Three essays on agriculture economics

by

Chao Li

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Dermot Hayes, Major Professor
Sergio Lence
Chad Hart
Keri Jacobs
Cindy Yu

Iowa State University

Ames, Iowa

2016

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

ACKNOWLEDGEMENTS

I hereby sincerely acknowledge my major professor, Dr. Dermot Hayes, for his wise guidance and unlimited supply of time and patience throughout the course of this research. His insights and words of encouragement always inspired me to complete my research. I would also like to thank my committee members Dr. Sergio Lence, Dr. Keri Jacbos, Dr. Chad Hart and Dr. Cindy Yu, for their constructive comments to help me complete this dissertation.

Finally, I would like to give my special thanks to my family for their understanding and love.

ABSTRACT

The main subjective of this dissertation is to analyze three issues of current interest in agricultural economics. Chapter 1 investigates the lead-lag relationships among soybean prices in U.S., Brazilian, and Chinese futures markets by using threshold co-integration methodologies. The empirical results indicate the influence of U.S. market in the long-term, and also show that overnight return of U.S. soybean futures and the daytime return of Chinese No. 1 soybean futures contemporaneously affect each other in the short-term. A weak temporal seasonal causality between U.S. and Brazilian soybean futures prices exists. Chapter 2 examines the impact of feedstock supply mechanisms under conditions of spatial monopoly on the supply of cellulose to the plants. The model shows that, in the absence of competition, the processor is indifferent between processor collection and supplier delivery, but that societal welfare is higher under supplier delivery. By using a repeated Nash Equilibrium, this paper shows that processor collection is first best for both incumbent and entrant. By comparing the slope of marginal cost curve for this monopsonistic processor with the slope of cost curve across other feedstocks, substantial quantities of other feedstocks may be required to meet the mandate. Chapter 3 investigates a change in the market power of the U.S. nitrogen fertilizer industry by examining the causal linkage between fertilizer, its main feedstock (natural gas), and output (corn) by using a Bayesian-based Kalman filter algorithm. The results of the time-varying estimation show that the U.S. nitrogen fertilizer price follows the value of its marginal productivity closer than its marginal cost of production, indicating a less competitive market structure. The estimation from the error correction model supports these results.

CHAPTER 1. PRICE DISCOVERY ON INTERNATIONAL SOYBEAN FUTURES MARKETS: A THRESHOLD CO-INTEGRATION APPROACH

Abstract

This paper investigates the lead-lag relationships among soybean prices in U.S., Brazilian, and Chinese futures markets. We focus on both long-run price co-movements and on short-run price relationships. Various co-integration methodologies and causality tests are applied to examine the changes in price relationships over time. The empirical results indicate the following: (*a*) the soybean futures market in the U.S. is still the most important and influential market, and the U.S. price, in the long-term, leads price changes in Brazil and China; (*b*) in the short-term, the overnight return of U.S. soybean futures and the daytime return of Chinese No. 1 soybean futures contemporaneously affect each other, but there is no significant causality between U.S. overnight return and the daytime return of Chinese No. 2 soybean futures; and, (*c*) a weak temporal seasonal causality between U.S. and Brazilian soybean futures price exists and more often than not Brazilian futures lead U.S. futures during the Brazilian growing season.

1.1 Introduction

The U.S., Brazil, and Argentina account for over 90% of the world's soybean exports. China, which imported 71.4 million tons of soybeans in 2014, is by far the largest importer and gets approximately 50% of its soybeans from the U.S. and 40% from Brazil.

The U.S., Brazil, and China all have active soybean futures. China has two different markets, one for non-GMO soybeans and the other for imported GMO soybeans. In the Dalian Commodity Exchange (DCE), the No. 1 contact is for non-GMO soybeans that are used for human consumption, and the No. 2 contract allows delivery of imported GMO soybean crops, which are used for soy oil and animal feed. Figure 1.1 shows soybean futures prices in all four markets. There is visual evidence of strong co-movements among these prices, and we investigate whether this co-movement is due to a stable long-run price relationship and examine the price lead-lag relationship across the four markets.

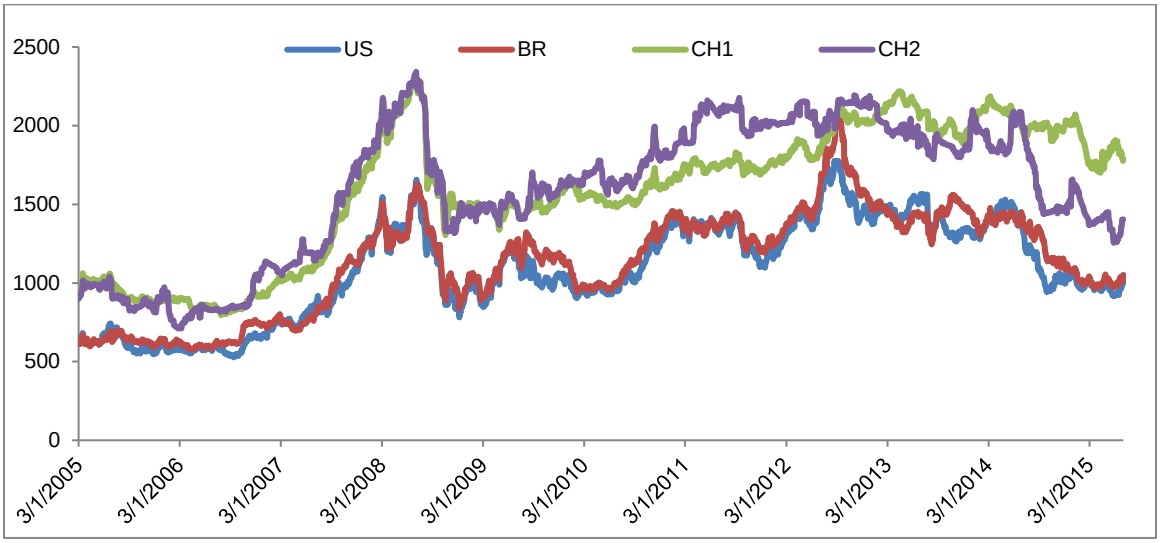


Figure 1.1: Soybean futures prices in the US, Brazil and China from 2005 to 2015

This paper is the first to investigate the long-run lead-lag relationship among the U.S., Brazilian, and Chinese markets, and explores the seasonal relationship between U.S. and Brazilian futures markets and the influence the Globex overnight trading platform in the U.S.

has on the daytime return of soybean futures in China.[1] Overnight trading in the U.S. and daytime trading in China occur contemporaneously, so we apply an autoregressive distributed-lag model to address this problem.

The paper is organized as follows: Section 1.1.1 provides a brief literature review. Section 1.2 describes the methodology that characterizes the price lead-lag relationships using linear and non-linear co-integration. Section 1.3 describes the data. Section 1.4 exhibits and explains empirical results of co-integration and demonstrates two short-run causality relationships. Section 1.5 presents conclusions.

1.1.1 Previous Theoretical Work

Granger (1981) introduced the most widely used methodology to study long-run price causality co-integration. He showed that two variables may have a long-run equilibrium relationship even if they are non-stationary. Engle and Granger (1987) extended this concept and showed that co-integrated variables can be represented by a vector error correction model (VECM) and provided test methodology for this framework. Balke and Fomby (1997) introduced the threshold concept to explain possible non-linear long-run equilibrium relationships. Hansen and Seo (2002) and Seo (2006) provided two methods to test a threshold and a method to estimate the parameters of a threshold vector error correction model (TVECM).

Wahab and Lashgari (1993) and Ghosh (1993) investigated the forecasting power of the S&P 500 index spot and futures prices changes using co-integration. Their results indicated a

[1] Globex is the electronic trading system in the U.S.

stable long-run equilibrium relationship between the index and its futures price. Chu et al. (1999) investigated the price discovery function in three S&P 500 index markets: the spot index, index futures and S&P Depositary Receipts markets. They found that the three price series are a co-integrated system with one long-run stochastic trend and the futures market serves the dominant price discovery function when the common stochastic trend is decomposed. Martens et al. (1998) applied a threshold error correction model to study index-futures arbitrage and found that the impact of futures market on the spot market is larger when the mispricing error is negative and that the impact of the mispricing error increases with the magnitude of that error.

Booth, Brockman, and Tse (1998) investigated the relationship between U.S. and Canadian wheat futures prices and showed that both of them are integrated of order one and that they are co-integrated. Fung et al. (2013) used daily data for 16 commodity futures contracts traded in China and the corresponding foreign markets to analyze price linkages among markets. They also studied the impact of Chinese futures daytime returns on the U.S. overnight returns using a regression of one return on the other; however, their study ignored the effect of the previous return on the current return. For soybeans, their results showed that the price causality between U.S. soybean futures contracts and both Chinese No. 1 and No. 2 contracts are statistically significant but not economically significant. Peri and Baldi (2010) employed the threshold co-integration approach to analyze the long-run relationship between vegetable oil prices and conventional diesel prices in the EU and suggested a two-regime threshold co-integration relationship for the rapeseed oil and diesel price pair. Natanelov et al. (2011) examined price linkages between crude oil futures and a series of agricultural

commodities futures using the VECM method and showed that co-movement of commodity prices is a temporal seasonal concept and should be treated accordingly.

In the literature regarding soybean futures price discovery between U.S., Brazilian, and Chinese markets, Han et al. (2013) examined the role that the DCE plays in the global discovery of soybean futures. They used a structural vector autoregressive model (SVAR) and VECM on the returns of the DCE and the CBOT soybean futures during trading and non-trading hours. The results indicate a bi-directional causality between the two markets with the CBOT leading the DCE. Han et al. did not include the Chinese No. 2 soybean futures contract due to a liquidity problem. Our paper employs more recent data to capture the role of Chinese No. 2 contracts and focuses on threshold co-integration analysis. Han et al. used a different approach than the one described below when rolling the price data when the nearby contract expires. This may help explain the difference between the results presented in this paper and those in Han et al. (2013). Christofoletti et al. (2012) examined the price linkage between soybean futures contracts in China, U.S., Brazil, and Argentina using VECM. The results indicated that the U.S. price has a dominant role. Liu et al. (2015) used a generalized autoregressive conditional heteroskedasticity (GARCH) model based on generalized error distribution (GED) and exponential GARCH-GED models and found that the spillover from CBOT soybean futures to DCE No. 1 soybean futures has weakened through time, indicating a more influential Chinese soybean market. Merener (2015) investigated how local supply shocks in the globally distributed production of commodities are incorporated into CME futures prices and found that CME soybean futures prices have become increasingly sensitive to supply shocks outside of the United States.

1.2 Theoretical Issues

1.2.1 Co-Integration and the Vector Error Correction Model

In time-series econometrics, a price series that has a stationary, invertible, ARMA representation after differencing d times, is said to be integrated of order d, denoted by $p_t \sim I(d)$. If both series x_t and y_t are $I(d)$ processes, their linear combination $\varepsilon_t = x_t - \beta y_t$ is also an $I(d)$ process. However, if there exists a vector, $[1, -\beta]$, such that $\varepsilon_t \sim I(d-b)$, where $b > 0$, then these two series are said to be co-integrated and the vector $[1, -\beta]$ is called the co-integrating vector. In appendices A1-A3, two types of unit root tests indicate that all futures prices studied in this paper can be characterized as $I(1)$ processes. Therefore, we concentrate our study on the case when $d = b = 1$. As a result, the co-integrated system can be simply characterized as a VECM

$$\Delta x_t = \mu + \Gamma x_t + \sum_{i=1}^{k} A_k \Delta x_{t-k} + v_t \tag{1.1}$$

where x_t is a $n \times 1$ vector of $I(1)$ processes, μ is a $n \times 1$ vector of constant, Γ and A_k are $n \times n$ coefficient matrices, and v_t is a $n \times 1$ vector of Gaussian white noise processes. Johansen (1988; 1991) demonstrated that the rank of matrix Γ represented the number of co-integration relationships in vector x_t. Thus, Johansen's co-integration test estimates matrix Γ through an unrestricted VAR and tests possible rejection of the restriction implied by the reduced rank of Γ. There are two test statistics, one using the trace and the other using the maximum eigenvalue, and inferences can be different.

The null hypothesis for the trace test is that the number of co-integrating vectors is less than or equal to r. The test statistic is given by

$$\lambda_{trace} = -T \sum_{i=r+1}^{n} \ln(1-\hat{\lambda}_i^2)$$

where T is the sample size actually used for estimation and $\widehat{\lambda}_i$ is the estimated values of the ordered eigenvalues from the estimated matrix. For the maximum eigenvalue test, the test statistic is given by

$$\lambda_{\max} = -T \ln(1-\hat{\lambda}_{r+1})$$

which tests the null hypothesis that the number of co-integrating vectors is exactly r against the alternative of $r+1$ co-integrating vectors.

When there is a co-integration relationship between time series, Granger causality can be tested by a Wald test. Specifically, a linear VECM of order $r+1$ can be compactly represented as

$$\Delta x_t = A' X_{t-1}(\beta) + v_t \tag{1.2}$$

with

$$X_{t-1}(\beta) = \begin{pmatrix} 1 \\ w_{t-1}(\beta) \\ \Delta x_{t-1} \\ \Delta x_{t-2} \\ \vdots \\ \Delta x_{t-r} \end{pmatrix}$$

where x_t is a n-dimensional $I(1)$ time series, which is co-integrated with one $n\times 1$ co-integrating vector β, $w_t(\beta)=\beta' x_{t-1}$ is the error correction term (ECT), and $A'=\begin{pmatrix}\mu & \alpha & A_1 & A_2 & \dots & A_r\end{pmatrix}$ is a $n\times(nr+2)$ matrix of coefficients.

Thus, the bi-variate co-integrated time series can be written as

$$\begin{bmatrix}\Delta x_{1,t}\\ \Delta x_{2,t}\end{bmatrix}=\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}+\begin{bmatrix}\alpha_1\\ \alpha_2\end{bmatrix}w_{t-1}(\beta)+\sum_{i=1}^{r}A_i\begin{bmatrix}\Delta x_{1,t-i}\\ \Delta x_{2,t-i}\end{bmatrix}+\begin{bmatrix}v_{1,t}\\ v_{2,t}\end{bmatrix} \tag{1.3}$$

where $w_{t-1}(\beta)=x_{1,t-1}-\beta x_{2,t-1}$ determines the ECT. The optimal length of lag r is determined by Akaike Information Criterion (AIC) or Schwarz Information Criterion (SIC).

Intuitively, parameter α_i measures the long-run causality relationship and parameter β characterizes the long-run equilibrium between these two series. By testing the null hypothesis of $\alpha_i=0$ against the alternative of $\alpha_i\neq 0$, three different results may be obtained: (*a*) $\alpha_1=0$ and $\alpha_2=0$; (*b*) $\alpha_1\neq 0$ and $\alpha_2\neq 0$; and (*c*) $\alpha_i\neq 0$ but $\alpha_j=0$. The first case indicates no co-integration exists, the second case indicates bi-directional long-run causality, and the last case indicates a unidirectional long-run causality relationship.

1.2.2 Threshold Co-Integration

The above traditional VECM assumes the adjustment process to the long-run equilibrium is continuous and linear. In reality, the influence of transaction costs, adjustment costs, or other market frictions makes it likely that movement toward the long-run equilibrium may only occur when the deviation from equilibrium exceeds a critical threshold level. It is also possible that the speed at which the system returns to long-run equilibrium differs under regimes. Balke and

Fomby (1997) introduced the concept of threshold co-integration to analyze this type of discrete adjustment process.

As an extension of model (1.2), we propose the following specification of a two-regime threshold co-integration model:

$$\Delta x_t = \begin{cases} A_1^{'} X_{t-1}(\beta) + v_t, \text{ if } |w_{t-1}(\beta)| \le \gamma \\ A_2^{'} X_{t-1}(\beta) + v_t, \text{ if } |w_{t-1}(\beta)| > \gamma \end{cases} \tag{1.4}$$

where one regime is close to the equilibrium regardless of the sign of ECT and the other regime is far from the equilibrium, and γ is the threshold parameter. Equation (4) can be rewritten as

$$\Delta x_t = A_1^{'} X_{t-1}(\beta) d_{1t}(\beta, \gamma) + A_2^{'} X_{t-1}(\beta) d_{2t}(\beta, \gamma) + v_t \tag{1.5}$$

with $d_{1t}(\beta, \gamma) = 1(|w_{t-1}| \le \gamma)$, $d_{2t}(\beta, \gamma) = 1(|w_{t-1}| > \gamma)$. The coefficient matrices A_1 and A_2 determine the dynamics in each regime. The advantage of this specification captures the idea that adjustment speed to long-run equilibrium would be lower when the deviation is banded. With the absolute value of w_{t-1} above or below the critical threshold value, this TVECM model allows all coefficients, except the co-integrating vector, β, to switch between two regimes. The estimated coefficients of w_{t-1} measure different adjustment speeds of price moving back towards the long-run equilibrium.

The hypothesis test for threshold co-integration involves four different cases: (*a*) co-integration with no threshold effect; (*b*) co-integration with a threshold effect; (*c*) no co-integration with a threshold effect; (*d*) no co-integration and no threshold effect. Thus, given the bi-variate Johansen co-integration test results, we apply two kinds of threshold tests, developed by Hansen and Seo (2002) and Seo (2006).

If two series have shown a co-integration relationship using Johansen's co-integration test, we further determine whether or not this co-integration is linear using the maximum likelihood method developed by Hansen and Seo (2002), which involves a joint grid search over the co-integrating vector β in the region $[\beta_L, \beta_U]$ and a threshold parameter γ in the region $[\gamma_L, \gamma_U]$. In our empirical applications, we set the number of grid searches for both threshold parameter and co-integrating vector at 300.

Since the threshold effect is only valid when $0 < P(|w_{t-1}| \leq \gamma) < 1$[2], it is imposed by assuming that

$$\pi_0 \leq P(|w_{t-1}| \leq \gamma) \leq 1 - \pi_0$$

where $\pi_0 > 0$ is a trimming parameter and set equal to 0.05.

Hanson and Seo (2002) test the null hypothesis of linear co-integration (no threshold effect) against the alternative hypothesis of threshold co-integration by developing two SupLM tests for a given or estimated β using a parametric bootstrap method to calculate asymptotic p-value. The first test is used when the true co-integrating vector β_0 is known to be a priori and the test statistic is denoted as

$$\text{SupLM}^0 = \sup_{\gamma \in [\gamma_L, \gamma_U]} LM(\beta_0, \gamma)$$

The second test is used when the true co-integrating vector is unknown and the test statistic is denoted as

[2] Otherwise, the model reduces to a linear co-integration model.

$$\text{SupLM}^0 = \sup_{\gamma \in [\gamma_L, \gamma_U]} LM(\widetilde{\beta}, \gamma)$$

where $\widetilde{\beta}$ is the null estimate of the co-integrating vector.

If two series fail to show a co-integration relationship by Johansen's co-integration test, we alternatively apply a supreme test developed by Seo (2006). This tests the null hypothesis of no co-integration against the alternative hypothesis of threshold co-integration using a Band-TVECM,

$$\Phi(L)\Delta x_t = \mu + \alpha_1 w_{t-1} 1(|\, w_{t-1} \,| \leq \gamma) + \alpha_2 w_{t-1} 1(|\, w_{t-1} \,| > \gamma) + v_t$$

where $t = 1, ..., n$, and $\Phi(L)$ is a qth-order polynomial in the lag operator defined as $\Phi(L) = I - \Phi_1 L^1 - \cdots - \Phi_q L^q$. When threshold parameter γ is fixed, the least-squares estimators for the coefficients are the OLS estimators. Thus, equation (1.5) can be specified as

$$\begin{aligned}\Delta x_t = {} & \mu(\gamma) + \alpha_1(\gamma) w_{t-1}(\beta) d_{1t}(\beta, \gamma) + \alpha_2(\gamma) w_{t-1}(\beta) d_{2t}(\beta, \gamma) \\ & + \Phi_1(\gamma)\Delta x_{t-1} + \cdots + \Phi_q(\gamma)\Delta x_{t-q} + v_t(\gamma)\end{aligned}$$

and the supreme Wald test statistic is defined as

$$\text{Sup}W = \sup_{\gamma \in [\gamma_L, \gamma_U]} W_n(\gamma)$$

where W_n is the Wald statistic from testing the null of no co-integration with a fixed threshold parameter γ.

1.2.3 Autoregressive Distributed-Lag Model

To study the temporal causal effect of soybean futures prices among different markets, we employ an autoregressive distributed lag model (ARDL). Specifically, the autoregressive distributed lag model of order p and q, ARDL(p, q), defined as follows:

$$y_t = c + \sum_{k=1}^{p} a_k y_{t-k} + \sum_{k=0}^{q} b_k x_{t-k} + \varepsilon_t$$

where y_t and x_t are stationary variables, and ε_t is white noise. More strictly, we assume that ε_t is stationary and independent of $x_t, x_{t-1}, \cdots$ and $y_t, y_{t-1}, \cdots$, so that this ARDL model can be estimated consistently using ordinary least squares. The estimated contemporaneous parameter coefficient b_0 is the impact multiplier that characterizes the temporal price relationships.

