

Probing Trading Activities in Commodity Futures Market via Volatility Modeling

Yunxi Yan^{†1,a}, Shiyu Hu^{†2}

¹Dalian No.24 High School Dalian, China

²Chongqing Nankai Secondary School Chongqing, China

[†]These authors contributed equally.

Abstract. In the context of the great fluctuation of the global financial market, it is particularly important to forecast the changing futures market. Inspired by the utilization of the Heterogeneous Autoregressive model of the Realized Volatility (HAR-RV) model in price volatility forecasting, the parameters in this model could be a fair indicator of trading activities in the market. To test this hypothesis, this HAR-RV model is applied to analyze some representative commodities futures in China. The HAR-RV model is created to forecast the future volatility patterns using existed fluctuations. We used the least square method to do the regression analysis and used the volatility of a day, a week, and a month as the variables to get the degree of influence of each variable on the future volatility, namely. Thus, we can tell whether the price of a futures commodity is more affected by short-term fluctuations or long-term fluctuations by comparing different. Preliminary results show different patterns between iron ore and soybean meal futures models, pinpointing the trading frequencies in the corresponding markets. The price of iron ore futures is largely influenced by weekly volatility, while the price of soybean meal futures is affected significantly by monthly fluctuation.

1 Introduction

Many commodities, such as crude oil and iron ore, are fundamentals in infrastructure construction and industrial development, and their price volatility poses significant influences on the economy. As a traditional asset class, commodities have again surfaced in the investment pool in the recent global macroeconomic reboot as described by the Merrill Lynch Clock. The futures market is an open market place where buyers and sellers trade standardized contracts collectively. The futures price is generally considered an unbiased proxy of the spot price. Its volatility is a fair reflection of trading behavior, given the market is well functioning, and the price is fully informative. Traders with different risk preferences, degrees of information, or trading frequencies often prefer different prices, and then they create price volatility. In this paper, the correlation between price volatility and trading frequency is fully discussed. A typical futures market is composed of a wide range of trading frequencies, namely high-frequency trading, intraday speculating, trend trading, and hedge trading, and different trading frequencies contribute to price volatilities in a different period.

Volatility is one of the prevailing features of financial markets and the most prolific topic in the financial research field. Considering volatility as an unobservable variable, Engle developed a fully specified conditional mean and conditional variance model, the Autoregressive Conditional Heteroskedastic (ARCH) Model to analyze

volatility [2]. Following that pioneered work, a voluminous literature has emerged for ARCH-GARCH related volatility research. Although ARCH-GARCH related models provide a way to explore the volatility in financial markets, they are not powerful enough to replicate some empirical features of financial markets, and model (especially stochastic volatility models) estimation is not easy. Considering the stock market and futures market can generate much intraday high-frequency data, Andersen and Bollerslev proposed to construct an observable proxy for the latent volatility to make up the weakness in ARCH-GARCH related models [3, 4]. The proxy is called Realized Volatility (RV), which aims to account for the main empirical features observed in the data and simplify model estimation. Muller put forward the Heterogeneous Market Hypothesis, which inspired researchers to develop an Autoregressive (AR) related model to incorporate both heterogeneities and realized volatilities in different time horizons [5]. This new model is labeled the Heterogeneous Autoregressive model of Realized Volatility (HAR-RV) [6]. HAR-RV model is easier for estimation and proved to perform well in volatility forecasting.

The remainder of the paper is organized as follows: In Section 2, the utilized sample, data, and HAR-RV model are discussed. Section 3 displays the result and discusses the model validation and the significance of differences in the HAR-RV model. A conclusion is obtained in the final section. following the example. Some components, such as multi-leveled equations, graphics, and tables are not prescribed, although the various table text styles are

^aguanghua.ren@gecademy.cn.

provided. The former will need to create these components, incorporating the applicable criteria that follow.

2 Data

We collected open and close price information for iron ore (I), soybean meal (M), and palm olein (P) from Jan 2nd, 2018 to Dec 23rd, 2020, in every 5-minute time window. Logarithmic prices are then computed and used for volatility calculation and modeling. The original data is commercially available from WIND.

