WORKING PAPER 7/2023 (STATISTICS)

Bayesian Predictive Distributions of Oil Returns Using Mixed Data Sampling Volatility Models

Audrone Virbickaite, Hoang Nguyen and Minh-Ngoc Tran

ISSN 1403-0586

Örebro University School of Business
SE-701 82 Örebro, Sweden

Bayesian Predictive Distributions of Oil Returns Using Mixed Data Sampling Volatility Models

Audronė Virbickaitė [(a),*], Hoang Nguyen[(b)], Minh-Ngoc Tran[(c)]

[*] Corresponding author

[(a)] Department of Quantitative Methods, CUNEF Universidad, calle Pirineos 55, 28040 Madrid, Spain, audrone.virbickaite@cunef.edu

[(b)] School of Business - Örebro University, Sweden

[(c)] Discipline of Business Analytics, The University of Sydney Business School and ACEMS

April 14, 2023

Abstract

This study explores the benefits of incorporating fat-tailed innovations, asymmetric volatility response, and an extended information set into crude oil return modeling and forecasting. To this end, we utilize standard volatility models such as Generalized Autoregressive Conditional Heteroskedastic (GARCH), Generalized Autoregressive Score (GAS), and Stochastic Volatility (SV), along with Mixed Data Sampling (MIDAS) regressions, which enable us to incorporate the impacts of relevant financial/macroeconomic news into asset price movements. For inference and prediction, we employ an innovative Bayesian estimation approach called the density-tempered sequential Monte Carlo method. Our findings indicate that the inclusion of exogenous variables is beneficial for GARCH-type models while offering only a marginal improvement for GAS and SV-type models. Notably, GAS-family models exhibit superior performance in terms of in-sample fit, out-of-sample forecast accuracy, as well as Value-at-Risk and Expected Shortfall prediction.

Keywords: ES, GARCH, GAS, log marginal likelihood, MIDAS, SV, VaR.

JEL Classification: C22, C52, C58, G32

1 Introduction

With the financialization of the commodity markets, it is now straightforward to include commodities in one's investment portfolio. According to the Futures Trading Association, oil is the most traded commodity in the world by far, and apart from being an excellent hedging instrument, it is also a major input in industrial production and transportation. Oil price and its volatility are one of the key indicators used to assess the overall condition of the economy, and as a result, it has attracted the attention of numerous scholars who have conducted research on the subject. For example, Hamilton (1996) shows that oil price shocks explain a large fraction of the variation in the U.S. macroeconomic aggregates. Alquist et al. (2013) find that oil volatility is related to business cycle fluctuations. Pan et al. (2017) show that high oil price uncertainty can lead to a decrease in future investments, which in turn might have a negative impact on the real economy. Kun (2019) shows how changes in the oil price trigger the regime shifts in the macro variables. Lee and Chiou (2011) demonstrate that high oil price volatility has a negative impact on stock market returns. Therefore, modeling and multi-step forecasting of oil return volatilities and entire distributions are of great importance.

Models for volatilities of financial assets date back nearly 40 years since the seminal works of Engle (1982); Bollerslev (1986) and Taylor (1982), who laid the foundations for the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) and Stochastic Volatility (SV)-type models, respectively. The GARCH model considers the conditional variance as a function of an observation-driven process while the SV model regards the volatility as a latent process. These models are the benchmark specifications among practitioners, policymakers, and researchers alike. Most of the volatility models are based on daily or even lower-frequency data, since such data is easily manageable and readily available, it can be obtained with little to no financial cost (unlike high-frequency data). More recently, a new class of data-driven models has gained its place next to the well-established GARCH and SV-type models, namely, the Generalized Autoregressive Score model (GAS, Creal et al., 2011, 2013), also known as the Dynamic Conditional Score model (Harvey and Chakravarty, 2008; Harvey, 2013). GAS models, same as GARCH-type models, are observation-driven, however, they are much more flexible in accommodating abrupt jumps and irregularities in the returns (Charles and Darné, 2017) and dealing with outliers (Harvey and

Sucarrat, 2014).

To account for the different characteristics of the financial returns, the volatility models (be it GARCH, GAS or SV-type) have been generalized in numerous ways. Engle and Patton (2007) suggest that a good volatility model should account for the asymmetric impact of innovations on volatility, namely “leverage effect”, excessive kurtosis of the unconditional return distribution, and the effect of exogenous variables. Generally, asymmetric volatility response and excessive kurtosis can be incorporated based on the model assumptions of volatility process and innovations, while the use of exogenous variables implies the extension of the available information set. It is natural that market participants adjust their expectations and uncertainty whenever new information is available. In many cases, the exogenous predictors are some economic factors that are observed at a lower frequency (monthly, quarterly, yearly) than the crude oil prices (daily). In order to account for the frequency mismatch, Ghysels et al. (2004) have proposed a mixed data sampling method or MIDAS in short. In the MIDAS setting, the effects of the lagged macro-finance variables are regularized by a polynomial weighting scheme to ensure a parsimonious model specification. Baumeister and Kilian (2012), for example, show that the explanatory variables related to the global oil market provide better forecasting of the real price of oil than the forecasts based on models without exogenous predictors.

The aforementioned models and their variants have been extensively employed to model and forecast crude oil returns and volatilities; the most prominent and recent works are briefly reviewed in Section 2. Our goal is to consolidate the abundant strands of literature and provide a unified analysis of the usefulness of exogenous predictors, fat tails, and volatility with leverage effect using GARCH, SV, and GAS-type models for crude oil return distribution modeling and multi-step forecasting. To this end, we compare the Bayesian predictive distributions of crude oil returns produced by competing fat-tailed volatility models with leverage effect in the context of an extended information set. To our knowledge, there is no previous research that compares the three model classes for oil return distribution modeling and forecasting, even more so in the MIDAS setting. In contrast to the majority of the related articles, we focus not on some specific moments, such as mean and variance, but on the entire distribution, which presents a more comprehensive account of the behavior of the financial returns. Related to our study, Chan and Grant (2016) use Bayes factors (BFs) to compare a variety of GARCH and SV models for commodity returns, including oil, petroleum products, and natural gas. The study concluded that, in general, the SV models outperformed their GARCH counterparts. To continue the discussion, we include the GAS-type models and account for the macro-financial news to forecast oil volatility.

For inference, model selection, and prediction we employ a novel Bayesian estimation strategy based on the Density Tempered Sequential Monte Carlo (DTSMC) sampler (Tran et al., 2014; Duan and Fulop, 2015). The DTSMC sampler provides not only the posterior distribution of model parameters but also the log marginal likelihoods (LML) as a by-product, which allows for consistent model comparison. Additionally, the DTSMC sampler is especially well suited for non-linear non-Gaussian state space models, such as SV-type specifications. Finally, the Bayesian estimation approach is also important because, as demonstrated in Johannes et al. (2014), the estimation uncertainty plays a crucial role in the return prediction: "(...) the Bayesian approach punishes needlessly complicated models and is often referred to as a *fully automatic Occam's razor* (...)".

By using daily crude oil prices spanning more than 22 years, we find that GAS-type models are better than the GARCH and SV-type models in terms of in-sample fit and out-of-sample forecast accuracy. Moreover, including macro-financial variables is more beneficial for GARCH-type models as compared to the GAS and SV-type specifications, both in-sample and out-of-sample. In addition, the adverse effects of the model complexity play an important role in the worsening out-of-sample forecasting performance of the SV models. Finally, in order to quantify the economic gains, we calculate several risk measures and find that overall the results favor GAS-family models.

The article is organized as follows. Section 2 briefly presents the selected literature review, Sections 3 and 4 outline the models and estimation, respectively. An empirical application is presented in Section 5, and Section 6 concludes.

2 Brief literature review

The existing literature that deals with crude oil return distribution modeling mainly covers four major, often overlapping, branches: (i) models with flexible distributional assumptions for the returns; (ii) models with flexible volatility dynamics, such as leverage effect; (iii) observation vs parameter-driven models; (iv) models using exogenous predictors for volatility modeling and forecasting. Next, we present a brief literature review of the most relevant articles for oil volatility and return distribution modeling in the context of GARCH, SV, and GAS models.

The majority of the studies concerning crude oil volatility modeling with MIDAS regressions employ GARCH-type models. Conrad et al. (2014), for example, consider a GARCH-MIDAS model for oil price volatility. The authors find that the movements in long-run oil market volatility can be predicted by various U.S. macroeconomic activity variables, divided into two major categories: the current state of the economy and forward-looking indicators; and find that downturns in U.S. economic activity are associated with increases in the long-run crude oil return volatility. Wei et al. (2017) use GARCH-MIDAS models for modeling and forecasting spot crude oil volatility and find that the economic policy uncertainty (EPU)-related indices comprehensively integrate the information contained in other determinants (such as supply and demand factors). Pan et al. (2017) employ a two-regime GARCH-MIDAS model and find that while the levels of the supply and demand variables have a negative effect on the oil price volatility, their uncertainties have a positive effect. Ma et al. (2019) employ a GARCH-MIDAS model and find that models with financial predictors as explanatory variables have the highest predictive accuracy of oil price volatility, and fundamental information, based on global crude oil supply and demand, becomes more important in the long-run forecasting horizons. Wang and Li (2021) extend the univariate GARCH-MIDAS to multivariate DCC-MIDAS to study the volatility spillover relationship between the crude oil market and three major financial markets of China. Finally, Bonnier (2022) presents a comprehensive literature review concerning articles that employ GARCH and Realized Volatility models with exogenous variables for crude oil volatility modeling and forecasting. The author also includes an extensive study of looking for the best predictor variables by comparing multiple GARCH-family models (GJR, EGARCH, log-GARCH, log-GARCH-X) for crude oil volatility forecasts (1, 2, 21, 63 days) and finds that OVX (CBOE Crude Oil Volatility Index) is a preferred predictor, with several exchange rates staying in the final model as well (neither supply/demand factors nor their volatilities were significant).

To our knowledge, there is no research that considers crude oil volatility modeling and forecasting using the SV-MIDAS approach. Nonetheless, other variants of SV models have been widely employed in the context of crude oil. Baum and Zerilli (2016), for example, consider several SV models to analyze the volatility of crude oil futures and find that SV models with jumps are more effective than SV models without jumps. Shang and Liu (2017) are the first to extend the SV model to incorporate the MIDAS effects, and Shang and Zheng (2021) employ a leverage SV-MIDAS model for modeling and forecasting the volatility of the Chinese and U.S. stock markets. They find that the leverage SV-MIDAS model produces superior point volatility forecasts as compared to the benchmark specifications. In a recent study, Chen et al. (2019) use multiple SV models to model the returns of spot crude oil and demonstrate that the conventional SV models with Gaussian innovations offer the most accurate out-of-sample prediction results. Lu et al. (2020) employ a Markov Switching MIDAS specification, based on the Realized Volatility models, for forecasting China's oil futures volatility using the OVX index. SV-MIDAS type specification has been considered by Nguyen and Virbickaite (2022), where the authors model the co-dependence between stock-oil returns via latent SV-type process, augmented with MIDAS regression.

