RESR Revista de Economia e Sociologia Rural

ISSN 1806-9479

ARTICLE

The impact of the USDA soybean crop condition reports on CBOT futures prices

O impacto da informação pública fornecida pelo USDA sobre a condição da lavoura de soja americana na expectativa de oferta e nos preços futuros na CBOT

Isadora Vercesi Bethlem[1], Roberto Arruda de Souza Lima[2], Lilian Maluf de Lima[2]

[1]Department of Agricultural Economics, Institute of Agriculture and Natural Resources (IARN), University of Nebraska-Lincoln (UNL), Lincoln (NE), United States. E-mail: ivercesibethlem2@unl.edu

[2]Departamento de Economia, Administração e Sociologia, Escola Superior "Luiz de Queiroz" (ESALQ), Universidade de São Paulo (USP), Piracicaba (SP), Brasil. E-mails: raslima@usp.br; lmlima@usp.br

How to cite: Bethlem, I. V., Lima, R. A. S., & Lima, L. M. (2023). The impact of the USDA soybean crop condition reports on cbot futures prices. Revista de Economia e Sociologia Rural, 61(2), e257641. https://doi.org/10.1590/1806-9479.2022.257641

Abstract: Soybean price formation in the Chicago Board of Trade (CBOT) is determined by many variables, with supply expectations being one of the most critical. The United States Department of Agriculture (USDA) publishes the Crop Progress report (CPR) weekly and, among other information, the public report contains an evaluation of current growing conditions in areas under soybean cultivation in the country. Agent awareness of crop conditions before harvest should affect their expectations of the soybean volume (supply) that will enter the market and should affect soybean futures contract prices, possibly in a predictable manner. This study is designed to examine this hypothesis by determining if the CPR's weekly release has a predictable impact on the following day's soybean futures contract price. Between 1995 and 2018, a 1% increase in soybean crop area evaluated as "good" and "excellent" (Condition variable) in the weekly CPR reduced soybean futures contract prices by 0.45% the day following the report's release and vice versa, and that the price trend ramped notably upward in 2008.

Keywords: *supply expectation, soybean, public information, linear regression, CPR.*

Resumo: A precificação da soja na bolsa de Chicago (CBOT) é determinada por diversas variáveis, sendo a expectativa de oferta uma das mais importantes. O Departamento de Agricultura dos Estados Unidos (USDA) publica semanalmente, ao longo da safra americana, o relatório de progresso das culturas, chamado de Crop Progress report (CPR). Entre outras informações, o relatório consiste na avaliação das condições das lavouras que estão em desenvolvimento no campo e, entre elas, a avaliação das áreas de soja no país. As informações sobre as condições da cultura de soja afetam as expectativas dos agentes em relação ao volume (oferta) de grãos que entrarão no mercado após a colheita e devem refletir nos preços dos contratos futuros da soja, possivelmente de uma maneira preditiva. O presente estudo visa examinar essa hipótese e entender se a divulgação semanal do CPR tem impacto previsível no preço dos contratos futuros da soja no dia seguinte à sua divulgação. Entre 1995 e 2018, estimou-se que a variação de 1% na avaliação das áreas consideradas como "boas" e "excelentes", entre um relatório e outro, variou, no sentindo contrário, os preços futuros em 0,45% Notou-se também que os preços atingiram um novo patamar de preço em 2008.

Palavras-chaves: *expectativa de oferta, soja, informação pública, regressão linear, CPR.*

1. Introduction

Commodity trading agents are motivated to reduce uncertainties in a market that tends to perfect competition, and all product supply information contributes to helping predict future prices (Frank, 1997). The futures market allows traders to speculate about future supply using current crop condition information, much of which is derived from surveys conducted by public agencies. This information influences decision made by commodity agents, both hedgers and speculators, throughout the world and should impact futures prices at the CBOT (Chicago Board of Trade).

In general, a futures contract's price is a reflection of the spot price plus the agents' expectations regarding factors that will affect that price at a future date. We highlight costs, volumes demanded and supplied, exchange rates, seasonality, and interest rates among these factors. This article examines the effect that soybean crop condition information contained in the CPR (Crop Progress report) has on soybean supply volume expectations and its concurrent impact on soybean futures contract prices.

In this context, the objective of the present article is to evaluate the impact that information provided by the USDA (United States Department of Agriculture) regarding the U.S. soybean crop's condition has on CBOT November soybean futures contract prices. November was selected as the date for the contract's stipulated product delivery because that month normally marks the end of the soybean harvest in the United States. This expected supply information is critical when making investments or securing protection (hedging) decisions.

