

Luo, Jiawen; Klein, Tony; Ji, Qiang; Hou, Chenghan

**Working Paper**

## Forecasting Realized Volatility of Agricultural Commodity Futures with Infinite Hidden Markov HAR Models

QMS Research Paper, No. 2019/10

**Provided in Cooperation with:**

Queen's University Belfast, Queen's Business School

*Suggested Citation:* Luo, Jiawen; Klein, Tony; Ji, Qiang; Hou, Chenghan (2019) : Forecasting Realized Volatility of Agricultural Commodity Futures with Infinite Hidden Markov HAR Models, QMS Research Paper, No. 2019/10, Queen's University Belfast, Queen's Management School, Belfast,  
<https://doi.org/10.2139/ssrn.3435054>

This Version is available at:

<https://hdl.handle.net/10419/271230>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



QUEEN'S  
UNIVERSITY  
BELFAST

MANAGEMENT  
SCHOOL

## QUEEN'S MANAGEMENT SCHOOL RESEARCH PAPER SERIES

QMS Research Paper 2019/10

# Forecasting Realized Volatility of Agricultural Commodity Futures with Infinite Hidden Markov HAR Models

Jiawen Luo – South China University of Technology

Tony Klein – Queen's University Belfast

Qiang Ji – Chinese Academy of Science

Chenghan Hou – Hunan University

9 August 2019

<https://ssrn.com/abstract=3435054>

Queen's University Belfast

Queen's Management School

<https://www.qub.ac.uk/mgt/>

Queen's Management School Research Paper Series

<http://ssrn.com/link/Queens-Management-BSR.html>

# Forecasting Realized Volatility of Agricultural Commodity Futures with Infinite Hidden Markov HAR Models<sup>☆</sup>

Jiawen Luo<sup>a</sup>, Tony Klein<sup>b</sup>, Qiang Ji<sup>c,\*</sup>, Chenghan Hou<sup>d</sup>

<sup>a</sup>South China University of Technology, Guangzhou, P. R. China

<sup>b</sup>Queen's Management School, Queen's University Belfast, United Kingdom

<sup>c</sup>Chinese Academy of Sciences, Beijing, P. R. China

<sup>d</sup>Hunan University, Changsha, P. R. China

---

## Abstract

We construct a set of HAR models with three types of infinite Hidden Markov regime switching structures. Particularly, jumps, leverage effects, and speculation effects are taken into account in realized volatility modeling. We forecast five agricultural commodity futures (Corn, Cotton, Indica Rice, Palm oil and Soybean) based on high frequency data from Chinese futures markets and evaluate the forecast performances with both statistical and economic evaluation measures. The statistical evaluation results suggest that HAR models with infinite Hidden Markov regime switching structures have better precision compared the benchmark HAR models based on the MZ- $R^2$ , MAFE, and MCS results. The economic evaluation results suggest that portfolios constructed with infinite Hidden Markov regime switching HARs achieve higher portfolio returns for risk averse investors compared to benchmark HAR model for short-term volatility forecasts.

*Keywords:* Agriculture commodity futures; Realized volatility forecasts; Infinite Hidden Markov switching process; HAR models;

*JEL classification:* C53; G17; Q14

---

## 1. Introduction & Literature review

Agricultural commodities, such as Corn, Wheat, Rice, and Soybean, are farming products and have been existing for centuries. For the same time, these commodities have been subject to trade and exchange. Only recently, those commodities are being traded centralized and in globalized markets. Volatility of agricultural commodity prices and sudden, unforeseen changes play an important role as they directly affect our daily lives. From a financialized perspective, forecasting volatility of agriculture commodity futures helps to assess and hedge risks associated with these contracts as well as to provide policy makers with tools to evaluate different scenarios.

---

<sup>☆</sup>We thank the editor Tao Hong and three anonymous referees for insightful comments and hints which led to a substantial improvement of the manuscript.

\*Corresponding author, mail: jqwxnjq@casipm.ac.cn.

Nerlove (1956) presents some early stage research on realized and expected prices as well as their elasticities of Cotton, Wheat, and Corn between 1909 and 1932. It is found that the response of supply to prices and price changes is only partly based on historical prices. Price expectation plays significant and vital role in supply management. This is of particular interest as it is early evidence that expected prices of future crop yields—to some degree similar to a futures price—are directly incorporated in crop management. Research on agricultural commodities is otherwise sparse but is being addressed with an increasing degree of financialization of these markets. This is outlined by Witt et al. (1987) who address cross-hedging applications and optimal hedge ratios. Krueger et al. (1988) analyzes the impact of trade policies on different categories of agricultural commodities and their effect on price stabilization for importable and exportable commodities in the U.S. Other countries have functioning derivative markets as well. The efficiency of commodity futures markets in the U.K. are analyzed with co-integration methods in Aulton et al. (1997) for example, where the Wheat market found to be efficiently functioning while Potatoes are not.

Regulation of agricultural commodities is a delicate topic in general as the aim for profits might clash with vital food supply. Those contrary positions become apparent in the food crisis of 2006 which is linked to the sharp rise of biofuels, in particular the first generation, which are fuels derived from crops grown on arable land. This directly impacts food security as less land and less crops are available for a growing population. With regard to this socio-economic impact, Escobar et al. (2009) call for regulation of biofuels to overcome the risks for food security. Schmidhuber (2007) finds evidence that policy changes with respect to biofuels should lead to a closer market integration between energy and agricultural commodities. Tyner (2010) finds a strong link between Ethanol and Corn prices after the financial crisis while before the crisis, Corn and crude oil linked significantly. Nazlioglu et al. (2013) find spillover effects between crude oil and major agricultural commodities (Wheat, Corn, Soybeans, and Sugar). A structural break is found in their relationship after the food crisis in 2006. Crude oil is now a major contributor to the volatility of agricultural commodities. This is also found for an earlier period in Natanelov et al. (2011) in a co-movement relationship of agricultural futures prices and crude oil. Policy developments are affecting these relationships; for low crude oil futures prices, biofuel policies buffer co-movements with Corn futures prices until the oil futures surpass a price threshold. Du et al. (2011) address spillover and linkage effects of crude oil markets to Corn and Wheat futures. After 2006, the spillover induced by dependence of these markets with ethanol production increases.

Forecasting volatility of commodity prices and their spillovers is of essential importance in order to identify and understand what factors drive different commodity markets (Nguyen & Walther, 2019). Also, co-movements play a vital role as outlined above. In recent literature, forecasting volatility has gained interest, oftentimes from a risk-

management perspective (Klein & Walther, 2016, Ji & Fan, 2016) or in terms of quantifying spillovers between markets (Nazlioglu et al., 2013, Todorova et al., 2014, Ji et al., 2018, Klein, 2018, Ma et al., 2019).

Having identified food crises and crude oil as a major contributor to the volatility of agricultural commodities, we expect highly volatile markets. By making use of high-frequency data at the 5-minute interval, we put particular focus on breakpoints and structural changes as well as sudden jumps in these volatilities. We model realized volatilities with the heterogeneous autoregressive (HAR) model of Corsi (2009). This HAR framework allows to model short, medium, and long-term dependencies among realized volatilities. Given its simplicity in application, the standard HAR gives rise to a plethora of model extensions covering different stylized facts. In this vein, we extend the standard HAR for jumps (Andersen et al., 2007), with jump thresholds (Corsi et al., 2010), or with signed jumps (Patton & Sheppard, 2015). Additionally, we make use of speculative behaviour and include volume and open interest as a predictor of future volatility (Bessembinder & Seguin, 1992). Accounting for recent advances, we also apply the centralized HAR of Bollerslev et al. (2018), where the long-run volatility is used to centralize volatilities.

The application of HAR-type models on agricultural commodities to forecast volatility is much less focused on than on other commodities such as crude oil (Souček & Todorova, 2013, Sévi, 2014, Wen et al., 2016, Degiannakis & Filis, 2017), precious metals (Caporin et al., 2015, Lyócsa & Molnár, 2016) or non-ferrous metals (Todorova et al., 2014, Zhu et al., 2017, Lyócsa et al., 2017). Forecasting realized volatilities with HAR models over longer time horizons becomes problematic if structural breaks are present in the observed data (Raggi & Bordignon, 2012, Yang & Chen, 2014). Ma et al. (2018a) and Ma et al. (2018b) apply regime switching HAR models on oil futures markets and find that regime switching increases the forecasting ability of HAR models. For oil futures, Ma et al. (2017) find that HAR models incorporating jumps yield better forecasting performance along with improvements of portfolio returns, where portfolio weights are calculated based on forecasting volatilities. Tian et al. (2016) use high-frequency data of Chinese agricultural commodity markets and find evidence for long memory and regime switching in form of a high and low volatility regime. Regime switching approaches yield better forecasting performances than their non-switching equivalent. Yang et al. (2017) employ so-called bagging and combination approaches to HAR models on Chinese agricultural commodity markets. It is again found that more sophisticated approaches surpass the forecasting quality of the base HAR models. Luo & Ji (2018) combine HAR-type models with dynamic correlation modelling (DCC, Engle, 2002) and detect a spillover from U.S. crude markets to Chinese agriculture commodities. These findings support earlier evidence of short-term spillovers in Jiang et al. (2017). We focus on those detected structural breaks and make use of infinite Hidden Markov switching (Hou, 2017). We compare in-sample

performance as well as forecasting ability of our base and switching HAR models for five agricultural commodities; Corn, Cotton, Indica Rice, Palm Oil, and Soybeans. In this manuscript, we employ the *infinite Hidden Markov switching* model (IHM) based on hierarchical Dirichlet processes introduced in Teh et al. (2006), the *double infinite Hidden Markov* Model (DIHM), and the IHM variant with constant conditional mean (IHMC). These three Markov-type frameworks are defined and described thoroughly in Hou (2017) and references therein. Volatility forecasts are carried out for aggregated one, five, and 22-day ahead forecasts, corresponding to daily, weekly, and monthly volatility. We test all 24 model variants<sup>1</sup> for intra-day prices of five agricultural commodities to gain a broad understanding of what drives these futures markets. This novel—and so far unique—approach of incorporating IHM variants reveals that regime switching Markov models offer an advantage in forecasting over the non-switching variants of the HAR class.

We contribute to the literature in the following aspects. Firstly, we focus on forecasting the realized volatilities of agricultural commodity futures in Chinese futures market for different time horizons (one, five, and 22-day ahead forecasts), which sheds some light on risk management, portfolio allocation, and pricing strategies for market participants in Chinese agricultural futures market. Secondly, we construct a set of volatility forecast models with structural breaks by combining three types of infinite Hidden Markov switching frameworks with conventional HAR models. Therefore, we can accommodate unknown breakpoints in volatilities due to uncertain factors such as policy changes, financial market crises, or investor sentiment in Chinese futures markets. We employ the precision-based algorithm of Chan & Jeliazkov (2009) which can simulate the parameters in distinct regimes efficiently and improve the computing speed greatly. The application of infinite Hidden Markov processes in modeling volatilities of agricultural commodities is our most important contribution as we are the first to thoroughly test this class of models in these markets. Finally, we employ both statistical and economic evaluation measures to evaluate the forecast precision as well as the economic significance in portfolio allocations for different volatility forecast models. Thus, we can provide a comprehensive scope for the comparisons of the volatility forecast models.

The remainder of this article is structured as follows. Section 2 presents the methodology and modelling framework. Section 3 gives an overview of the data and presents results of preliminary analyses and tests of daily returns, realized volatilities, and the speculative components for the five commodity futures markets focused on in this article. Section 4 discusses the results of the in-sample analysis as well as the performance of forecasting over different horizons. Section 5 provides robustness checks of the findings and Section 6 concludes this research.

---

<sup>1</sup>For each of the six HAR-type models, we apply the base as well as three switching variants, totalling 24 models.

## 2. Methodology

### 2.1. Realized Volatility

In order to calculate realized volatilities of intraday prices, we define returns as log-difference between two consecutive intraday prices:

$$r_{t,j} = \log P_{t,j} - \log P_{t,j-1}, \text{ for } j = 2, \dots, M, \quad (1)$$

where  $M$  denotes the number of 5-min prices per day. Hence, we obtain a sampling frequency of  $1/M$ . The realized volatility of day  $t$ , denoted  $RV_t$ , is then defined as

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

Barndorff-Nielsen & Sheppard (2004) introduce a measure of realized volatility which is robust against present jumps. This realized bi-power variation measure,  $BPV_t$ , is defined as

$$BPV_t = \mu_1^{-2} \sum_{j=1}^{M-1} |r_{t,j}| |r_{t,j+1}|$$

and used for identifying the jump component  $J_{t,\alpha}$ . We apply the definitions of Huang & Tauchen (2005) and calculate the jump component as follows

$$J_{t,\alpha} = I_{\{Z_t > \Phi_\alpha\}} (RV_t - BPV_t)^+, \quad (3)$$

where  $\Phi_\alpha$  refers to the critical value of the Gaussian distribution with

$$Z_t = \sqrt{M} \frac{1 - BPV_t \cdot RV_t^{-1}}{\sqrt{(\mu_1^{-4} + 2\mu_1^{-2} - 5) \max(1, TQ_t \cdot BPV_t^{-2})}} \quad (4)$$

and  $\mu_1 = \mathbb{E}(Z) = \sqrt{2/\pi}$ . The tri-power quarticity  $TQ_t$  is defined as

$$TQ_t = M \mu_{4/3}^{-3} \sum_{j=1}^{M-2} |r_{t,j}|^{4/3} |r_{t,j+1}|^{4/3} |r_{t,j+2}|^{4/3},$$

with  $\mu_p = 2^{p/2} \cdot \Gamma(1/2 \cdot (p+1)) \cdot \Gamma(1/2)$  (Andersen et al., 2007). Alternatives to this jump identification exist.

One commonly used approach that is robust against small and very high jumps is the *median RV* measures presented in Andersen et al. (2012). An additional measure, which is applied in this paper, is outlined in Corsi et al. (2010) and labelled *threshold bi-power variation (TBPV)*. It yields less biased estimates compared to standard, multipower variations of the continuous quadratic variation in finite samples (Corsi et al., 2010). The

TBPV is calculated as:

$$TBPV_t = \mu_1^{-2} \sum_{j=1}^{M-1} |r_{t,j}| |r_{t,j+1}| I_{\{|r_{t,j}|^2 \leq \vartheta_{t,j}\}} I_{\{|r_{t,j+1}|^2 \leq \vartheta_{t,j+1}\}},$$

where  $I_{\{\cdot\}}$  is the indicator function and  $\vartheta_{t,j}$  as a threshold variable for all  $t = 1, \dots, N$  and  $j = 1, \dots, M$ . For this threshold variable, it holds

$$\vartheta_{t,j} = c_\vartheta \cdot \hat{V}_{t,j},$$

where  $c_\vartheta$  is a scale-free constant while  $\hat{V}_{t,j}$  is the non-parameter recursive filter for calculating the local variance. Similar to Corsi et al. (2010), we set  $c_\vartheta = 3$ . Analogously to the calculation of the test statistic  $Z_t$  in Eq. (4), we define the *TBPV*-specific test statistic as follows:

$$Z_t^{TBPV} = \sqrt{M} \frac{1 - TBPV_t \cdot RV_t^{-1}}{\sqrt{(\mu_1^{-4} + 2\mu_1^{-2} - 5) \max(1, TTQ_t \cdot TBPV_t^{-2})}}, \quad (5)$$

with

$$TTQ_t = M \mu_{4/3}^{-3} \sum_{j=1}^{M-2} \prod_{k=0}^2 |r_{t,j+k}|^{4/3} I_{\{|r_{t,j+k}|^2 \leq \vartheta_{t,j+k}\}}.$$

Hence, we calculate the threshold jump component by

$$TJ_{t,\alpha} = I_{\{Z_t^{TBPV} > \Phi_\alpha\}} (RV_t - TBPV_t)^+. \quad (6)$$

Calculating the jump components  $J_{t,\alpha}$  and  $J_{t,\alpha}^{TBPV}$  in Eq. (3) and Eq. (6) yields the separation of the continuous components:

$$C_t = RV_t - J_{t,\alpha}, \quad (7)$$

$$TC_t = RV_t - TJ_{t,\alpha}. \quad (8)$$

In order to exploit possible leverage effects, Barndorff-Nielsen et al. (2010) separate the realized volatility in signed realized semivariances:

$$RS_t^+ = \sum_{j=1}^M I_{\{r_{t,j} > 0\}} r_{t,j}^2$$

$$RS_t^- = \sum_{j=1}^M I_{\{r_{t,j} < 0\}} r_{t,j}^2$$

which yields the signed jump variation  $\Delta J_t = RS_t^+ - RS_t^-$ .



## 2.2. Heterogeneous Autoregressive (HAR) Models

For all realized measures such as  $RV_t$ ,  $BPV_t$ ,  $TBPV_t$ ,  $C_t$ ,  $TC_t$ ,  $J_{t,\alpha}$ , and  $TJ_{t,\alpha}$  we calculate the average over a period of  $h$  days as

$$\overline{RM}_{h,t-1} = h^{-1} \sum_{j=1}^h RM_{t-j},$$

where  $RM$  is a placeholder for the aforementioned measures. We apply  $\overline{RM}_{1,t-1}$ ,  $\overline{RM}_{5,t-1}$ , and  $\overline{RM}_{22,t-1}$  as the daily, weekly, and monthly measures.

We follow Corsi (2009) in the definition of the standard HAR model:

$$RV_t = a_0 + a_d \overline{RV}_{1,t-1} + a_w \overline{RV}_{5,t-1} + a_m \overline{RV}_{22,t-1} + u_t \quad (9)$$

where  $u_t$  is an error term.