Finally, GAS models in the context of crude oil returns are also gaining popularity in recent years. Charles and Darné (2017) consider multiple GAS-type models for crude oil return volatility modeling and forecasting. The authors find that for out-of-sample forecasting, fat-tailed GARCH models with leverage effect are hard to beat. Xu and Lien (2022) employ the GAS model with student errors for multistep ahead oil volatility forecasting. The authors find that the GAS model overwhelmingly beats GARCH and EGARCH specifications for 1, 5, and 20-step ahead, but not 60-day ahead forecasts. To our knowledge, there are no research articles that consider GAS-MIDAS models for crude oil volatility modeling, nonetheless, GAS-MIDAS specifications have been employed in the multivariate setting. For example, Nguyen and Javed (2021) and Gong et al. (2022) independently propose to decompose the relation between financial returns into a short-term dependence, modeled via the GAS process, and a long-term dependence, modeled via MIDAS regression. Nguyen and Javed (2021) apply the proposed model on stock-bond returns, meanwhile Gong et al. (2022) investigate the co-dependence between multiple exchange rates.

3 Econometric framework

Next, we summarize the GARCH, GAS, and SV model classes and their variants that include the fat-tailed distributions, leverage effects, and MIDAS regressions. Denote the daily de-meaned log returns of crude oil as $r_t = \widetilde{r}_t - E[\widetilde{r}_t]$, where $\widetilde{r} = 100 \times \log(P_t/P_{t-1})$ is the daily log return for

$t = 1, \ldots, T$. In real data applications, the daily return expectation is replaced by its sample equivalent. Then the de-meaned log returns can be decomposed as $r_t = \sqrt{\sigma_t^2}\epsilon_t$, where σ_t^2 is the time-varying conditional variance and ϵ_t is the innovation term.

3.1 GARCH type models

Bollerslev (1986) proposed a GARCH model as a generalization of ARCH model by Engle (1982). The log returns are assumed to follow a normal distribution where the conditional variance σ_t^2 is a linear function of the past squared returns and lagged conditional variance. Engle and Rangel (2008) were the first to propose a multiplicative GARCH model such that the variance σ_t^2 is decomposed into two components, short and long-run. The multiplicative GARCH model was also considered by Engle et al. (2013), Asgharian et al. (2013), among many others. In such a model, the volatility has the short-term component g_t and a long-run component κ_τ, where τ is an indicator for monthly index, which is related to the asset frequency though $t = \tau L, \tau L{+}1, \ldots, (\tau{+}1)L$ where L is a number of trading days in a month. Conrad and Kleen (2020) suggest that the long-run component κ_τ can be affected by the exogenous variables sampled at lower frequencies via MIDAS regression. Then the general specification for the GARCH-MIDAS model, which can also accommodate fat-tailed innovations and leverage effect, is given by:

$$
\begin{aligned}
r_t &= \sqrt{\sigma_t^2}\epsilon_t = \sqrt{\kappa_\tau g_t}\epsilon_t, \quad \epsilon_t \sim \mathbf{F}, \\
g_t &= (1 - \alpha - 0.5\gamma - \beta) + (\alpha + \gamma \mathbf{I}_{\{\epsilon_{t-1}<0\}})\frac{r_{t-1}^2}{\kappa_\tau} + \beta g_{t-1}, \\
\kappa_\tau &= \exp\left(m + \sum_{j=1}^{N} \delta_j \left[\sum_{k=1}^{K_j} \phi_k(w_j) X_{j,\tau-k}\right]\right).
\end{aligned}
\tag{1}
$$

The short-term volatility component g_t follows a GJR GARCH process (Glosten et al., 1993) with the set of parameters (α, β, γ) such that $\alpha, \beta, \gamma > 0$ to ensure positive volatility and $0 < (\alpha + \beta + \gamma/2) < 1$ for a stationary condition. The short-term volatility fluctuates around some mean and depends on its past value with a persistence parameter β, the leverage effect of negative shock is captured through the parameter γ. In each period t, short-term volatility is updated through the past innovation ϵ_{t-1} that follows some distribution $\mathbf{F}$ with zero mean and unit variance. Secondly, a set of low-frequency variables X_j for $j = 1, \ldots, N$ helps to refine the long-term volatility component κ_τ in the spirit of MIDAS regression and filtering, see Engle et al. (2013). The effect of the j^{th} variable can be described through the size of a regression coefficient δ_j. The weighting function $\phi_k(w_j)$ creates a weighting scheme and regularizes for the effect of the last K_j observations of the explanatory variable X_j. We use the restricted beta function to smooth the effect of the explanatory variables,

$$\phi_k(w_j) = \frac{[1 - k/(K_j + 1)]^{w_j - 1}}{\sum_{l=1}^{K_j} [1 - l/(K_j + 1)]^{w_j - 1}}.$$

The model above nests several interesting specifications, considered in this article. When $\mathbf{F} \equiv \mathbf{N}(0, 1)$ and $\gamma = 0$, we have a standard GARCH-MIDAS model with Normal errors (Engle et al., 2013). When $\mathbf{F} \equiv \mathbf{Student}_\nu(0, 1)$ (Student-t with zero mean and unit variance) and $\gamma = 0$, we have a GARCH-MIDAS Student model. If the innovation distribution is a Student-t and $\gamma \neq 0$, we have a leverage GARCH-MIDAS specification (Conrad and Kleen, 2020), where a negative shock has a higher impact on the volatility than a positive shock. Finally, when $\delta_j = 0\,\forall j$ the GARCH MIDAS becomes the GJR GARCH model (Glosten et al., 1993).

3.2 GAS type models

The GAS models belong to a class of observation-driven models (the same as the GARCH-type). However, they are more efficient in accommodating abrupt jumps and irregularities in the returns (Charles and Darné, 2017) and dealing with outliers (Harvey and Sucarrat, 2014). In the GAS-type models, the conditional variance is updated based on the conditional score function of the last observation. This update mechanism is optimal in the sense that it leads to the minimum Kullback-Leibler divergence between the true conditional density and the model-implied conditional density (Blasques et al., 2015). We describe a variant of a GAS-MIDAS model, namely the one with an exponential link function discussed in Harvey and Sucarrat (2014). Such specification ensures positive variance and stationarity conditions are straightforward.

$$
\begin{aligned}
r_t &= \sqrt{\sigma_t^2}\epsilon_t = \sqrt{\kappa_\tau \exp(h_t)}\epsilon_t, \quad \epsilon_t \sim \mathbf{F}, \\
h_t &= (1-\beta) + \beta h_{t-1} + \alpha u_{t-1} + \gamma \operatorname{sign}(-\epsilon_{t-1})(u_{t-1} + \frac{1}{2}), \\
\kappa_\tau &= \exp\left(m + \sum_{j=1}^{N} \delta_j \left[\sum_{k=1}^{K_j} \phi_k(w_j) X_{j,\tau-k}\right]\right),
\end{aligned}
\tag{2}
$$

where $u_t = \dfrac{\partial \log p(r_t|h_t, \kappa_\tau)}{\partial h_t}$ is the conditional score of the observation at time t. Similarly to the GARCH-MIDAS model in (1), the same set of parameters (α, β, γ) governs the evolution of the short-term log volatility component h_t. Differently from the GARCH-MIDAS model above, no positivity constraints are necessary on the model parameters and the stationarity condition is $|\beta| < 1$. $\mathbf{F}$ is some (potentially non-normal) distribution with mean zero and unit variance. Finally, the MIDAS volatility component κ_τ is the same as in the GARCH-MIDAS model above.

The model nests several interesting specifications, considered in this article. When $\gamma = 0$ and $\epsilon_t \sim \mathbf{N}(0,1)$, then $u_t = \dfrac{\epsilon_t^2 - 1}{2}$, we have a GAS-MIDAS with normal errors model which is comparable to the Gaussian GARCH-MIDAS (the only difference being the exponential link function). The problem with the GARCH-MIDAS model specification is that some very large shocks (possible outliers) can have a very strong effect on volatility which would take time to go back to the "usual" levels, and also, affect the estimation of the model parameters. In other words, the impact of the shock is unbounded. An attractive feature of the GAS-type models is that if some non-Gaussian distribution is considered, then the impact of the large shocks is downweighted and has a smaller effect on the volatility evolution. Therefore, when $\gamma = 0$ and $\epsilon_t \sim \mathbf{Student}_\nu(0,1)$ (Student-t with zero mean and unit variance) we obtain the GAS-MIDAS model with Student errors and $u_t = \dfrac{1}{2}\left[\dfrac{(\nu+1)\epsilon_t^2}{(\nu-2)+\epsilon_t^2} - 1\right]$. In this case, the effect of the shock is bounded to $-\dfrac{1}{2} \leq u_t \leq \dfrac{\nu}{2}$. Finally, in order to introduce the leverage effect the parameter γ in (2) is left unrestricted resulting in the GAS-MIDAS model with leverage effect.

3.3 SV type models

Differently from the previously described GARCH and GAS specifications, the evolution of the variance equation in the SV models is parameter driven. In other words, it is a latent stochastic process that allows for greater flexibility in the short-term volatility component. The general formulation of the SV-MIDAS model can be written as follows:

$$
\begin{aligned}
r_t &= \sqrt{\sigma_t^2}\epsilon_t = \sqrt{\kappa_\tau \exp(h_t)}\epsilon_t, \quad \epsilon_t \sim \mathbf{F}, \\
h_t &= \beta h_{t-1} + \sigma_\eta \eta_t, \qquad \eta_t \sim \mathbf{N}(0,1), \\
\kappa_\tau &= \exp\left(m + \sum_{j=1}^{N} \delta_j \left[\sum_{k=1}^{K_j} \phi_k(w_j) X_{j,\tau-k} \right] \right).
\end{aligned}
\tag{3}
$$

Here the short-term log-volatility h_t follows a latent AR(1) process and the MIDAS volatility component remains the same as before. The asymmetric volatility response to a negative shock is captured via the negative correlation, $\mathrm{corr}(\epsilon_{t-1}, \eta_t) = \rho < 0$ which is equivalent to $\gamma > 0$ in GARCH-type and GAS-type models. Note that it is crucial to consider contemporaneous dependence (ϵ_{t-1}, η_t) as opposed to inter-temporal dependence (ϵ_t, η_t) between the error terms, otherwise, the latter specification is not consistent with the efficient market hypothesis (Harvey and Shephard, 1996; Yu, 2005). For a more convenient model representation, we can rewrite the model above with uncorrelated measurement and transition equation errors, as in Yu (2005). To this end, let $\eta_t = \rho\epsilon_{t-1} + \sqrt{1-\rho^2}w_t$ where $w_t \sim \mathbf{N}(0,1)$ and $\mathrm{corr}(\epsilon_t, w_{t+1}) = 0$. Then the evolution of short-term volatility can be written as:

$$
h_t = \beta h_{t-1} + \sigma_\eta \rho \epsilon_{t-1} + \sigma_\eta \sqrt{1-\rho^2} w_t. \tag{4}
$$

The model in (3)-(4) nests several interesting specifications, considered in this article. First, if $\mathbf{F} \equiv \mathbf{N}(0,1)$ and $\rho = 0$, w_t boils down to η_t and we have the standard SV-MIDAS model with Gaussian innovations. Next, even though there is evidence that the SV model with Normal errors is sufficient to create the excessive kurtosis in the real data (Carnero et al., 2004), we still consider an SV-MIDAS model with Student-t innovations. This can be easily done by replacing the Normal distribution of ϵ_t with a Student-t distribution with ν degrees of freedom. Finally, in order to include the leverage effect just allow the contemporaneous correlation ρ to be unrestricted, resulting in a leverage SV -MIDAS model.

