International Journal of Forecasting 33 (2017) 132–152

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Realized volatility forecasting of agricultural commodity futures using the HAR model with time-varying sparsity

Fengping Tian [a], Ke Yang [b], Langnan Chen [c,*]

[a] International School of Business and Finance, Sun Yat-sen University, Guangzhou 510275, China
[b] School of Economics and Commerce, South China University of Technology, Guangzhou 510006, China
[c] Lingnan College, Sun Yat-sen University, Guangzhou 510275, China

ARTICLE INFO

Keywords:
Realized volatility
Forecast
HAR model
Time-varying sparsity
Agricultural commodity futures

ABSTRACT

We develop a time-varying HAR model where both the predictors and the regression coefficients are allowed to change over time, and use it to forecast the realized volatility in the fast-growing agricultural commodity futures markets of China. The proposed model is constructed by incorporating all potential predictors in a time-varying HAR framework, and giving the independent normal-gamma autoregressive (NGAR) process priors to the regression coefficients. The out-of-sample forecast results show that the proposed HAR model with time-varying sparsity improves the forecast performances substantially relative to both the simple HAR model and more sophisticated HAR-type models in almost all cases. Finally, the forecast performance of the proposed model is robust to the alternative proxies of volatility.

1. Introduction

The measurement and forecasting of the volatility of agricultural commodity futures is crucial for agricultural production, resource allocation, and risk management. Traditionally, GARCH-type models where the volatility is assumed to be a latent factor and is estimated from the daily return series are commonly used for forecasting the conditional volatility. The advent of high-frequency data containing more intraday trading information has enabled the realized volatility (RV) based on the sum of squared intraday returns to be applied, thus making it a promising method for modeling and forecasting the volatility, as it allows a precise estimation of the daily volatility.

Nevertheless, almost none of the existing models for RV forecasting are able to capture the time-varying property of the regression coefficients and the predictors in model specifications simultaneously. This paper intends to fill this gap in the literature by developing a time-varying HAR model and using it to forecast the RV in a relatively under-researched market, namely Chinese agricultural commodity futures markets.

The RV constructed from the high-frequency data (Andersen & Bollerslev, 1998) is affected less by the measurement error, and thus less noisy. It allows us to treat the volatility as an observed variable and can be modeled directly, rather than being treated as a latent process in the GARCH-type models. Many properties of the RV have been well-documented in the recent literature. One of the most important properties of the RV is that it exhibits high persistence, as has been shown by Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold, and Labys (2001, 2003), and Barndorff-Nielsen and Shephard (2002), among many others. The volatility forecast models using the realized measures can be divided into two main approaches, as was suggested by Hansen and Lunde

* Corresponding author.
E-mail address: lnscln@mail.sysu.edu.cn (L. Chen).

http://dx.doi.org/10.1016/j.ijforecast.2016.08.002
0169-2070/

(2011): model-based volatility forecasts and reduced form forecasts. The model-based approach involves cases where the realized measure of volatility is included in a model for returns, such as a GARCH-type model that specifies the entire distribution of returns. The models that follow this approach, such as the GARCH-X model (Engle, 2002), the multiplicative error model (Engle & Gallo, 2006), the HEAVY model (Shephard & Sheppard, 2010) and the realized GARCH model (Hansen, Huang, & Shek, 2012), have convincingly shown the usefulness of the realized measure of volatility for modeling and forecasting the future conditional volatility. The reduced-form method relies on the daily time series of the realized volatility for forecasting the future latent volatility over some time spans. Each element of these time series can be viewed as a proxy for the latent volatility. Over recent decades, the reduced form time series models for the realized volatility have been used extensively to capture the persistence of the observed volatility series. For instance, Andersen et al. (2003), Giot and Laurent (2004), Lieberman and Phillips (2008) and Martens, Van Dijk, and De Pooter (2009) reported evidence of long memory and modeled the RV as a fractionally integrated process. As was suggested by Forsberg and Ghysels (2007) and Ghysels, Santa-Clara, and Valkanov (2006), the mixed data sampling approach (MIDAS) is also successful empirically at accounting for the observed strong serial dependence. In particular, inspired by the work of Müller et al. (1997), Corsi (2009) considered the volatility persistence as the effect of the sum of the heterogeneous components in the financial market, and developed an additive cascade model of the realized volatility summed at different time horizons. This cascade of heterogeneous volatility components results in a simple AR-type model that takes into account volatilities realized over different time horizons, and thus is called the heterogeneous-autoregressive (HAR) model. Although the HAR model for the RV is not classified formally as a long memory model, it regenerates the volatility persistence detected through the empirical analysis of the financial markets. The tractable estimation and superior forecast performance of the HAR model make it popular for forecasting the RV. Such HAR-type parameterizations have also been suggested by Andersen, Bollerslev, and Diebold (2007), Andersen, Bollerslev, and Huang (2011), Corsi, Mittnik, Pigorsch, and Pigorsch (2008) and McAleer and Medeiros (2008). The details of these models for the RV differ, but the general framework involves a dependent variable such as RV (or the log of RV) and the explanatory variables, such as lags of RV and other predictors.

However, there are several problems that we may encounter when using these reduced form time series approaches for modeling and forecasting the RV. First, the coefficients of the models for the RV forecast can change over time. Many studies have shown that there are structural breaks in the RV series and that the parameters change (Choi, Yu, & Zivot, 2010; Liu & Maheu, 2008; Raggi & Bordignon, 2012; Yang & Chen, 2014). The regression-based methods with constant coefficients are not able to capture such changes. Furthermore, the predictors for the model specification can potentially change over time. For example, the set of predictors for the RV forecast may have been different during the period of the financial crisis. Some variables might be useful for forecasting the RV in some periods but not others, implying that the forecasting model specifications might change over time.

Second, the number of potential predictors for the RV forecast can be large. Besides the six predictors with lags considered by Liu and Maheu (2009), there may be various other potential predictors for the RV forecast, such as the daily squared returns of Andersen and Bollerslev (1998), the range estimator of Brandt and Jones (2006), and the jumps detected by the $C - Tz$ of Corsi, Pirino, and Reno (2010). The existence of too many predictors can lead to a huge number of models. Liu and Maheu (2009) consider 72 model specifications and use the Bayesian model averaging (BMA) approach to forecast the RV, but the computations become complicated when so many models are investigated. In addition, none of the predictors of the RV dominate in forecasting power across all markets and the forecast horizons (see Liu & Maheu, 2009). Thus, selecting any one predictor for model specifications is risky. A natural way to avoid this risk is to incorporate all potential predictors in a model, but this method has the potential to cause over-fitting problems when too many predictors are included.

Finally, most of the existing studies on modeling and forecasting the volatility of agricultural commodity futures have focused primarily on developed markets. Little evidence on volatility forecasting in emerging markets has been documented, although a large volume of contracts is currently traded in emerging markets. Based on the Futures Industry Association (FIA) statistics in the USA, the trading volume in China's agricultural futures accounted for 58% of the global agricultural futures in 2011, and was about 292 trillion yuan at the end of 2014, implying that the Chinese agricultural commodity futures market has become the world's biggest market, with an increasing influence on pricing and portfolio decisions (e.g., Cheung & Miu, 2010). However, there is no literature on RV forecasting in this market.

We develop a new reduced form time series model for the RV by incorporating all potential predictors into a time-varying HAR framework, in order to avoid the model risk that is associated with selecting any one predictor for model specification, and applying the independent normal-gamma autoregressive (NGAR) processes priors proposed by Griffin and Brown (2010) and Kalli and Griffin (2014) to the regression coefficients, which allow both the regression coefficients and the predictors to change over time. Moreover, the NGAR process prior allows us to control the sparsity of the posterior distribution of the regression coefficients, and assumes that the regressors rarely jump in and out of the model, both of which points are important in order to avoid the over-fitting of the time-varying HAR model. We use the proposed model to forecast the RV in Chinese agricultural futures markets, and employ the model confidence set (MCS) to evaluate and compare the out-of-sample point forecast performances of the proposed model and the alternative HAR-type models that have been developed recently. We find that the proposed model appears to be the most accurate of all of the models in this study for forecasting the RV of Chinese agricultural commodity futures. Finally, the forecast performance of the proposed model is robust to the alternative proxies of volatility.

This paper is organized as follows. Section 2 develops the HAR model with time-varying sparsity. Section 3 presents the data, summary statistics, and out-of-sample forecast power of various potential predictors. Section 4 evaluates the performances of the proposed models relative to those of the competing models. Section 5 concludes.

2. The HAR model with time-varying sparsity

This section starts with a brief introduction to the realized volatility and its potential predictors, then describes the proposed model in detail, including model specifications and estimations.

2.1. Realized volatility

Over recent decades, the easy access to high-frequency financial data has led to the development of approaches for estimating the (latent) volatility of financial assets' prices. The most well-known and simplest realized measure is the realized variance, denoted by

$$\mathrm{RV}_t = \sum_{j=1}^{n} [p_{j,t} - p_{j-1,t}]^2 = \sum_{j=1}^{n} r_{j,t}^2, \tag{1}$$

where $p_{j,t}$ is the jth intraday log-price observation on day t, $r_{j,t} = p_{j,t} - p_{j-1,t}$ is the jth intraday log return, and n is the number of such returns within each trading day. In theory, this measure estimates the quadratic variation of the price process on day t consistently if the prices are observed without noise. In practice, the raw RV estimated from Eq. (1) may be an inconsistent estimator due to the autocorrelation in the high-frequency returns, caused by market microstructure noise (Bandi & Russell, 2006; Hansen & Lunde, 2006). The recent econometric literature suggests the use of 5 min return data in order to alleviate the effects of the noise. However, Zhang, Mykland, and Aït-Sahalia (2012) find that this is not an adequate solution to this problem.

As has been suggested in the recent literature, there are three types of realized estimators that are somewhat robust to noise: pre-averaging (Jacod, Li, Mykland, Podolskij, & Vetter, 2009), multiscale (Zhang et al., 2012; Zhang, 2006) and realized kernel (Barndorff-Nielsen, Hansen, Lunde, & Shephard, 2008). Following Liu and Maheu (2009), we focus on the kernel-based estimator proposed by Hansen and Lunde (2006), which is designed to capture the effects of autocorrelation on high frequency returns, induced by both the independent and dependent noise. This estimator is almost identical to the multiscale estimator of Zhang et al. (2012) (see Barndorff-Nielsen et al., 2008), and is also a special case of the flat-top realized kernel estimator suggested by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009). This kernel-based estimator takes the form

$$\mathrm{RV}_t^q = \sum_{j=1}^{n} r_{j,t}^2 + 2\sum_{w=1}^{q}\left(1 - \frac{w}{q+1}\right)\sum_{i=1}^{n-w} r_{i,t} r_{i+w,t}. \tag{2}$$

Following Hansen and Lunde (2006), we set $q = \left\lceil \frac{\omega}{(b-a)/n} \right\rceil$, where $\lceil x \rceil$ is the smallest integer that is greater than or equal to x, ω is the desired width of the lag window, and $b-a$ is the length of the sampling period. In this paper, we set $\omega = 5$, 10, and 15 min, $b - a = 240$ min, $n = 48$, and $q = 1$, 2 and 3.

2.2. Potential predictors

Since the daily squared returns are extremely noisy (Andersen & Bollerslev, 1998), we decided not to use them as a potential predictor for the RV forecast, but instead follow Engle and Gallo (2006) and consider the daily log-range estimator, defined as $\mathrm{Range}_t = \log(P_{H,t}/P_{L,t})$, where $P_{H,t}$ and $P_{L,t}$ are the intraday high and low price levels on day t, as the daily log-range is closer to being normally distributed and more efficient than the daily squared returns (Alizadeh, Brandt, & Diebold, 2002).

Several studies have considered the realized multipower variation with a range of different orders, such as 0.5, 1 and 1.5, as proposed by Barndorff-Nielsen and Shephard (2004) for the realized volatility models (Forsberg & Ghysels, 2007; Ghysels et al., 2006; Liu & Maheu, 2009). However, Corsi et al. (2010) suggest that the realized bipower variation is greatly upward biased in the presence of jumps in the finite samples, and therefore using measures based on the realized threshold multipower variation instead of those based on the realized multipower variation could improve the performances of the RV forecast. We therefore use the realized threshold multipower variation (TMPV) with orders 0.5, 1 and 1.5 and the realized threshold bipower variation (TBPV) as the potential predictors for the RV forecast, in addition to the daily ranges and the lagged RV. Since the bipower variation is also contaminated by the market microstructure noise, we employ an adjusted realized threshold bipower variation based on the staggered returns, as was suggested by Andersen et al. (2007) and Huang and Tauchen (2005), in order to decrease the correlation induced by the market microstructure noise. We quote the adjusted TBPV below as TBPV_t. The multipower variation measures are certainly contaminated by the market microstructure, but Ghysels and Sinko (2006) suggest that this is much harder to correct and may be less important empirically. Therefore, we do not make any adjustments for the realized threshold multipower variation. We quote the realized threshold multipower variation with orders 0.5, 1 and 1.5 as $\mathrm{TMPV}_t^{[0.5]}$, $\mathrm{TMPV}_t^{[1]}$ and $\mathrm{TMPV}_t^{[1.5]}$ below.

Since Corsi et al. (2010) showed that past jumps have a positive and significant impact on the future volatility, we also include both past jumps and the corresponding continuous component in the RV forecast as possible predictors. In this paper, we set the confidence level to 95%, and use the $C-Tz$ statistics proposed by Corsi et al. (2010) to estimate these two measures, referred to below as TJ_t and TC_t.

2.3. Model specification

2.3.1. Time-varying HAR model

The evidence in the previous studies has proved that the RV obtained from the intraday returns is long-range dependent (Andersen & Bollerslev, 1998; Andersen et al., 2001; Barndorff-Nielsen & Shephard, 2002). Based on the heterogeneous market hypothesis of Müller et al. (1997), Corsi (2009) approximates this long memory property of RV by using a long lagged autoregressive process, called the heterogeneous autoregressive (HAR) model. The simplest, logarithmic, version of HAR-RV model is defined as

$$\begin{aligned}\log(\mathrm{RV}_{t,h}) &= \alpha_0 + \alpha_1 \log(\mathrm{RV}_{t-1,1}) \\ &\quad + \alpha_2 \log(\mathrm{RV}_{t-5,5}) + \alpha_3 \log(\mathrm{RV}_{t-20,20}) \\ &\quad + u_{t,h},\ u_{t,h} \sim \mathrm{N}(0, \sigma_u^2),\end{aligned} \tag{3}$$

where $\mathrm{RV}_{t,h} = \frac{1}{h}\sum_{i=1}^{h} \mathrm{RV}_{t+i-1}$ is the h-step-ahead average RV, with $h = 1, 5$ and 20 being the daily, weekly and monthly volatility components, respectively. The HAR model is then a parsimonious AR model that has been reparameterized by imposing different sets of restrictions on the autoregressive coefficients of the AR model that can be estimated easily by simple OLS. The tractable estimation and superior performance of the HAR model make it a popular model for forecasting the RV.

However, as was discussed in Section 1, structural breaks in the RV series and the time-varying coefficients and set of predictors for the RV forecast mean that the HAR-type models with constant coefficients and fixed predictors cannot capture these time-varying properties. Furthermore, the existence of so many predictors whose forecast powers for the RV cannot dominate across markets and forecast horizons may lead to the model risk associated with selecting any one predictor. Incorporating all potential predictors into a model is a natural approach to avoiding this risk, but may result in potential over-fitting problems when too many predictors are included.

