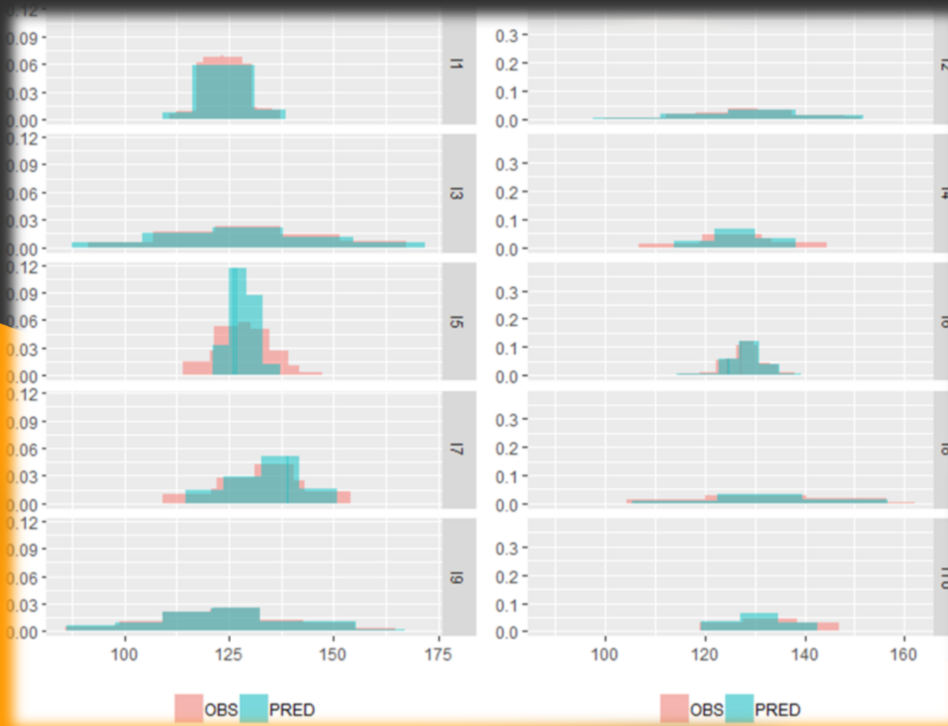
Hedging Agriculture Commodities Futures with Histogram data: A Markov Switching Volatility and Correlation model



Presented by Dr. Woraphon Yamaka
Center of Excellence in Econometric,
Faculty of Economics

