Analysis of Factors on the US Soybean Pricing in China's Import Market – Based on principal component analysis

Zhuoying Ye, Jinggui Ma[a]

College of Economics in Yangtze University, Jingzhou, Hubei, China

Abstract. Using quarterly data from 2002 to 2014, based on principal component analysis, this paper studies the factors affecting the US soybean pricing in China's import market, and results show that China's soybean import quantity from the US has positive impact on the US soybean pricing in China's import market, but its impact is very weak. China's gross domestic product, China's domestic soybean production price index, the yuan-dollar exchange rate, the yuan-Argentine peso exchange rate and the yuan-Brazilian real exchange rate have negative impact on the US soybean pricing in China's import market. According to this, there are some recommendations: maintain the growth of gross domestic product; improve soybean TFP to reduce soybean producer price index reasonably; take advantage of the exchange rate mechanism effectively to adjust the cost of soybean imports.
Keywords. soybean imports; principal component regression; soybean pricing

I Introduction

With the increasing demand of China's soybean imports, the world soybean import trade gradually concentrates to China. In 2012, the soybean imports of China are 58.38 million tons, accounting for 58.5 percent of world total soybean exports. In addition to the increasing total imports, China's soybean import structure has changed dramatically, from the US-led structure in the early 1990s changes gradually to today's competitive structure in which the three countries--the United States, Brazil and Argentina--compete with each other. With China's soybean imports from Brazil and Argentina growing, the US soybean market share in China's soybean import market fell from more than 90 percent in the early 1990s to around 45 percent in 2012[b]. Facing more competitors in China's soybean market, the US soybean pricing would inevitably be affected. So in the case of structural changes in China's soybean imports today, what factors would affect the US soybean pricing in China's soybean market specifically? How about the affected direction and extent of these factors?

Foreign scholars are more concerned about soybean seller's market power. These related literature mainly appeared in the later 1990s. Heien and Pick (1991) empirically analyze the import demand structure of soybeans and soybean meal through building demand equation, and find that the competition among different exporting countries in the soybean import market is fierce, so the

---

seller's market power does not exist[1]. Deodhar and Sheldon (1997) also address that the seller's market power does not exist in world soybean and soybean meal export market. Song et al. (2009) construct a partial equilibrium model, and do research on which party has more market power in China-US soybean trade, and finally find that China has greater market power[2].

Chinese scholars study soybean pricing mainly from three aspects. The first is the study of the international soybean prices and its influencing factors. Liyong Du (2013) finds that the most important factors affecting the long-term trend of the international soybean prices are biomass energy development and population growth[3]. The second is the study of the domestic soybean prices in China and its influencing factors. Jianwei Li and Hu Sang (2013) examine that domestic soybean production and import volume are not the Granger cause of domestic soybean price volatility, while the money supply (M2) is its Granger cause[4]. Finally is the research on relationship between international soybean prices and domestic soybean prices. Rui Wang and Zhuo Chen (2011) do empirical research through error correction model, and results show that China and the US soybean price in spot market and future market have long-term stable equilibrium relationship[5]. Research of Min Li et al (2012) establish VAR model and examine that international soybean prices is Granger cause of domestic soybean prices, not vice versa[6]. Xunbo Chen (2013) analyzes empirically through the use of VAR model, and finds that price volatility of imported soybean is the most important factor of the factors leading to price volatility of Chinese domestic soybean currently[7]. Changming Song et al (2013) find that China and the US soybean market have mutual influence on each other, but the influence of the US soybean market on Chinese soybean market is stronger[8]. Wei Si (2013) studies the seller's import market power though research on the United States, Brazil and Argentina soybean pricing in China's market, and finally finds that all three countries do not have seller's market power[9].

In summary, much work has been done on the seller's market power and soybean prices of world soybean export market. But for the purposes of these existing research, more attention are paid to the international market or the Chinese domestic market, studies on the US-China soybean are not much, and few studies add the roles of Brazil and Argentina; Secondly, in terms of content of research, studies on specific factors of US soybean pricing exported to China are scant. Based on the existing literature, this paper do further empirical research on the role of the US soybean in China's soybean import market, aims to study the influencing factors and their influencing extent of US soybean pricing in China's soybean import market, and to make some suggestions for China to seize initiative in soybean imports.

