Journal of Mathematical Finance, 2023, 13, 144-158
https://www.scirp.org/journal/jmf
ISSN Online: 2162-2442
ISSN Print: 2162-2434

The Impact of Investor Attention on China's Corn Futures Price

Lu Zhang, Yinpeng Zhang, Li Sun*, Junwei Cheng

Department of Applied Mathematics, School of Information Science and Technology, Shandong Agricultural University, Tai'an, China
Email: *sunlishi@hotmail.com

How to cite this paper: Zhang, L., Zhang, Y.P., Sun, L. and Cheng, J.W. (2023) The Impact of Investor Attention on China's Corn Futures Price. *Journal of Mathematical Finance*, **13**, 144-158.
https://doi.org/10.4236/jmf.2023.132009

Received: March 1, 2023
Accepted: May 20, 2023
Published: May 23, 2023

Abstract

With the spread of market sentiment of global economic recession, the market favors hedging varieties such as gold and agricultural products. This paper seeks to investigate the explanations of corn futures price movements from the perspective of investor attention. We employ a time series autoregressive model (VAR) to analyze the relationship between them. Empirical results indicate that investor attention is a Granger cause of changes in the corn futures price, generating both linear and nonlinear effects on corn futures price. Furthermore, out-of-sample analysis shows that introducing the Baidu index improves the model's prediction accuracy.

Keywords

Corn Futures Price, Baidu Index, VAR Model, Granger Causality, Investor Attention

1. Introduction

Agricultural futures trading has a history of 170 years and the earliest agricultural futures exchange is the Chicago futures exchange. China's agricultural product futures market was established in the 1990s. After many years of development, it initially covered the four major agricultural product futures that are grain, cotton, oil, and sugar. The value of the agricultural products futures market is reflected in three aspects:

1) It is a good guide for farmers to choose what kind of crops to grow and at what price to sell their crops.

2) The futures market effectively allows relevant enterprises to maintain and avoid risks. Most domestic corn and soybean business enterprises participate in hedging transactions and avoid the business risks caused by price fluctuations

through futures.

3) The price information of the futures market reflects the market expectation and is the scientific reference signal of national macro-control.

In agricultural products, corn has complex attributes. It has grain ration attributes, industrial crop attributes, and forge crop attributes. The relevant proportion of corn is shown in Figure 1 and Figure 2.

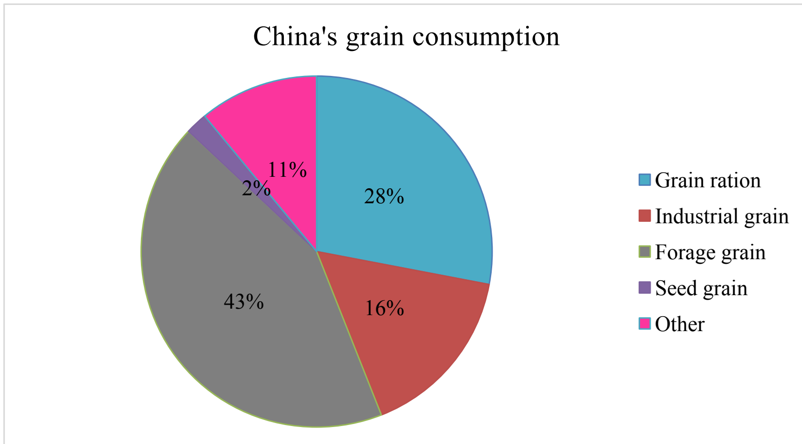
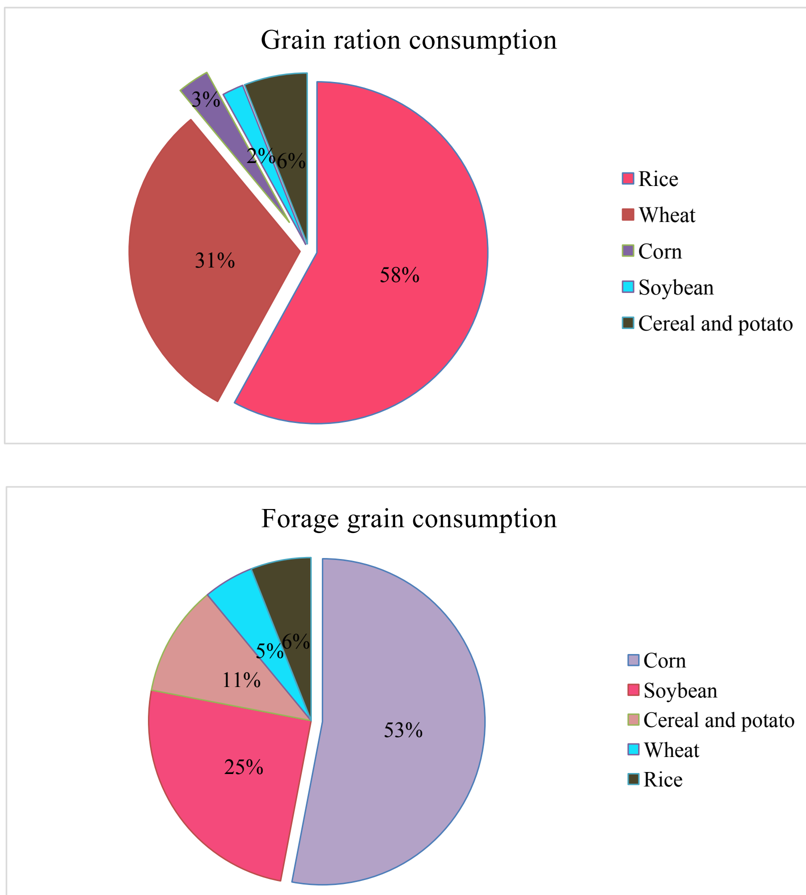