Agenda

Introduction

01 This paper is about the hedging strategy. How we determine the optimal weight in two assets namely Spot and Futures?

02 In this study, we aim forecast hedge ratio and portfolio weight in the agriculture commodity portfolio, say wheat.

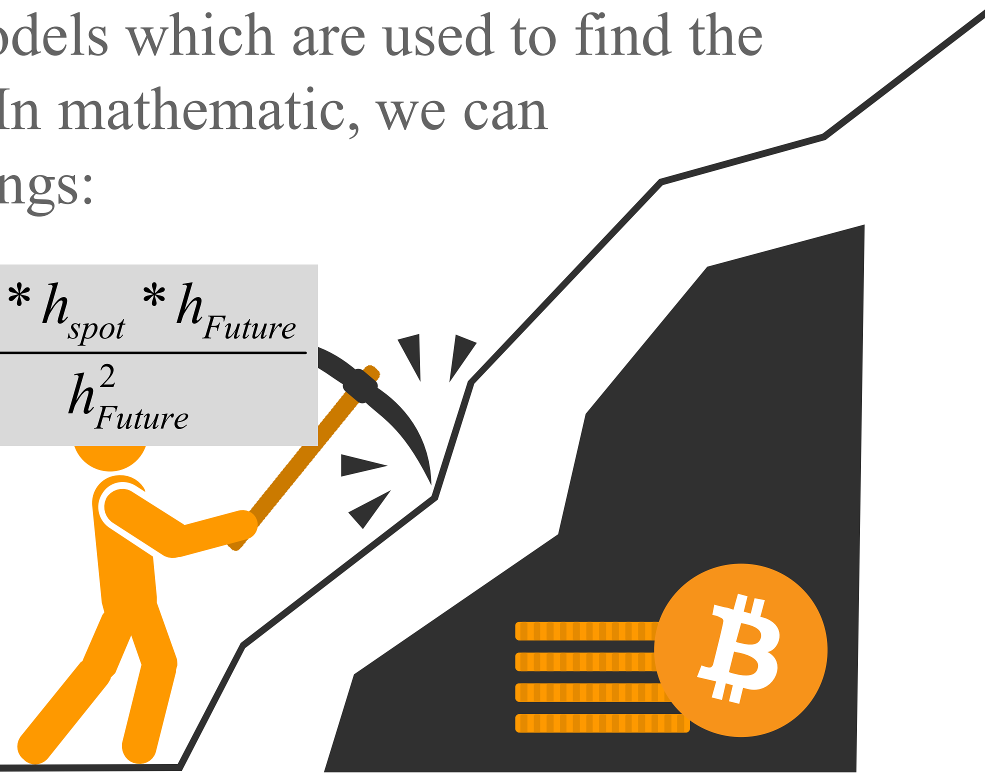
Introduction

03

There are various econometric models which are used to find the hedge ratio and portfolio weight. In mathematic, we can measure these twos as the followings:

$$Hedge-ratio=\frac{Cov(spot,Future)}{Var(Future)}=\frac{\rho * h_{spot} * h_{Future}}{h_{Future}^{2}}$$

$$Weight=\frac{h_{Future}-h_{spot,Future}}{h_{spot}-2h_{spot,Future}+h_{Future}},$$



Introduction

4 Various models are proposed to find the correlation and the volatility

Volatility Model

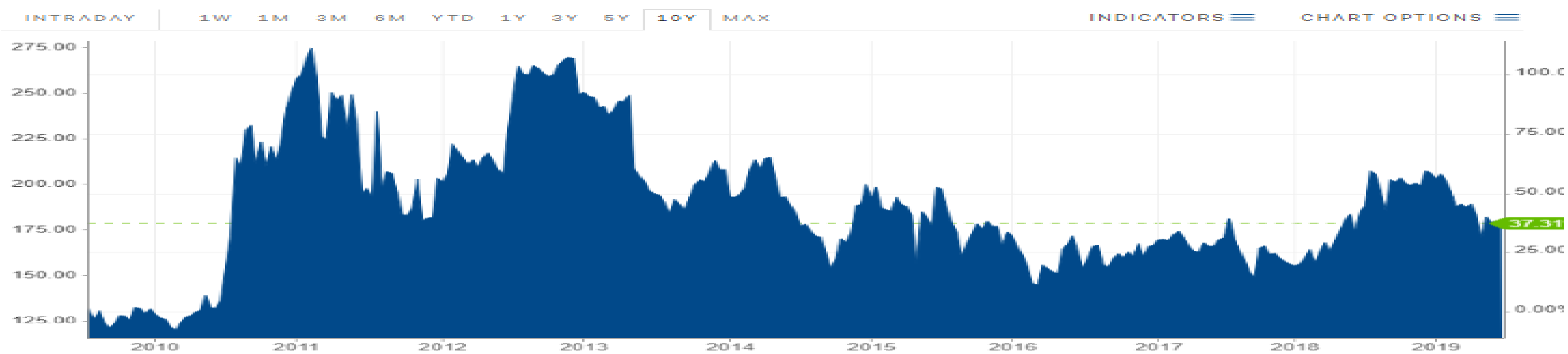
ARCH and GARCH types model, Stochastic Volatility

Correlation Model

CCC,DCC, Copula

What we are doing?

First of all, we want to improve those econometric models by incorporating the structural change of the Volatility and Correlation.



First Contribution

Volatility Model | ARCH and GARCH types model, Stochastic Volatility

+

Correlation Model | CCC, DCC, Copula

+

Structural Change Model | Markov Switching

First Contribution

The latest version of the model in nowadays is

Markov Switching Dynamic Copula-GARCH

The advantages of this model are

1. Capturing and explaining Structural Change in both volatility and correlation
2. Flexible to all the complicated dependence structures.

However, this model poses some limitations.

1. The model does not allow the GARCH parameters to be change between the regimes(states of economy)

First Contribution

Markov Switching Dynamic Copula-GARCH

$$h(U_1, U_2) = c_t(U_1, U_2 \left| \Theta_{t-1,s_t} \right.) \prod_{i=1}^{2} f_i(\Psi_{y_{it}}^{-1} \left| \theta \right.),$$

The flexible Markov Switching Dynamic Copula-GARCH

$$h(U_1, U_2) = c_t(U_1, U_2 \left| \Theta_{t-1,s_t} \right.) \prod_{i=1}^{2} f_i(\Psi_{y_{it}}^{-1} \left| \theta_{s_t} \right.),$$

Our first contribution is that we generalize the Markov Switching Dynamic Copula-GARCH

Second Contribution

In the general, we can use the daily data to observe the volatility and correlation between the assets.

However, the close daily price might not reflect the real behavior of data set. There exist some information during trading day. Regarding these information, we may lack of the real behavior of the data.