2 Research Design

2.1 Variable Selection

2.1.1 The dependent variable.

The dependent variable is the price of China's imports of soybeans from the United States. In the analysis, we adopt the US total soybean export amount to China divided by the total number in a specified period, that is, the average price level of China's soybean imports from the United States within a specified period to represent the dependent variable.

2.1.2 The explanatory variables.

(a) China's soybean import quantity from the US. China applies automatic licensing system to soybean imports, without the imposition of import quotas, so the number of China's soybean imports from the US represents China's market demand for US soybeans. More imports represent more market demand for US soybeans; therefore the US soybean prices may be higher. (b) China's real GDP. On behalf of China's economic development and their domestic aggregate demand,

more GDP would on the one hand means stronger national strength, so international influence may be greater, the discourse power may be relatively stronger, thus maybe one of the factors considered in US soybean pricing, on the other hand means larger total demand and may results in higher soybean prices, which may have an impact on the US soybean pricing. (c) China's domestic soybean producer price index. China's domestic soybean production cost is closely related to China's domestic soybean prices, which may also have an impact on the US soybean pricing. (d) Exchange rates. It includes the US dollar, the Argentine peso and the Brazilian real exchange rates. The US dollar exchange rate directly affects China's import and export soybean prices, the Argentine peso and the Brazilian real exchange rate represent behaviors of competitors of the US, which would have influence on soybean prices these competitors of the US would offer, thus may have an impact on the US soybean prices.

Table 1. Variable Interpretation Table

Variable	Letters	Interpretation
The price of China's soybean imports from the US	P	the price of China's soybean imports from the US (Yuan/Mt)
China's soybean import quantity from the US	Q	China's soybean imports from the US (Mt)
China's gross domestic product	GDP	China's real gross domestic product (million yuan)
China's domestic soybean producer price index	PPI	China's domestic soybean producer price index
the US dollar exchange rate	Em	The RMB against the US dollar
the Argentine peso exchange rate	Ea	The RMB against the Argentine peso
the Brazilian real exchange rate	Eb	The RMB against the Brazilian real

2.2 Data Description

In the process of data collection, most of the monthly data could be got, but some could not and just have quarterly data. Considering that data adjustment from low frequency to high frequency would inevitably lead to some distortion, so this paper takes quarterly data for research. If some could just get monthly data, we would adjust the monthly data to quarterly data. Data span is the first quarter of 2002 to the third quarter of 2014.

Data of P, Q variables come from the USDA' website, data of GDP, PPI variables come from the China's National Bureau of Statistics website. The Argentine peso against the US dollar exchange rate and the Brazilian real against the US dollar exchange rate data come from the China's National Bureau of Statistics website, international data, and the RMB against the US dollar exchange rate data come from the IMF website, then get exchange rate of the RMB against the Argentine peso exchange rate and the RMB against the Brazilian real exchange rate with calculation.

2.3 Model Specification

In order to eliminate the influence of dimensionless, logarithmic processing is adopted to traditional multiple linear model. Lag and time trend is added to the model too. Finally we construct the following double logarithmic econometric model to do empirical research. Model of the form is:

$$lnP_t = c + \sum_{j=1}^{n} \alpha_j lnF_{tj} + \theta lnP(-1) + \omega T + \mu_j \qquad (1)$$

Where in (1), t is time, Pt is soybean prices in time t, c is the intercept, αj, θ, ω are the parameters to be estimated, Ftj represents factor j in time t, which represents all of the explanatory variables, $lnP(-1)$ represents the lag, T represents time trend, μj is a random error term.

3 Empirical Tests

3.1 Stationary Test

In order to eliminate "spurious regression" or "dummy regression", stationary test prior analysis of regression is necessary. There are many methods for stationary test, here we adopt ADF test method to do unit root test for time series. Results are shown in Table 2.