Figure 1. The proportion of China's grain consumption in 2019.



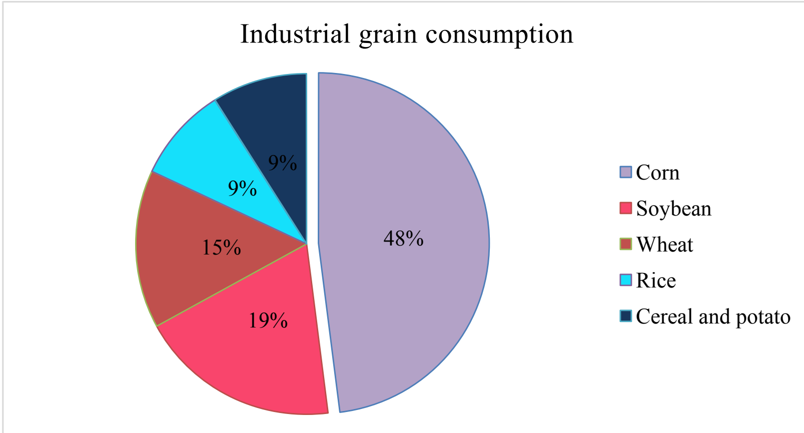


Figure 2. The proportion of grain varieties of China' grain ration, forage grain, and industrial grain in 2019.

Various prediction methods have been suggested, which can be classified into statistical and machine learning methods. Traditional statistical methods provide great prediction results under the linear assumption. For example, linear regression model [1], autoregressive integrated moving average model (ARIMA) [2], autoregressive conditional heteroskedasticity (ARCH) [3], generalized autoregressive conditional heteroskedasticity (GARCH) [4], and vector autoregressive regression (VAR) [5]. Because the financial time series is complicated, dynamic, and non-linear, various approaches which belong to the machine learning field have been employed to build the prediction model. Support vector machines (SVMs) [6] and deep learning type methods [7] [8] [9] [10] are widely used.

The influencing factors related to the price of corn futures include corn yield, climate factors, production cost, inventory, import and export volume, the price of soybean and wheat as substitutes, the development of animal husbandry, interest rate, and exchange rate, relevant national policies, crude oil, and so forth. It is challenging to establish an effective relationship between the corn price and quantifying the relevant influencing factors.

With the development of the Internet, Internet search data has vital timeliness and comprehensive coverage, and information comes from real economic individuals [11]-[15]. After realizing the advantages of Internet search volume, scholars have taken the Internet search index as a proxy variable for investors' attention. In this study, we seek to investigate the explanations of exchange rate movements from the perspective of investor attention, which is widely used in asset pricing and market efficiency. It is a relatively new perspective to study the price of China's agricultural futures market from the perspective of behavioral finance. It mainly analyzes the impact of investors' attention on the agricultural futures market and studies how the search volume of the Baidu index affects the agricultural futures market, which has specific reference significance for predicting agricultural futures market price.

The second section of this article is the data description, and an in-sample analysis is shown in Section 3. Based on the VAR model, we found that the Baidu index is the Granger reason for corn futures price, and a VAR model including the Baidu index and corn futures price is established. Parameter significance analysis shows that investors' attention significantly impacts corn futures prices. We also lead in other control variables that impact the price of corn futures in relevant studies, such as soybeans, pork, corn starch, and crude oil. The results show that investors' attention can still be an influential explanatory variable. The final out-of-sample prediction in Section 4 shows that the prediction accuracy of the VAR model is effectively improved after introducing the Baidu index. In the fourth section, we use the Baidu index of "corn starch" and "corn price", two keywords other than "corn", and carry out the Granger causality test. The numerical results show that the Baidu index obtained using similar keywords is still the Granger reason for the corn futures price.