2. Theoretical Foundation[1]

The weekly Crop Progress report produced by the United States Department of Agriculture's (USDA) National Agricultural Statistics Service (NASS) is the source of much of the data used in this study. The USDA is a government agency responsible for leadership in the development of laws, regulations, and information that address food, nutrition, agriculture, economic development, rural development, and the conservation of natural resources, among other public policy interests. Within the USDA, NASS conducts research to provide information used to understand and monitor U.S. agriculture. Several entities rely on the information and statistics provided by NASS, such as the U.S. Government when making policy, agricultural economists, agribusiness companies, and other agricultural market stakeholders.

The work carried out by NASS allows an organized flow of agricultural goods and services in terms of production, processing, and marketing that reduces the uncertainties and risks associated with these processes. The statistics NASS collects and distributes are used to project trends and ensure greater balance in the sector's economy.

Since 1986, NASS has published the Crop Progress report (CPR), a report that contains information derived from 5000 evaluators in 48 states responding to a NASS questionnaire that makes inquiries regarding area under cultivation, productivity, prices, wages, finances, and general crop condition in respect to U.S. production of fiber and food. In 1995, the CPR's methodology to gather and arrange information was changed to what is now in practice.

During the soybean season, a new CPR is released at 4 pm (ET) on the first working day of each week. It contains the information obtained from all evaluators responding to the questionnaires and is initially compiled at the state level and then compared with state data from the previous week to determine trends. These trends are compared with data from nearby regions and average historical data. The average of each state in a particular region is weighted according to the size of the area assessed. In the final CPR, a table is displayed with information from the current week of analysis, the previous week, the previous year, and the average of the five previous years.

In general, CPR allows monitoring of the country's principal agricultural activities, such as planting and harvesting, provides information about the crop's progress during different

[1] All information about CPR was obtained from the National Agricultural Statistics Service Information (NASS) website (National Agricultural Statistics Service, 2020a).

phenological stages, and an evaluation of crop condition as it relates to harvest yield. The CPR addresses 75% of the area planted with the main American commodities, providing information on soy, corn, cotton, sorghum, wheat, rice, peanuts, barley, and oat crops. This article will present only variables referring to soy, excluding other crop and livestock pasture information.

The CPR is often referred to as the "crop progress and condition report" because it presents objective data regarding crop progress and subjective data regarding crop condition. The report's "crop progress in the field" section includes information regarding soil moisture content, days where fieldwork was possible, the crop's phenological stage, the percentage of the area that was planted, and the percentage of the area that has been harvested. The more subjective data is in the report's crop condition section. This information is based on an evaluator's perception of growing conditions with an eye to eventual yield. The evaluator uses pre-established parameters and standards to make this evaluation, but the evaluation itself is ultimately subjective with the objective parameters used as a guide.

Evaluators rate crop conditions as "very poor," "poor," "fair," "good," and "excellent," and each category has its own implication on future supply. It is understood that (i) very poor signals an extreme loss of potential productivity, that is, production is partially or totally lost; (ii) poor indicates that there is a high potential for productivity loss due to parameters being outside critical levels, parameters such as soil moisture, pest infestation, and disease spread; (iii) fair indicates the culture is still below the condition considered normal, and a loss of productivity is a possibility but to an unknown extent; (iv) good indicates that the prospect of achieving the expected production is within the standard because there is adequate soil moisture and diseases, pests and invasive plants are under control; (v) and finally, excellent, which means that the perspective for elevated production is above normal, when, for example, the culture is suffering low or no stress and that the presence of pests, diseases, and invasive plants are insignificant.

Despite its subjective nature, crop condition information is used to evaluate the state of agricultural production in the country as a whole. If the condition pattern changes from what is expected at a particular time of the year, it often suggests a significant alteration of the amount to be harvested. Severe pattern changes, such as that caused by the drought in the American Midwest in 2012, increase future commodity supply speculation and may lead to extreme price volatility at the CBOT. This study aims to better understand the effects of the crop condition data on price formation on the CBOT and Globex trading sites.

Figure 1 is a crop condition summary example from one week's CPR, with the numbers representing the percent of all evaluators' estimations of the soybean crop's condition. The value of the change in the combined G and E values between "Last Week" and "This Week" is used in the study's model; in this instance, the change is negative 2 or -3,08% from the prior week.

(VP=Very Poor; P=Poor; F=Fair; G=Good; E=Excellent)

	This Week					Last Week					Last Year				
	VP	P	F	G	E	VP	P	F	G	E	VP	P	F	G	E
Soybean (%)	3	8	26	**50**	**13**	3	7	25	**52**	**13**	4	10	32	45	9

Figure 1. Weekly CPR soybean crop condition summary. Notes: source from NASS, USDA of the Weekly Crop Progress Report, specifically from day 9/14/2020 as an example (National Agricultural Statistics Service, 2020b).

The progress assessment is made available in each CPR between April and November, while the crop condition assessment is included from the beginning of June to the end of October when condition monitoring is relevant. In June, 50% of the soybean crop has emerged in at least

70% of the area under cultivation. As the season reaches its end in October, the pre-harvest CPR conditions assessment focuses exclusively on mature soybeans.

