Network

Rimini Centre for Economic Analysis Working Paper Series

wp 21-05

Modelling Volatility Cycles: The $(MF)^2$ GARCH Model

**Christian Conrad
Robert F. Engle**

RCEA aims to further independent, advanced research in Economics, Econometrics and related fields and to promote contact between economists, econometricians and scientists from other fields all over the world. Research at RCEA is conducted to enlighten scientific and public debate on economic issues, and not to advance any economic, political or social agenda. In this respect, RCEA mission is particularly concerned with raising awareness and stimulating discussion on the changes required to make capitalism sustainable along its economic, environmental, human, social and political dimensions

Modelling Volatility Cycles: The $(MF)^2$ GARCH Model*

Christian Conrad[†] and Robert F. Engle[‡]

First draft: January 30, 2019

This draft: March 14, 2021

Abstract

We suggest a multiplicative factor multi frequency component GARCH model which exploits the empirical fact that the daily standardized forecast errors of standard GARCH models behave counter-cyclical when averaged at a lower frequency. For the new model, we derive the unconditional variance of the returns, the news impact function and multi-step-ahead volatility forecasts. We apply the model to the S&P 500, the FTSE 100 and the Hang Seng Index. We show that the long-term component of stock market volatility is driven by news about the macroeconomic outlook and monetary policy as well as policy-related news. The new component model significantly outperforms the nested one-component (GJR) GARCH and several HAR-type models in terms of out-of-sample forecasting.

Keywords: Volatility forecasting, long- and short-term volatility, mixed frequency data, volatility cycles.

JEL Classification: C53, C58, G12

*We would like to thank Jörg Breitung, Zeno Enders, Jean-David Fermanian, Christian Francq, Christian Gourieroux, Onno Kleen, Enno Mammen, Jean Michel Zakoïan as well as seminar participants at the University of Hannover (January 2019), the University of Bonn (July 2019), CREST (January 2020) and IÉSEG Lille (January 2021) for helpful comments and suggestions.

[†]Corresponding author: Heidelberg University, Department of Economics and KOF Swiss Economic Institute, Email: christian.conrad@awi.uni-heidelberg.de.

[‡]New York University, Stern School of Business, Email: rengle@stern.nyu.edu.

1 Introduction

We document a new empirical fact for the volatility forecast errors of standard GARCH models. While the daily standardized forecast errors are essentially unpredictable, the standardized forecast errors become predictable when averaged at a lower frequency. Partly, this predictability arises because the GARCH-based forecasts tend to overestimate volatility during economic expansions and to underestimate volatility during recessions. Hence, the suitably averaged standardized forecast errors behave counter-cyclical.

Motivated by this empirical fact, we propose a new multiplicative volatility model. The model consists of a short-term GARCH component and a long-term component that is specified as a multiplicative error model (MEM) for the past forecast errors of the GARCH component. The long-term component exploits the predictability in the averaged standardized forecast errors of the GARCH component. Intuitively, the long-term component scales the GARCH volatility forecast up/down if the GARCH forecasts have under-/overestimated volatility in the recent past. Because we allow the long-term component to either evolve at the same frequency as the short-term component or at a lower frequency, we will refer to the new model as multiplicative factor multi frequency GARCH ((MF)² GARCH). Hence, the new specification belongs to the class of MIXed frequency Data Sampling (MIDAS) models.

The (MF)² GARCH is related to several GARCH-type specifications that model the conditional variance of stock returns as consisting of 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 on low-frequency macroeconomic or financial variables.

The properties of the (MF)² GARCH model clearly distinguish it from those 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 (MF)² GARCH explicitly specifies the long-term component as a function of the past forecast errors of the short-term component. Second, because the true economic drivers of long-term volatility are unknown and are likely to vary over time and potentially with the desired forecast horizon (see Conrad and Kleen, 2020), it is unlikely that a researcher

is able to correctly specify the long-term component in the GARCH-MIDAS in real time. Our specification avoids this problem and, instead, is based on a simple MEM equation for the long-term component. Third, our model is dynamically complete in the sense that it allows us to construct multi-step-ahead volatility forecasts. Again, this is in contrast to other multiplicative volatility models which either require strong assumptions (e.g. the zero slope Spline-GARCH) or forecasts of the explanatory variables (e.g. the GARCH-MIDAS-X in Conrad and Loch, 2015) for computing multi-step-ahead forecasts.

We obtain the following theoretical results: First, we derive the unconditional variance of the daily returns when the data generating process is an $(MF)^2$ GARCH. In contrast to simple GARCH models, the unconditional variance does not only depend 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 in the $(MF)^2$ GARCH the responsiveness to news changes with the level of volatility which is another important feature that distinguishes the $(MF)^2$ GARCH from other models in the GARCH family. Specifically, the 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 the conditional volatility. Our results show that the forecasts are much more flexible than forecasts from the nested GARCH model. The forecasts reflect both the current stance of the conditional variance as well as the prevailing volatility regime. In the short term, the forecast will approach the current level of the long-term component before it converges to the unconditional volatility in the long run. Finally, the $(MF)^2$ GARCH is straightforward to estimate by quasi-maximum likelihood.

Our empirical results provide strong support for the existence of a time-varying long-term component in the volatility of daily S&P 500, FTSE 100 and Hang Seng Index (HSI) returns. As expected, the long-term volatility component of the S&P 500 is closely related to U.S. macroeconomic conditions and volatility cycles in international stock markets are strongly aligned. We show that news about the macroeconomic outlook, monetary policy and policy-related news are important drivers of the long-term volatility component of the S&P 500. In addition, we document a new type of asymmetry for the short-run effect of major news events that are related to the macroeconomic outlook, corporate earnings and monetary policy: Good/bad news tends to decrease/increase short-term volatility, while long-term volatility increases in response to both types of news. As a result, the

news effect on the conditional volatility is also asymmetric: While good news reduces volatility for a few days, bad news increases volatility more persistently. In summary, our results 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), but also highlight that a model with a single explanatory variable in the long-term component will not fully capture the complex and time-varying relationship. Finally, we evaluate the out-of-sample forecast performance of the (MF)² GARCH against the nested GARCH and various HAR-type models (HAR, log-HAR and log-HAR with leverage). It turns out that the (MF)² GARCH clearly outperforms the nested GARCH at all forecast horizons. The (MF)² GARCH significantly beats the HAR-type models when the forecast horizon is beyond a quarter.

The paper is organized as follows. In Section 2 we present the empirical fact that volatility forecast errors of simple GARCH models are predictable. We introduce the model and discuss its properties in Section 3. The empirical results are presented in Section 4 and Section 5 concludes.

2 A New Empirical Fact of Volatility Forecast Errors

In this section, we provide evidence for a new empirical fact of volatility forecast errors: When aggregated to a monthly, quarterly or semi-annual frequency the standardized forecast errors of one-component GARCH models are predictable.

Throughout the paper, we use 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 represents 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 $r_{i,t}$.

In empirical research, the conditional heteroskedasticity in daily stock returns is commonly modeled as a GARCH process. Daily stock returns are written as

$$r_{i,t} = \sqrt{\tilde{h}_{i,t}} \zeta_{i,t}, \quad (1)$$

where $\tilde{h}_{i,t}$ denotes the conditional variance and the $\zeta_{i,t}$ are assumed to be *i.i.d.* with mean and variance equal to zero and one, respectively. For illustration, we estimate a GARCH

specification for a long time series of daily S&P 500 log-returns covering the January 1971 to June 2020 period.¹ Assuming that $\tilde{h}_{i,t}$ follows a GJR GARCH(1,1) (see Glosten et al., 1993), we obtain the following result:

$$\tilde{h}_{i,t} = 0.017 + \underset{(0.003)}{(0.022 + 0.110 \mathbf{1}_{\{r_{i-1,t} < 0\}})} r_{i-1,t}^2 + \underset{(0.009)}{0.905} \tilde{h}_{i-1,t}, \quad (2)$$

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

Several test statistics have been proposed for checking the adequacy of a GARCH specification. For example, Bollerslev (1986), 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. Under the alternative, GARCH specifications that include, for example, higher order terms are considered.

An alternative approach is to check whether equation (1) is misspecified in the sense that $\zeta_{i,t} = \sqrt{\tau_t} Z_{i,t}$, where the $Z_{i,t}$ are *i.i.d.* and τ_t represents an omitted multiplicative volatility component. If $n = 1$, the long-term component evolves at the same frequency as the daily returns. A low-frequency long-term component is obtained for $n > 1$. The daily returns, $r_{i,t}$, can be either stationary or non-stationary. For example, in the Spline-GARCH model it is assumed that $n = 1$ 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, $r_{i,t}/\sqrt{\tau_t}$, are assumed to follow a stationary GARCH process. LM tests for an omitted τ_t component have been proposed in Lundbergh and Teräsvirta (2002) and Amado and Teräsvirta (2017) for the case that $n = 1$ and in Conrad and Schienle (2020) for the case that $n \geq 1$. In addition, Conrad and Schienle (2020) allow for exogenous explanatory variables in the long-term component.

Both types of tests have in common that under the null hypothesis of a correctly specified model the squared standardized residuals, $\zeta_{i,t}^2 = r_{i,t}^2/\tilde{h}_{i,t}$, should be *i.i.d.*. However, if the long-term component is smoothly changing over time, it is difficult to detect a misspecification by investigating the behavior of the daily $\zeta_{i,t}^2$ (see Conrad and Schienle, 2020). Instead, we focus on the averaged squared standardized residuals within low-frequency

¹The data will be introduced and discussed in more detail in Section 4.1. See also Table 1.

period t :

$$\tilde{V}_t = \frac{1}{n} \sum_{i=1}^n \frac{r_{i,t}^2}{\tilde{h}_{i,t}}. \quad (3)$$

Note that $r_{i,t}^2$ is a conditionally unbiased proxy of the true but unobservable conditional variance and $\tilde{h}_{i,t}$ is a one-step-ahead forecast for the same quantity. If the GARCH model is correctly specified, the standardized volatility forecast errors, $r_{i,t}^2/\tilde{h}_{i,t}$, have an expected value of one and a variance of two if $\zeta_{i,t}$ is Gaussian. Hence, as noted in Patton (2011), the variance of the standardized forecast errors does not depend on the level of volatility. We think of $\tilde{V}_t$ as a measure for the *average* forecast performance of the GARCH model within period t .² Again, if the GARCH model is correctly specified, the $\tilde{V}_t$ are *i.i.d.* with an expected value of one.

Figure 1 shows the averaged standardized forecast errors based on the conditional variances from the GJR GARCH(1,1) in equation (2) at a quarterly (left panel) and a semi-annual (right panel) frequency. As expected, $\tilde{V}_t$ fluctuates around the value of one. However, it is evident that there are cycles in the averaged forecast errors, i.e. $\tilde{V}_t$ is below/above one for extended time periods. The figure – which also shows recession periods as grey shaded areas – suggests that the movements in $\tilde{V}_t$ are inversely related to the business cycle. That is, the GARCH model tends to overestimate volatility during expansions and to underestimate it during recessions. Next, we formally test for predictability in $\tilde{V}_t$. When running a regression of $\tilde{V}_t$ on four lagged values, we can strongly reject the null hypothesis of no predictability at the monthly, quarterly and semi-annual frequency. Note the similarity between those regressions and the Engle (1982) test for ARCH effects as well as the regression-based test for an omitted long-term component in Conrad and Schienle (2020). In Table 2, we exemplarily focus on the quarterly frequency and compare the predictability of $\tilde{V}_t$ with the predictability of the realized variance.³ We compute realized variances based on squared daily returns, i.e. $\widetilde{RV}_t = \sum_{i=1}^n r_{i,t}^2$. In Columns (1) and (3), we run regressions of $\widetilde{RV}_t$ and $\tilde{V}_t$ on four lagged values of the respective dependent variable. While the realized variance depends significantly on its first lag only, lags two to four are significant for $\tilde{V}_t$. This shows that $\tilde{V}_t$ is more persistent than $\widetilde{RV}_t$. Next, we add the median expectation from the Survey of Professional Forecasters for GDP growth

²Note the relation between $\tilde{V}_t$ and the relative squared error loss of Bollerslev and Ghysels (1996).

³The results for the monthly and semi-annual frequency are omitted for reasons of brevity but available upon request.

in quarter t based on information from quarter $t - 1$ to the regression.⁴ In line with the previous literature (see, for example, Schwert, 1989, and Paye, 2012, Berger et al., 2020), we find that $\widetilde{RV}_t$ is only weakly related to GDP growth (significant at the 15% level) when controlling for lagged volatility. In contrast, there is a highly significant and negative effect of GDP growth on the standardized forecast errors. This confirms the visual impression from Figure 1 that the volatility forecast errors behave counter-cyclical.

In order to show that the pattern that we observed in Figure 1 is not only specific to the S&P 500 but holds more generally, we additionally consider a European and an Asian stock market: the FTSE 100 and the Hang Seng Index (HSI). Daily return data for those stock markets are available from March 1990 onwards. For details, see Section 4.1. As before, we estimate GJR GARCH models for the daily returns of each stock market. Figure 2 shows the averaged standardized forecast errors for the FTSE, the HSI as well as the S&P 500 at a quarterly and semi-annual frequency. Again, the figure shows that there are distinct cycles in the averaged forecast errors. In addition, Figure 2 reveals that there is considerable co-movement in forecast errors across stock markets. This visual impression is confirmed by the cross-correlations between the averaged volatility forecast errors. For example, the contemporaneous correlation between the standardized volatility forecast errors of the S&P 500 and the FTSE 100 is 0.63/0.68 at the quarterly/semi-annual frequency.⁵

In summary, the evidence suggests that the simple GJR GARCH is misspecified and that the misspecification is detectable when averaging the volatility forecast errors over longer time periods.

3 The (MF)² GARCH Model

In this section, we introduce two parameterizations of the (MF)² GARCH model. Both specifications feature multiplicative volatility components. In the first parametrization the short- and the long-term component evolve at the daily frequency (Section 3.1). In

⁴We use expectations data because they are available in real time.