1.3 Data Description and Timelines

Our empirical analysis uses daily nominal prices of soybean futures contracts traded in the Chicago Mercantile Exchange (CME), the Dalian Commodity Exchange (DCE) and the Brazilian Mercantile and Futures Exchange (BM&F). All data is collected from a Bloomberg terminal, and the date range is from 03/01/2005 to 06/30/2015. A close-to-maturity method is employed to rollover data across contracts and all data is proportionally modified to eliminate the price jump across contracts.

Owing to differences in national holidays, data in all three markets are not automatically matched. We have eliminated mismatched data and the whole sample size is reduced to 2377 observations. We standardize the price quotation unit and convert all prices into the natural log of prices measured in U.S. cents per bushel.

The Chinese futures market is open from 9:00 a.m. to 11:30 a.m. and from 1:30 p.m. to 3:00 p.m. The trading hours in Brazil are from 9:00 a.m. to 3:15 p.m. The trading floor in the U.S. operates from 8:30 a.m. to 1:15 p.m., and the Globex overnight trading runs from 7:00 p.m. to 7:45 a.m. Figure 1.2 illustrates the timeline of these trading hours.

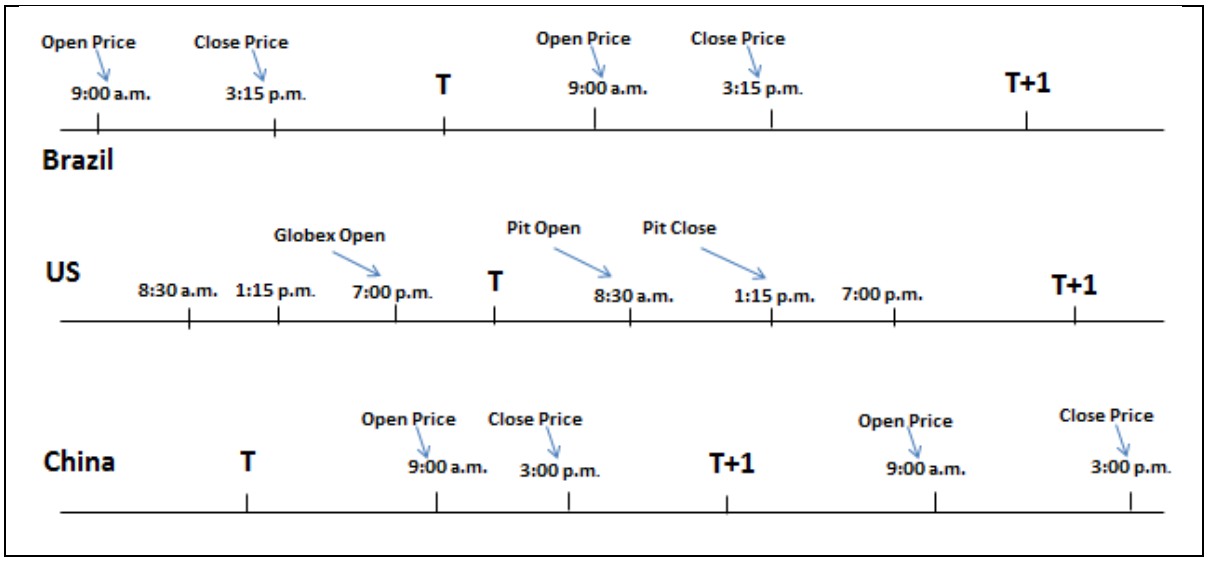


Figure 1.2: Timeline of the U.S., Brazilian, and Chinese futures markets

China is 13 hours ahead of the U.S. during the U.S. daylight saving period and 14 hours ahead during standard time. During daylight saving time, the Chinese market and U.S. Globex open at the same time, and the Chinese market closes 6.75 hours earlier than the Globex. During standard time, the Chinese market opens one hour later than U.S. and closes 5.75 hours earlier than the Globex. In other words, the Globex is always open when the Chinese market is open. Therefore, we do not need to adjust the data for this time change. The difference in daylight periods between the U.S. and Brazil does not impact the close-to-close return in each trading day in either market. Therefore, we do not need to account for the effect of daylight saving time in the empirical work.

In the following empirical applications, we focus on the close-to-close returns of soybean futures in each market, the open-to-close (daytime) returns of soybeans futures in the Chinese market, and the overnight return of soybean futures in the U.S. market. Detailed variable notations are provided in the appendix.

1.4 Empirical Results

1.4.1 Long-Run Lead-Lag Relationship

Here, we present results pertaining to the long-run causality relationship between soybean futures prices in the three examined markets. Given the stationary test results in appendices, all closing prices can be regarded as an $I(1)$ process in every sub-period. Thus, we apply the Johansen co-integration test to investigate whether there is a long-run linear relationship between the closing prices in each market. As shown in figure 1.1, the co-movements among the price series are quite strong from the beginning of the data period, but the relationship weakens after several years, suggesting a structural break in the relationship.

The traditional approach to test structural change would be picking an arbitrary sample breakpoint, often the midpoint of the sample, and using Chow's (1960) F-test. The result using this approach is very sensitive to the prior choice of break dates, and Hansen (2001) suggests that the Quandt-Likelihood Ratio (QLR) test is superior for detecting structural change with unknown timing. In our analysis, we are interested in whether the U.S. soybean futures market is still a world price leader, and we concentrate on the bi-variate causality relationship between soybean futures prices in the U.S. and elsewhere. Therefore, we employ this QLR structural change test and divide every pair-wise data sample into two sub-periods based on the test results. The results suggest that in July 2012 a structural break between soybean futures prices in the U.S. and Brazil occurred. Also, in October 2008 a structural break between the U.S. and Chinese No. 1 contracts occurred, and in August 2009 a structural break between U.S. and Chinese No. 2 contracts occurred as well.

Table 1.1 reports trace statistics (λ_{trace}) and maximum eigenvalue statistics ($\lambda_{\max}$) from the bi-variate Johansen test for each sub-period of the sample, and shows that a linear co-integration relationship between U.S. and Chinese No. 1 contracts does not exist in any period. Our result contradicts with Han et al. (2013) finding co-integration between U.S. and Chinese No. 1 contracts. A difference in data sample and modification approach may help explain the difference between the results in our paper and those in Han et al. (2013)—we employ more recent data to capture the price relationship and we proportionally modify futures prices to eliminate the influence of price jump across contracts when rolling price data when the nearby contract expires.

The underlying commodity for the Chinese No. 1 contract is non-GMO soybeans destined to be used for food. The underlying commodity for the U.S. and Chinese No. 2 contracts potentially contain GMO soybeans. Intuitively, the price relationship between U.S. and Chinese No. 2 soybean futures should be much closer than that between U.S. and Chinese No. 1 futures contracts. This is shown in panel C of table 1.1 with a co-integrating relationship prior to 2009 and for the entire period. The co-integration relationship also exists between the U.S. and Brazil.

Table 1.1: Bi-variate Johansen co-integration tests

Panel A: U.S. and Brazil soybean futures contract

	03/01/2005-07/31/2012		08/01/2012-05/30/2015		03/01/2005-06/30/2015	
	λ_{trace}	$\lambda_{\max}$	λ_{trace}	$\lambda_{\max}$	λ_{trace}	$\lambda_{\max}$
$r = 0$	19.91**	19.41***	9.96	7.83	26.03***	22.92***
$r \leq 1$	0.50	0.50	2.12	2.12	3.11	3.11

Table 1.1 continued

Analysis	Co-integrated		Not co-integrated		Co-integrated	
Panel B: U.S. and Chinese No. 1 soybean futures contract						
	03/01/2005-10/31/2008		11/01/2008-06/30/2015		03/01/2005-06/30/2015	
	λ_{trace}	$\lambda_{\max}$	λ_{trace}	$\lambda_{\max}$	λ_{trace}	$\lambda_{\max}$
$r = 0$	10.41	10.21	12.38	8.20	13.55	10.97
$r \leq 1$	0.19	10.21	4.18	4.18	2.58	2.58
Analysis	Not co-integrated		Not co-integrated		Not co-integrated	
Panel C: U.S. and Chinese No. 2 soybean futures contract						
	03/01/2005-08/31/2009		09/01/2009-06/30/2015		03/01/2005-06/30/2015	
	λ_{trace}	$\lambda_{\max}$	λ_{trace}	$\lambda_{\max}$	λ_{trace}	$\lambda_{\max}$
$r = 0$	20.32**	19.83***	10.28	9.25	26.74***	24.03***
$r \leq 1$	0.48	0.48	1.03	1.03	2.71	2.71
Analysis	Co-integrated		Not co-integrated		Co-integrated	

The VAR specification is estimated by applying up to 12 lags. The optimal lag length is determined by means of Schwarz information criterion (SIC). *,**,*** denotes statistical significance at the 10%, 5%, and 1% levels, respectively.

In order to find the direction of long-run lead-lag relationships, we apply a weak exogeneity test to the co-integrating pairs of prices. Table II shows the results of these tests. Parameter α in table 1.2 characterizes the long-run causality relationship in equation (3). When soybean futures price in Brazil is treated as the dependent variable, α is significantly positive. This suggests that U.S. soybean futures price leads the price in Brazil. For the prices of U.S. and Chinese No. 2 futures contracts, the test results indicate a bi-directional causality

relationship, implying the soybean futures prices in U.S. and Chinese No. 2 contracts are influenced by each other.

Table 1.2: Weak exogeneity test for co-integrated price pairs

Panel A: Lead-lag relationship between soybean futures in U.S. and Brazil			
03/01/2005-07/31/2012		03/01/2005-06/30/2015	
$H_0:\alpha_1=0$	$H_0:\alpha_2=0$	$H_0:\alpha_1=0$	$H_0:\alpha_2=0$
1.03	4.02**	0.01	8.64***
US $\Rightarrow$ *Brazil*		*US* $\Rightarrow$ *Brazil*	
Panel B: Lead-lag relationship between U.S. and Chinese No. 2 soybean futures			
03/01/2005-08/31/2009		03/01/2005-06/30/2015	
$H_0:\alpha_1=0$	$H_0:\alpha_2=0$	$H_0:\alpha_1=0$	$H_0:\alpha_2=0$
12.27***	4.13**	9.79***	8.53***
US $\Leftrightarrow$ *ChinaNo*2		*US* $\Leftrightarrow$ *ChinaNo*2	

*,**,*** denotes statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.3 presents the parameter estimates that characterize the long-run equilibrium relationship and the speed of adjustments to the long-run equilibrium. ECT_{t-1} is the error correction term of the VECM model, and its coefficient represents the adjustment speed to long-run equilibrium. ΔUS_t is the close-to-close return of soybean futures in the U.S. ΔBR_t is the close-to-close return of soybean futures in Brazilian market. $\Delta CH2_t$ is the close-to-close return of Chinese No. 2 soybean futures.

Table 1.3: Estimated results of linear error correction model (ECM)

	Panel A: U.S. and Brazilian soybean futures contracts			
	03/01/2005-07/31/2012		03/01/2005-06/30/2015	
	ΔUS_t	ΔBR_t	ΔUS_t	ΔBR_t
ECT_{t-1}	-0.0089	0.0138*	-0.0005	0.0149**
ΔUS_{t-1}	0.0565	0.192***	0.0485	0.1615***
ΔBR_{t-1}	-0.0685	-0.0774*	-0.0731	-0.0577*
Constant	-0.0002	0.0019*	0.0001	0.0033**
	$ECT_t = US_t - 1.008 * BR_t$		$ECT_t = US_t - 1.023 * BR_t$	
	Panel B: U.S. and Chinese No. 2 soybean futures contracts			
	03/01/2005-08/31/2009		03/01/2005-06/30/2015	
	ΔUS_t	$\Delta CH2_t$	ΔUS_t	$\Delta CH2_t$
ECT_{t-1}	-0.0296***	0.0122*	-0.0128***	0.0089**
ΔUS_{t-1}	0.0139	0.1143***	0.0102	0.1096***
$\Delta CH2_{t-1}$	0.0362	0.0199	0.0156	-0.0078
Constant	-0.0016*	0.0014*	-0.0061**	0.0046**
	$ECT_t = US_t - 0.955 * CH2_t$		$ECT_t = US_t - 1.014 * CH2_t$	

*,**,*** denotes statistical significance at the 10%, 5%, and 1% levels, respectively.

The implication of estimation results in table 1.3 coincide with the test results in table 1.2. The significance of estimated coefficients for ECT implies that U.S. soybean futures prices lead the price in Brazil and that the prices of U.S. and Chinese No. 2 futures influence each

other. The significant positive sign of ΔUS_{t-1} in each estimated equation suggests an increasing price change in other markets when the soybean futures price increases in the U.S. market. Finally, the adjustment speed to deviations from the long-run equilibrium is characterized by the magnitude of significant coefficient of ECT, and the long-run equilibrium relationship between prices is characterized by a co-integrating vector in ECT expression.

Turning to the analysis of a non-linear long-run causality relationship, we test whether or not the soybean futures prices in different markets are threshold co-integrated. For the pairs of linear co-integrated prices, the presence of a threshold is tested and estimated via the application of a SupLM test by Hansen and Seo (2002). This tests the null hypothesis of linear (Johansen) co-integration against the alternative hypothesis of threshold co-integration. Table 1.4 displays the test statistics and their bootstrapped p-values out of four data samples. Only the whole period pair of prices between U.S. and Chinese No. 2 soybean futures supports a threshold co-integration at a bootstrapped p-value of 0.035.

Table 1.4: Test of linear co-integration against threshold co-integration

	Test Statistic	P-value
US and Brazil (03/01/2005-07/31/2012)	13.66	0.384
US and Brazil (03/01/2005-06/30/2015)	16.10	0.181
US and CH No. 2 (03/01/2005-08/31/2009)	16.54	0.109
US and CH No. 2 (03/01/2005-06/30/2015)	20.70	0.035

Compared to the linear estimation in table 1.3 above, the co-integrating coefficient for this threshold co-integrated price pair decreases from 1.023 to 0.97. This again shows a strong co-movement between Chinese No. 2 soybean futures price and the U.S. soybean futures price.

The estimated critical threshold value is 0.26 cents per bushel, which divides the whole data set into two regimes. There are 86.9% observations that fall into the usual regime $|US_t - 0.97 * CH2_t| \leq 0.26$, while the remaining 13.1% of observations belong to the unusual regime $|US_t - 0.97 * CH2_t| > 0.26$. The estimated TVECM model is fully represented as

$$\Delta US_t = \begin{cases} -0.0016^{*} - 0.0084^{*} ECT_{t-1} + 0.0366 \Delta US_{t-1} + 0.0237 \Delta CH2_{t-1}, & |ECT_{t-1}| \leq 0.26 \\ 0.0158^{**} + 0.0394^{*} ECT_{t-1} - 0.0773^{*} \Delta US_{t-1} - 0.0512 \Delta CH2_{t-1}, & |ECT_{t-1}| > 0.26 \end{cases}$$

$$\Delta CH2_t = \begin{cases} 0.0016^{***} + 0.0083^{**} ECT_{t-1} + 0.1108^{***} \Delta US_{t-1} - 0.0162 \Delta CH2_{t-1}, & |ECT_{t-1}| \leq 0.26 \\ -0.0061 - 0.0153 ECT_{t-1} + 0.0986^{***} \Delta US_{t-1} + 0.0658 \Delta CH2_{t-1}, & |ECT_{t-1}| > 0.26 \end{cases}$$

where $ECT_t = US_t - 0.97 * CH2_t$

The adjustment parameters of ECT in the U.S. equation are -0.0084 and 0.0394 in the usual and unusual regime, respectively. This difference in the statistically significant magnitude of ECT coefficient indicates a faster adjustment speed toward long-run equilibrium when the absolute value of price deviation from equilibrium exceeds the critical threshold. In both the usual and the unusual regimes, the estimated coefficients of ΔUS_{t-1} are significantly different from zero in the Chinese No. 2 equation, while the estimated coefficients of $\Delta CH2_{t-1}$ are not statistically significant in the U.S. equation. This suggests that there is a significant short-run response of the Chinese No. 2 soybean futures price to the price change in the U.S. When the U.S. soybean futures price changes by 1%, the Chinese No. 2 soybean futures price changes by 0.1108% and 0.0986% in the same direction when deviation from equilibrium belongs to the usual and unusual regime, respectively. These short-run adjustment parameters provide evidence that prices in the U.S. typically lead prices for the Chinese No. 2 soybean futures contract.

Turning to the remaining five pairs of futures prices, which do not exhibit linear co-integration relations, we apply a SupWard test by Seo (2006) to test whether or not they are threshold co-integrated and to demonstrate parameter estimates. Specifically, we implement threshold co-integration analysis for the post-break period prices of U.S. and Brazilian soybean futures, the post-break period prices of U.S. and Chinese No. 2 soybean futures and each sub-period of U.S. and Chinese No. 1 soybean futures.

Table 1.5 shows the results of the test of no co-integration versus threshold co-integration. Two pairs of prices reject the no co-integration null hypothesis at a less-than 10% significant level. This provides evidence of threshold co-integration between these prices. The bootstrapped p-value is 0.001 and 0.076 for the pre-break period and the whole period of prices between U.S. and Chinese No. 1 contracts, respectively. The test results for the other three pairs of soybean futures prices are not significant at conventional levels.

Table 1.5: Test of no co-integration against threshold co-integration

	Test Statistic	P-value
US and Brazil (08/01/2012-05/30/2015)	15.20	0.103
US and CH No. 1 (03/01/2005-10/31/2008)	16.36	0.001
US and CH No. 1 (11/01/2008-06/30/2015)	19.39	0.551
US and CH No. 1 (03/01/2005-06/30/2015)	13.15	0.076
US and CH No. 2 (09/01/2009-06/30/2015)	10.67	0.316

The co-integrating coefficient is estimated as $\beta = 0.94$ for the pre-break period data of U.S. and Chinese No. 1 soybean futures, showing a strong responsiveness of the Chinese No. 1 contract price to the U.S. soybean futures price. The estimated threshold point is 0.09 cent

per bushel, which divides the observations into two regimes. Of the observations, 84.3% fall into the usual regime $|US_t - 0.94 * CH1_t| \leq 0.09$, and the remaining 15.7% of observations belong to the unusual regime $|US_t - 0.94 * CH1_t| > 0.09$. The estimated TVECM model is fully represented as

$$\Delta US_t = \begin{cases} 0.0011^{*} - 0.0134 ECT_{t-1} - 0.0135 \Delta US_{t-1} 0.0590 \Delta CH1_{t-1}, & |ECT_{t-1}| \leq 0.09 \\ 0.0035 - 0.0511 ECT_{t-1} + 0.1678 \Delta US_{t-1} - 0.3792^{**} \Delta CH1_{t-1}, & |ECT_{t-1}| > 0.09 \end{cases}$$

$$\Delta CH1_t = \begin{cases} 0.0002 - 0.0008 ECT_{t-1} + 0.2650^{***} \Delta US_{t-1} + 0.0850 \Delta CH1_{t-1}, & |ECT_{t-1}| \leq 0.09 \\ -0.0031^{**} + 0.0407^{**} ECT_{t-1} + 0.3589^{***} \Delta US_{t-1} - 0.3176^{***} \Delta CH1_{t-1}, & |ECT_{t-1}| > 0.09 \end{cases}$$

where $ECT_t = US_t - 0.94 * CH1_t$

The long-run adjustment parameters of ECT in Chinese No. 1 equations are significantly different from zero in the unusual regime, while the long-run adjustment parameters of ECT in U.S. equations are not statistically significant. This indicates that the U.S. soybean futures price drives the Chinese No.1 soybean futures price toward the equilibrium level. In particular, the adjustment parameters of ECT in the Chinese No. 1 equation are 0.0008 and 0.0407 in the usual and unusual regime, respectively. This difference in the magnitude of ECT coefficient demonstrates a faster adjustment speed toward long-run equilibrium when the absolute value of deviation from equilibrium exceeds the critical threshold. In both the usual and the unusual regimes, the estimated coefficients of ΔUS_{t-1} are significantly different from zero in the Chinese No. 1 equation. This suggests that there is a significant short-run response of the Chinese No. 1 soybean futures price to the price change in U.S. When the U.S. soybean futures price changes by 1%, the Chinese No. 1 soybean futures price changes by 0.2650% and 0.3589% in the same direction when deviation from equilibrium belongs to the usual and unusual regime, respectively. These short-run adjustment parameters provide evidence that

prices in the U.S. typically lead prices for the Chinese No. 1 soybean futures contract, this result coincides with the results from the long-run adjustment parameters of the ECT.

