

A DYNAMIC STATE-SPACE HAR MODEL

MIKE TSIONAS¹, AYA GHALAYINI², MARWAN IZZELDIN¹, AND LORENZO TRAPANI³

ABSTRACT. The Heterogeneous AutoRegressive model for the logs of Realised Volatility (HARL) has established itself as the benchmark specification for modelling and forecasting return volatility, owing to its parsimony and ability to capture the strong persistence typically observed in RV. To address potential concerns such as measurement errors, nonlinearities, and non-spherical residuals, numerous variants of the baseline HARL model have been developed in the literature. This paper contributes to this body of work by proposing a new class of dynamic state-space models with time-varying parameters. The parameter dynamics are assumed to follow an autoregressive process, with or without stochastic volatility, giving rise to two specifications: SHARP and SHARP-SV. Both models are designed to capture the evolving nature of return volatility and are estimated via Bayesian inference using Particle Gibbs sampling. Empirical applications to high-frequency data on SPY, sector ETFs, representative NYSE stocks, and the VIX index demonstrate that our proposed models on average outperform alternative HARL-based specifications in forecasting volatility, particularly at medium- and long-term horizons. An extensive Monte Carlo analysis further illustrates the advantages of our approach in terms of both estimation accuracy and predictive performance.

1. INTRODUCTION

The presence of time-varying parameters and time-varying volatility in economic and financial data (see, *inter alia*, [Baccheri and Corsi, 2021](#), and their literature review) is a very

¹LANCASTER UNIVERSITY MANAGEMENT SCHOOL, LANCASTER, UK

²COLLEGE OF BUSINESS AND SOCIAL SCIENCES, ASTON UNIVERSITY, BIRMINGHAM, UK

³UNIVERSITA' DEGLI STUDI DI PAVIA, PAVIA, ITALY, AND UNIVERSITY OF LEICESTER, LEICESTER, UK
E-mail addresses: m.tsionas@lancaster.ac.uk, a.ghalayini@aston.ac.uk,
m.izzeldin@lancaster.ac.uk, lorenzo.trapani@unipv.it.

Key words and phrases. Time-varying coefficients, HAR models, Particle Gibbs sampling, State Space models, volatility forecasting.

Acknowledgement. This paper was started (and virtually completed, although in a different form than the current one) when Mike was still alive; he proposed the new SHARP and SHARP-SV models, and developed the Bayesian, Particle Gibbs based, inferential procedure. We are grateful to the Editor, Michael Jansson, to the Guest Editor, Subal Kumbhakar, and to two anonymous Referees for very helpful and constructive feedback.

well-documented stylised fact. Models that fail to account for changes in the data generating process are liable to produce biased estimates and inaccurate predictions. Thus, it is essential to analyse and account for parameter behaviour in the statistical modelling of such data. The importance of taking such features into account has been emphasised in several studies, including those by [Koop and Potter \(2004\)](#), [D’Agostino et al. \(2013\)](#), [Clark and Ravazzolo \(2015\)](#), [Bekierman and Manner \(2018\)](#), [Chen et al. \(2018\)](#), and [Buccheri and Corsi \(2021\)](#). As far as time variation is concerned, the use of Time Varying Parameters (TVP) models has been long proposed as a more flexible (and, possibly, more realistic) alternative to models with abrupt breaks. Several applications have shown the superior forecasting ability of TVP models, also in the context of high-dimensional Vector AutoRegressive models; whilst a comprehensive literature review goes well beyond the scope of this paper, we refer to the contributions by [Doan et al. \(1984\)](#), [Sims \(1993\)](#), [Stock and Watson \(1996\)](#), and [Cogley and Sargent \(2005\)](#), as seminal papers, and to the more recent works by [Carriero et al. \(2019\)](#) and [Tsionas et al. \(2022\)](#), also for an up-to-date list of references. Alongside this strand of the literature, modelling time variation in the volatility is also a very important topic, e.g. in finance (owing to its significance in risk management, portfolio selection, and asset pricing), and also in macroeconomic forecasting (see, for example, the papers by [Clark, 2011](#); [Koop and Korobilis, 2013](#); [Carriero et al., 2015](#); [Clark and Ravazzolo, 2015](#); and [Koop et al., 2019](#) - where accounting for heteroskedasticity is shown to yield dramatic improvements in forecasting). In particular, in the context of financial econometrics, nonparametric approaches based on the so-called Realised Volatility (RV) are often preferred to parametric approaches, based e.g. on GARCH-type models ([Bollerslev, 1986](#)).

Since the seminal contributions by [Andersen and Bollerslev \(1998\)](#), [Barndorff-Nielsen and Shephard \(2002\)](#), and [Liu et al. \(2015\)](#), modelling realised volatility has advanced markedly, with the Heterogeneous Autoregressive (HAR) model by [Corsi \(2009\)](#) becoming the empirical workhorse. By employing daily, weekly, and monthly RV averages, HAR parsimoniously

captures persistence and heterogeneity across horizons. [Corsi \(2009\)](#) notes that modelling the logs of RV, as well as ensuring non-negativity, also improves forecast accuracy. We take this HAR-log (HARL) specification as the cornerstone of our contribution. Despite its success, the (linear) HARL can suffer from measurement error, nonlinearities, and autocorrelated, heteroskedastic residuals, motivating extensions such as regime-switching ([McAleer and Medeiros, 2008](#)), dynamic model averaging ([Wang et al., 2016](#)), and TVP variants ([Bekierman and Manner, 2018](#); [Chen et al., 2018](#); and [Buccheri and Corsi, 2021](#)).

We address the aforementioned limitations of the HARL framework by building on (and extending) the approaches mentioned above, accounting for nonlinearities, time variations in the parameters (and possible heteroskedasticity in the law of motion of the coefficients), and autocorrelation and heteroskedasticity in the residuals. Specifically, we present a novel dynamic state-space model with time-varying coefficients, assumed to follow an AutoRegressive (AR) process with or without Stochastic Volatility (SV) - henceforth, we refer to the former specification (that is, a model with TVP but no SV in the law of motion of the coefficients) as the SHARP model; and to the latter (i.e., a model with TVP and SV in the law of motion of the coefficients) as its SHARP-SV variant. An integral part of our methodology is the use, in both cases, of Bayesian inference with Particle Gibbs sampling, following the procedure by [Creal and Tsay \(2015\)](#), which allows for efficient computation of the latent variables (see also [Andrieu et al., 2010](#)). However, even though our main focus - as far as inference is concerned - is Bayesian, we also study in depth the dependence structure of our model. Indeed, the SHARP and SHARP-SV models produce observations - owing to the presence of time-varying parameters and stochastic volatility - whose dependence structure is highly complex. Still, we show that our proposed models generate observations which belong in a wide class of *weakly* dependent processes, which do not exhibit persistence or pseudo-long-memory behaviour.¹

¹In order not to overshadow the main results in this paper, we relegate these results in Section [E](#) in the Supplement.

Empirically, we embed these new features into the HARL framework to improve RV forecasts. Our results show that this approach outperforms standard methods in predicting the volatility of both index and individual stock returns. In an extensive empirical analysis, we examine daily RV from the SPY-ETF—a tradable U.S. market index—alongside 10 sector ETFs (2006–2023), 20 representative NYSE stocks (2000–2016), and the VIX index (2003–2023). Across all datasets, both the SHARP and SHARP-SV models significantly outperform the HARL model and its time-varying extensions, as confirmed by the predictive ability test of [Giacomini and White \(2006\)](#). While SHARP-SV provides a moderate forecasting edge over SHARP, both consistently belong in the Model Confidence Set of [Hansen et al. \(2011\)](#). Results are broadly confirmed via a Monte Carlo simulation, where we study the performance of our estimation methodology, and the forecasting accuracy of our proposed models (and of Bayesian estimation *cum* Particle Gibbs) under possible misspecification, using several Data Generating Processes (DGPs henceforth) exhibiting the typical features of financial data.

The remainder of the paper is organised as follows. Section 2 provides an overview of the Realised Variance measure and existing extensions to the HARL model with dynamic and time-varying parameters. In Section 3, we introduce our proposed models and provide details on the estimation method. In Section 4, we apply and evaluate the proposed models in forecasting stock volatility using real data. A comprehensive Monte Carlo study is reported in Section 5, where we assess the performance of our Bayesian estimator under correct specification (Section 5.1), and predictive accuracy under possible mis-specification (Section 5.2). Section 6 concludes. Further results are in the Supplement where we: report the detailed forecasting results and some misspecification analysis (Section A), and further Monte Carlo evidence (Section B); present the posterior derivations (Section C); provide details on the Particle Filtering algorithm (Section D); study the dependence structure of the SHARP and SHARP-SV models (Section E), and present technical lemmas (Section F) and proofs (Section G).

2. VOLATILITY MEASURE AND HARL FAMILY OF MODELS

Consider an asset whose log-price $\ln(P_s)$ follows the stochastic differential equation:

$$d\ln(P_s) = \mu_s ds + \sigma_s dW_s, \quad (2.1)$$

where μ_s denotes the drift, σ_s is the instantaneous volatility and W_t a standard Wiener process. The (latent) integrated variance for day t is defined as:

$$IV_t = \int_{t-1}^t \sigma_s^2 ds, \quad (2.2)$$

and its nonparametric, ex-post estimate based on the realised variance, RV_t , is calculated by aggregating intra-daily squared returns over a one-day horizon, t , using M sub-intervals:

$$RV_t = \sum_{j=1}^M r_{j,t}^2, \quad (2.3)$$

where $r_{j,t} = \ln(P_{(t-1)M+j}) - \ln(P_{(t-1)M+(j-1)})$ is the intra-day return of the j -th sub-interval within the t -th day, and $P_{(t-1)M+j}$ is the asset price at the start of the j -th interval computed as the average of the closing and opening prices of intervals $j - 1$ and j , respectively.² Several models can be used in order to forecast the daily RV; our benchmark model is the Heterogeneous Autoregressive (log) Realised Variance (HARL) of [Corsi \(2009\)](#) defined as:

$$RV_t^l = \beta_1 + \beta_2 RV_{t-1}^l + \beta_3 RV_{w,t-1}^l + \beta_4 RV_{m,t-1}^l + \nu_t, \quad \text{with } \nu_t \sim N(0, \sigma_v^2), \quad (2.4)$$

where RV_t^l denotes the log-transformation of daily RV_t , and $RV_{w,t}^l$ and $RV_{m,t}^l$ denote the weekly and monthly log-transformations of the RV realised at time t , respectively, computed

²We point out that [Zhang et al. \(2005\)](#) provide a discussion on optimising the sampling frequency for the estimation of RV_t . However, [Buccheri and Corsi \(2021\)](#) show that the relative forecast performance of models with time-varying coefficients, such as the ones discussed in this paper, is independent of the sampling frequency. Therefore, for conciseness, we use sub-intervals of length 300 seconds in constructing the daily RV series. The latter defines 78 intraday sub-intervals and combines balanced information from high-frequency data and microstructure effects [Andersen et al. \(2001\)](#).

over a recursive rolling window of fixed length (week or month) as $RV_{w,t}^l = \sum_{i=1}^5 RV_{t-i}^l / 5$ and $RV_{m,t}^l = \sum_{i=1}^{22} RV_{t-i}^l / 22$ respectively. As common wisdom would suggest, by using the log-transformation, the HARL model of equation (2.4), and indeed all the other variants estimated on the log series, are less affected by the huge peaks of RV. Nevertheless, the HARL model is liable to suffer, albeit less severely, from several issues such as measurement error, time variation in the parameters, etc... Hence, the literature has developed several extensions of the basic HARL, which we also use in our paper by way of comparison and which we briefly review here. [Bollerslev et al. \(2016\)](#) and [Bollerslev et al. \(2018\)](#) propose an extension (called the HARQ model) in order to deal with the estimation error of the RV, using the “realised quarticity” defined as $RQ_t = \frac{M}{3} \sum_{j=1}^M r_{j,t}^4$. This approach has the major advantage of being estimable using OLS, and it can be naturally extended to the HARL set-up (resulting in a model which we call the HARLQ specification). Further, again in order to address the measurement error, [Bekierman and Manner \(2018\)](#) - building on the original works by [Barndorff-Nielsen and Shephard, 2002](#) and [Bollerslev et al., 2016](#) - propose the HARSL model, which, in essence, is a state-space HARL model which assumes a latent Gaussian AR(1) process for the daily coefficient β_2 . The model is estimated using maximum likelihood with a standard Kalman filter, and it on average outperforms the HARL model in forecasting the RV. The success of the HARSL model is due to its ability to capture other sources of temporal variation in addition to the variance of the measurement error. However, [Bekierman and Manner \(2018\)](#) note that the maximum likelihood estimator of the model is inefficient, and allowing all coefficients to follow an AR process is computationally challenging to perform using their employed estimation method. Thus, in a related contribution, [Chen et al. \(2018\)](#) introduce a novel approach - based on a local linear smoothing method - to estimate the HARL model (called the TVCHAR model), which assumes time-varying coefficients of an unknown functional form. The TVCHAR specification is found to also outperform the benchmark HARL model, especially over longer forecasting horizons. In a similar vein, [Buccheri and](#)

[Corsi \(2021\)](#) propose the SHARK model, which accounts for time-varying coefficients and heteroskedastic error terms while also handling measurement errors. They find that their model produces moderate improvements in one-day-ahead forecasts but is more effective for longer-term forecasting. In [Table 2.1](#), we summarise the set of HARL versions that we consider in this paper to compare alongside our proposed models, described in the next section.

TABLE 2.1. Summary of competing models

Model Name	Author(s)	Model	Estimation Method
HARL (benchmark model)	Corsi (2009)	$RV_t^l = \beta_1 + \beta_2 RV_{t-1}^l + \beta_3 RV_{w,t-1}^l + \beta_4 RV_{m,t-1}^l + \nu_t; \quad \nu_t \sim N(0, \sigma_v^2)$	OLS
HARLQ	inspired by Bollerslev et al. (2016) & Buccheri and Corsi (2021)	$RV_t^l = \beta_1 + \beta_{2,t} RV_{t-1}^l + \beta_3 RV_{w,t-1}^l + \beta_4 RV_{m,t-1}^l + \nu_t; \quad \nu_t \sim N(0, \sigma_v^2)$ where $\beta_{2,t} = \beta_2 + \gamma \frac{\sqrt{RQ}_{t-1}}{RV_{t-1}^l}$	OLS
HARSL	Bekierman and Manner (2018)	$RV_t^l = \beta_1 + (\beta_2 + \lambda_t) RV_{t-1}^l + \beta_3 RV_{w,t-1}^l + \beta_4 RV_{m,t-1}^l + \nu_t; \quad \nu_t \sim N(0, \sigma_v^2)$ where $\lambda_t = \phi \lambda_{t-1} + \eta_t; \quad \eta_t \sim N(0, \sigma_\eta^2)$	Maximum Likelihood using Kalman filter
TVCHAR	Chen et al. (2018)	$RV_t^l = \beta_1(\tau_t) + \beta_2(\tau_t) RV_{t-1}^l + \beta_3(\tau_t) RV_{w,t-1}^l + \beta_4(\tau_t) RV_{m,t-1}^l + \nu_t; \quad \nu_t \sim N(0, \sigma_v^2)$ Using first-order Taylor expansion: $\beta_j(\tau_t) \approx \beta_j(\tau) + \beta_j'(\tau)(\tau_t - \tau)$ where $\beta_j'(\tau)$ is the first-order derivative of $\beta_j(\tau)$	Local Linear Method
∞	SHARK	$RV_t^l = \alpha_t + \nu_t; \quad \nu_t \sim N(0, h_t)$	Maximum Likelihood using Kalman filter
		$\alpha_t = \beta_{1,t} + \beta_{2,t} RV_{t-1}^l + \beta_{3,t} RV_{w,t-1}^l + \beta_{4,t} RV_{m,t-1}^l + \eta_t; \quad \eta_t \sim N(0, q_t)$	
		Let $f_t = (\beta_{1,t}, \beta_{2,t}, \beta_{3,t}, \beta_{4,t}, \log q_t)'$	
		The update rule is: $f_{t+1} = f_t + C s_t$ where s_t is a function of the Kalman filter prediction error and its covariance matrix.	
		$V_t = \frac{\sum_{j=1}^M r_{j,t}^4}{(\sum_{j=1}^M r_{j,t}^2)^2}$, a consistent estimate of the measurement error variance of RV^l , is selected as a proxy for h_t .	

3. METHODOLOGY

As the (short) literature review in the previous section demonstrates, research within the HARL family of models has been very active. In this section, we: propose two variants of the benchmark HARL model of equation (2.4), which account for the possible mis-specifications in the HARL set-up by considering a quite general specification for the law of motion of the time-varying coefficients (Section 3.1); briefly discuss (Bayesian) estimation (Section 3.2) and the priors we have chosen for our empirical exercise; and report the forecasting algorithm using both variants of the basic HARL model (Section 3.3).

3.1. The SHARP and SHARP-SV models. We describe the two variations of the proposed model. In order for the presentation not to be overshadowed by model complexity, here we present a simplified version of our model, where innovations are *i.i.d.* and Gaussian. However, the measurement equation (3.1) below is quite flexible, and the model can be extended to include autoregressive structure, deterministics such as a constant, (linear or non-linear) trends, or seasonal dummies. Also, in principle, other specifications than equation (3.5) below for time-varying heteroskedasticity are possible. For extensions, we refer to Section 5, where we show that using Bayesian estimation with the Particle Gibbs sampler within this simplified class of dynamic state-space models is sufficient to capture several features of financial time series, such as: correlation between innovations, presence of diffusive leverage and multiple regimes, and heavy tails. Further, as mentioned in the Introduction, in Section E in the Supplement, we derive a characterisation of the dependence structure of the observations y_t under a more complex version of the model below, which also nests (3.1)-(3.5). Consider the following state-space specification,

$$y_t = x_t' \beta_t + \nu_t, \tag{3.1}$$

where $1 \leq t \leq n$, and $\nu_t \sim i.i.d.\mathcal{N}(0, \sigma_v^2)$; we model the law of motion of β_t as

$$\beta_{j,t} = \alpha_j + \rho_j \beta_{j,t-1} + \varepsilon_{j,t}, \quad (3.2)$$

where $|\rho_j| < 1$ for $1 \leq j \leq k$. Whilst spelt out for a general model in terms of $\{y_t, x'_t\}$, in our context the measurement equation (3.1) describes the dependent variable $y_t = RV_t^l$ in terms of the covariates $x_t = (1, RV_{t-1}^l, RV_{w,t-1}^l, RV_{m,t-1}^l)'$ in (2.4), of the state vector, β_t , and of the disturbances, ν_t . The transition equation (3.2) describes the evolution of the coefficients over time, with the (stationary) AR process of the intercept also capturing the residuals autocorrelation in the measurement equation.

In the state equations (3.2), we use two different specifications, which characterise the SHARP and the SHARP-SV model respectively; in the former case (SHARP), we model the innovations as *i.i.d.* Gaussian, viz.

$$\varepsilon_{j,t} \sim i.i.d.\mathcal{N}(0, \sigma_{\varepsilon,j}^2), \quad (3.3)$$

independent across $1 \leq j \leq k$, whereas in the latter case (SHARP-SV), we assume a stochastic volatility (SV), equation-by-equation, process, i.e.

$$\varepsilon_{j,t} = h_{j,t}^{1/2} \eta_{j,t}, \quad (3.4)$$

$$\ln h_{j,t} = \gamma_j + \delta_j \ln h_{j,t-1} + u_{j,t}, \quad (3.5)$$

where $\{\eta_{j,t}, 1 \leq j \leq k, 1 \leq t \leq n\}$ is a mean zero, unit variance process independent across j and t , and independent of $\{u_{j,t}, 1 \leq j \leq k, 1 \leq t \leq n\}$, with $u_{j,t} \sim i.i.d.\mathcal{N}(0, \sigma_{u,j}^2)$ also independent across j , and $|\delta_j| < 1$. While the unconditional first and second moments of the stationary solution of (3.2) can be shown to be constant, its conditional second moment can change over time; hence, (3.4)-(3.5) allows for (a) conditional heteroskedasticity in the

main equation (3.1), through the intercept term (say $\beta_{1,t}$); and (b) also for conditional heteroskedasticity in the coefficients process. Model (3.4)-(3.5) is a standard (univariate) stochastic volatility model (see Taylor, 1982); we refer to Harvey et al. (1994) for extensions to multivariate settings, and to Kokoszka et al. (2025) for a functional version.

3.2. (Bayesian) estimation of SHARP and SHARP-SV models. We review the estimation technique employed, and the the priors used in our empirical exercise.

In order to estimate the model, we use a modified version of the sequential Monte Carlo method known as the Particle Gibbs (PG) sampler discussed in Andrieu et al. (2010). Whilst the details are in Section C of the Supplement, here we offer a bird's-eye view on the estimation algorithm. The latent variables in our model are $\lambda_t = (\beta'_t, h'_t)'$ where $\beta_t = (\beta_{1,t}, \dots, \beta_{k,t})'$ and $h_t = (h_{1,t}, \dots, h_{k,t})'$, whose prior can be described by $p(\lambda_t | \lambda_{t-1}, \theta)$, with θ a vector containing all the static parameters; the joint posterior is denoted as $p(\theta, \lambda_{1:T} | y_{1:T})$ and also studied in Section C of the Supplement; finally, in the PG sampler, we draw the structural parameters, as usual, from their posterior conditional distributions $p(\theta | \lambda_{1:T}, y_{1:T})$, reported again in Section C of the Supplement. By doing so, we can avoid mixture approximations or other Monte Carlo procedures that need considerable tuning and may not have good convergence properties. The latent variables can be integrated out of the joint posterior using the procedure by Creal and Tsay (2015), described in Section D of the Supplement. Our choice of conjugate priors is as follows:

$$\alpha_j, \gamma_j \sim \mathcal{N}(0, 1); \quad \rho_j, \delta_j \sim \mathcal{N}(0.5, 1) \mathbb{I}_{\rho_j \in (0,1)} \mathbb{I}_{\delta_j \in (0,1)}; \quad \sigma_{\varepsilon,j}^2 \sim \Gamma(6.5, 0.5); \quad \sigma_v^2 \sim \Gamma(6.5, 0.5),$$

for all $1 \leq j \leq k$, where $\Gamma(a, b)$ denotes the Gamma distribution with shape a and scale b . In the above set of priors: we restrict the estimation range of ρ and δ between $(0, 1)$, to

guarantee stationarity and also to reflect the belief that coefficients are positively autocorrelated; we note that the prior of γ and δ can be made more flexible in this exercise with a prior mean of 0; we choose the conjugate prior specification of the variances, $\sigma_{u,j}^2$ and σ_v^2 , to have a low mean, reflecting the prior belief that while the coefficients are time-varying, we do not expect high jumps in the magnitude of the coefficients from one day to another.³

3.3. Forecasting the log-RV using SHARP and SHARP-SV. To reduce the computational cost of the model, we implement an efficient strategy in our MCMC and particle filtering process. Instead of running the full MCMC and particle filtering for each observation, we perform these computations every 10 observations. During these runs, we use 100 particles for the particle filtering, and the estimated posterior statistics are then applied to the following 10 observations. Within these 10 observations, we estimate the states by running the particle filtering using 1,000 particles, so as to balance computational efficiency and model accuracy. We use an in-sample estimation window of approximately four years, and conduct 1,000 MCMC iterations, with the first 300 iterations discarded as burn-in. By only updating the MCMC estimates every 10 observations, and using the resulting estimates for the subsequent 10, we significantly reduce computational demands while maintaining forecasting capabilities.

We report below the algorithm for the forecasting of RV_t^l using both the SHARP and SHARP-SV models. Recall that n is the total number of observations, and let: is be the in-sample estimation window size (approximately four years), $nsim$ be the number of MCMC iterations, and $nburn$ the burn-in sample.

(1) For $T = is, is + 10, \dots, n - 1$

(a) For $i = 1, \dots, nsim$

³We fixed the hyperparameters primarily to allow for more straightforward implementation of the Gibbs sampler and reduce the computational burden. In a set of unreported experiments, we tried varying the hyperparameters, but we noted no material differences in the empirical findings, which seems to suggest that our results are not sensitive to the specific choice of hyperparameters.

- (i) Draw λ_t for $t = is - T + 1, \dots, T$, as illustrated in Section D of the Supplement.
 - (ii) Sample the parameters of equations (3.2) and (3.5) using their posterior distributions (see Section C of the Supplement).
 - (iii) Forecast the states in equation (3.2) and (3.5) then forecast $\widehat{RV}_{T+1,i}^l$ for $i = nburn + 1, \dots, nsim$ accordingly.
- (b) Estimate $\widehat{RV}_{T+1}^l = \frac{1}{nsim - nburn} \sum_{i=nburn+1}^{nsim} \widehat{RV}_{T+1,i}^l$
- (2) For $T = is + 1, is + 9, is + 11, \dots$
- (a) Draw λ_t for $t = is - T + 1, \dots, T$, as illustrated in Section D of the Supplement.
 - (b) Forecast the states in equation (3.2) and (3.5) then forecast $\widehat{RV}_{T+1}^l$ accordingly.

Forecasts of the Realised Variance are then computed based on the expectation of a log-normal distribution,⁴ as follows:

$$\widehat{RV}_{t+1} = \exp \left(\widehat{RV}_{t+1}^l + \frac{\hat{\omega}_{t+1|t}^2}{2} \right), \quad (3.6)$$

where:

$$\begin{aligned} \hat{\omega}_{t+1|t}^2 = & \hat{\sigma}_v^2 + \widehat{Var}(\beta_{1,t+1|t}) + (RV_t^l)^2 \times \widehat{Var}(\beta_{2,t+1|t}) \\ & + (RV_{w,t}^l)^2 \times \widehat{Var}(\beta_{3,t+1|t}) + (RV_{m,t}^l)^2 \times \widehat{Var}(\beta_{4,t+1|t}). \end{aligned}$$

Here, the first term in the expression of $\hat{\omega}_{t+1|t}^2$ - denoted as $\hat{\sigma}_v^2$ - is the variance of the measurement equation, whereas the subsequent terms represent the variance of each of the state equations entering through the coefficients in the measurement equation. In the SHARP model, $\widehat{Var}(\beta_{j,t+1|t}) = \hat{\sigma}_{\epsilon,j}^2$, for $j = 1, \dots, 4$; conversely, in the SHARP-SV model, $\widehat{Var}(\beta_{j,t+1|t}) = \exp \left(\ln(\widehat{h}_{j,t+1}) + \hat{\sigma}_{u,j}^2/2 \right)$.

⁴We note that we *always* compute forecasts based on the Gaussianity assumption, which can of course be mis-specified. As our simulations in Section 5.2 show, however, this does not impair the predictive accuracy of our methodology.

4. EMPIRICAL STUDY

In this section, we: describe the datasets employed in Section 4.1; analyse the estimated coefficients of our volatility model(s) in Section 4.2; and report a comprehensive out-of-sample forecasting analysis in Section 4.3. For brevity, we report only a selection of the main results; the full-blown set of results is in Section A of the Supplement.

4.1. Data. We use two datasets, and the VIX index. The first dataset comprises 4,451 trading days of SPY-ETF, representing the US stock market, and ten ETFs representing its ten economic sectors, spanning from January 3rd, 2006 to September 8th, 2023. The second dataset includes 4,277 trading days of twenty individual NYSE stocks, covering the period from January 3rd, 2000 to December 31st, 2016. We also use the VIX index, from July 1st, 2003 to December 29th, 2023, with a sample size of 5,161 observations. For both datasets and the VIX index, daily RV was computed from tick-level price observations obtained from TickWrite.⁵ The selected stocks represent different market sectors and vary in terms of market activity and volatility: the SPY serves as a widely recognized proxy for aggregate market behaviour; the sector ETFs and stocks offer insight into cross-sectional and sector-specific volatility dynamics; and the inclusion of the VIX adds a fundamentally different type of volatility measure (one that is anticipatory/forward-looking, rather than historical), which contrasts with the RV-based metrics - thus striking a balance between coverage, relevance, and comparability. The chosen sample period includes significant market events such as the

⁵TickWrite is a commercial database that provides data for futures, index, and equity markets. Tick data is sourced from NYSE's TAQ (Trade and Quote) database and is adjusted for ticker mapping, code filtering, price splits, and dividend payments. More information can be found at <https://www.tickdata.com/>. In the construction of our databases, the selection criteria for individual stocks were as follows: (a) only stocks continuously traded over the full sample period were considered; (b) in order to address liquidity and staleness biases, we selected stocks ranking in the top 15th percentile by trading volume, which also tend to fall in the bottom 20th percentile for zero returns; (c) in order to capture potential sector-specific intraday dynamics, we ensured representation from all GICS sectors, choosing two qualifying stocks per sector. The data were aggregated from the tick level using previous-tick interpolation, and sampled at 5-minute intervals, a standard frequency in the high-frequency literature that balances bias and variance (see Ait-Sahalia et al., 2005; and Hansen and Lunde, 2006).

2008 global financial crisis and the COVID-19 pandemic, providing a diverse range of market conditions. Table A.1 in Section A of the Supplement contains the descriptive statistics of RV for all datasets.

4.2. Analysis of temporal volatility coefficients. Using a rolling window approach with an estimation window of 1,000 daily observations (approximately four years), we plot the estimated coefficients from the SHARP and HARL models, used in the out-of-sample forecasting period for the realised variances of the SPY-ETF and the VIX respectively, in Figures 4.1 and 4.2 respectively.

We note that the estimated intercept $\hat{\beta}_{1,t}$ fluctuates significantly, dropping during calm periods and rising during crises (in fact, $\hat{\beta}_{1,t}$ tends to drop below its corresponding average during tranquil periods), while the intercept of the HARL model remains fixed; this suggests

Figure 4.1. Out-of-sample estimated coefficients, $\hat{\beta}_{tj}$, by SHARP and HARL models on SPY-ETF realised variance in the period January 1st, 2006 to September 8th, 2023.

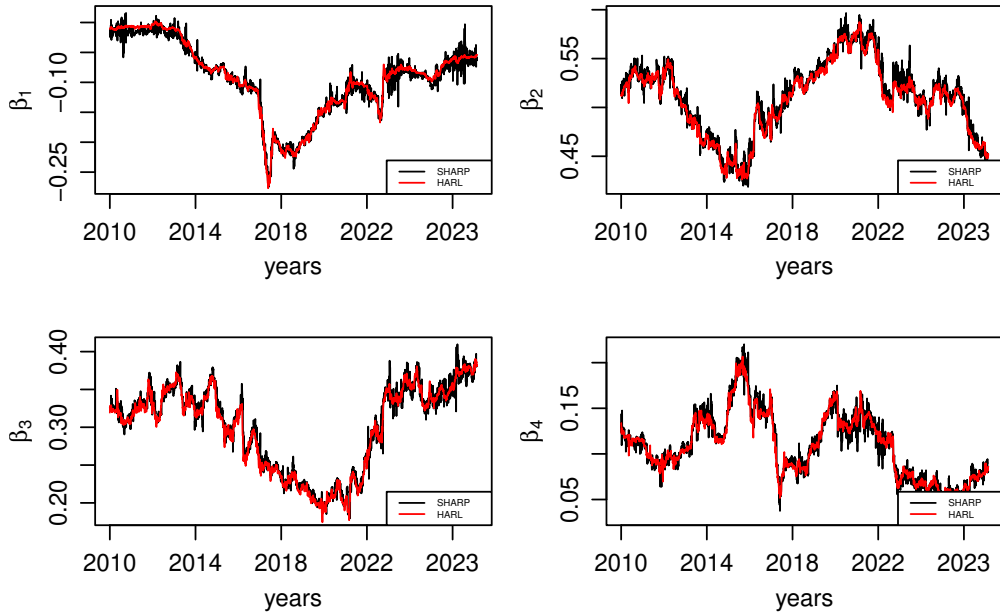
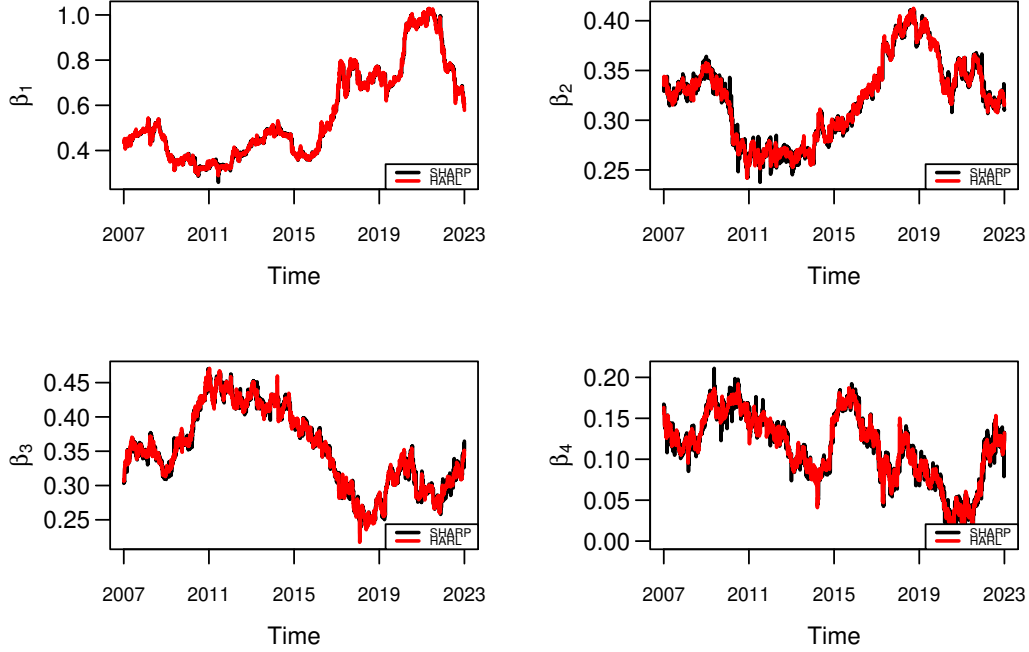


Figure 4.2. Out-of-sample estimated coefficients, $\hat{\beta}_{tj}$, by SHARP and HARL models on VIX realised variance in the period July 24th, 2007 to December 29th, 2023.



that SHARP better adapts to changing baseline levels of realised volatility. Indeed, SHARP shows varying persistence across daily, weekly, and monthly horizons, reflecting evolving investor responses at different frequencies. HARL, by contrast, imposes static weights, missing this flexibility. The daily coefficient becomes more prominent during periods of uncertainty, such as the COVID-19 period, whereas the weekly coefficient shows a decline during these times; further, the monthly coefficient increases during the COVID-19 period. This phenomenon during times of financial market uncertainty can be attributed to the “primacy” and “recency” effects; the current long-term conditions, reflected in the monthly average of RV^l , can be considered primary information, while recent information is represented by the daily RV^l . When predicting short-term (daily) volatility during periods of uncertainty, primary and recent information become more relevant than intermediate information, such as the weekly average volatility.

