Volatility forecasting with higher-order stochastic volatility models *

Md. Nazmul Ahsan† CIRANO
Jean-Marie Dufour ‡ McGill University
Gabriel Rodriguez-Rondon § McGill University

August 27, 2021

Abstract

We study the performance of higher-order stochastic volatility [SV(p)] models in forecasting volatility. This class of models provides more flexibility to represent volatility persistence and heavy tails and are natural extensions of the first-order model. The estimation of SV(p) models constitutes challenging problems due to the inherent difficulty of evaluating the likelihood function. Most of the proposed ones are inflexible, inefficient, and computationally expensive. We employ OLS-based winsorized ARMA estimator of **?** [which is computationally inexpensive and remarkably accurate], and present empirical applications relating to SV(p) models. Using different volatility proxies (the squared return of S&P 500 index and the realized volatility of S&P 500, FTSE100, NASDAQ100, N225, SSMI20 indices), we conduct two out-of-sample forecast experiments: (1) we forecast a moderately volatile period after the late-2000s financial crisis; (2) we forecast a highly volatile period, *i.e.*, the core financial crisis. We compare the accuracy of volatility forecasts among SV(p) models, GARCH models, and Heterogenous Autoregressive model of Realized Volatility (HAR-RV) models. The results suggest that SV(p) models perform better than other models in most cases. This finding holds even if a high volatility period (such as financial crisis) is included in the estimation sample or the forecasted sample. Formal prediction tests, *i.e.*, model confidence set procedure, also support these inferences. Our findings highlight the usefulness of higher-order SV models for volatility forecasting.

JEL Classification: C15, C22, C53, C58.

Keywords: Stochastic volatility, realized variance, volatility forecasting, high-frequency data.

* The authors thank Torben Andersen, Manabu Asai, Richard Baillie, Russell Davidson, Firmin Doko Tchatoka, John Galbraith, René Garcia, Byunguk Kang, Lynda Khalaf (Discussant), Jihyun Kim (Discussant), Mehdi Khorram (Discussant), Jinjing Liu, Richard Luger, Franck Moraux, Denis Pelletier (Discussant), Hashem Pesaran, Eric Renault, Lars Stentoft, Masaya Takano, Purevdorj Tuvaandorj, Aman Ullah, Pascale Valéry, Victoria Zinde-Walsh, as well as the seminar and conference participants for useful comments and constructive discussions. Earlier versions of this paper were presented at 2021 Joint Statistical Meetings, 7th RCEA Time Series Workshop, 60th Annual Southwestern Finance Association (SWFA) Conference, 21st IWH-CIREQ-GW Macroeconometric Workshop (Forecasting and Uncertainty), 2020 Bernoulli-IMS Symposium, 2019 French Econometrics Conference, 2019 NBER-NSF Time Series Conference, CityU HK 2019, IAAE 2019, AFFI 2019, Université Laval, 2019 BU Pi-day Conference, CESG 2018, ITISE 2018, IAAE 2018, CEA 2018, 2018 CIREQ-Montreal EC (Recent Advances in the Method of Moments), University of Bergamo, University of Southern California, 2018 New York Camp Econometrics, York University, CIREQ-McGill Lunch Seminar (2016, 2017), ESWC 2015, and 11th CIREQ PhD Students' Conference. This paper was previously circulating under titles "Volatility forecasting and option pricing with higher-order stochastic volatility models" and "Simple estimators for higher-order stochastic volatility models and forecasting."

† Centre interuniversitaire de recherche en analyse des organisations (CIRANO). Mailing address: 1130 Sherbrooke West, Suite 1400, Montreal, Quebec, H3A 2M8, Canada; e-mail: ahsann@cirano.qc.ca.

‡ William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 414, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 6071; FAX: (1) 514 398 4800; e-mail: jean-marie.dufour@mcgill.ca. Web page: http://www.jeanmariedufour.com

§ Ph.D. Candidate, Department of Economics, McGill University and Centre interuniversitaire de recherche en analyse des organisations (CIRANO). Mailing address: Department of Economics, McGill University. e-mail: gabriel.rodriguezrondon@mail.mcgill.ca. Wed page: https://grodriguezrondon.com

1 Introduction

Modelling and forecasting time-varying volatility of asset returns are pivotal in many areas of financial decision making, especially in investment theory, option pricing and risk management. Two main classes of parametric models have been proposed in the literature to estimate and forecast dynamic volatility: (1) GARCH-type models [**?**, **?**]; (2) stochastic volatility (SV) models [**?** (**?**, **?**)]. The main distinction between GARCH and SV models is that the variance process of the latter has an additional error term which captures the effect of new information coming to the market, so that conditional on the information set $\mathcal{F}_{t-1}$, volatility σ_t^2 is not known in SV models but rather an unobserved random variable.[1] In this paper, we consider higher-order stochastic volatility [SV(p)] models for forecasting volatility.

SV models may be preferable to their counterpart GARCH-type models for several reasons. *First*, SV models are discrete-time formulations of continuous-time diffusion processes used in theoretical finance for derivative pricing and portfolio optimization; see **?**, **?**. *Second*, SV models do not appear to require various *ad hoc* adjustments, like the addition of a random jump component or non-Gaussian innovations. These modifications improve the performance of standard GARCH models, but do not appear necessary with SV models; see **?**, **?**. *Third*, SV models often provide more accurate volatility forecasts than GARCH models, indicating that the time-varying volatility is better modelled as a latent stochastic process; see **?**, **?**, **?**, **?**. *Finally*, it is easy to derive the probabilistic properties (stationarity, ergodicity and mixing) of SV models than GARCH models; see **?**. In contrast, the stationarity of a GARCH process is difficult to establish; see **?**, **?**, **?**.

Despite these attractive features, the estimation of SV models is much more complicated than it is for GARCH-type models. In particular, due to the presence of latent variables, likelihood-based methods are difficult to apply, and statistical inference (estimation and testing) for SV models is quite challenging. Consequently, a variety of methods have been proposed to estimate SV(1) model, where the latent volatility process is modelled as a first-order autoregression. These include: quasi-maximum likelihood (QML) [**?**, **?**, **?**], the generalized method of moments (GMM) [**?**, **?**], the simulated method of moments (SMM) [**?**, **?**, **?**], Monte Carlo likelihood (MCL) [**?**], simulated maximum likelihood (SML) [**?**, **?** (**?**, **?**), **?**], the method based on linear representation [**?**], closed-form moment-based estimators [**?** (**?**, **?**), **?**], and Bayesian techniques based on Markov Chain Monte Carlo (MCMC) methods [**?**, **?**, **?**, **?**, **?**]. The vast majority of these are either computer-intensive and/or inefficient. [2]

[1] Several reviews of GARCH and SV literature are available; for GARCH, see **?**, and for SV, see **?**, **?**, and **?**. SV models are also common in macroeconomic modelling; see **?**, **?**, **?**, **?**, **?**, and **?**.

[2] Apart from the closed-form moment-based estimators, the above estimation methods are based on simulation techniques and/or numerical optimization. Simulation-based methods such as SML, MCL, SMM, and Bayesian MCMC methods [via the Metropolis-Hastings algorithm or the Gibbs sampler] are computer-intensive, inflexible across models, hard to implement in practice, and may converge very slowly; see **?**. Implementing these methods requires one to choose a sampling scheme, initial parameters, and an auxiliary model (which is largely conventional). The choice of initial parameter values for QML, GMM or MCMC plays a pivotal role in convergence. In particular, a poorly assigned prior may lead to a fragile Bayesian inference. In the context of GMM estimation, **?** pointed out that the criterion surface is highly irregular, so optimization often fails to converge in small samples, *e.g.*, **?** have documented a large number of non-converging GMM estimations. Further, GMM usually produces imprecise estimates due to an ill-conditioned weighting matrix. By contrast, the closed-form moment-based estimators are analytically tractable, computationally simple, and very easy to implement.

In an SV(p) model, the underlying latent volatility process follows an autoregressive process of order p. The estimation of SV(p) models is even more challenging than it is for an SV(1) model. Consequently, SV(p) models are rarely estimated in financial econometrics literature; exceptions are SMM of **?**, MCL of **?**, Bayes method of **?**. However, in line with these studies and empirical findings of this paper, motivations for SV(p) models are as follows: (1) SV(p) models are natural extensions of the basic SV(1) model, which can only generate geometrically decaying autocovariance function, whereas volatility process generically features persistent memory. (2) As pointed out by **?** and **?**, the latent volatility process of a multi-factor stochastic volatility (MFSV) model can be interpreted as a linear combination of latent and independent AR(1) processes which aggregate to an ARMA(p, q) process. So, the higher-order autoregressive terms in SV models naturally emerge from the aggregation process. (3) The empirical results of these studies suggest that higher-order models provide more flexibility to represent volatility persistence, heavy tails and may capture the effects of jumps as well. (4) Empirical evidence in this paper suggests that higher-order SV models may be preferable for both in-sample model fitting and out-of-sample volatility forecasting. Further, the higher-order SV models may be preferable for option pricing.

Due to its intrinsic complexity of SV(p) models, the work on the estimation of this class of models remains scarce, and most of the proposed ones are inflexible, computationally costly, and limited to low orders [see **?**, **?**, **?**]. Recently, **?** proposed a simple estimation method for SV(p) models by exploiting the non-Gaussian ARMA representation of these models.

The ARMA-SV method uses the moment structure of the logarithm of squared residual returns. This estimator is analytically tractable and computationally inexpensive. In particular, it can be readily implemented without using any numerical optimization, and it does not require one to choose an arbitrary initial parameter or an auxiliary model. The simple method may violate stationarity conditions in the presence of outliers or in small samples. To circumvent this problem, **?** suggested restricted estimation where the estimates are restrained on the space of acceptable parameter solutions by adjusting the eigenvalues that lie on or outside the unit circle.

Further, **?** also proposed an OLS-based winsorized version of the ARMA-SV (W-ARMA-SV) estimator, which substantially increases the probability of getting acceptable values and also improves efficiency. This computationally simple adjustment improves the stability and accuracy of the estimators. Indeed, in simulations, they show that W-ARMA-SV estimators improve the precision and uniformly superior to other estimators (including the Bayesian estimators proposed in this context) in terms of bias and RMSE. It is worth noting that, using the simple W-ARMA-SV estimator, one can perform a recursive-in-order calculation of the parameters of higher-order SV processes by exploiting the Durbin-Levinson-type (DL) algorithm. Under standard regularity assumptions, the W-ARMA-SV estimator is consistent and asymptotically normal when the fourth moment of the latent volatility process exists. Due to the $\sqrt{T}$-consistency, the W-ARMA-SV estimator can be effortlessly applied to very large samples, which are not rare in empirical finance. In these situations, estimators based on simulation technique and/or numerical optimization often require substantial computational effort to achieve convergence. So instead of using computationally costly estimators, one may prefer to use estimators that are available in analytical form.

We present empirical applications related to SV(p) models and the W-ARMA-SV estimator. Using different volatility proxies [the squared return of S&P 500 index and the realized volatility of S&P 500, FTSE100, NASDAQ100, N225, SSMI20 indices], we conduct two out-of-sample forecast experiments: (1) a moderately volatile period after the late-2000s financial crisis; (2) a highly volatile period, *i.e.*, the core financial crisis. We compare the accuracy of volatility forecasts among SV(p) models, GARCH models, and Heterogenous Autoregressive model of Realized Volatility (HAR-RV) models. The results suggest that SV(p) models perform better than other models in most cases. This finding holds even if a high volatility period (such as financial crisis) is included in the estimation sample or the forecasted sample. These inferences are not only based on a standard forecasting precision assessment [such as using MSE and MAE statistics] but also on formal prediction tests, using the MCS procedure of **?**. Our findings highlight the usefulness of higher-order SV models for volatility forecasting.

The W-ARMA-SV-OLS estimator considered in this paper can be interpreted as a *parsimonious moment-based* estimator where only a few (well chosen) moments are used. In a moment-based (or GMM) inference, using too many moments can be very costly from an estimation efficiency viewpoint as well as forecasting. Indeed, we show in our empirical applications that the W-ARMA-SV-OLS estimator exhibits the best performance in both forecasting and pricing, as well as maintain the numerical efficiency.

The paper proceeds as follows: Section 2 specifies SV(p) models and assumptions. Section 3 discusses simple estimators and the recursive prediction algorithm. Section 4 discusses the forecasting procedure. Section 5 assesses the forecasting performance. Section 6 concludes. The tables and other materials are provided in the Technical Appendix.

2 Framework

We consider a standard discrete-time SV process of order p, which is described below following **?**, **?** and **?**. Specifically, we say that a variable y_t follows a discrete-time SV(p) process if it satisfies the following assumption, where $t \in \mathbb{N}_0$, and $\mathbb{N}_0$ represents the non-negative integers.

Assumption 2.1. STOCHASTIC VOLATILITY OF ORDER p. *The process $\{y_t : t \in \mathbb{N}_0\}$ satisfies the equations*

$$y_t = \sigma_y \exp(w_t/2) z_t\,, \tag{2.1}$$

$$w_t = \sum_{j=1}^{p} \phi_j w_{t-j} + \sigma_v v_t\,, \tag{2.2}$$

where the vectors $(z_t, v_t)'$ are i.i.d. according to a $\mathrm{N}[0, I_2]$ *distribution, while $(\phi_1, \ldots, \phi_p, \sigma_y, \sigma_v)'$ are fixed parameters.*

We also make a stationarity assumption as follows.

Assumption 2.2. STATIONARITY. *The process $l_t = (y_t, w_t)'$ is strictly stationary.*

The last assumption entails that all the roots of the characteristic equation of the volatility process $[\phi(B)=0]$ lie outside the unit circle [*i.e.*, $\phi(z)\neq 0$ for $|z|\leq 1$], and $w_0 \sim N[0, \sigma_v^2/(1-\sum_{j=1}^{p}\phi_j^2)]$.

The SV(p) model consists of two stochastic processes, where y_t describes the dynamics of asset returns and $w_t := \log(\sigma_t^2)$ captures the dynamics of latent log volatilities. Usually the y_t's are residual returns, such that

$$y_t := r_t - \mu_r\,, \quad r_t := 100[\log(p_t) - \log(p_{t-1})]\,,$$

where μ_r is the mean of returns (r_t) and p_t is the raw prices of an asset.[3] The latent process w_t can be interpreted as a random flow of uncertainty shocks or new information in financial markets, while ϕ_j's capture the volatility persistence. This type of volatility model naturally fits into the theoretical framework of modern financial theory.

Let us now transform y_t by taking the logarithm of its squared value. We get in this way the following *measurement equation*:

$$\begin{aligned}\log(y_t^2) &= \log(\sigma_y^2) + w_t + \log(z_t^2) = \{\log(\sigma_y^2) + \mathbb{E}[\log(z_t^2)]\} + w_t + \{\log(z_t^2) - \mathbb{E}[\log(z_t^2)]\} \\ &= \mu + w_t + \epsilon_t \end{aligned} \tag{2.3}$$

where

$$\mu := \mathbb{E}[\log(y_t^2)] = \log(\sigma_y^2) + \mathbb{E}[\log(z_t^2)]\,, \quad \epsilon_t := \log(z_t^2) - \mathbb{E}[\log(z_t^2)]\,. \tag{2.4}$$

Note that this logarithmic transformation entails no information loss since the distribution of z_t is symmetric (see Remark 1 of **?**). Furthermore, even if v_t and z_t are not mutually independent, they are uncorrelated if the joint distribution of v_t and z_t is symmetric, that is $f(v_t, z_t) = f(-v_t, -z_t)$; see **?**.

Under the normality assumption for z_t, the errors ϵ_t are i.i.d. according to the distribution of a centered $\log(\chi_1^2)$ random variable [*i.e.*, ϵ_t has mean zero and variance $\mathbb{E}(\epsilon_t^2)$]. The cumulant generating function of $\log(\chi_1^2)$ distribution is:

$$\begin{aligned} M(s) &= \log\mathbb{E}\big[\exp\big(s\log(\chi_1^2)\big)\big] = \log\Big[\mathbb{E}\big(\chi_1^2\big)^s\Big] = \log\left[\frac{2^s\,\Gamma\big((1/2)+s\big)}{\Gamma(1/2)}\right] \\ &= s\log(2) + \log[\Gamma\big((1/2)+s\big)] - \log[\Gamma(1/2)]\,, \quad \text{for } s \geq 0\,, \end{aligned} \tag{2.5}$$

where $\Gamma(z) := \int_0^\infty x^{z-1}e^{-x}\,dx$ is the *gamma function*; see **?**. The m^{th} cumulant of the $\log(\chi_1^2)$ random variable is the m^{th} derivative of $M(s)$ evaluated at $s=0$. Thus, the corresponding cumulants (κ_m) and central moments ($\tilde{\mu}_m$) are:

$$\kappa_m = \begin{cases} \log(2) + \psi\big(\frac{1}{2}\big), & \text{if } m = 1 \\ \psi^{(m-1)}\big(\frac{1}{2}\big), & \text{if } m > 1 \end{cases}, \qquad \tilde{\mu}_m = \begin{cases} 0, & \text{if } m = 1 \\ \kappa_m + \sum_{j=1}^{m-2}\binom{m-1}{j}\kappa_{m-j}\tilde{\mu}_j, & \text{if } m > 1 \end{cases}, \tag{2.6}$$

where

$$\psi(z) := \frac{d}{dz}\log[\Gamma(z)] = \frac{\Gamma'(z)}{\Gamma(z)} \tag{2.7}$$

[3] It is noteworthy to mention that y_t is ordinarily the error term of any time series regression model, see for example **?**.

is the *digamma function* and

$$\psi^{(m)}(z) := \frac{d^m}{dz^m}\psi(z) = \frac{d^{m+1}}{dz^{m+1}}\log[\Gamma(z)] \tag{2.8}$$

is the *polygamma function* of order m [*i.e.*, the $(m+1)$-th order derivative of the logarithm of the *gamma function*].

