Effects of Price Volatility and Surging South American Soybean Production on Short-run Soybean Basis Dynamics

Rui Zhang*

And

Jack Houston

Paper presented at the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management
St. Louis, Missouri, April 18-19, 2005

*Zhang is a graduate student in the Department of Agricultural &Applied Economics, University of Georgia (ruizhang@uga.edu) and Houston is a professor in the Department of Agricultural & Applied Economics, University of Georgia.

Effects of Price Volatility and Surging South American Soybean Production on Short-run Soybean Basis Dynamics

Practitioner's Abstract

This study investigates the effects of South American production (SAP) and futures volatility on the soybean price dynamics in terms of their effects on the basis. The results of the econometric model showed that both South American production and futures volatility of the nearby contract have negative effects on the basis though in the forecast model, lagged values of these two factors failed to predict basis change in the future. If information about the change of the expected SAP or futures volatility is available, then the model can predict the changes in basis. This information would be helpful for hedgers to decide the time to lift their hedge.

Key Words: South American Soybean Production, Volatility of Futures Price, Basis

Introduction

The scenario considered in this study is that, following U.S. soybean harvest in November, farmers (or processors) may need to make a decision from three choices: to sell (buy) immediately after harvest in the spot market, to hedge in the futures market, or to wait for a more favorable spot price and sell (buy) at the spot market in later month. The last choice speculates in the spot market for farmers because farmers are holding long positions in spot market. To make such a decision, producers (or processors) need to compare the November spot price, the expected price they can get if they hedge and expected spot price in the future.

Under the efficient market assumption, futures price is an unbiased forecast of the spot price at the contract maturity, because all available information for the period has been incorporated. However, even though the futures market is efficient, futures price can be a poor forecast of the realized price in the maturity month, since many unforeseeable events can occur during the interim. Thus, if the futures price of a commodity is what the producers or processors are willing to accept, they are not guaranteed to that futures price when they need to sell or buy their stocks in the cash market, maybe not even a near one. Under the assumption that farmers and processors are risk averse, holding stocks or delaying order would be outside of their decision set.

Hedging in the futures market, on the contrary, could provide producers or processors with the tool to lock in the price they would accept profitably. By hedging in the futures market, holders of stocks or processors get a price that differs from what they expect only by the difference between the expected basis and actual basis. If the expected basis is a close forecast to the actual basis, hedgers get the price that they expect. If hedgers can predict the basis change, they can take advantage of this information to decide when to lift their hedge. Therefore, a more accurate basis forecast could help farmers or processors to make a better decision. Basis changes are more predicable than cash price changes based on two fundamental reasons. One is that cash price and futures price converge as the futures contract nears maturity. The second reason is that the theory of storage predicts that basis for storable goods is lower bounded by the negative sum of the marginal storage cost and time value. The potential value of basis forecasts to hedging decisions has been emphasized by Tomek. There has been a number of academic works targeting basis analysis or basis forecast[1]. However, as noted by Tomek and restated by Taylor, basis analysis and basis forecasting haven't been the subjects of many published studies.

A common practice to forecast basis is to use the average of historical basis as the expected basis. However, this naïve method fails to incorporate information available in the current year and provides the analysts with no understanding of the relationship between economic factors and basis. The purpose of this paper is to investigate the effects of economic factors on the soybeans intra-year basis behavior with the most up-to-date information included. In particular, we estimate an econometric model to evaluate the performance of the January and March soybeans basis. In addition to the commonly considered factors such as historical basis, interest rate, demand, and supply conditions, we are particularly interested in the effects of two new factors, South American soybeans supply and volatility of futures price.

South American soybeans supply is proposed as a candidate factor because, from the U.S. soybeans harvest in November to South American soybeans harvest in May, expectation for the soybeans harvest in South America could affect U.S. soybeans producers and processors storage and marketing decisions. This might in turn affect price and basis. Volatility of futures price implicates the uncertainty about futures. If the volatility increases in the period of concern, it suggests that uncertainty about future markets increases, inventory might be built up, and basis would be affected. The meanings of this study are twofold. First, it is another trial on the not-very-fruitful basis behavior area. Two new factors and their relationships with basis will be investigated. Secondly, if the effects are detected, they could be used to guide short-run basis forecast, at least qualitatively.