Fortunately, after the advent of the big data in the financial field, we can collect a tick data from various sources. (5-minute data). Therefore, this study we consider to use the tick data to find the optimal weight and hedge ratio of the agricultural spot and future.

Second Contribution

The advantage of Big Data

We gain more insight of the data behavior

What is the problem?

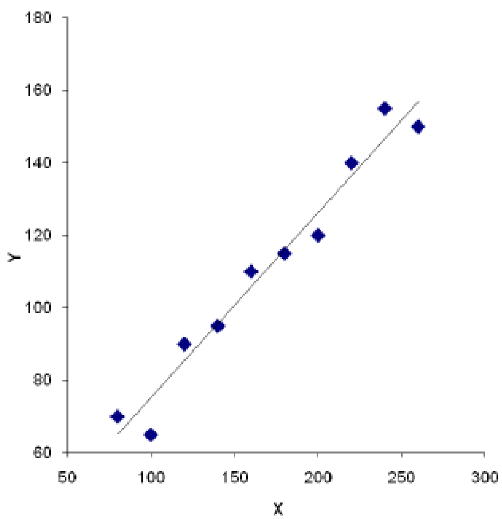
High dimensionality combined with large sample size creates issues such as heavy computational cost and algorithmic instability

Our second contribution is that we develop effective methods that can accurately deal with this high frequency data.

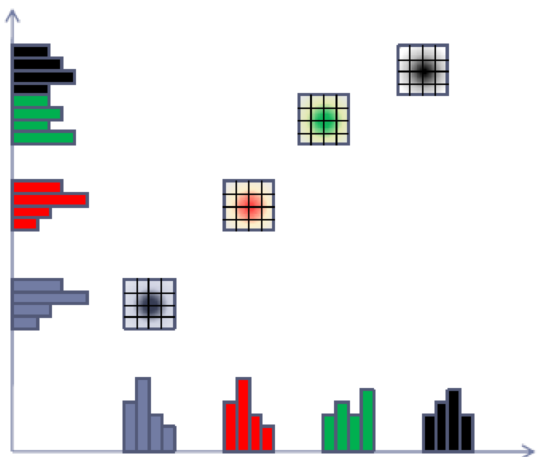
Second Contribution

To achieve second contribution, we consider use the histogram-valued data approach (Bilard and Deday, 2006)

Classic data



Histogram data



Contributions of The paper

Our first contribution is that we generalize the Markov Switching Dynamic Copula-GARCH

Our second contribution is that we develop effective methods that can accurately deal with this high frequency data.

We then apply this proposed model to fit the histogram-valued data for quantifying the hedge ratio and portfolio weight. In this study, we consider wheat spot and future as the empirical example.

Methodology

1) Find the representative data in the histogram data
2) Models: CCC, DCC, MS-DCC, MS-CCC, MS dynamic copula and our flexible MS dynamic copula
3) Hedge ratio and optimal portfolio weight

Methodology

Histogram value

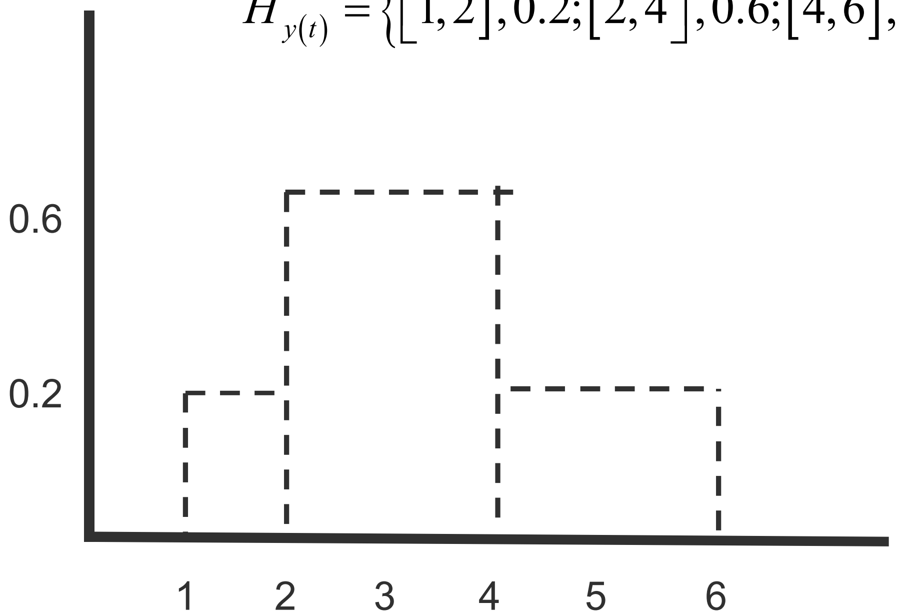
According to Dias and Brito (2013), we can define histogram-valued variables as follows:

