Capturing Volatility Persistence
A Dynamically Complete Realized EGARCH-MIDAS Model

Borup, Daniel; Jakobsen, Johan S.

Capturing Volatility Persistence: A Dynamically Complete Realized EGARCH-MIDAS Model

Daniel Borup and Johan S. Jakobsen

Capturing volatility persistence: A dynamically complete Realized EGARCH-MIDAS model*

Daniel Borup[†] Johan S. Jakobsen[**]

Abstract

We introduce extensions of the Realized Exponential GARCH model (REGARCH) that capture the evident high persistence typically observed in measures of financial market volatility in a tractable fashion. The extensions decompose conditional variance into a short-term and a long-term component. The latter utilizes mixed-data sampling or a heterogeneous autoregressive structure, avoiding parameter proliferation otherwise incurred by using the classical ARMA structures embedded in the REGARCH. The proposed models are dynamically complete, facilitating multi-period forecasting. A thorough empirical investigation with an exchange traded fund that tracks the S&P500 Index and 20 individual stocks shows that our models better capture the dependency structure of volatility. This leads to substantial improvements in empirical fit and predictive ability at both short and long horizons relative to the original REGARCH. A volatility-timing trading strategy shows that capturing volatility persistence yields substantial utility gains for a mean-variance investor at longer investment horizons.

Keywords: Realized Exponential GARCH; persistence; long memory; GARCH-MIDAS; HAR; realized kernel

JEL Classification: C10, C50, C51, C52, C53, C58

This version: February 7, 2019

*We thank Jim Gatheral (the Editor), two anonymous referees, Timo Teräsvirta, Asger Lunde, Peter Reinhard Hansen, Bent Jesper Christensen, Esther Ruiz Ortega, Charlotte Christiansen, Bezirgen Veliyev, Onno Kleen, Jorge Wolfgang Hansen and participants at research seminars at Aarhus University and the Frontiers of Factor Investing Conference (2018) at the Centre for Financial Econometrics, Asset Markets and Macroeconomic Policy, Lancaster University Management School, for useful comments and suggestions. We also thank Asger Lunde for providing cleaned high-frequency tick data. Programs for estimating our proposed models are available from the authors upon request. The authors acknowledge support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. Some of this research was carried out while D. Borup was visiting the Department of Economics, University of Pennsylvania, and the generosity and hospitality of the department is gratefully acknowledged. An earlier version of this paper was circulated under the title "Long-range dependence in the Realized (Exponential) GARCH framework".

[†]Corresponding author. CREATES, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark. Email: dborup@econ.au.dk, Tel.: +45 87165134.

[**]Copenhagen Business School and CREATES, Department of Economics and Business Economics, Aarhus University. Email: jsj@fi.dk.

I. Introduction

The Realized GARCH model (RGARCH) and Realized Exponential GARCH model (REGARCH) (Hansen, Huang, and Shek, 2012; Hansen and Huang, 2016) provide an advantageous structure for the joint modeling of stock returns and realized measures of their volatility. The models facilitate exploitation of granular information in high-frequency data by including realized measures, which constitute a much stronger signal of latent volatility than squared returns (Andersen, Bollerslev, Diebold, and Labys, 2001, 2003). Despite the empirical success of the R(E)GARCH models, these models do not adequately capture the dependence structure in volatility (both latent and realized) without proliferation in parameters. This dependence structure is typically characterized by a positive and slowly decaying autocorrelation function (long-range dependence) or a persistence parameter close to unity, known as the "integrated GARCH effect". Indeed, Hansen and Huang (2016) point out that the REGARCH does a good job at modeling returns, but falls short in describing the dynamic properties of the realized measure.

In this paper, we introduce parsimonious extensions of the REGARCH to capture this evident high persistence by means of a multiplicative decomposition of the conditional variance into a short-term and long-term component. The multiplicative decomposition was popularized by Feng (2004), Engle and Rangel (2008), and Engle, Ghysels, and Sohn (2013)), among others. This structure is particularly useful since it enables explicit modeling of a "baseline volatility", whose level arguably shifts over time, and is the basis around which short-term movements occur. This structure is appealing since it is intuitive and facilitates parsimonious specifications of a slow-moving component in volatility. Moreover, it allows for great flexibility as opposed to formal long-memory models employing, e.g., fractional integration. Whether the high persistence arises due to structural breaks, fractional integration or another source (see e.g. Lamoureux and Lastrapes (1990), Diebold and Inoue (2001), Hillebrand (2005), McCloskey and Perron (2013), and Varneskov and Perron (2017)) our proposed models are able to reproduce the high persistence of volatility observed in stock return data and alleviate the integrated GARCH effect. This plays an important role in stationarity of the short-term component and existence of the unconditional variance, but also provides a means to obtain improved multi-step forecasts by reducing the long-lasting impact of the short-term component and its innovations via faster convergence to the time-varying baseline volatility.

When specifying our models, we retain the dynamics of the short-term component like those from a first-order REGARCH, but model the long-term component either via mixed-data sampling (MIDAS) or a heterogeneous autoregressive (HAR) structure. Motivated by Engle et al. (2013), the former specifies the slow-moving component as a weighted average of weekly or monthly aggregates of the realized measure with the backward-looking window and weights estimated from the data. The latter is motivated by the simple, yet empirically successful HAR model of Corsi (2009), which approximates the dependencies in volatility by a simple additive cascade structure of a daily, weekly and monthly component of realized measures. Both our extensions introduce only two or three additional parameters, hence avoid parameter proliferation otherwise incurred by means of the classical ARMA structures embedded in the original REGARCH. Moreover, they remain dynamically complete. That is, the models fully characterize the dynamic properties of all variables included in the model. This property is especially relevant for forecasting purposes, since it allows for multi-period forecasting. This contrasts GARCH-X models, which only provide forecasts one period into the future, and related extensions including macroeconomic factors who rely on assumptions about the included variables' exogenous dynamics.

We apply our REGARCH-MIDAS and REGARCH-HAR to the exchange traded index fund, SPY, which tracks the S&P500 Index and 20 individual stocks and compare their performances to a quadratic REGARCH-Spline and a fractionally integrated REGARCH, the FloEGARCH (Vander Elst, 2015). We find that both our proposed models better capture the autocorrelation structure of latent and realized volatility relative to the original REGARCH, which is only able to capture the dependency over the very short term. This leads to substantial improvements in empirical fit (log-likelihood and information criteria) and predictive ability, particularly beyond shorter horizons, when benchmarked to the original REGARCH. We document, additionally, that the backward-looking horizon of the HAR specification is too short to sufficiently capture autocorrelation beyond approximately one month. While the REGARCH-Spline comes short relative to our proposals (with four-five extra parameters), the FloEGARCH performs well. It does, however, not perform better than our best-performing REGARCH-MIDAS specifications in-sample and lack predictive accuracy in the short-term. This leaves the REGARCH-MIDAS as a very attractive model for capturing volatility persistence in the REGARCH framework and improving forecasting performance at both short and long horizons. To assess the economic value of the improvements in predictive accuracy, we examine a volatility-timing strategy that uses each model's forecast as input to construct optimal portfolio weights. A risk-averse investor with mean-variance preferences who allocates funds into one risky asset and one risk-free asset would be willing to pay on average about 15 basis points per year, and for certain stocks as much as 40 basis points, to achieve the level of utility that is obtained by our REGARCH-MIDAS compared to the original REGARCH.

The remainder of the paper is laid out as follows. Section II introduces our extensions to the original REGARCH; the REGARCH-MIDAS and the REGARCH-HAR. Section III outlines the associated estimation procedure. Section IV summarizes our data set, examines the empirical fit and predictive ability of our proposed models, and introduces a procedure for generating multi-period forecasts. Section V concludes. Further empirical results and additional technical details are collected in the Supplementary Appendix available online. Programs for estimating our proposed models are available from the authors upon request.

A. *Related literature*

Our work builds on several strands of the literature on volatility modeling. We now briefly describe the primary strands most closely related to the present paper and provide a few exemplary contributions in each. The outset is the R(E)GARCH framework, but other models have been proposed to utilize information from realized measures. Notable innovations include the GARCH-X model (Engle, 2002), the multiplicative error model (Engle and Gallo, 2006), and the HEAVY model (Shephard and Sheppard, 2010). Within the class of GARCH models without realized measures, several contributions have been made to capture the evident volatility persistence. A few notable references include the Integrated GARCH (Engle and Bollerselv, 1986), the Fractionally Integrated (E)GARCH (Baillie, Bollerslev, and Mikkelsen, 1996; Bollerslev and Mikkelsen, 1996), FIAPARCH (Tse, 1998), regime-switching GARCH (Diebold and Inoue, 2001), HYGARCH (Davidson, 2004), the Spline-GARCH (Engle and Rangel, 2008), and the time-varying component GJR-GARCH (Amado and Teräsvirta, 2013). In the class of R(E)GARCH models, Vander Elst (2015) proposes a fractionally integrated REGARCH, whereas Huang, Liu, and Wang (2016) suggest adding weekly and monthly averages of a realized measure in the GARCH equation of the RGARCH.

Another strand of literature utilizes the idea of decomposing volatility, which originates from Engle and Lee (1999). This has primarily been used to empirically support countercyclicality in stock market volatility (see e.g. Engle et al. (2013) and Dominicy and Vander Elst (2015)). Inspired by the findings of Mikosch and Stărică (2004) which show that long-range dependence and the integrated GARCH effect may be explained by level shifts in the unconditional variance, Amado and Teräsvirta (2013) propose a multiplicative component version of the GJR-GARCH model for capturing volatility persistence. The MIDAS concept originally introduced in a regression framework (Ghysels, Santa-Clara, and Valkanov, 2004, 2005; Ghysels, Sinko, and Valkanov, 2007) has recently been successfully incorporated into the GARCH framework with the GARCH-MIDAS proposal of Engle et al. (2013). Conrad and Kleen (2018) indeed show formally that the autocorrelation function of squared returns is better captured by a multiplicative GARCH specification rather than its nested GARCH(1,1) model, arising from the persistence in the long-term component. Our innovation relative to this literature is to introduce the multiplicative component modeling of conditional variance into the popular R(E)GARCH framework in order to capture the evident volatility persistence.

II. Persistence in a multiplicative Realized EGARCH

Let $\{r_t\}$ denote a time series of returns, $\{x_t\}$ a time series of realized measures,[1] and $\{\mathscr{F}_t\}$ a filtration so that $\{r_t, x_t\}$ is adapted to $\mathscr{F}_t$. We define the conditional mean by $\mu_t = \mathbb{E}[r_t|\mathscr{F}_{t-1}]$ and the conditional variance by $\sigma_t^2 = \mathrm{Var}[r_t|\mathscr{F}_{t-1}]$. Our aim is to allow for more flexible dependence structures in the state-of-the-art specification of conditional variance provided by the REGARCH of Hansen and Huang (2016). To that end, we define

$$r_t = \mu_t + \sigma_t z_t,$$

where $\{z_t\}$ is an i.i.d. innovation process with zero mean and unit variance, and assume that the conditional variance can be multiplicatively decomposed into two

[1]For the remainder of this paper, we assume for clarity of exposition that x_t is one-dimensional, containing a single (potentially robust) realized measure consistently estimating integrated variance such as the realized variance or the realized kernel (Barndorff-Nielsen, Hansen, Lunde, and Shephard, 2008). This assumption is without loss of generality in the sense that additional realized measures (and their associated measurement equations) can be added, such as daily range or the realized quarticity (Bollerslev, Patton, and Quaedvlieg, 2016), rendering x_t a vector.

components as

$$\sigma_t^2 = h_t g_t. \tag{1}$$

We refer to h_t as the short-term component, supposed to capture day-to-day (high-frequency) fluctuations in the conditional variance. On the contrary, g_t is supposed to capture secular (low-frequency) movements in the conditional variance, henceforth referred to as the long-term component or baseline volatility. With the multiplicative decomposition in (1), we extend a daily REGARCH(1,1) to

$$r_t = \mu_t + \sigma_t z_t, \tag{2}$$

$$\log h_t = \beta \log h_{t-1} + \tau(z_{t-1}) + \alpha u_{t-1}, \tag{3}$$

$$\log x_t = \xi + \phi \log \sigma_t^2 + \delta(z_t) + u_t, \tag{4}$$

$$\log g_t = \omega + f(x_{t-2}, x_{t-3}, \ldots; \eta), \tag{5}$$

where $f(\cdot;\eta)$ is a $\mathscr{F}_{t-1}$-measurable function, which can be linear or non-linear. The equations are labeled as the "return equation", the "GARCH equation", the "measurement equation", and the "long-term equation", respectively. For identification purposes, we have omitted an intercept in (3). Note that the specification of the GARCH and measurement equations is a generalization of the logarithmic RGARCH in Hansen et al. (2012).[2] Our model framework applies, as a result, also to the nested RGARCH framework. Hansen and Huang (2016) document, moreover, considerable empirical superiority of REGARCH over RGARCH.

We facilitate level shifts in the baseline volatility via the function $f(\cdot;\eta)$, which takes as input past values of the realized measure. We make the dependence on η explicit in the function $f(\cdot;\eta)$, and prefer that it is low-dimensional. If $f(\cdot;\eta)$ is constant, we obtain the REGARCH as a special case. If $f(\cdot;\eta)$ is time-varying and persistent,

[2] Without multiplicative decomposition, the logarithmic RGARCH takes the form

$$\log h_t = \beta \log h_{t-1} + \alpha x_{t-1},$$
$$\log x_t = \xi + \phi \log h_t + \delta(z_t) + u_t,$$

such that by substitution we obtain

$$\log h_t = \tilde{\omega} + \tilde{\beta} \log h_{t-1} + \alpha \delta(z_{t-1}) + \alpha u_{t-1},$$

where $\tilde{\omega} = \alpha\xi$ and $\tilde{\beta} = \beta + \alpha\phi$. It is clear that the logarithmic RGARCH is nested in the REGARCH if the two leverage functions are proportional, $\tau(z_{t-1}) = \alpha\delta(z_{t-1})$, and that the coefficient on u_{t-1} determines the relative magnitude.

past information may assist in capturing the dependency structure of conditional variance better. We propose in the following sections two ways to parsimoniously formulate $f(\cdot;\eta)$ using non-overlapping weekly and monthly averages of the realized measure to be consistent with the idea of a slow-moving, low-frequency component.[3] We model low-frequency movements in conditional variance using (aggregates of) past information of the realized measure rather than tying it to macroeconomic state variables as in Engle et al. (2013) and Dominicy and Vander Elst (2015). This procedure renders the model in (2)-(5) complete with dynamic specifications of all variables included in the model. Consequently, multi-period forecasting can be conducted on the basis of the jointly estimated empirical dynamics. This contrasts specifications using exogenous information (from e.g. macroeconomic variables) that typically rely on additional assumptions on the dynamics of the exogenous variables (e.g. random walks (Dominicy and Vander Elst, 2015)), outside-generated forecasts (usually from a standard autoregressive specification) of the exogenous variables in the model (Conrad and Loch, 2015) or the assumption that the long-term component is constant for the forecasting horizon (Engle et al., 2013).[4] We do, however, emphasize that our proposed model accommodates well the inclusion of exogenous information such as (possibly low-frequency) macroeconomic variables if deemed appropriate.

Given the high persistence of the conditional variance (documented in the empirical section below), simply including additional lags in the ARMA structure embedded in the original REGARCH is not a viable solution, keeping parameter proliferation in mind (cf. Section IV). Instead, we utilize the multiplicative component structure which is both intuitively appealing and maintain parsimony. Whether high persistence of the conditional variance process arises due to structural breaks, fractional integration or any other source, the long-term component, if modeled accurately, facilitates high persistence in the REGARCH framework. That is, we do not explicitly take a stance on the reason for the presence of high persistence. Our models may be seen as a flexible alternative to a formal long-memory model (see e.g. Bollerslev and Mikkelsen (1996) and Vander Elst (2015)). This is motivated by Mikosch and

[3] The idea is to separate the effects of the realized measure into two, such that the day-to-day effects is (mainly) contained in the short-term component h_t via u_{t-1} and the long-term component captures the information contained in the realized measure further back in time. Excluding information in the realized measure on day $t-1$ from the function $f(\cdot;\eta)$ is consistent with the rolling-window formulations with realized variance in the GARCH-MIDAS framework of Engle et al. (2013).

[4] As suggested by a referee, a thorough comparison of the forecasting performance resulting from dynamically complete and incomplete specifications would be interesting for future research.

Stărică (2004), who show that the high persistence can be explained by level shifts in the unconditional variance (see also Diebold (1986) and Lamoureux and Lastrapes (1990)). On this basis, Amado and Teräsvirta (2013) propose a multiplicative decomposition of the GJR-GARCH model, where the "baseline volatility" changes deterministically according to the passage of time. We may, therefore, capture high persistence via the structure proposed above, when the long-term component in (5) is specified as a slow-moving baseline volatility around which stationary short-term fluctuations occur via the standard GARCH equation. Naturally, this interpretation (and the existence of the unconditional variance) depends on whether $|\beta| < 1$ holds in practice, which may be questionable on the basis of former evidence for the original REGARCH (confirmed in Section IV). However, this integrated GARCH effect is alleviated in our proposed models, where (estimated) β is notably below unity.

The leverage functions, $\tau(\cdot)$ and $\delta(\cdot)$, facilitate modeling of the dependence between return innovations and volatility innovations known to be empirically important (see e.g. Christensen, Nielsen, and Zhu (2010)). If the leverage functions are left out, the residuals $\hat{z}_t$ and $\hat{u}_t$ will be correlated and thereby at odds with the underlying assumptions (Hansen et al., 2012; Hansen and Huang, 2016). We adopt the quadratic form of the leverage functions based on the second-order Hermite polynomial,

$$\tau(z) = \tau_1 z + \tau_2(z^2 - 1),$$
$$\delta(z) = \delta_1 z + \delta_2(z^2 - 1).$$

The leverage functions have a flexible form and imply $\mathbb{E}[\tau(z)] = \mathbb{E}[\delta(z)] = 0$ when $\mathbb{E}[z] = 0$ and $\mathrm{Var}[z] = 1$. Thus, if $|\beta| < 1$, our identification restriction implies that $\mathbb{E}[\log h_t] = 0$ such that $\mathbb{E}\left[\log \sigma_t^2\right] = \mathbb{E}[\log g_t]$.[5] In the (Quasi-)Maximum Likelihood analysis below, we employ a Gaussian specification like Hansen and Huang (2016) with $z_t \sim N(0,1)$ and $u_t \sim N(0, \sigma_u^2)$, and z_t, u_t mutually and serially independent. We check the validity of this approach via a parametric bootstrap in Section III.

The return and GARCH equations are canonical in the GARCH literature. In the return equation, the conditional mean, μ_t, may be modeled in various ways including a GARCH-in-Mean specification or simply as a constant.[6] Following the latter

[5]The GARCH equation implies that $\log h_t = \beta^j \log h_{t-j} + \sum_{i=0}^{j-1} \beta^i \left[\tau(z_{t-1-i}) + \alpha u_{t-1-i}\right]$ such that $\log h_t$ has a stationary representation if $|\beta| < 1$.

[6]The mean is typically modeled as a constant since stock market returns generally are found to be close to serially uncorrelated, see, e.g., Ding, Granger, and Engle (1993). Sometimes the assumption of

approach, we estimate the constant $\mu_t = \mu$. In our multiplicative specification, the GARCH equation drives the dynamics of the high-frequency part of latent volatility. The dynamics are specified as a slightly modified version of the EGARCH model of Nelson (1991) (different leverage function) with the addition of the term αu_{t-1} that relates the latent volatility with the innovation to the realized measure. Hence, α represents how informative the realized measure is about future volatility. The persistence parameter β can be interpreted as the AR-coefficient in an AR(1) model for $\log h_t$ with innovations $\tau(z_{t-1}) + \alpha u_{t-1}$.

The measurement equation is the true innovation in the R(E)GARCH, which makes the model dynamically complete. The equation links the ex-post realized measure with the ex-ante conditional variance. Discrepancies between the two measures are expected, since the conditional variance (and returns) refers to a close-to-close market interval, whereas the realized measure is computed from a shorter, open-to-close market interval. Hence, the realized measure is expected to be smaller than the conditional variance on average. Additionally, the realized measure may be an imperfect measure of volatility. Therefore, the equation includes both a proportional, ξ, and an exponential, ϕ, correction parameter. The innovation term, u_t, can be seen as the true difference between ex-ante and ex-post volatility.

A. *The Realized EGARCH-MIDAS model*

Inspired by the GARCH-MIDAS model of Engle et al. (2013), we consider the following MIDAS specification of the long-term component

$$\log g_t = \omega + \lambda \sum_{k=1}^{K} \Gamma_k\left(\gamma\right) y_{t-1,k}^{(N)}, \tag{6}$$

where $\Gamma_k\left(\gamma\right)$ is a non-negative weighting function parametrized by the vector γ which satisfies the restriction $\sum_{k=1}^{K} \Gamma_k\left(\gamma\right) = 1$, and $y_{t,k}^{(N)} = \frac{1}{N}\sum_{i=1}^{N} \log x_{t-N(k-1)-i}$ is an N-day average of the logarithm of the realized measure. Hence, the value of N determines the frequency of the data feeding into the low-frequency component. We consider in the following $N = 5, 22$, corresponding to weekly and monthly averages.

By estimating γ, for a given weighting function and choice of K, the term $\sum_{k=1}^{K} \Gamma_k\left(\gamma\right) y_{t-1,k}$

zero mean, $\mu = 0$, is imposed for simplicity and may in fact generate better out-of-sample performance (Hansen and Huang, 2016). However, in option-pricing applications a GARCH-in-Mean specification is usually employed, see, e.g., Huang, Wang, and Hansen (2017).

acts as a filter, which extracts the empirically relevant information from past values of the realized measure with assigned importance given by λ. That is, the lag selection process is allowed to be data driven. In practice, we need to choose a value for K and a weighting scheme. Conventional weighting schemes are based on the exponential, exponential Almon lag, or the beta-weight specification. A detailed discussion can be found in Ghysels et al. (2007), who study the choice of weighting function in the context of MIDAS regression models. We employ in the following the two-parameter beta-weight specification defined by

$$\Gamma_k(\gamma_1,\gamma_2) = \frac{(k/K)^{\gamma_1-1}(1-k/K)^{\gamma_2-1}}{\sum_{j=1}^{K}(j/K)^{\gamma_1-1}(1-j/K)^{\gamma_2-1}} \tag{7}$$

due to its flexible form. We restrict $\gamma_2 > 1$, which ensures a monotonically decreasing weighting scheme when $\gamma_1 = 1$. We examine a single-parameter case in which we impose $\gamma_1 = 1$ (see Engle et al. (2013) and Asgharian, Christiansen, and Hou (2016) for a similar restriction) and a case where γ_1 is a free parameter. More rich structures for the weighting scheme can obviously be considered by introducing additional parameters, but we will not explore that route, since one important aim of the MIDAS models is parsimony. As long as the weighting function is reasonably flexible, the choice of lag length of the MIDAS component, K, is of limited importance if chosen reasonably large. The reason is that the estimated γ assigns the relevant weights to each lag simultaneously while estimating the entire model. Should one want to determine an 'optimal' K, we simply suggest to estimate the model for a range of values of K and choose that for which higher values lead to no sizeable gain in the maximized log-likelihood value.

B. The Realized EGARCH-HAR model

Inspired by Corsi (2009), we suggest the following HAR-specification of the long-term component

$$\log g_t = \omega + \gamma_1 \frac{1}{5}\sum_{i=1}^{5}\log x_{t-i-1} + \gamma_2 \frac{1}{22}\sum_{i=i}^{22}\log x_{t-i-1}. \tag{8}$$

The argument for this particular lag structure is motived by the heterogeneous market hypothesis (Müller et al., 1993), which suggests accounting for the heterogeneity in information arrival due to e.g. different trading frequencies of financial market participants. See Corsi (2009) for a more detailed discussion. This particular choice of lag structure including the lagged weekly and monthly average of the logarithm of the realized measure is intuitive and has been empirically successful, but is not data driven as opposed to the MIDAS lag structure. The lag structure can be seen as a special case of the step-function MIDAS specification in Forsberg and Ghysels (2007), which was, indeed, inspired by an unpublished version of Corsi (2009).

