Measuring the Influence of Commodity Fund Trading on Soybean Price Discovery

by

Gerald Plato and Linwood Hoffman

Suggested citation format:

Measuring the Influence of Commodity Fund Trading on Soybean Price Discovery

Gerald Plato

and

Linwood Hoffman*

Chicago, Illinois, April 16-17, 2007

* The authors are economists at the Economic Research Service, United States Department of Agriculture. The views expressed here are those of the authors and may not be attributed to the Economic Research Service or the United States Department of Agriculture.
Measuring the Influence of Commodity Fund Trading on Soybean Price Discovery

Practitioner’s Abstract

The increase in commodity fund trading in the agricultural commodity futures markets has raised concern that this trading is degrading the price discovery performance of these markets. We used the Beveridge-Nelson Decomposition procedure to estimate the price discovery performance of the soybean futures and spot markets. We found that the price discovery performance of the soybean futures market has improved along with the increased commodity fund trading. Our results indicated that a portion of the price discovered in the soybean futures market is passed to the spot market.

Keywords

price discovery, commodity funds, cointegration, Beveridge-Nelson decomposition

Introduction

There is concern that fund trading in commodity futures markets is degrading their price discovery performance. Investment in commodity-linked funds by pension funds, hedge funds and other institutional investors has been estimated to exceed $100 billion (Reuben, chairman, Commodity futures trading Commission (CFTC)). The price discovery concern stems from commodity fund trading that is not based on market fundamentals. The National Grain and Feed Association expressed the concern that traditional agricultural hedgers and producers cannot discern the degree to which market volume and prices reflect supply/demand in making business decisions (Feedstuffs, September 11, 2006, p. 24).

The CFTC responded to the price discovery concern by providing a supplement to their weekly Commitments of Traders (COT) report that reports a new fund category for 12 agricultural commodities called ‘Index Trader’. The supplemental report is provided under a two year pilot project that began in January 07. The supplemental report like the COT report provides Tuesday’s open interest on the following Friday at 3:30 pm, Eastern Time. The new category includes traders that are hedging exposure to risk from positions in commodity index funds, such as swap dealers. Traders in this category do not generally respond to commodity market fundamentals but may influence price, particularly, because they primarily take long positions. Reporting the new data will provide additional market transparency and may result in information about index hedging being more quickly incorporated into price.

We respond to the price discovery concern by measuring the influence of fund trading on the price discovery performance of the futures and spot markets for soybeans. We include spot prices in our analysis because the conventional wisdom is that price is largely discovered in
futures markets and then passed on to the corresponding spot markets. If correct then price discovery degradation in a futures market would decrease price discovery performance in the corresponding spot market(s).

Price discovery in the modern finance literature is defined as the incorporation of new information into permanent price. Alternative procedures are used in the price discovery literature to decompose price into permanent and transitory parts. The transitory part of price is the pricing error inherent in the price discovery process. We use the Beveridge-Nelson decomposition to divide spot and futures prices for soybeans into permanent and transitory parts (Beveridge and Nelson). Pricing performance is measured by the standard deviation of the transitory part to the standard deviation of price. An increase in this ratio is evidence of a decrease in price discovery performance.

Our price discovery analysis is made using daily spot and futures prices for soybeans for the 1982 through 2005 crop years. Price performance is estimated for crop years before and after commodity fund trading began. Swap option dealers began applying for and receiving hedging exemptions in order to avoid speculative open interest position limits in 1991 (Federal Register, p. 9 of 13). They were classified as hedgers along with traditional hedgers in the Commitment of Trader reports, although their long futures positions reflect investor demand for portfolio diversification in addition to that offered by equities and bonds rather than hedging risk from positions in physical commodities.

Previous Studies

We examined three studies that assessed the influence of commodity fund trading on price discovery using special daily tabulations of open interest by type of trader. We also examined a study that used the Beveridge-Nelson decomposition to examine price discovery in aluminum futures and spot markets.