$$H_{y(t)}=\left\{\left[\underline{I}_{y(t)1},\overline{I}_{y(t)1}\right],p_{t1};\left[\underline{I}_{y(t)2},\overline{I}_{y(t)2}\right],p_{t2};\ldots\left[\underline{I}_{y(t)n_t},\overline{I}_{y(t)n_t}\right],p_{tn_t}\right\}$$

where $\underline{I}_{y(t)i}$ and $\overline{I}_{y(t)i}$ represent the lower and upper bound of the $y(t)$ at each unit $i\in\{1,2,\ldots,n_j\}$ which are uniform distributed. p_{ti} is the frequency associated to the subinterval $\left[\underline{I}_{y(t)i},\overline{I}_{Y(t)i}\right]$ and $\sum_{i=1}^{n_j}p_{ij}=1$, n_t is the number of subintervals for the empirical distribution

Example of the histogram

$$H_{y(t)} = \{[1,2],0.2;[2,4],0.6;[4,6],0.2\}$$



Probability density function of the data

Methodology

Quantile function

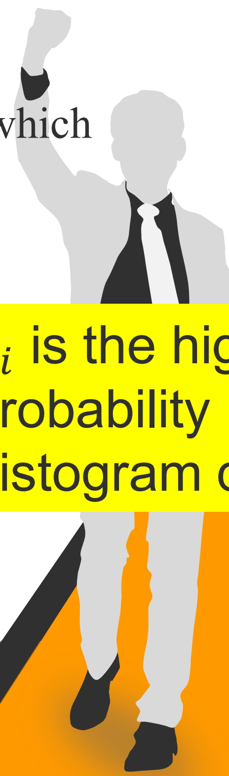
To find the single value data in each y(t), the quantile function, $\Psi^{-1}_{Y_t}$,which proposed in Irpino and Verde (2018) is employed.

$$\Psi^{-1}_{Y_{ti}} = \begin{cases} \underline{I}_{y(t)i1} + \frac{k_i}{w_{ti1}} a_{y(t)i1} & \text{if} \quad 0 \le k < w_{ti1} \\ \underline{I}_{y(t)i2} + \frac{t_i - w_{ti1}}{w_{ti2}} a_{y(t)i2} & \text{if} \quad w_{ti1} \le k_i < (w_{ti2}) \\ \vdots & \\ \underline{I}_{y(t)in_t} + \frac{t_i - (w_{tin_t})}{w_{tin}} a_{y(t)in_t} & \text{if} \quad (w_{tin_t}) \le k_i \le 1 \end{cases}$$

***

k_i

k_i is the highest probability in the histogram density



$a_{y(t)i} = \overline{I}_{y(t)i} - \underline{I}_{y(t)i}$ is range of each sub interval

$\Psi^{-1}_{Y_t}$ is the quantile function

Methodology

Markov Switching CCC-GARCH and DCC-GARCH

The Markov switching GARCH model is a nonlinear specification model which reflects different states of the volatilities namely high and low volatilities. The Markov-switching was introduced by Hamilton(1989). Changes in regimes can be estimated with the multiple regimes as (s_t) . The MS-DCC-GARCH (1,1) model can be written as

$$H_{t,s_t} = D_{t,s_t} R_{s_t} D_{t,s_t}$$

$$h_{t,s_t} = \alpha_{0,s_t} + \sum_{i=1}^{q} \alpha_{i,s_t} \varepsilon_{t-i,s_t}^2 + \sum_{j=1}^{p} \beta_{j,s_t} h_{t-j,s_t}$$

$$R_{s_t} = Q_{s_t}^{*-1} Q_{s_t} Q_{s_t}^{*-1}$$

$$Q_{s_t,t} = \left(1-\theta_{1,s_t} - \theta_{2,s_t}\right)\bar{Q}_{s_t,t} + \theta_{2,s_t} Q_{s_t,t-1} + \theta_{2,s_t} \varepsilon_{s_t,t-1} \varepsilon'_{s_t,t-1},$$

where $D_{t,s_t} = diag\left(h_{1t,s_t}, h_{2t,s_t}\right)$ is 2 x 2 regime dependent constant R_{s_t} correlation parameters of spot and futures returns.

