



WORKING PAPER 14/2025 (STATISTICS)

Volume-driven time-of-day effects in intraday volatility models

Igor Ferreira Batista Martins, Audronė Virbickaitė, Hoang Nguyen and Hedibert Freitas Lopes

ISSN 1403-0586

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

Volume-driven time-of-day effects in intraday volatility models

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 21, 2025

*Corresponding author. Email: igor.ferreira-batista-martins@oru.se. Address: Örebro University, Fakultetsgatan 1, 702 81, Örebro, Sweden. Igor Ferreira Batista Martins acknowledges financial support from the Jan Wallander and Tom Hedelius foundation and Tore Browalds foundation grant number BFv22-0005. A. 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.

Abstract

We propose a high-frequency stochastic volatility model that integrates persistent component, intraday periodicity, and volume-driven time-of-day effects. By allowing intraday volatility patterns to respond to lagged trading activity, the model captures economically and statistically relevant departures from traditional intraday seasonality effects. We find that the volume-driven component accounts for a substantial share of intraday volatility for futures data across equity indexes, currencies, and commodities. Out-of-sample, our forecasts achieve near-zero intercepts, unit slopes, and the highest R^2 values in Mincer-Zarnowitz regressions, while horse-race regressions indicate that competing forecasts add little information once our predictions are included. These statistical improvements translate into economically meaningful gains, as volatility-managed portfolio strategies based on our model consistently improve Sharpe ratios. Our results highlight the value of incorporating lagged trading activity into high-frequency volatility models.

Keywords: Intraday volatility, high-frequency, volume, periodicity.

JEL Classification: C11, C22, C53, C58

1 Introduction

Successful modeling and forecasting of financial volatility plays a critical role in risk management, portfolio allocation, derivative pricing, and market microstructure analysis. While daily and lower frequency volatility are largely driven by their own persistence, intraday volatility exhibits richer and more complex dynamics. In addition to the persistent component, high-frequency volatility follows a strong time-of-day pattern which is typically captured by a static dummy-based component forming a periodic effect, sometimes referred to as seasonal or diurnal (Engle and Sokalska, 2012; Stroud and Johannes, 2014; Rossi and Fantazzini, 2015; Bekierman and Gribisch, 2021; Martins and Lopes, 2025; Martins et al., 2025). However, such static structures cannot adapt to short-term deviations caused by market trading activity, and therefore miss an important source of intraday heterogeneity.

In order to address the limitations of previous approaches, we make the following methodological and empirical contributions. Methodologically, we propose a new high-frequency stochastic volatility model that embeds time-of-day effects within a volume-driven dynamic framework. The model decomposes time-of-day effects into two components: a static baseline that captures the recurring intraday pattern and a dynamic component that allows these periodic effects to adjust flexibly in response to deviations of lagged trading volume from its typical level. This volume-driven adaptation provides a parsimonious yet interpretable way to model economically meaningful departures from the periodic specification. As a result, volatility can respond dynamically to short-term market activities. We develop a Markov Chain Monte Carlo (MCMC) algorithm to estimate the proposed model and a particle filter to produce intraday volatility forecasts.

Empirically, we show the relevance and usefulness of this approach across multiple assets. First, our volume-driven component explains between 33% to 45% of intraday log-variance variation. Second, the dynamic time-of-day component significantly improves out-of-sample forecasting accuracy measured by squared and absolute losses. These improvements are statistically validated by Diebold-Mariano tests for predictive superiority. Third, beyond the statistical performance, the proposed model yields economically meaningful gains in volatility-managed portfolios highlighting the practical importance of volume-driven intraday heterogeneity in modern financial markets.

This paper relates to the literature highlighting the role of trading activity in volatility dynam-

ics, motivated by the well-established theoretical and empirical evidence linking trading volume and volatility. Classic frameworks such as the Mixture of Distributions Hypothesis (MDH) (Clark, 1973; Tauchen and Pitts, 1983; Andersen, 1996) and the Sequential Information Arrival Hypothesis (SIAH) (Copeland, 1976; Jones et al., 1994) indicate that trading volume reflects the rate of information flow into prices, which naturally generates a positive relation between volume and volatility. A detailed review of early literature on the return-volume relationship across financial markets is presented by Karpoff (1987).

MDH posits that volatility and trading volume are contemporaneously related, both driven by the rate of information flow arriving in the market. In this framework, contemporaneous trading volume serves as a proxy for the intensity of information arrivals, which in turn shapes volatility. Early evidence by Lamoureux and Lastrapes (1990) shows that when trading volume is included in volatility models, persistence parameters become small and statistically insignificant, suggesting that volume captures much of the serial correlation in volatility. In a stochastic volatility setting, Abanto-Valle et al. (2010) demonstrate that information about volatility can be inferred not only from returns but also from trading volume. Louhichi (2011) further decompose trading volume into the number and size of trades, showing that the well-documented positive relationship between volatility and volume is primarily driven by the number of trades. More recently, Slim and Dahmene (2016) confirm a strong positive contemporaneous relationship between total trading volume and realized volatility, particularly its continuous component. In line with the MDH, Ramos et al. (2020) show that salient information increases both trading volume and volatility, suggesting that the two are driven by a common information-based channel.

However, the contemporaneous focus of MDH overlooks potential lagged interactions between volume and volatility, which are central to the SIAH. According to SIAH, new information reaches investors sequentially, leading to a lead-lag relationship between trading volume and volatility. Trading volume tends to react first to new information, while volatility adjusts later as more market participants process the new information and respond to it. Multiple empirical findings support this view. Chiang et al. (2010) report that trading volume possesses strong predictive power for future volatility and that the relationship is nonlinear. Using HAR framework, Wang et al. (2015) show that lagged trading volume significantly improves forecasts of daily realized volatility. Similarly, Zheng et al. (2019) find that trading volume predicts both same-day and near-future peaks in

volatility. Yet, as Liu et al. (2023) point out, there remains no dominant conclusion regarding the forecasting role of volume, with roughly half of existing studies supporting SIAH and the other half consistent with MDH. More recent works show that volume remains a critical determinant of modern market dynamics, liquidity provision, price discovery, and volatility clustering even in today’s high-frequency, electronic trading environment (Rubia and Sanchis-Marco, 2013; Easley et al., 2016; Naimoli and Storti, 2019; Easley et al., 2021).

Our study contributes to the SIAH literature by providing novel high-frequency evidence on the predictive power of lagged trading volume for volatility. We show that lagged deviations of trading volume from its typical intraday pattern offer timely signals of changing market conditions. Importantly, the strength of this predictive relationship is heterogeneous: the estimated coefficients vary substantially across different trading hours, reflecting how the influence of information flow depends on which global markets are active. These findings extend the SIAH framework to the high-frequency domain and highlight the nonlinear nature of the volume-volatility relationship.

Closest to our setting is Bollerslev et al. (2018), who study the intraday interaction between volume and volatility around macroeconomic announcements. While their analysis highlights event-driven shifts, our focus is complementary. We concentrate on systematic time-of-day variation. Since each intraday observation is inherently tied to a trading interval, a volume-driven adjustment to these patterns provides a natural and broadly applicable channel through which trading activity shapes volatility throughout the trading day. Importantly, our framework is flexible enough to incorporate event-related effects in the spirit of Bollerslev et al. (2018), even though our emphasis here is on the time-of-day component.

More broadly, our work relates to a large literature on modeling intraday volatility. While it is possible to model the persistent component of intraday volatility using Generalized Autoregressive Conditional Heteroskedasticity (GARCH) as in Andersen and Bollerslev (1997, 1998); Andersen et al. (2007); Engle and Sokalska (2012), or realized volatility (RV) as in Dumitru et al. (2025), we opt to use a stochastic volatility (SV) approach as in Taylor (1986); Broto and Ruiz (2004). SV models treat volatility as an unobserved latent process, is consistent with continuous-time asset pricing, and generally yields smoother paths for volatility dynamics. Empirically, SV approaches have been shown to deliver superior out-of-sample performance relative to GARCH and RV, as documented in Stroud and Johannes (2014); Bekierman and Gribisch (2021); Martins and Lopes

(2025); Martins et al. (2025), among others.

The remainder of the paper is organized as follows. Section 2 introduces the data, the econometric model, and the Bayesian approach. Section 3 reports the in-sample evidence on the role of volume-driven time-of-day effects. Section 4 evaluates out-of-sample forecasting performance and presents the portfolio application. Section 5 concludes.

2 Data, econometric model and estimation approach

This section describes the data, econometric framework, and estimation strategy. We begin by introducing the futures contracts and intraday dataset used in this article. We then present our proposed stochastic volatility model that decomposes intraday return volatility into persistent, time-of-day, and Sunday opening components, with particular emphasis on a novel volume-driven time-of-day effect. Finally, we describe a Bayesian estimation approach that performs joint inference on model parameters and latent volatility states, relying on modern Bayesian forecasting techniques that have proven effective for nonlinear models (Martin et al., 2024).

2.1 Data

Our empirical analysis focuses on four highly liquid futures contracts: the E-mini S&P 500, the E-mini Nasdaq 100, the Euro FX futures, and West Texas Intermediate crude oil futures, or in short, S&P 500, Nasdaq, Euro and WTI. The E-mini equity index futures, traded electronically on CME Globex, are scaled-down versions of their full-sized counterparts, with notional values of 50 times and 20 times the underlying index for the S&P 500 and Nasdaq 100, respectively. The smaller size contracts lower margin requirements and trading costs, making E-minis the dominant vehicle for intraday trading and price discovery in U.S. equity index futures markets. The Euro FX futures contract, also traded on CME Globex, tracks the value of the Euro against the U.S. dollar. WTI crude oil futures, traded on NYMEX, are the most liquid energy futures contract worldwide and serve as the global benchmark for crude oil pricing. Together, these assets span equity, currency, and commodity markets, offering a rich variety of trading environments. All four contracts trade nearly around the clock, with sessions beginning on Sunday at 18:00 Eastern Time (ET) and ending on Friday at 17:00 ET, with one-hour break every day between 17:00-18:00. Historically, the E-mini

S&P 500 and Nasdaq 100 had an additional 15-minute daily maintenance break from 16:15 to 16:30 ET, but this restriction was lifted in June of 2021.

Our sample spans from January 3, 2016, to August 30, 2024. The data are split into an estimation sample ending on December 31, 2021, and an out-of-sample evaluation period covering the remaining observations. Our model uses 5-minute (log-)returns and trading volume in number of contracts. A 24-hour day consists of 288 five-minute intervals, although daylight saving time adjustments occasionally produce 23-hour or 25-hour days. Non-trading intervals are filled using the last available transaction price and assuming zero volume. One exception is when a trading break occurs. In such cases, all timestamps within the break are removed from our sample. Thus, most days consists of 276 observations. The data is constructed from continuous futures by linking successive most-liquid contracts from the CME group to form uninterrupted time series. Appendix A at the end of the manuscript contains the main descriptive statistics and plots related to our data.