2. Data

In this section, we will briefly describe the selection of data sources and variables. Firstly, the data used in this study are all from the WIND (https://www.wind.com.cn/) database, and we select weekly corn futures price data from January 2, 2017, to October 24, 2021. In order to eliminate the instability of time series, we use logarithm and difference data processing methods. Secondly, since Baidu is the largest search engine in China, which has extensive and professional characteristics in information collection, we choose the Baidu index (https://index.baidu.com/v2/index.html#/) as the proxy variable of investors' attention. The Baidu index directly reflects the size of specific keywords in Baidu web search. We selected several keywords closely related to the corn market, including "corn", "corn price", and so on. According to the Granger causality test results, we choose the keyword "corn" with the highest correlation with corn futures to conduct the follow-up research.

Finally, we select several closely related factors to carry out the following research, including pork price, soybean futures price, wheat futures price, international crude oil futures price, and cornstarch futures price. The reasons for selecting these factors are as follows. First, the impulse response model to study the factors affecting corn price fluctuations shows that pork price is the main factor affecting corn price. Moreover, corn accounts for more than 50% of hogs' feed costs. Secondly, wheat and soybean are the competitive commodities of corn, so the price change is bound to have a certain influence on corn. Third, corn starch is a by-product of corn, and the two are closely related. Fourthly, corn is the raw material of energy ethanol, and ethanol has a substitution relationship with crude oil, so international crude oil futures will also impact the return of corn futures.

Some basic descriptive statistics of corn futures price, investor attention, and the control variables are shown in **Table 1**. According to the data in **Table 1**, we can find that the time series of corn futures returns, investor attention on "corn", and relevant market variables are volatile.

Table 1. Descriptive statistic.

	Mean	Std. dev	Max	Min	Skewness	Kurtosis
Corn futures	0.0021	0.0131	0.0481	−0.0359	0.0674	3.6234
Investor attention	0.0005	0.1158	0.6047	−0.3809	0.6960	6.9043
Pork price	−0.0006	0.0279	0.1092	−0.0982	0.6959	6.2105
Soybean futures	0.0014	0.0235	0.1366	−0.1149	0.3318	9.5193
Wheat futures	0.0002	0.0074	0.0308	−0.0424	−1.6891	14.1867
Crude oil WTI	0.0017	0.1461	1.3914	−1.6349	−2.1797	96.4457
Corn starch futures	0.0023	0.0228	0.1162	−0.1224	0.1058	11.8756

The VAR model is one of the models dealing with the analysis and prediction of multiple related economic indicators. The VAR model is good at predicting the interrelated time series system and analyzing the dynamic impact of random disturbances on variable systems so as to explain the impact of various economic shocks on the formation of economic variables. In this paper, we will use the VAR model to analyze the corn futures price and the time series involved in Table 1.

VAR theory requires that each variable in the model is stationary. The ADF unit root test was applied to the time series of corn futures price and the Baidu index after transferring the data to log returns and proceeding with the difference transform. ADF test results shown in Table 2 indicated that VAR modeling analysis can be carried out.

3. In-Sample Analyses

In this section, we conduct an in-sample analysis. First, divide the data into two parts at the time of August 29, 2021, for in-sample and out-sample analysis. We use the Granger causality test and impulse response analysis to study the impact relations between the Baidu index and corn futures price. Finally, the VAR models based on multiple interrelated time series are established.

3.1. Granger Causality Test

The Granger causality test can be used to judge whether the change of one variable is the cause of the change of another variable. We conduct a basic Granger causality test to investigate the linear causality relationships between the corn futures price and investor attention. According to the AIC criterion, we set the lag length to be 2 in the Granger causality test. The results are shown in Table 3.

As shown in Table 3, the probability that investor attention is not the Granger cause of corn futures price is 0.0803, which obviously rejects the null hypothesis and admits that the change in investor attention will affect the change in corn futures price. At the same time, the probability that corn futures price is not the Granger cause of investor attention is 0.2334, which is to accept the null hypothesis, that is, corn futures price will not reverse the change of investor attention.

Table 2. ADF test results.

	Type	t-statistic
Corn futures price	None	−12.4031***
	Intercept	−12.6150***
	Trend and intercept	−12.5908***
Investor attention	None	−13.5861***
	Intercept	−13.5582***
	Trend and Intercept	−13.5307***

Note: None, Intercept, Trend and Intercept mean three types of the ADF test, respectively. The null hypothesis of the ADF test assumes that the time series has a unit root. *** represents the significance in 1% level.

Table 3. Granger causality test results.

Null hypothesis	Prob.
Investor attention does not Granger cause corn futures price	0.0803*
Corn futures price does not Granger cause investor attention	0.2334

Note: * represents the 10% level of significance.