For the whole period data of U.S. and Chinese No.1 soybean futures price, the co-integrating coefficient is estimated as $\beta = 0.98$, showing a strong responsiveness of the Chinese No. 1 contract price to the U.S. soybean futures price. The estimated threshold point is 0.41 cents per bushel, which divides the observations into two regimes. Of the observations, 92.5% fall into the usual regime $|US_t - 0.98 * CH1_t| \leq 0.41$, and the remaining 7.5% observations belong to the unusual regime $|US_t - 0.98 * CH1_t| > 0.41$. The estimated TVECM model is fully represented as

$$\Delta US_t = \begin{cases} -0.0011 - 0.0065 ECT_{t-1} + 0.0296 \Delta US_{t-1} - 0.0443 \Delta CH1_{t-1}, & |ECT_{t-1}| \leq 0.41 \\ 0.0073 + 0.0147 ECT_{t-1} - 0.2369^{***} \Delta US_{t-1} - 0.2729^{**} \Delta CH1_{t-1}, & |ECT_{t-1}| > 0.41 \end{cases}$$

$$\Delta CH1_t = \begin{cases} 0.0012^{***} + 0.0049^{**} ECT_{t-1} + 0.2189^{***} \Delta US_{t-1} - 0.0112 \Delta CH1_{t-1}, & |ECT_{t-1}| \leq 0.41 \\ 0.0181^{**} + 0.0387^{**} ECT_{t-1} - 0.0651 \Delta US_{t-1} + 0.1274 \Delta CH1_{t-1}, & |ECT_{t-1}| > 0.41 \end{cases}$$

where $ECT_t = US_t - 0.98 * CH1_t$

The long-run adjustment parameters of the ECT in the Chinese No. 1 equations are significantly different from zero in both the usual and unusual regimes, while the long-run adjustment parameters of ECT in the U.S. equations are not statistically significant. This indicates that the U.S. soybean futures price drives the Chinese No. 1 soybean futures price toward the equilibrium level. In particular, the adjustment parameters of the ECT in the Chinese No. 1 equation are 0.0049 and 0.0387 in the usual and unusual regimes, respectively. This difference in the magnitude of the ECT coefficients suggests a faster adjustment speed when the absolute value of the deviation from equilibrium exceeds the critical threshold. In the usual region, the estimated coefficient of ΔUS_{t-1} is significantly different from zero in the

Chinese No. 1 equation while the estimated coefficient of $\Delta CH1_{t-1}$ is not statistically significant in the U.S. equation. This suggests that there is a significant short-run response of the Chinese No. 1 soybean futures price to price changes in the U.S. When the U.S. soybean futures price changes by 1%, the Chinese No. 1 soybean futures price changes by 0.2189% in the same direction. These short-run adjustment parameters provide evidence that prices in the U.S. typically lead prices for the Chinese No. 1 soybean futures contract, coinciding with the results from the long-run adjustment parameters of the ECT.

In summary, the co-integration results demonstrate unidirectional long-run price causality from U.S. to Brazilian and U.S. to Chinese No. 1 soybean futures markets, and a bi-directional long-run causality relationship between U.S. and Chinese No. 2 soybean futures markets. However, the estimation results that indicate Chinese soybean futures prices leading the price of soybean futures in U.S. are only significant in the unusual sample regime, indicating that soybean futures prices in Brazil or China are still led by futures prices in the U.S.

1.4.2 Short-Run Causal Effect

We are interested in two different short-run causal effects. One is the contemporaneous effect of U.S. Globex overnight return on the daytime returns of soybean futures in China's market. The other is the seasonal harvest effect on U.S. and Brazilian soybean futures returns.

Since U.S. soybean futures contracts can be traded through the Globex overnight platform, the Chinese soybean futures daytime return may be affected by this synchronous trading. Our analysis of the short-run causal effect concentrates on how daytime returns in China are affected by the information content of U.S. overnight prices. For simplicity, we set $p = q = 1$ and derive the unstructured estimation of the ARDL(1,1) model by ordinary least

squares.[3] In particular, the impact of U.S. overnight returns on Chinese daytime returns is examined by the following regressions:

$$\Delta CH_t^D = \alpha_1 + \beta_1 * \Delta US_t^N + \eta_1 * \Delta US_{t-1}^N + \gamma_1 * \Delta CH_{t-1}^D + \varepsilon_t$$

$$\Delta US_t^N = \alpha_2 + \beta_2 * \Delta CH_t^D + \eta_2 * \Delta CH_{t-1}^D + \gamma_2 * \Delta US_{t-1}^N + \varepsilon_t$$

Table 1.6 reports the estimation results for different sub-period samples. It shows that U.S. soybean futures overnight returns and Chinese No. 1 soybean futures daytime returns significantly affect each other in five out of eleven years, while estimation results of U.S. soybean futures overnight returns and Chinese No. 2 soybean futures daytime returns are not statistically significant at conventional levels. These results indicate that information about U.S. Globex overnight trading influences the price change of Chinese No. 1 soybean futures contracts rather than No. 2 soybean futures contracts. This is mainly because Chinese No. 1 contracts are more active in the market. Thus, the short-run market price of No. 1 contracts would be more sensitive to the price in other markets due to no arbitrage theory. Therefore, the soybean futures price in Chinese No. 1 contracts are not only threshold co-integrated with those in the U.S. in the long-run, its short-run price change is also influenced by the information from overnight price changes in the U.S. market as well. Except for 2015, all significant coefficients of β_i s are positive, indicating that the price increase in U.S. overnight trading will stimulate the trading of Chinese No. 1 soybean futures contracts and tend to increase daytime return.

[3] Imposing no structure on the relationship of the coefficients of the lagged explanators may cause multicollinearity, leading to high variance of the coefficient estimates.

Table 1.6: U.S. overnight and Chinese daytime return

Year	Dependent Variable	Column A: U.S. and Chinese No. 1 Contract β_i	Column B: U.S. and Chinese No. 2 Contract β_i
2005	CH^D	0.2117***	0.1109
	US^N	0.5859***	0.1151
2006	CH^D	0.1390*	-0.0658
	US^N	0.1951*	-0.0359
2007	CH^D	0.0499	0.0208
	US^N	0.1125	0.0152
2008	CH^D	0.0759*	0.0530
	US^N	0.1933*	0.0939
2009	CH^D	0.0669	0.0694
	US^N	0.1656	0.0689
2010	CH^D	-0.0002	-0.0256
	US^N	-0.0001	-0.0339
2011	CH^D	0.0899*	-0.0071
	US^N	0.1514*	-0.0076
2012	CH^D	0.0717	-0.0225
	US^N	0.1516	-0.0835

Table 1.6 continued

2013	CH^D	-0.0424	0.1245
	US^N	-0.0514	0.0907
2014	CH^D	0.0084	-0.0356
	US^N	0.0063	-0.0572
2015	CH^D	-0.3805*	0.0916
	US^N	-0.1151*	0.0910

*,**,*** denotes statistical significance at the 10%, 5%, and 1% levels, respectively.

The other temporal causal effect we are interested in is the seasonal harvest effect on U.S. and Brazilian soybean futures. From the long-run analysis above, we conclude that soybean futures prices in the U.S. market lead those prices in the Brazilian market. In the short-run, however, this may not always be the case. The U.S. peak harvest period extends from May to October, while peak harvest period in Brazil extends from November to April. As a result, it is likely that the Brazilian soybean futures price leads the price in the U.S. in its harvest period when a strong seasonal effect exists. Table 1.7 illustrates the results of a causality test in each year, where expected seasonal causality exists in five out of ten years, suggesting a weak seasonal causal effect between U.S. and Brazilian soybean futures prices over time.

Table 1.7: Causality test of U.S. and Brazilian seasonal production effect

	Nov-April Period	
US vs. Brazil	Causality Test	
	$H_0: \alpha_1 = 0$	$H_0: \alpha_2 = 0$

Table 1.7 continued

Year 2005	7.34***	0.07
Analysis	*Brazil* $\Rightarrow$ *US*	
Year 2006	2.39	1.56
Analysis	-	
Year 2007	0.01	4.00**
Analysis	*US* $\Rightarrow$ *Brazil*	
Year 2008	3.37*	8.58**
Analysis	*Brazil* $\Leftrightarrow$ *US*	
Year 2009	1.81	0.77
Analysis	-	
Year 2010	3.51*	3.51*
Analysis	*Brazil* $\Leftrightarrow$ *US*	
Year 2011	0.16	7.62**
Analysis	*US* $\Rightarrow$ *Brazil*	
Year 2012	6.31**	0.29
Analysis	*Brazil* $\Rightarrow$ *US*	
Year 2013	0.25	2.58
Analysis	-	
Year 2014	5.28**	0.23
Analysis	*Brazil* $\Rightarrow$ *US*	

- indicates insignificant causality. *,**,*** denotes statistical significance at the 10%, 5%, and 1% levels, respectively.

1.5 Discussion and Conclusion

This paper offers a comprehensive study on price causality between soybean futures prices in different markets from 2005 to 2015. Both long-run and short-run price relations are examined by various time-series methods. The long-run empirical results in section 4 indicate that the U.S. soybean futures market is the most influential market, and soybean futures prices in other markets like Brazil and China are led by the price change in the U.S. However, some sub-period co-integration tests show that there is no directional causality between U.S. and Chinese No. 2 soybean futures prices, and that there is a unidirectional causality between U.S. and Chinese No. 1 soybean futures prices. Intuitively, the rapid growth of both Chinese soybean spot and futures markets make them more influential to futures prices in the world. Thus, the lead-lag relationship between U.S. and Chinese soybean futures has been changed in recent years. If we form a liquid trading strategy based on whole period lead-lag relationship between U.S. and Chinese soybean futures prices,[4] we find in appendix A4 that the return for trading strategy beats the real return in only four out of eleven years, showing that the directional causality between U.S. and Chinese soybean futures markets is not always robust.

This paper also investigates two types of temporal price causalities. One is the effect of overnight price changes of U.S. soybean futures through the Globex electronic trading system on the daytime return of Chinese soybean futures, the other is the seasonal harvest effect between U.S. and Brazilian soybean futures prices. The results indicate that the Globex

[4] The strategy is to purchase and sale of the Chinese soybean futures contract at time t when the U.S. soybean futures return was positive at time t-1. Otherwise, keep the asset and earn risk-free interest return.

overnight price change in soybean futures affects, to some extent, the daytime price of Chinese No. 1 soybean futures, but there is no significant evidence indicating that overnight trading affects Chinese No. 2 soybean futures. Moreover, half of empirical tests about seasonal harvest effect match with our intuitive expectation, indicating a week seasonal causality between U.S. and Brazilian soybean futures prices according to their harvest periods.

In general, we can conclude from this paper that the U.S. still plays an important role in the worldwide soybean market, and the price changes in U.S. soybean futures will affect the futures price in other markets like China and Brazil. However, with the development of soybean markets in Brazil and China, this long-run unidirectional price causality from the U.S. to Brazil or from the U.S. to China has been weakened, and the opposite direction of price causality has begun to emerge.

1.6 Appendix

Table 1.8: Unit root tests for U.S. and Brazil contracts in each sub-period

	03/01/2005-07/31/2012		08/01/2012-06/30/2015		03/01/2005-06/30/2015	
	ADF Test	PP Test	ADF Test	PP Test	ADF Test	PP Test
US	1.2643	-0.8794	0.4349	-1.5024	0.537	-1.8268
ΔUS	-29.5422***	-40.8397***	-23.1862***	-30.7071***	-34.4713***	-48.3848***
BR	1.5526	-0.7508	0.2979	-1.4684	0.6031	-1.87
ΔBR	-27.8188***	-37.3443***	-21.275***	-29.5627***	-32.3528***	-44.5508***

*,**,*** denotes statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.9: Unit root tests for U.S. and Chinese No. 1 contracts in each sub-period

	03/01/2005-10/31/2008		11/01/2008-06/30/2015		03/01/2005-06/30/2015	
	ADF Test	PP Test	ADF Test	PP Test	ADF Test	PP Test
US	0.6333	-0.961	0.1158	-1.8888	0.537	-1.8268
ΔUS	-21.1003***	-29.2323***	-26.9963***	-38.4172***	-34.4713***	-48.3848***
$CH1$	1.2178	-0.5078	0.41	-1.3797	1.1877	-1.5633
$\Delta CH1$	-17.3573***	-25.8864***	-28.2249***	-37.3029***	-31.5805***	-44.6086***
ΔUS^N	-18.7024***	-27.9519***	-25.0827***	-38.2538***	-31.2631***	-47.4886***
$\Delta CH1^D$	-13.518***	-25.9058***	-16.6432***	-27.8649***	-21.0603***	-37.7819***

*,**,*** denotes statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.10: Unit root tests for US and Chinese No. 2 contracts in each sub-period

	03/01/2005-08/31/2009		09/01/2009-06/30/2015		03/01/2005-06/30/2015	
	ADF Test	PP Test	ADF Test	PP Test	ADF Test	PP Test
US	0.8018	-1.0523	-0.0632	-1.4082	0.537	-1.8268
ΔUS	-23.1492***	-32.1139***	-25.2243***	-36.1708***	-34.4713***	-48.3848***
$CH2$	1.1567	-0.8125	-0.3902	-1.0234	0.6177	-1.7365
$\Delta CH2$	-31.3313***	-31.0776***	-26.1722***	-37.7651***	-33.3303***	-48.3886***
ΔUS^N	-20.7158***	-31.3953***	-23.3838***	-35.3627***	-31.2631***	-47.4886***
$\Delta CH2^D$	-13.8916***	-23.822***	-12.8969***	-20.0804***	-18.9965***	-32.6365***

*,**,*** denotes statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.11: Forecasted trading strategy vs. real return

Year	% of correct forecast direction	Strategy return	Real return
2005	48.74%	94.47%	104.73%
2006	46.90%	90.58%	117.01%
2007	56.28%	158.22%	168.04%
2008	52.79%	102.14%	83.23%
2009	50.22%	132.64%	116.55%
2010	50.44%	164.62%	109.95%
2011	50.00%	95.80%	109.63%
2012	48.00%	94.25%	105.61%
2013	53.78%	66.91%	96.07%
2014	42.61%	48.35%	76.74%
2015	62.50%	135.55%	92.09%

CHAPTER 2. THE SUPPLY CURVE FOR CELLULOSIC ETHANOL

Abstract

This paper examines the impact of feedstock supply mechanisms under conditions of spatial monopoly on the supply of cellulose to the plants and by implication on the supply of cellulosic ethanol. we show the minimization problem for cellulosic processors under three different collection mechanisms and provide optimal pricing rule and the optimal collection radius needed to meet feedstock supply requirements. These show that in the absence of competition the processor is indifferent between processor collection and supplier delivery, but that societal welfare is higher under supplier delivery. We then use a repeated Nash Equilibrium game to show that processor collection is first best for both incumbent and processor and is an effective deterrent against an entrant locating a plant within the draw area of the incumbent. We support the theoretical results with a numerical simulation showing the optimal premium and draw area under each mechanism. Third, we use the result of the simulation show the rate at which stover collection costs increase for a monopsonistic stover processor constrained to the original draw area. The slope of the marginal cost curve for this monopsonistic processor is then compared with the slope of the cost curve across other feedstocks. These results suggest that substantial quantities of these other feedstocks may be required to meet the mandate.

2.1 Introduction

Management at the first two commercial cellulosic ethanol producers have been willing to accept low feedstock supplier participation and a large draw area as an alternative to increasing the premium they pay for stover. This behavior is optimal if the plants are treated as monopsonistic buyers. The existing plants have also arranged to collect stover from some or all suppliers. This processor collection model is shown to be an effective way to deter future entrants from building plants within the draw area of the initial plants. As the cellulosic mandate causes more plants to come on line, these new plants will be close to, but will not overlap with, existing plants. All plants will be able to increase production within their original draw areas by increasing the premium offered for stover. However, the slope of the marginal cost function is steep due to monopsonistic competition. This situation was not anticipated by the original literature on the availability of cellulosic feedstock, where 100% participation was implicitly assumed. This low participation rate among stover producers puts in doubt the ability of the industry to meet the cellulosic mandate from agricultural residues alone and increases the likelihood that other more expensive feedstocks, such as perennial grasses will be required.

The cellulosic mandate written into the US renewable Fuel Standard (RFS) requires the use of almost 16 billion gallons of transportation fuel produced from grasses, trees, and agricultural and municipal waste (Bracmort 2015). If the law is implemented as originally proposed, mandated parties (blenders) will be required to purchase and use a specific quantity of cellulosic biofuels or to purchase Cellulosic Renewable Identification Numbers (CRINs) from another blender that has blended more cellulosic fuel than required. The market value of CRINs will rise to ensure that the combined value of the fuel and CRINs to cellulosic fuel processors is sufficient to cover the full production costs of the marginal fuel producer. This

law effectively creates a vertical demand curve for cellulosic ethanol at the quantity that is mandated for use each year. This means that the price and welfare implications of this policy will be determined at the intersection of the marginal cost curve and this mandated demand curve. As the annual mandate increases, the CRIN value will rise to ensure that the costs of the marginal supplier are fully met.

The existing literature on the availably of cellulosic feedstock, US DoE (2011), Ogden and Anderson (2011), Graham et al. (2007) and Archer and Johnson (2012), focused on whether sufficient cellulosic feedstocks would be physically available to meet the mandate. In a market based system, physical availability is not equal to quantity supplied. Individual owners of the feedstock must receive a price that is greater than their reservation price. Operators of cellulosic fuel plants may not be willing to pay a feedstock price that results in 100% participation.

Earlier work by Ogden and Anderson (2011) and Dumortier (2015) found that the supply of corn stover and other agricultural wastes would be sufficient to meet the mandate. However, these authors assumed either 100% participation or homogenous processors. The supply curve for cellulosic ethanol will be different than that for corn ethanol in two important ways. First, feedstock suppliers are not homogenous with respect to the price at which they will supply. Second, the cellulosic ethanol producer will, at least initially, have monopsony power in the draw area near their plant.

This study draws on the experience of two of the first large scale cellulosic ethanol processors, both of which are located in Iowa and use corn stover as a feedstock. One of these plants collects all of the stover and pays each farmer in the draw area the same price. The other uses a mixed collection model where it accepts some farmer delivered stover and collects the

rest at a distance from the plant. The experience with these plants indicates that, under the current price structure, less than 20% of farmers are willing to participate (Pieper 2015). Corn farmers with the lowest reservation price typically grow continuous corn and use animal manure as fertilizer. Farmers in a corn-soybean rotation who use chemical fertilizer on erodible soils are much less likely to participate. The first group of feedstock suppliers view stover as a waste product that can actually reduce yields if it is not removed. Suppliers with high reservation prices value the carbon and fertilizer value of the stover that is removed. As a result of this heterogeneity, and the low participation rate, both plants have collection areas that are larger than originally anticipated and which can reach as far as 50 miles from the plant (Pieper 2015; see also Swoboda 2014).

Cellulosic biofuel processors could increase the feedstock price in order to increase participation, but they are aware that they would be required to pay this additional price for those suppliers who are willing to deliver at the lower price. Instead, the processors have been willing to incur the additional transportation cost associated with very large draw areas. In essence, these processors are acting as monopsonistic buyers even though there is no barrier to entry other than the $200–$300 million construction cost.

The collection mechanism at both plants is unusual in that it depends on the use of plant-owned transportation equipment and labor to collect stover from corn fields. This processor collection mechanism runs counter to that used for grain where grain suppliers deliver to a central location.

The objective of this article is to examine the impact of feedstock supply mechanisms under conditions of spatial monopoly on the supply of cellulose to the plants and by implication

on the supply of cellulosic ethanol. The theoretical underpinnings of the model are based on French (1960) and Melvin L. Greenhut et. al. (1987).

This cost structure is provided for: (*a*) a processor who can perfectly price discriminate; (*b*) a collection mechanism where all feedstock suppliers receive the same price and the processor pays all collection costs (processor collection); and, (*c*) a feedstock supplier delivery mechanism with a single price at the plant (supplier delivery). The assumption of perfect price discrimination is unrealistic and is provided only as a benchmark against which the other two systems can be evaluated. It is shown that the processor is indifferent between processor collection and supplier delivery, total costs are lower under supplier delivery, and that welfare is higher under supplier delivery. It is then shown that processor collection is an effective tool to prevent new plants from locating within a collection area that overlaps with that of the original plant.

If existing processors have an effective tool to prevent poaching from within their draw areas, then the mandate can only be met with stover if existing processors offer higher premiums to increase participation to expand capacity within existing draw areas. The slope of the marginal cost curve in this situation is much steeper than the cross feedstock supply curve presented in Ogden and Anderson. In other words, the mandate will be met only if other, more expensive feedstocks are used.

The article proceeds as follows: first, we show the minimization problem for processors under each of the three different collection mechanisms and provide optimal pricing rule and the optimal collection radius needed to meet feedstock supply requirements. Second, we develop and prove three propositions related to the delivery system. These show that in the absence of competition the processor is indifferent between processor collection and supplier

delivery, but that societal welfare is higher under supplier delivery. We use a repeated Nash Equilibrium game to show that processor collection is first best for both incumbent and processor. We then show that processor collection is an effective deterrent against an entrant locating a plant within the draw area of the incumbent. We support the theoretical results with a numerical simulation showing the optimal premium and draw area under each mechanism. Third, we use the result of the simulation show the rate at which stover collection costs increase for a monopsonistic stover processor constrained to the original draw area. A key parameter is the degree to which stover suppliers will respond to higher prices. This parameter can be uncovered if we assume it is known to the processors and that they are optimizing based on this value. The slope of the marginal cost curve for this monopsonistic processor is then compared with the slope of the cost curve across other feedstock. These results suggest that substantial quantities of these other feedstock may be required to meet the mandate.