Irwin and Holt examined the effects of daily open interest changes on prices for 13 commodity futures markets including soybeans. CFTC provided the daily open interest data from a special daily tabulation of reportable trader positions from 4/4/1994 to 10/6/1994. They found that hedge funds and commodity trading advisors successfully traded based on market fundamentals. They were able to identify mispriced futures contracts, take positions in them, and trade out of them at a profit. Successful fundamental trading adds information to price. It increases the permanent part of price. Irwin and Holt found little evidence for momentum or herd trading based on past price changes. This type of trading would decrease price discovery performance by adding to the noise or transitory component of price.

Haigh, Hranaiova, and Overdahl examined the relationships between futures prices for natural gas and crude oil and the positions of several categories of traders in these two markets including managed money traders and large hedgers. Managed money traders include hedge funds. They also examined the relationships among the types of traders. A special daily CFTC tabulation of trader open interest positions from 8/4/2003 to 8/31/2004 was used. They found that managed money traders provide liquidity to large hedgers. Providing liquidity reduces price variability and the transitory price component by reducing the bid-ask bounce. They found no evidence that
managed money traders, trade on previous price changes, conditional on other types of traders. Such trading would add increase the transitory component of price.

Haigh, Harris, Overdahl, and Robe examined daily prices and daily open interest by type of trader for the crude oil futures market from the first week in January 2000 to the first week in May 2006. Again a special daily CFTC tabulation of trader open interest positions was used. They used an error correction model to show that the futures prices for nearby, one year, and two year maturity contracts became cointegrated in 2004. This result implies that futures contract prices for crude oil are being determined by information sharing between the nearby and more distant contracts. They attribute the cointegration to the increased index fund and hedge fund open interest in distant contracts which facilitates short hedging in the more distant contracts. They could have used their estimated error correction model to estimate the price discovery performance for each of the futures contracts before and after they became cointegrated.

Figuerola-Ferretti and Gilbert examined aluminum price discovery using the Beveridge-Nelson decomposition to decompose price into permanent and transitory parts. They were interested in determining if the introduction of aluminum futures trading improved aluminum spot price discovery. They found that it did as the aluminum futures market liquidity improved. They used monthly average prices from January 1970 through December 2003. Use of monthly average prices may influence results in that information available each day gets averaged over a month. Market participants make decisions based on the information available each day.

**Beveridge-Nelson Decomposition Procedures**

The multivariate Beveridge-Nelson decomposition procedures involves estimating an error correction model that contains futures and spot prices and transforming it into its dual moving average model. Coefficients and the common errors from both models are then used in the Beveridge-Nelson decomposition to separate futures and spot prices into permanent and transitory parts.

Equation 1 is the error correction model we used --it shows price adjustments toward equilibrium from new information.

\[
\Delta P_t = \mu + \alpha \beta' P_{t-1} + \Gamma_1 \Delta P_{t-1} + \Gamma_2 \Delta P_{t-2} + \ldots + \epsilon_t
\]

where:

\[
P_{t-1} = \begin{bmatrix} P_{f,t-1} \\ P_{s,t-1} \end{bmatrix}
\]

is a 2 by 1 vector that contains futures and cash prices for day t-1. Day t is the current day.

\[
\Delta P_{t-i} = \begin{bmatrix} \Delta P_{f,t-i} \\ \Delta P_{s,t-i} \end{bmatrix}
\]

is a 2 by 1 vector that contains price changes
from the previous day, \( i = 0, 1, 2, \ldots \)

\[ \mu = \begin{bmatrix} \mu_f \\ \mu_s \end{bmatrix} \] is a 2 by 1 vector that contains constant terms.

\[ \beta' = \begin{bmatrix} \beta_f & -\beta_s & \delta \end{bmatrix} \] is a 1 by 3 cointegration vector that contains the long run relationship between futures and cash prices. The two prices are in long run equilibrium with each other when:

\[ \beta' P_{t-1} = \begin{bmatrix} \beta_f & -\beta_s & \delta \end{bmatrix} \begin{bmatrix} p_{f,t-1} \\ p_{s,t-1} \\ 1 \end{bmatrix} = 0 \]

