WORKING PAPER 12/2025 (STATISTICS)

Fast and Slow Level Shifts in Intraday Stochastic Volatility

Igor F. B. Martins, Audronè Virbickaitè, Hoang Nguyen and Hedibert Freitas Lopes

ISSN 1403-0586

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

Fast and Slow Level Shifts in Intraday Stochastic Volatility

Igor F. B. Martins [(a)]*, Audronė Virbickaitė[(b)], Hoang Nguyen[(c)], Hedibert Freitas Lopes[(d)]

* Corresponding author

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

[(b)] Department of Quantitative Methods, CUNEF Universidad, Madrid, Spain

[(c)]Department of Management and Engineering, Linköping University, Sweden

[(d)] Insper Institute of Education and Research, São Paulo, Brazil,

November 7, 2025

*Corresponding author. Email: igor.ferreira-batista-martins@oru.se. Address: Örebro University, Fakultetsgatan 1, 702 81, Örebro, Sweden.

Abstract

This paper proposes a mixed-frequency stochastic volatility model for intraday returns that captures fast and slow level shifts in the volatility level induced by news from both low-frequency variables and scheduled announcements. A MIDAS component describes slow-moving changes in volatility driven by daily variables, while an announcement component captures fast event-driven volatility bursts. Using 5-minute crude oil futures returns, we show that accounting for both fast and slow level shifts significantly improves volatility forecasts at intraday and daily horizons. The superior forecasts also translate into higher Sharpe ratios using the volatility-managed portfolio strategy.

Keywords: Intraday volatility, high-frequency, announcements, MIDAS, oil, sparsity.

JEL Classification: C22, C52, C58, G32

1 Introduction

Understanding and forecasting volatility is central to modern financial econometrics as it has direct implications for risk management, option pricing, and portfolio allocation. Traditional intraday volatility models focus on persistence and intraday seasonal effects, which have proven to be effective in modeling short-term volatility dynamics. However, volatility in financial markets evolves over multiple time scales, reflecting both slow-moving changes in economic conditions and immediate responses to new information releases. As a result, traditional intraday models fall short in effectively describing volatility, since they either fail to capture long-run shifts in volatility or miss fast volatility spikes due to new information being released. Engle and Patton (2007) suggest that a good volatility model should include an extension of the available information set as market participants continuously revise their expectations and uncertainty according to the arrival of information. Proper representation of these multi-scale effects is crucial for improving modeling and forecasting of volatility.

In this paper, we propose a mixed-frequency stochastic volatility model for intraday returns that captures both slow and fast level shifts in volatility. The model extends the standard intraday stochastic volatility model by introducing two complementary components: a slow-moving component based on mixed-frequency data sampling (MIDAS, Ghysels et al., 2004) that allows low-frequency variables to change the volatility level, and an announcement component that accounts for abrupt volatility bursts due to scheduled announcements. By explicitly recognizing that volatility can shift at different speeds according to information sources, our model provides a unified framework for modeling both short- and long-term volatility dynamics.

Our contributions are both methodological and empirical. Methodologically, this paper extends traditional intraday stochastic volatility models by introducing fast and slow level shifts. Our fast level shift component is based on a large set of scheduled announcements with sparsity-inducing spike-and-slab priors that help to identify meaningful volatility responses while setting the effects of the uninformative events to zero. Our slow level shift component uses MIDAS to project the low-frequency variable into a high-frequency space allowing the model to account for changes in market conditions. For estimation, we develop a Markov chain Monte Carlo (MCMC) algorithm to obtain samples from the posterior distributions of model parameters and latent volatility states.

Empirically, we evaluate the performance of the proposed model using the crude oil market data, a setting well-suited for studying volatility level shifts given their sensitivity to both macroeconomic conditions and scheduled inventory announcements. Using 5-minute crude oil futures returns, paired with daily measures of market uncertainty as low-frequency inputs, and Bloomberg event calendar data with exact release timestamps, we show that modeling fast and slow volatility level shifts leads to superior forecasts at both 5-minute and day-ahead horizons. Our proposed model achieves superior performance in Mincer–Zarnowitz and horse-race regressions and delivers lower forecast errors compared to benchmark models in Diebold–Mariano tests. Moreover, volatility-managed portfolios based on our forecasts yield higher Sharpe ratios. The results show that modeling level shifts operating at different temporal frequencies is key to not only understanding the driving forces of the intraday volatility movements, but also to successful one- and multiple-step-ahead forecasting.

This paper contributes to a large literature on modeling intraday volatility. Early studies such as Andersen and Bollerslev (1997), Deo et al. (2006), Andersen et al. (2007), Engle and Sokalska (2012), and Rossi and Fantazzini (2015) highlight persistence and intraday seasonality as key features of volatility at high frequency. Persistence can be captured through the autoregressive component in GARCH, stochastic volatility (SV), or realized volatility (RV) models. Empirically, SV models have been shown to outperform GARCH and RV methods in out-of-sample forecasting, as documented by Stroud and Johannes (2014), Chan and Grant (2016), and Martins et al. (2025), among others. Therefore, in this paper we adopt the SV specification.

Closer to our setting are studies that model either fast or slow components of volatility. In addition to persistence and intraday seasonal patterns, Andersen and Bollerslev (1998), Stroud and Johannes (2014), and Martins and Lopes (2024) examine the effects of scheduled announcements, which represent fast level shifts in volatility. Andersen and Bollerslev (1998) and Stroud and Johannes (2014) assume that a fixed set of announcements always affects volatility, while Martins and Lopes (2024) consider a broader set and employ sparsity-inducing priors to estimate both the probability and the magnitude of each event's impact. However, these studies do not account for slow-moving changes in volatility driven by low-frequency variables, which are central to our framework.

Alternatively, Bekierman and Gribisch (2021) propose a mixed-frequency SV model for equities

that features a slow-moving component driven by a latent autoregressive process. The authors apply the model to a dataset of 30-minute returns, however, they do not consider exogenous information in either the fast or the slow level shifts. Our paper differs from theirs in several important ways. First, we incorporate an event component based on Bayesian variable selection, allowing us to identify volatility responses to scheduled announcements. Second, our slow-moving component is based on MIDAS regressions using daily variables, offering greater economic interpretability than their latent autoregressive process. Third, our empirical application uses 5-minute returns traded nearly 24 hours a day, compared with 30-minute intervals during regular trading hours in their study. This higher frequency allows us to capture short-lived volatility spikes that may be missed in lower-frequency frameworks. To our knowledge, no existing study combines the high-frequency stochastic volatility framework with a MIDAS-based level component and an explicit treatment of event-driven volatility shifts.

Our paper also relates to a large literature studying oil volatility and its determinants. Haugom et al. (2014) and Žikeš and Baruník (2015) combine RV with implied volatility indices such as VIX and OVX to forecast daily oil volatility finding that these indices improve forecast performance. Therefore, we consider VIX and OVX as potential sources of information for our slow level shift component. Several papers also study the potential effect of a type of announcement on oil volatility. Bu (2014) and Bjursell et al. (2015) show that U.S. Energy Information Administration (EIA) weekly crude oil inventory announcements impact the volatility of oil futures. Yang et al. (2023) studies the impact of monetary policy announcements in an SVAR framework. While Känzig (2021) and Schmidbauer and Rösch (2012) show that OPEC announcements affect oil volatility, Noguera-Santaella (2016) and Brandt and Gao (2019) indicate that geopolitical news has little to no impact on volatility for the last 20 years. Rather than selecting a small set of announcements *ex-ante*, we consider a large set of events and employ sparsity-inducing spike-and-slab priors to estimate both the probability that an event affects volatility and the size of its effect when included.

By combining these strands of research, our paper bridges the methodological literature on intraday and mixed-frequency volatility models with the empirical literature on oil market dynamics. We offer a unified approach that captures both fast and slow level shifts in volatility, linking high-frequency market reactions to scheduled announcements with slower-moving shifts driven by macro-financial conditions, and demonstrate its relevance through an application to crude oil prices,

which exhibit particularly rich volatility dynamics.

The rest of the paper is organized as follows. Section 2 introduces the stochastic volatility model with fast and slow level shifts. Section 3 details our Bayesian inference algorithm for the proposed model. The data is presented in Section 4, meanwhile the estimation results are in Section 5. Section 6 contains the forecasting and portfolio allocation applications. Finally, conclusions are drawn in Section 7.

2 Econometric model

Let $\mathbf{y} = \{y_t\}_{t=1}^T$ denote a series of de-meaned 5-minute (log-)returns, and let $\mathbf{h} = \{h_t\}_{t=1}^T$ represent the corresponding latent log-volatility process. We assume that returns have zero-mean and time-varying volatility:

$$y_t = \exp\left(\frac{h_t}{2}\right)\epsilon_t, \qquad \epsilon_t \sim \mathbf{N}(0,1) \text{ for } t = 1,\ldots,T, \tag{1}$$

where the log-variance h_t is decomposed into four additive components:

$$h_t = m_\tau + e_t + p_t + s_t, \tag{2}$$

with τ denoting the day that contains 5-minute intervals indexed by timestamp t [1]; and the first two terms, m_τ and e_t, capture the slow and fast level shifts in the volatility level, while p_t and s_t represent standard persistence and intraday seasonality components widely used in the high-frequency volatility literature.

We allow the daily volatility level, m_τ, to slowly evolve over time capturing shifts in volatility level according to changing macro-financial conditions as discussed in Haugom et al. (2014) and Žikeš and Baruník (2015), for example. Specifically, we define:

$$m_\tau = m_0 + \sum_{j=1}^{J} \delta_j \left[\sum_{l=1}^{L_j} \phi_l(w_j) X_{j,\tau-l}\right], \tag{3}$$

where $\{X_{j,\tau}\}_{j=1}^J$ denotes a set of J low-frequency variables at day τ; For each variable j, $\phi_l(w_j)$

[1] Therefore, $t = \tau K + 1, \ldots, \tau K + K$ with K representing the number of 5-minute observations in a day τ.

is a weighting function parameterized by w_j that assigns weights to daily lags $X_{j,\tau-l}$ of variable X_j, and δ_j measures the sensitivity of the daily volatility level from the available information of X_j. The number of lags L_j is determined a priori and exhibits little sensitivity to the estimation results. Following Virbickaitė et al. (2023), we employ a restricted beta weighting scheme:

$$\phi_l(w_j) = \frac{\left[1 - \frac{l}{L_j+1}\right]^{w_j-1}}{\sum_{m=1}^{L_j}\left[1 - \frac{m}{L_j+1}\right]^{w_j-1}}, \tag{4}$$

which enforces smoothness and non-negativity while ensuring weights sum to one. If the low-frequency variables are standardized, the unconditional long-run mean of log-volatility equals to m_0, and m_τ represents the conditional daily deviations induced by the low-frequency variables.