We overcome those constraints by first incorporating all potential predictors into a time-varying HAR model specification in order to avoid the model risk associated with selecting any one predictor. The proposed time-varying HAR model is as follows:

$$\begin{aligned}\log(\mathrm{RV}_{t,h}) &= \alpha_{0,t} + \alpha_{1,t} \log(\mathrm{RV}_{t-1,1}) \\ &\quad + \alpha_{2,t} \log(\mathrm{RV}_{t-5,5}) + \alpha_{3,t} \log(\mathrm{RV}_{t-20,20}) \\ &\quad + \alpha_{4,t} \log(\mathrm{TBPV}_{t-1,1}) + \alpha_{5,t} \log(\mathrm{TBPV}_{t-5,5}) \\ &\quad + \alpha_{6,t} \log(\mathrm{TBPV}_{t-20,20}) \\ &\quad + \cdots \\ &\quad + \alpha_{22,t} \log(\mathrm{TJ}_{t-1,1} + 1) + \alpha_{23,t} \log(\mathrm{TJ}_{t-5,5} + 1) \\ &\quad + \alpha_{24,t} \log(\mathrm{TJ}_{t-20,20} + 1) + u_{t,h} \\ &= \sum_{i=0}^{m=24} \alpha_{i,t} x_{i,t-1} + u_{t,h}, \\ &= X_{t-1}\alpha_t + u_{t,h}, \\ &\quad u_{t,h} \sim \mathrm{N}(0, \sigma_{u,t}^2),\ t = 1, \ldots, T, i = 0, \ldots, m,\end{aligned} \tag{4}$$

where $\alpha_t = [\alpha_{0,t}, \ldots \alpha_{24,t}]'$ is a vector of unknown regression coefficients for all regressors at time t, $X_{t-1} = [1, \log(\mathrm{RV}_{t-1,1}), \ldots, \log(\mathrm{TJ}_{t-20,20} + 1)]'$ is a vector of all predictors, including 1 (allowing for an intercept) and the logarithmic versions of daily, weekly and monthly components for the RV, TBPV, daily range, TMPV with orders of 0.5, 1 and 1.5, TC, and TJ + 1, respectively.

As was discussed by Gouriéroux and Jasiak (2006), Pitt, Chatfield, and Walker (2002) and Pitt and Walker (2005), the gamma autoregressive (GAR) process represents the complex nonlinear dynamics of various positive-valued time series, such as the financial volatility series, and is sufficiently flexible to accommodate both short and long memory. In addition, this process allows us to assign the correct stationary density to the initial observation, leading to an efficient likelihood estimation, and also enables us to obtain the autocorrelations of the series and point forecasts easily. Therefore, we utilize a GAR process for the time-varying variances of the innovation term $u_{t,h}$, $\sigma_{u,1}^2, \ldots, \sigma_{u,T}^2$, which can be defined by using the latent variables $\kappa_1^{\sigma_u}, \ldots, \kappa_{T-1}^{\sigma_u}$ through the recursion:

$$\sigma_{u,t}^2 \sim \mathrm{Ga}(\lambda^{\sigma_u} + \kappa_{t-1}^{\sigma_u}, \lambda^{\sigma_u}/(\mu^{\sigma_u}(1-\rho^{\sigma_u}))) \quad \text{and}$$
$$\kappa_{t-1}^{\sigma_u} | \sigma_{u,t-1}^2 \sim \mathrm{Pn}(\lambda^{\sigma_u}\rho^{\sigma_u}\sigma_{u,t-1}^2/(\mu^{\sigma_u}(1-\rho^{\sigma_u}))),$$

for $t = 2, \ldots, T$ and $\sigma_{u,1}^2 \sim \mathrm{Ga}(\lambda^{\sigma_u}, \lambda^{\sigma_u}/\mu^{\sigma_u})$, where $z \sim \mathrm{Ga}(a, b)$ denotes that z follows a gamma distribution with shape parameter a and mean a/b, and $z \sim \mathrm{Pn}(\mu)$ denotes that z follows a Poisson distribution with mean μ. This defines the first-order autoregressive model for $\sigma_{u,1}^2, \ldots, \sigma_{u,T}^2$, with autoregressive parameter ρ^{σ_u} and stationary distribution $\mathrm{Ga}(\lambda^{\sigma_u}, \lambda^{\sigma_u}/\mu^{\sigma_u})$. Following Kalli and Griffin (2014), we complete the specification of the volatility process $\sigma_{u,t}^2$ by considering the following priors for the parameters: $\lambda^{\sigma_u} \sim \mathrm{Ga}(3, 1)$, $p(\mu^{\sigma_u}) \propto (1 + \mu^{\sigma_u})^{-3/2}$, and finally $\rho^{\sigma_u} \sim \mathrm{Be}(38, 2)$.

This time-varying HAR model encompasses many other special HAR models in the recent literature, such as the logarithmic version of HAR-RV model of Corsi (2009), the logarithmic version of HAR-X model (Liu & Maheu, 2009; Tseng-Chan, Huimin, & Chin-Sheng, 2009), the logarithmic version of the HAR-$X-J$ model, and the HAR-CJ model of Andersen et al. (2007).

2.3.2. The NGAR process for $\alpha_{i,t}$

In the literature, the problem of over-fitting when the regression models have a large number of regressors is usually avoided by assuming that only a subset of the regressors are important for the prediction. For the proposed time-varying HAR model above, this assumption can be developed naturally as time-varying subsets of important regressors, which can be described in the prior by defining a stochastic process for $\alpha_{1,t}, \ldots, \alpha_{m,t}$ in Eq. (4) that allows for a subset of $\alpha_{1,t}, \ldots, \alpha_{m,t}$ to be equal or close to zero at time t, and allows this subset to change over time. We define the proportion of parameters $\delta = (\delta_1, \ldots, \delta_s)$ that are close to zero as the sparsity of δ, with a larger proportion being referred to as a greater sparsity. There are two interesting forms of sparsity in the proposed time-varying HAR model: (i) the sparsity of $\alpha_i = (\alpha_{i,1}, \ldots, \alpha_{i,T})$, which is the proportion of the time that $\alpha_{i,t}$ is close to zero; and (ii) the sparsity of $\alpha_{1,t}, \ldots, \alpha_{m,t}$, which is the proportion of the regression coefficients that are set close to zero at time t. The assumption of time-varying subsets of important regressors can be described by the time-varying sparsity of $\alpha_{1,t}, \ldots, \alpha_{m,t}$.

We describe these forms of sparsity by utilizing the independent NGAR process priors for the regression coefficients $\alpha_{1,t}, \ldots, \alpha_{m,t}$, which can be defined as

$$\alpha_{i,s} = \sqrt{\frac{\psi_{i,s}}{\psi_{i,s-1}}}\varphi_i \alpha_{i,s-1} + \eta_{i,s},$$

for $\eta_{i,s}|\psi_{i,s} \sim \mathrm{N}(0, (1-\varphi_i^2)\psi_{i,s}),\ s = 2, \ldots, T$, which is a normal AR(1) process conditional on $\psi_i = (\psi_{i,1}, \ldots, \psi_{i,T})$, and where $\psi_{i,s}$ is a first-order gamma autoregressive process that can be defined through the recursion using the latent variables $\kappa_{i,1}, \ldots, \kappa_{i,s-1}$ as follows:

$$\psi_{i,s}|\kappa_{i,s-1} \sim \mathrm{Ga}\left(\lambda_i + \kappa_{i,s-1}, \frac{\lambda_i}{\mu_i(1-\rho_i)}\right),$$
$$\kappa_{i,s-1}|\psi_{i,s-1} \sim \mathrm{Pn}\left(\frac{\rho_i \lambda_i \psi_{i,s-1}}{\mu_i(1-\rho_i)}\right),$$

with $\alpha_{i,1}|\psi_{i,1} \sim \mathrm{N}(0, \psi_{i,1}),\ \ \psi_{i,1} \sim \mathrm{Ga}(\lambda_i, \lambda_i/\mu_i)$, and where $\mathrm{Pn}(u)$ is a Poisson distribution with mean u. The NGAR process can be written as $\alpha_i \sim \mathrm{NGAR}(\lambda_i, \mu_i, \varphi_i, \rho_i)$.

The NGAR process prior allows us to control the two levels of sparsity for the posterior distribution of the regression coefficients, and assumes that the regressors rarely jump in and out of the model, both of which points are crucial in order to avoid over-fitting the time-varying HAR model. More specifically, this process prior assumes that $\alpha_{i,t}|\psi_{i,t}$ follow a normal distribution with a mean of zero and variance of $\psi_{i,t}$, where $\psi_{i,t}$ follows a gamma distribution. Thus, $\psi_{i,t}$ acts as the relevance of the ith regressor at time t, with a smaller value of $\psi_{i,t}$ suggesting greater shrinkage of $\alpha_{i,t}$. The sparsity parameter λ_i controls the proportion of the time that the ith regression coefficient spends close to zero. For a fixed μ_i, this proportion increases when the value of λ_i decreases and the regression coefficient $\alpha_{i,t}$ is close to zero for a larger proportion of observations. This introduces the first level of sparsity (at the level of time-varying regression coefficients). We use the prior $p(\lambda_i) \propto \lambda_i(0.5+\lambda_i)^{-4}$ to specify the sparsity parameter λ_i, which is a heavy-tailed prior with values of around one. The parameter μ_i plays the role of the overall relevance of the ith regressor with a smaller value of μ_i, implying that $\alpha_{i,t}$ is closer to zero for all t, since it controls the marginal variance of $\alpha_{i,t}$. We therefore choose the hierarchical prior, $\mu_i \sim \mathrm{Ga}(\lambda^*, \lambda^*/\mu^*),\ \ i = 0, \ldots, m$ to specify $\mu_1, \ldots \mu_m$. The parameter λ^* is the sparsity parameter for μ_i, which introduces the second level of sparsity (at the level of time-varying regressors), with a small value of λ^* implying that more μ_is are close to zero, which results in $\alpha_{i,t}$ being close to zero for all time t for more regressors, since μ_i is the variance of $\alpha_{i,t}$ under the stationary distribution. We select the prior $\lambda^* \sim \mathrm{Ex}(1/s^*)$ for λ^*, where the hyper-parameter s^* is the prior mean of λ^* and μ^* is given a heavy-tailed prior with prior mean b^* ($p(u^*) \propto (u^* + 2b^*)^{-3}$). The autocorrelation parameter φ_i controls the dependence between $\alpha_{i,t}$ and $\alpha_{i,t-1}$ conditional on $\psi_i = (\psi_{i,t} \ldots \psi_{i,T})$, and the autocorrelation parameter ρ_i controls the dependence between $\psi_{i,t-1}$ and $\psi_{i,t}$, with larger values of ρ_i leading to larger autocorrelations and favoring processes that spend longer periods either close to or far away from zero. We select the priors $\varphi_i \sim \mathrm{Be}(77.6, 2.4)$ and $\rho_i \sim \mathrm{Be}(77.6, 2.4)$ to the autocorrelation parameters φ_i and ρ_i that allow the processes for the regression coefficients and the relevant regressors to be strongly autocorrelated. This effectively excludes any models with regression coefficients that change rapidly over time.

2.4. Estimation method

We employ the MCMC method to fit the proposed HAR model with time-varying sparsity, labeled the HAR-TVS model. Due to the correlation between the process $(\alpha_i, \psi_i, \kappa_i)$ and its parameters $\theta_i = (\lambda_i, \mu_i, \rho_i, \varphi_i)$, and between the process $(\sigma_u^2, \kappa^{\sigma_u})$ and its parameters $\theta^{\sigma_u} = (\lambda^{\sigma_u}, \mu^{\sigma_u}, \rho^{\sigma_u})$, the standard Gibbs sampler that simulates from each full conditional distribution in turn may cause highly autocorrelated draws. Similarly to Kalli and Griffin (2014), our simulating plan involves jointly renovating (ψ_i, κ_i) with θ_i and $(\sigma_u^2, \kappa^{\sigma_u})$ with θ^{σ_u}, whilst integrating out a time-varying subset of $\alpha_1, \ldots, \alpha_m$. We therefore estimate the proposed model by using the Gibbs sampler with the adaptive Metropolis–Hastings random walk (AMHRW) step and the adaptive scale within adaptive Metropolis (ASWAM) step. The steps of the Gibbs sampler are detailed in Appendix A of Kalli and Griffin (2014). We summarize the algorithm as follows:

- Initialize the chain at $(\alpha_{(0)}, \psi_{(0)}, \kappa_{(0)}, \theta_{(0)}, \sigma^2_{u,(0)}, \kappa^{\sigma_u}_{(0)}, \theta^{\sigma_u}_{(0)}, \mu^*_{(0)}, \lambda^*_{(0)})$.
- At step $j = 1, \ldots n$,
  - Update ψ with the AMHRW step based on Atchadé and Rosenthal (2005);
  - Update κ with the AMHRW step;
  - Update σ_u^2 from a generalised inverse Gaussian distribution;
  - Update κ^{σ_u} with the AMHRW step;
  - Update θ, θ^{σ_u} and α with the ASWAM step based on the work of Haario, Saksman, and Tamminen (2001) and Atchadé and Fort (2010);
  - Update μ^* and λ^* with the AMHRW step.

3. Data and preliminary analysis

This section begins with a brief introduction to Chinese commodity futures markets, then presents the dataset and descriptive statistics. Finally, it discusses the out-of-sample forecast power of various potential predictors.

3.1. Chinese commodity futures markets

There are three currently commodity futures exchanges in China: Zhengzhou Commodity Exchange (ZCE), Dalian Commodity Exchange (DCE) and Shanghai Futures Exchange (SFE). SFE specializes in metals futures, while ZCE and DCE specialize in agricultural commodity futures, primarily wheat and soybeans, respectively. The FIA's Annual Report 2014 indicates that the international influence of the Chinese commodity futures exchanges has increased over recent decades. The three commodity exchanges are now ranked 9th, 10th, and 13th worldwide, respectively,

Table 1
The means of liquidity measures for agricultural commodity futures with different times to delivery.

	Measures	*Soybean*	*Cotton*	*Gluten wheat*	*Corn*	*Early indica rice*	*Palm*
Nearby month	Roll	0.6408	0.4328	0.4922	0.6511	0.5421	0.4421
	Zeros	0.2124	0.4054	0.3016	0.3426	0.2512	0.4167
	Amihud	0.9716	0.0286	2.1733	5.7278	0.0508	0.0296
One month	Roll	0.6483	0.4813	0.9805	1.4478	0.6154	0.4815
	Zeros	0.3822	0.5114	0.8869	0.5481	0.3122	0.5016
	Amihud	3.5824	0.2624	8.5692	11.8920	0.1563	0.2601
Two months	Roll	0.6648	0.4566	0.7865	0.9706	0.5472	0.4537
	Zeros	0.2818	0.4348	0.3940	0.4522	0.2758	0.4300
	Amihud	1.8726	0.0651	6.3100	11.2164	0.0851	0.0647
Three months	Roll	0.6463	0.4540	0.5045	0.6674	0.5460	0.4531
	Zeros	0.2139	0.4216	0.3226	0.3532	0.2713	0.4254
	Amihud	0.9728	0.0683	2.5104	6.5382	0.0518	0.0640
Four months	Roll	0.6468	0.4711	0.6418	0.9114	0.6123	0.4729
	Zeros	0.2247	0.4287	0.4822	0.4569	0.3115	0.4258
	Amihud	0.9775	0.1662	6.3405	7.8926	0.1178	0.1662
Five months	Roll	0.6518	0.4839	0.7559	1.1267	0.6790	0.4857
	Zeros	0.2318	0.5228	0.6482	0.4787	0.3992	0.5206
	Amihud	1.1440	0.7393	16.7140	8.0248	0.3246	0.7388

Note: 'Roll' represents the effective spread of Roll (1984) ($\times 10^3$); 'Zeros' the proportion of zero 5 min returns during a trading day; and 'Amihud' the illiquidity measure of Amihud (2002) ($\times 10^8$). The lower the value, the better the liquidity of the asset. The futures contracts are grouped according to their time to delivery.