⁵In addition, we investigated several other international stock markets and also U.S. industry portfolios. Again, we found strong evidence for predictability and co-movement in forecast errors. For further evidence on the counter-cyclical behavior of the volatility forecast errors of US industry portfolios see Conrad and Glas (2020).

the second parametrization, we allow for multiple frequencies, i.e. the short-term component evolves at the daily frequency, while the long-term component varies at a lower frequency (Section 3.2). Section 3.3 discusses quasi-maximum likelihood estimation and in Section 3.4 we propose several directions in which the (MF)² GARCH can be extended. In general and as motivated in Section 2, daily log-returns are defined as

$$r_{i,t} = \sigma_{i,t} Z_{i,t} = \sqrt{h_{i,t} \tau_t} Z_{i,t}. \quad (4)$$

We denote the information set on day i in period t by $\mathcal{F}_{i,t}$ and define $\mathcal{F}_t := \mathcal{F}_{n,t}$. Recall that $i = 1, \dots, n$. $\sigma_{i,t}^2$ denotes the conditional variance and the short- and long-term volatility components are given by $h_{i,t}$ and τ_t . We make the following assumption about the innovations $Z_{i,t}$.

Assumption 1. *Let $Z_{i,t}$ be i.i.d. with $\mathbf{E}[Z_{i,t}] = 0$ and $\mathbf{E}[Z_{i,t}^2] = 1$. Further, $Z_{i,t}^2$ is assumed to have a nondegenerate distribution and $\kappa = \mathbf{E}[Z_{i,t}^4] < \infty$.*

3.1 Daily long-term component

We now assume that $n = 1$, i.e. the low-frequency period consists of a single day. In this setting, the daily returns are given by $r_{1,t} = \sqrt{h_{1,t} \tau_t} Z_{1,t}$ and, hence, we can drop the first index and rewrite the returns as

$$r_t = \sigma_t Z_t = \sqrt{h_t \tau_t} Z_t. \quad (5)$$

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

$$h_t = (1 - \alpha - \beta) + \alpha \frac{r_{t-1}^2}{\tau_{t-1}} + \beta h_{t-1}. \quad (6)$$

Note that the driving variable in equation (6) is r_t^2 / τ_t . This distinguishes h_t from the daily conditional variance, $\tilde{h}_{i,t}$, in equation (2). Next, we introduce an assumption on the parameters of the short-term component.

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

Throughout the paper, for simplicity, we present all theoretical results for the symmetric GARCH(1,1). However, in the empirical application we also employ a GJR short-term

component that allows for asymmetry:

$$h_t = (1 - \alpha - \gamma/2 - \beta) + (\alpha + \gamma \mathbf{1}_{\{r_{t-1} < 0\}}) \frac{r_{t-1}^2}{\tau_{t-1}} + \beta h_{t-1}. \quad (7)$$

For the GJR short-term component Assumption 2 has to be adjusted accordingly. Theoretical results for a GJR short-term component can be found in Appendix B.

We now turn to the behavior of the scaled returns. We can either adjust returns by the long- or by the short-term component. If Assumptions 1 and 2 hold, then

$$\frac{r_t}{\sqrt{\tau_t}} = \sqrt{h_t} Z_t \quad (8)$$

follows a covariance stationary GARCH(1,1) with $\mathbf{E}[h_t Z_t^2] = \mathbf{E}[h_t] = 1$. It also follows from Assumptions 1 and 2 that $\mathbf{E}[\log(\alpha Z_t^2 + \beta)] < 0$ and, hence, $\sqrt{h_t} Z_t$ is strictly stationary and ergodic. If the GARCH(1,1) fully captures the conditional heteroskedasticity, then τ_t is equal to a constant and the multiplicative model reduces to a simple GARCH(1,1) for the daily returns. Next, consider the squared deGARCHed returns:

$$V_t = \frac{r_t^2}{h_t} = \tau_t Z_t^2. \quad (9)$$

If the GARCH component fully captures the conditional heteroskedasticity (i.e. τ_t is constant), then by Assumption 1 the V_t are *i.i.d.*. However, if τ_t is time-varying, there will be autocorrelation in V_t . 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 . We parameterize the long-term component as

$$\tau_t = \lambda_0 + \lambda_1 V_{t-1}^{(m)} + \lambda_2 \tau_{t-1}, \quad (10)$$

where

$$V_{t-1}^{(m)} = \frac{1}{m} \sum_{j=1}^m V_{t-j} = \frac{1}{m} \sum_{j=1}^m \frac{r_{t-j}^2}{h_{t-j}}. \quad (11)$$

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 average forecast performance of the GARCH component over the previous m days. Note that equation (10) can be written as a MEM(1, m) with the restriction that the m ‘ARCH’ coefficients are given by λ_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 $\lambda_1 + \lambda_2 < 1$. Next, we make an assumption regarding the parameters of the long-term component.

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 (5), (6) and (10) as $(MF)^2$ GARCH-rw- m , where ‘rw- m ’ stands for *rolling window* of length m . Whenever we replace the short-term GARCH component with a GJR GARCH component, we refer to the model as $(MF)^2$ GJR-rw- m .

Remark 1. *Note that there is trade-off in the choice of m . When m is ‘small,’ $V_{t-1}^{(m)}$ will be a noisy measure of the past forecast performance. On the other hand, when m gets ‘large,’ $V_{t-1}^{(m)}$ becomes more and more stale.*

Unconditional variance: 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 $(MF)^2$ GARCH-rw- m process, $(r_t)_{t \in \mathbb{Z}}$, is covariance stationary, then*

$$\left(\lambda_1 \frac{1}{m} (\alpha \kappa + \beta) + \lambda_2 (\alpha + \beta) \right) + \lambda_1 (\alpha \kappa + \beta) \frac{1}{m} \sum_{j=2}^m (\alpha + \beta)^{j-1} < 1. \quad (12)$$

The unconditional variance of the daily returns is given by

$$\mathbf{Var}[r_t] = \frac{\lambda_0 + \frac{\lambda_0}{1 - \lambda_1 - \lambda_2} (1 - \alpha - \beta) (\lambda_1 + \lambda_2) + \Delta_m}{1 - \left(\lambda_1 \frac{1}{m} (\alpha \kappa + \beta) + \lambda_2 (\alpha + \beta) \right) - \lambda_1 (\alpha \kappa + \beta) \frac{1}{m} \sum_{j=2}^m (\alpha + \beta)^{j-1}}, \quad (13)$$

where

$$\Delta_m = (1 - \alpha - \beta) \lambda_1 (\alpha + \beta) \frac{\lambda_0}{1 - \lambda_1 - \lambda_2} \left(\frac{m-1}{m} + \frac{1}{m} \sum_{j=2}^m \sum_{k=1}^{j-2} (\alpha + \beta)^k \right). \quad (14)$$

In general, the variance of the daily returns will depend on all the model parameters and on the fourth moment, κ , of the innovation. The fact that the unconditional variance depends on κ distinguishes the $(MF)^2$ GARCH-rw- m from standard GARCH models. Also, note that for the condition in equation (12) to be satisfied, it is not sufficient that each component is covariance stationary.

Figure 3 shows the annualized unconditional volatility, $\sqrt{252 \cdot \mathbf{Var}[r_t]}$, for $\kappa = 3$ (red), $\kappa = 5$ (blue), $\kappa = 7$ (green) and $m = 21, \dots, 126$. The model parameters are chosen according to the estimates from an (MF)² GARCH-rw-63 for daily S&P 500 returns over the January 1971 to June 2020 period (see Table 3). The figure shows that for a fixed value m the unconditional volatility increases in κ . As m increases the unconditional volatility decreases. The vertical dashed line indicates the unconditional volatility for $m = 63$.

When imposing the restriction $m = 1$, 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 would need to impose a restriction on the model parameters to ensure that τ_t is the long-run component (i.e. $\alpha + \beta < \lambda_1 + \lambda_2 < 1$). However, even under this restriction it turns out that empirically r_{t-1}^2/h_{t-1} is too noisy as a measure of past forecast accuracy. Hence, we will not employ this specification in the empirical application. For completeness, in the following corollary we derive the condition for the covariance stationarity of the daily returns when $m = 1$ and state their unconditional variance.

Corollary 1. *Let Assumptions 1-3 be satisfied. If $\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta) < 1$ the (MF)² GARCH-rw-1 process is covariance stationary and the unconditional variance of the returns is given by*

$$\mathbf{Var}[r_t] = \frac{\lambda_0 + \frac{\lambda_0}{1-\lambda_1-\lambda_2}(1-\alpha-\beta)(\lambda_1+\lambda_2)}{1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]}. \quad (15)$$

If the short-term component is constant (i.e. $\alpha = \beta = 0$), the expression in equation (15) reduces to $\mathbf{Var}[r_t] = \mathbf{E}[\tau_t] = \lambda_0/(1-\lambda_1-\lambda_2)$. If the long-term component is constant (i.e. $\lambda_1 = \lambda_2 = 0$), the expression in equation (15) reduces to $\mathbf{Var}[r_t] = \mathbf{E}[\lambda_0 h_t] = \lambda_0$. In the latter case, the (MF)² GARCH reduces to a simple GARCH with variance targeting, i.e. we can estimate λ_0 by the sample variance of the returns.

News impact curve: Following Engle and Ng (1993), we use the NIC to illustrate how the conditional volatility is updated in response to a shock. For a standard GARCH(1,1) with conditional variance $\tilde{h}_t = \alpha_0 + \alpha r_{t-1}^2 + \beta \tilde{h}_{t-1}$ the NIC is defined as

$$NIC_{t+1}^{GA} = \tilde{h}_{t+1}(r_t|\tilde{h}_t) = A_t^{GA} + \alpha r_t^2,$$

where $A_t^{GA} = \alpha_0 + \beta \tilde{h}_t$. That is, the NIC is a function of today's return and A_t^{GA} is known conditional on $\mathcal{F}_{t-1}$. While $\tilde{h}_t$ affects the intercept of the NIC, the marginal effect of r_t^2 is given by α .

Next, we consider the NIC of the (MF)² GARCH-rw- m :

$$\begin{aligned} NIC_{t+1}^{MF} &= \sigma_{t+1}^2(r_t|\tau_t, h_t) = \tau_{t+1}h_{t+1} \\ &= (\lambda_0 + \lambda_1 V_t^{(m)} + \lambda_2 \tau_t) \left((1 - \alpha - \beta) + \alpha \frac{r_t^2}{\tau_t} + \beta h_t \right) \\ &= A_t^{MF} + B_t^{MF} + \left(\lambda_1 \beta \frac{1}{m} + \lambda_2 \alpha \right) r_t^2, \end{aligned} \quad (16)$$

where

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

and

$$B_t^{MF} = \lambda_0 \alpha \frac{r_t^2}{\tau_t} + \lambda_1(1 - \alpha - \beta) \frac{1}{m} \frac{r_t^2}{h_t} + \lambda_1 \alpha \frac{1}{m} \frac{r_t^4}{\sigma_t^2} + \lambda_1 \alpha \frac{r_t^2}{\tau_t} \frac{1}{m} \sum_{j=1}^{m-1} \frac{r_{t-j}^2}{h_{t-j}}.$$

Thus, the NIC consists of three terms. 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 αr_t^2 in the simple GARCH. However, the term $B_{n,t}^{MF}$ shows that the marginal effect of r_t^2 now also depends on the current conditional variance, σ_t^2 , as well as h_t and τ_t individually.

Note that for $\lambda_1 = \lambda_2 = 0$ and $m = 1$, NIC_{t+1}^{MF} reduces to the NIC of the simple GARCH:

$$\begin{aligned} NIC_{t+1}^{MF} &= \sigma_{t+1}^2(r_t|h_t) = \lambda_0 h_{t+1} = \lambda_0(1 - \alpha - \beta) + \lambda_0 \beta h_t + \alpha r_t^2 \\ &= \alpha_0 + \beta \tilde{h}_t + \alpha r_t^2 = NIC_{t+1}^{GA}, \end{aligned}$$

where we use that $\tilde{h}_t = \lambda_0 h_t$ and $\mathbf{E}[\tilde{h}_t] = \alpha_0/(1 - \alpha - \beta) = \lambda_0 = \mathbf{E}[h_t]$.

Figure 4 shows the NIC as given in equation (16). We omit the intercept and plot $B_t^{MF} + (\lambda_1 \beta \frac{1}{m} + \lambda_2 \alpha) r_t^2$. We consider the case $m = 63$ (quarterly averaging) and choose the parameters as $\lambda_0 = 0.1$, $\lambda_1 = 0.56$, $\lambda_2 = 0.34$, $\alpha = 0.1$, $\beta = 0.88$. The short-term component is fixed at its unconditional expectation: $h_t = 1$. In addition, we 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 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). Note that for the given parameter choice: $\mathbf{E}[\tau_t] = 1$. That is the blue NIC represents a situation in which both the short- and long-term component are at their unconditional expectation. The green and red lines show that the marginal effect of the return becomes stronger/weaker when τ_t is below/above its expectation. That is, during a low volatility regime the conditional volatility is more sensitive to news than during a high volatility regime. This is reasonable because a large value of $|r_t|$ is to be expected during turbulent times but less so during tranquil times. This behaviour is also in line with what is empirically observed for option-implied volatility: “... , if VIX indicates the market is going to be volatile it may not react too much when the market is volatile. And if VIX indicates the market is expected to be less volatile, if we do get high volatility, VIX adjusts accordingly” (Rhoades, 2020, p.60).