According to the method for calculating the RV, presented by Andersen and Bollerslev [3], we divided the trading day into M segments, and the closing price in the trading day is denoted as. The RV of the trading day t is denoted as, which can be expressed as

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

Where $r_{t,i}$ denotes the logarithmic rate of return for the period of the trading day t, $r_{t,i}$ is magnified 100 times for easy observation, which is expressed as

$$r_{t,i} = (\ln P_{t,i} - \ln P_{t,i-1}) \times 100 \quad (2)$$

Furthermore, the weekly RV and the monthly RV of the trading day t are denoted as and, respectively. And their expression is defined as follows:

$$RV_t^* = \frac{\sum_{i=0}^{N-1} RV_{t-i}^d}{N} \quad (3)$$

where N is the number of trading days, and usually N=5 for and N=22 for. That means

$$RV_t^w = \frac{\sum_{i=0}^4 RV_{t-i}^d}{5} \quad (4)$$

$$RV_t^w = \frac{\sum_{i=0}^{21} RV_{t-i}^d}{22} \quad (5)$$

The average RV from the day t to (t+H) is defined as

$$\overline{RV}_{t,t+H} = \frac{1}{H} \sum_{i=1}^H RV_{t+i}^d \quad (6)$$

Besides, Andersen et al. found that the logarithmic form of the HAR model performs better than the linear HAR model [4]. Therefore, we used the logarithmic form of the HAR model to forecast the price volatility. The logarithmic form of the HAR-RV model is expressed as

$$\ln \overline{RV}_{t,t+H} = \ln \beta_0 + \ln \beta_d \overline{RV}_t^d + \ln \beta_w \overline{RV}_t^w + \ln \beta_m \overline{RV}_t^m + \varepsilon_t \quad (7)$$

3 Results and Discussion

We applied the HAR-RV model to simulate the next day's price volatility for all three commodities. The outcome showed high consistency with reality.

3.1 Model Validation

We first examined the distribution of RV. Statistically, the RVs for three commodities are nearly symmetrically scattered above and below the arithmetic mean value, which is approximately zero (0). This is quite common in a well-functioned market, indicating that price movement is random in the long term. Thus, RV2 can be treated as a proxy for the standard derivation, or the volatility of the price series, as shown in Figure 1.

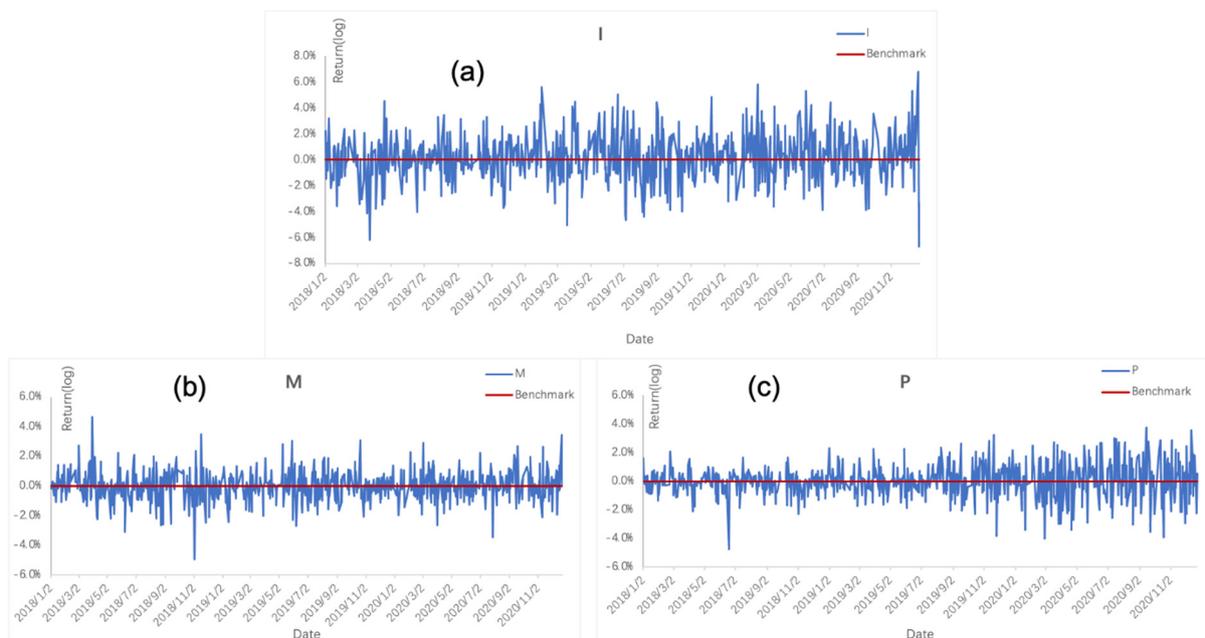


Fig. 1. RV values for iron ore future (I), soybean meal future(M), and palm olein future(P)