2.2 The Model

2.2.1 Single Processor Model

Our model assumes that corn stover suppliers are heterogeneous and have different reservation values for their feedstock. Only when price received by suppliers exceed their reservation value are they willing to sell. As a result, the participation rate varies depending on the net price received for the feedstock. As long as the plant offers a positive price, some farmers will supply.

Suppose each unit of land produces one unit of corn stover. Land is uniformly located along a line with one land unit per unit distance. The processor is located at one end of the line. The fraction of each land unit that supplies stover is given by the participation rate function

$s(p) = bp$, where p is the net price received by the farmer and parameter b characterizes supplier's willingness to supply[5].

Collection Mechanism I, Perfect Price Discrimination

The first collection mechanism allows for perfect price discrimination and processor collection. From the processor perspective, this is an ideal case where the processor can elicit each supplier's reservation value. We include this scenario because the outcomes mimic the radius the processor would choose if it did not behave in a monopsonistic fashion (Varian 2009). In this scenario there is no deadweight loss and the system reaches a competitive equilibrium (Varian 2009).

Let R be the distance between the processor and the marginal farmer and p is the highest reservation value in the collection region. Since each land unit supplies bp, the total supply received by the processor is bpR. The total requirement to reach capacity is Q and so, $Q = bpR$. This implies that $R = Q/bp$.

Here, the processor will pay each supplier their unique reservation value and the amount paid to each land unit is $bp^2 - \int_0^p bxdx = 0.5bp^2$. If processor lacks the ability to discriminate farmers, it has to pay each land unit bp^2. Thus, $\int_0^p bxdx = 0.5bp^2$ is the collection cost advantage

[5] An alternative model that allows for a circular draw area and which allows for an optimal response when 100% participation is reached is available from the authors. Key results are identical.

by the processor's price discrimination ability. In addition to the transportation cost, the total expenditure for processor to collect Q amount of corn stove is given as

$$Expenditure = \int_0^R (0.5bp^2 + bptr)dr = 0.5bp^2R + 0.5bptR^2$$

Substitute in $R = Q/bp$ to obtain

$$Expenditure = 0.5pQ + 0.5\frac{tQ^2}{bp}$$

Minimizing total expenditure respect to price provides optimal price and collection distance $p^* = \left(\frac{tQ}{b}\right)^{0.5}$ and $R^* = \left(\frac{Q}{bt}\right)^{0.5}$. The larger the capacity, the higher the required price and the longer the collection distance. Substituting the equilibrium price and distance to processor's expenditure function, the total expenditure for the processor is $(t/b)^{0.5}Q^{1.5}$. It is useful to separate equilibrium payments and equilibrium transportation costs for the processor, where

$$Payment = \int_0^{R^*} 0.5bp^2dr = 0.5bp^2R = 0.5(t/b)^{0.5}Q^{1.5},$$

$$Transportation\ cost = \int_0^{R^*} bptrdr = 0.5bptR^2 = 0.5(t/b)^{0.5}Q^{1.5}$$

As can be seen, half of money paid by processor is received by suppliers, and the other half is spent on transportation.

Collection Mechanism II, Processor Collection

In this model the processor is responsible for the cost of stover transportation and pays every supplier the same price. Given price p, the supply is $Q = bpR$. Thus, the required

distance to collect Q is $R = \frac{Q}{bP}$. If the distance between the supplier and processor is $r \le R$, the total amount paid by the processor for each unit corn stover is $p + tr$. Thus, total expenditure is given by

$$Expenditure = \int_0^R bp(p + tr)dr = bp^2R + 0.5bpt(R)^2$$

Substitute in $R = Q / bp$ to obtain

$$Expenditure = pQ + 0.5\frac{tQ^2}{bp}$$

Minimizing total expenditure with respect to price results in the optimal price and collection distance $p^* = \left(\frac{tQ}{2b}\right)^{0.5}$ and $R^* = \left(\frac{2Q}{bt}\right)^{0.5}$. The collection distance in this scenario is larger than under perfect price scenario because processor does not have the ability to discriminate. Substituting the equilibrium price and distance into the expenditure function, total expenditure is $\sqrt{2}\left(t / b\right)^{0.5} Q^{1.5}$.

Separating the equilibrium payments and equilibrium transportation costs, the payment to suppliers is

$$Payment = \int_0^{R^*} bp^2 dr = bp^2 R = \frac{\sqrt{2}}{2}(t / b)^{0.5} Q^{1.5} ,$$

and transportation cost is

$$Transportation\ cost = \int_0^{R^*} bptrdr = 0.5bptR^2 = \frac{\sqrt{2}}{2}(t / b)^{0.5} Q^{1.5}$$

Again, half of cost paid by the processor is received by farmers and the other half is spent on transportation.

Collection Mechanism III, Supplier Delivery

Under supplier delivery, the processor pays a price of p at the plant when stover is delivered. In this case, the net price received by a supplier r miles away is $p - tr$. Those who are closer to the processor will receive higher net prices and are more likely to participate. As a result, the participation rate is no longer constant within the collection area, that is $s(p) = b(p - tr)$. Assuming the marginal supplier is a distance R from the processor, the total supply received by the processor is

$$Q = \int_0^R b(p - tr)dr = bpR - 0.5btR^2$$

Thus, the price offered by processor can be expressed in term of Q as

$$p = \frac{Q + 0.5btR^2}{bR} = \frac{Q}{bR} + 0.5tR .$$

The total expenditure in this case is given by

$$Expenditure = \int_0^R b(p - tr)pdr = bp^2R - 0.5btpR^2 = \left(\frac{Q}{bR} + 0.5tR\right)Q$$

Minimizing total expenditure with respect to collection distance results in $R^* = \left(\frac{2Q}{bt}\right)^{0.5}$, which is the same as that for processor collection. Substitute this expression into $p = \frac{Q}{bR} + 0.5tR$ to obtain $p^* = \left(\frac{2tQ}{b}\right)^{0.5}$.

The expression for expenditure of the processor is:

Expenditure = $pQ = \left(\frac{Q}{bR} + 0.5tR\right)Q = \sqrt{2}\left(t/b\right)^{0.5} Q^{1.5}$, which is the same as the expression for processor collection case. Total transportation costs can be expressed as

$$\textit{Transportation cost} = \int_0^R b(p-tr)trdr = \frac{1}{2}btpR^2 - \frac{1}{3}bt^2R^3 = \frac{\sqrt{2}}{3}\left(t/b\right)^{0.5}Q^{1.5}$$

The model above indicates that the optimal collection distances are the same under processor collection and supplier delivery.

Proposition 1. If transportation costs are identical for processor and supplier, the biofuel processor is indifferent between processor collection and supplier delivery and the optimal collection distance and resulting collection expenditure are the same for both mechanisms.

Proof

In processor collection mechanism, the expenditure minimization problem for processor to collect Q amount of corn stover is expressed as

$$\min_{\{p\}} \; pQ + 0.5\frac{tQ^2}{bp}$$

The F.O.C is $Q - \frac{tQ^2}{2bp^2} = 0$, which implies the solution $p^* = \left(\frac{tQ}{2b}\right)^{0.5}$ and $R^* = \left(\frac{2Q}{bt}\right)^{0.5}$.

Substituting the equilibrium price and distance to processor's expenditure function, the total expenditure is given as $\sqrt{2}\left(t/b\right)^{0.5}Q^{1.5}$.

In the supplier delivery mechanism, the expenditure minimization problem for processor to collect Q is expressed as

$$\min_{\{R\}} \left(\frac{Q}{bR} + 0.5tR\right)Q$$

The F.O.C is $\frac{1}{2}tQ - \frac{Q^2}{bR^2} = 0$, which implies $R^* = \left(\frac{2Q}{bt}\right)^{0.5}$. Thus, the optimal collection distance between processor collection mechanism and supplier delivery mechanism are both

given as $\left(\frac{2Q}{bt}\right)^{0.5}$. Substitute this expression into $p=\frac{Q}{bR}+0.5tR$ to obtain optimal price $p^*=\left(\frac{2tQ}{b}\right)^{0.5}$. Total expenditures by the processor can be expressed as: $p^*Q=\sqrt{2}\left(t/b\right)^{0.5}Q^{1.5}$ which is the same as expenditure in the processor collection mechanism. **Q.E.D.**

The key reason makes processor is indifferent with processor collection and supplier delivery is because processor can't price discriminate. This makes the processor can't capture the supply efficiency, while the welfare for suppliers and society is higher in the supplier delivery mechanism.

Proposition 2. Supplier welfare is higher and total transportation costs are lower under supplier delivery.

Proof

In either mechanism, the welfare, or, payment to the supplier is the difference between the total expenditure of the processor and the cost of transportation. In the processor collection mechanism, payment to suppliers is $\int_0^{R^*} bp^2 dr = bp^2R = \frac{\sqrt{2}}{2}(t/b)^{0.5}Q^{1.5}$, which is half of the total expenditure. In supplier delivery mechanism, total expenditure $\left(\frac{Q}{bR}+0.5tR\right)Q=\sqrt{2}\left(t/b\right)^{0.5}Q^{1.5}$ is needed to collection Q amount of corn stover, while the transportation cost is $\int_0^R b(W-tr)trdr=\frac{1}{2}btWR^2-\frac{1}{3}bt^2R^3=\frac{\sqrt{2}}{3}\left(t/b\right)^{0.5}Q^{1.5}$. Thus, the payment to suppliers is

Payment = Expenditure – Transportation cost $=\frac{2\sqrt{2}}{3}\left(t/b\right)^{0.5}Q^{1.5}$,

which is two thirds of the total expenditure. As shown in Proposition 1 total expenditures are the same under both mechanisms. Participation under supplier delivery is highest near the plant. This reduces overall transportation costs The single price rule prohibits the plant from capturing the efficiencies associated with supplier delivery. Instead, these benefits are captured by suppliers located near the plant due to their locational advantage. Costs for processors are the same under both mechanisms and therefore societal welfare will be higher in the supplier delivery mechanism. **Q.E.D.**

Proposition 1 and 2 are similar to the spatial monopoly pricing results shown in Greenhut et. al. (1987) that a spatial monopolist's market boundaries and profit in the case of linear demand are identical for mill pricing and uniform delivered pricing, where mill pricing is similar to processor collection mechanism and uniform delivery pricing is similar to supplier delivery mechanism. Like the analysis in the standard spatial monopoly pricing model, the assumption of linear supply is critical in our model. The collection boundary for processor collection should be smaller than that of supplier delivery when supply function is more convex, and vice versa. These propositions help to explain why supplier delivery dominates in the grain system. This system leads to lower transportation costs and higher overall welfare. Why then has processor collection emerged for corn stover?

2.2.2 Collection Mechanism as a Barrier to Entry

This section examines the impact of potential competition on the choice of collection strategy and on the total collection costs of the initial processor and potential entrant. Assume two homogenous processors, one an incumbent and the other a potential entrant. Both minimize total collection costs while meeting their capacity constraint. The degree of competition will be influenced by the distance between processors. If cellulosic fuel prices are

high and there are no untapped draw areas, then the second processor may choose a location that puts them within the draw area of an existing processor. If this were to occur, then it is likely the two plants would compete to attract suppliers and that the premium and participation rate would increase.

In order to characterize the competition behavior between incumbent and entrant, we use a line model where incumbent and entrant be positioned at each end. The two processors are assumed to be identical in their production technology and capacity. They can choose either processor collection or supplier delivery. Thus, the optimization problem is the same for both. Each processor's pricing and collection strategies, and therefore total collection cost will depend on the distance between them. When the distance between the two processors is greater than or equal to two times the optimal single collection distance, there is no competition and supplier delivery and processor collection results in the same collection distance and expenditure for both incumbent and entrant.

Let d be the distance between two processors, competition exists when $d < 2R^* = 2\left(\frac{2Q}{bt}\right)^{0.5}$. In this situation, the entrant chooses its optimal collection mechanism based on the incumbent's strategy, then the incumbent reacts, and the entrant reacts again and so on. The four competitive outcomes in a Nash Equilibrium framework are:

		Entrant	
		Processor Collection	Supplier Delivery
Incumbent	Processor Collection	$\left(C_I^{PP}(d), C_E^{PP}(d)\right)$	$\left(C_I^{PS}(d), C_E^{SP}(d)\right)$
	Supplier Delivery	$\left(C_I^{SP}(d), C_E^{PS}(d)\right)$	$\left(C_I^{SS}(d), C_E^{SS}(d)\right)$

where d is the distance between processors, $C_i^{jk}(d)$ denotes the total cost in equilibrium for processor i using strategy j when the other processor using strategy k. Here $i = I, E$ represents Incumbent and Entrant, respectively, and $j, k = P$, S represents processor collection and supplier delivery, respectively. The optimal collection mechanism and total collection cost for plants in each case are derived by solving a repeated game Nash Equilibrium.

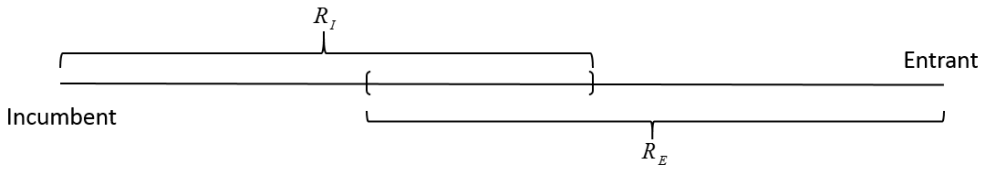
Specifically, we start with the incumbent's pricing strategy $S_I^{(1)} = (p_I^{(1)}, R_I^{(1)})$. Given this strategy, the entrant picks its best response while meeting its capacity needs at minimum cost, defined as $S_E^{(1)}(S_I^{(1)}) = (p_E^{(1)}, R_E^{(1)})$. The incumbent then has to increase its collection distance or collection price to $S_I^{(2)}(S_E^{(1)}) = (p_I^{(2)}, R_I^{(2)})$ because some suppliers located at the edge of the incumbents draw area will receive a higher net prices from entrant. The entrant has to respond to the incumbent's response which is defined as $S_E^{(2)}(S_I^{(2)}) = (p_E^{(2)}, R_E^{(2)})$. This procedure will continue until an equilibrium strategy $S_I^{(*)}(S_E^{(*)}) = (p_I^{(*)}, R_I^{(*)})$ and $S_E^{(*)}(S_I^{(*)}) = (p_E^{(*)}, R_E^{(*)})$ is achieved.

There are three different strategy combinations: (a) both incumbent and entrant use processor collection; (b) both incumbent and entrant use supplier delivery; (c) incumbent and entrant use a different mechanism.

Case 1: Both incumbent and entrant use processor collection.

Let (p_I, R_I) and (p_E, R_E) be the pricing strategy for incumbent and entrant, respectively. The resulting collection amounts are $Q_I = bp_I R_I$ and $Q_E = bp_E R_E$. Since both the incumbent and the entrant use processor collection, the one that offers the higher price will capture all suppliers in the overlapping collection area. If $p_I > p_E$, the incumbent will capture all

suppliers in the overlapping collection area and its optimal strategy should be the same as for the single processor case. That is $(p_I, R_E) = (p^*, R^*)$. Since both processors require the same amount of corn stover, $p_I > p_E$ implies that $R^* = R_I < R_E$. Thus, the distance between two processors is $d = R_I + R_E > 2R^*$. This contradicts the assumption that $d < 2R^*$ (this assumption ensures the existence of competition). On the other hand, if $p_I < p_E$, the entrant will capture all suppliers in the overlapping collection area and its optimal strategy will be the same as for the single processor case. That is $(p_E, R_E) = (p^*, R^*)$. Since both processors require the same amount of stover, $p_I < p_E$ implies that $R_I > R_E = R^*$, and the distance $d = R_I + R_E > 2R^*$. This again contradicts with the assumption that $d < 2R^*$. As a result, $p_I = p_E = p$ in equilibrium. The collection distance in this scenario can be illustrated as follows:



Therefore, the amount collected by incumbent is given as

$$Q_I = bp(d - R_E) + 0.5bp(R_I + R_E - d) = 0.5bpd + 0.5bp(R_I - R_E)$$

Similarly, the amount collected by entrant is given as

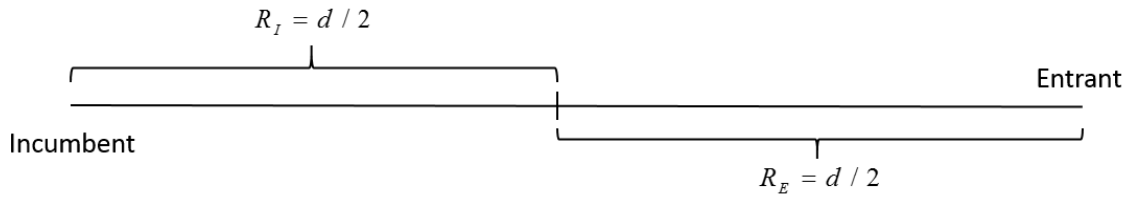
$$Q_E = bp(d - R_I) + 0.5bp(R_I + R_E - d) = 0.5bpd + 0.5bp(R_E - R_I)$$

From $Q_I = Q_E = Q$, it obtains to $R_I = R_E = R$. Thus, $Q = 0.5bpd$ implies that $p^* = \frac{2Q}{bd}$. The total cost for each processor is then expressed as

$$C_I^{PP} = \int_0^{d-R} bp(p+tr)dr + \frac{1}{2}\int_{d-R}^{R} bp(p+tr)dr$$

$$= bp^2(d-R) + \frac{bpt}{2}(d-R)^2 + \frac{1}{2}bp^2(2R-d) + \frac{1}{4}bpt(2R-d)^2.$$

Substituting $p^* = \frac{2Q}{bd}$ and minimize the total expenditure respect to R implies that $R^* = 0.5d$. Thus, each processor collects from those suppliers close to it and they both increase the collection price and participation to avoid a price war in the overlapping area. The equilibrium collection distance looks as follows



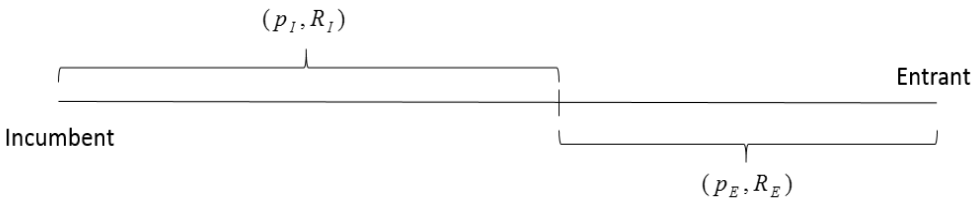
And total collection cost is given by

$$C_I^{PP} = \int_0^{d/2} bp(p+tr)dr = \frac{2Q^2}{bd} + \frac{tdQ}{4} \text{ and } C_E^{PP} = \int_0^{d/2} bp(p+tr)dr = \frac{2Q^2}{bd} + \frac{tdQ}{4}$$

Case 2: Both incumbent and entrant use supplier delivery mechanism

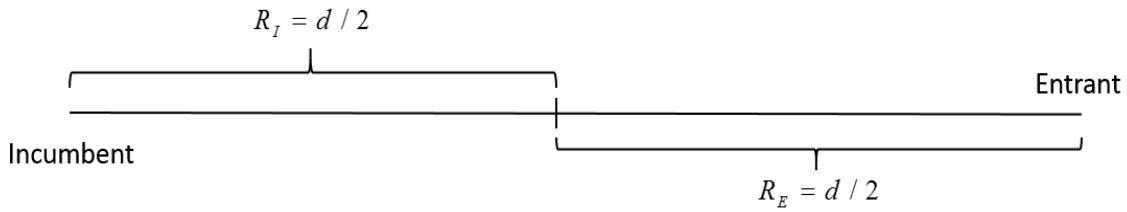
Let (p_I, R_I) and (p_E, R_E) be the pricing strategy for incumbent and entrant, respectively. If the corn stover producer locates r_I distance to incumbent and r_E distance to entrant ($r_I + r_E = d$) , he will receive $p_I - tr_I$ net price if he supplies to incumbent and receive $p_E - tr_E$

net price if he supplies to entrant. The supplier will always prefer supplying the processor from where he can receive higher net prices. As a result, the marginal supplier for both processors should be indifferent between supplying either incumbent or entrant. The collection distance for each processor can be represented in the following figure:



Since R_I and R_E are the collection distance for incumbent and entrant, respectively, the marginal supplier located R_I miles away from incumbent and R_E miles away from entrant should receive the same net prices from two processors. That is, $p_I - tR_I = p_E - tR_E$. Since $d = R_I + R_E$, it obtains the collection distance for incumbent is $R_I = \frac{d}{2} + \frac{p_I - p_E}{2t}$ and the collection distance for entrant is $R_E = \frac{d}{2} - \frac{p_I - p_E}{2t}$. Given the collection distance R_I and R_E, the amount collected by incumbent is $Q_I = \int_0^{R_I} b(p_I - tr)dr = bp_I R_I - 0.5btR_I^2$. Similarly, the amount collected by entrant is $Q_E = bp_E R_E - 0.5btR_E^2$.