2.2 Econometric model

We model univariate de-meaned 5-minute log-returns, y_t , as

$$y_t = \exp(h_t/2)\varepsilon_t, \quad \varepsilon_t \sim \mathbf{N}(0, 1), \quad \text{for } t = 1, \dots, T, \quad (1)$$

where h_t represents the conditional log-variance process of returns. We decompose the log-variance into five additive components,

$$h_t = m_0 + tod_t^p + tod_t^v + x_t + so_t, \quad (2)$$

where m_0 denotes the unconditional level of log-volatility, tod_t^p captures periodic time-of-day effects, tod_t^v captures volume-driven deviations in time-of-day effects, x_t is the persistent latent component, and so_t accounts for the effect of volatility shifts around the Sunday market opening ¹.

Let's denote k as the specific 5-minute intraday window corresponding to the timestamp t via the window indicator function $w(t) = k \in \{1, \dots, K\}$ and, since our assets are traded 23 hours a

¹The pronounced spike in volatility observed when markets reopen on Sundays reflects the accumulation of uncertainty and order imbalances over the weekend, see Martins et al. (2025) for empirical evidence regarding WTI.

day, $K = 276$. Let $\mathbf{I}_t$ be a K -dimensional vector that is 1 in the position $w(t)$ and zero otherwise. The periodic time-of-day component, tod_t^p , is represented as

$$tod_t^p = \mathbf{I}_t' \boldsymbol{\beta}^p \text{ with } \sum_{k=1}^K \beta_k^p = 0, \quad (3)$$

where the K -dimensional vector $\boldsymbol{\beta}^p = \{\beta_1^p, \dots, \beta_K^p\}'$ captures the fixed periodicity in intraday volatility. Since $\mathbf{I}_t$ includes all 5-minute windows of the 23 traded hours and our h_t specification in Equation (2) already includes the intercept m_0 , we impose the zero-sum restriction $\sum_{k=1}^K \beta_k^p = 0$ to ensure identification. This parameterization anchors the set of fixed intraday effects around zero, allowing m_0 to represent the average log-variance level.

Our main modeling novelty comes when modeling the volume-driven time-of-day effects tod_t^v . This component allows intraday effects to evolve over time by incorporating lagged trading activity. Consider the lagged trading window $k^* = w(t-1)$, let us denote the average log-volume for time-of-day k^* as $\bar{V}_{k^*}$ and, analogously, denote the time-of-day standard deviation as $\sigma_{V_{k^*}}$. The volume-driven time-of-day component tod_t^v is represented as

$$tod_t^v = v_{t-1} \mathbf{I}_t' \boldsymbol{\beta}^v \text{ with } v_{t-1} = \frac{V_{t-1} - \bar{V}_{k^*}}{\sigma_{V_{k^*}}}, \quad (4)$$

where v_{t-1} is the scaled deviation of lagged log-volume from its time-of-the day average $\bar{V}_{k^*}$ and the K -dimensional vector $\boldsymbol{\beta}^v$ captures how this deviation interacts with the intraday structure. The volume-driven component allows for the inclusion of new information to intraday volatility via lagged trading activity. When trading activity is unusually high or low, the volume-driven component allows the log-variance process to shift accordingly. Unlike $\boldsymbol{\beta}^p$, there is no need to impose a zero-sum restriction on $\boldsymbol{\beta}^v$, since v_{t-1} has mean-zero by construction. One could also consider v_{t-1} as a weighted sum of multiple lagged volume deviations. However, for simplicity, we consider deviations only from one lagged observation.

The fixed and dynamic time-of-day components can be viewed jointly as

$$tod_t = tod_t^p + tod_t^v = \mathbf{I}_t' (\boldsymbol{\beta}^p + v_{t-1} \boldsymbol{\beta}^v). \quad (5)$$

When volume is at its time-of-day average, $tod_t^v = 0$ and the specification collapses to the standard static form. However, when trading activity is unusually high or low, the volume-driven component allows more flexibility in the log-variance process to depart from the static baseline.

The persistent component follows a stochastic volatility process,

$$x_t = \phi x_{t-1} + \sigma_x \eta_t, \quad \eta_t \sim \mathbf{N}(0, 1), \quad (6)$$

with persistence parameter ϕ and innovation scale σ_x . This component captures the rate at which shocks to volatility decay.

Finally, we include adjustments for the Sunday market opening following Martins et al. (2025). This is modeled as

$$so_t = \mathbf{H}_t' \boldsymbol{\beta}^{so}, \quad (7)$$

where $\mathbf{H}_t$ is a m -dimensional indicator vector for the Sunday opening session and its lagged 5-minute trading windows, and the m -dimensional vector $\boldsymbol{\beta}^{so}$ captures the associated log-variance shift. Therefore, our model allows for shifts in volatility up to m lags after the Sunday opening event.

Equations (1) - (7) specify our model for intraday returns with volatility that combines a persistent latent factor, periodic and volume-driven time-of-day effects, and Sunday opening adjustments. The key innovation lies in the specification of the volume-driven time-of-day component, tod_t^v , which allows the strength of intraday periodic patterns to flexibly adjust to lagged market activity. This feature extends the traditional periodicity-based approach by linking intraday volatility directly to trading intensity, thereby capturing short-term deviations from static patterns.

Table 1 summarizes how our proposed specification nests several stochastic volatility models as special cases. These include (i) a vanilla SV model (SV), (ii) an SV with deterministic time-of-day effects (PSV), and (iii) an SV with deterministic time-of-day and Sunday-opening effects (PSVO). These benchmarks serve as natural points of comparison in our empirical analysis.

Table 1: Nested models

	Persistence (x_t)	Deterministic ToD (tod_t^p)	Sunday Open (so_t)	Volume ToD (tod_t^v)
Proposal	✓	✓	✓	✓
PSVO	✓	✓	✓	
PSV	✓	✓		
SV	✓			

The first column lists the model mnemonics, and the remaining columns indicate which components are included: a persistent stochastic volatility factor (x_t), deterministic time-of-day effects (tod_t^p), volume-driven time-of-day effects (tod_t^v) and Sunday-opening effects (so_t). Our proposed model includes all components, while the other specifications represent restricted variants.

2.3 Estimation approach

We adopt a Bayesian framework to estimate the model described in Section 2.2 by Equations (1) - (7). Our Bayesian approach simultaneously estimates all parameters and latent states, avoiding the need for potentially inefficient two-stage estimators and allowing for direct probabilistic statements on both parameters and latent volatility dynamics. Since the conditional likelihood involves high-dimensional latent states and cannot be integrated out in closed form, posterior inference relies on MCMC methods to generate draws from the joint posterior distribution of parameters and latent processes.

We follow the log-linearization strategy of Kim et al. (1998) to facilitate efficient sampling. Specifically, we rewrite (1) as

$$\log(y_t^2) = h_t + \log(\varepsilon_t^2), \quad (8)$$

and approximate $\log(\varepsilon_t^2)$ with a finite mixture of seven Gaussian components. The model can be expressed in a conditionally Gaussian state-space form. Intuitively, this transformation allows us to treat log-squared returns as a noisy signal of the latent log-variance. This representation is computationally convenient and enables the application of the Carter and Kohn (1994) and Frühwirth-Schnatter (1994) forward-filtering backward sampling (FFBS) algorithm for block updating of the latent states conditional on the mixture indicators.

Let $\Theta = (m_0, \phi, \sigma_x^2, \beta^p, \beta^v, \beta^{so}, \{x_t\}_{t=1}^T)$ be the set of model parameters and latent states. We employ standard proper, but not strongly informative, conjugate priors whenever possible. We assume Gaussian priors for m_0 , ϕ , β^p , β^v and β^{so} . We assume $m_0 \sim \mathbf{N}(\mu_{m_0} = \bar{y}^*, V_{m_0} = 2)$,

where $\bar{y}_t^* = \frac{1}{T} \sum_{t=1}^T \log(y_t^2) + 1.27$ reflecting that the average $\bar{y}_t^*$ should proxy the log-variance level while the constant 1.27 accounts for the expected value of $\log(\varepsilon_t^2)$, which is distributed as $\log-\chi_1^2$ (Ruiz, 1994). Due to the high persistence commonly observed in the stochastic volatility literature, we assume a prior $\phi \sim \mathbf{N}(\phi_0 = 0.95, V_{\phi_0} = 0.25)$. We take a conservative approach by assuming a prior belief of no time-of-day effects (neither fixed nor volume-driven) and no shift due to the Sunday opening while also allowing for reasonable variability by considering $\{\beta_k^p\}_{k=1}^{K-1} \sim \mathbf{N}(\beta_0^p = \mathbf{0}, \mathbf{V}_{\beta_0^p} = 0.5\mathbf{I}_{K-1})$, $\beta^v \sim \mathbf{N}(\beta_0^v = \mathbf{0}, \mathbf{V}_{\beta_0^v} = 0.5\mathbf{I}_K)$, and $\beta^{so} \sim \mathbf{N}(\beta_0^{so} = \mathbf{0}, \mathbf{V}_{\beta_0^{so}} = 4\mathbf{I}_m)$ where $\mathbf{I}$ is the identity matrix that corresponds to the dimensions of β^p , β^v and β^{so} . We assume an Inverse Gamma prior $\mathbf{IG}(\alpha_{\sigma_x} = 1, \beta_{\sigma_x} = 10)$ for σ_x^2 reflecting the restriction of non negative variance. Our choice of conditionally conjugate priors enables direct sampling from the conditional posterior distributions using a Gibbs sampler, see Appendix B at the end of the manuscript for details.

3 In-sample results

This section illustrates the volume-driven time-of-day effects on intraday volatility. We first quantify the relative importance of this channel through a variance decomposition of the latent log-variance, and then examine the dynamic effects of above-average volume on volatility throughout the trading day. We then turn to the remaining components of the log-variance decomposition described in Equation (2).

3.1 Volume-driven intraday heterogeneity

Table 2 reports the variance decomposition of the latent log-variance h_t , calculated at the posterior parameter modes, into its persistent, static and volume-driven time-of-day, and Sunday-opening components. We compute the decomposition as the ratio of the individual component contribution to the total variability of the estimated volatility.