Surging South American soybeans production

Argentina and Brazil are the two main soybean producing countries in South America. The fast growth of soybeans production in these two countries has made South America a major soybeans supplier in the world market (Figure 1). Their proportion in world soybeans production has increased from 5% in early 1970s to 43% in 2004/2005[2]. Their combined exports have increased from 2% of the world exports in the early 1970s to 41% in 2004/2005.[3] In the same time period, U.S. soybeans production and exports have dropped as a proportion of the in the world market from 73% and 96% to 38 % and 46%, respectively. The surging growth of South American soybeans production has changed the world from depending on one soybeans harvest to two major harvests, one in the U.S. in early November and one in Brazil and Argentina by late March.

Plato and Chambers studied the impact of the semiannual production on the season-average

[1] Philip Garcia and Dwight R. Sanders(1996), W.G. Tomek(1997), Binrong Jiang and Hayenga (1997), M. Taylor, *et al*.(2004). For a detailed literature review, see Binrong Jiang (1997).
[2],[4] Data source: Foreign Agricultural Service, official USDA estimates.

price received by U.S. soybean farmers. They found that the new forecasting model with South American soybeans production as a second explanatory variable provides a more accurate forecast than the old model including only stock-to-use ratio (SUR). They concluded that there is indeed a structural change in world soybean market.

Frechette estimated that change of November expectation about Brazil harvest in March would have a positive effect on the difference between near-contract and far-contract futures prices[4]. His explanation for the positive effect is that an expected larger Brazil harvest would stimulate consumption of U.S. stocks and result in a lower inventory level in the months before Brazil harvest. A lower inventory corresponds to a higher convenience yield, according to theory of storage, and therefore would raise the basis. In this paper, we attempt to test whether the change of the expected Brazil harvest has an impact on January and March basis. If such effect exists, then it could be incorporated to predict the basis changes in January and March.

Volatility of futures prices

The relationship between futures price volatility and underlying supply and demand conditions has been demonstrated by Kenyon as the Anderson-Danthine version. That is, the *ex ante* variance of futures prices depends on the interaction of expected demand and supply, with *ex ante* variance increasing for periods of relatively large uncertainty. Kenyon *et al*. studied the factors affecting agricultural futures price variances and found that volatility of March soybean futures price is positively related to futures price and production levels. The explanation for their findings is that when production and stock levels are relatively low compared to use, prices are higher and seem to be much more responsive to new information. Ng and Pirrong investigated the relationship between volatility and squared lagged basis to demonstrate that a strong link exists between fundamentals and industrial metal price volatility, with the reasoning that the adjusted basis is a parsimonious summary of supply and demand conditions.

Pindyck argued that one principal way that volatility affects prices is that it directly affects the marginal convenience yield. When prices are more volatile, it implies more volatile production and demand, which can lead to inventory build-ups and raise the spot price in the short run. With crude oil, heating oil and gasoline data, Pindyck found a positive relationship between volatility and basis.

However, Tilley and Campbell stated that futures must exceed the cash price by more than the net value of the storage costs and convenience yield because of the existence of a marginal risk-aversion factor (risk premium) defined by Brennan. If this is the case, with more volatile futures price, futures contract holders would be compensated by a higher risk premium, which results in a lower basis. In the economic model we construct, it is hypothesized that volatility of futures price will be related to the basis, but the sign of the relationship is not specified a *priori*. The sign estimated from the model we constructed could

[4] The difference is actually the basis defined in this study since Frechette used expiring November futures price and March futures price in November to calculate the difference.

be an indicator of the relative strength of the two opposite forces.

Data and Variables

The data used in this study are publicly available data from CBOT, FAS and ERS of USDA, and the Federal Reserve. The time period covers from November, 1975 to March, 2004.

1. Basis

Both spot price and futures price for the nearby contract month are required to calculate the basis (equals spot price minus futures price of a nearby contract on the same day). Since we don't have as long as a 30-year monthly soybean spot price from each year's November harvest to the following May harvest, we used the average of futures price at expiring month as the spot price for that month[5]. However, this limits our analysis to only two monthly basis, January basis and March basis as there is no December or February futures contracts for soybeans. The average of the same trading days' futures price for the nearby contract is used as futures price. Since the spot price is from expiring futures price, the basis used in this study should be close to the actual basis in the spots markets close to CBOT.