There are several extensions to the standard HAR model obtained by incorporating estimators for jumps or leverage effects. In addition to the standard HAR model, we apply the HAR-CJ model of Andersen et al. (2007) defined as

$$RV_t = a_0 + a_d \overline{C}_{1,t-1} + a_w \overline{C}_{5,t-1} + a_m \overline{C}_{22,t-1} + a_{dJ} \overline{J}_{1,t-1} + a_{wJ} \overline{J}_{5,t-1} + a_{mJ} \overline{J}_{22,t-1} + u_t, \quad (10)$$

the HAR-TCJ model (Corsi et al., 2010) defined as

$$\begin{aligned} RV_t = a_0 + a_d \overline{TC}_{1,t-1} + a_w \overline{TC}_{5,t-1} + a_m \overline{TC}_{22,t-1} \\ + a_{dJ} \overline{TJ}_{1,t-1} + a_{wJ} \overline{TJ}_{5,t-1} + a_{mJ} \overline{TJ}_{22,t-1} + u_t, \end{aligned} \quad (11)$$

and the HAR- $\Delta J$  model (Patton & Sheppard, 2015)

$$RV_t = a_0 + a_{d\Delta J} \overline{\Delta J}_{1,t-1} + a_d \overline{BPV}_{1,t-1} + a_w \overline{RV}_{5,t-1} + a_m \overline{RV}_{22,t-1} + u_t. \quad (12)$$

The significant connection between speculation activities and the market volatility has been addressed in literature (Bessembinder & Seguin, 1992, Chang et al., 2000, Brunetti et al., 2016). Therefore, we augment the standard HAR with the speculation variable  $X_{spec}$ , which is defined as

$$X_{spec} = \frac{Vol}{OI}, \quad (13)$$

and translates to the quotient of trading volume,  $Vol$ , and open interest,  $OI$ . Including this variable follows the assumption that speculation activity has some predictive power on future volatility (Bessembinder & Seguin, 1992). A large value of  $X_{spec}$  provides some evidence of increased speculation activity. The augmented HAR-spec model reads

$$RV_t = a_0 + a_d \overline{RV}_{1,t-1} + a_w \overline{RV}_{5,t-1} + a_m \overline{RV}_{22,t-1} + a_s X_{spec,t} + u_t. \quad (14)$$

As last model we apply the recently introduced centralized HAR model by Bollerslev et al. (2018). It is suggested that the unconditional variance level parameter (the constant  $a_0$ ) should be modelled variably. The intercept of the HAR model is replaced with a variable long-run volatility factor  $RV_t^{LR}$ , which equals the expanding sample mean of daily RV from the start of the sample up until day  $t$ . The model is defined as

$$RV_t - RV_t^{LR} = a_0 + a_d (\overline{RV}_{1,t-1} - RV_t^{LR}) + a_w (\overline{RV}_{5,t-1} - RV_t^{LR}) + a_m (\overline{RV}_{22,t-1} - RV_t^{LR}) + u_t. \quad (15)$$

### 2.3. The infinite Hidden Markov switching (IHMS) process

The HAR models defined in the previous section can be rewritten as general, linear regression model

$$y_t = \mathbf{X}_t \boldsymbol{\alpha} + u_t, \quad u_t \sim \mathcal{N}(0, \sigma^2), \quad (16)$$

where  $\mathcal{N}$  denotes the Gaussian distribution and  $\sigma^2$  is the variance of the residuals. The conventional HAR models outlined in Sec. 2.2 are specified with constant coefficients and variances. However, recent literature suggests that considering structural breaks improves the forecasting accuracy with HAR models as high-frequency data based realized volatilities of financial assets are subject to unknown structures (Bollerslev et al., 2015, ?). Therefore, we extend the HAR models with structural breaks by modeling coefficients and variance with infinite Hidden Markov switching processes.

As outlined before, three specifications of IHM-HAR models are implemented; the DIHM-HAR, the IHM-HAR, and the IHMC-HAR variants. Specifically, both IHM-HAR and DIHM-HAR allow the model coefficients and variances to change over time. The key difference between these two models appears in the latent infinite Markov switching processes that drive the time variation in the model parameters. The IHM-HAR implicitly restricts the changes in all model parameters to occur at the same time. However, forcing all parameters to change simultaneously might not be empirically plausible. To allow for richer dynamics, we consider a more general DIHM-HAR model in which switches in coefficients and variances are governed by two independent infinite Hidden Markov processes ( $s_t$  and  $z_t$ ). As many studies have shown that allowing time-varying variances is important in improving the out-of-sample forecasting performance, we also consider the IHMC-HAR in which only the variance is allowed to be changing over time. These three types of IHM-HAR models are defined in detail below.

For the DIHM-HAR model, we specify two independent infinite Markov regime switching processes to drive changes of the coefficients and variance, or in other words; the structural breaks in the coefficients and variances can occur at different time points.

Then, Eq. (16) can be rewritten with state-dependent coefficients and variances as

follows:

$$y_t = \mathbf{X}_t \boldsymbol{\alpha}_{s_t} + u_t, \quad u_t \sim \mathcal{N}(0, \sigma_{z_t}^2), \quad (17)$$

where the HAR coefficients  $\boldsymbol{\alpha}_{s_t}$  and variance  $\sigma_{z_t}^2$  are governed by two independent state variables  $s_t$  and  $z_t$ , respectively.

These regime indicators follow two infinite Hidden Markov processes in the following framework:

$$\boldsymbol{\alpha}_{s_t} \sim \mathcal{N}(\boldsymbol{\alpha}_0, \Sigma_0), \quad \sigma_{z_t}^2 \sim \mathcal{IG}(\phi_0, \nu_0), \quad (18)$$

$$s_t | s_{t-1}, \{\mathbf{p}_k^s\}_{k=1}^\infty \sim \mathbf{p}_{s_{t-1}}^s, \quad \text{and} \quad (19)$$

$$z_t | z_{t-1}, \{\mathbf{p}_k^z\}_{k=1}^\infty \sim \mathbf{p}_{z_{t-1}}^z, \quad \text{with} \quad (20)$$

$$\mathbf{p}_k^s | c^s, \rho^s, \boldsymbol{\pi}^s \sim \text{DP}(c^s, (1 - \rho^s)\boldsymbol{\pi}^s + \rho^s \delta_k), \quad (21)$$

$$\mathbf{p}_k^z | c^z, \rho^z, \boldsymbol{\pi}^z \sim \text{DP}(c^z, (1 - \rho^z)\boldsymbol{\pi}^z + \rho^z \delta_k), \quad (22)$$

$$\boldsymbol{\pi}^s | \gamma^s \sim \mathbf{SBP}(\gamma^s), \quad \text{and} \quad \boldsymbol{\pi}^z | \gamma^z \sim \mathbf{SBP}(\gamma^z). \quad (23)$$

In this framework, the coefficients  $\boldsymbol{\alpha}_{s_t}$ , and variances  $\sigma_{z_t}^2$  are sampled from two base distribution. In line with recent literature, we specify a Gaussian distribution (with parameters  $\boldsymbol{\alpha}_0$  and  $\Sigma_0$ ) for  $\alpha_i$  and an Inverse Gamma distribution (with parameters  $\phi_0$  and  $\nu_0$ ) for  $\sigma_i$ . The regime indicators  $s_t$  and  $z_t$  depend on the infinite state parameter  $\mathbf{p}_k^s$  and  $\mathbf{p}_k^z$  which are vectors of transition probabilities with infinite dimension.

Particularly, the infinite states parameter  $\mathbf{p}_k^s$  and  $\mathbf{p}_k^z$  follow a Dirichlet process (DP). The DP process is first proposed by Ferguson (1973) while we use the version of Fox et al. (2011). The DP process is determined by the positive concentration parameter  $c$  and the base distribution  $G_0$  which has the form  $G_0 = (1 - \rho)\pi + \rho\delta_i$ , where  $0 < \rho < 1$  denotes the sticky parameter that enhances the state self-transition probability and  $\delta_i$  is the degenerate probability measure sampled from  $G_0$ .

The parameter  $\pi$  in the DP process is obtained from a stick-breaking process (**SBP**) in Eq. (23) following Sethuraman (1994). The **SBP** can be described as dividing a unit-length stick for infinite times. Breaking the stick the  $i$ -th time is determined by  $V_i \sim \mathcal{B}(1, \gamma)$  where  $\mathcal{B}$  denotes the Beta distribution. Then,  $\pi_i = V_i \prod_{j=1}^{i-1} (1 - V_j)$  and we define  $\pi | \gamma \sim \mathbf{SBP}(\gamma)$  as the stick-breaking process.

For the IHM-HAR model, the variation coefficients and variances are governed by the same Markov state variable  $s_t$ , which follows an infinite Hidden Markov switching process

with

$$\boldsymbol{\alpha}_{s_t} \sim \mathcal{N}(\boldsymbol{\alpha}_0, \Sigma_0), \quad \sigma_{s_t}^2 \sim \mathcal{IG}(\phi_0, \nu_0), \quad (24)$$

$$s_t | s_{t-1}, \{\mathbf{p}_k^s\}_{k=1}^\infty \sim \mathbf{p}_{s_{t-1}}^s, \quad (25)$$

$$\mathbf{p}_k^s | c^s, \rho^s, \boldsymbol{\pi}^s \sim \text{DP}(c^s, (1 - \rho^s)\boldsymbol{\pi}^s + \rho^s \delta_k), \quad (26)$$

$$\boldsymbol{\pi}^s | \gamma^s \sim \mathbf{SBP}(\gamma^s). \quad (27)$$

For the IHMC-HAR model, the coefficients are assumed to be time-invariant and only the variances are driven by the infinite Hidden Markov switching process. The IHMC-HAR then reads

$$\boldsymbol{\alpha} \sim \mathcal{N}(\boldsymbol{\alpha}_c, \Sigma_c), \quad \sigma_{s_t}^2 \sim \mathcal{IG}(\phi_0, \nu_0), \quad (28)$$

$$s_t | s_{t-1}, \{\mathbf{p}_k^s\}_{k=1}^\infty \sim \mathbf{p}_{s_{t-1}}^s, \quad (29)$$

$$\mathbf{p}_k^s | c^s, \rho^s, \boldsymbol{\pi}^s \sim \text{DP}(c^s, (1 - \rho^s)\boldsymbol{\pi}^s + \rho^s \delta_k), \quad (30)$$

$$\boldsymbol{\pi}^s | \gamma^s \sim \mathbf{SBP}(\gamma^s). \quad (31)$$

This paper follows Song (2014) by imposing a second hierarchical structure on the base distribution. This setting has been shown to be helpful in improving model forecasting performance. It is worth noting that an infinite Hidden Markov switching model allows structural breaks to occur in each out-of-sample period. Without the second hierarchical prior, the parameters in the new born regimes are assumed to be generated from a pre-fix base distribution which might not be an attractive feature in practice. In the modeling setting with a hierarchical structure, the parameters specified the base distribution are unknown and updated based on the existing regime parameters. This allows the parameters from the new born regimes to learn from the existing regimes through this second hierarchical structure. The empirical importance of setting a second hierarchical structure on the base distribution has also been highlighted in many recent empirical studies (Maheu & Yang, 2016, Bauwens et al., 2017, Hou, 2017).

We follow this approach and set the second hierarchical priors as

$$\boldsymbol{\alpha}_0 \sim \mathcal{N}(\boldsymbol{\alpha}_{00}, \mathbf{B}_{00}), \quad \Sigma_0^{-1} \sim \mathcal{W}(\mathbf{A}_{00}, a_{00}), \quad (32)$$

$$\phi_0 \sim \mathcal{G}(q_{00}, b_{00}), \quad \nu_0 \sim \mathcal{E}(\lambda_{00}) \mathbb{1}(\nu_0 > n), \quad (33)$$

where  $\mathcal{W}(S, v)$  denotes the Wishart distribution with scale matrix  $S$  and degrees of freedom  $v$ ,  $\mathcal{E}(\lambda)$  refers to the exponential distribution with mean  $\lambda$ , and  $\mathbb{1}(\nu_0 > n)$  denotes the indicator function indicating that the support of the prior for  $\nu_0$  is restricted to be greater than  $n$ . Finally, the hyperparameters  $(c_s, c_z, \gamma_s, \gamma_z, \rho_s, \rho_z)$  are assumed to follow

independent priors

$$c^s \sim \mathcal{G}(w_s, \theta_s), \quad \gamma^s \sim \mathcal{G}(h_s, \eta_s), \quad \rho^s \sim \mathcal{B}(f_s, g_s), \quad (34)$$

$$c^z \sim \mathcal{G}(w_z, \theta_z), \quad \gamma^z \sim \mathcal{G}(h_z, \eta_z), \quad \rho^z \sim \mathcal{B}(f_z, g_z), \quad (35)$$

where  $\mathcal{G}(\kappa_1, \kappa_2)$  denotes the Gamma distribution with mean  $\kappa_1/\kappa_2$  and variance  $\kappa_1/\kappa_2^2$ .

#### 2.4. Estimations and priors

In this section we briefly summarize the Markov-chain Monte Carlo (MCMC) sampler used to estimate the IHM models. More details about the MCMC posterior sampler can be found in Hou (2017). The main difficulty for obtaining the posterior draws for the IHM models appears during the sampling of the state parameters. As the number of states for the IHM model is not bounded, the usual sampling method (Chib, 1996), which is designed to deal with the traditional Markov switching model, cannot be applied for the IHM model. To obtain the exact posterior samples for the state parameters, this paper follows the beam sampler approach proposed by Van Gael et al. (2008). The main idea of the beam sampler is to augment the parameter space by introducing a set of auxiliary variables which are sampled along with the other model parameters. These auxiliary variables do not change the marginal posterior distributions of the other parameters. Suppose that we introduce a set of auxiliary variables  $(u_1, \dots, u_T)$  which follow a conditional distribution with density

$$p(u_t | s_{t-1}, s_t) = \frac{p(u_t, s_t | s_{t-1})}{p(s_t | s_{t-1})} = \frac{\mathbb{1}(0 < u_t < p_{s_{t-1}, s_t})}{p_{s_{t-1}, s_t}}. \quad (36)$$

For simplicity, we suppress the rest of the conditioning variables. It can be seen that the joint density of  $(u_t, s_t)$  conditional on  $s_{t-1}$  is given by  $p(u_t, s_t | s_{t-1}) = \mathbb{1}(0 < u_t < p_{s_{t-1}, s_t})$ , which implies that the distribution obtained by marginalizing out  $u_t$  is  $p(s_t | s_{t-1}) = p_{s_{t-1}, s_t}$  which is exactly the same as the transition probability of the original model. Furthermore, it can be checked that the conditional density of  $s_t$  given  $(u_t, s_{t-1})$  is

$$p(s_t | u_t, s_{t-1}) = \frac{\mathbb{1}(0 < u_t < p_{s_{t-1}, s_t})}{\sum_i \mathbb{1}(0 < u_t < p_{s_{t-1}, i})}. \quad (37)$$

The finiteness of the set  $\{i : 0 < u_t < p_{s_{t-1}, i}^s\}$  implies that introducing the auxiliary variable  $u_t$  truncates the infinite number of state trajectories to a finite number. Conditional on the auxiliary variable  $u_t$ , the method of Chib (1996) can now be applied as usual to obtain posterior samples of the state parameters  $s_t$ . Given these state parameters, the other parameters are sampled using the standard method. For more discussion about the beam and posterior sampler for the IHM model, we refer to Song (2014), Hou (2017), and Van Gael et al. (2008).

For a more efficient comparison, we assume the same priors for parameters across all models. An informative prior is set for hyperparameters of the base distribution. In particular, we set the hyperparameters  $\beta_c = \beta_{00} = \mathbf{0}$ ,  $\mathbf{V}_c = \mathbf{B}_{00} = \mathbf{I}_{k_\beta}$ ,  $a_{00} = k_\beta + 10$ ,  $\mathbf{A}_{00} = 2a_{00}\mathbf{I}_{k_\beta}$ ,  $b_{00} = 5$ ,  $q_{00} = b_{00}^{-1}$ , and  $\lambda_{00} = n + 2$ . For the priors of the Hidden Markov parameters, we set  $w_s = w_z = 5$ ,  $\theta_s = \theta_z = 1$ , and  $h_s = h_z = \eta_s = \eta_z = 1$ . For the sticky parameters, we set  $f_s = f_z = 10$  and  $g_s = g_z = 1$  which implies a relatively high self-transition probability.

### 3. Data and sampling

With the rapid development of China’s economy and the gradual relaxation of regulations for foreign investors, the trading volume of its commodity futures market has increased significantly. There are three main futures exchanges for agricultural commodities in China: the Zhengzhou Commodity Exchange (ZCE), the Dalian Commodity Exchange (DCE), and the Shanghai Futures Exchange (SHFE). According to the 2018 statistics of the American Futures Industry Association (FIA) based on the number of contracts traded and/or cleared at 76 exchanges worldwide, 12 of the top 20 agricultural products with the largest trading volume in the world come from Chinese exchanges. Chinese commodity futures markets are gaining in global importance and prices have begun to affect global prices for commodities. Biofuels are gaining more and more attention as it is considered clean energy and an important renewable substitute for fossil fuels. Corn, Palm Oil, and Soybean are the main feedstock for the biofuel industry. Prices and volatilities of these commodities are shown to be correlated to fossil fuel markets.

In this paper, we use five agricultural commodity futures: Corn, Cotton, Indica Rice, Palm Oil, and Soybean. Soybean, Corn, and Palm Oil futures are traded at DCE, while Cotton and Indica Rice futures are traded at ZCE. Corn, Soybean, and Cotton futures are ranked among the top 20 agricultural agriculture futures contracts with the largest global trading volume. Indica Rice and Palm futures are two import agriculture commodity futures with high trading volumes in Chinese futures market.

We analyze their behavior and aim to accurately forecast the volatilities of these commodity futures to provide useful information and implications on derivative pricing and portfolio selection. A better understanding of volatility dynamics and forecastability benefits researchers, market participants, as well as policymakers.

We acquire high-frequency data in 5-min intervals from the *Wind Financial Database* for the time period from January 1, 2011 to December 31, 2015 for Corn, Cotton, Indica Rice, Palm Oil, and Soybean. After data pre-processing, we finally obtain 1256 daily observations for Corn, 1176 for Indica Rice, and 1214 for the other three commodities.

Tab. 1 presents descriptive statistics and preliminary tests of daily returns (Panel A), realized volatilities (Panel B), and the speculation variable (Panel C). In Panel A, it is found that the mean of all selected agriculture commodities in this paper is negative,

indicating the overall declining price trend during the sample period. Cotton traded at ZCE has the lowest mean return with  $-0.075$ , while Palm oil has the largest standard deviation with  $1.214$ . All return series are stationary which is verified by an ADF unit root test. Panel B of Tab. 1 shows that Palm oil has the largest mean of realized volatility with  $0.918$ , while Indica Rice has the largest standard deviation with  $2.924$ . All realized volatilities have a fat-tailed and leptokurtic distribution and thus, follow a non-normal distribution. This is additionally verified by significant Jarque-Bera statistics. Panel C presents findings for the speculation variable, the quotient of trading volume and open interest, to measure the speculative behavior. It is assumed that an increase (decrease) in the trading volume relative to the open interest indicates there is an increase (decrease) in the activity of the speculators. The statistics results in Panel C show that Palm oil has the largest mean of the speculation variable with  $1.086$ . It indicates that Palm oil tends to attract more speculation activity relative to other agricultural commodities, which is consistent with its largest mean of realized volatility among the five commodities. In addition to Tab. 1, daily returns, realized volatilities, and the speculation variable for all five commodities are visualized in Fig. 1, Fig. 2, and Fig. 3, respectively.