,

Methodology

Markov Switching Dynamic Copula GARCH

Different form MS-CCC-GARCH and MS-DCC-GARCH, the time varying dependence equation is constructed using ARMA(1, m) process as suggested by Patton(2006). The time varying equation in our model can be written as

$$\rho_{t,s_t} = \Lambda\left(\omega_{0,s_t} + \omega_{1,s_t}\rho_{t-1} + \omega_{2,s_t}\Gamma_{t-m}\right)$$

where ρ_{t,s_t} is time varying dependence parameter which is the linear correlation coefficient ρ and depend on state or regime s_t

Methodology

The hedging ratio

We measure the hedge ratio and also the optimal portfolio weight in terms of the expected value. Then, the expected hedge ratio and expected weight of foreign exchange rates spot/futures holding are as follows

$$\delta^* = \frac{Cov\left(\Delta\Psi^{-1}_{y_{Spot,t}}, \Delta\Psi^{-1}_{y_{Futures,t}}\right)}{Var\left(\Delta\Psi^{-1}_{y_{Futures,t}}\right)} = \frac{Cov\left(r_{\Psi^{-1}_{y_{Spot,t}}}, r_{\Psi^{-1}_{y_{Futures,t}}}\right)}{Var\left(\Delta r_{\Psi^{-1}_{y_{Futures,t}}}\right)}.$$

The optimal portfolio weight of commodity spot/futures holding is given by:

$$w_{\Psi^{-1}_{y_{Spot,t}}\Psi^{-1}_{y_{Future,t}}} = \frac{h_{\Psi^{-1}_{y_{Future,t}}} - h_{\Psi^{-1}_{y_{Spot,t}}\Psi^{-1}_{y_{Future,t}}}}{h_{\Psi^{-1}_{y_{Spot,t}}} - 2h_{\Psi^{-1}_{y_{Spot,t}}\Psi^{-1}_{y_{Future,t}}} + h_{\Psi^{-1}_{y_{Future,t}}}},$$