2. Futures volatility

Annualized volatility (e.g. Vol) for futures contract F_t was computed by the following commonly used formulas:

$$VOL_t = \left[250 \times \frac{1}{N-1} \sum_i (R_{it} - \overline{R})^2 \right]^{1/2} \times 100\%,$$

$$\begin{aligned} \textit{where } R_{it} &= \ln(F_{it} / F_{i-1t}), \\ \overline{R} &= \textit{average of } R_{it} \textit{ over n days}, \\ F_{it} &= \textit{price of futures contract at day } i, \\ t &= \textit{contract year and month.} \end{aligned}$$

N was decided by the trading days between two World Agricultural Supply and Demand Estimates (WASDE) monthly reports so that volatility reflects the information released from one report but not affected by the information from the next report[6]. Following the above procedure, we get October, November and December volatility for the next March futures contract, and October, January and February volatility for the May futures contract.

3. South American production

The annual soybeans production in Brazil and Argentina from 1976-2004 were obtained from FSA production supply online database. As for the November and January expectation of South American production (SAP) in the next harvest, we use the estimation from an ARIMA (2,1,0), which fit the historical data better than other time series models tested. For the March expectation of May harvest in South America, we use the true production as a proxy to the expectation, since March is the mid-harvest time and true production could then be estimated

[5] We tried the average of all trading days in the expiring month and the average of only the first 7 business days in the expiring month separately, but no qualitative difference emerged.

[6] For details about how N is decided, see Kenyon et al. (1987).

with a certain confidence.

4. Historical basis, Interest rate, Stock-to-use ratio and U.S. Inventory

Historical basis (HB) was calculated by taking the last 3 years' average. Treasury bills secondary market 3- and 6- month interest rate (I) from the Federal Reserve Board statistics releases and historical data website are used as the interest rate in this study. U.S. soybeans quarterly supply, disappearance, and ending stocks were provided by ERS of USDA. We calculate stock-to-use ratio (SUR) for Dec-Feb period and interpolate the SUR for Oct-Dec period. These two sets of SURs were used as the latest available SURs for January and March basis, respectively. For inventory, we use ending stocks in February directly, and interpolate December inventory from ending stocks in November and disappearance in Dec-Feb.

Econometric Models and Results

Under the framework of storage theory, we constructed the basis evaluation model as:

$$Basis_t = f\,(I_t, Vol_t, SUR_t, \text{lnestsap}, \text{lninv}_t), \qquad (1)$$

where the explanatory variables are as discussed above. We assume January basis and March basis are affected by the same set of variables, and we estimate them in the same model instead of estimating a separate model for each basis. An instrumental variable is used to capture the difference because of seasonality. In this way, the observations are doubled from 29 to 58, which adding more power to the model validation F-test and coefficients t-tests. Interest rate is used to identify the opportunity costs of carryover, such as storage costs. Volatility used in this model is the December volatility for January basis and February volatility for March basis. lnestsap and lninv are the natural logs of forecasted SAP and of U.S. ending stocks at time t. Historical basis is added to the economic model (1) to verify the year effect and seasonal patterns as well . The subscript t for interest rate, volatility, inventory and SUR, means that the latest available information at time t is used.

The estimated regression coefficients for Model 1 are presented in Table 1. The 3-month interest rate coefficient is negative (-1.66) for the regression of basis, which is consistent with the prediction of the storage theory that return from holding storable commodities for a certain period was positively related to the interest forgone (Fama and French, 1987). The higher the interest rate, the greater the basis declines given all the conditions are unchanged. The two new factors, expected SAP and volatility of the return of futures contract used in the basis calculation, both have negative coefficients. The coefficient for SAP is -3.585, which indicates that if the SAP is expected to have a 10% increase in the coming March-May harvest, basis in January and February would go down by about 0.36 cents per bushel. This estimation has the opposite sign as the effect of SAP on November basis estimated by Frechette. One reason for the disparity could be that in January and March, as an expected larger harvest is approaching, the probability of stock-out goes down and thus suppresses the demand for inventory. Although the larger expected SAP stimulated the consumption of soybeans after U.S. harvest[7] and the soybean stock is smaller in January than in November, it

[7] We did a simple regression of SUR on SAP and found a significant negative relationship between the two variables. It is

could be still larger than the desired inventory. A larger than desired inventory could result in a smaller convenience yield and a lower basis in January and March.