It is important to notice that the CPR makes no distinction between small and large producers, soil tillage and management techniques, technological level, and type of variety being grown, among other factors; therefore, it is not possible to infer a cause-and-effect relation regarding these variables. In other words, the CPR does not provide information that allows one to diagnose slow crop progress due to possible production system bottlenecks.

One of the CPR's main objectives is to provide data that can be used to make inferences about the future market supply and reach a more general understanding about the course of the season, information of significant importance to all economic agents that trade the commodity. Lehecka (2014) noted that weekly reports are a very important source of information on the development and condition of American crops and are among the most required publications if one is to accurately predict future commodity prices.

Leftwich (1974) further explains that a change in market price only occurs through joint actions by a large number of sellers. In the case of USDA reports, the direction of any change in the amount to be supplied is evident to all participants, thereby changing futures prices. This means that if an increase in supply is expected, a reduction in the future spot price is also expected, so there will be an immediate reduction in the futures price at the CBOT since, according to Pindyck & Rubinfeld (2006), one of the factors of supply curve displacement is expectation. If there are expectations, that is, if there is information that generates motivations for agents to believe that there is to be an increase in the supply of the product in the market, the supply curve will be shifted to the right and the equilibrium price reduced.

According to Ferguson (1972), because the commodity market operates in perfect or close to perfect competition, there is no change in price when only one producer increases his sales, as many participants are offering the same product. Therefore, producers assume that any change in their volume of sales will affect their revenue, but will be insignificant in the formation of market prices. These producers do not need to reduce the price to increase the quantity sold as they could already be practicing the equilibrium price. On the other hand, if they increase the price, they would partially or totally lose their market share. In other words, the demand curve for the product is infinitely elastic.

In a perfectly competitive market, the formation of prices is often based on the agents' average expectations about future events, making information extremely important. It can be said that price negotiation is based on probabilities associated with occurrences, such as potential political, economic, and climatic events (BM&F, 2006).

Milonas (1987) explains that the information provided by the USDA does not have an additional cost but only reduces uncertainty on the part of agents. The information especially benefits those that use this information to immediately adjust their positions as sellers or buyers, since any additional clue about the harvest situation and future supply directly affects futures contract prices. The author reasons that any indication of a potential increase in the quantity offered will lead to a drop in futures prices without necessarily changing spot market prices. However, if the expectation is for a smaller harvest, future and spot market prices will rise, since it will be preferable to store harvested products for future consumption. He concludes that agents look forward to the information released by the USDA to make decisions regarding their long or short positions, which will alter the supply and demand for futures contracts, thus changing their price.

If prices react to information announced in an efficient market, then the information is valid for all participants in that market (Campbell et al., 1997). Fortenbery & Summer (1993) state that estimated supply and demand are considered to be one of the most important sources of information that affect commodity trading. For this reason, the USDA maintains maximum confidentiality until the date of disclosure.

The value of public sector information has been the subject of long debate, especially due to the changes undergone by agriculture in recent decades and the growth of companies and private banks that provide information at relatively low prices (Good et al., 2006). According to a study conducted by Falk & Orazem (1984), the release of reports by the USDA causes a reaction in market prices, which suggests that the release of public information indeed generates social benefits. On the other hand, McKenzie (2008) developed a study that raises questions about the importance of the information provided by the USDA and whether its content can be predicted by private agents. The author hypothesizes that if the private sector can make predictions about the possible soybean supply daily, then the content of the weekly CPR is predictable *per se*. However, the author notes that even with this possibility, future prices continue to react to the release of CPR since the report is an extremely credible source.

A study by Good et al. (2006) concluded that the combination of information from both private agents and the USDA leads to a reduction in the variance regarding the error related to the USDA reports alone. That is, even if the content of USDA reports is not better than that of the reports provided by private agents, its disclosure will still induce a futures price reaction. In the same study, the authors also address areas that previous studies neglected: the total impact and relative impacts that the many relevant public information reports have on a commodity's futures price and the magnitude of these impacts.

The USDA works with the release of several reports related to agriculture; however, studies of the price effects of these have not addressed the impact of one report on others, the total impact of all the reports, and mostly focus on a report's qualitative effect rather than the magnitude of this effect. Using data from 1985 through 2003, Good et al. (2006) sought to rectify these previous omissions. They developed a methodology that estimates the total futures market impact of the often simultaneous release of six USDA hog and cattle information reports on futures market prices and each report's relative importance. One of the challenges of the work was to select the best modeling of time series for changes in future commodity prices, variables that usually do not have a normal and non-linear dynamic distribution of variance.

Lehecka (2014) carried out an empirical study to verify the economic value of the information reported by the USDA using two different methodologies: (i) effects of the information tested by examining different returns of the variables for the day of the release and the following days; (ii) changes in the conditions of the tested crops for a quick and rational price reaction. To this end, the author considered that analyses of the events in question, that is, the disclosures of the reports, are based on the idea that information is valuable to agents in a given market if, and only if, prices react to that information. The change in the perception of the supply, however, would not indicate the direction of movement of the supply itself, which would only be known when new information is released.