To capture fast level shifts in volatility caused by scheduled announcements, we introduce an event component:

$$e_t = \sum_{i=1}^{N} I_{it}\alpha_i, \tag{5}$$

where N is the total number of possible events, I_{it} is an indicator equal to one if event i occurs at time t, and α_i measures the corresponding impact of fast level shifts on log-variance. Since there is a big number of possible events but likely only a small subset is expected to significantly affect volatility, we impose a sparsity-inducing spike-and-slab prior on α_i. This prior structure enforces sparsity by setting uninformative announcement coefficients to zero, while retaining non-zero effects for influential events, thereby identifying volatility-relevant announcements in a data-driven way.

The remaining components in Equation (2) capture the persistence and intraday seasonality commonly exhibited in intraday volatility. To model short-run volatility persistence, we let p_t follow an AR(1) process:

$$p_t = \beta p_{t-1} + \sigma_\eta \eta_t, \qquad \eta_t \sim \mathbf{N}(0,1), \tag{6}$$

where β captures the persistence and σ_η scales the innovation variance. Intraday seasonality is represented by:

$$s_t = \sum_{k=1}^{K} H_{tk}\beta_k, \qquad \sum_{k=1}^{K}\beta_k = 0, \tag{7}$$

where H_{tk} is an indicator equal to one if time t corresponds to the k-th 5-minute interval of the

trading day, and β_k captures the average deviation in volatility associated with that interval. To ensure identification, we impose the zero-sum restriction $\sum_{k=1}^{K} \beta_k = 0$, which centers the seasonal pattern around zero.

Equations (1)–(7) define our high-frequency stochastic volatility model featuring both slow and fast level shifts in volatility levels. The slow level-shift component, m_τ, extends the high-frequency SV models of Stroud and Johannes (2014) and Martins and Lopes (2024) by allowing volatility to adapt to macro-financial information through a MIDAS-type structure. Conversely, the event-driven component, e_t, expands on the mixed-frequency SV model of Bekierman and Gribisch (2021) by introducing a sparsity-based identification of announcement effects. Together, these components provide a unified structure for analyzing volatility dynamics across multiple time scales.

Table 1 summarizes how our proposed specification (SSVA MIDAS) nests several stochastic volatility models as special cases. These include (i) a vanilla SV model (SV), (ii) an SV with intraday seasonality effects (SSV), (iii) an SV with with intraday seasonality and accounting for announcement effects considering a sparsity inducing approach (SSVA) or not accounting for sparsity by using a simple Gaussian prior (SSVAg). These benchmarks serve as natural points of comparison in our empirical analysis.

Table 1: Model typology

	Persistence (p_t)	Seasonal (s_t)	Fast shift (e_t)	Slow shift (m_τ)
SSVA MIDAS	✓	✓	✓	✓
SSVA	✓	✓	✓	
SSVAg	✓	✓	✓	
SSV	✓	✓		
SV	✓			

The first column lists the model mnemonics, and the remaining columns indicate which components are included: a persistent stochastic volatility component (p_t), intraday seasonality (s_t), fast level shifts due to scheduled events (e_t), and slow level shifts (m_τ). Our proposed model includes all components, while the other specifications represent restricted variants.

3 Bayesian inference

We adopt a Bayesian approach to estimate the model described in Section 2 by Equations (1) - (7). Our approach jointly estimates parameters and latent states which allows for direct probabilistic statements on both parameters and latent volatility dynamics. Since the model's likelihood involves high-dimensional latent states that cannot be integrated out in closed form, we rely on MCMC methods to generate draws from the joint posterior distribution of parameters and latent processes.

We employ the log-linearization approach of Kim et al. (1998) to facilitate likelihood computation. We rewrite Equation (1) as

$$\log y_t^2 = h_t + u_t, \tag{8}$$

where $u_t = \log \epsilon_t^2 \sim \log \chi^2(1)$. Following Kim et al. (1998), we approximate the distribution of u_t using a 7-component Gaussian mixture. Conditionally on the mixture indicators, our model can be expressed in a Gaussian state-space formulation which allows us to apply the forward-filtering backward sampling (FFBS) algorithm (Carter and Kohn, 1994; Frühwirth-Schnatter, 1994) for block updating of the latent states.

Having specified the likelihood in Equations (1) - (8), we next detail the prior structure completing the Bayesian formulation. Due to the large number of announcements, we consider Bayesian regularization technique by employing spike-and-slab priors for scheduled announcement coefficients $\{\alpha_i\}_{i=1}^N$. We model the spike as a Dirac delta on zero and the slab as a Normal distribution with a large variance, see Equation (9a), similarly to Mitchell and Beauchamp (1988). Although both distributions place zero within their support, the Dirac delta's mass concentrated entirely at zero implies that announcements with no effect on volatility are much more likely to originate from this component, whereas the Gaussian component allows for announcements to impact volatility in a relatively unrestricted manner. The probability of the announcement being drawn from the either the spike or the slab comes from the realizations of a Bernoulli distribution with a hyper-parameter γ:

$$\alpha_i | \pi_i \sim (1 - \pi_i)\delta_0 + \pi_i \mathbf{N}(0, \sigma_\alpha^2), \tag{9a}$$

$$\pi_i \sim \mathbf{Bern}(\gamma). \tag{9b}$$

We assume weakly informative priors for all remaining parameters. We attribute Gaussian priors for β, $\{\delta_j\}_{i=1}^{J}$ and $\{\beta_k\}_{k=1}^{K-1}$. Due to the high persistence commonly observed in the stochastic volatility literature, we assume a prior $\beta \sim \mathbf{N}(\beta_0 = 0.95, V_{\beta_0} = 0.5^2)$. We center all $\{\beta_k\}_{k=1}^{K-1}$ and $\{\delta_j\}_{i=1}^{J}$ on 0, implying a prior belief of no seasonal effects and no spill-over effects of exogenous low-frequency variables, but we allow for variability by considering a $\mathbf{N}(\boldsymbol{\beta}_{0k} = 0, \mathbf{V}_{\boldsymbol{\beta}_{0k}} = 0.5\mathbf{I}_{K-1})$ and $\mathbf{N}(\boldsymbol{\delta}_0 = 0, \mathbf{V}_{\boldsymbol{\delta}_0} = 2\mathbf{I}_J)$ prior distributions respectively. We assume Inverse Gamma priors $\mathbf{IG}(\alpha_{\sigma_\eta} = 5, \beta_{\sigma_\eta} = 1)$ and $\mathbf{IG}(\alpha_\alpha = 1, \beta_\alpha = 10)$ for σ_η^2 and σ_α^2 reflecting the restriction of non negative variance. We attribute a $\mathbf{Beta}(\alpha_\gamma = 1, \beta_\gamma = 1)$ prior for γ reflecting no prior knowledge if the announcement effect comes from the spike or the slab. The prior of w_j is assumed to be a uniform distribution $\mathbf{U}(1, 20)$. We also assume the mean level $m_0 \sim \mathbf{N}(\mu_{m_0} = y^*, V_{m_0} = 2)$ where $y^* = \frac{1}{T}\sum_{t=1}^{T} \log y_t^2 + 1.27$. Our choice of priors greatly simplifies the Gibbs sampling scheme detailed in Appendix A at the end of the manuscript.

4 Data

We use high-frequency data on West Texas Intermediate (WTI) crude oil futures traded on the New York Mercantile Exchange (NYMEX) covering the period from January 3, 2016 to August 30, 2024. The sample is divided into an in-sample period up to the last day of 2021 yielding more than 400 thousand observations and an out-of-sample period for the remainder of the sample consisting of more than 180 thousand observations. All timestamps are recorded in Eastern Time (New York). Trading begins at 18:00 on Sunday and continues until 17:00 on Friday, with a one-hour maintenance break each day between 17:00 and 18:00. A 24-hour day therefore consists of 288 five-minute intervals, although daylight saving time adjustments occasionally produce 23-hour or 25-hour days. To construct a continuous price series, we link the most liquid nearby futures contracts, applying a ratio adjustment at each rollover to correct for price differentials between expiring and new contracts. Non-trading intervals are filled using the last available transaction price to maintain a regular five-minute grid. One exception is when an official trading break occurs. In such cases, all timestamp during the break are removed from our sample. We then compute returns based on these adjusted prices. Finally, to be consistent with the zero-mean assumption in our econometric model, we de-mean returns using the in-sample mean only. Appendix B presents

some summary statistics of fast and slow moving components in the intraday volatility.

Information on scheduled macroeconomic and commodity-related announcements is obtained from the Bloomberg Economic Calendar. We focus on announcements pertaining to the United States (US), China (CH), and Germany (GE), with the latter serving as a proxy for the broader euro area. Two main event categories are considered: macroeconomic releases, such as inflation, output, and monetary policy indicators, and commodity-related announcements such as energy inventories. Only announcements with an explicit release timestamp are retained. Events lacking time information are excluded, while announcements whose release times do not coincide with a five-minute interval are rounded up to the appropriate five-minute mark to align with the frequency of the return data. To ensure sufficient recurrence, we keep only announcements that occur at least twice per year within the sample period. When multiple announcements share identical timestamps across the sample, we retain only the first occurrence to avoid perfect multicolinearity [2]. After this filtering procedure, the dataset contains 94 unique scheduled announcements spanning the three economies. A complete list of retained announcements is provided in the Appendix C.

To capture slow-moving shifts in the volatility level, we incorporate two daily financial volatility indices in scaled log-levels: the CBOE Crude Oil Volatility Index (OVX) and the CBOE Market Volatility Index (VIX). Both series are obtained from Bloomberg and are sampled at the daily frequency to match the mixed-frequency structure of the model. The OVX reflects the market's expectation of future oil price volatility implied by options on WTI crude oil futures, while the VIX represents expected stock market volatility derived from S&P 500 index options. Previous studies have shown that these measures contain valuable information for forecasting oil market volatility (Haugom et al., 2014; Žikeš and Baruník, 2015; Niu et al., 2022), motivating their inclusion as low-frequency drivers in our MIDAS component.

5 In-sample results

In this section, we analyze the in-sample estimation results and the role of the volatility components as specified in Equation (2). We put a special focus on the contribution of the distinct volatility components to the overall variation in the latent log-variance.

[2]For example, "GE CPI month-over-month (MoM)" and "GE CPI year-over-year (YoY)" always occur on the same timestamp over the sample.

5.1 Slow level-shifts

Table 2 reports the variance decomposition of the latent log-variance (h_t) at the posterior parameter modes, into its slow (m_τ) and fast level (e_t) shifts, persistent (p_t) and intraday seasonality (s_t) components[3]. Entries are reported as percentage shares.