3.2. The VAR Model

VAR is used to estimate the dynamic relationship of joint endogenous variables. It takes each endogenous variable of the system as the lag value function of all endogenous variables in the system to construct the model. In this subsection, we use the VAR model to analyze the relationships between the corn futures price and investor attention. The VAR model is shown in Equation (1).

$$Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \varepsilon_t . \tag{1}$$

In Equation (1), Y_t is the n-dimensional endogenous variable vector, which contains the corn futures price CF_t and the corresponding investor attention ATT_t, p is the lag length in the model, α_i is the coefficient for the lagged term, $i = 1, 2, \cdots, p$, ε_t is a 2-dimensional perturbed vector. Set $p = 2$, we have the results in Table 4.

Two columns in Table 4 present the coefficients of corn futures price and investor attention in Equation (1). Due to the significant performance of the first-order lag coefficient, that is to say, the change in the Baidu index will be immediately reflected in the corn futures price change.

The results in Table 5 indicate that there is no multicollinearity between variables. Also, the residual variance test proves that the VAR model does not have heteroscedasticity. These denote that the VAR model is feasible and can be used for further analysis.

Figure 3 and Figure 4 show the response function of corn futures price caused by the impact of investor attention, and vice versa. As can be seen from the impulse response diagram, the change of attention will definitely have an impact on corn futures price, and the impact is negative. The shock gradually decreases after reaching the maximum value in the second week and begins to stabilize from the sixth week. In addition, the impact of corn futures prices on attention can last for about five weeks.

3.3. Controlling Related Price

In this section, we will consider the influencing factors related to the corn futures price: the pork price, soybean futures price, wheat futures price, West Texas Intermediate crude oil futures price (WTI), and corn starch futures price. In this way, we can more comprehensively analyze the impact of investor attention on corn futures prices. The regression model used is shown in Equation (2).

Table 4. VAR analysis results.

	CF_t	ATT_t
CF_{t-1}	0.195802*** (0.064552)	−0.84236 (0.57068)
CF_{t-2}	−0.05342 (0.064387)	0.662358 (0.569219)
ATT_{t-1}	−0.01328* (0.007174)	0.159228** (0.063424)
ATT_{t-2}	−0.00767 (0.007215)	−0.11468* (0.06378)
Intercept	0.001815 (0.000847)	0.002034 (0.00749)
R^2	0.0593	0.0415

Note: *, **, *** denote significance at 10%, 5% and 1% level, respectively. ( ) represents the standard error.

Table 5. Tests for VAR model.

	CF_{t-1}	CF_{t-2}	ATT_{t-1}	ATT_{t-2}
Panel A: 1/VIF				
	0.9420	0.9539	0.9507	0.9533
Panel B: Correlation matrix				
CF_{t-1}	1.0000			
CF_{t-2}	0.1847	1.0000		
ATT_{t-1}	0.0047	−0.0697	1.0000	
ATT_{t-2}	−0.1175	0.0085	0.1351	1.0000
Panel C: VAR residual heteroscedasticity test				
No cross terms		0.8125 (0.5418)	With cross terms	0.6086 (0.9044)

Note: This table reports the results of VIF test, correlation analysis, and heteroscedasticity test. ( ) represents the standard error.

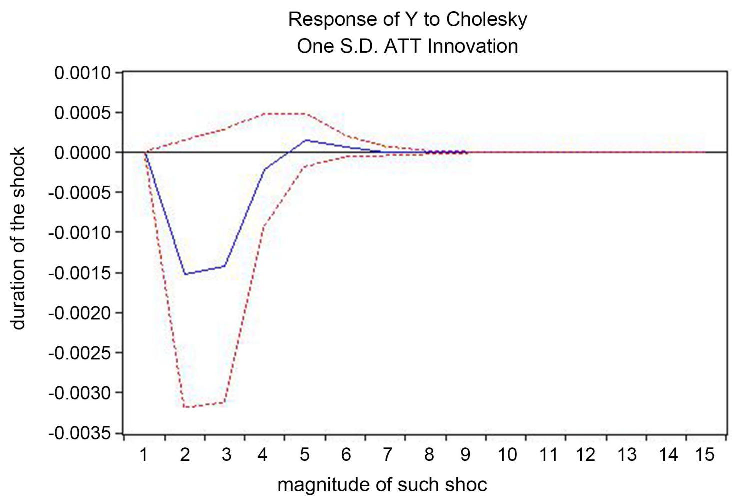


Figure 3. Response of corn futures price to the shock from investor attention.