In Tables 4.1 and 4.2, we report the correlations between estimated parameters in the SHARP and SHARP-SV models for the realised variances of the SPY-ETF and the VIX respectively. The tables indicate a negative correlation between the weekly and daily coefficients, as well as between the weekly and monthly coefficients, whereas the correlation between the daily and monthly coefficients is only weakly negative.

TABLE 4.1. Correlation matrix of β_{tj} parameters obtained by estimating the SHARP and SHARP-SV models on SPY-ETF realised variance in the period January 1st, 2000 to September 8th, 2023.

SHARP					$SHARP_{sv}$				
$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$		$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	
$\hat{\beta}_1$	1	-0.1740	0.7049	-0.0389	$\hat{\beta}_1$	1	-0.1431	0.6355	-0.0243
$\hat{\beta}_2$	-0.1740	1	-0.5772	-0.1763	$\hat{\beta}_2$	-0.1431	1	-0.4675	-0.1212
$\hat{\beta}_3$	0.7049	-0.5772	1	-0.4095	$\hat{\beta}_3$	0.6355	-0.4675	1	-0.3840
$\hat{\beta}_4$	-0.0388	-0.1763	-0.4095	1	$\hat{\beta}_4$	-0.0243	-0.1212	-0.3840	1

TABLE 4.2. Correlation matrix of β_{tj} parameters obtained by estimating the SHARP and SHARP-SV models on VIX realised variance in the period July 24th, 2007 to December 29th, 2023.

SHARP					$SHARP_{sv}$				
$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$		$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	
$\hat{\beta}_1$	1.0000	0.6318	-0.7545	-0.8420	$\hat{\beta}_1$	1.0000	0.6292	-0.7552	-0.8427
$\hat{\beta}_2$	0.6318	1.0000	-0.9344	-0.3613	$\hat{\beta}_2$	0.6292	1.0000	-0.9342	-0.3562
$\hat{\beta}_3$	-0.7545	-0.9344	1.0000	0.4049	$\hat{\beta}_3$	-0.7552	-0.9342	1.0000	0.4063
$\hat{\beta}_4$	-0.8420	-0.3613	0.4049	1.0000	$\hat{\beta}_4$	-0.8427	-0.3562	0.4063	1.0000

4.3. Out-of-sample comparative analysis. We report an out-of-sample analysis of the SHARP and SHARP-SV models, along with a set of competing models, in forecasting the daily, weekly, and monthly realised Variance (RV) of SPY-ETF, VIX, twenty NYSE individual stocks, and ten economic sector ETFs. We report the out-of-sample forecasts using

a rolling window approach with an estimation window of 1,000 daily observations (approximately four years).⁶ For the NYSE stocks dataset, the in-sample period covers 1,000 observations, approximately from January 2000 to December 2003. The out-of-sample period spans from January 2004 to December 2016. For the SPY-ETF and sector ETFs dataset, the in-sample period also uses the first 1,000 observations, approximately from January 2006 to December 2009, with the out-of-sample period from January 2010 to September 2023. For the VIX series, the in-sample period also uses the first 1,000 observations, approximately from July 2003 to July 2007, with the out-of-sample period from August 2007 to December 2023. In our horse-race, we evaluate the performance of our SHARP and SHARP-SV models, also in comparisons with other, existing set-ups. For example, the comparison between HARL and HARLQ investigates whether accounting for measurement error in the RV^l series improves the forecasts by HARL, as observed by [Bollerslev et al. \(2016\)](#) when replacing the HAR model with their HARQ specification. Additionally, comparing HARSL and HARLQ outlines whether modelling the daily coefficient as an AR(1) process provides better forecasting performance than the more restrictive yet straightforward HARLQ model. The analysis between TVCHAR, SHARK, and SHARP demonstrates which time-varying specification of the coefficients yields better forecasts when compared to each other and to the aforementioned models. Finally, the comparison between SHARP and SHARP-SV reveals whether the inclusion of the stochastic volatility (SV) feature in the AR process of the coefficients enhances the forecasts. These comparisons are designed to evaluate the robustness and forecasting accuracy of the proposed models in relation to existing benchmarks.

⁶Specifically, this involves continuously updating the dataset by removing the oldest observation and adding the most recent one, thus maintaining a constant window of 1,000 observations. While in real-time forecasting one should ideally jump 5 and 22 observations for weekly and monthly forecasts respectively, we adopted the aforementioned method for the sake of simplicity - especially when updating the state equations for longer horizons. This ensures continuous updating of the dataset with the latest data, thereby maintaining the robustness and uniformity of our forecasts. In an unreported set of experiments, we note that the relative performance of our SHARP and SHARP-SV models compared to their competitors remains unaltered even when using the alternative approach.

We compare the accuracy of each model in forecasting RV over three horizons: daily, weekly, and monthly. The forecasts are compared with their corresponding actual values $\exp(RV_t^l)$, $\exp(RV_{w,t}^l)$, and $\exp(RV_{m,t}^l)$, respectively. Since we perform direct forecasting for all the models, we do not report the SHARK recursive forecasting for weekly and monthly horizons to maintain consistency across the models.

In order to provide a comprehensive analysis of forecast accuracy, we use the following loss functions: (i) the Mean Squared Error (MSE); (ii) the Mean Absolute Error (MAE); (iii) the heteroskedasticity-adjusted version of the mean squared error (HMSE), designed to adjust for the variability in y_t , thus being less sensitive to periods of high volatility and providing a scale-invariant measure of forecast accuracy (e.g. [Wang et al., 2015](#)), defined as

$$HMSE = \frac{1}{n - is} \sum_{t=is+1}^n \left(1 - \frac{\hat{y}_t}{y_t}\right)^2,$$

where, as customary, we denote the actual and the predicted values as y_t and $\hat{y}_t$ respectively, and is is defined in Section 3.2; (iv) the heteroskedasticity-adjusted version of the Mean Absolute Error (HMAE), defined as

$$HMAE = \frac{1}{n - is} \sum_{t=is+1}^n \left|1 - \frac{\hat{y}_t}{y_t}\right|;$$

and (v) the Quasi-Likelihood (QLIKE) loss function, defined as

$$QLIKE = \frac{1}{n - is} \sum_{t=is+1}^n \left(\frac{\hat{y}_t}{y_t} - \ln \left(\frac{\hat{y}_t}{y_t} \right) - 1 \right),$$

which penalises under-predictions of volatility. In all cases, in order to facilitate comparisons, we report the *relative* loss measures (say RL_M , referring to model M), where we evaluate the relative loss compared to the benchmark HARL specification, viz.

$$RL_M = \frac{LF_M}{LF_{HARL}}, \tag{4.1}$$

where LF_M denotes the value of the relevant loss function of model M , and LF_{HARL} the value taken by the same loss function for the HARL model. We present the RL_M for SPY-ETF and VIX data in Tables 4.3 and 4.4, respectively, where we also use the Model Confidence Set methodology of Hansen et al. (2011).⁷

TABLE 4.3. Out-of-sample relative loss measure (4.1) of the models (with HARL being the benchmark) obtained by estimating the models on SPY-ETF realised variance over a rolling window of 1000 observations in the period January 3, 2006, to September 8, 2023.

Panel (a): Daily							
	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSv
MSE	1*	1.0631	0.9986*	0.9982*	1.0007*	1.0454	1.0479
MAE	1	1.0069	0.9970	1.0000	0.9998	0.9764	0.9624*
HMSE	1	0.9994	0.9712	1.0002	0.9787	0.8579	0.7121*
HMAE	1	0.9990	0.9845	0.9998	0.9890	0.9290	0.8549*
QLIKE	1	0.9981	0.9833	0.9999	0.9873	0.9163	0.8325*
Panel (b): Weekly							
MSE	1	0.9921	0.7744	0.9937		0.7138*	0.9561*
MAE	1	0.9975	0.9596	0.9991		0.7706	0.7022*
HMSE	1	0.9993	0.6262	0.9984		0.4452	0.2866*
HMAE	1	0.9994	0.7583	0.9992		0.6867	0.5550*
QLIKE	1	0.9999	0.6726	0.9989		0.5353	0.3820*
Panel (c): Monthly							
MSE	1	0.9916	0.8111	0.9973		0.6109	0.4890*
MAE	1	0.9967	0.8126	0.9990		0.6151	0.4381*
HMSE	1	0.9992	0.5094	0.9958		0.2632	0.1015*
HMAE	1	0.9993	0.6089	0.9975		0.5230	0.3183*
QLIKE	1	0.9995	0.4977	0.9968		0.3336	0.1379*

“*” indicates that the model is included in the $\hat{M}_{75\%}$ Model Confidence Set.

Figures 4.3 and 4.4 summarize the RL_M for the sector ETFs and the twenty NYSE individual stocks respectively; the full-blown set of numerical results is also reported in Tables A.2 and A.3 in Section A the Supplement, respectively.⁸

⁷In essence, this approach tests the null hypothesis that all models are equally effective against the alternative that a smaller subset of models is superior. We select a p -value threshold of 0.25 based on range statistics: models with p -values below the 0.25 threshold are excluded from the superior subset, denoted by $\hat{M}_{75\%}$. In Tables A.2 and A.3 in Section A of the Supplement, we also report the frequency with which each model is included in the $\hat{M}_{75\%}$ for the sector ETFs and individual NYSE stocks, respectively.

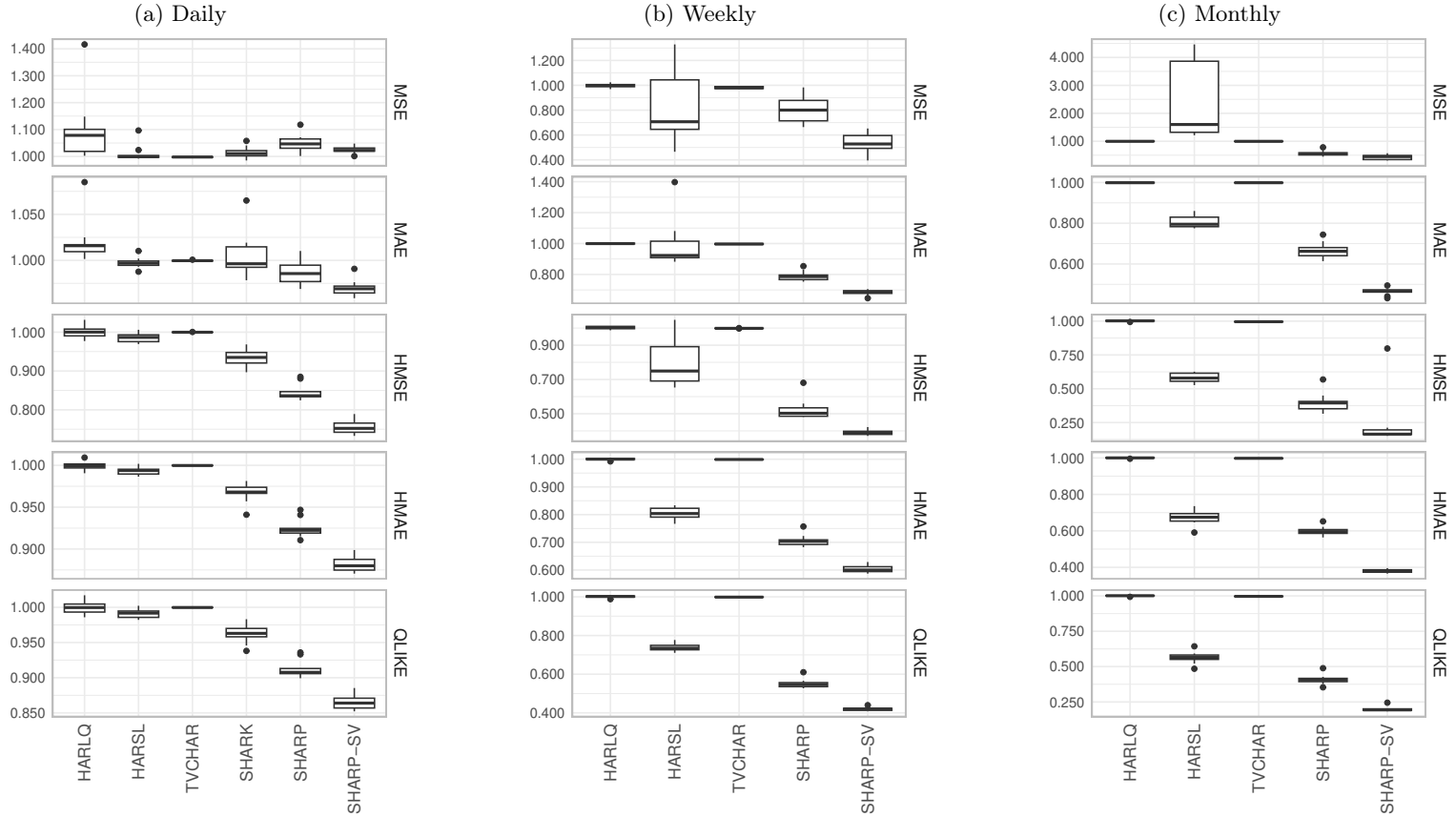
⁸In Section A of the Supplement, we also present detailed results for the individual ETFs (Tables A.4-A.8) and the individual stocks (Tables A.9-A.13).

TABLE 4.4. Out-of-sample relative loss measure (4.1) of the models (with HARL being the benchmark) obtained by estimating the models on VIX realised variance over a rolling window of 1,000 observations in the period July 1, 2003, to December 29, 2023.

Panel (a): Daily							
	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSV
MSE	1	1.0132	1.2956	1.0001	1.0086	0.9998	1.0055*
MAE	1	0.9970	1.3442	1.0001	1.0044	0.9708	0.9559*
HMSE	1	0.9399	4.7839	0.9995	1.0099	0.7337	0.6619*
HMAE	1	0.9818	1.6925	1.0000	1.0048	0.8622	0.8206*
QLIKE	1	0.9719	2.2312	0.9999	1.0057	0.8394	0.7929*
Panel (b): Weekly							
MSE	1	0.9919	7.0596	1.0000		0.8219	0.6069*
MAE	1	0.9924	0.8709	1.0000		0.9396	0.7443*
HMSE	1	0.9628	0.4948	0.9999		0.8002	0.3729*
HMAE	1	0.9861	0.5660*	1.0000		0.9051	0.6341
QLIKE	1	0.9777	0.4192*	0.9999		0.8137	0.4675
Panel (c): Monthly							
MSE	1	0.9916	1.2366	0.9999		0.7417	0.6994*
MAE	1	0.9971	0.4589*	0.9999		0.7727	0.8311
HMSE	1	0.9973	0.1573*	0.9996		0.3808	0.5881
HMAE	1	0.9964	0.3508*	0.9999		0.6599	0.8018
QLIKE	1	0.9918	0.1724*	0.9997		0.5090	0.6427

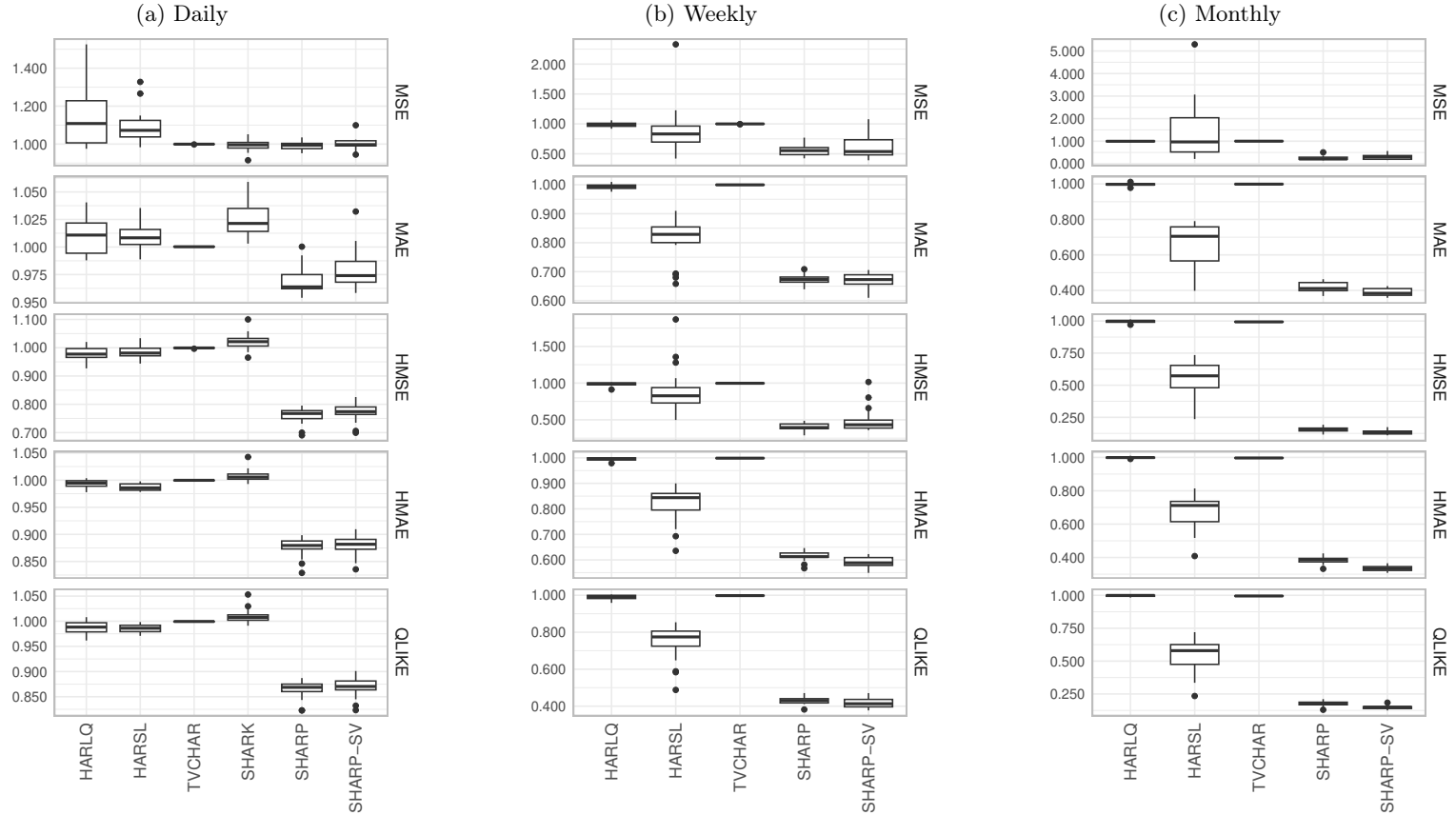
“*” indicates that the model is included in the $\hat{M}_{75\%}$ Model Confidence Set.

Figure 4.3. Box plots of relative loss functions comparison - using the ten-sectors ETFs dataset.



Box plots of relative loss measures for five loss functions (MSE, MAE, HMSE, HMAE, QLIKE) across three forecasting horizons (Daily, Weekly, and Monthly) for each model, using the HARL as the benchmark model. The models were estimated on the realised variance for each of the ten sector ETFs over a rolling window of 1,000 observations during the period from January 3, 2006, to September 8, 2023. One outlier point was removed from each series for the weekly and monthly forecasting horizons when computing the MSE for the HARSL model to prevent inflation of the measure. Applying the same principle to other models did not impact the MSE, so this adjustment was only done for the HARSL model.

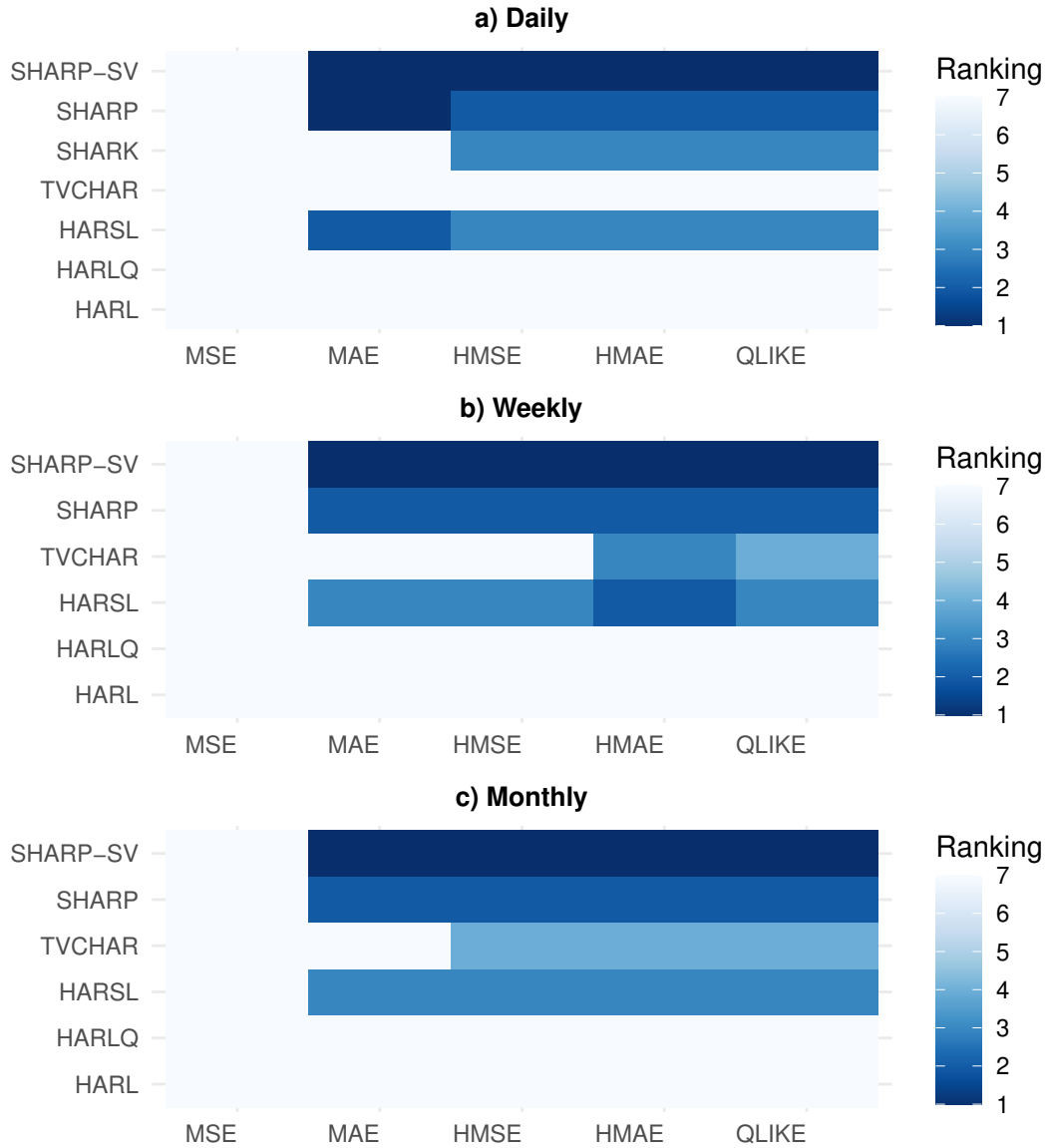
Figure 4.4. Box plots of relative loss functions comparison - using the twenty individual NYSE stocks dataset.



Box plots of relative loss measures for five loss functions (MSE, MAE, HMSE, HMAE, QLIKE) across three forecasting horizons (Daily, Weekly, and Monthly) for each model, using the HARL as the benchmark model. The models were estimated on the realised variance for each of the twenty NYSE individual stocks over a rolling window of 1,000 observations during the period from January 3, 2000, to December 31, 2016. One outlier point was removed from each series for the weekly and monthly forecasting horizons when computing the MSE for the HARSL model to prevent inflation of the measure. Applying the same principle to other models did not impact the MSE, so this adjustment was only done for the HARSL model.

We also use the *unconditional* predictive ability (uCPA) test by [Giacomini and White \(2006\)](#) to evaluate the out-of-sample predictions produced by the models. We perform pairwise tests of unconditional predictive ability over the full out-of-sample (OOS) period with a significance level of $\alpha = 0.05$. To ensure robustness, we conduct the same test using five different loss functions and examine the results across three forecasting horizons (see [Figure 4.5](#)).

Figure 4.5. Heatmap of Model Rankings Based on uCPA (Unconditional Conditional Predictive Ability)



Results of the uCPA test for out-of-sample forecasting performance between models, with each selected loss function obtained by estimating the models on SPY-ETF realised variance over a rolling window of 1000 observations during the period from January 1st, 2006 to September 8th, 2023. The rankings range from 1 (best) to 7 (least). A darker color indicates a better model performance. If two models have the same rank (i.e., color), they are considered equally good. The results are shown for three forecasting horizons in panels a, b, and c, representing daily, weekly, and monthly forecasts, respectively.

Overall, our results obtained using the RV^l of SPY-ETF, sector ETFs, individual stocks and the VIX are consistent throughout. For most individual stocks, incorporating some form of time-varying specification of the coefficients generally enhances the models forecasting accuracy, especially for longer forecasting horizons. With the exception of the MSE loss function in the case of daily forecasting, the SHARP and SHARP-SV models almost always exhibit the lowest relative loss (RL) using the SPY-ETF data and the average relative loss using both the sector ETFs and the twenty individual stocks data. A partial exception is encountered with the VIX, where, at lower frequencies, results are less clear-cut - although, at a daily frequency, our proposed models still deliver a superior forecasting performance than the other models. Introducing the SV feature to the AR process of the coefficients yields a further moderate improvement compared to the forecasts by the SHARP model. Thus, we find that the general specification of the time-varying coefficients in the SHARP-SV model leads to the most significant improvement in the forecasts of RV across all three forecasting horizons. These results are also robust across the five loss function measures used. As far as the other specifications are concerned, the HARSL model shows some moderate improvements - mainly at weekly and monthly horizons - suggesting that accounting for time variation in coefficients is more effective than simple corrections for measurement error. By contrast, HARLQ, TVCHAR, and SHARK generally fail to deliver meaningful improvements over the baseline HARL. Taken together, these results demonstrate that the flexibility of time-varying parameters, and especially the incorporation of stochastic volatility, is critical for achieving reliable forecast improvements across a wide cross-section of assets. Note also that our models, particularly the SHARP-SV model, are almost always included in the confidence set, while other models are frequently excluded with few exceptions.

In conclusion, we note that our model assumes Gaussianity of all innovations. Whilst assuming Gaussianity may not have an impact on the inference on coefficients, inference on the volatility process may be distorted in the presence of departures from normality such as

excess kurtosis or nonzero skewness.⁹ Seeing as our data are likely to exhibit such features, by way of misspecification analysis we carry out a test for Gaussianity inspired by [Koopman and Scharth \(2012\)](#). Whilst originally designed for a different specification, the test is based on a parametric bootstrap which we also use here, generating the pseudosamples using the estimated parameters and simulating all innovations from Gaussian distributions. Results in Tables [A.14](#) and [A.15](#) in Section [A.1](#) in the Supplement show that the assumption of Gaussianity is indeed most often rejected; on the other hand, our empirical results show some degree of robustness to such misspecification, which are also reinforced by our Monte Carlo analysis in Section [5.2](#).

5. MONTE CARLO STUDY

We evaluate the performance of our models (and of Bayesian estimation) via a comprehensive set of simulations. In Section [5.1](#), we assess the performance of the Particle Gibbs estimator under correct model specification; in Section [5.2](#), we evaluate predictive accuracy under a possibly misspecified model.

5.1. The performance of the Particle Gibbs estimator. We study the performance of Particle Gibbs estimation under correct model specification. In particular, we evaluate (a) the ability to recover true parameters and latent states, and (b) the comparative performance of SHARP estimation via Particle Gibbs versus a Kalman filter based estimation.

Data are generated according to equations [\(3.1\)](#)-[\(3.3\)](#), reported here for convenience

$$\begin{aligned} y_t &= x_t' \beta_t + \nu_t, \quad \nu_t \sim i.i.d. \mathcal{N}(0, \sigma_v^2), \\ \beta_{j,t} &= \rho_j \beta_{j,t-1} + \varepsilon_{j,t}, \quad \varepsilon_{j,t} \sim i.i.d. \mathcal{N}(0, \sigma_{\varepsilon,j}^2), \end{aligned}$$

⁹We are grateful to an anonymous Referee for pointing this out to us, and suggesting the testing approach used herein.

with $1 \leq t \leq n$. Recall that the vector $x_t = (1, RV_{t-1}^l, RV_{w,t-1}^l, RV_{m,t-1}^l)'$ includes the daily, weekly, and monthly (lagged) log-realized RV, and that the innovations ν_t and $\varepsilon_{t,j}$ are mutually independent. The parameter values are set as $\rho_j = 0.96$ for $1 \leq j \leq 4$, and $(\sigma_{\varepsilon,1}, \sigma_{\varepsilon,2}, \sigma_{\varepsilon,3}, \sigma_{\varepsilon,4}) = (0.15, 0.08, 0.08, 0.08)$; further, we set $\sigma_v = 0.02$ for the measurement error standard deviation. Initial values are set to $\beta_{1,1} = -0.5$, $\beta_{2,1} = 0.4$, $\beta_{3,1} = 0.3$, and $\beta_{4,1} = 0.15$, and are based on estimates from fitting a standard HARL model to the daily log-RV of the SPY index. These choices also align with decay structures documented in the literature (e.g. Andersen et al., 2007; and Corsi, 2009). We initialize the log-RV process y_t , by generating the first $1 \leq t \leq 22$ observations as *i.i.d.* $\mathcal{N}(0, 1)$. Our experiments are based on a sample size $n = 1,000$, and on 1,000 replications.