Remark 2. *Generalized autoregressive score (GAS) type volatility models have a similar property. For example, in the GAS-GARCH with t -distributed innovations the effect of extreme returns today on the volatility tomorrow is lower than in a standard GARCH framework, i.e. extreme observations are downweighted (see Creal et al., 2013).*

Forecasting volatility: Because the (MF)² 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 GARCH component can be simply computed as $\mathbf{E}[h_{t+s}|\mathcal{F}_t] = 1 + (\alpha + \beta)^{s-1}(h_{t+1} - 1)$. Forecasting the long-term component is slightly more involved:

1. $\mathbf{E}[\tau_{t+1}|\mathcal{F}_t] = \tau_{t+1}$

2. For $s = 2, \dots, m$ the forecasts can be recursively calculated as

$$\begin{aligned} \mathbf{E}[\tau_{t+s}|\mathcal{F}_t] &= \lambda_0 + (\lambda_1 \frac{1}{m} + \lambda_2) \mathbf{E}[\tau_{t+s-1}|\mathcal{F}_t] + \lambda_1 \frac{1}{m} \sum_{j=2}^{s-1} \mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t] \\ &\quad + \lambda_1 \frac{1}{m} \sum_{j=s}^m \frac{r_{t+s-j}^2}{h_{t+s-j}}. \end{aligned} \tag{17}$$

3. For $s > m$ the following recursion applies:

$$\mathbf{E}[\tau_{t+s}|\mathcal{F}_t] = \lambda_0 + (\lambda_1 \frac{1}{m} + \lambda_2) \mathbf{E}[\tau_{t+s-1}|\mathcal{F}_t] + \lambda_1 \frac{1}{m} \sum_{j=2}^m \mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t]. \quad (18)$$

Using these results, $\mathbf{E}[h_{t+s}\tau_{t+s}|\mathcal{F}_t]$ can be computed as follows:

Theorem 2. *Let Assumptions 1-3 and the constraint in equation (12) be satisfied. Then, in the $(MF)^2$ GARCH-rw-m the forecast of the conditional variance on day $t + s$, $s \geq 1$, can be computed as follows:*

$$1. \mathbf{E}[h_{t+1}\tau_{t+1}|\mathcal{F}_t] = h_{t+1}\tau_{t+1}$$

2. For $s = 2, \dots, m$ the forecasts can be recursively calculated as

$$\begin{aligned} \mathbf{E}[h_{t+s}\tau_{t+s}|\mathcal{F}_t] &= (1 - \alpha - \beta) \mathbf{E}[\tau_{t+s}|\mathcal{F}_t] + \lambda_0(\alpha + \beta) \mathbf{E}[h_{t+s-1}|\mathcal{F}_t] \\ &\quad + \left(\lambda_1 \frac{1}{m} (\alpha\kappa + \beta) + \lambda_2(\alpha + \beta) \right) \mathbf{E}[h_{t+s-1}\tau_{t+s-1}|\mathcal{F}_t] \\ &\quad + \lambda_1(\alpha + \beta) \mathbf{E}[h_{t+s-1}|\mathcal{F}_t] \frac{1}{m} \sum_{j=s}^m \frac{r_{t+s-j}^2}{h_{t+s-j}} \\ &\quad + (1 - \alpha - \beta) \lambda_1(\alpha + \beta) \frac{1}{m} \sum_{j=2}^{s-1} \mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t] \left(1 + \sum_{k=1}^{j-2} (\alpha + \beta)^k \right) \\ &\quad + \lambda_1(\alpha\kappa + \beta)(\alpha + \beta) \frac{1}{m} \sum_{j=2}^{s-1} (\alpha + \beta)^{j-2} \mathbf{E}[h_{t+s-j}\tau_{t+s-j}|\mathcal{F}_t]. \quad (19) \end{aligned}$$

3. For $s > m$ the following recursion applies:

$$\begin{aligned} \mathbf{E}[h_{t+s}\tau_{t+s}|\mathcal{F}_t] &= (1 - \alpha - \beta) \mathbf{E}[\tau_{t+s}|\mathcal{F}_t] + \lambda_0(\alpha + \beta) \mathbf{E}[h_{t+s-1}|\mathcal{F}_t] \\ &\quad + \left(\lambda_1 \frac{1}{m} (\alpha\kappa + \beta) + \lambda_2(\alpha + \beta) \right) \mathbf{E}[h_{t+s-1}\tau_{t+s-1}|\mathcal{F}_t] \\ &\quad + (1 - \alpha - \beta) \lambda_1(\alpha + \beta) \frac{1}{m} \sum_{j=2}^m \mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t] \left(1 + \sum_{k=1}^{j-2} (\alpha + \beta)^k \right) \\ &\quad + \lambda_1(\alpha\kappa + \beta)(\alpha + \beta) \frac{1}{m} \sum_{j=2}^m (\alpha + \beta)^{j-2} \mathbf{E}[h_{t+s-j}\tau_{t+s-j}|\mathcal{F}_t]. \quad (20) \end{aligned}$$

We illustrate the behavior of the volatility forecasts by an empirical example. Again, we rely on the parameter estimates from the $(MF)^2$ GARCH-rw-63 for daily S&P 500

returns (see Table 3). Figure 5 shows two panels. The conditional volatilities and long-components are plotted as red and blue lines, respectively. The unconditional volatility is represented by the black line. All quantities are annualized. We compute volatility forecasts from day $t = 153$ (indicated by the black vertical line) onwards. The green dotted line corresponds to the forecasts from the $(MF)^2$ GARCH-rw-63. For comparison, we also plot the volatility forecasts from a simple GARCH(1,1) as a dashed purple line. We rely on the estimates for α and β from Table 3 but impose the same unconditional variance as for the $(MF)^2$ GARCH-rw-63 and the same conditional variance at the forecast origin. That is, the s -step-ahead forecast is given by: $\mathbf{Var}[r_t] + (\alpha^{GA} + \beta^{GA})^{s-1}(h_{t+1}\tau_{t+1} - \mathbf{Var}[r_t])$, where α^{GA} and β^{GA} are the parameters of the one-component GARCH. We have chosen the forecast origin so that we can nicely illustrate the differences between the forecasts from the two models. In the upper/lower panel, $t = 153$ corresponds to August 12, 2011, and April 16, 2018, respectively. In the upper panel, the conditional volatility as well as the long-term component are close to the unconditional variance until there is a jump in volatility up to a level of almost 50%.⁶ A well known problem of simple GARCH models is that their volatility forecast are too persistent after such large shocks (see, for example, Klaassen, 2002). This is confirmed by the behavior of the dashed purple line. In contrast, the forecast from the $(MF)^2$ GARCH is less persistent and more in line with the actual behavior of the conditional volatility. This property is due to the lower persistence in the short-term component of the $(MF)^2$ GARCH. In the lower panel, volatility stayed below the unconditional volatility for an extended period of time. 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.⁷ Hence, the forecasts from the simple GARCH will directly converge to the unconditional volatility. Because the persistence in the simple GARCH is high, the forecast almost linearly declines towards the unconditional variance. The $(MF)^2$ GARCH forecast behaves very differently. In the short-run, the forecast converges towards the long-term component. That is, the forecast initially decreases below the unconditional volatility. Only in the very long-run, the $(MF)^2$ GARCH forecast converges towards the unconditional volatility. In summary, these

⁶The 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.

⁷The spike in volatility was associated with trade war fears and the imposition of tariffs on China.

examples illustrate that forecasts from the $(\text{MF})^2$ GARCH are more flexible than forecasts from the simple GARCH. The $(\text{MF})^2$ GARCH forecasts reflect both the current level of volatility, $(\sqrt{h_{t+1}\tau_{t+1}})$, as well as the prevailing volatility regime $(\sqrt{\tau_{t+1}})$. The $(\text{MF})^2$ GARCH forecasts capture the empirical observation that there are persistent cyclical movements of the conditional volatility around the unconditional volatility.

3.2 Low-frequency long-term component

For some applications, it is of interest to obtain a low-frequency measure of long-term volatility. For example, there is a literature that evaluates the predictive power of monthly or quarterly macroeconomic and financial variables for low-frequency measures of stock volatility (see, for example, Schwert, 1989, Paye, 2012, and Christiansen et al., 2012). For this purpose, we now specify a parametrization of the $(\text{MF})^2$ GARCH with $n > 1$. Using the notation that allows for multiple frequencies, we write the short-term volatility component as

$$h_{i,t} = (1 - \alpha - \beta) + \alpha \frac{r_{i-1,t}^2}{\tau_t} + \beta h_{i-1,t} \quad (21)$$

for $i = 2, \dots, n$. We use the convention that $r_{0,t} = r_{n,t-1}$ and $h_{0,t} = h_{n,t-1}$. Hence, for $i = 1$, we obtain

$$h_{1,t} = (1 - \alpha - \beta) + \alpha \frac{r_{n,t-1}^2}{\tau_{t-1}} + \beta h_{n,t-1}. \quad (22)$$

As before, we assume that Assumption 2 holds.

We now 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

$$V_t = \frac{1}{n} \sum_{i=1}^n \frac{r_{i,t}^2}{h_{i,t}} = \tau_t \frac{1}{n} \sum_{i=1}^n Z_{i,t}^2 = \tau_t \bar{Z}_t, \quad (23)$$

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.*. If we additionally assume that the $Z_{i,t}$ are standard normal, then the $\bar{Z}_t$ are χ^2 distributed. This suggests the following specification:

$$\tau_t = \lambda_0 + \lambda_1 V_{t-1} + \lambda_2 \tau_{t-1}. \quad (24)$$

Note that the low-frequency component τ_t still *measures* volatility in daily units. We will refer to this parametrization of the long-term component as $(\text{MF})^2$ GARCH-lf- n ,

where ‘lf’ stands for *low-frequency* and n refers to the choice of the low-frequency period: For monthly, quarterly or semi-annual low-frequency periods we set $n = 21$, $n = 63$ or $n = 126$. 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 $\mathbf{E}[V_t] = \lambda_0 / (1 - \lambda_1 - \lambda_2)$. Because $\mathbf{E}[\log(\lambda_1 \bar{Z}_t^2 + \lambda_2)] < 0$, it also follows that $V_t = \tau_t \bar{Z}_t$ is strictly stationary and ergodic.

The (MF)² GARCH with low-frequency long-term component can be used to extract the long-term component at a monthly or quarterly frequency and then to relate this component to explanatory variables. However, the fact that the long-term component is updated at the low-frequency only, introduces a discontinuity into the daily conditional variances. Hence, the daily returns will be non-stationary. Nevertheless, it can be shown that the vector of low-frequency returns $\mathbf{r}_t = (r_{1,t}, \dots, r_{n,t})'$ is stationary. In the following theorem we present the unconditional variances of the n days within low-frequency period t .

Theorem 3. *Let Assumptions 1-3 be satisfied. If the (MF)² GARCH process $(\mathbf{r}_t)_{t \in \mathbb{Z}}$ is covariance stationary, then $(\alpha + \beta)^{(n-1)}[\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)] < 1$. The unconditional variances of the n daily returns within low-frequency period t can be recursively computed as*

$$\mathbf{Var}[r_{1,t}] = \frac{\lambda_0(1 - \alpha - \beta) \left(1 + \sum_{i=1}^{n-1} (\alpha + \beta)^{n-i}\right) + \lambda_0(\alpha + \beta)^n + \frac{\lambda_0}{1 - \lambda_1 - \lambda_2} \Psi_n}{1 - (\alpha + \beta)^{n-1} [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]} \quad (25)$$

where

$$\begin{aligned} \Psi_n = & (1 - \alpha - \beta) \left[(\lambda_1 + \lambda_2) + \frac{1}{n} \lambda_1 \left((\alpha\kappa + \beta) \sum_{j=1}^{n-1} (n-j)(\alpha + \beta)^{n-1-j} \right. \right. \\ & \left. \left. + \sum_{j=1}^{n-1} \sum_{i=j}^{n-1} (\alpha + \beta)^{n-i} \right) + \lambda_2 \sum_{i=1}^{n-1} (\alpha + \beta)^{n-i} \right] \end{aligned} \quad (26)$$

and for $i = 2, \dots, n$

$$\mathbf{Var}[r_{i,t}] = \frac{\lambda_0}{1 - \lambda_1 - \lambda_2} (1 - \alpha - \beta) + (\alpha + \beta) \mathbf{Var}[r_{i-1,t}].$$

Because the long-term component is updated on the first day of each low-frequency period only, the unconditional variance, $\mathbf{Var}[r_{i,t}]$, jumps upwards on the first day of period t and then smoothly declines during the remaining days of period t until it jumps upwards

again on the first day of period $t+1$.⁸ The following corollary shows that the time-variation in the daily unconditional variances essentially disappears when n is large.

Corollary 2. *When n is sufficiently large, the unconditional variances are given by*

$$\mathbf{Var}[r_{i,t}] \approx \frac{\lambda_0}{1 - \lambda_1 - \lambda_2}$$

for $i = 1, \dots, n$.

Thus, when n is large, the unconditional variances are equal to the unconditional expectation of the long-term component. This is the case because the correlation between $h_{i,t}$ and τ_t converges to zero as n increases and, hence, $\mathbf{Var}[r_{i,t}] = \mathbf{E}[h_{i,t}\tau_t] = \mathbf{E}[h_{i,t}]\mathbf{E}[\tau_t] = \mathbf{E}[\tau_t]$.

3.3 Quasi-maximum likelihood estimation

We can estimate the new model using standard quasi-maximum likelihood asymptotic theory. 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 $\boldsymbol{\theta} = (\boldsymbol{\theta}'_{GA}, \boldsymbol{\theta}'_{MEM})'$. For simplicity, we use the notation for $n = 1$. The Gaussian quasi-loglikelihood function (omitting the constant) can be written as $L(\boldsymbol{\theta}|r_T, \dots, r_1) = \sum_{t=1}^T l_t$ with

$$l_t = -\frac{1}{2} \left[\ln(h_t) + \ln(\tau_t) + \frac{r_t^2}{h_t \tau_t} \right]. \quad (27)$$

The log-likelihood takes the same form as for the GARCH-MIDAS model with a daily explanatory variable (see, for example, Conrad and Schienle, 2020). Wang and Ghysels (2015) provide the asymptotic theory for the special case in which the explanatory variable is the sum of the squared returns of the last m days. We denote the first and second derivative of the log-likelihood by

$$\mathbf{s}_t(\boldsymbol{\theta}) = \frac{\partial l_t}{\partial \boldsymbol{\theta}} \quad \text{and} \quad \mathbf{d}_t(\boldsymbol{\theta}) = \frac{\partial^2 l_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}. \quad (28)$$

The first derivative can be written as

$$\mathbf{s}_t(\boldsymbol{\theta}) = \frac{1}{2} \left[\frac{r_t^2}{h_t \tau_t} - 1 \right] \left(\frac{1}{h_t} \frac{\partial h_t}{\partial \boldsymbol{\theta}} + \frac{1}{\tau_t} \frac{\partial \tau_t}{\partial \boldsymbol{\theta}} \right). \quad (29)$$

⁸Figure A.1 in Appendix A shows the unconditional variances $\mathbf{Var}[r_{i,t}]$, $i = 1, \dots, n$, of an (MF)² GARCH-lf-21 for three consecutive low-frequency periods with $\kappa = 3$ (red), $\kappa = 5$ (blue) or $\kappa = 7$ (green). The model parameters are chosen as $\alpha = 0.05$, $\beta = 0.85$, $\lambda_0 = 0.04$, $\lambda_1 = 0.3$, $\lambda_2 = 0.65$. Again, the unconditional variances are increasing in κ .