GARCH-MIDAS and GAS-MIDAS models are straightforward to estimate since both models and their variants are all observation-driven. The last three SV-MIDAS specifications (with Gaussian, Student errors, and leverage effect) face certain challenges because the short-term volatility is latent. However, by employing a novel Bayesian estimation strategy, based on Density Tempered Sequential Monte Carlo Sampler, parameters and latent states can be recovered without further complications. In the following section, we briefly describe the Bayesian inference algorithm for the models above.

4 Bayesian inference

In order to estimate and compare the volatility models with different assumptions of innovations and long/short-term volatility processes, we employ a novel Bayesian estimation approach - a density-tempered sequential Monte Carlo sampler of Tran et al. (2014) and Duan and Fulop (2015). The DTSMC sampler not only delivers posterior distributions of the model parameters but also provides an estimate of the log marginal likelihood which can be used for consistent model comparison.

4.1 Prior distributions

Denote $\Theta = \{\alpha, \beta, \gamma, m, \delta_{1:N}, \omega_{1:N}\}$ as the set of fixed parameters in the GARCH and GAS type models and $\Theta = \{\sigma_\eta, \beta, \rho, m, \delta_{1:N}, \omega_{1:N}\}$ as the set of fixed parameters in the SV type model. We use vague proper prior distributions for the model parameters. For the parameters that govern the long-term volatility κ_τ, we assume $m \sim \mathbf{N}(0,1)$, $\delta_j \sim \mathbf{N}(0,1)$, $\omega_j \sim \mathbf{U}(1,30)$ for $j = 1, \dots N$, see the discussion in Nguyen and Virbickaite (2022). Note that we standardize the low-frequency variables X_j so that the standard normal prior distribution of δ_j imposes the same prior effect of the explanatory variables. The uniform prior distribution of the weighting parameter ω allows for a flexible regularization scheme (from very loose to very tight). For the parameters that control the short-term volatility, we assume Gamma prior distributions, $\alpha \sim \mathbf{G}(1,10)$, $\beta \sim \mathbf{G}(10,1)$, and $\gamma \sim \mathbf{G}(1,10)$ in GARCH- and GAS-type models. In SV-type models, $\sigma_\eta \sim \mathbf{G}(0.5, 0.25)$, $\beta \sim \mathbf{B}(20, 1.5)$, and $\rho \sim \mathbf{U}(-1,1)$, similar to the prior distributions in Kastner and Frühwirth-Schnatter (2014). Finally, in the volatility models with Student-t innovations, the prior distribution for the degrees of freedom is $\nu \sim \mathbf{G}(1, 0.1)$ so that the prior mean of ν is 10.

4.2 Density tempered Sequential Monte Carlo

The DTSMC algorithm samples sequentially from the prior distribution $p(\Theta)$ to the posterior distribution $p(\Theta|r_{1:T}) \propto p(r_{1:T}|\Theta)p(\Theta)$ through a sequence of distributions,

$$\pi_i(\Theta) := p_i(\Theta|r_{1:T}) \propto p(r_{1:T}|\Theta)^{\gamma_i} p(\Theta),$$

where γ_i is a level temperature for $i = 0, \ldots, S$ such that $0 = \gamma_0 < \gamma_1 < \ldots < \gamma_S = 1$. The tempering sequence $\gamma_i = (i/S)^3$ for $i = 0, \ldots, S$ is chosen in order to smoothly interpolate between the prior and the posterior distribution. The conditional likelihood $p(r_{1:T}|\Theta)$ is calculated exactly in the GARCH and GAS models, and approximately in the SV model using a bootstrap particle filter.

To perform the estimation using the DTSMC sampler, we start by sampling a set of M weighted particles $\{W_0^j = 1/M, \Theta_0^j\}_{j=1}^M$ from the prior distribution $\pi_0(\Theta) = p(\Theta)$. Then, for each level of temperature γ_i, we perform reweighting, resampling, and Markov moves so that the distribution of particles approximates the interpolation distribution $\pi_i(\Theta)$.

In the reweighting step, a set of M weighted particles $\{W_{i-1}^j, \Theta_{i-1}^j\}_{j=1}^M$ obtained from the previous interpolation distribution $\pi_{i-1}(\Theta)$ are re-weighted,

$$\begin{aligned} w_i^j &= W_{i-1}^j p(r_{1:T}|\Theta_{i-1}^j)^{\gamma_i - \gamma_{i-1}}, \;\; j = 1, ..., M, \\ W_i^j &= \frac{w_i^j}{\sum_{s=1}^{M} w_i^s}, \;\; j = 1, ..., M. \end{aligned}$$

If the effective sample size (ESS) of the weighted particles is below some threshold, the particles are resampled and refreshed by a Markov kernel whose invariant distribution is $\pi_i(\Theta)$. We use a random walk proposal for the new value of Θ. An advantage of the DTSMC is that the Markov moves among particles are independent, hence the move can be run in parallel for each particle. As the sequence i progresses, the Markov moves are required at a lower rate. Finally, the log marginal likelihood is straightforward to compute as

$$\log p(r_{1:T}) = \sum_{i=1}^{S} \log\left(\sum_{j=1}^{M} w_i^j\right).$$

The difference of LMLs between two competing models is a log Bayes Factor (BF). The algorithm is described in detail in Appendix A.

5 Empirical application

This section presents the data and empirical results for modeling and forecasting crude oil return distributions using the previously described GARCH-, GAS-, and SV-MIDAS models.

Crude oil data

We consider daily spot prices (dollars per barrel) of West Texas Intermediate (WTI) available at U.S. Energy Information Administration website. The data is from January 4, 2000 to September 30, 2022 (5709 observations in total); the negative oil spot price that was observed on April 20, 2020 was removed. The log returns $\widetilde{r}_t$ (in %) are calculated from the daily price data P_t and then de-meaned by subtracting the overall sample mean. Due to negligible and mostly insignificant autocorrelations in the mean process, it is safe to assume that the mean is static. In the rest of the article whenever we refer to the log returns we are referring to the zero-mean returns. Figure 1 draws the daily WTI prices, log-returns, and 22-day rolling window realized volatility (RVroll), calculated as a sample standard deviation of the last 22 days, aligned at the right.

By looking at the evolution of crude oil prices, one can observe that during the 2008-09 global recession, the price of crude oil plummeted from around USD 145 per barrel in July 2008 to around USD 30 per barrel at the turn of 2009. There was another noticeable price decline at the end of 2014 - the beginning of 2015, known as the 2014 - 2016 oil price plunge. The experts explain that the plunge was primarily driven by supply factors. Another dramatic decrease in crude oil prices happened just at the beginning of the travel restrictions due to the COVID-19 pandemic. Apart from the negative oil price, observed on April 20, 2020, the crude oil prices were as low as USD 9 per barrel that same month. Finally, the most recent increase in the crude oil price happened at the end of February - the beginning of March 2022 as a result of Russia's invasion of Ukraine. The plots of the log returns and the rolling 22-day standard deviations confirm the increases in volatility during the aforementioned turbulent periods. Finally, the boxplot indicates a slight negative skewness of the distribution of the log returns and the presence of multiple outliers, an indication of fat-tailed behavior. Table 1 reports the main descriptive statistics for the crude oil log returns. We confirm that the log return data is negatively skewed, has extremely fat tails, and is non-normally distributed.

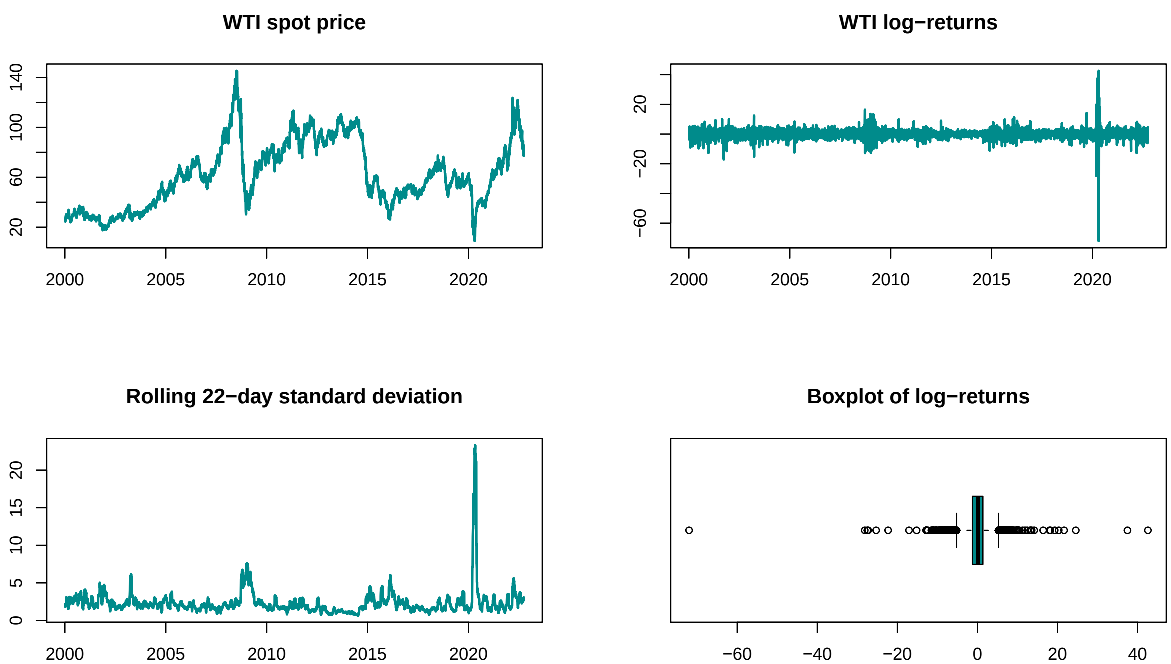

Figure 1: Crude oil spot prices, de-meaned log returns, rolling 22-day standard deviation (right-aligned), and a boxplot for the de-meaned log returns from January 4, 2000 to September 30, 2022.