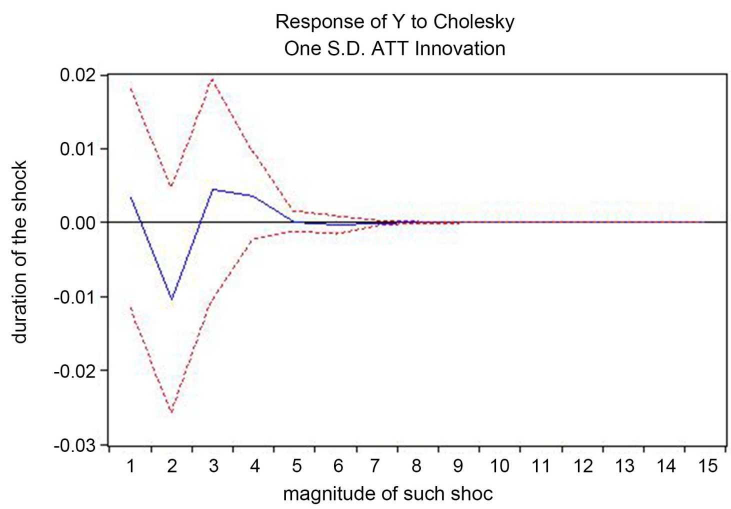


Figure 4. Response of investor attention to the shock from corn futures price.

$$CF_t = \sum_{i=1}^{2} \alpha_i CF_{t-i} + \sum_{i=1}^{2} \beta_i ATT_{t-i} + \sum_{i=1}^{2} \gamma_i C_{t-i} + \sum_{i=1}^{2} \delta_i ATT_{t-i} \cdot C_{t-i} + e_t , \tag{2}$$

where CF represents the five controlling factors previously introduced. The interaction term in Equation (2) stands for the multiplication of the investors' attention and the relevant controlling variables. The coefficient of the regression Model (2) is shown in **Table 6**. In this way, we analyse the importance of investors' attention by testing the significance of the parameters in the VAR model.

Results in **Table 7** imply that the interactive terms between attention and energy returns generate significantly negative effect on corn futures price at the first lag, which imply that the growth of investors' attention leads to a decline in returns. This result is consistent with Ref. [15]. And other interactive terms, that is, soybean, wheat, WIT and cornstarch, do not appear significant effects. If we fix the related control variable, investor attention shows significant negative impacts on corn futures price.

Table 6. The joint impact results.

	Pork	Soybean	Wheat	WTI	Cornstrach
	CF_t	CF_t	CF_t	CF_t	CF_t
CF_{t-1}	0.1993*** (0.0645)	0.1910*** (0.0645)	0.1975*** (0.0649)	0.1907*** (0.0646)	0.1926*** (0.0717)
CF_{t-2}	−0.0555 (0.0640)	−0.0527 (0.0643)	−0.0382 (0.0654)	−0.0459 (0.0647)	0.0097 (0.0719)
ATT_{t-1}	−0.0136* (0.0072)	−0.0153** (0.0074)	−0.0137* (0.0072)	−0.0133* (0.0071)	−0.0133* (0.0072)
ATT_{t-2}	−0.0077 (0.0072)	−0.0053 (0.0074)	−0.0079 (0.0072)	−0.0084 (0.0072)	−0.0084 (0.0071)
C_{t-1}	−0.0253 (0.0456)	−0.0157 (0.0376)	−0.1259 (0.1319)	−0.0122 (0.0091)	−0.0012 (0.0427)
C_{t-2}	0.0323 (0.0445)	−0.0415 (0.0377)	−0.0318 (0.1363)	0.0124 (0.0125)	−0.0786* (0.0419)
$ATT_{t-1} \cdot C_{t-1}$	0.5277* (0.2743)	0.6529 (0.4687)	0.2588 (1.7099)	−0.1506 (0.1608)	0.2625 (0.3202)
$ATT_{t-2} \cdot C_{t-2}$	−0.1552 (0.2731)	−0.3540 (0.4684)	0.4019 (1.6725)	0.1318 (0.1065)	−0.1073 (0.3304)
Intercept	0.0019 (0.0008)	0.0018 (0.0009)	0.0018 (0.0008)	0.0018 (0.0008)	0.0018 (0.0008)
R^2	0.0748	0.0778	0.0653	0.0716	0.078

Note: *, **, *** denote significance at 10%, 5% and 1% level, respectively. ( ) represents the standard error.

Table 7. Out-of-sample prediction results with forecast horizon equal to one.

	R^2_{OOS}	MSFE	MSFE_{adj}
VAR	0.0177	$2.8299e^{-4}$	1.4516*
Pork	0.0161	$2.8344e^{-4}$	1.6106*
WTI	−0.0037	$2.8915e^{-4}$	1.4929
Wheat	−0.0159	$2.9265e^{-4}$	0.9994
Cornstrach	−0.0193	$2.9363e^{-4}$	0.9064
Soybean	0.0177	$2.8297e^{-4}$	1.4257*

Note: VAR represents Model (3). Pork, WTI, Wheat, Starch and Soybean refer to the cases when the control variable in Equation (4) is controlled by the corresponding price, respectively.

In the in-sample analysis section, we use VAR theory to study the impact of investors' attention on corn futures price. Test results indicate that investor's attention is a Granger cause of carbon futures return and exhibit linear and nonlinear effects on corn futures price. In the next section, out-of-sample forecast section, we will use the data reserved for out-of-sample prediction, and test the generalization capability of the model.

