A Fuzzy Framework for Realized Volatility Prediction: Empirical Evidence from Equity Markets

Shafqat Iqbal[a], Štefan Lyócsa[a,b,*]

[a] *Faculty of Economics and Administration, Masaryk University, Brno, Czech Republic*
[b] *Institute of Economic Research, Slovak Academy of Sciences, Bratislava, Slovak Republic*

Abstract

This study introduces a fuzzy-based volatility forecasting model. The proposed approach, RV-FTS, applies a fuzzy c-means clustering algorithm to estimate (time-varying) c latent volatility states and their corresponding membership degrees. These memberships are used to construct a fuzzified volatility estimate as a weighted average of cluster centroids. The final volatility forecast is generated through an exponential weighted moving average mechanism that combines the most recent fuzzified volatility estimate and the previous forecast, governed by the given smoothing parameter ρ. The two hyperparameters are estimated using a rolling-window cross-validation approach. Our empirical study is based on volatility forecasts of 14 major stock market indices with more than 20 years of data. We predict one- to twenty-two-day-ahead volatility forecasts and compare the RV-FTS model with standard volatility model benchmarks, the GARCH, ARFIMA, AR and HAR models, and conditional combination forecasts. We find that in the short-term, day-ahead setting, the RV-FTS tends to outperform the GARCH, ARFIMA and AR model across most markets. When used for combination forecasts, the RV-FTS model contributes to greater forecast accuracy. Multiple-day ahead forecasts show that with increasing forecasting horizons, the utility of the RV-FTS model increases, as the conditional combinations that includes RV-FTS tends to outperform conditional combinations that do not include RV-FTS. Moreover, we find that for all markets, there are periods where the weight of the RV-FTS model in the conditional combination of eight models reaches 50% or more. These results show that the RV-FTS model offers competitive volatility forecasting advantages, particularly for longer forecast horizons.

Keywords: Volatility Forecasting, Realized Variance, Fuzzy Time Series, Fuzzy Clustering, Forecast Combination, HAR

1. Introduction

Traded assets in financial markets are subject to price changes, reflecting the inherent uncertainty in the value of future cash inflows for asset holders (e.g., capital gains, interest payments, and dividends). Understanding and predicting such price variation, i.e., volatility, can help investors in asset allocation tasks, thereby improving portfolio risk management or derivative pricing accuracy (Engle, 1982; Wu and Chiu, 2017; Saggi and Jain, 2018). Over the past two decades, several events led to extreme price variation in financial markets: the global financial

*Corresponding authors.
Email addresses: Shafqat.iqbal@econ.muni.cz (Shafqat Iqbal), stefan.lyocsa@econ.muni.cz (Štefan Lyócsa)

crisis in 2008 (Kock and Teräsvirta, 2014); the European debt crisis, Brexit, oil–price wars (Khan et al., 2023); the pandemic in early 2020 (Baek et al., 2020); (geo) political uncertainties since Russia's invasion of Ukraine (Lyócsa and Plíhal, 2022); and the sell-off of stocks in early 2025 associated with the announcement of U.S. tariffs. During such periods, accurate volatility models are in demand.

In this study, we present a new approach for volatility forecasting: a model based on fuzzy time series (FTS) and realized variance (RV), the RV-FTS model. Our empirical results are promising and suggest that the model not only provides competitive forecasts with common volatility model benchmarks but also that produces lower correlations with competing forecasts; this useful property, which we further utilize in combination forecasts, outperforms most models for most of the equity indices in our sample and for most forecasting horizons.

2. Literature review

2.1. Standard volatility models

Traditional parametric models are based on the Generalized AutoRegressive Conditional Heteroscedasticity (GARCH) approach introduced by Bollerslev (1986) and were based on the ARCH model of Engle (1982). The standard GARCH model captures the time-varying nature of volatility and clustering effects and depends on the assumed error distribution of heavy tails and, to some extent, long-memory properties. The popularity of the GARCH model was maintained as many extensions exist that account for other stylized facts. These models include asymmetric (E-GARCH Nelson (1991), GJR-GARCH Glosten et al. (1993), T-GARCH Zakoian (1994)), leverage (E-GARCH Nelson (1991)), or long-memory effects (e.g., component models Engle et al. (1999); Conrad and Kleen (2020)). Notable advancements within this class of models include the work of Engle et al. (2013), who included a mixed-frequency (MIDAS) component that allows the use of different covariates (realized volatility or macroeconomic variables) to drive volatility of the return process. A different approach was used by Hansen et al. (2012); Hansen and Huang (2016), who used a measurement equation that links latent and realized volatility.

The availability of high-frequency data has led to improved estimates of volatility (e.g., Andersen et al., 2001; Molnár, 2012) and, eventually, to improved volatility forecasts (e.g., Andersen et al., 2011; Horpestad et al., 2019; Lyócsa et al., 2021). These improvements were achieved mainly using the heterogeneous autoregressive (HAR) family of models, which map the persistent volatility process into a linear combination of past daily, weekly and monthly volatility levels (Corsi, 2009). The HAR models, which were estimated with high-frequency volatility estimates using a least-squares estimator, have become new benchmark models, with many extensions that usually improve upon the already competitive HAR model. For example, Andersen et al. (2012) argued that sudden price jumps distort the persistence of integrated volatility and that a jump-robust estimate[1] may be more suitable in auto-regressive volatility models (e.g., Corsi et al., 2010; Andersen et al., 2012; Patton and Sheppard, 2015). Building on the idea of the asymmetric volatility effect, (Patton and Sheppard, 2015) suggested decomposing the integrated variance into two components: variations due to price decreases and those owing to price increases. The two components may have different persistence levels, and within an HAR model framework, may yield different

[1]For an overview of multiple jump tests, see Maneesoonthorn et al. (2020).

predictive powers. Bollerslev et al. (2016) assumed that the unobserved integrated variance, such as the RV, is more persistent, as the latter may be influenced by measurement error, i.e., attenuation bias. It follows that a volatility model might be adjusted to reduce the weight of recently observed RV if the observation is found to have a high measurement error. Several other extensions of these models exist. The model introduced by Buccheri and Corsi (2021), known as the SHARK model, builds upon the ideas of Bollerslev et al. (2016) but also allows for time-variation of coefficients and conditional heteroscedasticity of volatility model residuals. Similarly, Cipollini et al. (2021) addressed the attenuation bias and heteroscedasticity of errors directly in an HAR model with GARCH errors.

A simple alternative to both GARCH and HAR models is the use of an exponential weighted moving average model (EWMA), introduced in a technical note by J.P. Morgan's (RiskMetrics, 1996), specifically in the context of risk management (Wong et al., 2016). Originally, an EWMA filter was applied for daily data. However, it can also be applied for realized volatilities (as well as for realized covariances, as in Bollerslev et al. (2018b)). The simplicity of the model comes at the cost of a fixed smoothing parameter; thus, the model might not react to changing market conditions (Ayele et al., 2017). Alternatively, one might cross-validate the smoothing parameter to improve the forecasting accuracy of the model Araneda (2021).

Standard EWMA, GARCH and HAR models might not be able to follow abrupt and nonlinear volatility regime transitions that can be caused by unpredictable exogenous shocks (e.g., geopolitical events, natural disasters) or other unexpected information, such as macroeconomic announcements, geopolitical events and/or investor sentiment (e.g., Audrino et al., 2020; Halousková and Lyócsa, 2025). A new area of volatility forecasting literature is focused on statistical learning methods that can adapt well to data-rich environments, where additional data are drawn from macroeconomic variables, other markets, text analyses of relevant sources or technical indicators.

In any early study, Baruník and Křehlík (2016) predicted the price variation of energy markets (crude and heating oil and natural gas) using the GARCH, HAR, and AutoRegressive Fractionally Integrated Moving Average (ARFIMA) models and a simple feedforward neural network. The HAR and ANN models performed well for most predicted measures of price variation in their study (e.g., realized variance, median realized variance), whereas the combined forecast from HAR and Artificial Neural Network (ANN) yielded only marginal improvements. (Lu et al., 2016) employed a hybrid approach to forecast volatility on the Chinese energy index (equity index), feeding GARCH model forecasts into an ANN model in one case and ANN forecasts into a GARCH model in another case. The former led to more accurate volatility forecasts. Christensen et al. (2023) applied a wide array of volatility forecasting models, including tree-based methods, linear regularization methods and standard feedforward NNs to data for individual U.S. stocks. In a setup with only realized price variation measures and/or returns of the given stock price (the vanilla models henceforth), the HAR model was difficult to outperform. After including other exogenous drivers, most machine learning models performed well, whereas the HAR models of Buccheri and Corsi (2021) (SHARK) and Bollerslev et al. (2016) (HAR-Q) underperformed, even lagging behind the vanilla HAR model. The volatility of the S&P market index was predicted by Chun et al. (2024) using 43 predictors (in addition to the usual realized measures), with the least absolute shrinkage and selection operator (LASSO) model outperforming ridge regression, the boosted regression tree, Bayesian model averaging and selection models and the signed jump HAR model of Patton and Sheppard (2015) as the benchmark. Unlike feedforward NN models,

long short-term memory (LSTM) models can address long-term dependencies in data, which are typically observed in long-memory volatility series. Recent works have adopted these techniques. For example, Jiao et al. (2022) used GARCH, support vector regression (SVR) and LSTM models to forecast crude oil volatility. Interestingly, the GARCH model did not perform worst, particularly in the vanilla setup (without additional variables like textual features or other risk factors). The volatility of the Huaxia 50 ETF index (Shanghai Stock Exchange 50 Index) was predicted using a wide battery of tests by Liu et al. (2024) with an LSTM model enhanced by high-frequency volatility series, technical indicators and time series parameters from ARFIMA, HAR, and HAR-Q models. This model outperformed competing models across multiple loss functions, including GARCH, HAR, ARFIMA, tree-based, regularized linear and support vector regression models.

While the proposed RV-FTS model is a separate class of models, given that we use high-frequency volatility estimates and a moving-average update, it is more closely related to a realized volatility AR and HAR model, than to a GARCH or ARFIMA class of models. In our study, we compare the forecasts of the RV-FTS model with forecasts from competing vanilla models that rely on returns and price variation measures only. Our setup thus follows those of Christensen et al. (2023) and Jiao et al. (2022), as we compare models that use the simplest possible information sets, i.e., only price data of the given market index of interest. The competing benchmark models that we use have been employed widely in the literature (e.g., Jiao et al., 2022; Christensen et al., 2023; Liu et al., 2024; Chun et al., 2024). Specifically, we use the GARCH Bollerslev (1986) model with flexible error distributions, the HAR model introduced by Corsi (2009) and its popular variations, the HAR-CJ model introduced by Andersen et al. (2012), the HAR-SV and HAR-SJ models proposed by Patton and Sheppard (2015), and long-memory ARFIMA models on realized variances with GARCH errors and flexible error distributions. To assess the forecasting utility of our approach against common benchmarks, we do not allow other variables to enter the feature space. We therefore also exclude machine learning models.

2.2. Fuzzy-based models in volatility forecasting

Fuzzy theory has been applied within the context of predicting market returns, as reported in a review by Cavalcante et al. (2016). More recently, Wang et al. (2023a) used a combination of a fuzzy-based model with a deep NN to predict returns of A shares in the Chinese stock market. Yolcu and Yolcu (2023) used several fuzzy-based models to predict returns of four market indices. These models outperformed their machine learning and standard econometric counterparts. The use of fuzzy clustering techniques is particularly appealing for volatility forecasting, as volatility is typically characterized by volatility clustering and long-memory. Specifically, volatility clustering is observed through regimes of high- and low-volatility periods, where fuzzy theory can be used to estimate such regimes and to estimate the overlapping membership degrees of a given trading day and its time series dynamics. The long-memory nature of volatility suggests that such regimes tend to remain stable over time. For example, a trading day may be categorized with a 0.75 degree of membership in the "high volatility" regime and a 0.25 degree of membership in the "moderate volatility" regime. Moreover, the membership degrees might change over time according to a given process with the goal of minimizing the sensitivity to short-term noninformative noise and allowing trading days to be assigned to numerous clusters with different membership degrees, which might lead to greater stability and likely greater accuracy (D'Urso et al., 2016; Maciel et al., 2016). Existing studies

have mostly followed the trends in volatility forecasting literature and used hybrid fuzzy GARCH models (e.g., Hung, 2011; Maciel et al., 2016; Dash and Dash, 2016; García and Kristjanpoller, 2019), incorporating fuzzy regime memberships to weight forecasts across different volatility models.

Given the existing fuzzy-based models, it is unclear to what extent the reported volatility forecast improvements can be attributed to the fuzzy-based models or the hybrid nature of the proposed models. Our *first* contribution is therefore methodological, as we design a volatility forecasting model (RV-FTS) that uses only one variable, the realized variance, and combines a FTS framework, fuzzy c-means algorithm and exponential weighted moving average (EWMA) to govern the transitions between volatility regimes. This approach extends Chen's FTS framework (Chen, 1996), which has been widely adopted as a standard in the time series forecasting literature because of its simplicity and effectiveness, particularly for assessing the predictive accuracy of FTS, fuzzy-based, or hybrid forecasting models (Chen and Tanuwijaya, 2011; Bas et al., 2021). The classic FTS model (Chen, 1996) has gained widespread acceptance as a reference method, prompting researchers to introduce subsequent modifications and improvements to enhance its forecasting performance. These advancements include the following: the integration of neural networks and FTS to increase accuracy and computational capabilities (Yu and Huarng, 2010); the adoption of various clustering techniques, such as hierarchical clustering, fuzzy c-means and k-means algorithms, to effectively handle data complexity and determine intervals systematically (Bang and Lee, 2011; Lu et al., 2014; Gupta and Kumar, 2023); the adoption of high-order FTS models to account for temporal dependencies (Chen, 2014); the incorporation of optimization techniques, such as particle swarm optimization, genetic algorithms or metaheuristics, into FTS models for parameter tuning (Chen and Chen, 2015; Chen et al., 2019); and hybrid FTS based on the fuzzy quantum optimization approach (Singh, 2021). Our *second* contribution is empirical in nature, as we compare the volatility forecasting accuracy of the RV-FTS model with that of the GARCH-S_U, RV-ARFIMA-S_U and HAR volatility models across 14 major stock market indices, forecasting horizons from 1–22 days, and two loss functions. Our results suggest that the RV-FTS model generates forecasts that are comparable with those of other benchmark models. Our *third* contribution pertains to the forecast combination literature, as we identify that the RV-FTS model is particularly useful when combined (using a simple average) with any other single benchmark volatility model forecasts. The resulting combination forecast model outperforms most individual models across most market indices, forecast horizons and loss functions.

3. Methodology

3.1. Data

We use daily high-frequency data for 14 stock market indices retrieved from the Oxford-Man Realized Library. Our sample period starts in early 2000 (see Table 1 for specific starting dates), and owing to the data availability constraints, ends for all indices on the 28^{th} of July 2021. The selected 14 markets represent most of the market capitalization from equity markets worldwide, and the dataset spans over 20 years, providing a sufficiently long observation window for analysis.

3.2. Realized volatility

In this study, we are interested in the forecasting accuracy of the h-day ahead variance. Let $P_{i,t}$ denote the intraday price, where $i = 0, 1, ..., M$ denotes the intraday time period and $t = 1, 2, ..., T$ corresponds to the trading days. The first price of the day is given by $P_{0,t}$, and the last price is $P_{M,t}$. M is the number of intraday observations, and, given the 5-minute calendar sampling scheme, $M = 78$. Let $r_{i,t} = ln(P_{i,t}) - ln(P_{i-1,t})$ denote the intraday return. The multiple-day ahead estimator of price variation is given by:

$$RV_{t,h} = h^{-1} \sum_{j=1}^{h} \left[\left(ln(P_{0,t+j}) - ln(P_{M,t+j-1}) \right)^2 + \sum_{i=1}^{M} r_{i,t+j-1}^2 \right] \tag{1}$$

The estimator consists of the standard intraday realized variance and an unbiased overnight price variation (e.g., Molnár, 2012), which must be accounted for in multiple-day ahead settings; $\left(ln(P_{0,t+j}) - ln(P_{M,t+j-1}) \right)^2$. In the following text we will used terms variance and volatility interchangeably, referring to the equation above.

3.3. Benchmark models

3.3.1. Generalized autoregressive models: GARCH-S_U

We follow the literature and use two popular GARCH models as benchmarks, albeit assuming a flexible return distribution. Let $R_t = \ln(P_{t,M}) - ln(P_{t-1,M})$ denote the daily close-to-close return. The GARCH-S_U specification is as follows:

$$\begin{aligned} R_t &= \mu + \epsilon_t \\ \epsilon_t &= \sigma_t \eta_t \\ \sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \\ \eta_t &\sim S_U(0, 1, \lambda, \delta) \end{aligned} \tag{2}$$

The estimated parameters include μ, ω, which are constants in the mean and variance equation and α, β, which describe the persistence of the volatility process. λ and δ are distribution parameters of the shocks η_t, which are assumed to follow the flexible Johnson's S_U distribution (Johnson, 1949b,a) instead of a normal distribution. In this framework, λ is the skewness, and δ is the kurtosis parameters. Both of these parameters are estimated from the data. The density is given by:

$$f(\eta; \lambda, \delta) = \frac{\delta}{\sqrt{2\pi(\eta^2 + 1)}} \exp\left(-\frac{1}{2} \left[\lambda + \delta \cdot \sinh^{-1}(\eta) \right]^2 \right) \tag{3}$$

Choi and Nam (2008) provided evidence that, for FX and stock returns, in an GARCH framework, such a flexible distribution might be more appropriate than normal and Student-t distributions are. More recent studies have applied Johnson's S_U distribution to model oil price volatility, as in Patra (2024), or to estimate value-at-risk and expected shortfalls, as in Lyócsa et al. (2024).