Macroeconomic variables

Selecting a particular set of explanatory variables that describe the dynamics of the crude oil price volatility is no easy task since crude oil is intricately interlaced with multiple aspects of the national and global economy. Some authors opt to consider as many macroeconomic/financial variables as possible and perform variable selection in a data-rich environment, see Ma et al. (2018) and Bonnier (2022), for example. Alternatively, we rely on the related research on crude oil volatility modeling

Table 1: Descriptive statistics.

	mean	Q2	sd	skew	kurt	min	max	JB	ADF	n
r_t	0.000	0.088	2.955	-2.048	82.686	-72.047	42.563	0.000	0.010	5709
RV	2.378	2.009	1.814	7.056	68.741	0.702	22.867	0.000	0.010	273
VIX	20.271	18.230	8.176	1.619	3.700	9.510	59.890	0.000	0.204	273
IGREA	9.377	3.215	70.524	0.408	-0.303	-162.569	188.500	0.014	0.354	273
GOP	0.850	1.103	3.512	-1.503	4.675	-15.071	8.617	0.000	0.010	273
EPU	123.940	114.654	45.276	1.392	3.108	57.203	350.460	0.000	0.081	273
NFCI	-0.347	-0.505	0.523	3.277	12.743	-0.788	2.659	0.000	0.228	273

Notes: Table reports the descriptive statistics for the daily log WTI returns (in %) and other explanatory variables (in monthly frequency) from January 4, 2000 to September 30, 2022. The JB reports the p-value for the Jarque-Bera test for Normality. The ADF reports the p-value for the augmented Dickey–Fuller test that a unit root is present.

(Conrad et al., 2014; Wei et al., 2017; Pan et al., 2017; Bonnier, 2022), and consider a selection of macroeconomic variables, that have been proven to be powerful indicators of changes in oil price volatility in the long-run. We have grouped such variables into three groups. We refer to the first group as *mechanical* volatility drivers, since the relationship between crude oil volatility and such variables has no economic meaning *per se*, i.e. the relationship is of a purely mechanical nature. The second group contains *fundamental* supply and demand drivers of the oil price volatility. Finally, the last group contains the *policy*-related indicators of the overall state of the economy.

1. *Mechanical* volatility drivers. **RV**: monthly realized volatility, calculated as a sum of squared log returns for each month. **OVX/VIX**: the Chicago Board Options Exchange's crude Oil Volatility (OVX) and S&P 500 Volatility (VIX) Indices. Since OVX data is available since 2007 only, we use VIX instead. VIX a popular measure of the stock market's expectation of volatility based on S&P 500 index options. VIX data is obtained from Bloomberg database. It is expected that the RV and VIX variables will have a positive effect on long-term crude oil volatility since the volatility of the financial returns tends to be highly persistent.

2. *Fundamental* variables. **GOP**: log growth rates of global oil production (thousands of barrels), as a proxy for oil supply, data is from the EIA website. **IGREA**: Index of Global Real Economic Activity (Kilian, 2009), in levels, as a proxy for oil demand. Data is from FRED. There is mixed evidence considering the effect of the supply/demand variables. Some authors find no effect (Wei et al., 2017), others find a positive effect (Kilian and Murphy, 2014; Wei et al., 2017; Ma et al., 2019), and others - a negative one (Pan et al., 2017).

3. *Policy* variables. **EPU**: economic policy uncertainty index, in levels. The EPU index is constructed based on newspaper articles regarding policy uncertainty from leading newspapers, therefore, an increase in the EPU is expected to have a positive effect on crude oil volatility. **NFCI**: The Chicago Fed's National Financial Conditions Index, in levels, as a monetary policy indicator. Positive values of the NFCI indicate financial conditions that are tighter than average, and an increase in the NFCI would be associated with an increase in crude oil volatility as well. Data for both variables are from FRED.

Table 1 presents the descriptive statistics for all explanatory variables. Even though VIX, IGREA, and NFCI do not reject the null hypothesis for unit root, it is common to use those variables in levels. Appendix B includes the time-series plots for the explanatory variables. Finally, Table 2 contains the sample correlations between the macroeconomic indicators. As expected, VIX is positively correlated with RV, the same as both policy uncertainty indices (EPU and NFCI). Finally, the levels of fundamental variables have a weak and negative correlation with the volatility of the crude oil returns. Note that in the estimation step, all variables have been scaled to have zero mean and unit variance to ensure stable samples from the prior. Therefore, the interpretation of the corresponding elasticities δ_j has to be adjusted accordingly.

Table 2: Correlation matrix of explanatory variables.

	RV	VIX	IGREA	GOP	EPU	NFCI
RV	1.000					
VIX	0.460	1.000				
IGREA	-0.150	-0.040	1.000			
GOP	-0.040	-0.170	0.180	1.000		
EPU	0.350	0.510	-0.270	-0.370	1.000	
NFCI	0.380	0.700	0.090	-0.110	0.260	1.000

Notes: Table reports the sample correlation coefficients between the monthly explanatory variables from January 4, 2000 to September 30, 2022 (273 observations).

5.1 Estimation results

For the in-sample results, we have used all available data, i.e. 5709 observations. Table 3 reports the log marginal likelihood estimates for GARCH, GAS and SV models with Normal and t innovations as well as the leverage effect in the volatility, with and without MIDAS effects. For the MIDAS regressions, the explanatory variable is the previous month's realized volatility. Note that the model fit is measured via log marginal likelihood, which accounts for model complexity and can be used for comparing even non-nested models. From the estimation results in Table 3, we can distinguish important findings in three dimensions: within-model dimension, across-model dimension, and MIDAS vs no MIDAS dimension.

First, in the within-model dimension, we look at each model class, GARCH, GAS and SV, individually. We can observe that fat tails and volatility with leverage effect play an important if not crucial role in the successful modeling of crude oil return distribution. Models that incorporate Student errors are consistently better than models with Gaussian errors, as well as models with leverage effect vs models with simple volatility dynamics. The results hold for all models in the non-MIDAS setting (first column), as well as for all models with MIDAS regressions (second column). Moreover, the differences in the LML are staggering: the log BFs between the most simple (volatility without leverage and Gaussian errors) and the most flexible (volatility with leverage and Student errors) models vary from 50 (in SV model class) to as much as 250 (in GAS model class). Interestingly, the inclusion of the fat-tailed distribution term and the leverage effect seems to be more important for observation-driven models (GARCH and GAS) as compared to the SV model. This is due to the fact the SV model's latent volatility specification allows for greater flexibility in the volatility dynamics.

Secondly, in the across-model dimension, we compare the corresponding volatility model classes. In particular, among the models with Gaussian errors, the SV-type model decidedly outperforms the GARCH and GAS-type models, both with and without MIDAS: the BFs vary from 160 to 190. Interestingly, among models with Student errors, GAS model, followed by GARCH model, perform better than the SV model, but only marginally. Finally, among the models with Student errors and leverage effects, the model ordering changes once again and now the GAS model is the best, followed by the SV and then GARCH model. The results emphasize the importance of fat-tailed distribution and leverage effect in the volatility: without these features, the more flexible SV model is clearly the best, but once both effects are considered, the GARCH, GAS, and SV models become comparable. The model order always stays the same for MIDAS and no MIDAS settings.

Finally, in the MIDAS vs non-MIDAS dimension, we compare the gains in the in-sample fit by using the RV as a predictor variable of the long-term volatility dynamics. The differences in the

LMLs (the log BFs) are represented by the third column in Table 3. We can see that incorporating an explanatory variable for long-term volatility dynamics is beneficial in the GARCH model class, independently from the distributional term or the volatility specification. However, the results are rather mixed for the GAS and SV-type models: there is only a marginal improvement in the usefulness of MIDAS regressions for these model classes. This can be explained by the fact that GAS and SV models are more flexible in the volatility dynamics as compared to the GARCH-type models, therefore, the inclusion of an additional explanatory variable is not necessarily always beneficial.

Table 3: Log marginal likelihoods and log Bayes factor of GARCH, GAS and SV models.

	without MIDAS	with MIDAS	log BF
GARCH Gaussian	-12798.42	-12791.99	6.43
GARCH Student	-12611.06	-12604.52	6.55
GARCH Student leverage	-12599.08	-12589.81	9.27
GAS Gaussian	-12829.09	-12830.75	-1.67
GAS Student	-12606.25	-12602.14	4.11
GAS Student leverage	-12583.74	-12580.65	3.09
SV Gaussian	-12637.34	-12636.24	1.11
SV Student	-12616.22	-12611.15	5.07
SV Student leverage	-12589.76	-12586.32	3.44

Notes: Table reports the log marginal likelihood of GARCH, GAS and SV models with and without MIDAS. In the MIDAS volatility models, the long-term volatility component is modelled using the RV as an explanatory variable and the restricted beta weighting function. The selected lag length is such that the marginal likelihood becomes insensitive to the its choice ($K = 24$).

Next, we wish to take a closer look at the effect of the exogenous explanatory variables on the long-term volatility dynamics. To this end, apart from the RV, we also consider the previously defined VIX, GOP, IGREA, EPU, and NFCI as an explanatory variable. The results for individual regressions are available in Table 4. We are interested not only in the in-sample fit in terms of the LML but also in the sign and significance of the estimated parameters that measure the effects of the macroeconomic indicators. Note that all models in Table 4 are with t innovations and leverage effect in the volatility.

First, in terms of the in-sample fit, MIDAS models with VIX as an explanatory variable are better than the MIDAS-RV models, for GARCH, GAS, and SV model classes. The rest of the explanatory macroeconomic variables do not present an increase in the LML. Overall, the universally best predictor variables are the *mechanical* drivers of the volatility, RV and VIX.

Secondly, the sign and significance of the explanatory variables are of interest to investors and policymaker, as it helps to understand how macroeconomic conditions affect long-term crude oil price volatility. The VIX coefficient is positive and significant, as expected, as it represents the *mechanical* driving force of the volatility. In other words, high past market volatility affects positively future long-run crude oil volatility. This effect can be attributed to market contagion, as the distress in the stock markets usually spreads to other markets, commodity markets included. A positive effect of the OVX volatility index on the long-run crude oil volatility was also found by Lu et al. (2020); Bonnier (2022), among others. The supply and demand variables have a positive and statistically significant effect (except for the demand variable in GAS and SV specifications) on the volatilities. In the previous research, the effect of these factors is contradicting. Kilian and Murphy (2014) demonstrate that supply and demand shocks have a positive and persistent effect on the oil price. Wei et al. (2017), for example, find that the demand factor is insignificant, meanwhile, the supply factor is positive and statistically significant in oil volatility modeling. Pan et al. (2017) find a negative and statistically significant effect of supply and demand factors on crude oil volatility. Ma et al. (2019) construct a predictor variable based on oil supply and demand and show that it has a positive effect on crude oil volatility. Finally, an increase in both *policy* variables is usually associated with an increase in the long-run crude oil volatility. Nonetheless, the estimated signs of the EPU coefficients are negative for all models, and is insignificant for the SV-MIDAS specification. This finding is in contrast to the previous research that found positive and statistically significant effects of EPU on crude oil price volatility (Wei et al., 2017; Ma et al., 2019; Bonnier, 2022). The tightening of the financial/monetary conditions, represented by the NFCI variable, results in increased uncertainty regarding crude oil price volatility.

5.2 Out-of-sample forecasting

In order to perform the out-of-sample forecast, we have retained the last 1000 observations and re-estimated all models. We then perform 1, 5, 10, and 20-step-ahead density forecasts by rolling the data. In order to evaluate the forecast accuracy, we calculate the value of the log predictive score (LPS), and the continuous rank probability score (CRPS), which is less sensitive to the outliers (Clark and Ravazzolo, 2015). Higher LPS and CRPS values indicate better-performing models.