Table 2. Results of Unit Root Test

Variable	ADF test statistic	1% critical value	5% critical value	10% critical value	P-value	Results
DLnP	-6.040804	-3.57131	-2.922449	-2.599224	0.0000	stationary ***
DLnQ	-4.9	-3.592462	-2.931404	-2.603944	0.0002	stationary ***
DLnGDP	-2.634202	-3.584743	-2.928142	-2.602225	0.0938	stationary *
DLnPPI	-4.211281	-3.588509	-2.929734	-2.603064	0.0018	stationary ***
DLnEm	-2.764116	-3.57131	-2.922449	-2.599224	0.0709	stationary *
DLnEa	-9.492168	-3.57131	-2.922449	-2.599224	0.0000	stationary ***
DLnEb	-5.024929	-3.57131	-2.922449	-2.599224	0.0001	stationary ***

Note: "*", "**" and "***" represent the rejection of the null hypothesis at 1%, 5% and 10% significance level respectively.

As can be seen from Table 2, although logarithm of the price of China's imports of soybeans from the United States, China's soybean import quantity from the US, China's gross domestic product, China's domestic soybean producer price index, the US dollar exchange rate, the Argentine peso exchange rate, and the Brazilian real exchange rate are non-stationary, but they are stationary after the first difference. So these seven logarithmic sequence have the same order, i.e., they are all I (1), on the basis of this, the regression analysis can be performed.

3.2 Principal Component Regression Analysis

To eliminate multicollinearity occurs between the explanatory variables, we calculate the correlation coefficients between the explanatory variables firstly. The matrix of correlation coefficients between the explanatory variables is shown in Table 3.

Table 3. The Matrix of Correlation Coefficients between the Explanatory Variables

	LnQ	**LnGDP**	**LnPPI**	**LnEM**	**LnEA**	**LnEB**
LnQ	1					
LnGDP	0.331253558	1				
LnPPI	0.236536527	0.935777126	1			
LnEM	0.283479782	0.95509193	0.933523772	1		
LnEA	0.184929938	0.878566674	0.835523394	0.90615534	1	
LnEB	-0.192394117	-0.17280405	-0.149350843	-0.130654366	0.166524438	1

As can be seen in table 3, not a few correlation coefficients are greater than 0.8. In this case, doing least squares regression directly is generally less effective. We could extract the first few (say m) principal components of these variables, and create a regression model that uses principal components as the explanatory variables, like the following

$$Y = a + b_1 Z_1 + b_2 Z_2 + b_3 Z_3 + \ldots\ldots + b_m Z_m \quad (m \text{ is less than or equal to the number of}$$

explanatory variables p) (2)

So p-ary is reduced to m-ary, the influence generated by the correlation between the variables is eliminated, what's more, the structure of the regression equation is simplified. But in most cases,

the economic meaning of the extracted principal components are not clear, so it usually use the inverse transformation after getting the principal component regression equation to translate the principal component regression equation into the regression equation of the original variables.

First, extract principal component for the raw data using Eviews software. Results are shown in Table 4 and Table 5.

Table 4. Results 1 of Principal Component Extraction

No.	Eigenvalues	Contribution rate	Cumulative contribution rate
The first principal component Z1	3.831222	0.6385	0.6385
The second principal component Z2	1.184446	0.1974	0.8359
The third principal component Z3	0.795539	0.1326	0.9685
The fourth principal component Z4	0.110293	0.0184	0.9869
The fifth principal component Z5	0.043107	0.0072	0.9941
The sixth principal component Z6	0.035394	0.0059	1

According to the output results, the six eigenvalues are 3.831222, 1.184446, 0.795539, 0.110293, 0.043107 and 0.035394 respectively. Cumulative contribution rate of the first two principal components has reached 83%, and the cumulative contribution rate of the first three principal components is more than 96%. At the same time, we get the relationship between each of the main components and variables, as is shown in Table 5.