3.3.2. Long-memory realized variance mdoels: RV-ARFIMA-GARCH-S_U

Given the long-memory properties of the realized volatility series, we can model RV_t directly as an ARFIMA process:

$$\begin{aligned} RV_{t,1} &= \mu + \nu_t \\ (1 - L)^d \nu_t &= \epsilon_t \end{aligned} \tag{4}$$

L is the lag operator, and $(1-L)^d$ is the fractional differencing operator equal to $\sum_{k=0}^{\infty} \frac{\Gamma(k-d)L^k}{\Gamma(-d)\Gamma(k+1)}$, where $\Gamma(.)$ is the gamma function and $d \in R$ is the parameter to be estimated. When $d = 0$, the process behaves as a white noise process. For $0 < d < 0.5$, the process is a long-memory process, likely for volatility time series, and $0.5 \leq d < 1$ leads to a nonstationary but mean-reverting process. Moreover, we allow the residuals ϵ_t to follow Eq. (2) and Eq. (3), which results in the RV-ARFIMA-GARCH-S_U model abbreviated into RF-ARFIMA-S_U in Tables[2].

3.3.3. Heterogeneous autoregressive models: AR, HAR, HAR-SV, HAR-SJ, HAR-CJ

Corsi (2009) showed that employing a linear combination of lagged daily, weekly and monthly realized volatility averages leads to typical volatility dynamics, i.e., long-memory and heavy tails, being reproduced. The HAR model is given as follows:

$$RV_{t,h} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{t,W} + \beta_3 RV_{t,M} + u_t, \tag{5}$$

where β represents the parameters to estimate, u_t represents the error terms, and the weekly (five-day) and monthly (22-day) volatility averages are $RV_{t,W} = 5^{-1}\sum_{i=1}^{5} RV_{t-i+1}$ and $RV_{t,M} = 22^{-1}\sum_{i=1}^{22} RV_{t-i+1}$, respectively. Given that our RV-FTS model uses only realized volatilities RV_t, we also compare the forecasting accuracy with that of a model using a similar information set, i.e., a restricted version of the HAR model with $\beta_2 = 0$ and $\beta_3 = 0$ and thus an AR model.

Among alternative variation of the HAR model, we also employ three that we can calculate using our data source, specifically the full version of the positive and negative semivariance model introduced by Patton and Sheppard (2015):

$$RV_{t,h} = \beta_0 + \beta_1 PV_t + \beta_2 NV_t + \beta_3 PV_{t,W} + \beta_4 NV_{t,W} + \beta_5 PV_{t,M} + +\beta_6 NV_{t,M} + u_t, \tag{6}$$

Here, $PV_t = \sum_{i=1}^{M} r_{i,t}^2 I(r_{i,t} \geq 0)$ is the positive semivariance, and $NV_t = \sum_{i=1}^{M} r_{i,t}^2 I(r_{i,t} < 0)$ is the negative semivariance. $I(.)$ is the signaling function, which returns a value of 1 if the condition applies and 0 otherwise. Weekly and monthly components are averages across the five- and 22-day periods.

Under the assumption of symmetric Brownian motion, the expected values of the positive and negative semivariances should be equal. Therefore, $SJ_t = PV_t - NV_t$, a signed jump, can be attributed to jump variation due to discontinuous price movements. We use the following HAR-SJ model specification where signed jumps are included:

$$RV_{t,h} = \beta_0 + \beta_1 SJ_t + \beta_2 CC_t + \beta_3 SJ_{t,W} + \beta_4 CC_{t,W} + \beta_5 SJ_{t,M} + \beta_6 CC_{t,M} + u_t, \tag{7}$$

The continuous component CC_t is estimated using the median realized variance[3] as given by Andersen et al. (2012):

$$CC_t = \frac{\pi}{6 - 4\sqrt{3} + \pi} \frac{M}{M-2} \sum_{j=2}^{M-1} \left(\text{med}\left(|r_{t,j-1}|, |r_{t,j}|, |r_{t,j+1}|\right)\right)^2 \tag{8}$$

A model with an alternative estimator of jump variation was given by Andersen et al. (2012), where the contribution of variation due to discontinuous price changes is given by $JC_t = (RV_t - CC_t) \times I((RV_t - CC_t) > 0)$ and the resulting model:

$$RV_{t,h} = \beta_0 + \beta_1 JC_t + \beta_2 CC_t + \beta_3 JC_{t,W} + \beta_4 CC_{t,W} + \beta_5 JC_{t,M} + \beta_6 CC_{t,M} + u_t, \tag{9}$$

[2]To estimate the GARCH and ARFIMA models, we use the rugarch library described in Ghalanos (2020)

[3]Recently, Kolokolov and Renò (2024) showed that such a multipower estimator might be biased in the presence of stale prices. However, given that we are working with broad market indices, such bias should not be an issue in our empirical application.

All the models in this section are estimated via weighted least squares, following Patton and Sheppard (2015); Clements and Preve (2021), with weights given by $w_t = \widetilde{RV}_{t,h}^{-1}$, where $\widetilde{RV}_{t,h}$ is the predicted variance from the OLS estimator, an approach that mitigates the impact of the heavy tails on the parameters β.

3.4. The RV-FTS Model

3.4.1. Fuzzy Time Series

The concept of FTS had been first presented in early works of (Song and Chissom, 1993, 1994). This approach uses fuzzy set theory of (Zadeh, 1965) to forecast time series data, particularly when the input contains linguistic or imprecise information. Subsequently, (Chen, 1996) modified the original (Song and Chissom, 1993) model, which laid the foundation for many advancements in FTS modeling. The basic steps of Chen's model can be summarized as: i) *Define and Partitioning the data set* - The universe of discourse for a given time series data is defined, and the data are divided into optimized, potentially overlapping intervals, ii) *Fuzzification* - The universe of discourse U is transformed into a sequence of fuzzy sets utilizing the membership functions and intervals that were established during the partitioning stage., iii) *Fuzzy Logical Relationships and Rules Extraction* - The fuzzy logical relationships (FLRs) are constructed by analyzing the fuzzy state relationships among fuzzy sets. On the basis of these established FLRs, fuzzy rules are derived, which describe the temporal dynamics among fuzzy sets, iv) *Defuzzification* During the final stage of the FTS framework, following the establishment of FLRs, fuzzy rules are extracted and matched with the current fuzzified input $\mathcal{F}_{t,h}$, the fuzzified output is defuzzified into a crisp forecast $\widehat{RV}_{t,h}$.

3.4.2. Proposed RV-FTS

The proposed RV-FTS model extends the classical FTS framework by integrating a rolling window mechanism, FCM clustering, and an EWMA approach for volatility forecasting. In the following, we present a step-by-step procedure and the algorithmic structure of the RV-FTS model.

Step 1: During this step, first, the realized volatility is computed(see Section 3.2), and the universe of discourse $\mathcal{U}$ is defined for realized variance data $\{RV_{1,h}, RV_{2,h}, \ldots, RV_{T-h,h}\}$, as follows:

$$\mathcal{U} = [\min\{RV_{t,h}\}_{t=1}^{T-h}, \max\{RV_{t,h}\}_{t=1}^{T-h}] \tag{10}$$

or

$$\mathcal{U} = \{RV_{1,h}, RV_{2,h}, \ldots, RV_{T-h,h}\}$$

Step 2: Next, the universe of discourse $\mathcal{U}$ is partitioned into C optimized, potentially overlapping intervals and transformed into a sequence of fuzzy sets $\{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_j\}$ utilizing the FCM clustering approach. In contrast to hard clustering techniques, FCM assigns partial membership degrees to each data point across multiple clusters simultaneously, reflecting the proportional association of each data point to every cluster. The membership degrees are derived by minimizing the objective function of FCM, which incorporates the weighted sum of the squared Euclidean distance between the data values and cluster centers (Bezdek et al., 1984).

$$J(\gamma_{\mathcal{B}}, V) = \sum_{j=1}^{c}\sum_{t=1}^{T} \gamma_{\mathcal{B}_j t}^{m} \left\| RV_{t,h} - g_j \right\|^2 \tag{11}$$

where $J(\gamma, V)$ is the objective function to be minimized, c represents the number of clusters, T represents the number of observations, $\gamma_{\mathcal{B}_j t}$ represents the membership information of the t^{th} observed value in the j^{th} cluster, $RV_{t,h}$ represents the t^{th} observed value, g_j represents the centroid of the j^{th} cluster, and m represents the fuzziness parameter, which is usually set at 2 and controls the degree of fuzziness.

Each value in the original realized volatility is mapped to the corresponding fuzzy set with varying membership degrees as $\gamma_{\mathcal{B}_j}(v_n) \in [0, 1]$. Let $\mathcal{U}$ denote the universe of discourse defined in Eq(), and fuzzy sets $\mathcal{B}_j$ defined over this domain can be written as:

$$\mathcal{B}_j = \left\{ \frac{\gamma_{\mathcal{B}_j}(RV_{1,h})}{RV_{1,h}} + \frac{\gamma_{\mathcal{B}_j}(RV_{2,h})}{RV_{2,h}} + \cdots + \frac{\gamma_{\mathcal{B}_j}(RV_{T-h,h})}{RV_{T-h,h}} \right\} \tag{12}$$

where $\gamma_{\mathcal{B}_j}(RV_{t,h}) \in [0, 1]$ is the membership function capturing the degree of fuzziness associated with $RV_{t,h}$. Each term $\frac{\gamma_{\mathcal{B}_j}(RV_{t,h})}{RV_{t,h}}$ is interpreted as "$RV_{t,h}$ belongs to B_j with membership information $\gamma_{\mathcal{B}_j}(RV_{t,h})$". The collection of all fuzzy sets $\mathcal{B}_j$ defined in $\mathcal{U}$ forms the FTS, $\mathcal{F}(t, h)$, given as follows:

$$F(t,h) = \left\{ \frac{\gamma_{\mathcal{B}_1}(RV_{t,h})}{RV_{t,h}}, \frac{\gamma_{\mathcal{B}_2}(RV_{t,h})}{RV_{t,h}}, \ldots, \frac{\gamma_{\mathcal{B}_c}(RV_{t,h})}{RV_{t,h}} \right\} \tag{13}$$

or

$$F(t,h) = \left\{ \frac{\gamma_{\mathcal{B}_j}(RV_{t,h})}{RV_{t,h}} \right\}_{j=1}^{c}$$

Within each window, FCM is applied to partition the realized volatility data into an optimal number of clusters through cross-validation to identify the latent structures within the volatility dynamics. This soft clustering technique is particularly suitable for modeling financial market volatility, where transitions are gradual and often overlap rather than abrupt (Maciel et al., 2016; Vilela et al., 2019). Additionally, the fuzzy clustering process within a rolling window allows the model to dynamically adjust to current market conditions, better reflecting the complex and nonlinear nature of financial markets more efficiently than conventional statistical models do (Dash and Dash, 2016; D'Urso et al., 2016).

Step 3: During this step, defuzzification is performed through an adaptive EWMA technique, which combines cluster centroids, membership degrees, and previous forecasts to anticipate the variance estimate within the rolling window. The adaptation and reframing of EWMA in this setting serves as a conceptual bridge between fuzzy logic and traditional time series by replacing fixed inputs and rule-based systems with fuzzified abstraction and recursive updating (Dash et al., 2015).

The RV-FTS model operates on the basis of a fuzzified variance estimate $\gamma_{\mathcal{B}_j t}$, which obtained through the FCM clustering algorithm. This procedure assigns varying membership degrees to each realized variance observation across multiple volatility states. The adaptive EWMA mechanism then transforms the fuzzified realized variance, cluster centroids and previous forecasts into crisp forecasts. This setup allows the model to adjust to recent market conditions while still incorporating past effects with decreasing weights. The fuzzified variance estimate $\tilde{v}_t$ and adaptive EWMA mechanism are expressed as follows:

$$\tilde{v}_t = \sum_{j=1}^{c} \gamma_{\mathcal{B}_j t}(RV_{t-h+1,h}) \cdot g_j \tag{14}$$

$$\widehat{RV}_{t+1,h} = \rho \tilde{v}_t + (1-\rho)\widehat{RV}_{t-h+1,h} \tag{15}$$

where $\widehat{RV}_{t+1,h}$ represents the *forecasted conditional variance* at time $t+1$, $\gamma_{\mathcal{B}_j}(RV_{t-h+1,h})$ represents the membership information for the most recent value for cluster j, which represents the degree of association between the most recent realized variance and each cluster (e.g., the volatility regime), $\rho \in (0,1)$ represents the smoothing factor, g_j represents the cluster centroid for the j-th cluster, and $\widehat{RV}_{t-h+1,h}$ represents the previous known forecast.

Algorithm 1: Proposed RV-FTS Model

Input: Opening and closing prices,
Realized daily variance over 5-minute intervals,
Parameters: number of clusters c, rolling window size W, smoothing factor ρ
Output: h-day ahead volatility forecasts $\widehat{RV}_{t,h}$ and performance metrics (e.g., MSE, QLIKE)

`/* Phase 1: Preprocessing */`

Compute realized variance using input price data (intraday returns and current day opening, and last day closing prices), as given by:

$$RV_{t,h} = h^{-1} \sum_{j=1}^{H} \left[\left(ln(P_{0,t+j}) - ln(P_{M,t+j-1}) \right)^2 + \sum_{i=1}^{M} r_{i,t+j-1}^2 \right]$$

`/* Phase 2: Initialization */`

Set model parameters: c (number of clusters), W (window size), ρ (smoothing factor); Split the $RV_{t-h+1,h}$ into rolling windows of size W;

`/* Phase 3: Fuzzification via FCM */`

foreach rolling window **do**

Initialize membership matrix γ_{jt} for $t = 1, \ldots, n$, $j = 1, \ldots, c$;

Optimize the FCM objective function:

$$J(\gamma_{\mathcal{B}}, V) = \sum_{j=1}^{c} \sum_{t=1}^{n} \gamma_{\mathcal{B}_j t}^{m} \left\| RV_{t-h+1} - g_j \right\|^2$$

Update centroids g_j and membership matrix $\gamma_{\mathcal{B}_{jt}}$ iteratively until convergence;

`/* Phase 4: Defuzzification and Forecasting */`

foreach forecasting step **do**

Compute the fuzzified estimate of variance:

$$\tilde{v}_t = \sum_{j=1}^{c} \gamma_{\mathcal{B}_j t}(RV_{t-h+1,h}) \cdot g_j$$

Compute the next forecast using EWMA-based defuzzification:

$$\widehat{RV}_{t-h+1,h} = \rho \tilde{v}_t + (1-\rho)\widehat{RV}_{t-h+1,h}$$

or equivalently,

$$\widehat{RV}_{t+1,h} = \rho \sum_{j=1}^{c} \gamma_{\mathcal{B}_j t}(RV_{t-h+1,h}) \cdot g_j + (1-\rho)\hat{RV}_{t-h+1,h}$$

Shift the rolling window forward and repeat;

`/* Phase 5: Model Evaluation */`

Compare the forecasted values $\widehat{RV}_{t,h}$ with the actual realized variance $RV_{t,h}$;

Calculate error metrics MSE and QLIKE;

return Forecasts $\widehat{RV}_{t,h}$ and evaluation metrics;

3.5. Forecast Combinations

The true data generating process (DGP) is unknown and is unlikely to be approximated by a single volatility model. In addition, major economic events might lead to changes in the DGP. A common strategy to mitigate such uncertainties in the forecasting literature is to rely on combination forecasts (Timmermann, 2006). Strategy is also commonly employed in the volatility forecasting literature (e.g., Kourentzes et al., 2019; Wang et al., 2023b). The core idea is similar[4] to that of variance reduction in a well-diversified portfolio. A suitable combination of unbiased but not perfectly correlated forecasts can lead to lower (squared) forecast errors.

In the empirical part of our research, we find that the RV-FTS model's performance is often competitive with that of most other benchmark models and that its forecasts are less correlated with forecasts from benchmark models. It follows that forecasts from the RV-FTS model are likely to improve existing forecasts in a combination forecasting framework. We therefore use a simple unconditional combination, an average between two models. Let $F_{t,B}$ and $F_{t,C}$ denote the benchmark and competing (RV-FTS) models forecasts, respectively. The combination we use is $2^{-1}(F_{t,B}+F_{t,C})$. These pairwise combinations are denoted as C-GARCH-S_U, ..., C-HAR-CJ, and C-HAR-SJ.

The resulting forecast improvements can be used to assess the extent to which the RV-FTS model helps specific volatility models achieve more accurate forecasts. However, the forecast improvements might, to some extent, be attributed to the averaging itself. We therefore also calculate three additional combination forecasts using a weighted average. *First*, a conditional average across all models except RV-FTS, with weights determined using a constrained regression similar to (Granger and Ramanathan, 1984):

$$\widehat{RV}_{t,h} = \sum_{m=1}^{M} \beta_m F_{t,m} + \epsilon_{t+1} \tag{16}$$

with $m = 1, 2, ..., M$ denoting all models except the RV-FTS model and $\sum_{m=1}^{M} \beta_m = 1$ and $\forall \beta_m \geq 0$. The constrained OLS is being estimated using a rolling sample of 252 observations (approx. 1 year). The resulting forecast is denoted as CC-Bench. *Second*, we use the same approach now with M^* models, that also includes forecasts from the RV-FTS model and is denoted as CC-FTS. The *third*, denoted as CC-FTS-Top follows from the previous conditional combination, but now we find a simple average of five forecasts that had the highest estimated weight implied by β_m.