Since $Q_I = Q_E = Q$, it implies that $bp_I R_I - 0.5btR_I^2 = bp_E R_E - 0.5btR_E^2$. Substituting the expressions of R_I and R_E into this equation, we get $\frac{(p_I - p_E)^2}{2t} = 0$, which implies $p_I = p_E = p$. Thus, $R_I = R_E = \frac{d}{2}$. The equilibrium collection distance looks as follows:



Substituting $R_I = R_E = \frac{d}{2}$ into the expression of Q to obtain $Q = \frac{bpd}{2} - \frac{bt}{2}\left(\frac{d}{2}\right)^2$. Thus, the price offered by processor is $p = \frac{2Q}{bd} + \frac{td}{4}$ and the total cost is:

$$C_I^{SS} = \int_0^{d/2} b(p - tr)\, pdr = \frac{2Q^2}{bd} + \frac{tdQ}{4}$$

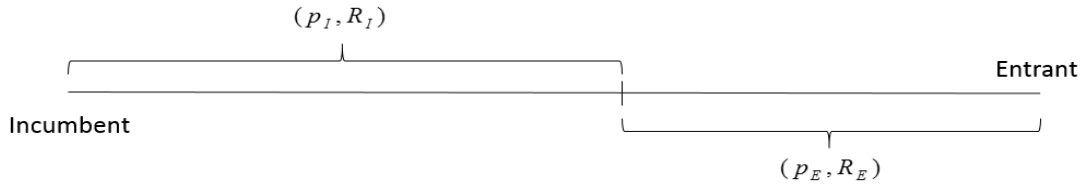
$$C_E^{SS} = \int_0^{d/2} b(p - tr)\, pdr = \frac{2Q^2}{bd} + \frac{tdQ}{4} \ .$$

This is the same as the case where both processors use processor collection. In this case, both processors increase the offer price so as to collect more from those suppliers close to them and to avoid the price war in overlapping area.

Case 3: one uses processor collection and the other uses supplier delivery

Without generality, we assume the incumbent use processor collection and entrant use supplier delivery mechanism. Let (p_I, R_I) and (p_E, R_E) be the pricing strategy for incumbent and entrant, respectively. For incumbent using processor collection mechanism, $Q_I = bp_I R_I$, which implies $p_I = \frac{Q_I}{bR_I}$. For the entrant who uses supplier delivery mechanism, the amount it collected is $Q_E = \int_0^{R_E} b(p_E - tr)dr = bp_E R_E - 0.5btR_E^2$, which implies that $P_E = \frac{Q_E + 0.5btR_E^2}{bR_E}$.

This is equivalent to $p_E = \frac{Q_E}{bR_E} + 0.5tR_E = \frac{Q_E}{b(d-R_I)} + 0.5t(d-R_I)$. Since the collection distances are R_I and R_E, the supplier located R_I miles away from processor collection plant and R_E miles away from supplier delivery plant should be indifferent between these two processors as he receives the same net prices from them. That is, $p_I = p_E - tR_E = p_E - t(d-R_I)$. Since $Q_I = Q_E = Q$, substituting p_I and p_E into this equation obtains $p_I = p_E - t(d-R_I) = \frac{Q}{b(d-R_I)} - \frac{1}{2}t(d-R_I) = \frac{Q}{bR_I}$. This relationship is equivalent to $\frac{Q}{b(d-R_I)} - \frac{Q}{bR_I} = \frac{1}{2}t(d-R_I) > 0$, which is $d - R_I < R_I$. Thus, $R_I > \frac{d}{2} > (d-R_I) > R_E$ must holds. Therefore, the collection radius for the incumbent who uses processor collection is always larger than the collection radius for entrant who uses supplier delivery.



Solving the above equation, the optimal collection distance can be expressed as function $R_I^*(d,Q)$ and $R_E^*(d,Q)$. And the prices offered to supplier in each mechanism can be expressed as $p_I^*(d,Q)$ and $p_E^*(d,Q)$. The total cost for processor collection incumbent is $C_I^{PS} = p_I^* Q + 0.5\frac{tQ^2}{bp_I^*}$. And the total cost for supplier delivery entrant is $C_E^{SP} = p_E^* Q$.

Proposition 3. The incumbent processor can use processor collection as a way to add costs to an entrant who is considering an overlapping collection area. In this case, both incumbent and

entrant will use processor collection to lower total expenditure when there is competition for suppliers.

Proof

We first show both incumbent and entrant will use processor collection mechanism under competition. Let R_p be the collection distance for processor using processor collection mechanism and R_s be the collection distance for processor using supplier delivery mechanism. As shown above, the total cost for processor using supplier delivery is a function of collection distance R_s,which can be expressed as $f(R_s) = p_s Q = \frac{Q^2}{bR_s} + \frac{tR_s Q}{2}$, where p_s is the price offered by processor using supplier delivery mechanism. The F.O.C of this equation is $f'(R_s) = -\frac{Q^2}{bR_s^2} + \frac{tQ}{2} = 0$, which implies $R_s = \pm\left(\frac{2Q}{bt}\right)^{0.5} = \pm R^*$. When $0 < R_s < R^*$, $f'(R_s) = -\frac{Q^2}{bR_s^2} + \frac{tQ}{2} < 0$, indicating shorter collection distance results in higher collection expenditure.

When the incumbent uses processor collection mechanism, the total cost for the entrant is $\frac{2Q^2}{bd} + \frac{tdQ}{4}$ if it also uses processor collection. This expenditure is the same as $f(\frac{d}{2})$. If the entrant uses supplier delivery, its total expenditure is $f(R_s^*) = \frac{Q^2}{bR_s^*} + \frac{tR_s^* Q}{2}$. Notice that the collection radius for processor using processor collection is always larger than the collection radius for processor using supplier delivery which implies $R_s^* < \frac{d}{2}$. Therefore, it must be true that $f(R_s^*) = \frac{Q^2}{bR_s^*} + \frac{tR_s^* Q}{2} > f(\frac{d}{2}) = \frac{2Q^2}{bd} + \frac{tdQ}{4}$. As a result, the entrant will always use

processor collection to lower it collection expenditure when the incumbent uses processor collection.

When the incumbent uses supplier delivery, the entrant's total collection expenditure is $\frac{2Q^2}{bd}+\frac{tdQ}{4}$ if it also use supplier delivery. If the entrant use processor collection, the expenditure is $p^*Q+0.5\frac{tQ^2}{bp^*}=\frac{Q^2}{bR_p^*}+\frac{tR_p^*Q}{2}=g(R_p)$. Similar to function $f(R_s^*)$, function $g(R_p)$ is decreasing in $(0,R^*)$. Since the collection radius for processor using collection processor is always larger than the collection radius for the processor using supplier delivery, $R^*\geq R_p^*>\frac{d}{2}$ holds. Thus, $g(R_p^*)=\frac{Q^2}{bR_p^*}+\frac{tR_p^*Q}{2}<\frac{2Q^2}{bd}+\frac{tdQ}{4}=g(\frac{d}{2})$. Thus, the entrant will always choose processor collection to lower its total collection expenditure when the incumbent uses supplier delivery.

In general, the best strategy for entrant is to use processor collection, no matter what strategy the incumbent uses. Given entrant's best response, the incumbent will always use processor collection. That is because the total cost under processor collection, $\frac{2Q^2}{bd}+\frac{tdQ}{4}$, is lower than the cost supplier delivery mechanism, $\frac{Q^2}{bR_s^*}+\frac{tR_s^*Q}{2}$, when $R_s^*<\frac{d}{2}$ holds. On the other hand, incumbent processor can use processor collection as a way to add costs to an entrant who is considering an overlapping collection area. By using the processor collection mechanism, the incumbent processor can increase the cost of entrant from a lower level to $\frac{2Q^2}{bd}+\frac{tdQ}{4}$.**Q.E.D.**

Proposition 3 may explain why has processor collection has emerged for corn stover. It also shows that the incumbent can increase the entrant's collection cost using processor collection. This increase in collection cost can be treated as the penalty for moving into the draw area of the incumbent. This penalty increases as the two plants get closer. The extra cost to the entrant occurs because both plants receive only one half of the stover sold in the area where they compete. As a result, both plants must increase premium in areas to attract suppliers.

2.3 From Theoretical Analysis to Practical Simulation

Realistically, the collection region for a processor will better resemble a circle than a line. However, the line model can be generalized to a circle model by thinking of the line as a radius connecting the center of the circle to a point on the circumference. Integrating the line from 0 to 2π, the quantity collected in the circle model can be obtained. Since the expenditure and collection distance are the same for processor collection and supplier delivery mechanism in the line model, it generalizes that the optimal collection radius (collection region) and collection expenditure are the same for processor collection and supplier delivery mechanism in the circle model as well.

A Numerical Example

Table 2.1 compares the optimal collection costs per ton across three collection mechanisms in a circle model and under different assumptions about the feedstock supplier's response to price change. These results are based on a transportation cost of $0.65 per ton per mile, a plant requirement for 300,000 MT/year of feedstock, and a 2 MT/acre corn stover removal rate—all of these values are based on Darr et al. (2013) which in turn are based on the

Du Pont plant located in central Iowa where corn is planted on one third of the total area (USDA NASS 2015).

Table 2.1: Comparison between different collection mechanisms

Comparison Between Different Collection Mechanism									
Response to price change	Perfect Price Discrimination			Processor Collection			Supplier Delivery		
	Radius	% of supply	Price	Radius	% of supply	Price	Radius	% of supply	Price
0.005	43.4	12.0%	24.0	54.7	7.5%	15.1	54.7	22.6%	45.3
0.0075	37.9	15.7%	20.9	47.8	9.9%	13.2	47.8	29.7%	39.6
0.01	34.5	19.0%	19.0	43.4	12.0%	12.0	43.4	35.9%	35.9
0.015	30.1	24.9%	16.6	37.9	15.8%	10.5	37.9	47.1%	31.4
0.02	27.3	30.2%	15.1	34.5	19.0%	9.5	34.5	57.1%	28.5
0.03	23.9	39.6%	13.2	30.1	24.9%	8.3	30.1	74.8%	24.9
0.04	21.7	47.9%	12.0	27.4	30.2%	7.5	27.4	90.6%	22.6
0.05	20.2	55.6%	11.1	25.4	35.0%	7.0	25.4	100%	21.0

Notes: 1. Collection radius is 15 miles when participation rate is 100%.

2. Participation rate for supplier delivery indicates the maximum value in the collection area.

The simulation results show that optimal collection price and collection radius are decreasing with the increase in suppliers' responses to price change. This table also provides evidence that the optimal collection radius between the processor collection and supplier

delivery are the same. Giving a 15% of land with this feedstock will participate and assuming that the processor is optimizing and knows the response rate, our estimation for parameter b is 0.015 and the optimal collection radius is 38 miles. The collection radius under 100% participation is 15 miles.

Given the deterrent effect of processor collection and a preference on behalf of both parties to avoid bidding up the price paid for stover, it can be assumed that the entrant will select a plant location that is close to, but not overlapping with the incumbent. Once all suitable locations have been used, any additional expansion will come from processors who increase capacity and supplier participation within their original draw area. In fact, the cost increase for entrant when the incumbent use processor collection is represented by

$$\frac{2Q^2}{bd}+\frac{tdQ}{4}-2\left(\frac{t}{2b}\right)^{1/2}Q^{3/2}=\left(\sqrt{\frac{2Q^2}{bd}}-\sqrt{\frac{tdQ}{4}}\right)^2.$$

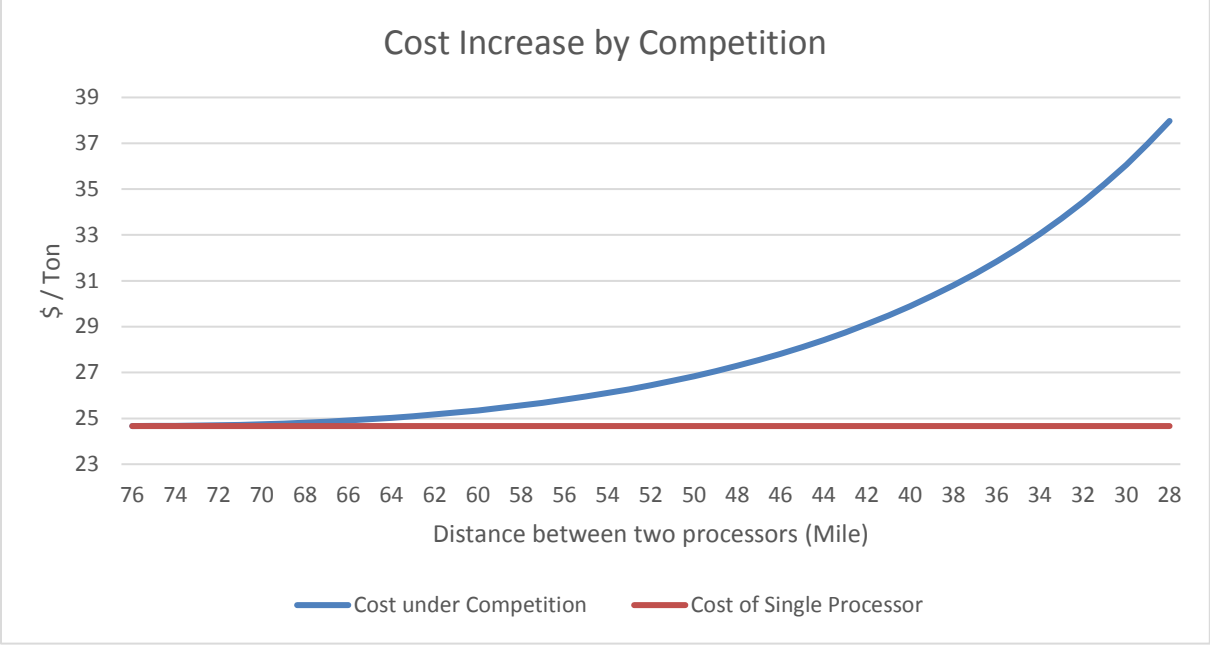


Figure 2.1: Collection cost increase as the distance between two processors change

Given the same parameter used in simulation above, figure 2.1 below illustrates how the collection cost change as the distance between two processors change. It demonstrates a significant cost increase when competition exists.

Marginal Cost for a Processor Expanding within an Existing Draw Area

As the mandate incentivizes additional production, competition for stover should ensure that all suitable locations are used. Further increases in ethanol produced from stover will therefore arise only if existing firms increase production from within the original draw area.

Figure 2.2 illustrates how costs increase as existing plants increase capacity, prices, and supplier participation within their original draw area. A doubling of capacity from 30 million gallons to 60 million gallons increases costs from $41.84 to $62.71—a 50% increase. Ogden and Anderson (2011) show that the marginal cost of ethanol made from agricultural residue rises from $2.80 per gallon of gasoline equivalent to $3.00 per gallon, at which time the 16 billion mandate (equal to 10.5 billion gallons of gasoline equivalent) is met. Expanding beyond the mandate, Ogden and Anderson (2011) also show that perennial grasses enter the mix at a price range of $3.30 to $3.80 and that pulpwood enters at $3.60. At $3.50 per gallon, production doubles from approximate 10 billion gallons of gasoline equivalent to 20 billion gallons of gasoline equivalent. The 16.6% increase from the $3 to the $3.50 double production using other feedstocks is much lower than the 50% increase in costs for monopsonistic stover processors who double production. This suggests that these alternative feedstocks will be used.

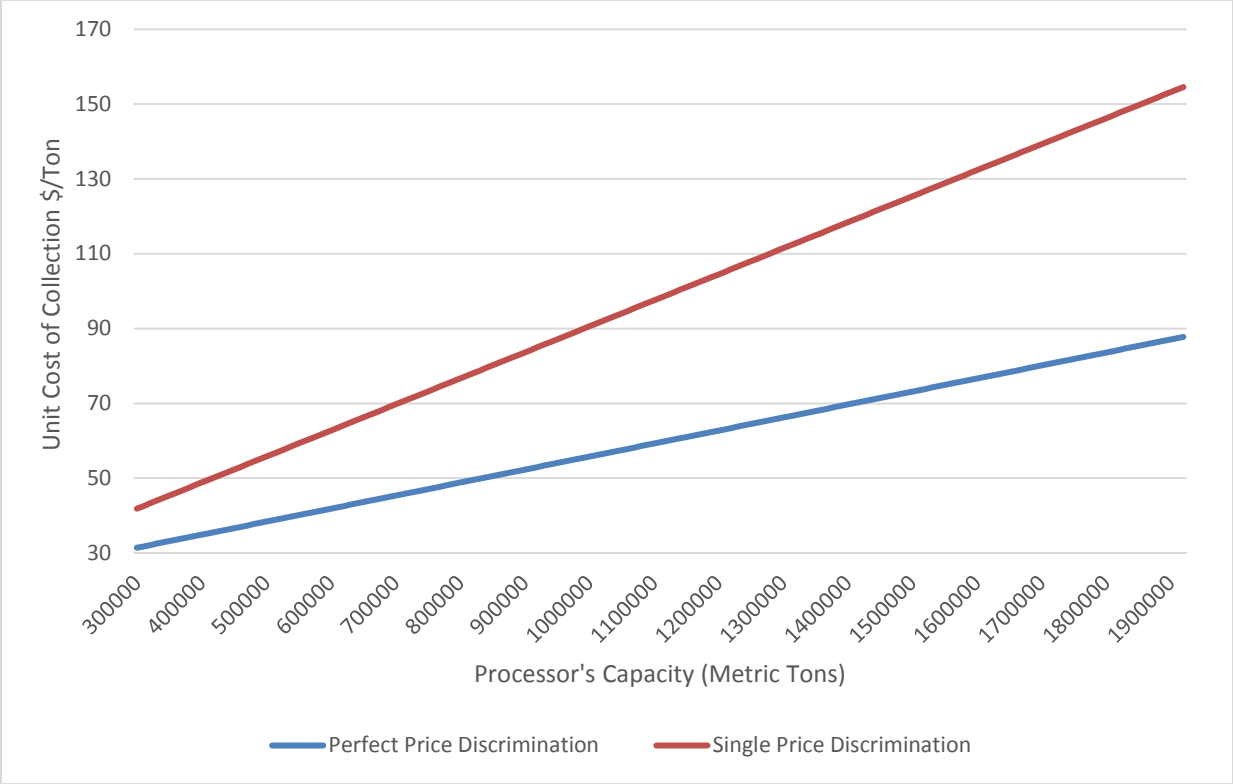


Figure 2.2: Marginal cost when the collection radius is fixed

2.4 Conclusion

The production of cellulosic energy involves bulky raw material, capital-intensive processing plants, and feedstock producers who are heterogeneous with respect to the price at which they are willing to sell raw material. These circumstances are very different to those that existed as the corn ethanol industry emerged. The first two commercial scale pilot plants have both shown a willingness to accept very large draw areas and low feedstock producer participation. They have also introduced a collection mechanism that relies on the use of plant-owned equipment, this runs counter to the mechanism used in the grain industry where grain producers deliver to plants. Results presented in this article show that plants are behaving rationally given their monopsonistic status. Results also show that the use of a processor

collection mechanism is a deterrent to new plants who are considering siting a plant with a draw area that overlaps the incumbent plant. The intuition behind this is that under supplier delivery, feedstock producers at the edge of the draw area for plants are easily poached. Under processor collection, these producers can only be poached if the entrant enters into a costly price war with the incumbent. If entrants locate away from the draw area for incumbent plants, then these plants will retain their monopsonistic status and will be able to expand capacity only if they increase feedstock prices so as to expand participation among feedstock suppliers. Taken together these results suggest that it will not be possible to meet the cellulosic mandate with stover alone and that other feedstock sources will be required.

CHAPTER 3. THE EXAMINATION OF MARKET POWER OF U.S. NITROGEN FERTILIZER INDUSTRY: A BAYESIAN BASED APPROACH

Abstract

This paper investigates a change in the market power of the U.S. nitrogen fertilizer industry by examining the causal linkage between fertilizer, its main feedstock (natural gas), and output (corn) price from 1999 to 2011. A time-varying parameter model has been established in our analysis and estimation is done by a Bayesian-based Kalman filter algorithm. We also utilize a single-equation error correction model to determine if there is a long-run equilibrium price relationship when the co-integrating vector is no longer constant over time. The results of the time-varying estimation show that the U.S. nitrogen fertilizer price follows the value of its marginal productivity closer than its marginal cost of production, indicating a less competitive market structure. The estimation from the error correction model supports these results.