Data Description

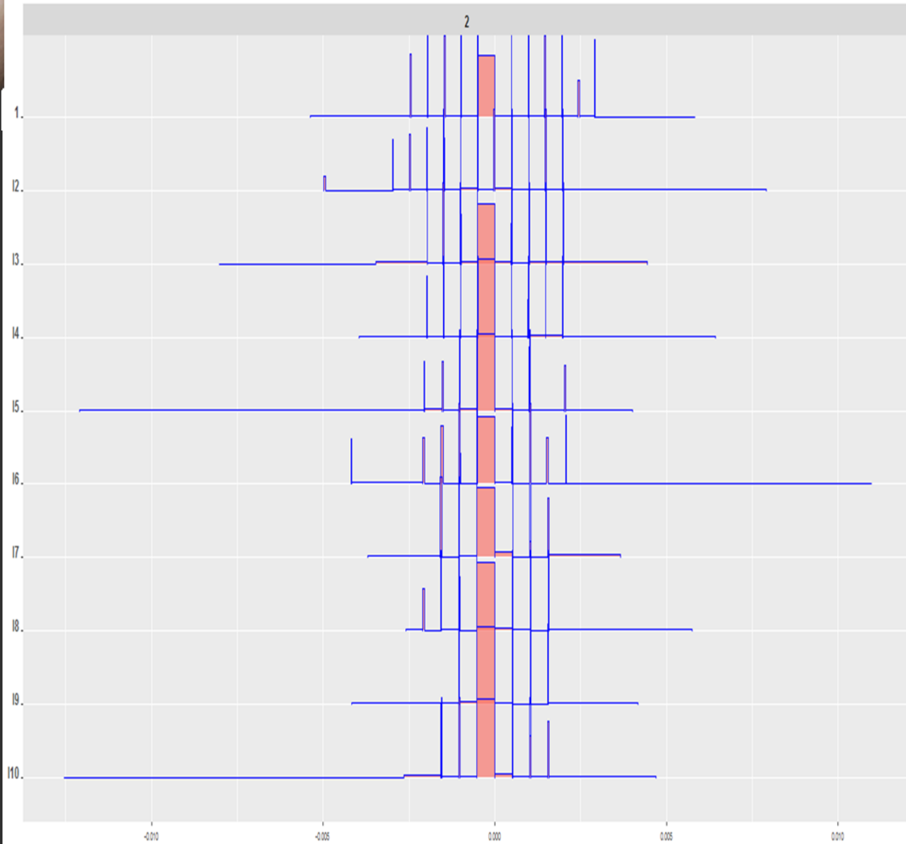
Descriptive Statistics

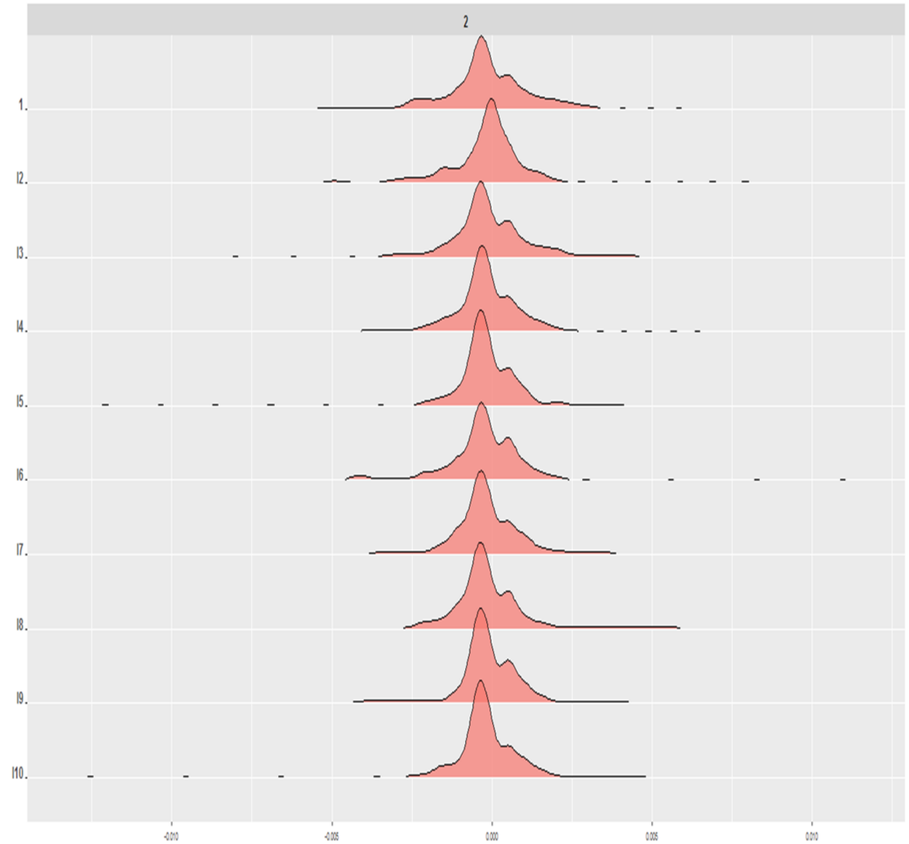
- In this study, we illustrate our model using wheat spot (wheat_s) and wheat futures (wheat_f).
- The data are high frequency 5-minute time series for the period from July 2017 to July 2018.
- Additionally, we transform these time series variables into return rate before estimation.

	wheat_s	**wheat_f**
Mean	-5.21E-07	-5.20E-07
Median	0.0000	0.0000
Maximum	0.0155	0.0782
Minimum	-0.0214	-0.0734
Std. Dev.	0.001084	0.0011
Skewness	0.069527	1.7713
Kurtosis	27.51	629.90
Jarque-Bera	1519482.	9.94E+08
MBF of Jarque-Bera	0.0000	0.0000
Unit root test	-258.65	-255.46
MBF of Unit root test	0.0000	0.0000

Histogram

Density





First 10 trading days

Results

Model Selection

In this study, we introduce nine models. It would be better to compare these models in order to find the best fit model for hedging application. Note that the GARCH (1,1) specification is used in all models for simplifying this model comparison. The comparison criteria considered here are AIC and BIC methods. According to the results, it is found that our flexible MS Dynamic Copula GARCH(1,1) shows the best fit model as the lowest AIC and BIC are obtained. Therefore, the further analysis of the hedge ratio is based on this model.

Table 2. AIC and BIC for model selection

Model	Wheat spot-futures	
	AIC	BIC
CCC-GARCH (1,1)	1528.93	1553.55
DCC-GARCH (1,1)	1514.00	1542.14
MS-CCC-GARCH (1,1)	2255.85	2312.06
MS-DCC-GARCH (1,1)	1370.74	1434.05
Copula GARCH(1,1)	1030.63	1062.32
MS-Copula GARCH(1,1)	1203.94	1207.46
Dynamic Copula GARCH(1,1)	1012.36	1022.92
MS Dynamic Copula GARCH(1,1)	1040.62	1072.32
Flexible MS Dynamic Copula GARCH(1,1)	**-29.60**	**40.74**