3.6. Forecasting framework

The forecasts are generated using a rolling window framework with 756 observations (approx. 3 years) to estimate all benchmark models. For the RV-FTS model, we need to estimate the number of clusters $\phi \in \{2, 3, 4, 6, 8, 10, 12, 14\}$ and the smoothing parameter $\rho \in \{0.10, 0.20, ..., 0.80, 0.90, 0.95, 0.975, 0.99\}$. Higher values of ρ are considered due to the greater persistence of the volatility process. The optimum parameters are cross-validated using additional 252 observations and the square loss function to select the preferred combination of ϕ and ρ values. It follows that the first forecast from the RV-FTS model is for the 1009^{th} trading day in the sample period of a given stock market index. The optimum parameters are re-estimated after each trading day.

[4] If the mean square error is the chosen loss function.

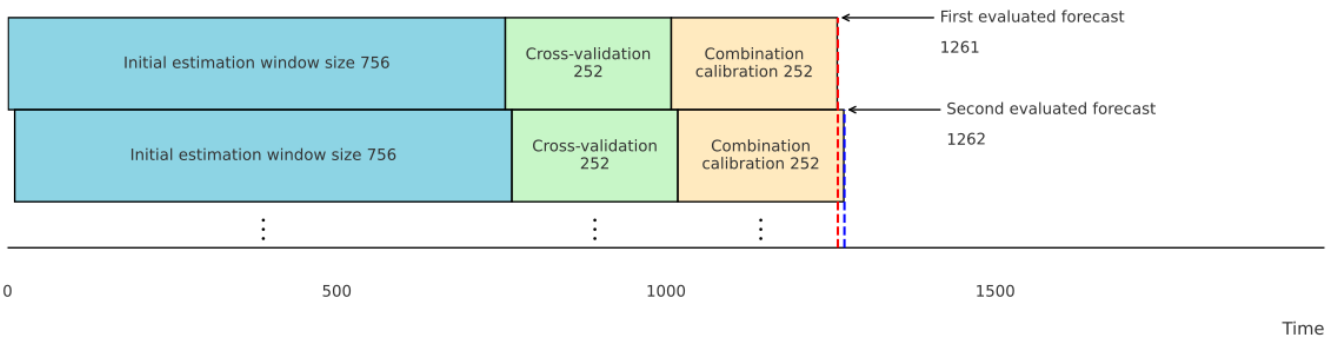

Figure 1: Forecasting and cross-validation framework for RV-FTS model

3.7. Forecast Evaluation

Given that we need 256 observations to estimate conditional combinations, ensuring comparisons across the same information sets between the models leads to the first evaluated forecast being the 1261^{th} observation. The following Figure 1 visualizes the forecasting framework showing the first and the second forecast.

More specifically, the RV-FTS model is benchmarked against 9 individual competing volatility models, with the HAR model representing the model against which we report relative losses. This comparison allows us to assess the relative merit of the RV-FTS model against different classes of standard volatility models. We then benchmark the 8 individual volatility models with their pairwise combination forecasts, where each competing model's forecast is averaged with the RV-FTS model's forecasts. This comparison allows us to assess the relative contribution of the RV-FTS model to the toolbox of an analyst using one of the standard volatility models. The three conditional combinations, CC-Bench, CC-FTS and CC-FTS-Top, allow us to assess the merit of RV-FTS in a broader set of volatility models. To facilitate comparisons across these models, we report the results for all 21 models in single tables.

We report the results for two loss functions, the mean square error (MSE) and one loss from the QLIKE family of functions, which have been shown to lead to consistent model rankings, even in the presence of a noisy variance proxy (Patton, 2011). Let $F_{t,h}$ be the h-head forecast from a given model; the MSE is given by:

$$L_{t,h}^{MSE} = \frac{1}{T}\sum_{t=1}^{T}\left(RV_{t,h} - F_{t,h}\right)^2, \tag{17}$$

and the QLIKE as:

$$L_{t,h}^{QLIKE} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{RV_{t,h}}{F_{t,h}} - \ln\left(\frac{RV_{t,h}}{F_{t,h}}\right) - 1\right), \tag{18}$$

Bollerslev et al. (2018a) noted that forecasting models can occasionally lead to implausibly large variance forecasts. To mitigate such issues, we follow the approach described by Bollerslev et al. (2018a). For each trading day, we substitute variance forecasts that are larger than the historical maximum before that trading day with that historical maximum [5] Moreover, unexpected market-relevant news often leads to extreme levels of price variation. Such unexpected events are unlikely to be predicted by any volatility model, resulting in higher squared errors for

[5]Note that Bollerslev et al. (2018b) replaced such forecasts with an expanding long-run sample mean of realized variances. However, we argue that relying on long-run variance might systematically underestimate a model's forecasts during times of high volatility when the model might correctly predict new maxima, i.e., during times when accurate volatility forecasts are most important.

such days. These errors are often several orders of magnitude greater than the median squared error is. It follows that even small changes in forecasting accuracy for such days can impact the ranking of models. We therefore report the average rank of each model by determining its rank for each trading day on the basis of the given loss value (one ranking for MSE and one for QLIKE). The model with the lowest loss is assigned a rank of 1. The average rank is not only the loss itself but can also be interpreted as an *average preference* for a specific model within a given loss function.

Apart from the descriptive comparison of the MSE and QLIKE losses and corresponding ranks, we employ the model confidence set (MCS) (Hansen et al., 2011) to compare ranks (based on either MSE or QLIKE) across all 21 forecasts in an iterative ('elimination') algorithm. Beginning with all the models, we test the null hypothesis of equal predictive ability (EPA) at the $(1-\alpha)$ confidence level. For this, we employ the T_{MAX} test and the 85% confidence level, which are standard for this test. If the null of the EPA test is rejected, the worst performing model is removed, and the test statistics are recalculated with the remaining models. The algorithm stops when the EPA hypothesis is not rejected, leading to the identification of the set of superior models[6].

4. Results and Discussions

4.1. Data characteristics

Volatility characteristics from annualized values of realized variances of the fourteen selected market indices can be found in Table 1. Aggregated descriptions of the variance series reveal the presence of standard stylized variance features, such as right-skewness, heavy-tails and long-memory properties. Even at $\rho(22)$, the autocorrelations range from 0.08 to 0.27. These values are much higher than expected, with an exponential decline in persistence. The results suggest that models that are able to capture such persistence (e.g., HAR) should achieve better forecasting performance.

Figure 2 shows the individual time series and highlights several major events that increased the uncertainty of almost all equity markets worldwide. Most notably, the global financial crisis before 2010 and the COVID-19 pandemic in 2020 appear to be greater in magnitude but much shorter in duration. Market-specific events are also visible. For example, the spike in volatility during the 2011 Fukushima disaster in Japan, the somewhat unexpected outcome of the 'Brexit' elections in the summer 2016 for the GBR and EU markets, or the sharp declines in 2015 and 2016 observed in the market in China[7].

These results confirm that realized volatilities in equity markets share common dynamics, an observation made by Bollerslev et al. (2018a) across other asset classes. However, markets are also exposed to local shocks. Given these volatility characteristics, we did not expect the RV-FTS model to perform equally across all markets. In fact, we expected this model would potentially perform worse in markets with higher volatility persistence. In these markets, models exploiting the long-memory features are expected to generate more accurate forecasts.

[6] We adopt the implementation introduced by Bernardi and Catania (2018).

[7] More detail on other components of the price variation, i.e., signed jumps, semivariances, and continuous and jump components of the overall variation in intraday prices, can be found in Table A1 in Appendix A.

4.2. Day-ahead market volatility forecasts

In this section, we demonstrate the forecasting performance of the RV-FTS model in day-ahead forecasts and compare the model to common volatility model benchmarks. Next, we evaluate the role of the RV-FTS model in a set of volatility models by exploring its usefulness in combination forecasts.

4.2.1. Day-ahead forecasts: Individual volatility models

In Figure 2, we display day-ahead forecasts from the RV-FTS model (red line), which captures the dynamics of the underlying volatility process (black line). High-volatility periods, such as the global financial crisis or pandemic, were captured by the model. For illustration purposes, Figure 3 presents the dynamics of the 'optimum' parameters of the RV-FTS model, as selected via cross-validation for the STOXX 50 market index, the number of clusters c and the persistence parameter ρ. The first row represents values for the day-ahead model. Both hyperparameters tend to vary greatly over time. At the onset of the pandemic in 2020, few volatility regimes (cluster) are selected. A similar trend was observed during the global financial crisis and in 2016 (Brexit elections). Such periods also tend to be accompanied by greater volatility persistence. During these periods, the ρ parameter in the RV-FTS model is higher. To some extent, similar trends were observed in early 2020 and 2016. Given the observed variation in these parameters over time, estimating (c, ρ), the only two parameters of the RV-FTS model, is important for forecasting purposes. Moreover, for multiple-day forecast horizons, as the middle and lower panels of Figure 3 show that for five- and twenty-two-day-ahead forecasts, different dynamics are observed for the optimum parameters.

The numerical results from the day-ahead forecasts can be found in Tables 2 and 3 under the MSE and QLIKE loss functions. In Panel A of Table 2, the first row shows the average value of the benchmark HAR model, with columns representing different markets. For example, the value of 10.3 corresponds to the MSE, which is annualized and scaled by 10^{-4}. The rows in Panels A and B represent improvements $(-)$ or deterioration $(+)$ in forecasts relative to the HAR model reported in the first row. The value of 34.5% in the second row implies that using the GARCH-S_U model led to a 34.5% higher average mean square error, i.e., for Australia, the standard GARCH-S_U model underperformed the HAR model considerably, which is often the case in our sample and in line with previous studies for US financial markets, as in Vortelinos (2017) or Horpestad et al. (2019); Lyócsa et al. (2021), where HAR models outperformed the GARCH class of models in day-ahead forecast settings.

Panel A of Table 2 presents the performance of individual models across all markets. In eight out of 14 markets, HAR-CJ yielded the lowest MSEs, followed by HAR-SV and FR-GARCH-S_U, with two and three cases, respectively. However, all three competing HAR models, HAR-SV, HAR-CJ and HAR-SJ, performed two times worse, i.e., leading to considerable model choice uncertainty. Since the RV-FTS model relies only on the most recent realized variance, similar to an AR model, it is surprising that it achieved the second lowest forecast error for three markets (French, UK and Japan) and never performed worst. In addition, it consistently outperformed the AR model (in 11/14 markets) and the GARCH-S_U model (in 10/14 markets), while remaining competitive with the FR-GARCH-S_U model (in 7/7 markets). With respect to the benchmark HAR and HAR-SV, HAR-CJ and HAR-SJ models, the RV-FTS model performed slightly worse in a day-ahead setting, outperforming these models in 5, 6, 4 and 6 of the 14 markets, respectively. In Panel A of Table 3, the results under the QLIKE loss also show that among the individual models, the HAR-CJ model is preferable. Under QLIKE, the RV-FTS always outperforms

the GARCH-S_U, RF-GARCH-S_U and AR models but rarely the HAR class models.

In Panels C and D of Tables 2 and 3, we report the average ranks of the individual models calculated across all 18 models (including forecast combinations). Specifically, for each day, we rank models from 1 (most accurate) to 18 (least accurate) according to the given loss. The reported values correspond to the average across time. For example, in Panel C of Table 2, we show the results for individual forecasting models. The value of 9.1 for the RV-FTS model means that out of all 18 models, on average, the individual RV-FTS model had a rank of 9.1 for the market in Austria (first column). Ranks are resistant to outliers, which are common for square error losses, and they also represent preferences for a specific model that can be compared across markets. For the RV-FTS model, ranks range from 8.8 (IND) to 9.8 (DEU), while those for the benchmark HAR are similar, ranging from 9.0 (AUS) to 10.2 (IND) For comparison, the least competitive volatility models, GARCH-S_U and AR, have average ranks ranging from 10.2 (KOR) to 13.9 (AUS) for GARCH-S_U and 10.8 (EUR) to 11.8 (HKG and IND) for AR. The top performing HAR-CJ model ranges in rank from 8.6 (AUS) to 9.3 (IND and HKG). Results under QLIKE (in Table 3) are very similar to those under MSE. The analysis shows that the day-ahead performance of the individual RV-FTS model is closer to that of the HAR model than to that of the previous generation (GARCH) of volatility models, especially under the rank losses when outlier effects are mitigated. We consider these findings encouraging for two reasons: i) RV-FTS relies only on one variable, and ii) it is not designed for rapid adaptation in the short term, giving HAR models a certain advantage in day-ahead settings.

4.2.2. Day-ahead forecasts: Combination forecasts

If the forecasts of the benchmark and RV-FTS models are unbiased but forecast losses are not perfectly correlated (e.g., Timmermann (2006)), the utility of the RV-FTS volatility model might increase when combined with benchmark forecasts, even if the forecast accuracy of the individual RV-FTS model is lower than that of the benchmark model. As the RV-FTS algorithm differs from the usual parametric benchmark HAR models, we expect that combining forecasts from the RV-FTS model with the usual benchmarks might lead to improved forecasts. We compare forecasts from individual models with those of their bivariate combination forecast counterparts, GARCH-S_U with C-GARCH-S_U (averages from GARCH-S_U and RV-FTS), RF-GARCH-S_U with C-RF-GARCH-S_U (averages from FR-GARCH-S_U and RV-FTS), etc. This comparison allows us to observe which individual models benefit from including the RV-FTS the most.

Under the MSE (Panel B of Table 2) loss, across all markets and individual models, bivariate combinations tended to improve forecasts in 80.6% of the cases. The utility of RV-FTS is model specific, and as expected, for the best performing individual models, HAR and HAR-CJ, adding the RV-FTS model improved the forecast in 60.7% of the cases in the day-ahead setting. On the other hand, the less competitive GARCH-S_U, RF-GARCH-S_U, AR, HAR-SV and HAR-SJ models benefited from the RV-FTS model the most, with improvements in 88.6% of the cases. The average rankings led to even stronger results, with RV-FTS improving individual model forecasts in 98.0% of the cases. Thus, when mitigating the effect of outliers that are typical for square error volatility losses, even the HAR and HAR-CJ models benefit from including the RV-FTS model. The results using the asymmetric QLIKE, which penalizes underestimation of the volatility more, are similar. However, the best-performing models, HAR and HAR-CJ, benefited from the inclusion of the RV-FTS model in approximately 35% of the cases. On the