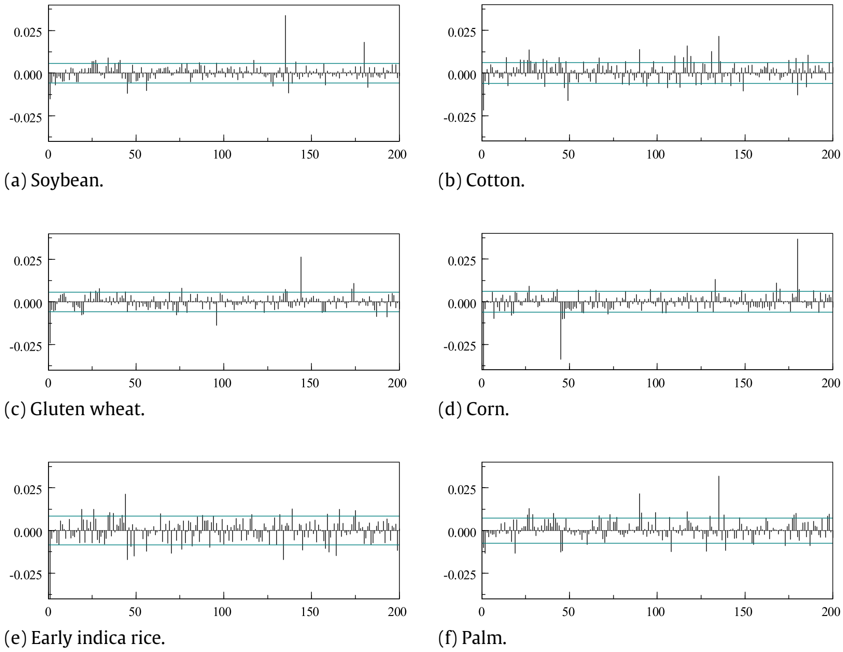
based on numbers of contracts traded and/or cleared. In 2014, 10 of the top 20 agricultural contracts by volume were traded on these China-based exchanges. The Dalian Commodity Exchange is now the second-largest soybean futures market in the world after CBOT, with a trading volume nearly four times greater than that of the Tokyo Commodity Exchange.

Futures trading in China is subject to two unique regulations: (i) the time-dependent margin rate for deposit, and (ii) limited trade on the nearby month contracts. Regarding the former, Peck (2008) presents the implication of this regulation, documenting an inverse U-shape trading volume in Chinese futures markets. For the latter, individual investors' positions are closed automatically by the clearing houses on the last trading day before a contract enters the delivery month. Both regulations were effective in the 1990s, with the aim of reducing and removing market manipulation and over-speculation for contracts that are close to delivery. Meanwhile, Chinese commodity futures markets are characterized by a lack of liquidity, as is evidenced by the large proportion of zero returns in the intraday data (Bekaert, Harvey, & Lundblad, 2007).

3.2. Data and descriptive statistics

We employ intraday data for the six agricultural commodity futures contracts collected from Wind Financial Terminal in Chinese markets, namely soybean, cotton, gluten wheat, corn, early indica rice and palm futures. Soybeans, corn and palm are traded in the Dalian Commodity Exchange, and cotton, gluten wheat and early indica rice are traded in the Zhengzhou Commodity Exchange. All sample periods end on 30 June 2014, and start on 1 August 2003 for soybean and gluten wheat futures, 28 May 2004 for cotton futures, 22 September 2004 for corn futures, 20 April 2009 for early indica rice futures, and 30 October 2007 for palm futures. Following Bollerslev, Litvinova, and Tauchen (2006), Corsi et al. (2008), and Degiannakis (2008), we use a 5 min sampling frequency to derive the 5 min intraday log returns for all futures, then employ these 5 min intraday log returns to obtain the realized volatility estimator proposed by Hansen and Lunde (2006, Eq. (2)) and the potential predictors described in Section 2.2. The 5 min intraday log returns are also used to estimate the daily returns (open to close prices), i.e., $r_t = \sum_{j=1}^{n} r_{j,t}$. In Chinese commodity futures markets, there are four hours of trading during a business day for all commodity futures, starting at 9:00 am and closing at 3:00 pm, with a two-hour break between 11:30 am and 1:30 pm.

Table 1 presents the means of the liquidity measures for the futures contracts with different times to delivery, at 5 min intervals. We describe the liquidities of six agricultural commodity futures contracts in Chinese markets by employing the effective spread of Roll (1984), the proportion of zero returns of Lesmond, Ogden, and Trzcinka (1999), and the illiquidity estimator of Amihud (2002), which are used widely in the literature. By comparing the values, we find that the liquidities for the futures contracts with three months to delivery are among the highest, while those for the contracts with longer or shorter times to maturity are lower. However, the shortest contracts, which are the active contracts with the delivery month closest to the trading day, show significant improvements in liquidity relative to those with three months to delivery. As the shortest contracts are the most liquid ones, we follow Yang, Bessler, and Leatham (2001) and select the futures prices from these. The total numbers of trading days are 2636 for soybean futures, 2439 for cotton futures, 2638 for gluten wheat futures, 2370 for corn futures, 1261 for early indica rice futures, and 1606 for palm futures, once we exclude weekends, public holidays, and days with incomplete data records. Fig. 1 presents the sample autocorrelation of the intraday returns for the six agricultural commodity futures. The intraday returns for all agricultural commodity futures are significantly autocorrelated (negatively) at a first order lag, and several high-order autocorrelations are also significant, implying that the patterns of autocorrelations in the intraday returns are richer than with an MA(1) structure.



(a) Soybean. (b) Cotton.

(c) Gluten wheat. (d) Corn.

(e) Early indica rice. (f) Palm.

Fig. 1. Autocorrelations for intraday returns.

This result indicates that the market microstructure noise may be a dependent type that can induce higher order dependence in the intraday returns. Therefore, we use a bias-correcting approach such as the kernel-based estimator to control such autocorrelation in the intraday returns.

Table 2 reports the descriptive statistics for the daily squared return, unadjusted RV and adjusted RV for $q = 1, 2$ and 3. Although the average daily squared return is a natural measure of the mean of the latent volatility, its large standard deviation makes it extremely noisy. For gluten wheat, corn, early indica rice and palm futures, the average difference between the mean of daily squared returns and the unadjusted RV is fairly large, while the adjusted RV provides an improvement, implying the existence of significant market structure effects. Therefore, we use the adjusted RV with $q = 2$ for gluten wheat and corn futures, the adjusted RV with $q = 3$ for early indica rice and palm futures, and the unadjusted RV for soybean and cotton futures. Fig. 2 displays the time series of $\log(\mathrm{RV}_t)$ for the six agricultural commodity futures.

3.3. Preliminary analysis

This subsection provides a brief discussion of the out-of-sample forecast powers of different potential predictors.

Taking soybean, corn and early indica rice futures as examples, Table 3 reports the values of the QLIKE loss function that are implied by a Gaussian likelihood for out-of-sample forecasting from the HAR model for these three futures. We use five prediction horizons, $h = 1, 5, 10, 15$ and 20, corresponding to one day, one week, two weeks, three weeks and one month, respectively. All of the HAR models have the common dependent variable of $\log(\mathrm{RV}_t)$, but different regressors, including RV, TBPV, daily range, TMPV with orders 0.5, 1 and 1.5, TC, and TJ. We find that no one specification dominates across markets or forecast horizons for the first seven predictors, and the forecast performance rankings for these predictors vary dramatically, depending on the data series and forecast horizon. For example, the daily range has the best performance of all of the predictors for soybean and early indica rice futures when $h = 1$, but only ranks fifth for corn futures at this forecast horizon, and second at the longer forecast horizons in most cases. Furthermore, although jumps have the worst performance in most cases, they perform better than TMPV with order 0.5 for early indica rice futures. Obviously, there is a model risk associated with selecting any one specification, and we therefore avoid this risk by incorporating all potential predictors into a time-varying HAR model specification.

4. Out-of-sample forecast performances

This section reports the out-of-sample forecast results, leaving the full sample estimation results using the proposed HAR-TVS model for the six agricultural commodity futures to be presented in Appendix A.

4.1. Competing models

We compare the out-of-sample point forecast performances of the proposed HAR-TVS model with those of the HAR-type models that are encompassed by the HAR-TVS model, the dynamic model averaging (DMA) of the logarithmic time-varying HAR model of Grassi and de Magistris (2015), labeled DMA-TV-HAR, and the Bayesian model averaging (BMA) of the logarithmic HAR model proposed by Liu and Maheu (2009), labeled BMA-HAR. The nested HAR-type models include the logarithmic version of the HAR-RV

Table 2
Summary statistics of the volatility measures.

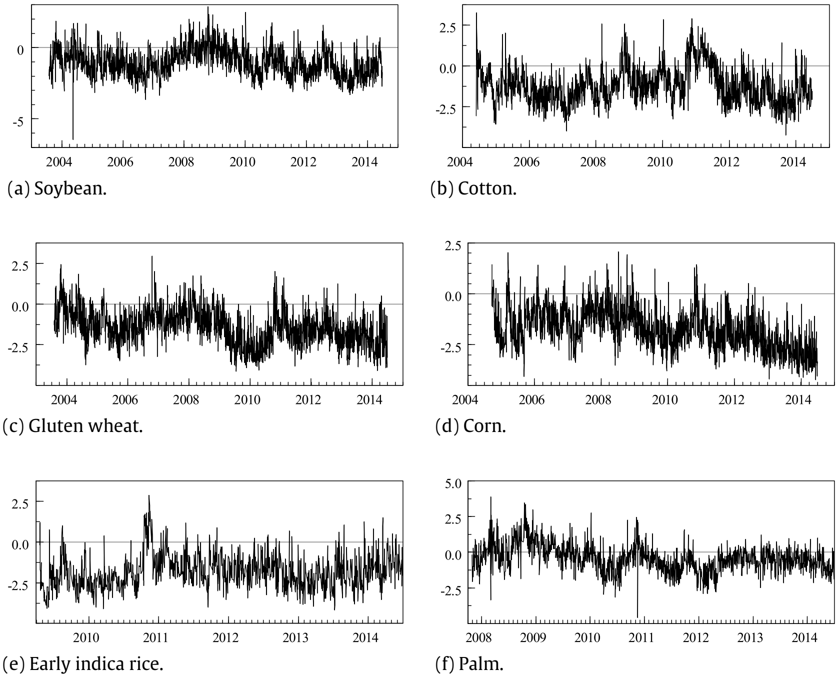
	Measures	Mean	Std.dev	Min.	Max.
Soybean	r^2	0.5973	1.3580	0.0000	23.2150
	RV	0.5827	0.8593	0.0016	17.3780
	RV^1	0.5645	0.8681	0.0008	17.5030
	RV^2	0.5591	0.8789	0.0005	16.4690
	RV^3	0.5584	0.9033	0.0004	16.0980
Cotton	r^2	0.6263	1.9008	0.0000	33.8870
	RV	0.6243	1.3601	0.0143	26.1830
	RV^1	0.5995	1.3421	0.0107	30.3350
	RV^2	0.5907	1.3705	0.0071	32.1290
	RV^3	0.5871	1.3745	0.0061	29.3750
Gluten wheat	r^2	0.4262	1.1394	0.0000	17.5250
	RV	0.4565	0.7419	0.0269	18.7310
	RV^1	0.4354	0.7541	0.0222	18.7310
	RV^2	0.4264	0.7648	0.0161	18.7140
	RV^3	0.4217	0.7776	0.0148	18.5410
Corn	r^2	0.3080	0.7915	0.0000	11.9560
	RV	0.3574	0.4878	0.0258	8.6574
	RV^1	0.3249	0.4944	0.0176	8.5041
	RV^2	0.3105	0.4973	0.0146	7.9281
	RV^3	0.3040	0.5100	0.0123	8.0403
Early indica rice	r^2	0.3185	0.9958	0.0000	15.9330
	RV	0.3848	0.8327	0.0264	13.7200
	RV^1	0.3601	0.8118	0.0191	14.2820
	RV^2	0.3481	0.8182	0.0166	15.6980
	RV^3	0.3432	0.8471	0.0154	17.6010
Palm	r^2	1.1087	2.4298	0.0000	38.8140
	RV	1.1690	1.9143	0.0249	33.8200
	RV^1	1.1483	2.1066	0.0173	47.2130
	RV^2	1.1281	2.1517	0.0101	50.3190
	RV^3	1.1231	2.1858	0.0103	49.0090

Note: r^2 is the daily squared return; RV is the unadjusted realized volatility; RV^1, RV^2 and RV^3 are adjusted RV^q for $q = 1, 2$ and 3 respectively, based on Eq. (2).

Table 3
Out-of-sample forecast powers of different potential predictors, $\log(RV_{t,h})$.

	Regressors	$h = 1$	$h = 5$	$h = 10$	$h = 15$	$h = 20$
Soybean	RV	0.2670 (2)	0.1662 (3)	0.1493 (5)	0.1397 (4)	0.1329 (4)
	TBPV	0.2783 (5)	0.1727 (6)	0.1519 (6)	0.1430 (6)	0.1368 (6)
	Range	0.2512 (1)	0.1596 (1)	0.1454 (2)	0.1366 (2)	0.1300 (1)
	$TMPV^{[0.5]}$	0.2802 (7)	0.1780 (7)	0.1577 (7)	0.1500 (7)	0.1436 (7)
	$TMPV^{[1]}$	0.2721 (4)	0.1691 (5)	0.1486 (4)	0.1400 (5)	0.1337 (5)
	$TMPV^{[1.5]}$	0.2719 (3)	0.1681 (4)	0.1473 (3)	0.1385 (3)	0.1324 (3)
	TC	0.2684 (6)	0.1637 (2)	0.1431 (1)	0.1357 (1)	0.1312 (2)
	TJ	0.3679 (8)	0.2486 (8)	0.2167 (8)	0.1974 (8)	0.1825 (8)
Corn	RV	0.2878 (2)	0.1580 (6)	0.1387 (6)	0.1359 (6)	0.1372 (6)
	TBPV	0.2931 (7)	0.1585 (7)	0.1390 (7)	0.1370 (7)	0.1376 (7)
	Range	0.2898 (5)	0.1355 (2)	0.1114 (2)	0.1044 (2)	0.1019 (2)
	$TMPV^{[0.5]}$	0.2878 (2)	0.1401 (3)	0.1144 (3)	0.1087 (3)	0.1051 (3)
	$TMPV^{[1]}$	0.2868 (1)	0.1343 (1)	0.1085 (1)	0.1030 (1)	0.1004 (1)
	$TMPV^{[1.5]}$	0.2885 (4)	0.1410 (4)	0.1175 (4)	0.1133 (4)	0.1122 (4)
	TC	0.2924 (6)	0.1533 (5)	0.1335 (5)	0.1314 (5)	0.1324 (5)
	TJ	0.3758 (8)	0.2710 (8)	0.2537 (8)	0.2490 (8)	0.2488 (8)
Early indica rice	RV	0.6036 (2)	0.2822 (2)	0.2017 (1)	0.1621 (1)	0.1456 (1)
	TBPV	0.6660 (6)	0.3058 (5)	0.2172 (5)	0.1752 (5)	0.1571 (5)
	Range	0.5961 (1)	0.2790 (1)	0.2041 (2)	0.1637 (2)	0.1485 (2)
	$TMPV^{[0.5]}$	0.6826 (8)	0.3392 (8)	0.2567 (8)	0.2130 (8)	0.1968 (8)
	$TMPV^{[1]}$	0.6541 (4)	0.3078 (6)	0.2248 (6)	0.1833 (6)	0.1688 (7)
	$TMPV^{[1.5]}$	0.6474 (3)	0.2968 (3)	0.2122 (3)	0.1718 (3)	0.1558 (4)
	TC	0.6594 (5)	0.3006 (4)	0.2130 (4)	0.1721 (4)	0.1548 (3)
	TJ	0.6744 (7)	0.3159 (7)	0.2314 (7)	0.1901 (7)	0.1685 (6)

Note: This table compares the out-of-sample forecast performances of different regressors using the soybean, corn and early indica rice futures data. The out-of sample period for all futures is from 4 January 2011 to 30 June 2014. The common model is the logarithmic version of the HAR-X model (for details, see Table 4), with X being RV, TBPV, daily range, $TMPV^{[0.5]}$, $TMPV^{[1]}$, $TMPV^{[1.5]}$, TC or TJ + 1. The values in brackets are the rankings of the forecast performances for these predictors.