Table 4: Estimation results for the MIDAS volatility models with leverage effect and t-innovations.

Variable	Model	m	α	β	γ	δ	ω	ν	LML
VIX	GARCH-MIDAS	1.720	0.043	0.901	0.062	0.333	25.050	6.574	-12588.516
		(1.551 ; 1.923)	(0.030 ; 0.058)	(0.882 ; 0.918)	(0.042 ; 0.081)	(0.238 ; 0.422)	(17.626 ; 29.503)	(5.718 ; 7.518)	
	GAS-MIDAS	1.571	0.162	0.980	0.076	0.330	23.664	6.643	-12576.368
		(1.426 ; 1.706)	(0.137 ; 0.190)	(0.974 ; 0.986)	(0.058 ; 0.094)	(0.220 ; 0.435)	(15.197 ; 29.429)	(5.815 ; 7.594)	
	SV-MIDAS	1.540	0.131	0.980	-0.453	0.357	24.877	9.402	-12580.822
		(1.405 ; 1.669)	(0.113 ; 0.152)	(0.974 ; 0.986)	(-0.550 ; -0.363)	(0.264 ; 0.451)	(17.493 ; 29.299)	(7.517 ; 11.762)	
GOP	GARCH-MIDAS	1.871	0.046	0.914	0.055	0.278	22.489	6.517	-12596.295
		(1.597 ; 2.218)	(0.031 ; 0.060)	(0.896 ; 0.929)	(0.035 ; 0.080)	(0.139 ; 0.426)	(13.301 ; 28.675)	(5.668 ; 7.497)	
	GAS-MIDAS	1.624	0.170	0.986	0.076	0.177	20.118	6.611	-12584.034
		(1.435 ; 1.821)	(0.147 ; 0.195)	(0.981 ; 0.991)	(0.058 ; 0.095)	(0.045 ; 0.313)	(7.142 ; 29.176)	(5.835 ; 7.535)	
	SV-MIDAS	1.501	0.116	0.989	-0.452	0.145	18.089	8.883	-12590.154
		(1.282 ; 1.710)	(0.104 ; 0.138)	(0.986 ; 0.992)	(-0.533 ; -0.369)	(0.014 ; 0.280)	(6.429 ; 27.248)	(7.374 ; 11.047)	
IGREA	GARCH-MIDAS	1.783	0.047	0.898	0.069	0.232	6.962	6.378	-12595.873
		(1.553 ; 2.074)	(0.033 ; 0.063)	(0.883 ; 0.912)	(0.050 ; 0.091)	(0.097 ; 0.370)	(1.477 ; 24.110)	(5.621 ; 7.324)	
	GAS-MIDAS	1.586	0.175	0.984	0.080	0.128	9.194	6.495	-12585.143
		(1.408 ; 1.771)	(0.150 ; 0.203)	(0.979 ; 0.989)	(0.063 ; 0.100)	(-0.069 ; 0.326)	(1.373 ; 26.699)	(5.729 ; 7.366)	
	SV-MIDAS	1.585	0.131	0.987	-0.494	0.012	14.061	8.975	-12590.329
		(1.401 ; 1.782)	(0.116 ; 0.149)	(0.982 ; 0.992)	(-0.563 ; -0.381)	(-0.144 ; 0.219)	(2.186 ; 26.536)	(7.479 ; 10.823)	
EPU	GARCH-MIDAS	1.868	0.049	0.900	0.064	-0.310	4.877	6.366	-12595.382
		(1.633 ; 2.189)	(0.035 ; 0.064)	(0.881 ; 0.918)	(0.043 ; 0.088)	(-0.491 ; -0.135)	(1.018 ; 17.976)	(5.659 ; 7.139)	
	GAS-MIDAS	1.626	0.171	0.986	0.077	-0.212	10.704	6.570	-12584.109
		(1.419 ; 1.830)	(0.146 ; 0.198)	(0.981 ; 0.991)	(0.059 ; 0.095)	(-0.452 ; -0.030)	(1.468 ; 26.491)	(5.785 ; 7.521)	
	SV-MIDAS	1.559	0.130	0.988	-0.441	-0.158	8.955	9.264	-12589.838
		(1.373 ; 1.769)	(0.113 ; 0.146)	(0.984 ; 0.992)	(-0.510 ; -0.359)	(-0.417 ; 0.053)	(1.725 ; 23.866)	(7.978 ; 11.119)	
NFCI	GARCH-MIDAS	1.766	0.048	0.892	0.067	0.290	24.176	6.491	-12595.643
		(1.588 ; 1.984)	(0.034 ; 0.065)	(0.873 ; 0.908)	(0.040 ; 0.093)	(0.190 ; 0.381)	(13.140 ; 29.619)	(5.745 ; 7.302)	
	GAS-MIDAS	1.597	0.175	0.981	0.082	0.250	18.932	6.558	-12582.705
		(1.441 ; 1.762)	(0.148 ; 0.204)	(0.975 ; 0.986)	(0.063 ; 0.101)	(0.114 ; 0.386)	(6.064 ; 28.889)	(5.757 ; 7.507)	
	SV-MIDAS	1.554	0.139	0.982	-0.435	0.218	16.686	9.601	-12589.062
		(1.373 ; 1.736)	(0.116 ; 0.160)	(0.975 ; 0.988)	(-0.518 ; -0.353)	(0.067 ; 0.366)	(4.035 ; 28.470)	(7.771 ; 11.939)	

Notes: The table reports the summary of the posterior estimations of the MIDAS volatility models with leverage effect and t-innovations. The long-term volatility component is modelled using one explanatory variable and the restricted beta weighting function. The selected lag length is such that the marginal likelihood becomes insensitive to the its choice ($K = 24$). The numbers in the brackets reveal the (5% - 95%) credible intervals.

The details for both score calculations are available in Appendix D. Table 5 presents the out-of-sample forecasting results. In particular, the top three lines of the table present the **average per observation** LPS and CRPS for the benchmark specifications: GARCH, GAS, and SV models with t-innovations, leverage effect in volatilities, and no MIDAS effects. The rest of the table contains the differences between the benchmark model and competing models, such that positive numbers mean better results.

First, by looking at the LPS and the CRPS of the benchmark models, we find that GAS models are consistently better than GARCH and SV models for all forecast horizons. Moreover, the differences are statistically significant for 1 and 20-step horizons at 10% level as compared to the GARCH models. SV model always stays in second place, with the GARCH model being generally the worst. This could be explained by the fact that GAS models are more flexible as compared to the GARCH-type models, and, being parameter-driven models, do not suffer the adverse effects of the model complexity, as the SV-type specifications.

Secondly, in terms of LPS (first four columns of Table 5) and CRPS (last four columns in the same table), including MIDAS regressions in the long-run volatility dynamics is most beneficial for GARCH-type models. In almost all cases, the short- and long-run predictive performance is improved as compared to the benchmark specification and the differences are almost always significant at 5% level. As for the GAS-type models, the usefulness of the exogenous regressors is less pronounced, and in general, the gain is more visible for shorter forecast horizons (1- and 5-step-ahead). Finally, in the context of the SV models, the inclusion of the MIDAS regressions does not tend to improve the out-of-sample forecast accuracy, except for the RV variable. The results are in line with the in-sample outcomes: GARCH-type models, being the least flexible in a sense, benefit the most from the exogenous information via MIDAS regressions. GAS-type models, which are more flexible than GARCH, in some instances are able to capture complex volatility dynamics even without the MIDAS regressions. Finally, for the SV-type models, which are the most flexible of the three, the inclusion of the MIDAS effects provides only marginal improvement in-sample and (almost) no improvement out-of-sample.

5.3 Risk management

In order to evaluate model performance in terms of economic gains, we perform a 1-step-ahead Value-at-Risk (VaR) and Expected Shortfall (ES) prediction. A q-level VaR is the q-level quantile of the return distribution (Duffie and Pan, 1997), and the q-level ES is the conditional expectation of returns that exceed the corresponding q-level VaR:

$$\begin{aligned} q &= Pr\left(r_t \leq \mathrm{VaR}_{q,t}\right), \\ \mathrm{ES}_{q,t} &= E\left[r_t \left| r_t \leq \mathrm{VaR}_{q,t}\right.\right]. \end{aligned}$$

We calculate the VaR and ES at the 5th and 1st percentiles of the return distribution, i.e., $q = 5\%$ and $q = 1\%$. Next, in order to evaluate the forecast accuracy of the associated risk measures, we employ several loss functions (Taylor, 2019):

$$\begin{aligned} \mathrm{IF} &= \sum_{t=1}^{T} \mathbf{I}\left(r_t < \mathrm{VaR}_{q,t}\right), \\ \mathrm{QS}_t &= (r_t - \mathrm{VaR}_{q,t})(q - \mathbf{I}\left(r_t < \mathrm{VaR}_{q,t}\right)), \\ \mathrm{ALS}_t &= -\log\left(\frac{q-1}{\mathrm{ES}_{q,t}}\right) - \frac{(r_t - \mathrm{VaR}_{q,t})(q - \mathbf{I}\left(r_t < \mathrm{VaR}_{q,t}\right))}{q\,\mathrm{ES}_{q,t}}. \end{aligned}$$

Table 5: Log predictive scores and continuous rank probability scores for out of sample forecasts.

	LPS				CRPS			
	$h = 1$	$h = 5$	$h = 10$	$h = 20$	$h = 1$	$h = 5$	$h = 10$	$h = 20$
GARCH	-2.405	-3.235	-3.656	-4.171	-1.774	-3.923	-5.721	-8.536
GAS	-2.382†	-3.209	-3.622	-4.110†	-1.744†	-3.845	-5.615	-8.334†
SV	-2.403	-3.231	-3.643	-4.132	-1.755	-3.877	-5.655	-8.350
GARCH MIDAS RV	0.005*	0.005	0.005	0.006	0.006*	0.019	0.025	0.028
GAS MIDAS RV	0.001	-0.000	-0.002	-0.006	0.001	0.005	0.006	-0.005
SV MIDAS RV	0.003	0.002	-0.000	-0.008	0.004	0.012	0.017	0.013
GARCH MIDAS VIX	0.002*	0.003*	0.004*	0.011*	0.003*	0.008*	0.012*	0.024*
GAS MIDAS VIX	0.001	0.001	-0.001	-0.004	0.000	0.001	-0.004	-0.005
SV MIDAS VIX	-0.006	-0.006	-0.008	-0.009	-0.003	-0.010	-0.020	-0.026
GARCH MIDAS GOP	0.001*	0.002*	0.006*	0.017*	0.001*	0.003*	0.006*	0.014*
GAS MIDAS GOP	-0.001	-0.001	-0.001	0.000	-0.001	-0.001	-0.001	0.001
SV MIDAS GOP	-0.003	-0.004	-0.005	-0.005	-0.001	-0.007	-0.011	-0.013
GARCH MIDAS IGREA	0.005*	0.006*	0.008*	0.013*	0.007*	0.015*	0.028*	0.046*
GAS MIDAS IGREA	-0.000	0.000	-0.000	-0.001	0.000	0.002	0.001	0.002
SV MIDAS IGREA	-0.006	-0.006	-0.008	-0.011	-0.002	-0.008	-0.014	-0.016
GARCH MIDAS EPU	0.004*	0.006*	0.010*	0.019*	0.006*	0.015*	0.027*	0.055*
GAS MIDAS EPU	-0.000	0.001	0.001	0.001	0.001	0.003	0.004	0.008
SV MIDAS EPU	-0.001	-0.001	-0.002	-0.004	0.001	0.000	-0.001	-0.001
GARCH MIDAS NFCI	0.003*	0.004*	0.006*	0.014*	0.004*	0.010*	0.017*	0.036*
GAS MIDAS NFCI	0.001	0.001	0.001	-0.001	0.001	0.002	-0.001	-0.001
SV MIDAS NFCI	-0.003	-0.004	-0.004	-0.004	-0.002	-0.008	-0.017	-0.024