Figure 5.1 summarizes the estimation accuracy for the four time-varying coefficients $\beta_{t,j}$ across three key metrics: normalized RMSE, 95% interval coverage, and interval width.¹⁰ As can be seen, using the SHARP specification with Particle Gibbs systematically achieves the lowest estimation error, with average normalised RMSEs of 0.428, 0.724, 0.894, and 0.949 for $\beta_{t,1}$ through $\beta_{t,4}$, respectively. Estimation via Kalman filter performs comparably well for the short-horizon coefficients $\beta_{t,1}$ and $\beta_{t,2}$, but it exhibits higher errors for $\beta_{t,3}$ and $\beta_{t,4}$ (1.101 and 1.570 respectively), reflecting its limited flexibility in tracking longer-term dynamics. Turning to interval calibration, using the Particle Gibbs estimator achieves reasonably balanced coverage across all coefficients, with coverage rates ranging from 0.707 to 0.781, accompanied by normalized interval widths of 0.774, 1.541, 2.165, and 2.367 for $\beta_{t,1}$ through $\beta_{t,4}$, respectively. In contrast, Kalman filter estimation yields wider intervals 1.153, 2.527, 2.677, and 2.770 for $\beta_{t,1}$ through $\beta_{t,4}$; while delivering better coverage for $\beta_{t,1}$ and $\beta_{t,2}$ (0.866 and 0.919 respectively), it still under-covers $\beta_{t,3}$ and $\beta_{t,4}$ (with coverages

¹⁰The *normalised* RMSE is defined by computing (the square root of) $\sum_{t=1}^T (\hat{\beta}_{t,j} - \beta_{t,j})^2 / T$, and then dividing it by $\sigma_{\varepsilon,j}$; we use the same normalisation also for coverage.

0.758 and 0.712 respectively).¹¹ Hence, Kalman filter estimation does not consistently yield better-calibrated coverage, especially for longer-horizon components, despite substantially wider intervals, which is reinforced by the variability in coverage across coefficients and replications (especially as far as higher-lag coefficients are concerned). Conversely, Particle Gibbs estimation exhibits tighter dispersion and fewer outliers across all metrics, suggesting robustness across simulated series.

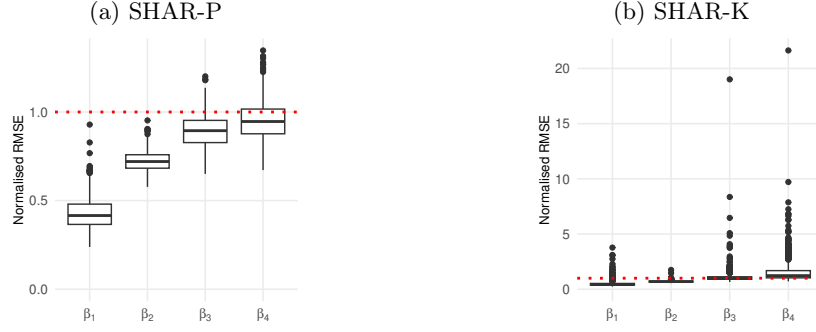
In Figure 5.2, we assess the recovery of fixed parameters persistence (ρ_j , $1 \leq j \leq 4$), innovation volatilities ($\sigma_{\varepsilon,j}$), and measurement error (σ_v). Similarly to the previous results, Kalman filter estimation recovers the persistence coefficients ρ_j accurately for short-horizon coefficients (0.958 and 0.954 on average across simulations for ρ_1 and ρ_2) but it underestimates persistence in β_3 and β_4 (with ρ_3 and ρ_4 being, on average, 0.895 and 0.873). Particle Gibbs estimation displays more balanced estimates across all lags (with ρ_j , on average, equal to 0.946, 0.927, 0.918, and 0.913 for $1 \leq j \leq 4$), despite having a slight downward bias. In terms of state innovation volatilities $\sigma_{\varepsilon,j}$, Kalman filter estimation yields average values of estimates (0.146, 0.079, 0.071, 0.071), closely matching the true values; Particle Gibbs also performs well, but with a larger downward bias, delivering average estimates (0.118, 0.069, 0.067, 0.067). Finally, turning to the measurement error variance σ_v , Kalman filter recovers the true value nearly exactly, whereas Particle Gibbs slightly overestimates (estimated values are 0.020 and 0.032 respectively). As before, Particle Gibbs corresponds to tight, symmetric distributions around the true values, with very few outliers, indicating robust and consistent performance; Kalman filter performs comparably, though it shows slightly greater dispersion in the longer-lag persistence estimates.

¹¹Naturally, the Kalman filter produces frequentist confidence intervals, based on the conditional state covariance matrices, while the particle filter generates Bayesian credible intervals from posterior samples. Although these intervals differ in interpretation, comparing their coverage and width offers insight into how each method quantifies uncertainty in recovering the latent states.

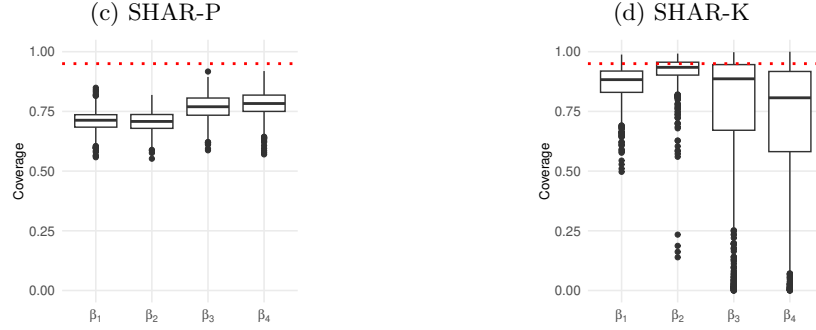
The distilled essence of our simulations is that the relative strengths of Kalman filter and Particle Gibbs are *complementary*. The former performs better at estimating fixed parameters, and it is also superior in terms of coverage for the “short-run” parameters - i.e. the intercept $\beta_{t,1}$ and the coefficient multiplying the daily log RV, $\beta_{t,2}$. On the other hand, the Particle Gibbs estimator performs better in terms of RMSE when estimating “long-run” parameters (i.e. $\beta_{t,3}$ and $\beta_{t,4}$, associated to the weekly and monthly log RVs), and it has - in general - less dispersion and fewer outliers across experiments relative to the Kalman filter. As we show in the next section,¹² these relative advantages of Particle Gibbs translate into superior predictive accuracy. Hence, while the Kalman filter remains useful for efficient inference on fixed parameters and near-term components, Particle Gibbs offers a more robust alternative when the primary objectives are accurate state recovery and predictive performance.

¹²See also Section B.2 in the Supplement.

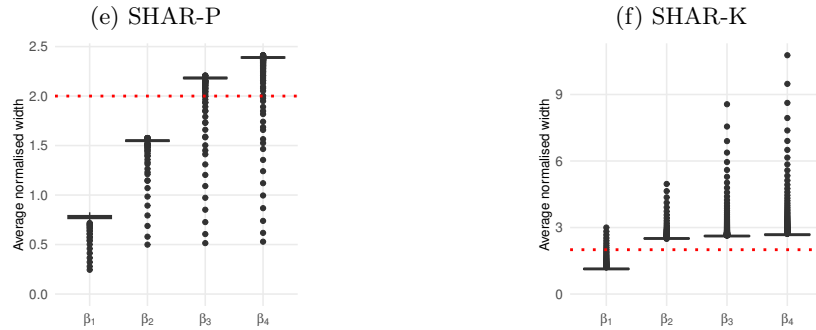
Figure 5.1. Time-Varying coefficients - estimation performance
(i) **Normalised RMSE**



(ii) **Coverage rate using the constructed 95% intervals for the true values**

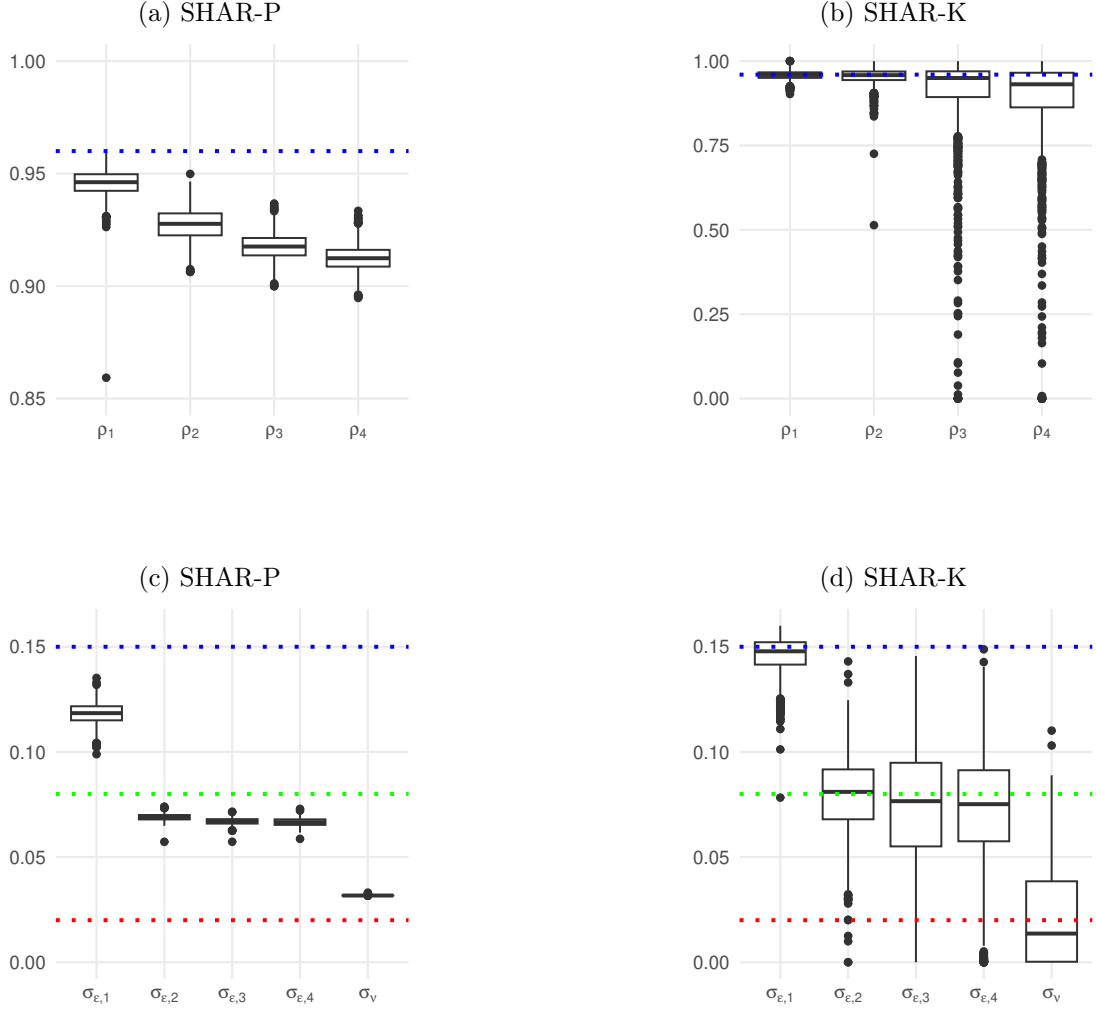


(iii) **Normalised width of the constructed 95% intervals**



Estimation results obtained under Particle Gibbs and Kalman filter are indicated as SHAR-P and SHAR-K respectively. The dotted red lines represent theoretical benchmark values (1 for RMSE, 0.95 for coverage, and 2 for interval width), corresponding to the expected variability of the latent coefficients under the DGP. Coverage values closer to one reflect correctly calibrated 95% intervals.

Figure 5.2. Fixed parameters - estimation performance



The blue dotted line in panels (a) and (b) indicates the true value of $\rho = 0.96$ used in the data-generating process (DGP). Recall that estimation results obtained under Particle Gibbs and Kalman filter are indicated as SHAR-P and SHAR-K respectively. Panel (c) and (d) show boxplots of the estimated innovation standard deviations $\sigma_{\epsilon,j}$ and the measurement error standard deviation σ_v under SHAR-P and SHAR-K, respectively; the blue dotted line corresponds to the true value of $\sigma_{\epsilon,1} = 0.15$, the green line marks the true value of $\sigma_{\epsilon,j}$, $j = 2, 3, 4$, and the red line marks the true value of the measurement error standard deviation $\sigma_v = 0.02$.

5.2. Forecasting accuracy with mis-specified models. In order to assess the robustness of our approach, we now investigate predictive accuracy under a possibly mis-specified model. We consider the following DGP, where we generate a series of returns along similar lines to [Jacquier et al. \(2004\)](#) and [Koopman and Scharth \(2012\)](#), viz.

$$y_t = \mu + \exp\left(\frac{h_t}{2}\right) \sqrt{\lambda_t} \epsilon_t + J_t \mathcal{Z}_t^y. \quad (5.1)$$

In (5.1): λ_t is a scaling factor which we model according to two schemes, with $\lambda_t = 1$ for all t , or drawn from an inverse-Gamma distribution, i.e.

$$\lambda_t \sim i.i.d. \mathcal{IG}\left(\frac{v}{2}, \frac{v}{2}\right); \quad (5.2)$$

J_t is an *i.i.d.* Bernoulli random variable representing the jump times with $P(J_t = 1) = \kappa$, and the jumps $\mathcal{Z}_t^y$ are generated as $\mathcal{Z}_t^y \sim i.i.d. \mathcal{N}(\mu_y, \sigma_y^2)$. We generate ϵ_t according to $\epsilon_t \sim i.i.d. \mathcal{N}(0, 1)$; this entails that, when (5.2) is used, the innovations $\sqrt{\lambda_t} \epsilon_t$ are *i.i.d.* with a Student's t distribution with v degrees of freedom. We extend the specification in (5.1) by following a similar approach to [Stroud and Johannes \(2014\)](#); we model the log of the diffusive, non-jump variance h_t as

$$h_t = (\mu_h + \gamma s_t) + x_{1,t} + x_{2,t}, \quad (5.3)$$

where: $\mu_h + \gamma s_t$ is a mean level depending on a regime variable $s_t \in \{0, 1\}$ following a Markov chain with transition probabilities $P(s_t = i | s_{t-1} = j) = p_{ij}$;¹³ $x_{1,t}$ is a “slow” variance component generated as

$$x_{1,t+1} = \phi_1 x_{1,t} + \sigma_1 u_{1,t}, \quad (5.4)$$

¹³We introduce a regime-switching dynamics in (5.3) following [Vo \(2009\)](#); this can be viewed as an extension of the model considered in [Stroud and Johannes \(2014\)](#).

with a “large” partial autocorrelation coefficient $0 < \phi_1 < 1$ and innovations $u_{1,t} \sim i.i.d.\mathcal{N}(0, 1)$; and $x_{2,t}$ is a “fast” variance component generated as

$$x_{2,t+1} = \phi_2 x_{2,t} + \sigma_2 \left(\rho \epsilon_t + \sqrt{1 - \rho^2} u_{2,t} \right) + J_t \mathcal{Z}_t^2, \quad (5.5)$$

with a “small” partial autocorrelation coefficient $0 < \phi_2 < \phi_1$, innovations $u_{2,t} \sim i.i.d.\mathcal{N}(0, 1)$, and jumps $\mathcal{Z}_t^2 \sim i.i.d.\mathcal{N}(\mu_2, \sigma_2^2)$. In (5.1), (5.4) and (5.5), $\{\epsilon_t, -\infty < t < \infty\}$, $\{\mathcal{Z}_t^y, -\infty < t < \infty\}$, $\{u_{1,t}, -\infty < t < \infty\}$, $\{u_{2,t}, -\infty < t < \infty\}$ and $\{\mathcal{Z}_t^2, -\infty < t < \infty\}$ are five mutually independent groups; hence, ρ in (5.5) represent a measure of diffusive leverage, via the correlation between the innovations in the returns DGP ϵ_t and the fast variance component $x_{2,t}$. The DGP in (5.1)-(5.5) thus considers a comprehensive specification designed to capture the key empirical features observed in financial market volatility, and in particular in the dynamics of the realised variance, and it builds - as well as on the references cited above - on several contributions in this literature including, *inter alia*, [Harvey and Shephard \(1996\)](#), [Jacquier et al. \(2004\)](#), [Chib et al. \(2006\)](#), [McAleer and Medeiros \(2008\)](#), and [Koopman and Scharth \(2012\)](#). Several models are nested within (5.1)-(5.5); in Table B.1 in the Supplement, we summarise five variants, each incorporating different empirical features. The baseline “SV” model assumes normal return innovations with two volatility factors; the “SVt” specification extends this by allowing for heavier tails; “SVL” considers the presence of a leverage effect between returns and volatility; “SVLJ” allows for jumps in both returns and volatility (as well as leverage effects); finally, “SVML” combines leverage effects with Markov-switching regimes. In Table B.2 in the Supplement, we report the values of the parameters used in our simulations, which largely follow the estimated values in Table 3 of [Stroud and Johannes \(2014\)](#), allowing our DGP to be more adherent to the features of real data.

In addition to the above five SV-based DGPs, we also consider a further DGP, based on a HAR-type model to account for correlations between state innovations in a SHARP model:

$$RV_t = \beta_{1,t} + \beta_{2,t}RV_{t-1} + \beta_{3,t}RV_{t-1}^{(w)} + \beta_{4,t}RV_{t-1}^{(m)} + \epsilon_t,$$

$$\beta_{j,t} = \phi_j \beta_{j,t-1} + \nu_{j,t}, \quad j = 1, 2, 3, 4,$$

with state innovations $\nu_t = (\nu_{1,t}, \nu_{2,t}, \nu_{3,t}, \nu_{4,t})' \sim \mathcal{N}(\mathbf{0}, \Sigma_\nu)$, $\phi = (0.92, 0.92, 0.90, 0.90)'$ and

$$\Sigma_\nu = \begin{bmatrix} 0.0400 & 0.0008 & -0.0010 & 0.0240 \\ 0.0008 & 0.0014 & -0.0010 & 0.0004 \\ -0.0010 & -0.0010 & 0.0050 & -0.0040 \\ 0.0240 & 0.0004 & -0.0040 & 0.0250 \end{bmatrix} \quad (5.6)$$

For each of the six DGPs mentioned above: we generate a synthetic realised volatility (RV) series of length $n = 1,440$, with the last 440 observations used for out-of-sample forecasting via a rolling window approach; we estimate our proposed models, SHARP and SHARP-SV, against “traditional” alternatives, including the HARL model, as well as more advanced benchmarks like HARSL and TVCHAR; and we compare the forecasting ability of the models mentioned above using standard loss functions (MSE, HMSE, HMAE, and QLIKE). Each experiment is based on 100 Monte Carlo simulations for the sake of computational efficiency. We report our main results in Figures B.1 and B.2 in the Supplement. The figures show that both the SHARP and SHARP-SV models consistently deliver the lowest loss function values across all DGPs, thereby demonstrating superior predictive accuracy. The only exception was the simplest SV-based DGP, where the HARLQ model performed marginally better on average but exhibited substantial outliers, making its performance less reliable across different simulations. Among the competing models, HARSL and TVCHAR showed more stable performance across DGPs; however, both the SHARP and SHARP-SV models

outperformed them on average, with lower mean loss function values in almost all cases. Further results are also in the Supplement. In Section B.2, along similar lines as in the previous Section 5.1, we assess the impact of the estimation technique on the performance of our proposed model - in essence, by estimating the SHARP model via MLE *cum* Kalman filtering (Table B.3). Results are decidedly worse, which reinforces the findings in Section 4: the superior forecasting ability associated with the SHARP model and with Particle Gibbs estimation holds also in a set of controlled scenarios using synthetic data. Further, in Table B.4 in Section B.3, we report computational times for the baseline SV DGP, showing that our Bayesian estimator remains feasible for large-scale applications.

6. CONCLUSION

This paper develops two new dynamic state-space models for realised volatility forecasting - SHARP and SHARP-SV - which extend the heterogeneous autoregressive framework of Corsi (2009) by allowing time-varying coefficients governed by autoregressive processes, with SHARP-SV further incorporating stochastic volatility in the coefficient dynamics. We follow Creal and Tsay (2015), using Gibbs particle filtering as a computationally efficient approach (Andrieu et al., 2010). We study the structure of our proposed models, characterising their dependence and deriving a set of asymptotic results which could be of independent interest. In a comprehensive set of empirical studies, we apply the SHARP and SHARP-SV models to the realised variance of the SPY index, sector ETFs, representative NYSE stocks, and the VIX index over long samples. Across all datasets, horizons, and loss functions (MSE, MAE, HMSE, HMAE, and QLIKE), SHARP and SHARP-SV on balance outperform HARL and its extensions, including HARSL (McAleer and Medeiros, 2008), HARLQ (Wang et al., 2016), TVCHAR (Bekierman and Manner, 2018), and SHARK (Buccheri and Corsi, 2021); these gains are confirmed by predictive ability tests (Giacomini and White, 2006) and Model Confidence Set analysis (Hansen et al., 2011). SHARP and SHARP-SV remain robust

even in the presence of non-Gaussian innovations. Monte Carlo evidence further validates these conclusions: under correct specification, Particle Gibbs accurately recovers model parameters and latent states, outperforming Kalman filtering in terms of RMSE and interval calibration for long-horizon coefficients. Under misspecification, we show that SHARP and SHARP-SV maintain superior predictive accuracy - under various forms of misspecification - compared with other, competing models, with SHARP-SV showing particular strength. These findings establish SHARP and SHARP-SV as versatile and effective tools for volatility forecasting, with direct implications for risk management, asset allocation, and derivative pricing.

REFERENCES

- Ait-Sahalia, Y., P. A. Mykland, and L. Zhang (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *The Review of Financial Studies* 18(2), 351–416.
- Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39(4), 885–905.
- Andersen, T. G., T. Bollerslev, and F. X. Diebold (2007, November). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *Review of Economics and Statistics* 89(4), 701–720.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* 96(453), 42–55.
- Andrieu, C., A. Doucet, and R. Holenstein (2010). Particle Markov Chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B* 72(3), 269–342.
- Barndorff-Nielsen, O. E. and N. Shephard (2002). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics* 17(5), 457–477.
- Bekierman, J. and H. Manner (2018). Forecasting realized variance measures using time-varying coefficient models. *International Journal of Forecasting* 34(2), 276–287.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31(3), 307–327.

- Bollerslev, T., A. J. Patton, and R. Quaadvlieg (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics* 192(1), 1–18.
- Bollerslev, T., A. J. Patton, and R. Quaadvlieg (2018). Modeling and forecasting (un)reliable realized covariances for more reliable financial decisions. *Journal of Econometrics* 207(1), 71–91.
- Buccheri, G. and F. Corsi (2021). HARK the SHARK: Realized volatility modeling with measurement errors and nonlinear dependencies. *Journal of Financial Econometrics* 19(4), 614–649.
- Carriero, A., T. E. Clark, and M. Marcellino (2015). Bayesian VARs: Specification choices and forecast accuracy. *Journal of Applied Econometrics* 30(1), 46–73.
- Carriero, A., T. E. Clark, and M. Marcellino (2019). Large Bayesian Vector Autoregressions with stochastic volatility and non-conjugate priors. *Journal of Econometrics* 212(1), 137–154.
- Chen, X. B., J. Gao, D. Li, and P. Silvapulle (2018). Nonparametric estimation and forecasting for time-varying coefficient realized volatility models. *Journal of Business & Economic Statistics* 36(1), 88–100.
- Chib, S., F. Nardari, and N. Shephard (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics* 134(2), 341–371.
- Clark, T. E. (2011). Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility. *Journal of Business & Economic Statistics* 29(3), 327–341.
- Clark, T. E. and F. Ravazzolo (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics* 30(4), 551–575.
- Cogley, T. and T. J. Sargent (2005). Drifts and volatilities: Monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics* 8(2), 262–302.
- Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7(2), 174–196.
- Creal, D. and R. S. Tsay (2015). High dimensional dynamic stochastic copula models. *Journal of Econometrics* 189(2), 335–345.
- D’Agostino, A., L. Gambetti, and D. Giannone (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics* 28(1), 82–101.
- Doan, T., R. Litterman, and C. Sims (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews* 3(1), 1–100.
- Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *Econometrica* 74(6), 1545–1578.

- Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics* 24(2), 127–161.
- Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica* 79(2), 453–497.
- Harvey, A., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *The Review of Economic Studies* 61(2), 247–264.
- Harvey, A. C. and N. Shephard (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics* 14(4), 429–434.
- Jacquier, E., N. G. Polson, and P. E. Rossi (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics* 122(1), 185–212.
- Kokoszka, P., N. Mohammadi, H. Wang, and S. Wang (2025). Functional diffusion driven stochastic volatility model. *Bernoulli* 31(2), 922–947.
- Koop, G. and D. Korobilis (2013). Large time-varying parameter VARs. *Journal of Econometrics* 177(2), 185–198.
- Koop, G., D. Korobilis, and D. Pettenuzzo (2019). Bayesian compressed Vector AutoRegressions. *Journal of Econometrics* 210(1), 135–154.
- Koop, G. and S. Potter (2004). Forecasting in dynamic factor models using Bayesian model averaging. *The Econometrics Journal* 7(2), 550–565.
- Koopman, S. J. and M. Scharth (2012). The analysis of stochastic volatility in the presence of daily realized measures. *Journal of Financial Econometrics* 11(1), 76–115.
- Liu, L. Y., A. J. Patton, and K. Sheppard (2015). Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *Journal of Econometrics* 187(1), 293–311.
- McAleer, M. and M. C. Medeiros (2008). A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. *Journal of Econometrics* 147(1), 104–119.
- Sims, C. A. (1993). A nine-variable probabilistic macroeconomic forecasting model. In *Business Cycles, Indicators and Forecasting*, pp. 179–212. University of Chicago Press.
- Stock, J. H. and M. W. Watson (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business & Economic Statistics* 14(1), 11–30.
- Stroud, J. R. and M. S. Johannes (2014). Bayesian modeling and forecasting of 24-hour high-frequency volatility. *Journal of the American Statistical Association* 109(508), 1368–1384.
- Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes - a study of the daily sugar prices 1961-75. *Time Series Analysis: Theory and Practice* 1, 203–226.

- Tsionas, M. G., M. Izzeldin, and L. Trapani (2022). Estimation of large dimensional time varying VARs using copulas. *European Economic Review* 141, 103952.
- Vo, M. T. (2009). Regime-switching stochastic volatility: Evidence from the crude oil market. *Energy Economics* 31(5), 779–788.
- Wang, X., C. Wu, and W. Xu (2015). Volatility forecasting: The role of lunch-break returns, overnight returns, trading volume and leverage effects. *International Journal of Forecasting* 31(3), 609–619.
- Wang, Y., F. Ma, Y. Wei, and C. Wu (2016). Forecasting realized volatility in a changing world: A dynamic model averaging approach. *Journal of Banking & Finance* 64, 136–149.
- Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 100(472), 1394–1411.

SUPPLEMENTARY MATERIAL FOR A DYNAMIC STATE-SPACE HAR MODEL

MIKE TSIONAS¹, AYA GHALAYINI², MARWAN IZZELDIN¹, AND LORENZO TRAPANI³

CONTENTS

A. Tables and further empirical evidence	2
A.1. Misspecification analysis	16
B. Further Monte Carlo results	20
B.1. Mnemonics, parameter values and figures for Section 5.2	20
B.2. Comparison with MLE <i>cum</i> Kalman filtering	23
B.3. Computational times	25
C. Posterior Derivation	26
D. Particle filtering within MCMC	31
E. The structure of the SHARP model	34
F. Technical lemmas	40
G. Proofs	51
References	56

¹LANCASTER UNIVERSITY MANAGEMENT SCHOOL, LANCASTER, UK

²COLLEGE OF BUSINESS AND SOCIAL SCIENCES, ASTON UNIVERSITY, BIRMINGHAM, UK

³UNIVERSITA' DEGLI STUDI DI PAVIA, PAVIA, ITALY, AND UNIVERSITY OF LEICESTER, LEICESTER, UK

E-mail addresses: m.tsionas@lancaster.ac.uk, a.ghalayini@aston.ac.uk,

m.izzeldin@lancaster.ac.uk, lorenzo.trapani@unipv.it.

Key words and phrases. Time-varying coefficients, HAR models, particle Gibbs sampling, state space models, volatility forecasting.

A. TABLES AND FURTHER EMPIRICAL EVIDENCE

TABLE A.1. Descriptive statistics for Realised Variance series

Stock names	Ticker	Sector	Mean	Median	St. Dev.	Skewness	Kurtosis	Min	Max
Panel (a): Individual NYSE stocks (January 3 rd , 2000 to December 31 st , 2016)									
Constellation Energy Group	AEE	Utilities	1.8275	1.0614	3.4532	14.5795	362.0346	0.1089	113.4878
Brown-Forman Corp.	BFB	Consumer Staples	1.9203	1.1517	4.8608	32.2490	1449.2850	0.0742	240.4141
BT Group plc (ADR)	BT	Communications Services	2.3113	1.1621	3.2277	4.7116	47.6503	0.1004	59.5677
Exelon Corp.	EXC	Utilities	2.6354	1.4288	4.7911	9.4223	161.6486	0.1585	130.8746
Freeport-McMoran	FCX	Materials	8.0344	4.3266	12.2786	5.6708	50.4463	0.3168	188.5795
General Dynamics	GD	Industrials	2.2370	1.2810	3.2592	6.2650	67.3871	0.0807	63.2822
General Electric	GE	Industrials	3.0201	1.3030	6.9818	10.3833	172.3605	0.1077	180.3886
The Home Depot	HD	Consumer Discretionary	3.1214	1.5733	4.9381	6.7444	83.0200	0.1557	103.4768
TECO Energy	HES	Energy	4.4738	2.5702	8.6746	12.7967	280.6983	0.2109	271.5113
Humana Inc.	HUM	Health Care	6.6787	2.6090	11.3665	4.4673	33.5169	0.2404	157.5287
IBM	IBM	Information Technology	2.0255	0.9862	3.5274	7.4383	92.6942	0.1019	71.2926
Coca-Cola	KO	Consumer Staples	1.5608	0.8355	2.5353	8.6219	138.4686	0.0456	58.8085
Marriott Int'l.	MAR	Consumer Discretionary	3.5370	1.7819	5.3635	5.4536	55.7000	0.1543	104.5781
Nucor Corp.	NUE	Materials	4.9098	2.7544	10.5817	13.8696	279.0002	0.3337	266.8244
Pfizer	PFE	Health Care	2.3324	1.3819	3.2242	6.4660	77.5016	0.1498	62.6970
AT&T	T	Communications Services	2.6549	1.1840	4.7673	9.4248	195.7955	0.1082	141.8456
Travelers -Travelers Group Inc	TRV	Financials	2.9683	1.1863	7.8664	15.2588	379.9201	0.1020	263.9287
Wells Fargo	WFC	Financials	4.3023	1.3299	12.1389	8.2667	94.9689	0.1036	226.6092
ExxonMobil	XOM	Energy	1.9866	1.1409	3.9555	15.8314	430.8596	0.1067	141.1297
Panel (b): Sectors ETFs (January 3 rd , 2006 to September 8 th , 2023)									
SPDR S&P 500 ETF Trust	SPY	US Market Index	0.9136	0.3515	2.3943	10.2282	155.3431	0.0102	59.8630
Apple Inc.	AAPL	Information Technology	5.2923	2.5776	7.8061	4.4878	37.8812	0.0791	126.1716
iShares US Real Estate ETF	IYR	Real Estate	2.1185	0.7179	5.4011	7.1884	74.4394	0.0610	89.8417
iShares US Telecommunications ETF	IYZ	Telecommunication	1.3845	0.6666	6.0169	44.1633	2464.7837	0.0801	347.1409
Materials Select Sector SPDR Fund	XLB	Materials	1.6320	0.7701	3.5186	9.4764	134.6835	0.0694	81.3017
Energy Select Sector SPDR Fund	XLE	Energy	2.4324	1.3051	4.9089	10.2001	162.7808	0.0721	123.5975
Financial Select Sector SPDR Fund	XLF	Financials	2.1276	0.7227	5.9090	10.7814	196.4172	0.0464	162.3466
Industrial Select Sector SPDR Fund	XLI	Industrials	1.2672	0.5728	2.9961	10.5326	168.8990	0.0356	71.4864
Consumer Staples Select Sector SPDR Fund	XLP	Consumer Staples	0.7733	0.3580	2.6630	18.8287	504.9143	0.0422	95.5998
Utilities Select Sector SPDR Fund	XLU	Utilities	1.2840	0.6695	3.4659	13.0561	242.6988	0.0918	88.8259
Health Care Select Sector SPDR Fund	XLV	Health Care	0.8953	0.4551	1.9639	10.1240	144.4243	0.0488	43.1275
Consumer Discretionary Select Sector SPDR Fund	XLY	Consumer Discretionary	1.3945	0.5658	3.1129	8.9484	123.1008	0.0291	66.5541
Panel (c): VIX (July 1 st , 2003 to December 29 th , 2023)									
Chicago Board Options Exchange's Volatility Index	VIX		22.3416	12.0669	44.4396	14.5572	344.9257	0.4354	1421.9150

The table presents the descriptive statistics for the realised variance series of selected individual NYSE stocks over January 3, 2000 to December 31, 2016 in Panel (a); sector ETFs over January 3, 2006 to September 8, 2023 in Panel (b); and the VIX series between July 1st, 2003 to December 29th, 2023 in Panel (c). It Includes stock names, tickers, sectors, and statistical measures: mean, median, standard deviation, skewness, kurtosis, minimum, and maximum.

