal{F}_t].\end{aligned}\tag{53}$$

■

## C Additional Figures

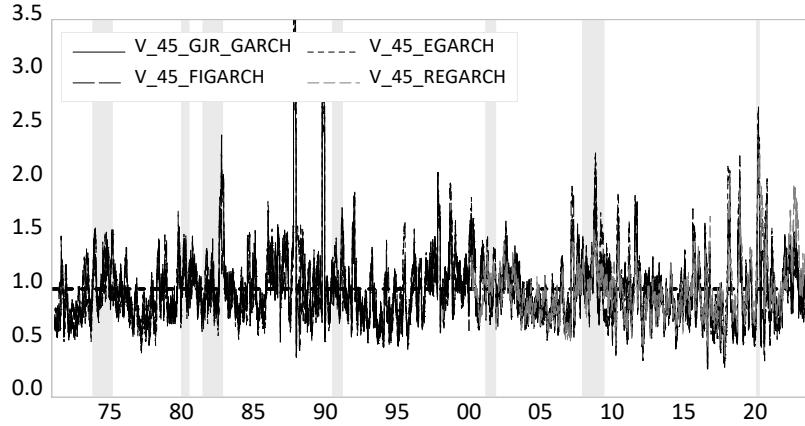


Figure A.1: A GJR-GARCH(1,1), an EGARCH(1,1), and a FIGARCH(1, $d$ ,1) are estimated for the S&P 500 using daily return data for the January 1971 to June 2023 period. A linear Realized GARCH(1,1) is estimated for the January 2000 to June 2023 period. The figure shows  $\tilde{V}_t^{(45)}$  for the GJR-GARCH, the EGARCH, the FIGARCH, and the Realized GARCH. Grey shaded areas represent NBER recession periods.

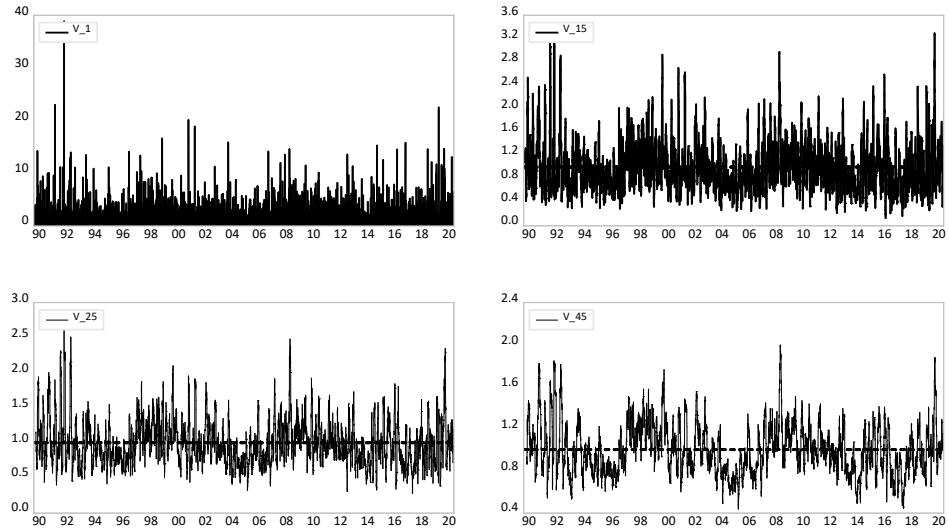


Figure A.2: A GJR-GARCH(1,1) is estimated using daily FTSE 100 return data for the January 1990 to June 2020 period. The figure shows  $\tilde{V}_t^{(m)}$  for  $m = 1$  (upper left panel),  $m = 15$  (upper right panel),  $m = 25$  (lower left panel), and  $m = 45$  (lower right panel).