Results

Estimation results of Flexible MS Dynamic Copula GARCH(1,1) by Histogram data

Table 3. MS-GARCH (1,1) by Histogram data

Coefficient	Wheat-spot	Wheat-futures
$\alpha_{0,1}$ (Regime 1)	0.0054(0.0548)[0.9952]	-0.0012(0.0003)[0.0003]
α_1 (Regime 1)	0.0615(0.2752)[0.9753]	0.0408(0.0104)[0.0005]
β_1 (Regime 1)	0.9269(0.2449)[0.0008]	0.6168(0.0301)[0.0000]
$\alpha_{0,2}$ (Regime 2)	0.0964(0.7021)[0.9906]	0.0446(0.0067)[0.0000]
α_1 (Regime 2)	0.0159(0.2294)[0.9976]	0.001(0.0062)[0.9871]
β_1 (Regime 2)	0.8178(1.4744)[0.8574]	0.5424(0.0503) [0.0000]

Table 4. Time varying Copula

Coefficient	Wheat
ω_{0,s_t} (Regime 1)	0.0992(0.0480)[0.1182]
ω_{1,s_t} (Regime 1)	-0.010(2.1450)[0.9999]
ω_{2,s_t} (Regime 1)	-0.1003(0.2134)[0.8954]
ω_{0,s_t} (Regime 2)	0.1984(0.030)[0.0000]
ω_{1,s_t} (Regime 2)	-0.0903(0.345)[0.9663]
ω_{2,s_t} (Regime 2)	0.8992(0.4903)[0.1860]

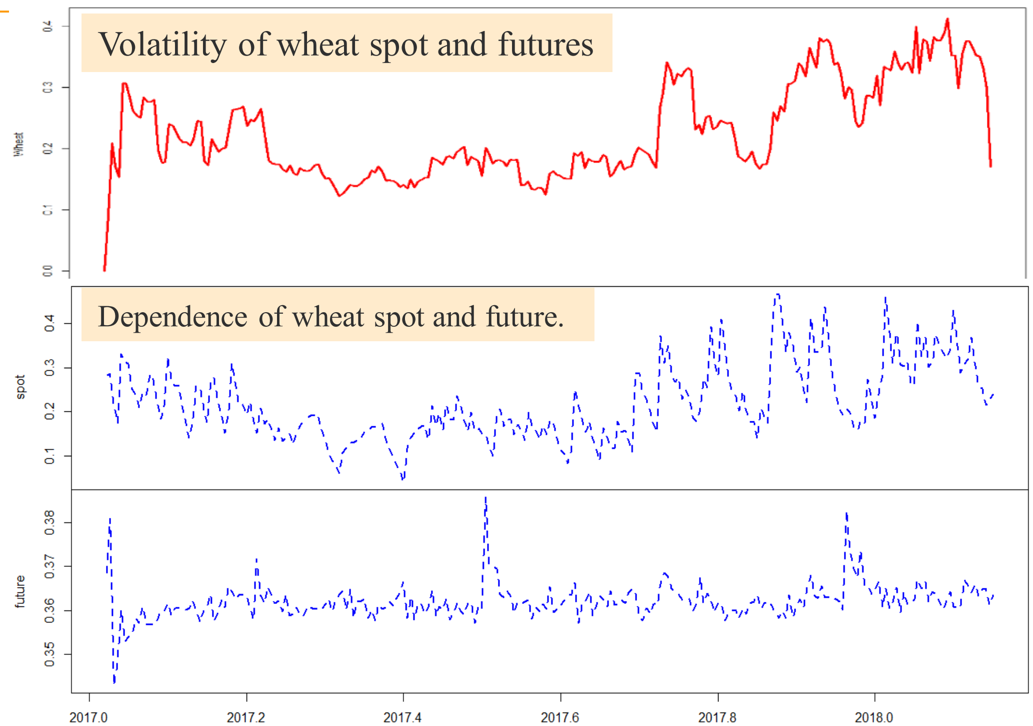
Note: () is standard error of the model. [] is Minimum Bayes factor

Results

Estimation results of Flexible MS Dynamic Copula GARCH(1,1) by Histogram data

Table 5 . Transition probability matrix

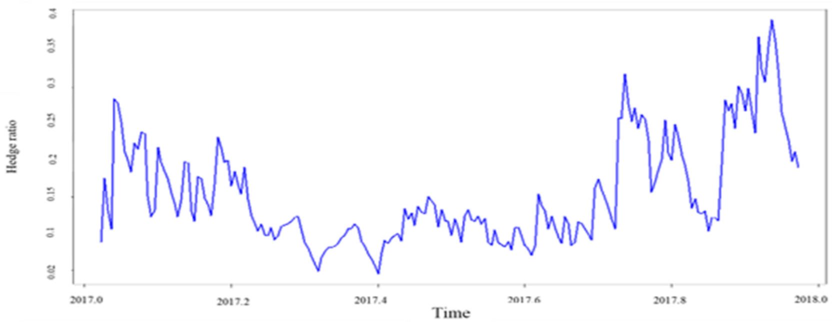
	Regime 1	Regime 2
Regime 1	0.9245	0.0755
Regime 2	0.0395	0.9605