For the proposed model, across almost all assets (with the exception of WTI), the volume-driven time-of-day component explains the largest share of variation, particularly for the Nasdaq where it accounts for more than 45 percent of the total. The static time-of-day factor and the persistent latent components also contribute meaningfully, while the Sunday-opening adjustment has a small percentage share overall which is consistent with its rare occurrence relative to the high-

Table 2: Variance decomposition

	Persistence (x_t)	Deterministic ToD (tod_t^p)	Sunday Open (so_t)	Volume ToD (tod_t^v)
Panel A: Nasdaq				
Proposal	25.34	29.14	0.10	45.42
PSVO	70.69	29.02	0.29	—
PSV	70.49	29.51	—	—
Panel B: S&P500				
Proposal	36.08	24.87	0.10	38.94
PSVO	75.73	23.98	0.29	—
PSV	76.38	23.62	—	—
Panel C: Euro				
Proposal	24.13	36.41	0.08	39.38
PSVO	57.92	41.89	0.19	—
PSV	58.29	41.71	—	—
Panel D: WTI				
Proposal	34.93	31.98	0.11	32.98
PSVO	65.11	34.51	0.38	—
PSV	65.56	34.44	—	—

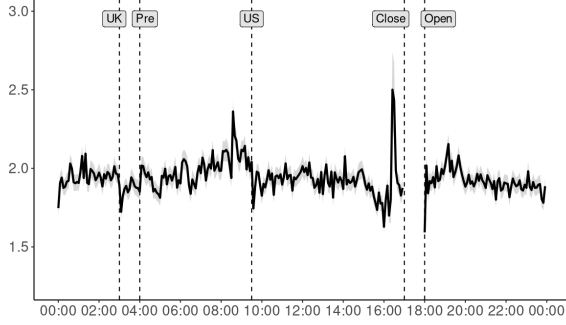
This table reports the variance decomposition of the log-variance (h_t) at the posterior parameter modes into the latent persistent component (x_t), periodic effect (tod_t^p), volume-driven time-of-day component (tod_t^v), and Sunday-opening adjustment (so_t). Entries are normalized percentage shares.

frequency sampling of intraday volatility. However, the Sunday open has an effect on volatility that is statistically significant and economically meaningful.

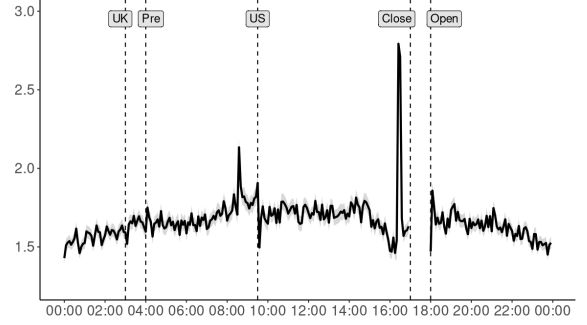
When comparing the variance decomposition with that of the restricted models, two notable features emerge: the reallocation of variability and the stability of the remaining components. In particular, once the volume-driven component is included, almost all variability is transferred from the persistent component. This leads to a dramatic reduction in the variance of the variance parameter (as reported in Appendix C), indicating a more stable volatility process. The improvement reflects a meaningful reallocation of explained variability from unobserved persistence to observed lagged trading volume, providing a more interpretable and data-driven structure. This, in turn, translates into notable gains in one-step-ahead forecasting accuracy, as demonstrated in the subsequent sections. Moreover, the share of explained variability remains largely stable, particularly for the deterministic time-of-day effects, suggesting that incorporating lagged trading volume enhances model flexibility without distorting the core volatility dynamics. A plot illustrating the component-wise volatility decomposition for a small subsample can be seen in Appendix E.

Figure 1 shows the posterior means of the volume-driven time-of-day effects in the standard deviation scale, $\exp(\beta^v/2)$, together with their 90 percent credible intervals, across Nasdaq, S&P 500, Euro, and WTI futures (Panels a–d). In all cases, $\exp(\beta^v/2)$ is positive and significantly above one, confirming that higher-than-typical lagged trading activity systematically amplifies volatility. Moreover, the evolution of the estimated coefficients clearly exhibits non-constant patterns, indicating that the effect of lagged volume on volatility varies over the trading day. This finding aligns with the broader body of research on the relationship between trading volume and asset returns. For example, McMillan (2007) highlight the nonlinear nature of this relationship and provide an extensive overview of related contributions in the literature.

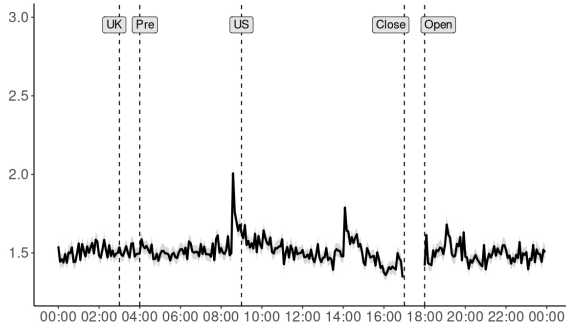
The figure also reveals an economically meaningful intraday structure. Almost all assets (except for the Euro) show a pronounced decline in the impact of above-average trading volume at the 18:00 ET market reopening following the daily trading break. This pattern can be attributed to two factors. First, the rise in volatility observed at reopening is largely captured by a spike in the periodic component (see Figure 2). Second, at that point, the most recent lagged-volume information comes from the previous trading session, which, from an economic perspective, should have only a limited influence on volatility at the start of a new session. Spikes are also evident



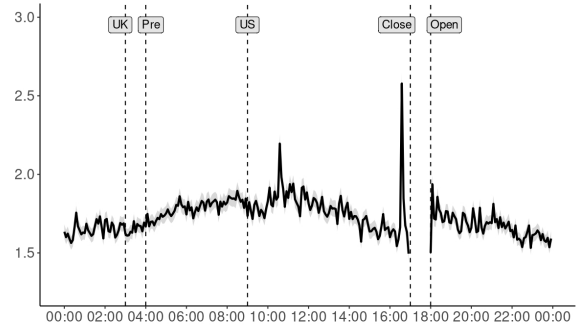
(a) Nasdaq



(b) S&P 500



(c) Euro



(d) WTI

Figure 1: Volume-driven time-of-day effects

The plots show the posterior means of the volume-driven time-of-day effects in the standard deviation scale, $\exp(\beta^v/2)$, with 90% credible intervals, for Nasdaq, S&P 500, Euro, and WTI futures (Panels a–d). Values above one indicate that above-average trading volume amplifies volatility. Vertical dashed lines represent the following daily events: UK stock market opening (UK), US opening auction (Pre), US opening (US), future's close (Close) and future's opening (Open).

around key market times: for Nasdaq, S&P 500, and WTI futures, effects vary sharply near the U.S. equity close (16:00 ET), while Euro futures display a sharp increase following the ECB fixing (8:15 ET). These patterns suggest that the interaction between intraday trading activity and volatility is closely tied to global market coordination, emphasizing the importance of a dynamic, volume-adjusted specification.

Overall, the in-sample analysis of volume-driven time-of-day effects shows that intraday fluctuations in trading activity play a central role in shaping high-frequency volatility across all four assets. These findings highlight the importance of accounting for lagged volume-related effects in volatility modeling.

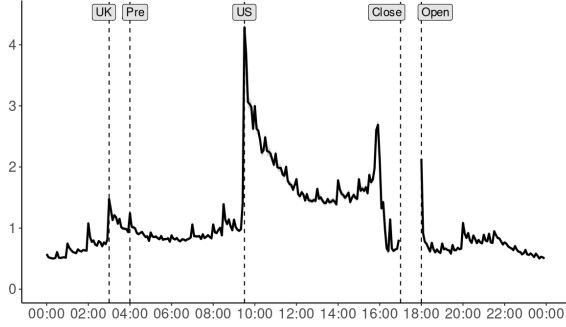
3.2 Static time-of-day, Sunday-opening and persistent volatility effects

Figures 2 and 3 illustrate the posterior means of the static time-of-day and Sunday-opening effects, on the standard deviation scale, $\exp(\beta^p/2)$ and $\exp(\beta^{so}/2)$, along with 90% credible intervals, for Nasdaq, S&P 500, Euro, and WTI futures (Panels a–d).

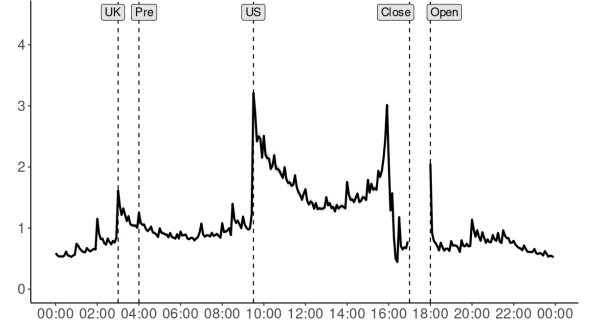
Panels a) and b) of Figure 2 display the well-documented U-shaped intraday pattern during regular U.S. trading hours (9:30 – 16:00 ET) for Nasdaq and S&P 500 futures (Andersen and Bollerslev, 1997, 1998; Stroud and Johannes, 2014; Martins et al., 2025). Panels c) and d) show similar U-shaped patterns for Euro and WTI futures, spanning from the London open (3:00 ET) to early U.S. market trading hours and the historical pit trading hours (9:00 to 14:30 ET), respectively. In all cases, peaks coincide with major market openings and closings, highlighting the role of scheduled trading activity in shaping intraday volatility.

Figure 3 indicates a strong Sunday-opening effect with volatility rising nearly to 1.5-1.75 times its baseline level in the first 5-minutes. However, this effect quickly dissipates within 30 minutes from the opening.