From (2.6), we get:

$$\mathbb{E}[\log(z_t^2)] = \kappa_1 = \log(2) + \psi(1/2) \simeq -1.2704\,, \tag{2.9}$$

$$\sigma_\epsilon^2 := \mathbb{E}(\epsilon_t^2) = \mathrm{Var}[\log(z_t^2)] = \tilde{\mu}_2 = \kappa_2 = \psi^{(1)}(1/2) = \pi^2/2\,, \tag{2.10}$$

$$\mathbb{E}(\epsilon_t^3) = \tilde{\mu}_3 = \kappa_3 = \psi^{(2)}(1/2)\,, \quad \mathbb{E}(\epsilon_t^4) = \tilde{\mu}_4 = \kappa_4 + 3\kappa_2^2 = \psi^{(3)}(1/2) + 3\sigma_\epsilon^2 = \pi^4 + 3\sigma_\epsilon^2\,; \tag{2.11}$$

see **?**, Chapter 6. The $\log(\chi_1^2)$ distribution is often approximated by a normal distribution with mean of -1.2704 and variance of $\pi^2/2$ [see **?**], or by a mixture distribution [**?**].

On setting

$$y_t^* := \log(y_t^2) - \mu\,, \tag{2.12}$$

the SV model (2.3) can be written as

$$y_t^* = w_t + \epsilon_t\,. \tag{2.13}$$

By combining (2.2) and (2.13), we see that the SV(p) model can be written in state-space form:

$$\text{State Transition Equation:} \quad w_t = \sum_{j=1}^{p}\phi_j w_{t-j} + v_t\,, \tag{2.14}$$

$$\text{Measurement Equation:} \quad y_t^* = w_t + \epsilon_t\,, \tag{2.15}$$

where w_t is a logarithm of latent daily volatility, y_t^* is a logarithm of the daily squared return corrected by its mean, where the variables v_t are i.i.d. $\mathrm{N}(0, \sigma_v^2)$, and the ϵ_t's are i.i.d. $\log(\chi_1^2)$; for further discussion of this representation, see **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**.

3 Simple ARMA-based estimation

Recently, **?** proposed simple estimators for SV(p) models given in assumptions 2.1 - 2.2 by exploiting the ARMA representation of the process y_t^*. They derived the following ARMA(p,p) representation for y_t^*:

$$y_t^* = \sum_{j=1}^{p}\phi_j y_{t-j}^* + \eta_t - \sum_{j=1}^{p}\theta_j\eta_{t-j} \tag{3.1}$$

with $\eta_t - \sum_{j=1}^{p}\theta_j\eta_{t-j} = v_t + \epsilon_t - \sum_{j=1}^{p}\phi_j\epsilon_{t-j}$, where the error processes $\{v_t\}$ and $\{\epsilon_t\}$ are mutually independent, the errors v_t are i.i.d. $N(0, \sigma_v^2)$, and the errors ϵ_t are i.i.d. according to the distribution of a $\log(\chi_1^2)$ random variable.

From the above expression, y_t^* has the following autocovariances:

$$\operatorname{cov}(y_t^*, y_{t-k}^*) := \gamma_{y^*}(k) = \begin{cases} \phi_1\gamma_{y^*}(k-1) + \cdots + \phi_p\gamma_{y^*}(k-p) + \sigma_v^2 + \sigma_\epsilon^2; & \text{if } k = 0, \\ \phi_1\gamma_{y^*}(k-1) + \cdots + \phi_p\gamma_{y^*}(k-p) - \phi_k\sigma_\epsilon^2; & \text{if } 1 \leq k \leq p, \\ \phi_1\gamma_{y^*}(k-1) + \cdots + \phi_p\gamma_{y^*}(k-p); & \text{if } k > p. \end{cases} \tag{3.2}$$

The above autocovariances yield the following closed-form expressions for SV parameters:

$$\boldsymbol{\phi_p} = \boldsymbol{\Gamma}^{-1}_{(\boldsymbol{p}+\boldsymbol{j}-\boldsymbol{1})}\gamma_{(\boldsymbol{p}+\boldsymbol{j})}, \quad j \geq 1 \tag{3.3}$$

$$\sigma_y = [\exp(\mu + 1.27)]^{1/2}, \tag{3.4}$$

$$\sigma_v = [\gamma_{y^*}(0) - \boldsymbol{\phi_p}^{'}\gamma_{(\boldsymbol{1})} - \pi^2/2]^{1/2}, \tag{3.5}$$

where $\boldsymbol{\phi_p} := (\phi_1, \ldots, \phi_p)^{'}$, $\gamma_{(\boldsymbol{p}+\boldsymbol{j})} := [\gamma_{y^*}(p+j), \ldots, \gamma_{y^*}(2p+j-1)]^{'}$ are vectors and $\boldsymbol{\Gamma}_{(\boldsymbol{p}+\boldsymbol{j}-\boldsymbol{1})}$ is a p-dimensional Toeplitz matrices such that

$$\boldsymbol{\Gamma}_{(\boldsymbol{p}+\boldsymbol{j}-\boldsymbol{1})} := \begin{bmatrix} \gamma_{y^*}(p+j-1) & \gamma_{y^*}(p+j-2) & \cdots & \gamma_{y^*}(j) \\ \gamma_{y^*}(p+j) & \gamma_{y^*}(p+j-1) & \cdots & \gamma_{y^*}(j+1) \\ \vdots & \vdots & & \vdots \\ \gamma_{y^*}(2p+j-2) & \gamma_{y^*}(2p+j-3) & \cdots & \gamma_{y^*}(p+j-1) \end{bmatrix}.$$

where p is the SV order, $\gamma_{y^*}(k) = \operatorname{cov}(y_t^*, y_{t-k}^*)$, with y_t^* and μ defined in (2.12).

Now, it is natural to estimate $\gamma_{y^*}(k)$ and μ by the corresponding empirical moments:

$$\hat{\gamma}_{y^*}(k) = \frac{1}{T-k}\sum_{t=1}^{T-k} y_t^* y_{t+k}^*, \qquad \hat{\mu} = \frac{1}{T}\sum_{t=1}^{T}\log(y_t^2), \tag{3.6}$$

where by construction y_t^* is a mean corrected process. Setting $j = 1$ in (3.3) and replacing theoretical moments by their corresponding empirical moments yield the following *simple ARMA-SV* estimator of the SV(p) coefficients:

$$\hat{\boldsymbol{\phi}}_{\boldsymbol{p}} = \hat{\boldsymbol{\Gamma}}^{-1}_{(\boldsymbol{k},\boldsymbol{p})}\hat{\gamma}_{(\boldsymbol{k},\boldsymbol{p})}, \tag{3.7}$$

$$\hat{\sigma}_y = [exp(\hat{\mu} + 1.27)]^{1/2}, \tag{3.8}$$

$$\hat{\sigma}_v = [\hat{\gamma}_{y^*}(0) - \hat{\boldsymbol{\phi}}_{\boldsymbol{p}}^{'}\hat{\gamma}_{(\boldsymbol{k},\boldsymbol{p})} - \pi^2/2]^{1/2}. \tag{3.9}$$

3.1 Restricted estimation

These simple estimators may yield a solution outside the admissible area, *i.e.*, some of the eigenvalues of the latent volatility process [it is an AR(p) process] may lie outside the unit circle or equal to unity. This issue can arise especially in small samples or in the presence of outliers. When this happens, a simple fix is projecting the estimate on the space of acceptable parameter solutions by altering the eigenvalues that lie on or outside the unit circle. The characteristic equation of the latent AR(p) process is given by $C(\lambda) = \lambda^p - \phi_1 \lambda^{p-1} - \cdots - \phi_p = 0$, and the stationary condition requires all roots lie inside the unit circle, *i.e.*, $|\lambda_i| < 1,\ i = 1, \cdots, p$. If the estimated parameters fail to satisfy this condition, then the restricted estimation can be done in the following two steps:

1. Given the estimated unstable parameters, we calculate the roots of the characteristic equation and restrict their absolute values to less than unity.

2. Given these restricted roots, we calculate the constrained parameters which ensure stationarity.

For example, in case of an SV(2) model, the characteristic equation of the latent volatility process is $C(\lambda) = \lambda^2 - \phi_1 \lambda - \phi_2 = 0$. It may have two types of roots: (i) if $\phi_1^2 + 4\phi_2 \geq 0$, then $C(\lambda)$ has two real roots, and these are given by $\lambda_{1,2} = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}$ and (ii) if $\phi_1^2 + 4\phi_2 < 0$ then $C(\lambda)$ has two complex roots, and these are given by $\lambda_{1,2} = \frac{\phi_1}{2} \pm i \frac{\sqrt{-(\phi_1^2 + 4\phi_2)}}{2}$. When the estimated polynomial coefficients produce an unstable solution, then we restrict the absolute value of the roots less than unity, *i.e.* $|\lambda_{1,2}| < 1$ or $|\lambda_{1,2}| = 1 - \Delta$ where Δ is a very small number. Given these restricted roots, we solve for restricted parameters which ensure the stationarity condition. These steps can be done very easily in MATLAB. In MATLAB, the **roots** function calculates the roots given the parameters, and the **poly** function calculates the parameters given the roots.

3.2 ARMA-based winsorized estimation

One can achieve better stability and efficiency of ARMA-SV estimator by using "winsorization" which exploits (3.3). Winsorization (censoring) substantially increases the probability of getting admissible values. From (3.3), it is easy to see that:

$$\phi_{\boldsymbol{p}} = \sum_{j=1}^{\infty} \omega_j \boldsymbol{\Gamma}^{-1}_{(\boldsymbol{p}+\boldsymbol{j}-1)} \gamma_{(\boldsymbol{p}+\boldsymbol{j})} \tag{3.10}$$

for any ω_j sequence with $\sum_{j=1}^{\infty} \omega_j = 1$. Using (3.10), we can define a more general class of estimators for $\phi_{\boldsymbol{p}}$ by taking a weighted average of several sample analogs of the ratio $\boldsymbol{\Gamma}^{-1}_{(\boldsymbol{p}+\boldsymbol{j}-1)} \gamma_{(\boldsymbol{p}+\boldsymbol{j})}$:

$$\tilde{\phi}_{\boldsymbol{p}} = \sum_{j=1}^{J} \omega_j \hat{\boldsymbol{\Gamma}}^{-1}_{(\boldsymbol{p}+\boldsymbol{j}-1)} \hat{\gamma}_{(\boldsymbol{p}+\boldsymbol{j})}, \tag{3.11}$$

where $1 \leq J \leq T - p$ with $\sum_{j=1}^{J} \omega_j = 1$ and T is the length of time series. We can expect that a sufficiently general class of weights may improve the efficiency of the ARMA-SV estimators.

Using (3.11), **?** proposed the OLS-based *winsorized ARMA-SV* (W-ARMA-SV) estimator of $\phi_{\boldsymbol{p}}$ as follows:

$$\hat{\phi}^{ols}_{\boldsymbol{p}} = (\bar{a}'\bar{a})^{-1}\bar{a}'\bar{e}, \tag{3.12}$$

where $\bar{a} = (\hat{\Gamma}_{(p)}\omega_1^{1/2}, \ldots, \hat{\Gamma}_{(p+J-1)}\omega_J^{1/2})'$ and $\bar{e} = (\hat{\gamma}_{(\boldsymbol{p}+1)}\omega_1^{1/2}, \ldots, \hat{\gamma}_{(\boldsymbol{p}+\boldsymbol{J})}\omega_J^{1/2})'$. Clearly, different OLS-based W-ARMA-SV can be generated by considering different weights $w_1, \ldots, w_J$. In our empirical applications, we focus on the case where the weights are equal ($\omega_j = 1/J$). Note that, in case of an SV(2), the W-ARMA-

SV-OLS (with equal weights) yields:

$$\hat{\phi}_1^{ols} = \frac{\sum_{j=1}^{J}[\hat{\gamma}_{y^*}(j+1)\hat{\gamma}_{y^*}(j+2) - \hat{\gamma}_{y^*}(j)\hat{\gamma}_{y^*}(j+3)][\hat{\gamma}_{y^*}(j+1)^2 - \hat{\gamma}_{y^*}(j)\hat{\gamma}_{y^*}(j+2)]}{\sum_{j=1}^{J}[\hat{\gamma}_{y^*}(j+1)^2 - \hat{\gamma}_{y^*}(j)\hat{\gamma}_{y^*}(j+2)]^2} \tag{3.13}$$

$$\hat{\phi}_2^{ols} = \frac{\sum_{j=1}^{J}[\hat{\gamma}_{y^*}(j+1)\hat{\gamma}_{y^*}(j+3) - \hat{\gamma}_{y^*}(j+2)^2][\hat{\gamma}_{y^*}(j+1)^2 - \hat{\gamma}_{y^*}(j)\hat{\gamma}_{y^*}(j+2)]}{\sum_{j=1}^{J}[\hat{\gamma}_{y^*}(j+1)^2 - \hat{\gamma}_{y^*}(j)\hat{\gamma}_{y^*}(j+2)]^2}. \tag{3.14}$$

The above simplification [simple regressions] follows from (3.3) with $p = 2$, which can be written as following:

$$\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \begin{bmatrix} \gamma_{y^*}(j+1) & \gamma_{y^*}(j) \\ \gamma_{y^*}(j+2) & \gamma_{y^*}(j+1) \end{bmatrix}^{-1} \begin{bmatrix} \gamma_{y^*}(j+2) \\ \gamma_{y^*}(j+3) \end{bmatrix} = \begin{bmatrix} \frac{\gamma_{y^*}(j+1)\gamma_{y^*}(j+2) - \gamma_{y^*}(j)\gamma_{y^*}(j+3)}{\gamma_{y^*}(j+1)^2 - \gamma_{y^*}(j)\gamma_{y^*}(j+2)} \\ \frac{\gamma_{y^*}(j+1)\gamma_{y^*}(j+3) - \gamma_{y^*}(j+2)^2}{\gamma_{y^*}(j+1)^2 - \gamma_{y^*}(j)\gamma_{y^*}(j+2)} \end{bmatrix}. \tag{3.15}$$

All these estimators are depend on J and for $J = 1$, they are equivalent to the simple ARMA-SV estimator which is given by (3.7).

3.3 Recursive estimation for SV(p) models

It is possible to estimate higher-order SV(p) models using a recursive Durbin-Levinson (DL) type estimation algorithm. For notational convenience, we use a different indexation for the autoregressive parameters of the volatility process [only for this section]. For example, the SV(p) parameters are now denoted by $\Theta_p^{SV} := (\{\phi_{p,j}\}_{j=1}^{p}, \sigma_{pv}, \sigma_y)'$.

The recursive estimation of the ARMA-SV estimator exploits extended Yule-Walker (EYW) equations of the observed process. When the MA order is fixed, the system of the EYW equations constitutes a nested Toeplitz system. A *Generalized Durbin-Levinson* algorithm for the ARMA-SV estimator for SV(p) model is useful when neither the AR order nor the MA order is known. We consider the case $i = p$, *i.e.*, the MA order is p, which also implies that the AR order is p.

For $i = 0$, use the Durbin-Levinson algorithm to calculate

$$\{\hat{\phi}_{p,j}^{(0)} \mid p \geq 1, \quad j = 1, \ldots, p\}.$$

For $i \geq 1$, calculate

$$\hat{\phi}_{p,0}^{(i-1)} = -1,$$

and

$$\hat{\phi}_{p,j}^{(i)} = \hat{\phi}_{p+1,j}^{(i-1)} - \frac{\hat{\phi}_{p+1,p+1}^{(i-1)}}{\hat{\phi}_{p,p}^{(i-1)}} \hat{\phi}_{p,j-1}^{(i-1)}, \quad \text{where } p \geq 1, \quad j = 1, \ldots, p,$$

$$\hat{\sigma}_y = [\exp(\hat{\mu} + 1.27)]^{1/2},$$

$$\hat{\sigma}_{pv} = [\hat{\gamma}_{y^*}(0) - \sum_{j=1}^{p} \hat{\phi}_{p,j} \hat{\gamma}_{y^*}(j) - \pi^2/2]^{1/2}.$$

This algorithm is the same as **?** algorithm [except for equations involving $\hat{\sigma}_y$ and $\hat{\sigma}_{pv}$] for calculating the extended sample autocorrelation function under the stationarity assumption.