III. Estimation

We estimate the models using (Quasi-)Maximum Likelihood (QML) consistent with the procedures in Hansen et al. (2012) and Hansen and Huang (2016). The log-likelihood function can be factorized as

$$\mathscr{L}(r,x;\theta) = \sum_{t=1}^{T} \ell_t(r_t, x_t;\theta) = \sum_{t=1}^{T} [\ell_t(r_t;\theta) + \ell_t(x_t|r_t;\theta)], \tag{9}$$

where $\theta = (\mu, \beta, \tau_1, \tau_2, \alpha, \xi, \phi, \delta_1, \delta_2, \omega, \eta, \sigma_u^2)'$ is the vector of parameters in (2)-(5), and $\ell_t(r_t;\theta)$ is the partial log-likelihood, measuring the goodness of fit of the return distribution. Given the distributional assumptions, $z_t \sim N(0,1)$ and $u_t \sim N(0,\sigma_u^2)$, and z_t, u_t mutually and serially independent, we have

$$\ell_t(r_t;\theta) = -\frac{1}{2}\left[\log 2\pi + \log\sigma_t^2 + z_t^2\right], \tag{10}$$

$$\ell_t(x_t|r_t;\theta) = -\frac{1}{2}\left[\log 2\pi + \log\sigma_u^2 + \frac{u_t^2}{\sigma_u^2}\right], \tag{11}$$

where $z_t = z_t(\theta) = (r_t - \mu)/\sigma_t$. We initialize the conditional variance process to be equal to its unconditional mean, i.e. $\log h_0 = 0$. Alternatively, one can treat $\log h_0$ as an unknown parameter and estimate it as in Hansen and Huang (2016), who show that the initial value is asymptotically negligible. To initialize the long-term component, $\log g_t$, at the beginning of the sample, we simply set past values of $\log x_t$ equal to $\log x_1$ for the length of the backward-looking horizon in the MIDAS-filter. This is done to avoid giving our proposed models an unfair advantage by utilizing more data than the benchmark REGARCH. To avoid inferior local optima in the numerical optimization, we perturb starting values and re-estimate the parameters for each perturbation.

A. Asymptotic properties

We document an MDS property of the score vector in Proposition 1 in the Supplementary Appendix. This is similar to the original REGARCH, leading Hansen and Huang (2016) to conjecture that the limiting distribution of the estimators is normal. We follow the same route and leave the development of the asymptotic theory for estimators of the REGARCH-MIDAS and REGARCH-HAR for future research. Hence, we conjecture that

$$\sqrt{T}(\hat{\theta}-\theta) \xrightarrow{d} N(0, TH^{-1}SH^{-1}), \tag{12}$$

where S is the limit of the outer-product of the scores and H is the negative limit of the Hessian matrix for the log-likelihood function. In practice, we rely on estimates of these two components in the sandwich formula for computing robust standard errors of the coefficients. The estimate $\hat{\theta}$ is obtained via QML.[7] To check the validity of this approach, we employ a parametric bootstrapping technique (Paparoditis and Politis, 2009) in the Supplementary Appendix and find that the in-sample distribution of estimated parameters for both the REGARCH, REGARCH-MIDAS and REGRACH-HAR is generally in agreement with a normal distribution. We also compared the bootstrapped standard errors with the robust QML standard errors computed from the sandwich-formula in (12). The standard errors were also quite similar, which in summary does not contradict the assertion that the QML approach and associated inferences are valid.

IV. Empirical results

In this section, we examine the empirical fit as well as the forecasting performance of the REGARCH-MIDAS and REGARCH-HAR, including an outline of the forecasting procedures involved with the proposed models. We mainly comment on the weekly REGARCH-MIDAS, since the empirical results are qualitatively similar for the monthly version.

[7] Han and Kristensen (2014) and Han (2015) conclude that inference for the QML estimator is quite robust to the level of persistence in covariates included in GARCH-X models, irrespective of them being stationary or not. Moreover, Francq and Thieu (2018) develop asymptotic theory for a wide class of asymmetric GARCH models with exogenous covariates, but not for EGARCH nor log-GARCH specifications.

A. Data

The full sample data set consists of daily close-to-close returns and the daily realized kernels (RK) of the SPY exchange traded fund that tracks the S&P500 Index and 20 individual stocks for the 2002/01-2013/12 period. In the computation of the realized kernel, we use tick-by-tick data, restrict attention to the official trading hours 9:30:00 and 16:00:00 New York time, and employ the Parzen kernel as in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011). See also Barndorff-Nielsen et al. (2008) and Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009) for additional details. For each stock, we remove short trading days where trading occurred in a span of less than 20,000 seconds (compared to typically 23,400 for a full trading day). We also remove data on February 27, 2007, which contains an extreme outlier associated with a computer glitch on the New York Exchange that day. This leaves a sample size for each stock of about 3,000 observations. Figure 1 depicts the evolution of returns, squared returns, realized kernel and the autocorrelation function (ACF) of the logarithm of the realized kernel for SPY.[8]

≪ Insert Figure 1 about here ≫

We estimate the fractional integration parameter d in the logarithm of the realized kernel with the two-step exact local Whittle estimator of Shimotsu (2010). Over the full sample all series satisfy $0.5 < d < 1$, suggesting that volatility is highly persistent.[9] This finding is supported by the slowly decaying ACF of the logarithm of the realized kernel for SPY. Moreover, we reject (Dickey-Fuller) unit root tests across all assets considered using both regular least-squares and instrumented persistence parameters, following the procedures in Hansen and Lunde (2014). See Supplementary Appendix for further details. Collectively, these findings motivate a modeling framework that is capable of capturing a high degree of persistence. Given the requirement that $|\beta| < 1$, this also motivates a framework that pulls β away from unity. This is where the proposed REGARCH-MIDAS and REGARCH-HAR prove useful.

[8] We include in the Supplementary Appendix a table with summary statistics of the daily returns and the logarithm of daily realized kernels for all 20 individual stocks.

[9] We estimated the parameters with $m = \lfloor T^q \rfloor$ for $q = \in \{0.50, 0.55, \dots, 0.80\}$, leading to no alterations of the conclusions obtained for $q = 0.65$. See also Wenger, Leschinski, and Sibbertsen (2017) for a comprehensive empirical study on long memory in volatility and the choice of estimator of d.

B. In-sample results

In this section, we examine the empirical fit of the proposed REGARCH-HAR and REGARCH-MIDAS using the full sample of observations for SPY and the 20 individual stocks. We start out by presenting some appropriate benchmarks.

B.1. Benchmark models

For comparative purposes, we estimate (using QML) two direct antecedents of the REGARCH-MIDAS and REGARCH-HAR proposed in this paper. The first is a REGARCH-Spline (REGARCH-S), with the only difference stemming from the specification of the long-term component. That is, we consider the quadratic spline formulation

$$\log g_t = \omega + c_0 \frac{t}{T} + \sum_{k=1}^{K} c_k \left(\max\left\{ \frac{t}{T} - \frac{t_{k-1}}{T}, 0 \right\} \right)^2,$$

where $\{t_0 = 0, t_1, t_2, \ldots, t_K = T\}$ denotes a partition of the time horizon T in $K+1$ equidistant intervals. Consequently, the smooth fluctuations in the long-term component arises from the (deterministic) passage of time instead of (stochastic) movements in the realized kernel as prescribed by the REGARCH-HAR and REGARCH-MIDAS.[10] The formulation of the long-term component originates from Engle and Rangel (2008) and is also examined in Engle et al. (2013), to which we refer for further details. The number of knots, K, is selected using the BIC information criterion.[11]

The second benchmark is the fractionally integrated REGARCH (FloEGRACH) of Vander Elst (2015), which incorporates fractional integration in the GARCH equation of the REGARCH in a similar vein to the development of the FI(E)GARCH model of Baillie et al. (1996) and Bollerslev and Mikkelsen (1996). The model, thus, explicitly incorporates long-memory via fractionally integrated polynomials in the ARMA structure defined via the parameter d. In contrast to our proposals and the REGARCH-S, the FloEGARCH does not incorporate a multiplicative component

[10] When the long-term component is specified as a deterministic component it follows that $\mathbb{E}[\log \sigma_t^2] = \log g_t$.

[11] In a similar spirit to the choice of K for the REGARCH-MIDAS, we apply the number of knots determined in the estimation on SPY uniformly in all subsequent analyses.

structure. Following Vander Elst (2015), we implement a FloEGARCH(1, d, 1),

$$
\begin{aligned}
r_t &= \mu + \sigma_t z_t, \\
\log \sigma_t^2 &= \omega + (1-\beta)L^{-1}(1-L)^{-d}(\tau(z_{t-1}) + \alpha u_{t-1}), \\
\log x_t &= \xi + \phi \log \sigma_t^2 + \delta(z_t) + u_t,
\end{aligned}
$$

where $(1-L)^d$ is the fractional differencing operator. The infinite polynomial can be written as

$$
(1-\beta)L^{-1}(1-L)^{-d} = \sum_{n=0}^{\infty} \left(\sum_{m=0}^{n} \beta^m \psi_{-d,n-m} \right) L^n,
$$

where $\psi_{-d,k} = \psi_{-d,k-1}\frac{k-1+d}{k}$ and $\psi_{-d,0} = 1$. In the implementation, we truncate the infinite sum at 1,000, similar to Bollerslev and Mikkelsen (1996) and Vander Elst (2015), and initialize the process similarly to Vander Elst (2015). For completeness, we also estimate a multiplicative component version of the EGARCH(1,1) model in Nelson (1991).

B.2. Results for the S&P500 Index

In Table 1, we report estimated parameters, their standard errors, and the associated maximized log-likelihood values for the models under consideration.[12]

≪ Insert Table 1 about here ≫

We derive a number of notable findings. First, the multiplicative component structures lead to substantial increases in the maximized log-likelihood value relative to the original REGARCH. It is worth noting that the null hypothesis of no MIDAS component, $\lambda = 0$ such that $f(\cdot;\eta) = 0$, renders γ_1 and γ_2 unidentified nuisance parameters. Hence, assessing the statistical significance of the differences in maximized log-likelihood values via a standard LR test and a limiting χ^2 distribution is infeasible. Instead, we construct a bootstrapped LR test (BLR), where the null is the REGARCH and the alternative is one of our proposed extensions. Specifically, we simulate 999 series using estimates for the null model (REGARCH) and compute the LR statistic for each simulated series. To avoid problems with local maxima in

[12] Given the importance of choosing the value of K, the maximum lag length of the MIDAS filter, large enough, we let data decide. A detailed analysis in the Supplementary Appendix shows that $K = 12$ and $K = 52$ are suitable choices for the monthly and weekly filters, respectively. We apply these uniformly in all subsequent analyses, including the individual stock results.

the estimation (of especially the alternative model) under the null hypothesis, we consider a wide grid of starting values by perturbation. Given this perturbation, numerical optimization is stable. Despite that the critical values of the bootstrap distributions are noticeably greater than those from a regular chi-square distribution with two-three degrees of freedom, we strongly reject the null of the REGARCH versus all versions of the REGARCH-MIDAS with a p-value of less than 1%. The log-likelihood improvements obtained from the REGARCH-HAR and our benchmark models are statistically significant at a 1% level as well.

We also nuance our interpretation of the log-likelihood gains by information criteria. The substantial increases in log-likelihood value by only a small increase in the number of parameters in the REGARCH-MIDAS and REGARCH-HAR lead to systematic improvements in information criteria. Despite the noticeably greater number of parameters in the REGARCH-S, the increase in the log-likelihood value is only comparable to that of the REGARCH-HAR, leading to a modest improvement in the AIC, only a slight improvement in the BIC, and even a worsening of the HQIC. The FloEGARCH comes closest to the REGARCH-MIDAS specifications, but is still short about seven log-likelihood points. Since it only introduces one additional parameter, the information criteria remain comparable to those of the REGARCH-MIDAS.[13] We have also considered higher-order versions of the original REGARCH(p,q), with $p,q \in \{1,\ldots,5\}$. The best fitting version, the REGARCH(5,5), provides a likelihood gain close to, but still less than the REGARCH-MIDAS models. This gain is, however, obtained with the inclusion of additional eight parameters, causing the information criteria to deteriorate.[14]

Secondly, we find that the single-parameter REGARCH-MIDAS performs similarly to the two-parameter version. Additionally, for the same number of parameters, the single-parameter REGARCH-MIDAS provides a considerable 16-point log-likelihood gain relative to the REGARCH-HAR. This suggests that the HAR formulation is too short-sighted to fully capture the conditional variance dynamics (despite providing a

[13] It is also noteworthy that the FloEGARCH attaches a positive weight to information four years in the past (1,000 daily lags), whereas the REGARCH-MIDAS only carries information from at most the last year. This suggests that the outperformance of the REGARCH-MIDAS relative to the FloEGARCH is somewhat conservative.

[14] It also stands out from Table 1 that the improvements in maximized value from all models under consideration arise from a better modeling of the realized measure and not returns, which comes as no surprise given the motivation behind their development and that the original REGARCH is already a very successful model in fitting returns while lacking adequate modelling of the realized measure, as put forward in Hansen and Huang (2016).

substantial gain relative to the original REGARCH) by using only the most recent month's realized kernels. The differences of the lag functions, as depicted in Figure 2, corroborate this point, by attaching a positive weight on observations further than a month in the past.

≪ Insert Figure 2 about here ≫

The cascade structure as evidenced in Corsi (2009) and Huang et al. (2016) of the HAR formulation is clear from the figure as well, leading to the conclusion that it constitutes a rather successful, yet suboptimal, approximation of the beta-lag function used in the MIDAS formulation.

In Figure 3, we depict the fitted conditional variance along with the long-term components of each multiplicative component model under consideration.

≪ Insert Figure 3 about here ≫

The long-term component of the REGARCH-MIDAS models appear smooth and do, indeed, resemble a time-varying baseline volatility. The long-term component in the REGARCH-HAR is less smooth in contrast to that from the REGARCH-S, which is excessively smooth. To elaborate on the pertinence of the long-term component within each model, we compute the variance ratio given by

$$\mathrm{VR} = \frac{\mathrm{Var}[\log g_t]}{\mathrm{Var}[\log h_t g_t]}, \tag{13}$$

which reveals how much of the variation in a model's fitted conditional variance can be attributed to the long-term component. Note that it is not a goodness-of-fit measure. Rather, it measures how much variation a given model finds it optimal to assign to the long-term component. The last row in Table 1 suggests that the long-term component's contribution is important with more than two-thirds of the variation for the REGARCH-HAR and REGARCH-MIDAS - noticeably larger than that for the REGARCH-S. Moreover, the monthly aggregation scheme for the realized kernel leads to a smoother slow-moving component and, by implication, a smaller VR ratio.

In terms of parameter estimates and associated standard errors, the values are very similar across the various REGARCH extensions for most of the intersection of parameters. The leverage effect appears to be supported in all model formulations, and estimated values of ϕ are less than unity with relatively small standard errors, consistent with the realized measure being computed from open-to-close data and conditional variance referring to the close-to-close period. Moreover, estimated λ is close to 0.9 and precisely estimated, suggesting that past information in the realized kernels are highly informative on conditional variance. The fractional integration parameter, d, is estimated to 0.65 in the FloEGARCH, confirming the high persistence in the conditional variance process also suggested by the summary statistics presented above. Note also that the parameters of the beta-weight function are imprecisely estimated when the restriction $\gamma_1 = 1$ is not imposed. The reason is that two almost identical weight structures may be obtained for two (possibly very) different combinations of γ_1 and γ_2, leaving the pair imprecisely estimated and hints at a potential identification issue.[15] This indeed motivates the restricted, single-parameter version also considered in this paper. Importantly, the estimated values of β are considerably smaller in our proposed models relative to the original REGARCH. A similar, but less pronounced result, is obtained for the REGARCH-S. This reduction in estimated β plays an important role in satisfying the condition that $|\beta| < 1$ and alleviating the integrated GARCH effect. This occurs intuitively since we enable a flexible level of the baseline volatility which the short-term component fluctuates around. Lastly, the measurement equations in the REGARCH-MIDAS and REGARCH-HAR have smaller estimated residual variances, σ_u^2, than the original REGARCH. This may indicate that the new models also provide a better empirical fit of the realized measure via the multiplicative component specifications proposed here.

Those conclusions for the SPY are echoed in our analysis of individual stocks, reported in the Supplementary Appendix. In summary, the REGARCH-MIDAS is the preferred model for all but two stocks when assessed by the likelihood gain. It also stands out that the weekly REGARCH-MIDAS consistently outperforms the REGARCH-HAR. This is generally the case for the monthly REGARCH-MIDAS as well, albeit with a few exceptions. These exceptions may relate to its crude aggregation scheme, which sacrifices too much fit of the autocorrelation structure in the short term for better fit in the long-term compared to the relatively short-sighted formulation in the REGARCH-HAR. Moreover, the estimated β is substantially smaller than unity across all stocks for the REGARCH-MIDAS and REGARCH-HAR models as opposed to the original REGARCH.

[15] For example, the weighting schemes are similar if (γ_1, γ_2) is set to either $(-0.5, 2)$, $(0, 7.5)$, or $(0.5, 12)$.

B.3. Autocorrelation function of conditional variance

In this section, we consider the implications of the REGARCH-HAR and REGARCH-MIDAS on the ACF of the conditional variance and the realized kernel relative to the original formulation in REGARCH. We depict in Figure 4 the simulated and sample ACF of the logarithm of the conditional variance, $\log\sigma_t^2$, for the REGARCH, REGARCH(5,5), REGARCH-HAR, single-parameter and two-parameter REGARCH-MIDAS, and FloEGARCH on SPY. The simulated ACF is obtained using the estimated parameters in Table 1 with a sample size of 3,750 (approximately 15 years), 10,000 Monte Carlo replications and the Gaussian specification of the error terms as in Hansen and Huang (2016). The sample ACF is based on the fitted conditional variance.

≪ Insert Figure 4 about here ≫

In general and for a given model, the closer the simulated and sample ACF are to each other, the larger is the degree of internal consistency between theoretical and actual model predictions of the dependency structure in conditional variance. This is especially relevant for economic applications of the models. For instance, in risk-management (e.g. Value-at-Risk) or option valuation, the researcher may rely on simulations from a given model to produce trajectories of conditional variances, see, e.g., Engle (2004). If the internal consistency is low, the autocorrelation property of the simulated variance processes are far from what the model tries to capture in data. We note that the original REGARCH is only able to capture the autocorrelation structure over the very short term. Moreover, the REGARCH(5,5) does not substantially improve upon the REGARCH. The simulated ACF of the REGARCH-HAR is closer to the sample ACF, but starts diverging at about lag 30. Only the REGARCH-MIDAS models and the FloEGARCH are capable of capturing the pattern of the autocorrelation structure over a long horizon. The monthly REGARCH-MIDAS, however, trades off some fit in the short term for improved accuracy in the long term by using a cruder aggregation scheme of the realized measure. Overall, this suggests that the multiplicative component structure used in the REGARCH-HAR and REGARCH-MIDAS constitutes a very appealing and parsimonious way of capturing high persistence in the REGARCH framework.[16]

[16]We also compared the simulated and sample ACFs of the logarithm of the realized kernel for each model to provide an insight into whether the models are able to capture the autocorrelation structure of the market realized variance. The conclusions are, expectedly, similar to the one in Figure 4, and are provided in the Supplementary Appendix.

C. Forecasting with the REGARCH-MIDAS and REGARCH-HAR

Denote by k, $k \geq 1$, the forecast horizon measured in days. Our aim is to forecast the conditional variance k days into the future. To that end, we note that for $k = 1$ one-step ahead forecasting can be easily achieved directly via the GARCH equation in (3). For multi-period forecasting ($k > 1$), we note that recursive substitution of the GARCH equation implies

$$\log h_{t+k} = \beta^k \log h_t + \sum_{j=1}^{k} \beta^{j-1} \left(\tau(z_{t+k-j}) + \alpha u_{t+k-j}\right),$$

such that

$$\log \sigma^2_{t+k} = \log h_{t+k} g_{t+k} = \beta^k \log h_t + \sum_{j=1}^{k} \beta^{j-1} \left(\tau(z_{t+k-j}) + \alpha u_{t+k-j}\right) + \log g_{t+k}.$$

Multi-period forecasts of $\log \sigma^2_{t+k}$ may then be obtained via

$$\log \sigma^2_{t+k|t} \equiv \mathbb{E}[\log \sigma^2_{t+k} | \mathscr{F}_t] = \beta^k \log h_t + \beta^{k-1} (\tau(z_t) + \alpha u_t) + \log g_{t+k|t}.$$

Consequently, the contribution of the short-term component to the forecast is easily computed with known quantities at time t, namely h_t, u_t, z_t. To obtain $g_{t+k|t}$, we generate recursively, using estimated parameters, the future path of the realized measure using the measurement equation in (4) and 10,000 simulations that re-samples from the empirical distributions of $\hat{z}_t$ and $\hat{u}_t$. It is worth noting that for multi-step forecast horizons a lower magnitude of β causes the forecast to converge more rapidly towards the baseline volatility, determined by (the forecast of) the long-term component. Because this baseline volatility is allowed to be time-varying, a lower magnitude of β is preferable since it generates more flexibility and reduces the long-lasting impact on the forecast from the most recent h_t and its innovation. By implication, the ability to generate reasonable forecasts of the long-term component is valuable, which strongly motivates the dynamic completeness of the models.[17]

Jensen's inequality stipulates that $\exp\{\mathbb{E}[\log \sigma^2_{t+k} | \mathscr{F}_t]\} \neq \mathbb{E}[\exp\{\log \sigma^2_{t+k}\} | \mathscr{F}_t]$ such that we need to consider the distributional aspects of $\log \sigma^2_{t+k|t}$ to obtain an unbiased forecast of $\sigma^2_{t+k|t}$. As a solution, we utilize a simulation procedure with empirical

[17] We found, indeed, that setting $g_{t+k|t} = g_t$ leads to notably inferior forecasting performance relative to the case that exploits the estimated dynamics of the realized kernel.

distributions of z_t and u_t. Using M simulations and re-sampling the estimated residuals, the resulting forecast of the conditional variance given by

$$\sigma^2_{t+k|t} = \frac{1}{M}\sum_{m=1}^{M} \exp\{\log \sigma^2_{t+k|t,m}\}$$

is unbiased. In the implementation, we estimate model parameters on a rolling basis with 10 years of data (2,500 observations) and leave the remaining (about 500) observations for (pseudo) out-of-sample evaluation. The empirical distribution of $\hat{z}_t$ and $\hat{u}_t$ is similarly obtained using the same historical window of observations. Forecasting with the REGARCH follows directly from the above with $\log g_{t+k|t} = \omega$.