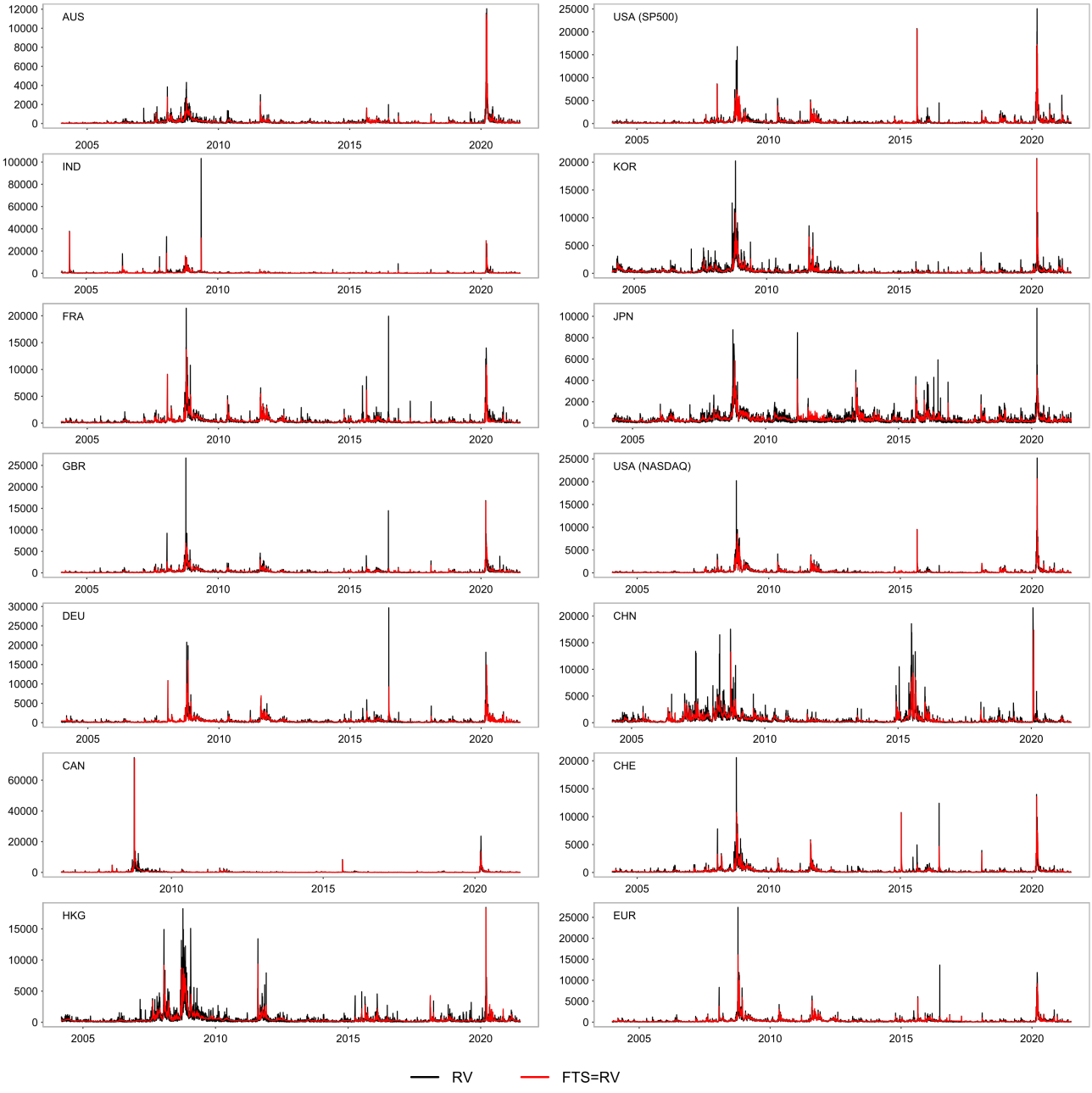

Figure 2: The realized variance series of stock market indices

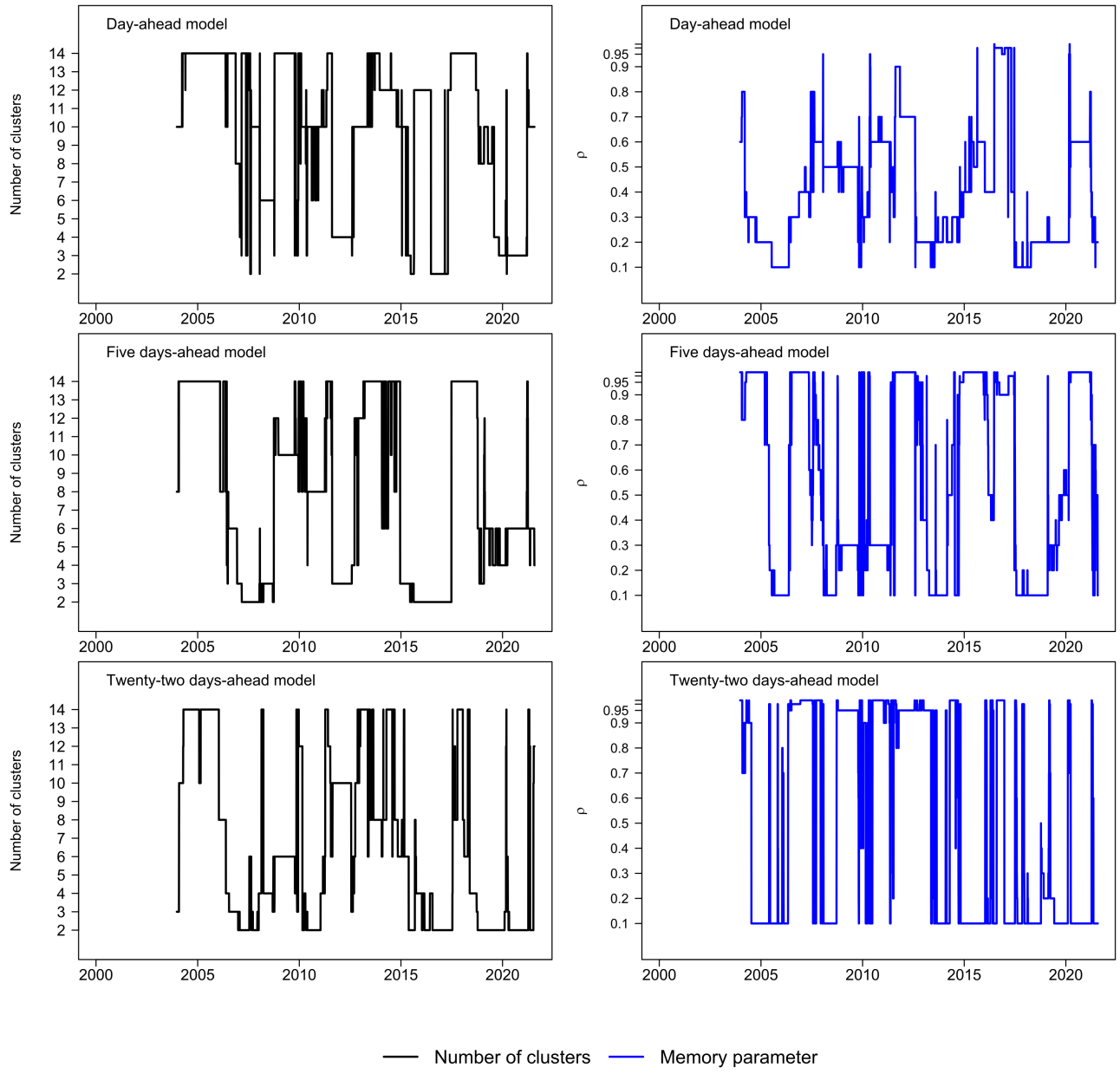

Figure 3: Optimum model parameters for RV-FTS: Case of STOXX 50

other hand, in terms of average ranks calculated from the QLIKE, the results are almost the same as those under the MSE, as inclusion of the RF-FTS model is favored in 96.9% of all cases.

While bivariate combinations help to identify which individual volatility forecasting models benefit the most from the RV-FTS model, some forecast improvements might be driven by combinations alone, i.e., combining two individual models (except RV-FTS) could lead to similar or greater forecast improvements. Moreover, it is unclear which models should be selected for a specific time and market. We therefore complement our analysis using three conditional combination forecasting models. The CC-Bench model combines forecasts from all eight individual models except RV-FTS, and weights are estimated following the

Both conditional forecast combinations generally outperform the benchmark HAR model, regardless of the loss (Table 2 and 3). As expected, forecast improvements are lower than those of the best individual or pairwise combination forecasts but are more stable and realistic given that prior knowledge about the best performing model is rarely available. A comparison of the CC-Bench and CC-FTS models results reveals that in a day-ahead setting, they perform similarly. However, CC-FTS-Top tends to outperform CC-Bench or CC-FTS in most (11/14 markets) cases. The utility of the RV-FTS model in conditional combinations is visualized in Figure 4, where we report its time-varying optimum weights. The weights across the 8 individual models sum to 1. Figure 4 reveals several interesting observations about the RV-FTS model. The blue line in this figure represents the day-ahead forecasts and shows considerable dynamics with periods of low and high weights, suggesting that the utility of the model changes over time and can occasionally reach values above 0.5.

To summarize the day-ahead results, we show that despite its simplicity, the RV-FTS volatility model can improve volatility forecasts, outperforming traditional benchmarks and complementing more advanced HAR models under both MSE and QLIKE losses. Moreover, occasionally, the RV-FTS model represents a preferred choice among all individual models, as indicated by the estimated weights in the CC-FTS conditional combination model.

Table 1: Characteristics of realized variance

Country	Mean	SD	SK	KU	$\rho(1)$	$\rho(5)$	$\rho(22)$	obs.	start
Australia	148.6	393.2	392.1	16.2	0.64	0.36	0.16	5446	2000-01-04
India	542.4	1928.4	1562.8	32.6	0.24	0.14	0.08	5351	2000-01-03
France	480.6	944.7	130.7	9.2	0.58	0.39	0.22	5499	2000-01-03
UK	313.5	747.2	390.7	15.3	0.49	0.36	0.18	5438	2000-01-04
Germany	527.1	1087.6	179.6	10.5	0.58	0.38	0.21	5463	2000-01-03
Canada	242.8	1319.5	2144.0	40.3	0.37	0.21	0.12	4809	2002-05-02
Hong Kong	489.8	1028.3	101.5	8.6	0.46	0.38	0.27	5281	2000-01-03
US (Nasdaq)	494.2	1100.6	135.4	9.4	0.58	0.37	0.21	5407	2000-01-03
South Korea	547.0	1235.3	295.3	13.1	0.47	0.28	0.20	5307	2000-01-04
Japan	400.3	545.8	75.9	6.8	0.57	0.38	0.19	5241	2000-02-02
US (SP 500)	304.0	791.8	306.7	13.7	0.66	0.43	0.24	5406	2000-01-03
China	538.2	1199.3	111.4	8.9	0.49	0.23	0.22	5208	2000-01-04
Swiss	318.4	763.7	177.5	10.8	0.60	0.37	0.19	5401	2000-01-04
Europe	430.3	858.4	233.7	11.4	0.61	0.41	0.22	5496	2000-01-03

Notes: Volatility characteristics are reported from annualized values. SD denotes standard deviation, SK skewness, KU kurtosis, and $\rho(\cdot)$ denotes auto-correlation of the given order. The market indices are All Ordinaries (Australia), BSE Sensex (India), CAC 40 (France), FTSE 225 (UK), DAX 30 (Germany), TSX Composite (Canada), Hang Seng (Hong Kong), KOSPI (South Korea), Nikkei 225 (Japan), S&P 500 (US), Shanghai Stock Exchange Composite Index (China), Swiss Market Index (Swiss), Stoxx 50 (Europe). Data for all markets end on 28^{th} July 2021.

Table 2: Day-ahead forecasts under MSE

Model	AUS	IND	FRA	GBR	DEU	CAN	HKG	USA-1	KOR	JPN	USA-2	CHN	CHE	EUR
Panel A: Loss comparison with individual volatility models relative to the HAR model														
HAR [MSE]	10.3	450.0	63.4	45.9	77.4	209.6	80.7	63.0	55.7	19.5	52.3	96.6	36.8	48.7
GARCH-S_U	34.5%	-16.6%	-7.8%	-1.3%	4.2%	-5.4%	15.5%	19.5%	-0.7%	153.4%	6.6%	17.7%	19.5%	2.6%
FR-GARCH-S_U	39.2%	-8.3%	-1.1%	-3.7%	5.5%	0.9%	1.1%	-0.5%	-2.1%	9.7%	-25.4%	12.5%	13.5%	1.9%
AR	17.5%	15.5%	-2.4%	11.9%	7.5%	32.1%	19.9%	12.8%	8.0%	6.9%	-9.1%	9.3%	13.1%	10.2%
HAR-SV	58.8%	-20.1%	-17.6%	8.4%	1.6%	6.4%	4.8%	0.2%	4.4%	7.7%	-29.9%	-4.9%	60.4%	7.6%
HAR-CJ	-10.6%	-22.2%	-3.6%	-6.5%	-19.0%	-13.8%	1.8%	34.5%	-6.2%	11.7%	10.3%	-3.0%	9.8%	-10.6%
HAR-SJ	10.4%	-15.6%	8.8%	1.9%	-14.7%	554.2%	12.1%	8.6%	5.8%	19.2%	-27.1%	3.1%	27.3%	-9.0%
RV-FTS	11.7%	-11.9%	-15.9%	-3.7%	0.7%	60.3%	7.5%	17.1%	7.8%	6.2%	-21.5%	7.7%	21.1%	-1.2%
Panel B: Loss comparison with combination forecast models relative to the HAR model														
C-GARCH-S_U	-0.9%	-18.5%	-20.3%	-10.0%	-7.3%	6.7%	-0.6%	1.8%	-4.9%	49.7%	-20.5%	4.8%	9.2%	-9.8%
C-FR-GARCH-S_U	0.2%	-16.3%	-17.9%	-11.3%	-5.0%	12.8%	-2.3%	-3.4%	-5.2%	3.2%	-30.5%	4.1%	8.2%	-9.2%
C-AR	10.7%	-2.3%	-13.4%	-4.3%	-0.8%	42.1%	7.6%	9.2%	2.3%	1.3%	-20.4%	2.5%	10.0%	0.9%
C-HAR	-5.4%	-8.3%	-12.5%	-8.3%	-4.6%	23.4%	0.8%	3.7%	-4.2%	0.7%	-15.8%	0.4%	5.3%	-3.2%
C-HAR-SV	32.1%	-18.2%	-21.2%	-8.0%	-7.3%	10.9%	-3.0%	-1.9%	-4.6%	-2.9%	-29.2%	-4.3%	20.2%	-1.5%
C-HAR-CJ	-10.9%	-19.4%	-13.3%	-10.2%	-15.5%	5.8%	-4.5%	15.3%	-7.9%	-0.2%	-7.1%	-2.9%	7.1%	-10.6%
C-HAR-SJ	-6.7%	-18.7%	-14.1%	-7.7%	-14.7%	153.8%	-1.4%	-2.0%	-5.9%	2.7%	-31.6%	-2.1%	13.2%	-10.4%
CC-Bench	-10.4%	-18.0%	-18.7%	-4.3%	-17.2%	37.8%	-3.2%	-4.9%	-11.2%	2.3%	-29.1%	-0.7%	20.3%	-14.5%
CC-FTS	-8.9%	-17.9%	-21.2%	-5.5%	-17.1%	39.2%	-2.6%	-5.2%	-11.5%	2.3%	-29.5%	-0.5%	20.3%	-14.6%
CC-FTS-Top	-10.0%	-19.6%	-20.6%	-9.6%	-13.4%	-9.9%	-5.8%	-4.5%	-11.8%	-4.1%	-22.6%	-4.8%	1.6%	-11.3%
Panel C: Average ranking of the individual forecast models														
HAR	9.0	10.2	9.7	9.4	9.8	9.4	9.4	9.5	9.5	9.4	9.4	9.3	9.6	9.3
GARCH-S_U	13.9	10.8	11.3	11.1	11.1	11.3	10.7	12.1	10.2	12.3	12.5	12.9	11.1	12.1
FR-GARCH-S_U	9.3	8.8	9.4	9.5	9.4	9.6	8.8	9.7	9.9	9.1	9.9	8.9	9.7	9.6
AR	10.9	11.8	11.1	11.3	11.2	11.4	11.8	11.1	11.7	11.1	11.0	11.6	11.1	10.8
HAR-SV	9.0	10.0	9.0	9.4	8.8	9.6	9.3	8.7	9.3	9.0	8.9	8.9	9.3	8.8
HAR-CJ	8.6	9.3	9.1	9.0	9.0	9.1	9.3	8.7	9.0	8.9	8.9	8.7	9.1	9.1
HAR-SJ	9.2	10.1	9.9	9.3	9.5	9.6	10.1	9.3	9.8	9.4	9.5	10.1	9.4	9.3
RV-FTS	9.1	8.8	9.4	9.5	9.8	9.3	9.0	9.7	9.2	9.5	9.2	9.0	9.4	9.4
Panel D: Average ranking of the combination forecast models														
C-GARCH-S_U	12.1	9.6	10.1	10.2	10.2	10.1	9.8	10.8	9.6	11.0	11.0	11.3	10.1	10.9
C-FR-GARCH-S_U	8.9	8.4	9.1	9.1	9.2	9.1	8.6	9.3	9.4	8.8	9.2	8.6	9.3	9.1
C-AR	9.9	10.5	10.2	10.4	10.5	10.5	10.7	10.4	10.5	10.2	10.0	10.3	10.1	10.1
C-HAR	8.9	9.3	9.3	9.3	9.6	9.2	9.1	9.4	9.2	9.2	9.1	8.9	9.4	9.2
C-HAR-SV	8.7	8.8	8.5	9.2	8.7	8.9	8.7	8.7	8.7	8.6	8.7	8.4	8.8	8.8
C-HAR-CJ	8.5	8.5	8.6	8.7	8.8	8.7	8.8	8.6	8.5	8.7	8.6	8.4	8.6	8.8
C-HAR-SJ	8.5	8.8	9.1	8.7	9.0	9.1	9.4	9.1	9.0	9.0	8.8	9.3	8.9	8.8
CC-Bench	8.9	9.1	9.2	8.9	8.9	8.7	9.1	8.7	9.4	9.0	9.0	8.9	9.0	9.0
CC-FTS	8.9	9.1	9.1	9.1	8.9	8.7	9.0	8.6	9.2	9.0	8.9	8.8	9.0	9.1
CC-FTS-Top	8.7	8.9	9.1	8.9	8.8	8.7	9.2	8.6	9.0	8.9	8.7	8.7	9.0	9.0

Notes: The values for row HAR in Panel A correspond to the MSE (Panel A). Values from remaining row of Panel A and B are % changes in MSE against the HAR model. Negative percentage changes are improvements. In Panel C and D we report the average rank. USA-1 corresponds to the S&P 500 market index and USA-2 to the NASDAQ 100 index.