(a) Soybean. (b) Cotton. (c) Gluten wheat. (d) Corn. (e) Early indica rice. (f) Palm.

Fig. 2. Time series of daily log-RV for the six agricultural commodity futures.

Table 4
The nested HAR-type models.

HAR-RV model of Corsi (2009)	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{RV}_{t-1,1}) + \alpha_2 \log(\mathrm{RV}_{t-5,5}) + \alpha_3 \log(\mathrm{RV}_{t-20,20}) + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-X models based on Liu and Maheu (2009) and Tseng-Chan et al. (2009)	
HAR-TBPV	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{TBPV}_{t-1,1}) + \alpha_2 \log(\mathrm{TBPV}_{t-5,5}) + \alpha_3 \log(\mathrm{TBPV}_{t-20,20}) + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-Range	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{Range}_{t-1,1}) + \alpha_2 \log(\mathrm{Range}_{t-5,5}) + \alpha_3 \log(\mathrm{Range}_{t-20,20}) + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-TMPV$^{[0.5]}$	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{TMPV}^{[0.5]}_{t-1,1}) + \alpha_2 \log(\mathrm{TMPV}^{[0.5]}_{t-5,5}) + \alpha_3 \log(\mathrm{TMPV}^{[0.5]}_{t-20,20}) + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-TMPV$^{[1]}$	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{TMPV}^{[1]}_{t-1,1}) + \alpha_2 \log(\mathrm{TMPV}^{[1]}_{t-5,5}) + \alpha_3 \log(\mathrm{TMPV}^{[1]}_{t-20,20}) + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-TMPV$^{[1.5]}$	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{TMPV}^{[1.5]}_{t-1,1}) + \alpha_2 \log(\mathrm{TMPV}^{[1.5]}_{t-5,5}) + \alpha_3 \log(\mathrm{TMPV}^{[1.5]}_{t-20,20}) + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-TC	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{TC}_{t-1,1}) + \alpha_2 \log(\mathrm{TC}_{t-5,5}) + \alpha_3 \log(\mathrm{TC}_{t-20,20}) + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-X-J models based on Andersen et al. (2007)	
HAR-RV-J	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{RV}_{t-1,1}) + \alpha_2 \log(\mathrm{RV}_{t-5,5}) + \alpha_3 \log(\mathrm{RV}_{t-20,20}) + \alpha_J TJ_{t-1} + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-TBPV-J	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{TBPV}_{t-1,1}) + \alpha_2 \log(\mathrm{TBPV}_{t-5,5}) + \alpha_3 \log(\mathrm{TBPV}_{t-20,20}) + \alpha_J TJ_{t-1} + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-Range-J	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{Range}_{t-1,1}) + \alpha_2 \log(\mathrm{Range}_{t-5,5}) + \alpha_3 \log(\mathrm{Range}_{t-20,20}) + \alpha_J TJ_{t-1} + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-TMPV$^{[0.5]}$-J	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{TMPV}^{[0.5]}_{t-1,1}) + \alpha_2 \log(\mathrm{TMPV}^{[0.5]}_{t-5,5}) + \alpha_3 \log(\mathrm{TMPV}^{[0.5]}_{t-20,20}) + \alpha_J TJ_{t-1} + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-TMPV$^{[1]}$-J	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{TMPV}^{[1]}_{t-1,1}) + \alpha_2 \log(\mathrm{TMPV}^{[1]}_{t-5,5}) + \alpha_3 \log(\mathrm{TMPV}^{[1]}_{t-20,20}) + \alpha_J TJ_{t-1} + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-TMPV$^{[1.5]}$-J	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{TMPV}^{[1.5]}_{t-1,1}) + \alpha_2 \log(\mathrm{TMPV}^{[1.5]}_{t-5,5}) + \alpha_3 \log(\mathrm{TMPV}^{[1.5]}_{t-20,20}) + \alpha_J TJ_{t-1} + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-TC-J	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{TC}_{t-1,1}) + \alpha_2 \log(\mathrm{TC}_{t-5,5}) + \alpha_3 \log(\mathrm{TC}_{t-20,20}) + \alpha_J TJ_{t-1} + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$
HAR-CJ model based on Andersen et al. (2007)	$\log(\mathrm{RV}_{t,h}) = \alpha_0 + \alpha_1 \log(\mathrm{TC}_{t-1,1}) + \alpha_2 \log(\mathrm{TC}_{t-5,5}) + \alpha_3 \log(\mathrm{TC}_{t-20,20}) + \alpha_{J_1} \log(TJ_{t-1,1} + 1) + \alpha_{J_2} \log(TJ_{t-5,5} + 1) + \alpha_{J_3} \log(TJ_{t-20,20} + 1) + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_u^2)$

model of Corsi (2009), the logarithmic version of the HAR-X model based on Liu and Maheu (2009) and Tseng-Chan et al. (2009), the logarithmic version of the HAR-X-J model, and the HAR-CJ model of Andersen et al. (2007), which are all summarized in Table 4.

The parameters of the TV-HAR model of Grassi and de Magistris (2015) are interpreted as time-varying weights for each component, and the model is defined as

$$
\begin{aligned}
&\log(\mathrm{RV}_{t,h}) = \alpha_{0,t} + \alpha_{1,t} \log(\mathrm{X}_{t-1,1}) + \alpha_{2,t} \log(\mathrm{X}_{t-5,5}) \\
&\qquad + \alpha_{3,t} \log(\mathrm{X}_{t-20,20}) + u_{t,h}, u_{t,h} \sim \mathrm{N}(0, \sigma_{u,t}^2) \\
&\alpha_{0,t} = \alpha_{0,t-1} + \eta_t^{\alpha_0}, \qquad \alpha_{1,t} = \alpha_{1,t-1} + \eta_t^{\alpha_1}, \\
&\alpha_{2,t} = \alpha_{2,t-1} + \eta_t^{\alpha_2}, \qquad \alpha_{3,t} = \alpha_{3,t-1} + \eta_t^{\alpha_3}.
\end{aligned} \tag{5}
$$

Here, $\eta_t = [\eta_t^{\alpha_0}, \eta_t^{\alpha_1}, \eta_t^{\alpha_2}, \eta_t^{\alpha_3}] \sim N(0, H_t)$, H_t is a 4×4 covariance matrix, and X denotes TBPV, daily range, TMPV$^{[0.5]}$, TMPV$^{[1]}$, TMPV$^{[1.5]}$, TC or TJ + 1. There are seven different single models with different predictors for the RV forecast. Since the model that is most relevant to the forecast can potentially change over time, we adopt the DMA, developed by Raftery, Kárný, and Ettler (2010), for these TV-HAR models with different predictors, which allow different single models to hold at each point in time. Let $L_t \in \{1, 2, \ldots, 7\}$ denote the model that applies at each time period, and $y_t = (\log(\mathrm{RV}_1), \ldots, \log(\mathrm{RV}_t))'$. DMA involves calculating $\Pr(L_t = k \mid y_{t-1})$ for $k = 1, \ldots, 7$ and averaging forecasts across single models using these probabilities, then selecting the single model with the highest value of $\Pr(L_t = k|y_{t-1})$ as the forecast. Therefore, the DMA-TV-HAR model used in this study allows the forecast model and coefficients to change over time simultaneously. The calculation of $\Pr(L_t = k|y_{t-1})$ is detailed by Koop and Korobilis (2012).

For the BMA method, based on Liu and Maheu (2009), we consider two sets of linear models. The first set of models is based on the logarithmic version of the HAR-X-J model proposed by Andersen et al. (2007), as detailed in Table 4, and the second set of models is based on autoregressive-type specifications. A summary of the specifications is provided in Table B.1, in Appendix B. A total of 106 different specifications enter the model averages in this paper. Given the information set $Y_N = \{y_1, \ldots, y_N\}$, inference on any parameter θ in BMA takes the following form:

$$p(\theta|Y_N) = \sum_{k=1}^{K} p(\theta|M_k, Y_N)p(M_k|Y_N), \tag{6}$$

where K stands for the total number of model specifications ($K = 106$ in this paper), $p(\cdot|Y_t)$ denotes the posterior distributions, and $p(\cdot|M_k, Y_t)$ denotes the posterior distributions under the assumption that M_k is the true model. $p(M_k|Y_N)$ is the probability of model M_k given the information set Y_N, which can be obtained using

$$p(M_k|Y_N) = \frac{p(Y_N|M_k)p(M_k)}{\sum_{i=1}^{K} p(Y_N|M_i)p(M_i)}, \tag{7}$$

where $p(M_k)$ is the prior model probability and $p(Y_N|M_k)$ is the marginal likelihood. It is more convenient to work with a period-by-period update of the model probability for the out-of-sample forecasts. Given Y_{N-1}, after observing a new observation y_N, the model probability can be updated as

$$p(M_k|y_N, Y_{N-1}) = \frac{p(y_N|Y_{N-1}, M_k)p(M_k|Y_{N-1})}{\sum_{i=1}^{K} p(y_N|Y_{N-1}, M_i)p(M_i|Y_{N-1})}, \tag{8}$$

where $p(y_N|Y_{N-1}, M_k)$ is the predictive likelihood value for model M_k based on Y_{N-1}, and $p(M_k|Y_{N-1})$ is the last period's model probability.

4.2. Forecast implementation

The forecast is implemented using a fixed rolling window that allows the parameters to be re-estimated on a daily basis. The fitted model is then used to generate h-step-ahead out-of-sample point forecasts of the average log volatility. In particular, we divide the sample period into two periods: an in-sample period and an out-of-sample period. The in-sample period ends on 31 December 2010 for all futures, but begins on different dates for different futures: 1 August 2003 for soybean and gluten wheat futures, 28 May 2004 for cotton futures, 22 September 2004 for corn futures, 20 April 2009 for early indica rice futures, and 30 October 2007 for palm futures; the out-of-sample period extends from 4 January 2011 to 30 June 2014 (a total of 844 trade days) for all futures. Then, we set the fixed window length N to 1792, 1595, 1794, 1526, 417 and 762 for soybean, cotton, gluten wheat, corn, early indica rice and palm futures, respectively. $(y_1, \ldots, y_N)$ are used to estimate each competing model, and each fitted model is then used to construct the first h-step-ahead forecast of the average log volatility. The second h-step-ahead forecast is based on $(y_2, \ldots, y_{N+1})$, and the last on $(y_{T-N-h+1}, \ldots, y_{T-h})$, where T is the size of the full sample. This procedure is repeated, until each model has produced 844, 840 and 825 point forecasts for $h = 1, 5$ and 20, respectively.

The out-of-sample point forecasts from the proposed HAR-TVS model and its nested HAR-type models are based on the posterior predictive mean, which can be obtained using particle filtering methods. We use $s^* = 0.1$ and $b^* = 0.1$ in all cases. The Bayesian MCMC is run for 110,000 iterations, of which 10,000 are discarded as an initial burn-in, which is sufficient to remove the dependence on the initial conditions, then every 25th draw is saved for summary posterior inference. The out-of-sample point forecasts from the DMA-TV-HAR model are obtained using approximations with the modified Kalman filter proposed by Koop and Korobilis (2012) and Raftery et al. (2010), which can lead to a quick real time forecast. Following the standard convention in the BMA, we select the following proper priors: $\alpha \sim N(0, 100I)$ and $\sigma_u^2 \sim IG(0.001/2, 0.001/2)$, and set the model probability to $p(M_k) = 1/K$, $k = 1, \ldots, K$, at observation 500, based on Eklund and Karlsson (2007). Thereafter, the model probability can be updated using Bayes' rules. The out-of-sample point forecasts are based on the posterior predictive mean, which can be obtained using the Gibbs sampler. The first 500 draws are discarded and the next 10,000 are collected for posterior inference and forecast.

4.3. Forecast results and evaluation

We compare the out-of-sample forecast performances of the proposed HAR-TVS model with those of the competing models by utilizing the model confidence set (MCS) approach developed by Hansen, Lunde, and Nason (2011). The MCS test procedure sequentially eliminates the model with worst performance from the full set of models M until the null hypothesis of equal forecast accuracy (EPA) is no longer rejected at the α significance level, and the set of models that survives at that point then forms the MCS. If a fixed significance level α is used at each step, the MCS, $\widehat{M}^*_{1-\alpha}$, contains the best models from M at the $(1-\alpha)$ confidence level. Following Hansen and Lunde (2011), we employ the range statistics $T_R = \max_{i,j\in M} \frac{|\bar{d}_{ij}|}{\sqrt{\widehat{var}(\bar{d}_{ij})}}$ and the semi-quadratic statistics $T_{SQ} = \sum_{i,j\in M} \frac{(\bar{d}_{ij})^2}{\widehat{var}(\bar{d}_{ij})}$ to test the null hypothesis of EPA, where $d_{ij,t}$ is the loss differential between two models in M, $\bar{d}_{ij} = \frac{1}{T_h}\sum_{t=N}^{T-h} d_{ij,t}$, and $\widehat{var}(\bar{d}_{ij})$ is an estimate of $\mathrm{var}(\bar{d}_{ij})$ that is obtained by using the stationary block bootstrap. The MCS test procedure assigns p-values to each model in the initial set. For a given model $i \in M$, the MCS p-value, $\hat{p}_i$, is the threshold confidence level that determines whether or not the model belongs to the MCS. It holds that $i \in \widehat{M}^*_{1-\alpha}$ only if $\hat{p}_i \geq \alpha$.

Table 5
MCS results of the point forecasts: 1- to 20-day-ahead forecasts, $\log(\mathrm{RV}_{t,h})$.

	$h=1$			$h=5$			$h=20$		
	MCS	T_R	T_{SQ}	MCS	T_R	T_{SQ}	MCS	T_R	T_{SQ}
Soybean	HAR-TMPV$^{[1]}$	0.114	0.127	HAR-Range-J	0.160	0.127	HAR-TVS	1.000	1.000
	HAR-RV-J	0.114	0.124	HAR-TVS	1.000	1.000	DMA-TV-HAR	0.193	0.198
	HAR-TVS	1.000	1.000	DMA-TV-HAR	0.161	0.154	BMA-HAR	0.155	0.160
	DMA-TV-HAR	0.380	0.384	BMA-HAR	0.218	0.207			
	BMA-HAR	0.442	0.445						
Cotton	HAR-TVS	1.000	1.000	HAR-TMPV$^{[0.5]}$	0.108	0.112	HAR-Range-J	0.110	0.112
	DMA-TV-HAR	0.128	0.124	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000
	BMA-HAR	0.224	0.230	DMA-TV-HAR	0.120	0.112	DMA-TV-HAR	0.437	0.441
				BMA-HAR	0.137	0.250	BMA-HAR	0.337	0.336
Gluten wheat	HAR-TVS	1.000	1.000	HAR-Range-J	0.215	0.212	HAR-Range-J	0.132	0.138
	DMA-TV-HAR	0.194	0.182	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000
	BMA-HAR	0.172	0.170	DMA-TV-HAR	0.408	0.410	DMA-TV-HAR	0.171	0.168
				BMA-HAR	0.223	0.220	BMA-HAR	0.158	0.155
Corn	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000
	DMA-TV-HAR	0.328	0.319	DMA-TV-HAR	0.175	0.172	DMA-TV-HAR	0.345	0.366
	BMA-HAR	0.270	0.274	BMA-HAR	0.201	0.208	BMA-HAR	0.280	0.283
Early indica rice	HAR-TVS	1.000	1.000	HAR-RV-J	0.106	0.108	HAR-TVS	1.000	1.000
	DMA-TV-HAR	0.511	0.512	HAR-TVS	0.684	0.688	DMA-TV-HAR	0.594	0.596
	BMA-HAR	0.240	0.247	DMA-TV-HAR	1.000	1.000	BMA-HAR	0.411	0.412
				BMA-HAR	0.259	0.287			
Palm	HAR-TMPV$^{[1]}$-J	0.237	0.244	HAR-TVS	0.629	0.627	HAR-TVS	1.000	1.000
	HAR-TVS	1.000	1.000	DMA-TV-HAR	1.000	1.000	DMA-TV-HAR	0.619	0.624
	DMA-TV-HAR	0.613	0.610	BMA-HAR	0.288	0.291	BMA-HAR	0.433	0.427
	BMA-HAR	0.527	0.522						

Notes: The table presents the MCS and the corresponding p-values for our selection of RV models. The p-values based on T_R and T_{SQ} are obtained from 10,000 block bootstraps. Models with underlines are included in both $\widehat{M}^*_{0.75}$ and $\widehat{M}^*_{0.90}$. Models without underlines are only included in $\widehat{M}^*_{0.90}$. Note that $\widehat{M}^*_{0.75} \subset \widehat{M}^*_{0.90}$. HAR-TVS is the proposed HAR model with time-varying sparsity.