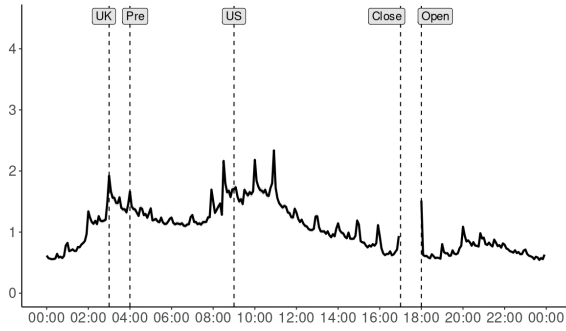
Finally, Figure 4 shows the posterior means of the log-variance on the standard deviation scale, $\exp(\mathbf{h}/2)$, for each asset reflecting the combined effect of the persistence, static and volume-driven time-of-day and Sunday-opening components. In all cases, the COVID19 pandemic is associated with a clear spike in volatility. For the Euro, however, the pandemic was not the peak volatility episode. Instead, the highest levels were observed in 2016–2017, plausibly reflecting the Brexit referendum, the U.S. presidential election, and elevated Eurozone political risk. Parameter estimation



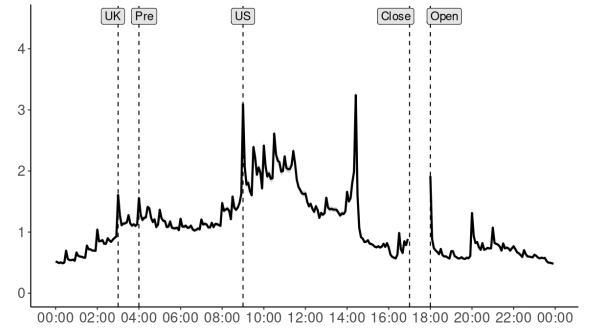
(a) Nasdaq



(b) S&P 500



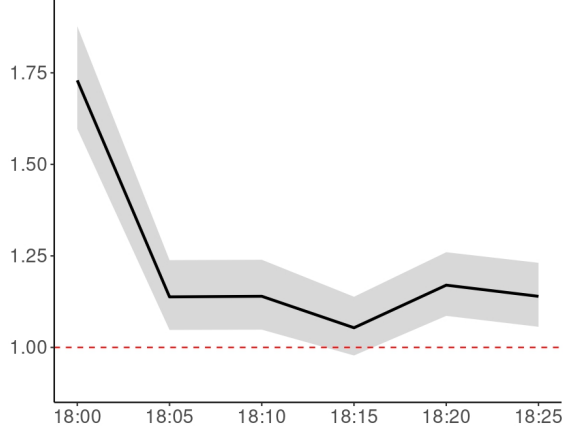
(c) Euro



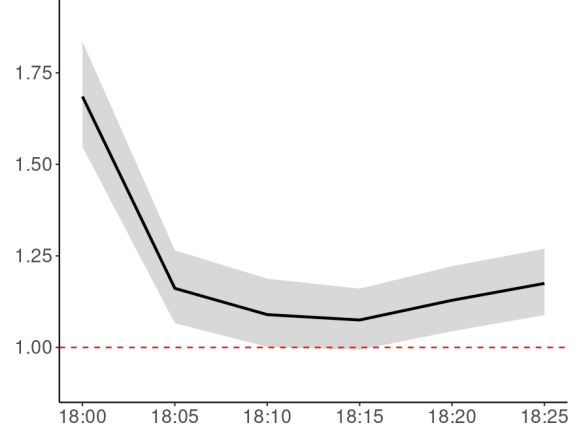
(d) WTI

Figure 2: Static time-of-day effects

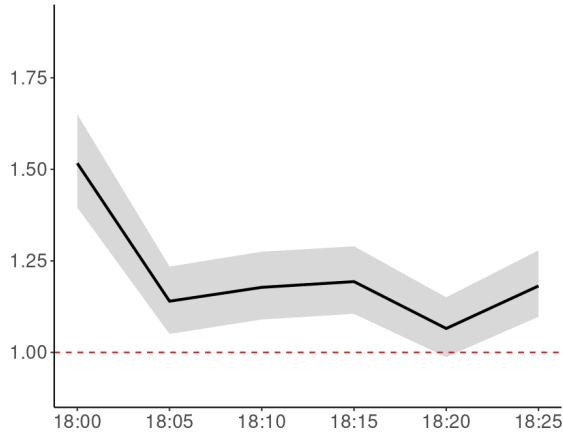
Posterior means of the static time-of-day effects in the standard deviation scale, $\exp(\beta^p/2)$, with 90% credible intervals, for Nasdaq, S&P 500, Euro, and WTI futures (Panels a–d). A value of 1.5 indicates that volatility is 1.5 times is baseline level. Vertical dashed lines represent the following daily events: UK stock market opening (UK), US opening auction (Pre), US opening (US), future's close (Close) and future's opening (Open).



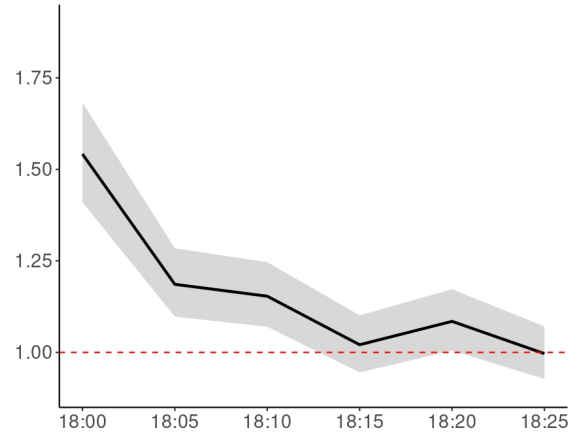
(a) Nasdaq



(b) S&P 500



(c) Euro



(d) WTI

Figure 3: Sunday-opening effects

Posterior means of the Sunday-opening effects in the standard deviation scale, $\exp(\beta^{so}/2)$, with 90% credible intervals, for Nasdaq, S&P 500, Euro, and WTI futures (Panels a–d).

results for the proposed model and SV-based counterparts can be found in Appendix C at the end of the manuscript.

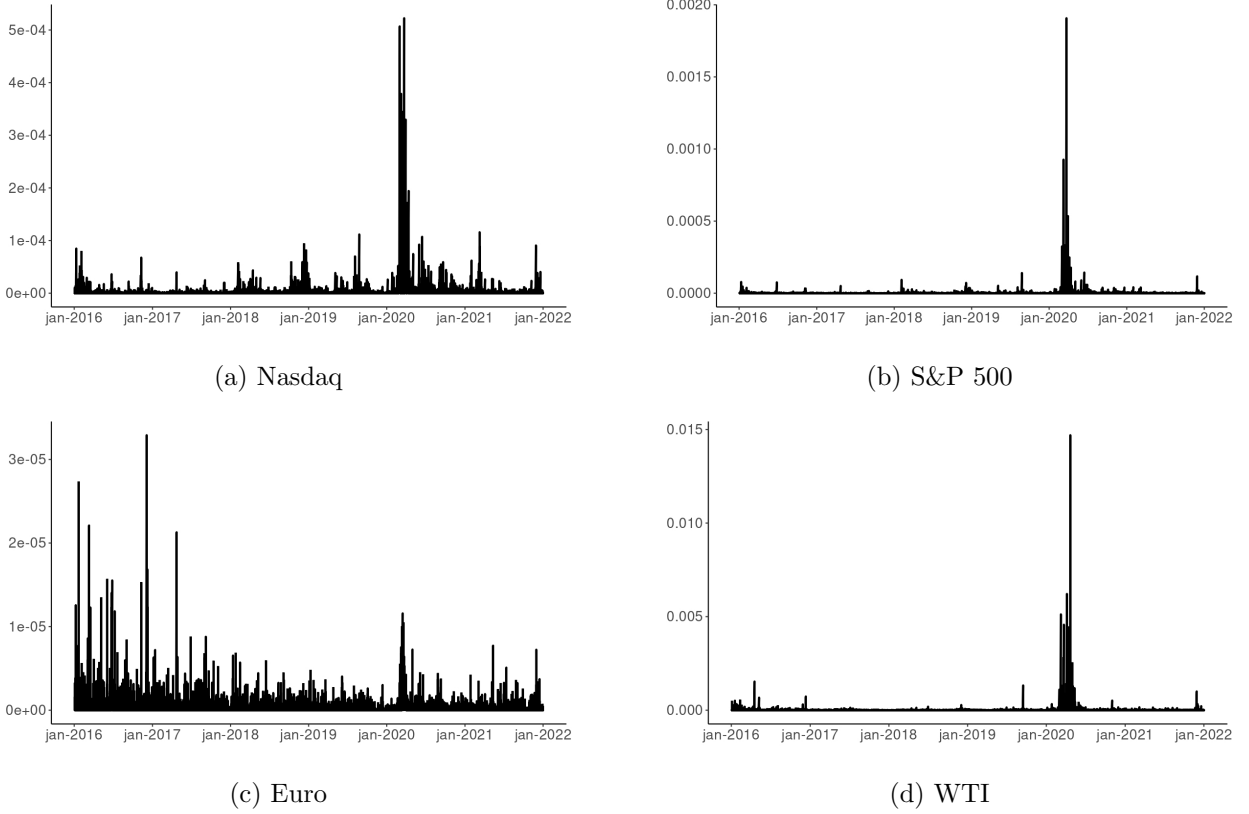


Figure 4: Estimated volatility

Estimated volatility calculated at the posterior parameter modes using the proposed stochastic volatility model, $\exp(\mathbf{h}/2)$, for Nasdaq, S&P 500, Euro, and WTI futures (Panels a–d).

4 Out-of-sample results and applications

Although achieving a good in-sample fit is important, the true measure of a model’s value lies in its predictive performance and practical applicability. Our results demonstrate that the proposed model performs exceptionally well on both fronts. We evaluate the one-step-ahead out-of-sample performance of the four stochastic volatility models described in Table 1, a GARCH(1,1) and a simple autoregressive realized volatility model (AR1–RV) for Nasdaq, S&P 500, Euro, and WTI crude oil during the period from January 1, 2022 to August 30, 2024. For the SV models specified in Table 1, we fix parameters on their in-sample posterior modes and rely on a particle filter

Gordon et al. (1993) to update latent states. For GARCH and AR1–RV, we fix parameters on their maximum likelihood estimates.

4.1 Volatility forecasts

We evaluate the one-step-ahead out-of-sample predictive performance using three complementary approaches. First, we consider the Mincer–Zarnowitz (MZ) regression, which provides a direct test of forecast unbiasedness and efficiency. Specifically, 5-minute realized volatility, calculated as the square root of the sum of squared one-minute returns during the corresponding 5-minute window, is regressed on a one-step-ahead forecast produced by a competing model:

$$rvol_t = \alpha_0 + \alpha_1 \widehat{vol}_{t|t-1}^{model} + \varepsilon_t, \quad \varepsilon_t \sim \mathbf{N}(0, \sigma^2),$$

where $\widehat{vol}_{t|t-1}^{model}$ denotes the forecasted volatility for period t made at time $t - 1$. Ideally, a model should produce $\alpha_0 = 0$ and $\alpha_1 = 1$ with high explanatory power as measured by R^2 .

Second, we implement a horse-race regression to directly compare the forecasts of the proposed specification against those of competing models. The regression is given by

$$rvol_t = \beta_0 + \beta_1 \widehat{vol}_{t|t-1}^{proposal} + (1 - \beta_1) \widehat{vol}_{t|t-1}^{competitor} + \varepsilon_t, \quad \varepsilon_t \sim \mathbf{N}(0, \sigma^2),$$

where the competitor is any alternative benchmark. If $\beta_1 = 1$, the proposal dominates, while $\beta_1 = 0$ implies that only the competitor contains all predictive information. Values of β_1 greater than one indicate that the competitor’s forecast contributes negatively once the proposal is included, further reinforcing the superiority of the proposed approach.

Finally, we complement these regression-based diagnostics with the Diebold–Mariano (DM) test of equal predictive ability, using both squared and absolute forecast errors. The DM statistic evaluates whether differences in forecast accuracy across models are statistically significant, with negative values favoring the proposal. Together, these metrics provide a comprehensive assessment of the forecasting performance and their statistical significance.