TABLE A.2. Summary Statistics of Out-of-Sample Relative Loss Measures for realised Variance of the Ten Sectors ETFs Dataset

Panel (a): Daily							
	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPs _v
MSE							
Mean	1.0000	1.0979	1.0105	0.9977	1.0147	1.0482	1.0244
Median	1.0000	1.0784	0.9991	0.9979	1.0106	1.0466	1.0227
M75%	3	1	6	7	3	1	1
MAE							
Mean	1.0000	1.0202	0.9973	0.9997	1.0051	0.9874	0.9701
Median	1.0000	1.0159	0.9969	0.9997	0.9963	0.9856	0.9691
M75%	0	0	0	0	0	0	10
HMSE							
Mean	1.0000	1.0006	0.9857	0.9998	0.9319	0.8448	0.7551
Median	1.0000	1.0003	0.9865	0.9998	0.9349	0.8362	0.7522
M75%	0	0	0	0	0	0	10
HMAE							
Mean	1.0000	0.9993	0.9932	0.9998	0.9670	0.9245	0.8816
Median	1.0000	0.9996	0.9939	0.9998	0.9681	0.9223	0.8799
M75%	0	0	0	0	0	0	10
QLIKE							
Mean	1.0000	0.9997	0.9910	0.9998	0.9621	0.9122	0.8650
Median	1.0000	0.9996	0.9918	0.9997	0.9630	0.9077	0.8640
M75%	0	0	0	0	0	0	10
Panel (b): Weekly							
MSE							
Mean	1.0000	1.0070	1.1231	0.9918		0.7547	0.5676
Median	1.0000	1.0071	0.8680	0.9923		0.7171	0.5687
M75%	0	0	0	0		1	10
MAE							
Mean	1.0000	1.0014	1.0813	0.9991		0.7821	0.6867
Median	1.0000	1.0019	0.9951	0.9990		0.7777	0.6840
M75%	0	0	0	0		0	10
HMSE							
Mean	1.0000	1.0029	0.9707	0.9990		0.5037	0.3933
Median	1.0000	1.0044	0.8820	0.9991		0.4925	0.3718
M75%	0	0	0	0		1	9
HMAE							
Mean	1.0000	1.0005	0.8475	0.9994		0.7049	0.6087
Median	1.0000	1.0004	0.8424	0.9994		0.7036	0.6063
M75%	0	0	0	0		0	10
QLIKE							
Mean	1.0000	1.0012	0.7982	0.9990		0.5473	0.4234
Median	1.0000	1.0023	0.7965	0.9990		0.5416	0.4253
M75%	0	0	0	0		0	10
Panel (c): Monthly							
MSE							
Mean	1.0000	0.9975	2.4021	0.9966		0.6455	0.4441
Median	1.0000	0.9978	1.5996	0.9971		0.6378	0.4139
M75%	0	0	0	0		1	10
MAE							
Mean	1.0000	0.9986	0.8084	0.9989		0.6255	0.4379
Median	1.0000	0.9988	0.7951	0.9988		0.6200	0.4355
M75%	0	0	0	0		0	10
HMSE							
Mean	1.0000	1.0022	0.5818	0.9959		0.2867	0.1295
Median	1.0000	1.0019	0.5791	0.9959		0.2746	0.1278
M75%	0	0	0	0		0	10
HMAE							
Mean	1.0000	1.0000	0.6722	0.9979		0.5272	0.3467
Median	1.0000	1.0002	0.6748	0.9980		0.5224	0.3442
M75%	0	0	0	0		0	10
QLIKE							
Mean	1.0000	1.0007	0.5640	0.9968		0.3326	0.1502
Median	1.0000	1.0011	0.5650	0.9967		0.3252	0.1463
M75%	0	0	0	0		0	10

This table presents the mean and median of out-of-sample relative loss measures using five loss functions (MSE, MAE, HMSE, HMAE, and QLIKE) over three forecasting horizons: daily (Panel A), weekly (Panel B), and monthly (Panel C). The models are estimated based on the realised variance of ten sector ETFs, using a rolling window of 1,000 observations, from January 3, 2006, to September 8, 2023. Additionally, the table reports the number of times out of the ten ETFs that the model was included in the $\hat{M}_{75\%}$ model confidence set for each horizon.

TABLE A.3. Summary Statistics of Out-of-Sample Relative Loss Measures for realised Variance of the Twenty NYSE Stocks Dataset

Panel (a): Daily							
	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPs _v
MSE							
Mean	1.0000	1.1470	1.2348	0.9995	0.9939	1.0141	1.0324
Median	1.0000	1.1084	1.1866	0.9995	0.9962	1.0114	1.0194
M75%	8	6	0	13	15	4	0
MAE							
Mean	1.0000	1.0098	1.0387	1.0003	1.0254	0.9739	0.9861
Median	1.0000	1.0108	1.0341	1.0003	1.0214	0.9719	0.9770
M75%	2	1	1	2	1	20	4
HMSE							
Mean	1.0000	0.9774	1.0202	0.9991	1.0228	0.7617	0.7748
Median	1.0000	0.9774	1.0131	0.9992	1.0214	0.7690	0.7826
M75%	0.0000	0.0000	0.0000	0.0000	0.0000	16	6.0000
HMAE							
Mean	1.0000	0.9933	0.9999	0.9996	1.0079	0.8775	0.8842
Median	1.0000	0.9945	1.0006	0.9997	1.0056	0.8803	0.8893
M75%	0	0	0	0	0	15	3
QLIKE							
Mean	1.0000	0.9876	1.0025	0.9994	1.0100	0.8651	0.8726
Median	1.0000	0.9883	1.0028	0.9995	1.0076	0.8695	0.8777
M75%	0	0	0	0	0	16	4
Panel (b): Weekly							
MSE							
Mean	1.0000	0.9840	1.4364	0.9988		0.6351	0.5829
Median	1.0000	0.9741	1.1875	0.9992		0.6204	0.5672
M75%	0	0	0	0		9	12
MAE							
Mean	1.0000	0.9930	0.8932	0.9996		0.6722	0.6952
Median	1.0000	0.9925	0.9043	0.9996		0.6703	0.6910
M75%	0	0	0	0		19	2
HMSE							
Mean	1.0000	0.9825	1.2008	0.9977		0.4516	0.4092
Median	1.0000	0.9869	1.0684	0.9977		0.4147	0.3903
M75%	0	0	0	0		7	17
HMAE							
Mean	1.0000	0.9944	0.8701	0.9988		0.5979	0.6141
Median	1.0000	0.9956	0.8882	0.9988		0.5950	0.6089
M75%	0	0	0	0		20	1
QLIKE							
Mean	1.0000	0.9882	0.8329	0.9981		0.4198	0.4276
Median	1.0000	0.9915	0.8521	0.9981		0.4142	0.4158
M75%	0	0	0	0		15	8
Panel (c): Monthly							
MSE							
Mean	1.0000	0.9953	0.7952	0.9998		0.2335	0.2808
Median	1.0000	0.9967	0.7190	0.9997		0.2073	0.2772
M75%	0	0	1	0		16	6
MAE							
Mean	1.0000	0.9972	0.6634	0.9990		0.4137	0.4387
Median	1.0000	0.9978	0.6916	0.9988		0.4099	0.4314
M75%	0	0	0	0		12	9
HMSE							
Mean	1.0000	0.9941	0.6068	0.9935		0.1543	0.1385
Median	1.0000	0.9988	0.6315	0.9931		0.1527	0.1328
M75%	0	0	0	0		5	15
HMAE							
Mean	1.0000	0.9980	0.6997	0.9966		0.3850	0.3602
Median	1.0000	0.9989	0.7375	0.9966		0.3867	0.3553
M75%	0	0	0	0		5	15
QLIKE							
Mean	1.0000	0.9956	0.5756	0.9947		0.1763	0.1560
Median	1.0000	0.9975	0.6125	0.9949		0.1759	0.1504
M75%	0	0	0	0		5	15

This table presents the mean and median of out-of-sample relative loss measures using five loss functions (MSE, MAE, HMSE, HMAE, and QLIKE) over three forecasting horizons: daily (Panel A), weekly (Panel B), and monthly (Panel C). The models are estimated based on the realised variance of twenty NYSE stocks, using a rolling window of 1,000 observations, from January 3, 2000, to December 31, 2016. Additionally, the table reports the number of times out of the twenty stocks that the model was included in the $\hat{M}_{75\%}$ model confidence set for each horizon.

TABLE A.4. Out-of-Sample Relative Loss Measure MSE using the Sectors ETFs dataset

	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSV
Panel (a): Daily							
IYR	1.0000	1.1481	1.0239	0.9955*	0.9851*	1.0412	1.0268
IYZ	1.0000*	1.0026*	1.0008*	0.9999*	1.0016*	1.0016*	1.0012*
XLB	1.0000	1.1015	0.9986*	0.9971*	1.0408	1.0671	1.0245
XLE	1.0000	1.0085	1.0966	0.9969*	1.0165	1.1181	1.0196
XLF	1.0000	1.0145	0.9922*	0.9990	0.9957	1.0073	1.0051
XLI	1.0000	1.0985	0.9967*	0.9988	1.0578	1.0587	1.0210
XLP	1.0000	1.4164	0.9996	0.9983*	1.0101	1.0521	1.0479
XLU	1.0000	1.0699	1.0041	0.9955*	1.0235	1.0711	1.0447
XLV	1.0000	1.0870	0.9957*	0.9975	1.0111	1.0291	1.0329
XLY	1.0000*	1.0320	0.9971*	0.9982*	1.0045*	1.0352	1.0203
Panel (b): Weekly							
IYR	1.0000	1.0127	1.6624	0.9884		0.7187	0.5995*
IYZ	1.0000	0.9859	0.7172	0.9928		0.6766*	0.6497*
XLB	1.0000	1.0024	0.8624	0.9937		0.6516	0.4428*
XLE	1.0000	1.0119	2.3435	0.9917		0.7657	0.5390*
XLF	1.0000	1.0013	1.0973	0.9908		0.6627	0.4767*
XLI	1.0000	1.0051	1.6651	0.9932		0.8403	0.5984*
XLP	1.0000	1.0045	0.7624	0.9918		0.9602	0.7202*
XLU	1.0000	1.0091	0.5359	0.9877		0.8493	0.6390*
XLV	1.0000	1.0247	0.7116	0.9935		0.7156	0.4923*
XLY	1.0000	1.0125	0.8736	0.9945		0.7060	0.5184*
Panel (c): Monthly							
IYR	1.0000	1.0035	4.3168	0.9949		0.6482	0.4610*
IYZ	1.0000	0.9953	1.5784	0.9975		0.6443	0.3509*
XLB	1.0000	0.9961	1.2161	0.9983		0.6024	0.3576*
XLE	1.0000	1.0050	1.6208	0.9943		0.6313	0.2479*
XLF	1.0000	1.0021	1.5257	0.9957		0.5608	0.3016*
XLI	1.0000	0.9882	2.5037	0.9971		0.6732	0.5613*
XLP	1.0000	0.9992	4.4646	0.9971		0.7901*	0.8337*
XLU	1.0000	0.9963	4.3259	0.9965		0.7320	0.4968*
XLV	1.0000	1.0009	1.2166	0.9974		0.6184	0.4635*
XLY	1.0000	0.9884	1.2525	0.9972		0.5540	0.3668*

This table presents the out-of-sample relative loss measure MSE over three forecasting horizons: daily, weekly, and monthly, for models estimated on the sector ETFs' realised variance over a rolling window of 1,000 observations for the period January 3, 2006, to September 8, 2023. The columns represent different models (HARL, HARLQ, HARSL, TVCHAR, SHARK, SHARP, SHARPSV), and the rows correspond to various sector ETFs. An asterisk (*) indicates that the model is included in the $\hat{M}_{75\%}$ model confidence set.

TABLE A.5. Out-of-Sample Relative Loss Measure MAE using the Sectors ETFs dataset

	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSV
Panel (a): Daily							
IYR	1.0000	1.0250	0.9950	0.9994	0.9917	0.9842	0.9672*
IYZ	1.0000	1.0090	0.9945	0.9998	0.9975	0.9819	0.9712*
XLB	1.0000	1.0153	0.9931	0.9993	1.0048	0.9871	0.9637*
XLE	1.0000	1.0015	1.0102	0.9998	0.9948	1.0094	0.9723*
XLF	1.0000	1.0108	0.9877	0.9992	0.9784	0.9689	0.9589*
XLI	1.0000	1.0167	0.9969	0.9996	1.0192	0.9874	0.9707*
XLP	1.0000	1.0850	0.9999	1.0007	1.0650	1.0102	0.9908*
XLU	1.0000	1.0171	1.0019	0.9992	1.0180	0.9972	0.9762*
XLV	1.0000	1.0165	0.9969	0.9997	0.9951	0.9753	0.9675*
XLY	1.0000	1.0052	0.9972	0.9998	0.9862	0.9727	0.9628*
Panel (b): Weekly							
IYR	1.0000	1.0029	1.0635	0.9986		0.7749	0.6867*
IYZ	1.0000	1.0014	0.9514	0.9994		0.7479	0.6801*
XLB	1.0000	1.0005	0.9379	0.9990		0.7561	0.6498*
XLE	1.0000	1.0007	0.9964	0.9997		0.8208	0.7112*
XLF	1.0000	1.0019	0.9937	0.9989		0.7512	0.6643*
XLI	1.0000	1.0019	1.1403	0.9988		0.7806	0.6813*
XLP	1.0000	0.9941	1.5826	0.9990		0.8486	0.7235*
XLU	1.0000	1.0025	1.1999	0.9986		0.7878	0.7062*
XLV	1.0000	1.0062	0.9668	0.9990		0.7681	0.6724*
XLY	1.0000	1.0024	0.9802	0.9996		0.7852	0.6918*
Panel (c): Monthly							
IYR	1.0000	1.0017	0.8570	0.9988		0.6228	0.4464*
IYZ	1.0000	0.9989	0.7826	0.9986		0.5957	0.4126*
XLB	1.0000	1.0001	0.7751	0.9986		0.5820	0.4074*
XLE	1.0000	0.9987	0.8195	0.9993		0.6535	0.4363*
XLF	1.0000	1.0008	0.7905	0.9988		0.5844	0.4041*
XLI	1.0000	0.9977	0.7997	0.9984		0.6172	0.4307*
XLP	1.0000	0.9962	0.8332	0.9984		0.7080	0.5005*
XLU	1.0000	0.9979	0.8601	0.9993		0.6521	0.4690*
XLV	1.0000	1.0006	0.7816	0.9994		0.6234	0.4369*
XLY	1.0000	0.9940	0.7847	0.9993		0.6156	0.4347*

This table presents the out-of-sample relative loss measure MAE over three forecasting horizons: daily, weekly, and monthly, for models estimated on the sector ETFs' realised variance over a rolling window of 1,000 observations for the period January 3, 2006, to September 8, 2023. The columns represent different models (HARL, HARLQ, HARSL, TVCHAR, SHARK, SHARP, SHARPSV), and the rows correspond to various sector ETFs. An asterisk (*) indicates that the model is included in the $\hat{M}_{75\%}$ model confidence set.

TABLE A.6. Out-of-Sample Relative Loss Measure HMSE using the Sectors ETFs dataset

	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSV
Panel (a): Daily							
IYR	1.0000	1.0105	0.9693	0.9997	0.9391	0.8345	0.7586*
IYZ	1.0000	1.0317	0.9875	0.9997	0.9367	0.8336	0.7443*
XLB	1.0000	0.9972	0.9754	0.9997	0.9194	0.8472	0.7679*
XLE	1.0000	0.9771	0.9974	0.9999	0.9239	0.8850	0.7891*
XLF	1.0000	1.0045	0.9746	1.0000	0.8988	0.8246	0.7458*
XLI	1.0000	1.0017	0.9856	0.9994	0.9505	0.8370	0.7415*
XLP	1.0000	0.9882	1.0061	1.0005	0.9684	0.8804	0.7634*
XLU	1.0000	1.0090	0.9761	0.9994	0.8968	0.8454	0.7665*
XLV	1.0000	0.9988	0.9934	0.9999	0.9332	0.8354	0.7409*
XLY	1.0000	0.9876	0.9920	1.0001	0.9525	0.8245	0.7328*
Panel (b): Weekly							
IYR	1.0000	1.0106	1.0279	0.9990		0.5121	0.3943*
IYZ	1.0000	1.0125	0.8069	0.9993		0.4770*	0.5499
XLB	1.0000	0.9971	0.7632	0.9985		0.4807	0.3654*
XLE	1.0000	0.9959	1.3362	0.9990		0.5377	0.4050*
XLF	1.0000	0.9976	0.8000	0.9993		0.4923	0.3687*
XLI	1.0000	1.0054	1.3414	0.9984		0.4817	0.3566*
XLP	1.0000	0.9873	1.1484	0.9998		0.5591	0.3750*
XLU	1.0000	1.0081	0.9572	0.9985		0.5210	0.4061*
XLV	1.0000	1.0110	0.7354	0.9993		0.4824	0.3491*
XLY	1.0000	1.0034	0.7908	0.9994		0.4927	0.3624*
Panel (c): Monthly							
IYR	1.0000	1.0096	0.6128	0.9968		0.2778	0.1306*
IYZ	1.0000	1.0049	0.5270	0.9956		0.3130	0.1383*
XLB	1.0000	1.0013	0.5740	0.9948		0.2607	0.1171*
XLE	1.0000	1.0026	0.6245	0.9957		0.2688	0.1264*
XLF	1.0000	1.0006	0.5492	0.9964		0.2423	0.1119*
XLI	1.0000	1.0010	0.5739	0.9950		0.2714	0.1169*
XLP	1.0000	0.9939	0.6161	0.9977		0.3875	0.1688*
XLU	1.0000	1.0046	0.6264	0.9949		0.2845	0.1355*
XLV	1.0000	1.0049	0.5295	0.9960		0.2910	0.1291*
XLY	1.0000	0.9987	0.5842	0.9962		0.2698	0.1204*

This table presents the out-of-sample relative loss measure HMSE over three forecasting horizons: daily, weekly, and monthly, for models estimated on the sector ETFs' realised variance over a rolling window of 1,000 observations for the period January 3, 2006, to September 8, 2023. The columns represent different models (HARL, HARLQ, HARSL, TVCHAR, SHARK, SHARP, SHARPSV), and the rows correspond to various sector ETFs. An asterisk (*) indicates that the model is included in the $\hat{M}_{75\%}$ model confidence set.

TABLE A.7. Out-of-Sample Relative Loss Measure HMAE using the Sectors ETFs dataset

	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSV
Panel (a): Daily							
IYR	1.0000	1.0022	0.9865	0.9997	0.9670	0.9204	0.8833*
IYZ	1.0000	1.0093	0.9941	0.9998	0.9717	0.9195	0.8766*
XLB	1.0000	0.9995	0.9895	0.9997	0.9667	0.9244	0.8882*
XLE	1.0000	0.9965	0.9943	1.0000	0.9692	0.9407	0.8988*
XLF	1.0000	1.0008	0.9896	1.0000	0.9570	0.9106	0.8744*
XLI	1.0000	0.9997	0.9936	0.9996	0.9747	0.9189	0.8733*
XLP	1.0000	0.9906	1.0018	1.0000	0.9813	0.9468	0.8867*
XLU	1.0000	1.0019	0.9896	0.9997	0.9410	0.9247	0.8879*
XLV	1.0000	0.9981	0.9976	0.9998	0.9665	0.9243	0.8760*
XLY	1.0000	0.9939	0.9956	0.9999	0.9745	0.9152	0.8706*
Panel (b): Weekly							
IYR	1.0000	1.0029	0.8418	0.9993		0.7094	0.6161*
IYZ	1.0000	1.0064	0.8710	0.9993		0.6871	0.6107*
XLB	1.0000	0.9992	0.8431	0.9991		0.6903	0.6019*
XLE	1.0000	0.9996	0.8825	0.9996		0.7222	0.6319*
XLF	1.0000	0.9994	0.8547	0.9999		0.6823	0.5994*
XLI	1.0000	1.0020	0.8308	0.9992		0.6886	0.5866*
XLP	1.0000	0.9921	0.8108	0.9997		0.7570	0.6200*
XLU	1.0000	1.0026	0.8728	0.9993		0.7045	0.6225*
XLV	1.0000	1.0009	0.8372	0.9995		0.7032	0.5999*
XLY	1.0000	0.9999	0.8303	0.9995		0.7041	0.5983*
Panel (c): Monthly							
IYR	1.0000	1.0020	0.6874	0.9983		0.5277	0.3529*
IYZ	1.0000	1.0013	0.6640	0.9976		0.5188	0.3389*
XLB	1.0000	1.0008	0.7020	0.9972		0.5020	0.3334*
XLE	1.0000	1.0001	0.7358	0.9980		0.5096	0.3529*
XLF	1.0000	0.9997	0.6969	0.9987		0.4889	0.3300*
XLI	1.0000	1.0004	0.6510	0.9973		0.5142	0.3293*
XLP	1.0000	0.9954	0.5908	0.9984		0.6236	0.3823*
XLU	1.0000	0.9999	0.6857	0.9979		0.5282	0.3586*
XLV	1.0000	1.0017	0.6471	0.9980		0.5330	0.3451*
XLY	1.0000	0.9989	0.6609	0.9980		0.5260	0.3434*

This table presents the out-of-sample relative loss measure HMAE over three forecasting horizons: daily, weekly, and monthly, for models estimated on the sector ETFs' realised variance over a rolling window of 1,000 observations for the period January 3, 2006, to September 8, 2023. The columns represent different models (HARL, HARLQ, HARSL, TVCHAR, SHARK, SHARP, SHARPSV), and the rows correspond to various sector ETFs. An asterisk (*) indicates that the model is included in the $\hat{M}_{75\%}$ model confidence set.

TABLE A.8. Out-of-Sample Relative Loss Measure QLIKE using the Sectors ETFs dataset

	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSV
Panel (a): Daily							
IYR	1.0000	1.0051	0.9822	0.9996	0.9644	0.9074	0.8681*
IYZ	1.0000	1.0171	0.9919	0.9997	0.9656	0.9063	0.8591*
XLB	1.0000	0.9983	0.9857	0.9997	0.9571	0.9119	0.8705*
XLE	1.0000	0.9910	0.9948	0.9999	0.9613	0.9332	0.8854*
XLF	1.0000	1.0033	0.9859	0.9999	0.9458	0.8993	0.8599*
XLI	1.0000	1.0010	0.9916	0.9996	0.9722	0.9057	0.8541*
XLP	1.0000	0.9859	1.0022	1.0001	0.9831	0.9359	0.8711*
XLU	1.0000	1.0053	0.9854	0.9994	0.9381	0.9135	0.8733*
XLV	1.0000	0.9982	0.9959	0.9998	0.9616	0.9080	0.8563*
XLY	1.0000	0.9916	0.9944	0.9999	0.9715	0.9010	0.8523*
Panel (b): Weekly							
IYR	1.0000	1.0055	0.8015	0.9988		0.5580	0.4347*
IYZ	1.0000	1.0086	0.8113	0.9989		0.5176	0.4366*
XLB	1.0000	0.9991	0.7793	0.9987		0.5284	0.4136*
XLE	1.0000	0.9969	0.8590	0.9990		0.5649	0.4416*
XLF	1.0000	1.0004	0.7915	0.9994		0.5272	0.4144*
XLI	1.0000	1.0040	0.8056	0.9987		0.5344	0.4047*
XLP	1.0000	0.9880	0.7611	0.9993		0.6101	0.4304*
XLU	1.0000	1.0048	0.8290	0.9985		0.5441	0.4317*
XLV	1.0000	1.0036	0.7651	0.9992		0.5390	0.4063*
XLY	1.0000	1.0011	0.7780	0.9992		0.5490	0.4201*
Panel (c): Monthly							
IYR	1.0000	1.0043	0.5826	0.9974		0.3364	0.1556*
IYZ	1.0000	1.0027	0.5491	0.9964		0.3266	0.1460*
XLB	1.0000	1.0011	0.5927	0.9962		0.3075	0.1398*
XLE	1.0000	1.0012	0.6434	0.9965		0.2989	0.1448*
XLF	1.0000	1.0011	0.5797	0.9974		0.2963	0.1376*
XLI	1.0000	1.0008	0.5567	0.9962		0.3232	0.1417*
XLP	1.0000	0.9931	0.4847	0.9976		0.4373	0.1817*
XLU	1.0000	1.0014	0.5656	0.9963		0.3366	0.1592*
XLV	1.0000	1.0028	0.5217	0.9972		0.3393	0.1487*
XLY	1.0000	0.9988	0.5644	0.9969		0.3238	0.1466*

This table presents the out-of-sample relative loss measure QLIKE over three forecasting horizons: daily, weekly, and monthly, for models estimated on the sector ETFs' realised variance over a rolling window of 1,000 observations for the period January 3, 2006, to September 8, 2023. The columns represent different models (HARL, HARLQ, HARSL, TVCHAR, SHARK, SHARP, SHARPSV), and the rows correspond to various sector ETFs. An asterisk (*) indicates that the model is included in the $\hat{M}_{75\%}$ model confidence set.

TABLE A.9. Out-of-Sample Relative Loss Measure MSE using the NYSE individual stocks dataset

	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSv
Panel (a): Daily							
AAPL	1.0000	1.0036	1.2112	0.9986*	0.9881*	1.0107	1.0107
AEE	1.0000	1.2857	1.6152	0.9995*	1.0211	1.0112	1.0112
BFB	1.0000	0.9973*	1.0084	0.9999	1.0054	1.0069	1.0075
BT	1.0000	0.9754*	1.0818	0.9989*	0.9819*	0.9946*	1.0115
EXC	1.0000*	1.2609	1.5486	0.9981*	1.0182*	1.0029*	1.0190
FCX	1.0000*	1.0400	1.0903	0.9996*	0.9966*	1.0170	1.0710
GD	1.0000*	0.9949*	1.1169	0.9992*	1.0050*	1.0012*	1.0147
GE	1.0000	1.1094	1.3427	0.9994*	1.0258	1.0115	1.0350
HD	1.0000*	1.2017	1.1631	0.9998*	0.9919*	1.0127	1.0144
HES	1.0000	1.3180	1.0918	0.9999	0.9539*	1.0398	1.0678
HUM	1.0000*	1.0075*	1.0340	0.9994*	0.9917*	1.0112	1.0120
IBM	1.0000*	1.2177	1.2112	0.9997*	1.0035*	1.0296	1.0351
KO	1.0000	0.9837*	1.0636	0.9993	0.9958	1.0127	1.0223
MAR	1.0000*	1.0157*	1.0808	1.0000*	0.9991*	1.0056*	1.0264
NUE	1.0000	1.0756	1.1920	0.9997	0.9148*	1.0091	1.0197
PFE	1.0000	1.1678	1.1812	0.9999	0.9712*	1.0138	1.0063
T	1.0000	1.5255	1.3692	0.9993	0.9559*	1.0082	1.0103
TRV	1.0000*	1.1448	1.5015	1.0000*	1.0517	1.0144	1.0330
WFC	1.0000	1.1073	1.2108	0.9998*	1.0395	1.0535	1.2002
XOM	1.0000	1.5081	1.5816	0.9995	0.9672*	1.0153	1.0208
Panel (b): Weekly							
AAPL	1.0000	0.9936	4.3356	0.9941		0.7829	0.5483*
AEE	1.0000	0.9602	1.0227	0.9991		0.8260	0.6612*
BFB	1.0000	0.9191	1.1888	0.9999		0.6486	0.6070*
BT	1.0000	0.9701	1.3175	0.9987		0.4584*	0.4871
EXC	1.0000	1.0164	0.9389	0.9979		0.4882*	0.5155
FCX	1.0000	0.9880	1.1780	0.9995		0.5705*	0.5617*
GD	1.0000	0.9674	1.5742	0.9988		0.5253	0.5061*
GE	1.0000	1.0395	2.4967	1.0005		0.7039	0.6795*
HD	1.0000	0.9551	1.1863	1.0004		0.6766	0.5780*
HES	1.0000	1.0256	0.5791	0.9997		0.4983*	0.5989
HUM	1.0000	1.0126	0.9346	1.0009		0.6845	0.7962*
IBM	1.0000	0.9888	1.3576	0.9990		0.6351	0.5727*
KO	1.0000	0.9602	1.0815	0.9993		0.6560	0.5365*
MAR	1.0000	1.0102	0.9034	1.0002		0.4774*	0.5378
NUE	1.0000	0.9566	0.6566	0.9964		0.5088*	0.5802
PFE	1.0000	0.9738	1.5651	0.9957		0.5202	0.5016*
T	1.0000	0.9985	0.8094	0.9997		0.5099*	0.5482
TRV	1.0000	0.9580	1.9222	1.0013		0.6231*	0.6657
WFC	1.0000	0.9744	1.0440	1.0010		0.6515*	0.7773
XOM	1.0000	1.0599	1.9201	0.9982		0.6178	0.5436*
Panel (c): Monthly							
AAPL	1.0000	0.9991	0.4168	0.9953		0.3238	0.2915*
AEE	1.0000	0.9816	1.7673	0.9992		0.3599*	0.3022
BFB	1.0000	0.9727	0.8989	0.9996		0.2516*	0.2651
BT	1.0000	1.0114	0.7191	0.9998		0.1815*	0.2245
EXC	1.0000	1.0109	0.7189	0.9997		0.2128*	0.2551
FCX	1.0000	1.0109	0.3414	1.0015		0.2019*	0.3042
GD	1.0000	1.0173	0.7697	1.0003		0.2072*	0.2523
GE	1.0000	0.9928	0.5624	1.0027		0.2060*	0.3020
HD	1.0000	0.9943	0.5853	1.0004		0.1927*	0.2651
HES	1.0000	0.9696	0.2312	1.0055		0.1533*	0.3004
HUM	1.0000	0.9906	0.8990	0.9949		0.1800	0.2581*
IBM	1.0000	1.0222	0.6825	0.9991		0.3099	0.2799*
KO	1.0000	0.9870	1.4590	1.0020		0.3139*	0.2604
MAR	1.0000	1.0110	0.3753*	1.0031		0.1805*	0.2860
NUE	1.0000	0.9732	0.2174	0.9975		0.2074*	0.2746
PFE	1.0000	1.0023	0.9399	0.9958		0.1986*	0.2391*
T	1.0000	1.0279	0.7504	0.9989		0.2814*	0.2812
TRV	1.0000	0.9473	2.3932	1.0041		0.2598*	0.3547
WFC	1.0000	0.9795	0.3168	0.9982		0.1758	0.3576*
XOM	1.0000	1.0034	0.8601	0.9980		0.2713	0.2615

This table presents the out-of-sample relative loss measure MSE over three forecasting horizons: daily, weekly, and monthly, for models estimated on the twenty NYSE stocks realised variance over a rolling window of 1,000 observations for the period January 3, 2000, to December 31, 2016. The columns represent different models (HARL, HARLQ, HARSL, TVCHAR, SHARK, SHARP, SHARPSV), and the rows correspond to various individual stocks. An asterisk (*) indicates that the model is included in the $\hat{M}_{75\%}$ model confidence set.