We include a constant term, \( \delta \), in the cointegration vector to account for location differences between futures and cash prices.  

\[ \alpha = \begin{bmatrix} \alpha_f \\ \alpha_s \end{bmatrix} \] is a 2 by 1 vector that contains the price responses to departures from long run equilibrium. \( \alpha_f \) and \( \alpha_s \) are called speed of adjustment coefficients.

\[ \Gamma_i = \begin{bmatrix} \Gamma_{f,f} & \Gamma_{f,s} \\ \Gamma_{s,f} & \Gamma_{s,s} \end{bmatrix} \] is a 2 by 2 matrix containing the short run price adjustments to previous price changes for day \( t - i \).

\[ \epsilon_1 = \begin{bmatrix} \epsilon_{f,f} & 0 \\ 0 & \epsilon_{s,s} \end{bmatrix} \] is a 2 by 2 matrix containing the current price responses to new information. \( \epsilon_{f,t} \) and \( \epsilon_{s,t} \) are highly correlated when using prices at the daily frequency making it difficult to assign the portion of price discovered in an individual market. Our estimated correlations between, \( \epsilon_{f,t} \) and \( \epsilon_{s,t} \) were larger than 0.90.

Equation 2 is the moving average model that corresponds to (is the dual of) the error correction model in equation 1.\(^2\)

\(^1\) Zivot shows the relationship between the constant term in the cointegration relationship, \( \delta \), and the constant term, \( \mu \), in equation 1 (Zivot, 2000, p. 793).
\( \Delta P_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \ldots = \mu + \psi(L) \varepsilon_t \)

\( \psi_i = \begin{bmatrix} \psi_{f,f,i} & \psi_{f,s,i} \\ \psi_{s,f,i} & \psi_{s,s,i} \end{bmatrix} \) is a 2 by 2 matrix containing the moving average coefficients for day \( t-i \). The moving average coefficients are price responses in the current period, \( t \), to unit errors for day \( t-i, i = 1, 2, \ldots \). They were estimated in this paper by calculating the impulse responses to unit values for \( \varepsilon_{f,t} \) and \( \varepsilon_{s,t} \) in the error correction model (equation 1). A detailed description of estimating the moving average coefficients is provided in the analysis and findings section. \( L \) is the lag operator for day \( t-i, i = 0, 1, 2, \ldots \).

Equation 3 is the first step in using the results from the error correction and moving average models to decompose price into permanent and temporary parts. \(^3\)

\( \psi(L) = \psi(1) + (1-L)\psi^*(L) \)

where: \( \psi(1) = \begin{bmatrix} \psi_{f,f} & \psi_{f,s} \\ \psi_{s,f} & \psi_{s,s} \end{bmatrix} = I_n + \psi_1 + \psi_2 + \ldots \) contains the sums of the moving average coefficients for periods \( t = 0, 1, 2 \ldots \) and are the permanent effects of 1 unit errors on price. \( I_n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \) is 2 by 2 identity matrix.

The factors \( \psi(1) + (1-L)\psi^*(L) \) are substituted for \( \psi(L) \) in equation 2 and are shown in equation 4.

\( \Delta P_t = \mu + [\psi(1) + (1-L)\psi^*(L)] \varepsilon_t \)

\( \Delta \) is substituted for \( (1-L) \) in equation 4 and shown in equation 5.

\( \Delta P_t = \mu + \psi(1) \varepsilon_t + \psi^*(L) \Delta \varepsilon_t \)

The first term is the permanent effect on price from one period to the next due to the deterministic trend. The second term is the permanent effect of the error, \( \varepsilon_t \), term on price. This part of the permanent effect is the stochastic trend. The third term is the transitory part of price.

Equation 5 is converted from price changes to levels and shown in equation 6.