Based on the QLIKE loss function, which belongs to Patton's (2011) family of robust loss functions that are robust to the noise of the volatility proxy, Table 5 presents the MCS as along with the corresponding p-values at the 10% and 25% significance levels. The p-values are obtained from 10,000 block bootstraps. In most cases, the only three models in the MCS are the proposed HAR-TVS model, the DMA-TV-HAR model and the BMA-HAR model, as the p-values are larger than 0.1, indicating that, in general, these three models that allow the regression model to change over time are superior to the other nested HAR-type models. This indicates the importance of allowing time-variation in the relevance of the regression coefficients. Of these three models, we find that the DMA-TV-HAR model belongs to the MCS in eleven cases out of eighteen and the BMA-HAR model belongs to the MCS in only half of the cases, while the HAR-TVS model belongs to the MCS in all cases, with its p-values being bigger than those of either the DMA-TV-HAR or BMA-HAR models in almost all cases, at one or very close to one at the 75% confidence level. This indicates that the proposed HAR-TVS model appears to be the most accurate for forecasting the RV of agricultural commodity futures in Chinese markets. These results do not differ significantly across futures or forecast horizons. The main reason for the proposed HAR-TVS model delivering the best forecast performance may be due to the fact that the proposed HAR-TVS model not only overcomes the model risk associated with selecting any individual HAR-type models, but also allows both the regression coefficients and the sparsity to vary over time, by giving NGAR process priors to the regression coefficients. In addition, there are two reasons why the BMA-HAR model does worse than the HAR-TVS model and the DMA-TV-HAR model. First, the BMA method is built on the Bayesian variable selection techniques, which consider all possible regression models explicitly and construct forecasts by weighted averages over all individual models, but the proposed BMA-HAR model here only considers two sets of linear models, including 106 different specifications. The computational demands can become daunting if we consider all possible models. Second, while some model specifications may contribute little to the forecast, these models are still included when using the BMA method, in spite of their poor forecast performances.

It is worth noting that, in a few cases, some nested HAR-type models are also included in the MCS, with a 90% confidence level. These models are dominated largely by the HAR models with the jump component as a potential predictor, since seven of the nine models have the jump terms. For example, the HAR-RV-J model belongs to the MCS at the 1-day horizon for soybean futures, and at the 5-day horizon for early indica rice futures. The HAR-Range-J model belongs to the MCS at the 5-day horizon for soybean and gluten wheat futures, and at the 20-day horizon for cotton and gluten wheat futures. These results indicate the importance of taking into account the jump component when forecasting the RV of agricultural commodity futures in Chinese markets.

4.4. Robustness check with alternative proxies of volatility

Following Fiszeder and Perczak (2016) and Wang, Wu, and Xu (2015), we employ two alternative proxies for the daily volatility in order to check the robustness of our forecast results: the kernel-based estimator based on a denser sampling frequency (1 min) and the non-flat-top realized kernel estimator. The latter, proposed by Barndorff-Nielsen et al. (2009), is defined as:

$$\mathrm{RV}_t^{FTRK} = \sum_{j=1}^{n} r_{j,t}^2 + \sum_{h=1}^{H} K\left(\frac{h}{H}\right)(\widehat{\gamma}_h + \widehat{\gamma}_{-h}), \tag{9}$$

where $\widehat{\gamma}_h = \frac{n}{n-h}\sum_{j=1}^{n} r_{j,t} r_{j-h,t}$, and $K(0) = 1, K(1) = 0$. Obviously, the right-hand side of Eq. (9) is the realized kernel correction of the market microstructure noise. The kernel-based estimator of Hansen and Lunde (2006) is a special case of the flat-top realized kernel estimator with general H, and $K(x) = 1$. Following Barndorff-Nielsen et al. (2009), we use the non-flat-top realized kernel estimator with the Parzen kernel, the optimal value of the bandwidth $H^* = c^*\xi^{4/5}n^{3/5}$, and $c^* = 3.5134$. The MCS results obtained by all forecast models using these two proxies for the daily volatility are reported in Tables C.1 and C.2 respectively. The insights yielded by these results do not contradict the findings discussed above. Therefore, the forecast performance of the time-varying HAR model with an NGAR process prior for the regression coefficients is robust to alternative proxies for the volatility.

5. Conclusion

We capture the time-varying property of the regression coefficients and the predictors for the RV forecast simultaneously by incorporating all potential predictors into the framework of a time-varying HAR model so as to avoid the model risk associated with selecting any one specification, and utilize the independent normal-gamma autoregressive processes priors that allow both the regression coefficients and the predictors to change over time. We then estimate the proposed model using the MCMC method, and produce out-of-sample volatility forecasts for agricultural commodity futures in China. Finally, we use the MCS test to compare its out-of-sample point forecast performances with those of various competing HAR-type models, including both the simple nested HAR-type models and more sophisticated HAR-type models that have been developed recently at various horizons.

The out-of-sample volatility forecast results show that the jump component is important for forecasting the RV in Chinese agricultural commodity futures markets, and that the DMA-TV-HAR and BMA-HAR models outperform the nested HAR-type models in general. More importantly, the proposed HAR model with time-varying sparsity appears to be the most accurate model for forecasting the RV of agricultural commodity futures of all of the models considered in this study. This is because the proposed HAR-TVS model not only overcomes the model risk associated with selecting a single HAR-type model, but also allows both the predictors and the regression coefficients of the forecast model to vary over time by giving NGAR process priors to the regression coefficients. In addition, the proposed model is robustness to alternative volatility proxies.

Acknowledgments

This research is supported by National Natural Science Foundation of China under Grant No. 71203067 and 71673089, National Social Science Foundation of China under Grant No. 15CJY004 and 14ZDA020, Humanities and Social Sciences Foundation of Ministry of Education of China under Grant Nos. 14YJCZH141 and 14YJA79004, and Philosophy and Social Science Development Foundation of Guangzhou under Grant No. 15Y21.

Appendix A. Full sample estimation results

This section presents the full-sample fit results using the proposed HAR model with time-varying sparsity for six Chinese agricultural commodity futures. This study aims to prove that the proposed model accounts for the time-varying effect of the regression coefficients adequately, and to show which predictors are important for forecasting the realized volatility of agricultural commodity futures over time. We focus on the full-sample fit results for the short forecast horizon ($h = 1$).

As is detailed in Section 2.3, a smaller value of $\psi_{i,t}$ indicates that the ith regressor is less important for predicting the realized volatility at time t. We plot the change in the posterior mean of $\sqrt{\psi_{i,t}}$ over time in order to show the importance of each predictor over time for all agricultural commodity futures. We also plot the change in the posterior mean of $\alpha_{i,t}$ over time in order to show the effect of each relevant predictor over time for all agricultural commodity futures. We expect the value of $\alpha_{i,t}$ to be very close to zero when the posterior mean of $\sqrt{\psi_{i,t}}$ is zero, implying that the predictor is not relevant. We also present 95% credible intervals (CI) for all plots. Since the posterior mean of $\alpha_{i,t}$ in all cases is seldom away from zero, we use the 95% CI as a set of plausible values for the regression coefficients.

The posterior mean and 95% CI of both $\sqrt{\psi_{i,t}}$ and $\alpha_{i,t}$ for soybean futures are illustrated in Figs. A.1 and A.2 respectively. Although all predictors are more or less relevant for forecasting the realized volatility of soybean futures, the value of the posterior mean of $\sqrt{\psi_{i,t}}$ for most predictors is never greater than 0.3. Of these twenty-four predictors, the three most relevant predictors of the realized volatility are the daily components of the range ($\mathrm{Range}_{t-1,1}$) and the weekly and the monthly components of TMPV with order 0.5 ($\mathrm{TMPV}_{t-5,5}^{[0.5]}$ and $\mathrm{TMPV}_{t-20,20}^{[0.5]}$). The relevance of these three predictors is fairly constant over time (around 0.5), but the values of their coefficients are not constant, as they fluctuate considerably. Taking $\mathrm{TMPV}_{t-20,20}^{[0.5]}$ as an example, its 95% CI band shows a downward trend between 2004 and 2007, followed by an upward trend until the end of 2008. We also find the relevance of the weekly components of jump ($\mathrm{TJ}_{t-5,5}$) to be obvious at the end of 2004 (jumps to around 0.35), the middle of 2009 (jumps to around 0.35) and the middle of 2013 (jumps to around 0.4). Its coefficient also fluctuates significantly during these periods, implying that the jump component is very important for forecasting the realized volatility of soybean futures for some periods.

The posterior mean and 95% CI of both $\sqrt{\psi_{i,t}}$ and $\alpha_{i,t}$ for cotton futures are displayed in Figs. A.3 and A.4

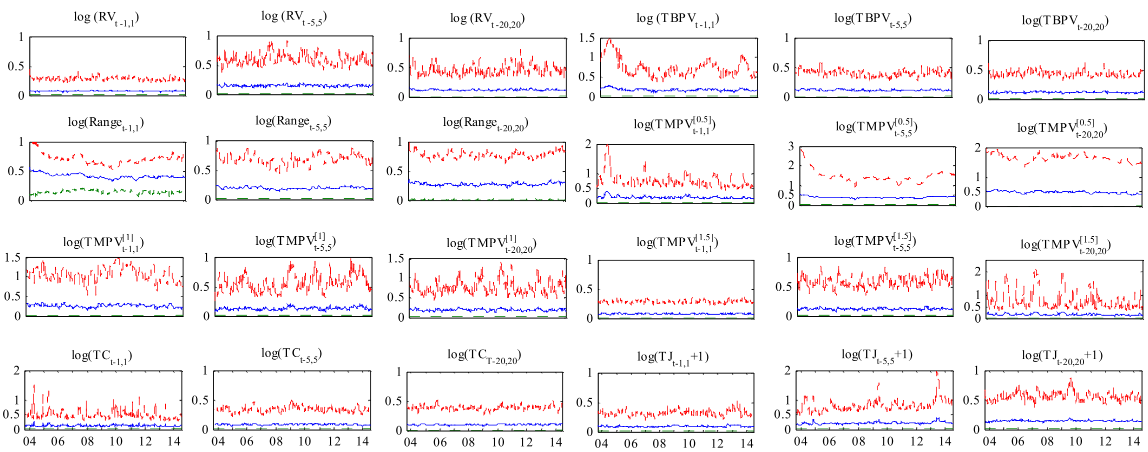


Fig. A.1. Soybean futures: posterior mean (solid) of the time-varying regression relevance, $\sqrt{\psi_{i,t}}$, with 95% CI (dotted).

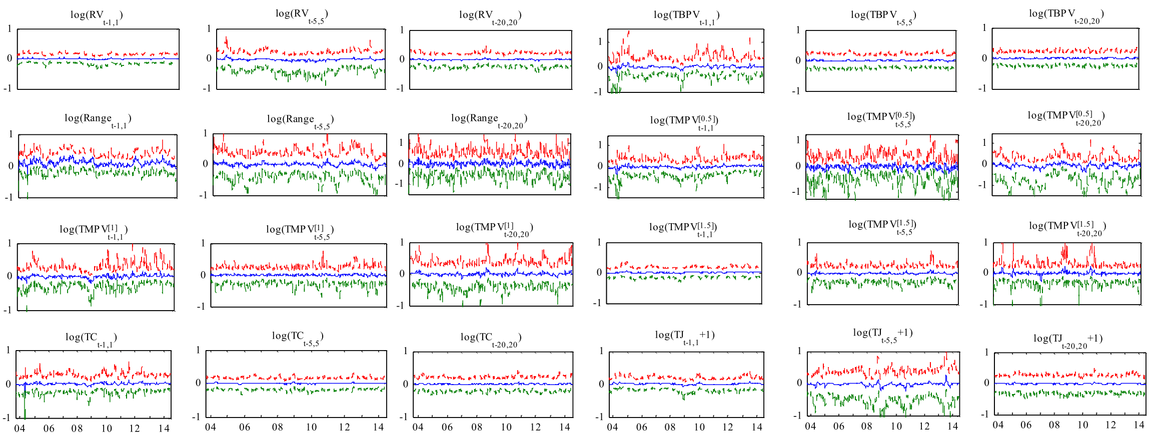


Fig. A.2. Soybean futures: posterior mean (solid) of the time-varying regression coefficients, $\alpha_{i,t}$, with 95% CI (dotted).

respectively, which show that the value of the posterior mean of $\sqrt{\psi_{i,t}}$ for thirteen predictors is very close to zero over time, indicating that these predictors are less important in the realized volatility forecast for cotton futures. These predictors are the daily components of RV ($\mathrm{RV}_{t-1,1}$), the daily, weekly and monthly components of TBPV ($\mathrm{TBPV}_{t-1,1}$, $\mathrm{TBPV}_{t-5,5}$ and $\mathrm{TBPV}_{t-20,20}$), the weekly components of the range ($\mathrm{Range}_{t-5,5}$), the daily and weekly components of TMPV of order 1 ($\mathrm{TMPV}^{[1]}_{t-5,5}$ and $\mathrm{TMPV}^{[1]}_{t-20,20}$), the daily, weekly and monthly components of TMPV of order 1.5 ($\mathrm{TMPV}^{[1.5]}_{t-1,1}$, $\mathrm{TMPV}^{[1.5]}_{t-5,5}$ and $\mathrm{TMPV}^{[1.5]}_{t-20,20}$), the weekly and monthly components of TC ($\mathrm{TC}_{t-5,5}$ and $\mathrm{TC}_{t-20,20}$), and the daily components of jump ($\mathrm{TJ}_{t-1,1}$). Of the eleven remaining predictors, the three most relevant predictors of the realized volatility are the daily components of the range ($\mathrm{Range}_{t-1,1}$), and the weekly and monthly components of TMPV of order 0.5 ($\mathrm{TMPV}^{[0.5]}_{t-5,5}$ and $\mathrm{TMPV}^{[0.5]}_{t-20,20}$). The value of the posterior mean of $\sqrt{\psi_{i,t}}$ for $\mathrm{Range}_{t-1,1}$ is higher at the end of 2004, then falls to around 0.3 and remains constant thereafter. Its coefficient is always positive, with higher values at the beginning of 2007 and the middle of 2008. The importance of the daily components of TMPV of order 0.5 decreases slightly between 2005 and 2008, then stabilises. Its coefficient is always negative, with the lowest value being in the middle of 2006. The relevance of the monthly components of TMPV of order 0.5 is fairly constant, but the values of its coefficient vary. We also find that the weekly and monthly components of jumps ($\mathrm{TJ}_{t-5,5}$ and $\mathrm{TJ}_{t-20,20}$) are more relevant than the weekly and monthly components of the realized volatility itself.