Secondly, we examined the validity of the HAR-RV model in these markets. The simulated RVs fit well within the tolerance of error, as shown in Figure 2. For iron ore and palm olein, over 70% of the simulation results fall in the $\pm 50\%$ range of the empirical data; for soybean meal,

over 50% of the simulation results fall in the $\pm 50\%$ range. Most of the deviations were accounted for the sudden intraday price movements or missing time windows for night-session.

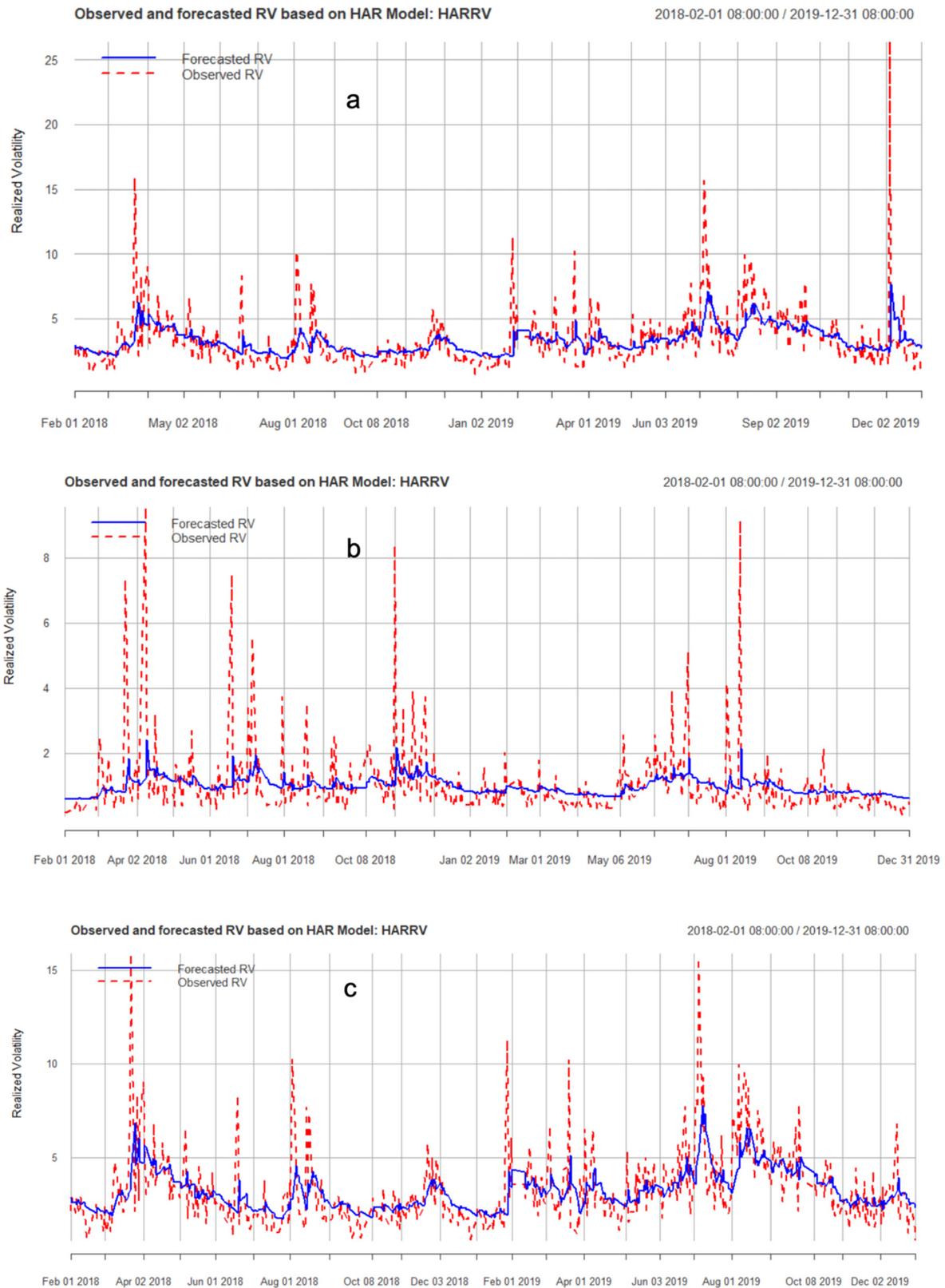


Fig.2. Simulation results for iron ore future (a), soybean meal future (b), palm olein future.

We also examined the significance of the parameters in the regression model. In the case of iron ore and palm olein, β_m is not significant, since P is larger than 0.1, and β_w has the largest value; while all the coefficients for the soybean meal are significant, with the largest value in β

m, as shown in Table 1.

Table I. HAR-RV results for three commodities

β_0	β_d	β_w	β_m	R-squared
0.90	0.17	0.45	0.13	0.2013
P<0.05	P<0.05	P<0.05	P>0.1	
0.40	0.12	0.22	0.24	0.08175
P<0.05	P<0.05	P<0.05	P<0.05	
0.69	0.21	0.48	0.11	0.263
P<0.05	P<0.05	P<0.05	P>0.1	

3.2 Data Analysis

The overall pattern of the coefficients in the HAR-RV model contains information about trading activities in these markets. Taking iron ore as an example, its β_m value is not significant, indicating the monthly price volatility does not contribute significantly to the next day's price volatility. The β_w value is the largest among all three, indicating the next day price volatility is most sensitive to the weekly price volatility change. Thus, the price movement is mostly accounted for weekly trading activities. The coefficients for palm olein showed a similar conclusion.

Moreover, the coefficients for soybean meal suggest a different trading presence. The β_m value is the largest, indicating the next day price volatility is most sensitive to the monthly price volatility change. Thus, the price movement is mostly accounted for monthly trading activities.

The difference arises from the underlying market. In the past two years, the price fluctuation ranges for iron ore, soybean meal, and palm olein are 92.96%, 20.01%, and 29.95%; the annualized standard deviations are 30.84%, 16.91%, and 18.96%, respectively. The futures markets of iron ore and palm olein are more volatile and drew more attention from short-term speculators.