Table 2: Variance decomposition

	m_τ	e_t	p_t	s_t
SSVA MIDAS^{ovx,vix}	39.07	0.68	20.28	39.98
SSVA MIDAS^{ovx}	38.62	0.69	20.71	39.98
SSVA MIDAS^{vix}	25.60	0.68	34.14	39.58

The table reports the variance decomposition of the log-variance, h_t, at the posterior modes into its four components: slow (m_τ) and fast (e_t) level shifts, latent persistence (p_t) and intraday seasonality (s_t). Entries are percentage shares.

Across all specifications, the slow-moving MIDAS component accounts for a substantial share of the variation in the latent log-variance. Its contribution ranges from 25.6% in the VIX-only specification to around 39% in the SSVA MIDAS^{ovx} and SSVA MIDAS^{ovx,vix} models, indicating that low-frequency shifts play a central role in capturing long-run movements in volatility. The announcement component contributes less than 1% of the total variance, a result consistent with the sparse timing of scheduled events relative to the high-frequency sampling of intraday volatility. The seasonal component exhibits a remarkably stable share across all models, nearly 40%, reflecting the robustness of intraday periodicity as a key feature of volatility dynamics.

The pattern across model specifications in Table 2 is informative regarding the role of the MIDAS variables in capturing low-frequency volatility movements. Specifications including OVX allocate more variation to the slow-moving component m_τ and correspondingly less to the short-run persistent term p_t. For instance, the share associated with m_τ rises from roughly 25% in the SSVA MIDAS^{vix} model to nearly 39% in the SSVA MIDAS^{ovx,vix} case, while the contribution of p_t declines from 34% to nearly 20%. This inverse relation reflects how the inclusion of informative low-frequency predictors allows the MIDAS component to absorb long-horizon variation that would otherwise be attributed to short-run persistence. As noted in Bekierman and Gribisch (2021), introducing a long-run component helps capture the long-memory behavior in intraday volatility

[3]We compute the total variability as the sum of the variances of each component plus their pairwise covariances; and the share of variability as the proportion of the component contribution to the total variability.

processes. When the MIDAS term is omitted, the model must mimic this long-memory feature by inflating the contribution associated with volatility persistence.

Figure 1 reports the evolution of the slow-moving component to illustrate how it captures daily shifts in volatility. For example, during the onset of the COVID-19 pandemic in early 2020, m_τ rises substantially, mirroring the sustained increase in market volatility observed in OVX and VIX measures. This pattern indicates that the slow-moving factor effectively tracks long-horizon risk dynamics that are not fully reflected in high-frequency announcement effects or intraday seasonal variation. Consequently, m_τ provides both a statistically robust and economically interpretable representation of low-frequency volatility shifts in the market.

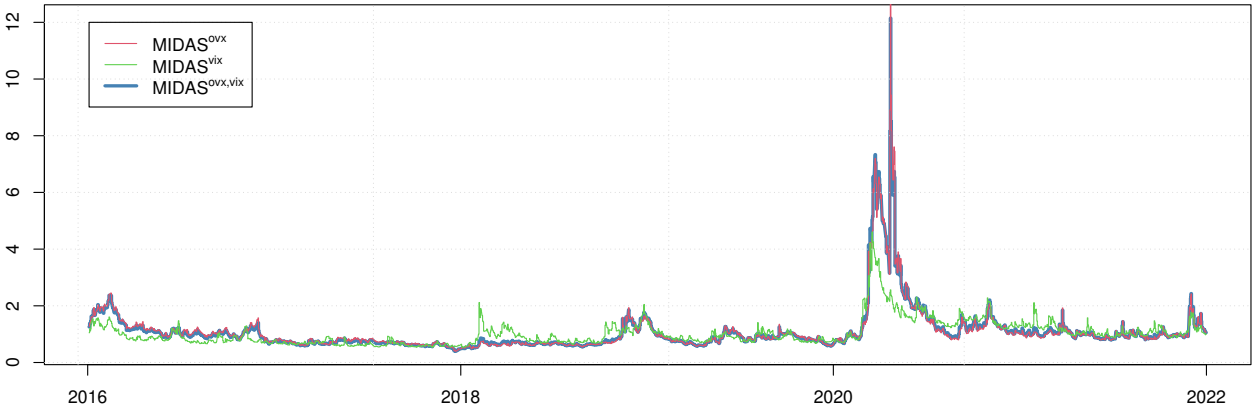

Figure 1: The evolution of MIDAS component in terms of standard deviation $\exp(m_\tau/2)$ for the in-sample period, estimated at the posterior modes of model parameters.

5.2 Fast level-shifts

While Table 2 reports the unconditional variance decomposition, Table 3 presents the corresponding results conditional on five-minute intervals containing at least one scheduled announcement. This conditioning reveals a sharp increase in the share of variance explained by the fast level-shift component (e_t), from below 1% in the unconditional decomposition to roughly 23% during the time windows when at least one event occurs. The announcement-driven component thus becomes a major source of short-term volatility variation, consistent with the idea that discrete information releases temporarily dominate intraday price dynamics. The relative importance of the slow-moving (m_τ) and persistent (p_t) components declines, while the intraday seasonal share remains stable near 30%.

Table 3: Conditional variance decomposition

	m_τ	e_t	p_t	s_t
SSVA MIDAS^{ovx,vix}	30.19	22.47	17.09	30.24
SSVA MIDAS^{ovx}	29.71	22.65	17.33	30.31
SSVA MIDAS^{vix}	19.27	22.75	27.19	30.79

This table reports the variance decomposition of the log-variance, h_t, into its four components when at least one event occurs: slow (m_τ) and fast (e_t) level shifts, latent persistence (p_t) and intraday seasonality (s_t). The model parameters are fixed in their posterior modes. Entries are percentage shares.

Naturally, the magnitude of these fast shifts depends on the specific set of announcements included in the model and their estimated importance. Rather than pre-selecting a small number of events or relying on a naive OLS-based identification, we employ a sparsity-inducing spike-and-slab approach allowing the data to determine which events meaningfully impact volatility. This approach accommodates a broad range of candidate announcements while ensuring that only the most influential events contribute to the fast level-shift component e_t.

Table 4 reports the estimated announcement effects obtained from both linear regression of log squared returns using the ordinary least squares (OLS) method and the spike-and-slab specification for the full SSVA MIDAS^{ovx,vix} model. The first two columns report conventional OLS coefficients and p-values, while the remaining columns show posterior information and inclusion probabilities from the spike-and-slab posterior. The contrast between the two methods is striking. The OLS estimates attribute significance to a relatively large number of announcement releases, including several without plausible economic links to oil-market volatility. For instance, the "U.S. Condition Cotton" event, not only is included with low p-value but also is estimated to have a negative impact on volatility. In contrast, the spike-and-slab procedure shrinks weak coefficients to zero, retaining only those with high posterior inclusion probabilities.

Using the SSVA MIDAS^{ovx,vix} model, only four announcements display non-zero posterior modes and inclusion probabilities near one: the U.S. Change in Nonfarm Payrolls, the FOMC Rate Decision, the DOE Crude Oil Inventories release, and the Sunday market open. The Nonfarm Payrolls and FOMC Rate Decision releases convey critical information about macroeconomic activity, monetary policy, and financial conditions. Both announcements tend to amplify uncertainty and risk revaluation, thereby increasing intraday oil price volatility being consistent with the ev-

Table 4: OLS and MCMC estimation results for the events

	OLS estimate	OLS p-value	Post. median	Post mode	Post. inclusion
GE GfK Consumer Confidence	-1.69	0.01	0.00	0.00	0.04
GE HCOB Germany Services PMI	1.13	0.03	0.00	0.00	0.02
GE ZEW Survey Current Situation	-1.31	0.04	0.00	0.00	0.18
US Change in Nonfarm Payrolls	2.01	0.00	1.24	1.11	1.00
US Condition Cotton	-1.08	0.09	0.00	0.00	0.12
US CPI MoM	0.92	0.17	0.61	-0.00	0.90
US DOE Crude Oil Inventories	2.83	0.00	2.29	2.27	1.00
US FOMC Rate Decision	1.53	0.05	1.18	1.27	1.00
US ISM Manufacturing	1.23	0.06	0.00	0.00	0.16
US U. of Michigan Sentiment	-0.17	0.72	-0.48	-0.00	0.84
Sunday Open	2.75	0.00	2.12	2.09	1.00

The table presents the OLS announcement coefficients and the corresponding p-values of regression $\log(y_t^2) \sim \log(y_{t-1}^2) + s_t + e_t$, the MCMC posterior announcement coefficients (median and mode) using the SSVA MIDAS^{ovx,vix} model and posterior inclusion probabilities.

idence in Basistha and Kurov (2015), López (2018), and Yang et al. (2023). On the supply side, the DOE Crude Oil Inventories release communicates short-term signals about market tightness and the supply side of oil with similar findings being reported in Bjursell et al. (2015). Finally, the Sunday open captures a structural calendar effect reflecting the systematic volatility spike that occurs when markets reopen after the weekend closure and assimilate information accumulated during the non-trading period. Appendix D at the end of the manuscript provides further details regarding the posterior distributions for the estimated event effects.

Taken together, these results show that fast level-shifts in intraday oil volatility are both sparse and economically interpretable. Only a small subset of high-impact announcements generates distinct volatility bursts, while the remainder of the process is governed by persistent dynamics, seasonal patterns, and slow-moving market-wide uncertainty. This decomposition highlights how the proposed model successfully disentangles the heterogeneous temporal drivers of oil market volatility.

5.3 Intraday seasonality, persistence and model parameters

Figure 2 shows the estimated intraday seasonal component of volatility on the standard deviation scale, $\exp\left(\frac{s_t}{2}\right)$, along with 90% credible intervals. The figure reveals a pronounced U-shaped

profile between 9:00 and 14:30 corresponding to the historical pit-trading hours of the NYMEX. Volatility peaks sharply at the opening and closing of this interval, reflecting bursts of trading activity and information arrival when market participation is highest, while mid-session periods exhibit a temporary decrease. This structure is a well-documented feature of intraday volatility and remains robust across specifications. In addition, the relative magnitude of volatility differs systematically across trading regions: the U.S. market exhibits the highest volatility levels, followed by the European session, with the Asian hours showing the lowest overall activity. These cross-regional differences highlight the dominant influence of U.S. market hours on global oil price formation.