Notes: The table reports the relative LPS and CRPS for the last $T = 1000$ observations. The first three lines report the LPS and CRPS of volatility models without an explanatory variable. The rest of the lines report the relative scores of the MIDAS model w.r.t. to the corresponding non-MIDAS model where positive numbers are better. † denotes that the corresponding model is significantly different from the GARCH model at 10% level based on the two-sided Diebold and Mariano (1995) test for the non-nested models. * denotes that the corresponding MIDAS model significantly outperforms the corresponding non-MIDAS model at the 5% level based on the one-sided Diebold and Mariano (1995) test for the nested models where the standard errors of the test statistics are computed with the Newey–West estimator (Clark and McCracken, 2001). All models are with t-innovations and leverage effects in the volatility.

The Indicator loss Function (IF) counts the number of exceptions where the crude oil return falls below the VaR threshold. Since the out-of-sample evaluation period is 1000 days, the expected number of exceptions is 50 for $q = 5\%$ and 10 for $q = 1\%$. QS is the quantile score loss function often used to evaluate the quality of the VaR forecast. ALS is the Asymmetric Laplace log score loss function that evaluates the joint forecast of VaR and ES. Both of these measures are strictly proper rules, meaning that the expected QS is minimized at the true quantile series, and the expected ALS is minimized at the true VaR and ES series (Fissler and Ziegel, 2016; Taylor, 2019).

Table 6: Risk measures.

	VaR	ES	IF	QS	ALS	VaR	ES	IF	QS	ALS
	q = 5%					q = 1%				
GARCH	-6.33	-8.62	41	0.48	3.08	-9.95	-12.45	12	0.19	3.70
GAS	-5.38	-7.34	49	0.46†	3.06	-8.48	-10.62	16	0.19	3.71†
SV	-4.67	-6.23	62	0.46	3.11	-7.15	-8.74	20	0.21	3.96†
GARCH MIDAS RV	-6.23	-8.45	43	0.47*	3.07*	-9.75	-12.13	14	0.18*	3.69
GAS MIDAS RV	-5.57	-7.58	46	0.46	3.05	-8.75	-10.94	16	0.18	3.68
SV MIDAS RV	-4.76	-6.30	62	0.46	3.11	-7.23	-8.77	19	0.20	3.94
GARCH MIDAS VIX	-6.36	-8.68	39	0.48*	3.08*	-10.02	-12.55	12	0.18	3.69
GAS MIDAS VIX	-5.39	-7.35	48	0.46	3.06	-8.49	-10.62	16	0.19	3.71
SV MIDAS VIX	-4.66	-6.20	64	0.47	3.12	-7.12	-8.69	21	0.21	3.99
GARCH MIDAS GOP	-6.20	-8.46	43	0.47	3.07	-9.76	-12.21	14	0.18	3.69
GAS MIDAS GOP	-5.34	-7.29	49	0.46	3.06	-8.42	-10.55	16	0.19	3.71
SV MIDAS GOP	-4.57	-6.08	65	0.47	3.13	-6.97	-8.49	22	0.21	4.02
GARCH MIDAS IGREA	-6.19	-8.41	42	0.47*	3.08*	-9.70	-12.09	14	0.19	3.70
GAS MIDAS IGREA	-5.30	-7.23	49	0.46	3.06	-8.35	-10.45	16	0.19	3.72
SV MIDAS IGREA	-4.65	-6.19	64	0.46	3.12	-7.10	-8.66	21	0.21	3.97
GARCH MIDAS EPU	-6.26	-8.53	43	0.48*	3.08*	-9.84	-12.32	13	0.19	3.70
GAS MIDAS EPU	-5.18	-7.08	50	0.46	3.06	-8.17	-10.24	17	0.19	3.73
SV MIDAS EPU	-4.55	-6.07	65	0.47	3.13	-6.97	-8.51	23	0.21	4.03
GARCH MIDAS NFCI	-6.23	-8.50	42	0.47*	3.08*	-9.82	-12.30	14	0.19	3.70
GAS MIDAS NFCI	-5.20	-7.10	50	0.46	3.06	-8.19	-10.26	17	0.19	3.73
SV MIDAS NFCI	-4.52	-6.03	64	0.47	3.13	-6.92	-8.45	23	0.21	4.02

The table reports the average VaR and ES together with the average of the IF, QS, and ALS loss functions. The expected exceptions at the 5% and 1% quantiles are 50 and 10 respectively. A smaller loss value means better. † denotes that the corresponding model is significantly different from the GARCH model at 10% level based on the two-sided Diebold and Mariano (1995) test for the non-nested models. * denotes that the corresponding MIDAS model significantly outperforms the corresponding non-MIDAS model at the 5% level based on the one-sided Diebold and Mariano (1995) test for the nested models where the standard errors of the test statistics are computed with the Newey–West estimator (Clark and McCracken, 2001).

The first five columns in Table 6 present the results for the risk prediction exercise for $q = 5\%$. Among the models with no MIDAS effects (first three lines), we find that GAS model produces the best results for $\text{VaR}_{5\%}$ and $\text{ES}_{5\%}$ prediction, and the improvement in $\text{VaR}_{5\%}$ case is statistically significant at 10% level as compared to the GARCH model. Next, we compare the benchmark non-MIDAS models with their corresponding counterparts that employ MIDAS regressions. Meanwhile, the improvement in GARCH-type models is statistically significant almost for all macroeconomic variables, the inclusion of exogenous regressors in SV-type models either has no effect or worsens the VaR and ES prediction performance. For GAS family models the MIDAS regressions present marginal yet statistically insignificant improvements as compared to the non-MIDAS benchmark specification. For the $q = 5\%$ percentile, GAS-type models are the best overall. The last five columns of Table 6 present the results for the risk prediction exercise for $q = 1\%$. Here the results are different than before: GARCH-type models, MIDAS and non-MIDAS, are generally the best, yet the benefits of MIDAS regressions are negligent for all three model classes. The results for $q = 1\%$ have to be regarded with caution, as the resulting sample of exceptions is very small.

6 Conclusions

The article proposes a comprehensive study that compares GARCH, GAS, and SV-type models for oil return distribution modeling and prediction. The models considered incorporate fat-tailed return distributions, asymmetric volatility response, and MIDAS regressions that allow long-term volatility to be affected by some external financial/macroeconomic variables. Inference and precision are carried out using a novel Bayesian estimation strategy, based on DTSMC.

In the in-sample results, we find overwhelming evidence that fat tails and asymmetric volatility response are of major importance for the in-sample model fit, independently of the model class (GARCH, GAS, SV) or the use of exogenous variables (MIDAS vs non-MIDAS). The inclusion of fat tails and leverage effect in the volatility is more important in observation-driven models (GARCH and GAS) as compared to the SV model. In comparing the model classes, we find that a simple SV model is decidedly better than the equivalent GARCH and GAS models, however, once the fat-tails and leverage effects are considered, the GARCH, GAS, and SV models become comparable, with GAS being the best. Finally, the inclusion of MIDAS effects benefits all GARCH-type models, but for GAS and SV the improvement is only marginal. Taking a closer look at the effects of the explanatory variables, we find that the universally best predictor variables are the mechanical drivers of the volatility, RV and VIX.

In the out-of-sample forecasting exercise, we find that GAS is always the best for all forecast horizons, SV is the second best and GARCH is the worst. The inclusion of MIDAS effects tends to improve the performance of GARCH-type models, which is not always true for the GAS-type models. Finally, SV-type models do not seem to benefit from exogenous sources of information in terms of out-of-sample prediction accuracy.

Finally, in order to evaluate the economic gains we perform risk prediction for 5 and 1% quantiles. We find that GAS model is the best overall for the 5% quantile, and the inclusion of MIDAS regressions is most beneficial to GARCH-type models. For 1% quantile prediction, GARCH model produced the best results, yet there is no apparent improvement in risk forecasting by including MIDAS regressions for any model class.

For future research projects, the GARCH, GAS and SV-type models, considered in this article, could be extended to incorporate asymmetric return distributions as well as more complex volatility dynamics, such as jumps. Another avenue for future research is to consider multivariate volatility models for the co-dependence modeling of oil and some stock market indices in MIDAS setting.

Acknowledgments

The first author is partially supported by the grant PID2020-113192GB-I00/AEI/10.13039/501100011033 (Mathematical Visualization: Foundations, Algorithms and Applications) from the Spanish State Research Agency (Ministerio de Ciencia e Innovación). Financial support from Jan Wallanders and Tom Hedelius Foundation (grant number BFV22-0005) is gratefully acknowledged by the second author. The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at NSC partially funded by the Swedish Research Council through grant agreement no. 2022-06725).

Appendix

A Density tempered Sequential Monte Carlo algorithm

Algorithm 1 The DTSMC for the volatility model

1. Sample $\Theta_0^j \sim p(\Theta)$, and set $W_0^j = 1/M$ for $j = 1...M$
2. **For** $i = 1, ..., K$,

Step 1: Reweighting: Compute the unnormalized weights

$$w_i^j = W_{i-1}^j \frac{p(r_{1:T}|\Theta_{i-1}^j)^{\gamma_t} p(\Theta_{i-1}^j)}{p(r_{1:T}|\Theta_{i-1}^j)^{\gamma_{i-1}} p(\Theta_{i-1}^j)} = W_{i-1}^j p(r_{1:T}|\Theta_{i-1}^j)^{\gamma_i - \gamma_{i-1}}, \;\; j = 1, ..., M \tag{5}$$

where the conditional likelihood $p(r_{1:T}|\Theta_{i-1}^j)$ is calculated exactly in the GARCH and GAS models, and approximately in the SV model using a particle filter. Then, set the new normalized weights,

$$W_i^j = \frac{w_i^j}{\sum_{s=1}^M w_i^s}, \;\; j = 1, ..., M. \tag{6}$$

Step 2: Compute the effective sample size (ESS):

$$\text{ESS} = \frac{1}{\sum_{j=1}^M \left(W_i^j\right)^2}. \tag{7}$$

if ESS $< cM$ for some $0 < c < 1$, **then**

(i) **Resampling**: Resampling from $\{\Theta_{i-1}^j\}_{j=1}^M$ using the weights $\{W_i^j\}_{j=1}^M$, and then set $W_i^j = 1/M$ for $j = 1...M$, to obtain the new equally-weighted particles $\{\Theta_i^j, W_i^j\}_{j=1}^M$.