Figs. A.5 and A.6 show the posterior mean and 95% CI of $\sqrt{\psi_{i,t}}$ and $\alpha_{i,t}$, respectively, for gluten wheat futures. Similarly to soybean futures, although all predictors are more or less relevant for forecasting the realized volatility of gluten wheat futures, the value of the posterior mean of $\sqrt{\psi_{i,t}}$ for most predictors is never greater than 0.3 over time. The three most relevant predictors for the realized volatility forecast are the monthly components of the range ($\mathrm{Range}_{t-20,20}$), the monthly components of TMPV of order 0.5 ($\mathrm{TMPV}^{[0.5]}_{t-20,20}$), and the monthly components of jumps ($\mathrm{TJ}_{t-20,20}$). The relevance of $\mathrm{Range}_{t-20,20}$ is higher at the end of 2003, then decreases until the middle of 2005, and stabilises thereafter. Overall, it has a negative coefficient, and reaches its lowest value at the beginning of 2010. The relevance of $\mathrm{TMPV}^{[0.5]}_{t-20,20}$ has a downward trend over time, and the value of its coefficients is generally negative, with the lowest value being at the beginning of

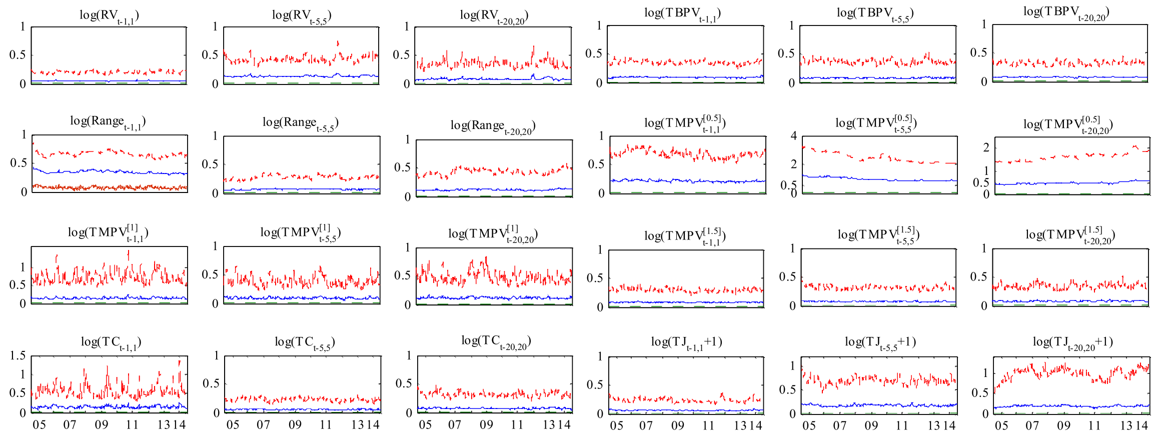


Fig. A.3. Cotton futures: posterior mean (solid) of the time-varying regression relevance, $\sqrt{\psi_{i,t}}$, with 95% CI (dotted).

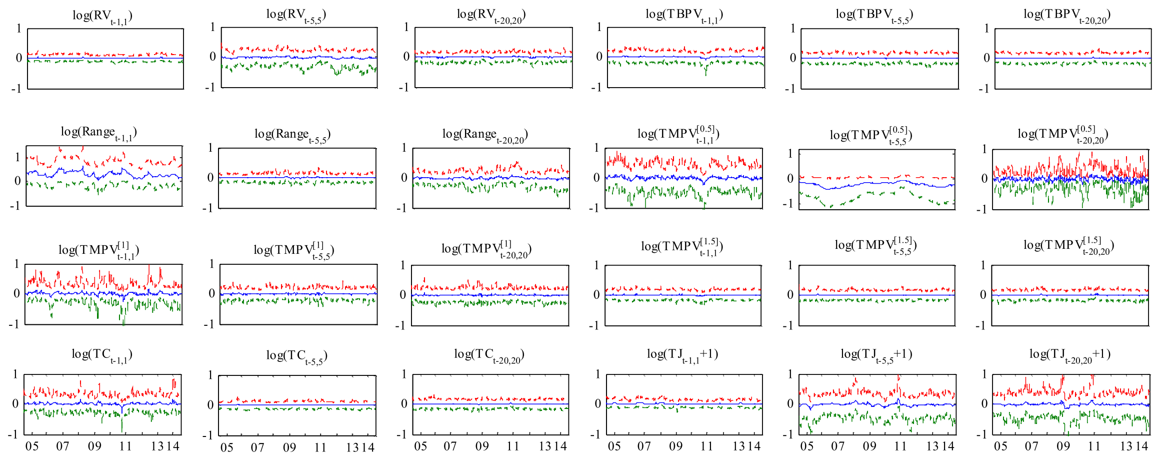


Fig. A.4. Cotton futures: posterior mean (solid) of the time-varying regression coefficients, $\alpha_{i,t}$, with 95% CI (dotted).

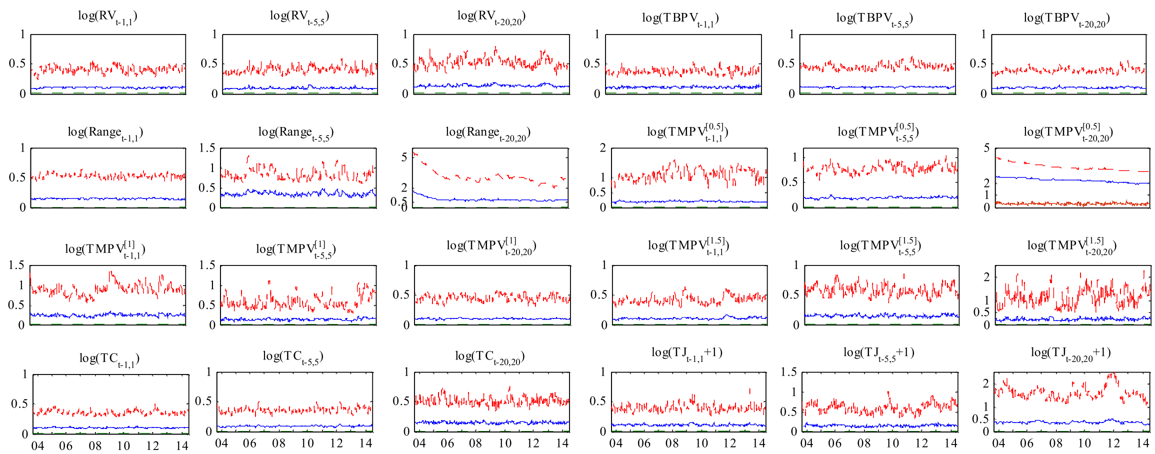


Fig. A.5. Gluten wheat futures: posterior mean (solid) of the time-varying regression relevance, $\sqrt{\psi_{i,t}}$, with 95% CI (dotted).

2010. The relevance of $\mathrm{TJ}_{t-20,20}$ is fairly constant over time (around 0.5), but the value of its coefficients is not constant, as they fluctuate, with a large jump at the beginning of 2007 (financial crisis).

Figs. A.7 and A.8 display the posterior mean and 95% CI of $\sqrt{\psi_{i,t}}$ and $\alpha_{i,t}$, respectively, for corn futures. The three most relevant predictors for the realized volatility prediction of corn futures are the daily components of the range ($\mathrm{Range}_{t-1,1}$) and the daily and monthly components of TMPV of order 0.5 ($\mathrm{TMPV}^{[0.5]}_{t-1,1}$ and $\mathrm{TMPV}^{[0.5]}_{t-20,20}$). The relevance of $\mathrm{Range}_{t-1,1}$ decreases slightly at the beginning, and stabilises thereafter. Overall, it has a

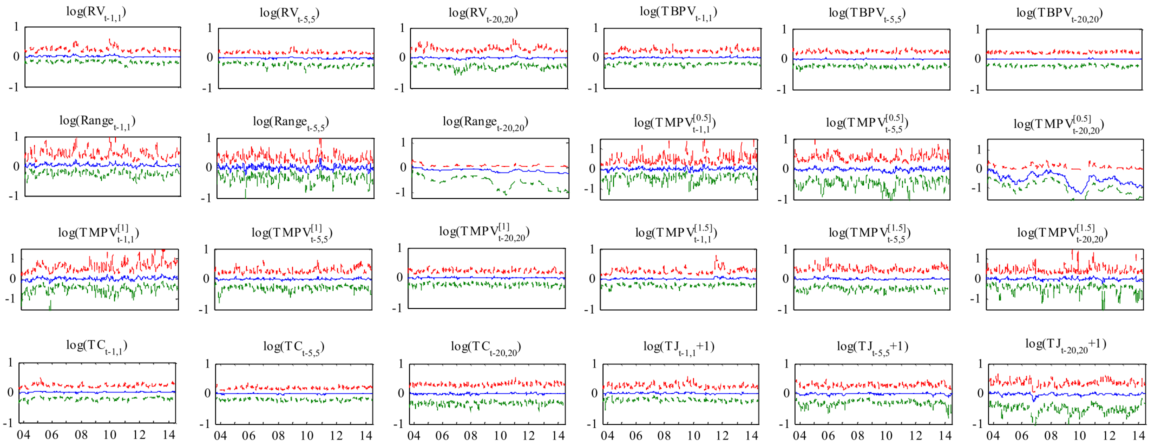


Fig. A.6. Gluten wheat futures: posterior mean (solid) of the time-varying regression coefficients, $\alpha_{i,t}$, with 95% CI (dotted).

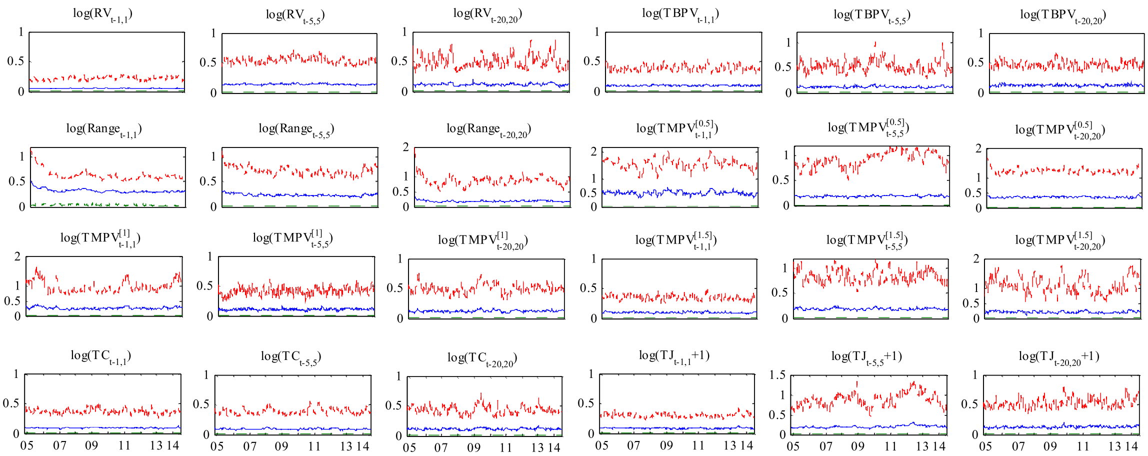


Fig. A.7. Corn futures: posterior mean (solid) of the time-varying regression relevance, $\sqrt{\psi_{i,t}}$, with 95% CI (dotted).

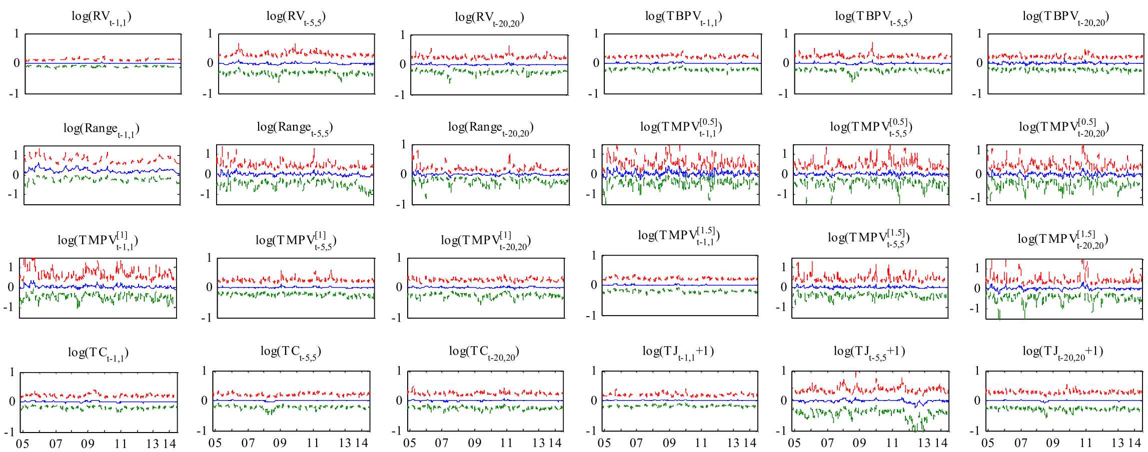


Fig. A.8. Corn futures: posterior mean (solid) of the time-varying regression coefficients, $\alpha_{i,t}$, with 95% CI (dotted).

positive coefficient with significant fluctuations. The relevance of $\mathrm{TMPV}^{[0.5]}_{t-1,1}$ fluctuates over time, similarly to its coefficients, and the sign of its coefficient turns out to be negative for some periods. The relevance of $\mathrm{TMPV}^{[0.5]}_{t-20,20}$ is fairly constant over time, but the value of its coefficients is not constant, as they fluctuate, with negative value for some periods. We also find that all components of jumps ($\mathrm{TJ}_{t-1,1}$, $\mathrm{TJ}_{t-5,5}$ and $\mathrm{TJ}_{t-20,20}$) are more relevant than the daily components of the realized volatility itself.

The posterior mean and 95% CI of both $\sqrt{\psi_{i,t}}$ and $\alpha_{i,t}$ for early indica rice futures are shown in Figs. A.9 and A.10 respectively. All predictors are more or less relevant

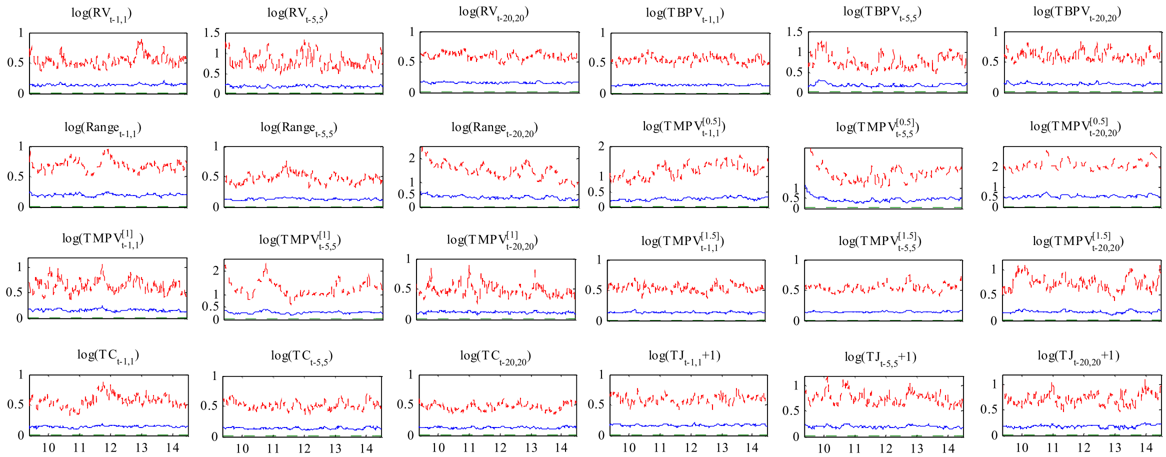


Fig. A.9. Early indica rice futures: posterior mean (solid) of the time-varying regression relevance, $\sqrt{\psi_{i,t}}$, with 95% CI (dotted).