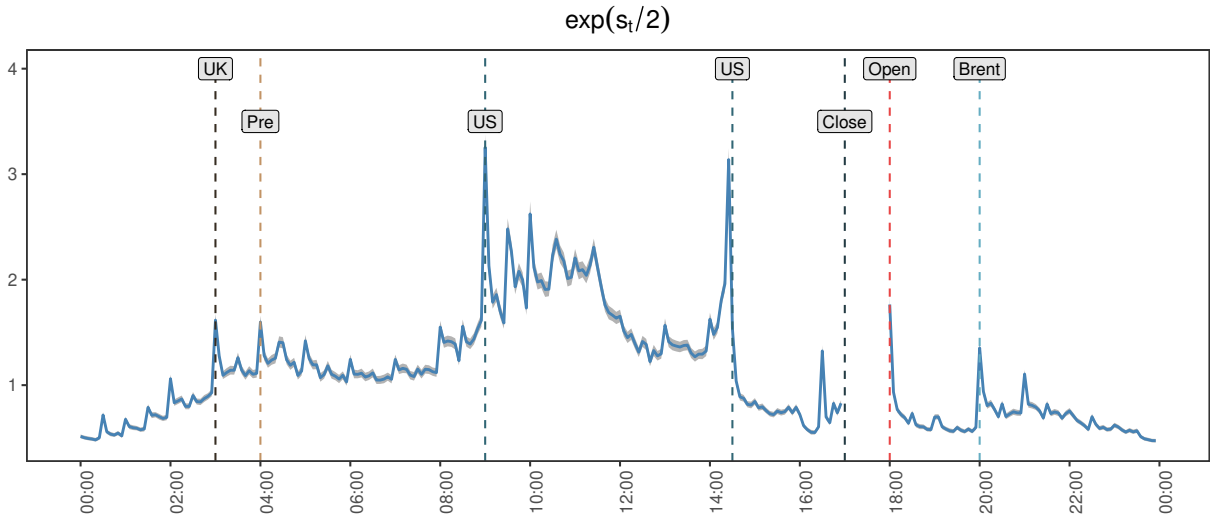

Figure 2: Posterior means of the intraday seasonality component in the standard deviation scale, $\exp\left(\frac{s_t}{2}\right)$, with 90% credible intervals. WTI Crude futures market opens at 18:00 and closes at 17:00 the next day. Brent Crude futures starts trading at 20:00. U.S. market trading time ranges from 9:00 to 14:30, with the pre-market open at 4:00.

Figure 3 presents the evolution of the estimated volatility, $\exp\left(\frac{h_t}{2}\right)$, over the in-sample period. The most notable surge occurs during the onset of the COVID-19 pandemic in early 2020, when uncertainty surrounding global economic activity and oil demand led to sharp and persistent increases in price variability.

The sizable share of variation attributed to the slow-moving component discussed in Section 5.1 is mirrored in the estimated model parameters reported in Table 5. The persistence of the latent autoregressive component, governed by the coefficient β, declines once the MIDAS term is

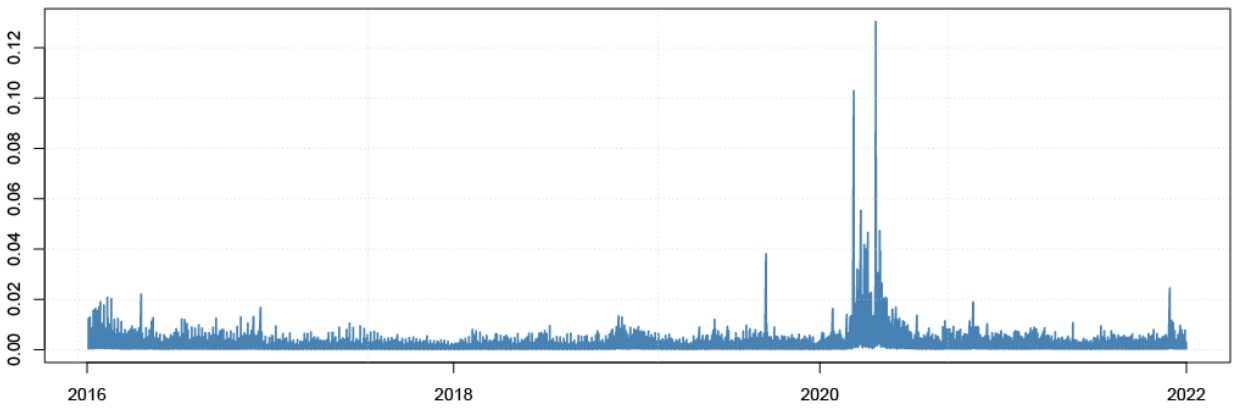

Figure 3: The evolution of volatility in standard deviation scale, $\exp\left(\frac{h_t}{2}\right)$, for the in-sample period estimated at the posterior modes of model parameters.

introduced. The two OVX-based specifications, SSVA MIDAS^{ovx,vix} and SSVA MIDAS^{ovx}, exhibit posterior medians of β below 0.8, compared with values above 0.96 in the models without MIDAS variables. The SSVA MIDAS^{vix} model lies between these cases, with $\beta \approx 0.93$. When the MIDAS term is omitted, the model compensates by inflating the autoregressive persistence parameter, effectively forcing p_t to mimic slow-moving variation. Consistent with this finding, Conrad and Kleen (2020) show a reduction in persistence following the incorporation of MIDAS component into the volatility model. The posterior median of the unconditional volatility level, m_0, is stable across specifications being near -14.2. The posterior event inclusion probability from the spike-and-slab prior γ is rather low (between 11-13%), supporting our use of a sparsity inducing prior.

Table 5: Posterior summary for the model parameters

	m_0	β	σ_η	γ	σ_α	δ_{ovx}	δ_{vix}
SSVA MIDAS^{ovx,vix}	-14.199	0.788	0.296	0.112	1.080	0.672	0.129
SSVA MIDAS^{ovx}	-14.202	0.798	0.283	0.116	1.072	0.763	–
SSVA MIDAS^{vix}	-14.153	0.928	0.128	0.123	1.052	–	0.648
SSVA	-14.192	0.973	0.072	0.132	1.067	–	–

The table presents the posterior MCMC medians for the parameters of the models considered.

The estimated MIDAS loadings further emphasize the economic relevance of low-frequency information in explaining intraday volatility dynamics. As shown in Table 5, the posterior median of the OVX loading, δ_{ovx}, is considerably larger than that of δ_{vix}, with values of 0.67 and 0.13, respectively, in the joint SSVA MIDAS^{ovx,vix} specification. This pattern suggests that oil-market specific volatility information embedded in the OVX carries stronger explanatory power for high-frequency

oil price volatility than broader equity-market uncertainty captured by the VIX. The relatively smaller yet positive δ_{vix} coefficient implies that general financial uncertainty still contributes to explaining low-frequency shifts, albeit to a lesser extent.

6 Forecasting and portfolio allocation

This section examines the out-of-sample performance and economic relevance of the proposed volatility models. Our evaluation period spans from 3 January 2022 to 30 August 2024, covering nearly 189 thousand 5-minute observations. Section 6.1 assesses the statistical accuracy of competing volatility forecasts using realized volatility constructed from 1-minute returns as a proxy for the latent volatility process. The analysis relies on Mincer–Zarnowitz (MZ) regressions (Mincer and Zarnowitz, 1969), horse-race regressions (HR), and Diebold–Mariano (DM) tests (Diebold and Mariano, 1995) for evaluating predictive performance across models and horizons. In Section 6.2, we translate these statistical improvements into economic terms by examining the implications for volatility-managed portfolio strategies. Following Moreira and Muir (2017), we evaluate how using our model-based forecasts to dynamically adjust portfolio exposure affects realized returns and risk-adjusted performance. Together, these two complementary exercises provide a comprehensive assessment of both the statistical and economic value of incorporating slow and fast level shifts in intraday volatility models.

6.1 Realized volatility forecasts

Evaluating volatility forecasts is inherently challenging because the volatility process is not observable. To assess predictive performance, we therefore rely on a high-frequency volatility proxy. We compute the realized volatility series as the square root of the sum of squared one-minute returns over non-overlapping five-minute intervals.

We use three complementary evaluation frameworks: Mincer–Zarnowitz regressions, horse-race regressions, and Diebold–Mariano tests. The MZ regressions assess the calibration and informational content of the forecasts by examining whether predicted volatilities are unbiased and correctly scaled relative to their realized counterparts. The HR regressions extend this framework to pairwise model comparisons, evaluating whether competing forecasts contain incremental predictive

information beyond the benchmark specification. Finally, the DM tests provide formal statistical comparisons of forecasting accuracy under squared and absolute loss functions. Results for all three evaluation procedures are reported jointly in Table 6, which is divided into Panel A (five-minute-ahead forecasts) and Panel B (day-ahead forecasts) to highlight potential differences across horizons.

Table 6: Forecasting results

	Mincer–Zarnowitz			Horse-Race		Diebold-Mariano	
Model	$\hat{\alpha}_0$	$\hat{\alpha}_1$	R^2	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	Squared	Absolute
Panel A: 5-min-ahead forecasts							
SSVA MIDAS^{ovx}	0.009	0.939	0.473	–	–	–	–
SSVA MIDAS^{ovx,vix}	0.009	0.931	0.472	1.117	24.645	-2.85***	-15.5***
SSVA MIDAS^{vix}	0.012	0.881	0.471	0.695	63.520	-4.72***	-20.77***
SSVA	0.013	0.889	0.471	0.594	71.937	-3.60***	-12.67***
SSVAg	0.013	0.882	0.470	0.614	75.858	-4.31***	-14.98***
SSV	0.013	0.884	0.455	0.698	104.285	-2.86***	-16.15***
SV	0.013	0.896	0.389	0.810	182.932	-6.85***	-53.89***
GARCH	0.018	0.779	0.356	0.758	243.396	-10.75***	-69.64***
AR1-RV	0.017	0.849	0.304	0.798	263.429	-7.45***	-74.37***
Panel B: Day-ahead forecasts							
SSVA MIDAS^{ovx}	-0.001	1.087	0.593	–	–	–	–
SSVA MIDAS^{ovx,vix}	0.000	1.032	0.585	0.801	2.979	-0.24	-1.40*
SSVA MIDAS^{vix}	0.003	0.811	0.371	1.063	19.997	-7.28***	-11.99***
SSVA	-0.091	4.930	0.205	1.082	29.506	-6.75***	-14.97***
SSVAg	-0.086	4.589	0.180	1.083	29.647	-7.13***	-15.96***
SSV	-0.214	10.079	0.348	1.080	29.767	-6.68***	-15.44***
SV	-0.139	7.095	0.302	1.078	29.371	-6.32***	-14.36***
GARCH	-0.027	2.886	0.436	1.016	23.811	-8.35***	-14.30***
AR1-RV	-0.221	12.855	0.065	1.082	31.332	-6.65***	-13.54***

The table reports combined results for Mincer–Zarnowitz regressions, Horse-race regressions, and Diebold–Mariano tests for one-step-ahead (Panel A) and one-day-ahead (Panel B) out-of-sample forecasting from 2022-01-03 to 2024-08-30.

Columns 2-4 report $\hat{\alpha}_0$, $\hat{\alpha}_1$, and R^2 of the Mincer–Zarnowitz regressions, $RV_i = \alpha_0 + \alpha_1 \widehat{vol}_{i|i-1}^{\,model} + \varepsilon_i$.

