Forecasting Volatility with Time-Varying Coefficient Regressions

Qifeng Zhu,1 Miman You,2 and Shan Wu3

1School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu, China
2School of Mathematics and Statistics, North China University of Water Resources and Electric Power, Zhengzhou, China
3School of Finance, Nanjing University of Finance and Economics, Nanjing, China

Correspondence should be addressed to Qifeng Zhu; zhuqifengcn@eyou.com

Received 22 January 2020; Accepted 30 March 2020; Published 1 May 2020

Abstract

We extend the heterogeneous autoregressive- (HAR-) type models by explicitly considering the time variation of coefficients in a Bayesian framework and comprehensively comparing the performances of these time-varying coefficient models and constant coefficient models in forecasting the volatility of the Shanghai Stock Exchange Composite Index (SSEC). The empirical results suggest that time-varying coefficient models do generate more accurate out-of-sample forecasts than the corresponding constant coefficient models. By capturing and studying the time series of time-varying coefficients of the predictors, we find that the coefficients (predictive ability) of heterogeneous volatilities are negatively correlated and the leverage effect is not significant or inverse during certain periods. Portfolio exercises also demonstrate the superiority of time-varying coefficient models.

1. Introduction

Volatility is the key input variable for risk assessment, asset pricing, and portfolio allocation models. The early classic models are GARCH-type [1–3] and stochastic volatility (SV) [4] models, and because of the unavailability of high-frequency data, these models are based on daily or weekly returns. The omission of the informative intraday data makes these parameter volatility models not preferred. Volatility is an unobservable variable, and Andersen and Bollerslev [5] and Andersen et al. [6] suggested the sum of daily squared returns as the proxy of volatility and named it as realized volatility (RV). It makes volatility ex-post-observable, and the accuracy of RV is much higher than the proxies that are based on low-frequency data, such as square returns and intradaily range.

Corsi [7] built the heterogeneous autoregressive (HAR) model according to the heterogeneous market hypothesis. It is an AR(22) model that is restricted by economic interpretations. The model has 3 regressors, lagged daily, weekly, and monthly RV whose coefficients measure the impacts of short-term, medium-term, and long-term investors, and it can be easily estimated using the ordinary least squares (OLS) method. In spite of its simplicity, the HAR model not only successfully captures the main features of volatility, such as long memory, multiscale behavior, and fat tails, but also produces smaller forecast errors than the GARCH-type and SV models.

Some researchers extend the HAR model by adding additional components to the basic HAR-RV model [8–12], such as negative returns, jump variation, signed semi-variances, and overnight returns, and these extended models improve the performances in volatility forecasting on different aspects. The availability of various modeling approaches for volatility forecasting leads to model uncertainty for both researchers and practitioners, and model averaging approaches (i.e., trimmed) mean combination, discount mean square prediction error (DMSPE) combining method, triangular weighting (TW) method, and the Bayesian model averaging) are empirically proved to be more effective in volatility forecasting than a single model in spot, futures, energy, and commodity markets [13–16].

It is well documented that leverage effect, volatility persistence, mean reversion, structural break, etc. are the typical features of the volatility process of asset prices [17]. Structural breaks in volatility imply that the coefficients of volatility models vary over time. In general, not only structural breaks but also noisy proxies, nonlinearity, and
model specification errors may result in time-varying coefficients. Some researchers studied the effect of neglected coefficient changes on the persistence and level of volatility [18–20]. They all reflect that it is reasonable to admit that the coefficients are time varying. Granger [10] proves that any nonlinear functional form can be replaced by a model that is linear in variables, but which has time-varying coefficients. As we know, fewer research studies relate time-varying coefficient methodology to the HAR-type models.

Our study makes 3 contributions in the field of volatility prediction. First, previous studies always use rolling or recursive window regressions to implement the out-of-sample volatility forecast, implying the coefficients of models as constant. Differently, we assume that the predictive abilities of the predictors as time varying and build time-varying coefficient (TVC) HAR-type models. The predetermined variables represent a kind of model uncertainty. We address this issue with the Bayesian model averaging approach. Inspired by the work by Raftery et al. [21], we introduce the forgetting factor to the state-space model, and the forgetting factor not only makes the coefficients evolve more efficiently at a reasonable speed, by reducing the impact of the obsolete data, but also simplifies the calculation of posterior distributions. Although Markov regime switching (MRS) admits the coefficients are time-varying, this method is ad hoc, neither systematic nor helpful in understanding the real changes of these coefficients over time. According to the degree of time variation in coefficients, we compare the performances of 3 types of models, constant-coefficients (CC), MRS, and TVC HAR-type models in volatility forecasting.

Second, by investigating the coefficient series of the TVC-HAR-type models, we find that the coefficients (predictive abilities) of heterogeneous volatilities are negatively correlated and the leverage effect is not significant or inverse during specific periods. Choi et al. [22] point out that, due to the presence of structural breaks, the persistence of volatility may be overstated. After a huge shock, the persistence of volatility and the impact of historical volatility on future volatility will be significantly weakened. Therefore, the predictive ability of historical volatility for future volatility may change over time. Partial correlation coefficients of the models’ coefficient series show that the predictive abilities of lagged short-, medium-, and long-term volatility for future volatility are negatively correlated. Negative returns always have a greater impact on volatility than positive returns in the stock market, and it is sometimes ascribed to a leverage effect. The “leverage parameter” (i.e., the coefficient of negative returns) is always assumed to be constant. Corsi et al. [9] find that stronger leverage effects are time varying and empirically related to higher volatility regimes. Negative returns increase future volatility of the SSEC on the whole, but by investigating the time series of “leverage parameter,” we find that the leverage effect is insignificant or inverse in specific periods. Fewer studies examine the time-varying predictive abilities of the variables in volatility forecasting.

Our last contribution to the literature is that we not only statistically evaluate the performances of the corresponding models but also examine the economic significance of these models. In evaluation exercises, most previous studies simply use the measures suggested by statisticians, such as mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percent error (MAPE), and we use the loss functions suggested by Patton [23], which are robust to market microstructure noise. Practically, investors care about the usefulness of a volatility forecast. Following Ferreira [24], we examine whether the volatility forecasts can be used to improve a mean-variance investor’s portfolio. A mean-variance investor will decide at the end of period $t$ to allocate his or her assets between stock index and risk-free asset in period $t+1$. According to the investor’s utility function, the optimal weight of the stock index in this portfolio depends on the forecast of volatility. The results show that considering the time variation of coefficients makes the mean-variance investor have a greater excess return and utility.