C.1. Forecast evaluation

Given the latent nature of the conditional variance, we require a proxy, $\hat{\sigma}^2_t$, of σ^2_t for forecast evaluation. To that end, we employ the adjusted realized kernel similarly to, e.g., Fleming, Kirkby, and Ostdiek (2001, 2003), Martens (2002), Koopman, Jungbacker, and Hol (2005), Bandi and Russell (2006) and Huang et al. (2016) given by $\hat{\sigma}^2_t = \kappa RK_t$, where $\kappa = \sum_{t=1}^{T} r^2_t / \sum_{t=1}^{T} RK_t$. The adjustment is needed since the realized measure is a measure of open-to-close variance, whereas the forecast generated by the REGARCH framework measures close-to-close variance. We compute κ on the basis of the out-of-sample period.[18] A second implication of using the realized kernel as proxy is that we implicitly restrict ourselves to the choice of robust loss functions (Hansen and Lunde, 2006; Patton, 2011) when quantifying the forecast precisions in order to obtain consistent ranking of forecasts. Let $L_{i,t+k}(\hat{\sigma}^2_{t+k}, \sigma^2_{t+k|t})$ denote the loss function for the i'th k-step ahead forecast. Two such robust functions are the Squared Prediction Error (SPE) and Quasi-Likelihood (QLIKE) loss function given as

$$L^{(\mathrm{SPE})}_{i,t+k}(\hat{\sigma}^2_{t+k}, \sigma^2_{t+k|t}) = (\hat{\sigma}^2_{t+k} - \sigma^2_{t+k|t})^2, \tag{14}$$

$$L^{(\mathrm{QLIKE})}_{i,t+k}(\hat{\sigma}^2_{t+k}, \sigma^2_{t+k|t}) = \frac{\hat{\sigma}^2_{t+k}}{\sigma^2_{t+k|t}} - \log\left(\frac{\hat{\sigma}^2_{t+k}}{\sigma^2_{t+k|t}}\right) - 1. \tag{15}$$

In both cases, a value of zero is obtained for a perfect forecast. The SPE (QLIKE) loss function penalizes forecast error symmetrically (asymmetrically), and the QLIKE

[18] As pointed out by a referee, an alternative approach would be to add squared overnight returns to the realized kernel. We find that the conclusions are unaltered.

often gives rise to more power in statistical forecast evaluation procedures, especially when comparing losses across different regimes (see e.g. Borup and Thyrsgaard (2017)). Given the objective of evaluating whether the REGARCH-MIDAS and REGARCH-HAR provide an improvement in forecasts relative to the REGARCH, we implement a Diebold-Mariano test (Diebold and Mariano, 1995) computed using Newey and West (1987) standard errors with a Bartlett kernel and data-dependent bandwidth based on an AR(1) approximaton Andrews (1991).[19] We perform the test against the alternative that the i'th forecast losses are smaller than the ones arising from the original REGARCH and evaluate the test statistic in the standard normal distribution. Moreover, we conduct a Model Confidence Set (MCS) procedure to identify the best performing set of models. See the explanation associated with Table 2 and Hansen, Lunde, and Nason (2011) for additional details.

C.2. Forecasting results

Figure 5 depicts Theil's U statistic in terms of the ratio of forecast losses on the SPY arising from forecasts generated by the original REGARCH to those from the REGARCH-HAR and the weekly REGARCH-MIDAS (single-parameter) on horizons $k = 1, \ldots, 22$. It depicts their associated statistical significance, too. Quantitatively and qualitatively similar results for the remaining MIDAS specifications are left out, but are available upon request.

$\ll$ Insert Figure 5 about here $\gg$

The figure convincingly concludes that both the REGARCH-HAR and REGARCH-MIDAS improve upon the forecasting performance of the original REGARCH for all forecast horizons. These improvements tend to grow as the forecast horizon increases from a few percentages to roughly 30-40% depending on the loss function. This indicates the usefulness of modeling a slow-moving component, particularly for forecasting beyond short horizons. In general, the improvements are statistically significant for all horizons, except for the shorter horizons in the REGARCH-MIDAS

[19] We acknowledge that the Diebold-Mariano test is technically not appropriate for comparing forecasts of nested models since the limiting distribution is non-standard under the null hypothesis (see e.g. Clark and McCracken (2001) and Clark and West (2007)). The adjusted mean squared errors of Clark and West (2007) or the bootstrapping procedure of Clark and McCracken (2015) are appropriate alterations to standard inferences. However, since we estimate our models on a rolling basis with a finite, fixed window size, the asymptotic framework of Giacomini and White (2006) provides a rigorous justification for proceeding with the Diebold-Mariano test statistic evaluated in a standard normal distribution. See also Diebold (2015) for a discussion.

case.[20] Identical conclusions hold for our 20 individual stocks, which are reported in the Supplementary Appendix. Having established the improvement upon the original REGARCH, we turn to a complete comparison of all our proposed models, the REGARCH-S and the FloEGARCH. Table 2 reports the percentage of stocks (including SPY) for which a given model is included in the MCS at an $\alpha = 10\%$ significance level.

≪ Insert Table 2 about here ≫

The inclusion frequency of our proposed REGARCH-MIDAS models are high and indicate superiority over all competing models in both the short-term and beyond. Interestingly, the cruder, monthly aggregation scheme dominates for longer horizons, whereas the finer, weekly scheme is preferred for short horizons. The REGARCH-S shows moderate improvement over the original REGARCH, but is less frequently included in the MCS compared to our proposed REGARCH-MIDAS and REGARCH-HAR. The FloEGARCH performs relatively bad for horizons 2,3,4 and 5, but is increasingly included in the MCS as the forecast horizon increases, reaching similar performance as the REGARCH-MIDAS models at monthly predictions. These findings indicate the usefulness of the flexibility obtained via the multiplicative component structure as opposed to, e.g., incorporating fractional integration as in the FloEGARCH.

C.3. Economic value of volatility timing

To assess the economic value of the improvements in predictive accuracy, we examine a volatility-timing strategy that uses each model's forecast as input to construct optimal portfolio weights. We consider a risk-averse investor with mean-variance preferences who allocates funds into one risky asset and one risk-free asset. The investor uses conditional mean-variance analysis to make allocation decisions each day with an investment horizon of k days into the future. For risk-aversion parameter A, this determines the weights on the risky asset between at time t by

$$w_{i,t+k|t} = \frac{\tilde{\mu}_{i,t+k|t} - \tilde{r}^f_{t+k}}{A\tilde{\sigma}^2_{i,t+k|t}},$$

[20]We have also examined the models' predictive ability of cumulative forecasts for a 5, 10, and 22 horizon. Consistent with the findings for the point forecasts, both the REGARCH-HAR and REGARCH-MIDAS provide substantial and statistically significant improvements relative to the original REGARCH.

where $\tilde{\mu}_{i,t+k|t}$ and $\tilde{\sigma}^2_{i,t+k|t}$ denote the cumulative forecasts of the stock return and variance, respectively, obtained from the i'th model, and $\tilde{r}^f_{t+k}$ the risk-free rate. The realized portfolio return is then

$$\tilde{r}^p_{i,t+k} = \tilde{r}^f_{t+k} + w_{i,t+k|t}(\tilde{r}_{t+k} - \tilde{r}^f_{t+k}).$$

To evaluate the incremental impact of the variance forecast, we assume the investor treats future returns as unpredictable and use a rolling mean over the same in-sample window used for generating the variance forecasts. This is consistent with approaches in, e.g., Fleming et al. (2001, 2003) and Bandi and Russell (2006). We proxy the daily risk-free rate by the T-Bill return which, over the number of trading days within the month, compounds to the one month T-Bill rate obtained from Kenneth French's data library. For $k = \{1, 5, 10, 22\}$ investment horizons and $P = 521 - k$ out-of-sample forecasts, the investor's realized average utility is

$$\bar{U}\left(\tilde{r}^p_{i,t+k}\right) = \frac{1}{P}\sum_{t=1}^{P}\left(\tilde{r}^p_{i,t+k} - \frac{A}{2}\left(\tilde{r}^p_{i,t+k} - \frac{1}{P}\sum_{t=1}^{P}\tilde{r}^p_{i,t+k}\right)^2\right).$$

To quantify the economic benefit relative to a benchmark model indexed by "0", here the original REGARCH, we use the performance fee metric put forward by Fleming et al. (2001, 2003). Denoted by Φ, the performance fee measures the amount (in basis points) the investor would be willing to pay to switch from the benchmark strategy to the competing strategy. In our context it, thus, captures the value of the forecasts from each of our considered models relative to those from the REGARCH seen from the perspective of an investor. The performance fee is the solution to $\bar{U}\left(\tilde{r}^p_{i,t+k} - \Phi\right) = \bar{U}\left(\tilde{r}^p_{0,t+k}\right)$ such that

$$\Phi = \bar{U}\left(\tilde{r}^p_{i,t+k}\right) - \bar{U}\left(\tilde{r}^p_{0,t+k}\right) \tag{16}$$

in the case of mean-variance preferences.[21] Figure 6 depicts the performance fee on SPY for two conventional levels of the risk-aversion parameter, $A = 2, 10$, and our set of models under consideration.

≪ Insert Figure 6 about here ≫

At short investment horizons, there is an economically small, yet positive gain from applying our models. The gain is monotonically increasing in the investment horizon, and at monthly horizons the gain is substantial for both risk-aversion levels. This is consistent with the intuitive finding above that capturing volatility persistence leads to substantial improvements in predictive accuracy at especially longer horizons. The REGARCH-HAR is the worst performing model, whereas the REGARCH-MIDAS fares comparably to the FloEGARCH. Interestingly, even though the REAGRCH-S was strongly outperformed in the statistical forecast evaluation, it is the best performing model in terms of economic value. We evaluate the statistical significance of the performance fees via a one-sided Diebold-Mariano type statistic and Newey and West (1987) standard errors as in C.1, see Engle and Colacito (2006) and Bandi, Russell, and Zhu (2008) for a similar approach. The p-values associated with the null hypothesis $\Phi = 0$ against the alternative $\Phi > 0$ indicate rejections at $k = 22$ for all models and both risk-aversion values, applying conventional significance levels of the test. The economic gain is statistically insignificant at shorter investment horizons, except for the REGARCH-S case, which remains significant at $k = 10$ at both risk-aversion values. These findings confirm the link between statistical and economic evaluation in that improved predicative accuracy leads to improved portfolio performance. This link is, however, not monotonic. In the Supplementary Appendix, we report the distribution of performance fees across all our individual stocks. Conclusions are identical to those from SPY, except that the REGARCH-S is no longer uniformly the best performing model on average. Moreover, the cross-sectional variation of performance fees is notably smaller for the REGARCH-MIDAS compared to both REGARCH-S and FloEGARCH with a similar average, showing

[21] Another frequently used measure to assess portfolio performance is the Sharpe ratio. However, this measure can easily be manipulated (Goetzmann, Ingersoll, Spiegel, and Welch, 2007). As such we also cmpute the manipulation-proof performance measure of Goetzmann et al. (2007) relative to the original REGARCH defined as

$$\Theta = \frac{1}{1-A}\left[\log\left(\frac{1}{P}\sum_{t=1}^{P}\left[\frac{\tilde{r}^{p}_{i,t+k}}{1+\tilde{r}^{f}_{t+k}}\right]^{1-A}\right) - \log\left(\frac{1}{P}\sum_{t=1}^{P}\left[\frac{\tilde{r}^{p}_{0,t+k}}{1+\tilde{r}^{f}_{t+k}}\right]^{1-A}\right)\right].$$

We find that the conclusions are identical.

indications of an appealing overall economic gain across all stocks. For certain stocks and high risk-aversion value, the performance fee can be as high as 70 basis points for REGARCH-S, 60 for FloEGARCH and 40 for REGARCH-MIDAS.

V. Concluding remarks

We introduce two extensions of the otherwise successful REGARCH model to capture the evident high persistence observed in stock return volatility series. Both extensions exploit a multiplicative decomposition of the conditional variance process into a short-term and a long-term component. The latter is modeled either using mixed-data sampling or a heterogeneous autoregressive structure, giving rise to the REGARCH-MIDAS and REGARCH-HAR, respectively. Both models lead to substantial in-sample improvements of the REGARCH with the REGARCH-MIDAS dominating the REGARCH-HAR. Evidently, the backward-looking horizon of the HAR specification is too short to adequately capture the autocorrelation structure of volatility for horizons longer than a month.

Our suggested models are dynamically complete, facilitating multi-period forecasting. Coupled with a lower estimated β and time-varying baseline volatility, we show in a forecasting exercise that the REGARCH-MIDAS and REGARCH-HAR lead to significant improvements in predictive ability of the REGARCH at both short and long horizons. A volatility-timing trading strategy shows that capturing volatility persistence yields substantial utility gains for a mean-variance investor at longer investment horizons.

Our proposed models enable an easy inclusion of additional realized measures, macroeconomic variables or event-related dummies (e.g. from policy announcements). Some additional questions remain for future research. On the empirical side, applications to other asset classes exhibiting high persistence such as commodities,[22] bonds or exchange rates, or the use of our proposed models in estimating the (term structure of) variance risk premia, or investigating the risk-return relationship via the return equation (see e.g. Christensen et al. (2010)) are of potential interest. On the theoretical side, development of a misspecification test for comparison of our models with the nested REGARCH and asymptotic properties of the QML estimator would prove very useful.

[22] See e.g. Lunde and Olesen (2013) for an application of the REGARCH to commodities.

References

AMADO, C. AND T. TERÄSVIRTA (2013): "Modelling volatility by variance decomposition," *Journal of Econometrics*, 175, 142–153.

ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2001): "The distribution of realized exchange rate volatility," *Journal of the American Statistical Association*, 96, 42–55.

——— (2003): "Modeling and forecasting realized volatility," *Econometrica*, 71, 579–625.

ANDREWS, D. W. K. (1991): "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica*, 59, 817–858.

ASGHARIAN, H., C. CHRISTIANSEN, AND A. J. HOU (2016): "Macro-finance determinants of the long-run stock-bond correlation: The DCC-MIDAS specification," *Journal of Financial Econometrics*, 14, 617.

BAILLIE, R. T., T. BOLLERSLEV, AND H. O. MIKKELSEN (1996): "Fractionally integrated generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 74, 3–30.

BANDI, F. M. AND J. F. RUSSELL (2006): "Separating microstructure noise from volatility," *Journal of Financial Economics*, 79, 655–692.

BANDI, F. M., J. F. RUSSELL, AND Y. ZHU (2008): "Using high-frequency data in dynamic portfolio choice," *Econometric Reviews*, 27, 163–198.

BARNDORFF-NIELSEN, O. E., P. R. HANSEN, A. LUNDE, AND N. SHEPHARD (2008): "Designing realized kernels to measure the ex post variation of equity prices in the presence of noise," *Econometrica*, 76, 1481–1536.

——— (2009): "Realised kernels in practice: trades and quotes," *Econometrics Journal*, 12, 1–32.

——— (2011): "Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading," *Journal of Econometrics*, 162, 149–169.

BOLLERSLEV, T. AND H. O. MIKKELSEN (1996): "Modeling and pricing long memory in stock market volatility," *Journal of Econometrics*, 73, 542–547.

BOLLERSLEV, T., A. J. PATTON, AND R. QUAEDVLIEG (2016): "Exploiting the errors: A simple approach for improved volatility forecasting," *Journal of Econometrics*, 192, 1–18.

BORUP, D. AND M. THYRSGAARD (2017): "Statistical tests for equal predictive ability across multiple forecasting methods," CREATES Research Paper.

CHRISTENSEN, B. J., M. O. NIELSEN, AND J. ZHU (2010): "Long memory in stock market volatility and the volatility-in-mean effect: The FIEGARCH-M model," *Journal of Empirical Finance*, 17, 460–470.

CLARK, T. E. AND M. MCCRACKEN (2001): "Tests of equal forecast accuracy and encompassing for nested models," *Journal of Econometrics*, 105, 85–110.

——— (2015): "Nested forecast model comparisons: a new approach to testing equal accuracy," *Journal of Econometrics*, 186, 160–177.

CLARK, T. E. AND K. D. WEST (2007): "Approximately normal tests for equal predictive accuracy in nested models," *Journal of Econometrics*, 138, 291–311.

CONRAD, C. AND O. KLEEN (2018): "Two are better than one: volatility forecasting using multiplicative component GARCH models," Working paper.

CONRAD, C. AND K. LOCH (2015): "Anticipating long-term stock market volatility," *Journal of Applied Econometrics*, 30, 1090–1114.

CORSI, F. (2009): "A simple approximate long-memory model of realised volatility," *Journal of Financial Econometrics*, 7, 174–196.

DAVIDSON, J. (2004): "Moment and memory properties of linear conditional heteroscedasticity models, and a new model," *Journal of Business and Economic Statistics*, 22, 16–29.

DIEBOLD, F. X. (1986): "Modeling the persistence of conditional variance: A comment," *Econometric Reviews*, 5, 51–56.

——— (2015): "Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of the Diebold-Mariano tests," *Journal of Business and Economic Statistics*, 33, 1–9.

DIEBOLD, F. X. AND A. INOUE (2001): "Long memory and regime switching," *Journal of Econometrics*, 105, 131–159.

DIEBOLD, F. X. AND R. S. MARIANO (1995): "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

DING, Z., C. W. GRANGER, AND R. F. ENGLE (1993): "A long memory property of stock market returns and a new model," *Journal of Empirical Finance*, 1, 83 – 106.

DOMINICY, Y. AND H. VANDER ELST (2015): "Macro-driven VaR forecasts: From very high to very low-frequency data," Working paper.

ENGLE, R. (2002): "New frontiers for ARCH models," *Journal of Applied Econometrics*, 17, 425–446.

——— (2004): "Risk and volatility: Econometric models and financial practice," *The American Economic Review*, 94, 405–420.

ENGLE, R. AND T. BOLLERSELV (1986): "Modeling the persistence in conditional variances," *Econometric Reviews*, 5, 1–50.

ENGLE, R. AND R. COLACITO (2006): "Testing and valuing dynamic correlations for asset allocation," *Journal of Business Economics and Statistics*, 24, 238–253.

ENGLE, R. F. AND G. M. GALLO (2006): "A multiple indicators model for volatility using intra-daily data," *Journal of Econometrics*, 131, 3–27.

ENGLE, R. F., E. GHYSELS, AND B. SOHN (2013): "Stock market volatility and macroeconomic fundamentals," *The Review of Economics and Statistics*, 95, 776–797.

ENGLE, R. F. AND G. LEE (1999): "A long-run and short-run component model of stock return volatility," *In R. F. Engle and H. White (eds.), Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger*, 475–497.

ENGLE, R. F. AND J. G. RANGEL (2008): "The Spline-GARCH model for low-frequency volatility and its global macroeconomic causes," *The Review of Financial Studies*, 21, 1187–1222.

FENG, Y. (2004): "Simultaneously modeling conditional heteroskedasticity and scale change," *Econometric Theory*, 20, 563–596.

FLEMING, J., C. KIRKBY, AND B. OSTDIEK (2001): "The economic value of volatility timing," *The Journal of Finance*, 56, 329–352.

——— (2003): "The economic value of volatility timing using "realized" volatility," *Journal of Financial Economics*, 67, 473–509.

FORSBERG, L. AND E. GHYSELS (2007): "Why do absolute returns predict volatility so well?" *Journal of Financial Econometrics*, 5, 31.

FRANCQ, C. AND L. Q. THIEU (2018): "QML inference for volatility models with covariates," *Econometric Theory*, Forthcoming, 1–36.

GHYSELS, E., P. SANTA-CLARA, AND R. VALKANOV (2004): "The MIDAS touch: mixed data sampling regression models," Working paper.

——— (2005): "There is a risk-return trade-off after all," *Journal of Financial Economics*, 76, 509 – 548.

GHYSELS, E., A. SINKO, AND R. VALKANOV (2007): "MIDAS regressions: further results and new directions," *Econometric Reviews*, 26, 53–90.

GIACOMINI, R. AND H. WHITE (2006): "Tests of conditional predictive ability," *Econometrica*, 74, 1545–1578.

GOETZMANN, W., J. INGERSOLL, M. SPIEGEL, AND I. WELCH (2007): "Portfolio performance manipulation and manipulation-proof performance measures," *The Review of Financial Studies*, 20, 1503–1546.

HAN, H. (2015): "Asymptotic properties of GARCH-X processes," *Journal of Financial Econometrics*, 13, 188–221.

HAN, H. AND D. KRISTENSEN (2014): "Asymptotic theory for the QMLE in GARCH-X Models with stationary and nonstationary covariates," *Journal of Business and Economic Statistics*, 32, 416–429.

HANSEN, P. R. AND Z. HUANG (2016): "Exponential GARCH modeling with realized measures of volatility," *Journal of Business and Economic Statistics*, 34, 269–287.

HANSEN, P. R., Z. HUANG, AND H. H. SHEK (2012): "Realized GARCH: a joint model for return and realized mesures of volatility," *Journal of Applied Econometrics*, 27, 877–906.

HANSEN, P. R. AND A. LUNDE (2006): "Consistent ranking of volatility models," *Journal of Econometrics*, 131, 97–121.

——— (2014): "Estimating the persistence and the autocorrelation function of a time series that is measured with error," *Econometric Theory*, 30, 60–93.

HANSEN, P. R., A. LUNDE, AND J. M. NASON (2011): "The model confidence set," *Econometrica*, 79, 453–497.

HILLEBRAND, E. (2005): "Neglecting parameter changes in GARCH models," *Journal of Econometrics*, 129, 121 – 138.

HUANG, Z., H. LIU, AND T. WANG (2016): "Modeling long memory volatility using realized measures of volatility: a realized HAR GARCH model," *Economic Modelling*, 52, 812–821.

HUANG, Z., T. WANG, AND P. R. HANSEN (2017): "Option pricing with the Realized GARCH model: an analytical approximation approach," *Journal of Futures Markets*, 37, 328–358.

KOOPMAN, S. J., B. JUNGBACKER, AND E. HOL (2005): "Forecasting daily variability of the S&P100 stock index using historical, realised and implied volatility measurements," *Journal of Empirical Finance*, 12, 445–475.

LAMOUREUX, C. G. AND W. D. LASTRAPES (1990): "Persistence in variance, structural change, and the GARCH model," *Journal of Business and Economic Statistics*, 8, 225–234.

LUNDE, A. AND K. V. OLESEN (2013): "Modeling and forecasting the distribution of energy forward returns," CREATES Research Paper, 2013-19.

MARTENS, M. (2002): "Measuring and forecasting S&P500 index-futures volatility using high-frequency data," *Jornal of Futures Markets*, 22, 497–518.

MCCLOSKEY, A. AND P. PERRON (2013): "Memory parameter estimation in the presence of level shifts and deterministic trends," *Econometric Theory*, 29, 1196–1237.

MIKOSCH, T. AND C. STĂRICĂ (2004): "Nonstationarities in financial time series, the long-range dependence, and IGARCH effects," *The Review of Economics and Statistics*, 86, 378–390.

MÜLLER, U. A. ET AL. (1993): "Fractals and intrinsic time - A challenge to econometricians," in *39th International AEA Conference on Real Time Econometrics, 14-15 October 1993, Luxembourg.*

NELSON, D. B. (1991): "Conditional heteroskedasticity in asset returns: A new approach," *Econometrica*, 59, 347–370.

NEWEY, W. K. AND K. D. WEST (1987): "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance marix," *Econometrica*, 55, 703–708.

PAPARODITIS, E. AND D. N. POLITIS (2009): "Resampling and subsampling for financial time series," *In "Handbook of financial time series", Springer*, 983–999.

PATTON, A. J. (2011): "Volatility forecast comparison using imperfect volatility proxies," *Journal of Econometrics*, 160, 246–256.

SHEPHARD, N. AND K. SHEPPARD (2010): "Realising the future: forecasting with high-frequency-based volatility (HEAVY) models," *Journal of Applied Econometrics*, 25, 197–231.

SHIMOTSU, K. (2010): "Exact local whittle estimation of fractional integration with unknown mean and time trend," *Econometric Theory*, 26, 501–540.

TSE, Y. (1998): "The conditional heteroskedasticity of the Yen-Dollar exchange rate," *Journal of Applied Econometrics*, 13, 49–55.

VANDER ELST, H. (2015): "FloGARCH: Realizing long memory and asymmetries in returns volatility," Working paper.

VARNESKOV, R. AND P. PERRON (2017): "Combining long memory and level shifts in modeling and forecasting the volatility of asset returns," *Quantitative Finance*, 1–23.

WENGER, K., C. LESCHINSKI, AND P. SIBBERTSEN (2017): "Long memory of volatility," Working paper.