Columns 5-6 report $\hat{\beta}_1$ and $t(\hat{\beta}_1)$ testing $H_0 : \beta_1 = 0$ from the Horse-race regressions $RV_i = \beta_0 + \beta_1 \widehat{vol}_{i|i-1}^{\,benchmark} + (1-\beta_1)\widehat{vol}_{i|i-1}^{\,competitor} + \varepsilon_i$.

Columns 7-8 report the Diebold–Mariano (DM) statistics compare predictive accuracy against the SSVA MIDAS^{ovx} benchmark under squared and absolute loss. Negative and statistically significant DM statistics indicate superior performance of the benchmark.

***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

We begin by examining the MZ regressions, which relate the realized volatility proxy to the model-based forecast according to

$$RV_i = \alpha_0 + \alpha_1 \widehat{vol}_{i|i-1}^{model} + \varepsilon_i,$$

where RV_i denotes realized volatility computed from high-frequency returns over forecast interval i, and $\widehat{vol}_{i|i-1}^{model}$ is the corresponding forecast generated using information available up to $i-1$. A well-calibrated forecast should yield an intercept α_0 close to zero, a slope α_1 near one, and a high coefficient of determination R^2, indicating both unbiasedness and accurate scaling.

The first three columns of Table 6 report the MZ regression results for both forecast horizons. Across the five-minute-ahead forecasts in Panel A, all variants of the SSVA models with the slow level shifts exhibit intercepts close to zero and the estimated slope coefficients ($\hat{\alpha}_1$) close to unity. The inclusion of slow-moving MIDAS components, particularly those based on the OVX or joint OVX–VIX specifications, leads to higher explanatory power, with R^2 values approaching 47%. In contrast, models that omit either the long-run MIDAS term display noticeably lower fit with the plain SV and GARCH models yielding R^2 values below 40% and 36%, respectively. These results indicate that high-frequency volatility forecasts benefit substantially from accounting for slow and fast level shift components.

Turning to the day-ahead forecasts in Panel B, the performance differences become even more pronounced. The SSVA MIDAS models again produce slope coefficients closest to one, while simpler specifications exhibit significant biases, with slopes often well above or well below unity. The R^2 values for the best-performing SSVA MIDAS^{ovx} and SSVA MIDAS^{ovx,vix} models exceed 0.58, compared with values below 0.45 for standard benchmarks. This widening performance gap across horizons emphasizes the value of incorporating low-frequency information when forecasting high-frequency volatility over longer periods. In short, while MIDAS terms improve precision even at very short horizons, their contribution to forecast accuracy becomes particularly evident at the daily level.

To further evaluate the relative forecasting performance of the competing models, we implement horse-race regressions that enable pairwise comparisons between our benchmark specification and

alternative forecasts. The HR regression is specified as

$$RV_i = \beta_0 + \beta_1 \widehat{vol}_{i|i-1}^{benchmark} + (1 - \beta_1)\widehat{vol}_{i|i-1}^{competitor} + \varepsilon_i,$$

where $\widehat{vol}_{i|i-1}^{benchmark}$ denotes the volatility forecast from the benchmark model and $\widehat{vol}_{i|i-1}^{competitor}$ represents the forecast from a competing specification. The parameter β_1 captures the relative informational content of the benchmark forecast. When $\beta_1 = 1$, the benchmark fully dominates, implying that the competitor adds no incremental predictive value. Conversely, $\beta_1 = 0$ indicates that the benchmark is entirely dominated by the competitor. Intermediate values of β_1 suggest partial complementarity between the two forecasts, whereas $\beta_1 > 1$ implies that the competitor's inclusion worsens forecast accuracy - reinforcing the superiority of the benchmark model. The associated t-statistic for $\hat{\beta}_1$ provides a formal test of whether the benchmark's informational content differs significantly from zero.

The middle columns of Table 6 report the HR regression results. In the five-minute-ahead forecasts (Panel A), the benchmark SSVA MIDAS^{ovx} model consistently dominates all competing alternatives. The estimated $\hat{\beta}_1$ coefficients range between 0.6 and 1.1, and are uniformly significant at the 1% level, indicating that the benchmark contains the vast majority of predictive information. Other SSVA MIDAS variants, such as those incorporating both OVX and VIX or only VIX, display the estimated $\hat{\beta}_1$ values somewhat below one, suggesting a limited degree of complementarity but confirming that information from the OVX-based specification remains dominant. As model structure becomes progressively simpler, first by removing the MIDAS term, then by excluding seasonality or announcement effects, the HR coefficients decline sharply, reflecting substantial information loss. The GARCH and AR1-RV benchmarks perform particularly poorly, with very low relative weights.

In the day-ahead forecasts reported in Panel B, the relative gains from the SSVA MIDAS specifications again become even more pronounced. The benchmark dominates most competitors, with $\hat{\beta}_1$ typically near or above unity. In contrast, competing models that omit the MIDAS structure or rely solely on high-frequency information exhibit weaker or statistically insignificant contributions. The results suggest that forecasts derived from the SSVA MIDAS framework contain nearly all the relevant information for predicting next-day volatility, leaving little or no incremental value for

simpler specifications. This dominance at longer horizons is consistent with the notion that MIDAS terms effectively capture persistent low-frequency components of volatility that govern multi-period dynamics. Taken together, the HR regression evidence complements the MZ findings by showing that competing models add virtually no predictive power at either intraday or daily horizons, with the gap widening as the forecast horizon increases.

Finally, we employ the Diebold–Mariano test to assess whether observed differences in forecast accuracy across models are statistically significant. The DM test compares two competing forecasts under a given loss function, here based on squared and absolute forecast errors, against the null hypothesis of equal predictive accuracy. Each model is tested against the benchmark SSVA MIDAS^{ovx} specification under the alternative that the competitor is less accurate. Negative and statistically significant DM statistics therefore indicate superior performance of the benchmark.

The last two columns of Table 6 report the DM statistics for both horizons. Across both the five-minute and day-ahead forecasts, all values are negative and nearly all are significant at 1%, confirming that the SSVA MIDAS^{ovx} model consistently delivers lower forecast errors than any competing specification. The differences are modest among MIDAS variants but substantial relative to simpler SV and GARCH benchmarks. Overall, the DM results reinforce the conclusions drawn from the MZ and HR regressions.

6.2 Volatility-managed portfolios

We complement our out-of-sample forecasting analysis by assessing whether the statistical gains achieved by the proposed volatility model translate into economically meaningful improvements in investment performance by constructing volatility-managed portfolios as in Moreira and Muir (2017). The intuition is that investors can improve risk-adjusted performance by scaling portfolio exposure inversely with the forecasted volatility of the underlying asset. Specifically, the volatility-managed portfolio return is defined as

$$y^*_{t+1} = \frac{c}{\widehat{vol}_{t+1|t}}\, y_{t+1},$$

where y^*_{t+1} denotes the managed portfolio return, y_{t+1} the raw portfolio return, and $\widehat{vol}_{t+1|t}$ the volatility forecast at time t. The constant c is chosen such that the unconditional variance of the

managed and non-managed portfolios is identical, ensuring comparability of risk across strategies.

The results, reported in Table 7, highlight the economic significance of improved volatility forecasts. The non-managed portfolio, based on a simple buy-and-hold strategy for WTI (first row), achieves an annualized mean return of approximately 5% and an annualized Sharpe ratio of 0.13. By design, all managed portfolios have the same unconditional standard deviation as the benchmark, allowing direct comparison of expected returns and risk-adjusted performance.

Table 7: Volatility-managed portfolio

	Mean	SR
WTI	0.053	0.132
SSVA MIDAS^{ovx}	0.213	0.531
SSVA MIDAS^{ovx,vix}	0.211	0.526
SSVA MIDAS^{vix}	0.214	0.532
SSVA	0.194	0.484
SSVAg	0.189	0.469
SSV	0.204	0.508
SV	0.097	0.242
GARCH	0.075	0.188
AR1-RV	0.101	0.252

1-step-ahead volatility-managed portfolio results for 2022-01-03 – 2024-08-30 out-of-sample period (189k observations). Mean is annualized average return, SR is the annualized Sharpe ratio.

Two main insights emerge. First, implementing volatility-managed strategies, even when volatility is modeled using a standard stochastic volatility specification, nearly doubles both the annualized mean return and the Sharpe ratio relative to the buy-and-hold benchmark. Second, enriching the volatility model with additional features such as intraday seasonality and the slow-moving MIDAS component leads to further gains, with annualized mean returns reaching up to 21% for the best-performing SSVA MIDAS specifications. These results indicate that the superior forecasting accuracy of our model translates directly into tangible portfolio benefits.

7 Conclusion

This paper develops a mixed-frequency stochastic volatility framework that simultaneously captures slow-moving and fast level shifts in intraday volatility. The slow-moving component, modeled via a MIDAS specification, effectively integrates information from daily market uncertainty measures, while the fast-moving component isolates transitory volatility spikes triggered by scheduled announcements. Our empirical analysis of 5-minute WTI futures returns shows that incorporating both types of level shifts substantially improves out-of-sample volatility forecasts across intraday and daily horizons. In particular, models including the slow-moving MIDAS term better capture long-memory behavior in volatility, reducing the need to inflate short-run persistence parameters, whereas the announcement component accurately isolates temporally sparse but economically significant volatility jumps.

Variance decomposition analyses highlight the economic and statistical relevance of these components. Across the full sample, the slow-moving MIDAS factor explains a substantial portion of the variation in the latent log-variance, ranging from roughly 26% in VIX-only specifications to nearly 39% when OVX is included, doubling the share of variance explained by the persistence component. Restricting the decomposition to intervals with at least one scheduled event reveals that the explained share of the announcement component rises sharply to more than 22% of variance, highlighting the importance of news events as a driving force of intraday oil volatility. In contrast, the MIDAS and persistence components decline proportionally, reflecting the temporary dominance of event-driven information.

Empirical evidence confirms that only a limited set of scheduled announcements, notably U.S. Change in Nonfarm Payrolls, FOMC rate decisions, and DOE crude oil inventories, show statistically and economically meaningful effects on high-frequency volatility. These events capture sudden shifts in expectations regarding energy demand, monetary policy, or supply conditions, while the MIDAS component tracks broader uncertainty in oil markets. Collectively, these findings suggest that intraday oil volatility is shaped by a combination of fast information shocks and low-frequency market sentiment, with each component contributing in a complementary manner.

The statistical improvements translate directly into economically meaningful gains. Volatility-managed portfolios constructed using our model-based forecasts achieve higher Sharpe ratios rel-

ative to both simple buy-and-hold strategies and forecasts from standard stochastic volatility or GARCH-type models. These results highlight the importance of jointly modeling multiple volatility drivers to capture the rich dynamics of high-frequency returns.