Table 5. Results 2 of Principal Component Extraction

Variable	Factor of the first principal component PC 1	Factor of the second principal component PC 2	Factor of the third principal component PC 3	Factor of the fourth principal component PC 4	Factor of the fifth principal component PC 5	Factor of the sixth principal component PC 6
LnQ	0.184744	-0.520856	0.829691	0.01813	-0.059093	-0.048693
LnGDP	0.501698	-0.029161	-0.059711	0.01786	0.842958	0.181622
LnPPI	0.488423	0.019737	-0.151061	0.738598	-0.236703	-0.369701
LnEM	0.501773	0.029355	-0.080621	-0.135624	-0.455128	0.717862
LnEA	0.468691	0.294651	0.052929	-0.619452	-0.136165	-0.537072
LnEB	-0.064529	0.799868	0.525288	0.227399	0.064956	0.155534

In empirical studies, the accumulative contribution rate of principal component is required to reach 70%. In order to improve the accuracy of the study, we sequentially select the first two principal components and the first three principal components respectively to do the research. According to Table 4, the relationship between the first three principal components and each variable are as follows:

$$Z1=0.18*LnQ+0.5*LnGDP+0.49*LnPPI+0.5*LnEm+0.47*LnEa-0.06*LnEb \quad (3)$$
$$Z2=-0.52*LnQ-0.03*LnGDP+0.02*LnPPI+0.03*LnEm+0.29*LnEa-0.8*LnEb \quad (4)$$
$$Z3=0.83*LnQ-0.06*LnGDP-0.15*LnPPI-0.08*LnEm+0.05*LnEa-0.53*LnEb \quad (5)$$

Select the first two principal components and the first three principal components respectively, and using model (1) to do regression analysis, and then we get two regression equations of principal components, as is shown in Table 6.

Table 6. Results of Principal Component Regression

Variable	(1)	(2)
C	1.687459**	1.676542**
	(2.13932)	(2.088117)
Z1	-0.19882***	-0.190121
	(-2.002902)	(-1.527716)

Z2	-0.100494*	-0.106705
	(-2.724309)	(-1.656251)
Z3		-0.00619***
		(-0.118232)
LNP(-1)	0.889338*	0.881538*
	(7.728487)	(6.59053)
@TREND	0.013199**	0.012744***
	(2.199099)	(1.774173)
F	75.17426	58.82445
Adjusted R^2	0.858257	0.855082
D.W.	1.790368	1.778585

Note: "", "**" and "***" represent the rejection of the null hypothesis at 1%, 5% and 10% significance level respectively. The corresponding values in parentheses are t values.*

According to Table 5, the two models generally fit well, and also pass the F test. Given a 10% significance level, the factor of each variable in the first model all reject the null hypothesis, while the factors of the first principal component and the second principal component in the second model do not reject the null hypothesis. What's more, the goodness of fit of the equation is also slightly lower. Therefore it is not suitable to add the third principal component. According to this, the standardized regression equation is:

$$LNP=1.69-0.2*Z1-0.1*Z2+0.89*LNP\ (-1)\ +0.01*T \tag{6}$$

Considering that the economic meaning of the extraction of principal components is not clear, we use the inverse transform, that is, to bring equation (3) (4) (5) to (6) respectively, and to translate the principal component regression equation into the regression equation of the original variables:

$$LnP=1.69+0.02*LnQ-0.1*LnGDP-0.1*LnPPI-0.1*LnEm-0.12*LnEa-0.07*LnEb+0.89*LNP\ (-1)\ +0.01*T \tag{7}$$