Let $\hat{\boldsymbol{\theta}}$ denote the quasi-maximum likelihood estimator, i.e. $\sum_t \mathbf{s}_t(\hat{\boldsymbol{\theta}}) = \mathbf{0}$, and $\boldsymbol{\theta}_0$ denote the vector of true parameters. Because $\mathbf{E}[Z_t^2 | \mathcal{F}_{t-1}] = 1$, 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. Hence, under standard regularity conditions we expect to have

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{D}^{-1} \boldsymbol{\Omega} \mathbf{D}^{-1}), \quad (30)$$

where $\boldsymbol{\Omega} = \boldsymbol{\Omega}(\boldsymbol{\theta}_0) = \mathbf{E}[\mathbf{s}_t(\boldsymbol{\theta}_0) \mathbf{s}_t(\boldsymbol{\theta}_0)']$ and $\mathbf{D} = \mathbf{D}(\boldsymbol{\theta}_0) = -\mathbf{E}[\mathbf{d}_t(\boldsymbol{\theta}_0)]$. The corresponding estimators are denoted by

$$\hat{\boldsymbol{\Omega}} = \boldsymbol{\Omega}(\hat{\boldsymbol{\theta}}) = \frac{1}{T} \sum_t \mathbf{s}_t(\hat{\boldsymbol{\theta}}) \mathbf{s}_t(\hat{\boldsymbol{\theta}})' \quad \text{and} \quad \hat{\mathbf{D}} = \mathbf{D}(\hat{\boldsymbol{\theta}}) = -\frac{1}{T} \sum_t \mathbf{d}_t(\hat{\boldsymbol{\theta}}).$$

We will base the standard errors on $\widehat{\mathbf{Var}}(\hat{\boldsymbol{\theta}}) = \hat{\mathbf{D}}^{-1} \hat{\boldsymbol{\Omega}} \hat{\mathbf{D}}^{-1}$. In a Monte-Carlo simulation, we confirmed that this procedure leads to valid inference.

3.4 Extensions: Modifications of the (MF)² GARCH-rw- m

(MF)² GARCH with Beta-weights

In equation (11), we assume that $V_{t-1}^{(m)}$ is based on the average of the last m standardized forecast errors. Imposing equal weights might be too restrictive. Alternatively, we can consider a weighted average of the form

$$V_{t-1}^{(m)} = \sum_{j=1}^m w_j(\omega) V_{t-j} = \sum_{j=1}^m w_j(\omega) \frac{r_{t-j}^2}{h_{t-j}}. \quad (31)$$

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 simply given by $1/m$ and, hence, we obtain the model with rolling window $V_t^{(m)}$. We will refer to this parametrization as (MF)² GARCH-bw- m , where ‘bw- m ’ stands for *Beta-weights* of length m .

Realized volatility MEM

As discussed before, when m is small the average standardized forecast error of the short-term component, $V_t^{(m)} = \frac{1}{m} \sum_{j=1}^m \frac{r_{t-j}^2}{h_{t-j}}$, will be a noisy proxy for the true forecast

error. Instead, if we observe daily realized variances, $RV_{i,t}$, we can base τ_t on the realized measure:

$$V_t^{(m,RV)} = \frac{1}{m} \sum_{j=1}^m \frac{RV_{t-j}}{h_{t-j}}.$$

Again, we can apply the MEM specification from equation (10). However, note that this specification is no longer dynamically complete.

3.5 Relation to other components models

Amado et al. (2019) provide a recent survey on multiplicative volatility models. They consider models that can be written as in equation (8). Essentially, those models have in common the specification of the right-hand side (i.e. h_t follows some GARCH process), but differ with respect to the specification of τ_t . For example, 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 (MF)² GARCH is closely related to the GARCH-MIDAS of Engle et al. (2013). In the latter model, τ_t depends on the lagged values of an explanatory variable and can evolve either at the same or at a lower frequency than the short-term component. If we rewrite the long-term component of the (MF)² GARCH-lf- n as a MEM(∞)

$$\tau_t = \frac{\lambda_0}{1 - \lambda_2} + \lambda_1 \sum_{j=0}^{\infty} \lambda_2^j V_{t-1-j},$$

it becomes clear that the (MF)² GARCH corresponds to a GARCH-MIDAS with an exponential weighting scheme and explanatory variable V_t . Thus, the (MF)² GARCH avoids the problem of selecting the most appropriate explanatory variable and, instead, takes a more reduced form approach.⁹ Further related models are the Realized EGARCH-MIDAS of Borup and Jakobsen (2019) and the realized-measure-based component model of Noureldin (2020). In contrast to the (MF)² GARCH, those specifications combine daily returns with high-frequency measures of volatility. While all the previous models are of the GARCH-type, there are also pure component MEM models. For examples, see the Composite-MEM for intraday volumes by Brownlees et al. (2011) or the Component-MEM for realized volatilities by Amendola et al. (2020).

⁹For a GARCH-MIDAS model with LASSO-based variable selection, see Fang et al. (2020).

4 Empirical Application

4.1 Stock market data

We use daily return data for the S&P 500, FTSE 100 and the HSI. For the S&P 500 we consider a long sample period starting in January 1971 and ending in June 2020. For the other stock markets the sample covers the period March 1990 to June 2020. 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 price on day t . Summary statistics for the daily returns are presented in Panel A of Table 1. To evaluate the forecast performance of the various model specifications, we employ realized variances based on intraday data. We obtained the realized measures from the Realized Library of the Oxford-Man Institute of Quantitative Finance (see Heber et al., 2009). 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). Data for realized variances are available for the January 2000 to June 2020 period. Panel B of Table 1 shows summary statistics for the realized variances.

4.2 The (MF)² GARCH in action

4.2.1 Full sample parameter estimates

We first present and discuss full sample parameter estimates for the (MF)² GARCH models. The short-term component is either specified as a symmetric or an asymmetric (GJR) GARCH model. We consider the (MF)² GARCH-rw- m with daily long-term component and the (MF)² GARCH-lf- n with low-frequency long-term component. From Section 3.4, we consider the model with Beta-weights (bw- m). We leave the $V_t^{(m,RV)}$ -based model for future research.

S&P 500

The first two rows in Table 3 (labeled as ‘ τ_t const.’) show the parameter estimates for the nested one-component (GJR) GARCH model. The parameter estimates for α , γ and β take typical values and indicate strong persistence in the GARCH component. The remaining models allow for a MEM structure in the long-term component. We first estimate the (MF)² GARCH-rw- m models, i.e. models with rolling window $V_t^{(m)}$. We observe that the persistence in the short-term (GJR) GARCH component decreases

once the long-term component is time-varying. This effect is the stronger, the smaller m is. On the other hand, the persistence in the long-term component tends to increase with m decreasing. When the short-term component is specified as a GJR GARCH, the estimates of λ_2 are somewhat smaller than when the short-term component is specified as a GARCH. Among the $(MF)^2$ GJR/GARCH-rw- m models, the BIC is minimized when choosing $m = 63$, which corresponds to a quarterly frequency.¹⁰ Next, we estimate the $(MF)^2$ GARCH-bw- m specification. In general, the parameter estimates are close to the ones from the $(MF)^2$ GARCH-rw- m models. For $m = 63$ the estimate of ω approaches the lower bound of one. That is, the optimal weighting scheme for $m = 63$ indeed appears to be flat, which provides further support for the $(MF)^2$ GARCH-rw-63 model. Only for $m = 126$ and $m = 252$ in combination with a GJR short-term component, the estimates of ω are significantly different from one. However, this may be the case because the true weighting scheme is flat for the first 63 lags and zero thereafter. In this case the the optimal choice of ω for $m > 63$ will be greater than one. Also the BIC favors the $(MF)^2$ GJR-bw-63.¹¹ Finally, we estimate the $(MF)^2$ GARCH-lf- n models. Those models feature a low-frequency long-term component. When specifying a semi-annual ($n = 126$) long-term component, the estimate for λ_1 is around 0.75 while the estimate of λ_2 is insignificant. That is, when n is large, a MEM(0,1) might be sufficient to capture the movements in the conditional expectation of V_t . In contrast, when τ_t evolves at a quarterly ($n = 63$) or a monthly ($n = 21$) frequency, λ_2 increases and becomes significant. That is, when the frequency of the long-term component increases, the estimates of λ_1/λ_2 tend to decrease/increase. Again, switching from a symmetric GARCH specification to an asymmetric one tends to decrease the estimate of λ_2 .

Among all models, the BIC favors the $(MF)^2$ GJR-rw-63. In general, models with a GJR short-term component are preferred to those with a symmetric GARCH component. All $(MF)^2$ GARCH models are preferred when compared to the nested GARCH models (although the differences in the BIC are small). Finally, note that the estimated

¹⁰We have estimated the $(MF)^2$ GARCH-rw- m on a grid of m values between 21 and 136. Figure A.2 in Appendix A shows the BIC as a function m . The figure confirms that the optimal choice of m is slightly above 60.

¹¹For completeness, Figure A.3 in Appendix A shows the estimated weights $w_j(\omega)$ for the $(MF)^2$ GJR-bw- m with $m = 252$ and $m = 126$. The figure illustrates that both choices of m lead to very similar weighting schemes (although the estimates of ω are quite different).

parameters from all models satisfy Assumptions 2 and 3. The parameters also satisfy the conditions in Theorem 1 and Theorem 3.

Figure 6 shows the annualized conditional volatility, $\sqrt{h_{i,t}\tau_t}$, and the long-term volatility $\sqrt{\tau_t}$ of the (MF)² GJR-rw-63. Figure 7 compares the long-term volatilities of the (MF)² GJR-rw-63 and the (MF)² GJR-lf-63. The long-term volatility of the latter model changes at the quarterly frequency only. First, both figures clearly show that there are prolonged cycles in long-term volatility. Thus, as expected, the behavior of $\sqrt{\tau_t}$ mirrors that of the averaged standardized forecast errors in Figure 1. Second, the daily long-term volatility component of the (MF)² GJR-rw-63 responds more quickly and more strongly to news and, hence, anticipates the behaviour of the low-frequency long-term component.

International stock markets

Table 4 shows the estimation results for the FTSE 100 and the HSI. Again, the table shows that according to the BIC the (MF)² GARCH models are preferred to the nested GARCH models. For the FTSE the (MF)² GJR-rw-63 achieves the lowest BIC and for the HSI the (MF)² GJR-rw-21 performs best. As for the S&P 500, there is no evidence that a Beta weighting scheme leads to an improvement relative to simply using a rolling window weighting scheme with equal weights.

4.2.2 What does the low-frequency long-term component capture?

Next, we relate the movements in the low-frequency long-term component of the (MF)² GJR-lf to the evolution of the macroeconomic environment. We employ the equity market volatility (EMV) trackers of Baker et al. (2019a) which relate financial volatility to news about the macroeconomy and policy. The EMV trackers are based on a counting of newspaper articles on stock market volatility and 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. The ‘Overall’ tracker is strongly correlated with the VIX and measures of realized volatility (see Baker et al., 2019a). In addition, we focus on three important sub-trackers: ‘Macroeconomic News & Outlook’, ‘Monetary Policy’ and ‘Policy Related’. For example, the macroeconomic outlook is mentioned in 72% of articles that count towards the ‘Overall’ EMV tracker. Note that the EMV trackers do not distinguish between good and bad news. We regress estimates of the long-term component τ_{t+1} from (MF)² GJR-lf models with monthly and

quarterly long-term components on the different EMV trackers while controlling for τ_t .¹² We standardize each tracker so that the estimated coefficients measure the effect of a one standard deviation change in the respective tracker. Table 5 presents estimation results.

Columns (1) and (3) show the coefficient estimates from a regression on the ‘Overall’ tracker at a monthly/quarterly frequency. There is strong evidence that the long-term component and, hence, the volatility forecast errors of the short-term component increase when there are more newspaper articles on stock market volatility. Because the number of newspaper articles increases with the level of volatility, this suggests that the forecast errors of the short-term component also increase with the level of volatility. In Column (2), we regress the long-term component on the three sub-trackers. We find that the ‘Macroeconomic News & Outlook’ and ‘Policy Related’ trackers increase the long-term component while the ‘Monetary Policy’ tracker reduces it. The latter finding is in line with Baker et al. (2019b) who report that news about monetary policy are more often good than bad news and tend to reduce future volatility on average. Our result suggests that news about monetary policy does not only reduce volatility but also volatility forecast errors. Column (4) confirms this finding at the quarterly frequency. Among the three sub-trackers, ‘Macroeconomic News & Outlook’ has the strongest association with long-term volatility at both frequencies. In Column (5), we additionally control for the one-quarter-ahead GDP forecast, $GDP_{t+1|t}^{SPF}$, from the SPF. As in Table 2, we find that the long-term component behaves counter-cyclical. This confirms the findings in Engle et al. (2013), Conrad and Loch (2015) and Conrad and Kleen (2020) using GARCH-MIDAS models. However, our results also suggest that GARCH-MIDAS models with a single explanatory variable will not be able to capture the various sources of news that drive long-term volatility.

4.2.3 Effects of major news events on conditional volatility

What are the instantaneous effects of *major* news events on the short- and long-term component of the (MF)² GARCH-rw- m ? To answer this question we focus on days with extreme returns in the S&P 500. Baker et al. (2019b) consider days with returns of at least $\pm 2.5\%$ and, based on newspaper articles on the following day, examine which events

¹²We do not rely on the models from Table 3 which were estimated using a fixed n . Instead, we reestimate the models using the actual number of calendar days, n_t , for each month/quarter. That is, the long-term component is updated on the first day of each calendar month/quarter.

caused those movements in U.S. stock prices. We rely on their classification and focus on three major news categories: ‘Macroeconomic News & Outlook’, ‘Corporate Earnings & Outlook’ and ‘Monetary Policy & Central Banking’.¹³ According to Baker et al. (2019b) these three news categories account for roughly 57% of all extreme returns during the 1946 to 2013 period. Again, the coding of the newspaper articles does not directly specify whether the news was good or bad. However, we can now infer the direction from the sign of the corresponding daily return.