4. Out-of-Sample Forecasts

We conduct an out-of-sample analysis on the following two models, which belong to the rolling window prediction method [13]. According to this prediction method, we can obtain the price of corn futures in the $t+h$ week, as shown in Equations (3) and (4).

$$\widehat{CF}_{t+h} = \hat{\alpha}_0 + \sum_{i=1}^{2} \hat{\alpha}_i CF_{t+h-i} + \sum_{i=1}^{2} \hat{b}_i ATT_{t+h-i}, \tag{3}$$

$$\widehat{CF}_{t+h} = \hat{\alpha}_1 + \sum_{i=1}^{2} \hat{\alpha}_i CF_{t+h-i} + \sum_{i=1}^{2} \hat{b}_i ATT_{t+h-i} + \sum_{i=1}^{2} \hat{c}_i C_{t+h-i} + \sum_{i=1}^{2} \hat{d}_i ATT_{t+h-i} \cdot C_{t+h-i}. \tag{4}$$

The coefficients of Equations (3) and (4) are obtained from the in-sample analysis and update in each prediction. The forecast horizon h represents the certain week, the investor would like to estimate. After h is given, the model will give the price after h weeks, using data from 0 to t.

We analyze and assess the accuracy of different predictive models by calculating out-of-sample R squared (R^2_{oos}), mean squared forecast error (MSFE), MSFE-adjusted statistic. The R^2_{oos} indicates the proportion of reduction in MSFE for using the predictive models compared with the benchmark model. A positive R^2_{oos} indicates that the forecasting performance of the predictive models outperforms the benchmark model. The R^2_{oos} is obtained from the following equation:

$$R^2_{oos} = 1 - \frac{\sum_{k=t+h}^{T} \left(R_k - \hat{R}_k\right)^2}{\sum_{k=t+h}^{T} \left(R_k - \bar{R}_k\right)^2}, \tag{5}$$

where T is the numbers of out-of-sample. R_k is the real value of carbon futures return. $\hat{R}_k$ contains the return forecasted by the above-mentioned Equations (3) and (4). $\bar{R}_k$ denotes the predicted return of the benchmark forecast model. The historical average model is the bench mark model in our out-sample analysis, which is also adopted by other researches [13]. The model is given by Equation (6).

$$\bar{R}_{k+1} = \frac{1}{k} \sum_{s=1}^{k} R_s, \tag{6}$$

where $\bar{R}_{k+1}$ represents the forecast value of the benchmark model, and R_s means the real value. As for the MSFE, it can obtain from Equation (7).

$$\text{MSFE} = \frac{\sum_{k=t+h}^{T} \left(R_k - \hat{R}_k\right)^2}{T - h + 1}. \tag{7}$$

The MSFE-adjusted statistic is a one-sided (upper-tail) measure to test the null hypothesis that the MSFE of the historical average model is less than or equal to that of the predictive models incorporated with investor attention. The alternative hypothesis is that the MSFE of the historical average model is more than that of the predictive models. Specifically, the MSFE-adjusted statistic is measured by the following equation:

$$\text{MSFE}_{\text{adj}} = \text{MSFE}_a - \text{MSFE}_b + \frac{\sum_{k=t+h}^{T}\left(\hat{R}_{k,a} - \hat{R}_{k,b}\right)^2}{T-h+1}, \tag{8}$$

where MSFE_a and MSFE_b denote the MSFE statistics of the predictive models incorporated investor attention and the historical average model, respectively. $\hat{R}_{k,a}$ and $\hat{R}_{k,b}$ represent the forecast value of carbon futures return of the predictive models and the historical average model, respectively.

4.1. One Period Ahead Forecast

In out-of-sample prediction, we set the length of the rolling window as 170 and $h=1$.

The model established with pork and soybean prices as control variables (2), its value is greater than 0. These indicate that these two predictive regression model always outperforms the historical average for any out-of-sample period.

4.2. Longer Horizon Forecasts

We set the prediction period as 2, 4, 8, and 12 to explore the forecast performance of investors' attention on corn futures price. These settings correspond to the prediction performance after half a month, one month, two months, and three months. The results are shown in **Table 8**.

Table 8. Out-of-sample prediction results with different forecast horizons.