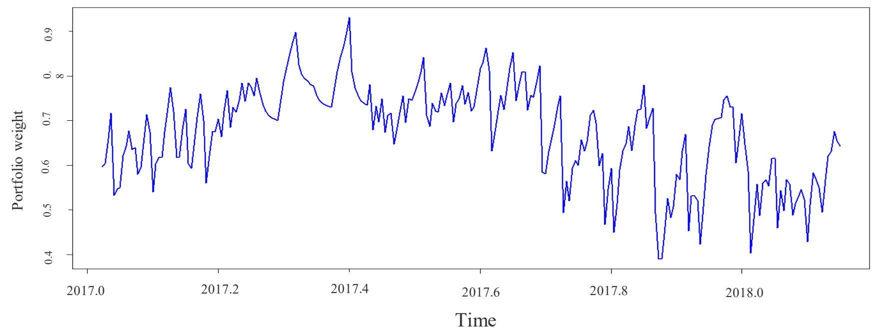
Results

Hedge ratios and portfolio weights

Hedge ratio



Portfolio weight



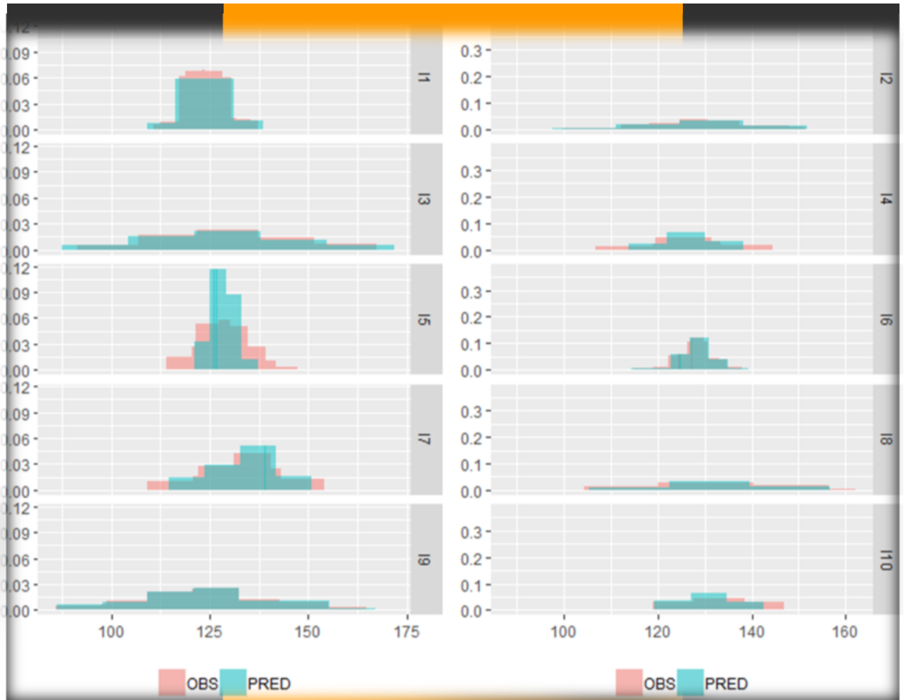
The results show the time varying hedging ratio and portfolio weight for wheat spot and futures. It can be seen that the time varying hedge ratios clearly change when new information arrives in the market. We find that the time-varying hedging ratio seem to response to the change of real economic situation. The hedging ratio is volatile over time. We observe that the highest average hedge ratio value is 0.17 meaning that, in order to minimize risk, a long (buy) position of one dollar in wheat spot should be hedged by a short (sell) position of $0.17 in wheat futures. Considering the optimal portfolio weights, we observe that it swings around 0.4-0.9. The value 0.4 would imply that investors should have more futures than spot, account for 60% of portfolio, to minimize risk without lowering expected returns. On the other hand, the value 0.9 would imply that investors should hold spot 90% in their portfolio.

Conclusion

- One of the main function of the future market is to provide a hedging mechanism. It is also a well-documented claim in the future market literature that the optimal hedge ratio should be time varying and not constant.

- This study uses the histogram value data as it contains a tone of information and reflects the real behavior of data set hence, we can gain more information of the prices.

- The empirical results shows that the flexible MS Dynamic Copula GARCH model performs the best in this data as the AIC and BIC are lower than other models.


l1
l2
l3
l4
l5
l6
l7
l8
l9
l10
0.3
0.2
0.1
0.0
100
125
150
175
100
120
140
160
OBS
PRED
OBS
PRED

Thank you