Therefore, for the speculators, the iron ore and palm olein futures are more suitable choices for short term investments, given that their volatility patterns are more significantly influenced by short term fluctuation. Furthermore, soybean meal futures are more likely used as long-term investments. Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.

4 Conclusion

The HAR-RV model allowed us to analyze the relation between volatilities of different time intervals and the overall price volatility. For a more volatile futures market, the short-term volatility contributes most to the overall price fluctuation. In this paper, we examine how differences in the HAR-RV model influence different kinds of futures and how iron ore and soybean meal futures are influenced by short-term and long-term factors.

However, since we only use the HAR-RV model to

examine the data, we do not know how the future samples would perform under other models. Thus, further experiments would be conducted to verify the conclusion.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity "Magnetization", or "Magnetization, M", not just "M". If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write "Magnetization (A/m)" or "Magnetization {A[m(1)]}", not just "A/m". Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)", not "Temperature/K".

References

1. Li W, Cheng Y, Fang Q (2020) Forecast on silver futures linked with structural breaks and day-of-the-week effect. *The North American Journal of Economics and Finance* 53:101192.
2. ENGLE, R.F., "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica* 50 (1982), 987-1008.
3. Andersen, T. G. and Bollerslev, T.: 1998, Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review* 39, 885–905.
4. Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P.: 2001, The distribution of exchange rate volatility, *Journal of the American Statistical Association* 96, 42–55.
5. Muller, U., Dacorogna, M., Dav, R., Pictet, O., Olsen, R. and Ward, J.: 1993, Fractals and intrinsic time - a challenge to econometricians. XXXIXth International AEA Conference on Real Time Econometrics, 14-15 Oct 1993, Luxembourg.
6. Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2), 174–196