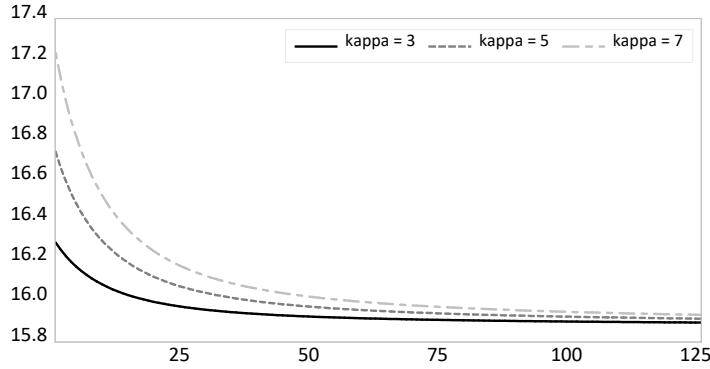


Figure A.3: The figure shows the annualized unconditional volatility,  $\sqrt{252 \cdot \text{Var}[r_t]}$ , of the MF2-GARCH-rw- $m$  with parameters  $\alpha = 0.02$ ,  $\gamma = 0.1$ ,  $\beta = 0.8$ ,  $\lambda_0 = 0.01$ ,  $\lambda_1 = 0.05$  and  $\lambda_2 = 0.94$  as a function of  $m$  for  $\kappa \in \{3, 5, 7\}$ .

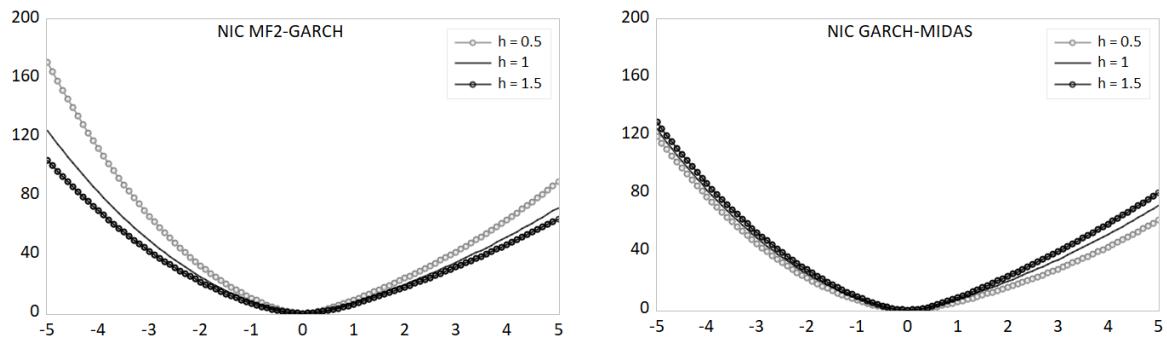


Figure A.4: The figure show the news impact functions for an MF2-GARCH-rw-1 (left panel) and a GARCH-MIDAS-RV (right panel) model with parameters as described in Section A.3. We fix  $\tau_t = \tau_t^{GM} = 1$ . The NICs are plotted as a function of the return,  $r_t$ , and for  $h_t \in \{0.5, 1, 1.5\}$ . As in Figure 2, the NICs are standardized and presented as annualized volatilities.