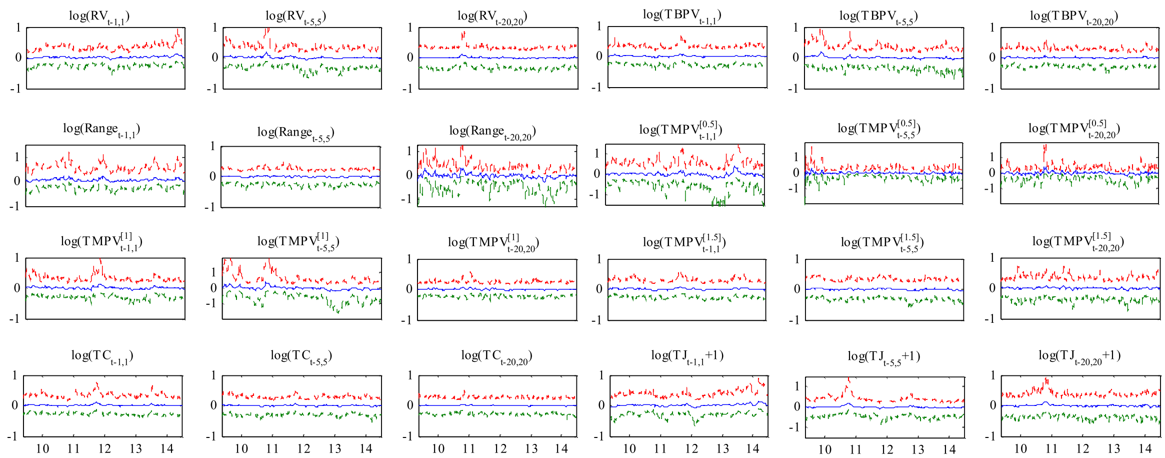


Fig. A.10. Early indica rice futures: posterior mean (solid) of the time-varying regression coefficients, $\alpha_{i,t}$, with 95% CI (dotted).

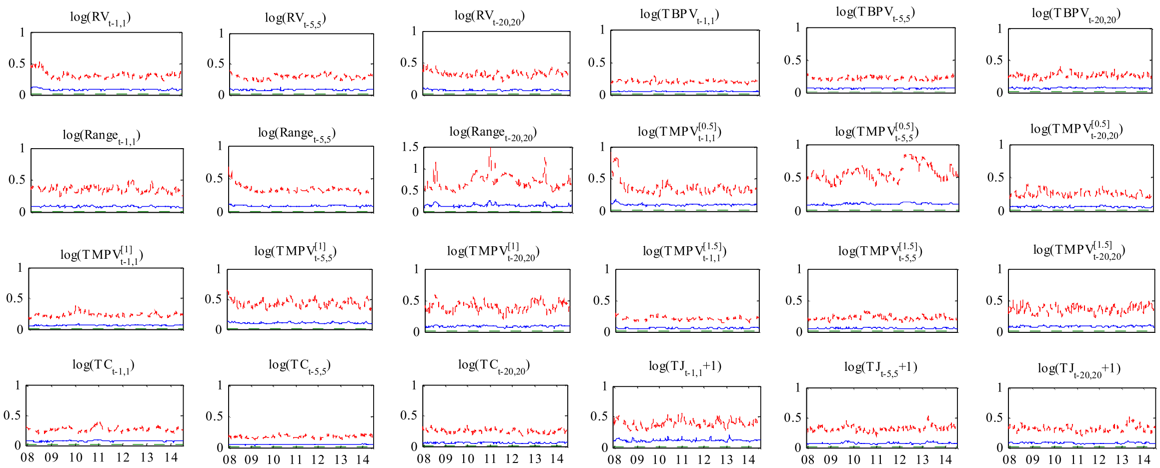


Fig. A.11. Palm futures: posterior mean (solid) of the time-varying regression relevance, $\sqrt{\psi_{i,t}}$, with 95% CI (dotted).

for forecasting the realized volatility of early indica rice futures. The value of the posterior mean of $\sqrt{\psi_{i,t}}$ for most predictors is never greater than 0.3. Of the twenty-four predictors, the ones that are more noticeably relevant (the posterior mean of $\sqrt{\psi_{i,t}}$ is greater than 0.3 in some periods) for forecasting the realized volatility of early indica rice futures are the weekly components of TBPV ($TBPV_{t-5,5}$), the monthly components of the range ($Range_{t-20,20}$),

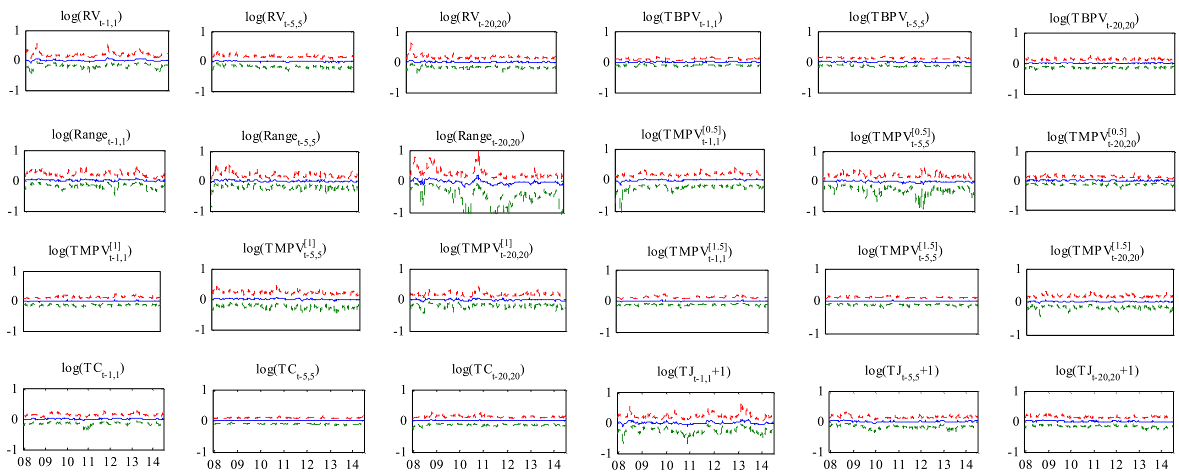


Fig. A.12. Palm futures: posterior mean (solid) of the time-varying regression coefficients, $\alpha_{i,t}$, with 95% CI (dotted).

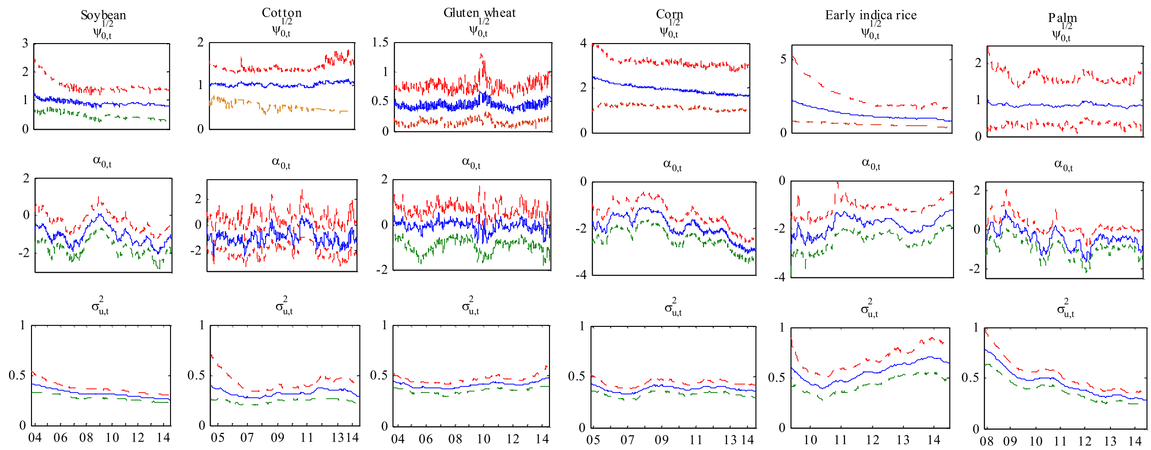


Fig. A.13. The posterior mean (solid line) and 95% CI (dashed line) of the relevance of the intercept $\sqrt{\psi_{0,t}}$, and the posterior mean (solid line) and 95% CI (dotted line) of the intercept $\alpha_{0,t}$, and the posterior mean (solid line) and 95% CI (dotted line) of the time-varying innovation volatility, σ_u^2.

the daily, weekly and monthly components of TMPV of order 0.5 ($\text{TMPV}_{t-1,1}^{[0.5]}$, $\text{TMPV}_{t-5,5}^{[0.5]}$ and $\text{TMPV}_{t-20,20}^{[0.5]}$), and the weekly components of TMPV of order 1.5 ($\text{TMPV}_{t-5,5}^{[1.5]}$). The relevance of $\text{TBPV}_{t-5,5}$ is higher at the end of 2009 than in all other periods, and its coefficient is also higher during this period. The posterior means of $\sqrt{\psi_{i,t}}$ for $\text{Range}_{t-20,20}$, $\text{TMPV}_{t-1,1}^{[0.5]}$, $\text{TMPV}_{t-20,20}^{[0.5]}$ and $\text{TMPV}_{t-5,5}^{[1.5]}$ are fairly constant (around 0.5), but the values of their coefficients are not constant, as they fluctuate over time. For $\text{Range}_{t-20,20}$, the sign of its coefficient turns to be negative for some periods, and reaches its peak value in the middle of 2009 and at the beginning of 2011. The importance of the monthly components of TMPV of order 0.5 decreases slightly at the end of 2009 and stabilises at around 0.5 thereafter, while the value of its coefficient is higher at the end of 2009.

The posterior mean and 95% CI of both $\sqrt{\psi_{i,t}}$ and $\alpha_{i,t}$ for palm futures are illustrated in Figs. A.11 and A.12, respectively. Compared to the other agricultural commodity futures, the relevance of all predictors for forecasting the realized volatility of palm futures is less obvious. The two most relevant predictors for the realized volatility forecast of corn futures are the monthly components of the range ($\text{Range}_{t-20,20}$) and the daily components of jumps ($\text{TJ}_{t-1,1}$).

Fig. A.13 shows the time-varying relevance, $\sqrt{\psi_{0,t}}$, the effect of the intercept, $\alpha_{0,t}$, and the behaviour of the innovation variance, σ_u^2, for all agricultural commodity futures over time. In all cases, the value of $\sqrt{\psi_{0,t}}$ is higher than those of most other predictors, implying that incorporating the intercept is very important when forecasting the realized volatility of agricultural commodity futures, and that the effect of the intercept, $\alpha_{0,t}$, on all agricultural commodity futures varies over time. The innovation variances for cotton, gluten wheat and corn futures are fairly constant over time, at around 0.5, whereas those for soybean and palm futures have a downward trend over time, and those for early indica rice futures decrease between the middle of 2009 and the middle of 2010, then increase until the beginning of 2014.

Appendix B. Model specifications for the Bayesian model averaging method

See Table B.1.

Table B.1
Model specifications.

Panel A: HAR-type

Model	RV	$TMPV^{[0.5]}$	$TMPV^{[1]}$	$TMPV^{[1.5]}$	Range	TBPV	TC
1	3	0	0	0	0	0	0
2	0	3	0	0	0	0	0
3	0	0	3	0	0	0	0
4	0	0	0	3	0	0	0
5	0	0	0	0	3	0	0
6	0	0	0	0	0	3	0
7	0	0	0	0	0	0	3
8	1	1	0	0	0	0	0
9	1	2	0	0	0	0	0
10	1	3	0	0	0	0	0
11	1	0	1	0	0	0	0
12	1	0	2	0	0	0	0
13	1	0	3	0	0	0	0
14	1	0	0	1	0	0	0
15	1	0	0	2	0	0	0
16	1	0	0	3	0	0	0
17	1	0	0	0	1	0	0
18	1	0	0	0	2	0	0
19	1	0	0	0	3	0	0
20	1	0	0	0	0	1	0
21	1	0	0	0	0	2	0
22	1	0	0	0	0	3	0
23	1	0	0	0	0	0	1
24	1	0	0	0	0	0	2
25	1	0	0	0	0	0	3
26	2	1	0	0	0	0	0
27	2	2	0	0	0	0	0
28	2	3	0	0	0	0	0
29	2	0	1	0	0	0	0
30	2	0	2	0	0	0	0
31	2	0	3	0	0	0	0
32	2	0	0	1	0	0	0
33	2	0	0	2	0	0	0
34	2	0	0	3	0	0	0
35	2	0	0	0	1	0	0
36	2	0	0	0	2	0	0
37	2	0	0	0	3	0	0
38	2	0	0	0	0	1	0
39	2	0	0	0	0	2	0
40	2	0	0	0	0	3	0
41	2	0	0	0	0	0	1
42	2	0	0	0	0	0	2
43	2	0	0	0	0	0	3
44	3	1	0	0	0	0	0
45	3	2	0	0	0	0	0
46	3	3	0	0	0	0	0
47	3	0	1	0	0	0	0
48	3	0	2	0	0	0	0
49	3	0	3	0	0	0	0
50	3	0	0	1	0	0	0
51	3	0	0	2	0	0	0
52	3	0	0	3	0	0	0
53	3	0	0	0	1	0	0
54	3	0	0	0	2	0	0
55	3	0	0	0	3	0	0
56	3	0	0	0	0	1	0
57	3	0	0	0	0	2	0
58	3	0	0	0	0	3	0
59	3	0	0	0	0	0	1
60	3	0	0	0	0	0	2
61	3	0	0	0	0	0	3

Panel B: AR-type

Model	RV	$TMPV^{[0.5]}$	$TMPV^{[1]}$	$TMPV^{[1.5]}$	Range	TBPV	TC
62	5	0	0	0	0	0	0
63	0	5	0	0	0	0	0
64	0	0	5	0	0	0	0
65	0	0	0	5	0	0	0
66	0	0	0	0	5	0	0
67	0	0	0	0	0	5	0
68	0	0	0	0	0	0	5
69	10	0	0	0	0	0	0
70	0	10	0	0	0	0	0
71	0	0	10	0	0	0	0
72	0	0	0	10	0	0	0
73	0	0	0	0	10	0	0
74	0	0	0	0	0	10	0
75	0	0	0	0	0	0	10
76	15	0	0	0	0	0	0
77	0	15	0	0	0	0	0
78	0	0	15	0	0	0	0
79	0	0	0	15	0	0	0
80	0	0	0	0	15	0	0
81	0	0	0	0	0	15	0
82	0	0	0	0	0	0	15
83	5	1	0	0	0	0	0
84	5	0	1	0	0	0	0
85	5	0	0	1	0	0	0
86	5	0	0	0	1	0	0
87	5	0	0	0	0	1	0
88	5	0	0	0	0	0	1
89	5	5	0	0	0	0	0
90	5	0	5	0	0	0	0
91	5	0	0	5	0	0	0
92	5	0	0	0	5	0	0
93	5	0	0	0	0	5	0
94	5	0	0	0	0	0	5
95	10	1	0	0	0	0	0
96	10	0	1	0	0	0	0
97	10	0	0	1	0	0	0
98	10	0	0	0	1	0	0
99	10	0	0	0	0	1	0
100	10	0	0	0	0	0	1
101	10	5	0	0	0	0	0
102	10	0	5	0	0	0	0
103	10	0	0	5	0	0	0
104	10	0	0	0	5	0	0
105	10	0	0	0	0	5	0
106	10	0	0	0	0	0	5

Notes: Panel A shows the HAR-type models. Each row indicates the regressors included in a model: 1 indicates a daily factor (e.g., $\log(RV_{t-1,1})$), 2 a daily and weekly factor (e.g., $\log(RV_{t-1,1})$, $\log(RV_{t-5,5})$), and 3 a daily, weekly and monthly factor (e.g., $\log(RV_{t-1,1})$, $\log(RV_{t-5,5})$, $\log(RV_{t-20,20})$), using the respective regressor in that column. Models 1–7 are the HAR-log specifications in logarithms of RV, $TMPV^{[0.5]}$, $TMPV^{[1]}$, $TMPV^{[1.5]}$, Range, TBPV or TC with a jump term TJ_{t-1}. Models 8–61 are the HAR models of mixtures of volatility terms with a jump term TJ_{t-1}. For instance, model 27 has regressors $X_t - 1 = [1, \log(RV_{t-1,1}), \log(RV_{t-5,5}), \log(TMPV^{[0.5]}_{t-1,1}), \log(TMPV^{[0.5]}_{t-5,5}), TJ_{t-1}]$. Panel B shows the AR-type models. Each row lists the number of lagged regressors from the respective regressor column. Models 62–82 are the AR specifications in logarithms of RV, $TMPV^{[0.5]}$, $TMPV^{[1]}$, $TMPV^{[1.5]}$, Range, TBPV or TC. Models 83–106 are the AR models of mixtures of volatility terms. All specifications include a jump term TJ_{t-1}. For example, model 101 includes 10 lags of daily RV, and five logs of daily $TMPV^{[0.5]}$, as well as a jump term TJ_{t-1}.