Table 3 reports the combined results across all assets. Columns 2–4 present the MZ regression estimates of α_0 , α_1 , and the associated R^2 . Columns 5–6 show the estimates of β_1 and its t -statistic

from the horse-race regressions. Columns 7–10 report average squared and absolute forecast errors (scaled by 10^{-3}) together with DM statistics relative to the proposal.

Table 3: Forecasting results

Model	Mincer–Zarnowitz			Horse–Race		Squared		Absolute	
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	R^2	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	Avg. $\times 10^{-3}$	DM	Avg. $\times 10^{-3}$	DM
Panel A: Nasdaq									
Proposal	-0.001	1.027	0.748	–	–	1.045	–	16.995	–
PSVO	0.005	0.901	0.553	1.167	396.075	1.880	-17.976	22.219	-106.798
PSV	0.006	0.896	0.546	1.153	402.542	1.912	-17.553	22.309	-105.609
SV	0.005	0.921	0.459	1.158	477.689	2.254	-21.750	24.771	-135.357
GARCH	0.005	0.913	0.421	1.107	503.270	2.414	-24.391	25.669	-135.500
AR1-RV	0.008	0.914	0.409	1.023	506.604	2.469	-22.543	25.759	-127.506
Panel B: S& P500									
Proposal	0.000	1.033	0.733	–	–	0.629	–	13.507	–
PSVO	0.005	0.896	0.550	1.214	375.134	1.070	-15.950	17.004	-99.825
PSV	0.005	0.894	0.544	1.197	380.797	1.086	-14.677	17.086	-99.179
SV	0.004	0.914	0.470	1.196	445.672	1.251	-17.575	18.729	-128.443
GARCH	0.003	0.949	0.433	1.134	466.908	1.330	-19.409	19.328	-129.089
AR1-RV	0.004	0.931	0.414	1.030	475.494	1.377	-19.013	19.881	-128.651
Panel C: Euro									
Proposal	0.000	0.999	0.652	–	–	0.157	–	7.581	–
PSVO	0.000	1.014	0.423	1.018	353.922	0.261	-16.876	8.624	-52.681
PSV	0.000	1.011	0.419	1.022	356.776	0.262	-16.867	8.635	-53.019
SV	0.001	0.994	0.373	1.001	390.275	0.283	-17.460	9.018	-66.215
GARCH	0.002	0.878	0.341	0.985	416.399	0.301	-19.840	9.614	-87.306
AR1-RV	0.002	0.983	0.293	0.965	443.572	0.320	-22.009	9.848	-91.742
Panel D: WTI									
Proposal	0.009	0.966	0.689	–	–	3.162	–	34.909	–
PSVO	0.012	0.896	0.470	0.900	373.737	5.406	-6.725	42.125	-71.654
PSV	0.013	0.888	0.453	0.905	387.348	5.578	-5.864	42.278	-71.132
SV	0.013	0.898	0.390	0.945	430.984	6.201	-6.702	45.334	-95.958
GARCH	0.018	0.779	0.356	0.934	471.039	6.875	-8.161	48.518	-109.681
AR1-RV	0.017	0.849	0.304	0.922	493.537	7.113	-7.960	49.060	-112.704

The table reports combined results for Mincer–Zarnowitz regressions, Horse-race regressions, and Diebold–Mariano tests for one-step-ahead out-of-sample forecasting exercise for January 1, 2022 to August 30, 2024. The volatility proxy employed is the 5-minute realized volatility.

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 8-10 report the Diebold–Mariano (DM) statistics compare predictive accuracy against the Proposal under squared and absolute loss. Negative and statistically significant DM statistics indicate superior performance of the benchmark.

The MZ results show that the proposed model is consistently the best performing across all

markets. In every case, $\hat{\alpha}_0$ is closer to zero and $\hat{\alpha}_1$ closer to unity than for competing models, and R^2 values are substantially higher. For instance, the R^2 ranges from 0.65 for Euro futures to 0.75 for Nasdaq, while the best competitor typically delivers values below 0.55. These results indicate that the proposal produces forecasts that are both the closest of being unbiased and more informative about realized volatility.

The horse-race regressions reinforce this finding. Across all assets, $\hat{\beta}_1$ lies very close to one and is in some cases above unity, implying that once the forecast from the proposal is included, the competitor adds little or even negative information. Taken together, the horse-race regressions show that the proposal systematically outperforms its restricted variants (including traditional stochastic volatility), GARCH, and RV-based benchmarks.

The DM tests provide formal evidence that these gains are statistically significant. Using both squared and absolute forecast errors, the null hypothesis of equal predictive accuracy is overwhelmingly rejected in favor of the proposal. The statistics are large in magnitude and consistently negative, indicating that the proposed model achieves strictly lower forecast losses across equity, currency, and commodity markets.

Overall, the evidence across all tests and assets consistently supports the superior performance of including lagged volume-driven time-of-day effects. As a robustness check, we have checked the forecasting performance using a different proxy for the volatility, namely, absolute 5-minute returns, and the obtained results are identical, see Appendix D at the end of the manuscript for details.

4.2 Volatility-managed portfolios

An important question is whether the statistical gains in volatility forecasting documented above translate into economically meaningful improvements. Moreira and Muir (2017) proposed a volatility-managed portfolio strategy that directly exploits time-variation in expected risk. The approach scales the portfolio’s exposure to the risky asset inversely with its forecasted volatility, thereby reducing leverage in high-volatility periods and increasing exposure in low-volatility periods. Intuitively, if volatility is predictable, then risk-adjusted returns, as measured by the Sharpe ratio, can be improved by dynamically adjusting risk rather than holding a constant exposure.

Formally, the return on the volatility-managed portfolio, y_t^{vm} , in each period t is given by

$$y_t^{vm} = \frac{c}{vol_{t|t-1}^{model}} y_t,$$

where, as before, $vol_{t|t-1}^{model}$ denotes the forecasted volatility for period t made at time $t - 1$, y_t is the asset return and c is a constant chosen to ensure y_t^{vm} and y_t have the same unconditional volatility².

Table 4 reports the performance of volatility-managed portfolios across Nasdaq, S&P 500, Euro and WTI futures. For each asset, we present average annualized returns (in percent) and annualized Sharpe ratios for the unmanaged (buy and hold) benchmark and for volatility-managed strategies based on competing volatility models.

Table 4: Volatility-managed portfolios

	Nasdaq		S&P 500		Euro		WTI	
	<i>Avg</i>	<i>Sharpe</i>	<i>Avg</i>	<i>Sharpe</i>	<i>Avg</i>	<i>Sharpe</i>	<i>Avg</i>	<i>Sharpe</i>
Not Managed	0.03	0.14	0.04	0.19	-0.03	-0.37	0.07	0.17
Proposal	0.79	3.28	0.64	3.52	0.00	0.05	0.32	0.81
PSVO	0.38	1.56	0.30	1.61	-0.02	-0.28	0.20	0.50
PSV	0.38	1.56	0.30	1.61	-0.03	-0.31	0.24	0.60
SV	0.22	0.90	0.21	1.15	-0.04	-0.52	0.12	0.31
GARCH	0.12	0.49	0.12	0.64	-0.05	-0.56	0.10	0.25
AR1-RV	0.22	0.93	0.14	0.74	-0.04	-0.46	0.12	0.31

The table reports annualized average returns (*Avg*) and Sharpe ratios for unmanaged portfolios and for volatility-managed strategies based on competing volatility models. Results are shown for Nasdaq, S&P 500, Euro, and WTI futures.

Two key patterns emerge. First, volatility management delivers large economic gains relative to the unmanaged portfolios: Sharpe ratios increase markedly, particularly for equity index futures. For example, in the S&P 500 and Nasdaq, the Sharpe ratios of the unmanaged portfolios are 0.19 and 0.13, respectively, but rise above 3.0 under our proposed specification. Second, across all assets, the proposal dominates competing approaches (PSVO, PSV, SV, GARCH, and AR1-RV), yielding consistently higher average returns and Sharpe ratios. These improvements reflect the proposal's ability to better time volatility and hence scale exposure more effectively.

The benefits are especially pronounced in equity index futures, where risk timing appears most valuable, but meaningful gains are also present in commodities (WTI) and currencies (Euro). In contrast, traditional SV and GARCH-based strategies yield relatively modest improvements,

²As explained in Moreira and Muir (2017), the choice of c does not affect the Sharpe ratio.

while AR1-RV often fail to deliver positive economic gains. Overall, Table 4 highlights that more accurate high-frequency volatility forecasts translate directly into superior portfolio performance when applied in a volatility-managed framework.

5 Conclusion

This paper develops a framework for modeling high-frequency volatility that jointly accounts for persistent behavior, static time-of-day adjustments, and volume-driven time-of-days effects. By conditioning intraday periodicity on past trading volume, our model captures timely and economically relevant deviations from static volatility patterns, providing a more accurate and microstructure-consistent description of intraday risk dynamics.

Across equity indexes, currency, and commodity futures, the volume-driven component explains a large share of intraday volatility variation, highlighting the importance of conditioning on trading activity. The model not only reproduces well-known U-shaped volatility patterns but also uncovers features associated with global trading activity. It indicates that the deviations from the time-of-day average volume are deeply tied to volatility. Our results complement the Sequential Information Arrival Hypothesis (SIAH) literature by showing that the lead-lag relationship between trading volume and volatility also holds in a high-frequency setting. Moreover, we find that the strength of this relationship is time-varying: information flows shape return volatility differently depending on which market is open at a given time. Nonetheless, the relationship remains consistently statistically significant.

Out-of-sample forecast comparisons provide strong evidence in favor of the proposed specification. In Mincer–Zarnowitz regressions, our model yields intercepts close to zero, slopes close to one, and the highest explanatory power across all assets. Horse-race regressions confirm that once our forecasts are included, those from competing models contribute little to none additional information. Moreover, Diebold–Mariano tests overwhelmingly reject equal predictive accuracy and favor our specification under both squared and absolute loss functions. The results are robust even when different volatility proxies are used. Together, these findings indicate that the model provides the most accurate forecasts of realized volatility among the alternatives considered.

From an economic perspective, these statistical improvements translate into superior portfolio

performance. When applied to volatility-managed investment strategies, our forecasts consistently improve Sharpe ratios relative to unmanaged benchmarks and competing volatility models. The benefits are most striking in equity index futures, where dynamic exposure based on our specification more than triples the Sharpe ratio compared to GARCH or standard stochastic volatility alternatives. Overall, our results suggest that lagged trading volume provides a critical dimension for understanding and forecasting intraday volatility.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to check for grammar errors and improve readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

A Appendix: Descriptive statistics and plots

Table 5 reports the summary statistics of five-minute log-returns and trading volumes for the Nasdaq, S&P 500, Euro, and WTI futures over the sample period from January 2016 to August 2024. As expected, returns exhibit near-zero means with substantial excess kurtosis and negative skewness (except for Euro), reflecting the leptokurtic and asymmetric nature of high-frequency financial returns. Unconditional volatility (as captured by standard deviation) is highest for WTI and lowest for the Euro. Trading volumes vary considerably across assets, with the S&P 500 showing the largest average volume, consistent with its market depth and liquidity.