Table 3: Day-ahead forecasts under QLIKE

Model	AUS	IND	FRA	GBR	DEU	CAN	HKG	USA-1	KOR	JPN	USA-2	CHN	CHE	EUR
Panel A: Loss comparison with individual volatility models relative to the HAR model														
HAR [QLIKE]	0.28	0.28	0.24	0.23	0.24	0.26	0.31	0.28	0.27	0.29	0.24	0.26	0.22	0.21
GARCH-S_U	24.0%	6.0%	20.6%	18.5%	16.3%	20.9%	9.0%	27.6%	8.8%	911.4%	39.7%	44.7%	23.0%	32.3%
FR-GARCH-S_U	4.5%	6.4%	22.6%	16.3%	10.8%	16.2%	9.4%	24.6%	16.1%	12.7%	28.9%	20.4%	17.4%	18.9%
AR	8.3%	20.7%	18.9%	19.4%	18.3%	22.8%	30.9%	13.3%	24.4%	15.9%	14.0%	20.7%	14.6%	16.2%
HAR-SV	-17.2%	-0.3%	5.9%	2.7%	-3.4%	87.8%	3.1%	5.2%	-2.6%	-3.1%	0.6%	-3.0%	0.9%	1.5%
HAR-CJ	-19.9%	-3.9%	-8.9%	-4.6%	-9.4%	28.0%	1.2%	7.7%	-7.3%	-4.3%	-3.9%	-2.7%	-11.0%	-4.7%
HAR-SJ	-15.9%	8.1%	-2.9%	-1.7%	1.5%	118.5%	4.6%	13.4%	2.4%	1.8%	21.3%	32.9%	15.5%	-2.6%
RV-FTS	-11.4%	-1.9%	9.7%	7.1%	8.2%	5.7%	5.2%	11.6%	4.3%	5.9%	8.0%	7.4%	7.6%	10.2%
Panel B: Loss comparison with combination forecast models relative to the HAR model														
C-GARCH-S_U	-2.6%	-7.8%	4.2%	3.9%	1.0%	3.9%	-0.1%	8.1%	0.9%	7.8%	12.0%	13.9%	6.7%	10.8%
C-FR-GARCH-S_U	-13.7%	-6.9%	5.9%	4.3%	0.7%	2.5%	1.6%	4.8%	5.2%	3.3%	5.3%	6.2%	6.4%	6.4%
C-AR	-8.4%	2.8%	6.4%	6.6%	7.5%	6.3%	11.2%	5.7%	7.4%	5.3%	4.8%	6.6%	4.2%	7.1%
C-HAR	-15.3%	-5.3%	-0.3%	0.6%	-0.4%	-0.8%	0.4%	1.5%	0.7%	0.8%	0.5%	1.3%	0.5%	1.8%
C-HAR-SV	-17.4%	-10.2%	-6.1%	0.9%	-5.0%	30.8%	-1.6%	-2.0%	-5.8%	-3.4%	-2.5%	-4.7%	-4.9%	1.3%
C-HAR-CJ	-19.1%	-12.4%	-7.5%	-4.2%	-7.3%	-3.0%	-2.8%	-1.3%	-7.0%	-4.1%	-3.2%	-2.9%	-7.0%	-2.5%
C-HAR-SJ	-20.9%	-9.4%	-5.8%	-4.3%	-6.8%	12.7%	-1.2%	-0.4%	-5.5%	-2.3%	0.1%	2.5%	-5.0%	-2.5%
CC-Bench	-20.0%	-11.7%	-7.6%	-5.1%	-9.8%	2.4%	-1.1%	-6.0%	-5.8%	-4.4%	-4.1%	-3.8%	-8.0%	-4.5%
CC-FTS	-19.0%	-11.9%	-7.6%	-3.6%	-9.4%	15.1%	-1.6%	-6.0%	-5.9%	-4.4%	-4.4%	-4.1%	-8.7%	-2.1%
CC-FTS-Top	-20.2%	-13.8%	-8.3%	-3.8%	-9.5%	-3.9%	-3.2%	-8.0%	-8.0%	-4.3%	-5.4%	-5.1%	-8.9%	-3.2%
Panel C: Average ranking of the individual forecast models														
HAR	9.0	10.2	9.6	9.4	9.8	9.4	9.4	9.5	9.5	9.4	9.4	9.2	9.6	9.3
GARCH-S_U	13.6	10.8	11.2	11.1	11.1	11.2	10.7	11.9	10.3	12.1	12.3	12.8	11.1	12.0
FR-GARCH-S_U	9.4	8.9	9.5	9.6	9.5	9.6	8.9	9.7	10.0	9.2	10.0	9.1	9.7	9.7
AR	11.0	11.7	11.0	11.3	11.1	11.4	11.8	11.1	11.6	11.1	11.0	11.6	11.0	10.8
HAR-SV	9.0	10.0	9.0	9.3	8.8	9.9	9.3	8.7	9.3	9.0	8.9	8.9	9.4	8.8
HAR-CJ	8.7	9.3	9.1	9.0	9.0	9.1	9.3	8.7	9.0	8.9	8.9	8.7	9.1	9.1
HAR-SJ	9.2	10.1	9.8	9.3	9.5	9.6	10.1	9.4	9.8	9.4	9.5	10.2	9.5	9.3
RV-FTS	9.2	8.9	9.4	9.6	9.8	9.3	9.1	9.8	9.2	9.5	9.2	9.1	9.4	9.4
Panel D: Average ranking of the combination forecast models														
C-GARCH-S_U	12.0	9.6	10.1	10.2	10.2	10.1	9.8	10.8	9.6	11.0	10.9	11.2	10.1	10.8
C-FR-GARCH-S_U	8.9	8.4	9.1	9.1	9.2	9.1	8.7	9.4	9.4	8.9	9.2	8.7	9.3	9.1
C-AR	9.9	10.4	10.2	10.4	10.5	10.5	10.7	10.4	10.4	10.2	10.0	10.3	10.1	10.1
C-HAR	9.0	9.3	9.3	9.3	9.6	9.2	9.1	9.4	9.2	9.3	9.1	8.9	9.4	9.2
C-HAR-SV	8.8	8.8	8.5	9.2	8.7	9.1	8.7	8.7	8.7	8.6	8.7	8.4	8.8	8.8
C-HAR-CJ	8.5	8.5	8.6	8.7	8.8	8.7	8.8	8.6	8.5	8.7	8.6	8.4	8.6	8.8
C-HAR-SJ	8.6	8.8	9.1	8.7	9.0	9.1	9.4	9.1	9.0	9.0	8.8	9.3	8.8	8.8
CC-Bench	8.9	9.1	9.2	8.9	8.9	8.7	9.1	8.7	9.4	9.0	9.0	8.8	9.0	8.9
CC-FTS	8.9	9.1	9.1	9.1	8.9	8.6	9.0	8.6	9.2	8.9	8.9	8.8	8.9	9.1
CC-FTS-Top	8.7	8.9	9.1	8.9	8.8	8.7	9.2	8.6	9.0	8.8	8.7	8.6	9.0	8.9

Notes: The values for row HAR in Panel A correspond to the QLIKE (Panel A). Values from remaining row of Panel A and B are % changes in QLIKE against the HAR model. Negative percentage changes are improvements. In Panel C and D we report the average rank. USA-1 corresponds to the S&P 500 market index and USA-2 to the NASDAQ 100 index.

4.3. Multiple day-ahead forecasts

In this section, we present results for multiple-day ahead forecasts, where the average volatility over the next two to twenty-two trading days is predicted either recursively (GARCH-S_U and RF-GARCH-S_U models) or directly (HAR, RV-FTS). The numerical results for five- and twenty-two-day-ahead forecasts can be found in Appendix B (see Tables B.1, B.2, B.3 and B.4) In Figures 5 (for MSE) and 6 (for QLIKE), we present the average ranks (y-axis) over the forecasting horizons (x-axis) for multiple model types. Each dot on the line indicates that the model in the given forecasting horizon belongs to the set of superior models. The first points on the plots correspond to the day-ahead analysis we reported in the previous section.

The black line shows the average rank of the best individual volatility forecasting model, except for the RV-FTS model. The blue line corresponds to the average rank of the best pairwise combination forecasts that include the RV-FTS model. The pairwise combination (blue lines) is preferred in 12 out of 14 countries across most forecasting horizons, with longer-forecasting horizons occasionally excluding the benchmark model from the set of superior models. These results hold for both the MSE and QLIKE (see Figure 6).

The red and green lines correspond to average ranks of the conditional combination forecasts, reflecting more realistic scenarios where the analyst is unaware of which model to use. The red line corresponds to CC-Bench, a conditional weighted average across seven individual models (excluding RV-FTS), and CC-FTS, a conditional weighted average across eight models that includes the RV-FTS model. With increasing forecast horizons, both conditional combinations tend to yield more accurate forecasts. For Australia, the US (S&P 500), India, South Korea, Japan, the US (NASDAQ), Germany, China and Canada, Switzerland, and Hong Kong, the CC-FTS tends to outperform the CC-Bench for multiple forecast horizons, whereas for the remaining countries, forecasts are still comparable between CC-FTS and CC-Bench. Overall, the findings suggests that RV-FTS can be a useful addition to volatility forecasting models, especially with increasing forecast horizons. Figure 4 shows that the dynamics of weights differ with forecasting horizon, with twenty-two-day-ahead forecasts occasionally estimated with weights greater than 0.8 or greater for the RV-FTS model. Generally, the periods of non-zero weights show that RV-FTS often considerably influences the CC-FTS model's forecasts.

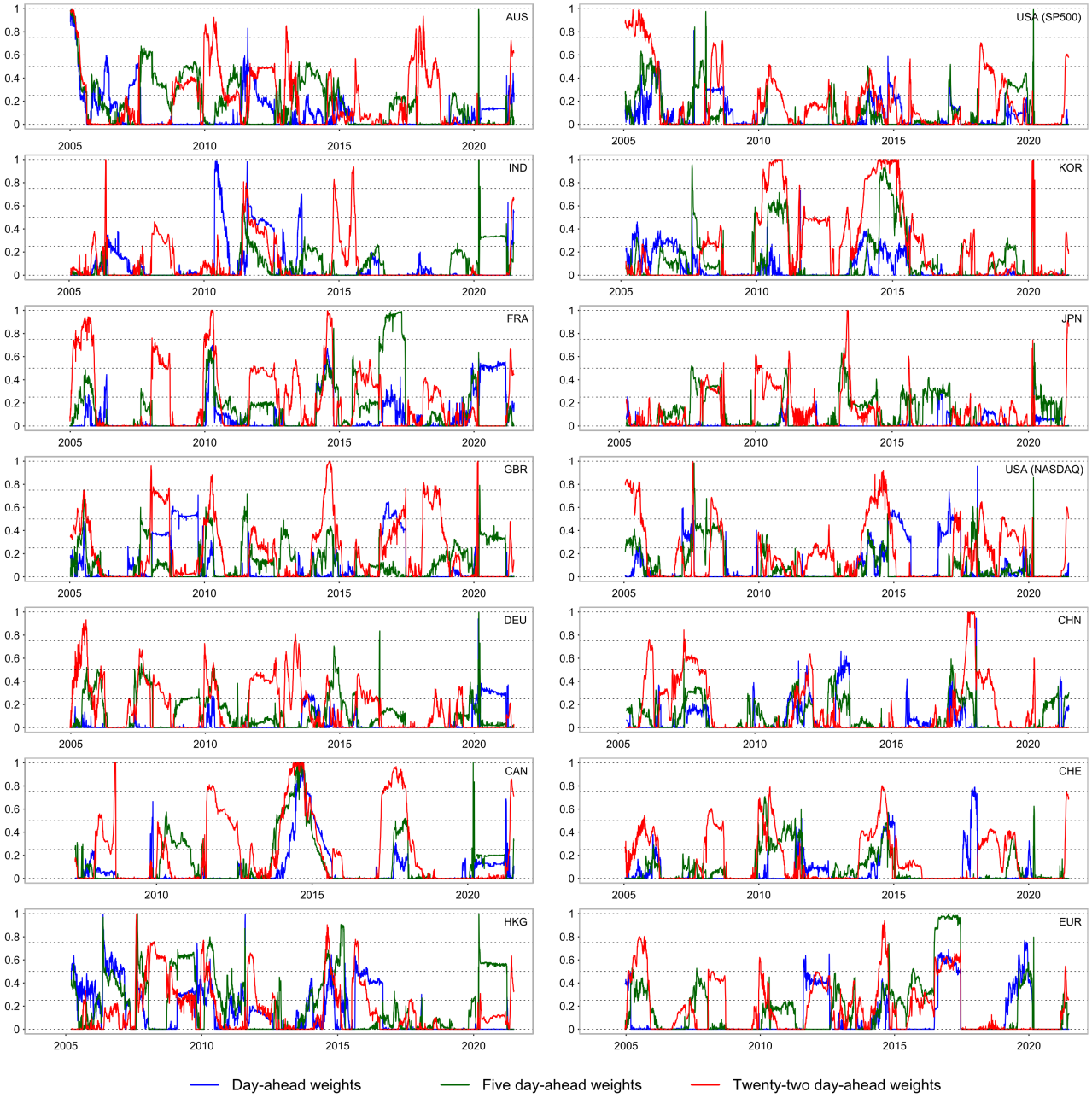

Figure 4: Optimum weights for RV-FTS of conditional combination model

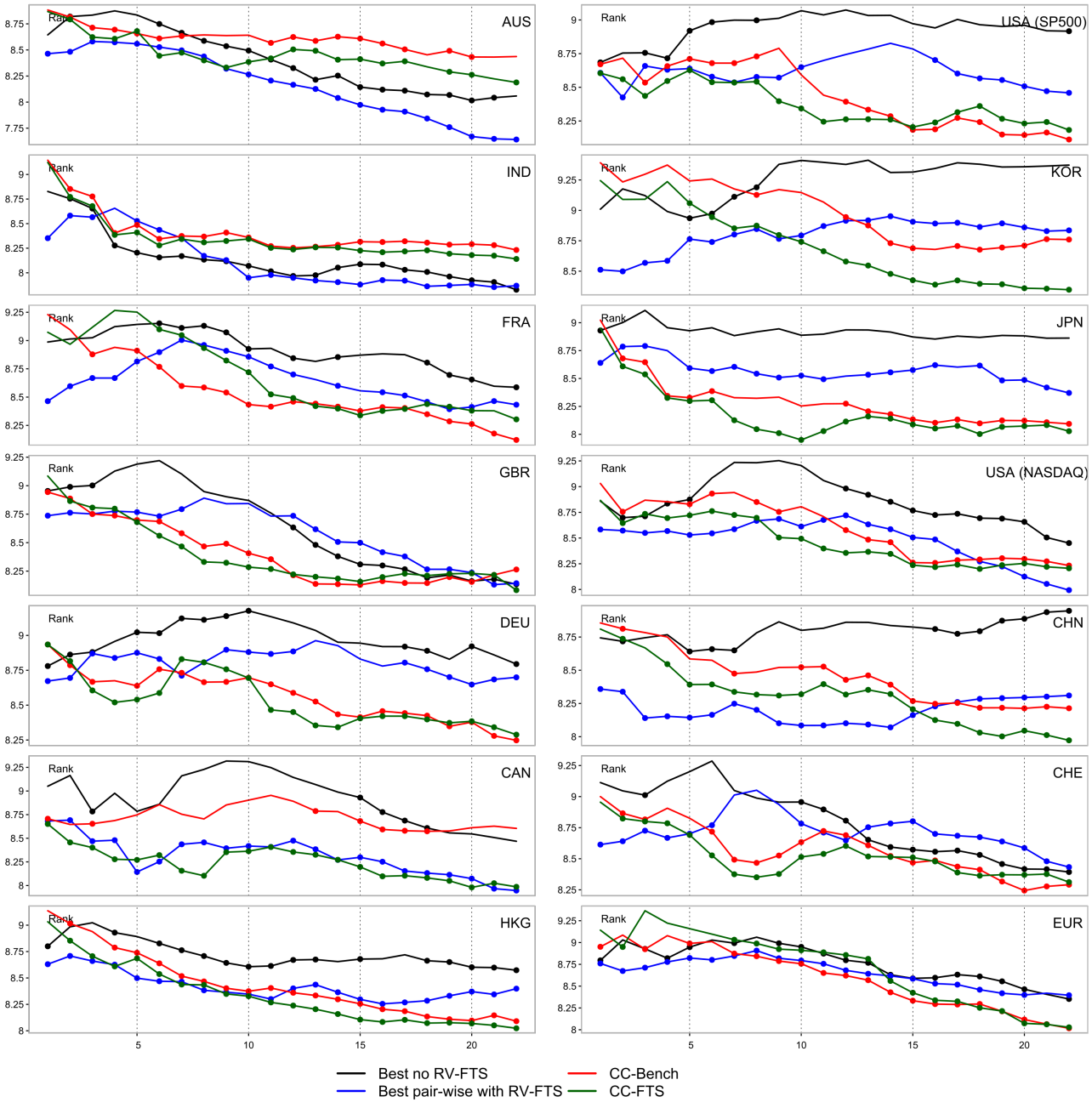

Figure 5: Average ranks (y-axis) of top-performing models under MSE across forecast horizons (x-axis)

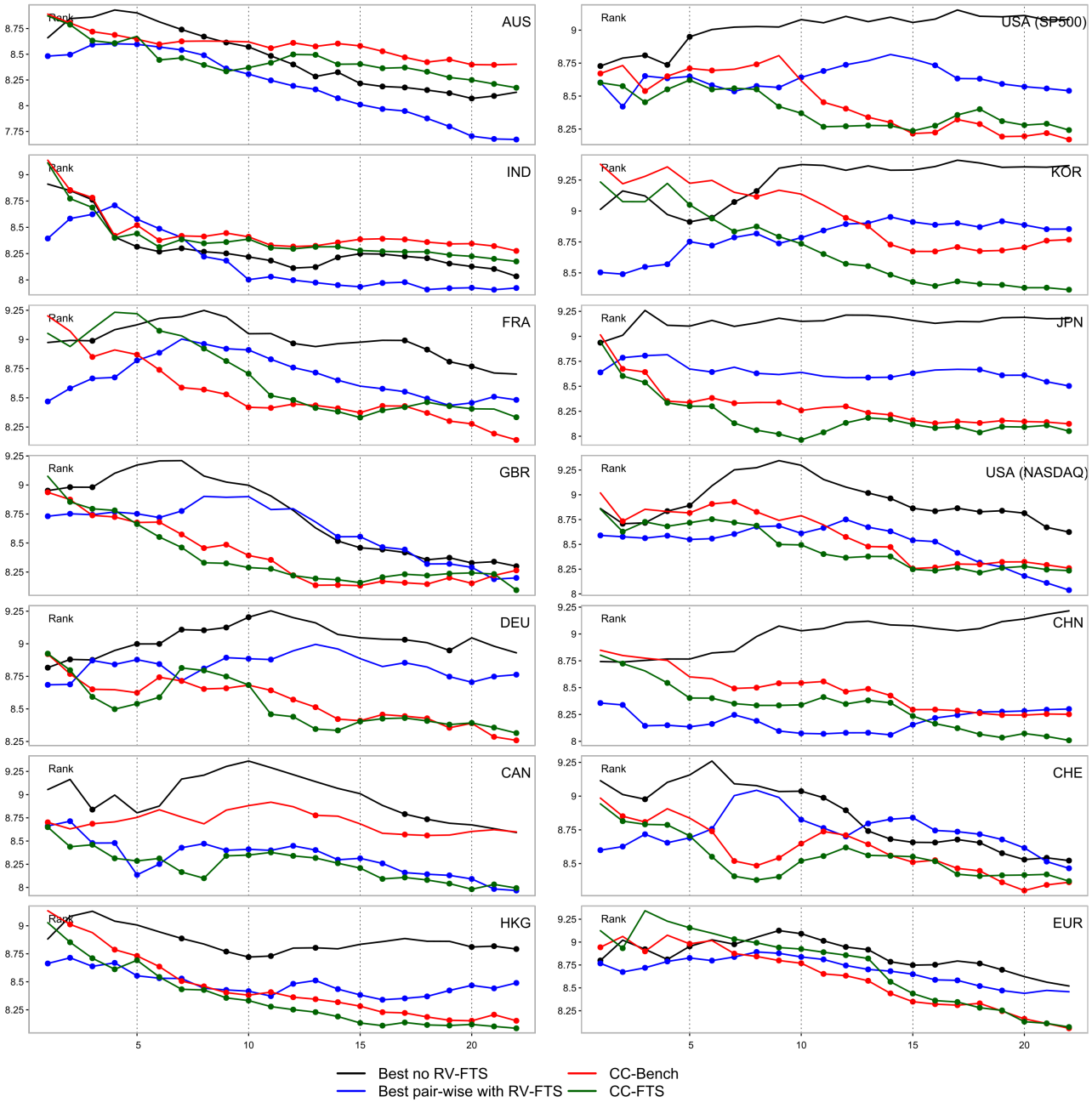

Figure 6: Average ranks (y-axis) of top-performing models under QLIKE across forecast horizons (x-axis)