In Table 6, we investigate how these extreme events affect the updating of the short- and long-run component. We rely on the full sample estimates of the short- and long-term component from the (MF)² GJR-rw-63 (i.e. the model with daily-long term component). In Columns (1) and (3), we regress r_t^2/τ_t and r_t^2/h_t on the three types of news whereby we distinguish between good and bad news. As implied by Columns (1) and (3), bad news typically has stronger effects than good news, i.e. leads to larger volatility forecast errors and, hence, a stronger updating of short- and long-run volatility. It is also interesting to compare the adjusted R^2 . If the short-term component fully captured the conditional heteroskedasticity, then the R^2 in Column (3) should be zero. In this case, τ_t would be constant and the R^2 in Column (1) should be high. As expected, the R^2 of 8% in Column (3) in combination with the R^2 of 20% in Column (1) imply that the short-term component does not fully anticipate the actual volatility. In Columns (2) and (4), we regress $\Delta h_{t+1} = h_{t+1} - h_t$ and $\Delta \tau_{t+1} = \tau_{t+1} - \tau_t$ on the news events. The regression results highlight another important difference between the short- and long-run component. The short-run component decreases in response to good news but increases in response to bad news (whereby the latter effect is much stronger). In contrast, the long-term component (almost symmetrically) increases in response to good and bad news. The latter finding is not in contrast with our findings in the previous subsection. While good macroeconomic conditions reduce the monthly/quarterly low-frequency long-term component, *extreme* events (whether good or bad) increase daily long-term volatility.

Next, we investigate the short-term/intermediate effect of major news events on the conditional volatility of the (MF)² GARCH. This effect will be driven by the interaction of the separate effects on the short- and long-run component. We run the following

¹³The indicator variables are obtained from: <https://stockmarketjumps.com/>

regression

$$\begin{aligned}
\sqrt{\frac{1}{S} \sum_{s=1}^S \sigma_{t+s}^2} = & b_1 + b_2 \text{Macro}_t \times \mathbf{1}_{(r_t > 0)} + b_3 \text{Macro}_t \times \mathbf{1}_{(r_t < 0)} \\
& + b_4 \text{Corp. Earn.}_t \times \mathbf{1}_{(r_t > 0)} + b_5 \text{Corp. Earn.}_t \times \mathbf{1}_{(r_t < 0)} \\
& + b_6 \text{Mon. Pol.}_t \times \mathbf{1}_{(r_t > 0)} + b_7 \text{Mon. Pol.}_t \times \mathbf{1}_{(r_t < 0)} \\
& \sqrt{\sigma_t^2} + e_t
\end{aligned} \tag{32}$$

for $S \in \{1, \dots, 21\}$. The upper left/right panels in Figure 8 show the effect of positive/negative macroeconomic news. The left panel shows that good news tends to significantly decrease the conditional volatility only during the following six days. This response is driven by the interaction of a decline of the short-term component which dominates over the first few days and an increase of the long-term component which dominates over the medium term.¹⁴ In contrast, negative news leads to a significant and much more sizable increase in the conditional variance. The effect is persistent over the following 21 days. We observe similar effects for news about corporate earnings and monetary policy. While good news reduces volatility only in the short-run, bad news has a much stronger and long-lasting effect.¹⁵

4.3 Out-of-sample forecasting

Finally, we evaluate the out-of-sample forecast performance of the (MF)² GARCH. We now exclusively focus on models with a daily long-term component. We reestimate all (MF)² GARCH models on a rolling window of observations. The parameter estimates are updated on a monthly basis. We then compute daily (1 day), monthly (21 days), quarterly (63 days), semi-annual (126 days) and annual (252 days) out-of-sample volatility forecasts for the January 2010 to June 2020 period. For the forecast evaluation, we proxy the true latent volatility by the daily realized variances, RV_t . We denote the cumulative realized variance over the next s days by $RV_{t+1,t+s} = \sum_{j=1}^s RV_{t+j}$ and the cumulative forecasts by

¹⁴Figure A.4 in Appendix A shows the separate effects on the short-term (upper panel) and the long-term (lower panel) component.

¹⁵Interestingly, for the simple GJR GARCH positive news significantly increases volatility on impact but has no significant effect thereafter (see Figure A.5 in Appendix A). That is, the initial impact of positive news has a counter-intuitive sign.

$\hat{\sigma}_{t+1,t+s|t}^2 = \sum_{j=1}^s \hat{\sigma}_{t+j|t}^2$. We use the squared error (SE),

$$SE(RV_{t+1,t+s}, \hat{\sigma}_{t+1,t+s|t}^2) = (RV_{t+1,t+s} - \hat{\sigma}_{t+1,t+s|t}^2)^2, \quad (33)$$

as the evaluation criterion.¹⁶ We compare the forecast performance of the (MF)² GARCH models with several competitor models. First, we consider the HAR model of Corsi (2009), the log-HAR and log-HAR with leverage of Corsi and Reno (2012). While the (MF)² GARCH models are purely based on daily returns, the HAR-type models make direct use of the realized variances and, hence, intraday data. HAR models are well known for their excellent forecast performance (see, for example, Bollerslev et al., 2018). Second, we use the nested GARCH and GJR GARCH models. Within the class of GARCH-type models, the GJR GARCH is well known for its outstanding forecasting abilities (see, for example, Hansen and Lunde, 2005, and Brownlees et al., 2012). Third, we employ two random walk type forecasts. The first forecast, Random Walk Type 1 (RW 1), simply scales the current realized variance, RV_t , by the appropriate horizon, i.e. uses the forecast $s \cdot RV_t$. Random Walk Type 2 (RW 2) uses the following forecast: $RV_{t-s+1,t}$. While RW 1 might perform well for short horizons, we expect that RW 2 might be competitive for long forecast horizons. Finally, we also compute the historical volatility forecast which we define as the appropriately scaled mean of the daily realized variances during the previous ten years. For each model and forecast horizon, we report the model performance as the ratio of the mean squared error (MSE) of the respective model and the mean squared error of historical volatility. For each forecast horizon and based on the squared forecast errors, we compute a model confidence set (MCS) as discussed in Hansen et al. (2011). We approximate the asymptotic distribution of the test statistic by block-bootstrapping and choose a confidence level of 85%.¹⁷

Table 7 shows the out-of-sample forecasting results for the S&P 500. The red-coloured numbers indicate which model achieves the lowest MSE ratio. The blue-shaded numbers indicate that the respective model is included in the MCS. First, note that for all forecast horizons an (MF)² GJR-rw model attains the lowest RMSE. Notably, the (MF)² GJR-rw-126 performs best for the quarterly, semi-annual and annual forecast horizon.

¹⁶Our results are robust to using the QLIKE instead of the SE loss.

¹⁷Specifically, we fit an AR(p) process to the squared error loss differences and choose the block length as the maximum of the forecast horizon and the maximal lag length that was selected by the AIC across all loss series.

At the shortest forecast horizon of one day the $(MF)^2$ GJR-rw-21 performs best. At all forecast horizons, the MCS includes $(MF)^2$ GJR-rw and $(MF)^2$ GJR-bw models. In contrast, models with symmetric short-term GARCH components are only included in the MCS for the monthly forecast horizon. This illustrates that allowing for asymmetry in the short-term component remains important when modelling volatility as a multiplicative two-component process. Moreover, the simple one-component GARCH and GJR GARCH models are never included in the MCS. Hence, the $(MF)^2$ GJR models are a clear improvement over their one-component counterparts. The $(MF)^2$ GJR-bw models are always included in the MCS, but they do not achieve a lower MSE than the models with rolling window $V_t^{(m)}$ (except for the one month horizon). Hence, modelling decaying weights in $V_t^{(m)}$ does not pay off in terms of improved forecast performance. The HAR-type models are included in the MCS for the monthly and quarterly forecast horizon but not for the semi-annual and annual forecast horizon. Finally, it is interesting to look at the size of the MSE ratios. The improvements in forecast performance relative to historical volatility can be up to 50%. For essentially all models, we observe the strongest improvements at the monthly horizon. Interestingly, the relative forecast improvement stays roughly constant for the $(MF)^2$ GJR-rw models when increasing the forecast horizon but becomes considerably weaker for HAR-type models.

The out-of-sample forecasting results for the FTSE 100 and the HSI can be found in Table 8. For brevity, we focus on the quarterly, semi-annual and annual forecast horizon. For both the FTSE 100 and the HSI, the $(MF)^2$ GJR-rw-126 achieves the lowest MSE ratio for all forecast horizons. Although the MCS for the quarterly forecast horizon now includes many models, the $(MF)^2$ GJR-rw/bw models dominate the nested (GJR) GARCH as well as the HAR models when further increasing the forecast horizon. For both stock markets, the $(MF)^2$ GJR-rw/bw specifications are the only models that are included in the MCS at the annual forecast horizon. Again, with increasing forecast horizon the improvements in relative forecast performance are much stronger for the $(MF)^2$ GJR-rw models than for the HAR-type models.

In summary, the $(MF)^2$ GJR models not only significantly improve upon existing models but also appear to be particularly suited for long-term volatility forecasting.

5 Conclusions

We suggest a multiplicative factor multi frequency component GARCH model. The new specification is based on the observation that the averaged standardized forecast errors of simple GARCH models are predictable. This predictability is due to counter-cyclical movements in financial volatility which are not adequately captured by simple GARCH models. In contrast, the $(MF)^2$ GARCH captures this property of financial volatility. We show that it is straightforward to compute multi-step-ahead volatility forecasts for the new model and illustrate that the conditional variance of the $(MF)^2$ GARCH responds asymmetrically to good and bad news. The response is asymmetric with respect to the size of the impact as well as with respect to the length of the impact. Finally, the empirical application shows that the new specification outperforms various competitor models in terms of out-of-sample forecast performance. The $(MF)^2$ GARCH model will be particularly useful in applications that require long-term forecasts of financial volatility such as long-run value-at-risk predictions or measurement of systemic risk (see Brownlees and Engle, 2017) and can also be used for policy analysis. As discussed in Engle et al. (2013), we can think of the realized variance as a noisy proxy of the long-term volatility component. Hence, it will be interesting to employ the estimate of the low-frequency long-term component when analyzing the link between financial volatility and financial crises or macroeconomic conditions (see, for example, Danielsson et al., 2018, and Berger et al., 2020).

References

- 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.
- Amado, C., Teräsvirta, T., 2013. Modelling volatility by variance decomposition. *Journal of Econometrics*, 175, 142-153.
- Amado, C., Teräsvirta, T., 2017. Specification and testing of multiplicative time-varying GARCH Models with applications. *Econometric Reviews*, 36, 421-446.

- 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>
- Baker, S.R., Bloom, N., Davis, S.J., Kost, K., 2019a. Policy news and stock market volatility. University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2019-53 . Available at SSRN: <http://dx.doi.org/10.2139/ssrn.3363862>
- Baker, S.R., Bloom, N., Davis, S.J., Sammon, M., 2019b. What triggers stock market jumps? Available at: <https://stockmarketjumps.com/>
- Berger, D., Dew-Becker, I., Giglio, S., 2020. Uncertainty shocks as second-moment news shocks. *Review of Economic Studies*, 87, 40–76.
- Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307-327.
- Bollerslev, T., Ghysels, E., 1996. Periodic autoregressive conditional heteroskedasticity, *Journal of Business & Economic Statistics*, 14, 139-151.
- Bollerslev, T., Hood, B., Huss, J., Pedersen, L.H., 2018. Risk everywhere: Modeling and managing volatility. *Review of Financial Studies*, 31, 2730-2773.
- Borup, D., Jakobsen, J.S., 2019. Capturing volatility persistence: a dynamically complete realized EGARCH-MIDAS model. *Quantitative Finance*, 19, 1839-1855.
- Brownlees, C., Cipollini, F., Gallo, G.M., 2011. Intra-daily volume modeling and prediction for algorithmic trading. *Journal of Financial Econometrics*, 9, 489-518.
- Brownlees, C., Engle, R.F., 2017. SRISK: A conditional capital shortfall measure of systemic risk. *Review of Financial Studies*, 30, 48-79.
- Brownlees, C., Engle, R.F., Bryan, K., 2012. A practical guide to volatility forecasting through calm and storm. *Journal of Risk*, 14, 3-22.
- Christiansen, C., Schmeling, M., Schrimpf, A., 2012. A comprehensive look at financial volatility prediction by economic variables. *Journal of Applied Econometrics*, 27, 956-977.

Conrad, C., Glas, A., 2020. 'Déjà vol' revisited: Survey forecasts of macroeconomic variables predict volatility in the cross-section of industry portfolios. Available at SSRN: <http://ssrn.com/abstract=3186567>.

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.

Conrad, C., Loch, K., 2015. Anticipating long-term stock market volatility. *Journal of Applied Econometrics*, 30, 1090-1114.

Conrad, C., Schienle, M., 2020. Testing for an omitted multiplicative long-term component in GARCH models. *Journal of Business & Economic Statistics*, 38, 229-242.

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

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.

Creal, D., Koopman, S.J., Lucas, A., 2013. Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28, 777-795.

Danielsson, J., Valenzuela, M., Zer, I., 2018. Learning from history: Volatility and financial crises. *Review of Financial Studies*, 31, 2774-2805.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987-1007.

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

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

- Engle, R.F., Ng, V.K., 1993. Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48, 1749-1778.
- 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.
- Fang, T., Lee, T.-H., Su, Z., 2020. Predicting the long-term stock market volatility: A GARCH-MIDAS model with variable selection. *Journal of Empirical Finance*, 58, 36-49.
- 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.
- Halunga, A.G., Orme, C. D., 2009. First-order asymptotic theory for parametric misspecification tests of GARCH models. *Econometric Theory*, 25, 364-410.
- 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.
- Hansen, P.R., Lunde, A., Nason, J. M., 2011. The model confidence set. *Econometrica*, 79, 453-497.
- Heber, G., Lunde, A., Shephard, N., Sheppard, K. K., 2009. Oxford-Man Institute's Realized Library. Oxford-Man Institute. University of Oxford.
- Klaassen, F., 2002. Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics*, 27, 363-394.
- Lundbergh, S., Teräsvirta, T., 2002. Evaluating GARCH models. *Journal of Econometrics*, 110, 417-435.
- Noureldin, D., 2020. Volatility prediction using a realized-measure-based component model. *Journal of Financial Econometrics*, forthcoming. Available at: <https://doi.org/10.1093/jjfinec/nbz041>
- Patton, A.J., 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160, 246-256.