Overall, the proposed framework offers a flexible and interpretable approach to intraday volatility modeling, providing both improved predictive accuracy and tangible economic benefits. The methodology is broadly applicable to other asset classes characterized by mixed-frequency information flows, highlighting its potential for general use in high-frequency financial econometrics.

Acknowledgments

Igor Ferreira Batista Martins acknowledges financial support from the Jan Wallander and Tom Hedelius foundation and Tore Browalds foundation grant number BFv22-0005. Audronė Virbickaitė is partially supported by the grant PID2022-138289NB-I00 from the Spanish State Research Agency (Agencia Estatal de Investigación - Ministerio de Ciencia e Innovación). Hoang Nguyen acknowledges financial support from the project "Improved Economic Policy and Forecasting with High-Frequency Data" (Dnr: E47/22) funded by the Torsten Söderbergs Foundation. Hedibert Freitas Lopes was partially supported by São Paulo Research Foundation (FAPESP) Grants 2013/00506-1, 2018/04654-9 and 2023/02538-0.

References

T. G. Andersen and T. Bollerslev. Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2-3):115–158, 1997.

T. G. Andersen and T. Bollerslev. Deutsche Mark—Dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. *The Journal of Finance*, 53(1): 219–265, 1998.

T. G. Andersen, T. Bollerslev, F. X. Diebold, and C. Vega. Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics*, 73(2):251–277, 2007.

A. Basistha and A. Kurov. The impact of monetary policy surprises on energy prices. *Journal of Futures Markets*, 35(1):87–103, 2015.

J. Bekierman and B. Gribisch. A mixed frequency stochastic volatility model for intraday stock market returns. *Journal of Financial Econometrics*, 19(3):496–530, 2021.

J. Bjursell, J. E. Gentle, and G. H. Wang. Inventory announcements, jump dynamics, volatility and trading volume in US energy futures markets. *Energy Economics*, 48:336–349, 2015.

M. W. Brandt and L. Gao. Macro fundamentals or geopolitical events? A textual analysis of news events for crude oil. *Journal of Empirical Finance*, 51:64–94, 2019.

H. Bu. Effect of inventory announcements on crude oil price volatility. *Energy Economics*, 46: 485–494, 2014.

C. K. Carter and R. Kohn. On Gibbs sampling for state space models. *Biometrika*, 81(3):541–553, 1994.

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

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

R. Deo, C. Hurvich, and Y. Lu. Forecasting realized volatility using a long-memory stochastic volatility model: Estimation, prediction and seasonal adjustment. *Journal of Econometrics*, 131 (1-2):29–58, 2006.

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

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

R. F. Engle and M. E. Sokalska. Forecasting intraday volatility in the US equity market. Multiplicative component GARCH. *Journal of Financial Econometrics*, 10(1):54–83, 2012.

S. Frühwirth-Schnatter. Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15(2):183–202, 1994.

E. Ghysels, P. Santa-Clara, and R. Valkanov. The MIDAS touch: Mixed data sampling regression models. *UCLA Working Paper*, 2004.

E. Haugom, H. Langeland, P. Molnár, and S. Westgaard. Forecasting volatility of the US oil market. *Journal of Banking & Finance*, 47:1–14, 2014.

D. R. Känzig. The macroeconomic effects of oil supply news: Evidence from OPEC announcements. *American Economic Review*, 111(4):1092–1125, 2021.

S. Kim, N. Shephard, and S. Chib. Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies*, 65(3):361–393, 1998.

D. Korobilis and D. Pettenuzzo. Adaptive hierarchical priors for high-dimensional vector autoregressions. *Journal of Econometrics*, 212(1):241–271, 2019.

R. López. The behaviour of energy-related volatility indices around scheduled news announcements: Implications for variance swap investments. *Energy Economics*, 72:356–364, 2018.

I. Martins and H. F. Lopes. What events matter for exchange rate volatility? *arXiv preprint arXiv:2411.16244*, 2024.

I. Martins, A. Virbickaitė, H. Nguyen, and H. F. Lopes. Volume-driven time-of-the-day effects in intraday volatility models. *Scandinavian Working Papers*, 2025.

J. A. Mincer and V. Zarnowitz. The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance*, pages 3–46. NBER, 1969.

T. J. Mitchell and J. J. Beauchamp. Bayesian variable selection in linear regression. *Journal of the American Statistical Association*, 83(404):1023–1032, 1988.

A. Moreira and T. Muir. Volatility-managed portfolios. *The Journal of Finance*, 72(4):1611–1644, 2017.

Z. Niu, F. Ma, and H. Zhang. The role of uncertainty measures in volatility forecasting of the crude oil futures market before and during the COVID-19 pandemic. *Energy Economics*, 112:106120, 2022.

J. Noguera-Santaella. Geopolitics and the oil price. *Economic Modelling*, 52:301–309, 2016.

E. Rossi and D. Fantazzini. Long memory and periodicity in intraday volatility. *Journal of Financial Econometrics*, 13(4):922–961, 2015.

H. Schmidbauer and A. Rösch. OPEC news announcements: Effects on oil price expectation and volatility. *Energy Economics*, 34(5):1656–1663, 2012.

J. R. Stroud and M. S. Johannes. Bayesian modeling and forecasting of 24-hour high-frequency volatility. *Journal of the American Statistical Association*, 109(508):1368–1384, 2014.

A. Virbickaitė, H. Nguyen, and M.-N. Tran. Bayesian predictive distributions of oil returns using mixed data sampling volatility models. *Resources Policy*, 86:104167, 2023.

Y. Yang, J. Zhang, and S. Chen. Information effects of monetary policy announcements on oil price. *Journal of Commodity Markets*, 30:100268, 2023.

F. Žikeš and J. Baruník. Semi-parametric conditional quantile models for financial returns and realized volatility. *Journal of Financial Econometrics*, 14(1):185–226, 2015.

A Appendix: MCMC algorithm

Let $\mathbf{\Theta} = (\beta, \sigma_\eta, \{p_t\}_{t=1}^T; \{\beta_k\}_{k=1}^{K-1}; \{\alpha_i\}_{i=1}^N, \{\pi_i\}_{i=1}^N, \sigma_a^2, \gamma; \{w_j\}_{j=1}^J, \{\delta_j\}_{j=1}^J)$ be the set of model parameters and latent states. And let $\mathbf{\Psi}$ be all the parameters in $\mathbf{\Theta}$ except the ones we sample from in a given step of the MCMC procedure. We employ a Bayesian approach and use MCMC methods to sample from the joint posterior distribution conditional on the observed data. The choice of conjugate priors allows for efficient and fast implementation of the Gibbs sampler. We log-linearize the model and approximate $u_t = \log \epsilon_t^2 \sim \log \chi^2(1)$ by a mixture of 7 Gaussian components following Kim et al. (1998). A summary of the equations required for sampling is as follows:

$$
\begin{aligned}
\log y_t^2 &= h_t + u_t, \text{ with } u_t \sim \mathbf{N}, (\mu_{c,t}, \sigma_{c,t}^2) \\
h_t &= m_\tau + p_t + s_t + e_t, \\
p_t &= \beta p_{t-1} + \sigma_\eta \eta_t, \text{ with } \eta_t \sim \mathbf{N}(0,1), \\
s_t &= \sum_{k=1}^{K} H_{tk} \beta_k, \\
e_t &= \sum_{i=1}^{N} I_{it} \alpha_i, \\
m_\tau &= \sum_{j=1}^{J} \delta_j \left[\sum_{l=1}^{L_j} \phi_l(w_j) X_{j,\tau-l} \right].
\end{aligned}
$$

1) In order to sample $\beta|\mathbf{\Psi}, \mathbf{y}$, we consider the persistent component as a linear regression,

$$
p_t = \beta p_{t-1} + \sigma_\eta \eta_t.
$$

Then the conditional posterior distribution of β is a conjugate Gaussian distribution due to a Gaussian prior and a Gaussian likelihood. We have that

$$
\begin{aligned}
&\beta|\mathbf{\Psi}, \mathbf{y} \sim \mathbf{N}(\hat{\beta}_0, \hat{V}_{\beta_0}), \text{where} \\
&\hat{V}_{\beta_0} = \left(\frac{1}{V_{\beta_0}} + \frac{1}{\sigma_\eta^2} \sum_{t=1}^{T} p_{t-1} p_{t-1} \right)^{-1}, \\
&\hat{\beta}_0 = \hat{V}_{\beta_0} \left(\frac{\beta_0}{V_{\beta_0}} + \frac{1}{\sigma_\eta^2} \sum_{t=1}^{T} p_t p_{t-1} \right).
\end{aligned}
$$

2) In order to sample $\sigma_\eta^2|\mathbf{\Psi}, \mathbf{y}$, we also use the linear representation of the persistent component above. Then the conditional posterior distribution of σ_η^2 is a conjugate inverse Gamma distribution due to an inverse Gamma prior and a Gaussian likelihood. We have that

$$\sigma_\eta^2|\mathbf{\Psi}, \mathbf{y} \sim \mathbf{IG}\left(\alpha_{\sigma_\eta} + \frac{T}{2}, \beta_{\sigma_\eta} + \frac{\sum_{t=1}^{T}(p_t - \beta p_{t-1})^2}{2}\right).$$

3) In order to sample the mixture component $\{c_t, \mu_{c,t}, \sigma_{c,t}^2\}_{t=1}^T|\mathbf{\Psi}$, we follow (Kim et al., 1998) and first sample $\{c_t\}_{t=1}^T$ from the posterior probability

$$\pi\left(c_t = c \mid \mathbf{\Psi}\right) \propto \mathbb{P}\left(c_t = c\right) \frac{1}{\sigma_c} \exp\left\{-\frac{\left(u_t - \mu_c\right)^2}{2\sigma_c^2}\right\},$$

where $\mathbb{P}\left(c_t = c\right)$ is the mixture weight of the c_th Gaussian component with mean μ_c and variance σ_c^2 (Kim et al., 1998). Set the $\{\mu_{c,t}, \sigma_{c,t}^2\}_{t=1}^T$ as the mean and variance of the c_t_th Gaussian component.

4) In order to sample $\{p_t\}_{t=1}^T|\mathbf{\Psi}, \mathbf{y}$, we rewrite the model in terms of measurement and state equations as,

$$\begin{aligned}\log y_t^2 - m_\tau - s_t - e_t &= p_t + u_t,\\ p_t &= \beta p_{t-1} + \sigma_\eta \eta_t.\end{aligned}$$

Then we follow Carter and Kohn (1994) and Frühwirth-Schnatter (1994) to sample the conditional posterior distribution of $\{p_t\}_{t=1}^T$ from a Gaussian linear state space model using the Forward Filtering Backward Sampling (FFBS) algorithm.