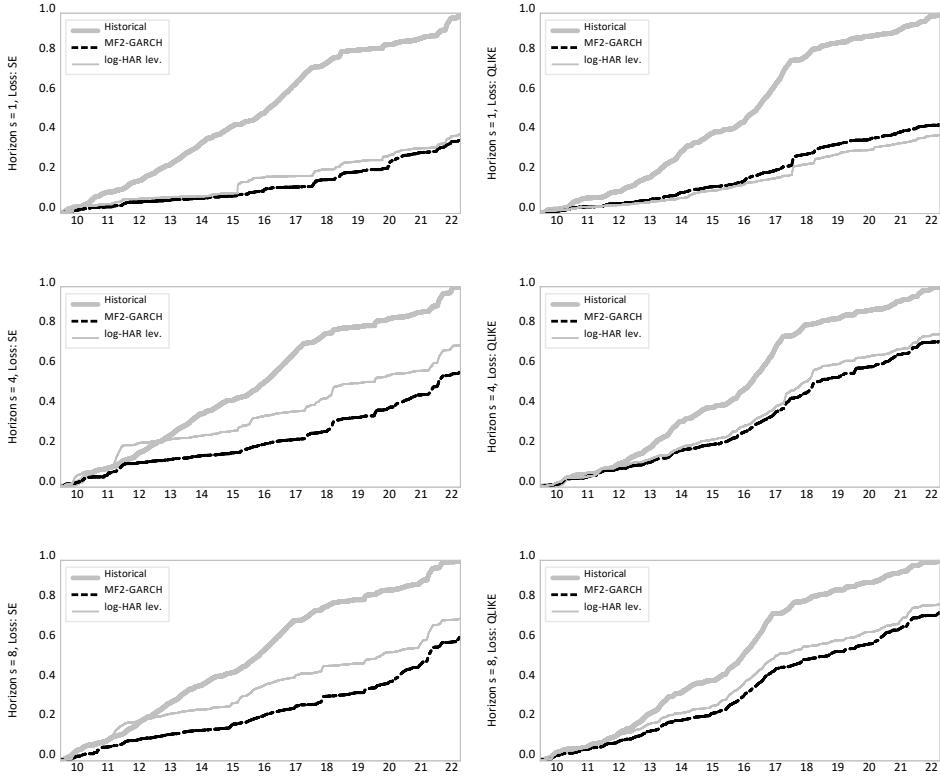


Figure A.5: Cumulated losses for forecast horizons of one month, four months and eight months for SE (left panels) and QLIKE (right panels) loss. All losses are normalized with the cumulated loss of the historical forecast at the end of the out-of-sample period.