3.1 Introduction

Nitrogen, phosphate, and potash play important roles in the ability of crops to develop proteins and enzymes, which in turn, help improve crop yields. Commercial fertilizer consumption increased rapidly before 1980 as more acreage was devoted to high-yield crop varieties and hybrids that responded favorably to more intensive fertilizer use. Since the mid-1980's, the consumption of phosphate and potash remains stable, while the consumption of nitrogen fertilizer has increased more rapidly due to the development of seed varieties with

more favorable yield responses to nitrogenous fertilizers. This increasing annual usage shows the importance of nitrogen fertilizer to U.S. agricultural production. In fact, U.S. farmers are moving away from using multiple-nutrient fertilizers toward using single-nutrient fertilizers or fertilizers with a high level of nutrient concentration. This is because single nutrient fertilizers with high nutrient concentration allow farmers to apply precise amounts of a specific nutrient for plant use at the least cost. Figure 3.1 shows that the annual usage for all types of nitrogen fertilizers has nearly quadrupled from 2.73 million tons to 12.84 million tons from 1960 to 2011; whereas the annual usage of phosphate and potash has only increased roughly two-fold during the same period.

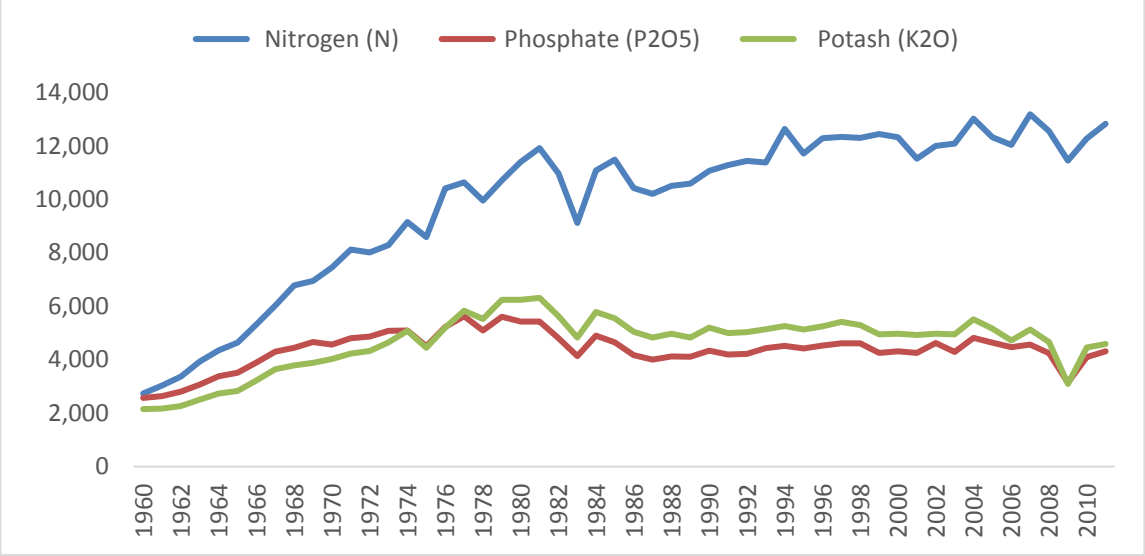


Figure 3.1: U.S. consumption of nitrogen, phosphate, and potash 1960–2011.

In 2011, nitrogen fertilizer accounted for 59% of total U.S. agricultural nutrient usage. Thus, research on price behaviors and the market structure of U.S. nitrogen fertilizer industry is pertinent; and yet, literature on this field is very limited.

Among all nitrogen fertilizer usage, corn production accounted for the largest share at over 45%. Figure 3.2 illustrates nitrogen use by main crops in the United States from 1964 to 2010, and indicates increasing nitrogen use by corn and wheat during that time.

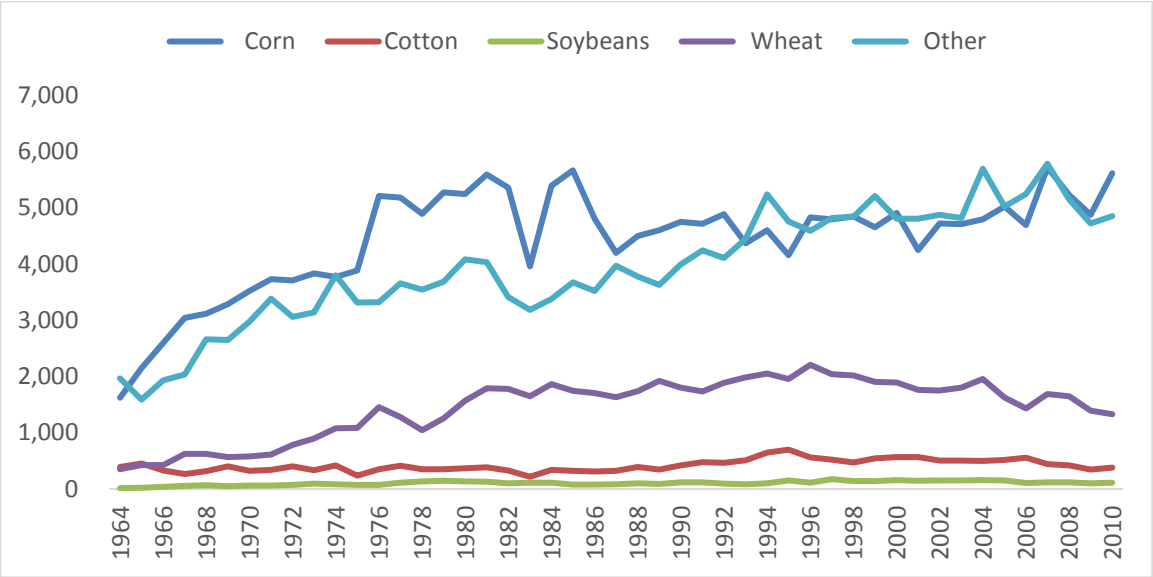


Figure 3.2: U.S. plant nitrogen use by corn, soybeans, cotton, and wheat, 1964–2010

In this paper, we are interested in investigating the market competitiveness of nitrogen fertilizer industry. In a market that is uncompetitive, sellers are able to price their product depending on the buyer's ability to pay. Thus, if the nitrogen fertilizer industry is uncompetitive, nitrogen fertilizer prices quoted by producers will be closely correlated with crop price—the key factor of a farmer's ability to pay. In competitive markets such as grains and meat, output prices tend to revert to production costs in the long run. Under perfect competition circumstances, firms face perfectly elastic demand and have no power to increase prices higher than the industry marginal cost.

Among all inputs used to produce nitrogen fertilizer, such as anhydrous ammonia and urea, natural gas is the single most important and accounts for the largest share of the cost structure. Depending on plant technology, 75% of total urea production cost is from natural gas, of which 25% is through direct use in the production process and 50% is through the use of ammonia as a feedstock, of which natural gas is also the most important cost component. This cost structure implies that, under a competitive market and with all else being equal, urea price changes should follow closely to natural gas price changes. However, figure 3.3 shows that natural gas prices have fallen to historically low levels since early 2009, due to the latest discovery of shell-rock natural gas reserves in the United States and improved technology to ensure a sustainable supply, while urea prices have increased after a sharp decrease in 2008. In fact, urea price follows closer to corn prices, both of which have increased in recent years.

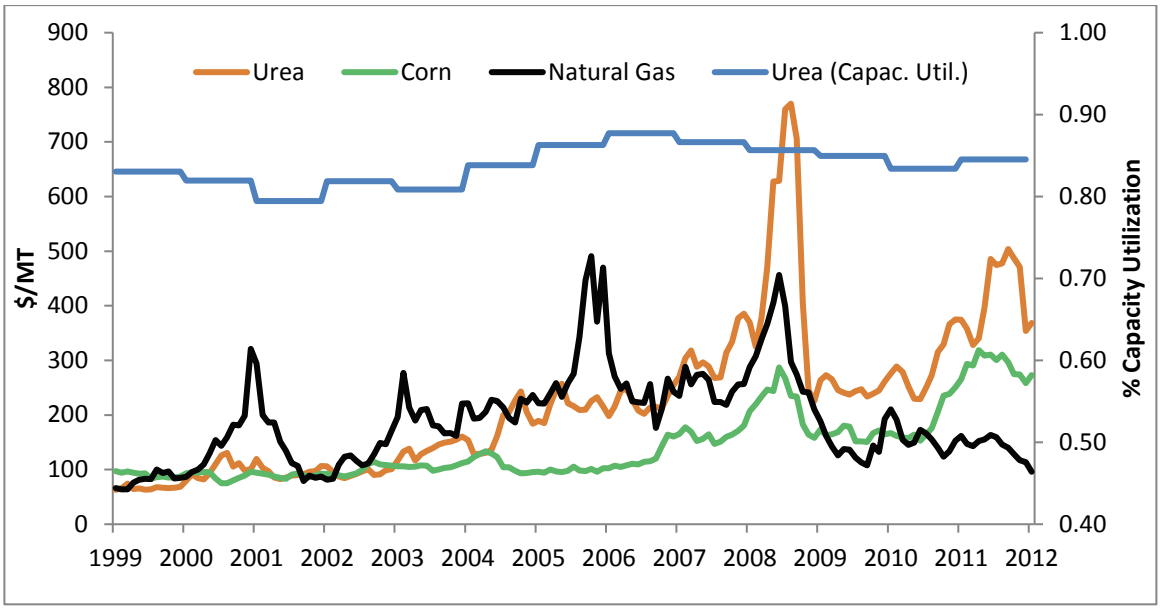


Figure 3.3: Monthly price of urea, natural gas, corn and % of capacity utilization

Another way to look at the problem is though the correlation coefficients between the mentioned price variables. We analyze the period 1999-2011 in two sub-periods, one from 1999 to 2005 and the other from 2006 to 2011 (Galbraith 2010). As table 3.1 shows, in the first period we observe a strong correlation between urea and natural gas prices and a weak correlation between urea and corn. The picture is exactly the opposite between 2006 and 2011; a very weak correlation of urea with respect to its own and main feedstock, but strong with one of its main demand factors.

Table 3.1: Correlation coefficients between urea, natural gas and corn

Period 1999-2005				**Period 2006-2011**			
	Natural Gas	Corn	Urea		Natural Gas	Corn	Urea
Natural Gas	1.0			Natural Gas	1.0		
Corn	0.3	1.0		Corn	-0.1	1.0	
Urea	0.8	0.3	1.0	Urea	0.4	0.6	1.0

Natural gas is also the main feedstock of ammonia and it accounts for about 85% of the total production cost. In this regard, ammonia price should follow natural gas even closer than urea. Yet, figure 3.4 shows that since 2009, ammonia price has increased and natural gas prices have decreased. The increasing corn price in the same time indicates that ammonia price is closer to corn price (a factor determine farmer's ability to pay) than to its main feedstock price.

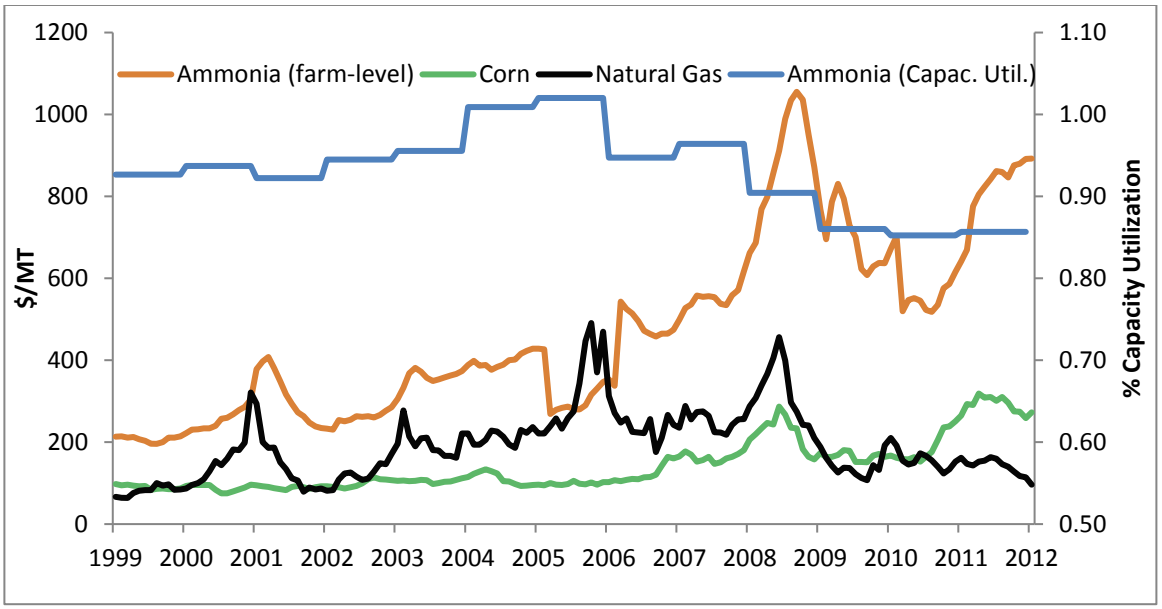


Figure 3.4: Monthly price of ammonia, natural gas, corn and % of capacity utilization

Table 3.2 compares the correlation coefficients between these commodities. There is positive correlation of 0.5 between natural gas and ammonia prices between 1999 and 2005, similar to that with respect to corn prices. However, from 2006 to 2011 we observe no correlation between ammonia and natural gas, but an increased correlation of 0.7 with corn.

Table 3.2: Correlation coefficients between ammonia, natural gas and corn

Period 1999-2005				**Period 2006-2011**			
	Natural Gas	Corn	Ammonia		Natural Gas	Corn	Ammonia
Natural Gas	1.0			Natural Gas	1.0		
Corn	0.3	1.0		Corn	-0.1	1.0	
Ammonia	0.5	0.5	1.0	Ammonia	0.0	0.7	1.0

One of the key purposes of this study is to analyze the competitiveness of the U.S. nitrogen fertilizer industry to determine whether the prices of nitrogen fertilizer, such as anhydrous ammonia and urea, follow the marginal cost of production (the competitive case) or the value of the marginal productivity of fertilizers in agriculture (the non-competitive case). The co-integration and vector error correction model (VECM) introduced by Granger (1981) and Engle and Granger (1987) has been the most widely used methodology to analyze long-run price causality. Our results, however, demonstrate an unstable causal relationship between nitrogen fertilizer price and its feedstock (natural gas) or output (corn) price over the sample period. This non-constancy of causality makes the application of standard Granger causality test inappropriate. In this paper, we investigate the time-varying effects of natural gas prices and corn prices on U.S. nitrogen fertilizer prices from 1999 to 2011 by using a Bayesian-based time-varying parameter approach.

In understanding the market power of the U.S. nitrogen fertilizer industry, the role played by capacity utilization cannot be overlooked. The capacity constraint on nitrogen production is a significant factor that affects the price-cost margin (a fundamental way to characterize market power of an industry) of the nitrogen industry in the short term. Nitrogen demand is quite inelastic since farmers must use a certain amount to enhance crop yields. Thus, even a small firm with available production capacity can earn great market power if market demand is greater than industry capacity and supplies of other firms are close to their individual capacity. However, the market power caused by this capacity constraint cannot be maintained with the capacity expansion in the medium term and long term.

We observe in figure 3.3 that, in the case of urea, capacity utilization oscillates around 83% from 1999 to 2005, and has decreased since 2006. Figure 3.4 shows that ammonia

capacity utilization has also decreased steadily since 2006. These observations in decreasing capacity utilization might be an indication of increasing market power in the U.S. fertilizer industry since a non-competitive industry has the incentive to exert its market power and underutilize its capacity in order to increase output price. Thus, the second contribution of this paper is to take into account the industry capacity utilization into an original price formulation model to analyze how the interrelation between different factors affect nitrogen fertilizer prices.

The last contribution of this paper is to utilize a single-equation error correction model to analyze both long-run and short-run adjustments to equilibrium relationship by constructing an error correction term in which a co-integrating vector is no longer constant as in the standard model, but instead time-varying estimates.

The paper proceeds as follows. Section 3.2 summarizes the literature review. Section 3.3 explains the methodology employed. Section 3.4 provides an empirical analysis, including data description and empirical results. Section 3.4 also exhibits the time-varying relationship between different nitrogen feedstock and product prices and fertilizer prices via a Bayesian-based Kalman filter and the estimation of a single-equation error correction model. Section 3.5 offers concluding remarks.

3.2 Related Literature

There have been very limited academic studies on the price behavior and market power of the U.S. nitrogen fertilizer industry. Huang (2007) analyzes the impacts of rising natural gas prices on U.S. ammonia price and supply and finds that further increases in natural gas prices in the United States would result in a continuous decrease in U.S. aggregate ammonia supply

that could make U.S. ammonia markets more vulnerable. Galbraith (2010) compares price variations of urea and ammonia in North American locations to that of natural gas and corn using an error correction model in two different periods: 2002–2005 and 2006–2009. He finds that fertilizer prices respond more to natural gas prices in the first period and corn prices in the second period. The main pitfall of this analysis is that it tested each factor individually, not accounting for the correlation between them. Humber (2014) utilizes a structural vector autoregressive (SVAR) model to determine the impact the 2010 merger between CF industries and Terra industries had on fertilizer prices. The counterfactual fertilizer prices generated suggest a 75% increase in fertilizer prices caused by the merger.

Carter and Kohn (1994) provide an algorithm to carry out Bayesian inference on a linear-state space model, in which states are generated efficiently using the Kalman filter. Frühwirth-Schnatter (1994) suggest a data augmentation algorithm to approximate posterior distribution and model likelihoods for a dynamic linear model. Jong and Shephard (1995) introduce a simulation smother, which draws from the multivariate posterior distribution of the disturbance of the model. Durbin and Koopman (2002) present a simpler and more efficient simulation smoother relative to that of Jong and Shephard (1995), in which only mean corrections for unconditional vectors are required. Petris et al. (2009) formally introduce the Bayesian-based Kalman filter, the state space model and dynamic linear model, as well as the procedure for estimating and forecasting. Koop et al. (2011) introduces a Markov Chain Monte Carlo (MCMC) algorithm that allows Bayesian inference in a time-varying co-integration model and combines the simulation smoother for state-space time-series models and the Gibbs sampling method for time-invariant VECM.

Empirically, both error correction models and time-varying parameter (TVP) models have been increasingly used in many fields of study. Hung-Gay Fung et al. (2013) use the daily data of 16 commodity futures contracts traded in China and the corresponding foreign markets to analyze price linkage between markets. Their results show that no significant causal relation was found in most of futures pairs. Arslanturk et al. (2011) use the rolling window and time-varying parameter estimation methods to analyze the Granger causality between tourism receipts and economic growth in a small open economy. Park et al. (2010) estimateU.S. gasoline demand from 1976 to 2008 using a time-varying co-integrating regression method. Balcilar et al. (2015) analyze the time-varying causality between spot and futures crude oil prices via a Markov-switching vector-error correction model, and find that the lead-lag relationship between the spot and futures oil markets existed only temporarily.

3.3 Methodology

3.3.1 Model Specification

The standard invariant-parameter Granger causality test in section 4 demonstrates that nitrogen fertilizer prices are co-integrated with natural gas prices in the pre-2006 sub-period; however, they are co-integrated with corn prices in the post-2006 sub-period. There is no Granger causal relationship between fertilizer prices and corn price or natural gas price in the full sample range. This implies an unstable causal relationship over the sample period, and the examination of time-varying relationship between fertilizer, corn, and natural gas prices is of interest.

Consider a model in which nitrogen fertilizer price is determined by both demand side force and supply side force for the given capacity utilization:

$$\text{price}=f(\text{demand, supply, capacity}) \tag{3.1}$$

Natural gas accounts for 75% and 85% of total production cost of urea and ammonia, respectively; thus, its price would be one of the most influential factors that affects nitrogen fertilizer price. Accordingly, natural gas prices are selected as the proxy variable for demand side force of the price system. Corn price is selected as the proxy variable for the supply side force of the price system since corn accounted for the largest share of nitrogen use among crops.

Consider a regression model in which coefficients are time varying rather than fixed:

$$Fertilizer_t = \beta_t^{(1)} + \beta_t^{(2)} Corn_t + \beta_t^{(3)} Gas_t + \beta_t^{(4)} Capacity_t + \varepsilon_t,\ \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{3.2}$$

where $Fertilizer_t$, $Corn_t$ and Gas_t are the monthly prices of nitrogen fertilizer, corn, and natural gas after a natural logarithm, respectively, and $Capacity_t$ is the fertilizer capacity utilization rate at time t.