3. Methodology

All CBOT November prices were collected using the Bloomberg Terminal (2020) -a software platform that allows extraction of historical prices at CBOT -, and the crop condition percentages were collected from the USDA through its NASS website (National Agricultural Statistics Service, 2020c). All prices and conditions are weekly, from the weeks during which crop conditions can be evaluated (usually from week 25 to week 40 of each year) from 1995 to 2018. Data collection began with the 1995 crop year: the year the USDA adopted a new methodology to determine crop conditions.

This study considers only prices for the November soybean futures contract. Of all the months for which soybean contracts have expiration dates, that is January, March, May, July, August, September, and November, only November's contract reflects the definitive soybean spot price at its expiration since it is the contract that expires nearest to the end or at the end of an average US soybean harvest (CME Group, 2013).

Trading of soybean futures contracts at the CBOT on the day of the report's 4 p.m. release can be conducted on the CME Group's Globex electronic exchange starting at 7 p.m. that day, but it was understood that the market would not have sufficient liquidity to properly reflect the report's impact during the Globex night trading session, 7 p.m. to 7:30 am the next day. For this reason, the CPR's impact on prices was determined using the natural logarithm of soybean futures prices at the 1:30 p.m. CBOT closing on the day following the report's release as the model's dependent variable.

Considering each report to be an observation, the period studied consists of 448 observations of the five crop condition evaluations ('very poor', 'poor', 'fair', 'good', and 'excellent), for a total of 2240 events. The percent of the total area classified under the different crop conditions varies in range. Over the analyzed period, the situation in which 50% or more of the area was within a single condition category occurred only for crops in the "good" category.

The area of crops evaluated as being in a specific condition fell within a particular range for all the studied period's years: the percent of the total area in 'very poor' condition always varied between 0% and 17% (with an upward trend during the harvest), 'poor' between 3% and 23% (with an upward trend during the harvest), 'fair' between 19% and 37% (with an upward trend during the harvest), 'good' between 25 and 62% (with a downward trend during the harvest), and 'excellent' between 3% and 21% (with a downward trend during the harvest).

The distribution of total crop area among crops in a specific condition category followed a pattern: crops in 'very poor' condition always covered the lowest proportion of total land under soybean cultivation, followed by 'poor,' 'excellent,' 'fair,' and the consistently higher proportion of cropland was in the 'good' condition. Figure 2 contains a box graph created using E-Views 9.5 software showing the percent of the total area cultivated with soybeans separated into particular crop conditions for the entire study period (1995 through 2018).

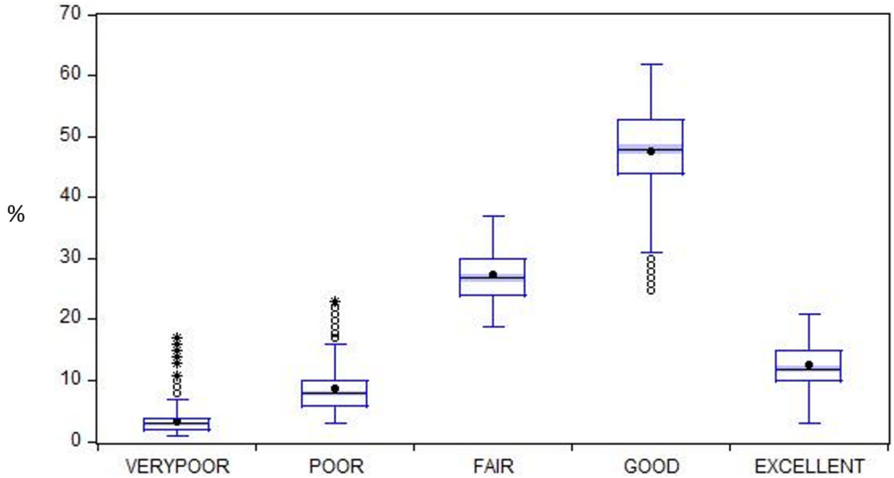

Figure 2. Box-plot graph for each CPR crop condition from 1995 through 2018. Notes: elaborated by the authors using data from the Crop Progress Reports (National Agricultural Statistics Service, 2020c).

It was found that the classification of the weekly report's total area in a generalized category of either 'very poor,' 'poor,' 'fair,' 'good,' or 'excellent' was not realistic. It was therefore decided to define the condition factor as being the sum of the percentages of the total area classified as 'good' and 'excellent' in the weekly report, considering this to be an analysis method that still captures the sensitivity of crop conditions variations. The basic premise is that these are the two conditions that will be most relevant as agents interpret the report. The sensitivity of the model consists of the variation of this sum which, if positive between one week and the following, automatically implies a drop in the sum of the other variables ('very poor', 'poor', and 'fair'), and vice versa. Therefore, the model's non-binary independent variable, the Condition variable, is the natural logarithm of the sum of the "good" and "excellent" percentage points.