Our data are 5-min intraday prices of the Shanghai Stock Exchange Composite Index (SSEC). HAR-RV and its 6 extensions are used to forecast 1-day-ahead volatility. Four robust loss functions, including mean square error (MSE) and quasi-likelihood (QL), which are robust to noisy proxy RV, are used to evaluate the performances of these volatility models. The average losses of the models considering the time variation of coefficients are always the least, indicating that they generate the most accurate out-of-sample volatility forecast among CC, MRS, and TVC models. The performances of MRS-type models are between the values of CC and TVC-type models.

There are 3 heterogeneous volatility components in the HAR model. Time-varying persistence of volatility results in the time-varying predictive ability of each component. By studying the series of time-varying coefficients, we find that the coefficients of the heterogeneous volatility components are negatively correlated, indicating their negatively correlated predictive abilities. We use our method to capture the time series of the “leverage parameter” which measures the correlation between price shocks and volatility and found that downside risk increases future volatility of the SSEC on the whole, but the leverage effect is not significant or inverse during certain periods. This finding is suggestive of the time-varying correlation between price and variance.

We use excess return and certainty equivalent return (CER) to measure the performances of these models in portfolio exercises. We find that the investors allocating their assets according to the forecast from TVC-type models always get more excess return and CER. In the Chinese stock market, which is a typical example of emerging markets, the investors with lower risk aversion coefficient have more excess return. Overall, the model that includes all the predictors and considers the time variation of coefficients outperforms the other models not only statistically but also economically.

The remainder of the paper is structured as follows. Section 2 provides the backgrounds of RV and HAR-type models and introduces our time-varying coefficients model. Section 3 is the empirical analysis, and we compare the performances of the time-varying coefficients HAR-type models with constant coefficients and Markov regime switching HAR-type models through statistical and economic measures and report our empirical findings. Section 4 concludes the study.
2. Methodology

2.1. Realized Volatility and the HAR Model. In an arbitrage-free market, consider an asset whose price process $p_t$ follows the continuous-time jump diffusion process and is given by the following stochastic differential equation [25]:

$$d \log(p_t) = \mu_t dt + \sigma_t dW_t + J_t dq_t, \quad 0 \leq t \leq T,$$

where $\mu_t$ is a continuous and locally bounded variation process, $W_t$ denotes a standard Brownian motion, $\sigma_t$ denotes the instantaneous volatility which is a stochastic process independent of $W_t$, $dq_t$ is a counting process whose intensity $\lambda t$ is time varying, $J_t$ denotes the size of a jump in the price process, $dq_t = 1$ refers to a jump at time $t$, and $dq_t = 0$ refers to no jump. The increment of the RV from $t-1$ to $t$ is defined as

$$QV_t = \int_{t-1}^{t} \sigma^2_d ds + \sum_{t-1 < t_i \leq t} J_i^2,$$

where the first term $\int_{t-1}^{t} \sigma^2_d ds$ is the integrated variance (IV) and the second term $\sum_{t-1 < t_i \leq t} J_i^2$ is called the jump variation. To calculate RV, which is defined by Andersen et al. [6], without loss of generality, we normalize the daily time interval to 1 and then divide it into $M$ periods. The length of each period is $\Delta = 1/M$. The $j$-th log return on the $t$-th day is $r_{t-j} = \log(p(t-1+j\Delta)/p(t-1+(j-1)\Delta))$, $j = 1, \ldots, M$. RV is the summation of the corresponding $I/\Delta$ high-frequency intraday squared returns:

$$RV_t = \sum_{i=1}^{M} r_{t-1,j}^2.$$

As emphasized by quadratic variation theory [5, 6], the realized variation converges in probability to the quadratic variation process as $M \rightarrow \infty$:

$$RV_t \rightarrow QV_t = \int_{t-1}^{t} \sigma^2_d ds + \sum_{t-1 < t_i \leq t} J_i^2.$$

For the HAR-C VAR model, Dacorogna et al. [26] argued that long and short-time horizon volatility propagates asymmetrically. Motivated by the HAR-C VAR model and RV, Corsi [7] build the HAR-RV model which is almost the standard model for volatility modeling and forecasting. There are 3 regressors in the HAR-RV model, lagged one-day volatility $RV_t$, lagged one-week volatility $RV_{t-7}$, and lagged one-month volatility $RV_{t-21}$, and here, $RV_{t-k} = (1/h) \sum_{j=0}^{h} RV_{t-j}$.

The HAR-RV model proposed by Corsi [7], which is a special case of the autoregressive (AR) model by imposing economically meaningful restrictions, is expressed as

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{t-4} + \beta_3 RV_{t-21} + \epsilon_{t+1}. \quad (5)$$

2.2. Extensions of the HAR Model. The simplicity of the model makes it easily extendable by adding additional predictors. Our choice of additional predictive variables is guided by previous academic studies. There are so many combinations of these predictors, so we build 2 types of models. The 1st type of the models only includes a single class of additional variables, and the 2nd type is a model which includes all the additional variables.

Asymmetry is a well-documented stylized fact about stock volatility, and it means that positive and negative shocks have different impacts on the volatility. Black [27] named it as a leverage effect. Just like the EGARCH and GJR-GARCH models, we add the negative returns as an additional predictor to the basic HAR model and build the LHAR-RV model, where ‘L’ stands for leverage.

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{t-4} + \beta_3 RV_{t-21} + \beta_4 I_{t} + \epsilon_{t+1}. \quad (6)$$

Andersen et al. [8] add jump variation to the basic HAR model and build the HAR-RV-J model. The empirical analyses of the equity index returns suggest that the volatility jump variation is highly important but less persistent than the integrated variation, extracting jump variation results in significant out-of-sample forecast improvements. The HAR-RV-J is expressed as

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{t-4} + \beta_3 RV_{t-21} + \beta_4 J_t + \epsilon_{t+1}, \quad (7)$$

where $J_t$ is an estimator of the jump variation. Andersen et al. [8] proved that, as the sampling frequency of the underlying returns increases, $J_t = \max (RV_t - BPV_t, 0) \rightarrow \sum_{t-1 < t_j \leq t} J_j^2$, where BPV is the bipower variation, $BPV_t = (\sqrt{2/\pi}) \sum_{j=1}^{\infty} |r_{t-j} + r_{t+j}|/ \sqrt{2t-2(j-1)^2}|$. The information flow about stocks is continuous, but major stock trading is limited in just a few hours a day. A lot of financial and economic events take place during close-to-open periods, and overnight information has an important influence on future volatility of stock markets. Emerging markets like China and Brazil are inevitably affected by mature markets such as the United States and Britain. Trading hours for the US stock market is the time when the Chinese stock market is closed. Overnight returns have been used as a proxy of overnight information flow to enhance forecast accuracy of volatility modes in recent literature [12, 28, 29]. Tseng et al. [28] argued that the impact of overnight returns on future volatility is also asymmetric. So, negative overnight returns are added to the basic HAR model, termed as LHAR-RV-OR.