$\beta_t^{(1)}, \beta_t^{(2)}, \beta_t^{(3)}$ and $\beta_t^{(4)}$ are time-varying parameters treated as stochastic state variables, for which the transition equation follows random walk with trends:

$$\begin{pmatrix} \beta_t^{(1)} \\ \beta_t^{(2)} \\ \beta_t^{(3)} \\ \beta_t^{(4)} \end{pmatrix} = \begin{pmatrix} \alpha_{t-1}^{(1)} \\ \alpha_{t-1}^{(2)} \\ \alpha_{t-1}^{(3)} \\ \alpha_{t-1}^{(4)} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_{t-1}^{(1)} \\ \beta_{t-1}^{(2)} \\ \beta_{t-1}^{(3)} \\ \beta_{t-1}^{(4)} \end{pmatrix} + \begin{pmatrix} v_t^{(1)} \\ v_t^{(2)} \\ v_t^{(3)} \\ v_t^{(4)} \end{pmatrix} \tag{3.3}$$

where $\alpha_t^{(1)}, \alpha_t^{(2)}, \alpha_t^{(3)}$ and $\alpha_t^{(4)}$ are possible time trends, which satisfies the random walk assumption:

$$\begin{pmatrix} \alpha_t^{(1)} \\ \alpha_t^{(2)} \\ \alpha_t^{(3)} \\ \alpha_t^{(4)} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_{t-1}^{(1)} \\ \alpha_{t-1}^{(2)} \\ \alpha_{t-1}^{(3)} \\ \alpha_{t-1}^{(4)} \end{pmatrix} + \begin{pmatrix} w_t^{(1)} \\ w_t^{(2)} \\ w_t^{(3)} \\ w_t^{(4)} \end{pmatrix} \tag{3.4}$$

All random error terms are assumed to be identically and independently normal distributed: $v_t^{(i)} \sim N(0, \sigma_{i,v}^2)$ for $i = 1,2,3,4$ and $w_t^{(i)} \sim N(0, \sigma_{i,w}^2)$ for $i = 1,2,3,4$. This I.I.D. random walk assumption is flexible and can be modified to impose more restrictions on the structure of the time variation. The time-varying parameter model characterized by equations (3.1) and (3.2) is in the form of dynamic linear model and can be estimated recursively from the updated information available at each time point t, by using a Bayesian–based Kalman filter algorithm.

3.3.2 Estimation Method

In order to estimate the time-varying parameters, we apply a Markov Chain Monte Carlo (MCMC) method to sequentially obtain samples of parameters from undated posterior conditional distribution by Gibbs sampling. The model above can be rewritten as a dynamic linear model specified by a normal prior distribution of the 4- dimensional state vector at time $t = 0$,

$$\beta_0 \sim N_4(m_0, C_0)$$

together with a pair of equations for each time $t \geq 1$,

$$\begin{aligned} y_t &= X_t \beta_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \\ \beta_t &= \alpha_{t-1} + \beta_{t-1} + v_t, \quad v_t \sim N(0, diag(\sigma_{1,v}^2, \sigma_{2,v}^2, \sigma_{3,v}^2, \sigma_{4,v}^2)) \\ \alpha_t &= \alpha_{t-1} + w_t, \quad w_t \sim N(0, diag(\sigma_{1,w}^2, \sigma_{2,w}^2, \sigma_{3,w}^2, \sigma_{4,w}^2)) \end{aligned}$$

where y_t is the dependent variable, nitrogen fertilizer price. X_t is the vector that contains all explanatory variables, β_t is a time-varying state vector with time trend α_t in its dynamic . We consider the commonly used inverse-gamma conjugate priors for the unknown variance of random errors. More specifically, we assume the inverse of the variances σ_ε^2, $\sigma_{1,v}^2, \sigma_{2,v}^2, \sigma_{3,v}^2$, $\sigma_{4,v}^2$ and $\sigma_{1,w}^2, \sigma_{2,w}^2, \sigma_{3,w}^2$, $\sigma_{4,w}^2$ have independent gamma prior with mean a and variance b. Given the observations $y_{1:T}$ and $X_{1:T}$, the conjugate posterior conditional distribution of unknown variances are given as:

$$(\sigma_\varepsilon^2)^{-1} \mid y_{1:T}, X_{1:T} \sim gamma(\frac{a^2}{b}+\frac{T}{2}, \frac{a}{b}+\frac{1}{2}\sum_{t=1}^{T}(y_t - X_t\beta_t)^2),$$

$$(\sigma_{1,v}^2)^{-1} \mid y_{1:T}, X_{1:T} \sim gamma(\frac{a^2}{b}+\frac{T}{2}, \frac{a}{b}+\frac{1}{2}\sum_{t=1}^{T}(\beta_t^{(1)} - \alpha_{t-1}^{(1)} - \beta_{t-1}^{(1)})^2),$$

$$(\sigma_{2,v}^2)^{-1} \mid y_{1:T}, X_{1:T} \sim gamma(\frac{a^2}{b}+\frac{T}{2}, \frac{a}{b}+\frac{1}{2}\sum_{t=1}^{T}(\beta_t^{(2)} - \alpha_{t-1}^{(2)} - \beta_{t-1}^{(2)})^2),$$

$$(\sigma_{3,v}^2)^{-1} \mid y_{1:T}, X_{1:T} \sim gamma(\frac{a^2}{b}+\frac{T}{2}, \frac{a}{b}+\frac{1}{2}\sum_{t=1}^{T}(\beta_t^{(3)} - \alpha_{t-1}^{(3)} - \beta_{t-1}^{(3)})^2),$$

$$(\sigma_{4,v}^2)^{-1} \mid y_{1:T}, X_{1:T} \sim gamma(\frac{a^2}{b}+\frac{T}{2}, \frac{a}{b}+\frac{1}{2}\sum_{t=1}^{T}(\beta_t^{(4)} - \alpha_{t-1}^{(4)} - \beta_{t-1}^{(4)})^2),$$

$$(\sigma_{1,w}^2)^{-1} \mid y_{1:T}, X_{1:T} \sim gamma(\frac{a^2}{b}+\frac{T}{2}, \frac{a}{b}+\frac{1}{2}\sum_{t=1}^{T}(\alpha_t^{(1)} - \alpha_{t-1}^{(1)})^2),$$

$$(\sigma_{2,w}^2)^{-1} \mid y_{1:T}, X_{1:T} \sim gamma(\frac{a^2}{b}+\frac{T}{2}, \frac{a}{b}+\frac{1}{2}\sum_{t=1}^{T}(\alpha_t^{(2)} - \alpha_{t-1}^{(2)})^2),$$

$$(\sigma_{3,w}^2)^{-1} \mid y_{1:T}, X_{1:T} \sim gamma(\frac{a^2}{b}+\frac{T}{2}, \frac{a}{b}+\frac{1}{2}\sum_{t=1}^{T}(\alpha_t^{(3)} - \alpha_{t-1}^{(3)})^2),$$

$$(\sigma_{4,w}^2)^{-1} \mid y_{1:T}, X_{1:T} \sim gamma(\frac{a^2}{b}+\frac{T}{2}, \frac{a}{b}+\frac{1}{2}\sum_{t=1}^{T}(\alpha_t^{(4)} - \alpha_{t-1}^{(4)})^2).$$

Letting $a = b = 1$ and using the posterior conditional distribution of random errors, the forward filtering backward sampling (FFBS) procedure can be implemented by the following steps:

1. Choose the initial value of the mean m_0 and distance from C_0 for the normal prior distribution of state vector at initiation time.

2. Draw initial value of $\sigma_{\varepsilon}^2, \sigma_{1,v}^2, \sigma_{2,v}^2, \sigma_{3,v}^2, \sigma_{4,v}^2$ and $\sigma_{1,w}^2, \sigma_{2,w}^2, \sigma_{3,w}^2, \sigma_{4,w}^2$ from their independent and identical inverse gamma priors.

3. Run Kalman filter and get sample of state variables via FFBS algorithm.

4. Update information of state variables obtained in step 3 and draw samples of $\sigma_{\varepsilon}^2, \sigma_{1,v}^2, \sigma_{2,v}^2$, $\sigma_{3,v}^2, \sigma_{4,v}^2$ and $\sigma_{1,w}^2, \sigma_{2,w}^2, \sigma_{3,w}^2, \sigma_{4,w}^2$ from their corresponding posterior inverse gamma distribution equation.

5. Substitute the sample obtained in step 4 into step 2 as the initial values for $\sigma_{\varepsilon}^2, \sigma_{1,v}^2, \sigma_{2,v}^2$, $\sigma_{3,v}^2, \sigma_{4,v}^2$ and $\sigma_{1,w}^2, \sigma_{2,w}^2, \sigma_{3,w}^2, \sigma_{4,w}^2$.

6. Repeat steps 1 through 5 for M times, obtaining a bootstrapped set of parameter estimates $\left\{(\hat{\alpha}_t^m, \hat{\beta}_t^m) : m = 1, ..., M\right\}$.

This parameter set enables us to construct a confidence interval for estimates and to analyze the interrelationship between prices. We run the algorithms a total of 40,000 times and discard the first 20,000 samples as the burn period, then use the subsequent sample to make a Bayesian inference about the parameters.

3.4 Empirical Analysis

3.4.1 Description of Data

The empirical analysis of this study is based on data for the period from January 1999 to December 2011 for a total of 156 observations. The Agricultural Marketing Service (AMS) reports farm-level average monthly retail prices of urea and anhydrous ammonia. Monthly natural gas spot price comes from the Henry Hub terminal in Louisiana, and standardized as

U.S. dollars per thousand cubic meters of gas. Corn price is obtained from Economic Research Service (ERS). All prices are converted to natural logarithms. Capacity utilization is calculated as the ratio between the fertilizer supply and the sum of individual plant capacity of all plants in the North America. Data for the supply of fertilizer products and plant capacity are obtained from International Fertilizer Industry Association (IFA) and International Fertilizer Development Center (IFDA), respectively.

To investigate the time-varying relationship between nitrogen fertilizer price and its main feedstock and product prices, we first perform the Augmented Dickey-Fuller (ADF) and Phillip-Perron (PP) unit root tests to all price series in different sample periods in order to determine whether they are stationary or not. The results in table 3.3 indicate that all price series are non-stationary in levels, but stationary for the first difference.

Table 3.3: ADF and PP unit root test

	99-Jan to 05-Dec		06-Jan to 11-Dec		09-Jan to 11-Dec	
	ADF Test	PP Test	ADF Test	PP Test	ADF Test	PP Test
Urea	0.97	-1.22	0.00	-2.40	0.45	-1.69
ΔUrea	-5.65***	-6.83***	-5.95***	-4.79***	-8.01***	-7.59***
Ammonia	0.87	-1.01	0.75	-1.18	1.23	-1.11
ΔAmmonia	-3.43***	-4.50***	-6.95***	-8.60***	-7.34***	-10.91***
Corn	0.01	-2.12	1.12	-1.35	0.93	-0.75
ΔCorn	-5.03***	-5.72***	-4.95***	-7.41***	-6.99***	-9.72***
Gas	0.81	-1.84	-0.60	-1.49	0.08	-2.67*
ΔGas	-6.00***	-7.36***	-5.74***	-8.78***	-8.42***	-11.53***

Note: (***), (**) and (*) represents stationary at 1%, 5% and 10% confidence level

3.4.2 Model Estimation and Empirical Results

To determine whether nitrogen fertilizer prices follow the marginal cost of production (the competitive case) or follow the value of the marginal productivity of fertilizers in agriculture (the non-competitive case), we first check the co-integration relation between price series to perform a Granger causality analysis. Since the analysis in the introduction implies a structural change in the causal relationship over the time period, we split our data into pre-2006 and post-2006 components and apply Johansen's co-integration test in each sub-period.

Table 3.4: Pair-wise Johansen co-integration test

		Urea		Ammonia	
		Corn	Gas	Corn	Gas
99-Jan to 05-Dec	$r = 0$	14.40	13.03	11.37	21.88**
	$r \leq 1$	1.20	1.60	1.92	2.97
06-Jan to 11-Dec	$r = 0$	24.06***	13.37	17.41*	10.17
	$r \leq 1$	1.50	1.78	3.71	1.89
99-Jan to 11-Dec	$r = 0$	12.46	12.05	15.94	14.78
	$r \leq 1$	0.86	4.53	0.67	3.59

Note: (***), (**) and (*) represents co-integration at 1%, 5% and 10% confidence level

Table 3.4 illustrates the maximum likelihood trace statistics (Johansen, 1991) of pair-wise Granger causality test in each sample component. Table 3.4 shows that both urea and ammonia prices are co-integrated with corn price in the post-2006 sample period, indicating a non-competitive market structure in which fertilizer prices follow the value of its marginal productivity in agriculture; whereas ammonia price is co-integrated with natural gas price in the pre-2006 sample period, which is an indication of competitive market structure in which

fertilizer prices follow the marginal cost of production. In addition, the full sample test results demonstrate that there is no Granger causal relationship between fertilizer prices and corn price or natural gas price in the full sample range.

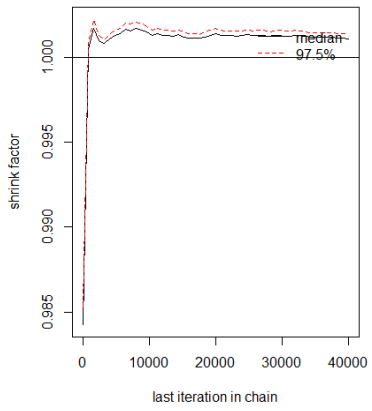
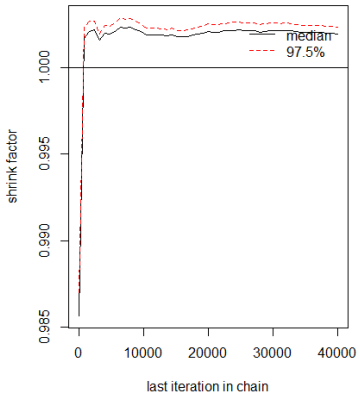
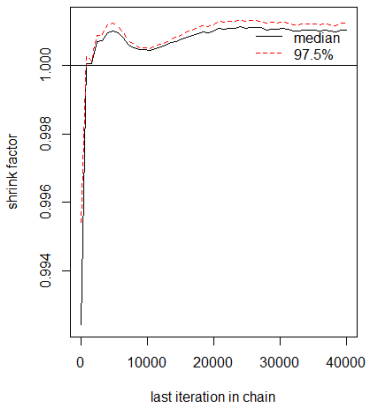
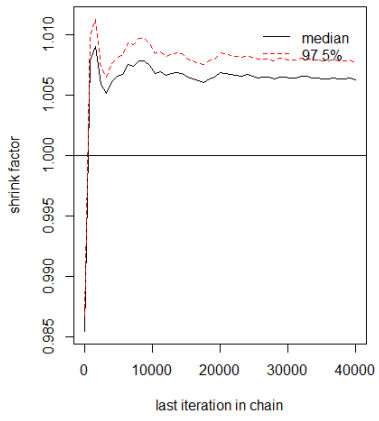
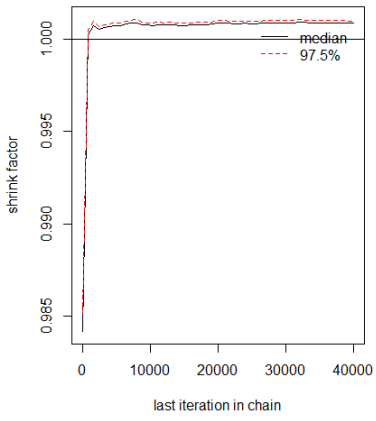
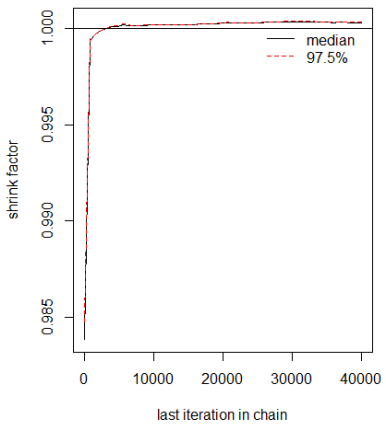
The results of the standard pair-wise Granger causality test in table 3.4 demonstrate an unstable causal relationship between nitrogen fertilizer price and its feedstock (natural gas) or output (corn) price over the sample period. This non-constancy of causality makes the application of standard Granger causality test inappropriate. Thus, we next investigate the time-varying effects of natural gas prices and corn prices on U.S. nitrogen fertilizer prices over the 1999–2011 sample period by using a Bayesian-based time-varying parameter approach.

The main advantage of this method is that it enables us to explore the structural change, when there is transition from one type causal relationship to another within the model, using information presented in the full sample. The rejection of the existence of a co-integrated relationship between nitrogen fertilizer price and its main feedstock or output price over the full sample period might be the result of structural changes in the long-run relationship that cannot be modelled explicitly by a traditional error correction model. Instead of splitting the sample into sub-periods according to pre-defined points, the time-varying parameter approach, which is based on a dynamic linear model, enables us to investigate the transitions of relationships between price series via full sample data information.

The critical issue for Markov Chain Monte Carlo (MCMC) methods in applications is how to determine when it is safe to stop sampling and use the samples to estimate characteristics of the distribution of interest. This paper applies Gelman and Rubin's convergence diagnostic on all posterior draws of parameters to ensure the convergence of MCMC items. The test result in table 3.5 demonstrates the evolution of Gelman and Rubin's

shrink factor as the number of iterations increases for posterior draws of parameters $\beta_t^{(2)}$, $\beta_t^{(3)}$ and $\beta_t^{(4)}$. These parameters characterize the time varying effect on nitrogen fertilizer prices. The evolutions for all impact parameters are close to 1, indicating convergence of these posterior draws.

Table 3.5: MCMC convergence test

Impact of Corn on Urea	**Impact of Gas on Urea**	**Impact of Capacity on Urea**
		
Impact of Corn on AA	**Impact of Gas on AA**	**Impact of Capacity on AA**
		

The impacts of corn price, natural gas price and capacity utilization on nitrogen fertilizer price are characterized by $\beta_t^{(2)}$, $\beta_t^{(3)}$ and $\beta_t^{(4)}$, respectively. The corresponding time trends are characterized by the evolution of $\alpha_t^{(2)}, \alpha_t^{(3)}$ and $\alpha_t^{(4)}$. Figure 3.5 shows posterior mean of estimation results of $\alpha_t^{(2)}, \alpha_t^{(3)}$ and $\alpha_t^{(4)}$ in ammonia equation. The posterior means of $\alpha_t^{(2)}$ and $\alpha_t^{(3)}$ are always close to zero, indicating no time trend of corn and natural gas price effect on ammonia price. But there is obvious downward trending for capacity utilization effect on ammonia price.

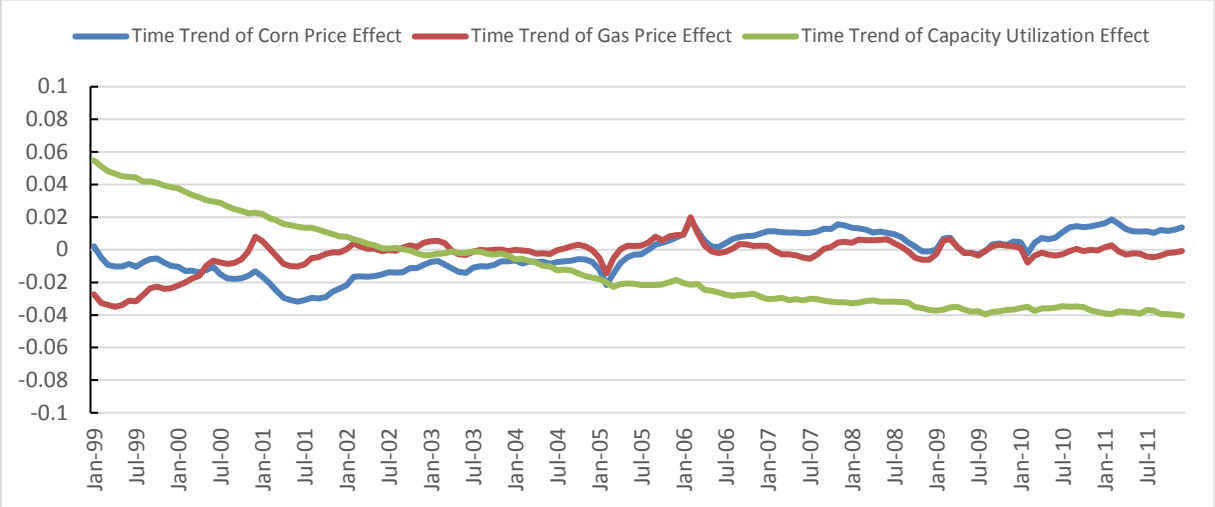


Figure 3.5: Time trend of effects on ammonia price

Similarly, figure 3.6 illustrates posterior mean of estimation results of $\alpha_t^{(2)}, \alpha_t^{(3)}$ and $\alpha_t^{(4)}$ in urea equation. There is no time trend of corn price effect on urea price as the posterior mean of $\alpha_t^{(2)}$ varies around zero. The natural gas price effect on urea price experiences slightly increasing trend, while the time trend of capacity utilization effect on ammonia price is significantly downward sloping.