*Insert Tab. 1 approximately here.*

*Insert Fig. 1-3 approximately here.*

Fig. 1 and Fig. 2 plot the daily returns and realized volatilities of the five agricultural commodity futures. The returns and volatilities exhibit great fluctuations in 2010, 2012, and throughout 2015. These periods correspond to the first listing of Chinese index futures in 2010, the Euro debt crisis in 2011, and the China's stock market crisis in 2015. Additionally, we plot the speculation variables for each futures series in Fig. 3. In these plots, we find evidence that high levels of these speculation variables correspond to market phases of high volatility for the respective commodity.

## 4. Results

### 4.1. In-sample estimations

This section presents the full-sample estimation results for the DIHM-HAR and IHM-HAR model. Fig. 4 and Fig. 5 show the posterior mean and posterior distribution of the number of regimes for the DIHM-HAR and IHM-HAR model, respectively. These figures show that each agricultural commodity is specific in its characteristics on posterior mean and distribution of regime numbers, implying its different regime uncertainty in model estimations. Different numbers of regimes intuitively reflect differing uncertainty in which the existence of more regimes usually indicates better capacity to depict extreme situations in market changes.

In Fig. 4 for the DIHM-HAR model, the posterior distributions of the number of regimes for coefficients  $\beta$  for Corn, Indica Rice, and Soybean present symmetric char-

acteristics with similar probability peaks around 30 regimes, though the shapes of their posterior mean of regime switching is quite different. The posterior distribution of the number of regimes for coefficients  $\beta$  for Cotton and Palm oil is a right-skewed, fat-tail distribution. The shape of the posterior distribution of number of regimes for variances for the five agricultural commodities are similar in the sense that the probability distribution decreases with an increased regime number, except for Soybean.

Fig. 5 shows a different situation of posterior distributions of regime numbers for the IHM-HAR model. The shapes of posterior distribution of regime numbers for the five agricultural commodities are relatively asymmetric compared with Fig. 4, except for Palm oil. Moreover, the shapes of posterior mean of regime switching for the five agricultural commodities present no obvious difference.

*Insert Fig. 4-5 approximately here.*

To better investigate the data classification into different regimes, the heat maps for the posterior probabilities in the DIHM-HAR and the IHM-HAR model for the five commodities are depicted in Fig. 6 and Fig. 7. Referring to Hou (2017), the lighter (darker) the color is in a cell  $(i, j)$  of a heat map, the higher (lower) the probability that time  $i$  and time  $j$  are clustered into the same regime. For the heat maps of Fig. 6 and Fig. 7, it is evident that the clustering of regimes for coefficients and the volatilities are quite different over time for each agricultural commodity. The posterior probabilities for the coefficients visualized in the left column of Fig. 6 show that the regimes are not likely to be recurrent over time, indicated by dark colors for each agricultural commodity. However, for the posterior probabilities of volatilities visualized in the right column of Fig. 6, the color distribution is quite different for each agricultural commodity. In the case of Indica Rice the color is light before 2015, indicating that regimes have no apparent switching over time. However, for Cotton and Soybean, the light regions for the posterior probabilities of volatilities are few, indicating their time-variable regime switching. The heat map in Fig. 7 based on the IHM-HAR model also shows that the regimes change frequently over time.

*Insert Fig. 6-7 approximately here.*

Exploring the effects of the speculation activity on the volatility of agriculture commodity futures, we analyze the mean and standard deviation of the time-varying coefficients of the DIHM-HAR-spec and IHM-HAR model which are given in Tab. 2 and Tab. 3, respectively. In short, the results suggest that the speculation variable exhibits strong and positively directed effects on the future volatility of commodity futures. The coefficient  $a_s$  is highly significant, particularly for Corn, Cotton, and Palm oil. Hence, we augment the conventional HAR framework with the speculation variable and quantify its usefulness for forecasting in what follows.



*Insert Tab. 2 and Tab. 3 approximately here.*

#### 4.2. Out-of-sample evaluation

To evaluate the out-of-sample forecasting performance of the different HAR variants, we divide each sample series into two parts: the in-sample that includes about 2/3 of observations and the out-of-sample that includes the remainder of 1/3 of the sample observations. We employ a recursive forecast method to obtain the short-term ( $h = 1$ ), mid-term ( $h = 5$ ), and long-term ( $h = 22$ ) out-of-sample forecasts, corresponding to one-day, one-week and one-month ahead forecasts.

For the 1-step-ahead forecasts, we obtain realized volatility forecasts by re-estimating each volatility forecast model with in-sample observations.

For the multi-step forecasts, there are two types of forecast methods: the direct forecasts and the iterated forecasts as suggested in Marcellino et al. (2006). Specifically, the multi-step, iterated forecast results are obtained by iterating the forecast model with the previous forecast results. However, the forecast precision would deteriorate if the misspecification problem exists as suggested by Andersen et al. (2003) and Sizova (2011). Therefore, we use the direct forecasts to obtain the multi-step ahead forecast results following Corsi (2009). The direct forecasts aim at forecasting the cumulative  $h$ -day ahead realized volatility

$$RV_{t+1:t+h,t} := \frac{1}{h} \sum_{i=1}^h RV_{t+i,t},$$

for  $h = 1, 5, 22$ , which means that the one-step ahead realized volatility is replaced with the accumulated realized volatility over the weekly or monthly horizon in the forecast models.

Both the forecast precision and the economic significance of various volatility forecast models are evaluated in this article. By combining three types of the infinite Hidden Markov switching methods and six different HAR base models, outlined in Section 2.2, we construct a total of 18 HAR models. The six non-switching HAR models are included as benchmark models. We evaluate the unbiasedness of the forecast model based on the Mincer-Zarnowitz- $R^2$  (MZ- $R^2$ ) and the statistical loss functions such as the Mean Absolute Forecast Error (MAFE). The forecast precision is evaluated using the Model Confident Set (MCS) method of Hansen et al. (2011). Furthermore, we assess the forecast quality for all models from a portfolio optimization perspective. Given the optimal weights of a portfolio based on a risky and a risk-free asset, we compute the excess return and the economic value in order to measure the performances of various volatility models in portfolio construction.

#### 4.3. Statistical evaluation

Firstly, we calculate the MZ- $R^2$  and the MAFE loss function to evaluate the out-of-sample forecast precision. The MZ- $R^2$  is the goodness-of-fit ( $R^2$ ) for the Mincer-Zarnowitz

regression

$$RV_{t+1:t+h} = \beta_0 + \beta_1 \hat{R}V_{t+1:t+h,t} + \varepsilon_t,$$

where  $\hat{R}V_{t+1:t+h,t}$  is the  $h$ -step-ahead forecast of the realized volatility based on time  $t$  and  $RV_{t+1:t+h}$  is the actual realized volatility. The higher the value of  $MZ-R^2$ , in our case when  $\beta_0 = 0$  and  $\beta_1 = 1$ , the better the forecast performance of the respective volatility model.

Given the size of the in-sample period  $T_1$  and the size of the out-of-sample period  $T_2$ , with  $T_1 + T_2 = T$ , the MAFE is defined as

$$MAFE = \frac{1}{T_2} \sum_{t=T_1+1}^T \left| RV_{t+1:t+h} - \hat{R}V_{t+1:t+h} \right|$$

In order to determine models which outperform in terms of forecasting quality, we apply the Model Confidence Set (MCS) of Hansen et al. (2011). The MCS identifies a set of models with preferable forecasting performance (the confidence set) from a set of all models

$$\mathcal{M}_0 = \{\mathcal{M}^i, i = 1, \dots, M\}.$$

For a given confidence level  $\alpha$ , we identify the MCS  $\mathcal{M}_{1-\alpha}^* \subseteq \mathcal{M}_0$  which includes models that outperform all other models in the complementary set  $\overline{\mathcal{M}_{1-\alpha}^*}$ . It holds that  $\mathcal{M}_{1-\alpha}^* \cup \overline{\mathcal{M}_{1-\alpha}^*} = \mathcal{M}_0$ . By defining the relative performance variable  $d_{ij,t} := L_{i,t} - L_{j,t}$  with  $i, j = 1, \dots, M$  and  $L_{i,t}$  being the loss function of the  $i$ -th Model,  $\mathcal{M}^i$ , the MCS method then removes the least-performing models based on the equal predictive accuracy (ECA) test with

$$H_{0,\mathcal{M}} = \mathbb{E}(d_{ij,t}) = 0, \text{ for } i, j = 1, \dots, M.$$

We implement the MCS using the stationary bootstrap outlined in Hansen et al. (2011). Firstly, we generate  $B$  bootstrap resamples given the loss functions and choose the block-length bootstrap parameter  $L = 16$  as in Hansen et al. (2011) and calculate the relative performance variable for each bootstrap resample. Based on the range statistic  $T_R$  of Hansen et al. (2011), we determine  $\mathcal{M}_{1-\alpha}^*$  in a two-step iterative procedure.<sup>2</sup>

We display the statistical evaluation results for the 1-step forecasts in Tab. 4 based on the  $MZ-R^2$ , the MAFE loss function and the MCS results. The results are consistent for selecting the best performing forecast models. For all five agriculture commodity futures, the HAR model combined with the independent Hidden Markov switching have better forecast precision compared with the benchmark HAR models. Particularly, for the sample of Cotton, Indica Rice, and Palm futures, the HAR models within the three

---

<sup>2</sup>For more detail on the MCS and stationary block bootstrap we refer the reader to Hansen et al. (2011) and Politis & Romano (1994).

types of infinite Hidden Markov switching outperform the corresponding benchmark HAR models, suggesting that considering structural breaks in coefficients and variance improves the forecast performances of the original HAR model. For Corn and Soybean, the IHMC model with constant coefficients cannot outperform the benchmark HAR model. As shown in Tab. 1, realized volatilities of Corn and Soybean futures have low standard deviations; hence, only considering the structural breaks in variances results in insufficient performance of the IHMC-HAR models for these two samples.

Moreover, we find that the specification of independent Markov processes (DIHM) for coefficients and variances in the HAR models achieves the best forecasting precision among each group of HAR models. It is worth noting that the HAR models with the speculation variable have better forecast accuracy than all other models, suggesting that the speculation variable plays an important role in volatility forecasts for agricultural commodity futures. Based on the MCS results, most of the HAR models with infinite Hidden Markov processes survive in the MCS at the  $\alpha = 25\%$  confidence level.

*Insert Tab. 4 approximately here.*

Tab. 5 and Tab. 6 display the statistical evaluation results for the mid-term (a week ahead) and long-term (a month ahead) forecast models. As shown in Tab. 5, the results of  $MZ-R^2$  and the MAFE loss function suggest that the HAR models with the DIHM and IHM structures outperform the corresponding benchmark HAR models for all the five agriculture commodity futures. Moreover, most of the DIHM-HAR type models and IHM-HAR type models can survive in the MCS at the 25% confident level, suggesting their significant improvement in the forecast precision compared to the benchmark models.

The long-term forecast results as shown in Tab.6 are similar to the mid-term forecast results. We also find superior forecast precision of the DIHM-HAR type models and IHM-HAR type models compared to the corresponding benchmark HAR models. All HAR models with DIHM and IHM structures can survive in the MCS at the 10% confident set for the long-term forecasts.

In addition, we find that including the speculation variable in the HAR models improve forecast accuracy of the original HAR models for both the mid-term and long-term forecasts.

*Insert Tab. 5 and Tab. 6 approximately here.*

In conclusion, the IHM-HAR and DIHM-HAR models achieve higher forecast precision for short-, mid-, and long-term forecasts compared with to the benchmark HAR models. Both IHM-HAR and DIHM-HAR models allow the coefficients and variances to be changing over time. As the time-varying coefficients capture the unknown structural breaks in the volatilities of agricultural commodity futures, both of these two models perform rather well for the out-of-sample forecasts in term of statistical evaluation. Par-

ticularly, the DIHM-HAR models perform better than the IHM-HAR models for the short- and long-term forecasts, suggesting that modeling structural changes in the coefficients and variances separately produces better forecast precision. However, the IHMC-HAR models perform relative worse compared with the benchmark HAR models in terms of forecast precision. Only variances are allowed to be time-varying and coefficients are set to be time-invariant in the IHMC-HAR. Thus, the time-variation in the volatilities cannot be fully captured by the IHMC-HAR models.

#### 4.4. The economic evaluation

The statistical evaluation focuses on comparing the forecast accuracy and quality of various volatility forecasting models. However, market participants care more about the economic gains from volatility forecasts. To evaluate their economic significance, we consider an investor with the following mean-variance utility:

$$U(R_p) = \mathbb{E}(R_p) - \frac{1}{2}\gamma\mathbb{V}(R_p),$$

where  $\gamma$  refers to the *risk aversion rate*,  $\mathbb{E}(R_p)$  denotes the expected value of the portfolio return  $R_p$ , and  $\mathbb{V}(R_p)$  denotes the portfolio variance.

The risk-averse investor allocates its budget to a portfolio containing a risky asset and a risk-free asset. According to Campbell & Thompson (2008) and Neely et al. (2014), the optimal weight which is allocated to the agricultural commodity futures, based on forecasts of time  $t + 1$ , should be

$$\hat{w}_t = \frac{1}{\gamma} \frac{\hat{r}_{t+1}}{\hat{\sigma}_{t+1}^2},$$

where  $\hat{r}_{t+1}$  is the forecasted excess return of one of the agricultural commodity futures and  $\hat{\sigma}_{t+1}^2$  is the forecasted variance. Based on Campbell & Thompson (2008) and Neely et al. (2014), we employ the moving average of the excess returns with a rolling window of 256 trading days (1-year) to proxy the forecasted excess return. As the volatility of risk-free asset is trivial compared to the risky asset, the forecasted variance of returns is the volatility forecast obtained by the models outlined and analyzed in previous sections.

Then, the portfolio return at time  $t + 1$  is given by

$$R_{p,t+1} = \hat{w}_t r_{t+1} + r_{t+1}^f$$

where  $r_{t+1}$  and  $r_{t+1}^f$  are the excess returns of agriculture futures and the risk-free return at time  $t + 1$ , respectively. In this paper, we employ the return of a 5-year Treasury bill in China as the return of the risk-free asset.

Moreover, for the direct multi-step forecasts, the results for the multi-step forecast models are averaged over all possible weekly (e.g., Monday to Monday, Tuesday to Tuesday, etc.) horizons and monthly horizons (e.g. the 1st day of a month, the 2nd day of a

month and so on). Thus, the actual returns corresponding to the long horizon forecast of realized volatility for portfolio optimization are the moving averages of returns with a weekly or monthly window such that

$$r_{t+1:t+h} = \frac{1}{h} \sum_{i=1}^h r_{t+i}.$$

In addition, we employ the economic value to assess the economic significance of various multivariate volatility models. The economic value is the value of  $\Delta$  such that

$$\sum_{t=T_1+1}^T U(r_{p,t}^k) = \sum_{t=T_1+1}^T U(r_{p,t}^l - \Delta)$$

for two different portfolios  $p_1$  and  $p_2$ . The greater  $\Delta$ , the more returns a risk-averse investor is willing to sacrifice to switch from model  $l$  to model  $k$ . Based on Fleming & Hernández-Hernández (2003), Bollerslev et al. (2016), and Callot et al. (2017), we use the quadratic utility with risk aversion to determine the economic values of various forecast models:

$$U(r_{p,t}^k, \gamma) = (1 + r_{p,t}) - \frac{\gamma}{2(1 + \gamma)}(1 + r_{p,t})^2.$$

In this paper, we consider two levels of risk aversion rates for the investor, the mild risk aversion rate  $\gamma = 1$  and the strong risk aversion rate  $\gamma = 10$ . Tab. 7 shows the results of economic evaluation in terms of the portfolio return and the economic values of various HAR models within the infinite Hidden Markov switching framework. For the sake of brevity, we only report the return of portfolio based on a mild risk aversion level  $\gamma = 1$ . We also display the economic values of each infinite Hidden Markov switching HAR models against the corresponding benchmark HAR model based on two types of risk aversion rates.

The economic evaluation results suggest the HAR models combined with the three types of infinite Hidden Markov switching methods achieve higher portfolio returns compared to the original HAR models. Particularly, including the speculation variables in the HAR models increases the portfolio returns for all types of agricultural commodity futures. Moreover, almost all infinite Hidden Markov switching HAR models have positive economic values against the corresponding benchmark HAR model, suggesting the utility of the risk-averse investor can be improved by switching from the benchmark HAR models to the infinite Hidden Markov switching HAR models for making investment decisions. The improvement is more significant for the investor with a mild risk-averse rate.

*Insert Tab. 7 approximately here.*

Tab. 8 and Tab. 9 present the results of the economic evaluation in terms of the mid-

term ( $h = 5$ ) and long-term ( $h = 22$ ) forecasts. The results among these two forecasting horizons are similar. For Palm Oil futures, almost all infinite Hidden Markov switching HAR models achieve positive economic values against the corresponding benchmark HAR models, which is consistent with the results of short-term forecasts. However, for the other agriculture commodity samples, only HAR models with IHMC structure outperform the corresponding HAR models for mid-term and long-term forecasts.

In conclusion, the HAR models combined with the three types of infinite Hidden Markov switching structure improve the portfolio return and achieve positive economic gains compared to the base HAR models for the short-term forecasts, while only the IHMC-HAR models outperform the corresponding HAR models in terms of the portfolio return and the economic value for the mid-term and long-term forecasts.

*Insert Tab. 8 and Tab. 9 approximately here.*

## 5. Robustness checks with alternative realized volatility estimators

Following Fiszeder & Perczak (2016) and ?, we conduct a robustness check of the proposed models with alternative realized volatility estimators. Hansen & Lunde (2006) suggest the realized kernel estimator to correct the effects of market microstructure noise on the estimators of realized volatility, which reads

$$RK_t(M) = \sum_{j=1}^M r_{t,j}^2 + 2 \sum_{w=1}^q \left(1 - \frac{w}{q+1}\right) \sum_{i=1}^{M-w} r_{t,i} r_{t,i+w}. \quad (38)$$

Following Hansen & Lunde (2006), we set

$$q = \left\lceil \frac{\omega}{(b-a)/M} \right\rceil,$$

where  $\lceil x \rceil$  denotes the smallest integer that is greater or equal to  $x$ ,  $\omega$  is the desired width of the lag window, and  $b - a$  is the length of the sampling period.

Then, we conduct the 1-step ahead forecasts with the realized kernel estimator. The statistical evaluation and the economic evaluation results are displayed in Tab. 10 and Tab. 11 respectively.

From Tab. 10 the statistical evaluation results suggest that most of the HAR models with infinite Hidden Markov models outperform the benchmark HAR models. Moreover, as shown in Tab. 11, the HAR models with infinite Hidden Markov switching structures perform better than the benchmark HAR models in terms of portfolio return and economic value. These results verify that our models have outstanding forecast performances and that these forecast results are robust to the choice of realized volatility estimators.