A. Figures

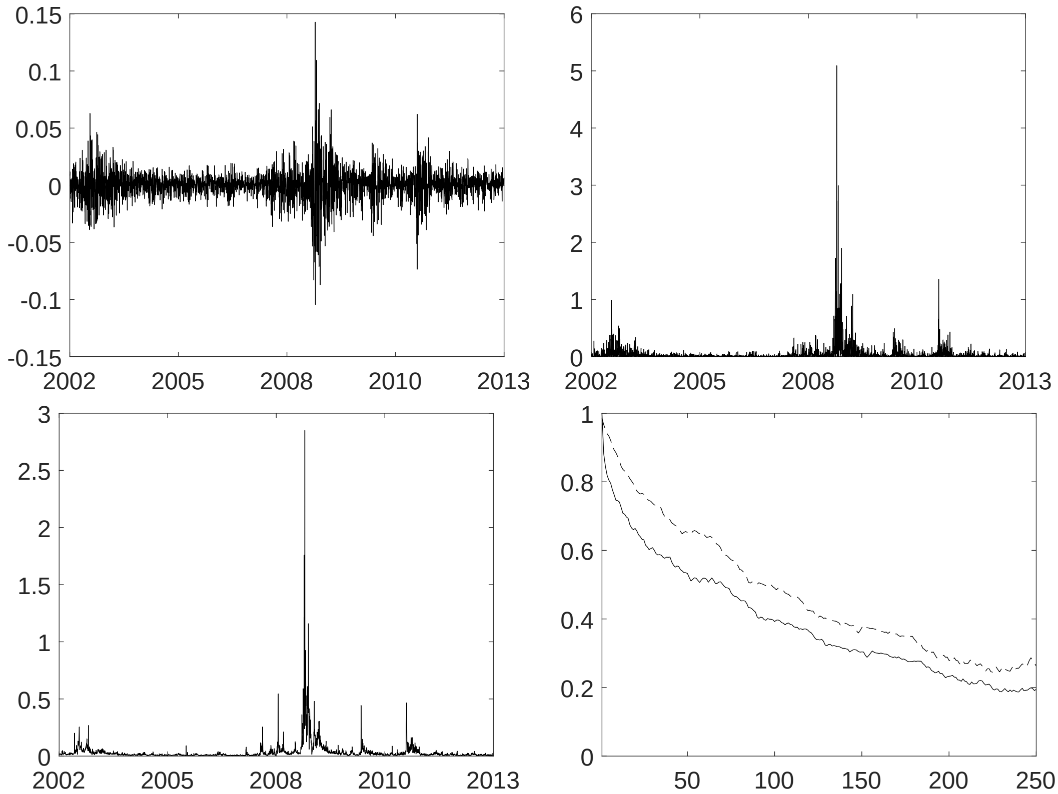

Figure 1: Summary statistics for SPY daily returns and realized kernel

This figure depicts the evolution of SPY daily returns (upper-left panel), annualized squared returns (upper-right panel), annualized realized kernel (lower-left panel), and autocorrelation function of the logarithm of the realized kernel (lower-right panel). The solid line indicates the conventional autocorrelation function, whereas the dashed line indicates the instrumented variable autocorrelation function of Hansen and Lunde (2014) using their preferred instruments (four through ten) and optimal combination.

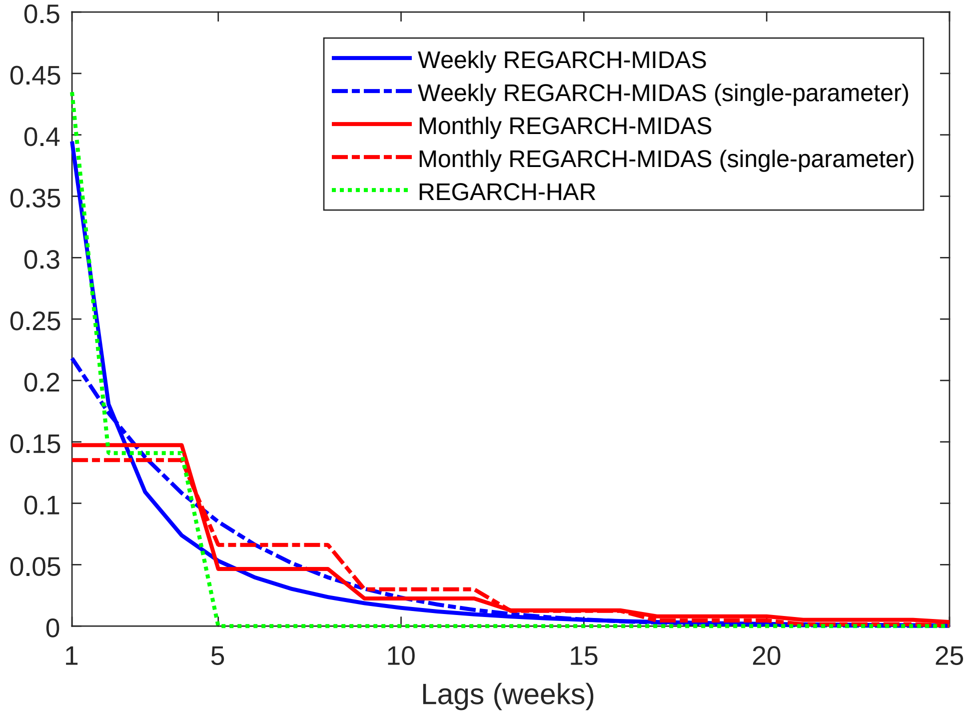

Figure 2: Estimated SPY weighting functions

This figure depicts the estimated weighting functions in our proposed models for SPY with $K = 52$ and $K = 12$ in the weekly and monthly REGARCH-MIDAS, respectively. Blue lines relate to the weekly REGARCH-MIDAS, red lines relate to the monthly REGARCH-MIDAS, and the green line to the REGARCH-HAR. Solid lines refer to the two-parameter weighting function, whereas dashed lines refer to the restricted, single-parameter weighting function.

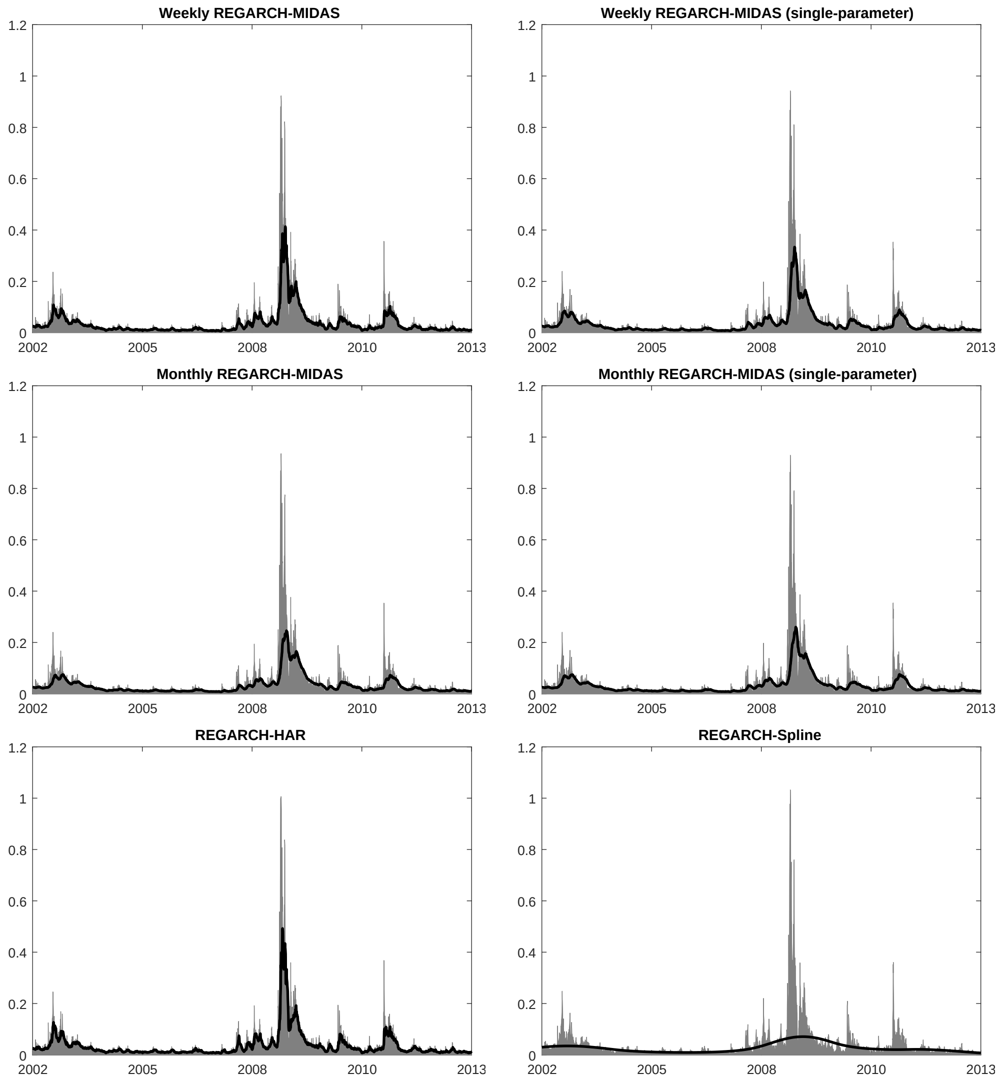

Figure 3: Fitted conditional variance and the long-term component

This figure depicts the evolution of the fitted annualized conditional variance together with its long-term component from the multiplicative REGARCH modifications in Table 1.

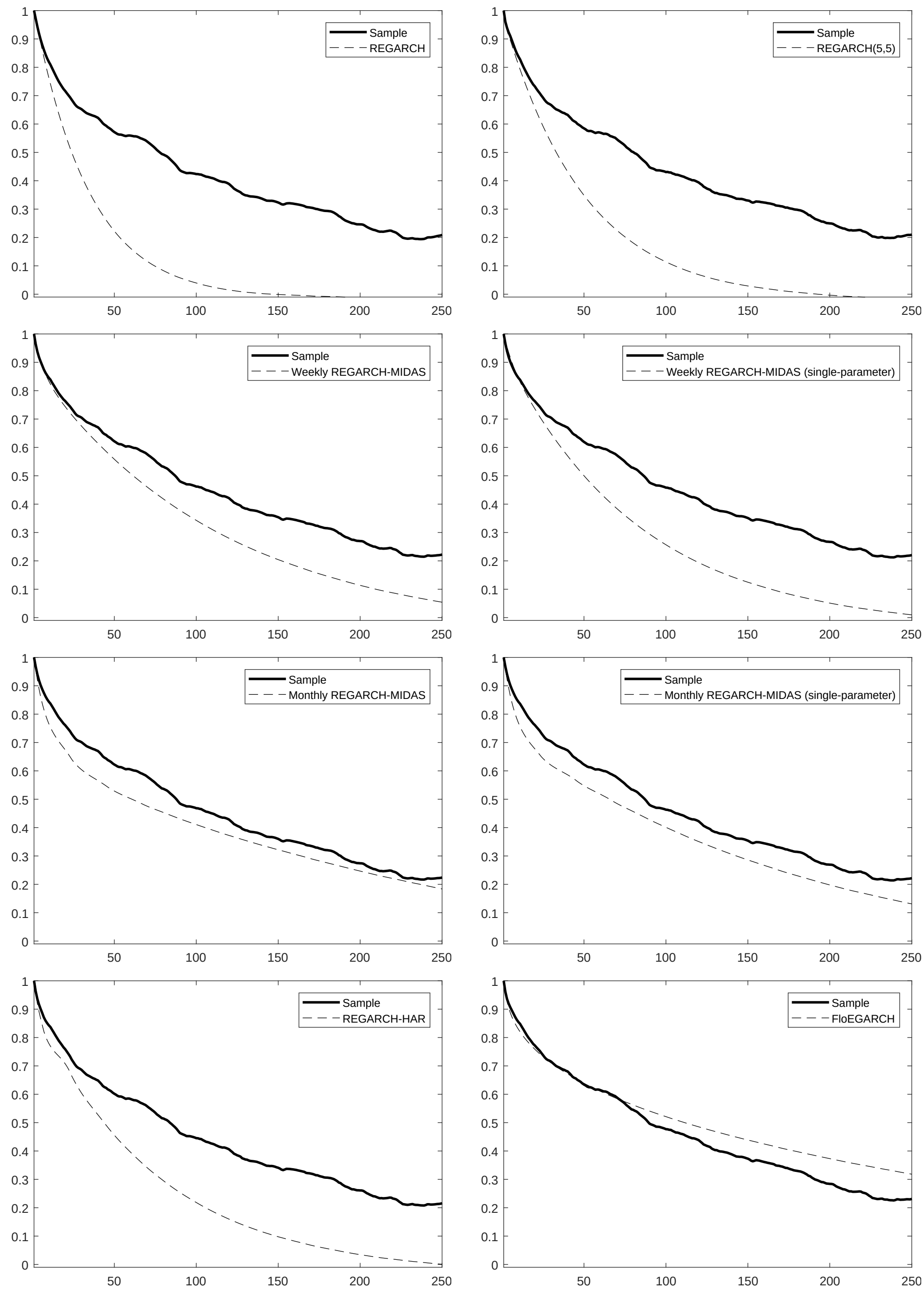

Figure 4: Simulated and sample autocorrelation function of $\log\sigma_t^2$

This figure depicts the simulated (dashed line) and sample (solid line) autocorrelation function of $\log\sigma_t^2$ for the REGARCH, REGARCH(5,5), REGARCH-MIDAS, REGARCH-HAR and the FloEGARCH. We use the estimated parameters for SPY reported in Table 1 and $K = 52$ ($K = 12$) for the weekly (monthly) REGARCH-MIDAS. See Section B.3 for additional details on their computation.

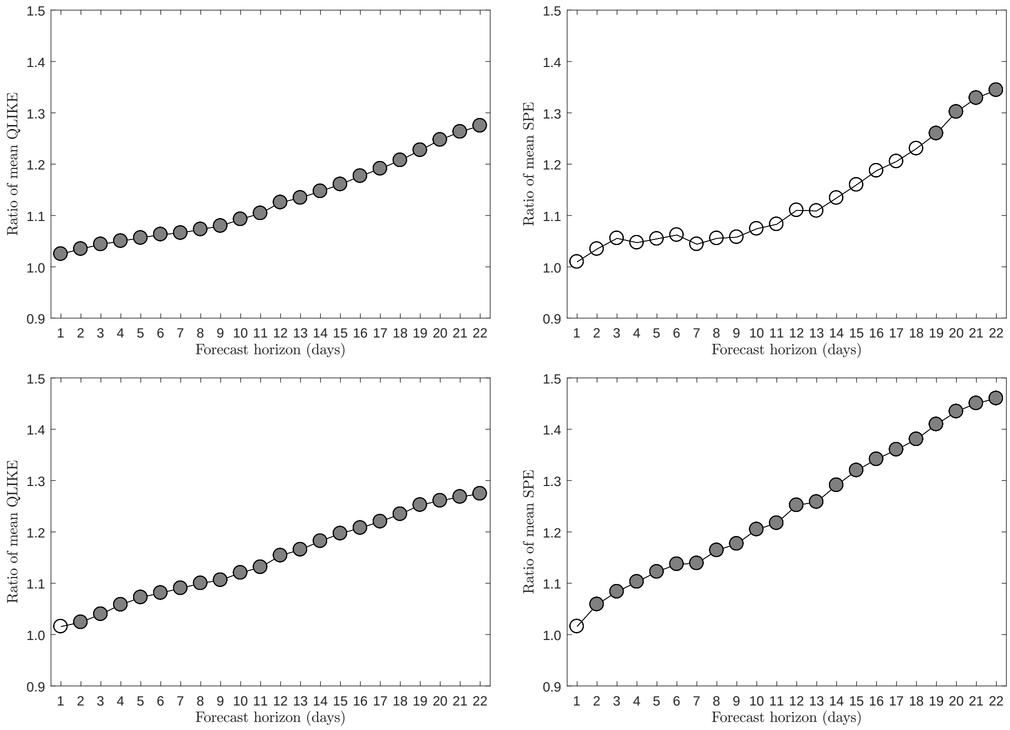

Figure 5: Forecast evaluation of REGARCH-MIDAS and REGARCH-HAR
This figure depicts the ratio of forecast losses of the REGARCH-MIDAS and REGARCH-HAR to the original REGARCH. Values exceeding unity indicate improvements in predictive ability of our proposed models. Full circles indicate whether difference in forecast loss (for a given forecast horizon) is significant on a 5% significance level using a Diebold-Mariano test for equal predictive ability. Empty circles indicate insignificance. See Section C.1 for additional details. The left panel uses the QLIKE loss function in (15), whereas the right panel uses the SPE loss function in (14). The upper panel reports results for the weekly single-parameter REGARCH-MIDAS and the lower panel for the REGARCH-HAR (results for the remaining REGARCH-MIDAS specifications are similar and are available upon request).

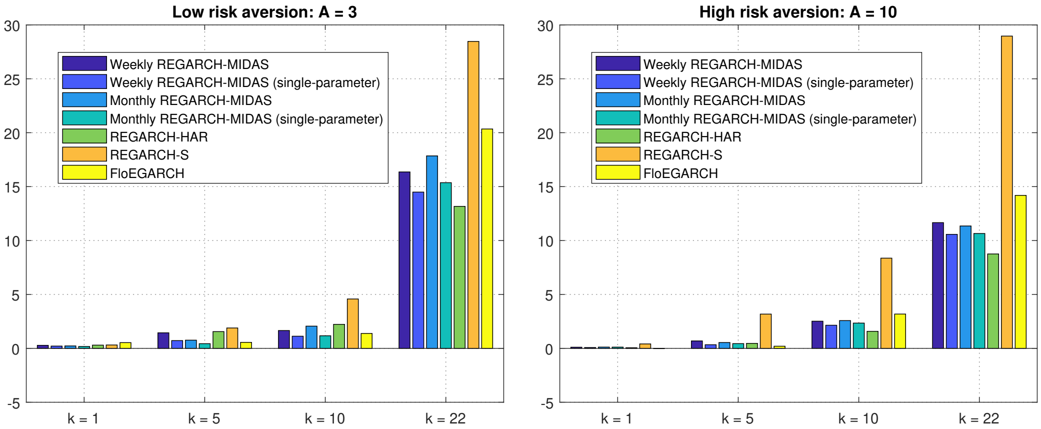

Figure 6: Economic value of volatility-timing strategy

This figure depicts the performance fee in annualized basis points as computed via (16) for the REGARCH-MIDAS, REGARCH-HAR, REGARCH-S and FLoEGARCH relative to the original REGARCH. The investment horizon is set to $1, 5, 10$, and 22 days as indicated by the x-axis. The left figure contains results for a low risk-aversion parameter, $A = 3$, and the right figure for a high risk-aversion parameter, $A = 10$.

B. Tables

Table 1: Full sample results for SPY

This table reports estimated parameters, robust standard errors (using the sandwich formula and reported in parentheses), number of parameters (p), information criteria, variance ratio from (13) and partial ($\mathscr{L}^p$), as well as full maximized log-likelihood value ($\mathscr{L}$) for each model under consideration. 5% critical values of the bootstrapped LR (BLR) statistic are reported in parenthesis and the associated p-values in square brackets. Results for the REGARCH-MIDAS are for $K = 52$ ($K = 12$) in the weekly (monthly) case.

	EGARCH	REGARCH	REGARCH-MIDAS (weekly)	REGARCH-MIDAS (weekly) (single-parameter)	REGARCH-MIDAS (monthly)	REGARCH-MIDAS (monthly) (single-parameter)	REGARCH-HAR	REGARCH-S	FloEGARCH
μ	0.020(0.0135)	0.016(0.0127)	0.015(0.0144)	0.015(0.0143)	0.016(0.0143)	0.016(0.0143)	0.014(0.0144)	0.024(0.0143)	0.015(0.0101)
β	0.981(0.0025)	0.972(0.0036)	0.761(0.0166)	0.842(0.0118)	0.872(0.0098)	0.880(0.0094)	0.734(0.0180)	0.943(0.0058)	0.176(0.0274)
α	0.121(0.0144)	0.338(0.0225)	0.337(0.0274)	0.329(0.0250)	0.324(0.0239)	0.324(0.0238)	0.355(0.0270)	0.324(0.0216)	0.370(0.0226)
ξ		−0.265(0.0267)	−0.271(0.0269)	−0.270(0.0269)	−0.269(0.0269)	−0.269(0.0269)	−0.272(0.0268)	−0.264(0.0264)	−0.274(0.0267)
σ_u^2		0.155(0.0058)	0.150(0.0057)	0.150(0.0057)	0.150(0.0057)	0.150(0.0057)	0.151(0.0057)	0.153(0.0057)	0.150(0.0057)
τ_1	−0.138(0.0118)	−0.148(0.0074)	−0.170(0.0084)	−0.166(0.0081)	−0.164(0.0079)	−0.163(0.0079)	−0.171(0.0085)	−0.150(0.0075)	−0.170(0.0083)
τ_2		0.040(0.0049)	0.047(0.0055)	0.045(0.0053)	0.044(0.0051)	0.044(0.0051)	0.047(0.0056)	0.041(0.0051)	0.051(0.0054)
δ_1		−0.113(0.0083)	−0.115(0.0083)	−0.115(0.0083)	−0.115(0.0083)	−0.115(0.0083)	−0.114(0.0083)	−0.112(0.0084)	−0.115(0.0082)
δ_2		0.049(0.0059)	0.051(0.0060)	0.050(0.0059)	0.050(0.0059)	0.050(0.0059)	0.051(0.0060)	0.050(0.0062)	0.051(0.0059)
ϕ		0.962(0.0253)	0.968(0.0167)	0.969(0.0187)	0.970(0.0198)	0.970(0.0201)	0.964(0.0163)	0.961(0.0232)	0.969(0.0231)
ω	0.058(0.1632)	−0.089(0.1255)	0.243(0.0397)	0.225(0.0458)	0.222(0.0499)	0.213(0.0515)	0.235(0.0386)	0.175(0.2366)	−0.092(0.1670)
λ			0.947(0.0298)	0.906(0.0327)	0.914(0.0426)	0.888(0.0397)			
γ_1			0.025(0.4475)		−0.518(0.7866)		0.294(0.0429)		
γ_2			6.337(6.3325)	12.545(1.8349)	2.063(2.8431)	8.508(1.4365)	0.620(0.0461)		
d									0.649(0.0142)
p	5	11	14	13	14	13	13	18	12
$\mathscr{L}$		−5,623.55	−5,577.52	−5,578.02	−5,577.52	−5,578.23	−5,595.10	−5,589.44	−5,584.65
$\mathscr{L}^p$	−4,213.84	−4,148.71	−4,159.70	−4,157.43	−4,156.60	−4,156.10	−4,160.22	−4,148.62	−4,162.54
BLR			92.05	91.07	92.06	90.64	56.90	68.22	77.81
(crit. value)			(27.98)	(24.40)	(23.45)	(20.58)	(18.13)	(24.35)	(10.43)
[p-value]			[< 0.01]	[< 0.01]	[< 0.01]	[< 0.01]	[< 0.01]	[< 0.01]	[< 0.01]
AIC		11,269.10	11,183.04	11,182.03	11,183.04	11,182.46	11,216.20	11,214.87	11,193.29
BIC		11,335.24	11,267.23	11,260.20	11,267.22	11,260.62	11,294.37	11,323.11	11,265.45
HQIC		11,423.39	11,379.41	11,364.37	11,379.40	11,364.79	11,398.53	11,467.34	11,361.60
VR			0.73	0.65	0.61	0.59	0.74	0.40	

Table 2: Model Confidence Set evaluation for individual stocks and SPY

This table reports the percentage of stocks and SPY for which a given model is included in the Model Confidence Set (MCS) on a 10% significance level and one-step and multi-step ahead forecasting. For a fixed significance level, α, the procedure identifies the MCS, $\hat{M}^*_{\alpha}$, from the set of competing models, M_0, which contains the best models with $1-\alpha$ probability (asymptotically as the length of the out-of-sample window approaches infinity). The procedure is conducted recursively based on an equivalence test for any $M \subseteq M_0$ and an elimination rule, which identifies and removes a given model from M in case of rejection of the equivalence test. The equivalence test is based on pairwise comparisons using a conventional t-statistic, $S_{i,j}$, for all $i,j \in M$ and the range statistic $T_M = \max_{i,j\in M}\{|S_{ij}|\}$, where the eliminated model is identified by $\text{argmax}_{i\in M}\sup_{j\in M}\{S_{ij}\}$. Following Hansen et al. (2011), we implement the procedure using a block bootstrap and 10^5 replications. Shaded grey numbers indicate the highest inclusion frequency for each forecast horizon.