Paye, B. S., 2012. ‘Déja Vol’: Predictive regressions for aggregate stock market volatility using macroeconomic variables. *Journal of Financial Economics*, 106, 527-546.

Rhoades, Russell, 2020. The VIX traders’s handbook. *Harriman House*.

Schwert, G.W., 1989. Why does stock market volatility change over time? *The Journal of Finance*, 44, 1115-1153.

Wang, F., Ghysels, E., 2015. Econometric analysis of volatility component models. *Econometric Theory*, 31, 362-393.

Figures

Figure 1: A GJR GARCH(1,1) is estimated for the S&P 500 using daily return data for the January 1971 to June 2020 period. The figure shows the average volatility forecast errors, $\tilde{V}_t = n^{-1} \sum_{i=1}^n r_{i,t}^2 / \tilde{h}_{i,t}$, at a quarterly (left panel) and semi-annual (right panel) frequency. Grey shaded areas represent NBER recession periods.

Figure 2: A GJR GARCH(1,1) is estimated for three international stock markets (FTSE 100, HSI and S&P 500) using daily return data for the July 1990 to June 2020 period. The figure shows the average volatility forecast errors, $\tilde{V}_t = n^{-1} \sum_{i=1}^n r_{i,t}^2 / \tilde{h}_{i,t}$, at a quarterly (left panel) and semi-annual (right panel) frequency. Grey shaded areas represent NBER recession periods.

Figure 3: The figure shows the annualized unconditional volatility, $\sqrt{252 \cdot \mathbf{Var}[r_t]}$, of the daily returns of the $(\text{MF})^2$ GARCH-rw- m as a function of $m = 21, \dots, 126$ for $\kappa = 3$ (red), $\kappa = 5$ (blue) and $\kappa = 7$ (green). The model parameters are chosen according to the estimates for the $(\text{MF})^2$ GARCH-rw-63 in Table 3.

Figure 4: News impact function for $(\text{MF})^2$ GARCH-rw-63. The parameters are: $\lambda_0 = 0.1$, $\lambda_1 = 0.56$, $\lambda_2 = 0.34$, $\alpha = 0.1$, $\beta = 0.88$. We plot $B_{n,t}^{MF} + (\lambda_1 \beta_n^1 + \lambda_2 \alpha) r_{n,t}^2$ as a function of the return, $r_{n,t}$, and for $\tau_t = 1$ (blue), $\tau_t = 1.5$ (red), $\tau_t = 0.5$ (green). The short-term component is fixed at its unconditional expectation: $h_{n,t} = 1$.

Figure 5: The figure shows the volatility forecasts from an $(MF)^2$ GARCH-rw-63 that was estimated for S&P 500 returns (see Table 3). The estimated conditional volatilities and long-term components are plotted as red and blue lines, respectively. The unconditional volatility is represented by the black vertical line. We compute volatility forecasts from day $t = 153$ (indicated by the black vertical line) onwards. The green dotted line corresponds to the forecasts from the $(MF)^2$ GARCH-rw-63. The dashed purple line shows the volatility forecasts from a simple GARCH(1,1). We impose the same conditional volatility at the forecast origin and the same unconditional volatility. In the upper/lower panels $t = 153$ corresponds to August 12, 2011 and April 16, 2018, respectively. All quantities are annualized.

Figure 6: The figure shows the estimated conditional volatilities, $\sqrt{h_t \cdot \tau_t}$, (red line) and long-term volatilities, $\sqrt{\tau_t}$, (black line) from the $(MF)^2$ GJR-rw-63 model in Table 3. All quantities are annualized.

Figure 7: The figure shows the estimated long-term volatilities, $\sqrt{\tau_t}$, for the $(MF)^2$ GJR-rw-63 (red line) and the $(MF)^2$ GJR-lf-63 (blue line) models. The dashed black line corresponds to the unconditional volatility of the MEM component of the $(MF)^2$ GJR-lf-63 model. For the parameter estimates see Table 3. All quantities are annualized.

Figure 8: The figure shows the effect of positive (left panels) and negative (right panels) news about the macro economy (upper panels), corporate earnings (middle panels) and monetary policy (lower panels) on the conditional volatility of the (MF)² GJR-rw-63 (see equation (32) and Table 3). The effect is plotted for 1 up to 21 days after the news event.

Tables

Table 1: Summary statistics: returns and realized volatilities.

	mean	sd	skewness	kurtosis	min	max	AC(1)
Panel A: Returns							
S&P 500	7.11	1.09	-1.05	28.46	-22.90	10.64	-0.02
FTSE 100	3.92	1.09	-0.25	10.04	-10.14	9.48	0.01
HSI	8.40	1.54	-0.13	12.08	-14.73	17.25	0.02
Panel B: Realized volatilities							
S&P 500	14.01	0.66	3.68	27.94	0.13	10.00	0.81
FTSE 100	14.71	0.62	3.78	31.09	0.12	10.30	0.74
HSI	18.56	0.77	3.34	20.65	0.28	8.56	0.61

Notes: Panels A and B show summary statistics for the returns and the realized volatilities of the S&P 500, the FTSE 100 and the HSI. The first column (mean) presents the mean of the annualized daily returns and the annualized daily realized volatilities. The remaining columns report summary statistics for the daily (non-annualized) returns/volatilities. We report 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 2020. The return data for the FTSE 100 and the HSI cover the period March 1990 to June 2020. Daily realized volatilities are available for the period January 2000 to June 2020.

Table 2: Predictability of realized variance and volatility forecast errors.

	(1) $\widetilde{RV}_t$	(2) $\widetilde{RV}_t$	(3) $\widetilde{V}_t$	(4) $\widetilde{V}_t$
Lag 1	0.313*** (0.106)	0.297*** (0.100)	0.065 (0.063)	0.020 (0.065)
Lag 2	0.024 (0.050)	-0.002 (0.058)	0.242*** (0.058)	0.196*** (0.058)
Lag 3	0.024 (0.039)	0.021 (0.039)	0.258*** (0.083)	0.257*** (0.080)
Lag 4	-0.002 (0.017)	0.001 (0.017)	-0.094** (0.048)	-0.084* (0.046)
GDP ^{SPF} _{<i>t</i> <i>t</i>-1}		-7.084 (4.397)		-2.102*** (0.494)
adj. R^2	0.085	0.089	0.122	0.159

Notes: The table presents quarterly regressions for the 1972Q1 to 2020Q2 period. Quarterly realized variances are defined as $\widetilde{RV}_t = \sum_{i=1}^n r_{i,t}^2$. V_t is defined as in equation (3). GDP^{SPF}_{*t*|*t*-1} is the median expectation from the Survey of Professional Forecasters for GDP growth in quarter t based on information from quarter $t - 1$. The numbers in the columns show the parameter estimates and HAC standard errors in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Table 3: (MF)² GARCH models: S&P 500.

τ_t	α	γ	β	λ_0	λ_1	λ_2	ω	LLF	BIC
GARCH									
const.	0.089*** (0.015)		0.897*** (0.015)	1.193*** (0.250)				-15764.76	2.532
const.	0.022*** (0.005)	0.114*** (0.020)	0.902*** (0.014)	1.008*** (0.140)				-15639.02	2.512
(MF) ² GARCH-rw- m									
$m = 126$	0.096*** (0.015)		0.869*** (0.019)	0.006** (0.003)	0.028* (0.016)	0.966*** (0.018)		-15742.16	2.530
$m = 126$	0.011* (0.006)	0.142*** (0.023)	0.868*** (0.018)	0.017 (0.014)	0.097 (0.088)	0.885*** (0.102)		-15591.25	2.506
$m = 63$	0.098*** (0.015)		0.853*** (0.022)	0.005 (0.003)	0.030 (0.022)	0.965*** (0.024)		-15737.75	2.529
$m = 63$	0.004 (0.007)	0.158*** (0.021)	0.843*** (0.019)	0.015* (0.008)	0.100** (0.050)	0.885*** (0.057)		-15572.62	2.503
$m = 21$	0.100*** (0.016)		0.839*** (0.034)	0.003** (0.001)	0.018* (0.010)	0.979*** (0.011)		-15747.18	2.531
$m = 21$	0.002 (0.007)	0.170*** (0.022)	0.817*** (0.026)	0.004** (0.002)	0.034*** (0.013)	0.962*** (0.014)		-15580.74	2.505
(MF) ² GARCH-bw- m									
$m = 252$	0.097*** (0.015)		0.859*** (0.028)	0.005*** (0.002)	0.025*** (0.009)	0.971*** (0.009)	5.011 (3.242)	-15742.48	2.531
$m = 252$	0.000 (0.008)	0.165*** (0.022)	0.825*** (0.030)	0.007** (0.003)	0.049*** (0.017)	0.944*** (0.019)	12.418*** (4.288)	-15579.45	2.505
$m = 126$	0.097*** (0.016)		0.861*** (0.030)	0.005*** (0.002)	0.024*** (0.009)	0.971*** (0.010)	1.967 (1.952)	-15741.76	2.530
$m = 126$	0.001 (0.009)	0.165*** (0.023)	0.827*** (0.032)	0.008** (0.003)	0.054** (0.022)	0.938*** (0.025)	5.115*** (1.952)	-15579.07	2.505
$m = 63$	0.098*** (0.015)		0.853*** (0.022)	0.005 (0.006)	0.030 (0.030)	0.964*** (0.035)	1.000 (1.214)	-15737.75	2.530
$m = 63$	0.004 (0.010)	0.158*** (0.027)	0.843*** (0.045)	0.015 (0.026)	0.100 (0.162)	0.885*** (0.188)	1.000 (1.952)	-15572.62	2.504
(MF) ² GARCH-lf- n									
$n = 126$	0.094*** (0.014)		0.880*** (0.017)	0.266*** (0.100)	0.749*** (0.228)	0.000 (0.276)		-15745.29	2.530
$n = 126$	0.019*** (0.005)	0.128*** (0.021)	0.885*** (0.016)	0.227*** (0.076)	0.767*** (0.175)	0.000 (0.206)		-15605.57	2.509
$n = 63$	0.097*** (0.014)		0.868*** (0.017)	0.138*** (0.043)	0.539*** (0.127)	0.328*** (0.123)		-15744.13	2.530
$n = 63$	0.013** (0.006)	0.144*** (0.022)	0.868*** (0.017)	0.148*** (0.046)	0.750*** (0.136)	0.099 (0.138)		-15588.71	2.506
$n = 21$	0.097*** (0.015)		0.855*** (0.028)	0.045*** (0.015)	0.211** (0.097)	0.744*** (0.103)		-15750.93	2.531
$n = 21$	0.006 (0.007)	0.162*** (0.022)	0.837*** (0.022)	0.055*** (0.018)	0.363*** (0.090)	0.581*** (0.094)		-15587.37	2.506

Notes: The table reports estimation results for (GJR) GARCH models, (MF)² GARCH-rw- m , (MF)² GARCH-bw- m and (MF)² GARCH-lf- n models. All models are estimated using daily return data for the period January 1971 to June 2020. 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 4: $(MF)^2$ GARCH: International stock markets.

τ_t	α	γ	β	λ_0	λ_1	λ_2	ω	LLF	BIC
Panel A: FTSE 100									
GJR GARCH									
const.	0.009 (0.007)	0.134*** (0.016)	0.905*** (0.012)	0.945*** (0.126)				-9497.04	2.461
$(MF)^2$ GARCH-rw- m									
$m = 126$	0.000 (0.008)	0.157*** (0.016)	0.883*** (0.015)	0.015** (0.007)	0.081** (0.037)	0.902*** (0.042)		-9470.62	2.457
$m = 63$	0.000 (0.008)	0.164*** (0.016)	0.868*** (0.017)	0.006** (0.003)	0.044** (0.017)	0.949*** (0.019)		-9465.96	2.456
$m = 21$	0.000 (0.008)	0.171*** (0.018)	0.853*** (0.023)	0.003*** (0.001)	0.027*** (0.009)	0.970*** (0.010)		-9469.41	2.457
$(MF)^2$ GARCH-bw- m									
$m = 252$	0.000 (0.009)	0.165*** (0.018)	0.866*** (0.021)	0.005*** (0.002)	0.036*** (0.009)	0.958*** (0.009)	8.141*** (2.344)	-9468.89	2.458
$m = 126$	0.000 (0.008)	0.165*** (0.018)	0.867*** (0.021)	0.006*** (0.002)	0.041*** (0.009)	0.953*** (0.009)	3.150** (1.265)	-9468.13	2.457
$m = 63$	0.000 (0.008)	0.164*** (0.016)	0.868*** (0.016)	0.006* (0.004)	0.044* (0.024)	0.949*** (0.028)	1.000* (0.560)	-9465.96	2.457
$(MF)^2$ GARCH-lf- n									
$n = 126$	0.000 (0.007)	0.151*** (0.017)	0.895*** (0.013)	0.218*** (0.059)	0.758*** (0.180)	0.000 (0.185)		-9478.34	2.459
$n = 63$	0.000 (0.007)	0.163*** (0.016)	0.879*** (0.013)	0.159*** (0.047)	0.794*** (0.142)	0.031 (0.146)		-9471.96	2.457
$n = 21$	0.000 (0.008)	0.170*** (0.019)	0.862*** (0.023)	0.048*** (0.015)	0.346*** (0.091)	0.602*** (0.091)		-9471.79	2.457
Panel B: HSI									
GJR GARCH									
const.	0.034*** (0.006)	0.069*** (0.012)	0.918*** (0.009)	2.032*** (0.323)				-11943.91	3.154
$(MF)^2$ GARCH-rw- m									
$m = 126$	0.016** (0.007)	0.096*** (0.017)	0.891*** (0.017)	0.117** (0.054)	0.376*** (0.144)	0.563*** (0.163)		-11923.04	3.151
$m = 63$	0.000 (0.007)	0.118*** (0.019)	0.854*** (0.027)	0.160*** (0.060)	0.601*** (0.149)	0.321* (0.165)		-11908.17	3.147
$m = 21$	0.004 (0.007)	0.141*** (0.018)	0.788*** (0.031)	0.014** (0.006)	0.055*** (0.016)	0.939*** (0.018)		-11904.94	3.146
$(MF)^2$ GARCH-bw- m									
$m = 252$	0.000 (0.011)	0.136*** (0.018)	0.787*** (0.035)	0.026 (0.037)	0.104 (0.150)	0.884*** (0.167)	14.155 (11.763)	-11902.04	3.147
$m = 126$	0.000 (0.019)	0.136*** (0.019)	0.788*** (0.035)	0.026 (0.094)	0.104 (0.373)	0.884** (0.418)	6.826 (14.361)	-11902.12	3.147
$m = 63$	0.000 (0.009)	0.133*** (0.019)	0.795*** (0.037)	0.043 (0.041)	0.167 (0.149)	0.813*** (0.168)	2.337*** (0.902)	-11901.93	3.147
$(MF)^2$ GARCH-lf- n									
$n = 126$	0.031*** (0.006)	0.080*** (0.015)	0.905*** (0.015)	0.548** (0.215)	0.711** (0.308)	0.000 (0.362)		-11938.44	3.155
$n = 63$	0.027*** (0.007)	0.092*** (0.017)	0.889*** (0.018)	0.294*** (0.100)	0.668*** (0.176)	0.180 (0.185)		-11932.01	3.154
$n = 21$	0.013* (0.007)	0.128*** (0.021)	0.825*** (0.031)	0.133*** (0.044)	0.475*** (0.089)	0.459*** (0.094)		-11918.94	3.150
Notes: The table reports estimation results for (GJR) GARCH models, $(MF)^2$ GARCH-rw- m , $(MF)^2$ GARCH-bw- m and $(MF)^2$ GARCH-lf- n models. All models are estimated using daily return data for the period March 1990 to June 2020. 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 5: Determinants of low-frequency long-term volatility.