TABLE A.10. Out-of-Sample Relative Loss Measure MAE using the NYSE individual stocks dataset

	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSv
Panel (a): Daily							
AAPL	1.0000	1.0031	1.0435	0.9999	1.0240	0.9574*	0.9590
AEE	1.0000	1.0314	1.0681	1.0005	1.0312	0.9791	0.9770*
BFB	1.0000	0.9879	1.0105	1.0001	1.0436	0.9602	0.9571*
BT	1.0000	0.9999	1.0101	0.9998	1.0144	0.9598*	0.9675
EXC	1.0000	1.0265	1.0634	1.0001	1.0375	0.9682*	0.9765
FCX	1.0000	0.9923	1.0341	1.0004	1.0259	0.9870*	1.0225
GD	1.0000	0.9892	1.0341	0.9999	1.0135	0.9619*	0.9633
GE	1.0000	1.0276	1.1006	1.0005	1.0556	0.9904*	1.0225
HD	1.0000	1.0146	1.0221	1.0001	1.0153	0.9706*	0.9742
HES	1.0000*	1.0170*	1.0086*	1.0005*	1.0166*	0.9949*	1.0152*
HUM	1.0000	0.9894	1.0066	1.0004	1.0166	0.9588*	0.9629
IBM	1.0000	1.0202	1.0343	1.0004	1.0221	0.9736*	0.9802
KO	1.0000	0.9950	1.0211	1.0004	1.0031	0.9678*	0.9726
MAR	1.0000	0.9925	1.0064	1.0003	1.0130	0.9668*	0.9770
NUE	1.0000	1.0117	1.0310	1.0006	1.0081	0.9834*	0.9997
PFE	1.0000	1.0100	1.0204	1.0002	1.0114	0.9680	0.9665*
T	1.0000	1.0284	1.0459	1.0001	1.0207	0.9745*	0.9771
TRV	1.0000	1.0072	1.0968	1.0007	1.0591	0.9732*	0.9889
WFC	1.0000*	1.0117	1.0599	1.0002*	1.0342	1.0028*	1.0782
XOM	1.0000	1.0405	1.0567	1.0002	1.0431	0.9800*	0.9849
Panel (b): Weekly							
AAPL	1.0000	0.9988	0.9554	0.9987		0.6760*	0.6835
AEE	1.0000	0.9854	0.9159	0.9998		0.6896*	0.7013
BFB	1.0000	0.9760	0.9702	1.0000		0.6838*	0.6902
BT	1.0000	0.9942	0.8850	0.9991		0.6274*	0.6587
EXC	1.0000	0.9998	0.8751	0.9990		0.6612*	0.6905
FCX	1.0000	0.9901	0.7366	1.0003		0.6772*	0.7253
GD	1.0000	0.9926	0.9553	0.9992		0.6595*	0.6736
GE	1.0000	1.0090	0.9745	0.9999		0.6959*	0.7214
HD	1.0000	0.9845	0.8893	0.9991		0.6640*	0.6868
HES	1.0000	0.9993	0.7517	1.0001		0.6624*	0.7171
HUM	1.0000	1.0023	0.8534	0.9999		0.7085	0.7057*
IBM	1.0000	0.9924	0.8956	0.9995		0.6697*	0.6821
KO	1.0000	0.9883	0.9080	1.0001		0.6562*	0.6607
MAR	1.0000	0.9914	0.8427	0.9998		0.6436*	0.6858
NUE	1.0000	0.9973	0.7295	0.9994		0.6749*	0.7256
PFE	1.0000	0.9956	0.9625	0.9988		0.6475*	0.6579
T	1.0000	0.9842	0.9006	0.9997		0.6709*	0.6946
TRV	1.0000	0.9851	0.9536	1.0003		0.6672*	0.6950
WFC	1.0000	0.9986	0.8090	1.0002		0.6834*	0.7496
XOM	1.0000	1.0099	0.9936	0.9994		0.6938*	0.6915*
Panel (c): Monthly							
APL	1.0000	1.0012	0.6104	0.9977		0.4084	0.4673*
AEE	1.0000	0.9915	0.8143	0.9986		0.4631	0.4135*
BFB	1.0000	0.9889	0.8114	0.9989		0.4564*	0.4241*
BT	1.0000	0.9994	0.7257	0.9973		0.4015*	0.3996
EXC	1.0000	1.0029	0.6730	0.9997		0.4099*	0.4230
FCX	1.0000	1.0011	0.4107	1.0002		0.3683	0.5125*
GD	1.0000	1.0068	0.7564	0.9979		0.4168*	0.3947
GE	1.0000	0.9947	0.6020	0.9996		0.3893*	0.4389
HD	1.0000	0.9970	0.6682	0.9975		0.4100*	0.4387
HES	1.0000	0.9894	0.4828	1.0015		0.3794*	0.4954
HUM	1.0000	0.9894	0.6308	0.9985		0.3854	0.4554*
IBM	1.0000	1.0038	0.7101	0.9990		0.4474	0.4176*
KO	1.0000	0.9952	0.7864	1.0000		0.4431*	0.3995
MAR	1.0000	0.9983	0.5998	0.9997		0.4070*	0.4628
NUE	1.0000	0.9968	0.4504	0.9994		0.3870	0.4611*
PFE	1.0000	1.0001	0.7722	0.9985		0.4112	0.3910*
T	1.0000	1.0116	0.7204	0.9983		0.4224*	0.4115
TRV	1.0000	0.9781	0.7905	1.0002		0.4423*	0.4616
WFC	1.0000	0.9973	0.4758	0.9983		0.3708	0.5083*
XOM	1.0000	1.0011	0.7775	0.9984		0.4539	0.3970

This table presents the out-of-sample relative loss measure MAE over three forecasting horizons: daily, weekly, and monthly, for models estimated on the twenty NYSE stocks realised variance over a rolling window of 1,000 observations for the period January 3, 2000, to December 31, 2016. The columns represent different models (HARL, HARLQ, HARSL, TVCHAR, SHARK, SHARP, SHARPSV), and the rows correspond to various individual stocks. An asterisk (*) indicates that the model is included in the $\hat{M}_{75\%}$ model confidence set.

TABLE A.11. Out-of-Sample Relative Loss Measure HMSE using the NYSE individual stocks dataset

	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSV
Panel (a): Daily							
AAPL	1.0000	1.0035	1.0177	0.9987	1.0582	0.7012*	0.7023*
AEE	1.0000	0.9826	1.0634	0.9996	1.0074	0.7861*	0.7861*
BFB	1.0000	0.9271	1.0091	1.0001	1.1000	0.7338	0.7269*
BT	1.0000	0.9953	1.0206	0.9984	1.0019	0.7789*	0.7909
EXC	1.0000	0.9706	1.0826	1.0001	1.0327	0.7518*	0.7535*
FCX	1.0000	0.9679	1.0076	0.9995	1.0293	0.7897*	0.8416
GD	1.0000	0.9764	1.0143	0.9984	1.0178	0.7488*	0.7501
GE	1.0000	0.9941	1.0463	0.9987	1.0245	0.7733*	0.8065
HD	1.0000	1.0068	1.0056	0.9987	1.0002	0.7772*	0.7893
HES	1.0000	0.9380	1.0232	1.0000	1.0332	0.7974*	0.8213
HUM	1.0000	0.9469	1.0001	0.9989	1.0202	0.6911*	0.6924*
IBM	1.0000	1.0204	0.9853	0.9994	1.0013	0.7693*	0.7790
KO	1.0000	1.0018	1.0482	0.9995	0.9845	0.7659*	0.7777
MAR	1.0000	0.9596	1.0066	0.9990	1.0093	0.7516*	0.7583
NUE	1.0000	0.9734	1.0015	0.9994	1.0227	0.7813*	0.8043
PFE	1.0000	0.9785	1.0611	1.0000	1.0105	0.7806	0.7736*
T	1.0000	0.9792	1.0119	0.9987	1.0564	0.7687*	0.7967
TRV	1.0000	0.9757	1.0364	1.0001	1.0515	0.7359*	0.7471
WFC	1.0000	0.9431	0.9736	0.9964	0.9651	0.7688*	0.8016
XOM	1.0000	1.0069	0.9880	0.9990	1.0287	0.7833*	0.7970
Panel (b): Weekly							
AAPL	1.0000	1.0055	2.1047	0.9964		0.3876	0.3501*
AEE	1.0000	0.9784	1.3924	0.9989		0.5704	0.4651*
BFB	1.0000	0.9122	1.2190	1.0000		0.5524	0.5259*
BT	1.0000	1.0020	1.1675	0.9966		0.3994	0.3930*
EXC	1.0000	0.9761	0.9910	0.9986		0.4739	0.4316*
FCX	1.0000	0.9758	0.6892	0.9989		0.4075*	0.4510
GD	1.0000	0.9907	1.3790	0.9970		0.3948	0.3618*
GE	1.0000	1.0065	1.8214	0.9971		0.4828	0.4041*
HD	1.0000	0.9857	0.8359	0.9963		0.6037	0.3848*
HES	1.0000	0.9609	0.9796	0.9986		0.4677*	0.4678*
HUM	1.0000	1.0030	0.4975	0.9967		0.2882	0.3595*
IBM	1.0000	1.0103	0.8138	0.9985		0.3764	0.3641*
KO	1.0000	1.0040	0.8541	0.9981		0.3644*	0.3645*
MAR	1.0000	0.9652	1.1494	0.9971		0.3919	0.3863*
NUE	1.0000	0.9784	0.7545	0.9976		0.4219*	0.4419
PFE	1.0000	0.9917	1.1545	0.9980		0.4506	0.3835*
T	1.0000	0.9973	0.8504	0.9973		0.3479*	0.3529
TRV	1.0000	0.9775	0.8621	0.9979		0.3624*	0.3629*
WFC	1.0000	0.9881	0.7160	0.9966		0.3842*	0.3877*
XOM	1.0000	0.9960	3.1356	0.9978		0.5060	0.3967*
Panel (c): Monthly							
AAPL	1.0000	1.0061	0.6012	0.9924		0.1424	0.1301*
AEE	1.0000	0.9748	0.7351	0.9953		0.1890	0.1337*
BFB	1.0000	0.9708	0.7813	0.9963		0.1925	0.1534*
BT	1.0000	1.0122	0.7818	0.9923		0.1660	0.1428*
EXC	1.0000	1.0037	0.5507	0.9939		0.1675*	0.1394
FCX	1.0000	0.9954	0.2497	0.9945		0.1160	0.1948*
GD	1.0000	1.0019	0.8288	0.9902		0.1571	0.1120*
GE	1.0000	1.0011	0.5337	0.9924		0.1518	0.1231*
HD	1.0000	0.9994	0.6286	0.9927		0.1477*	0.1319
HES	1.0000	0.9865	0.3992	0.9956		0.1494*	0.1989
HUM	1.0000	0.9723	0.6700	0.9924		0.1508	0.1674*
IBM	1.0000	1.0090	0.6204	0.9931		0.1661	0.1127*
KO	1.0000	0.9910	0.6343	0.9954		0.1561	0.1162*
MAR	1.0000	0.9731	0.5389	0.9920		0.1448*	0.1348
NUE	1.0000	0.9892	0.3769	0.9932		0.1300	0.1618*
PFE	1.0000	0.9987	0.7326	0.9957		0.1537	0.1195*
T	1.0000	1.0023	0.6711	0.9929		0.1517	0.1161*
TRV	1.0000	0.9931	0.6560	0.9943		0.1556*	0.1174
WFC	1.0000	1.0024	0.3916	0.9933		0.1261	0.1540*
XOM	1.0000	0.9989	0.7542	0.9930		0.1720	0.1098

This table presents the out-of-sample relative loss measure HMSE over three forecasting horizons: daily, weekly, and monthly, for models estimated on the twenty NYSE stocks realised variance over a rolling window of 1,000 observations for the period January 3, 2000, to December 31, 2016. The columns represent different models (HARL, HARLQ, HARSL, TVCHAR, SHARK, SHARP, SHARPSV), and the rows correspond to various individual stocks. An asterisk (*) indicates that the model is included in the $\hat{M}_{75\%}$ model confidence set.

TABLE A.12. Out-of-Sample Relative Loss Measure HMAE using the NYSE individual stocks dataset

	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSv
Panel (a): Daily							
AAPL	1.0000	1.0019	1.0036	0.9993	1.0217	0.8466*	0.8482
AEE	1.0000	0.9970	1.0011	0.9999	1.0039	0.8913*	0.8928
BFB	1.0000	0.9823	1.0015	0.9999	1.0427	0.8546	0.8482*
BT	1.0000	1.0009	0.9984	0.9992	1.0004	0.8899*	0.8981
EXC	1.0000	0.9934	0.9942	0.9999	1.0076	0.8744*	0.8746*
FCX	1.0000	0.9873	1.0061	0.9997	1.0140	0.8934*	0.9183
GD	1.0000	0.9907	1.0031	0.9992	1.0058	0.8700*	0.8709
GE	1.0000	1.0011	1.0008	0.9993	1.0051	0.8842*	0.8984
HD	1.0000	0.9975	0.9972	0.9993	0.9947	0.8878*	0.8937
HES	1.0000	0.9778	1.0011	0.9999	1.0165	0.8993*	0.9113
HUM	1.0000	0.9798	0.9983	0.9997	1.0103	0.8296*	0.8295*
IBM	1.0000	1.0036	0.9925	0.9998	0.9979	0.8867*	0.8922
KO	1.0000	0.9982	1.0003	0.9997	0.9929	0.8777*	0.8864
MAR	1.0000	0.9866	1.0020	0.9996	1.0054	0.8794*	0.8835
NUE	1.0000	0.9894	0.9993	0.9997	1.0084	0.8883*	0.8995
PFE	1.0000	0.9952	1.0029	0.9997	1.0028	0.8811*	0.8834
T	1.0000	0.9938	1.0004	0.9993	1.0196	0.8791*	0.8864
TRV	1.0000	0.9909	1.0055	0.9997	1.0024	0.8597*	0.8654
WFC	1.0000	0.9966	0.9952	0.9991	0.9997	0.8793*	0.8986
XOM	1.0000	1.0014	0.9940	0.9995	1.0068	0.8982*	0.9042
Panel (b): Weekly							
AAPL	1.0000	1.0008	0.8641	0.9983		0.5492*	0.5721
AEE	1.0000	0.9954	0.9153	0.9994		0.6274*	0.6339
BFB	1.0000	0.9789	0.9554	0.9996		0.6121*	0.6170
BT	1.0000	1.0007	0.8799	0.9983		0.5918*	0.6095
EXC	1.0000	0.9932	0.8957	0.9990		0.6152*	0.6305
FCX	1.0000	0.9898	0.6762	0.9991		0.6119*	0.6684
GD	1.0000	0.9966	0.9286	0.9983		0.5837*	0.5918
GE	1.0000	1.0034	0.8676	0.9983		0.5912*	0.6043
HD	1.0000	0.9940	0.8900	0.9982		0.5982*	0.6077
HES	1.0000	0.9890	0.7956	0.9993		0.6208*	0.6599
HUM	1.0000	1.0010	0.6928	0.9988		0.5671*	0.5520
IBM	1.0000	0.9999	0.8766	0.9990		0.6088*	0.6126
KO	1.0000	0.9994	0.8724	0.9990		0.5867*	0.5962
MAR	1.0000	0.9902	0.8863	0.9987		0.5885*	0.6107
NUE	1.0000	0.9884	0.7548	0.9988		0.6090*	0.6516
PFE	1.0000	0.9979	0.9329	0.9989		0.5883*	0.5918
T	1.0000	0.9957	0.8977	0.9986		0.5826*	0.5949
TRV	1.0000	0.9922	0.8914	0.9987		0.5819*	0.5891
WFC	1.0000	0.9979	0.8037	0.9983		0.5824*	0.6060
XOM	1.0000	1.0001	0.9141	0.9988		0.6075*	0.6084*
Panel (c): Monthly							
AAPL	1.0000	1.0025	0.6759	0.9964		0.3688	0.3654*
AEE	1.0000	0.9961	0.7913	0.9981		0.4247	0.3533*
BFB	1.0000	0.9905	0.8450	0.9981		0.4217	0.3604*
BT	1.0000	1.0026	0.7715	0.9954		0.3889	0.3573*
EXC	1.0000	0.9999	0.7164	0.9974		0.4066*	0.3598
FCX	1.0000	0.9971	0.4216	0.9967		0.3322	0.4456*
GD	1.0000	1.0003	0.8034	0.9948		0.3890	0.3215*
GE	1.0000	1.0004	0.6609	0.9960		0.3782	0.3339*
HD	1.0000	0.9989	0.7403	0.9957		0.3830*	0.3521
HES	1.0000	0.9960	0.5639	0.9974		0.3754*	0.4343
HUM	1.0000	0.9887	0.6561	0.9970		0.3659	0.3964*
IBM	1.0000	1.0024	0.7404	0.9967		0.4082	0.3389*
KO	1.0000	0.9978	0.7347	0.9973		0.3905	0.3375*
MAR	1.0000	0.9929	0.6942	0.9964		0.3894*	0.3694
NUE	1.0000	0.9969	0.5309	0.9967		0.3546	0.3979*
PFE	1.0000	0.9992	0.7948	0.9978		0.3845	0.3273*
T	1.0000	0.9990	0.7555	0.9960		0.3830	0.3258*
TRV	1.0000	0.9957	0.7535	0.9964		0.3897*	0.3294
WFC	1.0000	1.0015	0.5826	0.9957		0.3590	0.3715*
XOM	1.0000	1.0010	0.7605	0.9966		0.4074	0.3271

This table presents the out-of-sample relative loss measure HMAE over three forecasting horizons: daily, weekly, and monthly, for models estimated on the twenty NYSE stocks realised variance over a rolling window of 1,000 observations for the period January 3, 2000, to December 31, 2016. The columns represent different models (HARL, HARLQ, HARSL, TVCHAR, SHARK, SHARP, SHARPSV), and the rows correspond to various individual stocks. An asterisk (*) indicates that the model is included in the $\hat{M}_{75\%}$ model confidence set.

TABLE A.13. Out-of-Sample Relative Loss Measure QLIKE using the NYSE individual stocks dataset

	HARL	HARLQ	HARSL	TVCHAR	SHARK	SHARP	SHARPSv
Panel (a): Daily							
AAPL	1.0000	1.0015	1.0065	0.9991	1.0298	0.8238*	0.8262
AEE	1.0000	0.9935	1.0114	0.9996	1.0053	0.8838*	0.8837*
BFB	1.0000	0.9675	1.0029	0.9998	1.0531	0.8458	0.8417*
BT	1.0000	0.9998	1.0012	0.9989	0.9982	0.8747*	0.8830
EXC	1.0000	0.9865	1.0013	0.9996	1.0080	0.8620*	0.8628*
FCX	1.0000	0.9790	1.0070	0.9996	1.0162	0.8827*	0.9112
GD	1.0000	0.9848	1.0049	0.9989	1.0071	0.8570*	0.8579
GE	1.0000	0.9984	1.0066	0.9990	1.0066	0.8690*	0.8854
HD	1.0000	0.9965	0.9988	0.9991	0.9950	0.8751*	0.8811
HES	1.0000	0.9617	1.0068	0.9998	1.0187	0.8885*	0.9014
HUM	1.0000	0.9689	0.9984	0.9995	1.0114	0.8236*	0.8245
IBM	1.0000	1.0081	0.9906	0.9995	0.9977	0.8734*	0.8790
KO	1.0000	0.9962	1.0056	0.9996	0.9911	0.8681*	0.8759
MAR	1.0000	0.9755	1.0028	0.9993	1.0045	0.8625*	0.8674
NUE	1.0000	0.9784	0.9979	0.9996	1.0103	0.8813*	0.8925
PFE	1.0000	0.9906	1.0114	0.9996	1.0033	0.8703*	0.8700*
T	1.0000	0.9879	1.0028	0.9991	1.0251	0.8667*	0.8764
TRV	1.0000	0.9852	1.0114	0.9998	1.0112	0.8436*	0.8494
WFC	1.0000	0.9887	0.9904	0.9985	0.9954	0.8700*	0.8948
XOM	1.0000	1.0031	0.9907	0.9993	1.0114	0.8811*	0.8873
Panel (b): Weekly							
AAPL	1.0000	0.9997	0.8926	0.9975		0.3819*	0.3910
AEE	1.0000	0.9882	0.9152	0.9988		0.4552	0.4490*
BFB	1.0000	0.9587	0.9636	0.9994		0.4577*	0.4573*
BT	1.0000	1.0008	0.8739	0.9975		0.3959*	0.4099
EXC	1.0000	0.9846	0.8613	0.9984		0.4318*	0.4432
FCX	1.0000	0.9820	0.5475	0.9988		0.4163*	0.4762
GD	1.0000	0.9933	0.9378	0.9975		0.3997*	0.3976*
GE	1.0000	1.0039	0.8464	0.9976		0.4204*	0.4220*
HD	1.0000	0.9898	0.8201	0.9971		0.4284	0.4155*
HES	1.0000	0.9764	0.7263	0.9988		0.4409*	0.4747
HUM	1.0000	1.0028	0.5887	0.9980		0.3821	0.3772*
IBM	1.0000	1.0021	0.8178	0.9985		0.4120*	0.4144
KO	1.0000	0.9988	0.8248	0.9984		0.3961*	0.4043
MAR	1.0000	0.9777	0.8579	0.9978		0.4008*	0.4162
NUE	1.0000	0.9776	0.6398	0.9981		0.4249*	0.4642
PFE	1.0000	0.9939	0.9181	0.9981		0.4108	0.4034*
T	1.0000	0.9932	0.8365	0.9980		0.3882*	0.4005
TRV	1.0000	0.9824	0.8286	0.9981		0.3946*	0.4002
WFC	1.0000	0.9943	0.7088	0.9973		0.4029*	0.4244
XOM	1.0000	0.9990	0.9495	0.9983		0.4275	0.4145*
Panel (c): Monthly							
AAPL	1.0000	1.0018	0.5713	0.9949		0.1711	0.1616*
AEE	1.0000	0.9883	0.6945	0.9960		0.2095	0.1468*
BFB	1.0000	0.9832	0.7684	0.9966		0.2112	0.1597*
BT	1.0000	1.0061	0.6891	0.9940		0.1747	0.1491*
EXC	1.0000	1.0008	0.5700	0.9951		0.1907*	0.1569
FCX	1.0000	0.9951	0.2442	0.9956		0.1301	0.2168*
GD	1.0000	1.0003	0.7322	0.9923		0.1811	0.1280*
GE	1.0000	1.0008	0.5201	0.9939		0.1770	0.1417*
HD	1.0000	0.9998	0.6004	0.9935		0.1707*	0.1517
HES	1.0000	0.9924	0.3856	0.9960		0.1596*	0.2065
HUM	1.0000	0.9823	0.5192	0.9938		0.1649	0.1845*
IBM	1.0000	1.0041	0.6247	0.9943		0.1921	0.1330*
KO	1.0000	0.9956	0.6255	0.9959		0.1788	0.1334*
MAR	1.0000	0.9835	0.5588	0.9938		0.1741*	0.1568
NUE	1.0000	0.9875	0.3483	0.9949		0.1441	0.1799*
PFE	1.0000	0.9986	0.6958	0.9959		0.1748	0.1337*
T	1.0000	0.9963	0.6395	0.9943		0.1770	0.1354*
TRV	1.0000	0.9928	0.6328	0.9949		0.1866*	0.1377
WFC	1.0000	1.0013	0.4198	0.9941		0.1591	0.1781*
XOM	1.0000	1.0007	0.6717	0.9951		0.1988	0.1288

This table presents the out-of-sample relative loss measure QLIKE over three forecasting horizons: daily, weekly, and monthly, for models estimated on the twenty NYSE stocks realised variance over a rolling window of 1,000 observations for the period January 3, 2000, to December 31, 2016. The columns represent different models (HARL, HARLQ, HARSL, TVCHAR, SHARK, SHARP, SHARPSV), and the rows correspond to various individual stocks. An asterisk (*) indicates that the model is included in the $\hat{M}_{75\%}$ model confidence set.

A.1. Misspecification analysis. Our specification (3.1)-(3.5) is based on assuming (*i.i.d.*) Gaussian innovations throughout. A stylised fact of Bayesian inference is that, typically, coefficient inference is robust to such an assumption; on the other hand, however, inference on the volatility process can be affected if the assumption of Gaussianity is not satisfied by the innovations, with posterior inference on volatility becoming invalid.¹ Whilst results in our empirical application seem to suggest some degree of robustness, in this section - by way of misspecification analysis - we report a test for the normality of innovations. Our test is constructed in a similar way to the one proposed (although for a different model) in Koopman and Scharth (2012) - namely, we use a parametric bootstrap method where we compare the skewness and kurtosis of the observations y_t against the skewness and kurtosis of a pseudosample $y_{t,b}$ generated using Gaussian innovations.

More specifically, we have generated the pseudosamples $\{y_{t,b}, 1 \leq b \leq B\}$ using

$$(A.1) \quad y_{t,b} = x'_{t,b} \beta_{t,b} + v_{t,b},$$

$$(A.2) \quad \beta_{j,t,b} = \hat{\alpha}_j + \hat{\rho}_j \beta_{j,t-1,b} + \varepsilon_{j,t,b},$$

$$(A.3) \quad \varepsilon_{j,t,b} = h_{j,t,b} \eta_{j,t,b},$$

$$(A.4) \quad \ln h_{j,t,b} = \hat{\gamma}_j + \hat{\delta}_j \ln h_{j,t-1,b} + u_{j,t,b},$$

with $v_{t,b}$, $\varepsilon_{j,t,b}$, $\eta_{j,t,b}$ and $u_{j,t,b}$ all generated as zero mean, *i.i.d.* Gaussian with variance given by the corresponding estimates - similarly, $\hat{\alpha}_j$, $\hat{\rho}_j$, $\hat{\gamma}_j$ and $\hat{\delta}_j$ refer to the relevant estimates, and we initialise setting $\beta_{j,0,b}$ equal to the full sample OLS estimator. The bootstrap scheme constructed above lends itself to constructing an omnibus test, for the null hypothesis of the correct specification of the distribution of all the innovations $\{v_t\}$, $\{\varepsilon_{j,t,b}, 1 \leq j \leq k\}$ and $\{u_{j,t,b}, 1 \leq j \leq k\}$ simultaneously, which - along the same lines as Koopman and Scharth (2012) - we test by testing separately for the null hypotheses that (a) the skewness and (b)

¹We are grateful to an anonymous Referee for bringing this point to our attention, and also for suggesting the test by Koopman and Scharth (2012).

the kurtosis of the observations y_t are compatible with the DGP in (A.1)-(A.4). We report results in Tables A.14 and A.15 for the NYSE stocks and the ETF sectors respectively, using $B = 499$; as can be seen, the null hypothesis of Gaussian innovations is almost always rejected. This can be read in conjunction with our results in Section 5, where the impact of a non Gaussian distribution is studied (see in particular “DGP SVt”).

TABLE A.14. Bootstrap-Based Normality Test Results - Twenty NYSE Stocks Dataset

	Daily				Weekly				Monthly			
	SHARP		SHARPSv		SHARP		SHARPSv		SHARP		SHARPSv	
	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis
AAPL	0.3592 (0.0000)	3.5029 (0.0000)	0.3595 (0.0000)	3.5038 (0.0000)	0.4098 (0.0000)	3.4611 (0.0000)	0.4111 (0.0000)	3.4668 (0.0000)	0.5080 (0.0000)	3.4949 (0.0000)	0.5076 (0.0000)	3.4938 (0.0000)
AEE	1.1269 (0.0000)	5.7151 (0.0000)	1.1271 (0.0000)	5.7176 (0.0000)	1.4542 (0.0000)	6.7707 (0.0000)	1.4545 (0.0000)	6.7716 (0.0000)	1.7248 (0.0000)	7.8640 (0.0000)	1.7264 (0.0000)	7.8725 (0.0000)
BFB	1.0050 (0.0000)	5.1443 (0.0000)	1.0046 (0.0000)	5.1402 (0.0000)	1.0324 (0.0000)	4.3939 (0.0000)	1.0321 (0.0000)	4.3921 (0.0000)	1.2555 (0.0000)	4.7105 (0.0000)	1.2564 (0.0000)	4.7153 (0.0000)
BT	0.8292 (0.0000)	4.3635 (0.0000)	0.8299 (0.0000)	4.3653 (0.0000)	1.0518 (0.0000)	4.7751 (0.0000)	1.0516 (0.0000)	4.7731 (0.0000)	1.1766 (0.0000)	5.0465 (0.0000)	1.1767 (0.0000)	5.0454 (0.0000)
EXC	1.0795 (0.0000)	5.1980 (0.0000)	1.0807 (0.0000)	5.2029 (0.0000)	1.3941 (0.0000)	6.2808 (0.0000)	1.3921 (0.0000)	6.2698 (0.0000)	1.6080 (0.0000)	6.9652 (0.0000)	1.6088 (0.0000)	6.9714 (0.0000)
FCX	0.6848 (0.0000)	3.8039 (0.0000)	0.6847 (0.0000)	3.8021 (0.0000)	0.8076 (0.0000)	3.9583 (0.0000)	0.8078 (0.0000)	3.9609 (0.0000)	0.8584 (0.0000)	4.0473 (0.0000)	0.8584 (0.0000)	4.0478 (0.0000)
GD	0.9479 (0.0000)	4.8276 (0.0000)	0.9483 (0.0000)	4.8274 (0.0000)	1.2264 (0.0000)	5.5295 (0.0000)	1.2263 (0.0000)	5.5288 (0.0000)	1.5378 (0.0000)	6.3040 (0.0000)	1.5382 (0.0000)	6.3043 (0.0000)
GE	1.1534 (0.0000)	4.9968 (0.0000)	1.1538 (0.0000)	4.9985 (0.0000)	1.3655 (0.0000)	5.3101 (0.0000)	1.3662 (0.0000)	5.3141 (0.0000)	1.4688 (0.0000)	5.3583 (0.0000)	1.4694 (0.0000)	5.3617 (0.0000)
HD	0.9472 (0.0000)	4.3307 (0.0000)	0.9485 (0.0000)	4.3351 (0.0000)	1.2109 (0.0000)	4.6988 (0.0000)	1.2108 (0.0000)	4.6960 (0.0000)	1.4177 (0.0000)	5.0224 (0.0000)	1.4177 (0.0000)	5.0201 (0.0000)
HES	0.6077 (0.0000)	3.9705 (0.0000)	0.6071 (0.0000)	3.9677 (0.0000)	0.7678 (0.0000)	4.2944 (0.0000)	0.7676 (0.0000)	4.2950 (0.0000)	0.8397 (0.0000)	4.4109 (0.0000)	0.8420 (0.0000)	4.4218 (0.0000)
HUM	0.9103 (0.0000)	4.1359 (0.0000)	0.9101 (0.0000)	4.1348 (0.0000)	1.0577 (0.0000)	4.4133 (0.0000)	1.0577 (0.0000)	4.4126 (0.0000)	1.3774 (0.0000)	5.2413 (0.0000)	1.3769 (0.0000)	5.2401 (0.0000)
IBM	1.1413 (0.0000)	5.4573 (0.0000)	1.1421 (0.0000)	5.4607 (0.0000)	1.5331 (0.0000)	6.5966 (0.0000)	1.5350 (0.0000)	6.6045 (0.0000)	1.8440 (0.0000)	7.7030 (0.0000)	1.8454 (0.0000)	7.7113 (0.0000)
KO	1.0664 (0.0000)	5.6224 (0.0000)	1.0662 (0.0000)	5.6242 (0.0000)	1.3972 (0.0000)	6.7393 (0.0000)	1.3980 (0.0000)	6.7437 (0.0000)	1.7121 (0.0000)	7.8873 (0.0000)	1.7127 (0.0000)	7.8903 (0.0000)
MAR	0.7571 (0.0000)	3.4387 (0.0000)	0.7568 (0.0000)	3.4376 (0.0000)	0.9834 (0.0000)	3.7516 (0.0000)	0.9836 (0.0000)	3.7526 (0.0000)	1.1492 (0.0000)	3.9617 (0.0000)	1.1495 (0.0000)	3.9637 (0.0000)
NUE	0.7744 (0.0000)	4.1587 (0.0000)	0.7727 (0.0000)	4.1501 (0.0000)	0.9209 (0.0000)	4.5132 (0.0000)	0.9199 (0.0000)	4.5062 (0.0000)	0.9729 (0.0000)	4.5252 (0.0000)	0.9729 (0.0000)	4.5240 (0.0000)
PFE	0.8494 (0.0000)	4.5775 (0.0000)	0.8510 (0.0000)	4.5849 (0.0000)	1.0607 (0.0000)	4.9848 (0.0000)	1.0618 (0.0000)	4.9905 (0.0000)	1.3384 (0.0000)	5.7667 (0.0000)	1.3387 (0.0000)	5.7672 (0.0000)
T	1.0738 (0.0000)	5.1614 (0.0000)	1.0740 (0.0000)	5.1632 (0.0000)	1.3595 (0.0000)	6.0616 (0.0000)	1.3585 (0.0000)	6.0557 (0.0000)	1.6091 (0.0000)	6.7064 (0.0000)	1.6098 (0.0000)	6.7094 (0.0000)
TRV	1.1588 (0.0000)	4.9155 (0.0000)	1.1595 (0.0000)	4.9173 (0.0000)	1.3581 (0.0000)	5.3843 (0.0000)	1.3586 (0.0000)	5.3871 (0.0000)	1.4854 (0.0000)	5.5865 (0.0000)	1.4874 (0.0000)	5.5953 (0.0000)
WFC	1.0539 (0.0000)	4.0078 (0.0000)	1.0545 (0.0000)	4.0115 (0.0000)	1.1876 (0.0000)	4.1503 (0.0000)	1.1880 (0.0000)	4.1516 (0.0000)	1.2435 (0.0000)	4.1045 (0.0000)	1.2444 (0.0000)	4.1069 (0.0000)
XOM	0.8984 (0.0000)	4.9802 (0.0000)	0.8980 (0.0000)	4.9781 (0.0000)	1.1041 (0.0000)	5.4871 (0.0000)	1.1036 (0.0000)	5.4839 (0.0000)	1.2179 (0.0000)	5.8473 (0.0000)	1.2197 (0.0000)	5.8576 (0.0000)

The table reports the skewness, kurtosis, and bootstrap p-values for testing the normality of the predicted residuals across individual stocks under two models: SHARP and SHARP-SV.