Horizon	REGARCH	REGARCH-MIDAS (weekly)	REGARCH-MIDAS (weekly) (single-parameter)	REGARCH-MIDAS (monthly)	REGARCH-MIDAS (monthly) (single-parameter)	REGARCH-HAR	REGARCH-S	FloEGARCH
Panel A: QLIKE loss function								
k = 1	0.57	0.67	0.81	0.52	0.57	0.71	0.38	0.76
k = 2	0.29	0.67	0.71	0.52	0.52	0.67	0.38	0.05
k = 3	0.33	0.67	0.71	0.52	0.57	0.67	0.43	0.10
k = 4	0.38	0.76	0.81	0.71	0.67	0.67	0.43	0.14
k = 5	0.43	0.71	0.76	0.81	0.76	0.67	0.52	0.38
k = 10	0.29	0.67	0.76	0.81	0.76	0.67	0.52	0.76
k = 15	0.29	0.71	0.86	0.90	0.76	0.71	0.52	0.81
k = 22	0.33	0.76	0.76	0.90	0.90	0.71	0.76	0.86
Panel B: Squared Predition Error loss function								
k = 1	0.43	0.57	0.71	0.57	0.67	0.52	0.52	0.86
k = 2	0.29	0.76	0.81	0.67	0.71	0.62	0.48	0.19
k = 3	0.33	0.76	0.81	0.67	0.67	0.67	0.43	0.10
k = 4	0.33	0.76	0.81	0.67	0.57	0.62	0.52	0.19
k = 5	0.29	0.76	0.81	0.76	0.76	0.71	0.52	0.33
k = 10	0.10	0.71	0.81	0.81	0.86	0.76	0.62	0.76
k = 15	0.24	0.71	0.86	0.90	0.86	0.76	0.76	0.81
k = 22	0.14	0.71	0.76	0.86	0.86	0.67	0.67	0.86

Supplementary Appendix for

Capturing volatility persistence: A dynamically complete Realized EGARCH-MIDAS model

Abstract

This Supplementary Appendix provides results, in this order, on the score function of the log-likelihood functions associated with the REGARCH-MIDAS and REGARCH-HAR, the choice of the lag length of the MIDAS filter for the use in our empirical analysis, the autocorrelation function of the realized kernel of the models under consideration as well as summary statistics and in- and out-of-sample results for our sample of 20 individual stocks.

A. Score function

Since the scores define the first order conditions for the maximum-likelihood estimator and facilitate direct computation of standard errors for the coefficients, we present closed-form expressions for the scores in the following. To simplify notation, we write $\tau(z) = \tau' a(z)$ and $\delta(z) = \delta' b(z)$ with $a(z) = b(z) = \left(z, z^2 - 1\right)'$, and let $\dot{a}_{z_t} = \partial a(z_t)/\partial z_t$ and $\dot{b}_{z_t} = \partial b(z_t)/\partial z_t$. In addition, we define $\theta_1 = (\beta, \tau_1, \tau_2, \alpha)'$, $\theta_2 = (\xi, \phi, \delta_1, \delta_2)'$, $m_t = (\log h_t, a(z_t)', u_t)'$, and $n_t = (1, \log \sigma_t^2, b(z_t)')'$.

Proposition 1 (Scores)**.** *The scores,* $\frac{\partial \ell}{\partial \theta} = \sum_{t=1}^{T} \frac{\partial \ell_t}{\partial \theta}$, *are given from*

$$\frac{\partial \ell_t}{\partial \theta} = \begin{pmatrix} B(z_t, u_t)\dot{h}_{\mu,t} - \left[z_t - \delta' \frac{u_t}{\sigma_u^2} \dot{b}_{z_t}\right] \frac{1}{\sigma_t} \\ B(z_t, u_t)\dot{h}_{\theta_1,t} \\ B(z_t, u_t)\dot{h}_{\theta_2,t} + \frac{u_t}{\sigma_u^2} n_t \\ B(z_t, u_t)\dot{h}_{\omega,t} + D(z_t, u_t)\dot{g}_{\omega,t} \\ B(z_t, u_t)\dot{h}_{\eta,t} + D(z_t, u_t)\dot{g}_{\eta,t} \\ \frac{1}{2} \frac{u_t^2 - \sigma_u^2}{\sigma_u^4} \end{pmatrix}, \tag{A.1}$$

and

$$A(z_t) = \frac{\partial \log h_{t+1}}{\partial \log h_t} = \left(\beta - \alpha\phi\right) + \frac{1}{2}\left(\alpha\delta' \dot{b}_{z_t} - \tau' \dot{a}_{z_t}\right) z_t, \tag{A.2}$$

$$B(z_t, u_t) = \frac{\partial \ell_t}{\partial \log h_t} = -\frac{1}{2}\left[(1 - z_t^2) + \frac{u_t}{\sigma_u^2}\left(\delta' \dot{b}_{z_t} z_t - 2\phi\right)\right], \tag{A.3}$$

$$C(z_t) = \frac{\partial \log h_{t+1}}{\partial \log g_t} = -\alpha\phi + \frac{1}{2}\left(\alpha\delta' \dot{b}_{z_t} - \tau' \dot{a}_{z_t}\right) z_t, \tag{A.4}$$

$$D(z_t, u_t) = \frac{\partial \ell_t}{\partial \log g_t} = -\frac{1}{2}\left[(1 - z_t^2) + \frac{u_t}{\sigma_u^2}\left(\delta' \dot{b}_{z_t} z_t - 2\phi\right)\right]. \tag{A.5}$$

Furthermore, we have

$$\begin{aligned}
\dot{h}_{\mu,t+1} &= \frac{\partial \log h_{t+1}}{\partial \mu} = A(z_t)\dot{h}_{\mu,t} + \left(\alpha\delta'\dot{b}_{z_t} - \tau'\dot{a}_{z_t}\right)\frac{1}{\sigma_t}, &\text{(A.6)}\\
\dot{h}_{\theta_1,t+1} &= \frac{\partial \log h_{t+1}}{\partial \theta_1} = A(z_t)\dot{h}_{\theta_1,t} + m_t, &\text{(A.7)}\\
\dot{h}_{\theta_2,t+1} &= \frac{\partial \log h_{t+1}}{\partial \theta_2} = A(z_t)\dot{h}_{\theta_2,t} + \alpha n_t, &\text{(A.8)}\\
\dot{h}_{\omega,t+1} &= \frac{\partial \log h_{t+1}}{\partial \omega} = A(z_t)\dot{h}_{\omega,t} + C(z_t), &\text{(A.9)}\\
\dot{h}_{\eta,t+1} &= \frac{\partial \log h_{t+1}}{\partial \eta} = A(z_t)\dot{h}_{\eta,t} + C(z_t)\dot{g}_{\eta,t}, &\text{(A.10)}
\end{aligned}$$

where $\dot{g}_{\eta,t}$ depends on the specification of $f(\cdot;\eta)$ and is presented below in the proof.

By corollary, the score function is a Martingale Difference Sequence (MDS), provided that $\mathbb{E}[z_t|\mathscr{F}_{t-1}] = 0$, $\mathbb{E}\left[z_t^2|\mathscr{F}_{t-1}\right] = 1$, $\mathbb{E}[u_t|z_t,\mathscr{F}_{t-1}] = 0$, and $\mathbb{E}\left[u_t^2|z_t,\mathscr{F}_{t-1}\right] = \sigma_u^2$, which is useful for future analysis of the asymptotic properties of the QML estimator.[1]

Proof: First, consider $A(z_t) = \partial \log h_{t+1}/\partial \log h_t$ and $C(z_t) = \partial \log h_{t+1}/\partial \log g_t$. From $z_t = \frac{r_t-\mu}{\sigma_t}$, it can easily be shown that

$$\frac{z_t}{\log h_t} = \frac{z_t}{\log g_t} = -\frac{1}{2}z_t. \tag{A.11}$$

From $u_t = \log x_t - \phi \log \sigma_t^2 - \delta(z_t)$, we find

$$\begin{aligned}
\frac{\partial u_t}{\partial \log h_t} &= -\delta'\frac{\partial b(z_t)}{\partial z_t}\frac{\partial z_t}{\log h_t} - \phi = -\delta'\dot{b}_{z_t}\frac{\partial z_t}{\log h_t} - \phi, &\text{(A.12)}\\
\frac{\partial u_t}{\partial \log g_t} &= -\delta'\frac{\partial b(z_t)}{\partial z_t}\frac{\partial z_t}{\log g_t} - \phi = -\delta'\dot{b}_{z_t}\frac{\partial z_t}{\log g_t} - \phi. &\text{(A.13)}
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\frac{\partial \tau(z_t)}{\partial \log h_t} &= \tau'\frac{\partial a(z_t)}{\partial z_t}\frac{\partial z_t}{\log h_t} = \tau'\dot{a}_{z_t}\frac{\partial z_t}{\log h_t}, &\text{(A.14)}\\
\frac{\partial \tau(z_t)}{\partial \log g_t} &= \tau'\frac{\partial a(z_t)}{\partial z_t}\frac{\partial z_t}{\log g_t} = \tau'\dot{a}_{z_t}\frac{\partial z_t}{\log g_t}. &\text{(A.15)}
\end{aligned}$$

[1] These are the same conditions as in Hansen and Huang (2016) and we refer the reader hereto for further details.

Inserting the above components in the following expressions for $A(z_t)$ and $C(z_t)$

$$A(z_t) = \frac{\partial \log h_{t+1}}{\partial \log h_t} \quad = \quad \beta + \frac{\partial \tau(z_t)}{\partial \log h_t} + \alpha \frac{\partial u_t}{\partial \log h_t}, \tag{A.16}$$

$$C(z_t) = \frac{\partial \log h_{t+1}}{\partial \log g_t} \quad = \quad \frac{\partial \tau(z_t)}{\partial \log g_t} + \alpha \frac{\partial u_t}{\partial \log g_t}, \tag{A.17}$$

yields

$$A(z_t) \quad = \quad (\beta - \alpha\phi) + \frac{1}{2}\left(\alpha\delta' \dot{b}_{z_t} - \tau' \dot{a}_{z_t}\right) z_t, \tag{A.18}$$

$$C(z_t) \quad = \quad -\alpha\phi + \frac{1}{2}\left(\alpha\delta' \dot{b}_{z_t} - \tau' \dot{a}_{z_t}\right) z_t. \tag{A.19}$$

Next, we turn to $B(z_t, u_t) = \partial \ell_t / \partial \log h_t$ and $D(z_t, u_t) = \partial \ell_t / \partial \log g_t$. The terms $\log h_t$ and $\log g_t$ enter the log-likelihood contribution at time t directly due to $\log \sigma_t^2 = \log h_t + \log g_t$ and indirectly through z_t^2 and u_t^2. Thus, we have

$$B(z_t, u_t) \quad = \quad -\frac{1}{2}\left[1 + \frac{\partial z_t^2}{\partial \log h_t} + \frac{1}{\sigma_u^2} 2u_t \frac{\partial u_t}{\partial \log h_t}\right], \tag{A.20}$$

$$D(z_t, u_t) \quad = \quad -\frac{1}{2}\left[1 + \frac{\partial z_t^2}{\partial \log g_t} + \frac{1}{\sigma_u^2} 2u_t \frac{\partial u_t}{\partial \log g_t}\right]. \tag{A.21}$$

We note that

$$\frac{\partial \ell_t}{\partial \log g_t} = \frac{\partial \ell_t}{\partial \log h_t} = -z_t^2. \tag{A.22}$$

Combining the different expressions yields

$$B(z_t, u_t) \quad = \quad -\frac{1}{2}\left[(1 - z_t^2) + \frac{u_t}{\sigma_u^2}\left(\delta' \dot{b}_{z_t} z_t - 2\phi\right)\right], \tag{A.23}$$

$$D(z_t, u_t) \quad = \quad -\frac{1}{2}\left[(1 - z_t^2) + \frac{u_t}{\sigma_u^2}\left(\delta' \dot{b}_{z_t} z_t - 2\phi\right)\right]. \tag{A.24}$$

Now, we turn to the derivatives of $\log h_{t+1}$ with respect to the different parameters. For $\dot{h}_{\mu,t+1} = \partial h_{t+1} / \partial \mu$, we have

$$\dot{h}_{\mu,t+1} = \beta \frac{\partial \log h_t}{\partial \mu} + \frac{\partial \tau(z_t)}{\partial \mu} + \alpha \frac{\partial u_t}{\partial \mu}, \tag{A.25}$$

where

$$\begin{aligned}
\frac{\partial\tau(z_t)}{\partial\mu} &= \frac{\partial\tau(z_t)}{\partial z_t}\frac{\partial z_t}{\partial\mu} = \tau'\dot{a}_{z_t}\left[-\frac{1}{2}z_t\frac{\partial\log h_t}{\partial\mu} - \frac{1}{\sigma_t}\right], && \text{(A.26)}\\
\frac{\partial u_t}{\partial\mu} &= -\phi\frac{\partial\log h_t}{\partial\mu} - \delta'\dot{b}_{z_t}\frac{\partial z_t}{\partial\mu}\\
&= -\phi\frac{\partial\log h_t}{\partial\mu} - \delta'\dot{b}_{z_t}\left[-\frac{1}{2}z_t\frac{\partial\log h_t}{\partial\mu} - \frac{1}{\sigma_t}\right]. && \text{(A.27)}
\end{aligned}$$

Inserting (A.26) and (A.27) in (A.25) and rearranging yields

$$\begin{aligned}
\dot{h}_{\mu,t+1} &= \left[(\beta-\alpha\phi) + \frac{1}{2}\left[\alpha\delta'\dot{b}_{z_t} - \tau'\dot{a}_{z_t}\right]z_t\right]\frac{\partial\log h_t}{\partial\mu} + \left[\alpha\delta'\dot{b}_{z_t} - \tau'\dot{a}_{z_t}\right]\frac{1}{\sigma_t}\\
&= A(z_t)\dot{h}_{\mu,t} + \left[\alpha\delta'\dot{b}_{z_t} - \tau'\dot{a}_{z_t}\right]\frac{1}{\sigma_t}. && \text{(A.28)}
\end{aligned}$$

For $\dot{h}_{\theta_1,t+1} = \partial h_{t+1}/\partial\theta_1$, we have

$$\dot{h}_{\theta_1,t+1} = \beta\frac{\partial\log h_t}{\partial\theta_1} + \frac{\partial\tau(z_t)}{\partial\theta_1} + \alpha\frac{\partial u_t}{\partial\theta_1} + (\log h_t, z_t, z_t^2-1, u_t)'. \tag{A.29}$$

However, we remember that $\tau(z_t)$ and u_t only depend on θ_1 through $\log h_t$ such that we can reduce the first three terms into one

$$\begin{aligned}
\dot{h}_{\theta_1,t+1} &= \frac{\partial\log h_{t+1}}{\partial\log h_t}\frac{\partial\log h_t}{\partial\theta_1} + (\log h_t, z_t, z_t^2-1, u_t)'\\
&= A(z_t)\dot{h}_{\theta_1,t} + m_t. && \text{(A.30)}
\end{aligned}$$

For $\dot{h}_{\theta_2,t+1} = \partial h_{t+1}/\partial\theta_2$, $\dot{h}_{\omega,t+1} = \partial h_{t+1}/\partial\omega$ and $\dot{h}_{\eta,t+1} = \partial h_{t+1}/\partial\eta$, we obtain

$$\begin{aligned}
\dot{h}_{\theta_2,t+1} &= \frac{\partial\log h_{t+1}}{\partial\log h_t}\frac{\partial\log h_t}{\partial\theta_2} + \alpha(1, \log\sigma_t^2, z_t, z_t^2-1)'\\
&= A(z_t)\dot{h}_{\theta_2,t} + n_t, && \text{(A.31)}\\
\dot{h}_{\omega,t+1} &= \frac{\partial\log h_{t+1}}{\partial\log h_t}\frac{\partial\log h_t}{\partial\omega} + \frac{\partial\log h_{t+1}}{\partial\log g_t}\frac{\partial\log g_t}{\partial\omega}\\
&= A(z_t)\dot{h}_{\omega,t} + C(z_t), && \text{(A.32)}\\
\dot{h}_{\eta,t+1} &= \frac{\partial\log h_{t+1}}{\partial\log h_t}\frac{\partial\log h_t}{\partial\eta} + \frac{\partial\log h_{t+1}}{\partial\log g_t}\frac{\partial\log g_t}{\partial\eta}\\
&= A(z_t)\dot{h}_{\eta,t} + C(z_t)\dot{g}_{\eta,t}, && \text{(A.33)}
\end{aligned}$$

respectively. Finally, we turn to the scores. The parameter μ enters the log-likelihood contribution at time t through $\log h_t$, z_t, and u_t^2 such that

$$\begin{aligned}
\frac{\partial \ell_t}{\partial \mu} &= -\frac{1}{2}\dot{h}_{\mu,t} - \frac{z_t^2}{\partial \mu} - \frac{1}{2}\frac{1}{\sigma_u^2}\frac{\partial u_t^2}{\partial \mu} \\
&= \frac{\partial \ell_t}{\partial \log h_t}\frac{\partial \log h_t}{\partial \mu} - \left[z_t - \delta' \frac{u_t}{\sigma_u^2}\dot{b}_{z_t}\right]\frac{1}{\sigma_t} \\
&= B(z_t, u_t)\dot{h}_{\mu,t} - \left[z_t - \delta' \frac{u_t}{\sigma_u^2}\dot{b}_{z_t}\right]\frac{1}{\sigma_t}. \qquad \text{(A.34)}
\end{aligned}$$

Since θ_1 only enters the log-likelihood contribution at time t indirectly through $\log h_t$, an application of the chain-rule yields

$$\frac{\partial \ell_t}{\partial \theta_1} = B(z_t, u_t)\dot{h}_{\theta_1,t}. \qquad \text{(A.35)}$$

The parameter vector θ_2 also enters through u_t^2,

$$\frac{\partial \ell_t}{\partial \theta_2} = B(z_t, u_t)\dot{h}_{\theta_2,t} + \frac{u_t}{\sigma_u^2} n_t. \qquad \text{(A.36)}$$

The parameters ω and η enter through $\log h_t$ and $\log g_t$,

$$\begin{aligned}
\frac{\partial \ell_t}{\partial \omega} &= B(z_t, u_t)\dot{h}_{\omega,t} + D(z_t, u_t)\dot{g}_{\omega,t}, \\
\frac{\partial \ell_t}{\partial \eta} &= B(z_t, u_t)\dot{h}_{\eta,t} + D(z_t, u_t)\dot{g}_{\eta,t}. \qquad \text{(A.37)}
\end{aligned}$$

The parameter σ_u^2 only enters directly in the log-likelihood contribution such that

$$\frac{\partial \ell_t}{\partial \sigma_u^2} = \frac{1}{2}\frac{u_t^2 - \sigma_u^2}{\sigma_u^2}. \qquad \text{(A.38)}$$

Stacking the above scores,

$$\frac{\partial \ell_t}{\partial \theta} = \left(\frac{\partial \ell_t}{\partial \mu}, \frac{\partial \ell_t}{\partial \theta_1'}, \frac{\partial \ell_t}{\partial \theta_2'}, \frac{\partial \ell_t}{\partial \omega}, \frac{\partial \ell_t}{\partial \eta'}, \frac{\partial \ell_t}{\partial \sigma_u^2}\right)', \qquad \text{(A.39)}$$

yields the result in Proposition 1.

A.1. Derivatives specific to the long-run component

In the REGARCH-HAR, we have $\eta = (\gamma_1, \gamma_2)'$ such that

$$\dot{g}_{\eta,t} = \begin{pmatrix} \frac{1}{5}\sum_{i=1}^{5} \log x_{t-i-1} \\ \frac{1}{22}\sum_{i=1}^{22} \log x_{t-i-1} \end{pmatrix}. \tag{A.40}$$

In the two-parameter REGARCH-MIDAS, we have $\eta = (\lambda, \gamma_1, \gamma_2)'$ such that

$$\dot{g}_{\eta,t} = \begin{pmatrix} \sum_{k=1}^{K} \pi_k(\gamma_1,\gamma_2) y_{t-1,k} \\ \sum_{k=1}^{K} \frac{(\gamma_1-1)\left(1-\frac{k}{K}\right)^{\gamma_2-1}\left(\frac{k}{K}\right)^{\gamma_1-1}\sum_{j=1}^{K}\left(1-\frac{j}{K}\right)^{\gamma_2-1}\left(\frac{k}{K}-\left(\frac{j}{K}\right)^{-1}\right)\left(\frac{j}{K}\right)^{\gamma_{1,i}-1}}{\left[\sum_{j=1}^{K}\left(\frac{j}{K}\right)^{\gamma_1-1}\left(1-\frac{j}{K}\right)^{\gamma_2-1}\right]^2} y_{t-1,k} \\ \sum_{k=1}^{K} \frac{(\gamma_2-1)\left(1-\frac{k}{K}\right)^{\gamma_2-1}\left(\frac{k}{K}\right)^{\gamma_1-1}\sum_{j=1}^{K}\left(1-\frac{j}{K}\right)^{\gamma_2-1}\left(1-\frac{k}{K}-\left(1-\frac{j}{K}\right)^{-1}\right)\left(\frac{j}{K}\right)^{\gamma_1-1}}{\left[\sum_{j=1}^{K}\left(\frac{j}{K}\right)^{\gamma_1-1}\left(1-\frac{j}{K}\right)^{\gamma_2-1}\right]^2} y_{t-1,k} \end{pmatrix}. \tag{A.41}$$

In the single-parameter REGARCH-MIDAS, we have $\eta = (\lambda, \gamma_2)'$ such that

$$\dot{g}_{\eta,t} = \begin{pmatrix} \sum_{k=1}^{K} \pi_k(\gamma_1,\gamma_2) y_{t-1,k} \\ \sum_{k=1}^{K} \frac{(\gamma_2-1)\left(1-\frac{k}{K}\right)^{\gamma_2-1}\sum_{j=1}^{K}\left(1-\frac{j}{K}\right)^{\gamma_2-1}\left(1-\frac{k}{K}-\left(1-\frac{j}{K}\right)^{-1}\right)}{\left[\sum_{j=1}^{K}\left(1-\frac{j}{K}\right)^{\gamma_2-1}\right]^2} y_{t-1,k} \end{pmatrix}. \tag{A.42}$$

B. Checking validity of the asymptotic distribution of the estimators

To check the validity of the assumed asymptotic distribution of the estimators, we employ a parametric bootstrapping technique (Paparoditis and Politis, 2009). We use 999 replications, a sample size of 2,500 observations (approximately 10 years, similar to the size of the rolling in-sample window used in the forecasting exercise below), and re-sampled residuals from the empirical distribution to bootstrap observations. Figure 1 depicts the empirical standardized distribution of a subset of the estimated parameters.

It stands out that the in-sample distribution of the estimated parameters for both the REGARCH, REGARCH-MIDAS and REGRACH-HAR is generally in agreement with a standard normal distribution. We also compared the bootstrapped standard errors with the robust QML standard errors computed from the sandwich-formula, which are reported in the empirical section below. The standard errors were quite similar, which in conjunction with Figure 1 does not contradict the assertion that the QML approach and associated inferences are valid. We do, however, note that the QML standard errors are slightly smaller on average relative to the bootstrapped standard errors, causing us to be careful in not putting too much weight on the role of standard errors in the interpretation of the results in the paper.