*Insert Tab. 10 and Tab. 11 approximately here.*

Moreover, we employ another alternative realized volatility estimator, the *medRV* estimator proposed by Andersen et al. (2012), which is a jump-robust estimators of integrated variance. The statistical and economic evaluation results are shown in Tab. 12 and Tab. 12, respectively. According to these tables, the HAR models with infinite Hidden Markov models outperform the benchmark HAR models in terms of the statistical and economic evaluation criteria. Therefore, the forecast results are robust to the choice of realized volatility estimator.

*Insert Tab. 12 and Tab. 13 approximately here.*

## 6. Conclusions

We construct a set of HAR models with three types of infinite Hidden Markov regime switching structures. Particularly, jumps, leverage effects, and speculation effects are taken into account in realized volatility modeling. We forecast five agricultural commodity futures (Corn, Cotton, Indica Rice, Palm Oil, and Soybean) based on high frequency data obtained from Chinese futures markets. We analyze regime switching behavior with infinite Hidden Markov variants of different base HAR models. We compare the forecast precision of various models based on the  $MZ-R^2$ , the MAFE, and the MCS tests and further evaluate the economic significance of these forecast models in terms of portfolio return and economic value.

With respect to the statistical evaluation, the HAR models combined with double Hidden Markov processes (DIHM) and a single Hidden Markov process (IHM) have better forecast precision when compared with non-switching benchmark HAR models. These results hold for short-term, mid-term, and long-term forecasts. The HAR models with speculation variable have better forecast accuracy than all other models, suggesting the speculation variable plays an important role in volatility forecasts for agricultural commodity futures.

With respect to economic evaluation, HAR models combined with three types of infinite Hidden Markov switching structures improve the portfolio return and achieve positive economic gains compared to the corresponding benchmark HAR models for short-term forecasts. The IHMC-HAR model outperforms the corresponding benchmark models in terms of portfolio return and economic value for the mid-term and long-term forecasts. In addition, the improvements in the portfolio performances of the HAR models with infinite Hidden Markov switching structures are higher for mildly risk-averse investors.

In conclusion, we focus on introducing different Markov switching structures to improve the efficiency and accuracy of high-frequency volatility forecast models. We compare the forecast performance based on statistical and economic evaluation methods, which provide a comprehensive way to identify suitable forecast models for practical application.

We provide a new analytical tool for market participants to model and forecast volatility of agricultural commodity futures. This has great economic significance for investors to generate suitable portfolio strategies to quantify and control risk as well as for policy makers to determine adequate policies to ensure the stability of financial markets.



## References

- Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility. *Review of Economics and Statistics*, *89*, 701–720. doi:10.1162/rest.89.4.701.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, *71*, 579–625. doi:10.1111/1468-0262.00418.
- Andersen, T. G., Dobrev, D., & Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, *169*, 75–93. doi:10.1016/j.jeconom.2012.01.011.
- Aulton, A. J., Ennew, C. T., & Rayner, A. J. (1997). Efficiency Tests of Futures Markets for UK Agricultural Commodities. *Journal of Agricultural Economics*, *48*, 408–424. doi:10.1111/j.1477-9552.1997.tb01162.x.
- Barndorff-Nielsen, O. E., Kinnebrock, S., & Shephard, N. (2010). Measuring Downside Risk – Realized Semivariance. In *Volatility and Time Series Econometrics* (pp. 117–136). Oxford University Press. doi:10.1093/acprof:oso/9780199549498.003.0007.
- Barndorff-Nielsen, O. E., & Sheppard, N. (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics*, *2*, 1–37. doi:10.1093/jjfinec/nbh001.
- Bauwens, L., Carpentier, J.-F., & Dufays, A. (2017). Autoregressive Moving Average Infinite Hidden Markov-Switching Models. *Journal of Business & Economic Statistics*, *35*, 162–182. doi:10.1080/07350015.2015.1123636.
- Bessembinder, H., & Seguin, P. J. (1992). Futures-Trading Activity and Stock Price Volatility. *The Journal of Finance*, *47*, 2015–2034. doi:10.1111/j.1540-6261.1992.tb04695.x.
- Bollerslev, T., Hood, B., Huss, J., & Pedersen, L. H. (2018). Risk Everywhere: Modeling and Managing Volatility. *The Review of Financial Studies*, *31*, 2729–2773. doi:10.1093/rfs/hhy041.
- Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, *192*, 1–18. doi:10.1016/j.jeconom.2015.10.007.
- Bollerslev, T., Xu, L., & Zhou, H. (2015). Stock return and cash flow predictability: The role of volatility risk. *Journal of Econometrics*, *187*, 458–471. doi:10.1016/j.jeconom.2015.02.031.
- Brunetti, C., Büyükkahin, B., & Harris, J. H. (2016). Speculators, Prices, and Market Volatility. *Journal of Financial and Quantitative Analysis*, *51*, 1545–1574. doi:10.1017/S0022109016000569.
- Callot, L. A. F., Kock, A. B., & Medeiros, M. C. (2017). Modeling and Forecasting Large Realized Covariance Matrices and Portfolio Choice. *Journal of Applied Econometrics*, *32*, 140–158. doi:10.1002/jae.2512.

- Campbell, J. Y., & Thompson, S. B. (2008). Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies*, *21*, 1509–1531. doi:10.1093/rfs/hhm055.
- Caporin, M., Ranaldo, A., & Velo, G. G. (2015). Precious metals under the microscope: a high-frequency analysis. *Quantitative Finance*, *15*, 743–759. doi:10.1080/14697688.2014.947313.
- Chan, J. C.-C., & Jeliazkov, I. (2009). MCMC Estimation of Restricted Covariance Matrices. *Journal of Computational and Graphical Statistics*, *18*, 457–480. doi:10.1198/jcgs.2009.08095.
- Chang, E., Chou, R. Y., & Nelling, E. F. (2000). Market volatility and the demand for hedging in stock index futures. *Journal of Futures Markets*, *20*, 105–125. doi:10.1002/(SICI)1096-9934(200002)20:2<105::AID-FUT1>3.0.CO;2-Q.
- Chib, S. (1996). Calculating posterior distributions and modal estimates in Markov mixture models. *Journal of Econometrics*, *75*, 79–97. doi:10.1016/0304-4076(95)01770-4.
- Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, *7*, 174–196. doi:10.1093/jjfinec/nbp001.
- Corsi, F., Pirino, D., & Renò, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, *159*, 276–288. doi:10.1016/j.jeconom.2010.07.008.
- Degiannakis, S., & Filis, G. (2017). Forecasting oil price realized volatility using information channels from other asset classes. *Journal of International Money and Finance*, *76*, 28–49. doi:10.1016/j.jimonfin.2017.05.006.
- Du, X., Yu, C. L., & Hayes, D. J. (2011). Speculation and volatility spillover in the crude oil and agricultural commodity markets: A Bayesian analysis. *Energy Economics*, *33*, 497–503. doi:10.1016/j.eneco.2010.12.015.
- Engle, R. (2002). Dynamic Conditional Correlation. *Journal of Business & Economic Statistics*, *20*, 339–350. doi:10.1198/073500102288618487.
- Escobar, J. C., Lora, E. S., Venturini, O. J., Yáñez, E. E., Castillo, E. F., & Almazan, O. (2009). Biofuels: Environment, technology and food security. *Renewable and Sustainable Energy Reviews*, *13*, 1275–1287. doi:10.1016/j.rser.2008.08.014.
- Ferguson, T. S. (1973). A Bayesian Analysis of Some Nonparametric Problems. *The Annals of Statistics*, *1*, 209–230. doi:10.1214/aos/1176342360.
- Fiszeder, P., & Perczak, G. (2016). Low and high prices can improve volatility forecasts during periods of turmoil. *International Journal of Forecasting*, *32*, 398–410. doi:10.1016/j.ijforecast.2015.07.003.
- Fleming, W. H., & Hernández-Hernández, D. (2003). An optimal consumption model with stochastic volatility. *Finance and Stochastics*, *7*, 245–262. doi:10.1007/s007800200083.

- Fox, E., Sudderth, E. B., Jordan, M. I., & Willsky, A. S. (2011). Bayesian Nonparametric Inference of Switching Dynamic Linear Models. *IEEE Transactions on Signal Processing*, *59*, 1569–1585. doi:10.1109/TSP.2010.2102756.
- Hansen, P. R., & Lunde, A. (2006). Consistent ranking of volatility models. *Journal of Econometrics*, *131*, 97–121. doi:10.1016/j.jeconom.2005.01.005.
- Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The Model Confidence Set. *Econometrica*, *79*, 453–497. doi:10.3982/ECTA5771.
- Hou, C. (2017). Infinite hidden markov switching VARs with application to macroeconomic forecast. *International Journal of Forecasting*, *33*, 1025–1043. doi:10.1016/j.ijforecast.2017.06.006.
- Huang, X., & Tauchen, G. (2005). The Relative Contribution of Jumps to Total Price Variance. *Journal of Financial Econometrics*, *3*, 456–499. doi:10.1093/jjfinec/nbi025.
- Ji, Q., Bouri, E., Roubaud, D., & Shahzad, S. J. H. (2018). Risk spillover between energy and agricultural commodity markets: A dependence-switching CoVaR-copula model. *Energy Economics*, *75*, 14–27. doi:10.1016/j.eneco.2018.08.015.
- Ji, Q., & Fan, Y. (2016). Evolution of the world crude oil market integration: A graph theory analysis. *Energy Economics*, *53*, 90–100. doi:10.1016/j.eneco.2014.12.003.
- Jiang, H., Todorova, N., Roca, E., & Su, J.-J. (2017). Dynamics of volatility transmission between the U.S. and the Chinese agricultural futures markets. *Applied Economics*, *49*, 3435–3452. doi:10.1080/00036846.2016.1262517.
- Klein, T. (2018). Trends and contagion in WTI and Brent crude oil spot and futures markets - The role of OPEC in the last decade. *Energy Economics*, *75*, 636–646. doi:10.1016/j.eneco.2018.09.013.
- Klein, T., & Walther, T. (2016). Oil Price Volatility Forecast with Mixture Memory GARCH. *Energy Economics*, *58*, 46–58. doi:10.1016/j.eneco.2016.06.004.
- Krueger, A. O., Schiff, M., & Valdés, A. (1988). Agricultural Incentives in Developing Countries: Measuring the Effect of Sectoral and Economywide Policies. *The World Bank Economic Review*, *2*, 255–271. doi:10.1093/wber/2.3.255.
- Luo, J., & Ji, Q. (2018). High-frequency volatility connectedness between the US crude oil market and China’s agricultural commodity markets. *Energy Economics*, *76*, 424–438. doi:10.1016/j.eneco.2018.10.031.
- Lyócsa, S., & Molnár, P. (2016). Volatility forecasting of strategically linked commodity ETFs: gold-silver. *Quantitative Finance*, *16*, 1809–1822. doi:10.1080/14697688.2016.1211799.
- Lyócsa, S., Molnár, P., & Todorova, N. (2017). Volatility forecasting of non-ferrous metal futures: Covariances, covariates or combinations? *Journal of International Financial Markets, Institutions and Money*, *51*, 228–247. doi:10.1016/j.intfin.2017.08.005.
- Ma, F., Wahab, M., Huang, D., & Xu, W. (2017). Forecasting the realized volatility of the oil futures market: A regime switching approach. *Energy Economics*, *67*, 136–145. doi:10.1016/j.eneco.2017.08.004.

- Ma, F., Wei, Y., Chen, W., & He, F. (2018a). Forecasting the volatility of crude oil futures using high-frequency data: further evidence. *Empirical Economics*, *55*, 653–678. doi:10.1007/s00181-017-1294-6.
- Ma, F., Wei, Y., Liu, L., & Huang, D. (2018b). Forecasting realized volatility of oil futures market: A new insight. *Journal of Forecasting*, *37*, 419–436. doi:10.1002/for.2511.
- Ma, Y., Ji, Q., & Pan, J. (2019). Oil financialisation and volatility forecast: Evidence from multidimensional predictors. *Journal of Forecasting*, (p. for.2577). doi:10.1002/for.2577.
- Maheu, J. M., & Yang, Q. (2016). An infinite hidden Markov model for short-term interest rates. *Journal of Empirical Finance*, *38*, 202–220. doi:10.1016/j.jempfin.2016.06.006.
- Marcellino, M., Stock, J. H., & Watson, M. W. (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, *135*, 499–526. doi:10.1016/j.jeconom.2005.07.020.
- Natanelov, V., Alam, M. J., McKenzie, A. M., & Van Huynenbroeck, G. (2011). Is there co-movement of agricultural commodities futures prices and crude oil? *Energy Policy*, *39*, 4971–4984. doi:10.1016/j.enpol.2011.06.016.
- Nazlioglu, S., Erdem, C., & Soytas, U. (2013). Volatility spillover between oil and agricultural commodity markets. *Energy Economics*, *36*, 658–665. doi:10.1016/j.eneco.2012.11.009.
- Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the Equity Risk Premium: The Role of Technical Indicators. *Management Science*, *60*, 1772–1791. doi:10.1287/mnsc.2013.1838.
- Nerlove, M. (1956). Estimates of the Elasticities of Supply of Selected Agricultural Commodities. *Journal of Farm Economics*, *38*, 496. doi:10.2307/1234389.
- Nguyen, D. K., & Walther, T. (2019). Modeling and Forecasting Commodity Market Volatility with Long-term Economic and Financial Variables. *Journal of Forecasting*, *forth.*, . doi:10.1002/for.2617.
- Patton, A. J., & Sheppard, K. (2015). Good Volatility, Bad Volatility: Signed Jumps and The Persistence of Volatility. *Review of Economics and Statistics*, *97*, 683–697. doi:10.1162/REST-12-00503.
- Politis, D. N., & Romano, J. P. (1994). The Stationary Bootstrap. *Journal of the American Statistical Association*, *89*, 1303–1313. doi:10.1080/01621459.1994.10476870.
- Raggi, D., & Bordignon, S. (2012). Long memory and nonlinearities in realized volatility: A Markov switching approach. *Computational Statistics & Data Analysis*, *56*, 3730–3742. doi:10.1016/j.csda.2010.12.008.
- Schmidhuber, J. (2007). Biofuels: An Emerging Threat to Europe’s Food Security? *Notre Europe Policy Paper 27*, (pp. 1–40).

- Sethuraman, J. (1994). A Constructive Definition of Dirichlet Priors. *Statistica Sinica*, 4, 639–650.
- Sévi, B. (2014). Forecasting the volatility of crude oil futures using intraday data. *European Journal of Operational Research*, 235, 643–659. doi:10.1016/j.ejor.2014.01.019.
- Sizova, N. (2011). Integrated variance forecasting: Model based vs. reduced form. *Journal of Econometrics*, 162, 294–311. doi:10.1016/j.jeconom.2011.02.004.
- Song, Y. (2014). Modelling Regime Switching and Structural Breaks with an Infinite Hidden Markov Model. *Journal of Applied Econometrics*, 29, 825–842. doi:10.1002/jae.2337.
- Souček, M., & Todorova, N. (2013). Realized volatility transmission between crude oil and equity futures markets: A multivariate HAR approach. *Energy Economics*, 40, 586–597. doi:10.1016/j.eneco.2013.08.011.
- Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2006). Hierarchical Dirichlet Processes. *Journal of the American Statistical Association*, 101, 1566–1581. doi:10.1198/016214506000000302.
- Tian, F., Yang, K., & Chen, L. (2016). Realized Volatility Forecasting of Agricultural Commodity Futures Using Long Memory and Regime Switching. *Journal of Forecasting*, . doi:10.1002/for.2443.
- Todorova, N., Worthington, A., & Souček, M. (2014). Realized volatility spillovers in the non-ferrous metal futures market. *Resources Policy*, 39, 21–31. doi:10.1016/j.resourpol.2013.10.008.
- Tyner, W. E. (2010). The integration of energy and agricultural markets. *Agricultural Economics*, 41, 193–201. doi:10.1111/j.1574-0862.2010.00500.x.
- Van Gael, J., Saatci, Y., Teh, Y. W., & Ghahramani, Z. (2008). Beam sampling for the infinite hidden Markov model. In *Proceedings of the 25th international conference on Machine learning - ICML '08* (pp. 1088–1095). New York, New York, USA: ACM Press. doi:10.1145/1390156.1390293.
- Wen, F., Gong, X., & Cai, S. (2016). Forecasting the volatility of crude oil futures using HAR-type models with structural breaks. *Energy Economics*, 59, 400–413. doi:10.1016/j.eneco.2016.07.014.
- Witt, H. J., Schroeder, T. C., & Hayenga, M. L. (1987). Comparison of analytical approaches for estimating hedge ratios for agricultural commodities. *Journal of Futures Markets*, 7, 135–146. doi:10.1002/fut.3990070204.
- Yang, K., & Chen, L. (2014). Realized Volatility Forecast: Structural Breaks, Long Memory, Asymmetry, and Day-of-the-Week Effect. *International Review of Finance*, 14, 345–392. doi:10.1111/irfi.12030.
- Yang, K., Tian, F., Chen, L., & Li, S. (2017). Realized volatility forecast of agricultural futures using the HAR models with bagging and combination approaches. *International Review of Economics & Finance*, 49, 276–291. doi:10.1016/j.iref.2017.01.030.

Zhu, X.-H., Zhang, H.-W., & Zhong, M.-R. (2017). Volatility forecasting in Chinese nonferrous metals futures market. *Transactions of Nonferrous Metals Society of China*, 27, 1206–1214. doi:10.1016/S1003-6326(17)60141-9.

## 7. Tables

Table 1: Descriptive Statistics of the agricultural commodities of this study; Corn, Cotton, Indica Rice, Palm, and Soybeans.