Coefficient of futures volatility is around -.19, showing that when the futures volatility increases by 1%, the basis is expected to decline by about 0.2 cents per bushel. As we have introduced above, volatility have been predicted to have two opposing effects on basis. One opinion is that the higher uncertainty about futures supply and demand conditions associated with the higher volatility could induce a higher convenience yield and leads to a higher basis (Pindyck). The other view is that because futures price is the sum of spot price, marginal storage cost and risk premium minus marginal convenience yield on stock, according to the theory of storage, a higher risk premium provoked by higher futures volatility would result in a lower basis. It is clear from our estimation that soybeans basis is negatively related to the volatility of nearby futures contract, and the convenience yield effect is weaker than the risk premium effect.

Because December and February volatility were calculated using part of futures data in January and March, which were also used to calculate January and March basis, the Durbin–Wu–Hausman test was performed to investigate whether volatility is endogenous or not. One month lagged volatility of the same futures contract was used as the instrument for each December and February volatility. The test result is not significant and we conclude that the December and February volatility could be treated as exogenous variables and the estimation from the model (1) above are consistent.

The coefficient of lninv, natural log of U.S.ending stock, is not significant at 0.1 significance level. Considering SUR was calculated from disappearance and the same ending stock information, it is intuitive that a high correlation exists between the two variables[8]. Thus we dropped the variable lninv and refitted the model with the results showing in table 2. In the modified model, the coefficients of all variables remain the same sign as in the original fitted model. The magnitudes of coefficients are almost the same, except for SAP and SUR. However, we can at least predict the direction of the basis change as SAP and SUR changes.

Model (1) is more of an economic evaluation model rather than a forecast model in that all the independent variables except history basis are known at the same time as, or only with a very short lag, as the dependent variable. From model (1), we can predict the change of basis once we know the variation in SAP, SUR and interest rate. This information might help hedges to choose a more profitable time to lift the hedge. However, private information might be needed to make an accurate prediction of the changes in independent variables in the future period.

To test whether the public information on the same set of independent variables in November can forecast the changes of January and March basis, we fit the following two basis forecasting models:

---

estimated that for 1% increase in SAP, SUR goes down by 0.006.

[8] The correlation is 0.48 between the SUR and lninv with p value 0.0002.

$$\text{Basis}_{Jan} = g^1\ (I_{Nov3}, \text{Vol}^{Mar}_{Nov}, \text{SUR}_{Nov}, \text{lnestsap}, \text{lninv}_{Nov,}\ \text{HB}) \quad (2)$$
$$\text{Basis}_{Mar} = g^2\ (I_{Nov6}, \text{Vol}^{May}_{Nov}, \text{SUR}_{Nov}, \text{lnestsap}, \text{lninv}_{Nov,}\ \text{HB}) \quad (3)$$

The results in Table 3 show that only interest rate and historical basis are consistently significant for January basis. Table 4 shows that only interest rate is significant for March basis. It seems that the values of economic factors, such as Vol and SUR, considered in model (1) in November are not close expectations of their values in January and March. The effects of expected SAP cannot be detected with information available up to November.

Conclusions

In this study, we have focused on investigating the effects of South American production (SAP) and nearby futures volatility on the basis. Since we used expiring futures price as spot price, our analysis has been limited to January and March basis and the basis should be close to the basis of the spot markets located near the CBOT.

The expected SAP constructed from ARIMA model and real historical data negatively affect the basis. Although the magnitudes of the effect are not exactly the same for models with and without U.S.ending stock, the negative sign of the effect is consistent. If the expectation about the South American harvests increases (or decreases) by 10%, the basis in January and March is predicted to decrease by 0.2 to 0.4 cents per bushel. The reason behind the negative effect could be that higher expected SAP reduces the possibility of stock out, and thus the demand for inventory, which results in a lower convenience yield and lower basis.

The results from our model show that higher futures volatility would lead to a lower basis. In particular, it is estimated that when futures volatility increases by 1%, the basis is expected to decline by about 0.2 cents per bushel. Because there are two opposite predictions of the effects in the literature, we concluded that from the soybeans data we have and the model we have estimated, the risk premium effect is stronger than the convenience yield effect.

The model estimated in this study can be used to predict the change of basis once the variation in SAP, SUR and interest rate is known. This information might help hedges to choose a more profitable time to lift the hedge. But, variations of the same economic factors in November failed to predict the changes in basis, implying that information up to November couldn't be used as good expectation of the realized value in January and March. Therefore, good expectation and awareness of the change of the expected South American production, the futures volatility and other regressors in the model are important to the practical use of the model.