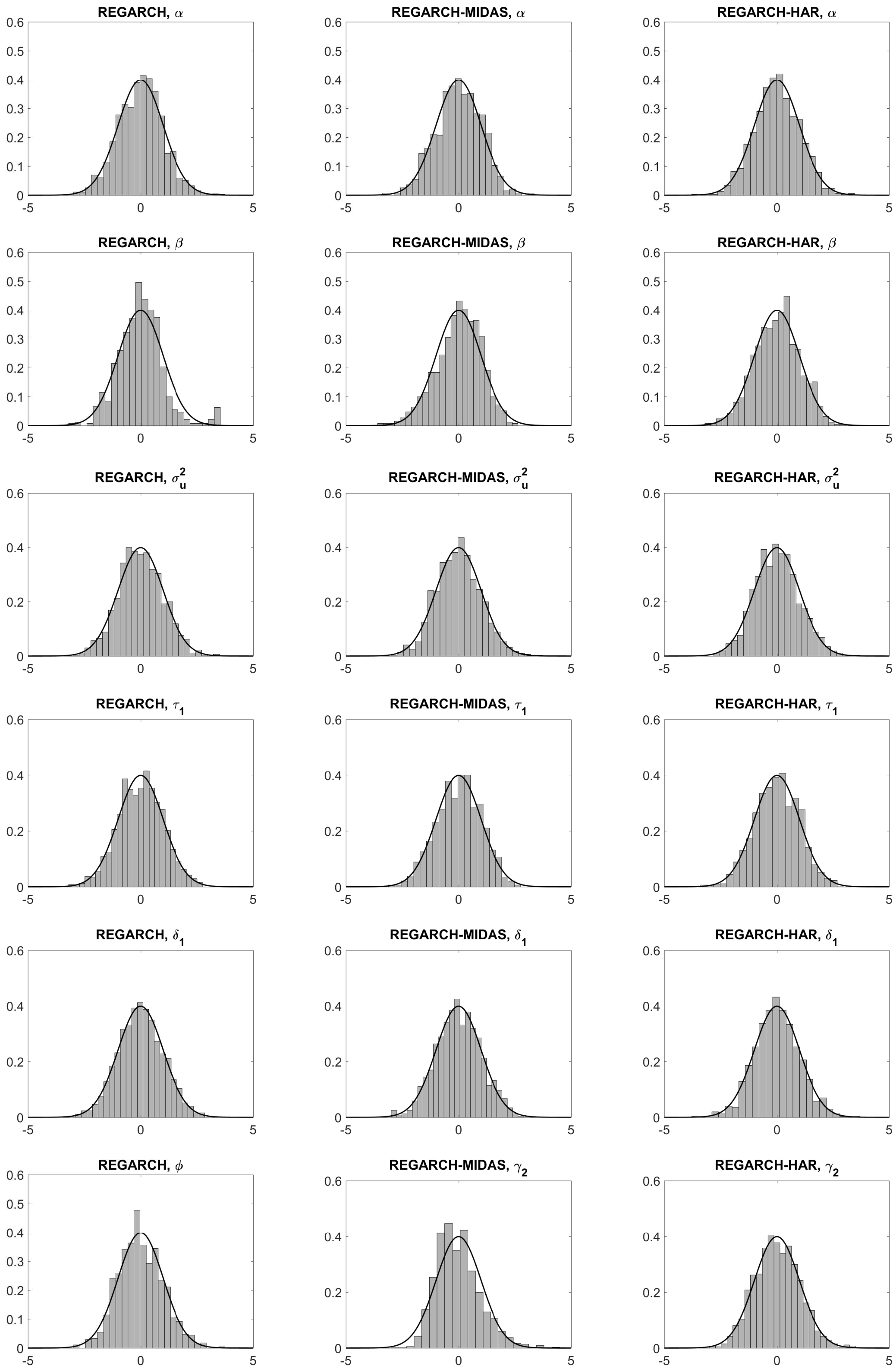

Figure 1: Standardized empirical distribution of estimated parameters
This figure depicts the standardized empirical distribution of a subset of the model parameters using a parametric bootstrap with resampling of the empirical residuals from the estimation on SPY (Paparoditis and Politis, 2009). We use 999 bootstrap replications and a sample size of 2,500 observations. The left column depicts results for the original REGARCH, the middle column for the weekly, single-parameter REGARCH-MIDAS, and the right column for the REGARCH-HAR.

C. Fractional integration parameter d and unit root tests

We estimate the fractional integrated parameter d in the logarithm of the realized kernel with the two-step exact local Whittle estimator of Shimotsu (2010). Over the full sample all series have $d > 0.5$, suggesting that volatility is highly persistent.[2] This finding is supported by the slowly decaying ACF of the logarithm of the realized kernel for SPY. Since the conventional ACF may be biased for the unobserved ACF of the logarithm conditional variance due to the presence of measurement errors,[3] we also compute the instrumented ACF proposed by Hansen and Lunde (2014). We use the authors' preferred specification with multiple instruments (four through ten) and optimal combination. The instrumented ACF show a similar pattern as the conventional ACF, but points toward an even higher degree of persistence. We also conducted a (Dickey-Fuller) unit root test across all assets considered using the instrumented persistence parameter (cf. Table 1).

The (biased) conventional least square estimates point to moderate persistence and strong rejection of a unit root. The persistence parameter is, as expected, notably higher when using the instrumented variables estimator of Hansen and Lunde (2014), however the null hypothesis of a unit root remains rejected for all assets. Collectively, these findings motivate a modeling framework that is capable of capturing a high degree of persistence. Given the requirement that $|\beta| < 1$, this also motivates a framework that pulls β away from unity. This is where the proposed REGARCH-MIDAS and REGARCH-HAR prove useful.

[2] We estimated the parameters with $m = \lfloor T^q \rfloor$ for $q = \in \{0.50, 0.55, \ldots, 0.80\}$, leading to no alterations of the conclusions obtained for $q = 0.65$. See also Wenger, Leschinski, and Sibbertsen (2017) for a comprehensive empirical study on long memory in volatility and the choice of estimator of d.

[3] The element of microstructure noise is, arguably, low, given the construction of the realized kernel, however sampling error may still be present, causing the differences in the conventional and instrumented ACF.

Table 1: Persistence analysis of realized kernel

This table reports estimated autoregressive persistence parameters, π, unit root tests, DF, and the fractional integration order of the logarithm of the realized kernel. The first column contains the conventional least squares estimator, whereas the following two columns contain the instrumented variables estimator from Hansen and Lunde (2014) using the first lag as instrument and their preferred specification (four through ten) with optimal combination, respectively. The following three columns contain the Dickey-Fuller unit root test using each estimate of the persistence parameter. The 1%, 5% and 10% critical values are -20.7, -14.1 and -11.3, respectively (see Fuller (1996), Table 10.A.1). The last column contains the fractional integrated parameter d, which is estimated using the two-step exact local Whittle estimator of Shimotsu (2010) and bandwidth choice of $m = \lfloor T^{0.65} \rfloor$. The standard error of these estimates is approximately 0.04.

	π_{OLS}	π_1	$\pi_{4:10}$	DF_{OLS}	DF_1	$\text{DF}_{4:10}$	d
SP500	0.883	0.959	0.985	-354.3	-124.8	-45.8	0.66
AA	0.865	0.961	0.985	-405.3	-116.6	-44.8	0.64
AIG	0.919	0.966	0.990	-242.4	-103.1	-30.5	0.64
AXP	0.926	0.980	0.992	-222.7	-59.0	-23.5	0.70
BA	0.847	0.956	0.987	-458.6	-131.7	-37.4	0.64
CAT	0.866	0.949	0.988	-400.6	-151.9	-35.6	0.67
DD	0.856	0.952	0.983	-431.6	-143.8	-51.8	0.63
DIS	0.866	0.956	0.986	-401.6	-132.3	-41.4	0.66
GE	0.904	0.969	0.990	-287.6	-93.8	-30.6	0.68
IBM	0.870	0.959	0.983	-389.5	-122.3	-52.0	0.65
INTC	0.869	0.951	0.985	-395.4	-148.0	-45.5	0.63
JNJ	0.852	0.955	0.988	-443.6	-134.3	-37.0	0.68
KO	0.836	0.953	0.985	-492.7	-140.4	-45.7	0.63
MMM	0.833	0.940	0.981	-499.5	-178.2	-57.3	0.64
MRK	0.815	0.942	0.983	-552.7	-174.7	-49.6	0.61
MSFT	0.857	0.951	0.981	-429.7	-146.7	-56.3	0.63
PG	0.818	0.937	0.980	-546.2	-188.9	-58.5	0.61
VZ	0.861	0.961	0.987	-414.7	-118.0	-38.5	0.67
WHR	0.823	0.938	0.986	-528.4	-186.6	-41.5	0.58
WMT	0.844	0.957	0.985	-467.4	-127.8	-43.8	0.65
XOM	0.878	0.954	0.980	-366.7	-137.8	-59.4	0.68

D. Choice of lag length, K

This section details a procedure for choosing a uniform value for the lag length of the MIDAS filter, K, or picking an optimal value.

As noted in the paper, the REGARCH-HAR utilizes by construction lagged information equal to four weeks (approximately one month) to describe the dynamics of the realized measure, whereas the REGARCH-MIDAS allows the researcher to explore and subsequently choose a suitable lag length, possibly beyond four weeks. For the original two-parameter setting as well as the single-parameter setting, Figure 2 depicts the estimated lag weights and associated maximized log-likelihood values of the weekly REGRACH-MIDAS on SPY for a range of K starting with four lags up to 104 lags (approximately two years).

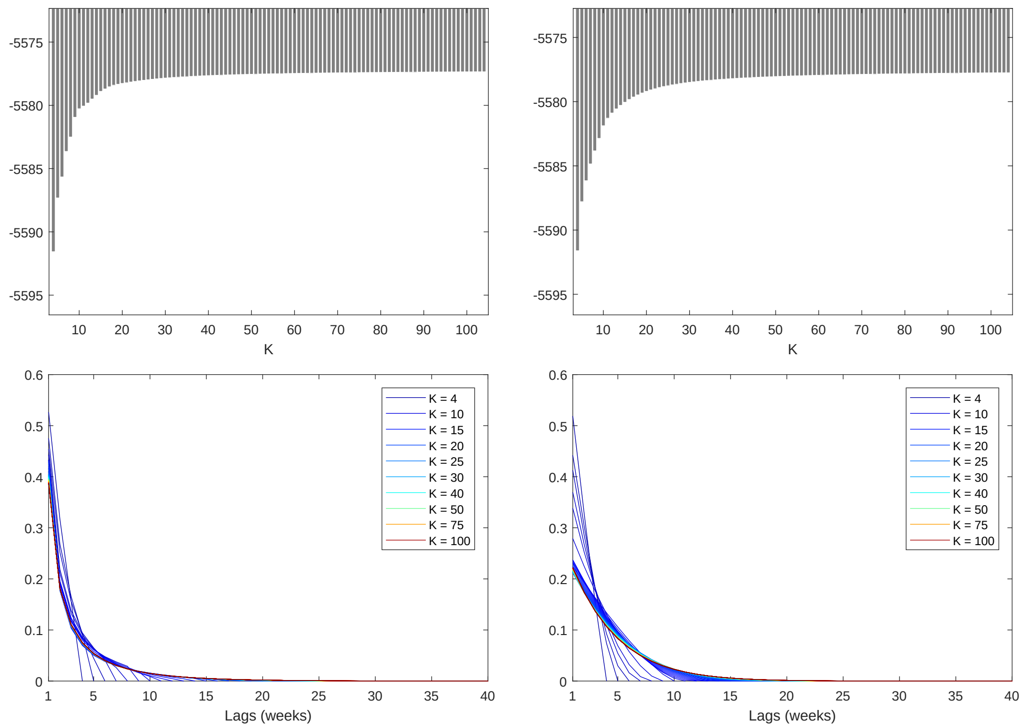

Figure 2: Lag length, K, for weekly REGARCH-MIDAS

This figure depicts in the upper panel the maximized log-likelihood values for SPY in the weekly two-parameter setting (left panel) and weekly single-parameter setting (right panel) for $K = 4, \ldots, 104$ weeks. The lower panel depicts the estimated lag function for a range of values of K.

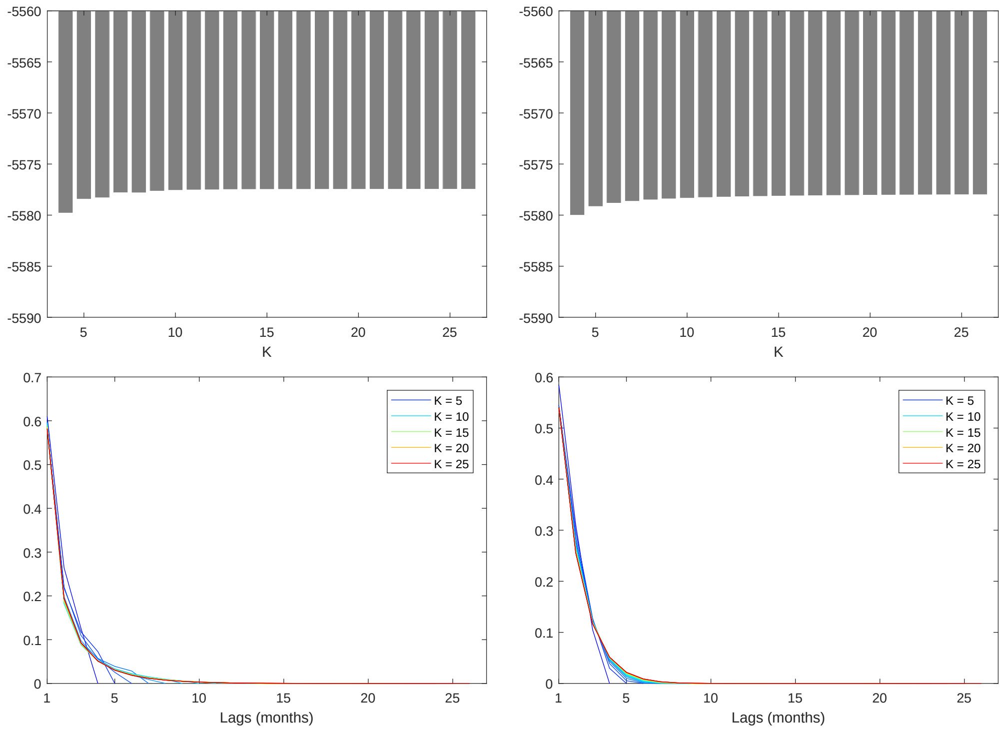

Figure 3: Lag length, K, for monthly REGARCH-MIDAS

This figure depicts in the upper panel the maximized log-likelihood values for SPY in the monthly two-parameter setting (left panel) and monthly single-parameter setting (right panel) for $K = 4, \ldots, 104$ weeks. The lower panel depicts the estimated lag function for a range of values of K.

The figure yields a number of interesting insights. First, the maximized log-likelihood values and associated patterns are very similar across the single-parameter and two-parameter case. The maximized log-likelihood values initially increase until lag 25-50, after which the values reach a ceiling. This observation is corroborated by the estimated lag functions in the lower panel of the figure. Their patterns show that recent information matters the most with the information content decaying to zero for lags approximately equal to 20 in the two-parameter setting and 25 in the single-parameter setting. Hence, based on the figure we may conclude that information up to half a year in the past is most important for explaining the dynamics of the conditional variance. This is generally supported by a similar analysis using monthly averages rather than weekly in the MIDAS component, but the monthly specification seems to indicate that additional past information is relevant (cf. Figure 3).

Second, a REGARCH-MIDAS with information only up to the past four weeks provides only a slightly greater log-likelihood value than the REGARCH-HAR (cf. Table 1 in the paper). This indicates that the step-function approximation in the REGARCH-HAR does a reasonable job at capturing the information content up to four weeks in the past. Collectively, however, these findings also suggest that the information lag in the REGARCH-HAR is too short. Based on these findings, we proceed in the paper with a value of $K = 52$ for the weekly MIDAS and $K = 12$ for the monthly MIDAS uniformly in all subsequent analyses, including the individual stock results. Note that we choose K larger than what the initial analysis suggests for the weekly specification, since we want consistency between the weekly and monthly specifications and greater flexibility when applying the choice to the individual stocks. We do, however, emphasize that it is free for the researcher to optimize over the choice of K for each individual asset to achieve an even better fit.

E. Autocorrelation function of realized kernel

This section contains the autocorrelation function of the logarithm realized kernel as observed in the market versus the one simulated from the models under consideration. Additional details can be found in the paper.

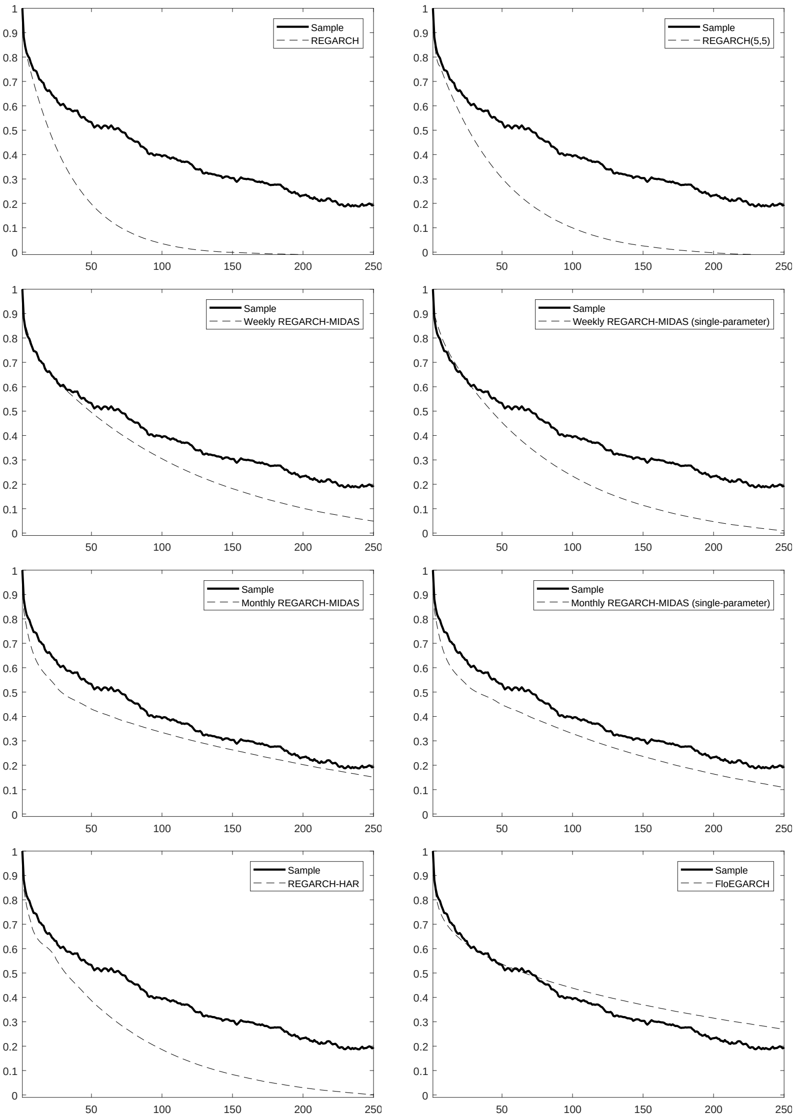

Figure 4: Simulated and sample autocorrelation function of $\log \mathbf{RK}_t$
This figure depicts the simulated (dashed line) and sample (solid line) autocorrelation function of $\log \mathrm{RK}_t$ for the REGARCH, REGARCH(5,5), REGARCH-MIDAS, REGARCH-HAR and the FloEGARCH. We use the estimated parameters for SPY reported in Table 1 in the paper and $K = 52$ ($K = 12$) for the weekly (monthly) REGARCH-MIDAS. See the paper for additional details on their computation.

F. Empirical results for individual stocks

This section presents in-sample and out-of-sample results for our parallel analysis on 20 individual stocks as referred to in the paper. Additional details are provided in the table text.

Table 2: Summary statistics for daily returns and realized kernel

This table reports summary statistics for the daily returns and the logarithm of the realized kernel. Daily returns and the realized kernel are in percentages. Robust skewness and kurtosis are from Kim and White (2004).

		Return							Log(RK)		
	No. of obs.	Mean	Std. Dev.	Skew	Robust Skew	Ex. Kurt.	Robust Ex. Kurt.	Median	Mean	Std. Dev.	Median
SP500	3,020	0.02	1.32	0.07	-0.08	11.67	1.03	0.08	-0.35	1.00	-0.50
AA	3,004	0.00	2.73	0.23	0.02	8.92	0.99	0.00	1.13	0.86	0.98
AIG	2,999	-0.00	4.55	1.42	0.01	54.58	1.17	-0.03	1.07	1.30	0.88
AXP	2,994	0.07	2.44	0.55	0.04	11.12	1.07	0.02	0.60	1.18	0.38
BA	2,996	0.07	1.89	0.23	0.01	4.07	0.84	0.07	0.60	0.82	0.47
CAT	2,998	0.07	2.09	0.11	0.03	5.06	0.92	0.06	0.73	0.82	0.59
DD	2,995	0.04	1.78	-0.04	0.01	5.68	0.88	0.04	0.51	0.85	0.38
DIS	2,997	0.07	1.91	0.51	-0.02	6.76	0.88	0.06	0.55	0.88	0.38
GE	3,005	0.02	1.99	0.38	0.03	10.30	1.06	0.00	0.40	1.05	0.22
IBM	2,996	0.03	1.53	0.20	0.01	6.87	0.87	0.02	0.10	0.83	-0.05
INTC	3,016	0.03	2.20	-0.22	-0.00	6.09	0.90	0.04	0.85	0.80	0.74
JNJ	2,997	0.03	1.16	-0.28	0.03	20.06	0.95	0.02	-0.28	0.86	-0.43
KO	2,999	0.04	1.24	0.32	-0.02	11.96	0.92	0.04	-0.10	0.81	-0.22
MMM	2,992	0.05	1.45	-0.06	0.02	5.54	0.97	0.06	0.13	0.80	0.03
MRK	2,994	0.03	1.80	-1.21	0.04	24.18	0.87	0.03	0.38	0.85	0.26
MSFT	3,016	0.03	1.81	0.37	0.02	8.94	0.96	0.00	0.46	0.81	0.32
PG	2,998	0.04	1.14	-0.02	0.01	6.74	0.92	0.03	-0.18	0.77	-0.30
VZ	2,995	0.03	1.57	0.34	-0.03	7.37	0.90	0.05	0.31	0.89	0.14
WHR	2,992	0.07	2.52	0.40	0.03	5.14	0.96	-0.01	1.01	0.86	0.92
WMT	3,001	0.03	1.31	0.30	-0.03	5.57	0.88	0.02	0.05	0.80	-0.09
XOM	3,001	0.05	1.60	0.34	-0.01	13.37	0.81	0.07	0.24	0.86	0.14

Table 3: REGARCH

This table reports full-sample estimated parameters, information criteria as well as full maximized log-likelihood value for the original REGARCH.