Table 5: Summary statistics

	Mean	Std	Q2	Min	Max	Skew	Kurt	n_{ins}	n_{oos}
Panel A: Nasdaq									
Return (percent)	0.000	0.084	0.000	-4.304	2.828	-0.953	85.482	420713	189182
Volume (1,000 contracts)	1.693	2.758	0.509	0.000	44.085	3.130	16.748	–	–
Panel B: S&P 500									
Return (percent)	0.000	0.069	0.000	-4.194	2.782	-1.199	139.143	420746	189182
Volume (1,000 contracts)	5.444	9.793	1.369	0.000	275.823	4.823	50.257	–	–
Panel C: Euro									
Return (percent)	0.000	0.029	0.000	-1.045	1.531	0.246	61.937	425077	190208
Volume (1,000 contracts)	0.644	0.896	0.386	0.000	44.343	6.304	91.629	–	–
Panel D: WTI									
Return (percent)	0.000	0.196	0.000	-41.388	30.118	-14.365	5672.061	424924	189443
Volume (1,000 contracts)	1.297	2.155	0.563	0.000	93.090	5.659	70.710	–	–

The table reports summary statistics of the 5-minute log-returns (in percentage) and trading volumes (1,000 contracts) for two market indices (Nasdaq and S&P 500), exchange rate (Euro) and commodity (WTI) from January 3, 2016 to August 30, 2024. n_{ins} and n_{oos} stand for the number of observations in-sample and out-of-sample.

Figure 5 complements the descriptive statistics table by illustrating the joint dynamics of absolute returns (in black, scale on the left) and trading volumes (1,000 contracts, in gray, inverted scale in the right). Both series display pronounced clustering, particularly around major market stress episodes such as the 2020 volatility spike. The mirror plotting permits to illustrate the close co-movement between return volatility and trading activity: periods of heightened volatility tend to coincide with elevated trading volumes.

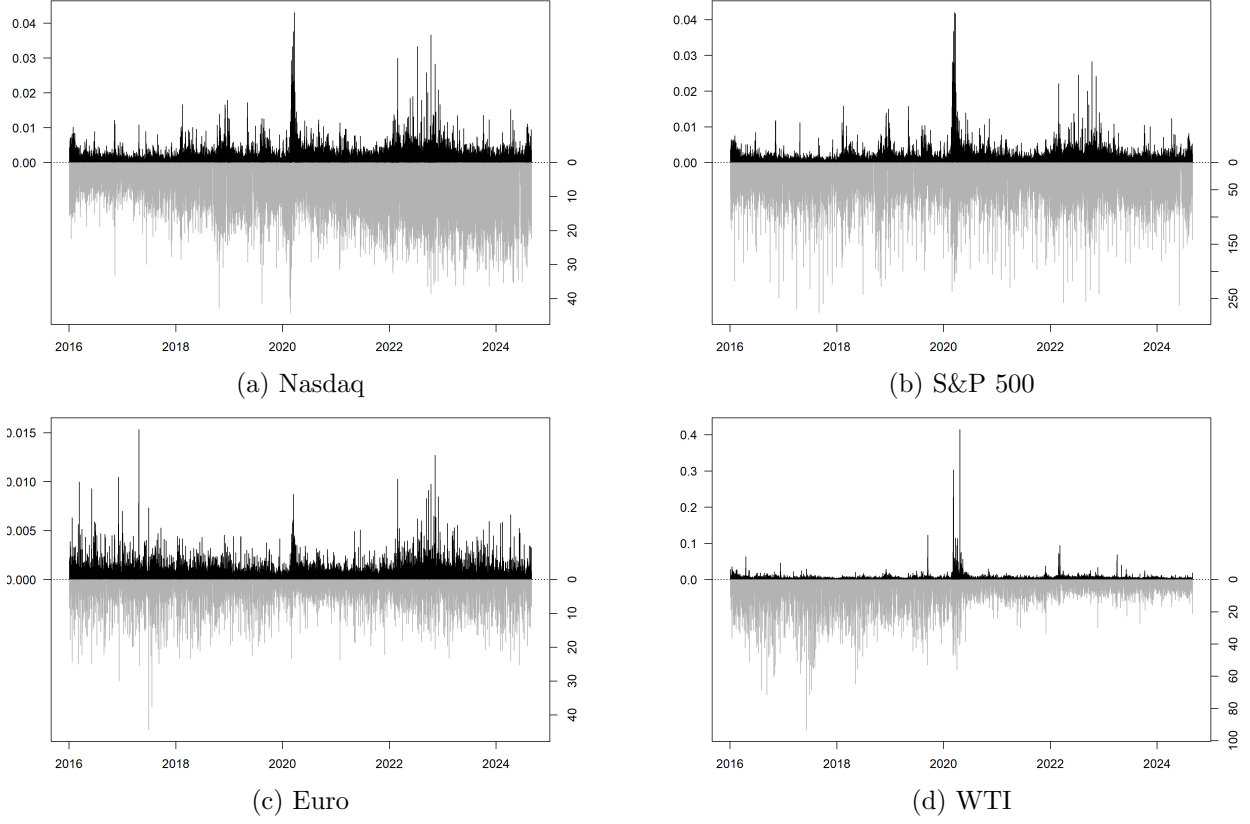


Figure 5: Descriptive plots

The plots draw the evolution of absolute 5-minute log-returns, in black (scale on the left) and volume, in gray (1,000 contracts, inverted scale on the right) from January 3, 2016 to August 30, 2024 for Nasdaq, S&P 500, Euro and WTI (Panels a-d).

B Appendix: MCMC algorithm

Our proposed model described in Section 2.2 is defined by Equations (1) - (7). To facilitate efficient sampling, we apply log-linearization to Equation (1) and approximate $u_t = \log(\varepsilon_t^2)$ with a finite mixture of seven Gaussian components (Kim et al., 1998). A summary of the equations required for sampling as follows,

$$\begin{aligned}\log(y_t^2) &= h_t + u_t \text{ with } u_t \sim \mathbf{N}(\mu_{c,t}, \sigma_{c,t}^2), \\ h_t &= m_0 + x_t + tod_t^p + tod_t^v + so_t, \\ x_t &= \phi x_{t-1} + \sigma_x \eta_{x,t} \text{ with } \eta_{x,t} \sim \mathbf{N}(0, 1), \\ tod_t^p &= \mathbf{I}_t' \boldsymbol{\beta}^p \text{ with } \sum_{k=1}^K \beta_k^p = 0, \\ tod_t^v &= v_{t-1} \mathbf{I}_t' \boldsymbol{\beta}^v, \\ so_t &= \mathbf{H}_t' \boldsymbol{\beta}^{so}.\end{aligned}$$

Let $\boldsymbol{\Theta} = (m_0, \phi, \sigma_x^2, \boldsymbol{\beta}^p, \boldsymbol{\beta}^v, \boldsymbol{\beta}^{so}, \{x_t\}_{t=1}^T)$ be the set of model parameters and latent states. And let $\boldsymbol{\Psi}$ be all the parameters in $\boldsymbol{\Theta}$ except the ones we sample from in a given step of the MCMC chain. For each iteration the MCMC scheme samples from the following full conditional posterior distributions.

1) Sample $\phi | \boldsymbol{\Psi}, \mathbf{y}$ from a conjugate Gaussian distribution. We consider the persistent equation as a linear regression, $x_t \sim \phi x_{t-1}$. Then the conditional posterior distribution of ϕ is a conjugate Gaussian distribution given a Gaussian prior and a Gaussian likelihood. We have that

$$\begin{aligned}\phi | \boldsymbol{\Psi}, \mathbf{y} &\sim \mathbf{N}(\hat{\phi}, \hat{V}_\phi), \text{ where} \\ \hat{V}_\phi &= \left(\frac{1}{V_{\phi_0}} + \frac{1}{\sigma_x^2} \sum_{t=1}^T x_{t-1} x_{t-1} \right)^{-1}, \\ \hat{\phi} &= \hat{V}_\phi \left(\frac{\phi_0}{V_{\phi_0}} + \frac{1}{\sigma_x^2} \sum_{t=1}^T x_t x_{t-1} \right).\end{aligned}$$

2) Sample $\sigma_x^2 | \boldsymbol{\Psi}, \mathbf{y}$ from a conjugate inverse Gamma distribution. We use the linear representation of the persistent component. The conditional posterior distribution of σ_x^2 is a conjugate inverse

Gamma distribution due to an inverse Gamma prior and a Gaussian likelihood. We have that

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

3) Sample the mixture component $\{c_t, \mu_{c,t}, \sigma_{c,t}^2\}_{t=1}^T | \Psi$ following (Kim et al., 1998). We first sample $\{c_t\}_{t=1}^T$ from the posterior probability

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

where $\mathbb{P}(c_t = c)$ is the mixture weight of the c -th Gaussian component with mean μ_c and variance σ_c^2 (Kim et al., 1998). Then 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) Sample $\{x_t\}_{t=1}^T | \Psi, \mathbf{y}$ following in Carter and Kohn (1994) and Frühwirth-Schnatter (1994). We rewrite the model in terms of measurement and state equations as,

$$\begin{aligned} \log y_t^2 - m_0 - tod_t^p - tod_t^v - so_t &= x_t + u_t, \\ x_t &= \phi x_{t-1} + \sigma_x \eta_{x,t}. \end{aligned}$$

We sample the conditional posterior distribution of $\{x_t\}_{t=1}^T$ from a Gaussian linear state space model using the Forward Filtering Backward Sampling (FFBS) algorithm.