A second independent variable is binary, with 0 representing information from 1995 to 2007, and 1 representing information from 2008 through 2018. This division in the study period was adopted due to a change in the price trend in 2008 and because it was the year that presented the best fit for the model.

The analysis employed multiple linear regression by the method of ordinary least squares (OLS) based on Gujarati (1995). This method explains the impact of the explanatory variable (independent) on the explained variable (dependent).

Table 1 indicates that the years 2008 through 2018 make up 47.54% of the sample, while the other years (1995 through 2007) make up the other 52.45% of the sample. Also, it shows that the mean for the Condition variable was 60.46 (percentage points). For the continuous variable, the coefficient of variation (CV—a division of the standard deviation by the mean of the variable) indicates that the Condition series varies about 14.98% on their average. This value is within the acceptable range of data variability for this variable. According to Pimentel-Gomes (1984) when studying the CVs of several agricultural tests, he proposed the following classification: low when less than 10%, medium when between 10 and 20%, high when between 20 and 30%, and very high when greater than 30%.

Table 1. Data regarding the averages and other indicators for descriptive statistical analysis of the information obtained.

Variables	Description	Mean	Standard Deviation	Coefficient of Variation (CV)
Condition	Sum of percentage points of crops considered "good" and "excellent" for each week	60.46	9.056	0.1498
Binary	0 if the information regards the years between 1995 and 2007, and 1 if it regards years between 2008 and 2018	0.4754	0.4999	-

Notes: elaborated by the authors using research data (2020). The continuous variable Condition is in absolute numbers.

Modeling was undertaking using the following Equation 1:

$$\ln Y = \alpha + \beta_1 \ln X_1 + \beta_2 X_2 + \varepsilon \quad (1)$$

in which, lnY corresponds to the natural logarithm of the November soybean contract's closing price on the day after the release of the CPR for each release week of the respective year; α, , are the estimated parameters of the proposed model; lnX_1 refers to the quantitative variable called "Condition," which represents the natural logarithm of the sum of the 'good' and 'excellent' condition values from the weekly CPR; X_2 refers to the binary variable, where

X_2 = 1 if the years are from 2008 through 2018 and X_2 = 0 if the years are from 1995 to 2007; ε matches the random error term. N (0.1) distribution is assumed.

According to Wooldridge (2009), the interpretation of coefficient β_1 is given as the variation, in percent, of the dependent variable (price) when the explanatory variable varies by one percent about its prior week's value, since both are presented through the natural logarithm. Thus, it can be considered the elasticity.

Assuming Wooldridge (2009), interpretation of coefficient β_2 is given according to Equation 2:

$$\Delta\hat{Y}\% = 100 \times [exp(\beta_2\ ) - 1] \quad (2)$$

4. Results and Discussion

E-Views 9.5 software was used to perform all econometric procedures (estimates of the model coefficients and generation of related tests). The White test was used to check for heteroscedasticity, that is, to check if there is a change in the variance of the estimated model's residuals (model error) (Wooldridge, 2009). The null hypothesis that the error variance is constant (homoscedasticity) was rejected at a one percent significance level, confirming the existence of heteroscedasticity in the initially adjusted model (White Test).

The Durbin-Watson test (DW) was used to detect positive first-order autocorrelation of the residuals. The result of the test showed DW = 0.1304 and, for our sample (n=448), the critical values found were 1.789 (upper value) and 1.748 (lower value), that is, autocorrelation was detected (0.1304 < 1.748). The Bresch-Godfrey test (BG) was also used and, although this test also detects negative kth-order autocorrelation, it helped to reassure DM test results on the positive first-order autocorrelation of the residuals. The value obtained was BG=403.11 at a 1% significance level, rejecting the null hypothesis (H_o) of the absence of autocorrelation.

Based on the results confirming both heteroscedasticity and autocorrelation, the Newey-West test (NW) was used to correct the model. According to Gujarati (1995), this method is an extension of White's robust correction, and it is only valid for samples big enough, which is our case (n=448). This method corrects standard errors in the Ordinary Least Squares (OLS) regression for autocorrelation as well as heteroscedasticity.

Also, the DW test used before and after the corrections applied by the Newey-West method presented the same results but, according to Gujarati (1995), this should not be of any concern as the NW method already considers this when correcting the OLS standard errors.

The presence of multicollinearity between the explanatory variables was discarded, and variable presented a variance inflation factor (VIF) value of less than 10; variable is binary with a value of zero for the 1995 to 2007 period, referred to as the omitted period, and 1 for the period from 2008 through 2018.