\( P_t = P_0 + \mu t + \psi(1)(\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_t) + \psi^*(L) \varepsilon_t \)

---

\(^2\) Engle and Granger proved that cointegrated variables could be specified as a vector autoregressive model, an error correction model, or a vector moving average model.

\(^3\) Mills shows the algebraic steps in deriving the right hand side of equation 3 (p. 61).
Equation 6 is the multivariate Beveridge-Nelson decomposition. $P_0$ is the price level in day zero. The error, $\varepsilon_{o}$, for period zero is assumed to equal zero.\footnote{Rossi and Zivot provide detailed derivations of the univariate and multivariate Beveridge-Nelson decomposition. Evans and Reichlin derived the multivariate Beveridge-Nelson decomposition.}

The first three terms in equation 6 are used to estimate the permanent part of price for each period, $t$. $\varepsilon_1$ is used for period $t=1$, $\varepsilon_1 + \varepsilon_2$ for period $t=2$, $\varepsilon_1 + \varepsilon_2 + \varepsilon_3$ for period $t=3$, etc. The transitory part of price for each period $t$, the last term in equation 6, is calculated as a residual.

**Data**

We examine daily soybean spot and futures prices from September 1, 1982 through August 31, 2006. The data include the 1982 through 2005 crop years. The spot prices are central Illinois prices and the futures prices are the closing prices on the Chicago Board of Trade. Futures prices were taken from the closest to maturity contract (the nearby futures contact) excluding the delivery month.

Figure 1 shows the standard deviation of the daily futures and spot soybean price changes by crop year. The two prices are very similar. The daily standard deviation over all the crop years is 8.8 cents per bushel for the futures price and 9.0 for the spot price. The two standard deviations have a 0.98 correlation. The average futures price over all the crop years is $6.12 per bushel for the futures price and $6.01 for the spot price.

The price standard deviations tend to peak above 10 cents a bushel for two years and then return to more average levels. Seven of the 25 crop years have daily price standard deviations above 10 cents bushel. Three of them are followed by another crop year with price standard deviations above 10 cents per bushel then the standard deviation of price returns to 8 cents per bushel or less.

The data were separated into three periods for our price discovery examination. The periods were crop years (I) 1982 through 1989, (II) 1990 through 1999, and (III) 2000 through 2005. There was little or no commodity fund trading in Period (I), moderate trading in period (II), and heaviest trading in period (III).

Daily price standard deviations were 8.0 and 7.8 cents per bushel for futures and spot prices in period II. The corresponding numbers were 9.4 and 9.8 cents per bushel for period I and 9.4 and 9.91 cents per bushel for period III. A larger daily price standard deviation does not necessarily result in a larger transitory portion of price. The transitory percentage of price would decrease if the permanent percentage of price increased due to improved price discovery performance.

**Analysis and Findings**

Our analysis followed the procedures discussed in the section on the Beveridge-Nelson decomposition. However, before the multivariate Beveridge-Nelson procedure can be applied the futures and spot prices and price changes must be examined for unit roots (nonstationary).
Prices (price levels) must be nonstationary. Price differences must be stationary. The futures and spot prices must also be examined for cointegration. Cointegration is also a requirement for using the multivariate Beveridge-Nelson decomposition.

We used the method proposed and used by Yang, Bessler, and Leatham to account for the interest cost of storage, a major storage cost component. Spot and futures prices will not be cointegrated if storage cost is also nonstationary. They estimated a cash (spot) equivalent futures price by deducting interest cost from the futures price and converted their futures and spot prices to natural logs. We used their method.

First, we examined the price changes for nonstationary using the augmented Dickey-Fuller tests. The test for nonstationary was done on combined data for all three periods –crop years 1982 through 2005. The null hypothesis of unit roots in the price changes from one day to the next was rejected at the 1 percent significance level for the soybean spot price, futures price, and spot equivalent futures price. The test was run with and without a constant term in the augmented Dickey-Fuller equation and with both a constant term and linear trend included. We conclude that the price changes are stationary.