Agriculture commodity	Mean	Standard Deviation	Skewness	Kurtosis	Jarque-Bera	LB Q(5)	LB Q(10)	LB Q(20)	ADF
Panel A: Daily Returns									
Corn	-0.030	1.182	-4.89	105.69	556889.74***	84.06	143.90	190.14*	-31.58***
Cotton	-0.075	1.101	-2.93	44.17	87492.75***	5.23	11.12	28.90	-25.83***
Indica Rice	-0.005	0.733	1.95	31.51	40586.82***	6.86	10.45	18.94	-23.00***
Palm Oil	-0.057	1.214	0.31	6.60	674.21***	2.56	3.80	10.88	-24.28***
Soybean	-0.018	0.877	-0.08	6.70	692.03***	2.57	11.17	20.29	-25.18***
Panel B: Realized Volatility									
Corn	0.231	0.306	4.85	37.70	67943.83***	1186.67***	1932.48***	2791.40***	-14.24***
Cotton	0.637	1.030	4.80	36.76	62300.92***	1376.24***	2232.77***	3626.86***	-13.06***
Indica Rice	0.686	2.924	19.59	483.47	11386890.41***	34.27***	53.89***	77.56***	-21.64***
Palm Oil	0.918	1.895	16.33	331.85	5524222.36***	95.35***	112.65***	139.77***	-21.47***
Soybean	0.459	0.509	3.62	22.42	21719.07***	761.81***	1202.67***	1734.76***	-15.40***
Panel C: Speculation Variable									
Corn	0.309	0.228	2.31	12.53	5869.32***	1338.78***	2137.89***	3027.59***	-13.38***
Cotton	0.884	1.195	2.63	9.53	3558.62***	5030.28***	9443.85***	17071.90***	-4.91***
Indica Rice	0.291	0.330	4.19	28.89	36277.70***	2157.24***	3888.59***	5768.60***	-11.57***
Palm Oil	1.086	0.516	0.97	4.33	278.19***	2556.19***	4404.98***	6830.32***	-10.01***
Soybean	0.600	0.433	2.35	12.11	5312.79***	1900.82***	2842.42***	3866.14***	-11.19***

*Note:* This table reports the statistical summary of the daily returns (Panel A), realized volatility (Panel B), and the speculation variable (Panel C) of each agricultural commodity future. The speculation variable is the ratio between the open interest and the trading volume. \*\*\*, \*\*, and \* refer to the level of significance of 1%, 5%, and 10%, respectively.

Table 2: Time-varying coefficients for the DIHM-HAR-spec model.

	Corn		Cotton		Indica Rice		Palm		Soybean	
	mean	std	mean	std	mean	std	mean	std	mean	std
$a_0$	0.0407	0.0390	0.0930	0.0979	0.0906	0.0554	0.2847	0.1412	0.0306	0.0158
$a_d$	0.0711	0.0705	0.0687	0.0756	0.0772	0.0491	0.1338	0.1012	0.0573	0.0435
$a_w$	0.1081	0.0674	0.0024	0.0526	0.0966	0.0975	0.0830	0.0690	0.1559	0.1052
$a_m$	0.1936	0.1105	0.2907	0.1328	0.2129	0.1554	0.0210	0.0397	0.2567	0.0704
$a_s$	0.3092	0.0434	0.2037	0.0528	0.2274	0.1939	0.2355	0.0413	0.3293	0.1950

*Note:* This table shows the mean and standard deviation (std) of the time-varying coefficients for the constant, the daily, weekly, and monthly volatility variables, and the speculation variable in the DIHM-HAR-Spec model.

Table 3: Time-varying coefficients for the IHM-HAR-spec model.

	Corn		Cotton		Indica Rice		Palm		Soybean	
	mean	std	mean	std	mean	std	mean	std	mean	std
$a_0$	0.1109	0.0857	0.2424	0.1693	0.2509	0.1913	0.4312	0.1791	0.2406	0.1581
$a_d$	0.0285	0.0116	0.0830	0.0765	0.0639	0.0444	0.0748	0.0514	0.0521	0.0202
$a_w$	0.0422	0.0278	0.0404	0.0327	0.0219	0.0241	0.0325	0.0116	0.0158	0.0468
$a_m$	0.0538	0.0631	0.1098	0.0688	0.0964	0.0762	0.0483	0.0240	0.0966	0.0436
$a_s$	0.2784	0.0382	0.1244	0.0250	0.1322	0.0654	0.1898	0.0532	0.2008	0.1061

*Note:* This table shows the mean and standard deviation (std) of the time-varying coefficients for the constant, the daily, weekly, and monthly volatility variables, and the speculation variable in the IHM-HAR-Spec model.

Table 4: Forecasting precision results for 1-step forecasts.

	Corn		Cotton		Indica Rice		Palm Oil		Soybean	
	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE
DIHM-HAR	0.345	0.174**	0.219	0.398**	0.006	1.231**	0.032	0.819**	0.177	0.320**
IHM-HAR	0.361	0.176**	0.181	0.406**	0.011	1.194**	0.025	0.793**	0.182	0.319**
IHMC-HAR	0.372	0.180**	0.175	0.408**	0.005	1.250**	0.026	0.822**	0.155	0.333
HAR	0.357	0.175**	0.159	0.445	0.002	1.271**	0.024	0.889	0.170	0.325
DIHM-HARCJ	0.358	0.174**	0.188	0.399**	0.004	1.215**	0.046	0.794**	0.184	0.316**
IHM-HARCJ	0.359	0.175**	0.185	0.404**	0.021	1.173**	0.035	0.782**	0.191	0.316**
IHMC-HARCJ	0.368	0.180**	0.182	0.405**	0.003	1.253*	0.049	0.779**	0.175	0.329
HARCJ	0.351	0.176**	0.146	0.453	0.002	1.393*	0.028	0.865	0.194	0.322**
DIHM-HARTCJ	0.347	0.172**	0.226	0.397**	0.013	1.203**	0.024	0.825	0.206	0.314**
IHM-HARTCJ	0.342	0.176**	0.188	0.403**	0.022	1.178**	0.040	0.783**	0.197	0.314**
IHMC-HARTCJ	0.338	0.179**	0.179	0.403**	0.003	1.260	0.045	0.78**	0.174	0.330
HARTCJ	0.354	0.177**	0.150	0.451**	0.001	1.307**	0.030	0.864	0.200	0.323**
DIHM-HARΔJ	0.377	0.172**	0.221	0.391**	0.007	1.218**	0.012	0.838	0.204	0.313**
IHM-HARΔJ	0.364	0.175**	0.209	0.395**	0.009	1.195**	0.024	0.793**	0.208	0.313**
IHMC-HARΔJ	0.377	0.178**	0.203	0.398**	0.004	1.246*	0.029	0.814	0.160	0.331
HARΔJ	0.363	0.173**	0.170	0.440	0.004	1.271*	0.031	0.875	0.207	0.317**
DIHM-HARspec	0.356	0.173**	0.213	0.395**	0.005	1.230**	0.015	0.812*	0.163	0.324**
IHM-HARspec	0.359	0.175**	0.179	0.398**	0.016	1.193**	0.028	0.791**	0.176	0.319**
IHMC-HARspec	0.361	0.185	0.173	0.410*	0.006	1.237**	0.027	0.814	0.153	0.334
HARspec	0.352	0.175**	0.160	0.424*	0.002	1.276**	0.024	0.885	0.168	0.324
DIHM-HARc	0.345	0.173**	0.206	0.394**	0.005	1.255*	0.017	0.802**	0.172	0.320**
IHM-HARc	0.361	0.173**	0.173	0.409*	0.012	1.198**	0.025	0.788**	0.176	0.318**
IHMC-HARc	0.377	0.178**	0.171	0.425*	0.007	1.249**	0.028	0.823*	0.162	0.332
HARc	0.356	0.174**	0.160	0.440	0.002	1.277*	0.025	0.886	0.172	0.325

*Note:* This table reports the evaluation results of the forecast precision based on MZ- $R^2$ , MAFE, and MCS results for the 1-step forecasts. The MAFE results marked with \* and \*\* refer to corresponding MCS p-values of > 10% and > 25%, respectively.

Table 5: Forecasting precision results for 5-step forecasts.

	Corn		Cotton		Indica Rice		Palm Oil		Soybean	
	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE
DIHM-HAR	0.627	0.087*	0.410	0.210**	0.052	0.877**	0.057	0.555	0.728	0.119*
IHM-HAR	0.797	0.071**	0.800	0.145**	0.058	0.832**	0.210	0.483**	0.783	0.109**
IHMC-HAR	0.487	0.148	0.210	0.357	0.017	1.047	0.034	0.775	0.178	0.277
HAR	0.496	0.131	0.233	0.375	0.008	1.003	0.038	0.768	0.177	0.254
DIHM-HARCJ	0.659	0.084**	0.413	0.224*	0.070	0.840**	0.024	0.566	0.620	0.128
IHM-HARCJ	0.763	0.073**	0.766	0.152**	0.051	0.842**	0.167	0.493**	0.734	0.111**
IHMC-HARCJ	0.500	0.149	0.142	0.385	0.023	1.094	0.068	0.684	0.215	0.270
HARCJ	0.500	0.132	0.325	0.376	0.005	1.172**	0.068	0.735	0.203	0.249
DIHM-HARTCJ	0.602	0.093*	0.416	0.212*	0.086	0.837**	0.047	0.549	0.589	0.131
IHM-HARTCJ	0.760	0.072**	0.790	0.147**	0.067	0.800**	0.219	0.476**	0.755	0.112**
IHMC-HARTCJ	0.458	0.154	0.172	0.367	0.009	1.096	0.053	0.707	0.203	0.271
HARTCJ	0.495	0.133	0.219	0.398	0.003	1.066*	0.083	0.729	0.208	0.248
DIHM-HARΔJ	0.607	0.095*	0.440	0.210*	0.095	0.847**	0.066	0.546	0.697	0.125*
IHM-HARΔJ	0.789	0.070	0.809	0.147**	0.070	0.824**	0.275	0.467**	0.734	0.114*
IHMC-HARΔJ	0.492	0.148	0.196	0.358	0.018	1.053	0.039	0.760	0.188	0.274
HARΔJ	0.506	0.131	0.217	0.379	0.019	0.981	0.053	0.746	0.193	0.249
DIHM-HARspec	0.652	0.087*	0.420	0.208**	0.051	0.886**	0.066	0.537	0.637	0.127
IHM-HARspec	0.794	0.071**	0.744	0.162**	0.039	0.858**	0.258	0.467**	0.754	0.108**
IHMC-HARspec	0.488	0.150	0.195	0.362	0.020	1.051	0.035	0.774	0.183	0.277
HARspec	0.499	0.131	0.248	0.346	0.006	1.014	0.042	0.764	0.180	0.254
DIHM-HARc	0.569	0.096*	0.651	0.186*	0.051	0.887**	0.043	0.555	0.726	0.123
IHM-HARc	0.691	0.080**	0.660	0.181*	0.031	0.897**	0.122	0.503*	0.807	0.105**
IHMC-HARc	0.505	0.151	0.223	0.369	0.031	1.013	0.036	0.778	0.186	0.272
HARc	0.498	0.132	0.240	0.373	0.008	1.007	0.042	0.760	0.181	0.254

*Note:* This table reports the evaluation results of the forecast precision based on MZ- $R^2$ , MAFE, and MCS results for the 5-step forecasts. The MAFE results marked with \* and \*\* refer to corresponding MCS p-values of > 10% and > 25%, respectively.

Table 6: Forecasting precision results for 22-step forecasts.

	Corn		Cotton		Indica Rice		Palm Oil		Soybean	
	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE
DIHM-HAR	0.962	0.028**	0.905	0.067**	0.397	0.469**	0.753	0.224**	0.870	0.052**
IHM-HAR	0.946	0.031**	0.922	0.060**	0.552	0.329**	0.752	0.187**	0.940	0.042**
IHMC-HAR	0.583	0.133	0.034	0.373	0.023	0.911	0.176	0.594	0.326	0.297
HAR	0.572	0.106	0.089	0.348	0.046	0.818	0.205	0.600	0.329	0.203
DIHM-HARCJ	0.952	0.030**	0.879	0.074**	0.427	0.451**	0.746	0.233*	0.818	0.054**
IHM-HARCJ	0.953	0.031**	0.887	0.068**	0.554	0.334**	0.732	0.192**	0.939	0.041**
IHMC-HARCJ	0.529	0.141	0.039	0.394	0.030	0.978	0.194	0.589	0.365	0.276
HARCJ	0.570	0.108	0.132	0.349	0.036	0.810	0.215	0.591	0.343	0.198
DIHM-HARTCJ	0.953	0.031**	0.843	0.082**	0.333	0.505**	0.779	0.225*	0.835	0.055**
IHM-HARTCJ	0.956	0.029**	0.912	0.065**	0.683	0.289**	0.819	0.184**	0.933	0.042**
IHMC-HARTCJ	0.540	0.138	0.046	0.383	0.061	0.922	0.167	0.591	0.351	0.282
HARTCJ	0.572	0.107	0.085	0.379	0.014	0.858	0.219	0.583	0.340	0.204
DIHM-HARΔJ	0.953	0.030**	0.885	0.071**	0.381	0.472**	0.799	0.194**	0.868	0.051**
IHM-HARΔJ	0.950	0.031**	0.931	0.062**	0.602	0.306**	0.768	0.201**	0.938	0.043**
IHMC-HARΔJ	0.588	0.133	0.032	0.375	0.025	0.911	0.170	0.595	0.330	0.296
HARΔJ	0.570	0.107	0.084	0.346	0.069	0.810	0.209	0.597	0.321	0.204
DIHM-HARspec	0.967	0.027**	0.893	0.069**	0.347	0.489**	0.812	0.197**	0.886	0.049**
IHM-HARspec	0.956	0.030**	0.936	0.058**	0.609	0.311**	0.796	0.191**	0.930	0.043**
IHMC-HARspec	0.583	0.131	0.034	0.373	0.035	0.885	0.176	0.593	0.323	0.302
HARspec	0.574	0.107	0.114	0.325	0.036	0.827	0.208	0.598	0.345	0.204
DIHM-HARc	0.940	0.035**	0.942	0.066**	0.419	0.451**	0.801	0.196**	0.892	0.048**
IHM-HARc	0.906	0.039*	0.880	0.080**	0.590	0.316**	0.801	0.176**	0.944	0.041**
IHMC-HARc	0.603	0.121	0.074	0.378	0.027	0.901	0.200	0.571	0.405	0.259
HARc	0.592	0.111	0.101	0.366	0.039	0.827	0.214	0.595	0.343	0.203

Note: This table reports the evaluation results of the forecast precision based on MZ- $R^2$ , MAFE, and MCS results for the 22-step forecasts. The MAFE results marked with \* and \*\* refer to corresponding MCS p-values of > 10% and > 25%, respectively.

Table 7: Results of the economic evaluation for 1-step forecasts.

	Corn			Cotton			Indica Rice			Palm Oil			Soybean		
	$r$	EV		$r$	EV		$r$	EV		$r$	EV		$r$	EV	
		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$
DIHM-HAR	20.381	1.176	0.127	14.388	2.404	0.260	2.804	-0.822	-0.089	5.707	0.656	0.071	5.862	0.151	0.016
IHM-HAR	20.677	1.469	0.159	12.984	0.965	0.104	3.758	0.199	0.021	5.532	0.479	0.052	5.729	0.016	0.002
IHMC-HAR	25.739	6.659	0.722	18.113	6.299	0.681	4.332	3.679	0.397	5.461	0.406	0.044	6.559	0.858	0.093
HAR	19.223	0.000	0.000	12.041	0.000	0.000	3.655	0.000	0.000	5.058	0.000	0.000	5.712	0.000	0.000
DIHM-HARCJ	21.562	1.798	0.195	14.410	2.257	0.244	4.242	1.104	0.119	5.516	0.356	0.038	5.890	0.206	0.022
IHM-HARCJ	20.642	0.892	0.097	12.496	0.287	0.031	4.650	0.459	0.050	5.725	0.568	0.061	5.740	0.055	0.006
IHMC-HARCJ	25.131	5.498	0.596	17.162	5.123	0.554	10.243	8.481	2.706	5.876	0.720	0.078	6.550	0.876	0.095
HARCJ	19.772	0.000	0.000	12.205	0.000	0.000	3.877	0.000	0.000	5.163	0.000	0.000	5.686	0.000	0.000
DIHM-HARTCJ	21.405	1.170	0.127	15.050	2.601	0.281	4.462	0.331	0.036	5.566	0.347	0.037	5.988	0.115	0.012
IHM-HARTCJ	21.274	1.024	0.111	13.133	0.626	0.068	3.852	-0.576	-0.062	5.655	0.436	0.047	5.764	-0.111	-0.012
IHMC-HARTCJ	26.017	5.943	0.645	18.075	5.761	0.623	-3.655	3.205	-0.755	5.773	0.555	0.060	6.668	0.805	0.087
HARTCJ	20.287	0.000	0.000	12.516	0.000	0.000	3.804	0.000	0.000	5.222	0.000	0.000	5.874	0.000	0.000
DIHM-HARDJ	20.402	2.344	0.254	14.576	2.722	0.294	3.992	0.858	0.093	5.393	-0.084	-0.009	5.979	0.387	0.042
IHM-HARDJ	20.151	2.104	0.228	11.970	0.028	0.003	3.991	0.382	0.041	5.465	-0.011	-0.001	5.638	0.041	0.004
IHMC-HARDJ	24.330	6.399	0.694	6.543	-6.170	-0.666	4.589	3.283	0.354	5.436	-0.042	-0.004	6.702	1.121	0.121
HARDJ	18.081	0.000	0.000	11.927	0.000	0.000	4.096	0.000	0.000	5.474	0.000	0.000	5.597	0.000	0.000
DIHM-HARspec	21.217	1.364	0.148	14.526	1.439	0.155	4.534	1.323	0.143	5.347	0.384	0.041	5.944	0.168	0.018
IHM-HARspec	20.971	1.109	0.120	13.236	0.132	0.014	3.851	0.330	0.036	5.529	0.570	0.061	5.767	-0.012	-0.001
IHMC-HARspec	27.350	7.655	0.831	19.004	6.129	0.663	2.953	2.847	0.307	5.465	0.503	0.054	6.641	0.874	0.094
HARspec	19.863	0.000	0.000	13.105	0.000	0.000	3.667	0.000	0.000	4.965	0.000	0.000	5.778	0.000	0.000
DIHM-HARc	21.286	1.685	0.183	16.518	3.429	0.371	3.822	0.394	0.042	5.330	0.347	0.037	5.850	0.218	0.023
IHM-HARc	21.154	1.564	0.170	15.436	2.321	0.251	3.930	0.175	0.019	5.585	0.607	0.065	5.739	0.106	0.011
IHMC-HARc	25.277	5.808	0.630	15.247	2.122	0.229	4.076	2.411	0.260	5.412	0.430	0.046	6.429	0.805	0.087
HARc	19.624	0.000	0.000	13.190	0.000	0.000	3.630	0.000	0.000	4.985	0.000	0.000	5.634	0.000	0.000

Note: This table shows the results of economic evaluation in terms of the portfolio return and the economic values of various HAR models for the 1-step forecasts, where  $r$  refers to the portfolio return given a mild risk aversion rate ( $\gamma = 1$ ) and EV denotes the economic values of different HAR models with infinite Hidden Markov switching structures against the corresponding benchmark HAR model.



Table 8: Results of the economic evaluation for 5-step forecasts.