Division of the entire 1995 through 2018 period allows comparison between periods, an approach sufficient to discard the presence of multicollinearity. The division was made in 2008 because different period input tests indicated that the price trend before 2008 was distinctly different than the trend from 2008 on. The separation of the data in these two periods (before 2008 and after 2008) was also conducted by Bain & Fortenbery (2016), since, according to them, the first "period captures the market before the effects of the Great Recession and before the general increase in commodity price levels" and the second period "captures each market from the start of the Great Recession forward and reflects higher average commodity prices compared with the earlier period".

The United Nations Food and Agriculture Organization (Food and Agriculture Organization of the United Nations, 2011) reported that the increase in prices of agricultural commodities since 2006, which reached peaks in 2008 and 2011, generated inflationary pressures and concerns at global economic levels regarding food security. Kaufman (2010) adds that the low risk and high return from investments in agricultural commodities led to an intense capital flow into the futures market for these products, most notably in 2008.

According to a report in the Brazilian Center for Advanced Studies in Applied Economics (CEPEA) publication “Agromensal,” for December 2008 as an analysis of the year, prices of soybean for the first semester of that year reached a record high at the time. According to the report, among the main reasons for the increases are the rise in oil prices, the entry of funds into the commodities market, speculative purchases, and heated demand for soybeans and their by-products (Centro de Estudos Avançados em Economia Aplicada, 2008).

Through June of that year, prices for the soy complex’s goods continually increased, particularly due to climatic conditions in the United States that resulted in flooding and delayed planting and because Argentine producers protested the taxation of exported products. Even though the price for soybeans and their by-products futures contracts fluctuated downward over the following months, the year ended with a sharp year-over-year price increase as drought caused crop failure in South America adding to earlier climatic and political issues.

Garcia et al. (2012) explain some of the 2008 price increase as a result of the increase in soy production input prices (due to the increase in the price of petroleum) and the devaluation of the U.S. dollar. Clarissa Black (2015) notes that the increase in commodity prices since 2008 is unique in terms of scope, duration, and magnitude and is the only price spike that simultaneously involved three groups of commodities: agricultural, metal, and fuel.

The residuals’ behavior is similar to that of a normal distribution even though the Jarque-Bera test led to the rejection of the null hypothesis of normality of errors (asymmetry values = 0.5296; kurtosis = 2.846; sample size (n) = 448; JB statistic = 21.38 approx., with zero probability of significance, rejecting the null hypothesis of normality). According to Oliveira (2014), even if this hypothesis is not validated, a correct inference can be made if the sample is large enough to apply the law of large numbers. Although the residuals’ histogram shown in Figure 3 is similar to a normal distribution, it presents asymmetry on the right, which according to Hoffmann (2006) is expected for socioeconomic series.

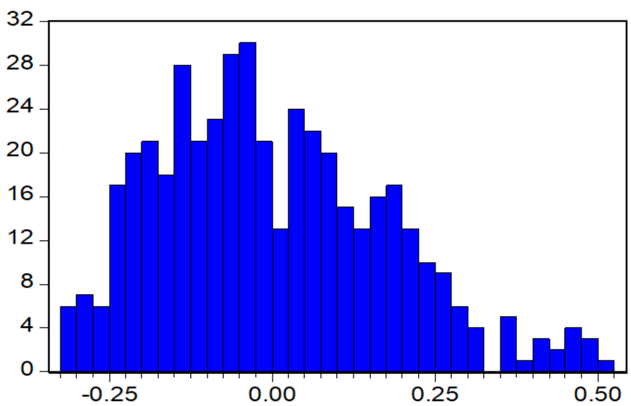

Figure 3. Residuals histogram. Notes: elaborated by the authors using research data (2020).

Figure 4 represents the series of residuals, that is, the difference between the observed and the adjusted series. The more this blue line "oscillates," the greater the evidence that the part not explained by the model is random.

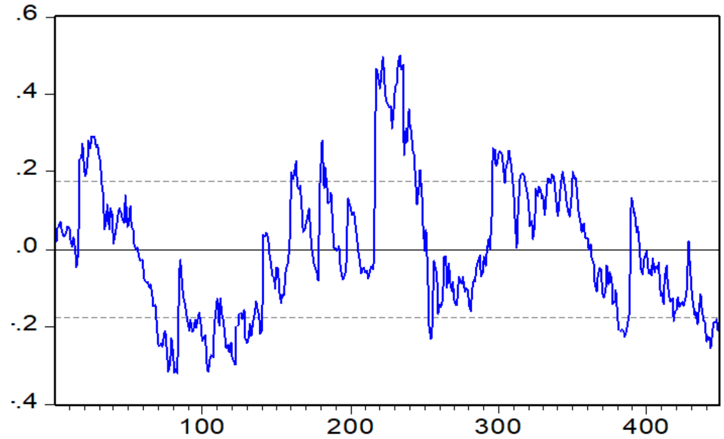

Figure 4. Regression's residuals. Notes: elaborated by the authors using research data (2020).