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{t-4} + \beta_3 RV_{t-21} + \beta_4 OR_{t} + \epsilon_{t+1}, \quad (8)$$

RV which is an even function of high-frequency intraday returns (sum of squared returns) neglects the fact that the impacts of positive and negative returns on future volatilities are different. Barndorff-Nielson et al. [30] developed a new measure, realized semivariance, which relates to positive and negative intraday returns, and named them as “good volatility” and “bad volatility,” respectively. Their empirical study shows that the findings produce significantly better performance in out-of-sample volatility forecasting. Suppose the price process $p_t$ follows (1), they decompose $RV_t$ into signed realized semivariances $R^+_s$ and $R^-_s$, where
and $I(\cdot)$ is the indicator function. According to equations (3), (9), and (10), $RV_t = RS_t^+ + RS_t^-$. They proved that as the sampling frequency $M \rightarrow \infty$,

$$RS_t^+ \rightarrow \frac{1}{2} \sum_{t-1}^{t} \sigma_s^2 ds + \sum_{t-1 < s \leq t} f_s^2 I(J_s > 0),$$

(11)

$$RS_t^- \rightarrow \frac{1}{2} \sum_{t-1}^{t} \sigma_s^2 ds + \sum_{t-1 < s \leq t} f_s^2 I(J_s < 0).$$

(12)

We replace the predictor $RV_t$ in (5) as $RS_t^+$ and $RS_t^-$ to test whether the positive and negative semivariances have asymmetric influences on future volatility and build the AHAR-RV model, where “A” stands for asymmetric:

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{t-4} + \beta_3 RV_{t-21} + \beta_4 V_t + \epsilon_{t+1}. $$

(13)

The model (13) encompasses the basic HAR model [7] by setting $\beta_1 = \beta_2$. More importantly, they defined signed jump variation which measures the difference between positive and negative jump variations. It is positive/negative when the price is dominated by upward/downward jumps:

$$\Delta J_t = RS_t^+ - RS_t^- \rightarrow \sum_{t-1 < s \leq t} f_s^2 I(J_s > 0) - \sum_{t-1 < s \leq t} f_s^2 I(J_s < 0).$$

(14)

The simplicity of signed jump variation is that it does not need to know or estimate the corresponding jump variation which may be noisy. To investigate the predictive ability of signed jump variation, Barndoff-Neilsen et al. [30] extend the HAR model by adding signed jump variation to the model and name it as HAR-RV-SJ, here ‘SJ’ stands for signed jump:

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{t-4} + \beta_3 RV_{t-21} + \beta_4 V_t + \epsilon_{t+1}. $$

(15)

Because $RV_t$ and $\Delta J_t$ can be linearly represented by $RS_t^+$ and $RS_t^-$, equation (15) and equation (13) are equivalent, and the empirical study also verifies that. We do no exhibit the results corresponding to the model HAR-RV-SJ, equation (15).

The sequential information arrival hypothesis [31] and the noise trading hypothesis [32] both imply that a causal relation between return volatility and trading volume exists. Lamoureux and Lastrapes [33] find volume, a proxy of information flow, to be strongly significant when it is inserted into the ARCH variance process. Le and Zurbruegg [34] propose to introduce trading volume into ARCH-type models to improve volatility forecasting, and the empirical research results are robust to different measures of volatility and trading volume. Wang and Huang [35] find that daily integrated variance (IV) is positively related to trading volume, but jump variation reveals a negative relationship with trading volume and contains some “public information.” These researches all proved that trading volume can be exploited for forecasting volatility purposes. According to Lamoureux and Lastrapes’ approach [33], we use the turnover ratio as the proxy of trading volume which represents the arrival of information and build the HAR-RV-V model, where “V” stands for volume:

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{t-4} + \beta_3 RV_{t-21} + \beta_4 V_t + \epsilon_{t+1}. $$

(16)

According to $RV_t = RS_t^+ + RS_t^-$ and equation (14), if predictors $RV_t, RS_t^+, RS_t^-$, and $\Delta J_t$ are included in a model, it will result in perfect collinearity. It is verified in our empirical analysis; these coefficients cannot be estimated by OLS and out time-variation co-efficients method. Because $RS_t^+$ and $RS_t^-$ contain the information of $RV_t$ and $\Delta J_t$, to avoid perfect collinearity, we build the model HAR-RV-ALL that contains all the predictors mentioned above except $RV_t$ and $\Delta J_t$:

$$RV_{t+1} = \beta_0 + \beta_1 RV_{t-4} + \beta_2 RV_{t-21} + \beta_3 RV_{t-21} + \beta_4 V_t + \epsilon_{t+1}. $$

(17)

2.3. Time-Varying Coefficient Regression. An important purpose of our study is to evaluate whether the time-varying coefficient paradigm can be used to improve volatility forecasting. Suppose $k$ is the number of predictive variables that a constant-coefficient model contains (including the intercept term); to reduce the risk of model selection, we build and estimate $2^k - 1$ time-varying coefficient submodels that are associated with all the combinations of the predictive variables, and the model that does not contain any predictor is not considered. Then, we use the Bayesian criterion to calculate the posterior probability weight of each submodel (similar to the study by Avramov [36]), and the final predictive mean of volatility is the weighted average of predictive means of all submodels. We first focus on describing the characteristics of a time-varying coefficient submodel that contains a given selection of predictive variables, and then we introduce the Bayesian model selection criterion.

It is natural to relate the time-varying coefficients to the state-space model, so we build the time-varying coefficient submodel as

$$RV_{t+1} = F_t^T \theta_t + \epsilon_t, \quad \epsilon_t \sim N(0, V),$$

(18)

$$\theta_t = \theta_{t-1} + \eta_t, \quad \eta_t \sim N(0, W_t),$$

(19)

where $F_t$ is the $k \times 1$ vector of all the predictors. Suppose $\theta_t = (\theta_{t,1}, \theta_{t,2}, \ldots, \theta_{t,k})^T$ is the $k \times 1$ vector of time-varying
coefficients which determine the impact of $F_t$ on $RV_{t+1}$ at time $t$, then $\theta_t$ evolves according to equation (19). For example, for equation (6), there are 5 predictors in the model, and we build $2^5-1 = 31$ time-varying coefficient submodels. $F_t$ can be any combinations of the predictive variables, such as $(1, RV_t, RV_{t-4}, RV_{t-21}, I(r_t < 0))^T$ and $(1, RV_t)^T$, where 1 denotes the constant term and $\theta_t = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_{21}, \hat{\beta}_{32}, \hat{\beta}_{t}, \hat{\beta}_{t-4}, \hat{\beta}_{t-21})^T$ or $\theta_t = (\hat{\beta}_0, \hat{\beta}_1)^T$. In observational equation (18), the innovations $\epsilon_t$ are distributed as $\epsilon_t \sim N(0, V_t)$. $V_t$ denotes the unknown conditional variance; similarly, $W_t$ denotes the time-varying conditional variance matrix in state equation (19), and the specification is similar to the Bayesian framework for stochastic volatility suggested by Johannes, Korteweg, and Polson [37]. Equation (19) suggests that $\theta_t$ follows a random walk process that can capture various variations of coefficients.