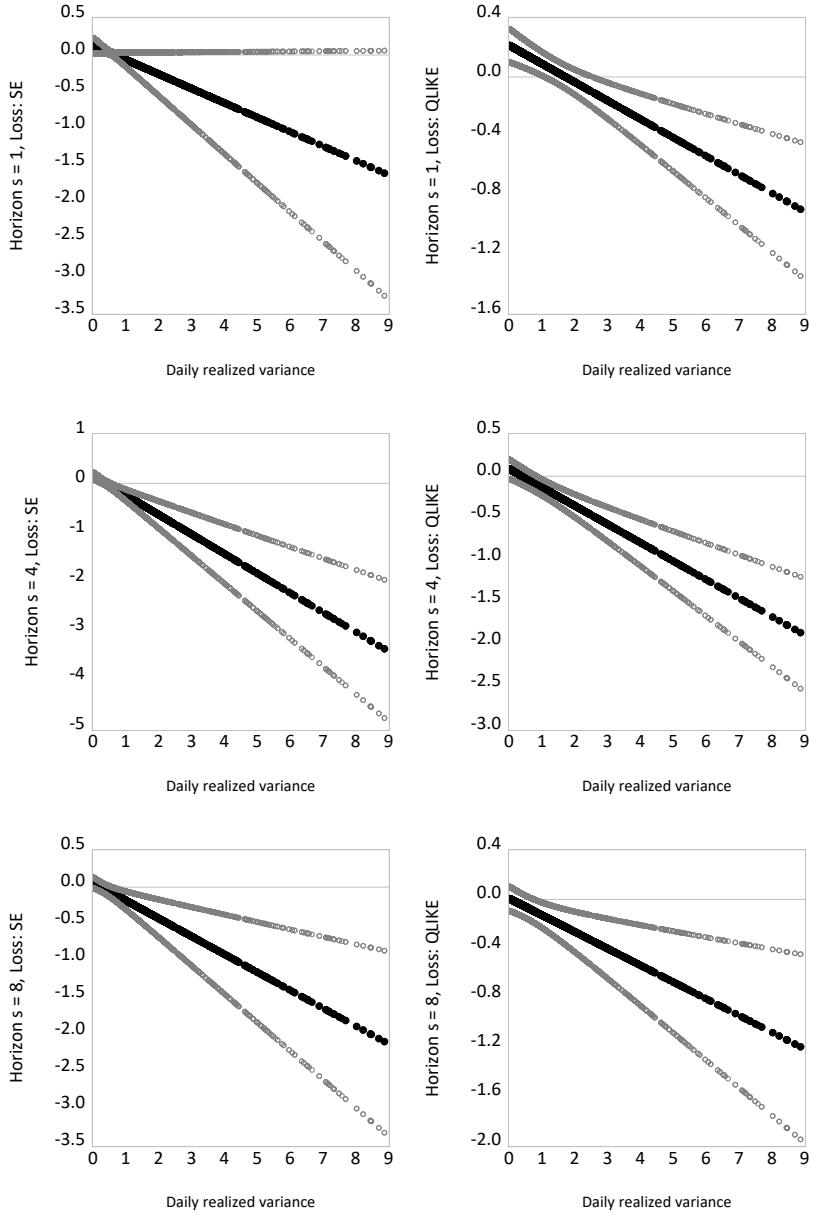


Figure A.6: The figure shows the results of a Giacomini and White (2006) test of conditional predictive ability. We compare the forecasts of the MF2-GARCH and the log-HAR by regressing the loss difference on a constant and the realized variance on the day the forecast was made. The figure shows the estimate of the conditional mean (and a 95% confidence interval) as a function of  $RV_t$ . Negative values for the predicted conditional mean of the loss difference imply that the MF2-GARCH outperforms the log-HAR. Forecast horizons are of one month (upper panels), four months (middle panels) and eight months (lower panels) for squared error (SE, left panels) and QLIKE (right panels) loss differences.

## D Additional Tables

Table A.1: Monte Carlo parameter estimates and standard errors.

	$\alpha$	$\gamma$	$\beta$	$\lambda_0$	$\lambda_1$	$\lambda_2$
true	0.02	0.10	0.8	0.02	0.05	0.94
<b>Panel A: <math>Z_t \sim \mathcal{N}(0, 1)</math></b>						
$m = 21, T = 5,000$						
Bias	-0.0018	0.0018	-0.0062	0.0051	0.0018	-0.0046
RMSE	{0.0107}	{0.0182}	{0.0387}	{0.0126}	{0.0161}	{0.0207}
ASE	(0.0117)	(0.0191)	(0.0426)	(0.0142)	(0.0180)	(0.0238)
$m = 21, T = 10,000$						
Bias	-0.0011	0.0009	-0.0015	0.0020	0.0003	-0.0014
RMSE	{0.0078}	{0.0135}	{0.0258}	{0.0074}	{0.0101}	{0.0126}
ASE	(0.0080)	(0.0132)	(0.0272)	(0.0076)	(0.0110)	(0.0138)
$m = 63, T = 5,000$						
Bias	-0.0025	0.0019	-0.0063	0.0193	0.0194	-0.0295
RMSE	{0.0107}	{0.0182}	{0.0361}	{0.0524}	{0.0705}	{0.0938}
ASE	(0.0117)	(0.0191)	(0.0386)	(0.0549)	(0.0805)	(0.1050)
$m = 63, T = 10,000$						
Bias	-0.0017	0.0011	-0.0024	0.0095	0.0112	-0.0161
RMSE	{0.0077}	{0.0134}	{0.0247}	{0.0325}	{0.0512}	{0.0665}
ASE	(0.0082)	(0.0132)	(0.0255)	(0.0356)	(0.0589)	(0.0756)
<b>Panel B: <math>Z_t \sim t(5)</math></b>						
$m = 21, T = 5,000$						
Bias	-0.0009	0.0020	-0.0110	0.0071	0.0048	-0.0090
RMSE	{0.0148}	{0.0323}	{0.0610}	{0.0267}	{0.0262}	{0.0367}
ASE	(0.0174)	(0.0322)	(0.0819)	(0.0293)	(0.0382)	(0.0503)
$m = 21, T = 10,000$						
Bias	-0.0011	0.0014	-0.0048	0.0028	0.0017	-0.0034
RMSE	{0.0106}	{0.0216}	{0.0402}	{0.0096}	{0.0149}	{0.0184}
ASE	(0.0110)	(0.0219)	(0.0414)	(0.0101)	(0.0167)	(0.0203)
$m = 63, T = 5,000$						
Bias	-0.0021	0.0016	-0.0093	0.0236	0.0267	-0.0399
RMSE	{0.0144}	{0.0315}	{0.0544}	{0.0599}	{0.0874}	{0.1149}
ASE	(0.0188)	(0.0320)	(0.0606)	(0.0658)	(0.0948)	(0.1244)
$m = 63, T = 10,000$						
Bias	-0.0017	0.0014	-0.0051	0.0113	0.0175	-0.0235
RMSE	{0.0104}	{0.0215}	{0.0365}	{0.0329}	{0.0749}	{0.0891}
ASE	(0.0110)	(0.0219)	(0.0382)	(0.0622)	(0.1088)	(0.1380)