Having made these analyses, the results from the final model estimated through multiple linear regression using the ordinary least squares method are shown in Table 2.

Table 2. Results of the estimated coefficients for the regression model (by Newey-West method)

Exogenous variables	Coefficients	*t*-statistics	*p*-value	Percent variation
Constant	8.1849	23.1070*	0.0000	-
Condition	-0.4446	-5.1400*	0.0000	-0.45
Binary	0.6420	17.1157*	0.0000	90.02
R-square				***0.7647***
Observations				***448***
F Statistic				***723.419****

Notes: elaborated by the authors using research data (2020). Single asterisks (*) represent significance at the 1% level.

It is observed that if there is a one percent increase in the Condition variable's value (G+E), keeping all other variables constant, the price of soy in the futures market decreases by 0.45%. In other words, the impact of the information contained in the CPR leads to the agents' expectations increasing regarding soybean supply at harvest and a concurrent reduction in the futures contract price of soybeans, as expected by the adjustment of economic equilibrium after the supply curve's shift to the right.

For the Binary variable, it is observed that average prices from 2008 through 2018 are 90.02% higher than average prices practiced between 1995 and 2007. The model error, that is, changes in the futures price not explained by the changes in the "Condition" and "Binary" variables, likely represents the effect of changes in other variables on the agents' expectations and can be assumed to justify commodity futures price volatility. These other variables include costs related to soy production, storage and transportation, deterioration while in storage, international customs barriers, technical regulation modifications, changes in world demand due to interest and incomes, and interaction between the soybean and corn markets.

It is important to note that these results refer to speculative prices at the CBOT and do not necessarily portray the behavior of prices in years outside the study's period.

There was a significant increase in prices that begins in 2008 and does not retreat to prior levels. This behavior suggests that the price trend in place since 1995 had been altered, with an even more accentuated price spike in 2012.

The model's estimated coefficient of variance (R^2) being 76.5% (Table 2) indicates that the CPR's release was likely responsible for 76.5% of the November soybean futures contract's price change from one week to the other in those years and weeks understudy, while 23.5% of any price change is explained by other factors' effects. This result makes it possible to refine Lehecka's (2014) finding that the impact of the CPR's release on derivatives' short and medium-term prices is predictable by showing that an approximate value for that impact on the day following its release can be calculated, at least for the November soybean futures contract. It also demonstrates the importance of the CPR once the market is kept with the same crop condition information throughout the whole week and it is deeply affected by its release.

Using this R^2 value, traders' responses to changes in the CPR's condition evaluation can be quantified irrespective of other variables, such as inflation, demand, and transportation constraints. Our modeling of the effect of the CPR's release on the November soybean futures contract averaged over the years from 1995 through 2018 resulted in an estimated negative 0.45% change in soybean November futures prices for each one percent increase in the total CPR good and excellent crop condition ratings.

5. Conclusion

Soybeans are of fundamental importance to humanity's food supply, both directly and as a source of animal feed and biofuel. Soybean and its derivatives are traded on commodity exchanges by agents interested in the actual product and by speculators who find in these markets a way to profit from price variations. The Chicago Board of Trade (CBOT), the most important of the commodity futures and options exchanges, provides the global benchmark for soybean price formation. This study sought to verify the impact of information provided by the United States Department of Agriculture in their weekly Crop Progress and conditions report in areas cultivated with soybean on futures prices at the CBOT, more specifically on November expiration futures contracts.

It was found that information provided by the USDA report is of significant importance to agents operating at the commodity exchange. It was also found that a major break in the soybean price trend over the historical series analyzed (1995 through 2018) occurred in 2008, with other price peaks reached in the following years. Despite this, it is important to understand that price behavior throughout the studied historical series considered only soybean futures contract prices, a data source for this study, and not the final price for soy when sold in the spot market. The inclusion of variables separating the period under study into two time periods (1995 to 2007 and 2008 through 2018) was to better adjust the model and to allow for a more complete understanding of the futures contract's price evolution.

If this study's findings are confirmed over time and the assumptions and methodologies applied are determined to be appropriate, then the results from our analysis should be acceptably accurate and repeatable, although certainly in need of refinement, and give high-frequency commodities traders focused on the soybean futures market another data point to incorporate in their algorithms. To improve real-time predictive accuracy, for future studies that set out to refine this study's methodologies, incorporation of more variables and inclusion of more recent data are suggested.