Appendix C. The results of a robustness check

See Tables C.1 and C.2.

Table C.1
MCS results of point forecasts: 1- to 20-day-ahead forecasts, $\log(\mathrm{RV}_{t,h})$ based on a 1 min sampling frequency.

	$h = 1$			$h = 5$			$h = 20$		
	MCS	T_R	T_{SQ}	MCS	T_R	T_{SQ}	MCS	T_R	T_{SQ}
Soybean	HAR-RV-J	0.122	0.130	HAR-TMPV$^{[1.5]}$-J	0.108	0.112	HAR-Range-J	0.110	0.112
	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000
	DMA-TV-HAR	0.368	0.364	DMA-TV-HAR	0.143	0.144	DMA-TV-HAR	0.193	0.198
	BMA-HAR	0.440	0.442	BMA-HAR	0.130	0.133	BMA-HAR	0.155	0.160
Cotton	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000	HAR-Range-J	0.127	0.130
	DMA-TV-HAR	0.130	0.126	DMA-TV-HAR	0.139	0.140	HAR-TVS	1.000	1.000
	BMA-HAR	0.216	0.219	BMA-HAR	0.168	0.170	DMA-TV-HAR	0.189	0.192
							BMA-HAR	0.260	0.263
Gluten wheat	HAR-TVS	1.000	1.000	HAR-Range-J	0.216	0.208	HAR-Range-J	0.165	0.147
	DMA-TV-HAR	0.215	0.207	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000
	BMA-HAR	0.186	0.188	DMA-TV-HAR	0.429	0.427	DMA-TV-HAR	0.186	0.185
				BMA-HAR	0.217	0.216	BMA-HAR	0.160	0.163
Corn	HAR-TVS	1.000	1.000	HAR-RV-J	0.106	0.104	HAR-TVS	1.000	1.000
	DMA-TV-HAR	0.336	0.338	HAR-TVS	1.000	1.000	DMA-TV-HAR	0.327	0.328
	BMA-HAR	0.215	0.217	DMA-TV-HAR	0.180	0.181	BMA-HAR	0.264	0.270
				BMA-HAR	0.214	0.210			
Early indica rice	HAR-TVS	1.000	1.000	HAR-RV-J	0.106	0.108	HAR-TVS	1.000	1.000
	DMA-TV-HAR	0.473	0.482	HAR-TVS	0.684	0.688	DMA-TV-HAR	0.527	0.530
	BMA-HAR	0.227	0.225	DMA-TV-HAR	1.000	1.000	BMA-HAR	0.410	0.415
				BMA-HAR	0.259	0.287			
Palm	HAR-TMPV$^{[1]}$-J	0.232	0.239	HAR-TVS	0.845	0.849	HAR-TVS	1.000	1.000
	HAR-TVS	1.000	1.000	DMA-TV-HAR	1.000	1.000	DMA-TV-HAR	0.604	0.615
	DMA-TV-HAR	0.622	0.619	BMA-HAR	0.242	0.249	BMA-HAR	0.427	0.426
	BMA-HAR	0.520	0.521						

Notes: The table presents the MCS with the corresponding p-values for our selection of RV models. The p-values are obtained based on 10,000 block bootstraps. The underlined models are included in both $\widehat{M}^*_{0.75}$ and $\widehat{M}^*_{0.90}$. The models without underlines are only included in $\widehat{M}^*_{0.90}$. Note that $\widehat{M}^*_{0.75} \subset \widehat{M}^*_{0.90}$. HAR-TVS is the proposed HAR model with time-varying sparsity.

Table C.2
MCS results of point forecasts: 1- to 20-day-ahead forecasts, $\log(\mathrm{RV}^{FTRK}_{t,h})$.

	$h = 1$			$h = 5$			$h = 20$		
	MCS	T_R	T_{SQ}	MCS	T_R	T_{SQ}	MCS	T_R	T_{SQ}
Soybean	HAR-RV-J	0.106	0.112	HAR-TVS	0.872	0.885	HAR-TVS	1.000	1.000
	HAR-TVS	1.000	1.000	DMA-TV-HAR	1.000	1.000	DMA-TV-HAR	0.647	0.658
	DMA-TV-HAR	0.338	0.352	BMA-HAR	0.360	0.354	BMA-HAR	0.410	0.405
	BMA-HAR	0.415	0.416						
Cotton	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000	HAR-Range-J	0.132	0.138
	DMA-TV-HAR	0.133	0.129	DMA-TV-HAR	0.156	0.155	HAR-TVS	1.000	1.000
	BMA-HAR	0.259	0.250	BMA-HAR	0.162	0.174	DMA-TV-HAR	0.175	0.169
							BMA-HAR	0.162	0.164
Gluten wheat	HAR-TMPV$^{[1]}$-J	0.205	0.209	HAR-TVS	0.808	0.814	HAR-TVS	1.000	1.000
	HAR-TVS	1.000	1.000	DMA-TV-HAR	1.000	1.000	DMA-TV-HAR	0.600	0.613
	DMA-TV-HAR	0.613	0.612	BMA-HAR	0.526	0.529	BMA-HAR	0.527	0.521
	BMA-HAR	0.537	0.548						
Corn	HAR-RV-J	0.126	0.131	HAR-TMPV$^{[1.5]}$-J	0.108	0.112	HAR-Range-J	0.110	0.112
	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000
	DMA-TV-HAR	0.375	0.379	DMA-TV-HAR	0.148	0.144	DMA-TV-HAR	0.426	0.434
	BMA-HAR	0.442	0.407	BMA-HAR	0.132	0.140	BMA-HAR	0.359	0.360
Early indica rice	HAR-TVS	1.000	1.000	HAR-RV-J	0.104	0.103	HAR-TVS	1.000	1.000
	DMA-TV-HAR	0.542	0.533	HAR-TVS	0.872	0.891	DMA-TV-HAR	0.629	0.622
	BMA-HAR	0.328	0.327	DMA-TV-HAR	1.000	1.000	BMA-HAR	0.474	0.470
				BMA-HAR	0.426	0.428			
Palm	HAR-TMPV$^{[1]}$-J	0.108	0.112	HAR-Range-J	0.102	0.100	HAR-TVS	1.000	1.000
	HAR-TVS	1.000	1.000	HAR-TVS	1.000	1.000	DMA-TV-HAR	0.493	0.500
	DMA-TV-HAR	0.146	0.154	DMA-TV-HAR	0.449	0.462	BMA-HAR	0.387	0.388
	BMA-HAR	0.238	0.232	BMA-HAR	0.275	0.282			

Notes: The table presents the MCS with the corresponding p-values for our selection of RV models. The p-values are obtained based on 10,000 block bootstraps. The underlined models are included in both $\widehat{M}^*_{0.75}$ and $\widehat{M}^*_{0.90}$. The models without underlines are only included in $\widehat{M}^*_{0.90}$. Note that $\widehat{M}^*_{0.75} \subset \widehat{M}^*_{0.90}$. HAR-TVS is the proposed HAR model with time-varying sparsity.

References

Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance*, *57*(3), 1047–1091.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56.

Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 885–905.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics*, *89*(4), 701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, *96*(453), 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*(2), 579–625.

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

Atchadé, Y., & Fort, G. (2010). Limit theorems for some adaptive MCMC algorithms with subgeometric kernels. *Bernoulli*, *16*(1), 116–154.

Atchadé, Y. F., & Rosenthal, J. S. (2005). On adaptive Markov chain Monte Carlo algorithms. *Bernoulli*, *11*(5), 815–828.

Bandi, F. M., & Russell, J. R. (2006). Separating microstructure noise from volatility. *Journal of Financial Economics*, *79*(3), 655–692.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, *76*, 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *The Econometrics Journal*, *12*(3), C1–C32.

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, *17*(5), 457–477.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, *2*(1), 1–37.

Bekaert, G., Harvey, C. R., & Lundblad, C. (2007). Liquidity and expected returns: Lessons from emerging markets. *Review of Financial Studies*, *20*(6), 1783–1831.

Bollerslev, T., Litvinova, J., & Tauchen, G. (2006). Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics*, *4*(3), 353–384.

Brandt, M. W., & Jones, C. S. (2006). Volatility forecasting with range-based EGARCH models. *Journal of Business and Economic Statistics*, *24*(4), 470–486.

Cheung, C. S., & Miu, P. (2010). Diversification benefits of commodity futures. *Journal of International Financial Markets, Institutions and Money*, *20*(5), 451–474.

Choi, K., Yu, W. C., & Zivot, E. (2010). Long memory versus structural breaks in modeling and forecasting realized volatility. *Journal of International Money and Finance*, *29*(5), 857–875.

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

Corsi, F., Mittnik, S., Pigorsch, C., & Pigorsch, U. (2008). The volatility of realized volatility. *Econometric Reviews*, *27*(1–3), 46–78.

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

Degiannakis, S. (2008). ARFIMAX and ARFIMAX-TARCH realized volatility modeling. *Journal of Applied Statistics*, *35*(10), 1169–1180.

Eklund, J., & Karlsson, S. (2007). Forecast combination and model averaging using predictive measures. *Econometric Reviews*, *26*(2–4), 329–363.

Engle, R. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics*, *17*(5), 425–446.

Engle, R. F., & Gallo, G. M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, *131*(1), 3–27.

Fiszeder, P., & Perczak, G. (2016). Low and high prices can improve volatility forecasts during periods of turmoil. *International Journal of Forecasting*, *32*(2), 398–410.

Forsberg, L., & Ghysels, E. (2007). Why do absolute returns predict volatility so well? *Journal of Financial Econometrics*, *5*(1), 31–67.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, *131*(1), 59–95.

Ghysels, E., & Sinko, A. (2006). Comment on Hansen and Lunde's JBES paper. *Journal of Business and Economic Statistics*, *24*(2), 192–194.

Giot, P., & Laurent, S. (2004). Modelling daily value-at-risk using realized volatility and ARCH type models. *Journal of Empirical Finance*, *11*(3), 379–398.

Gouriéroux, C., & Jasiak, J. (2006). Autoregressive gamma processes. *Journal of Forecasting*, *25*(2), 129–152.

Grassi, S., & de Magistris, P. S. (2015). It's all about volatility of volatility: Evidence from a two-factor stochastic volatility model. *Journal of Empirical Finance*, *30*, 62–78.

Griffin, J. E., & Brown, P. J. (2010). Inference with normal-gamma prior distributions in regression problems. *Bayesian Analysis*, *5*(1), 171–188.

Haario, H., Saksman, E., & Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli*, 223–242.

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

Hansen, P. R., & Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business and Economic Statistics*, *24*(2), 127–161.

Hansen, P., & Lunde, A. (2011). Forecasting volatility using high frequency data. In *Oxford handbook of economic forecasting*. Oxford University Press.

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

Huang, X., & Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, *3*(4), 456–499.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., & Vetter, M. (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and their Applications*, *119*(7), 2249–2276.

Kalli, M., & Griffin, J. E. (2014). Time-varying sparsity in dynamic regression models. *Journal of Econometrics*, *178*(2), 779–793.

Koop, G., & Korobilis, D. (2012). Forecasting inflation using dynamic model averaging. *International Economic Review*, *53*(3), 867–886.

Lesmond, D. A., Ogden, J. P., & Trzcinka, C. A. (1999). A new estimate of transaction costs. *Review of Financial Studies*, *12*(5), 1113–1141.

Lieberman, O., & Phillips, P. C. (2008). Refined inference on long memory in realized volatility. *Econometric reviews*, *27*(1–3), 254–267.

Liu, C., & Maheu, J. M. (2008). Are there structural breaks in realized volatility? *Journal of Financial Econometrics*, *6*(3), 326–360.

Liu, C., & Maheu, J. M. (2009). Forecasting realized volatility: a Bayesian model-averaging approach. *Journal of Applied Econometrics*, *24*(5), 709–733.

Martens, M., Van Dijk, D., & De Pooter, M. (2009). Forecasting S&P 500 volatility: Long memory, level shifts, leverage effects, day-of-the-week seasonality, and macroeconomic announcements. *International Journal of Forecasting*, *25*(2), 282–303.

McAleer, M., & Medeiros, M. C. (2008). A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. *Journal of Econometrics*, *147*(1), 104–119.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Olsen, R. B., Pictet, O. V., & von Weizsäcker, J. E. (1997). Volatilities of different time resolutions—analyzing the dynamics of market components. *Journal of Empirical Finance*, *4*(2), 213–239.

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

Peck, A. E. (2008). The development of futures markets in China: Evidence of some unique trading characteristics. In B. Goss (Ed.), *Debt, risk, and liquidity in futures markets* (pp. 46–74). New York: Routledge, Oxon.

Pitt, M. K., Chatfield, C., & Walker, S. G. (2002). Constructing first order stationary autoregressive models via latent processes. *Scandinavian Journal of Statistics*, *29*(4), 657–663.

Pitt, M. K., & Walker, S. G. (2005). Constructing stationary time series models using auxiliary variables with applications. *Journal of the American Statistical Association*, *100*(470), 554–564.

Raftery, A. E., Kárný, M., & Ettler, P. (2010). Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill. *Technometrics*, *52*(1), 52–66.

Raggi, D., & Bordignon, S. (2012). Long memory and nonlinearities in realized volatility: A Markov switching approach. *Computational Statistics and Data Analysis*, *56*(11), 3730–3742.

Roll, R. (1984). A simple implicit measure of the effective bid—ask spread in an efficient market. *The Journal of Finance*, *39*(4), 1127–1139.

Shephard, N., & Sheppard, K. (2010). Realising the future: forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, *25*(2), 197–231.

Tseng-Chan, T., Huimin, C., & Chin-Sheng, H. (2009). Modeling jump and continuous components in the volatility of oil futures. *Studies in Nonlinear Dynamics and Econometrics*, *13*(3), 1–30.

Wang, X., Wu, C., & Xu, W. (2015). Volatility forecasting: The role of lunch-break returns, overnight returns, trading volume and leverage effects. *International Journal of Forecasting*, *31*(3), 609–619.

Yang, J., Bessler, D. A., & Leatham, D. J. (2001). Asset storability and price discovery of commodity futures markets: A new look. *Journal of Futures Markets*, *21*, 279–300.

Yang, K., & Chen, L. (2014). Realized volatility forecast: structural breaks, long memory, asymmetry, and day-of-the-week effect. *International Review of Finance*, *14*(3), 345–392.

Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, *12*(6), 1019–1043.

Zhang, L., Mykland, P. A., & Aït-Sahalia, Y. (2012). A tale of two time scales. *Journal of the American Statistical Association*, *100*(472), 1394–1411.

Fengping Tian has a Ph.D. in Management from Lingnan College at Sun Yat-Sen University and is a lecturer in Finance in International School of Business and Finance at Sun Yat- sen University.

Her research interests lie mainly in financial engineering and financial econometrics.

Ke Yang gained his Ph.D. in Finance from Lingnan College at Sun Yat-Sen University and is an Associate Professor in Finance in School of Economics and Commerce at South China University of Technology.

His research interests include financial risk forecasting, financial econometrics and time series.

Langnan Chen received his Ph.D. in Finance from Xiamen University and is a Professor in Finance and economics in Lingnan College at Sun Yat-sen University.

His research interests lie mainly in financial economics and monetary economics.