4 Forecasting with SV(p) models

As discussed earlier, SV(p) models can be written as a linear state-space model without losing any information. The state-space representation of SV(p) models is given by

$$\begin{aligned} y_t^* &= w_t + \epsilon_t, \\ w_t &= \sum_{j=1}^{p} \phi_j w_{t-j} + v_t, \end{aligned} \tag{4.1}$$

where the distribution ϵ_t is approximated by a normal distribution with mean 0 and variance $\pi^2/2$. Using the standard notations of **?**, the model defined in (4.1) can be rewritten as following:

$$\begin{aligned} y_t &= A' x_t + H' \xi_t + w_t, \\ \xi_{t+1} &= F \xi_t + \nu_{t+1}, \end{aligned} \tag{4.2}$$

with $y_t = y_t^*$, $A' = 0$, $x_t = 1$, $H' = (1, 0, \ldots, 0)$ is a $1 \times p$ vector, $w_t = \epsilon_t$, $R = \mathbb{E}(w_t w_t') = \pi^2/2$,

$$\xi_t = \begin{bmatrix} w_t \\ w_{t-1} \\ w_{t-2} \\ \vdots \\ w_{t-p+1} \end{bmatrix}, \quad F = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad \nu_{t+1} = \begin{bmatrix} v_{t+1} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad Q = \mathbb{E}(\nu_t \nu_t') = \begin{bmatrix} \sigma_v^2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

where F and Q are $p \times p$ matrices, and ξ_t are ν_{t+1} are $p \times 1$ vectors. Now using (4.2), the Kalman filter can be applied as follows:

- Initialization:

$$\begin{aligned} \hat{\xi}_{1|0} &= \mathbb{E}(\xi_1) = \mathbf{0}_{(p\times 1)}, \\ \mathbf{P}_{1|0} &= \mathbb{E}([(\xi_1 - \mathbb{E}(\xi_1)][(\xi_1 - \mathbb{E}(\xi_1)]') = diag[\sigma_v^2, \ldots, \sigma_v^2]_{(p\times p)}, \end{aligned} \tag{4.3}$$

  where $\mathbf{P}_{1|0}$ is the MSE associated with $\hat{\xi}_{1|0}$.

- Sequential updating:

$$\begin{aligned} \hat{\xi}_{t|t} &= \hat{\xi}_{t|t-1} + \mathbf{P}_{t|t-1} H (H' \mathbf{P}_{t|t-1} H + R)^{-1} \times (y_t - H' \hat{\xi}_{t|t-1}), \\ \mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} H (H' \mathbf{P}_{t|t-1} H + R)^{-1} \mathbf{P}_{t|t-1} H'. \end{aligned} \tag{4.4}$$

- In-sample prediction:

$$
\begin{aligned}
\hat{\xi}_{t+1|t} &= F\hat{\xi}_{t|t-1} + F\mathbf{P}_{t|t-1}H(H'\mathbf{P}_{t|t-1}H+R)^{-1}\times(y_t - H'\hat{\xi}_{t|t-1}),\\
\mathbf{P}_{t+1|t} &= F\mathbf{P}_{t|t}F' + Q.
\end{aligned}
\tag{4.5}
$$

Given (4.5), the forecast of y_{t+1} and the MSE of forecast error are given by

$$
\begin{aligned}
\hat{y}_{t+1|t} &= H'\hat{\xi}_{t+1|t},\\
\mathbb{E}([y_{t+1}-\hat{y}_{t+1|t}][y_{t+1}-\hat{y}_{t+1|t}]') &= H'\mathbf{P}_{t+1|t}H + R.
\end{aligned}
\tag{4.6}
$$

- Out-of-sample h-step-ahead forecasting:

$$
\begin{aligned}
\hat{\xi}_{T+h|T} &= F^h\hat{\xi}_{T|T},\\
\hat{y}_{T+h|T} &= H'\hat{\xi}_{T+h|T} = H'F^h\hat{\xi}_{T|T}.
\end{aligned}
\tag{4.7}
$$

The h-step-ahead forecast is computed by (4.7) with the simple estimates plugged in.

5 Volatility forecast performance

In this section, we demonstrate two empirical applications of SV(p) models. *First*, we examine the fit of SV(p) models with real data to see the empirical evidence of this type of parametric models. *Second*, we further extend our analysis and compare the forecast performance of three typical volatility models in out-of-sample forecasting experiments, models include the GARCH-type, SV-type and high-frequency realized volatility based models.

We evaluate the volatility forecast performance amongst GARCH, SV, and realized volatility based models.[4] Volatility has long been modeled and forecasted using GARCH models because of the earlier discussed complexity of SV models. We considered several popular GARCH-type models in our experiments, these include: GARCH models of **?**, Exponential GARCH (EGARCH) models of **?** and GJR models of **?**. For the details of these models and their forecast equations, see appendix A.2.

We also consider high-frequency based Heterogenous Autoregressive model of Realized Volatility (HAR-RV) models of **?**. In this study, we use a logarithmic version of HAR-RV model since the logarithmic transformation of RV appears approximately Gaussian; see **?**, **?**. The HAR-RV model takes into account the long memory feature of realized volatility, and among the models proposed to forecast realized volatility, it stands out because of its simplicity (for details, see appendix A.4).

For SV(p) models, we exploit the state-space representation in (2.14-2.15) and calculate forecasts based on the Kalman filter. The SV(p) parameters are computed using our simple method, where we used (3.12) with $J = 50$ and fixed the value of J before any estimations. Given the simple estimates, we calculate

[4] Realized volatility (RV), is a model free volatility, received much attention among the financial economists and econometricians as an accurate measure of the true latent volatility under the ideal market assumption [**?**, **?**] and it can be used as a proxy for true latent volatility (for details, see appendix A.3).

the forecasts of SV(p) models through the Kalman filter. For the details of this forecasting procedure and out-of-sample forecasting equations, see appendix 4.

We use three loss measures to evaluate the forecast accuracy. These include MSE, MAE, and R2LOG. MSE and MAE are the mean squared error and mean absolute error, respectively, and R2LOG is the logarithmic loss function proposed by **?** and can penalize volatility forecast asymmetry in high and low level of volatility. These loss measures are defined as follows:

$$MSE: l_t = (\hat{\sigma}_t^2 - h_{t|t-k}^2)^2, \quad MAE: l_t = |\, \hat{\sigma}_t^2 - h_{t|t-k}^2 \,|, \quad R2LOG: l_t = (\log \hat{\sigma}_t^2 - \log h_{t|t-k})^2,$$

where $\hat{\sigma}_t^2$ is an unbiased *ex-post* proxy of conditional variance (such as squared return or realized volatility) and $h_{t|t-k}$ is a volatility forecast based on $t-k$ information set where $k > 0$.

Using the above loss functions, we also compute the model confidence set (MCS) procedure proposed by **?**. The model confidence set involves a sequence of tests for equal predictive ability (EPA) hypothesis. Given a model set $\mathcal{M}_0$, which contains m competing forecast models, the null hypothesis is that all models in $\mathcal{M}_0$ have equal predictive accuracy. If the null hypothesis is rejected at a given confidence level α, then the worst performing model in $\mathcal{M}_0$ is eliminated. After that, the EPA test is repeated until the null hypothesis is accepted. When the null hypothesis is accepted, the remainder composes $1-\alpha$ confidence set, $\hat{\mathcal{M}}_{1-\alpha}^*$.

We now briefly discuss how it is implemented. Define the relative loss differential between models by

$$d_{i,j,t} = l_{i,t} - l_{j,t}, \quad \text{for all} \;\; i,j \in \mathcal{M}, \quad t = 1, \ldots, T,$$

be the simple loss of model i relative to any other model j at time t. Using the loss differential between competing models, the MCS procedure tests the EPA hypothesis in two alternative ways.

$$H_0: \mu_{ij} = 0 \quad \text{for all} \;\; i,j \in \mathcal{M} \quad \text{and} \quad H_A: \mu_{ij} \neq 0 \quad \text{for some} \;\; i,j \in \mathcal{M} \tag{5.1}$$

or

$$H_0: \mu_{i,\cdot} = 0 \quad \text{for all} \;\; i \in \mathcal{M} \quad \text{and} \quad H_A: \mu_{i,\cdot} \neq 0 \quad \text{for some} \;\; i \in \mathcal{M} \tag{5.2}$$

where $\mu_{ij} = \mathbb{E}(d_{i,j})$ and $\mu_{i,\cdot} = \mathbb{E}(d_{i,\cdot})$. The two statistics, used in the model confidence set test, are expressed as follows:

$$MCS\_T_{R,\mathcal{M}} = \max_{i,j \in \mathcal{M}} |\, t_{i,j} \,| \quad \text{and} \quad MCS\_T_{\max,\mathcal{M}} = \max_{i \in \mathcal{M}} t_{i,\cdot}, \tag{5.3}$$

where

$$t_{i,j} = \frac{\bar{d}_{i,j}}{\sqrt{\widehat{\text{Var}}(\bar{d}_{i,j})}}, \quad t_{i,\cdot} = \frac{\bar{d}_{i,\cdot}}{\sqrt{\widehat{\text{Var}}(\bar{d}_{i,\cdot})}},$$

$$\bar{d}_{i,\cdot} = m^{-1} \sum_{j \in \mathcal{M}} \bar{d}_{i,j}, \quad \bar{d}_{i,j} = T^{-1} \sum_{t=1}^{T} d_{i,j,t} \quad \text{for} \;\; i,j \in \mathcal{M},$$

while $\widehat{\text{Var}}(\bar{d}_{i,\cdot})$ and $\widehat{\text{Var}}(\bar{d}_{i,j})$ are bootstrap estimates of $\text{Var}(\bar{d}_{i,\cdot})$ and $\text{Var}(\bar{d}_{i,j})$, respectively. In our calculations, we perform a block-bootstrap using a block length of 12 days and 10000 bootstrap replications. The first statistic, $t_{i,j}$, is used in the well-known test for comparing two forecasts; see **?** and **?**, while the second one, $t_{i,\cdot}$, is used in **?**, **?**, and **?**.

We conduct two out-of-sample forecast experiments using different volatility proxy:

1. **Design 1 (Moderate volatility regimes)**: In this setting, we consider a sample period from September 01, 2005 to August 31, 2010. The in-sample is from September 01, 2005 to August 31, 2008 and the out-of-sample is from September 01, 2008 to August 31, 2010. We forecast a moderately volatile period but the in-sample contains the most volatile part of the late-2000s financial crisis.

2. **Design 2 (High volatility regimes)**: In this design, we consider a sample period, from January 01, 2005 to December 31, 2009. The in-sample is from January 01, 2005 to December 31, 2007 and the out-of-sample is from January 01, 2008 to December 31, 2009. The out-of-sample includes a highly volatile period, *i.e.*, the late-2000s financial crisis (Subprime mortgage crisis / United States housing bubble).

In both designs, we consider a sample of five years that split into three years span of in-sample and two years span of out-of-sample. Three years span for the in-sample window is adequate for finding the most accurate volatility forecasts; see **?**.

Within the SV and GARCH framework, the key element is the specification for conditional variance. Parametric SV and GARCH models utilize daily returns (typically squared returns) to extract information about the current level of volatility, and this information is used to form expectations about the next period's volatility. Although the squared return is a noisy measure, it is a conditionally unbiased estimator of the daily conditional variance. In contrast, **?** suggest that realized volatility (which is based on cumulative intraday squared returns) is a more accurate proxy for true latent volatility. Therefore, we examine out-of-sample volatility forecasts across competing models using different loss functions as well as the MCS procedure with squared return and realized volatility proxies.

We computed out-of-sample forecasts using rolling (moving) window method and computed for a range of forecast horizons which are 1-day, 2-day, 1-week, 2-week, 3-week and 1-month. In this rolling forecasts setup, an initial sample using data from $t = 1, \ldots, T$ is used to determine a window width T, to estimate the models, and to form $h-$step ahead out-of-sample forecasts starting at time T. Then the window is moved ahead one time period, the models are re-estimated using data from $t = 2, \ldots, T+1$, and $h-$step ahead out-of-sample forecasts are produced starting at time $T+1$. This process is repeated until no more $h-$step ahead forecasts can be computed.

5.1 Forecasting squared return

Now we consider the daily squared return as volatility proxy and evaluate the volatility forecast performance among GARCH, SV and HAR-RV models using the S&P 500 index. The high-frequency RV estimates and prices of S&P 500 index are obtained from the Oxford-Man Institute's Realized Library; see **?**. The raw prices p_t are converted to returns by the transformation $r_t = 100[\log(p_t) - \log(p_{t-1})]$. The returns are converted to residual returns by $y_t = r_t - \hat{\mu}_r$ where $\hat{\mu}_r$ is the sample average of returns. Note that, y_t^2 is the volatility proxy at time t.

We consider a modified version of the HAR-RV model [that defined in (A.4)], where the daily squared return is the dependent variable and realized volatilities are independent variables. We are using additional information from the high-frequency data to forecast the squared return. However, there is a problem in measuring the realized volatility from high-frequency data. The high-frequency estimate of realized volatility may be very unstable because of the market microstructure noise, which captures a mixture of frictions inherent in the trading mechanism: bid-ask bounces, discreteness of price changes, different price impacts due to differences in trade sizes, slow response of prices to a block trade, strategic component of the order flow, inventory control effects, etc. The choice of RV estimator is important. We consider other RV estimates, including: realized bi-power variation (BV) [**?**], realized semi-variance (RSV) [**?**], realized kernel (RK) [**?** (**?**, **?**)], median realized volatility (MedRV) [**?**], two-scale realized kernels (TSRK) [**?**]. We also consider subsampled RV, BV, and RSV. Subsampling, introduced by **?**, is a simple way to improve the efficiency of sparse-sampled estimators.[5]

For each forecast experiment, we compute forecasts from three SV models, eleven GARCH-type models, and nine HAR-RV type models. The eight GARCH-type models are: GARCH$(1,1)$, GARCH$(1,2)$, GARCH$(2,1)$, GARCH$(2,2)$, GARCH$(3,3)$, EGARCH$(1,1)$, EGARCH$(2,2)$, EGARCH$(3,3)$, GJR$(1,1)$, GJR$(2,2)$ and GJR$(3,3)$.

For S&P 500, in design 1, the sample period is from September 01, 2005 to August 31, 2010 and the number of observations is $T = 1258$. The in-sample is from September 01, 2005 to August 31, 2008 ($T = 753$) and the out-of-sample is from September 01, 2008 to August 31, 2010 ($T = 505$). In design 2, the sample period is from January 01, 2005 to December 31, 2009 and the number of observations is $T = 1259$. The in-sample is from January 01, 2005 to December 31, 2007 ($T = 754$) and the out-of-sample is from January 01, 2008 to December 31, 2009 ($T = 505$).

Tables A1-A2 report summary statistics of daily variables, high-frequency RV estimates, and their logarithms. Using our out-of-sample forecasts, we calculate forecast evaluation measures, *i.e.*, MSE, MAE and R2LOG. Tables A3-A8 presents the main results of our forecasting experiments. For easy comparison, we report the relative MSE, the relative MAE and the relative R2LOG of forecast error. These are relative to the reference model HAR-RV5, and hence, values smaller than unity indicate better forecast performance than the HAR-RV5 model. We also report the MCS p-value for the corresponding model.

We check the stationarity of SV(p) estimates and find that the W-ARMA-SV-OLS ($J = 50$) estimator produces a few unstable solutions for the SV(3) model. In these cases, the unstable estimates are modified by applying the restricted estimation (proposed in Section 3.1). We report both the unrestricted and restricted SV(3) forecasts in Table A3-A8. However, these forecasts are qualitatively similar. Note that, in the

[5]Subsampling involves using a variety of "grids" of prices sampled at a given frequency to obtain a collection of realized measures, which are then averaged to yield the "subsampled" version of the estimator. For example, 5-minute RV can be computed using prices sampled at 10:30, 10:35, etc. and can also be computed using prices sampled at 10:31, 10:36, etc.

realized volatility forecasting application below, we do not find any unstable parameter estimates of SV models.

In design 1, Tables A3-A5, when we forecast a moderately unstable period after the core financial crisis, the forecasting performance of higher-order SV models [especially, the SV(3) model (unrestricted or restricted)] are superior to all other volatility models. This result holds across different forecast horizons, different evaluation measures and based on MCS. According to MCS, the SV(3) model dominates all other competing models, except for 1- and 2-weeks horizon as per MSE loss function. Several HAR-RV models performed well according to MSE, but these results are undermined by their performance in terms of MAE and R2LOG. Except for one- and two-day horizons, the forecasting performance of GARCH-type models is poor for all models, according to all loss measures and across different forecast horizons.