Specifying a normal distribution as prior for $\theta_0$ and an inverse-gamma distribution for the observational variance $V_0$, results in a conjugate Bayesian analysis. It means that the prior distribution of $\theta$ is $\theta \sim N(0, V_0)$, where $V_0$ is from the inverse-gamma distribution $IG(\kappa, \nu_0)$, specifically, it is assumed that $\epsilon_t \sim N(0, V_t)$, and the existence of so many predictive variables results in a huge number of time-varying coefficient submodels, and it the less the forgetting factor $\lambda$ is, the relatively larger shock hits to the coefficients, the higher the evolution speed of these coefficients is.

In empirical studies, due to our ex-ante selection of the set of predictive variables, there is uncertainty about whether the variables have significant predictive abilities in volatility forecasting during specific periods. The predetermined variables represent a kind of model uncertainty. We address this issue with the Bayesian model averaging approach which has been applied to volatility prediction and performs well [16, 42]. For example, Lyócsa et al. find that combination forecasts, especially with univariate specifications or Bayesian model averaging, conclusively outperform the benchmark in forecasting the volatility of nonferrous metal futures [16]. We can build $2^5-1$ submodels with predetermined $k$ predictors, and their posterior probabilities are updated day by day according to the Bayesian approach. The final predictive density of $RV_{t+1}$ is the weighted average of the predictive densities of all time-varying coefficient submodels, and the probability of each submodel conditional on the current information $\phi_t$ depends on the corresponding predictive likelihood. The details about the Bayesian model averaging approach are given in the Appendix.

By applying the time-varying coefficient regression method to the previous HAR-type models, we get the time-varying coefficient HAR-type models. For example, we improve the basic HAR model in equation (5) as the time-varying coefficient HAR (TVC-HAR) model and improve the HAR-RV-J model in equation (7) as the TVC-HAR-RV-J model; similarly, we improve all the other models.

3. Empirical Study

3.1. Data Description. Many studies focus on the developed stock markets, such as the US, UK, and Japan. Differently, we intend to study forecasting volatility of the Chinese stock market which is a typical example of the emerging markets. Despite its higher-than-average return, empirical studies show that the Chinese stock market is counter-cyclical and more volatile than the mature markets. On January 1, 2018, MSCI included China A-shares in the Emerging Markets (EM) Index and the ACWI (All Country World Index) Index. The weight of Chinese A-shares in the EM Index is 33.00% from the data as of March 29, 2019. Exploring and forecasting volatility of the Chinese stock market are very meaningful for risk management, asset pricing, and global asset allocation.

Our data are 5-minute intraday high-frequency data of the SSEC published by Wind Information Co., Ltd, and there are 48 observations on each trading day. Our data cover from November 8, 1999, to April 23, 2018, and 4465 trading days in total, approximately 18 years. The SSEC index consists of all listed stocks (including A shares and B shares) on the Shanghai Stock Exchange (SSE), and it reflects the overall performance of the Chinese stock market. Figure 1 presents the time series plot of these variables for...
the SSEC. According to panel A of Figure 1, the RV of the SSEC is clustering, and it was soaring during the 2008 subprime crisis and 2015-16 the Chinese stock market turbulence, accompanied by fast-rising risk. The absolute values of these variables become greater during financial crises, and all the variables are correlated but clearly contain different predictive information. It is hard to directly distinguish their predictive abilities.

Table 1 presents some statistics for our data. \( J_t, RS_t^+, \) and \( RS_t^- \) are all smaller than that of RV. It makes sense, as \( J_t, RS_t^+, \) and \( RS_t^- \) are the components extracted from RV. The augmented Dickey–Fuller (ADF) tests (without constant and trend) show that all the variables reject the null hypothesis of unit root at a 1% significance level, implying that all these variables are stationary and will not result in spurious regressions. The Ljung–Box Q statistic for serial correlation up to 22 lags shows that RV of the SSEC displays a significant level of autocorrelation, implying RV of the SSEC exhibits persistence.

### 3.2. In-Sample Estimation Results

As mentioned above, we use the constant and time-varying coefficient models to forecast the RV of the SSEC. We first estimate the coefficients of the models listed above, expressed in equations (5)–(8), (13), (16), and (17), and then evaluate whether the variables are significant in forecasting future volatility. The variables in the models are autocorrelated and heteroscedastic, so the Newey–West covariance correction is employed. Table 2 exhibits the in-sample estimation results for the models.

From Table 2, we find that adjusted \( R^2 \) of the basic HAR model is the lowest, and for the models with only one additional predictor, the additional predictors are all significant for future volatility except positive semivariances, but for the last model that includes all the predictors, the predictors are significant except jump variation. It illustrates that the information content of jump variation is encompassed by other variables. It is consistent with the previous findings by Andersen et al. [8]. By comparing the adjusted \( R^2 \) of model (6) and model (13), we find that the explanatory ability of negative semivariances is not superior to that of negative returns for the Chinese stock market. This is different to the American market [30]. There are 2 reasons. First, negative semivariances do not contain overnight information, and the Chinese stock market is significantly influenced by the overnight information, especially the information from the American market. Second, the Chinese stock market is dominated by individual investors that rarely get the high-frequency intraday trading information, and low-frequency daily returns are always used to make investment decisions. The model that contains all of the predictors, expressed by equation (17), has the greatest coefficient of determination and greatly improves the predictive ability of the basic HAR model. The coefficient of determination raises from 0.4915 to 0.5755 which is significantly greater than that of the models with only one additional predictor.