TABLE A.15. Bootstrap-Based Normality Test Results - Ten Sectors ETFs Dataset and the VIX index

	Panel (a): Daily				Panel (b): Weekly				Panel (c): Monthly			
	SHARP		SHARPs _v		SHARP		SHARPs _v		SHARP		SHARPs _v	
	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis
IYR	0.8815 (0.0000)	4.7333 (0.0000)	0.8795 (0.0000)	4.7233 (0.0000)	1.0624 (0.1440)	5.2604 (0.0000)	1.0618 (0.6520)	5.2595 (0.0000)	1.1155 (0.0000)	5.2189 (0.0000)	1.1154 (0.7180)	5.2157 (0.0000)
IYX	0.8726 (0.0000)	5.9455 (0.0000)	0.8727 (0.0000)	5.9448 (0.0000)	0.9337 (0.0160)	5.5910 (0.0000)	0.9347 (0.7520)	5.5961 (0.0000)	0.8783 (0.0000)	5.2193 (0.0000)	0.8808 (0.0000)	5.2291 (0.0000)
XLB	0.4368 (0.0000)	3.5668 (0.6120)	0.4373 (0.0000)	3.5679 (0.5360)	0.5520 (0.0000)	3.7004 (0.0000)	0.5515 (0.0000)	3.6990 (0.0020)	0.5965 (0.0000)	3.6035 (0.0000)	0.5946 (0.0000)	3.5933 (0.0000)
XLE	0.5663 (0.0000)	3.9685 (0.0500)	0.5669 (0.0000)	3.9695 (0.0240)	0.6869 (0.0000)	4.2335 (0.0000)	0.6870 (0.0000)	4.2336 (0.0000)	0.6071 (0.0000)	3.9528 (0.0000)	0.6061 (0.0000)	3.9449 (0.0000)
XLF	0.5012 (0.0460)	3.4568 (0.0000)	0.5019 (0.0180)	3.4574 (0.0000)	0.5721 (0.0000)	3.4147 (0.0000)	0.5722 (0.0000)	3.4151 (0.0000)	0.5485 (0.0000)	3.1686 (0.0040)	0.5494 (0.0000)	3.1707 (0.0000)
XLI	0.7897 (0.0000)	4.7142 (0.0000)	0.7889 (0.0000)	4.7096 (0.0000)	0.9408 (0.0000)	4.8088 (0.0000)	0.9411 (0.0000)	4.8097 (0.0000)	0.9948 (0.0000)	4.7314 (0.0000)	0.9940 (0.0000)	4.7311 (0.0000)
XLP	0.6311 (0.0000)	4.2179 (0.0000)	0.6314 (0.0000)	4.2197 (0.0000)	0.7551 (0.0000)	4.4299 (0.0000)	0.7564 (0.0000)	4.4352 (0.0000)	0.7720 (0.0000)	4.3059 (0.0000)	0.7756 (0.0000)	4.3269 (0.8560)
XLU	1.0861 (0.0000)	6.1167 (0.0000)	1.0881 (0.0000)	6.1312 (0.0000)	1.3057 (0.0000)	6.9122 (0.0000)	1.3037 (0.0000)	6.8988 (0.0000)	1.3265 (0.0000)	6.8526 (0.0000)	1.3278 (0.1700)	6.8547 (0.0000)
XLV	0.9301 (0.0000)	6.2353 (0.0000)	0.9324 (0.0000)	6.2492 (0.0000)	1.1953 (0.0000)	7.5702 (0.0000)	1.1962 (0.0000)	7.5764 (0.0000)	1.2117 (0.0000)	7.6673 (0.0000)	1.2049 (0.0000)	7.6106 (0.0000)
XLY	0.8189 (0.0000)	4.5379 (0.0000)	0.8190 (0.0000)	4.5380 (0.0000)	0.9637 (0.0000)	4.9407 (0.0000)	0.9635 (0.0000)	4.9389 (0.0000)	0.9440 (0.0000)	4.7792 (0.0000)	0.9419 (0.0000)	4.7609 (0.0240)
SPY	0.5465 (0.0000)	3.2629 (0.0000)	0.5463 (0.0000)	3.2640 (0.0000)	0.6418 (0.0000)	3.2217 (0.0000)	0.6419 (0.0220)	3.2219 (0.0060)	0.6531 (0.0000)	2.9830 (0.9280)	0.6553 (0.0000)	2.9943 (0.0000)
VIX	0.30133 (0.0000)	3.40041 (0.0000)	0.30070 (0.0000)	3.39824 (0.0000)	0.2975 (0.0000)	3.3507 (0.0000)	0.2975 (0.0000)	3.3507 (0.0000)	0.1059 (0.0000)	3.0037 (0.0000)	0.1059 (0.0000)	3.0000 (0.0000)

The table reports the skewness, kurtosis, and bootstrap p-values for testing the normality of the predicted residuals across the sectors ETFs (and, in the last line, the VIX index) under two models: SHARP and SHARP-SV. The bootstrap procedure assesses whether the distribution of residuals deviates significantly from normality.

B. FURTHER MONTE CARLO RESULTS

B.1. Mnemonics, parameter values and figures for Section 5.2. In Table B.1, we summarise five variants of the main model, each incorporating different empirical features.

TABLE B.1. Model specifications and mnemonics

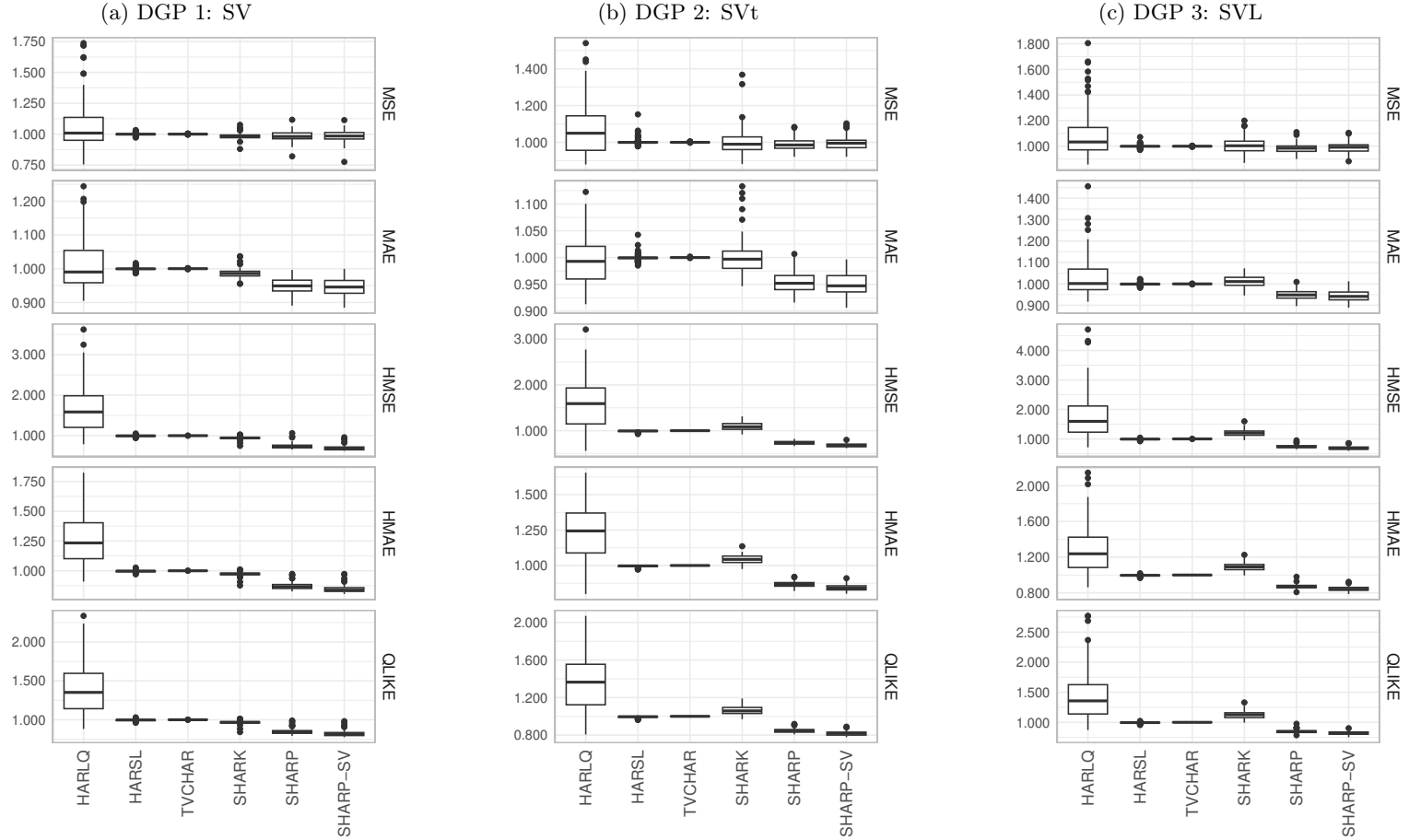
DGP	Distribution of ϵ_t	Leverage effect	Jumps	Multi-regime
SV	Gaussian - $\lambda_t = 1$	NO - $\rho = 0$	NO - $\kappa = 0$	NO - $\gamma = 0$
SVt	Student's t - $\lambda_t \sim \mathcal{IG}(v/2, v/2)$	NO - $\rho = 0$	NO - $\kappa = 0$	NO - $\gamma = 0$
SVL	Gaussian - $\lambda_t = 1$	YES - $\rho \neq 0$	NO - $\kappa = 0$	NO - $\gamma = 0$
SVJ	Gaussian - $\lambda_t = 1$	YES - $\rho \neq 0$	YES - $\kappa > 0$	NO - $\gamma = 0$
SVML	Gaussian - $\lambda_t = 1$	YES - $\rho \neq 0$	NO - $\kappa = 0$	YES - $\gamma \neq 0$

We report the values of the parameters used in our simulations, which - as mentioned in Section 5.2 in the main paper - are based on Table 3 of [Stroud and Johannes \(2014\)](#). Compared to [Stroud and Johannes \(2014\)](#), however, we do not consider seasonal dummies, and we therefore use higher values for the volatilities parameters σ_1 and σ_2 in (5.5) and (5.4), to account for such omitted factors; further, as mentioned above, we allow for regime switching via $\gamma > 0$.

TABLE B.2. Parameter Values for Different Stochastic Volatility Specifications

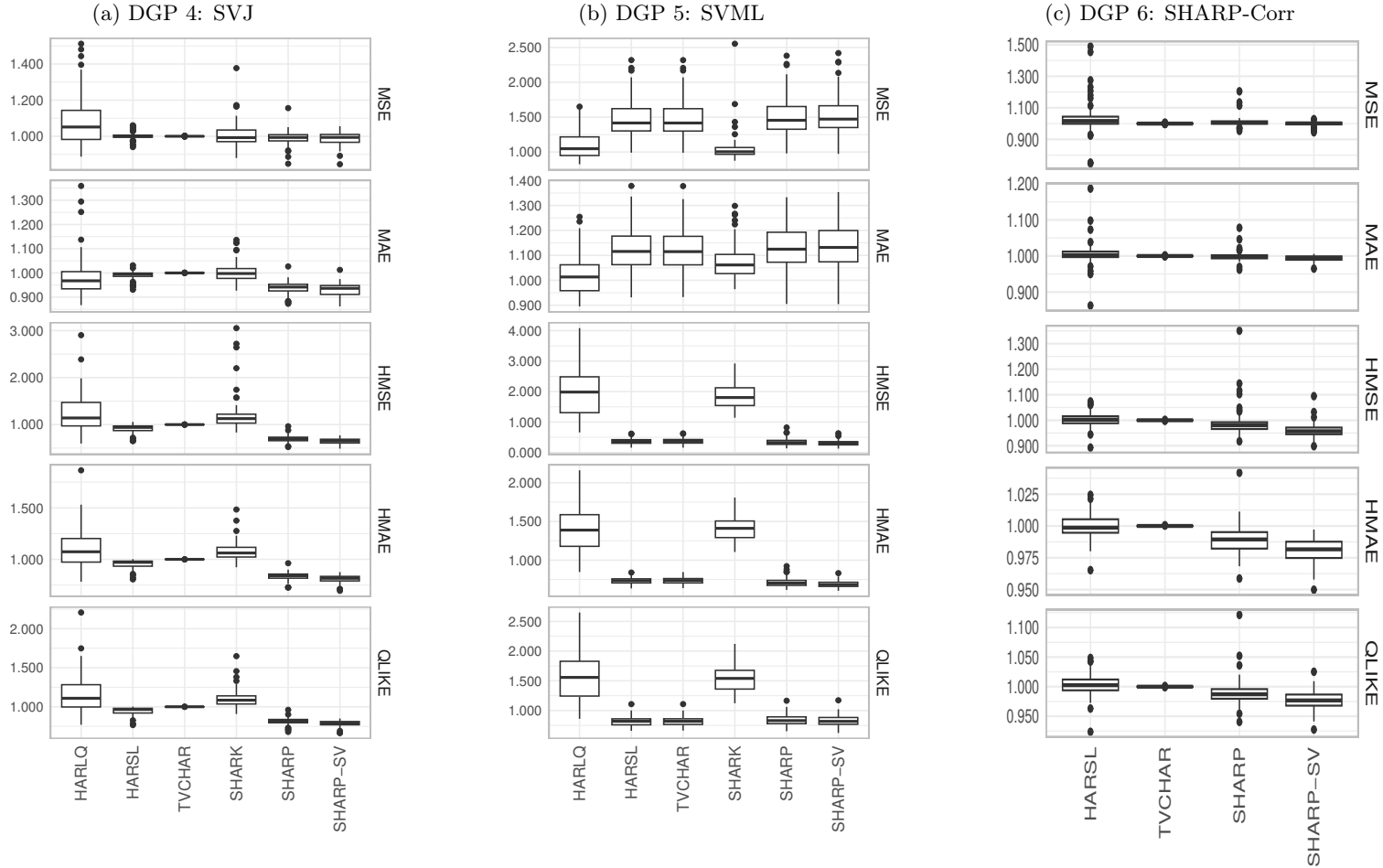
	SV	SVt	SVL	SVJ	SVML
μ	0.0001	0.0000	0.0000	0.0000	0.0000
σ	0.0600	0.0600	0.0600	0.0600	0.0600
ϕ_1	0.9998	0.9998	0.9998	0.9998	0.9998
σ_1	0.0330	0.0330	0.0330	0.0330	0.0330
ϕ_2	0.9270	0.9270	0.9270	0.9270	0.9270
σ_2	0.3000	0.3000	0.3000	0.3000	0.3000
ρ			-0.0950	-0.1360	-0.1360
v		20.0000			
κ				0.0042	
μ_y				-0.0070	
σ_y				0.2020	
μ_v				0.8160	
σ_v				1.2200	
γ					0.1000
p_{11}					0.9500
p_{22}					0.9000

Figure B.1. Box plots of relative loss functions comparison - using the hundred simulated series for each DGP.



Box plots of relative loss measures for five loss functions (MSE, MAE, HMSE, HMAE, QLIKE) across different DGPs for each model, using the HARL as the benchmark model. The models were estimated using a Monte Carlo study with 1,440 simulated observations per iteration, where the first 1,000 observations were used for in-sample estimation, and the last 440 observations were reserved for out-of-sample forecasting evaluation.

Figure B.2. Box plots of relative loss functions comparison - using the hundred simulated series for each DGP.



Box plots of relative loss measures for five loss functions (MSE, MAE, HMSE, HMAE, QLIKE) across different DGPs for each model, using the HARL as the benchmark model. The models were estimated using a Monte Carlo study with 1,440 simulated observations per iteration, where the first 1,000 observations were used for in-sample estimation, and the last 440 observations were reserved for out-of-sample forecasting evaluation. Note that we exclude HARLQ and SHARK from DGP 6, as both require RQ in the estimated model, and in DGP 6 we only simulate RV series directly without simulating intraday return series by design.

B.2. Comparison with MLE *cum* Kalman filtering. We complement our Monte Carlo exercise in Section 5.2 of the main paper by assessing the impact of our Bayesian estimation (using Particle Gibbs) on the performance of our proposed models, using - as an alternative - a Maximum Likelihood Estimator *cum* Kalman filter to estimate the SHARP model.² We note from the outset that the code executes without any numerical optimisation issues; however, contrary to the case where our proposed Bayesian estimator is employed, when using MLE plus Kalman filtering, the model exhibits a tendency to over-forecast the dependent variable at certain time points. In order to investigate this behaviour, we examined the components of the forecasted values, including the bias correction term applied to account for the log-normality of the data; our analysis revealed that the source of the overestimation is not the bias correction *per se*, but rather the forecasted log-volatility component, which in some cases was significantly inflated.

We include a summary table, Table B.3 offering a snapshot of the model’s performance. The table reports the mean, standard deviation, minimum, and maximum of the mean squared error (MSE) computed under each of the six data-generating processes (DGPs). Results from alternative loss functions yielded even worse performance. To ensure that the reported statistics are not unduly influenced by outliers, we excluded extreme forecasts by comparing them to the maximum observed value in the actual data sample.

Results are decidedly worse when using Kalman filtering, which indicates that Particle Gibbs sampling is not merely one of several possible estimation techniques for our proposed SHARP and SHARP-SV models - it is in fact integral to their formulation and practical feasibility: the models are specified as nonlinear state-space systems with time-varying parameters

²Routines have been coded in R. Implementation uses the KFAS package to construct the state space model dynamically for each estimation window. At each step, the log-likelihood function is defined to optimise the AR(1) parameters, the measurement variance, and the drift terms directly. The Kalman filter is used to obtain the filtered state estimates, which are then used to forecast one-step-ahead coefficients and predict the dependent variable using a bias-corrected exponential transformation. This forecasting process is embedded in a rolling window framework and parallelised using FOREACH and DOPARALLEL to improve computational efficiency. Further details are available upon request.

(and stochastic volatility in the case of SHARP-SV), and Particle Gibbs allows to perform Bayesian inference jointly over latent states and parameters, maintaining computational tractability while accommodating the full complexity of the models. Without this approach, the simultaneous estimation of latent volatilities and time-varying coefficients - central to capturing the evolving structure of financial volatility - appears to be significantly more limited or infeasible. All these findings confirm that our proposed models using Bayesian approach effectively captures the key features of financial time series, offering robust and practical improvements over traditional, competing estimation techniques.

TABLE B.3. Relative Forecast Error Summary (MSE) Compared to HARL Benchmark Model

	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5	DGP 6
Mean	4.4712	2.1888	3.0356	2.3119	3.2865	3.2865
Standard Deviation	1.8192	0.5894	1.2811	0.7680	1.0230	1.0230
Minimum	1.7362	1.2205	0.9966	1.2480	1.7077	1.7077
Maximum	9.8005	3.7561	8.3339	4.9219	7.5148	7.5148

This table reports summary statistics of the forecast mean squared error (MSE) of the Kalman filter-based time-varying parameters model. Each value represents the ratio of the model's MSE to that of HARL, averaged across forecasts generated under each of the six data-generating processes (DGPs). The results are computed after excluding all (extreme) predicted values, defined as exceeding the maximum observed value in the actual data.

B.3. Computational times. We report computational times for the basic SV DGP in Table B.4; the results broadly support our claim that the Bayesian inference method remains feasible for large-scale applications.³

TABLE B.4. The average computational time (in seconds) to produce forecasts for 440 out-of-sample observations in a series of the first DGP "SV" using different HPC configurations. Some models were run in parallel, while others were executed serially.

Model	Execution Time (s)	Node Type	Parallelization	Number of Parameters*
HARSL	1159	**Unknown (serial execution)**	No	7
TVCHAR	180	**Unknown (serial execution)**	No	9
SHARK	1697	Intel Ivy Bridge/Haswell/Broadwell (16 cores, 64GB RAM, 10Gb low latency)	Yes	15
SHARP	10222	Intel Skylake/Cascade Lake (40 cores, 192GB RAM, 25Gb low latency)	Yes	13
SHARP-SV	10914	Intel Skylake/Cascade Lake (40 cores, 192GB RAM, 25Gb low latency)	Yes	21

* denotes the number of parameters required per out-of-sample observation.

³Whilst, for the sake of a concise presentation, we only consider one DGP, unreported experiments showed that the computational times, when considering the other DGPs, exhibit essentially the same pattern as in Table B.4.

C. POSTERIOR DERIVATION

This appendix provides the detailed posterior derivations used in the estimation of the volatility models. These derivations are crucial for understanding the underlying statistical methods and ensuring the robustness of our results. We will walk through each step of the derivation, providing context and highlighting key aspects.

Our SHARP-SV model reads as follows:

$$\begin{aligned} y_t &= x_t' \beta_t + v_t; \quad v_t \sim \mathcal{N}(0, \sigma_v^2), \quad t = 1, \dots, n \\ \beta_{tj} &= \alpha_j + \rho_j \beta_{t-1,j} + \varepsilon_{tj}; \quad \varepsilon_{tj} | h_{tj} \sim \mathcal{N}(0, h_{tj}), \quad j = 1, \dots, k, t = 1, \dots, n \\ \ln h_{tj} &= \gamma_j + \delta_j \ln h_{t-1,j} + u_{tj}; \quad u_{tj} \sim \mathcal{N}(0, \sigma_{uj}^2), \quad j = 1, \dots, k, t = 1, \dots, n \end{aligned}$$

In what follows we will utilise the following symbols:

$$\begin{aligned} \textbf{Let} \quad \lambda_t &= [\beta_t', \ln h_t']' \\ \textbf{where} \quad \beta_t &= [\beta_{t1}, \dots, \beta_{tk}] \quad \textbf{and} \quad \ln h_t = [\ln h_{t1}, \dots, \ln h_{tk}] \\ \textbf{Let} \quad \theta &= \{\theta_1, \dots, \theta_k\}, \quad \textbf{where} \quad \theta_j = \{\alpha_j, \rho_j, \gamma_j, \delta_j, \sigma_{uj}\}, \quad j = 1, \dots, k, \end{aligned}$$

We assume the following prior for the parameters:

$$\begin{aligned} \alpha_j, \gamma_j &\sim \mathcal{N}(0, 1), \quad j = 1, \dots, k, \\ \rho_j, \delta_j &\sim \mathcal{N}(0.5, 1) \mathbb{I}_{\rho_j \in (0,1)} \mathbb{I}_{\delta_j \in (0,1)}, \quad j = 1, \dots, k, \\ \sigma_{uj}^2 &\sim \Gamma(6.5, 0.5), \quad j = 1, \dots, k \\ \sigma_v^2 &\sim \Gamma(6.5, 0.5) \end{aligned}$$

Equivalently, the prior can be written as:

$$\begin{aligned}
p(\alpha_j) &\propto \exp(-\frac{n}{2}\alpha_j^2), \quad j = 1, \dots, k \\
p(\gamma_j) &\propto \exp(-\frac{n}{2}\gamma_j^2), \quad j = 1, \dots, k \\
p(\rho_j) &\propto \exp(-\frac{n}{2}(\rho_j - 0.5)^2)\mathbb{I}_{\rho_j \in (0,1)}, \quad j = 1, \dots, k \\
p(\delta_j) &\propto \exp(-\frac{n}{2}(\delta_j - 0.5)^2)\mathbb{I}_{\delta_j \in (0,1)}, \quad j = 1, \dots, k \\
p(\sigma_{uj}) &\propto \sigma_{uj}^{-(\underline{n}+1)} \exp\left(-\frac{q}{2\sigma_{uj}^2}\right), \quad j = 1, \dots, k, \quad \underline{n} = 6.5 \quad \text{and} \quad \underline{q} = 1 \\
p(\sigma_v) &\propto \sigma_{vj}^{-(\underline{n}+1)} \exp\left(-\frac{q}{2\sigma_{vj}^2}\right), \quad \underline{n} = 6.5 \quad \text{and} \quad \underline{q} = 1
\end{aligned}$$

The posterior distribution is derived using Bayes' theorem. This step is fundamental as it combines prior information with the likelihood of the observed data to form the posterior distribution. It is thus derived as follows:

$$\begin{aligned}
P(\theta, \sigma_v, \{\lambda\}_{t=1}^T | Y) &\propto (\sigma_v^{-n}) \exp(-\frac{1}{2\sigma_v^2} \sum_{t=1}^n (y_t - x_t' \beta_t)^2) \\
&(\prod_{i=1}^k \prod_{t=1}^n h_{it}^{-1}) \exp(-\sum_{i=1}^k \sum_{t=1}^n \frac{1}{2h_{it}^2} (\beta_{it} - \alpha_i - \rho_i \beta_{i,t-1})^2) \\
&(\prod_{i=1}^k \sigma_{u_i}^{-n}) \exp(-\sum_{i=1}^k \frac{1}{2\sigma_{u_i}^2} \sum_{t=1}^n (\ln h_{it} - \gamma_{it} - \delta_{it} \ln h_{i,t-1})^2) \\
&\prod_{j=1}^k p(\sigma_{u_i}) \prod_{i=1}^k p(\alpha_i) \prod_{i=1}^k p(\gamma_i) \prod_{i=1}^k p(\rho_i) \prod_{i=1}^k p(\delta_i)
\end{aligned}$$

In order to draw the posterior for each parameter, we use a mixed estimator approach within Gibbs sampling. This method involves drawing each parameter sequentially within each Monte Carlo iteration. Below, we demonstrate this process for each parameter.

$$(1) \text{ For } \gamma_i \left| \{ \{\theta \setminus \{\gamma_i\}\}, \sigma_v, \{\lambda\}_{t=1}^T, Y \} :
\right.$$

$$\begin{aligned}
& \text{let } Z = \begin{bmatrix} \ln h_{i,2} - \delta_i \ln h_{i,1} \\ \cdot \\ \cdot \\ \ln h_{i,T} - \delta_i \ln h_{i,T-1} \end{bmatrix} \text{ and } u_i = \begin{bmatrix} u_{i,2} \\ \cdot \\ \cdot \\ u_{i,T} \end{bmatrix} \\
& \text{then, } \begin{bmatrix} Z \\ 0 \end{bmatrix} = \begin{bmatrix} 1_{(n \times 1)} \\ 1 \end{bmatrix} \gamma_i + \begin{bmatrix} u_i \\ \eta \end{bmatrix}; \begin{bmatrix} u_i \\ \eta \end{bmatrix} \sim N(0, \begin{bmatrix} \sigma_i I_n & \\ & 1 \end{bmatrix}) \\
& \text{where } n = T - 1 \\
& \Rightarrow \gamma_i \left| \left\{ \{\theta \setminus \{\gamma_i\}\}, \{\lambda\}_{t=1}^T, Y \right\} \right. \sim N\left(\frac{(\ln h_{i,T} - \ln h_{i,1})}{1 + \sigma_i^2}, \frac{\sigma_i^2}{1 + \sigma_i^2}\right)
\end{aligned}$$

(2) For $\delta_i \left| \left\{ \{\theta \setminus \{\delta_i\}\}, \sigma_v, \{\lambda\}_{t=1}^T, Y \right\} \right.:$

$$\begin{aligned}
& \text{let } Z = \begin{bmatrix} \ln h_{i,2} - \gamma_i \\ \cdot \\ \cdot \\ \ln h_{i,T} - \gamma_i \end{bmatrix}, \kappa = \begin{bmatrix} \ln h_{i,1} \\ \cdot \\ \cdot \\ \ln h_{i,T-1} \end{bmatrix} \text{ and } u_i = \begin{bmatrix} u_{i,2} \\ \cdot \\ \cdot \\ u_{i,T} \end{bmatrix} \\
& \text{then, } \begin{bmatrix} Z \\ 0.5 \end{bmatrix} = \begin{bmatrix} \kappa \\ 1 \end{bmatrix} \delta_i + \begin{bmatrix} u_i \\ \eta \end{bmatrix}; \begin{bmatrix} u_i \\ \eta \end{bmatrix} \sim N(0, \begin{bmatrix} \sigma_i I_n & \\ & 1 \end{bmatrix}) \\
& \text{where } n = T - 1
\end{aligned}$$

$$\Rightarrow \delta_i \left| \left\{ \{\theta \setminus \{\delta_i\}\}, \{\lambda\}_{t=1}^T, Y \right\} \right. \sim N\left(\frac{\kappa' Z + 0.5 \sigma_i^2}{\kappa' \kappa + \sigma_i^2}, \frac{\sigma_i^2}{\kappa' \kappa + \sigma_i^2}\right)$$

(3)

$$\begin{aligned}
& P\left(\sigma_{u_i} \left| \{\theta \setminus \{\sigma_{u_i}\}\}, \sigma_v, \{\lambda\}_{t=1}^T, Y \right.\right) \\
& \propto \\
& \sigma_{u_i}^{-n} \exp\left(-\frac{1}{2\sigma_{u_i}^2} \sum_{t=1}^n (\ln h_{it} - \gamma_i - \delta_i \ln h_{i,t-1})^2\right) \sigma_{u_i}^{-(n+1)} \exp\left(-\frac{q}{2\sigma_{u_i}^2}\right) \\
& \propto \\
& \sigma_{u_i}^{-(n+n+1)} \exp\left(-\frac{1}{2\sigma_{u_i}^2} \left(\sum_{t=1}^n (\ln h_{it} - \gamma_i - \delta_i \ln h_{i,t-1})^2 + \underline{q}\right)\right)
\end{aligned}$$

$$\implies \frac{\sum_{t=1}^n (\ln h_{it} - \gamma_i - \delta_i \ln h_{i,t-1})^2 + q}{\sigma_{u_i}^2} \left| \{ \{\theta \setminus \{\sigma_i\}\}, \{\lambda\}_{t=1}^T, Y \} \sim \chi^2(n + \underline{n} + 3) \right.$$

(4) For $\alpha_i \left| \{ \{\theta \setminus \{\alpha_i\}\}, \{\lambda\}_{t=1}^T, Y \} \right.:$

$$\text{let } Z = \begin{bmatrix} \beta_{i,2} - \rho_i \beta_{i,1} \\ \cdot \\ \cdot \\ \beta_{i,T} - \rho_i \beta_{i,T-1} \end{bmatrix}, H_i = \begin{bmatrix} h_{i,1} \\ \cdot \\ \cdot \\ h_{i,T} \end{bmatrix}, \text{ and } \epsilon_i = \begin{bmatrix} \epsilon_{i,2} \\ \cdot \\ \cdot \\ \epsilon_{i,T} \end{bmatrix}$$

$$\text{then, } \begin{bmatrix} Z \\ 0 \end{bmatrix} = \begin{bmatrix} 1_{(n \times 1)} \\ 1 \end{bmatrix} \alpha_i + \begin{bmatrix} \epsilon_i \\ \eta \end{bmatrix}; \begin{bmatrix} u_i \\ \eta \end{bmatrix} \sim N(0, \begin{bmatrix} H_i & \\ & 1 \end{bmatrix})$$

where $n = T - 1$

$$\implies \alpha_i \left| \{ \{\theta \setminus \{\alpha_i\}\}, \{\lambda\}_{t=1}^T, Y \} \sim N\left(\frac{\sum_{t=2}^T \frac{Z_{i,t}}{h_{i,t}}}{\sum_{t=2}^T \frac{1}{h_{i,t}} + 1}, \frac{1}{\sum_{t=2}^T \frac{1}{h_{i,t}} + 1}\right) \right.$$

(5) For $\rho_i \left| \{ \{\theta \setminus \{\rho_i\}\}, \sigma_v, \{\lambda\}_{t=1}^T, Y \} \right.:$

$$\text{let } Z = \begin{bmatrix} \beta_{i,2} - \alpha_i \\ \cdot \\ \cdot \\ \beta_{i,T} - \alpha_i \end{bmatrix}, \kappa = \begin{bmatrix} \beta_{i,1} \\ \cdot \\ \cdot \\ \beta_{i,T-1} \end{bmatrix}, H_i = \begin{bmatrix} h_{i,1} \\ \cdot \\ \cdot \\ h_{i,T} \end{bmatrix},$$

$$\text{and } \epsilon_i = \begin{bmatrix} \epsilon_{i,2} \\ \cdot \\ \cdot \\ \epsilon_{i,T} \end{bmatrix}$$

$$\text{then, } \begin{bmatrix} Z \\ 0.5 \end{bmatrix} = \begin{bmatrix} \kappa \\ 1 \end{bmatrix} \rho_i + \begin{bmatrix} \epsilon_i \\ \eta \end{bmatrix}; \begin{bmatrix} u_i \\ \eta \end{bmatrix} \sim N(0, \begin{bmatrix} H_i & \\ & 1 \end{bmatrix})$$

where $n = T - 1$

$$(6) \quad \Rightarrow \rho_i \left| \left\{ \{\theta \setminus \{\rho_i\}\}, \{\lambda\}_{t=1}^T, Y \right\} \sim N \left(\frac{\sum_{t=2}^T \frac{Z_{i,t}}{h_{i,t}} + 0.5}{\sum_{t=2}^T \frac{1}{h_{i,t}} + 1}, \frac{1}{\sum_{t=2}^T \frac{1}{h_{i,t}} + 1} \right) \right.$$

$$P \left(\sigma_v \left| \{\theta\}, \{\lambda\}_{t=1}^T, Y \right. \right)$$

$$\propto$$

$$\sigma_v^{-n} \exp \left(-\frac{1}{2\sigma_v^2} \sum_{t=1}^n (y_t - x'_t \beta_t)^2 \right) \quad \sigma_v^{-(\underline{n}+1)} \exp \left(-\frac{q}{2\sigma_v^2} \right)$$

$$\propto$$

$$\sigma_v^{-(n+\underline{n}+1)} \exp \left(-\frac{1}{2\sigma_i^2} \left(\sum_{t=1}^n (y_t - x'_t \beta_t)^2 + \underline{q} \right) \right)$$

$$\Rightarrow \frac{\sum_{t=1}^n (y_t - x'_t \beta_t)^2 + \underline{q}}{\sigma_v^2} \left| \left\{ \{\theta\}, \{\lambda\}_{t=1}^T, Y \right\} \sim \chi^2(n + \underline{n} + 3) \right.$$

D. PARTICLE FILTERING WITHIN MCMC

Particle filtering is a simulation-based algorithm that sequentially approximates continuous marginal distributions using discrete distributions. This is performed by using a set of support points called “particles” and probability masses; see [Creal \(2012\)](#) for a review. The PG sampler draws a single path of the latent or state variables from this discrete approximation. As the number of particles M goes to infinity, the PG sampler draws from the exact full conditional distribution. The advantage of the algorithm is that it allows for drawing paths of the state variables in large blocks.