5. Conclusions and Future Work

In our study, we propose an RV-FTS volatility model inspired by fuzzy logic and the persistence of the volatility process. Our approach involves separating the observed realized volatility into time-varying c sets (clusters) and employing a simple exponential moving average process that models the dynamics of the membership of the next period volatility into the clusters. We compare the volatility forecasting accuracy of the RV-FTS model using data from 14 stock market indices: AUS, USA (S&P 500), IND, KOR, FRA, JPN, GBR, USA (NASDAQ), DEU, CHN, CAN, CHE, HKG, and EUR. Apart from the proposed RV-FTS model, we use seven benchmark models: GARCH-S_U, ARFIMA-GARCH-S_U (RF-GARCH-S_U), AR, HAR, HAR-CJ (based on (Andersen et al., 2012)), HAR-SV and HAR-SJ (based on (Patton and Sheppard, 2015)). We find that across all markets and longer-forecast horizons, the RV-FTS model almost always outperforms the GARCH-S_U and RF-GARCH-S_U models and the AR model, whereas it performs less competitively against the HAR and HAR-CJ models in the short day-ahead forecasting setting.

However, we find that the proposed model can be used to complement even the best individual forecasting model in a pairwise combination forecast. Moreover, conditional combinations across all the models show that for all the markets and forecast horizons, there are periods where the RV-FTS model receives considerable weights in forming the conditional forecast. Our key observations are that with increasing forecasting horizon, the utility of the RV-FTS model tends to increase, which we attribute to the fact that volatility regimes are persistent, a feature captured by clusters in the RV-FTS framework.

We contribute to the volatility forecasting literature by designing a RV-FTS volatility model, which, despite its simplicity, leads to accurate volatility forecasts that often outperform several benchmark model forecasts. Our study expands the class of fuzzy-based models in volatility forecasting (e.g., Hung, 2011; Maciel et al., 2016; Dash and Dash, 2016; García and Kristjanpoller, 2019). We also contribute to the combination forecasting literature and demonstrate that RV-FTS is useful, particularly in conjunction with other models, as it improves forecasts even with the most accurate individual models. Finally, the clusters in the RV-FTS correspond to different volatility periods, which improves the interpretability of the model.

The limitation of this model is the need to estimate two hyperparameters, the number of clusters, c, and the smoothing parameter, ρ, which is carried out using rolling window cross-validation approach. Note that HAR model also uses information about past five- and twenty-two-day volatility, and HAR-CJ also continuous and jump components. The later model often performs the best in the short term and compared to RV-FTS utilizes additional information that can be decomposed from the realized volatility. On the other hand, the proposed RV-FTS model relies on only the realized volatility, similar to an AR model that almost always outperforms other models. The natural extension of the RV-FTS will be to use an expanded information set to either create clusters (volatility regimes) or govern the dynamics of the memberships, which will be explored in future research.

Acknowledgement

This work was supported by MSCA-Fellow 6 MUNI (CZ.02.01.01/00/22_010/0007541) and APVV-22-0472.

Data availability statement

The data is available upon request from the corresponding author at stefan.lyocsa@econ.muni.cz.

References

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001). The distribution of realized stock return volatility. Journal of financial economics, 61(1):43–76.

Andersen, T. G., Bollerslev, T., and Huang, X. (2011). A reduced form framework for modeling volatility of speculative prices based on realized variation measures. Journal of Econometrics, 160(1):176–189.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. Journal of Econometrics, 169(1):75–93.

Araneda, A. A. (2021). Asset volatility forecasting: The optimal decay parameter in the ewma model. arXiv preprint arXiv:2105.14382.

Audrino, F., Sigrist, F., and Ballinari, D. (2020). The impact of sentiment and attention measures on stock market volatility. International Journal of Forecasting, 36(2):334–357.

Ayele, A. W., Gabreyohannes, E., and Tesfay, Y. Y. (2017). Macroeconomic determinants of volatility for the gold price in ethiopia: the application of garch and ewma volatility models. Global Business Review, 18(2):308–326.

Baek, S., Mohanty, S. K., and Glambosky, M. (2020). Covid-19 and stock market volatility: An industry level analysis. Finance research letters, 37:101748.

Bang, Y.-K. and Lee, C.-H. (2011). Fuzzy time series prediction using hierarchical clustering algorithms. Expert Systems with Applications, 38(4):4312–4325.

Baruník, J. and Křehlík, T. (2016). Combining high frequency data with non-linear models for forecasting energy market volatility. Expert Systems with Applications, 55:222–242.

Bas, E., Yolcu, U., and Egrioglu, E. (2021). Intuitionistic fuzzy time series functions approach for time series forecasting. Granular computing, 6(3):619–629.

Bernardi, M. and Catania, L. (2018). The model confidence set package for r. International Journal of Computational Economics and Econometrics, 8(2):144–158.

Bezdek, J. C., Ehrlich, R., and Full, W. (1984). Fcm: The fuzzy c-means clustering algorithm. Computers & geosciences, 10(2-3):191–203.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of econometrics, 31(3):307–327.

Bollerslev, T., Hood, B., Huss, J., and Pedersen, L. H. (2018a). Risk everywhere: Modeling and managing volatility. The Review of Financial Studies, 31(7):2729–2773.

Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. Journal of Econometrics, 192(1):1–18.

Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2018b). Modeling and forecasting (un) reliable realized covariances for more reliable financial decisions. Journal of Econometrics, 207(1):71–91.

Buccheri, G. and Corsi, F. (2021). Hark the shark: Realized volatility modeling with measurement errors and nonlinear dependencies. Journal of Financial Econometrics, 19(4):614–649.

Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P., and Oliveira, A. L. (2016). Computational intelligence and financial markets: A survey and future directions. Expert Systems with Applications, 55:194–211.

Chen, M.-Y. (2014). A high-order fuzzy time series forecasting model for internet stock trading. Future Generation Computer Systems, 37:461–467.

Chen, M.-Y. and Chen, B.-T. (2015). A hybrid fuzzy time series model based on granular computing for stock price forecasting. Information Sciences, 294:227–241.

Chen, S.-M. (1996). Forecasting enrollments based on fuzzy time series. Fuzzy sets and systems, 81(3):311–319.

Chen, S.-M. and Tanuwijaya, K. (2011). Multivariate fuzzy forecasting based on fuzzy time series and automatic clustering techniques. Expert Systems with Applications, 38(8):10594–10605.

Chen, S.-M., Zou, X.-Y., and Gunawan, G. C. (2019). Fuzzy time series forecasting based on proportions of intervals and particle swarm optimization techniques. Information Sciences, 500:127–139.

Choi, P. and Nam, K. (2008). Asymmetric and leptokurtic distribution for heteroscedastic asset returns: the su-normal distribution. Journal of Empirical finance, 15(1):41–63.

Christensen, K., Siggaard, M., and Veliyev, B. (2023). A machine learning approach to volatility forecasting. Journal of Financial Econometrics, 21(5):1680–1727.

Chun, D., Cho, H., and Ryu, D. (2024). Volatility forecasting and volatility-timing strategies: A machine learning approach. Research in International Business and Finance, page 102723.

Cipollini, F., Gallo, G. M., and Otranto, E. (2021). Realized volatility forecasting: Robustness to measurement errors. International Journal of Forecasting, 37(1):44–57.

Clements, A. and Preve, D. P. (2021). A practical guide to harnessing the har volatility model. Journal of Banking & Finance, 133:106285.

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

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics, 7(2):174–196.

Corsi, F., Pirino, D., and Reno, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. Journal of Econometrics, 159(2):276–288.

Dash, R. and Dash, P. K. (2016). An evolutionary hybrid fuzzy computationally efficient egarch model for volatility prediction. Applied Soft Computing, 45:40–60.

Dash, R., Dash, P. K., and Bisoi, R. (2015). A differential harmony search based hybrid interval type2 fuzzy egarch model for stock market volatility prediction. International Journal of Approximate Reasoning, 59:81–104.

D'Urso, P., De Giovanni, L., and Massari, R. (2016). Garch-based robust clustering of time series. Fuzzy Sets and Systems, 305:1–28.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica: Journal of the econometric society, pages 987–1007.

Engle, R. F. et al. (1999). A long-run and short-run component model of stock return volatility. Cointigration, Causality, and forecating, pages 475–497.

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

García, D. and Kristjanpoller, W. (2019). An adaptive forecasting approach for copper price volatility through hybrid and non-hybrid models. Applied soft computing, 74:466–478.

Ghalanos, A. (2020). Introduction to the rugarch package.(version 1.3-1). Manuscript, http://cran. r-project. org/web/packages/rugarch. Accessed, 11(8).

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

Granger, C. W. and Ramanathan, R. (1984). Improved methods of combining forecasts. Journal of forecasting, 3(2):197–204.

Gupta, K. K. and Kumar, S. (2023). K-means clustering based high order weighted probabilistic fuzzy time series forecasting method. Cybernetics and Systems, 54(2):197–219.

Halousková, M. and Lyócsa, Š. (2025). Forecasting us equity market volatility with attention and sentiment to the economy. arXiv preprint arXiv:2503.19767.

Hansen, P. R. and Huang, Z. (2016). Exponential garch modeling with realized measures of volatility. Journal of Business & Economic Statistics, 34(2):269–287.

Hansen, P. R., Huang, Z., and Shek, H. H. (2012). Realized garch: a joint model for returns and realized measures of volatility. Journal of Applied Econometrics, 27(6):877–906.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. Econometrica, 79(2):453–497.

Horpestad, J. B., Lyócsa, Š., Molnár, P., and Olsen, T. B. (2019). Asymmetric volatility in equity markets around the world. The North American Journal of Economics and Finance, 48:540–554.

Hung, J.-C. (2011). Applying a combined fuzzy systems and garch model to adaptively forecast stock market volatility. Applied Soft Computing, 11(5):3938–3945.

Jiao, X., Song, Y., Kong, Y., and Tang, X. (2022). Volatility forecasting for crude oil based on text information and deep learning pso-lstm model. Journal of Forecasting, 41(5):933–944.

Johnson, N. L. (1949a). Bivariate distributions based on simple translation systems. Biometrika, 36(3/4):297–304.

Johnson, N. L. (1949b). Systems of frequency curves generated by methods of translation. Biometrika, 36(1/2):149–176.

Khan, M. H., Ahmed, J., Mughal, M., and Khan, I. H. (2023). Oil price volatility and stock returns: Evidence from three oil-price wars. International Journal of Finance & Economics, 28(3):3162–3182.

Kock, A. B. and Teräsvirta, T. (2014). Forecasting performances of three automated modelling techniques during the economic crisis 2007–2009. International Journal of Forecasting, 30(3):616–631.

Kolokolov, A. and Renò, R. (2024). Jumps or staleness? Journal of Business & Economic Statistics, 42(2):516–532.

Kourentzes, N., Barrow, D., and Petropoulos, F. (2019). Another look at forecast selection and combination: Evidence from forecast pooling. International Journal of Production Economics, 209:226–235.

Liu, G., Zhuang, Z., and Wang, M. (2024). Forecasting the high-frequency volatility based on the lstm-hit model. Journal of Forecasting, 43(5):1356–1373.

Lu, W., Yang, J., Liu, X., and Pedrycz, W. (2014). The modeling and prediction of time series based on synergy of high-order fuzzy cognitive map and fuzzy c-means clustering. Knowledge-Based Systems, 70:242–255.

Lu, X., Que, D., and Cao, G. (2016). Volatility forecast based on the hybrid artificial neural network and garch-type models. Procedia Computer Science, 91:1044–1049.

Lyócsa, Š., Molnár, P., and Vỳrost, T. (2021). Stock market volatility forecasting: Do we need high-frequency data? International Journal of Forecasting, 37(3):1092–1110.

Lyócsa, Š. and Plíhal, T. (2022). Russia's ruble during the onset of the russian invasion of ukraine in early 2022: The role of implied volatility and attention. Finance Research Letters, 48:102995.

Lyócsa, Š., Plíhal, T., and Vỳrost, T. (2024). Forecasting day-ahead expected shortfall on the eur/usd exchange rate: The (i) relevance of implied volatility. International Journal of Forecasting, 40(4):1275–1301.

Maciel, L., Gomide, F., and Ballini, R. (2016). Evolving fuzzy-garch approach for financial volatility modeling and forecasting. Computational Economics, 48:379–398.

Maneesoonthorn, W., Martin, G. M., and Forbes, C. S. (2020). High-frequency jump tests: Which test should we use? Journal of Econometrics, 219(2):478–487.

Molnár, P. (2012). Properties of range-based volatility estimators. International Review of Financial Analysis, 23:20–29.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica: Journal of the econometric society, pages 347–370.

Patra, S. (2024). An empirical analysis of the volume-volatility nexus in crude oil markets under structural breaks: Implications for forecasting. International Review of Economics & Finance, 94:103434.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics, 160(1):246–256.

Patton, A. J. and Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. Review of Economics and Statistics, 97(3):683–697.

RiskMetrics (1996). RiskMetrics Technical Document. J.P. Morgan/Reuters, New York, fourth edition. Fourth Edition.

Saggi, M. K. and Jain, S. (2018). A survey towards an integration of big data analytics to big insights for value-creation. Information Processing & Management, 54(5):758–790.

Singh, P. (2021). Fqtsfm: A fuzzy-quantum time series forecasting model. Information Sciences, 566:57–79.

Song, Q. and Chissom, B. S. (1993). Fuzzy time series and its models. Fuzzy sets and systems, 54(3):269–277.

Song, Q. and Chissom, B. S. (1994). Forecasting enrollments with fuzzy time series—part ii. Fuzzy sets and systems, 62(1):1–8.

Timmermann, A. (2006). Forecast combinations. Handbook of economic forecasting, 1:135–196.

Vilela, L. F., Leme, R. C., Pinheiro, C. A., and Carpinteiro, O. A. (2019). Forecasting financial series using clustering methods and support vector regression. Artificial Intelligence Review, 52:743–773.

Vortelinos, D. I. (2017). Forecasting realized volatility: Har against principal components combining, neural networks and garch. Research in international business and finance, 39:824–839.

Wang, W., Lin, W., Wen, Y., Lai, X., Peng, P., Zhang, Y., and Li, K. (2023a). An interpretable intuitionistic fuzzy inference model for stock prediction. Expert Systems with Applications, 213:118908.

Wang, X., Hyndman, R. J., Li, F., and Kang, Y. (2023b). Forecast combinations: An over 50-year review. International Journal of Forecasting, 39(4):1518–1547.

Wong, Z. Y., Chin, W. C., and Tan, S. H. (2016). Daily value-at-risk modeling and forecast evaluation: The realized volatility approach. The Journal of Finance and Data Science, 2(3):171–187.

Wu, C.-C. and Chiu, J. (2017). Economic evaluation of asymmetric and price range information in gold and general financial markets. Journal of International Money and Finance, 74:53–68.

Yolcu, O. C. and Yolcu, U. (2023). A novel intuitionistic fuzzy time series prediction model with cascaded structure for financial time series. Expert Systems with Applications, 215:119336.

Yu, T. H.-K. and Huarng, K.-H. (2010). A neural network-based fuzzy time series model to improve forecasting. Expert Systems with Applications, 37(4):3366–3372.

Zadeh, L. A. (1965). Fuzzy sets. Information and Control.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. Journal of Economic Dynamics and control, 18(5):931–955.