### 3.3. Out-of-Sample Forecast

To compare the out-of-sample performances of constant coefficient, MRS, and time-varying coefficient HAR-type models, we consider the loss functions which are robust to market microstructure noise [23]. They are defined as
Table 1: Summary statistics of the variables.

|       | RV | $|r| I (r < 0)$ | $I$ | $|OR| I (OR < 0)$ | $RS_1^I$ | $RS_2^I$ | $V_1$ |
|-------|----|---------------|----|----------------|-----------|----------|-------|
| Mean  | 1.721e-04 | 5.440e-03 | 1.920e-05 | 2.678e-03 | 8.803e-05 | 8.414e-05 | 2.130 (%) |
| Std. dev. | 2.760e-04 | 1.046e-02 | 6.013e-05 | 5.035e-03 | 1.508e-04 | 1.445e-04 | 1.499 |
| Min   | 3.675e-06 | 0.000 | 0.000 | 1.288e-06 | 4.877e-10 | 3.041e-01 | 0.05 |
| Max   | 3.905e-03 | 9.256e-02 | 2.019e-03 | 7.105e-02 | 2.523e-03 | 1.914e-03 | 1.039e01 (|
| Ljung-Box lag(22) | 4.029e05 | 4.269e04 | 2.679e07 | 1.438e05 | 7.460e05 | 3.211e05 | 4.565e03 |

**Indicates that the null hypothesis is rejected at the 1% level.

Table 2: Estimation results for the models with constant coefficients.

<table>
<thead>
<tr>
<th></th>
<th>HAR-RV</th>
<th>LHAR-RV</th>
<th>HAR-RV-J</th>
<th>LHAR-RV-OR</th>
<th>Aثار-RV</th>
<th>HAR-RV-V</th>
<th>HAR-RV-ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>1.641e-05***</td>
<td>-2.388e-06***</td>
<td>1.606e-05***</td>
<td>5.105e-06</td>
<td>1.631e-05***</td>
<td>-1.093e-05*</td>
<td>-4.851e-05***</td>
</tr>
<tr>
<td>(3.651e-06)</td>
<td>(6.002e-06)</td>
<td>(3.943e-06)</td>
<td>(4.741e-06)</td>
<td>(3.804e-06)</td>
<td>(5.804e-06)</td>
<td>(8.545e-06)</td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>3.848e-01***</td>
<td>2.555e-01***</td>
<td>3.959e-01***</td>
<td>3.332e-01***</td>
<td>3.140e-01***</td>
<td>4.765e-02</td>
<td>2.759e-01***</td>
</tr>
<tr>
<td>(4.743e-02)</td>
<td>(4.228e-02)</td>
<td>(5.632e-02)</td>
<td>(4.527e-02)</td>
<td>(4.527e-02)</td>
<td>(4.765e-02)</td>
<td>(8.545e-06)</td>
<td></td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>3.132e-01***</td>
<td>3.328e-01***</td>
<td>3.027e-01***</td>
<td>2.715e-01***</td>
<td>3.249e-01***</td>
<td>3.033e-01***</td>
<td>2.759e-01***</td>
</tr>
<tr>
<td>(8.183e-02)</td>
<td>(7.666e-02)</td>
<td>(7.509e-02)</td>
<td>(7.775e-02)</td>
<td>(7.663e-02)</td>
<td>(8.363e-02)</td>
<td>(7.222e-02)</td>
<td></td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>2.475e-01***</td>
<td>2.194e-01***</td>
<td>2.481e-01***</td>
<td>2.292e-01***</td>
<td>2.413e-01***</td>
<td>1.973e-01***</td>
<td>1.373e-01***</td>
</tr>
<tr>
<td>(5.007e-02)</td>
<td>(4.861e-02)</td>
<td>(5.272e-02)</td>
<td>(5.289e-02)</td>
<td>(4.927e-02)</td>
<td>(5.461e-02)</td>
<td>(4.606e-02)</td>
<td></td>
</tr>
<tr>
<td>$\beta_{1(r,c&lt;0)</td>
<td>r,c}$</td>
<td>6.336e-03***</td>
<td>(1.272e-03)</td>
<td>4.205e-01***</td>
<td>(1.395e-01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>8.651e-01***</td>
<td>(1.565e-03)</td>
<td>1.010e-01***</td>
<td>(5.273e-03)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{1cR}$</td>
<td>-1.105e-01</td>
<td>(7.774e-02)</td>
<td>1.895e-01**</td>
<td>(9.104e-02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1^R$</td>
<td>7.936e-01***</td>
<td>(1.293e-01)</td>
<td>2.951e-01**</td>
<td>(1.161e-01)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1^R$</td>
<td>1.961e-05***</td>
<td>(3.835e-06)</td>
<td>2.485e-05***</td>
<td>(4.546e-06)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.4915</td>
<td>0.5434</td>
<td>0.4970</td>
<td>0.5143</td>
<td>0.5213</td>
<td>0.4994</td>
<td>0.5755</td>
</tr>
</tbody>
</table>

The terms in parenthesis represent the Newey–West robust standard error. ***, **, and * indicate that the null hypothesis is rejected at the 1%, 5%, and 10% level, respectively.

The loss function values of the TVC-type models are less than those of the corresponding MRS-type and CC-type models, implying that the TVC models provide more accurate out-of-sample volatility forecast. The inclusion of all the predictors makes the HAR-RV-ALL-type models outperform the other models except when $b = 1$, and the loss of HAR-RV-ALL-type models is greater than that of the LHR-RV-type models. Due to the unknown time and intensity of a

$$L(\hat{\sigma}^2, h; b) = \begin{cases} \frac{1}{(b + 1)(b + 2)} \left( (\hat{\sigma}^{2b+4} - h^{b+2}) \right), & b \neq -1, -2 \\ -\frac{1}{b + 1} h^{b+1} \left( \hat{\sigma}^2 - h \right), & b = -1, -2 \\ h - \hat{\sigma}^2 + \hat{\sigma}^2 \log \frac{2}{h}, & b = 1 \\ \frac{\hat{\sigma}^2}{h} - \log \frac{\hat{\sigma}^2}{h} - 1, & b = 2 \end{cases}$$

(20)
3.4.1. Time-Varying Leverage Effect.

In our method, the coefficients of leverage parameter are related to the choice of loss function, but for the models that include the same predictors, TVC-type models do generate the most accurate forecasts. These results verify the importance of explicitly considering the time variation of coefficients when modeling and forecasting volatility.

3.4. Economic Interpretation. It is well documented that the volatility of equity returns exhibits asymmetry and persistence. It is very convenient to study the time variation of leverage effect and the strength of persistence in the time-varying coefficient regression framework. In this section, we first study the time-varying leverage effect and its relationship to the stock market cycle by capturing the time variation of the "leverage parameter." Further, we study the correlation of predictive abilities of the heterogeneous structure volatility. It is related to the persistence of volatility. Finally, we examine the economic value of the volatility forecast by using a volatility timing strategy.