Second, we repeated the procedure for price levels. We could not reject the null hypothesis of a unit root at the 5 percent level when a constant term was not included and when both a constant term and a linear trend term were included in the augmented Dickey-Fuller equation. However, we did reject the null hypothesis at the 5 percent level when a constant term was included. Although, the results are mixed we chose to accept the null hypothesis that the spot, futures price, cash equivalent futures price are nonstationary in levels.

Nonstationary and price discovery are closely related in that the permanent part of price is nonstationary. Nonstationary results from the accumulation of the permanent part of the errors as shown in equation 6. Part of the influence of each error, as shown by the second term in equation 6, does not dissipate. For example, part of the error, $e_1$, in equation 6 influences the price levels in all following periods. Since the permanent effects of errors do not dissipate the two prices do not revert to a mean –a condition required for stationary.

The Johansen trace and maximum eigenvalue tests for cointegration were done for the spot equivalent futures and spot prices for each of the three periods. With two variables the test results suggest that there is either 1 cointegration relationship or none. Both tests for each of the three periods indicated 1 cointegration vector at the 5 percent significance level. The tests were based on the assumption of a constant term in the cointegration vector and a constant term in the error correction equation.

---

5 Price discovery analysis can be done using the univariate Beveridge-Nelson decomposition if nonstationary is found but cointegration is not.
The estimated cointegrating relationships, $\beta$, and the speed of adjustment coefficients, $\alpha$, to departures from long run equilibrium shown below were estimated using the Johansen method.  

$$
\begin{bmatrix}
\beta_f \\
\beta_s \\
\delta
\end{bmatrix} = 
\begin{bmatrix}
1 \\
-1 \\
-0.01
\end{bmatrix} 
\begin{bmatrix}
1 \\
-0.93 \\
-0.46
\end{bmatrix} 
\begin{bmatrix}
1 \\
-1 \\
-0.02
\end{bmatrix}
$$

$$
\begin{bmatrix}
\alpha_f \\
\alpha_s
\end{bmatrix} = 
\begin{bmatrix}
-0.02 \\
0.08
\end{bmatrix} 
\begin{bmatrix}
-0.02 \\
0.03
\end{bmatrix} 
\begin{bmatrix}
-0.01 \\
0.06
\end{bmatrix}
$$

We tested the hypothesis, that the futures price is an unbiased predictor of the spot price except for location differences. This hypothesis is tested by determining if the coefficient for the spot price, $\beta_s$, differs from -1. We did not reject the unbiased predictor hypothesis at the 5 percent level for periods I and III, crop years 82 through 1989 and crop years 2000 through 2005, respectively. We did reject it for period II crop years 1990 through 1999.

The speed of adjustment coefficients, $\alpha_s$ are significantly different from zero for the spot price but not so for the spot equivalent futures price. This result suggest that price is discovered in the futures market and passed to the spot market via the the long term equilibrium relationship. Incorporating information more slowly increases transitory to total price variability.

We used the error correction model to estimate the moving average coefficients in equation 2. The procedure involves separately setting the errors, $\epsilon_{ft}$ and $\epsilon_{st}$ equal to 1 and calculating $\Delta P_t$ for $t = 1, 2, \ldots$ Errors in period 2 and following periods are set equal to zero. Each $\Delta P_t$ is the moving average coefficients for period $t$, $\psi_t$. The procedure is continued until the sum of the moving average coefficients, $\psi(1)$, converge to a limit. Since the price changes and the

---

6 * ** *** indicate significant difference at the 10, 5, and 1 percent levels. The test for $\beta_s$ is for difference from -1.

The test for $\delta$ is for difference from zero. $\beta_f$ is normalized to equal one.
The estimated long run permanent effects of a one-unit and one-time change in each error term, \( \varepsilon_{f,t} \) and \( \varepsilon_{s,t} \), in the error correction model are shown below.