Table A.1: Daily returns and realized volatility components

	Panel A: Returns							Panel D: Signed jump						
	Mean	SD	SK	KU	$\rho(1)$	$\rho(5)$	$\rho(22)$	Mean	SD	SK	KU	$\rho(1)$	$\rho(5)$	$\rho(22)$
Australia	0.00	0.9	6.3	-0.8	0.00	0.00	-0.01	6.2	220.7	506.8	13	-0.03	0.08	0.00
India	0.00	1.5	9.2	-0.4	0.04	0.03	0.01	24	256.1	443.4	-7.7	-0.07	-0.04	-0.01
France	0.00	1.4	5.9	-0.2	-0.01	-0.02	0.01	7.7	167.9	120.8	3.7	-0.09	0.04	0.07
UK	0.00	1.2	7.2	-0.3	-0.02	-0.01	0.02	11.4	455.8	709.7	19.7	-0.11	0.09	0.00
Germany	0.00	1.5	6.4	-0.2	0.01	-0.02	0.02	12.4	280.4	168.2	1.3	0.03	-0.01	0.01
Canada	0.00	1.1	16.8	-1.2	-0.03	0.01	0.00	-4.2	1041.3	3650.9	-55.9	-0.11	0.06	0.00
Hong Kong	0.00	1.5	7.7	-0.1	-0.01	-0.03	-0.02	8.7	185.2	198.9	-2.9	-0.08	-0.03	-0.07
US (Nasdaq)	0.00	1.6	6.7	-0.2	-0.06	-0.01	-0.01	1.1	217.7	620.9	2.7	-0.13	-0.01	0.00
South Korea	0.00	1.5	7.6	-0.6	0.02	0	-0.02	10.3	155.3	114.4	4.5	0.03	0.06	0.01
Japan	0.00	1.5	6.3	-0.4	-0.03	0.02	-0.01	7.9	207.7	131.3	5.4	-0.05	0.03	0.01
US (SP 500)	0.00	1.2	10.7	-0.4	-0.11	0.01	0.01	2.9	238.1	92.3	1.4	-0.15	0.01	0.06
China	0.00	1.5	5.1	-0.4	0.03	0.00	0.01	-6.6	236.8	45.6	-1	-0.11	-0.02	-0.01
Swiss	0.00	1.2	8.3	-0.3	0.02	0.01	0.01	6.4	249.5	1022.9	22.9	-0.12	0.12	0.00
Europe	0.00	1.4	5.8	-0.2	-0.01	-0.03	0.02	17.7	421.6	540.9	15.4	-0.06	0.07	0.01
	Panel B: Continuous component							Panel E: Negative semi-variance						
Australia	63.8	161.2	256.9	13.1	0.51	0.28	0.13	37	88.3	294.5	13.8	0.75	0.44	0.22
India	173.6	492	643.9	21.6	0.37	0.15	0.08	167.3	313.5	176.3	9.5	0.61	0.38	0.25
France	164.8	302.8	146.8	9.5	0.63	0.38	0.25	154	293.7	117.1	8.7	0.73	0.43	0.26
UK	147.4	313.9	127.2	9.3	0.55	0.32	0.22	98.6	195.9	114.6	8.8	0.75	0.47	0.32
Germany	196	369.1	156.4	9.2	0.69	0.43	0.29	170.4	315.5	100.4	7.5	0.76	0.49	0.33
Canada	109	1059.4	4104.7	61.9	0.11	0.13	0.06	60.3	178.7	126.9	9.9	0.8	0.5	0.39
Hong Kong	117.3	225.8	288.2	13.4	0.52	0.34	0.27	50.6	91.5	337.2	13.2	0.66	0.35	0.25
US (Nasdaq)	156.4	331.4	147.1	9.5	0.57	0.38	0.25	162.5	351.4	202.1	10.4	0.62	0.39	0.27
South Korea	145.2	270.5	155.3	9.4	0.69	0.43	0.32	134	251.7	106.6	8.1	0.73	0.54	0.35
Japan	124	209.5	96.1	8	0.64	0.42	0.19	71	134.6	167.5	10.5	0.66	0.42	0.18
US (SP 500)	136.9	341.6	230.1	11.7	0.62	0.43	0.28	122.2	316.2	156.1	10.4	0.72	0.46	0.32
China	202.3	354.4	76.3	6.8	0.56	0.34	0.29	106.9	199.3	129.9	8	0.61	0.35	0.32
Swiss	105.8	241.9	211.9	11.6	0.65	0.37	0.2	95	177.9	95.5	7.9	0.73	0.47	0.31
Europe	189.7	359.9	99.6	8.2	0.64	0.38	0.23	154	286.4	78.5	7.3	0.73	0.44	0.29
	Panel C: Jump component							Panel F: Positive semi-variance						
Australia	96.8	288.5	550.1	19.4	0.5	0.26	0.11	70	248.4	779.8	23.4	0.41	0.27	0.1
India	203.9	778	1224.4	30.3	0.28	0.07	0.04	197.5	505.1	585.5	19.1	0.44	0.17	0.11
France	183.3	346.8	199.3	11.1	0.54	0.33	0.22	172.5	325.5	143.5	9.5	0.64	0.39	0.25
UK	207.7	597.4	585.1	19.6	0.35	0.27	0.11	158.8	527.7	793.7	23.3	0.25	0.27	0.1
Germany	233.8	461.3	143.9	9.3	0.57	0.39	0.26	208.3	410.9	126.9	8.7	0.57	0.41	0.28
Canada	153.5	1124.7	3221.7	52.4	0.25	0.16	0.07	104.8	409.3	396.5	16.9	0.38	0.26	0.18
Hong Kong	192.8	346.1	198.8	10.9	0.58	0.42	0.34	126	229.6	250.4	11.7	0.54	0.37	0.28
US (Nasdaq)	151.3	332	430.1	15.6	0.48	0.28	0.2	157.5	317	354.7	12.9	0.63	0.36	0.27
South Korea	166.8	334.6	225.7	11.4	0.59	0.32	0.24	155.6	295.9	145.7	9.2	0.66	0.45	0.29
Japan	184.8	330	142.2	9.6	0.56	0.34	0.17	131.9	271.5	194.4	11.5	0.43	0.3	0.15
US (SP 500)	154.5	395.7	268.6	12.9	0.54	0.38	0.22	139.8	359.7	162.3	10.4	0.57	0.41	0.27
China	291.1	496.8	58.5	6.2	0.63	0.34	0.31	195.7	353.7	54.2	6	0.61	0.32	0.31
Swiss	122.9	372.6	426.1	17.2	0.52	0.33	0.1	112.1	321.5	609.9	20.1	0.43	0.34	0.11
Europe	243	596.7	440	16.2	0.44	0.32	0.14	207.4	547.1	537.4	18.1	0.38	0.31	0.14

Notes: SD, SK, KU denote standard deviation, skewness and kurtosis and $\rho(.)$ auto-correlation of the given order.

Appendix A. Additional data characteristics

Appendix B. Multiple day-ahead forecasts

Table B.1: Five day-ahead forecasts under MSE

Model	AUS	IND	FRA	GBR	DEU	CAN	HKG	USA-1	KOR	JPN	USA-2	CHN	CHE	EUR
Panel A: Loss comparison with individual volatility models relative to the HAR model														
HAR [MSE]	6.0	127.4	34.3	22.5	40.3	91.0	35.3	35.1	29.2	11.3	26.2	55.1	24.9	29.5
GARCH-S_U	50.2%	-10.7%	-7.3%	-2.0%	-1.6%	-16.2%	5.2%	48.0%	-3.1%	143.2%	53.0%	5.7%	-0.1%	-4.8%
FR-GARCH-S_U	56.8%	-6.2%	6.4%	-6.1%	-0.6%	-16.2%	3.1%	6.2%	-2.7%	3.9%	1.0%	2.4%	-8.2%	-0.4%
AR	12.9%	-0.7%	3.8%	2.2%	5.8%	-13.6%	36.7%	12.8%	16.9%	15.7%	-3.6%	12.7%	-0.5%	2.0%
HAR-SV	-2.3%	-21.8%	-2.9%	4.0%	-15.0%	-10.4%	-0.3%	0.6%	-1.0%	11.3%	-15.0%	-24.3%	18.3%	-3.5%
HAR-CJ	-2.4%	-22.4%	61.3%	3.3%	-19.1%	-19.7%	-5.8%	20.2%	12.0%	10.7%	0.1%	-18.1%	19.3%	-4.4%
HAR-SJ	-17.1%	-24.1%	20.1%	-1.9%	-16.9%	-30.0%	-0.1%	1.4%	10.1%	15.8%	-10.4%	-14.1%	13.3%	-6.1%
RV-FTS	-3.8%	-10.7%	-4.8%	-4.1%	-12.2%	2.7%	-3.2%	19.9%	-1.6%	4.7%	12.8%	3.9%	-1.9%	-2.9%
Panel B: Loss comparison with combination forecast models relative to the HAR model														
C-GARCH-S_U	3.7%	-14.3%	-11.5%	-10.5%	-15.0%	-15.1%	-6.9%	24.9%	-10.2%	58.5%	24.0%	-2.8%	-6.3%	-11.6%
C-FR-GARCH-S_U	1.0%	-12.4%	-9.4%	-12.3%	-15.8%	-14.1%	-6.3%	5.4%	-12.2%	-2.5%	-3.4%	-3.0%	-12.4%	-11.2%
C-AR	-4.6%	-8.3%	-15.4%	-14.2%	-12.7%	-12.2%	-3.4%	4.4%	-7.7%	-5.4%	-8.5%	-5.7%	-15.4%	-13.4%
C-HAR	-10.0%	-5.8%	-7.1%	-6.0%	-12.0%	-9.6%	-8.3%	2.0%	-6.1%	-1.9%	0.3%	-3.1%	-7.2%	-6.7%
C-HAR-SV	-8.1%	-20.6%	-11.6%	-6.4%	-19.4%	-13.0%	-10.7%	2.8%	-7.7%	-1.2%	-11.2%	-17.2%	-1.1%	-9.1%
C-HAR-CJ	-10.9%	-20.9%	6.0%	-7.7%	-23.2%	-19.9%	-10.6%	5.1%	-4.6%	1.6%	-4.2%	-13.0%	-8.4%	-9.2%
C-HAR-SJ	-16.9%	-21.9%	-0.5%	-7.6%	-20.9%	-23.7%	-9.6%	1.7%	-4.5%	2.9%	-8.9%	-14.5%	-3.7%	-12.6%
CC-Bench	-17.6%	-21.7%	-18.4%	-13.5%	-21.3%	-26.5%	-11.6%	-11.7%	-21.0%	-9.8%	-11.2%	-22.4%	-18.2%	-20.7%
CC-FTS	-15.0%	-22.6%	-18.6%	-14.4%	-23.0%	-29.2%	-12.3%	-5.4%	-21.1%	-8.7%	-11.2%	-22.1%	-17.4%	-20.6%
CC-FTS-Top	-12.1%	-19.7%	-15.2%	-14.9%	-22.6%	-28.1%	-13.1%	-7.3%	-19.8%	-8.0%	-18.8%	-19.8%	-16.8%	-19.6%
Panel C: Average ranking of the individual forecast models														
HAR	9.1	10.3	9.5	9.5	9.9	9.8	9.2	9.6	8.9	9.2	9.3	9.7	9.6	9.3
GARCH-S_U	14.0	10.0	10.9	10.7	10.7	11.2	10.4	12.0	9.8	12.7	12.5	12.5	10.8	11.8
FR-GARCH-S_U	8.8	8.2	9.3	9.2	9.3	9.7	8.9	9.4	9.9	8.9	9.6	8.6	9.4	9.1
AR	11.4	12.1	11.5	11.6	11.5	12.2	12.3	11.5	11.9	11.0	11.1	12.3	11.5	11.1
HAR-SV	9.1	9.7	9.1	9.5	9.0	9.3	9.8	9.0	9.6	9.5	9.2	9.3	9.4	8.9
HAR-CJ	8.9	9.2	9.3	9.2	9.2	8.8	9.8	8.9	9.4	9.5	8.9	8.8	9.2	9.3
HAR-SJ	9.1	9.6	9.7	9.4	9.3	9.5	10.5	9.1	9.9	9.7	9.4	10.0	9.4	9.4
RV-FTS	9.4	10.0	9.8	10.0	10.2	9.5	9.2	9.8	9.5	9.7	9.7	9.2	9.9	9.9
Panel D: Average ranking of the combination forecast models														
C-GARCH-S_U	12.0	9.7	9.9	9.9	9.9	10.0	9.5	10.7	9.3	11.1	11.0	10.9	9.9	10.6
C-FR-GARCH-S_U	8.6	8.5	9.0	9.0	9.3	9.2	8.5	9.3	9.4	8.6	9.3	8.5	9.2	9.0
C-AR	10.0	10.9	10.2	10.4	10.3	11.0	10.7	10.3	10.5	9.9	9.8	10.7	10.3	10.0
C-HAR	8.9	10.0	9.3	9.5	9.7	9.2	8.9	9.5	9.0	9.1	9.1	9.1	9.5	9.2
C-HAR-SV	8.8	9.3	8.8	9.3	8.9	8.8	8.8	8.7	8.9	8.9	8.8	8.6	9.0	8.8
C-HAR-CJ	8.6	8.9	8.8	8.8	8.9	8.1	8.9	8.6	8.8	9.0	8.5	8.1	8.7	8.9
C-HAR-SJ	8.6	9.1	9.1	8.8	9.0	8.8	9.4	8.8	9.0	9.0	8.6	9.1	8.9	8.9
CC-Bench	8.7	8.5	8.9	8.7	8.6	8.7	8.7	8.7	9.2	8.3	8.8	8.6	8.8	9.0
CC-FTS	8.7	8.4	9.3	8.7	8.5	8.3	8.7	8.6	9.1	8.3	8.7	8.4	8.7	9.2
CC-FTS-Top	8.3	8.6	8.7	8.7	8.7	8.8	8.7	8.6	8.8	8.5	8.7	8.6	8.8	8.7

Notes: The values for row HAR in Panel A correspond to the MSE (Panel A). Values from remaining row of Panel A and B are % changes in MSE against the HAR model. Negative percentage changes are improvements. In Panel C and D we report the average rank. USA-1 corresponds to the S&P 500 market index and USA-2 to the NASDAQ 100 index.

Table B.2: Five day-ahead forecasts under QLIKE

Model	AUS	IND	FRA	GBR	DEU	CAN	HKG	USA-1	KOR	JPN	USA-2	CHN	CHE	EUR
Panel A: Loss comparison with individual volatility models relative to the HAR model														
HAR [MSE]	0.15	0.24	0.18	0.17	0.17	0.21	0.17	0.23	0.16	0.16	0.21	0.22	0.19	0.16
GARCH-S_U	59.9%	-1.2%	18.2%	12.0%	10.9%	19.1%	11.4%	21.6%	12.0%	770.6%	32.9%	34.1%	10.7%	25.9%
FR-GARCH-S_U	35.4%	2.2%	33.5%	12.6%	10.6%	18.3%	17.7%	24.5%	30.7%	19.9%	27.8%	27.7%	10.0%	22.1%
AR	42.4%	22.1%	30.3%	30.8%	28.9%	36.7%	54.4%	19.6%	41.1%	22.1%	17.3%	37.4%	25.0%	26.1%
HARSV	1.9%	-8.2%	-1.3%	3.6%	-1.5%	7.7%	9.0%	-2.7%	1.6%	1.4%	6.0%	-3.5%	-4.2%	7.0%
HARCJ	8.5%	5.6%	-4.8%	-4.8%	3.4%	21.7%	11.5%	6.5%	0.1%	-1.7%	-0.3%	-6.4%	-9.1%	-2.6%
HARSJ	-1.4%	-9.8%	-0.4%	3.5%	-4.5%	28.0%	11.9%	3.1%	1.9%	-2.0%	1.2%	3.0%	-1.6%	27.2%
RV-FTS	10.3%	7.7%	12.1%	11.0%	11.5%	17.6%	5.2%	17.9%	8.7%	12.5%	16.7%	8.6%	12.9%	15.1%
Panel B: Loss comparison with combination forecast models relative to the HAR model														
C-GARCH-S_U	18.9%	-3.7%	5.8%	0.0%	0.7%	7.4%	-1.0%	10.0%	2.5%	13.7%	14.4%	10.2%	2.6%	9.8%
C-FR-GARCH-S_U	5.2%	-1.2%	8.9%	2.7%	1.6%	7.1%	1.7%	11.5%	10.7%	7.6%	10.0%	7.3%	4.2%	8.1%
C-AR	4.7%	3.2%	4.0%	3.1%	4.3%	8.5%	10.5%	3.6%	8.3%	2.4%	1.4%	7.5%	3.5%	4.1%
C-HAR	-0.7%	-2.7%	-0.6%	-1.2%	-1.4%	22.6%	-2.2%	1.5%	0.0%	1.1%	-0.1%	-1.2%	0.0%	0.4%
C-HAR-SV	-2.1%	-9.9%	-4.9%	-0.8%	-5.8%	-1.6%	-3.8%	-4.4%	0.7%	-1.3%	-0.9%	-5.2%	-5.6%	0.3%
C-HAR-CJ	-3.1%	-10.7%	-5.9%	-7.8%	-7.4%	-5.5%	-4.0%	-5.0%	-4.7%	-2.5%	-1.5%	-7.0%	2.8%	-3.0%
C-HAR-SJ	-6.6%	-12.8%	-4.9%	-8.8%	-7.7%	-3.1%	-3.7%	-4.6%	-4.4%	-3.4%	-2.1%	-3.9%	-7.0%	-2.4%
CC-Bench	-5.8%	-14.7%	-7.1%	-11.1%	-11.8%	-4.1%	-4.9%	-9.6%	25.6%	-7.9%	-4.5%	-7.9%	-11.3%	-2.1%
CC-FTS	-6.7%	-16.8%	-4.1%	-10.4%	-12.6%	-1.8%	-5.9%	-9.6%	6.7%	-8.0%	-4.6%	-8.4%	-13.2%	-2.6%
CC-FTS-Top	-5.9%	-15.3%	-8.6%	-9.7%	-11.3%	-9.0%	-6.0%	-10.2%	-6.6%	-6.7%	-6.8%	-7.2%	-11.1%	-5.7%
Panel C: Average ranking of the individual forecast models														
HAR	9.1	10.3	9.4	9.5	9.8	9.7	9.1	9.5	8.9	9.2	9.3	9.6	9.6	9.2
GARCH-S_U	13.7	10.0	10.8	10.7	10.7	11.1	10.4	11.8	9.8	12.4	12.3	12.4	10.8	11.6
FR-GARCH-S_U	8.9	8.3	9.4	9.3	9.4	9.7	9.0	9.5	10.0	9.1	9.7	8.8	9.5	9.2
AR	11.5	12.0	11.5	11.6	11.5	12.2	12.3	11.4	11.9	11.0	11.1	12.3	11.5	11.2
HAR-SV	9.1	9.7	9.1	9.5	9.0	9.3	9.8	9.0	9.6	9.5	9.2	9.3	9.4	9.0
HAR-CJ	8.9	9.2	9.3	9.2	9.1	8.8	9.8	8.9	9.4	9.5	8.9	8.8	9.2	9.3
HAR-SJ	9.1	9.6	9.6	9.4	9.3	9.6	10.5	9.2	9.9	9.7	9.4	10.0	9.4	9.4
RV-FTS	9.5	10.0	9.9	10.0	10.3	9.6	9.3	9.9	9.6	9.8	9.8	9.2	10.0	9.9
Panel D: Average ranking of the combination forecast models														
C-GARCH-S_U	11.8	9.6	9.9	9.9	9.9	10.0	9.5	10.6	9.3	11.0	10.9	10.9	9.9	10.5
C-FR-GARCH-S_U	8.6	8.6	9.0	9.0	9.3	9.2	8.6	9.3	9.5	8.7	9.3	8.5	9.2	9.0
C-AR	10.0	10.8	10.2	10.4	10.3	11.0	10.7	10.3	10.4	9.9	9.8	10.7	10.2	10.0
C-HAR	9.0	10.0	9.3	9.5	9.7	9.2	8.9	9.5	9.0	9.1	9.1	9.1	9.5	9.2
C-HAR-SV	8.9	9.2	8.8	9.3	8.9	8.8	8.8	8.7	8.9	9.0	8.8	8.6	9.0	8.8
C-HAR-CJ	8.6	8.9	8.9	8.8	8.9	8.1	8.9	8.6	8.8	9.0	8.5	8.1	8.7	8.9
C-HAR-SJ	8.6	9.1	9.1	8.8	9.0	8.8	9.4	8.8	9.0	9.1	8.6	9.1	8.9	8.9
CC-Bench	8.6	8.5	8.9	8.7	8.6	8.8	8.7	8.7	9.2	8.3	8.8	8.6	8.8	9.0
CC-FTS	8.7	8.4	9.2	8.7	8.5	8.3	8.7	8.6	9.0	8.3	8.7	8.4	8.7	9.2
CC-FTS-Top	8.4	8.6	8.7	8.7	8.7	8.8	8.7	8.6	8.8	8.5	8.7	8.6	8.7	8.6