As mentioned in [Creal and Tsay \(2015\)](#), the PG sampler is a standard Gibbs sampler but defined on an extended probability space where a particle filter generates all the random variables. [Chopin and Singh \(2015\)](#) analysed the theoretical properties of the PG sampler, and showed that the sampler is uniformly ergodic. Unlike the standard particle filter, the PG sampler involves a “conditional” resampling algorithm in the last step. Namely, for draws from the particle filter to be a valid Markov transition kernel on the extended probability space, the state variables drawn at the previous iteration must have a positive sampling probability [Andrieu et al. \(2010\)](#). The conditional resampling step within the PG forces the pre-existing path to survive the particle filter’s resampling steps. We use the conditional multinomial resampling algorithm from [Andrieu et al. \(2010\)](#), although other resampling algorithms exist (see, for example, [Fearnhead et al., 2010](#) and [Chopin and Singh, 2015](#)).

Suppose the posterior is $p(\theta, \lambda_{1:T} | \mathbf{y}_{1:T})$ where $\lambda_{1:T}$ denotes the latent variables whose prior can be described by $p(\lambda_t | \lambda_{t-1}, \theta)$. In the PG sampler, we can draw the structural parameters $\theta | \lambda_{1:T}, \mathbf{y}_{1:T}$ as usual, from their posterior conditional distributions. This is important because, in this way, we can avoid mixture approximations or other Monte Carlo procedures that need considerable tuning and may not have good convergence properties. Suppose we have $\lambda_{1:T}^{(1)}$ from the previous iteration. The particle filtering procedure consists of two phases, forward and backward filtering.

Phase I: Forward filtering [Andrieu et al. \(2010\)](#).

For $t = 1, \dots, T$

- Draw a proposal $\lambda_t^{(m)}$ from an importance density $q(\lambda_t | \lambda_{t-1}^{(m)}, y_t)$, $m = 2, \dots, M$.
- Compute the importance weights:

$$(D.1) \quad w_t^{(m)} = \frac{p(y_t | \lambda_t^{(m)}) p(\lambda_t^{(m)} | \lambda_{t-1}^{(m)}, y_t)}{q(\lambda_{it} | \lambda_{t-1}^{(m)}, y_t)}, \quad m = 1, \dots, M.$$

- Normalise the weights: $\tilde{w}_t^{(m)} = \frac{w_t^{(m)}}{\sum_{m=1}^M w_t^{(m)}}$, $m = 1, \dots, M$.
- Re-sample, conditionally, the particles $\{\lambda_t^{(m)}, m = 1, \dots, M\}$ with probabilities $\{\tilde{w}_t^{(m)}, m = 1, \dots, M\}$.

In the original PG sampler, the particles are stored for $t = 1, \dots, T$ and a single trajectory is sampled using the probabilities from the last iteration. An improvement upon the original PG by drawing the path of the latent variables from the particle approximation is using the backwards sampling algorithm of [Godsill et al. \(2004\)](#). In the forward pass, we store the normalised weights and particles then we draw a path of the latent variables as we detail below (the draws are from a discrete distribution).

Phase II: Backward filtering ([Chopin and Singh, 2015](#); [Godsill et al., 2004](#)).

- At time $t = T$ draw a particle $\lambda_T^* = \lambda_T^{(m)}$.
- Compute the backward weights: $w_{t|T}^{(m)} \propto \tilde{w}_t^{(m)} p(\lambda_{t+1}^* | \lambda_t^{(m)}, \theta)$.
- Normalise the weights: $\tilde{w}_{t|T}^{(m)} = \frac{w_{t|T}^{(m)}}{\sum_{m=1}^M w_{t|T}^{(m)}}$, $m = 1, \dots, M$.
- Draw a particle $\lambda_{it}^* = \lambda_t^{(m)}$ with probability $\tilde{w}_{t|T}^{(m)}$.

Therefore, $\lambda_{1:T}^* = \{\lambda_1^*, \dots, \lambda_T^*\}$ is a draw from the full conditional distribution. When the state vector dimension is large, we can draw $\lambda_{i,1:T}$, conditional on all other paths $\lambda_{-i,1:T}$ that are not path i . Therefore, we can draw from the full conditional distribution

$p(\lambda_{i,1:T}|\lambda_{-i,1:T}, \mathbf{y}_{1:T}, \theta)$. The backwards step often results in dramatic improvements in computational efficiency and strictly dominates the original PG (Chopin and Singh, 2015). For example, Creal and Tsay (2015) find that $M = 100$ particles is enough.

There remains the problem of selecting an importance density $q(\lambda_t|\lambda_{t-1}, y_t)$. We use an importance density implicitly defined by $\lambda_{it} = a_i + \sum_{p=1}^P b_{i,p} \lambda_{i,t-1}^p + \Omega_i \xi_{it}$ where ξ_{it} follows a standard normal distribution (although a student t-distribution maybe more appropriate in other applications). That is, we use polynomials in $\lambda_{i,t-1}$ of order P . We select the parameters a_i, b_i and Ω_i during the burn-in phase (using $P = 1$ and $P = 2$) so that the weights $\{\tilde{w}_{it}^{(m)}, m = 1, \dots, M\}$ and $\{\tilde{w}_{t|T}^{(m)}, m = 1, \dots, M\}$ are approximately not too far from a uniform distribution.

E. THE STRUCTURE OF THE SHARP MODEL

We begin by studying the dependence structure of the following model

$$(E.1) \quad y_t = x_t' \beta_t + v_t,$$

$$(E.2) \quad (\beta_t - \mu) = \mathbf{\Gamma} (\beta_{t-1} - \mu) + \varepsilon_t,$$

$$(E.3) \quad \varepsilon_t = \mathbf{H}_t^{1/2} \eta_t,$$

$$(E.4) \quad \text{vech}(\ln \mathbf{H}_t - \mathbf{C}) = \mathbf{A} \text{vech}(\ln \mathbf{H}_{t-1} - \mathbf{C}) + u_t,$$

where x_t in (E.1) is a $k \times 1$ vector of exogenous regressors, and $\mathbf{C}$ in (E.4) is a symmetric matrix. The structure of model (E.1)-(E.4) has not been analysed in the literature yet; Aue et al. (2009) analyse the dependence structure of the stationary solution of (E.3)-(E.4), considering a different notion of decomposability than the one we use herein (see Definition E.1 below) and without studying the impact of the initial conditions; Kokoszka et al. (2025) study a *univariate* version of the stationary solution of (E.3)-(E.4), again without assessing the impact of the initial condition.

In order to make a closer comparison with our SHARP-SV model, we will also consider, separately, the dynamic version of (E.1), namely

$$(E.5) \quad y_t = \sum_{j=1}^p \beta_{j,t} y_{t-j} + v_t,$$

where, with reference to (E.2), we define $\beta_t = (\beta_{1,t}, \dots, \beta_{p,t})'$. Model (E.2)-(E.5) encompasses the SHARP-SV model of equations (3.1)-(3.5) in the main paper as a special case - and, therefore, also the SHARP model. Equations (E.2)-(E.5) generalise our framework in the main paper in at least three directions. Firstly, we do not assume Gaussian innovations in any equation, and we refer to Assumptions E.1-E.4 below for further remarks on this - although, as we show below, we do need the existence of the Moment Generating Function

of (some of) the innovations in a sufficiently large neighbourhood of zero. Secondly, in the law of motion in (E.2), we allow for a full-fledged VAR(1) specification, thus allowing for feedback in the coefficient dynamics. Thirdly and finally, we do not impose a diagonal structure in the dynamic covariance matrix $\mathbf{H}_t$. Equations (E.2) and (E.4) nest the equation-by-equation specification in (3.2) and (3.5) respectively.⁴ We would like to point out that the analysis of the dependence structure (and of the related asymptotics) for model (E.2)-(E.5) is entirely novel in the literature. Indeed, in a series of related contributions, Professor Lajos Horváth and his co-authors have studied a related model, i.e. the Random Coefficient AutoRegressive (RCAR) model, but virtually all results are derived for the case where $p = 1$ in (E.5), and with *independent* random coefficients β_t - see, *inter alia*, Aue et al. (2006) and Horváth and Trapani (2016); Horváth and Trapani (2023) study the case $p > 1$ under stationarity, but again all their arguments rely heavily on the assumption of independent β_t , and cannot be immediately extended to the case of serially dependent random coefficients.

Henceforth, we let $\|\cdot\|$ denote the Euclidean norm of a vector, and the induced norm when applied to a matrix; when using the $\mathcal{L}_\nu$ -norm of a vector or matrix valued random variable X , we define it as $|X|_\nu = (E \|X\|^\nu)^{1/\nu}$.

We begin with the following definition of weak dependence.

Definition E.1. *The r -dimensional sequence $\{m_t, -\infty < t < \infty\}$ forms an $\mathcal{L}_\nu$ -decomposable Bernoulli shift if and only if it holds that $m_t = g(\eta_t, \eta_{t-1}, \dots)$, where: $g(\cdot) : \mathbb{S}^\infty \rightarrow \mathbb{R}^r$ is a non random measurable function; $\{\eta_t, -\infty < t < \infty\}$ is an i.i.d. sequence with values in a measurable space $\mathbb{S}$; $E(m_t) = 0$ and $|m_t|_\nu < \infty$; and $|m_t - m_{t,\ell}^*|_\nu \leq c_0 \ell^{-a}$, for some $c_0 > 0$ and $a > 0$, where $m_{t,\ell}^* = g(\eta_t, \dots, \eta_{t-\ell+1}, \eta_{t-\ell,t,\ell}^*, \eta_{t-\ell-1,t,\ell}^*, \dots)$, with $\{\eta_{s,t,\ell}^*, -\infty < s, \ell, t < \infty\}$ i.i.d. copies of η_0 independent of $\{\eta_t, -\infty < t < \infty\}$. Further, if $|m_t - m_{t,\ell}^*|_\nu \leq c_0 \rho^\ell$ for some $c_0 > 0$ and $|\rho| < 1$, the sequence $\{m_t, -\infty < t < \infty\}$ forms an $\mathcal{L}_\nu$ -decomposable Bernoulli shift with exponential rate.*

⁴We are grateful to an anonymous Referee for suggesting these generalisations.

The concepts of Bernoulli shift and decomposability appeared first in [Ibragimov \(1962\)](#); we also refer to the works by [Wu \(2005\)](#), [Liu and Lin \(2009\)](#), and [Berkes et al. \(2011\)](#), for further results and insights on this form of dependence. Bernoulli shifts have proven a convenient way to model dependent time series, mainly due to their generality and to the fact that it is much easier to verify whether a sequence forms a decomposable Bernoulli shift than e.g. verifying mixing conditions. [Aue et al. \(2009\)](#), [Liu and Lin \(2009\)](#) and [Barigozzi and Trapani \(2022\)](#) provide various theoretical results, and numerous examples of data which satisfy Definition [E.1](#), including ARMA models, ARCH/GARCH sequences, and other nonlinear time series models (such as e.g. random coefficient autoregressive models and threshold autoregressive models).

Henceforth, we define the bounds $O(\ell^{-a})$ and $O(\rho^\ell)$ as the *rate function* of the Bernoulli shift.

We begin by studying [\(E.1\)](#)-([E.4](#)). We require the following assumptions.

Assumption E.1. *It holds that: (i) x_t is an $\mathcal{L}_{p'}$ -decomposable Bernoulli shift, for some $p' > 2$ with rate function $O(\ell^{-a})$ for some $a > 0$; (ii) v_t is an $\mathcal{L}_{p''}$ -decomposable Bernoulli shift, for some $p'' > 2$ with rate function $O(\ell^{-c})$ for some $c > 0$; (iii) $\{x_t, -\infty < t < \infty\}$, $\{\beta_t, -\infty < t < \infty\}$ and $\{v_t, -\infty < t < \infty\}$ are three mutually independent groups.*

Assumption E.2. *It holds that: (i) $\|\Gamma\| < 1$; (ii) β_0 is independent of $\{\varepsilon_t, -\infty < t < \infty\}$ with $|\beta_0|_p < \infty$ for some $p > 2$.*

Assumption E.3. *It holds that: (i) $\|\mathbf{A}\| < 1$; (ii) u_t is i.i.d. with mean zero and $E \exp(\tau \|u_0\|) < \infty$ for some $\tau > 0$; (iii) $|\ln \mathbf{H}_0|_q < \infty$ for all $q > 0$.*

Assumption E.4. *It holds that: (i) η_t is an $\mathcal{L}_{p(1+\epsilon)}$ -decomposable Bernoulli shift, for some $p > 2$, $\epsilon > 0$ and with rate function $O(\ell^{-b})$ for some $b > 0$.*

Some comments on Assumptions [E.1-E.4](#) are in order. In Assumption [E.1](#), we require that x_t be weakly dependent; the value of a in the rate function is specified below, but - in order

to study the dependence structure of y_t - it suffices to have $a > 0$. Assumption E.3 does not require Gaussianity of the innovation u_t , but by part (ii) heavy tails are ruled out; in principle, it would be possible to accommodate this important potential feature of the data by modifying (E.3), moving e.g. from an exponential GARCH-type specification to a power-type one. Finally, note that in Assumptions E.3 and E.4 we do not require independence between $\mathbf{H}_t$ and η_t .

The following results state that y_t is weakly dependent in the sense of Definition E.1. Consider (E.1)-(E.4), and define the stationary version of y_t as

$$(E.6) \quad \bar{y}_t = x_t' \bar{\beta}_t + v_t,$$

where $\bar{\beta}_t$ is the stationary solution to (E.2). Let $\tilde{p} = \min \{p, p', p''\}$.

Theorem E.1. *We assume that Assumptions E.1-E.4 are satisfied, with $\tau > (3/\sqrt{2}) \tilde{p}\epsilon / (1 + \epsilon)$ in Assumption E.3(ii). Then it holds that $\bar{y}_t$ is an $\mathcal{L}_{\tilde{p}}$ -decomposable Bernoulli shift with rate function $O(\min \{\ell^{-a}, \ell^{-b}, \ell^{-c}\})$.*

Theorem E.1 is, to the best of our knoweldge, the first full-fledged study of the dependence structure of (the stationary solution of) a model with time-varying parameters and possible conditional heteroskedasticity in the innovations of the law of motion of the parameters.

Turning to the dynamic case of (E.2)-(E.5), let

$$(E.7) \quad \mathbf{B}_t = \begin{bmatrix} \bar{\beta}_{1,t} & \bar{\beta}_{2,t} & \cdot & \cdot & \bar{\beta}_{p,t} \\ 1 & 0 & & & 0 \\ 0 & 1 & & & \cdot \\ \cdot & 0 & & & \cdot \\ 0 & & & & 0 \end{bmatrix},$$

and consider the following assumption, which extends parts of Assumptions E.1-E.4, where we use $x_t = (y_{t-1}, \dots, y_{t-p})'$.

Assumption E.5. *It holds that: (i) $E \ln \|\mathbf{B}_0\| < 0$; (ii) $|\mathbf{B}_0|_{p'''} \leq \bar{\kappa} < 1$ for all $p''' > 0$; (iii) x_0 is independent of $\{\mathbf{B}_t, -\infty < t < \infty\}$, and such that $|x_0|_{\tilde{p}} < \infty$.*

Theorem E.2. *We assume that the assumptions of Theorem E.1 and Assumption E.5 are satisfied. Then it holds that $\bar{y}_t$, defined as the stationary solution of (E.2)-(E.5), is an $\mathcal{L}_{\tilde{p}}$ -decomposable Bernoulli shift with exponential rate function.*

As mentioned above, the result in Theorem E.2 is entirely novel in the literature. Partnering the result in the theorem with the Hájek-Rényi inequality in the proof of Corollary E.1, essentially all asymptotic results for limits involving averages of y_t can be (readily) derived. Indeed, using exactly the same method of proof with minor variations, the same results can be derived for virtually all models in the HAR class.

The theorems above immediately yields the following results for the partial sums of y_t .

Corollary E.1. *We assume that the assumptions of Theorems E.1 or E.2 are satisfied with both $a > 1$ and $b > 1$. Then, for any $\gamma \in \mathbb{R}$ and $T \in \mathbb{N}$, it holds that*

$$(E.8) \quad \left| \sum_{t=\gamma+1}^{\gamma+T} y_t \right|_{\tilde{p}}^{\tilde{p}} \leq C_{\tilde{p}} T^{\tilde{p}/2},$$

where $C_{\tilde{p}}$ is a constant that depends only on $\tilde{p}$ and on the sequence $\{y_t, -\infty < t < \infty\}$. Also

$$(E.9) \quad \max_{1 \leq k \leq T} \left| \sum_{t=\gamma+1}^{\gamma+T} y_t \right|_r^r \leq C_{\tilde{p},r} T^{r/2},$$

for all $2 < r < \tilde{p}$, where $C_{\tilde{p},r}$ is a constant that depends only on $\tilde{p}$, r and on the sequence $\{y_t, -\infty < t < \infty\}$.

Corollary E.2. *We assume that the assumptions of Theorem E.1 or E.2 are satisfied with both $a > 2$ and $b > 2$. Then, on a suitably enlarged probability space, for each T there exist two independent Wiener processes with the same long-run variance as y_t , say $\{W_{T,1}(k), 1 \leq k \leq T/2\}$ and $\{W_{T,2}(k), 1 \leq k \leq T/2\}$, such that*

$$(E.10) \quad \max_{1 \leq k \leq T/2} \frac{1}{k^{1/\tilde{p}}} \left| \sum_{t=1}^k y_t - W_{T,1}(k) \right| = O_P(1),$$

$$(E.11) \quad \max_{T/2 < k \leq T} \frac{1}{(T-k)^{1/\tilde{p}}} \left| \sum_{t=k+1}^T y_t - W_{T,2}(T-k) \right| = O_P(1).$$

The results in Corollaries E.1 and E.2 are, naturally, outside the scope of this paper, where *Bayesian* estimation is proposed as the main tool to carry out inference. Nevertheless, Corollaries E.1 and E.2 completely characterize the whole asymptotics associated with SHARP and SHARP-SV models, and, *mutatis mutandis*, with the whole class of HAR models.

F. TECHNICAL LEMMAS

Henceforth, “ $\stackrel{\mathcal{D}}{=}$ ” denotes equality in distribution, and $\mathbf{I}$ is an identity matrix of conformable dimensions with the other matrices involved in the same expression.

Some of our lemmas are similar to the ones in [Aue et al. \(2009\)](#), with the crucial differences that we use a different definition of weak dependence (and, thus, need to adapt our proofs thereto), and the further presence of equation (E.2).

We begin with a definition of the matrix exponential for a *symmetric* $m \times m$ matrix $\mathbf{B}$

$$\begin{aligned} \text{(F.1)} \quad \exp(\mathbf{B}) &= \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{B}^k \\ \text{(F.2)} \quad &= \lim_{n \rightarrow \infty} \left(\mathbf{I} + \frac{1}{n} \mathbf{B} \right)^n; \end{aligned}$$

for the ease of exposition, we collect some well-known facts on the matrix exponential in the following lemma (see also Lemma B.3 in [Aue et al., 2009](#), although we provide more details on the perturbation bound in part (iv) of our lemma).

Lemma F.1. *Let $\mathbf{B}$ be an $m \times m$ symmetric matrix. Then it holds that*

- (i): $\exp(\mathbf{B})$ is positive definite;
- (ii): for any two real numbers a and b , $\exp(a\mathbf{B})\exp(b\mathbf{B}) = \exp((a+b)\mathbf{B})$, and $\exp(\mathbf{B})\exp(-\mathbf{B}) = I_m$;
- (iii): $(\exp(\mathbf{B}))^{1/2} = \exp(\mathbf{B}/2)$;
- (iv): for any $m \times m$ matrix $\mathbf{E}$, it holds that

$$\|\exp(\mathbf{B} + \mathbf{E}) - \exp(\mathbf{B})\| \leq \|\mathbf{E}\| \exp(\|\mathbf{B}\|) \exp(\|\mathbf{E}\|);$$

- (v): $\|\exp(\mathbf{B})\| \leq \exp(\|\mathbf{B}\|)$.

Proof. Parts (i) and (ii) are well-known (see e.g. Corollary 11.1.6 in [Bernstein, 2009](#)). Part (iii) can be readily shown by using (F.2). As far as part (iv) is concerned, equation (3.5) in

Van Loan (1977) yields

$$\|\exp(\mathbf{B} + \mathbf{E}) - \exp(\mathbf{B})\| \leq \|\mathbf{E}\| \|\exp(\mathbf{B})\| \exp(\|\mathbf{E}\|);$$

further, by (F.1)

$$\|\exp(\mathbf{B})\| = \left\| \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{B}^k \right\| \leq \sum_{k=0}^{\infty} \frac{1}{k!} \|\mathbf{B}\|^k = \exp(\|\mathbf{B}\|),$$

whence the desired result obtains. Part (v) also follows. $\square$

We now present a (local) reverse Jensen's inequality, which refines Lemma B4 in Aue et al. (2009).

Lemma F.2. *Given a random variable X for which $E \exp(\tilde{\tau} |X|) < \infty$, it holds that there exist a $\tau' < \tilde{\tau}$ such that*

$$E \exp(\tau |X|) \leq \exp(2\tau E |X|)$$

for all $0 < \tau < \tau'$.

Proof. Lemma B4 in Aue et al. (2009) states that there exists a $\tau' < \tau$ such that

$$E \exp(\tau |X|) \leq 1 + 2\tau E(|X|),$$

for all $0 < \tau < \tau'$. By the elementary inequality $1 + y \leq \exp(y)$, valid for all y , it now follows that

$$E \exp(\tau |X|) \leq 1 + 2\tau E(|X|) \leq \exp(2\tau E(|X|)).$$

$\square$

We are now ready to present the main lemmas needed for our proofs.

Lemma F.3. *We assume that Assumption E.3 is satisfied. Then it holds that equation (E.4) admits the following unique, non-anticipative, strictly stationary and ergodic solution*

$$\text{vech}(\ln \bar{\mathbf{H}}_t - \mathbf{C}) = \sum_{s=0}^{\infty} \mathbf{A}^s u_{t-s},$$

which converges absolutely a.s. with probability 1.

Proof. The existence and the properties of the stationary solution follow from routine arguments given the sufficient condition in Assumption E.3(i) (see e.g. Brandt, 1986). Here, we show the absolutely summability of $\text{vech}(\ln \bar{\mathbf{H}}_t - \mathbf{C})$. It holds that

$$\|\text{vech}(\ln \bar{\mathbf{H}}_0 - \mathbf{C})\| \leq \sum_{s=0}^{\infty} \|\mathbf{A}\|^s \|u_{-s}\|.$$

On account of Assumption E.3(ii), the Borel-Cantelli Lemma entails that for all $\nu > 0$ and all $-\infty < s < \infty$, $\|u_s\| = O_{a.s.}(|s|^{1/\nu})$. Hence, there exist two random variables C_0 and s_0 such that

$$\begin{aligned} \sum_{s=0}^{\infty} \|\mathbf{A}\|^s \|u_{-s}\| &= \sum_{s=0}^{s_0-1} \|\mathbf{A}\|^s \|u_{-s}\| + \sum_{s=s_0}^{\infty} \|\mathbf{A}\|^s \|u_{-s}\| \\ &\leq \sum_{s=0}^{s_0-1} \|\mathbf{A}\|^s \|u_{-s}\| + C_0 \sum_{s=s_0}^{\infty} \|\mathbf{A}\|^s |s|^{1/\nu}. \end{aligned}$$

The first term is obviously $O_P(1)$; further, by Assumption E.3(i) it holds that

$$\sum_{s=s_0}^{\infty} \|\mathbf{A}\|^s |s|^{1/\nu} \leq \sum_{s=0}^{\infty} \|\mathbf{A}\|^s |s|^{1/\nu} \leq c_0,$$

whence the desired result obtains. □

Lemma F.4. *We assume that Assumption E.3 is satisfied, with part (ii) holding for some $\tau > (3/\sqrt{2})\nu$. Then it holds that $\bar{\mathbf{H}}_t^{1/2}$ is an $\mathcal{L}_\nu$ -decomposable Bernoulli shift with exponential rate function for all $\nu \geq 1$.*

Proof. Define

$$(F.3) \quad \bar{\mathbf{S}}_t = \mathbf{C} + \text{math} \left[\sum_{s=0}^{\infty} \mathbf{A}^s u_{t-s} \right],$$

where $\text{math}(\cdot)$ is the inverse of the operator $\text{vech}(\cdot)$ - that is, for any matrix $\mathbf{B}$, $\mathbf{B} = \text{math}(\text{vech}(\mathbf{B}))$. Clearly, the symmetry of $\mathbf{C}$ entails that $\bar{\mathbf{S}}_t$ is a symmetric matrix. Define also the coupling construction

$$\mathbf{S}_{t,\ell} = \mathbf{C} + \text{math} \left[\sum_{s=0}^{\ell} \mathbf{A}^s u_{t-s} \right] + \text{math} \left[\sum_{s=\ell+1}^{\infty} \mathbf{A}^s u_{t-s,\ell} \right],$$

where $\{u_{t,\ell}, -\infty < t < \infty\}$ is an *i.i.d.* sequence independent of $\{u_t, -\infty < t < \infty\}$ and such that $\{u_{t,\ell}, -\infty < t < \infty\} \stackrel{\mathcal{D}}{=} \{u_t, -\infty < t < \infty\}$, and let

$$(F.4) \quad \mathbf{H}_{t,\ell} = \text{vech}(\mathbf{S}_{t,\ell}).$$

Note, finally, that $\bar{\mathbf{H}}_t = \exp(\bar{\mathbf{S}}_t)$ and $\mathbf{H}_{t,\ell} = \exp(\mathbf{S}_{t,\ell})$. It holds that

$$\begin{aligned} & \left\| \bar{\mathbf{H}}_t^{1/2} - \mathbf{H}_{t,\ell}^{1/2} \right\| \\ &= \left\| \exp \left[\frac{1}{2} (\bar{\mathbf{S}}_t \pm \mathbf{S}_{t,\ell}) \right] - \exp \left[\frac{1}{2} (\mathbf{S}_{t,\ell}) \right] \right\| \\ &\leq \left\| \bar{\mathbf{S}}_t - \mathbf{S}_{t,\ell} \right\| \exp \left(\frac{1}{2} \|\mathbf{S}_{t,\ell}\| \right) \exp \left(\frac{1}{2} \|\bar{\mathbf{S}}_t - \mathbf{S}_{t,\ell}\| \right), \end{aligned}$$

having used Lemma F.1(iv). Hence, using Hölder's inequality and the Cauchy-Schwartz inequality, it holds that, for all $\nu \geq 1$

$$\begin{aligned} & \left| \bar{\mathbf{H}}_t^{1/2} - \mathbf{H}_{t,\ell}^{1/2} \right|_{\nu} \\ &\leq \left| \bar{\mathbf{S}}_t - \mathbf{S}_{t,\ell} \right|_{3\nu} \left| \exp \left(\frac{1}{2} \|\mathbf{S}_{t,\ell}\| \right) \right|_{3\nu} \left| \exp \left(\frac{1}{2} \|\bar{\mathbf{S}}_t - \mathbf{S}_{t,\ell}\| \right) \right|_{3\nu}. \end{aligned}$$

Using the inequality $\|\mathbf{A}\| \leq \sqrt{2} \|\text{vech}(\mathbf{A})\|$, we now have

$$\begin{aligned} & \left| \bar{\mathbf{S}}_t - \mathbf{S}_{t,\ell} \right|_{3\nu} \\ & \leq \sqrt{2} \left| \sum_{s=\ell+1}^{\infty} \mathbf{A}^s u_{t-s} - \sum_{s=\ell+1}^{\infty} \mathbf{A}^s u_{t-s,\ell} \right|_{3\nu} \leq 2\sqrt{2} \sum_{s=\ell+1}^{\infty} \|\mathbf{A}\|^s |u_0|_{3\nu} \leq c_0 \|\mathbf{A}\|^\ell, \end{aligned}$$

having used Assumptions E.3(i)-(ii). Also,

$$\begin{aligned} & \left| \exp \left(\frac{1}{2} \|\bar{\mathbf{S}}_t - \mathbf{S}_{t,\ell}\| \right) \right|_{3\nu} \\ & \leq \left| \exp \left(\frac{1}{\sqrt{2}} \|\text{vech}(\bar{\mathbf{S}}_t - \mathbf{S}_{t,\ell})\| \right) \right|_{3\nu} = \left| \exp \left(\frac{1}{\sqrt{2}} \left\| \sum_{s=\ell+1}^{\infty} \mathbf{A}^s (u_{t-s} - u_{t-s,\ell}) \right\| \right) \right|_{3\nu} \\ & \leq \left| \exp \left(\frac{1}{\sqrt{2}} \sum_{s=\ell+1}^{\infty} \|\mathbf{A}^s\| \|u_{t-s} - u_{t-s,\ell}\| \right) \right|_{3\nu} = \left[E \exp \left(\frac{3\nu}{\sqrt{2}} \sum_{s=\ell+1}^{\infty} \|\mathbf{A}\|^s \|u_{t-s} - u_{t-s,\ell}\| \right) \right]^{1/3\nu} \\ & = \left[\prod_{s=\ell+1}^{\infty} E \exp \left(\frac{3\nu}{\sqrt{2}} \|\mathbf{A}\|^s \|u_{t-s} - u_{t-s,\ell}\| \right) \right]^{1/3\nu} \\ & \leq \left[\prod_{s=\ell+1}^{\infty} \exp(c_1 \|\mathbf{A}\|^s) \right]^{1/3\nu} \leq \exp \left(c_2 \sum_{s=\ell+1}^{\infty} \|\mathbf{A}\|^s \right) \leq c_3, \end{aligned}$$

having used the independence of the u_t s in the fourth line, and Lemma F.2 in the fifth line.