5) In order to sample $\{\beta_k\}_{k=1}^{K-1}|\mathbf{\Psi}, \mathbf{y}$, we rewrite the log-linearized model as a linear regression,

$$\begin{aligned}\log y_t^2 - m_\tau - p_t - e_t &= \sum_{k=1}^{K} H_{tk}\beta_k + u_t,\\ &= \sum_{k=1}^{K-1}(H_{tk} - H_{tK})\beta_k + u_t,\end{aligned}$$

Then the conditional posterior distribution of $\{\beta_k\}_{k=1}^{K-1}$ is a conjugate Gaussian distribution due to

a Gaussian prior and a Gaussian likelihood. We have that,

$$\{\beta_k\}_{k=1}^{K-1}|\mathbf{\Psi},\mathbf{y} \sim \mathbf{N}(\hat{\boldsymbol{\beta}}_k, \hat{\mathbf{V}}_{\beta_k}),$$
$$\hat{\mathbf{V}}_{\beta_k} = \left(\mathbf{V}_{\beta_{0k}}^{-1} + \sum_{t=1}^{T} \frac{1}{\sigma_{c_t}^2}\tilde{\mathbf{H}}_t'\tilde{\mathbf{H}}_t\right)^{-1},$$
$$\hat{\boldsymbol{\beta}}_k = \hat{\mathbf{V}}_{\beta_k}\left(\mathbf{V}_{\beta_{0k}}^{-1}\boldsymbol{\beta}_{0k} + \sum_{t=1}^{T} \frac{1}{\sigma_{c_t}^2}\tilde{\mathbf{H}}_t'(\log y_t^2 - m_\tau - p_t - e_t - \mu_{c,t})\right),$$

where $\tilde{\mathbf{H}}_t = (H_{t,1} - H_{t,K}, \ldots, H_{t,K-1} - H_{t,K})$, and set $\beta_K = -\sum_{k=1}^{K-1}\beta_k$.

6) In order to sample $\sigma_a^2|\mathbf{\Psi},\mathbf{y}$, we consider all announcements with non-zero size effect $\alpha_i|\pi_i = 1 \sim N(0,\sigma_\alpha^2)$. Therefore, the conditional posterior distribution of σ_a^2 is an inverse Gamma distribution due to an inverse Gamma prior and a Gaussian likelihood. We have that

$$\sigma_\alpha^2|\mathbf{\Psi},\mathbf{y} \sim \mathbf{IG}\left(\alpha_\sigma + \frac{\sum_{i=1}^{N}\mathbf{I}_{(\pi_i=1)}}{2}, \beta_\sigma + \frac{\sum_{i=1}^{N}\alpha_i^2}{2}\right).$$

7) In order to sample $\gamma|\mathbf{\Psi},\mathbf{y}$, we consider the likelihood of all announcements with non-zero size effect $\pi_i|\gamma \sim \mathbf{Bern}(\gamma)$. Hence, the conditional posterior distribution of γ is a conjugate Beta distribution due to a Beta prior and a Binomial likelihood. We have that

$$\gamma|\mathbf{\Psi},\mathbf{y} \sim \mathbf{Beta}\left(\alpha_\gamma + \sum_{i=1}^{N}\mathbf{I}_{(\pi_i=1)}, \beta_\gamma + \sum_{i=1}^{N}\mathbf{I}_{(\pi_i=0)}\right).$$

8) In order to sample $\{\pi_i,\alpha_i\}_{i=1}^N|\mathbf{\Psi},\mathbf{y}$, we consider the hierarchical spike and slap prior $\alpha_i|\pi_i \sim (1-\pi_i)\delta_0 + \pi_i\mathbf{N}(0,\sigma_\alpha^2)$. Then following Korobilis and Pettenuzzo (2019), the conditional posterior distribution of π_i is a conjugate Bernoulli distribution due to a Bernoulli prior and a Beta likelihood. And the conditional posterior distribution of $\{\alpha_i\}_{i=1}^N$ is a conjugate Gaussian distribution for $\pi_i = 1$ due to a Gaussian prior and a Gaussian likelihood. We have that

$$\pi_i|\mathbf{\Psi},\mathbf{y} \sim \mathbf{Bern}\left(\frac{\frac{\gamma\phi(0;0,\sigma_\alpha^2)}{\phi(0;\hat{\alpha}_i,\hat{V}_i)}}{\frac{\gamma\phi(0;0,\sigma_\alpha^2)}{\phi(0;\hat{\alpha}_i,\hat{V}_i)} + 1 - \gamma}\right),$$

where $\hat{\alpha}_i, \hat{V}_i$ is the conditional posterior mean and variance of α_i from

$$\begin{aligned}
\alpha_i|\mathbf{\Psi}, \mathbf{y} &\sim \mathbf{N}(\hat{\alpha}_i, \hat{V}_i), \\
\hat{V}_i &= \left(\frac{1}{\sigma_\alpha^2} + \sum_{t=1}^{T} \frac{1}{\sigma_{c,t}^2} \mathbf{I}_{i,t}^2 \right)^{-1}, \\
\hat{\alpha}_i &= \hat{V}_i \left(\sum_{t=1}^{T} \frac{1}{\sigma_{c,t}^2} \mathbf{I}_{i,t} (\log y_t^2 - m_\tau - p_t - s_t - \mathbf{I}_{-i,t}' \alpha_{-i} - \mu_{c,t}) \right),
\end{aligned}$$

where $\mathbf{I}_{-i,t}$ and $\alpha_{-i,t}$ are the $\mathbf{I}_t$ and α vectors without the $i_$th element.

9) In order to sample $\{\delta_j\}_{j=1}^J|\mathbf{\Psi}, \mathbf{y}$, we rewrite the log-linearized model as a linear regression,

$$\log y_t^2 - p_t - s_t - e_t - m_0 = \boldsymbol{\delta}' \tilde{\mathbf{X}}_\tau + u_t.$$

Then the conditional posterior distribution of $\{\delta_j\}_{j=1}^J$ is a Gaussian distribution due to a Gaussian prior and a Gaussian likelihood. We have that

$$\begin{aligned}
\boldsymbol{\delta}|\mathbf{\Psi}, \mathbf{y} &\sim \mathbf{N}(\hat{\boldsymbol{\delta}}, \hat{\mathbf{V}}_{\boldsymbol{\delta}}), \\
\hat{\mathbf{V}}_{\boldsymbol{\delta}} &= \left(\mathbf{V}_{\boldsymbol{\delta}_0}^{-1} + \sum_{t=1}^{T} \frac{1}{\sigma_{c,t}^2} \tilde{\mathbf{X}}_\tau' \tilde{\mathbf{X}}_\tau \right)^{-1}, \\
\hat{\boldsymbol{\delta}} &= \hat{\mathbf{V}}_{\boldsymbol{\delta}} \left(\mathbf{V}_{\boldsymbol{\delta}_0}^{-1} \boldsymbol{\delta}_0 + \sum_{t=1}^{T} \frac{1}{\sigma_{c,t}^2} \tilde{\mathbf{X}}_\tau' (\log y_t^2 - p_t - s_t - e_t - m_0 - \mu_{c,t}) \right),
\end{aligned}$$

where $\tilde{\mathbf{X}}_\tau = [\tilde{X}_{1,\tau}, \ldots, \tilde{X}_{J,\tau}]$ where $\tilde{X}_{j,\tau} = \sum_{l=1}^{L_j} \phi_l(w_j) X_{j,\tau-l}$.

10) In order to sample $m_0|\mathbf{\Psi}, \mathbf{y}$, we rewrite the log-linearized model as a linear regression,

$$\log y_t^2 - p_t - s_t - e_t - \tilde{\mathbf{X}}_\tau' \boldsymbol{\delta} = m_0 + u_t$$

Then the conditional posterior distribution of m_0 is a Gaussian distribution due to a Gaussian prior

and a Gaussian likelihood. We have that

$$
\begin{aligned}
m_0|\boldsymbol{\Psi}, \mathbf{y} &\sim \mathbf{N}(\hat{m}_0, \hat{V}_{m_0}), \\
\hat{V}_{m_0} &= \left(V_{m_0}^{-1} + \sum_{t=1}^{T} \frac{1}{\sigma_{c,t}^2} \right)^{-1}, \\
\hat{m}_0 &= \hat{V}_{m_0} \left(V_{m_0}^{-1} \mu_{m_0} + \sum_{t=1}^{T} \frac{1}{\sigma_{c,t}^2} (\log y_t^2 - p_t - s_t - e_t - \mu_{c,t} - \tilde{\mathbf{X}}_\tau \boldsymbol{\delta}) \right).
\end{aligned}
$$

11) In order to sample $\{w_j\}_{j=1}^J|\boldsymbol{\Psi}, \mathbf{y}$, we apply a Metropolis-Hastings step with an independence proposal. We draw $w_j^{(*)} \sim \mathbf{N}(w_j, \sigma_w^2)$ for $j = 1, \ldots, J$ and accept with the probability

$$
\min \left\{ 1, \frac{\pi(w_j^{(*)}|\boldsymbol{\Psi}, \mathbf{y})}{\pi(w_j|\boldsymbol{\Psi}, \mathbf{y})} \right\}
$$

where

$$
\pi(w_j|\boldsymbol{\Psi}, \mathbf{y}) \propto \exp \left(-\frac{1}{2} \sum_{t=1}^{T} \frac{1}{\sigma_{c,t}^2} (\log y_t^2 - p_t - s_t - e_t - \mu_{c,t} - \tilde{\mathbf{X}}_\tau \boldsymbol{\delta} - m_0) \right).
$$

This step is necessary when the lags of exogenous variables are chosen to be greater than 1.

B Appendix: Exploratory data analysis

The returns are calculated using the log difference of 5-minute prices. We remove returns outside active trading hours. However, the instances of stagnant prices are kept in order to reproduce realistic market conditions for forecasting exercises. We are left with 614,367 5-minute return observations. Figure 4 panel (a) draws the absolute returns series and the locally weighted scatterplot smoothing (LOWESS) in red color. The long term level of the volatility represented by the red line varies through time and hits the highest level during the COVID-19 pandemic. The average full-sample annualized return is around 3.63%.

Panel (b) in Figure 4 exhibits the seasonal pattern of the intraday volatility as the average of absolute returns for each 5-minute mark over the entire sample. The gap in the series at the 17:00-18:00 mark is when the markets are closed. The spikes in volatility observed right before(after) the market (Asian, U.S., United Kingdom) closes(opens) are a well-known stylized features of intraday volatility, that happen due to a variety of factors: accumulated uncertainty due to no trading during the break; order imbalance when traders adjust their positions based on new information; and, possibly, lower liquidity meanwhile the market participants adjust. This feature is especially pronounced on Sundays, see panel (c) in Figure 4. Since the markets open on Sundays after the weekend, there is an extra amount of accumulated uncertainty. Therefore, in order to accommodate this sharp increase in volatility, we have created an additional event called "Sunday Open". Unusual increase in intraday volatility on Sundays at 18:00 is not unique to oil volatility only and has been documented for other asset classes, such as S&P500 and Nasdaq indices and EUR/USD exchange rate, see Martins et al. (2025) for example.