Notes: HAR [QLIKE] row in Panel A reports the QLIKE values. Other rows in Panels A and B show percentage changes in MSE compared to the HAR benchmark. Panels C and D report average ranks. USA-1 is the S&P 500 and USA-2 is the NASDAQ 100 index.

Table B.3: Twenty-two day-ahead forecasts under MSE

Model	AUS	IND	FRA	GBR	DEU	CAN	HKG	USA-1	KOR	JPN	USA-2	CHN	CHE	EUR
Panel A: Loss comparison with individual volatility models relative to the HAR model														
HAR [MSE]	10.4	58.7	31.2	28.6	36.4	56.9	31.4	34.3	35.4	11.3	35.6	36.1	20.9	26.0
GARCH-S_U	-23.4%	-0.6%	-2.4%	-34.7%	-1.6%	-18.0%	-14.0%	32.5%	-23.2%	93.9%	6.2%	5.7%	-5.5%	2.7%
FR-GARCH-S_U	-27.2%	-11.0%	-8.1%	-39.4%	-10.8%	-25.5%	1.0%	-8.4%	-25.9%	-8.8%	-29.2%	-2.3%	-14.5%	-9.3%
AR	-21.3%	1.8%	3.5%	-30.6%	1.7%	-16.6%	15.5%	-3.5%	-9.3%	3.2%	-26.8%	16.0%	-0.2%	2.6%
HARSV	11.0%	-11.8%	-9.8%	20.7%	7.6%	11.9%	6.7%	-3.9%	5.0%	14.3%	-23.5%	-23.4%	21.3%	12.0%
HARCJ	-24.7%	3.9%	-9.8%	-19.5%	2.7%	20.5%	-0.7%	-2.2%	32.7%	15.8%	-26.1%	-17.5%	45.7%	4.8%
HARSJ	-19.9%	-13.8%	-4.8%	-20.2%	9.0%	21.5%	21.0%	-6.4%	11.7%	14.4%	-19.0%	-20.0%	33.5%	9.7%
RV-FTS	-3.2%	15.1%	13.9%	-20.9%	20.0%	2.4%	-4.5%	32.1%	-0.8%	20.4%	4.9%	-0.5%	21.8%	16.8%
Panel B: Loss comparison with combination forecast models relative to the HAR model														
C-GARCH-S_U	-20.3%	1.7%	-2.8%	-33.7%	0.4%	-14.3%	-14.6%	26.8%	-17.2%	39.6%	3.0%	-5.2%	1.9%	-0.4%
C-FR-GARCH-S_U	-24.6%	-4.7%	-6.9%	-37.2%	-8.5%	-21.9%	-7.7%	0.9%	-22.2%	-4.4%	-22.7%	-11.2%	-5.8%	-6.1%
C-AR	-22.5%	-4.7%	-5.9%	-34.5%	-2.1%	-16.3%	-10.9%	-0.4%	-17.8%	-3.6%	-20.3%	-10.4%	-4.1%	-3.3%
C-HAR	-7.2%	1.9%	0.6%	-23.1%	3.3%	-2.6%	-7.6%	8.9%	-2.7%	1.2%	-13.0%	-3.3%	2.0%	-0.3%
C-HAR-SV	-21.4%	-11.1%	-11.2%	-18.8%	0.4%	8.2%	0.6%	-1.0%	-2.9%	8.1%	-17.6%	-22.1%	12.1%	-0.5%
C-HAR-CJ	-25.8%	4.5%	-10.3%	-25.8%	2.1%	5.4%	-4.7%	2.7%	10.7%	7.1%	-20.5%	-15.2%	17.0%	-2.1%
C-HAR-SJ	-26.5%	-10.9%	-6.3%	-27.0%	1.9%	11.2%	5.5%	-1.5%	0.3%	8.7%	-13.7%	-19.1%	18.7%	1.0%
CC-Bench	-37.9%	-17.3%	-19.4%	-45.3%	-17.8%	-29.0%	-23.2%	-12.9%	-32.1%	-20.8%	-39.2%	-29.0%	-18.1%	-20.4%
CC-FTS	-38.3%	-16.9%	-19.4%	-44.9%	-18.2%	-28.8%	-22.0%	-13.0%	-32.4%	-20.2%	-38.9%	-28.6%	-18.0%	-20.6%
CC-FTS-Top	-31.4%	-16.3%	-16.2%	-37.8%	-11.6%	-18.1%	-18.3%	-11.2%	-25.5%	-12.4%	-33.7%	-23.5%	-15.1%	-12.9%
Panel C: Average ranking of the individual forecast models														
HAR	9.9	11.3	10.0	10.6	10.3	10.9	9.6	10.4	9.5	9.6	9.8	10.6	10.4	9.8
GARCH-S_U	13.6	9.4	10.5	10.4	10.1	10.5	9.9	11.5	9.4	13.2	12.2	12.1	10.1	11.3
FR-GARCH-S_U	8.1	7.8	8.6	8.1	8.8	8.5	8.6	8.9	9.7	8.9	8.5	8.9	8.4	8.4
AR	11.8	12.3	11.7	12.2	11.7	12.7	12.5	12.1	12.0	10.7	11.5	12.8	11.9	11.4
HAR-SV	10.0	10.2	10.0	10.5	9.5	10.1	10.2	9.5	10.5	9.6	9.8	9.6	10.1	10.0
HAR-CJ	9.8	10.3	9.7	10.6	9.8	10.0	10.5	9.5	10.2	10.1	10.4	9.5	10.1	9.6
HAR-SJ	9.8	10.4	10.4	10.1	9.9	10.0	11.1	9.7	10.5	9.9	10.3	9.7	9.9	9.8
RV-FTS	8.8	9.2	9.6	9.4	10.1	9.1	9.4	9.5	9.7	9.5	9.2	9.0	9.7	10.1
Panel D: Average ranking of the combination forecast models														
C-GARCH-S_U	11.6	8.9	9.7	9.1	9.6	9.1	9.1	10.3	8.8	11.5	10.9	10.4	9.5	10.5
C-FR-GARCH-S_U	7.6	7.9	8.4	8.1	8.7	7.9	8.4	8.5	9.0	8.4	8.0	8.4	8.4	8.4
C-AR	9.9	10.9	10.2	10.3	10.3	11.0	10.6	10.5	10.1	9.3	9.9	10.7	10.3	10.2
C-HAR	8.8	10.0	9.2	9.3	9.6	9.3	8.9	9.5	8.9	8.8	8.9	9.3	9.4	9.4
C-HAR-SV	8.9	9.0	9.2	9.3	9.1	8.7	9.0	8.8	8.9	8.8	8.7	8.5	9.1	9.4
C-HAR-CJ	8.6	9.1	8.9	9.0	9.3	8.8	9.3	8.7	9.0	9.2	9.1	8.3	9.2	9.1
C-HAR-SJ	8.6	9.2	9.4	8.9	9.1	8.7	9.6	8.8	9.0	8.9	8.9	8.6	9.1	9.2
CC-Bench	8.4	8.2	8.1	8.3	8.2	8.6	8.1	8.1	8.8	8.1	8.2	8.2	8.3	8.0
CC-FTS	8.2	8.1	8.3	8.1	8.3	8.0	8.0	8.2	8.3	8.0	8.2	8.0	8.3	8.0
CC-FTS-Top	8.6	8.6	9.0	8.7	8.6	9.1	8.5	8.5	8.7	8.5	8.6	8.4	8.9	8.6

Notes: The values for row HAR in Panel A correspond to the MSE (Panel A). Values from remaining row of Panel A and B are % changes in MSE against the HAR model. Negative percentage changes are improvements. In Panel C and D we report the average rank. USA-1 corresponds to the S&P 500 market index and USA-2 to the NASDAQ 100 index.

Table B.4: Twenty-two day-ahead forecasts under QLIKE

Model	AUS	IND	FRA	GBR	DEU	CAN	HKG	USA-1	KOR	JPN	USA-2	CHN	CHE	EUR
Panel A: Loss comparison with individual volatility models relative to the HAR model														
HAR [QLIKE]	0.28	0.33	0.28	0.27	0.28	0.46	0.18	0.30	0.20	0.16	0.48	0.24	0.32	0.27
GARCH-S_U	3.2	-9.8	-6.4	-2.1	-12.9	2.1	-7.9	1.8	-3.1	658.7	-27.9	15.2	-11.0	-10.3
FR-GARCH-S_U	20.7	2.5	37.4	-0.1	2.7	10.2	23.0	25.5	26.5	27.4	-17.5	35.3	-4.4	8.3
AR	31.6	15.5	20.9	24.5	18.3	23.4	53.5	15.3	34.4	16.2	-21.0	41.1	12.5	15.1
HARSV	25.2	-4.2	2.3	4.3	6.0	11.1	4.5	6.6	436.2	19.2	-20.1	-6.5	-11.3	1.2
HARCJ	16.8	-2.6	-6.8	14.3	-4.0	7.8	7.1	21.1	5.2	4.7	-18.9	-4.8	-10.1	-2.8
HARSJ	5.8	-7.0	3.1	-0.5	2.8	7.3	9.5	18.5	6.6	28.3	-19.9	-6.8	-8.1	21.0
RV-FTS	7.6	-3.6	2.3	-2.1	8.3	13.8	3.2	10.6	4.6	11.5	-16.6	3.3	5.8	9.1
Panel B: Loss comparison with combination forecast models relative to the HAR model														
C-GARCH-S_U	-9.4	-11.6	-11.5	-11.1	-9.8	0.5	-9.4	-1.7	-6.3	8.5	-30.6	0.0	-9.9	-10.6
C-FR-GARCH-S_U	6.0	-6.5	6.8	-8.8	-2.0	4.7	0.8	8.4	5.2	6.3	-22.7	5.3	-4.6	1.3
C-AR	5.3	-2.9	-1.6	-4.1	1.9	5.6	4.7	-0.7	2.8	-2.5	-28.7	2.8	-1.5	-0.6
C-HAR	-3.1	-7.3	-6.3	-9.3	-2.9	-1.3	-5.6	-2.5	-5.6	-3.4	-29.0	-6.2	-5.7	-4.3
C-HAR-SV	-2.5	-13.3	-8.4	-7.3	-2.0	4.3	-5.5	-2.7	-5.4	-2.1	-25.5	-10.8	-12.8	-4.4
C-HAR-CJ	0.3	-9.4	-12.1	-11.1	-5.7	1.6	-4.3	-1.9	-5.8	-3.4	-26.4	-10.5	-9.6	-6.4
C-HAR-SJ	-3.2	-13.0	-8.0	-11.5	-4.1	1.7	-4.8	-2.3	-5.5	-3.7	-25.1	-11.8	-11.8	-5.9
CC-Bench	-24.1	-19.7	-16.6	-14.6	-13.3	-7.4	-15.7	-10.0	-12.3	-13.9	-36.6	-14.7	-20.2	-23.7
CC-FTS	-25.0	-20.1	-16.5	-18.2	-14.7	-9.0	-16.6	-8.9	-15.0	-15.5	-35.5	-15.4	-10.6	-22.6
CC-FTS-Top	-10.7	-15.6	-12.7	-13.0	-12.2	-4.7	-10.6	-3.2	-9.4	-8.6	-33.6	-13.1	-17.8	-13.2
Panel C: Average ranking of the individual forecast models														
HAR	9.9	11.3	10.0	10.6	10.2	10.8	9.5	10.4	9.5	9.6	9.8	10.6	10.4	9.8
GARCH-S_U	13.4	9.4	10.5	10.3	10.0	10.4	9.8	11.3	9.4	13.0	12.1	12.0	10.1	11.2
FR-GARCH-S_U	8.1	8.0	8.7	8.3	8.9	8.6	8.8	9.1	9.9	9.2	8.6	9.2	8.5	8.5
AR	11.9	12.3	11.7	12.3	11.8	12.7	12.5	12.1	12.0	10.7	11.5	12.7	11.9	11.4
HAR-SV	10.1	10.3	10.0	10.5	9.5	10.1	10.1	9.5	10.5	9.6	9.8	9.6	10.0	10.0
HAR-CJ	9.9	10.3	9.7	10.5	9.8	10.0	10.4	9.5	10.1	10.1	10.3	9.5	10.0	9.6
HAR-SJ	9.9	10.4	10.4	10.1	9.9	9.9	11.1	9.7	10.4	9.8	10.3	9.7	9.9	9.7
RV-FTS	8.8	9.3	9.7	9.5	10.2	9.2	9.5	9.6	9.8	9.6	9.3	9.0	9.8	10.1
Panel D: Average ranking of the combination forecast models														
C-GARCH-S_U	11.4	8.9	9.7	9.1	9.6	9.1	9.1	10.2	8.9	11.4	10.8	10.3	9.4	10.5
C-FR-GARCH-S_U	7.7	7.9	8.5	8.2	8.8	8.0	8.5	8.5	9.1	8.5	8.0	8.5	8.5	8.5
C-AR	9.9	10.8	10.1	10.3	10.2	11.0	10.5	10.4	10.1	9.3	9.9	10.7	10.2	10.1
C-HAR	8.8	9.9	9.2	9.3	9.6	9.3	8.8	9.5	8.9	8.8	8.9	9.2	9.3	9.3
C-HAR-SV	8.9	9.0	9.1	9.2	9.0	8.7	8.9	8.8	8.9	8.8	8.7	8.5	9.1	9.4
C-HAR-CJ	8.6	9.1	8.9	9.0	9.3	8.8	9.2	8.7	9.0	9.2	9.0	8.3	9.1	9.1
C-HAR-SJ	8.6	9.1	9.4	8.8	9.1	8.6	9.5	8.8	8.9	8.8	8.9	8.6	9.1	9.1
CC-Bench	8.4	8.3	8.1	8.3	8.3	8.6	8.2	8.2	8.8	8.1	8.3	8.3	8.4	8.1
CC-FTS	8.2	8.2	8.3	8.1	8.3	8.0	8.1	8.2	8.4	8.1	8.2	8.0	8.4	8.1
CC-FTS-Top	8.6	8.6	9.0	8.7	8.6	9.1	8.4	8.5	8.7	8.5	8.5	8.4	8.9	8.6

Notes: The values for row HAR in Panel A correspond to the QLIKE (Panel A). Values from remaining row of Panel A and B are % changes in QLIKE against the HAR model. Negative percentage changes are improvements. In Panel C and D we report the average rank. USA-1 corresponds to the S&P 500 market index and USA-2 to the NASDAQ 100 index.