	Panel A: $h = 2$			Panel B: $h = 4$		
	R^2_{OOS}	MSFE	MSFE_{adj}	R^2_{OOS}	MSFE	MSFE_{adj}
VAR	0.0204	2.8601 e^{-4}	1.4810*	0.0321	2.8871e^{-4}	1.6246*
Pork	0.0214	2.8572 e^{-4}	1.6880**	0.0365	2.8740e^{-4}	1.9045**
Crude oil WTI	0.0163	2.8720 e^{-4}	2.0112**	0.0371	2.8722e^{-4}	1.7549**
Wheat	−0.0079	2.9428 e^{-4}	1.0726	0.0037	2.9717e^{-4}	1.1980
Cornstarch	−0.0175	2.9707 e^{-4}	0.8626	−0.0027	2.9908e^{-4}	1.0853
Soybean	0.0150	2.8759e^{-4}	1.4252*	0.0308	2.8910e^{-4}	1.6568**
	Panel C: $h = 8$			Panel D: $h = 12$		
	R^2_{OOS}	MSFE	MSFE_{adj}	R^2_{OOS}	MSFE	MSFE_{adj}
VAR	0.0327	2.9996e^{-4}	1.5985*	0.0375	3.1032e^{-4}	1.6149*
Pork	0.0394	2.9787e^{-4}	1.8909**	0.0483	3.0684e^{-4}	2.0246**
Crude oil WTI	0.0274	3.0160e^{-4}	1.5312*	0.0334	3.1165e^{-4}	1.6055*
Wheat	0.0083	3.0752e^{-4}	1.2507	0.0045	3.2097e^{-4}	1.2569
Cornstarch	0.0011	3.0974e^{-4}	1.0891	0.0022	3.2171e^{-4}	1.0832
Soybean	0.0502	2.9452e^{-4}	1.7872**	0.0370	3.1050e^{-4}	1.5951*

Note: *, **, *** denote significance at 10%, 5% and 1% level, respectively. VAR represents Model (3). Pork, WTI, Wheat, Starch and Soybean refer to the cases when the control variable in Equation (4) is controlled by the corresponding price, respectively.

It can be concluded from **Table 8** that the prediction model with relevant market variables is better than the benchmark model regardless of the short-term or more extended period, and is positive in the models created with pork price, soybean futures price, and crude oil futures price, which indicate the performance of these three prediction models is higher than the benchmark model. It should be noted that the prediction powers may be weaker than the benchmark model in the out-of-sample forecast, even if they display an excellent explanatory power on corn futures prices.

5. Robustness Checks

This study believes that investor attention is an important variable that can explain and predict the returns of corn futures. However, all results are calculated based on the Baidu index keyword "corn". To ensure the study's accuracy, we conducted robustness tests to explore whether the experimental results would produce different conclusions due to the change of keywords.

We search other keywords, *i.e.* "corn price" and "corn starch" during the same period from January 02, 2017 to August 29, 2021. Then, we adopt the same data pre-processing method, conduct the Granger causality test for different keywords and establish the VAR model in order to show that corn futures returns are still significantly affected by investor attention. The results of correlation regression are shown in **Table 9** and **Table 10**. The results of the Granger causality test show that the change of keywords does not change the explanatory power of attention. As can be seen from the VAR results, investor attention can indeed

Table 9. Granger causality and VAR analyses for robustness checks (keyword = corn price).

Panel A: Granger causality test result for robustness check		
Investor attention is not the Granger cause of corn futures price		2.0603*
Corn futures price is not the Granger cause of investor attention		1.8878
Panel B: VAR analysis		
	Y_t	ATT_t
Y_{t-1}	0.19358***	0.97161
Y_{t-2}	−0.03092	1.65493*
Y_{t-3}	−0.00244	−1.35596
Y_{t-4}	0.02660	0.08266
ATT_{t-1}	0.00035	0.20158***
ATT_{t-2}	−0.01220**	−0.16870***
ATT_{t-3}	−0.00337	−0.10712*
ATT_{t-4}	0.00109	−0.05431
intercept	0.00144*	−0.00967
R^2	0.0743	0.1352

Note: *, **, *** denote significance at 10%, 5% and 1% level, respectively.

Table 10. Granger causality and VAR analyses for robustness checks (keyword = cornstarch).

Panel A: Granger causality test result for robustness check		
Investor attention is not the Granger cause of corn futures price		2.7480*
Corn futures price is not the Granger cause of investor attention		3.8201**
Panel B: VAR analysis		
	Y_t	ATT_t
Y_{t-1}	0.20057***	0.91855
Y_{t-2}	−0.02412	1.56713*
ATT_{t-1}	−0.00036	0.25309***
ATT_{t-2}	−0.01233***	−0.18423***
intercept	0.00166**	−0.01228
R^2	0.0709	0.1043

Note: *, **, *** denote significance at 10%, 5% and 1% level, respectively.

have a significant impact on corn futures returns. Here, we find that the lag is not the same as before, which is not surprising, but probably because different keywords give investors different feelings. The above results show that investor attention is a pricing factor that can not be ignored in the corn futures market, which deserves more attention.

6. Conclusions

In this paper, we combine the studies of agricultural economics and behavioral finance and provide significant evidence to show investors' attention impacts the price of corn futures. The Granger causality test and the significance test of coefficients in the VAR method indicate that the Baidu index, treated as the proxy variable of investors' attention, can be used to analyze the corn futures price. It is a relatively new perspective to study the price of China's agricultural futures market from the perspective of behavioral finance. This work has specific reference significance for the prediction of agricultural futures market prices.