Finally, panels (d) and (e) in Figure 4 show the average absolute value of intraday returns over the selected days that contain some of the scheduled announcements, for example, "U.S. FOMC Rate Decision Lower Bound", and "U.S. DOE U.S. Crude Oil Inventories". We expect these announcements to have an impact on intraday volatility, as there is economic justification for that.

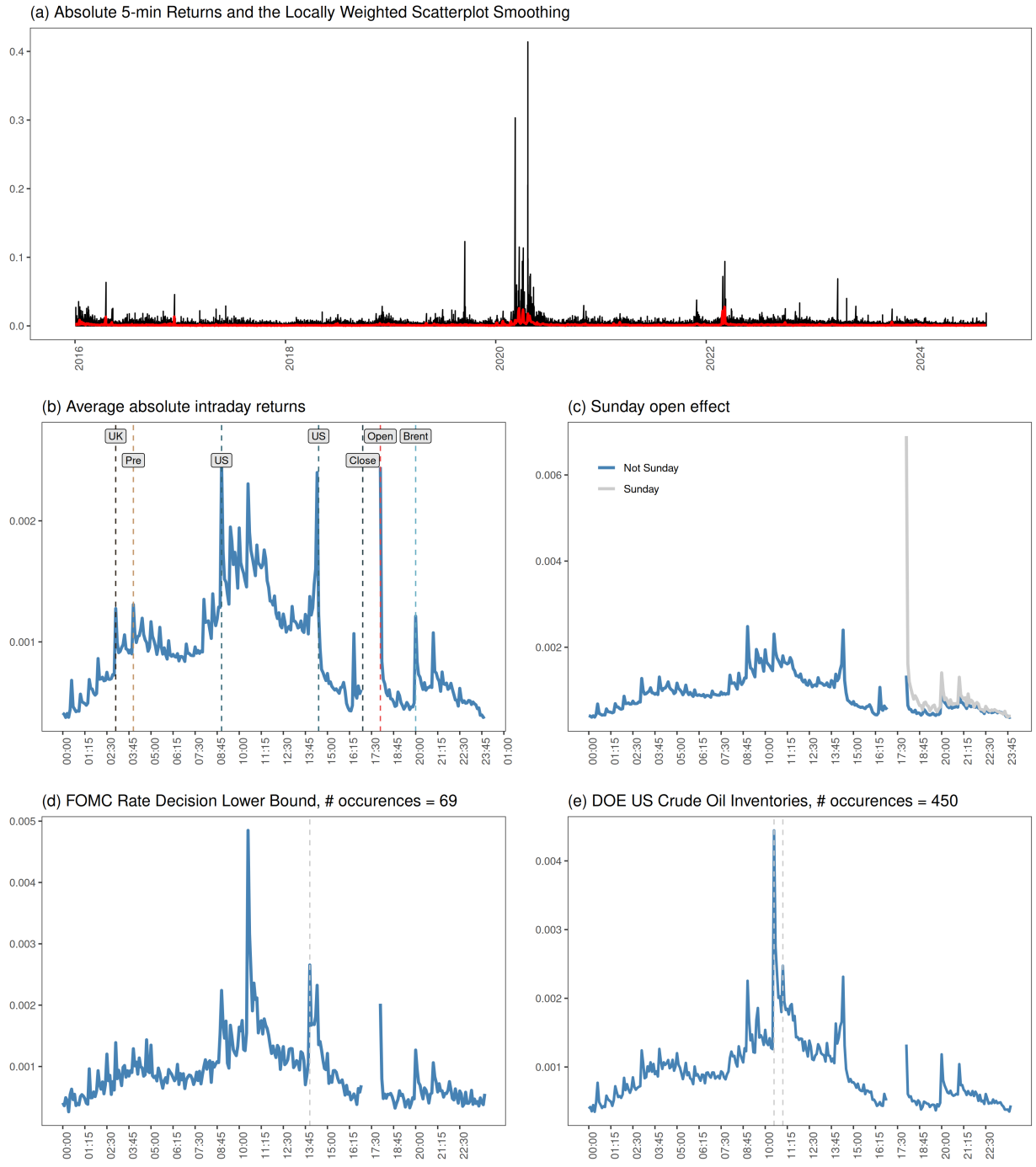

Figure 4: Slow and fast components using the summary statistics of the absolute intraday returns.

C Macroeconomic and commodity-related announcements

Table 8: Scheduled announcement list

Event	Count	Event	Count
CH BoP Current Account Balance	65	US Empire Manufacturing	103
CH Caixin China PMI Mfg	98	US Employment Cost Index	35
CH Caixin China PMI Services	100	US Existing Home Sales	104
CH CPI YoY	90	US Export Inspections - Soybeans	452
CH Exports YoY CNY	84	US FHFA House Price Index MoM	104
CH Foreign Reserves	75	US FOMC Meeting Minutes	70
CH Imports YoY	84	US FOMC Rate Decision (Lower Bound)	69
CH Industrial Production YTD YoY	93	US GDP Annualized QoQ	102
CH Industrial Profits YoY	60	US House Price Purchase Index QoQ	35
CH Money Supply M2 YoY	99	US Household Change in Net Worth	34
CH Non-manufacturing PMI	74	US Housing Starts	103
CH Swift Global Payments CNY	104	US Import Price Index MoM	104
GE CPI MoM	209	US Industrial Production MoM	103
GE Exports SA MoM	104	US Initial Jobless Claims	452
GE Factory Orders MoM	104	US ISM Manufacturing	104
GE GDP NSA YoY	79	US ISM Services Index	104
GE GfK Consumer Confidence	104	US JOLTS Job Openings	104
GE HCOB Germany Construction PMI	104	US Kansas City Fed Manf. Activity	103
GE HCOB Germany Manufacturing PMI	206	US Leading Index	104
GE HCOB Germany Services PMI	208	US MBA Mortgage Applications	444
GE IFO Business Climate	105	US MNI Chicago PMI	104
GE Import Price Index MoM	104	US Monthly Budget Statement	103
GE Industrial Production SA MoM	104	US NAHB Housing Market Index	104
GE PPI MoM	104	US New Home Sales	104
GE Retail Sales MoM	102	US NFIB Small Business Optimism	104
GE Unemployment Change (000's)	105	US Nonfarm Productivity	68
GE Wholesale Price Index MoM	104	US Pending Home Sales MoM	104
GE ZEW Survey Current Situation	104	US Personal Income	102
US ADP Employment Change	102	US Philadelphia Fed Business Outlook	104
US Agriculture Prices Received	104	US Poultry Slaughter LW YOY%	104
US Baker Hughes U.S. Rig Count	452	US PPI Final Demand MoM	104
US Business Inventories	103	US Progress - Spring Wheat Harvest	70
US Cattle on Feed-Placements YOY%	104	US Retail Sales Advance MoM	103
US Challenger Job Cuts YoY	104	US Richmond Fed Manufact. Index	104
US Change in Nonfarm Payrolls	102	US S&P Global US Manufacturing PMI	208
US Chicago Fed Nat Activity Index	103	US S&P Global US Services PMI	209
US Cold Storage- Beef	104	US Total Net TIC Flows	103
US Condition - Cotton	193	US Trade Balance	104
US Condition - Winter Wheat	170	US U. of Mich. Sentiment	208
US Conf. Board Consumer Confidence	104	US USDA Hogs & Pigs Inventory	33
US Consumer Credit	103	US USDA Quarterly All Wheat Stock	35
US CPI MoM	102	US USDA Red Meat Production	104
US Current Account Balance	34	US USDA Soybean Crush	104
US Dallas Fed Manf. Activity	104	US USDA Total Milk Production	104
US DOE U.S. Crude Oil Inventories	450	US WASDE Total Wheat Production	104
US Durable Goods Orders	208	US Wholesale Inventories MoM	197
US EIA Natural Gas Storage Change	451	Sunday Open	443

A complete list of scheduled macroeconomic and commodity-related announcements from the U.S., Germany, and China and the total number of occurrences from 2016-01-03 till 2024-08-30 (94 events in total). The data is obtained from the Bloomberg Economic Calendar. We keep only announcements that occur at least twice per year within the sample period. When multiple announcements share identical timestamps across the sample, we retain only the first occurrence to avoid perfect multicolinearity.

D Appendix: Posterior distributions of model parameters

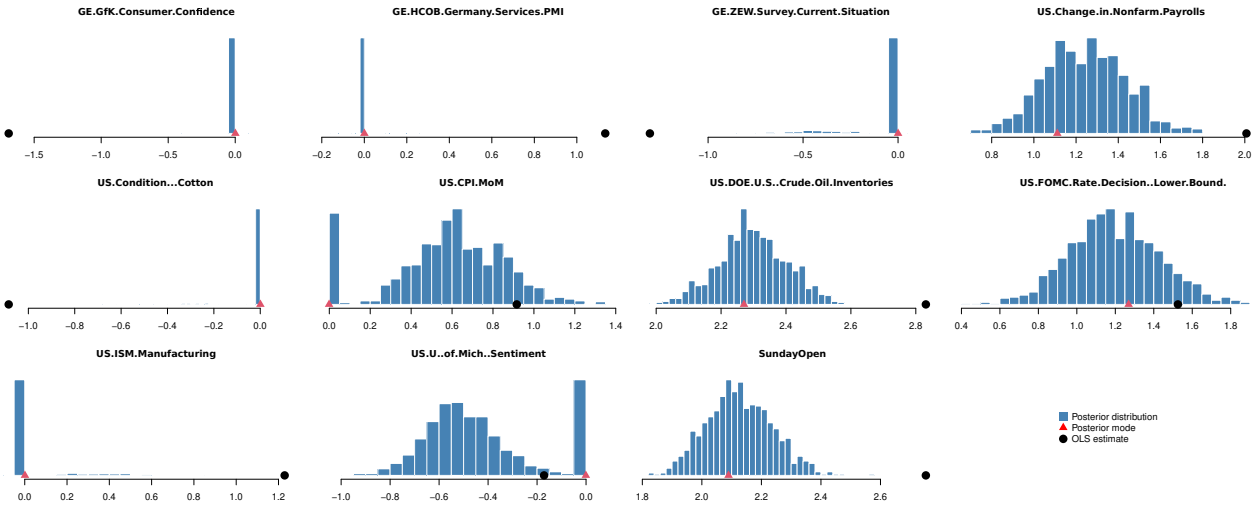

Figure 5: Posterior distributions for the selected events for the SSVA MIDAS^{ovx,vix} model. Red triangle is the posterior MCMC mode, black dot is the OLS estimate.