REFERENCES

Fama, E. F. and French, K.R. (1987) Commodity futures prices: some evidence on forecast power, premiums, and the theory of storage, *The Journal of Business*, 60, 55-73.

Frechette, D.L. (1997) The dynamics of convenience and the Brazilian soybean boom. *American Journal of Agricultural Economics,* 79, 1108-1118.

Garcia, P., and Sanders, D.R. (1996) Ex ante basis risk in the live hog futures contract: Has hedgers' risk increased? *The Journal of Futures Markets,* 16, 421-440.

Jiang, B.R. (1997) Corn and soybean basis behavior and forecasting: fundamental and alternative approaches, PhD dissertation, Iowa State University.

Kenyon, David, etc. (1987) Factors affecting agricultural futures price variance, *The Journal of Futures Markets*, 7, 73-91.

Ng, V.K. and Pirrong, S.C. (1994) Fundamentals and volatility: Storage, spreads, and the dynamics of metals prices. *The Journal of Business*, 67, 203-230.

Pindyck, R.S. (2004) Volatility and commodity price dynamics. *The Journal of Futures Markets*, 24, 1029-1047.

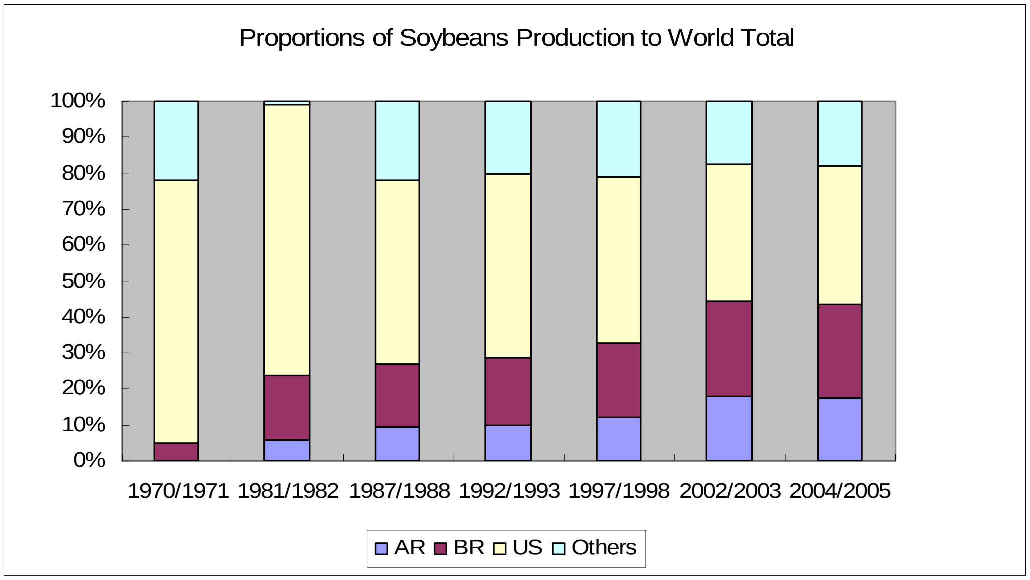
Plato, G. and Chambers, W. How does structural change in the global soybean market affect the U.S. Price? *Oil Crops Special Report*, Apr. 2004, USDA.