	AA	AIG	AXP	BA	CAT	DD	DIS	GE	IBM	INTC
μ	0.015	-0.017	0.051	0.074	0.074	0.041	0.053	0.021	0.029	0.026
β	0.972	0.972	0.987	0.977	0.973	0.972	0.981	0.982	0.974	0.972
α	0.355	0.606	0.394	0.322	0.376	0.413	0.356	0.418	0.438	0.478
ξ	-0.518	-0.296	-0.385	-0.446	-0.590	-0.207	-0.327	-0.342	-0.375	-0.285
σ_u^2	0.136	0.201	0.148	0.135	0.130	0.147	0.146	0.153	0.129	0.129
τ_1	-0.054	-0.086	-0.085	-0.062	-0.056	-0.076	-0.076	-0.063	-0.072	-0.051
τ_2	0.039	0.039	0.039	0.036	0.016	0.023	0.022	0.029	0.014	0.021
δ_1	-0.061	-0.049	-0.065	-0.054	-0.069	-0.071	-0.079	-0.045	-0.063	-0.038
δ_2	0.063	0.042	0.060	0.076	0.043	0.048	0.047	0.049	0.037	0.036
ϕ	1.055	0.855	0.993	1.046	1.105	0.961	0.981	0.985	0.962	0.924
ω	1.544	1.513	1.135	1.041	1.155	0.770	1.067	0.865	0.549	1.353
$\mathscr{L}$	-7,883.37	-8,493.81	-7,122.76	-7,000.25	-7,237.72	-6,751.76	-6,965.93	-6,832.49	-6,176.02	-7,343.71
AIC	15,788.74	17,009.63	14,267.52	14,022.51	14,497.45	13,525.52	13,953.85	13,686.98	12,374.05	14,709.42
BIC	15,854.83	17,075.70	14,333.57	14,088.57	14,563.51	13,591.57	14,019.91	13,753.08	12,440.11	14,775.55
HQIC	15,942.92	17,163.77	14,421.62	14,176.62	14,651.58	13,679.63	14,107.98	13,841.17	12,528.17	14,863.68

	JNJ	KO	MMM	MRK	MSFT	PG	VZ	WHR	WMT	XOM
μ	0.030	0.029	0.044	0.023	0.035	0.026	0.028	0.074	0.017	0.042
β	0.976	0.973	0.967	0.971	0.968	0.961	0.979	0.963	0.979	0.967
α	0.359	0.399	0.357	0.257	0.447	0.373	0.334	0.287	0.307	0.357
ξ	-0.123	-0.150	-0.342	-1.017	-0.381	-0.159	-0.185	-1.112	-0.259	-0.291
σ_u^2	0.151	0.144	0.147	0.191	0.137	0.151	0.153	0.171	0.137	0.123
τ_1	-0.066	-0.065	-0.080	-0.034	-0.044	-0.059	-0.064	-0.040	-0.035	-0.087
τ_2	0.038	0.027	0.010	0.001	0.013	0.025	0.039	0.024	0.028	0.041
δ_1	-0.031	-0.053	-0.071	-0.057	-0.036	-0.053	-0.061	-0.043	-0.028	-0.105
δ_2	0.058	0.064	0.041	0.010	0.033	0.056	0.059	0.068	0.058	0.050
ϕ	0.961	0.949	1.067	1.450	1.003	1.084	1.009	1.344	1.120	1.082
ω	-0.071	0.112	0.437	0.974	0.900	0.021	0.530	1.538	0.350	0.554
$\mathscr{L}$	-5,426.50	-5,694.49	-6,291.30	-7,470.15	-6,832.84	-5,651.66	-6,419.04	-8,210.50	-5,951.18	-6,113.27
AIC	10,875.01	11,410.98	12,604.59	14,962.29	13,687.67	11,325.33	12,860.09	16,443.00	11,924.36	12,248.54
BIC	10,941.07	11,477.05	12,670.64	15,028.34	13,753.80	11,391.39	12,926.14	16,509.04	11,990.44	12,314.61
HQIC	11,029.13	11,565.12	12,758.68	15,116.40	13,841.94	11,479.46	13,014.20	16,597.08	12,078.52	12,402.69

Table 4: Weekly REGARCH-MIDAS

This table reports full-sample estimated parameters, information criteria, variance ratio as well as full maximized log-likelihood value for the weekly two-parameter REGARCH-MIDAS. Results are for $K = 52$.

	AA	AIG	AXP	BA	CAT	DD	DIS	GE	IBM	INTC
μ	0.011	-0.016	0.056	0.078	0.071	0.044	0.063	0.025	0.033	0.031
β	0.649	0.577	0.710	0.606	0.553	0.623	0.520	0.817	0.612	0.602
α	0.392	0.589	0.410	0.360	0.432	0.455	0.429	0.452	0.479	0.531
ξ	-0.487	-0.291	-0.392	-0.441	-0.572	-0.202	-0.322	-0.333	-0.378	-0.276
σ_u^2	0.133	0.191	0.144	0.132	0.126	0.144	0.142	0.149	0.127	0.125
τ_1	-0.062	-0.094	-0.096	-0.076	-0.062	-0.090	-0.089	-0.068	-0.086	-0.061
τ_2	0.044	0.050	0.048	0.047	0.022	0.025	0.029	0.035	0.015	0.018
δ_1	-0.063	-0.048	-0.065	-0.054	-0.069	-0.071	-0.080	-0.046	-0.064	-0.040
δ_2	0.061	0.047	0.059	0.075	0.042	0.047	0.045	0.050	0.037	0.032
ϕ	1.039	0.872	0.999	1.044	1.093	0.959	0.976	0.978	0.963	0.920
ω	0.538	0.361	0.402	0.448	0.547	0.224	0.340	0.362	0.392	0.331
λ	0.900	1.117	0.973	0.908	0.873	0.999	0.993	0.945	0.981	1.038
γ_1	-0.156	-0.577	-0.041	-0.214	-0.822	-0.758	-0.866	2.003	-0.971	-0.296
γ_2	6.538	1.481	6.938	6.994	2.134	1.000	1.130	27.928	1.004	4.883
$\mathscr{L}$	-7,836.88	-8,367.58	-7,081.18	-6,956.80	-7,187.14	-6,716.86	-6,912.97	-6,791.63	-6,144.97	-7,280.34
AIC	15,701.75	16,763.15	14,190.36	13,941.60	14,402.28	13,461.72	13,853.93	13,611.26	12,317.94	14,588.67
BIC	15,785.86	16,847.24	14,274.42	14,025.67	14,486.36	13,545.79	13,938.01	13,695.38	12,402.01	14,672.84
HQIC	15,897.98	16,959.33	14,386.49	14,137.75	14,598.44	13,657.86	14,050.09	13,807.49	12,514.09	14,785.01
VR	0.82	0.87	0.89	0.85	0.85	0.82	0.89	0.80	0.83	0.84

	JNJ	KO	MMM	MRK	MSFT	PG	VZ	WHR	WMT	XOM
μ	0.034	0.032	0.045	0.025	0.038	0.028	0.031	0.067	0.021	0.051
β	0.627	0.562	0.587	0.552	0.613	0.547	0.649	0.554	0.536	0.598
α	0.411	0.453	0.415	0.297	0.492	0.416	0.384	0.315	0.342	0.382
ξ	-0.128	-0.149	-0.340	-1.002	-0.365	-0.158	-0.177	-1.055	-0.254	-0.288
σ_u^2	0.147	0.141	0.143	0.188	0.133	0.147	0.150	0.164	0.133	0.120
τ_1	-0.076	-0.071	-0.086	-0.039	-0.048	-0.068	-0.071	-0.048	-0.045	-0.102
τ_2	0.049	0.036	0.018	0.002	0.015	0.031	0.048	0.032	0.036	0.046
δ_1	-0.029	-0.053	-0.070	-0.057	-0.038	-0.053	-0.060	-0.045	-0.027	-0.105
δ_2	0.058	0.065	0.041	0.010	0.033	0.055	0.058	0.067	0.056	0.049
ϕ	0.958	0.936	1.060	1.437	0.986	1.071	1.000	1.315	1.117	1.084
ω	0.103	0.149	0.324	0.712	0.391	0.132	0.192	0.861	0.225	0.279
λ	0.949	1.017	0.873	0.647	0.951	0.863	0.938	0.701	0.855	0.851
γ_1	1.471	-0.875	-0.638	-0.563	-0.477	-0.957	0.001	-0.034	-0.776	-1.116
γ_2	48.950	1.000	3.334	4.365	3.124	1.000	11.393	7.918	2.321	1.000
$\mathscr{L}$	-5,395.67	-5,652.84	-6,256.25	-7,438.83	-6,785.22	-5,604.99	-6,390.01	-8,138.73	-5,910.60	-6,069.31
AIC	10,819.34	11,333.68	12,540.50	14,905.66	13,598.45	11,237.98	12,808.02	16,305.46	11,849.19	12,166.62
BIC	10,903.42	11,417.77	12,624.56	14,989.73	13,682.61	11,322.06	12,892.09	16,389.52	11,933.29	12,250.72
HQIC	11,015.50	11,529.86	12,736.61	15,101.79	13,794.78	11,434.14	13,004.16	16,501.58	12,045.39	12,362.82
VR	0.83	0.84	0.80	0.83	0.82	0.80	0.84	0.81	0.87	0.80

Table 5: Weekly REGARCH-MIDAS (single-parameter)

This table reports full-sample estimated parameters, information criteria, variance ratio as well as full maximized log-likelihood value for the weekly single-parameter REGARCH-MIDAS. Results are for $K = 52$.

	AA	AIG	AXP	BA	CAT	DD	DIS	GE	IBM	INTC
μ	0.011	-0.017	0.057	0.078	0.071	0.044	0.064	0.025	0.033	0.032
β	0.711	0.623	0.790	0.656	0.584	0.710	0.554	0.742	0.638	0.654
α	0.390	0.592	0.408	0.359	0.433	0.451	0.431	0.459	0.482	0.532
ξ	-0.491	-0.301	-0.392	-0.442	-0.573	-0.204	-0.321	-0.333	-0.378	-0.276
σ_u^2	0.133	0.191	0.144	0.132	0.126	0.144	0.142	0.149	0.127	0.125
τ_1	-0.062	-0.095	-0.095	-0.076	-0.061	-0.089	-0.089	-0.068	-0.085	-0.060
τ_2	0.043	0.049	0.046	0.046	0.022	0.025	0.029	0.036	0.016	0.018
δ_1	-0.063	-0.048	-0.065	-0.054	-0.069	-0.071	-0.079	-0.046	-0.063	-0.040
δ_2	0.061	0.047	0.059	0.075	0.043	0.048	0.045	0.049	0.037	0.032
ϕ	1.041	0.876	1.000	1.045	1.093	0.959	0.975	0.978	0.962	0.920
ω	0.563	0.406	0.409	0.459	0.571	0.256	0.358	0.356	0.398	0.353
λ	0.878	1.078	0.959	0.889	0.844	0.945	0.964	0.964	0.942	1.014
γ_2	22.519	26.382	17.205	27.678	40.966	27.351	39.739	19.960	44.999	25.944
$\mathscr{L}$	-7,838.40	-8,373.86	-7,081.88	-6,957.85	-7,190.46	-6,720.73	-6,915.24	-6,791.92	-6,148.72	-7,283.34
AIC	15,702.80	16,773.72	14,189.76	13,941.71	14,406.93	13,467.46	13,856.48	13,609.84	12,323.44	14,592.68
BIC	15,780.90	16,851.80	14,267.82	14,019.78	14,485.01	13,545.53	13,934.55	13,687.95	12,401.51	14,670.83
HQIC	15,885.01	16,955.88	14,371.88	14,123.85	14,589.08	13,649.59	14,038.63	13,792.06	12,505.58	14,774.99
VR	0.80	0.86	0.87	0.83	0.84	0.78	0.88	0.84	0.82	0.82

	JNJ	KO	MMM	MRK	MSFT	PG	VZ	WHR	WMT	XOM
μ	0.034	0.032	0.045	0.024	0.038	0.028	0.031	0.068	0.021	0.050
β	0.620	0.617	0.634	0.657	0.689	0.567	0.691	0.631	0.582	0.610
α	0.411	0.453	0.414	0.293	0.492	0.420	0.381	0.315	0.343	0.388
ξ	-0.129	-0.150	-0.343	-1.012	-0.364	-0.157	-0.177	-1.052	-0.256	-0.287
σ_u^2	0.147	0.141	0.144	0.188	0.134	0.148	0.150	0.164	0.134	0.120
τ_1	-0.076	-0.071	-0.086	-0.039	-0.048	-0.068	-0.071	-0.048	-0.045	-0.103
τ_2	0.049	0.036	0.017	0.002	0.015	0.031	0.048	0.032	0.036	0.046
δ_1	-0.029	-0.053	-0.070	-0.057	-0.038	-0.054	-0.060	-0.044	-0.027	-0.105
δ_2	0.058	0.064	0.041	0.010	0.033	0.055	0.058	0.067	0.056	0.048
ϕ	0.958	0.939	1.064	1.446	0.986	1.073	1.001	1.313	1.119	1.082
ω	0.104	0.148	0.332	0.723	0.407	0.125	0.197	0.871	0.229	0.288
λ	0.953	0.973	0.838	0.625	0.920	0.820	0.923	0.691	0.828	0.820
γ_2	37.685	35.949	35.219	28.038	24.932	46.647	28.890	20.578	39.253	53.373
$\mathscr{L}$	-5,395.73	-5,657.00	-6,258.62	-7,442.47	-6,789.21	-5,609.35	-6,390.15	-8,141.40	-5,915.05	-6,072.96
AIC	10,817.46	11,340.01	12,543.24	14,910.94	13,604.43	11,244.70	12,806.31	16,308.79	11,856.10	12,171.92
BIC	10,895.53	11,418.09	12,621.29	14,989.00	13,682.59	11,322.78	12,884.37	16,386.85	11,934.20	12,250.01
HQIC	10,999.61	11,522.17	12,725.34	15,093.06	13,786.74	11,426.86	12,988.44	16,490.90	12,038.29	12,354.10
VR	0.83	0.83	0.78	0.79	0.79	0.78	0.83	0.79	0.86	0.79

Table 6: Monthly REGARCH-MIDAS

This table reports full-sample estimated parameters, information criteria, variance ratio as well as full maximized log-likelihood value for the monthly two-parameter REGARCH-MIDAS. Results are for $K = 12$.

	AA	AIG	AXP	BA	CAT	DD	DIS	GE	IBM	INTC
μ	0.011	-0.018	0.057	0.078	0.071	0.044	0.064	0.025	0.034	0.032
β	0.820	0.742	0.861	0.837	0.825	0.864	0.832	0.843	0.891	0.812
α	0.383	0.590	0.404	0.345	0.414	0.432	0.390	0.445	0.451	0.515
ξ	-0.492	-0.306	-0.391	-0.441	-0.567	-0.201	-0.335	-0.333	-0.379	-0.276
σ_u^2	0.133	0.193	0.144	0.132	0.127	0.144	0.143	0.149	0.127	0.126
τ_1	-0.062	-0.092	-0.094	-0.073	-0.063	-0.086	-0.087	-0.069	-0.078	-0.060
τ_2	0.041	0.048	0.044	0.042	0.020	0.023	0.026	0.034	0.014	0.018
δ_1	-0.063	-0.048	-0.065	-0.054	-0.070	-0.071	-0.080	-0.046	-0.063	-0.040
δ_2	0.061	0.049	0.059	0.075	0.044	0.048	0.044	0.050	0.036	0.033
ϕ	1.042	0.881	0.999	1.044	1.089	0.957	0.987	0.978	0.966	0.920
ω	0.577	0.405	0.404	0.467	0.565	0.241	0.361	0.355	0.390	0.349
λ	0.864	1.080	0.959	0.869	0.842	0.951	0.951	0.952	0.897	1.009
γ_1	-0.583	-1.605	0.025	1.398	-0.865	-0.266	-0.878	0.424	0.800	1.028
γ_2	4.112	1.000	5.262	13.651	3.036	2.728	1.933	8.113	6.714	9.921
$\mathscr{L}$	-7,842.06	-8,377.55	-7,081.13	-6,963.39	-7,202.89	-6,721.18	-6,928.36	-6,792.40	-6,153.55	-7,292.92
AIC	15,712.13	16,783.10	14,190.26	13,954.77	14,433.77	13,470.37	13,884.72	13,612.79	12,335.11	14,613.85
BIC	15,796.24	16,867.19	14,274.33	14,038.85	14,517.86	13,554.44	13,968.80	13,696.91	12,419.18	14,698.01
HQIC	15,908.35	16,979.28	14,386.39	14,150.92	14,629.94	13,666.51	14,080.88	13,809.03	12,531.26	14,810.18
VR	0.72	0.82	0.83	0.73	0.70	0.65	0.76	0.78	0.60	0.74

	JNJ	KO	MMM	MRK	MSFT	PG	VZ	WHR	WMT	XOM
μ	0.034	0.032	0.045	0.024	0.038	0.028	0.031	0.067	0.021	0.049
β	0.850	0.827	0.863	0.864	0.826	0.831	0.875	0.745	0.852	0.873
α	0.388	0.434	0.380	0.277	0.476	0.400	0.359	0.314	0.325	0.361
ξ	-0.127	-0.149	-0.344	-1.006	-0.367	-0.159	-0.180	-1.032	-0.256	-0.294
σ_u^2	0.148	0.142	0.144	0.188	0.134	0.149	0.151	0.165	0.134	0.121
τ_1	-0.074	-0.069	-0.084	-0.038	-0.050	-0.065	-0.069	-0.048	-0.040	-0.095
τ_2	0.045	0.033	0.013	0.001	0.014	0.028	0.044	0.031	0.032	0.044
δ_1	-0.030	-0.053	-0.070	-0.057	-0.039	-0.053	-0.060	-0.045	-0.028	-0.105
δ_2	0.058	0.064	0.041	0.010	0.033	0.056	0.058	0.067	0.057	0.049
ϕ	0.958	0.941	1.071	1.439	0.988	1.080	1.005	1.301	1.120	1.089
ω	0.082	0.141	0.330	0.730	0.407	0.118	0.205	0.874	0.225	0.297
λ	0.914	0.962	0.783	0.595	0.906	0.794	0.877	0.686	0.807	0.761
γ_1	-1.427	-1.388	2.524	3.325	0.013	-0.931	1.769	3.777	1.456	-0.651
γ_2	1.000	1.000	18.021	21.725	5.094	1.783	13.707	33.579	13.546	2.319
$\mathscr{L}$	-5,404.34	-5,662.62	-6,265.13	-7,446.52	-6,791.44	-5,623.29	-6,395.34	-8,141.29	-5,921.83	-6,087.31
AIC	10,836.68	11,353.24	12,558.27	14,921.04	13,610.88	11,274.58	12,818.68	16,310.58	11,871.65	12,202.61
BIC	10,920.76	11,437.32	12,642.32	15,005.10	13,695.05	11,358.66	12,902.75	16,394.64	11,955.75	12,286.71
HQIC	11,032.84	11,549.41	12,754.38	15,117.17	13,807.22	11,470.75	13,014.82	16,506.70	12,067.85	12,398.81
VR	0.67	0.70	0.58	0.63	0.69	0.59	0.68	0.74	0.72	0.54

Table 7: Monthly REGARCH-MIDAS (single-parameter)

This table reports full-sample estimated parameters, information criteria, variance ratio as well as full maximized log-likelihood value for the monthly single-parameter REGARCH-MIDAS. Results are for $K = 12$.

	AA	AIG	AXP	BA	CAT	DD	DIS	GE	IBM	INTC
μ	0.012	-0.017	0.057	0.078	0.071	0.045	0.065	0.025	0.034	0.032
β	0.824	0.737	0.865	0.836	0.831	0.871	0.841	0.845	0.893	0.812
α	0.383	0.591	0.404	0.345	0.413	0.431	0.387	0.445	0.450	0.515
ξ	-0.492	-0.311	-0.391	-0.441	-0.568	-0.201	-0.337	-0.333	-0.378	-0.276
σ_u^2	0.134	0.192	0.144	0.132	0.127	0.144	0.143	0.149	0.127	0.126
τ_1	-0.062	-0.093	-0.094	-0.073	-0.063	-0.085	-0.087	-0.069	-0.078	-0.060
τ_2	0.041	0.048	0.044	0.042	0.019	0.023	0.026	0.034	0.014	0.018
δ_1	-0.063	-0.047	-0.065	-0.054	-0.070	-0.071	-0.080	-0.046	-0.063	-0.040
δ_2	0.061	0.048	0.059	0.075	0.044	0.048	0.044	0.050	0.036	0.033
ϕ	1.042	0.883	0.999	1.045	1.090	0.957	0.988	0.978	0.966	0.920
ω	0.587	0.431	0.406	0.466	0.576	0.252	0.373	0.356	0.390	0.349
λ	0.855	1.054	0.955	0.871	0.828	0.930	0.933	0.950	0.894	1.009
γ_2	12.772	21.575	9.843	11.378	13.169	8.224	10.921	11.252	7.447	9.787
$\mathscr{L}$	-7,842.51	-8,381.29	-7,081.41	-6,963.39	-7,203.34	-6,721.53	-6,929.11	-6,792.42	-6,153.56	-7,292.92
AIC	15,711.02	16,788.57	14,188.81	13,952.79	14,432.67	13,469.06	13,884.22	13,610.84	12,333.13	14,611.85
BIC	15,789.13	16,866.65	14,266.87	14,030.86	14,510.75	13,547.13	13,962.29	13,688.95	12,411.20	14,690.00
HQIC	15,893.23	16,970.74	14,370.94	14,134.93	14,614.83	13,651.19	14,066.36	13,793.06	12,515.27	14,794.16
VR	0.72	0.82	0.83	0.73	0.70	0.64	0.76	0.78	0.59	0.74

	JNJ	KO	MMM	MRK	MSFT	PG	VZ	WHR	WMT	XOM
μ	0.034	0.033	0.045	0.024	0.038	0.028	0.031	0.067	0.021	0.049
β	0.856	0.829	0.856	0.849	0.834	0.837	0.871	0.742	0.850	0.880
α	0.387	0.434	0.381	0.277	0.474	0.399	0.360	0.314	0.326	0.361
ξ	-0.127	-0.149	-0.345	-1.016	-0.367	-0.159	-0.180	-1.037	-0.256	-0.293
σ_u^2	0.148	0.142	0.144	0.188	0.134	0.149	0.151	0.165	0.134	0.121
τ_1	-0.073	-0.069	-0.085	-0.038	-0.050	-0.065	-0.069	-0.048	-0.041	-0.095
τ_2	0.045	0.033	0.013	0.001	0.014	0.028	0.044	0.030	0.032	0.044
δ_1	-0.030	-0.052	-0.070	-0.058	-0.039	-0.053	-0.060	-0.044	-0.028	-0.105
δ_2	0.058	0.064	0.041	0.010	0.033	0.056	0.058	0.067	0.057	0.049
ϕ	0.958	0.940	1.072	1.448	0.988	1.080	1.005	1.304	1.120	1.089
ω	0.077	0.140	0.329	0.728	0.409	0.116	0.203	0.874	0.224	0.304
λ	0.889	0.938	0.792	0.602	0.901	0.772	0.883	0.686	0.809	0.739
γ_2	14.140	14.689	10.161	10.076	9.315	11.923	9.798	15.579	11.051	10.146
$\mathscr{L}$	-5,405.05	-5,663.48	-6,265.40	-7,447.15	-6,791.78	-5,623.97	-6,395.39	-8,141.55	-5,921.85	-6,087.94
AIC	10,836.10	11,352.96	12,556.81	14,920.31	13,609.55	11,273.94	12,816.77	16,309.10	11,869.70	12,201.87
BIC	10,914.18	11,431.04	12,634.86	14,998.37	13,687.71	11,352.02	12,894.84	16,387.16	11,947.79	12,279.96
HQIC	11,018.25	11,535.13	12,738.91	15,102.43	13,791.87	11,456.10	12,998.90	16,491.21	12,051.89	12,384.05
VR	0.65	0.70	0.59	0.65	0.69	0.58	0.68	0.74	0.72	0.52

Table 8: REGARCH-HAR

This table reports full-sample estimated parameters, information criteria, variance ratio as well as full maximized log-likelihood value for the REGARCH-HAR.