Our future research remains into two parts: one is to use the feature extraction or feature selection method of machine learning to analyze whether investors' attention has an impact on the price of agricultural products; the other is to use the current popular deep learning method to analyze the prediction accuracy of agricultural product prices before and after the consideration of investors' attention.

Funding

This work was supported by the National Natural Science Foundation of China (52275262 and 11701337), and the National Natural Science Foundation of Shandong Province (ZR2022MA009).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Ge, Y. and Wu, H.X. (2020) Prediction of Corn Price Fluctuation Based on Multiple Linear Regression Analysis Model under Big Data. *Neural Computing & Applications*, **32**, 16843-16855. https://doi.org/10.1007/s00521-018-03970-4

[2] Dou, Z.W., Ji, M.-X., Wang, M. and Shao, Y.-N. (2022) Price Prediction of Pu'er Tea Based on ARIMA and BP Models. *Neural Computing & Applications*, **34**, 3495-3511. https://doi.org/10.1007/s00521-021-05827-9

[3] Osah, T.T. and Mollick, A.V. (2023) Stock and Oil Price Returns in International Markets: Identifying Short and Long-Run Effects. *Journal of Economics and Finance*, **47**, 116-141. https://doi.org/10.1007/s12197-022-09602-x

[4] Kristjanpoller, W. and Minutolo, M.C. (2018) A Hybrid Volatility Forecasting Framework Integrating GARCH, Artificial Neural Network, Technical Analysis and Principal Components Analysis. *Expert Systems with Application*, **109**, 1-11. https://doi.org/10.1016/j.eswa.2018.05.011

[5] Han, S., Zhang, B.-S., Tang, X. and Guo, K.-Q. (2017) The Relationship between International Crude Oil Prices and China's Refined Oil Prices Based on a Structural VAR Model. *Petroleum Science*, **14**, 228-235. https://doi.org/10.1007/s12182-016-0139-9

[6] Ouahilal, M., Mohajir, M.E., Chahhou, M. and Mohajir, B.E.E. (2017) A Novel Hybrid Model Based on Hodrick-Prescott Filter and Support Vector Regression Algorithm for Optimizing Stock Market Price Prediction. *Journal of Big Data*, **4**, Article No. 31. https://doi.org/10.1186/s40537-017-0092-5

[7] Banerjee, T., Sinha, S. and Choudhury, P. (2022) Long Term and Short Term Forecasting of Horticultural Produce Based on the LSTM Network Model. *Applied Intelligence*, **52**, 9117-9147. https://doi.org/10.1007/s10489-021-02845-x

[8] Diqi, M., Hiswati, M.E. and Nur, A.S. (2022) StockGAN: Robust Stock Price Prediction Using GAN Algorithm. *International Journal of Information Technology*, **14**, 2309-2315. https://doi.org/10.1007/s41870-022-00929-6

[9] Wu, H., Wu, H., Zhu, M., *et al.* (2017) A New Method of Large-Scale Short-Term Forecasting of Agricultural Commodity Prices: Illustrated by the Case of Agricultural Markets in Beijing. *Journal of Big Data*, **4**, Article No. 1. https://doi.org/10.1186/s40537-016-0062-3

[10] Kumar, G., Jain, S. and Singh, U.P. (2021) Stock Market Forecasting Using Computational Intelligence: A Survey. *Archives of Computational Methods in Engineering*, **28**, 1069-1101. https://doi.org/10.1007/s11831-020-09413-5

[11] Xu, J.L. and Hsu, Y.L. (2022) The Impact of News Sentiment Indicators on Agricultural Product Prices. *Computational Economics*, **59**, 1645-1657. https://doi.org/10.1007/s10614-021-10189-4

[12] Zhang, Y., Li, Y. and Shen, D. (2022) Investor Attention and the Carbon Emission Markets in China: A Nonparametric Wavelet-Based Causality Test. *Asia-Pacific Financial Markets*, **29**, 123-137. https://doi.org/10.1007/s10690-021-09348-2

[13] Zhang, Y.P., Chen, Y., Wu, Y. and Zhu, P.P. (2022) Investor Attention and Carbon Return: Evidence Form the EU-ETS. *Economic Reaearch-Ekonomska Istraživanja*,

35, 709-727. https://doi.org/10.1080/1331677X.2021.1931914

[14] Zhang, Y.J., Chu, G. and Shen, D.H. (2022) The Role of Investor Attention in Predicting Stock Prices: The Long Short-Term Memory Networks Perspective. *Finance Research Letters*, **38**, Article ID: 101484. https://doi.org/10.1016/j.frl.2020.101484

[15] Han, L.Y., Li, Z.Y. and Yin, L.B., (2017) The Effects of Investor Attention on Commodity Futures Markets. *Journal of Futures Markets*, **37**, 1031-1049. https://doi.org/10.1002/fut.21853