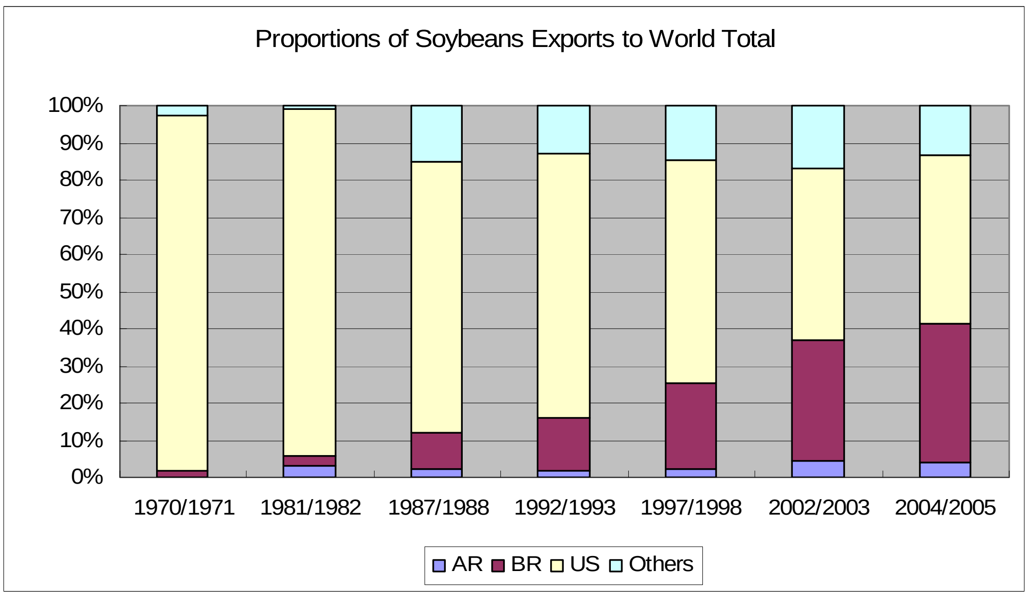
Taylor, M. ect. (2004) Incorporating current information into historical-average-based forecasts to improve crop price basis forecasts, *NCR-134 Conference on Applied Commodity Price Analysis*, *Forecasting, and Market Risk Management*.

Tilley, D.S. and Campbell, S.K. (1988) Performance of the weekly Gulf-Kansas city hard-red winter wheat basis, *American Journal of Agricultural Economics,* 70, 929-935.

Tomek, W.G. (1997), Commodity futures prices as forecasts, *Review of Agricultural Economics,* 19, 23-44.

Figure 1. World distribution of soybeans production and exports





Source: Foreign Agricultural Service, Official USDA Estimates

Table 1. Economic evaluation model with inventory included: Soybeans basis

Independent Variables	Coef.	Robust Std. Err.	t statistics	Pr> |t|
Constant	-21.7484	57.5816	-0.38	0.707
Interest rate	-1.661	0.1713	-9.69	0.000
Vol_t	-0.1828	0.0724	-2.53	0.015
lnestsap	-3.585	1.4386	-2.49	0.016
SUR	-4.4122	1.3717	-3.22	0.002
hisbs	-0.1922	0.0968	-1.99	0.053
season	-2.0508	1.485	-1.38	0.173
lninv	5.2575	3.8241	1.37	0.175
R-squared	0.7974	RMSE	2.9156	

Table 2. Economic evaluation model without U.S.inventory: Soybeans basis

Independent Variables	Coef.	Robust Std. Err.	t statistics	Pr> |t|
Constant	50.7675	19.8054	2.56	0.013
Interest rate	-1.5925	0.1395	-11.41	0.000
Vol_t	-0.2256	0.0684	-3.3	0.002
lnestsap	-1.8876	0.8161	-2.31	0.025
SUR	-3.1066	1.2469	-2.49	0.016
hisbs	-0.2100	0.0955	-2.2	0.033
season	-2.6426	1.3617	-1.94	0.058
R-squared	0.7896	RMSE	2.8898	

Table 3. Forecast model for January soybean basis

Independent	Data:Yr1976-Yr2004			Data:Yr1976-Yr1996		
Variables	Coef.	t stat	P>|t|	Coef.	t-stat	P>|t|
Constant	-198.0883	-2.1	0.047	-20.0067	-0.15	0.882
Interest rate	-2.2662	-8.49	0.000	-2.3263	-7.4	0.000
Vol_Nov	0.0527	0.39	0.703	-0.0960	-0.46	0.655
lnestsap	-2.4909	-0.91	0.373	-3.7213	-0.98	0.343
SUR_Nov	-1.1364	-0.7	0.491	-0.6476	-0.33	0.750
lninvnov	11.9179	2.29	**0.032**	4.7498	0.85	**0.409**
hisbs	-0.4915	-3.31	0.003	-0.5463	-3.6	0.003

Table 4. Forecast model for March soybean basis

Independent Variables	Coef.	Robust Std. Err.	t statistics	P>|t|
Constant	43.1711	103.1476	0.42	0.680
Interest rate	-2.0484	0.3394	-6.03	0.000
Vol_Nov	-0.0527	0.0859	-0.61	0.546
lnestsap	-0.6246	2.9244	-0.21	0.833
SUR_Nov	1.4182	2.7855	0.51	0.616
lninvnov	-1.4491	5.9768	-0.24	0.811
hisbs	-0.1363	0.1315	-1.04	0.311
R-squared	0.7095			