(ii) **Markov move**: Parallel for each $j = 1,...,M$, move the samples Θ_i^j for N steps:

(a) Sample Θ_i^{j*} from the random walk proposal density $q(\Theta_i^{j*}|\Theta_i^j)$.

(b) Compute the proposal likelihood $p(r_{1:T}|\Theta_i^{j*})$ in GARCH and GAS models. In the SV model, we approximate the proposal likelihood with a correlated pseudomarginal method, see details in Deligiannidis et al. (2018).

(c) Set $\Theta_i^j = \Theta_i^{j*}$ with the probability

$$\min\left(1, \frac{p(r_{1:T}|\Theta_i^{j*})^{\gamma_i} p(\Theta_i^{j*})}{p(r_{1:T}|\Theta_i^j)^{\gamma_i} p(\Theta_i^j)} \frac{q(\Theta_i^j|\Theta_i^{j*})}{q(\Theta_i^{j*}|\Theta_i^j)}\right), \tag{8}$$

otherwise keep Θ_i^j, u_i^j unchanged.

end

3. The log of marginal likelihood estimate is

$$\log\widehat{p}(r_{1:T}) = \sum_{i=1}^K \log\left(\sum_{j=1}^M w_i^j\right). \tag{9}$$

B Macroeconmic variables

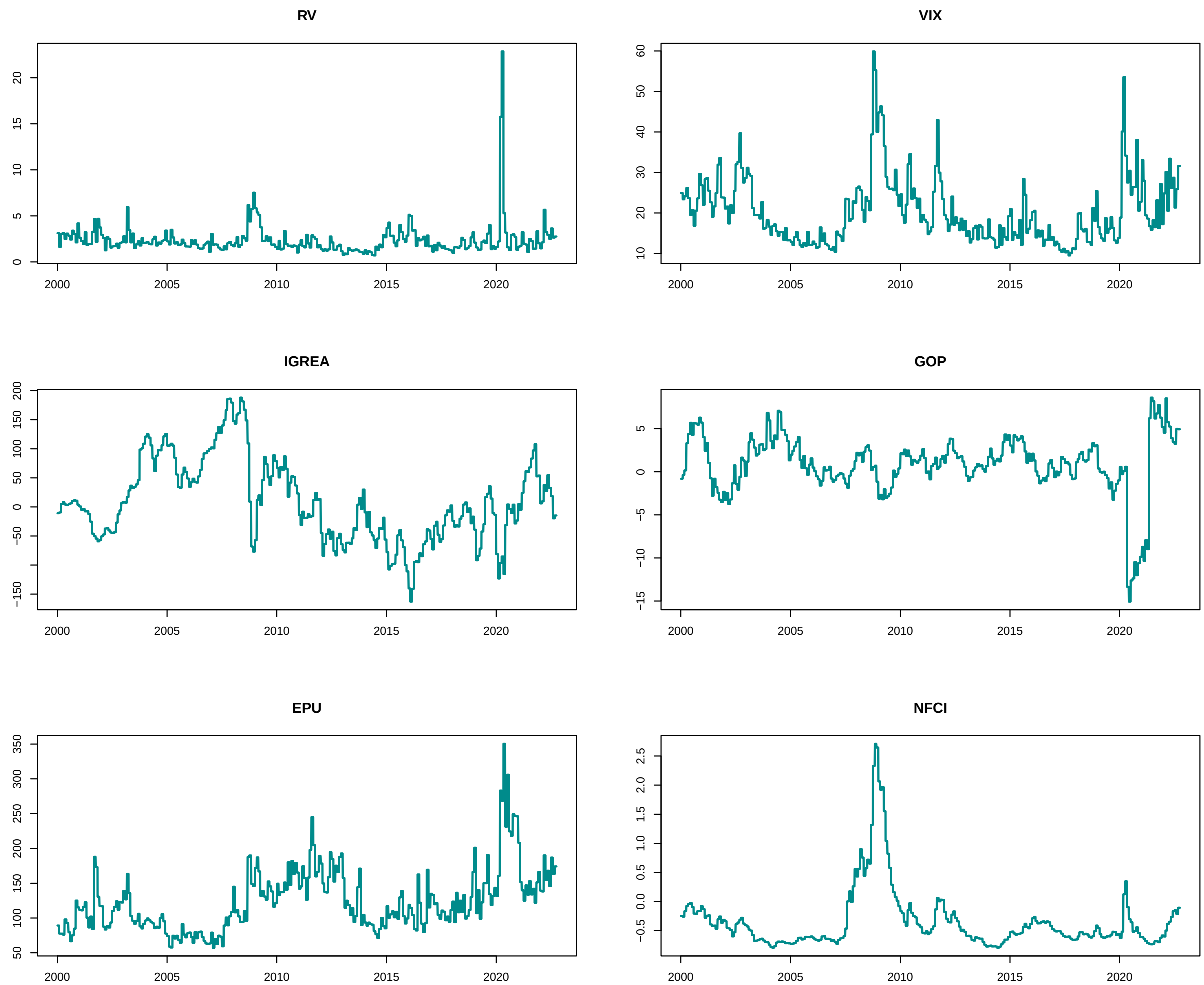

Figure 2: Plots for the macroeconomic variables from January 4, 2000 to September 30, 2022.

C Volatility components

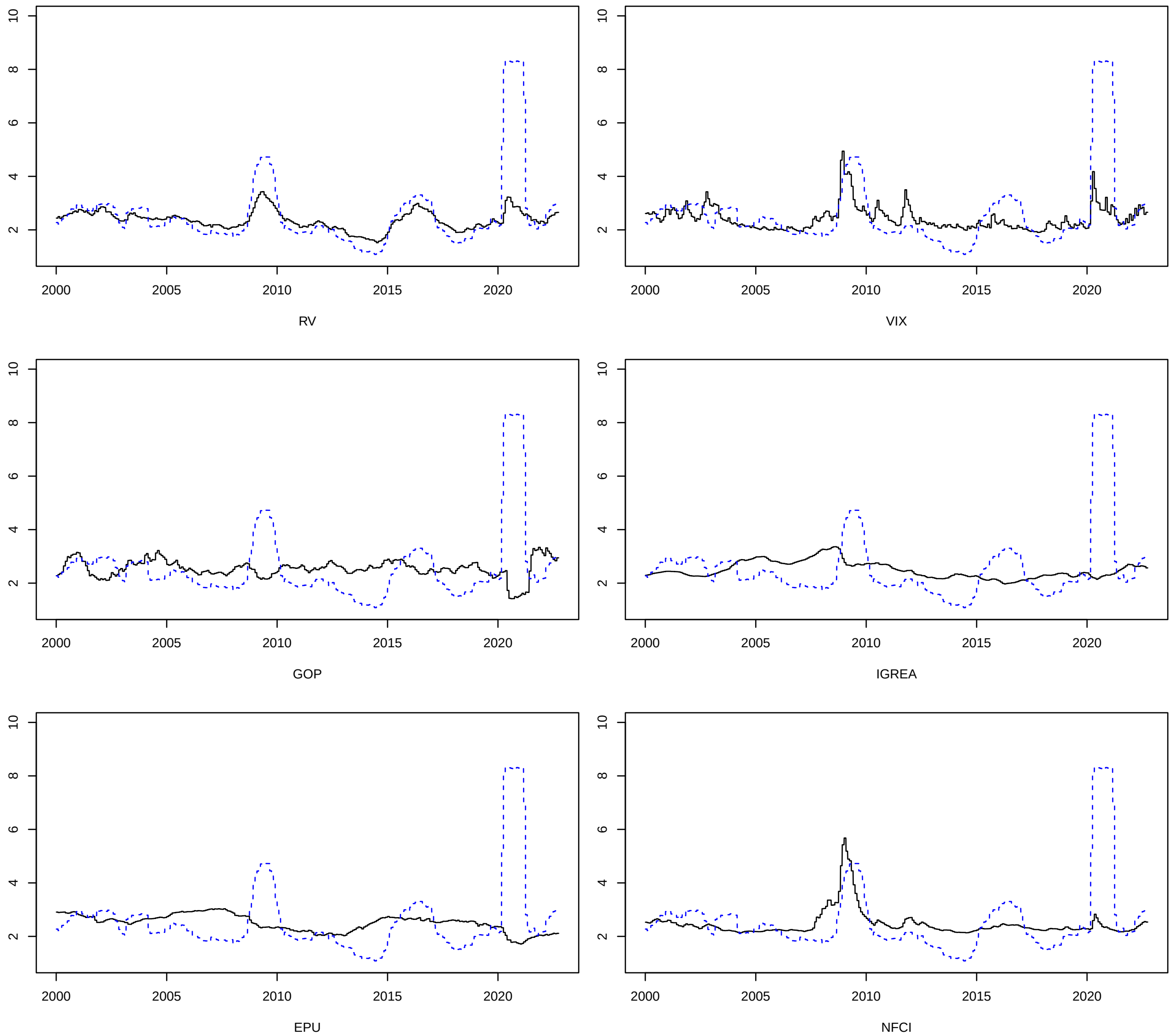

Figure 3: Long-term volatility components $\sqrt{\kappa_\tau}$ obtained from the GAS-MIDAS models with t-innovations and leverage effect.

The figures show the long-term volatility components obtained from the GAS-MIDAS models with t-innovations and leverage effect in the black solid line in comparison to the 1Y rolling realized volatility (calculated as a sample standard deviation of the daily returns in the last 12 months) in the blue dashed line.

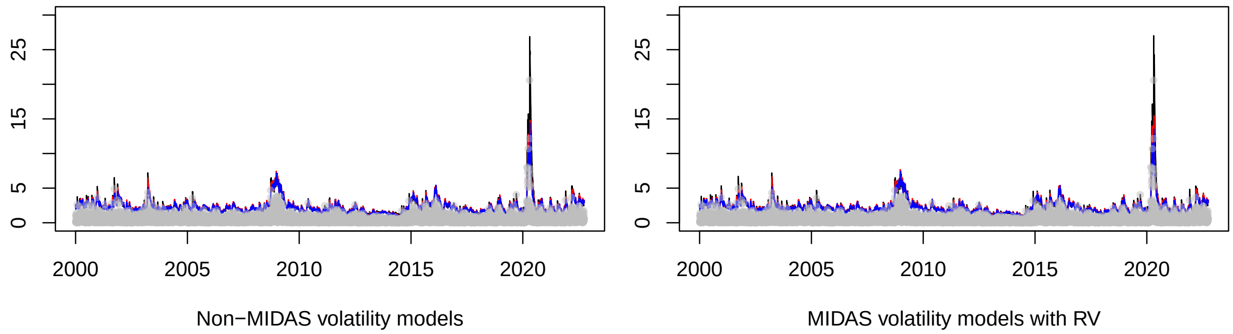

Figure 4: Total volatility components σ_t.

The figures show the total volatility components obtained from the GARCH, GAS, SV models in the black, red, and blue solid lines respectively. The scaled absolute returns are the grey dots.