In design 2, Tables A6-A8, when we forecast a highly volatile period, *i.e.*, the core financial crisis, SV(p) models perform better than other competing models in most cases (this holds across different forecast horizons and different evaluation measure) while HAR-RV models perform better than GARCH models. The SV(3) model produces the superior forecast in terms of MSE criteria in horizon 1-day, 2-day, 3-weeks, and 1-month. SV(p) models [SV(3) or SV(2)] are the top forecasting models based on MCS p-value when using MAE and R2LOG. Performances of GARCH and HAR-RV models are poor according to R2LOG. These models may produce asymmetric forecast errors because R2LOG heavily penalized asymmetry in a high and low level of volatility. However, HAR-RV models perform better than GARCH models according to RMSE, while GARCH models outperform HAR-RV models according to MAE and R2LOG. This implies that GARCH models produce large forecast errors because MSE heavily penalized any outlier.

In both settings, among HAR models, those that based on the subsampled version of RV estimators produces identical forecasts, implies that subsampling is not improving any forecast performance. Further, the performance of HAR models is inferior among all models in long-horizon. Note that in both designs, the financial crisis is included either in-sample or out-of-sample. During this time, the financial market is unstable, and the high-frequency RV estimators are affected by large market microstructure noise. The forecasting performance of HAR-RV models may be affected by these noisy RV estimators.

From Tables A3-A8, we can see that except for a few instances, higher-order SV models perform better than GARCH-type and HAR-RV type models not only in all evaluation measures but also across different volatility regimes and horizons. In both of our out-of-sample experiments, higher-order SV models also outperform the first-order SV model. This finding suggests that additional lag terms in the latent volatility equation are essential for forecasting volatility.

5.2 Forecasting realized volatility

In previous forecast experiments, we forecast the daily squared return, which is a noisy proxy for the true latent volatility. Now, we consider realized volatility as a volatility proxy since it is an accurate measure of the true latent volatility; see **?**, **?**. In this section, we compare the performance of three SV models to the HAR-RV model. Several estimators have been proposed in realized volatility literature, but following the results of **?**, we use the 5-minute RV which is constructed from five-minute intraday returns. In the case of SV(p) models, we replace the squared return by realized volatility (squared return is considered as the observed process in SV models) and then estimate the models by our W-ARMA-SV estimator.

We consider five assets: S&P 500, FTSE100, NASDAQ100, N225, SSMI20 indices. Their 5-minute realized volatilities are sourced from the Oxford-Man Institute's Realized Library. The main results of these forecast experiments are reported in Tables A9-A14. For easy comparison, we report the relative MSE, the relative MAE, and the relative R2LOG of forecast error. These are relative to the HAR-RV model, and hence, values smaller than unity indicate better forecast performance than the HAR-RV model. We also report the MCS p-value for the corresponding model and highlight the best model by boldface color font.

In design 1, Tables A9-A11, when we forecast a moderately volatile period after the financial crisis, in most cases, higher-order SV models [SV(2) or SV(3)] provide superior forecasts. This finding is consistent across different evaluation measures. Out of 30 cases (across five assets and six forecast horizons), SV models delivered the best forecast performance in almost all cases (according to MSE, MAE, R2LOG), except for the 1-week ahead forecasting of SSMI20 volatility in terms of MSE. The forecasting performances of higher-order SV models are ranked top in 80%, 97%, and 97% of cases according to MSE, MAE, R2LOG, respectively. If we consider only longer horizons (2-week, 3-week, 1-month) then these winning percentages of SV(p) models are increased to 87%, 100% and 100%, while for short horizons (1-day, 2-day, 1-week) these percentages are 73%, 93%, and 93%. One out of ninety cases (across five assets, six forecast horizons, and three loss measures), the HAR-RV model produces the best forecasting performance.

Between the SV(2) and SV(3) model, the performance of the SV(2) model is better in short forecast horizons but SV(3) is better in long horizons. This finding tells us that additional lag term is essential for forecasting realized volatility in long horizons. Compared to all other models, the forecasting performance of the SV(3) model is getting better as the forecast horizon increases. Note that the MCS p-values of other competing models declined significantly in long horizons. The performance of HAR-RV is clearly poor compared to SV(p) models in long horizons since the relative loss measures of SV(p) models are now lower.

In design 2, Tables A12-A14, when we forecast highly volatile periods such as a crisis or expansion, the ranking of models are similar to design 1. Out of ninety cases (across five assets, six forecast horizons, and three loss measures), HAR-RV model produced the best forecasts in 3% of cases, whereas SV(p) models delivered best forecasts in 86% of cases.

In both designs, our findings suggest that SV(p) models are better in forecasting realized volatility. So fitting non-parametric volatility measures in traditional parametric models can provide better forecasting performance. This also tells us that the HAR-RV model is not capturing the proper mean dynamics that comes from the moving average part of the market microstructure noise during the financial crisis. As pointed out by **?**, if several factors influence the dynamics of RV, then RV follows an ARMA-type process. In this study, within a parametric SV framework, we model realized volatility as a non-Gaussian ARMA process [see(3.1)].

6 Conclusion

We consider higher-order SV models for forecasting volatility, in which we depart from the existing literature that mostly uses first-order persistence in latent volatility. We estimate this class of models by the simple W-ARMA estimator proposed by **?**. Given the simple estimates with the procedure discussed in Section 4, forecasting volatility is straightforward.

Empirically, we find that higher-order SV models' forecasting performance is superior to GARCH and HAR-RV models. This finding holds even if a high volatility period (such as financial crisis) is included in the estimation sample or the forecasted sample. These inferences are not only based on a standard forecasting precision assessment but also on formal prediction tests. These findings highlight the usefulness of higher-order SV models for volatility forecasting.

Volatility forecasting with higher-order stochastic volatility models

Md. Nazmul Ahsan[†] Jean-Marie Dufour[‡] Gabriel Rodriguez-Rondon[§]

August 27, 2021

Technical Appendix

[†]Centre interuniversitaire de recherche en analyse des organisations (CIRANO). Mailing address: 1130 Sherbrooke West, Suite 1400, Montreal, Quebec, H3A 2M8, Canada; e-mail: ahsann@cirano.qc.ca.

[‡]William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 414, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 6071; FAX: (1) 514 398 4800; e-mail: jean-marie.dufour@mcgill.ca. Web page: http://www.jeanmariedufour.com

[§]Ph.D. Candidate, Department of Economics, McGill University and Centre interuniversitaire de recherche en analyse des organisations (CIRANO). Mailing address: Department of Economics, McGill University. e-mail: gabriel.rodriguezrondon@mail.mcgill.ca. Wed page: https://grodriguezrondon.com

A.1 Tables

Table A1. Summary statistics of full sample of experiment - 1

S&P 500 index, September 01, 2005 to August 31, 2010, $T = 1258$								
Series	Mean	Std. Dev.	Kurtosis	Skewness	Range	Max	Min	LB(10)
y	0.00	0.67	11.39	-0.17	8.83	4.63	-4.20	46.9
y^2	0.45	1.44	99.59	8.58	21.41	21.41	0.00	1117.5
y^*	0.00	2.67	4.80	-0.93	19.56	6.11	-13.45	473.3
RV5	0.00	0.00	108.51	8.05	0.01	0.01	0.00	4275.8
RV5-SS	0.00	0.00	108.51	8.05	0.01	0.01	0.00	4275.8
BV5	0.00	0.00	91.24	7.73	0.01	0.01	0.00	4362.1
BV5-SS	0.00	0.00	91.24	7.73	0.01	0.01	0.00	4362.1
MedRV	0.00	0.00	76.43	7.43	0.00	0.00	0.00	4457.5
TSRK	0.00	0.00	152.67	9.47	0.01	0.01	0.00	3729.8
RK	0.00	0.00	54.33	6.35	0.01	0.01	0.00	4273.5
RSV5	0.00	0.00	79.42	7.12	0.00	0.00	0.00	3706.0
RSV5-SS	0.00	0.00	79.42	7.12	0.00	0.00	0.00	3706.0
Log-RV5	-4.12	0.52	3.18	0.60	3.17	-2.11	-5.28	7569.5
Log-RV5-SS	-4.12	0.52	3.18	0.60	3.17	-2.11	-5.28	7569.5
Log-BV5	-4.21	0.52	3.23	0.62	3.35	-2.22	-5.57	7831.3
Log-BV5-SS	-4.21	0.52	3.23	0.62	3.35	-2.22	-5.57	7831.3
Log-MedRV	-4.51	0.56	3.22	0.54	3.45	-2.54	-5.99	7470.3
Log-TSRK	-4.14	0.51	3.31	0.67	3.19	-2.08	-5.27	8126.4
Log-RK	-4.20	0.57	3.09	0.40	3.44	-2.29	-5.73	5562.8
Log-RSV5	-4.47	0.57	3.02	0.48	3.48	-2.45	-5.93	6017.0
Log-RSV5-SS	-4.47	0.57	3.02	0.48	3.48	-2.45	-5.93	6017.0

Notes:

1. $y_t = r_t - \hat{\mu}_r$ is the residual return, y_t^2 is the squared of residual return and y_t^* is the residual of log squared of residual return.
2. RV5 is the 5-minute realized variance, BV5 is the 5-minute bi-power variation, RSV5 is the 5-minute realized semi-variance, RK is the realized kernel, TSRK is the two-scale realized kernels, and MedRV is the median realized volatility.
3. SS denotes the use of 1-minute subsamples in the calculation of realized volatility estimators.
4. LB(10) is the heteroskedasticity-corrected Ljung-Box statistics with 10 lags. The critical values for LB(10) are: 15.99 (10%), 18.31 (5%), and 23.21 (1%).

Table A2. Summary statistics of full sample of experiment - 2

S&P 500 index, January 01, 2005 to December 31, 2009, $T = 1259$								
Series	Mean	Std. Dev.	Kurtosis	Skewness	Range	Max	Min	LB(10)
y	0.00	0.65	12.78	-0.18	8.83	4.62	-4.20	53.7
y^2	0.42	1.43	102.64	8.78	21.39	21.39	0.00	1181.7
y^*	0.00	2.69	6.45	-1.14	24.16	6.22	-17.94	466.2
RV5	0.00	0.00	111.55	8.21	0.01	0.01	0.00	4483.6
RV5-SS	0.00	0.00	111.55	8.21	0.01	0.01	0.00	4483.6
BV5	0.00	0.00	93.70	7.88	0.01	0.01	0.00	4547.5
BV5-SS	0.00	0.00	93.70	7.88	0.01	0.01	0.00	4547.5
MedRV	0.00	0.00	77.39	7.50	0.00	0.00	0.00	4571.3
TSRK	0.00	0.00	160.47	9.74	0.01	0.01	0.00	4028.3
RK	0.00	0.00	55.33	6.44	0.01	0.01	0.00	4425.0
RSV5	0.00	0.00	82.90	7.32	0.00	0.00	0.00	3957.9
RSV5-SS	0.00	0.00	82.90	7.32	0.00	0.00	0.00	3957.9
Log-RV5	-4.18	0.53	3.34	0.77	3.17	-2.11	-5.28	8036.3
Log-RV5-SS	-4.18	0.53	3.34	0.77	3.17	-2.11	-5.28	8036.3
Log-BV5	-4.27	0.53	3.40	0.80	3.35	-2.22	-5.57	8215.8
Log-BV5-SS	-4.27	0.53	3.40	0.80	3.35	-2.22	-5.57	8215.8
Log-MedRV	-4.57	0.56	3.37	0.71	3.45	-2.54	-5.99	7726.2
Log-TSRK	-4.20	0.51	3.46	0.85	3.19	-2.08	-5.27	8598.9
Log-RK	-4.25	0.57	3.28	0.53	3.53	-2.29	-5.82	5826.7
Log-RSV5	-4.53	0.57	3.17	0.65	3.48	-2.45	-5.93	6520.8
Log-RSV5-SS	-4.53	0.57	3.17	0.65	3.48	-2.45	-5.93	6520.8

Notes:

1. $y_t = r_t - \hat{\mu}_r$ is the residual return, y_t^2 is the squared of residual return and y_t^* is the residual of log squared of residual return.
2. RV5 is the 5-minute realized variance, BV5 is the 5-minute bi-power variation, RSV5 is the 5-minute realized semi-variance, RK is the realized kernel, TSRK is the two-scale realized kernels, and MedRV is the median realized volatility.
3. SS denotes the use of 1-minute subsamples in the calculation of realized volatility estimators.
4. LB(10) is the heteroskedasticity-corrected Ljung-Box statistics with 10 lags. The critical values for LB(10) are: 15.99 (10%), 18.31 (5%), and 23.21 (1%).

Table A3. Forecasting squared return: relative MSE and associated MCS p-value during moderate volatility regimes

	1 – *day*			2 – *day*			1 – *week*			2 – *week*			3 – *week*			1 – *month*		
S&P 500	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
SV(1)	0.196	0.12	0.06	0.749	0.16	0.07	1.114	0.35	0.33	1.162	0.39	0.27	1.116	0.31	0.28	1.064	0.07	0.22
SV(2)	0.194	0.12	0.06	0.742	0.16	0.07	1.124	0.33	0.32	1.160	0.26	0.24	1.117	0.07	0.18	1.062	0.07	0.24
SV(3)	**0.180****	1.00	1.00	**0.676****	1.00	1.00	0.982*	0.65	0.83	1.026*	0.64	0.83	**0.985****	1.00	1.00	**0.967****	1.00	1.00
SV(3)*	0.185	0.12	0.07	0.696	0.16	0.08	1.014*	0.65	0.66	1.040*	0.64	0.67	0.998*	0.95	0.94	0.973**	0.97	0.95
GARCH(1,1)	0.197	0.12	0.06	0.750	0.16	0.07	1.104	0.35	0.34	1.102	0.49	0.40	1.056	0.39	0.52	1.005*	0.82	0.82
GARCH(1,2)	0.205	0.12	0.06	0.777	0.16	0.07	1.144	0.33	0.27	1.132	0.49	0.31	1.076	0.36	0.38	1.017*	0.63	0.72
GARCH(2,1)	0.197	0.12	0.06	0.750	0.16	0.07	1.104	0.35	0.37	1.102	0.49	0.37	1.056	0.36	0.38	1.005*	0.82	0.82
GARCH(2,2)	0.202	0.12	0.06	0.767	0.16	0.07	1.130	0.33	0.30	1.122	0.49	0.35	1.068	0.36	0.38	1.012*	0.82	0.82
GARCH(3,3)	0.204	0.12	0.06	0.775	0.16	0.07	1.143	0.33	0.28	1.131	0.49	0.33	1.075	0.36	0.38	1.015*	0.82	0.80
GJR(1,1)	0.214	0.12	0.05	0.811	0.13	0.07	1.193	0.23	0.21	1.173	0.26	0.21	1.111	0.31	0.33	1.043	0.38	0.51
GJR(2,2)	0.214	0.12	0.04	0.811	0.13	0.07	1.193	0.23	0.18	1.173	0.26	0.19	1.111	0.31	0.29	1.044	0.38	0.46
GJR(3,3)	0.214	0.06	0.03	0.813	0.07	0.07	1.196	0.07	0.13	1.175	0.07	0.16	1.113	0.07	0.24	1.046	0.07	0.39
EGARCH(1,1)	0.215	0.06	0.03	0.816	0.07	0.07	1.201	0.07	0.11	1.180	0.07	0.12	1.117	0.07	0.19	1.049	0.07	0.32
EGARCH(2,2)	0.215	0.06	0.03	0.815	0.07	0.07	1.198	0.07	0.12	1.179	0.07	0.13	1.117	0.07	0.20	1.049	0.07	0.29
EGARCH(3,3)	0.214	0.12	0.04	0.813	0.13	0.07	1.196	0.23	0.16	1.176	0.07	0.15	1.114	0.07	0.22	1.048	0.07	0.34
HAR-RV5	1.000	0.06	0.02	1.000	0.07	0.07	1.000*	0.65	0.77	1.000*	0.64	0.83	1.000*	0.95	0.94	1.000**	0.97	0.95
HAR-RV5-SS	1.000	0.06	0.02	1.000	0.07	0.07	1.000*	0.65	0.66	1.000*	0.64	0.83	1.000*	0.95	0.94	1.000**	0.97	0.95
HAR-BV5	1.114	0.06	0.02	1.097	0.07	0.07	1.018*	0.65	0.52	1.024*	0.64	0.67	1.038	0.39	0.43	1.047	0.38	0.50
HAR-BV5-SS	1.114	0.06	0.03	1.097	0.07	0.07	1.018*	0.65	0.46	1.024*	0.64	0.59	1.038	0.36	0.38	1.047*	0.54	0.58
HAR-MedRV	1.016	0.06	0.03	0.977	0.07	0.07	0.973*	0.65	0.83	0.999*	0.64	0.83	1.015*	0.54	0.76	1.035*	0.54	0.60
HAR-TSRK	1.135	0.06	0.02	1.111	0.07	0.07	1.043	0.44	0.42	1.028*	0.64	0.58	1.034*	0.54	0.60	1.046	0.07	0.39
HAR-RK	1.076	0.06	0.02	1.067	0.07	0.07	1.022*	0.65	0.46	1.058	0.49	0.43	1.038	0.36	0.38	1.026*	0.63	0.69
HAR-RSV5	0.941	0.06	0.02	0.960	0.07	0.07	0.973*	0.65	0.83	0.996*	0.64	0.83	0.990**	0.95	0.95	1.007*	0.82	0.82
HAR-RSV5-SS	0.941	0.06	0.02	0.960	0.16	0.07	**0.973****	1.00	1.00	**0.996****	1.00	1.00	0.990**	0.95	0.95	1.007*	0.82	0.82