**Notes:** The table shows the Bias, the root mean squared error (RMSE) and the asymptotic standard error (ASE) for each parameter of the MF2-GARCH-rw- $m$  across  $R = 1,000$  Monte Carlo replications.

Table A.2: Monte Carlo parameter estimates and forecast losses for misspecified models.

	$\alpha$	$\gamma$	$\beta$	$\lambda_0$	$\lambda_1$	$\omega$	SE	QLIKE
<b>Panel A: <math>m = 21</math></b>								
GJR-GARCH	0.0356	0.0468	0.9233	2.0633			6.4933	6.2025
GARCH-MIDAS-RV	0.0282	0.0858	0.8372	0.6820	0.6468	43.7895	3.3679	3.4614
<b>Panel B: <math>m = 63</math></b>								
GJR-GARCH	0.0279	0.0583	0.9030	2.0086			9.0311	9.8326
GARCH-MIDAS-RV	0.0209	0.1003	0.7968	0.6290	0.6791	14.0709	1.8025	1.8887

**Notes:** The table shows the mean parameter estimates across  $R = 1,000$  Monte Carlo replications.  $Z_t$  is standard normal and  $T = 5,000$ . The true data generating process is an MF2-GARCH-rw- $m$  with parameters as in Table A.1. For the GARCH-MIDAS-RV,  $\lambda_0$  corresponds to  $c$  and we set  $K = 252$ . The column SE (QLIKE) shows the ratio of the average SE (QLIKE) loss of the model in the respective row and the average SE (QLIKE) of the MF2-GARCH.

Table A.3: Determinants of long-term volatility.

	dependent variable: $\tau_{n,t} - \tau_{n,t-1}$		
	(1)	(2)	(3)
Overall $_t$	0.515*** (0.051)		
Macro News $_t$		0.267* (0.144)	
Monetary Policy $_t$		0.205** (0.104)	
Fiscal Policy $_t$		0.096 (0.096)	
Inflation $_t$			0.212** (0.102)
Interest Rates $_t$			0.317*** (0.115)
adj. $R^2$	0.224	0.242	0.219

**Notes:** Monthly regressions are for the January 1985 to June 2023 period.  $\tau_{n,t}$  is taken from the MF2-GARCH-rw-63 model for daily S&P 500 returns from Table 2. The EMV-trackers Overall $_t$ , Macro News $_t$ , Monetary Policy $_t$ , Fiscal Policy $_t$ , Inflation $_t$ , Interest Rates $_t$  are from Baker et al. (2021) (see Section A.5). The numbers in the columns show the parameter estimates and HAC standard errors in parentheses. Each tracker has been standardized so that the estimated coefficients measure the effect of a one standard deviation change in the respective tracker. \*\*\*, \*\* and \* indicate significance at the 1%, 5% and 10% level.