5) Sample $m_0 | \Psi, \mathbf{y}$ from a conjugate Gaussian distribution. We rewrite the log-linearized model as a linear regression,

$$\log y_t^2 - x_t - tod_t^p - tod_t^v - so_t = m_0 + u_t.$$

The conditional posterior distribution of m_0 is a Gaussian distribution due to a Gaussian prior and a Gaussian likelihood. We have that

$$\begin{aligned} m_0 | \Psi, \mathbf{y} &\sim \mathbf{N}(\hat{\mu}_{m_0}, \hat{V}_{m_0}), \\ \hat{V}_{m_0} &= \left(V_{m_0}^{-1} + \sum_{t=1}^T \frac{1}{\sigma_{c,t}^2} \right)^{-1}, \\ \hat{\mu}_{m_0} &= \hat{V}_{m_0} \left(V_{m_0}^{-1} \mu_{m_0} + \sum_{t=1}^T \frac{1}{\sigma_{c,t}^2} (\log y_t^2 - x_t - tod_t^p - tod_t^v - so_t - \mu_{c,t}) \right). \end{aligned}$$

6) Sample $\{\beta_k^p\}_{k=1}^{K-1}|\Psi, \mathbf{y}$ from a conjugate Gaussian distribution. We rewrite the log-linearized model as a linear regression,

$$\begin{aligned}\log y_t^2 - m_0 - tod_t^v - so_t &= \mathbf{I}_t' \boldsymbol{\beta}^p + u_t, \\ &= \sum_{k=1}^{K-1} (I_{t,k} - I_{t,K}) \beta_k^p + u_t.\end{aligned}$$

Then the conditional posterior distribution of $\{\beta_k^p\}_{k=1}^{K-1}$ is a conjugate Gaussian distribution due to a Gaussian prior and a Gaussian likelihood. We have that,

$$\begin{aligned}\{\beta_k^p\}_{k=1}^{K-1}|\Psi, \mathbf{y} &\sim \mathbf{N}(\hat{\boldsymbol{\beta}}^p, \hat{\mathbf{V}}_{\beta^p}), \\ \hat{\mathbf{V}}_{\beta^p} &= \left(\mathbf{V}_{\beta_0^p}^{-1} + \sum_{t=1}^T \frac{1}{\sigma_{c_t}^2} \tilde{\mathbf{I}}_t' \tilde{\mathbf{I}}_t \right)^{-1}, \\ \hat{\boldsymbol{\beta}}^p &= \hat{\mathbf{V}}_{\beta^p} \left(\mathbf{V}_{\beta_0^p}^{-1} \boldsymbol{\beta}_0^p + \sum_{t=1}^T \frac{1}{\sigma_{c_t}^2} \tilde{\mathbf{I}}_t' (\log y_t^2 - m_0 - tod_t^v - so_t - \mu_{c,t}) \right),\end{aligned}$$

where $\tilde{\mathbf{I}}_t = (I_{t,1} - I_{t,K}, \dots, I_{t,K-1} - I_{t,K})$, and set $\beta_K^p = -\sum_{k=1}^{K-1} \beta_k^p$.

7) Sample $\beta^v|\Psi, \mathbf{y}$ from a conjugate Gaussian distribution. We rewrite the log-linearized model as a linear regression,

$$\log y_t^2 - m_0 - tod_t^p - so_t = v_{t-1} \mathbf{I}_t' \boldsymbol{\beta}^v + u_t.$$

Then the conditional posterior distribution of $\boldsymbol{\beta}^v$ is a conjugate Gaussian distribution due to a Gaussian prior and a Gaussian likelihood.

$$\begin{aligned}\boldsymbol{\beta}^v|\Psi, \mathbf{y} &\sim \mathbf{N}(\hat{\boldsymbol{\beta}}^v, \hat{\mathbf{V}}_{\beta^v}), \\ \hat{\mathbf{V}}_{\beta^v} &= \left(\mathbf{V}_{\beta_0^v}^{-1} + \sum_{t=1}^T \frac{1}{\sigma_{c_t}^2} v_{t-1}^2 \mathbf{I}_t' \mathbf{I}_t \right)^{-1}, \\ \hat{\boldsymbol{\beta}}^v &= \hat{\mathbf{V}}_{\beta^v} \left(\mathbf{V}_{\beta_0^v}^{-1} \boldsymbol{\beta}_0^v + \sum_{t=1}^T \frac{1}{\sigma_{c_t}^2} v_{t-1} \mathbf{I}_t' (\log y_t^2 - m_0 - tod_t^p - so_t - \mu_{c,t}) \right).\end{aligned}$$

8) Sample $\beta^{so}|\Psi, \mathbf{y}$ from a conjugate Gaussian distribution. We rewrite the log-linearized model as a linear regression,

$$\log y_t^2 - m_0 - tod_t^p - tod_t^v = \mathbf{H}_t' \boldsymbol{\beta}^{so} + u_t.$$

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

$$\begin{aligned}\beta^{so}|\Psi, \mathbf{y} &\sim \mathbf{N}(\hat{\beta}^{so}, \hat{\mathbf{V}}_{\beta^{so}}), \\ \hat{\mathbf{V}}_{\beta^{so}} &= \left(\mathbf{V}_{\beta_0^{so}}^{-1} + \sum_{t=1}^T \frac{1}{\sigma_{c_t}^2} \mathbf{H}_t' \mathbf{H}_t \right)^{-1}, \\ \hat{\beta}^{so} &= \hat{\mathbf{V}}_{\beta^{so}} \left(\mathbf{V}_{\beta_0^{so}}^{-1} \beta_0^{so} + \sum_{t=1}^T \frac{1}{\sigma_{c_t}^2} \mathbf{H}_t' (\log y_t^2 - m_0 - tod_t^p - tod_t^v - \mu_{c,t}) \right).\end{aligned}$$

C Additional estimation results

Table 6 reports the posterior modes of the estimated model parameters for the proposed model and its restricted specifications across the four asset classes. Across all assets, the persistence parameter ϕ is close to unity, indicating a high degree of volatility persistence typical of high-frequency financial data. Since there are 276 coefficients in β^p and β^v vectors, and 6 more in β^{so} vector, we report the minimum and maximum posterior modes for each vector. Interestingly, the inclusion of volume-driven time-of-day effects considerably reduces the variance of the volatility process σ_x , resulting in more stable volatility dynamics. The rest of the parameters remain relatively similar across models. Overall, the results confirm that incorporating trading volume and time-of-day information enhances model flexibility without distorting the core volatility structure.

Table 6: Estimated model parameters

	m_0	ϕ	σ_x	$\min(\beta^f)$	$\max(\beta^f)$	$\min(\beta^{so})$	$\max(\beta^{so})$	$\min(\beta^v)$	$\max(\beta^v)$
Panel A: Nasdaq									
Proposal	-15.848	0.997	0.006	-1.375	2.904	0.101	1.098	0.956	1.792
PSVO	-15.688	0.986	0.051	-1.359	2.737	0.594	2.368	-	-
PSV	-15.684	0.986	0.052	-1.358	2.728	-	-	-	-
SV	-15.657	0.975	0.132	-	-	-	-	-	-
Panel B: S&P 500									
Proposal	-16.204	0.996	0.007	-1.610	2.332	0.144	1.047	0.715	1.589
PSVO	-16.076	0.986	0.056	-1.305	2.209	0.508	2.378	-	-
PSV	-16.068	0.985	0.058	-1.308	2.211	-	-	-	-
SV	-16.073	0.974	0.138	-	-	-	-	-	-
Panel C: Euro									
Proposal	-17.191	0.996	0.004	-1.209	1.691	0.120	0.847	0.665	1.438
PSVO	-17.077	0.970	0.039	-1.210	1.953	0.236	1.567	-	-
PSV	-17.079	0.970	0.040	-1.209	1.945	-	-	-	-
SV	-17.077	0.969	0.071	-	-	-	-	-	-
Panel D: WTI									
Proposal	-14.686	0.998	0.005	-1.445	2.345	-0.005	0.855	0.903	1.904
PSVO	-14.178	0.975	0.065	-1.428	2.239	0.338	2.503	-	-
PSV	-14.176	0.975	0.068	-1.436	2.236	-	-	-	-
SV	-14.193	0.965	0.148	-	-	-	-	-	-

The table reports the posterior modes of the estimated model parameters for the proposed model and its restricted specifications across the four asset classes.

D Robustness check

As a robustness check for the forecasting exercise, we employ a different proxy for the volatility: instead of the square root of the realized variance, we consider the absolute value of the 5-minute return, see Table 7. The model ordering remains the same, with the Proposal being vastly superior as compared to its restricted variants and other standard benchmarks.

Table 7: Forecasting results: Robustness check

Model	Mincer–Zarnowitz			Horse–Race		Squared		Absolute	
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	R^2	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	Avg. $\times 10^{-3}$	DM	Avg. $\times 10^{-3}$	DM
Panel A: Nasdaq									
Proposal	-0.006	0.943	0.490	-	-	2.823	-	31.850	-
PSVO	0.004	0.765	0.309	1.384	293.043	3.964	-21.907	37.305	-105.919
PSV	0.004	0.760	0.305	1.358	295.675	3.996	-21.371	37.390	-104.369
SV	0.004	0.771	0.250	1.253	319.475	4.218	-22.786	38.696	-112.132
GARCH	0.004	0.768	0.231	1.156	324.796	4.323	-23.751	39.333	-109.830
AR1-RV	0.006	0.773	0.228	1.051	321.870	4.275	-21.013	38.278	-85.750
Panel B: S&P 500									
Proposal	-0.004	0.941	0.468	-	-	1.675	-	24.882	-
PSVO	0.003	0.763	0.307	1.430	276.387	2.281	-20.414	28.661	-102.956
PSV	0.003	0.760	0.302	1.400	278.412	2.298	-19.280	28.744	-102.199
SV	0.003	0.768	0.255	1.296	299.577	2.406	-20.060	29.639	-110.099
GARCH	0.001	0.802	0.238	1.167	297.640	2.430	-20.008	30.009	-106.839
AR1-RV	0.002	0.789	0.229	1.046	300.188	2.455	-19.788	30.064	-96.058
Panel C: Euro									
Proposal	-0.003	0.930	0.409	-	-	0.389	-	13.140	-
PSVO	0.000	0.850	0.215	1.118	253.551	0.510	-17.474	14.369	-58.896
PSV	0.000	0.846	0.213	1.122	255.654	0.512	-17.483	14.379	-59.105
SV	0.000	0.829	0.188	1.062	269.859	0.527	-17.227	14.582	-62.531
GARCH	0.001	0.730	0.171	1.047	288.372	0.565	-21.770	15.496	-94.749
AR1-RV	0.001	0.818	0.147	0.978	292.442	0.549	-19.088	14.924	-66.439
Panel D: WTI									
Proposal	-0.008	0.924	0.466	-	-	7.584	-	57.715	-
PSVO	0.007	0.753	0.246	1.087	294.715	11.082	-10.445	67.913	-93.122
PSV	0.008	0.744	0.235	1.080	301.904	11.248	-8.954	68.051	-91.973
SV	0.009	0.744	0.197	1.074	320.596	11.656	-9.047	69.765	-102.277
GARCH	0.013	0.648	0.182	1.045	344.360	12.847	-11.621	75.028	-130.779
AR1-RV	0.013	0.702	0.153	0.989	345.214	12.381	-9.697	72.466	-107.081

The table reports combined results for Mincer–Zarnowitz regressions, Horse-race regressions, and Diebold–Mariano tests for one-step-ahead out-of-sample forecasting exercise for January 1, 2022 to August 30, 2024. The volatility proxy employed is the absolute value of 5-minute returns.