According to the regression equation of the original variables (7), we could draw the following conclusions: China's soybean import quantity from the US has a positive impact on the price of China's soybean imports from the US, but the impact of it is weak, when China's soybean import quantity from the US moves 1%, the price of imported soybeans from the US would move 0.02% in the same direction; China's GDP has a negative effect on the price of China's soybean imports from the US, when China's GDP moves 1%, the price of imported soybeans from the US would move 0.1% in the inverse direction, which indicates that the improve of China's comprehensive national strength would help to ease the rising trend of the US soybean prices; China's domestic soybean producer price index has a negative impact on the price of China's soybean imports from the US, when China's domestic soybean producer price index changes 1%, the price of imported soybeans from the US would change 0.1% in the inverse direction, which indicates that China's own lower soybean production cost would dampen the rising trend of the US soybean prices; exchange rate factors have a negative impact on the price of China's soybean imports from the US, when the US dollar exchange rate, the Argentine peso exchange rate and the Brazilian real exchange rate changes 1%, the price of China's soybean imports from the US would change 0.1%, 0.12% and 0.07% in the inverse direction respectively, which indicates that not only the price of currency of United States itself, but also the price of its competitors' currency would also affect the price of China's soybean imports from the US by affecting its competitors' soybean prices; lagged variable has a significant positive impact on the price of China's soybean imports from the US, which indicates that the price of China's soybean imports from the US in the prior period has a significant impact on the price of China's soybean imports from the US in the present period; the time trend variable has significant positive impact on the price of China's soybean imports from the US, which shows that other possible variables have a significant positive impact on the price of China's soybean imports from the US.

4 Conclusions and Recommendations

Results of principal component regression show that the number of China's soybean imports from the United States has a positive impact on the price of China's soybean imports from the United States, but the impact of it is weak. China's gross domestic product, China's domestic soybean production price index and exchange rate factors have a negative impact on the price of China's soybean imports from the United States. The demand for soybeans is rigid and soybean import quantity from the US is difficult to control, considering that the impact of it is weak, we could try to control other influencing factors.

First, maintain the growth of gross domestic product. While academia and practical circle advocate not pursuing growth in gross domestic product blindly, and also call for the transformation of economic development mode, but GDP growth in the volume should still continue. Second, increase total factor productivity of soybean, and to reduce the soybean producer price index reasonably. Domestic soybean industry has an urgent need of technological breakthroughs and institutional innovation, and the increase of scientific and technological investment, thus to improve TFP of soybeans, and reduce production costs and processing costs of soybeans to win more adjustment room for domestic soybean prices. Finally, take advantage of the exchange rate mechanism effectively to adjust the cost of soybean imports. Implement exchange rate adjustment mechanism according to the impact generated by changes of exchange rate to change the cost of imported soybeans from different countries to the extent possible, thus to reduce the impact on China's soybean market generated by volatility of the price of our soybean imports from the US.

References

1. D. Heien, D. Pick, The Structure of International Demand for Soybean Products, Southern Journal of Agricultural Economics, 1991(1): 137-143.
2. S. Y., Deodhar, I. M. Sheldon, Market Power in the Export World Market for Soy meal Exports, Journal of Agricultural and Resource Economics, 1997,22 (1): 78-86
3. Liyong Du, Analysis of factors on volatility of soybean prices in international market, Journal of Prices Monthly, 2013(7): 9-12.
4. Jianwei Li, Hu Sang, Analysis of factors and suggestions on volatility of soybean prices, Journal of Prices Monthly, 2013(1): 31-33.
5. Rui Wang, Zhuo Chen, Empirical analysis of soybean price relations between China and the US--based on statistic of four future markets and spot markets, Journal of Wuhan Finance Monthly, 2011(10): 30-33.
6. Min Li, Gucheng Li, Zhongchao Feng, Research on integration of domestic and international soybean prices based on the VAR model, Journal of World Agriculture, 2012(4): 42-47.
7. Xunbo Chen, Impact of volatility of imported soybean prices on domestic market--empirical analysis based on the VAR model, Journal of Price Theory and Practice, 2013(8): 83-84.
8. Changming Song, Chongguang Li, Juan Xu, Research on integration of Chinese and American agricultural markets and price transmission mechanism--illustrated by the example of soybean market, Journal of World Economy Study, 2013(3): 35-40+88.
9. Wei Si, Meng Zhang, soybean import market of China: competitive structure and market power.Journal of Chinese Rural Economy, 2013(8).