Table A.4: Out-of-sample forecasting: S&P 500 – high volatility regime.

	1m	2m	3m	4m	5m	6m	7m	8m
<b>Panel A: relative RMSE</b>								
Historical Forecast	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
GJR-GARCH	1.428	1.454	1.098	1.100	1.001	0.943	<b>0.898</b>	<b>0.856</b>
MF2-GARCH-rw	<b>0.856</b>	<b>0.896</b>	<b>0.881</b>	<b>0.878</b>	<b>0.869</b>	<b>0.885</b>	0.911	0.902
log-HAR lev.	0.957	1.090	1.131	1.156	1.142	1.137	1.109	1.083
GARCH-MIDAS $m = 21$	1.542	1.611	1.917	1.981	1.935	2.072	2.147	2.210
GARCH-MIDAS $m = 63$	1.454	2.066	2.082	2.141	2.171	2.293	2.415	2.413
<b>Panel B: relative QLIKE</b>								
Historical Forecast	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
GJR-GARCH	0.848	0.986	0.981	0.979	0.946	0.948	0.933	0.913
MF2-GARCH-rw	<b>0.738</b>	<b>0.874</b>	<b>0.882</b>	<b>0.880</b>	<b>0.887</b>	<b>0.905</b>	<b>0.926</b>	<b>0.923</b>
log-HAR lev.	0.756	0.950	0.985	1.015	0.998	1.011	1.004	0.995
GARCH-MIDAS $m = 21$	0.836	0.996	1.060	1.052	1.061	1.087	1.118	1.132
GARCH-MIDAS $m = 63$	0.837	1.015	1.043	1.042	1.061	1.093	1.124	1.139

**Notes:** See notes of Table 4. A day falls in a high volatility regime if the realized variance on that day is above the 70th percentile of the distribution of realized variances.

## References (cited only in the Appendix)

- Amado, C., Silvennoinen, A., Teräsvirta, T., 2019. Models with multiplicative decomposition of conditional variances and correlations. In J. Chevallier, S. Goutte, D. Guerreiro, S. Saglio and B. Sanhaji (eds.) *Financial Mathematics, Volatility and Covariance Modelling*, Volume 2. Routledge.
- Amendola, A., Candila, V., Cipollini, F., Gallo, G.M., 2020. Doubly multiplicative error models with long- and short-run components. Available at:  
<https://arxiv.org/abs/2006.03458>
- Brownlees, C., Cipollini, F., Gallo, G.M., 2011. Intra-daily volume modeling and prediction for algorithmic trading. *Journal of Financial Econometrics*, 9, 489-518.
- Borup, D., Jakobsen, J.S., 2019. Capturing volatility persistence: a dynamically complete realized EGARCH-MIDAS model. *Quantitative Finance*, 19, 1839-1855.
- Francq, C., Zakoïan, J.-M., 2007. Quasi-maximum likelihood in GARCH processes when some coefficients are equal to zero. *Stochastic Processes and their Applications* 117, 1265-1284.
- Francq, C., Zakoïan, J.-M., 2009. Testing the nullity of GARCH coefficients: Correction of the standard tests and relative efficiency comparisons. *Journal of the American Statistical Association* 104, 313-324.
- Francq, C., Kandji, B.M., Zakoian, J.-M., 2023. Infercence on GARCH-MIDAS models without any small-order moment. *Econometric Theory*. Forthcoming. DOI: <https://doi.org/10.1017/S0266466623000142>
- Noureddin, D., 2022. Volatility prediction using a realized-measure-based component model. *Journal of Financial Econometrics*, 20, 76-104.