Notes:

1. The sample period is from September 01, 2005 to August 31, 2010 and the number of observations is $T = 1258$. The in-sample is from September 01, 2005 to August 31, 2008 ($T = 753$) and the out-of-sample is from September 01, 2008 to August 31, 2010 ($T = 505$). The in-sample include most volatile part of late-2000s financial crisis.
2. HAR stands for Heterogenous Autoregressive model, RV5 is the 5-minute realized variance, BV5 is the 5-minute bi-power variation, RSV5 is the 5-minute realized semi-variance, RK is the realized kernel, TSRK is the two-scale realized kernels, and MedRV is the median realized volatility.
3. SS denotes the use of 1-minute subsamples in the calculation of realized volatility estimators.
4. The forecasts of SV(3)* model is based on the restricted W-ARMA-SV-OLS estimation.
5. These are relative to the reference model HAR-RV5 and values smaller than unity indicate better forecast performance than the HAR-RV5 model.
6. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
7. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
8. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

Table A4. Forecasting squared return: relative MAE and associated MCS p-value during moderate volatility regimes

	$1-day$			$2-day$			$1-week$			$2-week$			$3-week$			$1-month$		
S&P 500	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
SV(1)	0.307	0.00	0.00	0.670	0.00	0.00	0.796	0.28	0.18	0.816	0.00	0.06	0.814	0.34	0.45	0.770	0.10	0.32
SV(2)	0.293	0.47	0.45	0.639*	0.81	0.75	0.787	0.41	0.33	0.799*	0.66	0.62	0.807*	0.65	0.64	0.762*	0.92	0.92
SV(3)	0.291*	0.60	0.60	**0.634****	1.00	1.00	**0.764****	1.00	1.00	0.784*	0.66	0.62	0.796*	0.65	0.64	0.779	0.10	0.32
SV(3)*	**0.290****	1.00	1.00	0.636*	0.81	0.75	0.767	0.47	0.47	**0.779****	1.00	1.00	**0.787****	1.00	1.00	**0.759****	1.00	1.00
GARCH(1, 1)	0.372	0.00	0.00	0.816	0.00	0.00	0.989	0.00	0.00	1.000	0.00	0.00	1.013	0.00	0.00	0.957	0.00	0.00
GARCH(1, 2)	0.354	0.00	0.00	0.777	0.00	0.00	0.945	0.00	0.00	0.943	0.00	0.00	0.949	0.00	0.00	0.893	0.00	0.00
GARCH(2, 1)	0.372	0.00	0.00	0.816	0.00	0.00	0.989	0.00	0.00	1.000	0.00	0.00	1.013	0.00	0.00	0.957	0.00	0.00
GARCH(2, 2)	0.351	0.00	0.00	0.772	0.00	0.00	0.940	0.00	0.00	0.941	0.00	0.00	0.951	0.00	0.00	0.893	0.00	0.00
GARCH(3, 3)	0.354	0.00	0.00	0.777	0.00	0.00	0.946	0.00	0.00	0.945	0.00	0.00	0.950	0.00	0.00	0.893	0.00	0.00
GJR(1, 1)	0.324	0.00	0.00	0.712	0.00	0.00	0.866	0.00	0.00	0.859	0.00	0.01	0.859	0.00	0.03	0.802	0.03	0.09
GJR(2, 2)	0.324	0.00	0.00	0.713	0.00	0.00	0.867	0.00	0.00	0.860	0.00	0.00	0.859	0.00	0.01	0.803	0.03	0.05
GJR(3, 3)	0.327	0.00	0.00	0.717	0.00	0.00	0.872	0.00	0.00	0.865	0.00	0.00	0.864	0.00	0.00	0.808	0.00	0.00
EGARCH(1, 1)	0.326	0.00	0.00	0.717	0.00	0.00	0.872	0.00	0.00	0.866	0.00	0.00	0.866	0.00	0.00	0.810	0.00	0.00
EGARCH(2, 2)	0.329	0.00	0.00	0.724	0.00	0.00	0.877	0.00	0.00	0.872	0.00	0.00	0.873	0.00	0.00	0.817	0.00	0.00
EGARCH(3, 3)	0.327	0.00	0.00	0.720	0.00	0.00	0.876	0.00	0.00	0.869	0.00	0.00	0.871	0.00	0.00	0.816	0.00	0.00
HAR-RV5	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00
HAR-RV5-SS	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00
HAR-BV5	1.017	0.00	0.00	1.023	0.00	0.00	1.003	0.00	0.00	1.013	0.00	0.00	1.018	0.00	0.00	1.024	0.00	0.00
HAR-BV5-SS	1.017	0.00	0.00	1.023	0.00	0.00	1.003	0.00	0.00	1.013	0.00	0.00	1.018	0.00	0.00	1.024	0.00	0.00
HAR-MedRV	0.966	0.00	0.00	0.983	0.00	0.00	0.968	0.00	0.00	0.991	0.00	0.00	0.998	0.00	0.00	1.012	0.00	0.00
HAR-TSRK	1.032	0.00	0.00	1.027	0.00	0.00	1.018	0.00	0.00	1.018	0.00	0.00	1.024	0.00	0.00	1.027	0.00	0.00
HAR-RK	1.032	0.00	0.00	1.041	0.00	0.00	1.022	0.00	0.00	1.026	0.00	0.00	1.024	0.00	0.00	1.016	0.00	0.00
HAR-RSV5	0.970	0.00	0.00	0.979	0.00	0.00	0.981	0.00	0.00	0.990	0.00	0.00	0.994	0.00	0.00	1.009	0.00	0.00
HAR-RSV5-SS	0.970	0.00	0.00	0.979	0.00	0.00	0.981	0.00	0.00	0.990	0.00	0.00	0.994	0.00	0.00	1.009	0.00	0.00

Notes:

1. The sample period is from September 01, 2005 to August 31, 2010 and the number of observations is $T = 1258$. The in-sample is from September 01, 2005 to August 31, 2008 ($T = 753$) and the out-of-sample is from September 01, 2008 to August 31, 2010 ($T = 505$). The in-sample include most volatile part of late-2000s financial crisis.
2. HAR stands for Heterogenous Autoregressive model, RV5 is the 5-minute realized variance, BV5 is the 5-minute bi-power variation, RSV5 is the 5-minute realized semi-variance, RK is the realized kernel, TSRK is the two-scale realized kernels, and MedRV is the median realized volatility.
3. SS denotes the use of 1-minute subsamples in the calculation of realized volatility estimators.
4. The forecasts of SV(3)* model is based on the restricted W-ARMA-SV-OLS estimation.
5. These are relative to the reference model HAR-RV5 and values smaller than unity indicate better forecast performance than the HAR-RV5 model.
6. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
7. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
8. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

Table A5. Forecasting squared return: relative R2LOG and associated MCS p-value during moderate volatility regimes

	$1-day$			$2-day$			$1-week$			$2-week$			$3-week$			$1-month$		
S&P 500	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
SV(1)	0.558	0.00	0.00	0.755	0.00	0.00	0.823	0.00	0.01	0.807	0.00	0.01	0.782	0.01	0.06	0.783	0.02	0.10
SV(2)	0.489	0.12	0.05	0.692	0.54	0.42	0.756	0.19	0.20	0.752	0.08	0.13	0.744	0.03	0.12	0.751*	0.74	0.74
SV(3)	0.477	0.12	0.09	0.683	0.54	0.42	0.742	0.19	0.20	0.734	0.08	0.13	0.737	0.03	0.12	0.757	0.02	0.10
SV(3)*	**0.476****	1.00	1.00	**0.682****	1.00	1.00	**0.739****	1.00	1.00	**0.726****	1.00	1.00	**0.724****	1.00	1.00	**0.739****	1.00	1.00
GARCH(1, 1)	0.796	0.00	0.00	1.151	0.00	0.00	1.231	0.00	0.00	1.212	0.00	0.00	1.174	0.00	0.00	1.133	0.00	0.00
GARCH(1, 2)	0.755	0.00	0.00	1.091	0.00	0.00	1.166	0.00	0.00	1.146	0.00	0.00	1.107	0.00	0.00	1.063	0.00	0.00
GARCH(2, 1)	0.796	0.00	0.00	1.151	0.00	0.00	1.231	0.00	0.00	1.212	0.00	0.00	1.174	0.00	0.00	1.133	0.00	0.00
GARCH(2, 2)	0.753	0.00	0.00	1.088	0.00	0.00	1.164	0.00	0.00	1.145	0.00	0.00	1.109	0.00	0.00	1.066	0.00	0.00
GARCH(3, 3)	0.757	0.00	0.00	1.093	0.00	0.00	1.169	0.00	0.00	1.149	0.00	0.00	1.110	0.00	0.00	1.064	0.00	0.00
GJR(1, 1)	0.662	0.00	0.00	0.957	0.00	0.00	1.022	0.00	0.00	1.000	0.00	0.00	0.962	0.00	0.00	0.914	0.01	0.01
GJR(2, 2)	0.663	0.00	0.00	0.959	0.00	0.00	1.024	0.00	0.00	1.002	0.00	0.00	0.964	0.00	0.00	0.916	0.01	0.00
GJR(3, 3)	0.672	0.00	0.00	0.971	0.00	0.00	1.036	0.00	0.00	1.014	0.00	0.00	0.974	0.00	0.00	0.926	0.00	0.00
EGARCH(1, 1)	0.673	0.00	0.00	0.974	0.00	0.00	1.041	0.00	0.00	1.020	0.00	0.00	0.981	0.00	0.00	0.934	0.00	0.00
EGARCH(2, 2)	0.679	0.00	0.00	0.988	0.00	0.00	1.051	0.00	0.00	1.032	0.00	0.00	0.994	0.00	0.00	0.950	0.00	0.00
EGARCH(3, 3)	9.811	0.00	0.00	14.473	0.00	0.00	15.225	0.00	0.00	15.150	0.00	0.00	14.703	0.00	0.00	14.289	0.00	0.00
HAR-RV5	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00
HAR-RV5-SS	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00
HAR-BV5	0.979	0.00	0.00	0.990	0.00	0.00	0.991	0.00	0.00	0.991	0.00	0.00	1.000	0.00	0.00	1.002	0.00	0.00
HAR-BV5-SS	0.979	0.00	0.00	0.990	0.00	0.00	0.991	0.00	0.00	0.991	0.00	0.00	1.000	0.00	0.00	1.002	0.00	0.00
HAR-MedRV	0.947	0.00	0.00	0.967	0.00	0.00	0.972	0.00	0.00	0.981	0.00	0.00	0.982	0.00	0.00	0.989	0.00	0.00
HAR-TSRK	0.983	0.00	0.00	0.994	0.00	0.00	1.001	0.00	0.00	1.003	0.00	0.00	1.004	0.00	0.00	1.006	0.00	0.00
HAR-RK	1.023	0.00	0.00	1.019	0.00	0.00	1.011	0.00	0.00	1.008	0.00	0.00	1.004	0.00	0.00	1.007	0.00	0.00
HAR-RSV5	0.958	0.00	0.00	0.976	0.00	0.00	0.994	0.00	0.00	1.001	0.00	0.00	1.000	0.00	0.00	1.005	0.00	0.00
HAR-RSV5-SS	0.958	0.00	0.00	0.976	0.00	0.00	0.994	0.00	0.00	1.001	0.00	0.00	1.000	0.00	0.00	1.005	0.00	0.00

Notes:

1. The sample period is from September 01, 2005 to August 31, 2010 and the number of observations is $T = 1258$. The in-sample is from September 01, 2005 to August 31, 2008 ($T = 753$) and the out-of-sample is from September 01, 2008 to August 31, 2010 ($T = 505$). The in-sample include most volatile part of late-2000s financial crisis.
2. HAR stands for Heterogenous Autoregressive model, RV5 is the 5-minute realized variance, BV5 is the 5-minute bi-power variation, RSV5 is the 5-minute realized semi-variance, RK is the realized kernel, TSRK is the two-scale realized kernels, and MedRV is the median realized volatility.
3. SS denotes the use of 1-minute subsamples in the calculation of realized volatility estimators.
4. The forecasts of SV(3)* model is based on the restricted W-ARMA-SV-OLS estimation.
5. These are relative to the reference model HAR-RV5 and values smaller than unity indicate better forecast performance than the HAR-RV5 model.
6. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
7. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
8. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

Table A6. Forecasting squared return: relative MSE and associated MCS p-value during high volatility regimes

	1 − *day*			2 − *day*			1 − *week*			2 − *week*			3 − *week*			1 − *month*		
S&P 500	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
SV(1)	0.199	0.10	0.07	0.753	0.15	0.08	1.122	0.37	0.34	1.169	0.12	0.22	1.123	0.11	0.20	1.071	0.06	0.15
SV(2)	0.197	0.10	0.07	0.744	0.15	0.08	1.132	0.30	0.28	1.165	0.12	0.26	1.123	0.05	0.13	1.068	0.06	0.17
SV(3)	**0.181****	1.00	1.00	**0.677****	1.00	1.00	0.982*	0.67	0.74	1.027*	0.73	0.89	**0.974****	1.00	1.00	**0.956****	1.00	1.00
SV(3)*	0.186	0.13	0.07	0.695	0.15	0.08	1.017*	0.67	0.69	1.043*	0.73	0.63	1.002*	0.67	0.78	0.982*	0.68	0.72
GARCH(1, 1)	0.197	0.13	0.07	0.743	0.15	0.08	1.097	0.43	0.38	1.096*	0.67	0.54	1.051	0.49	0.39	1.003*	0.47	0.72
GARCH(1, 2)	0.206	0.10	0.07	0.773	0.15	0.07	1.142	0.30	0.25	1.129	0.42	0.34	1.075	0.49	0.38	1.018*	0.42	0.63
GARCH(2, 1)	0.197	0.13	0.07	0.743	0.15	0.08	1.097	0.43	0.37	1.096	0.48	0.44	1.051*	0.54	0.53	1.003*	0.47	0.72
GARCH(2, 2)	0.203	0.10	0.07	0.763	0.15	0.08	1.128	0.37	0.32	1.121	0.48	0.40	1.067	0.49	0.39	1.013*	0.47	0.72
GARCH(3, 3)	0.205	0.10	0.07	0.771	0.15	0.08	1.140	0.30	0.27	1.125	0.48	0.38	1.070	0.49	0.39	1.014*	0.47	0.72
GJR(1, 1)	0.215	0.10	0.05	0.811	0.10	0.06	1.196	0.12	0.19	1.176	0.12	0.18	1.115	0.13	0.30	1.049	0.15	0.43
GJR(2, 2)	0.216	0.10	0.04	0.811	0.10	0.06	1.197	0.12	0.15	1.177	0.12	0.16	1.115	0.13	0.25	1.049	0.15	0.37
GJR(3, 3)	0.216	0.10	0.03	0.813	0.10	0.06	1.199	0.12	0.12	1.179	0.12	0.14	1.117	0.11	0.21	1.051	0.13	0.31
EGARCH(1, 1)	0.217	0.10	0.03	0.817	0.10	0.06	1.205	0.12	0.07	1.184	0.09	0.09	1.123	0.05	0.14	1.055	0.06	0.21
EGARCH(2, 2)	0.217	0.10	0.03	0.815	0.10	0.06	1.202	0.12	0.09	1.182	0.09	0.10	1.121	0.05	0.15	1.054	0.06	0.23
EGARCH(3, 3)	0.216	0.10	0.03	0.813	0.10	0.06	1.201	0.12	0.10	1.180	0.12	0.12	1.119	0.11	0.18	1.053	0.13	0.27
HAR-RV5	1.000	0.10	0.02	1.000	0.10	0.06	1.000*	0.67	0.69	1.000*	0.73	0.89	1.000*	0.67	0.78	1.000*	0.68	0.72
HAR-RV5-SS	1.000	0.10	0.02	1.000	0.10	0.06	1.000*	0.67	0.74	1.000*	0.80	0.92	1.000*	0.54	0.71	1.000*	0.47	0.72
HAR-BV5	1.113	0.10	0.02	1.097	0.10	0.06	1.017*	0.67	0.55	1.024*	0.73	0.69	1.037	0.49	0.39	1.043	0.42	0.55
HAR-BV5-SS	1.113	0.10	0.02	1.097	0.10	0.06	1.017*	0.67	0.46	1.024*	0.73	0.63	1.037	0.49	0.39	1.043	0.42	0.54
HAR-MedRV	1.013	0.10	0.02	0.975	0.10	0.06	**0.972****	1.00	1.00	0.999*	0.80	0.92	1.015*	0.54	0.59	1.031	0.42	0.56
HAR-TSRK	1.140	0.10	0.02	1.114	0.10	0.06	1.044	0.44	0.41	1.029*	0.73	0.60	1.034*	0.54	0.46	1.043	0.15	0.43
HAR-RK	1.073	0.10	0.02	1.065	0.10	0.06	1.022*	0.67	0.46	1.057	0.48	0.44	1.036	0.49	0.39	1.024*	0.42	0.61
HAR-RSV5	0.947	0.10	0.02	0.963	0.10	0.06	0.974*	0.67	0.74	0.996*	0.80	0.92	0.990*	0.67	0.78	1.006*	0.47	0.72
HAR-RSV5-SS	0.947	0.10	0.02	0.963	0.15	0.07	0.974**	0.95	0.95	**0.996****	1.00	1.00	0.990*	0.82	0.82	1.006*	0.47	0.72