3.4.1. Time-Varying Leverage Effect. For equity returns, negative shocks have a greater impact on future volatility than positive shocks. The stylized fact is named as "leverage effect" by Black [27]. Many volatility models specify that volatility is affected asymmetrically by positive and negative shocks, i.e., Glosten et al. [46], Engle and Ng [47], Harvey and Shephard [48], Bakshi et al. [49], and Wu and Hou [50]. Bandi and Renô [51] point out that the correlation between shocks to prices should not be assumed constant and provide a nonparametric estimation to evaluate the variation of the leverage effect. They find that stronger leverage effects are empirically related to higher volatility regimes. Patton and Sheppard [1] use signed semivariances, new estimators proposed by Barndorff-Nielsen et al. [42], which are calculated by signed high-frequency returns, and find that negative realized semivariance has a more significant impact on future volatility than positive realized semivariance and extracting the positive and negative realized semivariances from RV significantly improves the performances of the volatility forecast. Negative returns and negative semivariances measure the downside risk of assets; similarly, the positive returns and positive semivariances measure the upside risk of assets, but the latter variables are not significant for forecasts of volatility. According to the view of Bandi and Renô [51], leverage is defined as a function of spot volatility level, and the impact of downside risk on future volatility is time varying. In our method, the coefficients of these predictors are allowed to vary over time, and the posterior point estimator of the coefficient $\theta_t$ is

$$E(\theta_t | F_t) = \sum_{i=1}^{2^{k-1}} E(\theta_{i,t} | F_t) p(M_i | F_t),$$

where $E(\theta_{i,t} | F_t)$ is the conditional expectation of the coefficient $\theta_t$ from the submodel $M_i$.

So, it is very convenient to capture the variation of the "leverage parameter" (i.e., the time-varying coefficient of negative returns, $\beta_{1t(r<0)r|\tau}$ in equation (6)). Figure 2 plots the time series of the "leverage parameter". Panel A is the plot of $\beta_{1t(r<0)r|\tau}$, and Panel B is the plot of $\tilde{\beta}_{1t}$; the coefficients of negative realized semivariances RV$^{-}$ are given in equation (13). For a better understanding of the variation, the historical prices of SSEC are also exhibited.

From Figure 2, we observe that, in most cases, the coefficient of negative returns and negative realized semivariances, $\beta_{1t(r<0)r|\tau}$ and $\tilde{\beta}_{1t}$ are greater than zero, indicating that downside risk, which is measured by negative returns and negative realized semivariances, increases future volatility of the SSEC. However, there are some special cases. For Panel A, $\beta_{1t(r<0)r|\tau}$ is less than 0 in two time periods, from 2006-11-04 to 2007-01-25 and from 2014-12-02 to 2015-01-22; for Panel B, $\tilde{\beta}_{1t}$ is less than 0 in the period, from 2014-12-05 to 2015-01-27, indicating that the leverage effect is inverted, and downside risk reduces future volatility of the SSEC.

The information contained in negative returns and negative semivariances is different. Although negative returns do not contain the information of high-frequency intraday returns, they contain the information of overnight returns. The mean of semivariances is 59.74% of the mean.

### Table 3: Out-of-sample forecast.

<table>
<thead>
<tr>
<th></th>
<th>HAR-RV</th>
<th>LHAR-RV</th>
<th>HAR-RV-J</th>
<th>LHRHAROV</th>
<th>AHRARV</th>
<th>HAR-RV-V</th>
<th>HAR-RV-ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b = -2(QL)$</td>
<td>CC</td>
<td>0.203</td>
<td>0.197</td>
<td>0.202</td>
<td>0.181</td>
<td>0.209</td>
<td>0.181</td>
</tr>
<tr>
<td></td>
<td>MRS</td>
<td>0.201</td>
<td>0.196</td>
<td>0.202</td>
<td>0.184</td>
<td>0.223</td>
<td>0.179</td>
</tr>
<tr>
<td></td>
<td>TVC</td>
<td>0.198</td>
<td>0.193</td>
<td>0.199</td>
<td>0.178</td>
<td>0.195</td>
<td>0.178</td>
</tr>
<tr>
<td>$b = -1$</td>
<td>CC</td>
<td>4.608e-05</td>
<td>3.963e-05</td>
<td>4.516e-05</td>
<td>4.230e-05</td>
<td>4.484e-05</td>
<td>4.230e-05</td>
</tr>
<tr>
<td></td>
<td>TVC</td>
<td>4.455e-05</td>
<td>3.823e-05</td>
<td>4.461e-05</td>
<td>4.081e-05</td>
<td>4.316e-05</td>
<td>4.079e-05</td>
</tr>
<tr>
<td>$b = 0(MSE)$</td>
<td>CC</td>
<td>4.994e-08</td>
<td>3.762e-08</td>
<td>4.884e-08</td>
<td>4.873e-08</td>
<td>4.831e-08</td>
<td>4.873e-08</td>
</tr>
<tr>
<td></td>
<td>MRS</td>
<td>4.953e-08</td>
<td>3.714e-08</td>
<td>4.840e-08</td>
<td>4.843e-08</td>
<td>4.830e-08</td>
<td>4.849e-08</td>
</tr>
<tr>
<td></td>
<td>TVC</td>
<td>4.915e-08</td>
<td>3.702e-08</td>
<td>4.834e-08</td>
<td>4.783e-08</td>
<td>4.760e-08</td>
<td>4.793e-08</td>
</tr>
<tr>
<td>$b = 1$</td>
<td>CC</td>
<td>2.530e-11</td>
<td>1.81e-11</td>
<td>2.459e-11</td>
<td>2.687e-11</td>
<td>2.460e-11</td>
<td>2.688e-11</td>
</tr>
<tr>
<td></td>
<td>MRS</td>
<td>2.496e-11</td>
<td>1.771e-11</td>
<td>2.411e-11</td>
<td>2.589e-11</td>
<td>2.455e-11</td>
<td>2.702e-11</td>
</tr>
<tr>
<td></td>
<td>TVC</td>
<td>2.435e-11</td>
<td>1.721e-11</td>
<td>2.281e-11</td>
<td>2.676e-11</td>
<td>2.332e-11</td>
<td>2.647e-11</td>
</tr>
</tbody>
</table>
square negative returns. Overall, negative returns contain more information than negative semivariances, and it is consistent with the previous results (Table 2) that the $R^2$ coefficient of equation (6) is greater than that of equation (13). So, it is not difficult to understand the phenomenon that the periods, when the coefficients of negative semivariances and returns are less than 0, are related but different.

We also compare the results from time-varying coefficient models with the OLS results. For the whole sample, from Table 2, we know that the OLS estimation of the coefficient $\beta_{I(r<0)}$ in equation (6) is $6.336e-03$ (Newey–West adjusted std. dev.: $1.272e-03$, $p$ value: $1.076e-06$), indicating that the SSEC exhibits significantly ordinary leverage effect on the whole. For equation (6), the OLS estimations of the coefficient of $\beta_{I(r<0)}$ are $-4.888e-03$ (Newey–West adjusted std. dev.: $1.677e-03$, $p$ value: $0.00526$) and $0.3400$ (Newey–West adjusted std. dev.: $0.7608e-03$, $p$ value: $0.6581$) during the 1st and 2nd special periods. The results are evidently different to the whole sample and challenge the classic stylized fact of volatility.