	Corn			Cotton			Indica Rice			Palm Oil			Soybean		
	$r$	EV		$r$	EV		$r$	EV		$r$	EV		$r$	EV	
		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$
DIHM-HAR	9.238	-4.151	-0.448	10.140	-3.650	-0.394	3.454	-3.203	-0.345	4.589	0.317	0.034	4.529	-0.032	-0.003
IHM-HAR	8.335	-5.053	-0.546	10.766	-3.975	-0.429	6.992	-0.689	-0.074	4.921	0.655	0.071	4.490	-0.072	-0.008
IHMC-HAR	17.104	3.891	0.421	14.834	7.896	0.853	3.038	3.106	0.335	4.447	0.174	0.019	4.806	0.247	0.027
HAR	13.328	0.000	0.000	10.336	0.000	0.000	3.151	0.000	0.000	4.274	0.000	0.000	4.561	0.000	0.000
DIHM-HARCJ	9.171	-4.192	-0.453	9.993	-3.847	-0.415	3.521	-3.133	-0.338	2.556	-1.720	-0.186	4.471	-0.099	-0.011
IHM-HARCJ	8.545	-4.818	-0.520	10.842	-3.672	-0.397	3.208	-4.144	-0.447	4.609	0.400	0.043	4.505	-0.064	-0.007
IHMC-HARCJ	16.846	3.660	0.396	14.755	7.176	0.776	8.732	8.317	0.898	4.508	0.296	0.032	4.815	0.249	0.027
HARCJ	13.302	0.000	0.000	10.362	0.000	0.000	3.132	0.000	0.000	4.215	0.000	0.000	4.569	0.000	0.000
DIHM-HARTCJ	8.280	-5.100	-0.551	10.445	-3.706	-0.400	3.209	-3.428	-0.369	5.213	1.002	0.108	4.514	-0.197	-0.021
IHM-HARTCJ	8.077	-5.289	-0.571	10.897	-4.245	-0.458	3.715	-3.933	-0.424	4.495	0.276	0.030	4.510	-0.201	-0.022
IHMC-HARTCJ	16.960	3.777	0.409	14.626	6.712	0.725	3.300	2.799	0.302	4.437	0.216	0.023	4.788	0.080	0.009
HARTCJ	13.302	0.000	0.000	10.744	0.000	0.000	3.185	0.000	0.000	4.222	0.000	0.000	4.709	0.000	0.000
DIHM-HAR $\Delta$ J	9.723	-3.371	-0.364	10.139	-3.229	-0.349	2.941	-3.171	-0.342	4.466	0.169	0.018	4.522	-0.034	-0.004
IHM-HAR $\Delta$ J	8.307	-4.793	-0.518	10.838	-3.755	-0.406	3.768	-3.599	-0.388	4.778	0.486	0.052	4.504	-0.052	-0.006
IHMC-HAR $\Delta$ J	16.667	3.738	0.405	14.680	7.363	0.796	3.265	3.083	0.332	4.411	0.113	0.012	4.784	0.231	0.025
HAR $\Delta$ J	13.040	0.000	0.000	10.496	0.000	0.000	3.202	0.000	0.000	4.298	0.000	0.000	4.556	0.000	0.000
DIHM-HARspec	9.382	-3.875	-0.419	10.247	-4.614	-0.498	4.042	-2.307	-0.249	4.628	0.383	0.041	4.545	-0.030	-0.003
IHM-HARspec	8.271	-4.989	-0.539	10.870	-5.104	-0.551	2.193	-5.289	-0.570	4.733	0.490	0.053	4.440	-0.136	-0.015
IHMC-HARspec	17.032	3.954	0.428	14.910	6.466	0.699	3.214	3.169	0.342	4.440	0.193	0.021	4.799	0.227	0.024
HARspec	13.199	0.000	0.000	11.634	0.000	0.000	3.108	0.000	0.000	4.249	0.000	0.000	4.575	0.000	0.000
DIHM-HARc	9.914	-3.447	-0.373	10.838	-3.824	-0.413	0.265	-5.619	-0.605	5.397	1.174	0.127	4.539	0.038	0.004
IHM-HARc	8.681	-4.687	-0.506	11.377	-4.453	-0.481	3.944	-3.125	-0.337	4.562	0.329	0.035	4.476	-0.026	-0.003
IHMC-HARc	16.734	3.540	0.383	13.916	4.515	0.488	3.059	2.386	0.257	4.433	0.198	0.021	4.670	0.171	0.018
HARc	13.305	0.000	0.000	11.843	0.000	0.000	3.046	0.000	0.000	4.237	0.000	0.000	4.501	0.000	0.000

Note: This table shows the results of economic evaluation in terms of the portfolio return and the economic values of various HAR models for the 5-step forecasts, where  $r$  refers to the portfolio return given a mild risk aversion rate ( $\gamma = 1$ ) and EV denotes the economic values of different HAR models with infinite Hidden Markov switching structures against the corresponding benchmark HAR model.

Table 9: Results of the economic evaluation for 22-step forecasts.

	Corn			Cotton			Indica Rice			Palm Oil			Soybean		
	$r$	EV		$r$	EV		$r$	EV		$r$	EV		$r$	EV	
		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$
DIHM-HAR	6.517	-1.734	-0.187	9.383	-1.539	-0.166	2.932	-0.841	-0.091	4.065	0.197	0.021	4.153	-0.135	-0.015
IHM-HAR	6.418	-1.834	-0.198	9.338	-1.702	-0.184	2.829	-1.062	-0.115	4.016	0.147	0.016	4.125	-0.163	-0.018
IHMC-HAR	10.021	1.845	0.199	13.452	4.421	0.478	2.881	0.964	0.104	4.041	0.173	0.019	4.580	0.297	0.032
HAR	8.223	0.000	0.000	9.966	0.000	0.000	2.799	0.000	0.000	3.870	0.000	0.000	4.286	0.000	0.000
DIHM-HARCJ	6.562	-1.549	-0.167	9.378	-1.305	-0.141	2.867	-0.784	-0.085	4.047	0.148	0.016	4.141	-0.175	-0.019
IHM-HARCJ	6.409	-1.702	-0.184	9.312	-1.577	-0.170	3.020	-0.837	-0.090	4.055	0.157	0.017	4.119	-0.198	-0.021
IHMC-HARCJ	10.334	2.310	0.250	12.009	3.055	0.330	1.069	-1.035	-0.112	4.055	0.156	0.017	4.585	0.274	0.030
HARCJ	8.085	0.000	0.000	9.887	0.000	0.000	2.822	0.000	0.000	3.900	0.000	0.000	4.315	0.000	0.000
DIHM-HARTCJ	6.505	-1.613	-0.174	9.339	-1.853	-0.200	2.968	-0.824	-0.089	4.079	0.211	0.023	4.145	-0.224	-0.024
IHM-HARTCJ	6.379	-1.741	-0.188	9.333	-1.924	-0.208	3.046	-0.809	-0.087	4.085	0.217	0.023	4.126	-0.244	-0.026
IHMC-HARTCJ	10.783	2.767	0.299	13.597	4.492	0.486	3.340	1.604	0.173	4.027	0.158	0.017	4.774	0.412	0.044
HARTCJ	8.092	0.000	0.000	10.248	0.000	0.000	2.840	0.000	0.000	3.870	0.000	0.000	4.367	0.000	0.000
DIHM-HARDJ	6.433	-1.819	-0.196	9.366	-1.705	-0.184	2.990	-0.961	-0.104	4.069	0.192	0.021	4.147	-0.138	-0.015
IHM-HARDJ	6.415	-1.837	-0.198	9.331	-1.779	-0.192	3.000	-0.987	-0.106	4.062	0.184	0.020	4.125	-0.160	-0.017
IHMC-HARDJ	10.001	1.824	0.197	13.344	4.281	0.463	2.931	0.962	0.104	4.044	0.166	0.018	4.583	0.303	0.033
HARDJ	8.223	0.000	0.000	9.970	0.000	0.000	2.829	0.000	0.000	3.880	0.000	0.000	4.283	0.000	0.000
DIHM-HARspec	6.385	-1.764	-0.190	9.361	-2.725	-0.294	2.902	-0.951	-0.103	4.078	0.223	0.024	4.134	-0.170	-0.018
IHM-HARspec	6.380	-1.769	-0.191	9.322	-2.782	-0.300	3.015	-0.854	-0.092	4.078	0.223	0.024	4.133	-0.171	-0.018
IHMC-HARspec	9.952	1.878	0.203	13.827	3.720	0.402	2.774	0.858	0.093	4.021	0.164	0.018	4.634	0.335	0.036
HARspec	8.122	0.000	0.000	10.981	0.000	0.000	2.739	0.000	0.000	3.858	0.000	0.000	4.303	0.000	0.000
DIHM-HARc	6.517	-1.574	-0.170	9.809	-4.347	-0.469	2.750	-0.830	-0.090	4.060	0.192	0.021	4.144	-0.084	-0.009
IHM-HARc	6.462	-1.629	-0.176	9.726	-4.457	-0.481	2.929	-0.682	-0.074	4.085	0.217	0.023	4.124	-0.105	-0.011
IHMC-HARc	9.282	1.247	0.135	12.489	-0.193	-0.021	2.708	0.590	0.064	4.000	0.131	0.014	4.307	0.080	0.009
HARc	8.065	0.000	0.000	13.134	0.000	0.000	2.601	0.000	0.000	3.870	0.000	0.000	4.227	0.000	0.000

Note: This table shows the results of economic evaluation in terms of the portfolio return and the economic values of various HAR models for the 22-step forecasts, where  $r$  refers to the portfolio return given a mild risk aversion rate ( $\gamma = 1$ ) and EV denotes the economic values of different HAR models with infinite Hidden Markov switching structures against the corresponding benchmark HAR model.

Table 10: Robustness check for the 1-step forecasting precision results.

	Corn		Cotton		Indica Rice		Palm Oil		Soybean	
	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE
DIHM-HAR	0.201	0.187**	0.109	0.403**	0.024	0.605**	0.019	0.813**	0.083	0.371**
IHM-HAR	0.195	0.195	0.093	0.428	0.038	0.604**	0.034	0.800**	0.088	0.369**
IHMC-HAR	0.216	0.196	0.089	0.403**	0.028	0.610**	0.032	0.808**	0.076	0.390
HAR	0.203	0.194*	0.083	0.471	0.034	0.621**	0.023	0.866*	0.073	0.385
DIHM-HARCJ	0.223	0.187**	0.112	0.401**	0.007	0.635**	0.045	0.793**	0.075	0.373**
IHM-HARCJ	0.201	0.193	0.095	0.419*	0.011	0.632**	0.036	0.803**	0.095	0.371**
IHMC-HARCJ	0.225	0.195*	0.082	0.405*	0.000	0.651	0.044	0.802**	0.095	0.387
HARCJ	0.212	0.194**	0.068	0.504	0.009	0.692	0.028	0.848**	0.109	0.386
DIHM-HARTCJ	0.214	0.186**	0.090	0.401**	0.005	0.629**	0.028	0.810**	0.105	0.370**
IHM-HARTCJ	0.187	0.192*	0.102	0.418*	0.020	0.615**	0.037	0.798**	0.093	0.370**
IHMC-HARTCJ	0.205	0.193**	0.108	0.398**	0.000	0.647	0.042	0.804**	0.092	0.387
HARTCJ	0.211	0.195**	0.095	0.489	0.003	0.646**	0.031	0.847*	0.098	0.385
DIHM-HARΔJ	0.216	0.187**	0.216	0.398**	0.007	0.643**	0.016	0.828*	0.081	0.372**
IHM-HARΔJ	0.198	0.193	0.099	0.423*	0.009	0.642**	0.024	0.807**	0.091	0.370**
IHMC-HARΔJ	0.217	0.196	0.097	0.396**	0.001	0.640**	0.028	0.812**	0.098	0.387
HARΔJ	0.200	0.196	0.092	0.488	0.002	0.660	0.036	0.837**	0.098	0.382**
DIHM-HARspec	0.223	0.188**	0.109	0.404*	0.013	0.628**	0.020	0.818**	0.059	0.375**
IHM-HARspec	0.203	0.193	0.106	0.407*	0.012	0.634**	0.038	0.805**	0.081	0.371**
IHMC-HARspec	0.207	0.199	0.103	0.404*	0.014	0.623**	0.038	0.807**	0.068	0.393
HARspec	0.209	0.191**	0.095	0.458	0.008	0.662*	0.027	0.851**	0.070	0.384
DIHM-HARc	0.202	0.190**	0.093	0.399**	0.010	0.609**	0.016	0.815**	0.082	0.373**
IHM-HARc	0.187	0.192**	0.082	0.415*	0.031	0.606**	0.029	0.795**	0.071	0.368**
IHMC-HARc	0.220	0.191**	0.090	0.455	0.034	0.608**	0.035	0.803**	0.077	0.384*
HARc	0.203	0.191**	0.080	0.456	0.037	0.619**	0.024	0.870*	0.074	0.387

Note: Robustness check of the statistical evaluation for 1-step forecast based on the realized kernel estimator. The MAFE results marked with \* and \*\* refer to corresponding MCS  $p$ -values of  $> 10\%$  and  $> 25\%$ , respectively.

Table 11: Robustness check for the economic evaluation for 1-step forecasts.

	Corn			Cotton			Indica Rice			Palm Oil			Soybean		
	$r$	EV		$r$	EV		$r$	EV		$r$	EV		$r$	EV	
		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$
DIHM-HAR	18.688	0.582	0.063	13.872	0.582	0.063	2.868	0.582	0.063	6.249	5.383	0.582	0.063	18.688	0.582
IHM-HAR	16.818	-0.194	-0.021	12.656	-0.194	-0.021	2.919	-0.194	-0.021	5.995	4.908	-0.194	-0.021	16.818	-0.194
IHMC-HAR	25.589	4.121	0.444	20.301	4.121	0.444	2.903	4.121	0.444	6.648	6.098	4.121	0.444	25.589	4.121
HAR	15.004	0.000	0.000	10.508	0.000	0.000	2.942	0.000	0.000	5.681	4.602	0.000	0.000	15.004	0.000
DIHM-HARCJ	17.245	0.618	0.067	13.824	0.618	0.067	2.979	0.618	0.067	6.910	5.125	0.618	0.067	17.245	0.618
IHM-HARCJ	16.497	0.473	0.051	12.437	0.473	0.051	2.818	0.473	0.051	6.788	4.947	0.473	0.051	16.497	0.473
IHMC-HARCJ	23.779	2.757	0.297	19.144	2.757	0.297	2.582	2.757	0.297	7.557	5.713	2.757	0.297	23.779	2.757
HARCJ	14.540	0.000	0.000	10.918	0.000	0.000	2.830	0.000	0.000	6.180	-6.712	0.000	0.000	14.540	0.000
DIHM-HARTCJ	18.423	0.373	0.040	14.805	0.373	0.040	3.232	0.373	0.040	6.719	5.228	0.373	0.040	18.423	0.373
IHM-HARTCJ	17.075	-0.233	-0.025	12.408	-0.233	-0.025	2.878	-0.233	-0.025	6.702	5.001	-0.233	-0.025	17.075	-0.233
IHMC-HARTCJ	25.764	3.184	0.343	19.826	3.184	0.343	3.552	3.184	0.343	7.600	5.925	3.184	0.343	25.764	3.184
HARTCJ	15.140	0.000	0.000	11.312	0.000	0.000	2.947	0.000	0.000	6.067	4.687	0.000	0.000	15.140	0.000
DIHM-HARΔJ	17.312	1.763	0.190	13.824	2.975	0.322	2.885	1.763	0.190	6.188	5.175	1.763	0.190	17.312	1.763
IHM-HARΔJ	16.136	1.396	0.150	12.417	1.396	0.150	2.939	1.396	0.150	6.289	4.915	1.396	0.150	16.136	1.396
IHMC-HARΔJ	23.101	4.069	0.438	29.219	4.069	0.438	2.834	4.069	0.438	6.698	5.728	4.069	0.438	23.101	4.069
HARΔJ	13.029	0.000	0.000	10.742	0.000	0.000	2.703	0.000	0.000	6.562	4.381	0.000	0.000	13.029	0.000
DIHM-HARspec	17.988	0.988	0.106	15.125	0.988	0.106	2.928	0.988	0.106	6.154	5.176	0.988	0.106	17.988	0.988
IHM-HARspec	17.150	0.728	0.078	12.649	0.728	0.078	2.915	0.728	0.078	6.214	5.013	0.728	0.078	17.150	0.728
IHMC-HARspec	26.172	5.224	0.563	20.401	5.224	0.563	2.830	5.224	0.563	6.857	5.880	5.224	0.563	26.172	5.224
HARspec	14.104	0.000	0.000	11.998	0.000	0.000	2.917	0.000	0.000	5.654	4.708	0.000	0.000	14.104	0.000
DIHM-HARc	20.951	0.410	0.044	17.853	0.410	0.044	2.953	0.410	0.044	6.230	5.279	0.410	0.044	20.951	0.410
IHM-HARc	17.489	-0.407	-0.044	14.720	-0.407	-0.044	3.078	-0.407	-0.044	6.319	4.803	-0.407	-0.044	17.489	-0.407
IHMC-HARc	23.599	2.049	0.221	16.858	2.049	0.221	2.979	2.049	0.221	6.552	5.840	2.049	0.221	23.599	2.049
HARc	15.903	0.000	0.000	12.781	0.000	0.000	2.946	0.000	0.000	5.654	4.534	0.000	0.000	15.903	0.000

Note: Robustness check of the economic evaluation for 1-step forecast based on the realized kernel estimators, where  $r$  refers to the portfolio return given a mild risk aversion rate ( $\gamma = 1$ ) and EV denotes the economic values of different HAR models with infinite Hidden Markov switching structures against the corresponding benchmark HAR model.

Table 12: Robustness check for the 1-step forecasting precision results based on the MedRV estimator.