Notes:

1. The sample period is from January 01, 2005 to December 31, 2009 and the number of observations is $T = 1259$. The in-sample is from January 01, 2005 to December 31, 2007 ($T = 754$) and the out-of-sample is from January 01, 2008 to December 31, 2009 ($T = 505$). In this setting, we forecast a highly volatile period.
2. HAR stands for Heterogenous Autoregressive model, RV5 is the 5-minute realized variance, BV5 is the 5-minute bi-power variation, RSV5 is the 5-minute realized semi-variance, RK is the realized kernel, TSRK is the two-scale realized kernels, and MedRV is the median realized volatility.
3. SS denotes the use of 1-minute subsamples in the calculation of realized volatility estimators.
4. The forecasts of SV(3)* model is based on the restricted W-ARMA-SV-OLS estimation.
5. These are relative to the reference model HAR-RV5 and values smaller than unity indicate better forecast performance than the HAR-RV5 model.
6. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
7. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
8. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

Table A7. Forecasting squared return: relative MAE and associated MCS p-value during high volatility regimes

	$1-day$			$2-day$			$1-week$			$2-week$			$3-week$			$1-month$		
S&P 500	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
SV(1)	0.313	0.00	0.00	0.661	0.00	0.00	0.797	0.23	0.15	0.825	0.00	0.05	0.831	0.24	0.37	0.801	0.02	0.42
SV(2)	0.300	0.26	0.25	0.632*	0.72	0.67	0.789	0.31	0.24	0.807*	0.67	0.62	0.824*	0.65	0.66	0.792*	0.80	0.80
SV(3)	**0.297****	1.00	1.00	**0.626****	1.00	1.00	**0.763****	1.00	1.00	0.792*	0.67	0.62	0.803*	0.81	0.81	0.799*	0.62	0.72
SV(3)*	0.297*	0.70	0.70	0.627*	0.77	0.77	0.765	0.41	0.41	**0.787****	1.00	1.00	**0.800****	1.00	1.00	**0.784****	1.00	1.00
GARCH(1, 1)	0.339	0.00	0.00	0.719	0.00	0.00	0.881	0.00	0.01	0.903	0.00	0.01	0.924	0.00	0.01	0.882	0.00	0.01
GARCH(1, 2)	0.330	0.00	0.00	0.701	0.00	0.00	0.862	0.00	0.01	0.871	0.00	0.01	0.884	0.00	0.01	0.842	0.00	0.01
GARCH(2, 1)	0.339	0.00	0.00	0.719	0.00	0.00	0.881	0.00	0.01	0.903	0.00	0.01	0.924	0.00	0.01	0.882	0.00	0.01
GARCH(2, 2)	0.328	0.00	0.00	0.697	0.00	0.00	0.858	0.00	0.01	0.870	0.00	0.01	0.886	0.00	0.01	0.843	0.00	0.01
GARCH(3, 3)	0.329	0.00	0.00	0.698	0.00	0.00	0.858	0.00	0.01	0.868	0.00	0.01	0.881	0.00	0.01	0.838	0.00	0.01
GJR(1, 1)	0.319	0.00	0.00	0.677	0.00	0.00	0.832	0.04	0.04	0.835	0.00	0.05	0.843	0.21	0.24	0.801	0.02	0.42
GJR(2, 2)	0.319	0.03	0.02	0.678	0.01	0.02	0.832	0.04	0.03	0.835	0.00	0.05	0.843	0.21	0.20	0.801	0.02	0.42
GJR(3, 3)	0.320	0.00	0.00	0.679	0.00	0.00	0.834	0.00	0.01	0.837	0.00	0.02	0.845	0.00	0.06	0.802	0.00	0.16
EGARCH(1, 1)	0.321	0.00	0.00	0.682	0.00	0.00	0.838	0.00	0.01	0.841	0.00	0.01	0.850	0.00	0.01	0.807	0.00	0.03
EGARCH(2, 2)	0.323	0.00	0.00	0.685	0.00	0.00	0.841	0.00	0.01	0.844	0.00	0.01	0.852	0.00	0.01	0.810	0.00	0.01
EGARCH(3, 3)	0.322	0.00	0.00	0.683	0.00	0.00	0.839	0.00	0.01	0.842	0.00	0.01	0.850	0.00	0.01	0.808	0.00	0.01
HAR-RV5	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.01	1.000	0.00	0.01	1.000	0.00	0.01	1.000	0.00	0.01
HAR-RV5-SS	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.01	1.000	0.00	0.01	1.000	0.00	0.01	1.000	0.00	0.01
HAR-BV5	1.022	0.00	0.00	1.025	0.00	0.00	1.004	0.00	0.01	1.015	0.00	0.01	1.019	0.00	0.01	1.022	0.00	0.01
HAR-BV5-SS	1.022	0.00	0.00	1.025	0.00	0.00	1.004	0.00	0.01	1.015	0.00	0.01	1.019	0.00	0.01	1.022	0.00	0.01
HAR-MedRV	0.971	0.00	0.00	0.985	0.00	0.00	0.969	0.00	0.01	0.991	0.00	0.01	0.999	0.00	0.01	1.009	0.00	0.01
HAR-TSRK	1.042	0.00	0.00	1.033	0.00	0.00	1.021	0.00	0.01	1.022	0.00	0.01	1.028	0.00	0.01	1.028	0.00	0.01
HAR-RK	1.021	0.00	0.00	1.033	0.00	0.00	1.018	0.00	0.01	1.023	0.00	0.01	1.019	0.00	0.01	1.011	0.00	0.01
HAR-RSV5	0.979	0.00	0.00	0.989	0.00	0.00	0.987	0.00	0.01	0.994	0.00	0.01	0.997	0.00	0.01	1.010	0.00	0.01
HAR-RSV5-SS	0.979	0.00	0.00	0.989	0.00	0.00	0.987	0.00	0.01	0.994	0.00	0.01	0.997	0.00	0.01	1.010	0.00	0.01

Notes:

1. The sample period is from January 01, 2005 to December 31, 2009 and the number of observations is $T = 1259$. The in-sample is from January 01, 2005 to December 31, 2007 ($T = 754$) and the out-of-sample is from January 01, 2008 to December 31, 2009 ($T = 505$). In this setting, we forecast a highly volatile period.
2. HAR stands for Heterogenous Autoregressive model, RV5 is the 5-minute realized variance, BV5 is the 5-minute bi-power variation, RSV5 is the 5-minute realized semi-variance, RK is the realized kernel, TSRK is the two-scale realized kernels, and MedRV is the median realized volatility.
3. SS denotes the use of 1-minute subsamples in the calculation of realized volatility estimators.
4. The forecasts of SV(3)* model is based on the restricted W-ARMA-SV-OLS estimation.
5. These are relative to the reference model HAR-RV5 and values smaller than unity indicate better forecast performance than the HAR-RV5 model.
6. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
7. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
8. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

Table A8. Forecasting squared return: relative R2LOG and associated MCS p-value during high volatility regimes

	$1-day$			$2-day$			$1-week$			$2-week$			$3-week$			$1-month$		
S&P 500	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
SV(1)	0.526	0.00	0.00	0.710	0.00	0.00	0.797	0.00	0.02	0.828	0.00	0.00	0.809	0.00	0.02	0.844	0.00	0.01
SV(2)	0.469	0.30	0.23	0.657*	0.47	0.60	0.738	0.30	0.38	0.760	0.23	0.16	0.754	0.24	0.23	0.782	0.32	0.20
SV(3)	0.462	0.38	0.38	0.655*	0.47	0.60	0.728	0.30	0.38	0.720	0.26	0.26	0.703*	0.60	0.60	0.730	0.32	0.20
SV(3)*	**0.461****	1.00	1.00	**0.654****	1.00	1.00	**0.725****	1.00	1.00	**0.717****	1.00	1.00	**0.700****	1.00	1.00	**0.725****	1.00	1.00
GARCH(1, 1)	0.641	0.00	0.00	0.918	0.00	0.00	1.016	0.00	0.00	1.025	0.00	0.00	1.002	0.00	0.00	0.983	0.00	0.00
GARCH(1, 2)	0.615	0.00	0.00	0.882	0.00	0.00	0.975	0.00	0.00	0.979	0.00	0.00	0.952	0.00	0.00	0.928	0.00	0.00
GARCH(2, 1)	0.641	0.00	0.00	0.918	0.00	0.00	1.016	0.00	0.00	1.025	0.00	0.00	1.002	0.00	0.00	0.983	0.00	0.00
GARCH(2, 2)	0.614	0.00	0.00	0.879	0.00	0.00	0.973	0.00	0.00	0.979	0.00	0.00	0.955	0.00	0.00	0.932	0.00	0.00
GARCH(3, 3)	0.612	0.00	0.00	0.878	0.00	0.00	0.971	0.00	0.00	0.975	0.00	0.00	0.948	0.00	0.00	0.923	0.00	0.00
GJR(1, 1)	0.564	0.00	0.00	0.809	0.00	0.00	0.893	0.00	0.00	0.892	0.00	0.00	0.863	0.00	0.00	0.834	0.05	0.05
GJR(2, 2)	0.565	0.00	0.00	0.811	0.00	0.00	0.895	0.00	0.00	0.893	0.00	0.00	0.865	0.00	0.00	0.835	0.00	0.01
GJR(3, 3)	0.569	0.00	0.00	0.816	0.00	0.00	0.900	0.00	0.00	0.898	0.00	0.00	0.869	0.00	0.00	0.839	0.00	0.01
EGARCH(1, 1)	0.576	0.00	0.00	0.826	0.00	0.00	0.914	0.00	0.00	0.912	0.00	0.00	0.885	0.00	0.00	0.855	0.00	0.00
EGARCH(2, 2)	0.582	0.00	0.00	0.833	0.00	0.00	0.921	0.00	0.00	0.918	0.00	0.00	0.893	0.00	0.00	0.861	0.00	0.00
EGARCH(3, 3)	0.577	0.00	0.00	0.828	0.00	0.00	0.914	0.00	0.00	0.913	0.00	0.00	0.886	0.00	0.00	0.856	0.00	0.00
HAR-RV5	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00
HAR-RV5-SS	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00
HAR-BV5	0.988	0.00	0.00	0.995	0.00	0.00	0.997	0.00	0.00	0.999	0.00	0.00	1.004	0.00	0.00	1.002	0.00	0.00
HAR-BV5-SS	0.988	0.00	0.00	0.995	0.00	0.00	0.997	0.00	0.00	0.999	0.00	0.00	1.004	0.00	0.00	1.002	0.00	0.00
HAR-MedRV	0.951	0.00	0.00	0.966	0.00	0.00	0.972	0.00	0.00	0.984	0.00	0.00	0.987	0.00	0.00	0.986	0.00	0.00
HAR-TSRK	0.995	0.00	0.00	1.001	0.00	0.00	1.010	0.00	0.00	1.011	0.00	0.00	1.012	0.00	0.00	1.012	0.00	0.00
HAR-RK	1.018	0.00	0.00	1.012	0.00	0.00	1.004	0.00	0.00	1.005	0.00	0.00	0.999	0.00	0.00	1.000	0.00	0.00
HAR-RSV5	0.973	0.00	0.00	0.988	0.00	0.00	1.003	0.00	0.00	1.006	0.00	0.00	1.005	0.00	0.00	1.008	0.00	0.00
HAR-RSV5-SS	0.973	0.00	0.00	0.988	0.00	0.00	1.003	0.00	0.00	1.006	0.00	0.00	1.005	0.00	0.00	1.008	0.00	0.00

Notes:

1. The sample period is from January 01, 2005 to December 31, 2009 and the number of observations is $T = 1259$. The in-sample is from January 01, 2005 to December 31, 2007 ($T = 754$) and the out-of-sample is from January 01, 2008 to December 31, 2009 ($T = 505$). In this setting, we forecast a highly volatile period.
2. HAR stands for Heterogenous Autoregressive model, RV5 is the 5-minute realized variance, BV5 is the 5-minute bi-power variation, RSV5 is the 5-minute realized semi-variance, RK is the realized kernel, TSRK is the two-scale realized kernels, and MedRV is the median realized volatility.
3. SS denotes the use of 1-minute subsamples in the calculation of realized volatility estimators.
4. The forecasts of SV(3)* model is based on the restricted W-ARMA-SV-OLS estimation.
5. These are relative to the reference model HAR-RV5 and values smaller than unity indicate better forecast performance than the HAR-RV5 model.
6. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
7. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
8. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

Table A9. Forecasting realized volatility: relative MSE and associated MCS p-value during moderate volatility regimes

	1 − *day*			2 − *day*			1 − *week*			2 − *week*			3 − *week*			1 − *month*		
S&P 500	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000**	0.95	0.95	1.000*	0.69	0.69	1.000*	0.81	0.84	1.000	0.24	0.31	1.000	0.21	0.23	1.000	0.15	0.13
SV(1)	**0.991****	1.00	1.00	**0.972****	1.00	1.00	0.986*	0.81	0.84	0.969	0.24	0.31	0.949	0.21	0.23	0.893	0.15	0.13
SV(2)	0.996**	0.95	0.95	0.986*	0.69	0.69	0.981*	0.81	0.84	0.963	0.24	0.31	0.944	0.21	0.23	0.877	0.15	0.13
SV(3)	1.087	0.16	0.17	1.059	0.24	0.32	**0.972****	1.00	1.00	**0.931****	1.00	1.00	**0.881****	1.00	1.00	**0.833****	1.00	1.00
FTSE100	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000*	0.54	0.54	1.000	0.46	0.44	1.000*	0.91	0.92	1.000*	0.89	0.83	1.000	0.37	0.41	1.000	0.27	0.13
SV(1)	1.004*	0.54	0.54	0.943*	0.89	0.89	1.009	0.41	0.66	**0.975****	1.00	1.00	**0.953****	1.00	1.00	0.950	0.27	0.14
SV(2)	**0.975****	1.00	1.00	**0.941****	1.00	1.00	**0.996****	1.00	1.00	0.981*	0.89	0.83	0.955*	0.85	0.78	0.947	0.27	0.26
SV(3)	1.046	0.27	0.21	1.019	0.46	0.47	1.001*	0.91	0.92	0.982*	0.89	0.83	0.958*	0.85	0.78	**0.943****	1.00	1.00
NASDAQ100	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000*	0.82	0.82	1.000*	0.93	0.95	1.000*	0.69	0.68	1.000	0.19	0.36	1.000	0.38	0.53	1.000	0.18	0.26
SV(1)	1.018*	0.80	0.79	1.001*	0.93	0.95	0.989*	0.69	0.68	1.020	0.12	0.16	1.021	0.14	0.22	0.984	0.09	0.14
SV(2)	**0.987****	1.00	1.00	**0.996****	1.00	1.00	**0.976****	1.00	1.00	0.985	0.19	0.36	0.976	0.38	0.53	0.943	0.18	0.26
SV(3)	1.087	0.29	0.25	1.070*	0.60	0.64	1.011*	0.69	0.68	**0.970****	1.00	1.00	**0.965****	1.00	1.00	**0.920****	1.00	1.00
N225	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.38	0.32	1.000	0.16	0.15	1.000**	0.99	0.99	1.000	0.30	0.15	1.000*	0.76	0.73	1.000*	0.78	0.70
SV(1)	0.943*	0.67	0.67	0.956	0.16	0.15	1.060	0.26	0.53	1.061	0.30	0.13	1.051	0.23	0.26	1.024	0.21	0.29
SV(2)	**0.937****	1.00	1.00	**0.922****	1.00	1.00	0.998**	0.99	0.99	0.940	0.30	0.15	0.973*	0.94	0.94	0.980*	0.78	0.70
SV(3)	1.203	0.32	0.19	1.084	0.16	0.15	**0.996****	1.00	1.00	**0.907****	1.00	1.00	**0.973****	1.00	1.00	**0.975****	1.00	1.00
SSMI20	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000*	0.65	0.60	1.000*	0.80	0.77	**1.000****	1.00	1.00	1.000*	0.66	0.55	1.000	0.37	0.27	1.000	0.14	0.12
SV(1)	0.968*	0.65	0.60	**0.978****	1.00	1.00	1.003**	0.98	0.97	0.984*	0.66	0.55	0.958	0.37	0.27	0.943	0.14	0.12
SV(2)	**0.963****	1.00	1.00	0.990*	0.80	0.77	1.006**	0.98	0.97	0.974*	0.66	0.55	0.948	0.37	0.27	0.924	0.26	0.26
SV(3)	1.128	0.29	0.24	1.127	0.43	0.46	1.011**	0.98	0.97	**0.961****	1.00	1.00	**0.932****	1.00	1.00	**0.912****	1.00	1.00