3.4.2. Correlation of the Heterogeneous Coefficients. In the basic HAR model, the heterogeneous structure volatility, $RV_p$, $RV_{t-4}$, and $RV_{t-21}$, correspond to the market expectations of the next period’s volatility which are based on the observations of past realized volatilities and their marginal contributions to future volatility are measured by their coefficients. Frijns et al. [53] point out that the structure of volatility may change due to the time-varying behavior of heterogeneous investors. Since the coefficients of heterogeneous predictors are time varying, by studying the correlations of the heterogeneous coefficients, we can find the mutual influence of the predictive ability of these variables. In multivariate correlation analysis, because of the interaction between the variables, the simple correlation coefficient cannot reflect the real correlation between the two predictors. Therefore, we use the partial correlation coefficient to measure the correlation between the coefficients of
3.4.3. Portfolio Exercises. In evaluation exercises, it is usual to use the statistical loss functions mentioned above; but practically, to check the economic value of a volatility forecast for investors, we should use the loss functions which are related to the utility functions of investors. Following Ferreira [24] and Santa-Clara and Neely et al. [55], among others, we consider a mean-variance investor who will decide at the end of period $t$ to allocate his or her assets between stock index and risk-free asset in period $t+1$. The portfolio return is given by

$$R_{p,t+1} = w_{t+1} r_{t+1} + R_{f,t+1},$$

(22)

where $w_{t+1}$ is the weight of the stock index in this portfolio, so $1 - w_{t+1}$ is the weight of risk-free bills; $r_{t+1}$ is the return of the stock index over risk-free rate; $R_{f,t+1}$ is the risk-free rate. We use certainty equivalent return (CER), which incorporates individual investor risk preferences into financial decisions, as the investor’s utility function to evaluate the performance of this portfolio. CER can be interpreted as the lowest risk-free rate that an investor would accept rather than holding a given portfolio. The mean-variance investor’s expected utility function, CER, is expressed as

$$U(R_{p,t}) = CER(R_{p,t}) = E(R_{p,t}) - \frac{1}{2} \gamma \text{Var}(R_{p,t})$$

$$= E(w_{t+1} r_{t+1} + R_{f,t+1}) - \frac{1}{2} \gamma \text{Var}(w_{t+1} r_{t+1} + R_{f,t+1}),$$

(23)

where $E(R_{p,t})$ and $\text{Var}(R_{p,t})$, respectively, denote the expected return and the variance of the portfolio and $\gamma$ is the relative risk aversion coefficient.

At the end of day $t$, by maximizing the utility function $U(R_{p,t})$, the investor optimally allocates the weight of stock index in day $t+1$:

$$w_t = \frac{1}{\gamma} \left( \frac{\bar{r}_{t+1}}{\bar{\sigma}_{t+1}} \right) = \frac{1}{\gamma} \left( \frac{\bar{r}_{t+1}}{\bar{\sigma}_{t+1}} \right),$$

(24)

where $\bar{r}_{t+1}$ and $\bar{\sigma}_{t+1}$, respectively, denote the mean and volatility forecast of the index excess returns. Following Santa-Clara and Neely et al. [55], the historical average of excess returns $r_{t+1}$ is the benchmark which is hard to exceed, and we use $\bar{r}_{t+1}$ as the mean forecast, $\bar{r}_{t+1} = \bar{r}_{t+1}$. The weight $w_t$ is constrained between 0 and 1.5 (inclusive). It is a financial constraint that precludes short sales and more than 50% leverage. From equation (24), the optimal weight of stock index $w_t$ is inversely proportional to risk aversion coefficient $\gamma$ and the greater risk aversion coefficient is, the lower weight of stock index that the investor would allocate in his or her portfolio.

From equations (23) and (24), we know that CER is determined by 3 parameters: risk-free rate $R_{f,t+1}$, risk aversion coefficient $\gamma$, and the forecast of volatility $\bar{\sigma}_{t+1}$. We use 3-month Shanghai Interbank Offered Rate (SHIBOR) as the risk-free rate $R_{f,t+1}$ and our data are from October 9, 2006, when the Chinese central bank put forward SHIBOR, to April 23, 2018. For robustness purposes, we set $\gamma$ as 3 and 6. $R_{f,t+1}$ and $\gamma$ are both known variables and volatility is the only input variable for this portfolio, so the performance of this portfolio is determined by the effectiveness of the volatility forecast.

Table 4 shows the portfolio performance measures (annualized) from 2006-10-9 to 2018-4-23, including the mean of excess returns ($R$) and CER. Under the premise of fixed risk aversion coefficient, all the average annualized excess returns and CER of the portfolio formed by the forecasts of volatility from constant coefficient models are smaller than that of the corresponding time-varying coefficient models, indicating time-varying coefficient strategy resulting in portfolios with better performances. For equation (17), the mean of excess returns ($R$) formed by the time varying or constant coefficient strategy is 26.079% or 25.420% during the same period, and the average return of the SSEC index is just 2.015%, implying that the investor gets significant excess returns by considering the realized volatility of the SSEC index. We find that these strategies successfully reduce the weight of the stock index during the stock market crash in 2009 and 2015-16 and increase the weight during calm periods. For the models that include only 1 additional variable, the best model is equation (6), indicating that the lagged negative return is the most important additional variable for asset allocation, and it not only significantly affected future volatility, but also the excess returns of the portfolio. Equation (17), which contains the most additional variables, performs best, indicating that the corresponding additional variables do improve portfolio performance. For the same strategy, the less the risk aversion.
returns and improving the utility of a certain risk aversion model. In portfolio exercises, we also demonstrate that, statistically, the time-varying coefficient models generate more accurate forecasts than the MRS and CER generated by the LHAR-RV, AHAR-RV, and HAR-RV-ALL models do not always outperform the CC-type models, i.e., the MRS-type coefficient models for forecasting the realized volatility of energy, precious metal, and foreign exchange markets.