D Density forecast

We compare the volatility forecast using the log predictive density score (LPS) and the continuous rank probability score (CRPS) of the posterior predictive distribution of the volatility forecast. Let's denote $r^o_{t+1:t+h} = \sum_{i=1}^h r_{t+i}$ as the cumulative returns during the h-step ahead at time t.

The LPS of the posterior predictive distribution is computed as,

$$\begin{aligned}
\text{LPS}_h &= \frac{1}{T_1 - T_0 - h + 1} \sum_{t=T_0}^{T_1-h} \left[\log p(r^o_{t+1:t+h} | \mathbf{r}_{1:t}) \right] \\
&= \frac{1}{T_1 - T_0 - h + 1} \sum_{t=T_0}^{T_1-h} \left[\log \int_{\boldsymbol{\Theta}} p(r^o_{t+1:t+h} | \boldsymbol{\Theta}, \mathbf{r}_{1:t}) p(\boldsymbol{\Theta} | \mathbf{r}_{1:t}) d\boldsymbol{\Theta} \right] \\
&\approx \frac{1}{T_1 - T_0 - h + 1} \sum_{t=T_0}^{T_1-h} \left[\log \sum_{j=1}^{M} p(r^o_{t+1:t+h} | \boldsymbol{\Theta}^j, \mathbf{r}_{1:t}) \right],
\end{aligned}$$

where $\boldsymbol{\Theta}^j$ for $j = 1, ..., M$ are posterior particles obtained from the in-sample data.

Gneiting and Raftery (2007) propose the CRPS as a probability score to compare the density forecasts. Clark and Ravazzolo (2015) noted the CRPS rewards more for the predictive distribution that is closer to the realization and less sensitive to outliers than the LP.

$$\text{CRPS}_h = \frac{1}{T_1 - T_0 - h + 1} \sum_{t=T_0}^{T_1-h} \left[-E_p \left| r_{t+1:t+h|t} + r^o_{t+1:t+h} \right| - 0.5 E_p \left| r_{t+1:t+h|t} - r'_{t+1:t+h|t} \right| \right],$$

where p is the predictive density of the variable $r_{t+1:t+h|t}$, and $(r_{t+1:t+h|t}, r'_{t+1:t+h|t})$ are independent random draws from the predictive density p. We use the Monte Carlo method to simulate 50,000 draws from the predictive density p and compute the expectation. Smaller CRPS means a better model for out-of-sample forecasts.

References

Alquist, R., Kilian, L. and Vigfusson, R. J. (2013), Forecasting the price of oil, *in* 'Handbook of economic forecasting', Vol. 2, Elsevier, pp. 427–507.

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

Baum, C. F. and Zerilli, P. (2016), 'Jumps and stochastic volatility in crude oil futures prices using conditional moments of integrated volatility', *Energy Economics* **53**, 175–181.

Baumeister, C. and Kilian, L. (2012), 'Real-time forecasts of the real price of oil', *Journal of business & economic statistics* **30**(2), 326–336.

Blasques, F., Koopman, S. J. and Lucas, A. (2015), 'Information-theoretic optimality of observation-driven time series models for continuous responses', *Biometrika* **102**(2), 325–343.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of econometrics* **31**(3), 307–327.

Bonnier, J.-B. (2022), 'Forecasting crude oil volatility with exogenous predictors: As good as it GETS?', *Energy Economics* **111**, 106059.

Carnero, M. A., Peña, D. and Ruiz, E. (2004), 'Persistence and kurtosis in garch and stochastic volatility models', *Journal of Financial Econometrics* **2**(2), 319–342.

Chan, J. C. and Grant, A. L. (2016), 'Modeling energy price dynamics: GARCH versus stochastic volatility', *Energy Economics* **54**, 182–189.

Charles, A. and Darné, O. (2017), 'Forecasting crude-oil market volatility: Further evidence with jumps', *Energy Economics* **67**, 508–519.

Chen, L., Zerilli, P. and Baum, C. F. (2019), 'Leverage effects and stochastic volatility in spot oil returns: A Bayesian approach with VaR and CVaR applications', *Energy Economics* **79**(3), 111–129.

Clark, T. E. and McCracken, M. W. (2001), 'Tests of equal forecast accuracy and encompassing for nested models', *Journal of econometrics* **105**(1), 85–110.

Clark, T. E. and Ravazzolo, F. (2015), 'Macroeconomic forecasting performance under alternative specifications of time-varying volatility', *Journal of Applied Econometrics* **30**(4), 551–575.

Conrad, C. and Kleen, O. (2020), 'Two are better than one: Volatility forecasting using multiplicative component garch-midas models', *Journal of Applied Econometrics* **35**(1), 19–45.

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

Creal, D., Koopman, S. J. and Lucas, A. (2011), 'A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations', *Journal of Business and Economic Statistics* **29**(4), 552–563.

Creal, D., Koopman, S. J. and Lucas, A. (2013), 'Generalized Autoregressive Score models with applications', *Journal of Applied Econometrics* **28**(5), 777–795.

Deligiannidis, G., Doucet, A. and Pitt, M. K. (2018), 'The correlated pseudomarginal method', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **80**(5), 839–870.

Diebold, F. X. and Mariano, R. S. (1995), 'Comparing predictive accuracy', *Journal of Business & Economic Statistics* **13**(3), 134–144.

Duan, J.-C. and Fulop, A. (2015), 'Density-tempered marginalized sequential Monte Carlo samplers', *Journal of Business & Economic Statistics* **33**(2), 192–202.

Duffie, D. and Pan, J. (1997), 'An overview of value at risk', *Journal of derivatives* **4**(3), 7–49.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation', *Econometrica: Journal of the econometric society* pp. 987–1007.

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

Engle, R. F. and Patton, A. J. (2007), What good is a volatility model?, *in* 'Forecasting volatility in the financial markets', Elsevier, pp. 47–63.

Engle, R. F. and Rangel, J. G. (2008), 'The spline-garch model for low-frequency volatility and its global macroeconomic causes', *Review of Financial Studies* **21**(3), 1187–1222.

Fissler, T. and Ziegel, J. F. (2016), 'Higher order elicitability and osband's principle', *The Annals of Statistics* **44**(4), 1680–1707.

Ghysels, E., Santa-Clara, P. and Valkanov, R. (2004), 'The MIDAS touch: Mixed data sampling regression models', *UCLA: Finance* pp. 1–33.

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

Gneiting, T. and Raftery, A. E. (2007), 'Strictly proper scoring rules, prediction, and estimation', *Journal of the American statistical Association* **102**(477), 359–378.

Gong, Y., Ma, C. and Chen, Q. (2022), 'Exchange rate dependence and economic fundamentals: A Copula-MIDAS approach', *Journal of International Money and Finance* **123**.

Hamilton, J. D. (1996), 'This is what happened to the oil price-macroeconomy relationship', *Journal of monetary economics* **38**(2), 215–220.

Harvey, A. C. (2013), *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*, Vol. 52, Cambridge University Press.

Harvey, A. C. and Shephard, N. (1996), 'Estimation of an asymmetric stochastic volatility model for asset returns', *Journal of Business & Economic Statistics* **14**(4), 429–434.

Harvey, A. and Chakravarty, T. (2008), 'Beta-t-(E)GARCH', *Cambridge Working Papers in Economics* **CWPE 0840**, 1–42.

Harvey, A. and Sucarrat, G. (2014), 'EGARCH models with fat tails, skewness and leverage', *Computational Statistics and Data Analysis* **76**, 320–338.

Johannes, M. S., Korteweg, A. and Polson, N. G. (2014), 'Sequential learning, predictability, and optimal portfolio returns', *The Journal of Finance* **69**(2), 611–644.

Kastner, G. and Frühwirth-Schnatter, S. (2014), 'Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models', *Computational Statistics & Data Analysis* **76**, 408–423.

Kilian, L. (2009), 'Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market', *American Economic Review* **99**(3), 1053–1069.

Kilian, L. and Murphy, D. P. (2014), 'The role of inventories and speculative trading in the global market for crude oil', *Journal of Applied econometrics* **29**(3), 454–478.

Kun, S. (2019), 'Unveiling the factors of oil versus non-oil sources in affecting the global commodity prices: A combination of threshold and asymmetric modeling approach', *Energy* **176**, 272–280.

Lee, Y. H. and Chiou, J. S. (2011), 'Oil sensitivity and its asymmetric impact on the stock market', *Energy* **36**(1), 168–174.

Lu, X., Ma, F., Wang, J. and Wang, J. (2020), 'Examining the predictive information of CBOE OVX on china's oil futures volatility: Evidence from MS-MIDAS models', *Energy* **212**.

Ma, F., Liu, J., Wahab, M. I. and Zhang, Y. (2018), 'Forecasting the aggregate oil price volatility in a data-rich environment', *Economic Modelling* **72**, 320–332.

Ma, Y., Ji, Q. and Pan, J. (2019), 'Oil financialization and volatility forecast: Evidence from multidimensional predictors', *Journal of Forecasting* **38**(6), 564–581.

Nguyen, H. and Javed, F. (2021), 'Dynamic relationship between stock market and bond market: A GAS MIDAS copula approach', *Working Paper, Örebro University School of Business, Örebro* **15**, 1–45.

Nguyen, H. and Virbickaite, A. (2022), 'Modeling stock-oil co-dependence with dynamic stochastic MIDAS copula models', *Scandinavian Working Papers in Economics* **5**.

Pan, Z., Wang, Y., Wu, C. and Yin, L. (2017), 'Oil price volatility and macroeconomic fundamentals: A regime switching GARCH-MIDAS model', *Journal of Empirical Finance* **43**, 130–142.

Shang, Y. and Liu, L. (2017), 'An extension of stochastic volatility model with mixed frequency information', *Economics Letters* **155**, 144–148.

Shang, Y. and Zheng, T. (2021), 'Mixed-frequency SV model for stock volatility and macroeconomics', *Economic Modelling* **95**, 462–472.

Taylor, J. W. (2019), 'Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric Laplace distribution', *Journal of Business & Economic Statistics* **37**(1), 121–133.

Taylor, S. J. (1982), 'Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961-75', *Time series analysis: theory and practice* **1**, 203–226.

Tran, M.-N., Scharth, M., Pitt, M. K. and Kohn, R. (2014), 'Annealed important sampling for models with latent variables', *arXiv preprint arXiv:1402.6035* .

Wang, H. and Li, S. (2021), 'Asymmetric volatility spillovers between crude oil and china's financial markets', *Energy* **233**, 121168.

Wei, Y., Liu, J., Lai, X. and Hu, Y. (2017), 'Which determinant is the most informative in forecasting crude oil market volatility: Fundamental, speculation, or uncertainty?', *Energy Economics* **68**, 141–150.

Xu, Y. and Lien, D. (2022), 'Forecasting volatilities of oil and gas assets: A comparison of GAS, GARCH, and EGARCH models', *Journal of Forecasting* **41**(2), 259–278.

Yu, J. (2005), 'On leverage in a stochastic volatility model', *Journal of Econometrics* **127**(2), 165–178.