6. References

Bain, R., & Fortenbery, T. R. (2016). Impact of crop condition reports on national and local wheat markets. *Journal of Agricultural and Applied Economics*, *49*, 97-119. http://dx.doi.org/10.1017/aae.2016.31

Black, C. (2015). *The price of soy in the last 10 years*. FEE. Retrieved in 2021, May 12, from http://panoramainternacional.fee.tche.br/article/o-preco-da-soja-no-ultimo-decenio/

Bloomberg Terminal. (2020). *Bloomberg Professional Services.* Retrieved 2020, February 20, from https://www.bloomberg.com/professional/solution/bloomberg-terminal/

BM&F. (2006). *Future price formation.* Retrieved in 2021, May 12, from https://mbafinancasebanking.files.wordpress.com/2014/03/mat2-para-leitura.pdf

Campbell, J. Y., Lo, A. W., & Mackinlay, A. C. (1997). *The econometrics of financial markets.* Princeton: Princeton University Press.

Centro de Estudos Avançados em Economia Aplicada – CEPEA. (2008). *Agromensal: soybean december.* Retrieved in 2021, May 12, from https://www.cepea.esalq.usp.br/upload/revista/pdf/0022499001465840559.pdf

CME Group. (2013). *Self-study guide to hedging with grain and oilseed futures and options.* Retrieved in 2021, May 12, from http://www.cmegroup.com/trading/agricultural/files/grain-oilseed-hedgers-guide.pdf

Falk, B., & Orazem, P. F. (1984). A theory of futures market responses to government crop forecasts. *Economic Staff Paper Series*, 34(150), XX-XX.

Ferguson, C. E. (1972). *Microeconomic theory.* (3rd ed.). Homewood: Irwin-Dorsey.

Food and Agriculture Organization of the United Nations – FAO. (2011). *The state of agricultural commodity markets: high food prices and the food crisis – experiences and lessons learned.* Retrieved in 2021, May 12, from http://www.fao.org/fileadmin/user_upload/ISFP/High_food_prices.pdf

Fortenbery, T. R., & Summer, D. A. (1993). The effects of USDA reports in futures and options markets. *Journal of Futures Markets*, *13*(2), 157-173.

Frank, R. H. (1997). *Microeconomics and Behavior* (3rd ed.). Nova York: McGraw-Hill.

Garcia, P., Mallory, M., & Trujillo-Barrera, A. (2012). Volatility spillovers in U.S. crude oil, ethanol, and corn Futures Markets. *Journal of Agricultural and Resource Economics*, *37*(2), 247-262.

Good, D. L., Irwin, S. H., & Isengildina, O. (2006). The value of USDA situation and outlook information in hog and cattle markets. *Journal of Agricultural and Resource Economics*, *31*(2), 262-282.

Gujarati, D. N. (1995). *Basic econometrics.* (3rd ed.). Nova York: McGraw-Hill.

Hoffmann, R. (2006). *Statistics for economists*. (4th ed.). São Paulo: Thompson.

Kaufman, F. (2010). The food bubble: how wall street starved millions and got away with it. *Harper's Magazine*. p. 27-34. Retrieved 2022, July 13, from https://www.democracynow.org/2010/7/16/the_food_bubble_how_wall_street

Leftwich, R. H. (1974). *The pricing system and resource allocation* 4th ed.). São Paulo: Editor Livraria Pioneira.

Lehecka, G. V. (2014). The Value of USDA Crop Progress and Condition Information: reactions of corn and soybean futures markets. *Journal of Agricultural and Resource Economics*, *39*(1), 88-105.

McKenzie, M. A. (2008). Pre-harvest price expectations for corn: the information content of USDA reports and new-crop futures. *American Journal of Agricultural Economics*, *90*(2), 351-366.

Milonas, N. T. (1987). The Effects of USDA crop announcements on commodity prices. *Journal of Futures Markets*, *7*(5), 571-589.

National Agricultural Statistics Service – NASS. (2020a). Retrieved 2020, July 8, from https://www.nass.usda.gov/index.php

National Agricultural Statistics Service – NASS. (2020b). *Data & statistics.* Washington, DC: NASS. Retrieved 2020, June 5, from https://www.nass.usda.gov/Data_and_Statistics/index.php

National Agricultural Statistics Service – NASS. (2020c). *Crop progress: report of day 9/14/2020*. Washington, DC: NASS. Retrieved 2020, June 5, from https://downloads.usda.library.cornell.edu/usda-esmis/files/8336h188j/qr46rp789/2r36vm941/prog3820.pdf

Oliveira, M. T. (2014). *Distância psíquica e seus efeitos sobre o fluxo de exportações dos estados brasileiros* (Ph.D. thesis). Universidade de Coimbra, Coimbra.

Pimentel-Gomes, F. (1984). *Modern statistics in agricultural research* (3rd ed.). Piracicaba: Potafos.

Pindyck, R. S., & Rubinfeld, D. L. (2006). *Microeconomics.* (8th ed.). Upper Saddle River: Pearson.

Wooldridge, J. M. (2009). *Introduction to econometrics: a modern approach.* (4th ed.). São Paulo: Cengage Learning.

Submitted on: October 24, 2021.
Accepted on: April 05, 2022.
Journal of Economic Literature (JEL) classification: C13, Q02, Q11 e Q13.