Notes:

1. The sample period is from September 01, 2005 to August 31, 2010 where the in-sample is from September 01, 2005 to August 31, 2008 and the out-of-sample is from September 01, 2008 to August 31, 2010. The in-sample include the most volatile part of the late-2000s financial crisis.
2. HAR stands for the Heterogenous Autoregressive model, and we used 5-minute RV following the results of **?**.
3. These are relative to the reference model HAR-RV and values smaller than unity indicate better forecast performance than HAR-RV model.
4. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
5. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
6. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

Table A10. Forecasting realized volatility: relative MAE and associated MCS p-value during moderate volatility regimes

	$1-day$			$2-day$			$1-week$			$2-week$			$3-week$			$1-month$		
S&P 500	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.01	0.01	1.000	0.10	0.06	1.000	0.18	0.13	1.000	0.15	0.17	1.000	0.05	0.05	1.000	0.05	0.06
SV(1)	0.936*	0.89	0.89	0.950	0.48	0.48	0.916	0.29	0.26	0.849*	0.85	0.79	0.808	0.17	0.18	0.775	0.21	0.23
SV(2)	**0.934****	1.00	1.00	**0.944****	1.00	1.00	0.905	0.36	0.36	0.847*	0.85	0.79	0.807	0.15	0.18	0.771	0.18	0.23
SV(3)	0.968	0.04	0.08	0.983	0.10	0.07	**0.895****	1.00	1.00	**0.843****	1.00	1.00	**0.785****	1.00	1.00	**0.745****	1.00	1.00
FTSE100	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.22	0.29	1.000	0.12	0.12	1.000	0.05	0.02	1.000	0.05	0.03	1.000	0.01	0.00	1.000	0.01	0.00
SV(1)	0.986	0.44	0.44	0.987	0.12	0.16	0.960	0.05	0.04	0.862	0.29	0.20	0.826	0.04	0.02	0.789	0.04	0.04
SV(2)	**0.971****	1.00	1.00	**0.950****	1.00	1.00	0.928	0.25	0.25	0.854	0.31	0.31	0.820	0.04	0.03	0.783	0.04	0.04
SV(3)	1.015	0.17	0.23	0.952*	0.92	0.92	**0.902****	1.00	1.00	**0.848****	1.00	1.00	**0.807****	1.00	1.00	**0.771****	1.00	1.00
NASDAQ100	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.32	0.39	1.000	0.10	0.11	1.000	0.22	0.12	1.000	0.10	0.04	1.000	0.13	0.08	1.000	0.11	0.05
SV(1)	**0.971****	1.00	1.00	0.966	0.12	0.17	0.944	0.27	0.19	0.883	0.11	0.07	0.866	0.17	0.18	0.859	0.11	0.05
SV(2)	0.976*	0.80	0.80	**0.943****	1.00	1.00	0.932	0.30	0.30	0.868	0.18	0.18	0.848*	0.72	0.72	0.832	0.11	0.10
SV(3)	1.003	0.18	0.24	0.965	0.23	0.23	**0.920****	1.00	1.00	**0.860****	1.00	1.00	**0.847****	1.00	1.00	**0.822****	1.00	1.00
N225	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.06	0.03	1.000	0.01	0.01	1.000	0.05	0.09	1.000	0.09	0.06	1.000	0.13	0.05	1.000	0.09	0.05
SV(1)	0.968	0.07	0.07	0.952	0.02	0.02	0.937	0.05	0.09	0.935	0.09	0.08	0.915	0.13	0.07	0.898	0.11	0.18
SV(2)	**0.950****	1.00	1.00	**0.927****	1.00	1.00	**0.915****	1.00	1.00	0.880	0.32	0.32	0.868*	0.50	0.50	0.865**	0.97	0.97
SV(3)	1.060	0.06	0.03	1.003	0.02	0.02	0.944	0.05	0.09	**0.866****	1.00	1.00	**0.860****	1.00	1.00	**0.865****	1.00	1.00
SSMI20	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000*	0.49	0.54	1.000*	0.56	0.66	1.000	0.17	0.13	1.000	0.01	0.01	1.000	0.04	0.01	1.000	0.04	0.02
SV(1)	0.995*	0.49	0.54	0.990*	0.56	0.66	0.975	0.17	0.13	0.890	0.02	0.02	0.861	0.05	0.05	0.831	0.13	0.10
SV(2)	**0.987****	1.00	1.00	0.984**	0.98	0.98	0.974	0.17	0.13	0.891	0.01	0.01	0.868	0.04	0.03	0.824	0.13	0.10
SV(3)	1.025	0.40	0.37	**0.984****	1.00	1.00	**0.934****	1.00	1.00	**0.862****	1.00	1.00	**0.839****	1.00	1.00	**0.813****	1.00	1.00

Notes:

1. The sample period is from September 01, 2005 to August 31, 2010 where the in-sample is from September 01, 2005 to August 31, 2008 and the out-of-sample is from September 01, 2008 to August 31, 2010. The in-sample include the most volatile part of the late-2000s financial crisis.
2. HAR stands for the Heterogenous Autoregressive model, and we used 5-minute RV following the results of **?**.
3. These are relative to the reference model HAR-RV and values smaller than unity indicate better forecast performance than HAR-RV model.
4. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
5. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
6. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

Table A11. Forecasting realized volatility: relative R2LOG and associated MCS p-value during moderate volatility regimes

	$1-day$			$2-day$			$1-week$			$2-week$			$3-week$			$1-month$		
S&P 500	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.01	0.02	1.000	0.07	0.05	1.000	0.07	0.08	1.000	0.27	0.22	1.000	0.16	0.11	1.000	0.07	0.07
SV(1)	0.921	0.46	0.46	0.933	0.09	0.09	0.939	0.07	0.09	0.887	0.51	0.37	0.853	0.16	0.19	0.782	0.24	0.27
SV(2)	**0.906****	1.00	1.00	**0.915****	1.00	1.00	0.909	0.39	0.39	0.878	0.51	0.39	0.852	0.16	0.19	0.783	0.23	0.27
SV(3)	0.990	0.04	0.09	0.973	0.07	0.07	**0.896****	1.00	1.00	**0.861****	1.00	1.00	**0.809****	1.00	1.00	**0.743****	1.00	1.00
FTSE100	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.02	0.02	1.000	0.00	0.01	1.000	0.00	0.01	1.000	0.04	0.02	1.000	0.02	0.00	1.000	0.04	0.01
SV(1)	0.959	0.03	0.03	0.977	0.00	0.01	0.940	0.00	0.01	0.873	0.04	0.03	0.834	0.02	0.00	0.786	0.04	0.02
SV(2)	**0.923****	1.00	1.00	**0.927****	1.00	1.00	0.893*	0.68	0.68	0.849	0.11	0.11	0.812	0.02	0.01	0.770	0.04	0.02
SV(3)	1.009	0.02	0.02	0.964	0.14	0.14	**0.885****	1.00	1.00	**0.829****	1.00	1.00	**0.784****	1.00	1.00	**0.744****	1.00	1.00
NASDAQ100	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.13	0.13	1.000	0.06	0.10	1.000	0.05	0.09	1.000	0.19	0.16	1.000	0.18	0.10	1.000	0.14	0.07
SV(1)	**0.944****	1.00	1.00	0.959	0.06	0.10	0.968	0.05	0.09	0.925	0.19	0.16	0.912	0.18	0.10	0.857	0.14	0.12
SV(2)	0.952*	0.66	0.66	**0.936****	1.00	1.00	0.934	0.40	0.40	0.896	0.36	0.36	0.874	0.20	0.20	0.816	0.29	0.29
SV(3)	0.999	0.08	0.08	0.964	0.06	0.10	**0.920****	1.00	1.00	**0.884****	1.00	1.00	**0.858****	1.00	1.00	**0.801****	1.00	1.00
N225	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.01	0.03	1.000	0.07	0.03	1.000	0.09	0.05	1.000	0.10	0.07
SV(1)	0.942	0.03	0.03	0.951	0.00	0.00	0.970	0.01	0.03	0.945	0.07	0.05	0.937	0.09	0.06	0.918	0.10	0.08
SV(2)	**0.897****	1.00	1.00	**0.914****	1.00	1.00	**0.902****	1.00	1.00	0.859*	0.89	0.89	**0.843****	1.00	1.00	**0.834****	1.00	1.00
SV(3)	1.128	0.00	0.00	1.102	0.00	0.00	0.967	0.03	0.03	**0.857****	1.00	1.00	0.861	0.35	0.35	0.854	0.31	0.31
SSMI20	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.17	0.10	1.000	0.01	0.10	1.000	0.08	0.06	1.000	0.11	0.08	1.000	0.00	0.01	1.000	0.02	0.00
SV(1)	0.969	0.17	0.10	0.984	0.01	0.10	0.979	0.08	0.06	0.931	0.11	0.08	0.887	0.02	0.02	0.832	0.02	0.01
SV(2)	**0.948****	1.00	1.00	**0.957****	1.00	1.00	0.955	0.09	0.09	0.919	0.11	0.08	0.877	0.00	0.01	0.813	0.02	0.01
SV(3)	1.015	0.17	0.10	1.007	0.01	0.10	**0.915****	1.00	1.00	**0.881****	1.00	1.00	**0.830****	1.00	1.00	**0.772****	1.00	1.00

Notes:

1. The sample period is from September 01, 2005 to August 31, 2010 where the in-sample is from September 01, 2005 to August 31, 2008 and the out-of-sample is from September 01, 2008 to August 31, 2010. The in-sample include the most volatile part of the late-2000s financial crisis.
2. HAR stands for the Heterogenous Autoregressive model, and we used 5-minute RV following the results of **?**.
3. These are relative to the reference model HAR-RV and values smaller than unity indicate better forecast performance than HAR-RV model.
4. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
5. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
6. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

Table A12. Forecasting realized volatility: relative MSE and associated MCS p-value during high volatility regimes

	$1-day$			$2-day$			$1-week$			$2-week$			$3-week$			$1-month$		
S&P 500	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000*	0.94	0.95	1.000*	0.69	0.72	1.000*	0.84	0.89	1.000	0.23	0.31	1.000	0.18	0.23	1.000	0.13	0.13
SV(1)	**0.994****	1.00	1.00	**0.974****	1.00	1.00	0.992*	0.84	0.89	0.981	0.23	0.31	0.967	0.18	0.23	0.914	0.13	0.13
SV(2)	1.002*	0.94	0.95	0.989*	0.69	0.72	0.989*	0.84	0.89	0.974	0.23	0.31	0.960	0.18	0.23	0.898	0.13	0.13
SV(3)	1.082	0.18	0.21	1.061	0.26	0.35	**0.980****	1.00	1.00	**0.940****	1.00	1.00	**0.895****	1.00	1.00	**0.852****	1.00	1.00
FTSE100	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000*	0.67	0.63	1.000	0.46	0.39	**1.000****	1.00	1.00	1.000*	0.90	0.93	1.000*	0.70	0.69	1.000	0.18	0.17
SV(1)	0.997*	0.67	0.63	0.949*	0.93	0.93	1.015	0.46	0.67	**0.988****	1.00	1.00	**0.970****	1.00	1.00	0.968	0.18	0.17
SV(2)	**0.977****	1.00	1.00	**0.948****	1.00	1.00	1.004*	0.93	0.93	0.993*	0.90	0.93	0.972*	0.88	0.85	0.963	0.20	0.20
SV(3)	1.058	0.18	0.15	1.024	0.46	0.42	1.008*	0.93	0.93	0.991*	0.90	0.93	0.973*	0.88	0.85	**0.960****	1.00	1.00
NASDAQ100	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000*	0.78	0.79	1.000**	0.99	0.99	1.000*	0.71	0.77	1.000	0.22	0.40	1.000	0.37	0.56	1.000	0.19	0.31
SV(1)	1.014*	0.78	0.79	**0.994****	1.00	1.00	0.992*	0.71	0.77	1.033	0.12	0.16	1.038	0.13	0.22	1.004	0.07	0.13
SV(2)	**0.976****	1.00	1.00	0.996**	0.99	0.99	**0.981****	1.00	1.00	0.992	0.22	0.40	0.988	0.37	0.56	0.957	0.19	0.31
SV(3)	1.089	0.33	0.25	1.083*	0.54	0.59	1.020*	0.65	0.69	**0.977****	1.00	1.00	**0.976****	1.00	1.00	**0.934****	1.00	1.00
N225	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.41	0.33	1.000	0.17	0.16	**1.000****	1.00	1.00	1.000	0.32	0.23	1.000**	0.97	0.97	1.000*	0.88	0.88
SV(1)	0.948*	0.62	0.62	0.966	0.17	0.16	1.072	0.25	0.50	1.080	0.32	0.18	1.071	0.23	0.29	1.044	0.19	0.28
SV(2)	**0.942****	1.00	1.00	**0.929****	1.00	1.00	1.008**	0.99	0.98	0.960	0.32	0.23	0.991**	0.97	0.97	0.999*	0.77	0.85
SV(3)	1.210	0.30	0.17	1.092	0.17	0.16	1.006**	0.99	0.98	**0.927****	1.00	1.00	**0.990****	1.00	1.00	**0.994****	1.00	1.00
SSMI20	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}	RMSE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000*	0.62	0.59	1.000*	0.74	0.74	**1.000****	1.00	1.00	1.000*	0.72	0.70	1.000	0.28	0.31	1.000	0.10	0.11
SV(1)	**0.958****	1.00	1.00	**0.985****	1.00	1.00	1.015*	0.85	0.86	0.997*	0.62	0.61	0.977	0.28	0.31	0.970	0.10	0.11
SV(2)	0.961*	0.83	0.83	0.998*	0.63	0.73	1.022*	0.83	0.82	0.986*	0.72	0.70	0.961	0.28	0.31	0.945	0.27	0.27
SV(3)	1.128	0.23	0.18	1.115	0.39	0.42	1.016*	0.85	0.86	**0.973****	1.00	1.00	**0.948****	1.00	1.00	**0.934****	1.00	1.00

Notes:

1. The sample period is from January 01, 2005 to December 31, 2009 where the in-sample is from January 01, 2005 to December 31, 2007 and the out-of-sample is from January 01, 2008 to December 31, 2009. In this setting, we forecast a highly volatile period.
2. HAR stands for the Heterogenous Autoregressive model, and we used 5-minute RV following the results of **?**.
3. These are relative to the reference model HAR-RV and values smaller than unity indicate better forecast performance than HAR-RV model.
4. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
5. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
6. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

Table A13. Forecasting realized volatility: relative MAE and associated MCS p-value during high volatility regimes