	dependent variable: τ_{t+1}				
	(1)	(2)	(3)	(4)	(5)
	monthly	monthly	quarterly	quarterly	quarterly
Overall _t	0.095*** (0.012)		0.147*** (0.015)		
Macro News _t		0.071** (0.031)		0.151*** (0.062)	0.130** (0.062)
Monetary Policy _t		-0.039** (0.019)		-0.077** (0.042)	-0.068 (0.042)
Policy Related _t		0.061** (0.024)		0.068* (0.043)	0.075* (0.043)
GDP ^{SPF} _{t+1 t}					-0.035** (0.015)
τ_t	0.876*** (0.029)	0.870*** (0.027)	0.714*** (0.032)	0.702*** (0.015)	0.676*** (0.039)
adj. R^2	0.924	0.924	0.823	0.825	0.830

Notes: Monthly and quarterly regressions are for the January 1985 to June 2020 period. τ_{t+1} is taken from an (MF)² GARCH-lf model with monthly/quarterly long-term component. The EMV-trackers Overall_t, Macro News_t, Monetary Policy_t, Policy Related_t are from Baker et al. (2019a) (see Section 4.2.2). GDP^{SPF}_{t|t-1} is the median expectation from the Survey of Professional Forecasters for GDP growth in quarter $t+1$ based on information from quarter t . The numbers in the columns show the parameter estimates and HAC standard errors in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Table 6: News effects on daily short- and long-term component.

	(1)	(2)	(3)	(4)
	r_t^2/τ_t	Δh_{t+1}	r_t^2/h_t	$\Delta\tau_{t+1}$
Macro _t × $\mathbf{1}_{(r_t > 0)}$	6.803*** (0.808)	-0.369** (0.153)	7.362*** (1.098)	0.020*** (0.003)
Macro _t × $\mathbf{1}_{(r_t < 0)}$	8.137*** (0.991)	2.529*** (0.748)	17.301*** (4.836)	0.016*** (0.003)
Corporate Earnings _t × $\mathbf{1}_{(r_t > 0)}$	7.571*** (0.915)	-0.398* (0.241)	7.267*** (0.837)	0.016*** (0.004)
Corporate Earnings _t × $\mathbf{1}_{(r_t < 0)}$	8.047*** (1.843)	0.885** (0.402)	5.738** (2.753)	0.017*** (0.004)
Monetary Policy _t × $\mathbf{1}_{(r_t > 0)}$	6.132*** (0.875)	-1.118** (0.466)	13.809*** (2.257)	0.020*** (0.004)
Monetary Policy _t × $\mathbf{1}_{(r_t < 0)}$	11.841*** (2.676)	2.463*** (0.639)	14.868*** (4.365)	0.024*** (0.006)
adj. R^2	0.202	0.063	0.082	0.054

Notes: Daily regressions ($n = 1$) are for the January 1973 to June 2020 period. τ_{t+1} and h_{t+1} are taken from the (MF)² GARCH-rw-63 estimates in Table 3. The numbers in the columns show the parameter estimates and HAC standard errors in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Table 7: Out-of-sample forecasting: S&P 500 – relative RMSE.

	forecast horizon				
	daily	monthly	quarterly	semi-annual	annual
(MF) ² GARCH-rw-126	0.793	0.716	0.723	0.729	0.743
(MF) ² GJR-rw-126	0.751	0.693	0.672	0.657	0.688
(MF) ² GARCH-rw-63	0.789	0.707	0.719	0.731	0.744
(MF) ² GJR-rw-63	0.741	0.684	0.692	0.687	0.705
(MF) ² GARCH-rw-21	0.786	0.707	0.728	0.742	0.752
(MF) ² GJR-rw-21	0.734	0.684	0.703	0.701	0.715
(MF) ² GARCH-bw-252	0.847	0.807	0.783	0.742	0.748
(MF) ² GJR-bw-252	0.737	0.684	0.703	0.700	0.714
(MF) ² GARCH-bw-126	0.818	0.785	0.779	0.780	0.779
(MF) ² GJR-bw-126	0.737	0.684	0.702	0.699	0.713
GARCH	0.809	0.781	0.851	0.910	0.876
GJR	0.782	0.788	0.806	0.810	0.771
HAR	0.771	0.711	0.759	0.800	0.875
log-HAR	0.760	0.687	0.740	0.781	0.846
log-HAR leverage	0.756	0.713	0.729	0.784	1.123
RW 1	0.907	1.278	1.520	1.730	1.761
RW 2	0.907	0.771	0.857	0.858	0.793

Notes: The out-of-sample period is January 2010 to June 2020. The first in-sample period is January 1971 to December 2009. Using a rolling window scheme, all models are re-estimated on a monthly basis (i.e. every 21 days). Numbers reported are the RMSE losses for each model and forecast horizon relative to the respective RMSE of historical volatility. Red-colored numbers indicate the model with the lowest RMSE. Blue-shaded numbers indicate that the respective model is included in the 85% model confidence set.

Table 8: Out-of-sample forecasting: FTSE 100 and HSI – relative RMSE.

	forecast horizon					
	quarterly		semi-annual		annual	
	FTSE	HSI	FTSE	HSI	FTSE	HSI
(MF) ² GARCH-rw-126	0.721	0.382	0.715	0.456	0.743	0.522
(MF) ² GJR-rw-126	0.672	0.310	0.635	0.334	0.641	0.341
(MF) ² GARCH-rw-63	0.719	0.378	0.702	0.450	0.744	0.539
(MF) ² GJR-rw-63	0.680	0.325	0.644	0.343	0.658	0.373
(MF) ² GARCH-rw-21	0.719	0.415	0.707	0.515	0.752	0.641
(MF) ² GJR-rw-21	0.689	0.355	0.656	0.399	0.674	0.461
(MF) ² GARCH-bw-252	0.691	0.487	0.657	0.465	0.674	0.439
(MF) ² GJR-bw-252	0.685	0.353	0.651	0.392	0.665	0.447
(MF) ² GARCH-bw-126	0.681	0.403	0.661	0.406	0.698	0.454
(MF) ² GJR-bw-126	0.685	0.355	0.650	0.394	0.665	0.450
GARCH	0.842	0.443	0.892	0.602	0.957	0.858
GJR	0.753	0.403	0.728	0.494	0.719	0.569
HAR	0.744	0.403	0.756	0.429	0.822	0.536
log-HAR	0.740	0.355	0.755	0.412	0.871	0.679
log-HAR leverage	0.725	0.347	0.751	0.399	1.174	0.653
RW 1	1.884	2.070	2.073	2.418	2.175	2.828
RW 2	0.827	0.417	0.832	0.471	0.842	0.554

Notes: The out-of-sample period is January 2010 to June 2020. The first in-sample period is January 1971 to December 2009. Using a rolling window scheme, all models are re-estimated on a monthly basis (i.e. every 21 days). Numbers reported are the RMSE losses for each model and forecast horizon relative to the respective RMSE of historical volatility. Red-colored numbers indicate the model with the lowest RMSE. Blue-shaded numbers indicate that the respective model is included in the 85% model confidence set.

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

$$\begin{aligned}
\mathbf{E}[h_t\tau_t] &= ((1 - \alpha - \beta) + (\alpha Z_{t-1}^2 + \beta)h_{t-1}) \times \left(\lambda_0 + \lambda_1 \frac{1}{m} \sum_{j=1}^m Z_{t-j}^2 \tau_{t-j} + \lambda_2 \tau_{t-1} \right) \\
&= (1 - \alpha - \beta) (\lambda_0 + \lambda_1 \mathbf{E}[\tau_t] + \lambda_2 \mathbf{E}[\tau_t]) \\
&\quad + \lambda_0(\alpha + \beta) \mathbf{E}[h_{t-1}] + \lambda_1 \frac{1}{m} (\alpha\kappa + \beta) \mathbf{E}[h_{t-1}\tau_{t-1}] + \lambda_2(\alpha + \beta) \mathbf{E}[h_{t-1}\tau_{t-1}] \\
&\quad + \lambda_1(\alpha + \beta) \frac{1}{m} \sum_{j=2}^m \mathbf{E}[h_{t-1}Z_{t-j}^2 \tau_{t-j}]. \tag{34}
\end{aligned}$$

Note that we can write

$$h_{t-1} = (1 - \alpha - \beta) \left(1 + \sum_{k=1}^{j-2} \prod_{s=1}^k (\alpha Z_{t-1-s}^2 + \beta) \right) + h_{t-j} \prod_{k=1}^{j-1} (\alpha Z_{t-1-k}^2 + \beta). \tag{35}$$

Using equation (35), we can express the expectation $\mathbf{E}[h_{t-1}Z_{t-j}^2\tau_{t-j}]$ as

$$\begin{aligned}
\mathbf{E}[h_{t-1}Z_{t-j}^2\tau_{t-j}] &= (1 - \alpha - \beta) \mathbf{E}[\tau_t] \left(1 + \sum_{k=1}^{j-2} (\alpha + \beta)^k \right) \\
&\quad + (\alpha + \beta)^{j-2} (\alpha\kappa + \beta) \mathbf{E}[h_{t-j}\tau_{t-j}]. \tag{36}
\end{aligned}$$

Plugging this into equation (34) leads to

$$\begin{aligned}
\mathbf{E}[h_t\tau_t] &= (1 - \alpha - \beta) \frac{\lambda_0}{1 - \lambda_1 - \lambda_2} + \lambda_0(\alpha + \beta) \\
&\quad + (1 - \alpha - \beta) \lambda_1(\alpha + \beta) \mathbf{E}[\tau_t] \left(\frac{m-1}{m} + \frac{1}{m} \sum_{j=2}^m \sum_{k=1}^{j-2} (\alpha + \beta)^k \right) \\
&\quad + \left(\lambda_1 \frac{1}{m} (\alpha\kappa + \beta) + \lambda_2(\alpha + \beta) \right) \mathbf{E}[h_{t-1}\tau_{t-1}] \\
&\quad + \lambda_1(\alpha\kappa + \beta)(\alpha + \beta) \frac{1}{m} \sum_{j=2}^m (\alpha + \beta)^{j-2} \mathbf{E}[h_{t-j}\tau_{t-j}]. \tag{37}
\end{aligned}$$

If r_t is covariance stationary, then

$$\left(\lambda_1 \frac{1}{m} (\alpha\kappa + \beta) + \lambda_2(\alpha + \beta) \right) - \lambda_1(\alpha\kappa + \beta) \frac{1}{m} \sum_{j=2}^m (\alpha + \beta)^{j-1} < 1. \tag{38}$$

It then follows that $\mathbf{Var}[r_t]$ is given by equation (13).

■

Proof of Corollary 1. If r_t is covariance stationary, we have

$$\begin{aligned}
\mathbf{E}[\tau_t h_t] &= \mathbf{E} \left[\left(\lambda_0 + \lambda_1 \frac{r_{t-1}^2}{h_{t-1}} + \lambda_2 \tau_{t-1} \right) \left((1 - \alpha - \beta) + \alpha \frac{r_{t-1}^2}{\tau_{t-1}} + \beta h_{t-1} \right) \right] \\
&= \lambda_0(1 - \alpha - \beta) + \lambda_0 \alpha \mathbf{E} \left[\frac{r_{t-1}^2}{\tau_{t-1}} \right] + \lambda_0 \beta \mathbf{E}[h_{t-1}] \\
&\quad + \lambda_1(1 - \alpha - \beta) \mathbf{E} \left[\frac{r_{t-1}^2}{h_{t-1}} \right] + \alpha \lambda_1 \mathbf{E} \left[\frac{r_{t-1}^4}{h_{t-1} \tau_{t-1}} \right] + \lambda_1 \beta \mathbf{E}[r_{t-1}^2] \\
&\quad + \lambda_2(1 - \alpha - \beta) \mathbf{E}[\tau_{t-1}] + \alpha \lambda_2 \mathbf{E}[r_{t-1}^2] + \beta \lambda_2 \mathbf{E}[\tau_{t-1} h_{t-1}] \\
&= \lambda_0 + (1 - \alpha - \beta)(\lambda_1 + \lambda_2) \mathbf{E}[\tau_t] + [\lambda_1(\alpha \kappa + \beta) + \lambda_2(\alpha + \beta)] \mathbf{E}[\tau_t h_t], \quad (39)
\end{aligned}$$

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 (15).