	AA	AIG	AXP	BA	CAT	DD	DIS	GE	IBM	INTC
μ	0.011	-0.017	0.056	0.077	0.071	0.043	0.063	0.024	0.032	0.030
β	0.638	0.578	0.811	0.583	0.537	0.591	0.508	0.801	0.568	0.565
α	0.396	0.585	0.420	0.367	0.434	0.465	0.438	0.459	0.486	0.538
ξ	-0.495	-0.310	-0.391	-0.445	-0.576	-0.203	-0.322	-0.337	-0.377	-0.280
σ_u^2	0.134	0.191	0.145	0.133	0.127	0.145	0.143	0.150	0.127	0.126
τ_1	-0.062	-0.094	-0.095	-0.076	-0.061	-0.089	-0.089	-0.069	-0.086	-0.060
τ_2	0.045	0.050	0.046	0.048	0.022	0.026	0.029	0.036	0.016	0.019
δ_1	-0.063	-0.047	-0.065	-0.055	-0.069	-0.070	-0.080	-0.046	-0.064	-0.040
δ_2	0.061	0.048	0.060	0.075	0.043	0.048	0.044	0.051	0.037	0.033
ϕ	1.043	0.881	0.998	1.047	1.096	0.958	0.976	0.980	0.959	0.922
ω	0.572	0.418	0.429	0.465	0.571	0.256	0.360	0.377	0.398	0.366
γ_1	0.321	0.391	0.033	0.373	0.465	0.497	0.506	0.000	0.597	0.453
γ_2	0.552	0.677	0.898	0.510	0.381	0.453	0.457	0.920	0.356	0.551
$\mathscr{L}$	-7,847.46	-8,374.10	-7,095.31	-6,972.04	-7,196.44	-6,731.86	-6,927.53	-6,804.48	-6,156.54	-7,298.11
AIC	15,720.92	16,774.20	14,216.63	13,970.07	14,418.88	13,489.73	13,881.07	13,634.95	12,339.08	14,622.23
BIC	15,799.03	16,852.28	14,294.69	14,048.14	14,496.96	13,567.79	13,959.14	13,713.06	12,417.15	14,700.38
HQIC	15,903.13	16,956.37	14,398.75	14,152.21	14,601.04	13,671.86	14,063.21	13,817.17	12,521.22	14,804.54
VR	0.82	0.86	0.84	0.85	0.85	0.82	0.88	0.80	0.84	0.84

	JNJ	KO	MMM	MRK	MSFT	PG	VZ	WHR	WMT	XOM
μ	0.034	0.032	0.044	0.025	0.039	0.028	0.030	0.068	0.020	0.049
β	0.570	0.558	0.554	0.494	0.585	0.521	0.599	0.531	0.500	0.561
α	0.419	0.458	0.423	0.303	0.502	0.421	0.394	0.317	0.349	0.389
ξ	-0.130	-0.149	-0.339	-0.998	-0.362	-0.157	-0.180	-1.066	-0.254	-0.288
σ_u^2	0.148	0.141	0.145	0.189	0.134	0.148	0.152	0.165	0.135	0.120
τ_1	-0.076	-0.072	-0.085	-0.039	-0.046	-0.068	-0.072	-0.047	-0.045	-0.103
τ_2	0.050	0.037	0.018	0.002	0.016	0.031	0.049	0.032	0.037	0.046
δ_1	-0.029	-0.053	-0.069	-0.056	-0.038	-0.053	-0.061	-0.044	-0.028	-0.105
δ_2	0.059	0.065	0.042	0.010	0.033	0.055	0.059	0.066	0.056	0.049
ϕ	0.959	0.936	1.058	1.433	0.982	1.069	1.001	1.321	1.117	1.082
ω	0.106	0.148	0.331	0.719	0.408	0.126	0.200	0.881	0.228	0.287
γ_1	0.480	0.499	0.457	0.356	0.423	0.500	0.437	0.246	0.472	0.556
γ_2	0.474	0.480	0.388	0.275	0.501	0.329	0.485	0.435	0.361	0.273
$\mathscr{L}$	-5,407.30	-5,659.23	-6,269.06	-7,449.17	-6,796.98	-5,613.81	-6,404.37	-8,143.54	-5,922.70	-6,077.50
AIC	10,840.60	11,344.46	12,564.13	14,924.33	13,619.96	11,253.63	12,834.74	16,313.08	11,871.40	12,181.01
BIC	10,918.68	11,422.54	12,642.18	15,002.40	13,698.11	11,331.71	12,912.80	16,391.13	11,949.49	12,259.10
HQIC	11,022.75	11,526.62	12,746.23	15,106.46	13,802.27	11,435.78	13,016.87	16,495.18	12,053.58	12,363.19
VR	0.84	0.84	0.80	0.84	0.82	0.80	0.85	0.81	0.87	0.81

Table 9: REGARCH-Spline

This table reports full-sample estimated parameters, information criteria as well as full maximized log-likelihood value for the REGARCH-Spline. Results are for $K = 6$.

	AA	AIG	AXP	BA	CAT	DD	DIS	GE	IBM	INTC
μ	-0.017	0.048	0.059	0.088	0.075	0.047	0.068	0.023	0.035	0.017
β	0.933	0.864	0.943	0.936	0.941	0.932	0.942	0.935	0.948	0.913
α	0.359	0.588	0.389	0.342	0.376	0.417	0.355	0.428	0.440	0.487
ξ	-0.510	-0.274	-0.379	-0.399	-0.591	-0.199	-0.328	-0.326	-0.375	-0.275
σ_u^2	0.134	0.195	0.144	0.133	0.128	0.144	0.143	0.150	0.128	0.127
τ_1	-0.058	-0.075	-0.088	-0.069	-0.058	-0.079	-0.081	-0.065	-0.073	-0.055
τ_2	0.039	0.045	0.040	0.043	0.016	0.023	0.024	0.031	0.013	0.018
δ_1	-0.063	-0.039	-0.064	-0.055	-0.069	-0.069	-0.078	-0.045	-0.062	-0.037
δ_2	0.062	0.048	0.058	0.076	0.042	0.047	0.046	0.050	0.036	0.034
ϕ	1.052	0.864	0.994	1.000	1.108	0.953	0.980	0.973	0.959	0.919
ω	1.684	1.435	1.726	2.015	1.375	1.398	2.009	1.860	1.632	2.769
$\mathscr{L}$	-7,854.94	-8,390.56	-7,086.92	-6,972.56	-7,214.19	-6,718.88	-6,931.05	-6,795.12	-6,156.17	-7,300.44
AIC	15,745.89	16,817.12	14,209.84	13,981.11	14,464.37	13,473.76	13,898.10	13,626.24	12,348.33	14,636.88
BIC	15,854.03	16,925.23	14,317.92	14,089.21	14,572.48	13,581.85	14,006.20	13,734.39	12,456.43	14,745.10
HQIC	15,998.18	17,069.35	14,462.00	14,233.30	14,716.59	13,725.94	14,150.30	13,878.54	12,600.52	14,889.32
VR	0.51	0.79	0.75	0.62	0.50	0.56	0.63	0.68	0.40	0.61

	JNJ	KO	MMM	MRK	MSFT	PG	VZ	WHR	WMT	XOM
μ	0.035	0.033	0.047	0.039	0.050	0.021	0.033	0.067	0.020	0.049
β	0.950	0.932	0.938	0.926	0.917	0.930	0.953	0.905	0.947	0.947
α	0.353	0.404	0.353	0.283	0.449	0.371	0.343	0.307	0.310	0.343
ξ	-0.123	-0.147	-0.351	-0.902	-0.383	-0.143	-0.177	-0.988	-0.252	-0.295
σ_u^2	0.148	0.142	0.144	0.190	0.134	0.149	0.152	0.167	0.135	0.121
τ_1	-0.068	-0.067	-0.081	-0.031	-0.047	-0.059	-0.064	-0.045	-0.036	-0.087
τ_2	0.040	0.028	0.012	0.002	0.013	0.025	0.040	0.027	0.028	0.043
δ_1	-0.030	-0.052	-0.070	-0.052	-0.037	-0.053	-0.060	-0.044	-0.028	-0.104
δ_2	0.057	0.064	0.040	0.011	0.032	0.055	0.058	0.069	0.057	0.050
ϕ	0.964	0.941	1.077	1.349	1.002	1.087	1.004	1.273	1.113	1.094
ω	0.962	0.738	0.936	1.168	2.001	0.169	1.366	1.490	1.164	1.040
$\mathscr{L}$	-5,403.06	-5,663.08	-6,270.65	-7,434.19	-6,796.46	-5,629.11	-6,401.72	-8,162.50	-5,927.05	-6,090.50
AIC	10,842.13	11,362.17	12,577.29	14,904.38	13,628.92	11,294.23	12,839.43	16,361.01	11,890.10	12,217.00
BIC	10,950.23	11,470.28	12,685.37	15,012.46	13,737.14	11,402.34	12,947.52	16,469.08	11,998.23	12,325.13
HQIC	11,094.33	11,614.40	12,829.44	15,156.55	13,881.35	11,546.45	13,091.61	16,613.15	12,142.36	12,469.26
VR	0.46	0.59	0.40	0.63	0.57	0.42	0.44	0.61	0.53	0.35

Table 10: FloEGARCH

This table reports full-sample estimated parameters, information criteria as well as full maximized log-likelihood value for the FloEGARCH.

	AA	AIG	AXP	BA	CAT	DD	DIS	GE	IBM	INTC
μ	0.017	-0.009	0.045	0.061	0.072	0.035	0.042	0.021	0.025	0.010
β	0.195	0.121	0.195	0.133	0.104	0.179	0.080	0.117	0.161	0.131
α	0.400	0.589	0.418	0.373	0.436	0.460	0.424	0.473	0.484	0.536
ξ	-0.476	-0.293	-0.389	-0.440	-0.566	-0.205	-0.327	-0.332	-0.378	-0.272
σ_u^2	0.134	0.192	0.144	0.132	0.127	0.144	0.142	0.149	0.128	0.126
τ_1	-0.065	-0.096	-0.098	-0.078	-0.066	-0.089	-0.095	-0.072	-0.084	-0.063
τ_2	0.041	0.048	0.045	0.045	0.020	0.025	0.029	0.037	0.015	0.020
δ_1	-0.063	-0.047	-0.067	-0.057	-0.069	-0.071	-0.082	-0.047	-0.064	-0.042
δ_2	0.060	0.048	0.060	0.076	0.043	0.047	0.045	0.050	0.036	0.033
ϕ	1.036	0.872	0.999	1.041	1.092	0.964	0.985	0.979	0.975	0.923
ω	1.393	1.076	1.299	1.365	0.950	0.961	1.510	0.980	0.825	1.898
d	0.633	0.620	0.678	0.658	0.673	0.645	0.673	0.678	0.672	0.644
$\mathscr{L}$	-7,843.29	-8,370.57	-7,082.26	-6,960.83	-7,195.15	-6,722.47	-6,922.21	-6,791.17	-6,159.11	-7,291.01
AIC	15,710.58	16,765.13	14,188.52	13,945.66	14,414.29	13,468.95	13,868.42	13,606.34	12,342.22	14,606.02
BIC	15,782.68	16,837.21	14,260.57	14,017.72	14,486.36	13,541.01	13,940.48	13,678.44	12,414.28	14,678.16
HQIC	15,878.78	16,933.28	14,356.63	14,113.79	14,582.44	13,637.07	14,036.55	13,774.54	12,510.34	14,774.31

	JNJ	KO	MMM	MRK	MSFT	PG	VZ	WHR	WMT	XOM
μ	0.030	0.029	0.042	0.029	0.029	0.024	0.028	0.088	0.014	0.040
β	0.093	0.121	0.137	0.108	0.158	0.138	0.156	0.092	0.113	0.150
α	0.430	0.458	0.413	0.298	0.498	0.425	0.396	0.317	0.361	0.413
ξ	-0.130	-0.147	-0.343	-1.010	-0.368	-0.159	-0.178	-1.062	-0.258	-0.292
σ_u^2	0.147	0.141	0.144	0.188	0.134	0.148	0.151	0.165	0.134	0.122
τ_1	-0.078	-0.072	-0.088	-0.041	-0.051	-0.068	-0.071	-0.051	-0.045	-0.101
τ_2	0.049	0.035	0.015	0.002	0.015	0.031	0.049	0.030	0.035	0.049
δ_1	-0.030	-0.053	-0.070	-0.057	-0.039	-0.055	-0.060	-0.044	-0.029	-0.106
δ_2	0.058	0.065	0.041	0.010	0.033	0.056	0.059	0.065	0.057	0.050
ϕ	0.955	0.939	1.073	1.448	0.998	1.080	1.001	1.325	1.116	1.078
ω	0.300	0.367	0.466	0.827	1.307	0.307	0.726	1.171	0.643	0.832
d	0.692	0.674	0.655	0.666	0.641	0.643	0.671	0.618	0.681	0.656
$\mathscr{L}$	-5,399.76	-5,660.47	-6,259.27	-7,442.64	-6,796.38	-5,620.41	-6,398.68	-8,152.11	-5,922.81	-6,097.09
AIC	10,823.52	11,344.93	12,542.53	14,909.29	13,616.76	11,264.82	12,821.36	16,328.22	11,869.62	12,218.18
BIC	10,895.59	11,417.01	12,614.58	14,981.34	13,688.90	11,336.89	12,893.42	16,400.27	11,941.70	12,290.26
HQIC	10,991.66	11,513.09	12,710.63	15,077.40	13,785.05	11,432.97	12,989.49	16,496.32	12,037.79	12,386.35

Table 11: Difference in maximized log-likelihood relative to REGARCH

This table reports the differences in the maximized log-likelihood values for our proposed models and the REGARCH-Spline and FloEGARCH relative to the original REGARCH. Positive values indicate improvements in empirical fit. We report results for all 20 individual stocks and include SPY for comparative purposes. Gray shaded areas indicate the model with the highest likelihood gain relative to the REGARCH.

	REGARCH-MIDAS (weekly)	REGARCH-MIDAS (weekly) (single-parameter)	REGARCH-MIDAS (monthly)	REGARCH-MIDAS (monthly) (single-parameter)	REGARCH-HAR	REGARCH-S	FloEGARCH
SP500	46.0	45.5	46.0	45.3	28.5	34.1	38.9
AA	46.5	45.0	41.3	40.9	35.9	28.4	40.1
AIG	126.2	120.0	116.3	112.5	119.7	103.3	123.2
AXP	41.6	40.9	41.6	41.4	27.4	35.8	40.5
BA	43.5	42.4	36.9	36.9	28.2	27.7	39.4
CAT	50.6	47.3	34.8	34.4	41.3	23.5	42.6
DD	34.9	31.0	30.6	30.2	19.9	32.9	29.3
DIS	53.0	50.7	37.6	36.8	38.4	34.9	43.7
GE	40.9	40.6	40.1	40.1	28.0	37.4	41.3
IBM	31.1	27.3	22.5	22.5	19.5	19.9	16.9
INTC	63.4	60.4	50.8	50.8	45.6	43.3	52.7
JNJ	30.8	30.8	22.2	21.5	19.2	23.4	26.7
KO	41.6	37.5	31.9	31.0	35.3	31.4	34.0
MMM	35.0	32.7	26.2	25.9	22.2	20.7	32.0
MRK	31.3	27.7	23.6	23.0	21.0	36.0	27.5
MSFT	47.6	43.6	41.4	41.1	35.9	36.4	36.5
PG	46.7	42.3	28.4	27.7	37.8	22.5	31.3
VZ	29.0	28.9	23.7	23.7	14.7	17.3	20.4
WHR	71.8	69.1	69.2	68.9	67.0	48.0	58.4
WMT	40.6	36.1	29.4	29.3	28.5	24.1	28.4
XOM	44.0	40.3	26.0	25.3	35.8	22.8	16.2
SP500	46.0	45.5	46.0	45.3	28.5	34.1	38.9
Mean	47.4	44.8	39.1	38.5	35.7	33.5	39.0

Table 12: Estimated β across various REGARCH specifications

This table reports estimated β for our proposed models and the REGARCH-Spline and FloEGARCH relative to the original REGARCH. We report results for all 20 individual stocks and include SPY for comparative purposes.

	REGARCH	REGARCH-MIDAS (weekly)	REGARCH-MIDAS (weekly) (single-parameter)	REGARCH-MIDAS (monthly)	REGARCH-MIDAS (monthly) (single-parameter)	REGARCH-HAR	REGARCH-S	FloEGARCH
SP500	0.972	0.761	0.842	0.872	0.880	0.734	0.943	0.176
AA	0.972	0.649	0.711	0.820	0.824	0.638	0.933	0.195
AIG	0.972	0.577	0.623	0.742	0.737	0.578	0.864	0.121
AXP	0.987	0.710	0.790	0.861	0.865	0.811	0.943	0.195
BA	0.977	0.606	0.656	0.837	0.836	0.583	0.936	0.133
CAT	0.973	0.553	0.584	0.825	0.831	0.537	0.941	0.104
DD	0.972	0.623	0.710	0.864	0.871	0.591	0.932	0.179
DIS	0.981	0.520	0.554	0.832	0.841	0.508	0.942	0.080
GE	0.982	0.817	0.742	0.843	0.845	0.801	0.935	0.117
IBM	0.974	0.612	0.638	0.891	0.893	0.568	0.948	0.161
INTC	0.972	0.602	0.654	0.812	0.812	0.565	0.913	0.131
JNJ	0.976	0.627	0.620	0.850	0.856	0.570	0.950	0.093
KO	0.973	0.562	0.617	0.827	0.829	0.558	0.932	0.121
MMM	0.967	0.587	0.634	0.863	0.856	0.554	0.938	0.137
MRK	0.971	0.552	0.657	0.864	0.849	0.494	0.926	0.108
MSFT	0.968	0.613	0.689	0.826	0.834	0.585	0.917	0.158
PG	0.961	0.547	0.567	0.831	0.837	0.521	0.930	0.138
VZ	0.979	0.649	0.691	0.875	0.871	0.599	0.953	0.156
WHR	0.963	0.554	0.631	0.745	0.742	0.531	0.905	0.092
WMT	0.979	0.536	0.582	0.852	0.850	0.500	0.947	0.113
XOM	0.967	0.598	0.610	0.873	0.880	0.561	0.947	0.150
Mean	0.973	0.612	0.657	0.838	0.840	0.590	0.932	0.136

Table 13: Forecast evaluation for individual stocks

This table summarizes the k-steps ahead predictive ability of the REGARCH-HAR and weekly single-parameter REGARCH-MIDAS benchmarked against the original REGARCH. Statistical significance of the differences in forecast losses is assessed by means of the Diebold-Mariano test for equal predictive ability using a HAC estimator, following state-of-the art good practice by using the data-dependent bandwidth selection by Andrews (1991) based on an AR(1) approximation and a Bartlett kernel. For each forecast horizon and stock we categorize the outcome of the test according to the size of the p-value and report the number of stocks falling into each category in the table. For instance, the last row in the left Panel A indicates that for the 22-steps ahead forecast, the weekly REGARCH-MIDAS outperforms REGARCH 15/20 times at the 1% significance level and 1/20 times at the 10% significance level (but not at the 5% and 1% level), whereas 2/20 times the difference in forecast performance was insignificant and 2/20 times the REGARCH was performing the best. As another measure of improvement, we also report the median ratio of forecast loss of our proposed models relative to the original REGARCH. A number above unity indicates superior performance of our proposed model.

Panel A: QLIKE loss function

Weekly REGARCH-MIDAS							REGARCH-HAR						
Horizon	1%	5%	10%	Insign.	REGARCH	Median loss ratio	Horizon	1%	5%	10%	Insign.	REGARCH	Median loss ratio
k = 1	5	2	3	5	5	1.02	k = 1	4	1	1	5	9	1.01
k = 2	6	3	2	5	4	1.03	k = 2	3	2	1	7	7	1.02
k = 3	9	2	2	4	3	1.05	k = 3	4	4	1	4	7	1.02
k = 4	9	1	1	6	3	1.05	k = 4	7	0	3	2	8	1.02
k = 5	9	1	1	5	4	1.05	k = 5	7	1	2	4	6	1.03
k = 10	11	2	2	1	4	1.11	k = 10	10	1	2	2	5	1.09
k = 15	13	2	0	1	4	1.22	k = 15	12	1	1	2	4	1.15
k = 22	15	0	1	2	2	1.33	k = 22	13	1	0	4	2	1.26

Panel B: Squared Predition Error loss function

Weekly REGARCH-MIDAS							REGARCH-HAR						
Horizon	1%	5%	10%	Insign.	REGARCH	Median loss ratio	Horizon	1%	5%	10%	Insign.	REGARCH	Median loss ratio
k = 1	6	2	1	4	7	1.01	k = 1	3	4	1	1	11	1.00
k = 2	10	2	1	4	3	1.04	k = 2	4	5	2	1	8	1.02
k = 3	9	4	1	4	2	1.08	k = 3	6	5	0	2	7	1.03
k = 4	9	5	0	4	2	1.08	k = 4	7	3	0	3	7	1.04
k = 5	9	4	1	3	3	1.09	k = 5	8	2	1	2	7	1.04
k = 10	13	3	0	2	2	1.17	k = 10	12	1	0	0	7	1.11
k = 15	14	1	2	1	2	1.30	k = 15	12	1	0	1	6	1.19
k = 22	15	2	0	1	2	1.48	k = 22	11	1	1	1	6	1.32

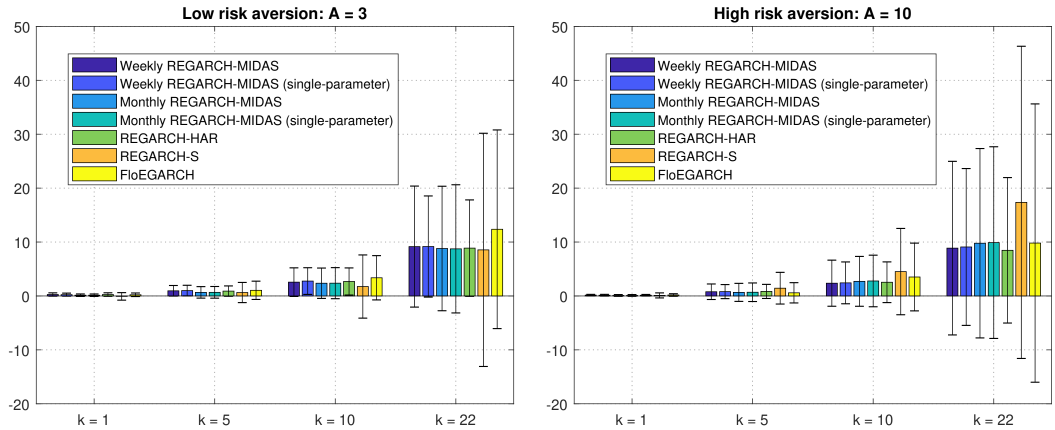

Figure 5: Economic value of volatility-timing strategy

This figure depicts the average performance fee in annualized basis points for the REGARCH-MIDAS, REGARCH-HAR, REGARCH-S and FLoEGARCH relative to the original REGARCH across all 20 individual stocks. The error bars represent +/− two cross-sectional standard errors. The investment horizon is set to 1, 5, 10, and 22 days as indicated by the x-axis. The left figure contains results for a low risk-aversion parameter, $A = 3$, and the right figure for a high risk-aversion parameter, $A = 10$.

References

ANDREWS, D. W. K. (1991): "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica*, 59, 817–858.

FULLER, W. A. (1996): *Introduction to statistical time series*, Wiley, second ed.

HANSEN, P. R. AND Z. HUANG (2016): "Exponential GARCH modeling with realized measures of volatility," *Journal of Business and Economic Statistics*, 34, 269–287.

HANSEN, P. R. AND A. LUNDE (2014): "Estimating the persistence and the autocorrelation function of a time series that is measured with error," *Econometric Theory*, 30, 60–93.

KIM, T.-H. AND H. WHITE (2004): "On more robust estimation of skewness and kurtosis," *Finance Research Letters*, 1, 56–73.

PAPARODITIS, E. AND D. N. POLITIS (2009): "Resampling and subsampling for financial time series," *In "Handbook of financial time series", Springer*, 983–999.

SHIMOTSU, K. (2010): "Exact local whittle estimation of fractional integration with unknown mean and time trend," *Econometric Theory*, 26, 501–540.

WENGER, K., C. LESCHINSKI, AND P. SIBBERTSEN (2017): "Long memory of volatility," Working paper.