**Appendix**

According to West and Harrison [40], we calculate the posterior probability of the system coefficients and the observational variance after a new observation of $RV$. The prior we used in 2.3 at time $t=0$ results in completely conjugate Bayesian analysis. Based on the information up to time $t$, the posteriors of $V$ and coefficients $\theta_{t-1}$ follows inverse gamma and student’s $t$ distribution:

$$V|\mathcal{F}_t \sim t_\nu\left(\frac{n_t}{2} \frac{n_t S_t}{2}\right).$$

$$\theta_{t-1}|\mathcal{F}_t \sim T_\eta\left(\bar{m}_t, C_t\right),$$

where $V$ is the observational variance, $n_t$ is the degree of freedom, $S_t$ is the point estimator of $V$, $m_t$ is a vector which denotes the mean of the estimate of coefficients $\theta_{t-1}$ conditional on $\Phi_t$ and $C_t$ denotes the estimator for the covariance matrix of $\theta_{t-1}$. As mentioned in Section 2.3, by introducing the forgetting factor $\lambda$, the distribution of $\theta_{t-1}|\mathcal{F}_t$ is simplified to $T_\eta\left[\bar{m}_t, \lambda/(1-\lambda)C_t\right]$. According to equation (19), by integrating the conditional density of $RV_{t+1}$ over $V$ and $\theta$, we get the forecast distribution of $RV_{t+1}$:

$$RV_{t+1}|\mathcal{F}_t \sim T_\eta\left(\bar{R}_{t+1}V_{t+1}, Q_{t+1}\right),$$

where $\bar{R}_{t+1} = F_t^T m_t$, $Q_{t+1} = F_t^T R_t F_t + S_t$, and $R_t$ is the variance of prior distribution of $\theta_t|\mathcal{F}_t$.

Once we observe $RV_{t+1}$, we can calculate the error in prediction as $e_{t+1} = RV_{t+1} - \bar{RV}_{t+1}$. The posteriors about $\theta_{t+1}$ and $V$ follow $t_\nu\left((n_t+1)/2, (n_t+1)S_{t+1}/2\right)$ and $T_\eta\left(m_{t+1}, C_{t+1}\right)$, where

$$n_{t+1} = n_t + 1,$$

and

$$S_{t+1} = S_t + \frac{S_t}{n_t+1} \left(\frac{e_{t+1}^2}{Q_{t+1}} - 1\right).$$

### Discrete Dynamics in Nature and Society 11

Table 4: Portfolio performance measures, 2006-10-9 to 2018-4-23.

<table>
<thead>
<tr>
<th></th>
<th>HAR-RV</th>
<th>LHAR-RV</th>
<th>HAR-RV-J</th>
<th>LHAR-RV-OR</th>
<th>AHAR-RV</th>
<th>HAR-RV-V</th>
<th>HAR-RV-ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma = 3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MRS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TVC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MRS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TVC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma = 6$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MRS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TVC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.167</td>
<td>16.150</td>
<td>4.666</td>
<td>4.901</td>
<td>10.076</td>
<td>5.159</td>
<td>17.905</td>
</tr>
<tr>
<td></td>
<td>CER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MRS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TVC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table reports the performance measures for an investor with the risk aversion coefficient of 3 or 6, and we show the mean of excess returns ($R$) and CER of the portfolios formed by the forecast of volatility.

4. Conclusions

Structural breaks, noisy proxies, and model specification errors indicate that the instability of coefficients represents an important challenge in volatility forecasting. We extend the HAR-type model by explicitly considering the time variation of coefficients and apply these models in forecasting realized volatility of the SSEC. The empirical results demonstrate that, statistically, the time-varying coefficient models generate more accurate forecasts than the MRS and constant coefficient models. In portfolio exercises, we also find that our models are helpful in generating more excess returns and improving the utility of a certain risk aversion investor.

In our time-varying coefficient regression framework, we find the predictive abilities of the three heterogeneous volatility components to be negatively correlated and quite different during calm and turbulent periods. We evaluate the variation of the leverage effect by capturing the time-varying “leverage parameter”; downside risk increases future volatility of the SSEC index on the whole, but the leverage effect is insignificant or inverse during bull markets. Our findings indicate the importance of considering time variation of coefficients, suggesting that practitioners and market regulators should treat volatility of the stock market with a dynamic perspective. For further enhancements, future research may explore the potential of time-varying coefficient models for forecasting the realized volatility of energy, precious metal, and foreign exchange markets.
\[ m_{t+1} = m_t + A_{t+1}e_{t+1}, \quad (A.7) \]
\[ C_{t+1} = R_t - A_{t+1}A_{t+1}'Q_{t+1}, \quad (A.8) \]

where \( A_{t+1} \) is the adoptive vector \( A_{t+1} \equiv (R_tF_t/Q_{t+1}) \) (see West and Harrison [40]).

After we get the forecast of volatility from each submodel, we combine these forecasts through the Bayesian model averaging approach. Suppose \( M_t \) denotes a submodel which contains a certain choice of predictors from a set of \( n = 2^k - 1 \) candidates, then \( \text{RV}_{t+1,i} \), which denotes the point estimator of \( \text{RV}_{t+1} \) from \( M_i \) is
\[ \text{RV}_{t+1,i} = E(\text{RV}_{t+1} | M_i, F_t) = F_t^T m_i. \quad (A.9) \]

The prior weight of each submodel is set as
\[ p(M_i | F_t) = 1/(2^k - 1), \quad i = 1, 2, \ldots, 2^k - 1. \]

As a new observation of \( \text{RV} \) arrives, probabilities of submodels are updated using the Bayesian recursions:
\[ p(M_i | F_{t-1}) \propto p(M_i | F_{t-1}) p(\text{RV}_{t+1,i} | M_i, F_{t-1}), \quad (A.10) \]

where \( p(\text{RV}_{t+1,i} | F_{t-1}) = \sum_{i=1}^{2^k-1} p(\text{RV}_{t+1,i} | M_i, F_{t-1}) p(M_i | F_{t-1}) \).

\[ p(M_i | F_{t-1}) \propto \frac{1}{\sqrt{Q_{t,i}}} T_{n,i} \left( \frac{\text{RV}_t - \text{RV}_{t,i}}{\sqrt{Q_{t,i}}} \right), \quad (A.11) \]

where \( \text{RV}_{t,i} \) and \( \sqrt{Q_{t,i}} \) are the point estimators of the mean and variance of the predictive density of \( \text{RV} \) from submodel \( M_i \), \( T_{n,i} \) is the density of Student’s \( t \) distribution with degree of freedom \( n_{t,i} \).

\( \text{RV}_{t+1,i} \), which is the predictive mean of \( \text{RV}_{t+1} \) conditional on \( \Phi_0 \), is a weighted average of predictive mean of each submodel:
\[ \text{RV}_{t+1} = \sum_{i=1}^{2^k-1} E(\text{RV}_{t+1} | M_i, F_t) p(M_i | F_t) \]
\[ \quad = \sum_{i=1}^{2^k-1} \text{RV}_{t+1,i} p(M_i | F_t). \quad (A.12) \]

Data Availability
The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest
The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments
This study was funded by the Fundamental Research Funds for the Central Universities of China (Grant no. JBK1607118).

References