	$1-day$			$2-day$			$1-week$			$2-week$			$3-week$			$1-month$		
S&P 500	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.05	0.02	1.000	0.07	0.06	1.000	0.29	0.27	1.000	0.25	0.25	1.000	0.09	0.10	1.000	0.11	0.11
SV(1)	0.945*	0.67	0.67	0.952*	0.53	0.53	0.919*	0.62	0.67	0.862*	0.85	0.90	0.831	0.44	0.40	0.812	0.31	0.26
SV(2)	**0.939****	1.00	1.00	**0.947****	1.00	1.00	0.914*	0.86	0.86	**0.859****	1.00	1.00	0.828	0.44	0.40	0.805	0.31	0.26
SV(3)	0.968	0.06	0.11	0.989	0.07	0.06	**0.912****	1.00	1.00	0.860**	0.95	0.95	**0.812****	1.00	1.00	**0.782****	1.00	1.00
FTSE100	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.19	0.23	1.000	0.13	0.19	1.000	0.12	0.06	1.000	0.11	0.04	1.000	0.02	0.01	1.000	0.06	0.02
SV(1)	0.991	0.26	0.26	0.992	0.13	0.19	0.964	0.12	0.09	0.892	0.20	0.13	0.859	0.06	0.03	0.834	0.16	0.11
SV(2)	**0.972****	1.00	1.00	**0.958****	1.00	1.00	0.942	0.36	0.36	0.885	0.22	0.22	0.855	0.06	0.03	0.829	0.16	0.11
SV(3)	1.026	0.12	0.15	0.967*	0.64	0.64	**0.923****	1.00	1.00	**0.876****	1.00	1.00	**0.842****	1.00	1.00	**0.819****	1.00	1.00
NASDAQ100	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.32	0.33	1.000	0.20	0.18	1.000	0.33	0.26	1.000	0.17	0.06	1.000	0.30	0.20	1.000	0.13	0.09
SV(1)	**0.971****	1.00	1.00	0.963	0.32	0.27	0.940*	0.70	0.63	0.904	0.17	0.07	0.886	0.46	0.39	0.902	0.13	0.09
SV(2)	0.971**	0.98	0.98	**0.950****	1.00	1.00	0.937*	0.70	0.63	0.889	0.17	0.12	0.872*	0.68	0.68	0.871	0.15	0.15
SV(3)	0.998	0.28	0.25	0.976	0.32	0.27	**0.929****	1.00	1.00	**0.879****	1.00	1.00	**0.870****	1.00	1.00	**0.862****	1.00	1.00
N225	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.11	0.08	1.000	0.05	0.03	1.000	0.06	0.12	1.000	0.15	0.09	1.000	0.13	0.06	1.000	0.16	0.13
SV(1)	0.970	0.35	0.35	0.967	0.05	0.03	0.953	0.06	0.12	0.954	0.15	0.10	0.941	0.13	0.06	0.935	0.16	0.15
SV(2)	**0.962****	1.00	1.00	**0.944****	1.00	1.00	**0.931****	1.00	1.00	0.904	0.31	0.31	0.895	0.27	0.27	0.902*	0.88	0.88
SV(3)	1.075	0.04	0.03	1.009	0.05	0.03	0.948	0.45	0.45	**0.889****	1.00	1.00	**0.881****	1.00	1.00	**0.900****	1.00	1.00
SSMI20	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}	RMAE	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000*	0.84	0.83	1.000**	0.99	0.99	1.000	0.11	0.10	1.000	0.02	0.01	1.000	0.09	0.03	1.000	0.11	0.05
SV(1)	0.991**	0.95	0.95	0.998**	0.99	0.99	0.986	0.11	0.10	0.911	0.02	0.01	0.879	0.11	0.11	0.863	0.15	0.15
SV(2)	**0.991****	1.00	1.00	1.000**	0.98	0.99	0.991	0.08	0.10	0.912	0.02	0.01	0.886	0.09	0.07	0.849	0.40	0.40
SV(3)	1.036	0.24	0.25	**0.996****	1.00	1.00	**0.945****	1.00	1.00	**0.883****	1.00	1.00	**0.861****	1.00	1.00	**0.844****	1.00	1.00

Notes:

1. The sample period is from January 01, 2005 to December 31, 2009 where the in-sample is from January 01, 2005 to December 31, 2007 and the out-of-sample is from January 01, 2008 to December 31, 2009. In this setting, we forecast a highly volatile period.
2. HAR stands for the Heterogenous Autoregressive model, and we used 5-minute RV following the results of **?**.
3. These are relative to the reference model HAR-RV and values smaller than unity indicate better forecast performance than HAR-RV model.
4. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
5. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
6. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

Table A14. Forecasting realized volatility: relative R2LOG and associated MCS p-value during high volatility regimes

	1 − *day*			2 − *day*			1 − *week*			2 − *week*			3 − *week*			1 − *month*		
S&P 500	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.00	0.00	1.000	0.01	0.02	1.000	0.43	0.34	1.000	0.39	0.35	1.000	0.18	0.11	1.000	0.18	0.10
SV(1)	0.925	0.25	0.25	0.922	0.40	0.40	0.933	0.43	0.48	0.914	0.54	0.44	0.915	0.18	0.11	0.870	0.18	0.10
SV(2)	**0.899****	1.00	1.00	**0.913****	1.00	1.00	**0.920****	1.00	1.00	0.908	0.54	0.44	0.897	0.18	0.11	0.850	0.18	0.10
SV(3)	0.968	0.00	0.02	0.981	0.01	0.02	0.921**	0.98	0.98	**0.885****	1.00	1.00	**0.844****	1.00	1.00	**0.788****	1.00	1.00
FTSE100	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.01	0.01	1.000	0.00	0.02	1.000	0.01	0.14	1.000	0.07	0.10	1.000	0.03	0.04	1.000	0.05	0.06
SV(1)	0.970	0.01	0.01	0.964	0.00	0.02	0.972	0.01	0.14	0.958	0.07	0.10	0.940	0.03	0.04	0.912	0.05	0.06
SV(2)	**0.929****	1.00	1.00	**0.931****	1.00	1.00	**0.942****	1.00	1.00	0.938	0.15	0.15	0.918	0.03	0.04	0.889	0.05	0.06
SV(3)	1.052	0.01	0.01	1.008	0.00	0.02	0.954*	0.62	0.62	**0.915****	1.00	1.00	**0.886****	1.00	1.00	**0.860****	1.00	1.00
NASDAQ100	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.14	0.10	1.000	0.21	0.14	1.000	0.43	0.42	1.000	0.14	0.14	1.000	0.07	0.06	1.000	0.06	0.06
SV(1)	**0.937****	1.00	1.00	0.952	0.21	0.14	0.958	0.43	0.42	0.960	0.14	0.14	0.952	0.07	0.06	0.931	0.06	0.06
SV(2)	0.944*	0.64	0.64	**0.934****	1.00	1.00	0.941**	0.97	0.97	0.916	0.27	0.27	0.893	0.07	0.06	0.851	0.10	0.10
SV(3)	1.015	0.08	0.05	0.980	0.21	0.14	**0.940****	1.00	1.00	**0.897****	1.00	1.00	**0.865****	1.00	1.00	**0.825****	1.00	1.00
N225	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.05	0.03	1.000	0.02	0.05	1.000	0.36	0.36	1.000	0.41	0.38	1.000*	0.67	0.62	1.000*	0.72	0.73
SV(1)	0.970	0.05	0.04	0.984	0.02	0.05	1.012	0.02	0.14	1.009	0.12	0.15	1.042	0.12	0.19	1.060	0.10	0.23
SV(2)	**0.936****	1.00	1.00	**0.953****	1.00	1.00	**0.951****	1.00	1.00	0.916*	0.84	0.84	**0.924****	1.00	1.00	**0.947****	1.00	1.00
SV(3)	1.182	0.00	0.00	1.136	0.00	0.00	1.005	0.26	0.34	**0.913****	1.00	1.00	0.936*	0.67	0.64	0.965*	0.72	0.73
SSMI20	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}	RR2LOG	p^M_{MCS}	p^R_{MCS}
HAR-RV	1.000	0.18	0.18	1.000	0.26	0.30	1.000	0.35	0.35	1.000	0.29	0.28	1.000	0.02	0.04	1.000	0.02	0.03
SV(1)	**0.957****	1.00	1.00	0.984	0.26	0.30	0.996	0.17	0.18	0.990	0.11	0.13	0.959	0.02	0.04	0.933	0.02	0.03
SV(2)	0.958*	0.92	0.92	**0.973****	1.00	1.00	0.982	0.35	0.35	0.963	0.29	0.28	0.918	0.02	0.04	0.867	0.16	0.16
SV(3)	1.061	0.07	0.05	1.048	0.26	0.21	**0.950****	1.00	1.00	**0.929****	1.00	1.00	**0.879****	1.00	1.00	**0.843****	1.00	1.00

Notes:

1. The sample period is from January 01, 2005 to December 31, 2009 where the in-sample is from January 01, 2005 to December 31, 2007 and the out-of-sample is from January 01, 2008 to December 31, 2009. In this setting, we forecast a highly volatile period.
2. HAR stands for the Heterogenous Autoregressive model, and we used 5-minute RV following the results of **?**.
3. These are relative to the reference model HAR-RV and values smaller than unity indicate better forecast performance than HAR-RV model.
4. p^M_{MCS} and p^R_{MCS} are associated with $MCS\_T_{\max,\mathcal{M}} = \max_{i\in\mathcal{M}} t_{i,\cdot}$ and $MCS\_T_{R,\mathcal{M}} = \max_{i,j\in\mathcal{M}} \mid t_{i,j} \mid$, respectively.
5. The forecasts in superior model sets $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are defined by the average of $p_{MCS} \geq 0.95$ and the average of $p_{MCS} \geq 0.50$, respectively.
6. The forecasts in $\hat{\mathcal{M}}^*_{5\%}$ and $\hat{\mathcal{M}}^*_{50\%}$ are identified by two and one asterisks, respectively. Boldface color font highlights the best model.

A.2 Forecasting with GARCH models

GARCH Model: The generalized autoregressive conditional heteroskedastic (GARCH) model is an extension of the ARCH model by **?**. If a series exhibits volatility clustering, this suggests that past variances might be predictive of the current variance. The GARCH(p,q) model is an autoregressive moving average model for conditional variances, with p GARCH coefficients associated with lagged variances, and q ARCH coefficients associated with lagged squared innovations or lagged squared residual returns. The GARCH(p,q) model of residual return is

$$
\begin{aligned}
y_t &= \sigma_t z_t, \quad z_t \sim i.i.d \;\; N(0,1),\\
\sigma_t^2 &= \omega + \beta_1 \sigma_{t-1}^2 + \ldots + \beta_p \sigma_{t-p}^2 + \alpha_1 y_{t-1}^2 + \ldots + \alpha_q y_{t-q}^2,
\end{aligned}
$$

where y_t is the residual return observed at time t and σ_t is the corresponding volatility. For stationarity and positivity, the GARCH model has the following constraints:

- $\omega > 0$,
- $\beta_i \geq 0$, $\alpha_j \geq 0$
- $\sum_{i=1}^{p} \beta_i + \sum_{j=1}^{q} \alpha_j < 1$.

The h-step-ahead forecast of the GARCH$(1,1)$ model is computed according to:

$$
\begin{aligned}
&\hat{\sigma}_{t+h|t}^2 = \hat{\omega} + \hat{\beta}_1 \hat{\sigma}_{t+h-1|t}^2 + \hat{\alpha}_1 \hat{y}_{t+h-1|t}^2,\\
&\hat{y}_{t+h|t}^2 = \hat{\sigma}_{t+h|t}^2 \quad if \quad h > 0,\\
&\hat{y}_{t+h|t}^2 = y_{t+h}^2 \quad \hat{\sigma}_{t+h|t}^2 = \sigma_{t+h}^2 \quad if \quad h \leq 0.
\end{aligned}
$$

EGARCH Model: The exponential GARCH (EGARCH) model was developed by **?**. It is a GARCH variant that models the logarithm of the conditional variance process. In addition to modeling the logarithm, the EGARCH model has additional leverage terms to capture asymmetry in volatility clustering. The EGARCH(p,q) model has p GARCH coefficients associated with lagged log variance terms, q ARCH coefficients associated with the magnitude of lagged standardized innovations, and q leverage coefficients associated with signed, lagged standardized innovations. The form of the EGARCH(p,q) model is

$$
\begin{aligned}
y_t &= \sigma_t z_t, \quad z_t \sim i.i.d \;\; N(0,1),\\
\log \sigma_t^2 &= \omega + \sum_{i=1}^{p} \beta_i \log \sigma_{t-i}^2 + \sum_{j=1}^{q} \alpha_j \left(|z_{t-j}| - \mathbb{E}\left(|z_{t-j}|\right)\right) + \sum_{j=1}^{q} \gamma_j z_{t-j},
\end{aligned}
$$

where $z_t := y_t \sigma_t^{-1}$ and to ensure stationarity, all roots of the GARCH coefficient polynomial, $(1 - \beta_1 L - \ldots - \beta_p L^p)$, must lie outside the unit circle. The h-step-ahead forecast of the EGARCH$(1,1)$ model is computed according to:

$$
\log \hat{\sigma}_{t+h|t}^2 = \hat{\omega} + \hat{\beta}_1 \log \hat{\sigma}_{t+h-1|t}^2 + \hat{\alpha}_1 \left(|\hat{z}_{t+h-1|t}| - \mathbb{E}\left(|\hat{z}_{t+h-1|t}|\right)\right) + \gamma_1 \hat{z}_{t+h-1|t}.
$$

GJR Model: The GJR-GARCH, or just GJR, model of **?** allows the conditional variance to respond differently to the past negative and positive innovations. The GJR(p,q) model has p GARCH coefficients associated with lagged variances, q ARCH coefficients associated with lagged squared innovations, and q leverage coefficients associated with the square of negative lagged innovations. The GJR(p,q) model may be expressed as:

$$y_t = \sigma_t z_t, \quad z_t \sim i.i.d \;\; N(0,1),$$
$$\log \sigma_t^2 = \omega + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \sum_{j=1}^{q} (\alpha_j + \gamma_j I_{[y_{t-j}<0]}) y_{t-j}^2,$$

where the indicator function $I[y_{t-j} < 0]$ equals 1 if $y_{t-j} < 0$, and 0 otherwise. Thus, the leverage coefficients are applied to negative innovations, giving negative changes additional weight. For stationarity and positivity, the GJR model has the following constraints:

- $\omega > 0$
- $\beta_i \geq 0$, $\alpha_j \geq 0$
- $\alpha_j + \gamma_j \geq 0$
- $\sum_{i=1}^{p} \beta_i + \sum_{j=1}^{q} (\alpha_j + \frac{1}{2}\gamma_j) < 1$

The GARCH model is nested in the GJR model. If all leverage coefficients are zero, then the GJR model reduces to the GARCH model. The recursive formula for the h-step-ahead forecast of the GJR-GARCH$(1,1)$ model is calculated as:

$$\hat{\sigma}_{t+h|t}^2 = \hat{\omega} + \left(\hat{\alpha}_1 + \frac{\hat{\gamma}_1}{2} + \hat{\beta}_1\right) \hat{\sigma}_{t+h-1|t}^2.$$

A.3 Realized volatility

Let $p_t = \log S_t$ denote the logarithmic of price where S_t is the observed price (at time t) and $r_t = p_t - p_{t-1}$ denote the continuously compounded return from time $t-1$ to t. Assume that the logarithmic price process, p_t, could belong to the class of continuous-time jump diffusion processes,

$$dp_t = \mu_t d_t + \sigma_t dW_t + \tilde{\kappa}_t dq_t, \quad 0 \leq t \leq T \tag{A.1}$$

where μ_t is a continuous and locally bounded variation process and σ_t is a stochastic volatility process; W_t is the standard Brownian motion; dq_t is a counting process such that $dq_t = 1$ represents a jump at time t (and $dq_t = 0$ if no jump) with jump intensity λ_t. If p_{t-} denotes the price immediately prior to the jump at time t, then $\tilde{\kappa}_t = \Delta p_t = p_t - p_{t-}$. The process p_t is consists of a continuous component and a pure jump component. The quadratic variation (QV) of this process is defined by

$$[r,r]_t = \int_0^t \sigma_s^2 dW_s + \sum_{0<s\leq t} \tilde{\kappa}_s^2, \tag{A.2}$$

where the first component, called integrated volatility, comes from the continuous component of (A.1) and the second term is the contribution from discrete jumps. In the absence of jumps, the second term on the right-hand side disappears and the quadratic variation is simply equal to the integrated volatility (IV).

Now, define the intraday return, r_{t_j} , as the difference between two logarithmic prices,

$$r_{t_j} = p_{t_j} - p_{t_{j-1}},$$

where t_j denotes the j-th intraday observation on the t-th day. Let Δ denote the discrete intraday sample period of length, $t_j - t_{j-1}$. The realized volatility (RV) is defined as the sum of squared intraday returns,

$$RV_t = \sum_{j=1}^{n} r_{t_j}^2,$$

where n is the number of Δ-returns during the t-th time horizon (such as a trading day) and is assumed to be an integer. **?** showed that RV is a natural estimator for the QV. Furthermore, The realized volatility satisfies

$$\lim_{\Delta \to 0} RV_t = \int_0^t \sigma_s^2 dW_s + \sum_{0<s\leq t} \tilde{\kappa}_s^2, \tag{A.3}$$

which means that RV_t is a consistent estimator of the QV.

A.4 Heterogenous Autoregressive model of Realized Volatility

Heterogenous Autoregressive model of Realized Volatility (HAR-RV) model proposed by **?**. In financial markets, either traders are perceived to be heterogeneous in the sense of a different horizon of investments [**?**] or information arrival is heterogeneous [**?**]. HAR-RV model takes into account the long memory feature, and among the models proposed to forecast volatility, it stands out in terms of performance and simplicity.

A generalized version of HAR-RV model that we used here is as follows:

$$\log RV_{t+1}^{(d)} = c + \beta^{(d)} \log RV_t^{(d)} + \beta^{(w)} \log RV_t^{(w)} + \beta^{(m)} \log RV_t^{(m)} + u_{t+1}^d \tag{A.4}$$

where

$$\log RV_t^{(w)} = \frac{1}{5} \sum_{j=0}^{4} \log RV_{t-j}^{(d)},$$

$$\log RV_t^{(m)} = \frac{1}{22} \sum_{j=0}^{21} \log RV_{t-j}^{(d)}.$$

This class of models can be estimated with ordinary least squares. For the details of forecasting in HAR-RV model, see **?**.