	Corn		Cotton		Indica Rice		Palm Oil		Soybean	
	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE	MZ- $R^2$	MAFE
DIHM-HAR	0.377	0.142**	0.159	0.290**	0.000	0.675**	0.189	0.371*	0.218	0.240**
IHM-HAR	0.360	0.144*	0.186	0.290**	0.001	0.654**	0.191	0.367*	0.220	0.235**
IHMC-HAR	0.387	0.145*	0.175	0.294**	0.000	0.708**	0.244	0.364*	0.219	0.251**
HAR	0.373	0.144*	0.167	0.311*	0.000	0.749**	0.233	0.370*	0.218	0.238**
DIHM-HARCJ	0.417	0.139**	0.209	0.297*	0.000	0.666**	0.266	0.364*	0.196	0.243*
IHM-HARCJ	0.387	0.143*	0.259	0.290**	0.003	0.642**	0.179	0.374*	0.227	0.234**
IHMC-HARCJ	0.410	0.146*	0.215	0.306	0.000	0.650**	0.253	0.372*	0.217	0.254**
HARCJ	0.396	0.142**	0.182	0.332	0.000	0.823*	0.219	0.380*	0.216	0.240**
DIHM-HARTCJ	0.395	0.139**	0.208	0.286**	0.001	0.643**	0.248	0.368*	0.238	0.237**
IHM-HARTCJ	0.097	0.239	0.227	0.293**	0.002	0.65**	0.218	0.366*	0.229	0.234**
IHMC-HARTCJ	0.383	0.142**	0.238	0.292**	0.000	0.645**	0.237	0.377*	0.236	0.253
HARTCJ	0.404	0.141**	0.198	0.323	0.000	0.781**	0.212	0.384*	0.242	0.240**
DIHM-HAR $\Delta$ J	0.419	0.137**	0.162	0.291**	0.001	0.635**	0.233	0.372*	0.216	0.234**
IHM-HAR $\Delta$ J	0.385	0.141**	0.197	0.292**	0.001	0.643**	0.210	0.368*	0.222	0.235**
IHMC-HAR $\Delta$ J	0.417	0.144**	0.172	0.298*	0.000	0.666**	0.254	0.366*	0.235	0.241**
HAR $\Delta$ J	0.403	0.140**	0.166	0.313	0.000	0.674**	0.265	0.369*	0.238	0.237**
DIHM-HARspec	0.428	0.139**	0.157	0.289**	0.000	0.679**	0.221	0.366*	0.238	0.234**
IHM-HARspec	0.392	0.141**	0.155	0.286**	0.000	0.649**	0.186	0.366*	0.203	0.235**
IHMC-HARspec	0.399	0.150	0.168	0.296**	0.000	0.694**	0.231	0.366*	0.234	0.242**
HARspec	0.401	0.140**	0.168	0.290**	0.000	0.758**	0.234	0.370*	0.225	0.237**
DIHM-HARc	0.363	0.139**	0.161	0.288**	0.001	0.654**	0.189	0.367*	0.234	0.235**
IHM-HARc	0.372	0.142**	0.157	0.289**	0.003	0.628**	0.177	0.366*	0.221	0.235**
IHMC-HARc	0.393	0.145*	0.170	0.304*	0.000	0.715**	0.245	0.362**	0.236	0.240**
HARc	0.373	0.143**	0.167	0.306*	0.000	0.758**	0.236	0.370*	0.234	0.238**

Note: Robustness check of the statistical evaluation for 1-step forecast based on the realized kernel estimator. The MAF results marked with \* and \*\* refer to corresponding MCS  $p$ -values of  $> 10\%$  and  $> 25\%$ , respectively.

Table 13: Robustness check for the economic evaluation for 1-step forecasts based on the MedRV estimator.

	Corn			Cotton			Indica Rice			Palm Oil			Soybean		
	$r$	EV		$r$	EV		$r$	EV		$r$	EV		$r$	EV	
		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$		$\gamma = 1$	$\gamma = 10$
DIHM-HAR	24.040	1.586	0.172	16.287	2.327	0.252	4.477	0.342	0.037	6.226	0.339	0.037	7.875	0.376	0.041
IHM-HAR	22.775	0.280	0.030	14.503	0.495	0.053	3.978	-0.469	-0.051	6.079	0.191	0.021	7.223	-0.286	-0.031
IHMC-HAR	28.032	5.666	0.615	24.654	11.108	1.203	4.606	3.501	0.377	6.294	0.409	0.044	8.849	1.370	0.148
HAR	22.486	0.000	0.000	14.012	0.000	0.000	4.091	0.000	0.000	5.888	0.000	0.000	7.505	0.000	0.000
DIHM-HARCJ	24.036	1.778	0.193	16.150	1.561	0.169	6.693	2.473	0.267	6.053	0.090	0.010	8.013	0.444	0.048
IHM-HARCJ	23.045	0.762	0.083	14.212	-0.442	-0.048	5.343	0.307	0.033	6.067	0.106	0.012	7.217	-0.364	-0.039
IHMC-HARCJ	27.649	5.436	0.590	20.951	6.582	0.712	4.986	3.632	0.391	6.414	0.456	0.049	9.041	1.495	0.161
HARCJ	22.318	0.000	0.000	14.608	0.000	0.000	4.343	0.000	0.000	5.962	0.000	0.000	7.576	0.000	0.000
DIHM-HARTCJ	24.404	1.351	0.147	16.953	1.822	0.197	6.009	1.000	0.108	6.097	0.107	0.012	7.597	0.444	0.048
IHM-HARTCJ	69.746	47.962	5.269	15.298	0.125	0.013	4.107	2.546	-1.545	6.125	0.137	0.015	6.986	-0.178	-0.019
IHMC-HARTCJ	28.119	5.141	0.558	21.626	6.747	0.730	4.972	2.364	0.255	6.417	0.431	0.047	8.412	1.275	0.138
HARTCJ	23.150	0.000	0.000	15.147	0.000	0.000	4.182	0.000	0.000	5.988	0.000	0.000	7.161	0.000	0.000
DIHM-HAR $\Delta$ J	23.297	2.158	0.234	15.288	0.702	0.076	4.621	0.506	0.054	-0.973	-6.713	-0.723	7.365	0.477	0.051
IHM-HAR $\Delta$ J	21.864	0.718	0.078	14.173	-0.437	-0.047	4.569	-0.882	-0.095	5.487	-0.002	0.000	6.995	0.099	0.011
IHMC-HAR $\Delta$ J	26.453	5.392	0.585	13.996	-1.011	-0.109	3.755	1.632	0.176	6.046	0.565	0.061	8.194	1.321	0.143
HAR $\Delta$ J	21.163	0.000	0.000	14.571	0.000	0.000	5.033	0.000	0.000	5.486	0.000	0.000	6.896	0.000	0.000
DIHM-HARspec	24.723	1.884	0.204	17.139	1.845	0.199	4.998	1.252	0.135	5.948	0.088	0.009	7.185	0.206	0.022
IHM-HARspec	23.410	0.545	0.059	14.687	-0.673	-0.073	0.559	-3.804	-0.410	6.075	0.217	0.023	7.054	0.074	0.008
IHMC-HARspec	29.830	7.069	0.768	20.999	5.870	0.635	4.694	4.780	0.515	6.305	0.449	0.048	8.103	1.141	0.123
HARspec	22.868	0.000	0.000	15.338	0.000	0.000	4.223	0.000	0.000	5.858	0.000	0.000	6.982	0.000	0.000
DIHM-HARc	24.284	1.338	0.145	19.198	4.419	0.478	13.815	8.931	0.964	5.957	0.113	0.012	7.054	0.218	0.023
IHM-HARc	22.910	-0.099	-0.011	18.203	3.376	0.365	5.090	0.690	0.074	6.003	0.160	0.017	6.917	0.079	0.009
IHMC-HARc	28.160	5.301	0.575	16.268	1.338	0.145	4.887	2.956	0.318	6.239	0.398	0.043	7.971	1.152	0.124
HARc	22.994	0.000	0.000	14.957	0.000	0.000	4.097	0.000	0.000	5.843	0.000	0.000	6.840	0.000	0.000

Note: Robustness check of the economic evaluation for 1-step forecast based on the realized kernel estimators, where  $r$  refers to the portfolio return given a mild risk aversion rate ( $\gamma = 1$ ) and EV denotes the economic values of different HAR models with infinite Hidden Markov switching structures against the corresponding benchmark HAR model.

8. Figures

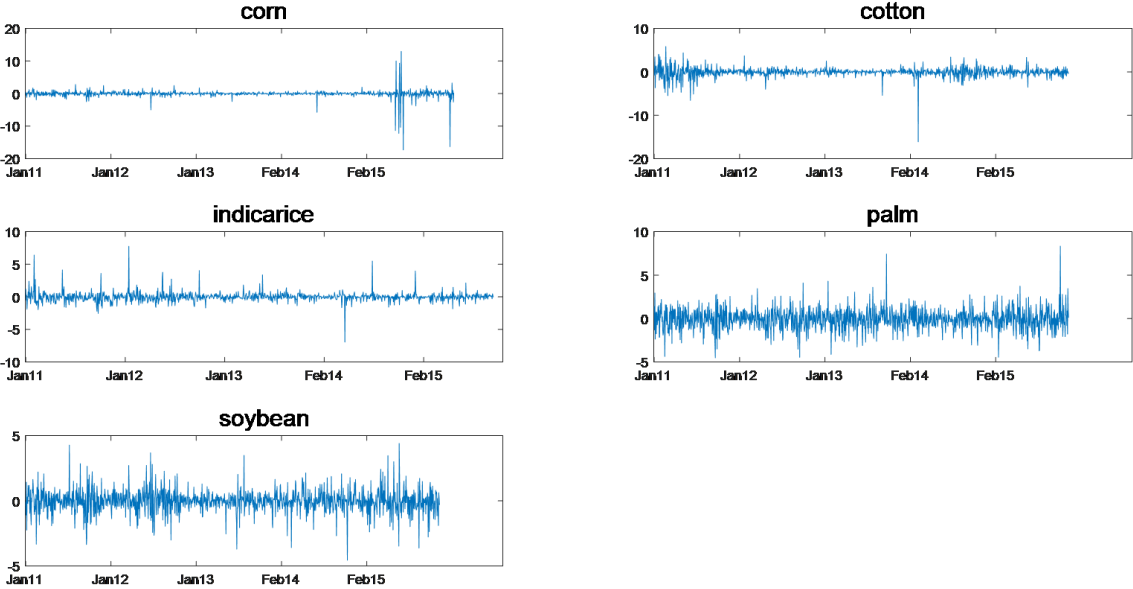


Figure 1: Daily returns of the agricultural commodity futures from January 1, 2011 to December 31, 2015.

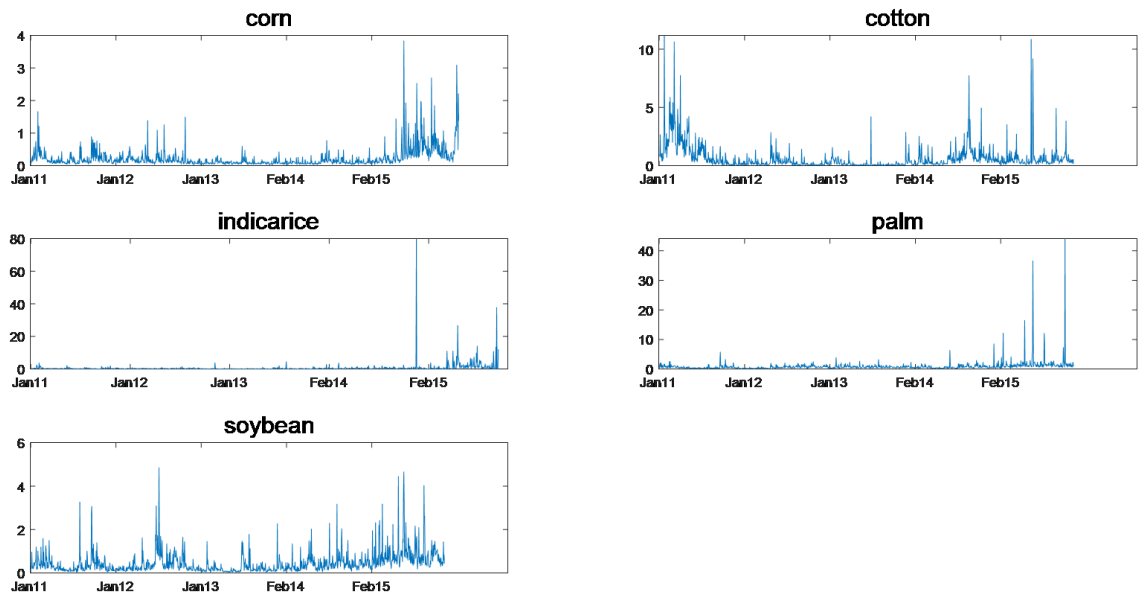


Figure 2: Daily realized volatilities of the agricultural commodity futures from January 1, 2011 to December 31, 2015.

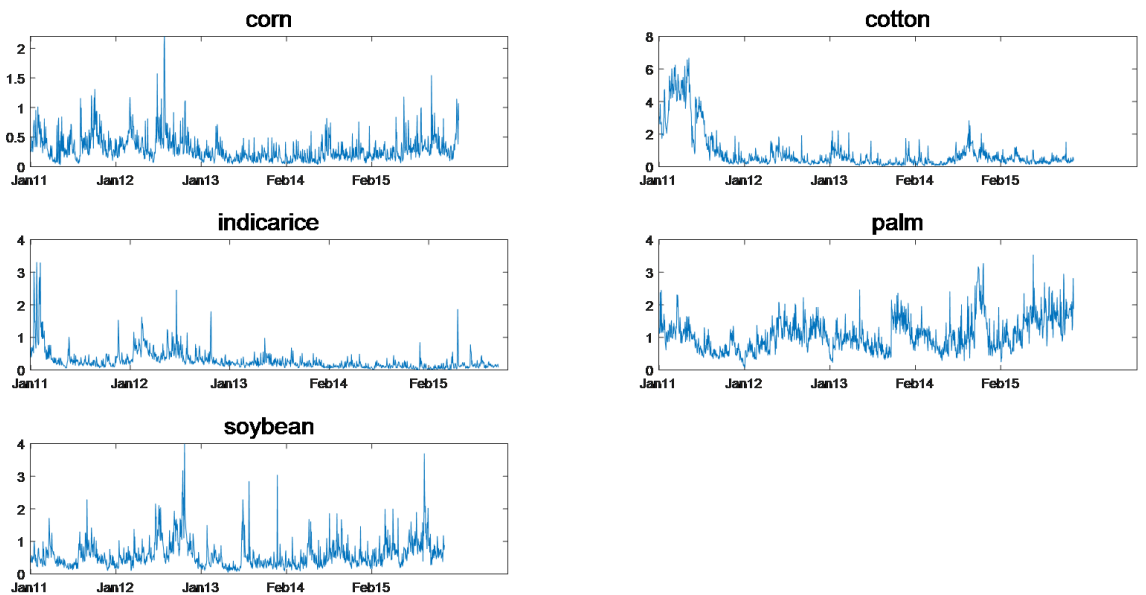
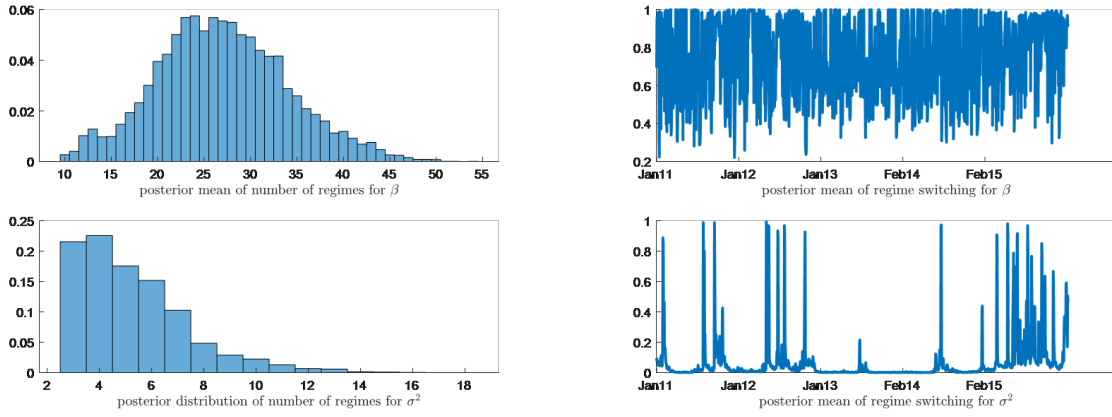
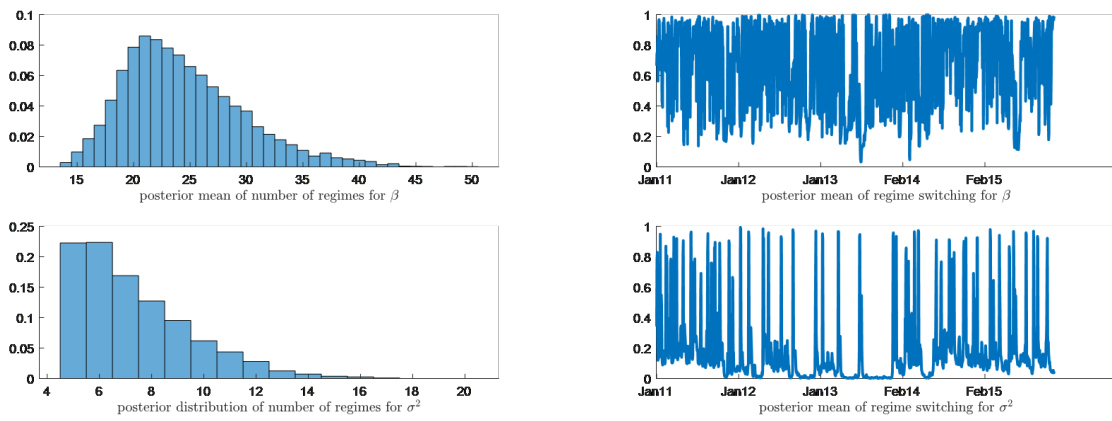


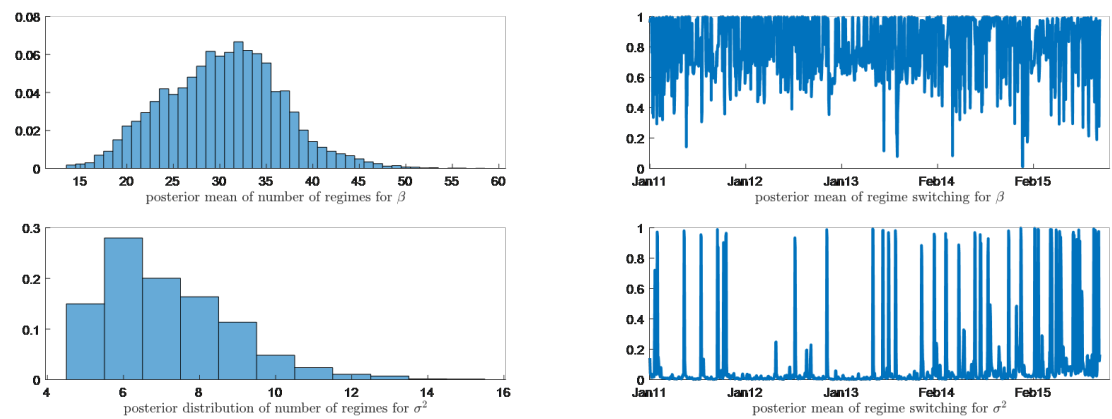
Figure 3: Speculation variable of the agricultural commodity futures from January 1, 2011 to December 31, 2015.