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

$$\bar{h}_t = (1 - \alpha - \beta) \left(1 + \sum_{i=1}^{\infty} \prod_{j=1}^i (\alpha Z_{t-j}^2 + \beta) \right) \quad (40)$$

and

$$\bar{\tau}_t = \lambda_0 \left(1 + \sum_{k=1}^{\infty} \prod_{j=1}^k (\lambda_1 Z_{t-j}^2 + \lambda_2) \right). \quad (41)$$

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

$$\psi_i = \prod_{j=1}^i (\alpha Z_{t-j}^2 + \beta) \quad \text{and} \quad \delta_k = \prod_{j=1}^k (\lambda_1 Z_{t-j}^2 + \lambda_2). \quad (42)$$

It follows that

$$\mathbf{E}[\psi_i \delta_k] = [\lambda_1(\alpha \kappa + \beta) + \lambda_2(\alpha + \beta)]^k (\alpha + \beta)^{i-k} \text{ for } k \leq i \quad (43)$$

and

$$\mathbf{E}[\psi_i \delta_k] = [\lambda_1(\alpha \kappa + \beta) + \lambda_2(\alpha + \beta)]^i (\lambda_1 + \lambda_2)^{k-i} \text{ for } i < k. \quad (44)$$

Thus, we can write

$$\mathbf{E}[h_t \tau_t] = \lambda_0(1 - \alpha - \beta) \mathbf{E} \left[\left(\sum_{i=0}^{\infty} \psi_i \right) \left(\sum_{k=0}^{\infty} \delta_k \right) \right], \quad (45)$$

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

$$\begin{aligned} \mathbf{E} \left(\sum_{i=0}^{\infty} \psi_i \right) \left(\sum_{k=0}^{\infty} \delta_k \right) &= \sum_{k=0}^{\infty} \sum_{i=k}^{\infty} \mathbf{E}[\psi_i \delta_k] + \sum_{i=0}^{\infty} \sum_{k=i+1}^{\infty} \mathbf{E}[\psi_i \delta_k] \\ &= \sum_{k=0}^{\infty} [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]^k \sum_{i=0}^{\infty} (\alpha + \beta)^i \\ &\quad + \sum_{i=0}^{\infty} [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]^i (\lambda_1 + \lambda_2) \sum_{k=0}^{\infty} (\lambda_1 + \lambda_2)^k \\ &= \frac{1}{[1 - (\alpha + \beta)](1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)])} \\ &\quad + \frac{(\lambda_1 + \lambda_2)}{[1 - (\lambda_1 + \lambda_2)](1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)])}. \end{aligned} \quad (46)$$

In summary, we get

$$\begin{aligned} \mathbf{E}[h_t \tau_t] &= \frac{\lambda_0}{(1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)])} \\ &\quad + \frac{\lambda_0(1 - \alpha - \beta)(\lambda_1 + \lambda_2)}{[1 - (\lambda_1 + \lambda_2)](1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)])} \\ &= \frac{\lambda_0 + (1 - \alpha - \beta)(\lambda_1 + \lambda_2) \frac{\lambda_0}{[1 - (\lambda_1 + \lambda_2)]}}{1 - [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)]}. \end{aligned} \quad (47)$$

■

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

$$\begin{aligned}
\mathbf{E}[h_{t+s}\tau_{t+s}|\mathcal{F}_t] &= \mathbf{E}[(1 - \alpha - \beta) + (\alpha Z_{t+s-1}^2 + \beta)h_{t+s-1}) \\
&\quad \times \left(\lambda_0 + \lambda_1 \frac{1}{m} \sum_{j=1}^{s-1} Z_{t+s-j}^2 \tau_{t+s-j} + \lambda_1 \frac{1}{m} \sum_{j=s}^m \frac{r_{t+s-j}^2}{h_{t+s-j}} + \lambda_2 \tau_{t+s-1} \right) \Big| \mathcal{F}_t] \\
&= (1 - \alpha - \beta) \left(\lambda_0 + (\lambda_1 \frac{1}{m} + \lambda_2) \mathbf{E}[\tau_{t+s-1}|\mathcal{F}_t] + \lambda_1 \frac{1}{m} \sum_{j=2}^{s-1} \mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t] \right) \\
&\quad + (1 - \alpha - \beta) \lambda_1 \frac{1}{m} \sum_{j=s}^m \frac{r_{t+s-j}^2}{h_{t+s-j}} + \lambda_0(\alpha + \beta) \mathbf{E}[h_{t+s-1}|\mathcal{F}_t] \\
&\quad + \left(\lambda_1 \frac{1}{m} (\alpha\kappa + \beta) + \lambda_2(\alpha + \beta) \right) \mathbf{E}[h_{t+s-1}\tau_{t+s-1}|\mathcal{F}_t] \\
&\quad + \lambda_1(\alpha + \beta) \mathbf{E}[h_{t+s-1}|\mathcal{F}_t] \frac{1}{m} \sum_{j=s}^m \frac{r_{t+s-j}^2}{h_{t+s-j}} \\
&\quad + \lambda_1(\alpha + \beta) \frac{1}{m} \sum_{j=2}^{s-1} \mathbf{E}[h_{t+s-1} Z_{t+s-j}^2 \tau_{t+s-j}|\mathcal{F}_t]. \tag{48}
\end{aligned}$$

A conditional version of equation (36) is given by

$$\begin{aligned}
\mathbf{E}[h_{t+s-1} Z_{t+s-j}^2 \tau_{t+s-j}|\mathcal{F}_t] &= (1 - \alpha - \beta) \mathbf{E}[\tau_{t+s-j}|\mathcal{F}_t] \left(1 + \sum_{k=1}^{j-2} (\alpha + \beta)^k \right) \\
&\quad + (\alpha + \beta)^{j-2} (\alpha\kappa + \beta) \mathbf{E}[h_{t+s-j} \tau_{t+s-j}|\mathcal{F}_t]. \tag{49}
\end{aligned}$$

Plugging this into equation (48) leads to

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

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

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

Again, using equation (36) the forecasts can be recursively calculated as

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

■

Proof of Theorem 3. We use the following representations:

$$h_{1,t} = (1-\alpha-\beta) \left(1 + \sum_{i=1}^{n-1} \prod_{j=i+1}^n (\alpha Z_{j,t-1}^2 + \beta) \right) + h_{1,t-1} \prod_{i=1}^n (\alpha Z_{i,t-1}^2 + \beta)$$

and

$$\tau_t = \lambda_0 + \left(\lambda_1 \frac{1}{n} \sum_{i=1}^n Z_{i,t-1}^2 + \lambda_2 \right) \tau_{t-1}.$$

It follows that

$$\begin{aligned}
\mathbf{E}[h_{1,t}\tau_t] &= \lambda_0(1 - \alpha - \beta) \left(1 + \sum_{i=1}^{n-1} (\alpha + \beta)^{n-i} \right) \\
&+ \lambda_0(\alpha + \beta)^n \\
&+ (1 - \alpha - \beta) \left[(\lambda_1 + \lambda_2) + \right. \\
&\quad \left. + \frac{1}{n} \lambda_1 \left(\sum_{i=1}^{n-1} (\alpha + \beta)^{n-i} \right) \right. \\
&\quad \left. + (\alpha\kappa + \beta)(\alpha + \beta)^{n-2} + \sum_{i=2}^{n-1} (\alpha + \beta)^{n-i} \right. \\
&\quad \left. + (\alpha\kappa + \beta)(\alpha + \beta)^{n-2} + (\alpha\kappa + \beta)(\alpha + \beta)^{n-3} + \sum_{i=3}^{n-1} (\alpha + \beta)^{n-i} \right. \\
&\quad \left. \vdots \right. \\
&\quad \left. + (\alpha\kappa + \beta)(\alpha + \beta)^{n-2} + (\alpha\kappa + \beta)(\alpha + \beta)^{n-3} + \dots + (\alpha\kappa + \beta)(\alpha + \beta) + (\alpha\kappa + \beta) \right) \\
&+ \lambda_2 \sum_{i=1}^{n-1} (\alpha + \beta)^{n-i} \Big] \mathbf{E}[\tau_t] \\
&+ (\alpha + \beta)^{n-1} [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)] \mathbf{E}[h_{1,t-1}\tau_{t-1}].
\end{aligned}$$

If $r_{1,t}$ is covariance stationary, then

$$(\alpha + \beta)^{n-1} [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)] < 1.$$

From recursively plugging in $\mathbf{E}[h_{1,t-k}\tau_{t-k}]$, for $k = 1, \dots$, it follows that $\mathbf{E}[h_{1,t}\tau_t]$ will be finite if $(\alpha + \beta)^{n-1} [\lambda_1(\alpha\kappa + \beta) + \lambda_2(\alpha + \beta)] < 1$. Solving for $\mathbf{E}[h_{1,t}\tau_t]$ yields the unconditional variance of $r_{1,t}$.

For $i > 1$, we obtain

$$\begin{aligned}
\mathbf{E}[h_{i,t}\tau_t] &= \mathbf{E}[(1 - \alpha - \beta) + (\alpha Z_{i-1,t}^2 + \beta)h_{i-1,t})\tau_t] \\
&= (1 - \alpha - \beta) \frac{\lambda_0}{1 - \lambda_1 - \lambda_2} + (\alpha + \beta) \mathbf{E}[h_{i-1,t}\tau_t].
\end{aligned}$$

These expectations can be recursively computed.

■

Proof of Corollary 2. When n is sufficiently large, the denominator in equation (25) is

essentially 1. The term Ψ_n is then approximately given by

$$\begin{aligned}\Psi_n &\approx (1 - \alpha - \beta) \left[(\lambda_1 + \lambda_2) + \frac{\lambda_1}{1 - \alpha - \beta} + \frac{\lambda_2}{1 - \alpha - \beta} \right] \\ &= \lambda_1 + \lambda_2,\end{aligned}$$

where we use that the term $\frac{1}{n} \sum_{j=1}^{n-1} (n-j)(\alpha + \beta)^{n-1-j}$ converges to zero.

■

Online Appendix

A Additional Figures

Figure A.1: The figure shows the unconditional variances $\mathbf{Var}[r_{i,t}]$, $i = 1, \dots, n$, for three consecutive low-frequency periods of an $(\text{MF})^2$ GARCH-M ($n = 21$) and $\kappa = 3$ (red), $\kappa = 5$ (blue) and $\kappa = 7$ (green). The model parameters are chosen as $\alpha = 0.05$, $\beta = 0.85$, $\lambda_0 = 0.04$, $\lambda_1 = 0.3$, $\lambda_2 = 0.65$.

Figure A.2: We estimate $(\text{MF})^2$ GARCH-rw- m models for daily S&P 500 returns for the period January 1971 to June 2020. We begin with $m = 21$ and then increase m by steps of 5. The figure shows the BIC as a function of m .

Figure A.3: The figure shows the estimated weighting schemes for the $(MF)^2$ GJR-bw- m models with $m = 252$ and $m = 126$ that were estimated for the S&P 500 (see Table 3).

Figure A.4: The figure shows the effect of positive (left panels) and negative (right panels) news about the macro economy on the short-term component (upper panels), and the long-term component (lower panels) of the $(MF)^2$ GJR-rw-63. The effect is plotted for 1 up to 21 days after the news event.

Figure A.5: The figure shows the effect of positive (left panels) and negative (right panels) news about the macro economy (upper panels), corporate earnings (middle panels) and monetary policy (lower panels) on the conditional volatility of a GJR GARCH. The effect is plotted for 1 up to 21 days after the news event.

B Theoretical Results for (MF)² GJR-rw- m

Assumption 4. *The parameters of the short-term GJR component satisfy the conditions $\alpha > 0$, $\alpha + \gamma > 0$, $\beta > 0$ and $\alpha + \gamma/2 + \beta < 1$.*

Theorem 4. *Let Assumptions 1, 3, and 4 be satisfied. If the (MF)² GJR-rw- m process, $(r_t)_{t \in \mathbb{Z}}$, is covariance stationary, then*

$$\left(\lambda_1 \frac{1}{m} ((\alpha + \gamma/2)\kappa + \beta) + \lambda_2 (\alpha + \gamma/2\beta) \right) + \lambda_1 ((\alpha + \gamma/2)\kappa + \beta) \frac{1}{m} \sum_{j=2}^m (\alpha + \gamma/2 + \beta)^{j-1} < 1.$$

The unconditional variance of the daily returns is given by

$$\text{Var}[r_t] = \frac{\lambda_0 + \frac{\lambda_0}{1-\lambda_1-\lambda_2} (1 - \alpha - \gamma/2 - \beta)(\lambda_1 + \lambda_2) + \Delta_m}{1 - \left(\lambda_1 \frac{1}{m} ((\alpha + \gamma/2)\kappa + \beta) + \lambda_2 (\alpha + \gamma/2 + \beta) \right) - \lambda_1 ((\alpha + \gamma/2)\kappa + \beta) \frac{1}{m} \sum_{j=2}^m (\alpha + \gamma/2 + \beta)^{j-1}},$$

where

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

The proof is similar as the Proof of Theorem 1.

If h_t is given by the GJR GARCH, the forecast of the short-term component can be computed as $\mathbf{E}[h_{t+s}|\mathcal{F}_t] = 1 + (\alpha + \gamma/2 + \beta)^{s-1}(h_{t+1} - 1)$. The forecasts of the long-term component are the same as before.

Theorem 5. *Let Assumptions 1, 3, 4 and the constraint in equation (12) be satisfied. Then, in the (MF)² GJR-rw- m the forecast of the conditional variance on day $t+s$, $s \geq 1$, can be computed as follows:*

$$1. \quad \mathbf{E}[h_{t+1}\tau_{t+1}|\mathcal{F}_t] = h_{t+1}\tau_{t+1}$$

2. *For $s = 2, \dots, m$ the forecasts can be recursively calculated as*

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

3. For $s > m$ the following recursion applies:

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

The proof is similar as the Proof of Theorem 2.