By the same token (recalling the distributional equivalence between u_t and $u_{t,\ell}$)

$$\begin{aligned} \text{(F.5)} \quad & \left| \exp \left(\frac{1}{2} \|\mathbf{S}_{t,\ell}\| \right) \right|_{3\nu} \\ & \leq \left| \exp \left(\frac{1}{\sqrt{2}} \left\| \sum_{s=0}^{\infty} \mathbf{A}^s u_{t-s} \right\| \right) \right|_{3\nu} \leq \left[E \exp \left(\frac{3\nu}{\sqrt{2}} \sum_{s=0}^{\infty} \|\mathbf{A}\|^s \|u_{t-s}\| \right) \right]^{1/3\nu} \\ & = \left[\prod_{s=0}^{\infty} E \exp \left(\frac{3\nu}{\sqrt{2}} \|\mathbf{A}\|^s \|u_{t-s}\| \right) \right]^{1/3\nu} \leq \left[\prod_{s=0}^{\infty} \exp(c_1 \|\mathbf{A}\|^s) \right]^{1/3\nu} \\ & = \exp \left(c_2 \sum_{s=0}^{\infty} \|\mathbf{A}\|^s \right) \leq c_3. \end{aligned}$$

The desired result now follows. $\square$

Define now the stationary counterpart to ε_t in (E.3)

$$\bar{\varepsilon}_t = \bar{\mathbf{H}}_t^{1/2} \eta_t.$$

Lemma F.5. *We assume that Assumptions E.3 and E.4 are satisfied, with $\tau > (3/\sqrt{2}) p\epsilon / (1 + \epsilon)$ in Assumption E.3(ii). Then it holds that $\bar{\varepsilon}_t$ is an $\mathcal{L}_p$ -decomposable Bernoulli shift with rate function $O(\ell^{-b})$.*

Proof. Consider the construction

$$\varepsilon_{t,\ell} = \mathbf{H}_{t,\ell}^{1/2} \eta_{t,\ell},$$

where $H_{t,\ell}$ is defined in (F.4) and $\eta_{t,\ell}$ is defined similarly. Then it holds that

$$\begin{aligned} & |\bar{\varepsilon}_t - \varepsilon_{t,\ell}|_p \\ &= \left| \bar{\mathbf{H}}_t^{1/2} \eta_t - \mathbf{H}_{t,\ell}^{1/2} \eta_{t,\ell} \right|_p = \left| \left(\bar{\mathbf{H}}_t^{1/2} \pm \mathbf{H}_{t,\ell}^{1/2} \right)^{1/2} \eta_t - \mathbf{H}_{t,\ell}^{1/2} \eta_{t,\ell} \right|_p \\ &\leq \left| \left(\bar{\mathbf{H}}_t^{1/2} - \mathbf{H}_{t,\ell}^{1/2} \right) \eta_t \right|_p + \left| \mathbf{H}_{t,\ell}^{1/2} (\eta_t - \eta_{t,\ell}) \right|_p \\ &\leq \left| \bar{\mathbf{H}}_t^{1/2} - \mathbf{H}_{t,\ell}^{1/2} \right|_{p\epsilon/(1+\epsilon)} |\eta_t|_{p(1+\epsilon)} + \left| \mathbf{H}_{t,\ell}^{1/2} \right|_{p\epsilon/(1+\epsilon)} |\eta_t - \eta_{t,\ell}|_{p(1+\epsilon)} \\ &\leq c_0 \|\mathbf{A}\|^\ell + c_1 \left| \mathbf{H}_{t,\ell}^{1/2} \right|_{p\epsilon/(1+\epsilon)} \ell^{-b}, \end{aligned}$$

having used Lemma F.4. We now show that $\left| \mathbf{H}_{t,\ell}^{1/2} \right|_{p\epsilon/(1+\epsilon)} \leq c_0$; indeed, seeing as $\mathbf{H}_{t,\ell} \stackrel{\mathcal{D}}{=} \bar{\mathbf{H}}_t$, this is tantamount to showing that $\left| \bar{\mathbf{H}}_t \right|_{p\epsilon/(1+\epsilon)}$ is finite. It holds that

$$\begin{aligned} & \left| \bar{\mathbf{H}}_t \right|_{p\epsilon/(1+\epsilon)} \\ &= \left| \exp \left(\frac{1}{2} \bar{\mathbf{S}}_t \right) \right|_{p\epsilon/(1+\epsilon)} \leq \left| \exp \left(\frac{1}{2} \|\bar{\mathbf{S}}_t\| \right) \right|_{p\epsilon/(1+\epsilon)} \\ &\leq \left| \exp \left(\frac{1}{\sqrt{2}} \left\| \text{vech } \mathbf{C} + \sum_{s=0}^{\infty} \mathbf{A}^s u_{t-s} \right\| \right) \right|_{p\epsilon/(1+\epsilon)} \leq \left| c_0 \exp \left(\frac{1}{\sqrt{2}} \sum_{s=0}^{\infty} \|\mathbf{A}\|^s \|u_{t-s}\| \right) \right|_{p\epsilon/(1+\epsilon)} \end{aligned}$$

$$= \left| c_0 \prod_{s=0}^{\infty} E \exp \left(\frac{p\epsilon}{\sqrt{2}(1+\epsilon)} \|\mathbf{A}\|^s \|u_{t-s}\| \right) \right|^{(1+\epsilon)/p\epsilon} \leq c_1,$$

by similar arguments as above (note the use of Lemma F.1(v) in the second line). The desired result now obtains. $\square$

Define now the stationary counterpart to β_t , viz.

$$\bar{\beta}_t = \mu + \sum_{s=0}^{\infty} \Gamma^s \bar{\varepsilon}_{t-s}.$$

Lemma F.6. *We assume that Assumptions E.2-E.4 are satisfied, with $\tau > (3/\sqrt{2}) p\epsilon / (1+\epsilon)$ in Assumption E.3(ii). Then it holds that $\bar{\beta}_t$ is an $\mathcal{L}_p$ -decomposable Bernoulli shift with rate function $O(\ell^{-b})$.*

Proof. Lemma F.5 entails that we can write

$$\bar{\varepsilon}_t = g(\zeta_t, \zeta_{t-1}, \dots),$$

with $g : \mathbb{S}^{s \times \infty} \rightarrow \mathbb{R}^k$ a nonrandom function, $\mathbb{S}$ a measurable space and $\{\zeta_t, -\infty < t < \infty\}$ an *i.i.d.* sequence taking values in $\mathbb{R}^s$ for some $s \geq 1$. Let

$$\bar{\beta}_{t,\ell} = \mu + \sum_{s=0}^{\infty} \Gamma^s \bar{\varepsilon}_{t-s,\ell},$$

where

$$\bar{\varepsilon}_{t-s,\ell} = g(\zeta_{t-s}, \dots, \zeta_{t-s-\ell}, \zeta'_{t-s-\ell-1}, \zeta'_{t-s-\ell-2}, \dots),$$

with $\{\zeta'_t, -\infty < t < \infty\}$ an *i.i.d.* sequence with $\zeta'_t \stackrel{\mathcal{D}}{=} \zeta_t$ whenever $s < \ell$, and

$$\bar{\varepsilon}_{t-s,\ell} = g(\zeta'_{t-s}, \zeta'_{t-s-1}, \dots),$$

otherwise. Then, by similar arguments as above

$$\begin{aligned}
& \left| \bar{\beta}_t - \bar{\beta}_{t,\ell} \right|_p \\
& \leq \sum_{s=0}^{\ell} \|\Gamma\|^s |\bar{\varepsilon}_{t-s} - \bar{\varepsilon}_{t-s,\ell}|_p + \sum_{s=\ell+1}^{\infty} \|\Gamma\|^s |\bar{\varepsilon}_{t-s} - \bar{\varepsilon}_{t-s,\ell}|_p \\
& \leq c_0 |\bar{\varepsilon}_0 - \bar{\varepsilon}_{0,\ell}|_p + c_1 \|\Gamma\|^\ell \leq c_2 \max \left\{ \ell^{-b}, \|\Gamma\|^\ell \right\},
\end{aligned}$$

which concludes the proof. $\square$

Lemma F.7. *We assume that Assumption E.3 is satisfied, with part (ii) holding for some $\tau > (3/\sqrt{2})\nu$. Then it holds that*

$$\left| \mathbf{H}_t^{1/2} - \bar{\mathbf{H}}_t^{1/2} \right|_\nu \leq c_0 \|\mathbf{A}\|^t,$$

for all $\nu \geq 1$.

Proof. By routine calculations

$$\text{vech}(\ln \mathbf{H}_t - \mathbf{C}) = \mathbf{A}^t \text{vech}(\ln \mathbf{H}_0 - \mathbf{C}) + \sum_{s=0}^{t-1} \mathbf{A}^s u_{t-s};$$

hence we define

$$\mathbf{S}_t = \ln \mathbf{H}_t = \mathbf{C} [\mathbf{I} - \mathbf{A}^t] + \text{math} [\mathbf{A}^t \text{vech}(\ln \mathbf{H}_0 - \mathbf{C})] + \text{math} \left[\sum_{s=0}^{t-1} \mathbf{A}^s u_{t-s} \right].$$

It now holds that

$$\begin{aligned}
& \left\| \mathbf{H}_t^{1/2} - \bar{\mathbf{H}}_t^{1/2} \right\| \\
& \leq \left\| \mathbf{S}_t - \bar{\mathbf{S}}_t \right\| \exp \left(\frac{1}{2} \left\| \bar{\mathbf{S}}_t \right\| \right) \exp \left(\frac{1}{2} \left\| \mathbf{S}_t - \bar{\mathbf{S}}_t \right\| \right),
\end{aligned}$$

where $\bar{S}_t$ is defined in (F.3). Thus

$$\begin{aligned} & \left| \mathbf{H}_t^{1/2} - \bar{\mathbf{H}}_t^{1/2} \right|_{\nu} \\ & \leq \left| \mathbf{S}_t - \bar{\mathbf{S}}_t \right|_{3\nu} \left| \exp \left(\frac{1}{2} \|\bar{\mathbf{S}}_t\| \right) \right|_{3\nu} \left| \exp \left(\frac{1}{2} \|\mathbf{S}_t - \bar{\mathbf{S}}_t\| \right) \right|_{3\nu}. \end{aligned}$$

We now have

$$\left| \mathbf{S}_t - \bar{\mathbf{S}}_t \right|_{3\nu} \leq c_0 \left(\|\mathbf{A}\|^t + \|\mathbf{A}\|^t |\ln \mathbf{H}_0|_{3\nu} + \left| \sum_{s=t}^{\infty} \mathbf{A}^s u_{t-s} \right|_{3\nu} \right) \leq c_0 \|\mathbf{A}\|^t,$$

by a similar logic to the above. We already know from (F.5) that $\left| \exp \left(\frac{1}{2} \|\bar{\mathbf{S}}_t\| \right) \right|_{3\nu}$ is bounded; finally, following the same passages as above, it can also be shown that $\left| \exp \left(\frac{1}{2} \|\mathbf{S}_t - \bar{\mathbf{S}}_t\| \right) \right|_{3\nu}$ is bounded. $\square$

Lemma F.8. *We assume that Assumptions E.2-E.4 are satisfied, with $\tau > (3/\sqrt{2}) p\epsilon / (1 + \epsilon)$ in Assumption E.3(ii). Then it holds that*

$$\left| \beta_t - \bar{\beta}_t \right|_p \leq c_0 \max \{ \|\mathbf{\Gamma}\|^t, \|\mathbf{A}\|^t \}.$$

Proof. Note that

$$\beta_t = \mu + \mathbf{\Gamma}^t (\beta_0 - \mu) + \sum_{s=0}^{t-1} \mathbf{\Gamma}^s \varepsilon_{t-s},$$

and

$$\left\| \beta_t - \bar{\beta}_t \right\| \leq \|\mathbf{\Gamma}\|^t (\|\beta_0\| + \|\mu\|) + \sum_{s=0}^{t-1} \|\mathbf{\Gamma}\|^s \|\varepsilon_{t-s} - \bar{\varepsilon}_{t-s}\| + \sum_{s=t}^{\infty} \|\mathbf{\Gamma}\|^s \|\bar{\varepsilon}_{t-s}\|.$$

Hence, we have

$$\left| \beta_t - \bar{\beta}_t \right|_p \leq c_0 \|\mathbf{\Gamma}\|^t + c_1 \|\varepsilon_t - \bar{\varepsilon}_t\|_p.$$

Finally

$$\|\varepsilon_t - \bar{\varepsilon}_t\|_p$$

$$= \left| \mathbf{H}_t^{1/2} \eta_t - \overline{\mathbf{H}}_t^{1/2} \eta_t \right|_p \leq |\eta_t|_{p(1+\epsilon)} \left| \mathbf{H}_t^{1/2} - \overline{\mathbf{H}}_t^{1/2} \right|_{p\epsilon/(1+\epsilon)} \leq c_0 \|\mathbf{A}\|^t,$$

by Lemma F.7. □

We now show that the stationary solution is a good approximation of y_t . The first result pertains to (E.1)-(E.4).

Lemma F.9. *We assume that Assumptions E.1-E.4 are satisfied, with $\tau > (3/\sqrt{2}) p\epsilon / (1 + \epsilon)$ in Assumption E.3(ii). Then it holds that*

$$|y_t - \bar{y}_t|_{\min\{p, p'\}} \leq c_0 \max \{ \|\Gamma\|^t, \|\mathbf{A}\|^t \}.$$

Proof. The proof follows immediately upon writing

$$y_t - \bar{y}_t = x'_t (\beta_t - \bar{\beta}_t),$$

and using Lemma F.8. □

The next lemma extends the result above to model (E.2)-(E.5).

Lemma F.10. *We assume that the assumptions of Theorem E.1 and Assumption E.5 are satisfied. Then it holds that $|y_t - \bar{y}_t|_{\bar{p}}$ - with $\bar{y}_t$ defined as the stationary solution of (E.2)-(E.5) - drifts exponentially fast to zero as $t \rightarrow \infty$.*

Proof. Consider the recursive solution

$$\mathbf{Y}_t = \left(\prod_{j=0}^{t-1} \mathbf{B}_j^* \right) \mathbf{Y}_0 + \sum_{k=0}^{t-1} \left(\prod_{j=t-k+1}^t \mathbf{B}_j^* \right) \mathbf{E}_{t-k},$$

where $\mathbf{B}_j^*$ is defined in the same way as $\mathbf{B}_j$ but with $\beta_{j,t}$ instead of $\bar{\beta}_{j,t}$. It holds that

$$\mathbf{Y}_t - \bar{\mathbf{Y}}_t$$

$$\begin{aligned}
&= \left(\prod_{j=0}^{t-1} \mathbf{B}_j^* \right) \mathbf{Y}_0 + \sum_{k=t}^{\infty} \left(\prod_{j=t-k+1}^t \mathbf{B}_j \right) \mathbf{E}_{t-k} \\
&\quad + \sum_{k=0}^{t-1} \left[\left(\prod_{j=t-k+1}^t \mathbf{B}_j^* \right) - \left(\prod_{j=t-k+1}^t \mathbf{B}_j \right) \right] \mathbf{E}_{t-k} \\
&= I + II + III.
\end{aligned}$$

Finally, by Assumption [E.5\(iii\)](#), we have

$$\left| \left(\prod_{j=0}^{t-1} \mathbf{B}_j^* \right) \mathbf{Y}_0 \right|_{\tilde{p}} \leq |\mathbf{Y}_0|_{\tilde{p}} \prod_{j=0}^{t-1} |\mathbf{B}_j^*|_{\tilde{p}t} \leq c_0 \exp(-c_1 t).$$

Similar arguments as in the proof of Theorem [E.2](#) also yield that $|II|_{\tilde{p}} \leq c_0 \exp(-c_1 t)$.

Finally, the same result can be shown for III upon using Lemma [F.8](#). $\square$

G. PROOFS

Proof of Theorem E.1. The proof is very similar to the proofs in the previous section. Define the coupling constructions $x_{t,\ell}$ and $\bar{\beta}_{t,\ell}$, and

$$y_{t,\ell} = x'_{t,\ell} \bar{\beta}_{t,\ell} + \nu_{t,\ell}.$$

Then we have

$$\begin{aligned} & |\bar{y}_t - y_{t,\ell}|_{\tilde{p}} \\ & \leq |x'_{t,\ell} (\bar{\beta}_t - \bar{\beta}_{t,\ell})|_{\tilde{p}} + |\bar{\beta}_t (x_t - x_{t,\ell})|_{\tilde{p}} + |\nu_t - \nu_{t,\ell}|_{\tilde{p}} \\ & = |x_{t,\ell}|_{\tilde{p}} |\bar{\beta}_t - \bar{\beta}_{t,\ell}|_{\tilde{p}} + |\bar{\beta}_t|_{\tilde{p}} |x_t - x_{t,\ell}|_{\tilde{p}} + |\nu_t - \nu_{t,\ell}|_{\tilde{p}}, \end{aligned}$$

whence the final result obtains. □

Proof of Theorem E.2. The proof is similar to those of Lemmas B.6 and B.7 in [Horváth and Trapani \(2023\)](#), with the (important) difference that the random coefficients β_t are serially dependent. Consider the following Markovian form of (E.5)

$$(G.1) \quad \mathbf{Y}_t = \mathbf{B}_t \mathbf{Y}_{t-1} + \mathbf{E}_t,$$

where $\mathbf{Y}_t = (y_t, \dots, y_{t-p+1})'$, $\mathbf{E}_t = (\nu_t, 0, \dots, 0)'$ and $\mathbf{B}_t$ is defined in (E.7). Let

$$(G.2) \quad \bar{\mathbf{Y}}_t = \sum_{k=0}^{\infty} \left(\prod_{j=t-k+1}^t \mathbf{B}_j \right) \mathbf{E}_{t-k};$$

and consider

$$\inf_{n \in \mathbb{N}} E \left(\frac{1}{n+1} \ln \left\| \prod_{j=-n}^0 \mathbf{B}_j \right\| \right) \leq \inf_{n \in \mathbb{N}} E \left(\frac{1}{n+1} \sum_{j=-n}^0 \ln \|\mathbf{B}_j\| \right) = E \ln \|\mathbf{B}_0\| < 0,$$

by stationarity and Assumption E.5(i). Then, by Brandt (1986), $\bar{\mathbf{Y}}_t$ converges a.s. and it is the unique strictly stationary solution of (G.1). Consider now

$$\bar{\mathbf{Y}}_{t,\ell} = \sum_{k=0}^{\ell} \left(\prod_{j=t-k+1}^t \mathbf{B}_j \right) \mathbf{E}_{t-k} + \sum_{k=\ell+1}^{\infty} \left(\prod_{j=t-k+1}^t \tilde{\mathbf{B}}_j \right) \mathbf{E}'_{t-k},$$

where $\mathbf{E}'_{t-k}$ is a coupling of $\mathbf{E}_{t-k}$, and $\tilde{\mathbf{B}}_j$ is defined so as to be equal to $\mathbf{B}_j$ whenever $j \leq \ell$ and to its coupling $\mathbf{B}'_j$ for $j \geq \ell + 1$. Seeing as

$$\|\bar{\mathbf{Y}}_t - \bar{\mathbf{Y}}_{t,\ell}\| \leq \sum_{k=\ell+1}^{\infty} \prod_{j=t-k+1}^t \|\tilde{\mathbf{B}}_j\| \|\mathbf{E}_{t-k} - \mathbf{E}'_{t-k}\|,$$

it follows that

$$\begin{aligned} & |\bar{\mathbf{Y}}_t - \bar{\mathbf{Y}}_{t,\ell}|_{\tilde{p}} \\ & \leq c_0 \sum_{k=\ell+1}^{\infty} \prod_{j=t-k+1}^t |\tilde{\mathbf{B}}_j|_{\tilde{p}k} = c_0 \sum_{k=\ell+1}^{\infty} \exp \left(\sum_{j=t-k+1}^t \ln |\tilde{\mathbf{B}}_j|_{\tilde{p}k} \right). \end{aligned}$$

Using Assumption E.5(ii), let $\gamma = \ln \bar{\kappa} < 0$, and rewrite the above as

$$\begin{aligned} & |\bar{\mathbf{Y}}_t - \bar{\mathbf{Y}}_{t,\ell}|_{\tilde{p}} \\ & \leq c_0 \sum_{k=\ell+1}^{\infty} \exp \left(\sum_{j=t-k+1}^t \left(\ln |\tilde{\mathbf{B}}_j|_{\tilde{p}k} \pm \gamma/2 \right) \right) \\ & \leq c_1 \exp(\gamma\ell/2) \sum_{k=\ell+1}^{\infty} \exp \left(\sum_{j=t-k+1}^t \left(\ln |\tilde{\mathbf{B}}_j|_{\tilde{p}k} - \gamma/2 \right) \right) \leq c_2 \exp(-c_3\ell). \end{aligned}$$

□

Proof of Corollary E.1. Recall $\bar{y}_t$ defined in (E.6). It holds that

$$\left| \sum_{t=1}^T y_t \right|_{\tilde{p}}^{\tilde{p}} \leq c_0 \left(\left| \sum_{t=1}^T \bar{y}_t \right|_{\tilde{p}}^{\tilde{p}} + \left| \sum_{t=1}^T (y_t - \bar{y}_t) \right|_{\tilde{p}}^{\tilde{p}} \right).$$

Theorem E.1 entails that $\bar{y}_t$ is an $\mathcal{L}_{\tilde{p}}$ -decomposable Bernoulli shift, where recall that $\tilde{p} > 2$; hence, by Proposition 4.1 in Berkes et al. (2011), it follows that $\left| \sum_{t=1}^T \bar{y}_t \right|_{\tilde{p}}^{\tilde{p}} \leq c_{\tilde{p}} T^{\tilde{p}/2}$, for some $c_{\tilde{p}} < \infty$ depending only on $\tilde{p}$. Note also that, using Minkowski's inequality

$$\left| \sum_{t=1}^T (y_t - \bar{y}_t) \right|_{\tilde{p}} \leq \sum_{t=1}^T |y_t - \bar{y}_t|_{\tilde{p}} \leq c_0 |y_t - \bar{y}_t|_{\min\{p', p\}} \leq c_0 \sum_{t=1}^T \max \{ \|\mathbf{A}\|^t, \|\mathbf{\Gamma}\|^t \} \leq c_1,$$

where we have used Lemma F.9 (or Lemma F.10) in the last passage. The desired result now follows immediately. $\square$

Proof of Corollary E.2. We only show (E.10); (E.11) can be shown using exactly the same logic, *mutatis mutandis*. Let $0 < \zeta < 1/2$ be a yet unspecified number. It holds that

$$\begin{aligned} & \max_{1 \leq k \leq T/2} \frac{1}{k^\zeta} \left| \sum_{t=1}^k y_t - W_{T,1}(k) \right| \\ & \leq \max_{1 \leq k \leq T/2} \frac{1}{k^\zeta} \left| \sum_{t=1}^k \bar{y}_t - W_{T,1}(k) \right| + \max_{1 \leq k \leq T/2} \frac{1}{k^\zeta} \left| \sum_{t=1}^k (y_t - \bar{y}_t) \right|, \end{aligned}$$

where recall that $\bar{y}_t$ is defined in (E.6). Using the results in Aue et al. (2014), it holds that

$$\max_{1 \leq j \leq k} \left| \sum_{t=1}^j \bar{y}_t - W_{T,1}(j) \right| = O_{a.s.}(k^\zeta),$$

for some $\zeta < 1/2$. This entails that

$$\left| \sum_{t=1}^k \bar{y}_t - W_{T,1}(k) \right| = O_{a.s.}(k^\zeta);$$

in turn, this entails that there exist two random variables C and k_0 such that, for all $k \geq k_0$

$$\left| \sum_{t=1}^k \bar{y}_t - W_{T,1}(k) \right| \leq C k^\zeta.$$

Therefore we may write

$$\begin{aligned}
& \max_{1 \leq k \leq T/2} \frac{1}{k^\zeta} \left| \sum_{t=1}^k \bar{y}_t - W_{T,1}(k) \right| \\
& \leq \max_{1 \leq k \leq k_0-1} \frac{1}{k^\zeta} \left| \sum_{t=1}^k \bar{y}_t - W_{T,1}(k) \right| + \max_{k_0 \leq k \leq T/2} \frac{1}{k^\zeta} \left| \sum_{t=1}^k \bar{y}_t - W_{T,1}(k) \right| \\
& \leq \max_{1 \leq k \leq k_0-1} \frac{1}{k^\zeta} \left| \sum_{t=1}^k \bar{y}_t - W_{T,1}(k) \right| + C_0.
\end{aligned}$$

It is now not hard to see that

$$\mathbb{P} \left(\max_{1 \leq k \leq k_0-1} \frac{1}{k^\zeta} \left| \sum_{t=1}^k \bar{y}_t - W_{T,1}(k) \right| + C_0 < \infty \right) = 1,$$

whence it follows that

$$\max_{1 \leq k \leq T/2} \frac{1}{k^\zeta} \left| \sum_{t=1}^k \bar{y}_t - W_{T,1}(k) \right| = O_P(1).$$

Also, letting $\lceil a \rceil$ denote the largest integer after a , by standard arguments we receive

$$\begin{aligned}
& \mathbb{P} \left(\max_{1 \leq k \leq T/2} \frac{1}{k^\zeta} \left| \sum_{t=1}^k (y_t - \bar{y}_t) \right| > \epsilon \right) \\
& \leq \mathbb{P} \left(\max_{0 \leq \ell \leq \lceil \ln(T/2) \rceil} \max_{\exp(\ell) \leq k \leq \exp(\ell+1)} \frac{1}{k^\zeta} \left| \sum_{t=1}^k (y_t - \bar{y}_t) \right| > \epsilon \right) \\
& \leq \sum_{\ell=0}^{\lceil \ln(T/2) \rceil} \mathbb{P} \left(\max_{\exp(\ell) \leq k \leq \exp(\ell+1)} \frac{1}{k^\zeta} \left| \sum_{t=1}^k (y_t - \bar{y}_t) \right| > \epsilon \right) \\
& \leq \sum_{\ell=0}^{\lceil \ln(T/2) \rceil} \mathbb{P} \left(\max_{\exp(\ell) \leq k \leq \exp(\ell+1)} \left| \sum_{t=1}^k (y_t - \bar{y}_t) \right| > \epsilon \exp(\zeta(\ell+1)) \right) \\
& \leq \sum_{\ell=0}^{\lceil \ln(T/2) \rceil} \mathbb{P} \left(\sum_{t=1}^{\exp(\ell+1)} |y_t - \bar{y}_t| > \epsilon \exp(\zeta(\ell+1)) \right)
\end{aligned}$$

$$\begin{aligned}
&\leq \epsilon^{-1} \sum_{\ell=0}^{\lceil \ln(T/2) \rceil} \exp(-\zeta(\ell+1)) \sum_{t=1}^{\exp(\ell+1)} \mathbb{E} |y_t - \bar{y}_t| \\
&\leq c_0 \epsilon^{-1} \sum_{\ell=0}^{\lceil \ln(T/2) \rceil} \exp(-\zeta(\ell+1)) \left(\sum_{t=1}^{\exp(\ell+1)} \max \{ \|\mathbf{A}\|^t, \|\mathbf{\Gamma}\|^t \} \right) \leq c_1,
\end{aligned}$$

for all $\zeta > 0$, having used Bonferroni inequality in the third line, and Lemma F.9 (or Lemma F.10) in the last line. These passages entail that

$$\max_{1 \leq k \leq T/2} \frac{1}{k^\zeta} \left| \sum_{t=1}^k (y_t - \bar{y}_t) \right| = O_P(1),$$

for all $\zeta > 0$. Finally, using Lemma 1 in Characiejus et al. (2025) and Theorem 2.1 in Berkes et al. (2014), it follows that $\zeta = 1/\tilde{p}$. The independence between $\{W_{T,1}(k), 1 \leq k \leq T/2\}$ and $\{W_{T,2}(k), 1 \leq k \leq T/2\}$ follows directly from the fact that the proofs in Aue et al. (2014) use the blocking argument. $\square$

REFERENCES

- Andrieu, C., A. Doucet, and R. Holenstein (2010). Particle Markov Chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B* 72(3), 269–342.
- Aue, A., S. Hörmann, L. Horváth, and M. Hušková (2014). Dependent functional linear models with applications to monitoring structural change. *Statistica Sinica*, 1043–1073.
- Aue, A., S. Hörmann, L. Horváth, and M. L. Reimherr (2009). Break detection in the covariance structure of multivariate time series models. *Annals of Statistics* 37(6B), 4046–4087.
- Aue, A., L. Horváth, and J. Steinebach (2006). Estimation in random coefficient autoregressive models. *Journal of Time Series Analysis* 27(1), 61–76.
- Barigozzi, M. and L. Trapani (2022). Testing for common trends in nonstationary large datasets. *Journal of Business & Economic Statistics* 40(3), 1107–1122.
- Berkes, I., S. Hörmann, and J. Schauer (2011). Split invariance principles for stationary processes. *Annals of Probability* 39(6), 2441–2473.
- Berkes, I., W. Liu, and W. B. Wu (2014). Komlós–Major–Tusnády approximation under dependence. *Annals of Probability* 42(2), 794–817.
- Bernstein, D. S. (2009). *Matrix mathematics: Theory, facts, and formulas*. Princeton university press.
- Brandt, A. (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Advances in Applied Probability* 18(1), 211–220.
- Characiejus, V., P. Kokoszka, and X. Meng (2025). Estimation of the long-run variance of nonlinear time series with an application to change point analysis. *Journal of Time Series Analysis*.
- Chopin, N. and S. S. Singh (2015). On particle Gibbs sampling. *Bernoulli* 21(3), 1855–1883.
- Creal, D. (2012). A survey of sequential Monte Carlo methods for economics and finance. *Econometric Reviews* 31(3), 245–296.
- Creal, D. and R. S. Tsay (2015). High dimensional dynamic stochastic copula models. *Journal of Econometrics* 189(2), 335–345.
- Fearnhead, P., O. Papaspiliopoulos, G. O. Roberts, and A. Stuart (2010). Random-weight particle filtering of continuous time processes. *Journal of the Royal Statistical Society: Series B* 72(4), 497–512.

- Godsill, S. J., A. Doucet, and M. West (2004). Monte Carlo smoothing for nonlinear time series. *Journal of the American Statistical Association* 99(465), 156–168.
- Horváth, L. and L. Trapani (2016). Statistical inference in a random coefficient panel model. *Journal of Econometrics* 193(1), 54–75.
- Horváth, L. and L. Trapani (2023). Changepoint detection in heteroscedastic random coefficient autoregressive models. *Journal of Business & Economic Statistics* 41(4), 1300–1314.
- Ibragimov, I. A. (1962). Some limit theorems for stationary processes. *Theory of Probability & Its Applications* 7(4), 349–382.
- Kokoszka, P., N. Mohammadi, and H. Wang (2025). Dependence properties of stochastic volatility models. *Journal of Time Series Analysis* 46(3), 421–431.
- Koopman, S. J. and M. Scharth (2012). The analysis of stochastic volatility in the presence of daily realized measures. *Journal of Financial Econometrics* 11(1), 76–115.
- Liu, W. and Z. Lin (2009). Strong approximation for a class of stationary processes. *Stochastic Processes and their Applications* 119(1), 249–280.
- Stroud, J. R. and M. S. Johannes (2014). Bayesian modeling and forecasting of 24-hour high-frequency volatility. *Journal of the American Statistical Association* 109(508), 1368–1384.
- Van Loan, C. (1977). The sensitivity of the matrix exponential. *SIAM Journal on Numerical Analysis* 14(6), 971–981.
- Wu, W. B. (2005). Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences* 102(40), 14150–14154.