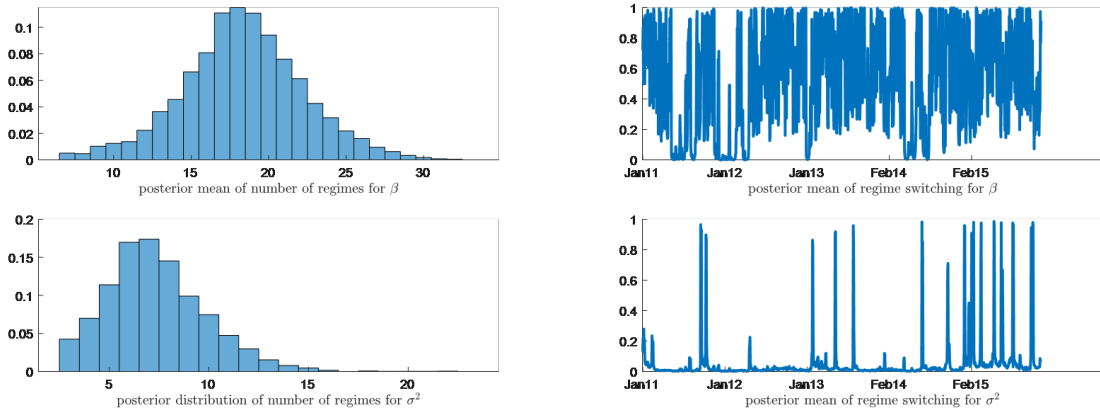
(a) Corn



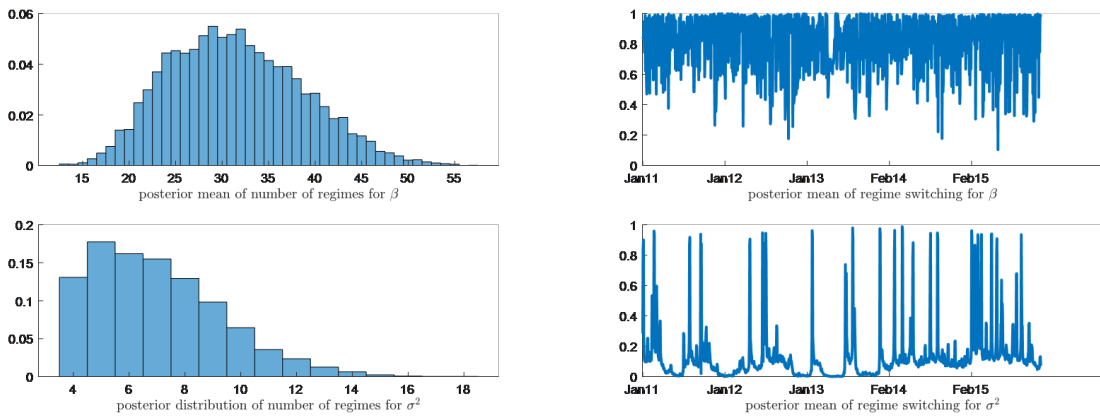
(b) Cotton



(c) Indica Rice

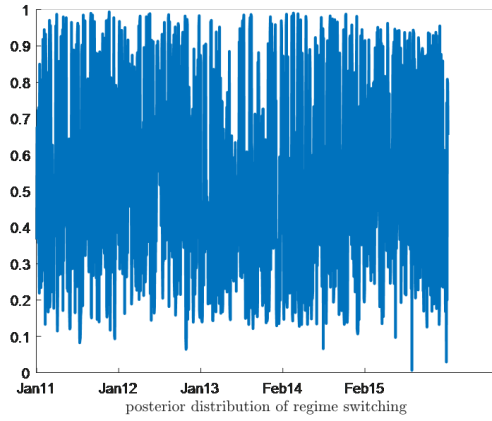
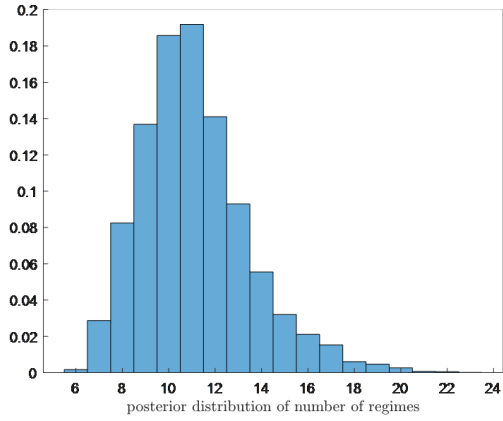


(d) Palm

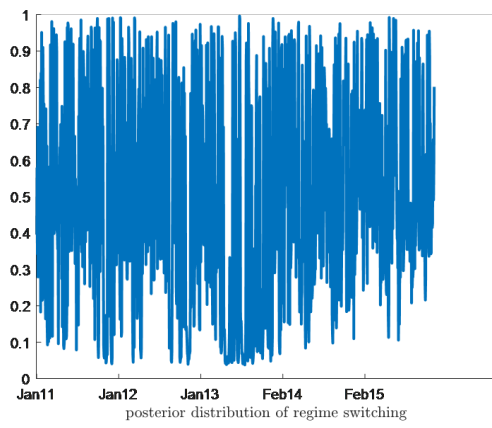
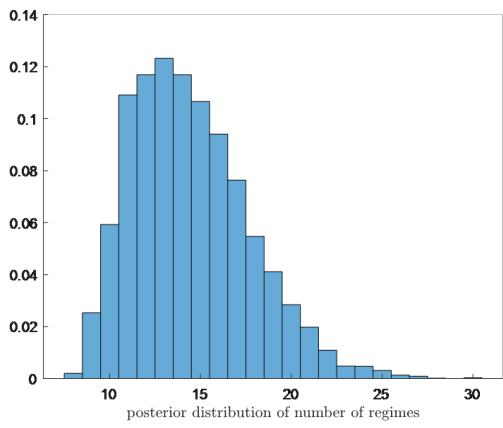


(e) Soybeans

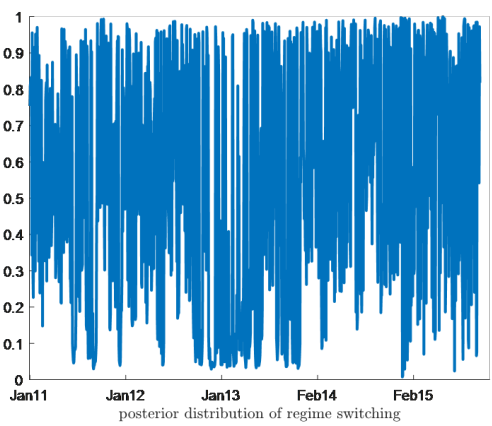
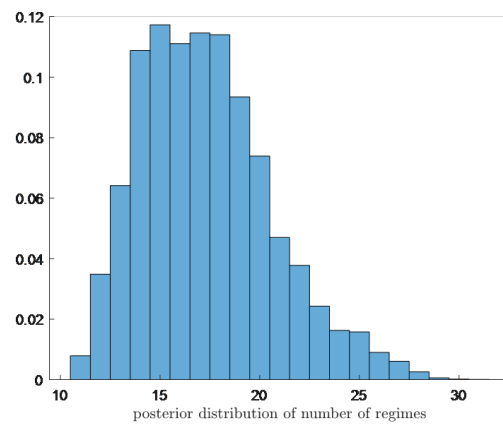
Figure 4: The posterior mean and distribution for the DIHM-HAR model.



(a) Corn

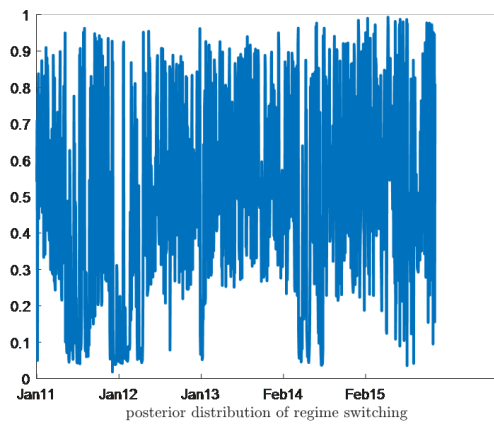
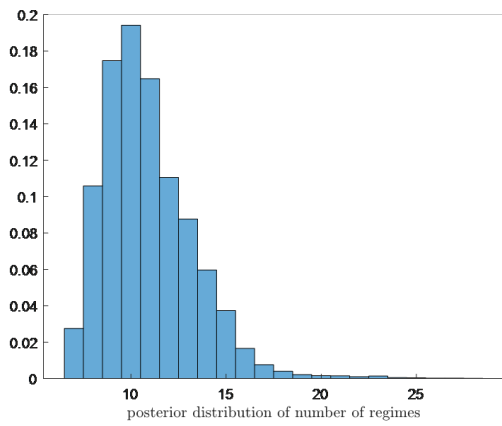


(b) Cotton

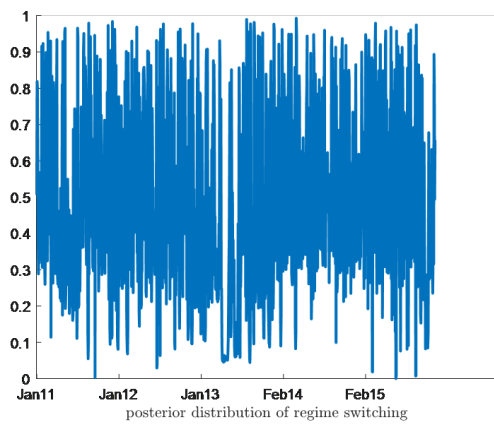
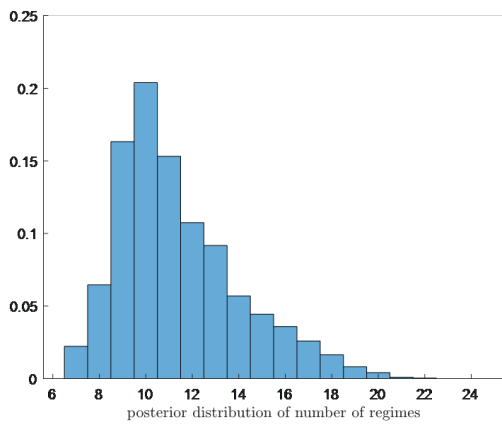


(c) Indica Rice





(d) Palm



(e) Soybean

Figure 5: The posterior mean and distribution for the IHM-HAR model.

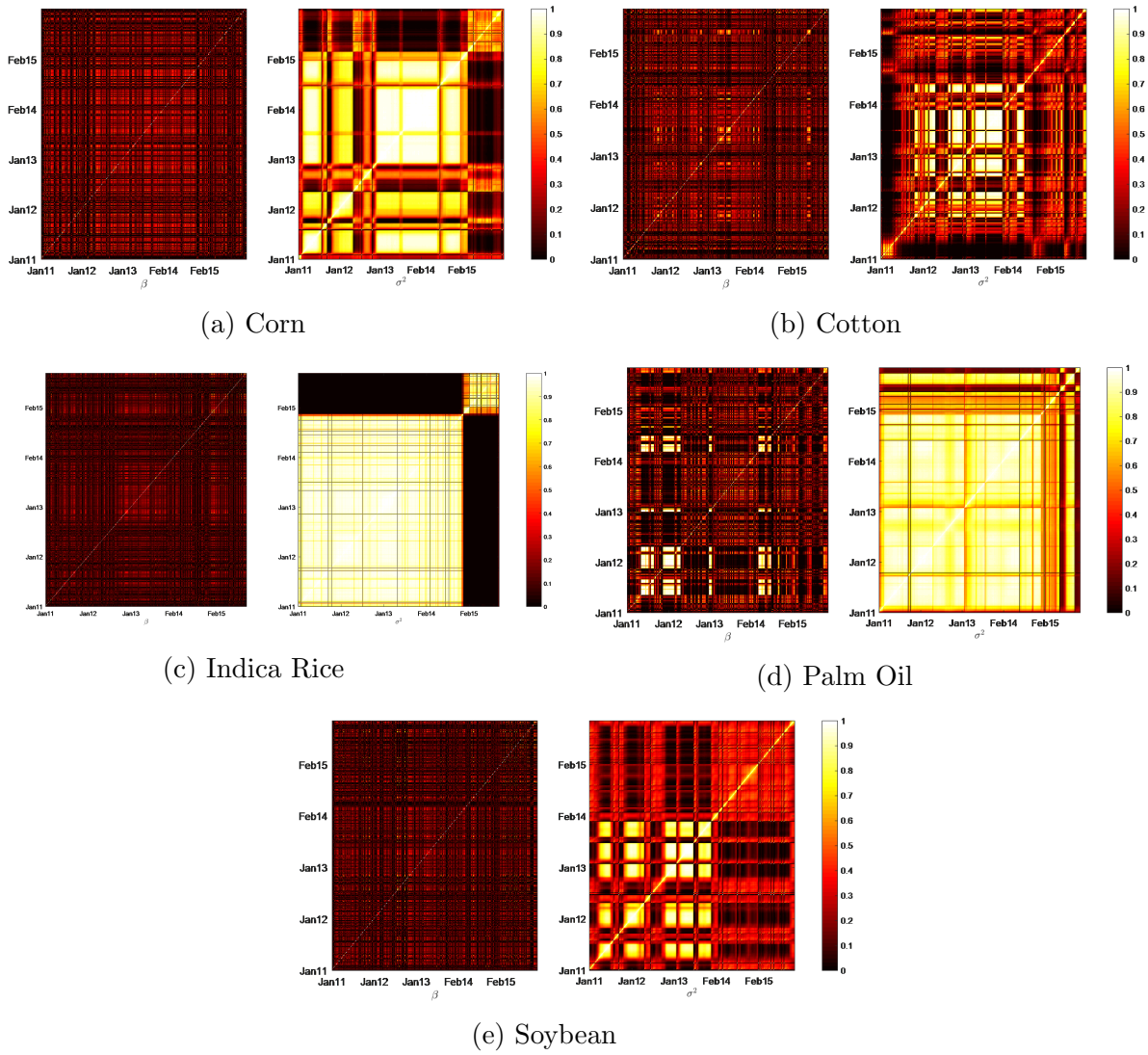
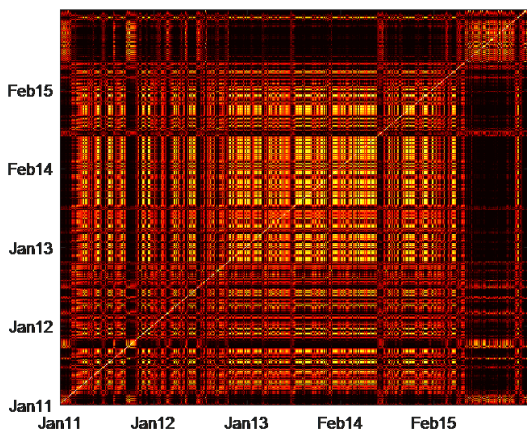
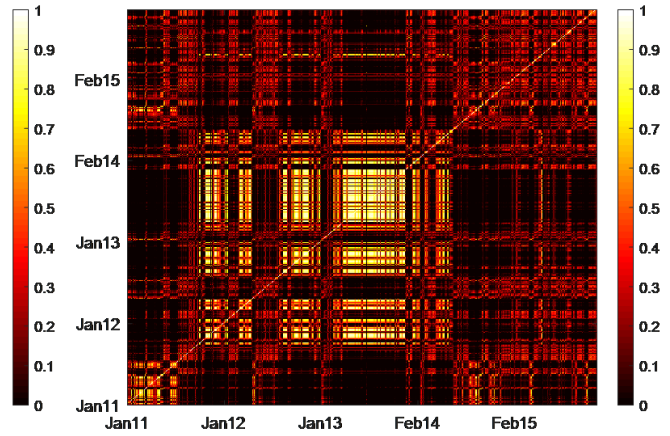


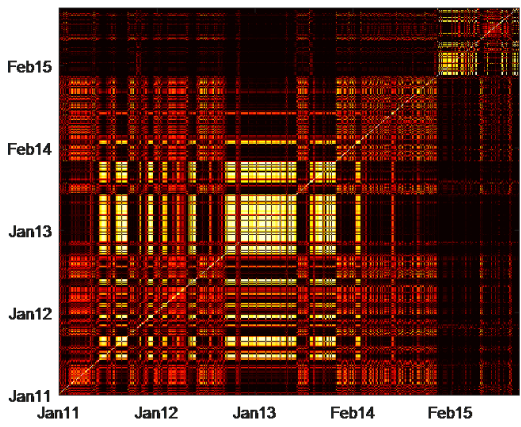
Figure 6: Heat maps for coefficients  $\beta$  (left plot of each subplot (a) - (e)) and variances  $\sigma^2$  (right plot of each subplot (a) - (e)) in the DIHM-HAR model for (a) Corn, (b) Cotton, (c) Indica Rice, (d) Palm Oil, and (e) Soybean.



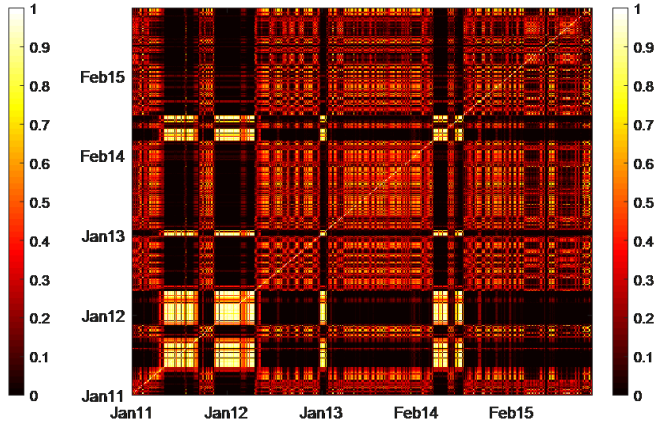
(a) Corn



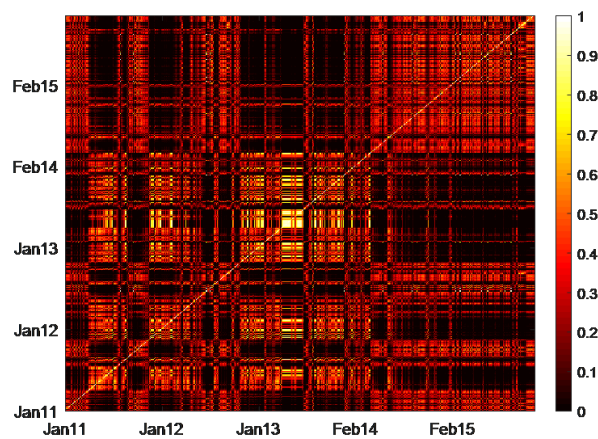
(b) Cotton



(c) Indica Rice



(d) Palm Oil



(e) Soybeans

Figure 7: Heat map in the IHM-HAR model for different agricultural commodity futures: (a) Corn, (b) Cotton, (c) Indica Rice, (d) Palm Oil, and (e) Soybeans.