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 8-10 report the Diebold–Mariano (DM) statistics compare predictive accuracy against the Proposal under squared and absolute loss. Negative and statistically significant DM statistics indicate superior performance of the benchmark.

E Volatility components

Figure 6 presents the total and component-wise volatility comparison between the proposed model and its restricted variant, PSVO, for a small subsample (two weeks) to facilitate readability. The proposed model is shown in gray, and the restricted PSVO model in black. The top plot displays the total volatility on a standard deviation scale. We observe that the estimated overall volatility is very similar across both models, confirming that both specifications capture the same underlying process.

The bottom four plots show the component-wise volatility decomposition for both models. The most striking differences appear in the persistence component, x_t . The proposed model estimates a smooth path for persistence, while the restricted model produces extremely noisy trajectories. These discrepancies arise because the proposed model incorporates the volume-driven component, which captures much of the variability through lagged log-volume shown in the bottom right plot. When this component is omitted, all variability is absorbed by the persistence component, making it more volatile and noisy, leading to less precise forecasts.

The remaining volatility components - namely, the periodic time-of-day (tod_t^p) and Sunday-open (so_t) effects - are nearly identical across both models. This further confirms that incorporating lagged trading volume enhances model flexibility without distorting the core volatility dynamics. In essence, the addition of the lagged volume to the model redistributes hard-to-predict uncertainty from the persistence term to the more predictable deterministic lagged-volume component.

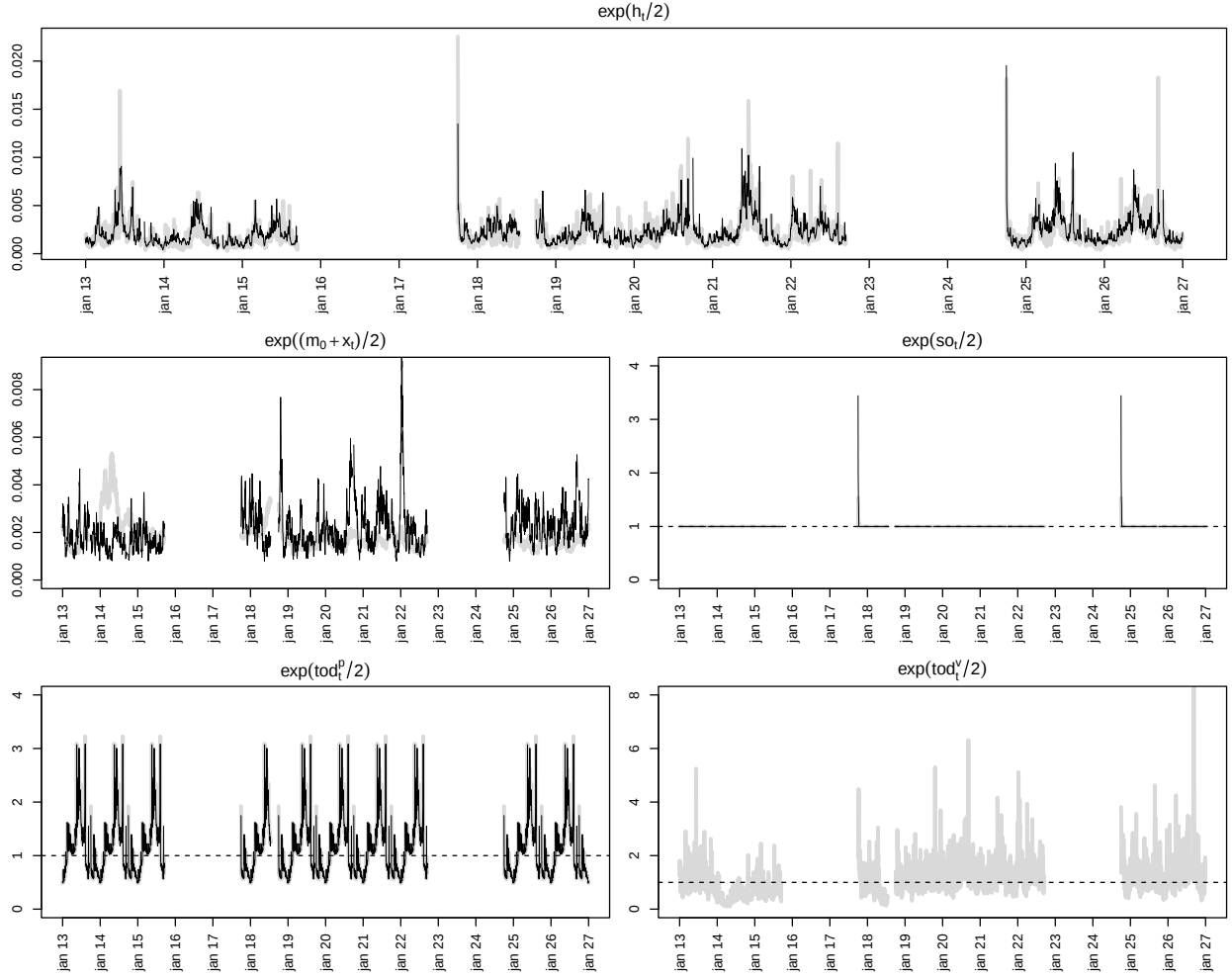


Figure 6: Estimated volatility for the proposed model (thick gray line) and the restricted PSVO variant (thin black line) for 2 weeks in January 2016. The top plot draws the overall volatility in the standard deviation scale, the rest of the plots draw the volatility components, also in the standard deviation scale.

References

- C. A. Abanto-Valle, H. S. Migon, and H. F. Lopes. Bayesian modeling of financial returns: A relationship between volatility and trading volume. *Applied Stochastic Models in Business and Industry*, 26(2):172–193, 2010.
- T. G. Andersen. Return volatility and trading volume: An information flow interpretation of stochastic volatility. *The Journal of Finance*, 51(1):169–204, 1996.
- 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.
- 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.
- T. Bollerslev, J. Li, and Y. Xue. Volume, volatility, and public news announcements. *The Review of Economic Studies*, 85(4):2005–2041, 2018.
- C. Broto and E. Ruiz. Estimation methods for stochastic volatility models: A survey. *Journal of Economic surveys*, 18(5):613–649, 2004.
- C. K. Carter and R. Kohn. On Gibbs sampling for state space models. *Biometrika*, 81(3):541–553, 1994.
- T. C. Chiang, Z. Qiao, and W.-K. Wong. New evidence on the relation between return volatility and trading volume. *Journal of Forecasting*, 29(5):502–515, 2010.
- P. K. Clark. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, pages 135–155, 1973.

- T. E. Copeland. A model of asset trading under the assumption of sequential information arrival. *The Journal of Finance*, 31(4):1149–1168, 1976.
- A. M. H. Dumitru, R. Hizmeri, and M. Izzeldin. Forecasting the realized variance in the presence of intraday periodicity. *Journal of Banking & Finance*, 170:107342, 2025.
- D. Easley, M. L. De Prado, and M. O’Hara. Discerning information from trade data. *Journal of Financial Economics*, 120(2):269–285, 2016.
- D. Easley, M. López de Prado, M. O’Hara, and Z. Zhang. Microstructure in the machine age. *The Review of Financial Studies*, 34(7):3316–3363, 2021.
- 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.
- N. J. Gordon, D. J. Salmond, and A. F. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *IEE proceedings F (radar and signal processing)*, volume 140, pages 107–113. IET, 1993.
- C. M. Jones, G. Kaul, and M. L. Lipson. Transactions, volume, and volatility. *The Review of Financial Studies*, 7(4):631–651, 1994.
- J. M. Karpoff. The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis*, 22(1):109–126, 1987.
- 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.
- C. G. Lamoureux and W. D. Lastrapes. Heteroskedasticity in stock return data: Volume versus GARCH effects. *The Journal of Finance*, 45(1):221–229, 1990.
- M. Liu, W.-C. Choo, C.-C. Lee, and C.-C. Lee. Trading volume and realized volatility forecasting: Evidence from the China stock market. *Journal of Forecasting*, 42(1):76–100, 2023.

- W. Louhichi. What drives the volume–volatility relationship on Euronext Paris? *International Review of Financial Analysis*, 20(4):200–206, 2011.
- G. M. Martin, D. T. Frazier, W. Maneesoonthorn, R. Loaiza-Maya, F. Huber, G. Koop, J. Maheu, D. Nibbering, and A. Panagiotelis. Bayesian forecasting in economics and finance: A modern review. *International Journal of Forecasting*, 40(2):811–839, 2024.
- I. Martins and H. F. Lopes. What events matter for exchange rate volatility? *The Quarterly Review of Economics and Finance*, 104:102073, 2025.
- I. Martins, A. Virbickaitė, H. Nguyen, and H. F. Lopes. Fast and slow level shifts in intraday stochastic volatility. *Scandinavian Working Papers No 2025:12*, pages 1–37, 2025.
- D. G. McMillan. Non-linear forecasting of stock returns: Does volume help? *International Journal of Forecasting*, 23(1):115–126, 2007.
- A. Moreira and T. Muir. Volatility-managed portfolios. *The Journal of Finance*, 72(4):1611–1644, 2017.
- A. Naimoli and G. Storti. Heterogeneous component multiplicative error models for forecasting trading volumes. *International Journal of Forecasting*, 35(4):1332–1355, 2019.
- S. B. Ramos, P. Latoeiro, and H. Veiga. Limited attention, salience of information and stock market activity. *Economic Modelling*, 87:92–108, 2020.
- E. Rossi and D. Fantazzini. Long memory and periodicity in intraday volatility. *Journal of Financial Econometrics*, 13(4):922–961, 2015.
- A. Rubia and L. Sanchis-Marco. On downside risk predictability through liquidity and trading activity: A dynamic quantile approach. *International Journal of Forecasting*, 29(1):202–219, 2013.
- E. Ruiz. Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of Econometrics*, 63(1):289–306, 1994.
- S. Slim and M. Dahmene. Asymmetric information, volatility components and the volume–volatility relationship for the CAC40 stocks. *Global Finance Journal*, 29:70–84, 2016.

- 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.
- G. E. Tauchen and M. Pitts. The price variability-volume relationship on speculative markets. *Econometrica*, pages 485–505, 1983.
- S. J. Taylor. *Modelling financial time series*. Chichester: Wiley, 1986.
- X. Wang, C. Wu, and W. Xu. Volatility forecasting: The role of lunch-break returns, overnight returns, trading volume and leverage effects. *International Journal of Forecasting*, 31(3):609–619, 2015.
- Z. Zheng, J. Gui, Z. Qiao, Y. Fu, H. E. Stanley, and B. Li. New dynamics between volume and volatility. *Physica A: Statistical Mechanics and its Applications*, 525:1343–1350, 2019.