<table>
<thead>
<tr>
<th>Period I</th>
<th>Period II</th>
<th>Period III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crop years</td>
<td>1982-89</td>
<td>1990-99</td>
</tr>
</tbody>
</table>

\[
\psi(1) = \begin{bmatrix}
1.63 & -0.52 \\
1.63 & -0.52 \\
\end{bmatrix} \begin{bmatrix}
0.62 & 0.43 \\
0.62 & 0.43 \\
\end{bmatrix} \begin{bmatrix}
1.10 & -0.12 \\
1.10 & -0.12 \\
\end{bmatrix}
\]

The elements in each row sum to about one suggesting that the permanent net effects of the two error terms, \( \varepsilon_{f,t} \) and \( \varepsilon_{s,t} \), on permanent price may are close to one. The rows in \( \psi(1) \) are equal for each period because there is one less cointegrating relationship than the number of variables. There are 2 variables in our analysis and 1 cointegration relationship for each period. This outcome is expressed as \( n – r = 1 \) in the price discovery literature where \( n \) is the number of variables and \( r \) is the number of cointegrating relationships (Hasbrouck, Haug).

The estimated long run permanent effects, the estimated errors from the error correction and moving average models, and the estimated constant term for each period were used in equation 6 to estimate permanent and transitory prices. We estimated the ratio of the standard deviation of transitory price and actual price for each period. The results are shown in figure 2.

The ratio of the standard deviation of transitory to actual price for soybean futures declined over the three periods from 0.28 to 0.21 and then to 0.07. These results indicate increasing price discovery performance for soybean futures. They do not support the hypothesis that commodity fund trading in soybeans has decreased price discovery performance.

The ratio of the standard deviation of transitory to actual price for the soybean spot price declined from period I to period II from 0.63 to 0.30 and increased to 0.62 in period III. The larger ratio for spot price each period supports the earlier result that a portion of price is first discovered in the futures market then passed to the spot market via the cointegrating relationship.

---

7 The cointegrating relationship \( \beta' P_{t-1} = \begin{bmatrix} \beta_f & -\beta_s & \delta \end{bmatrix} \begin{bmatrix} p_{f,t-1} \\ p_{s,t-1} \\ 1 \end{bmatrix} = e_t \) and \( e_t \) is a stationary variable.
The large increase in the ratio from period II to III may have been due to Hurricane Katrina that disrupted the basis between futures and spot price.

Summary

We used the multivariate Beveridge-Nelson decomposition to examine the influence of commodity fund trading on price discovery in the soybean futures and spot markets. We found that the price discovery performance of the soybean futures market was at its best during the recent years when commodity fund trading was largest. We found evidence that a portion of price is first discovered in the soybean futures market and then passed to the soybean spot market. Unlike for the soybean futures price we estimated that the price performance of the spot soybean price was not at its best during recent years. This may have been due to the transportation difficulties from Hurricane Katrina that disrupted soybean exports and the link between futures and spot prices.

The procedures used in this study can be used to supplement studies that use daily Commitment of Trader data to examine whether or not the different types of traders contribute to price discovery. These studies did not estimate the overall price discover performance as this study does. Updating the Irwin and Holt study using recent commitment of trader data for the agricultural futures markets would address the concern about the effects of the different types of traders on price discovery in these markets. In addition, using the procedures used in this study to examine the futures and spot prices in the agricultural markets would address the concern that commodity funds may be degrading the overall price performance of the agricultural markets.

Our procedures can be modified to examine price discovery in the livestock futures and spot markets although the two prices are not held together by the cost of storage. Cointegration is rejected for these markets. For these markets the univariate Beveridge-Nelson decomposition would be used. The parameters for the decomposition would be taken from estimated auto regressive intergraded moving average (ARIMA) models of each price.
References


Rossi, Eduardo. Cointegration, University of Pavia, Pavia, Italy http://economia.unipv.it/pagp/pagine_personali/erossi/cointegration.pdf


Figure 2. Soybean Price: Percent Transitory by crop year