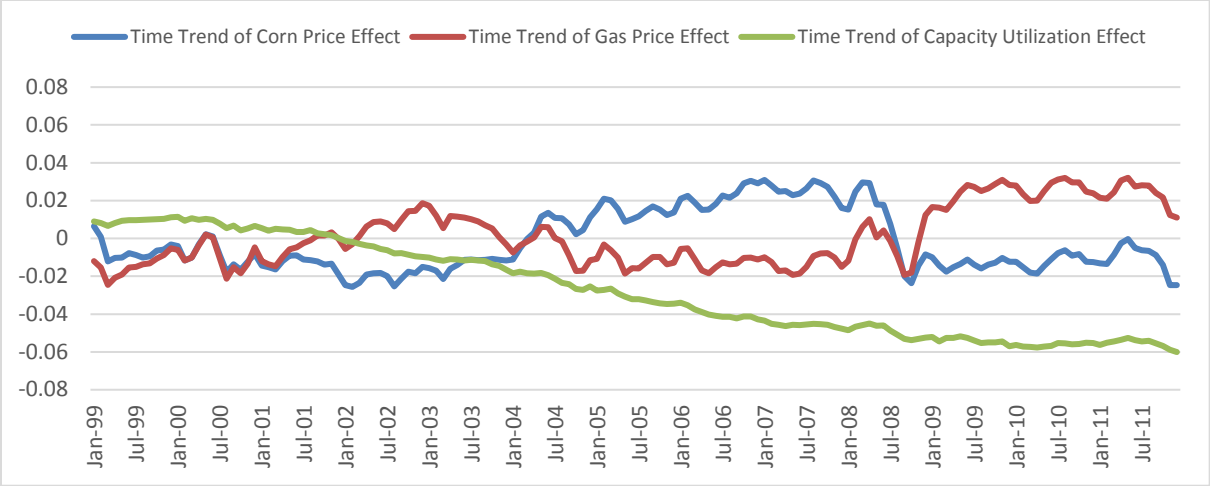


Figure 3.6: Time trend of effects on urea price

Taking into account these time trends for the time varying effect on ammonia and urea prices, the de-trended estimation results of parameter $\beta_t^{(2)}$, $\beta_t^{(3)}$ and $\beta_t^{(4)}$ in each equation are illustrated from figure 3.7 to 3.12.

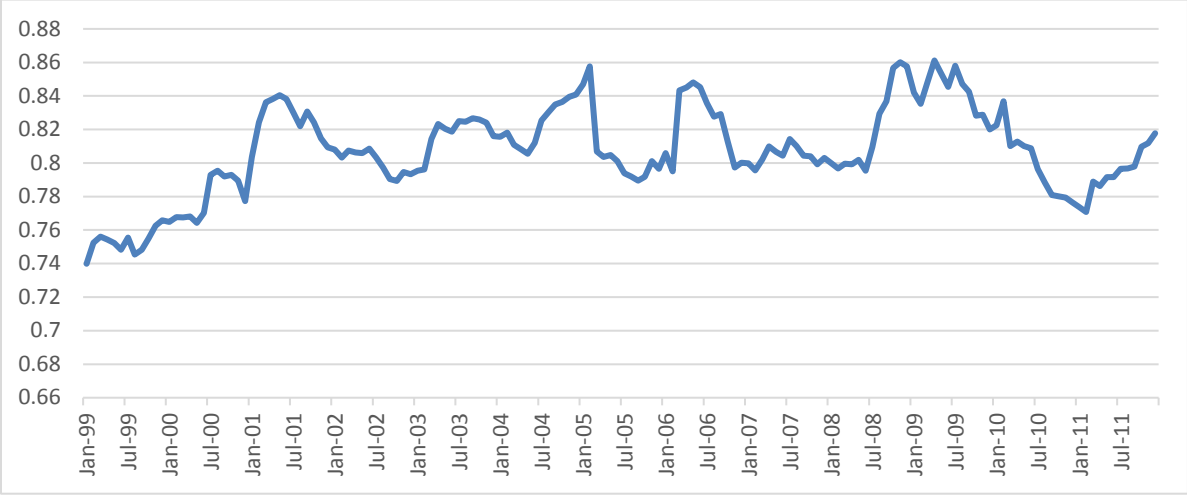


Figure 3.7: Impact of corn price on ammonia price

Figure 3.7 illustrates the de-trended posterior mean of time-varying estimations on the impact of corn price on ammonia price. It shows that the effect of corn prices on ammonia prices has increased over time—the estimated coefficient increased from 0.739 to 0.817 over the sample period with a peak of 0.861.

Figure 3.8 shows de-trended estimation of the time-varying posterior mean impact of natural gas prices on ammonia prices when considering the effect of capacity utilization. It shows a decreasing trend of the estimated effects of natural gas prices on ammonia prices—the coefficient decreased from 0.359 to 0.273 over the sample period with a trough of 0.184.

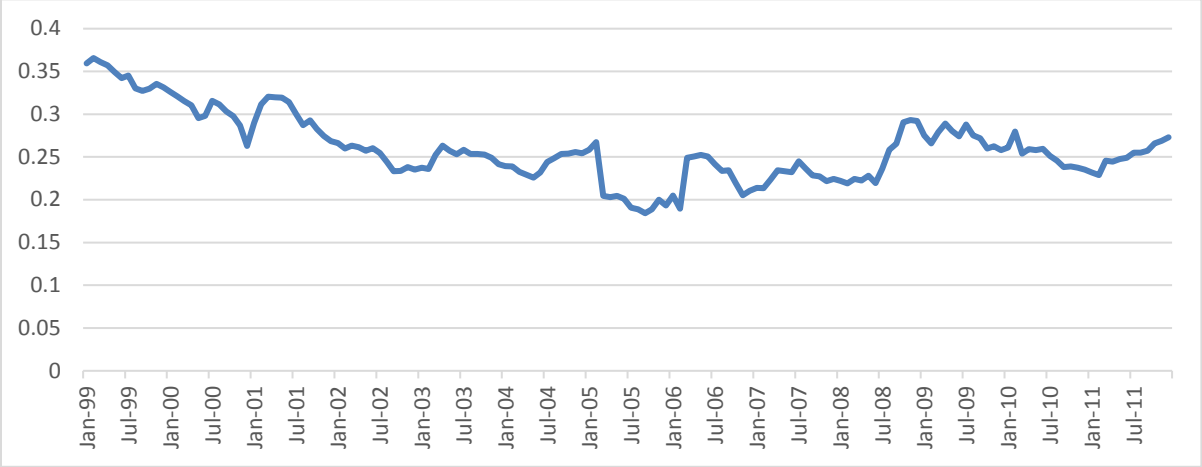


Figure 3.8: Impact of natural gas price on ammonia price

Both the increasing effect of corn price on ammonia price demonstrated in figure 3.7 and the declining effect of natural gas price on ammonia price demonstrated in figure 3.8 indicate that nitrogen fertilizer prices have been following more closely to the value of marginal productivity rather than marginal cost of production, implying a stronger market power within the fertilizer industry over time. In addition, the posterior mean effect of capacity utilization on ammonia price in figure 3.9 provides another support to the increasing market power in

ammonia market. It illustrates steady increasing impact of ammonia capacity utilization on its price. In an uncompetitive market, supplier tends to restrict its capacity utilization to lower quantity and increase price-cost margin. Thus, the higher impact of capacity utilization on prices, the higher market power caused by this capacity constraint.

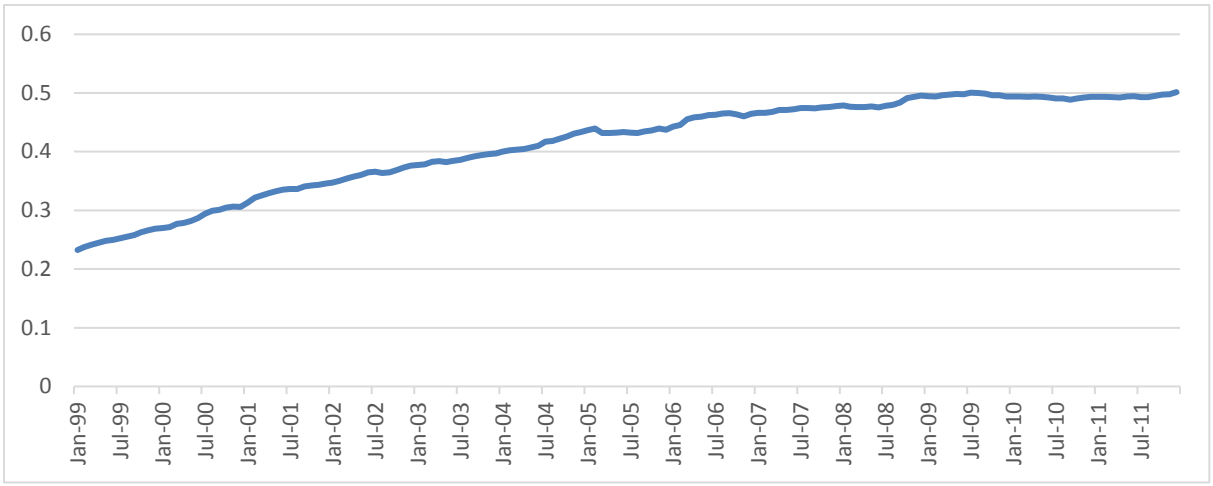


Figure 3.9: Impact of capacity utilization on ammonia price

Turn to the model of urea price, figure 3.10 shows the de-trended posterior mean of time-varying impact of corn prices on urea prices, when considering the effect of capacity utilization. After removing the effect of time trend, it still shows a convincing increasing impact of corn prices on urea prices over time, maintaining a high level in recent years. The time-varying estimated coefficient increased steadily from 0.363 to a peak of 0.653, and remained in a range between 0.539 and 0.653 since December, 2008.

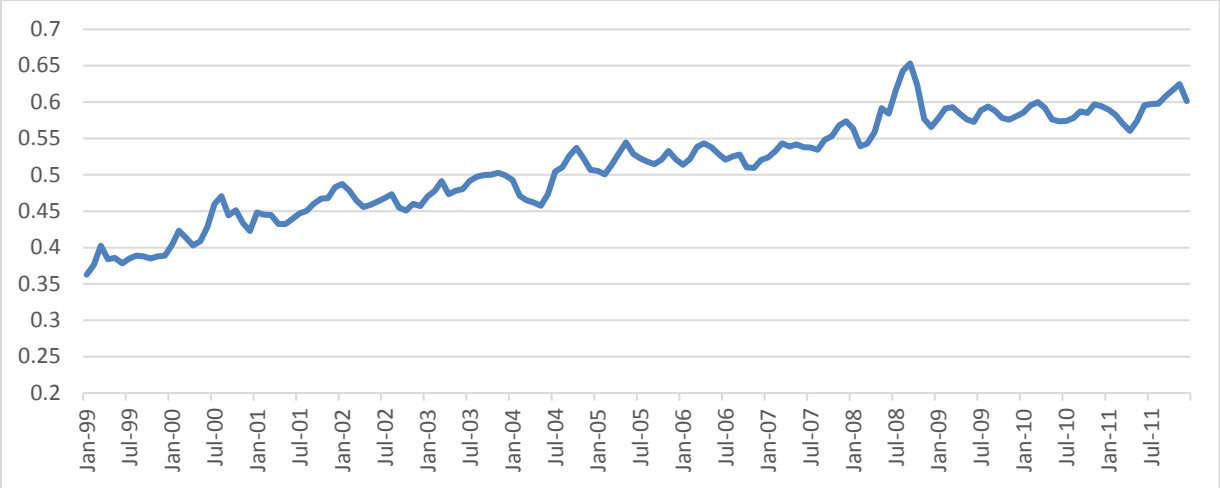


Figure 3.10: Impact of corn price on urea price

Figure 3.11 illustrates the de-trended estimation of the time-varying posterior mean impact of natural gas prices on urea prices when considering the effect of capacity utilization. It shows that the impact of natural gas prices on urea prices have been oscillating from around 0.20 from 1999 to mid-2008, then dropping below that level to the trough of 0.124 and increasing very slowly since year 2011.

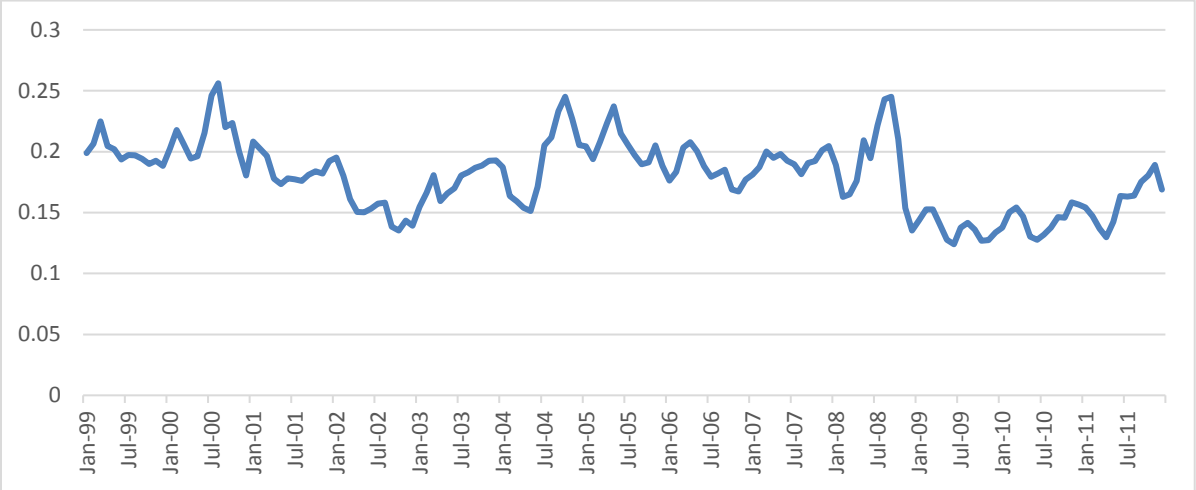


Figure 3.11: Impact of natural gas price on urea price

Both the increasing posterior mean effect of corn price on urea price demonstrated in figure 3.10 and the declining posterior mean effect of natural gas price on urea price demonstrated in figure 3.11 indicate that nitrogen fertilizer prices have been following more closely to the value of marginal productivity rather than marginal cost of production, implying a stronger market power within the urea industry over time. Again, the posterior mean effect of capacity utilization on urea price in figure 3.12 provides another support to the increasing market power in urea market. It illustrates steady increasing impact of urea capacity utilization on its price. This strong impact of capacity utilization on commodity prices indicates an increasing market power in urea market.

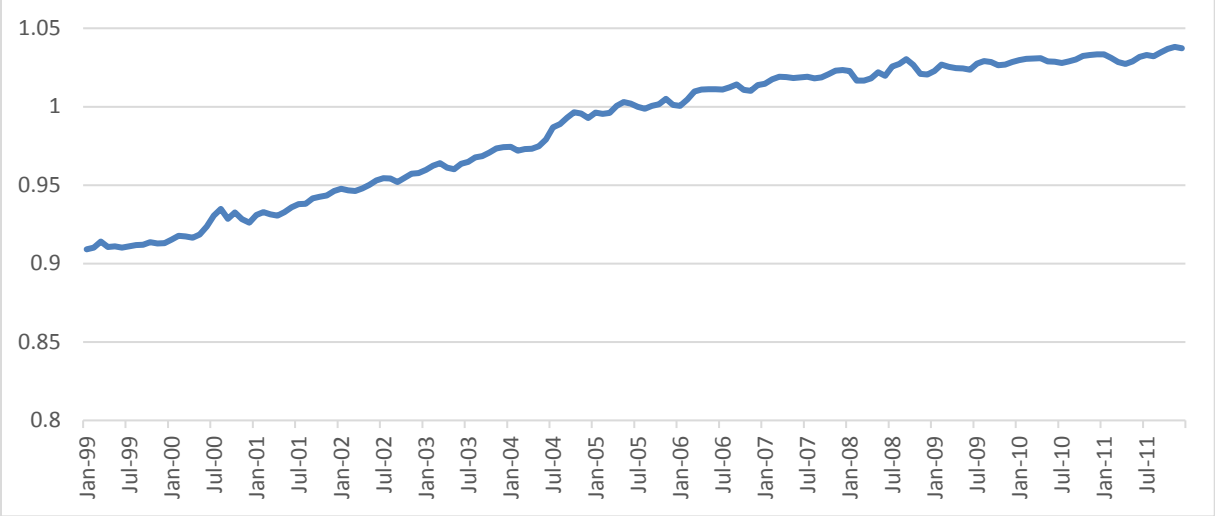


Figure 3.12: Impact of capacity utilization on urea price

In summary, the estimation results of the time-varying parameter model for both ammonia and urea prices demonstrate an increasing effect of corn prices on nitrogen fertilizer prices over the years and a decreasing effect of natural gas on nitrogen fertilizer prices. This variation in causal relations coincides with the increasing impact of capacity utilization of the industry on

its commodity price. These time-varying changes in price causal relationships can be explained by the increasing degree of the non-competitive market structure in the U.S. nitrogen fertilizer industry—its market concentration index, one measurement of market power, has also increased steadily in recent years.[6]

3.4.3 Single-Equation Error Correction Model

The standard invariant parameter Granger causality test performed in section 3.2 demonstrated that there is no co-integration relationship between nitrogen fertilizer price and its main feedstock and output price over the full sample period. The co-integrating vector remains constant over time in a standard co-integration model, but the results in section 3.2 show that the relationships between fertilizer prices and corn and natural gas prices are varied over time. As a result, the long-run equilibrium relationship between price series cannot be explained by the model with a constant co-integrating vector, since it is impossible to take into account the structural change in the model appropriately.

The next thing we are interested in examining is whether there is a long-run equilibrium relationship between fertilizer price and the price of its demand and supply factors. We utilize the single-equation error correction model (SSECM) to examine the long-run relationship between these three components. In our model, the error correction term is obtained from the Bayesian estimates of the time-varying regression model and can be represented as

---

[6] For urea, the 4-firm concentration index has increased from 68% to 84% over the sample period; and the 4-firm concentration index of the ammonia industry has increased from 48% to 77% over sample period.

$$ECT_t = Fertilizer_t - \hat{\beta}_t^{(1)} - \hat{\beta}_t^{(2)} Corn_t - \hat{\beta}_t^{(3)} Gas_t - \hat{\beta}_t^{(4)} Capacity_t$$

where $\hat{\beta}_t^{(i)}$, $i = 1, 2, 3, 4$ are estimates from the time-varying regression model in equation (3.2); $Fertilizer_t$, $Corn_t$ and Gas_t are monthly prices of nitrogen fertilizer, corn, and natural gas, respectively, and $Capacity_t$ is the fertilizer capacity utilization rate at time t. The error correction model is expressed as

$$\Delta Fertilizer_t = \phi_0 + \phi_1 ECT_{t-1} + \phi_2 \Delta Corn_t + \phi_3 \Delta Gas_t + \varepsilon_t$$

where ϕ_1 characterizes the long-run adjustment between the three price series back to their long-run equilibrium relationship, ϕ_2 and ϕ_3 measure the short-run adjustment of corn and natural gas price changes on fertilizer price.

Table 3.6: Results of error correction model

Variables	Model 1 (Urea)	Model 2 (AA)
Constant	0.0097	0.0081*
ECT_{t-1}	-24.8005***	-20.2709***
$\Delta Corn_t$	0.2885**	0.1972**
ΔGas_t	0.1202**	0.1327***

Note: (***), (**) and (*) represents significant at 1%, 5% and 10% level

The estimation results are reported in table 3.6. First, all coefficients are significant at the 5% level, except the constant term in the model, indicating a good explanation of this error correction model. Second, the estimated coefficients of the error correction term are both negative and significant, which is an indication of the existence of a co-integrated relationship.

Third, the coefficient of corn price changes is significantly larger than that of natural gas price changes, indicating a greater impact of corn price on fertilizer price. This supports the results we obtained in section 3.2 by using the time-varying parameter model.

3.5 Conclusion

This paper utilizes a time-varying parameter approach to analyze the price causal relationship between U.S. nitrogen fertilizer and its main feedstock and crop product, while considering capacity utilization of the industry. Our empirical results from the time-varying parameter approach show that the price of corn, the largest share of nitrogen use, has become more influential in affecting nitrogen fertilizer price over the years, while the effect of natural gas, the main feedstock of nitrogen fertilizer production, has been decreasing in recent years. Combining this result with increasing market concentration and decreasing capacity utilization, we conclude that the degree of non-competitiveness in the U.S. nitrogen fertilizer industry has increased over the years. The analysis of a long-run and short-run adjustment to the equilibrium relationship from the error correction model provides support for our conclusion.

REFERENCES

Archer, D.W., and J.M.F. Johnson. 2012. “Evaluating Local Crop Residue Biomass Supply: Economic and Environmental Impacts.” *BioEnergy Research* 5: 699–712.

Arslanturk, Y., M. Balcilar, and Z.A. Ozdemir. 2011. “Time-varying linkages between tourism receipts and economic growth in a small open economy.” *Economic Modelling* 28: 664–671.

Balcilar, M., H. Gungor, and S. Hammoudeh. 2015. “The time-varying causality between spot and futures crude oil prices: A regime switching approach.” *International Review of Economics & Finance* 40: 51-71.

Balke, N. S., & Fomby, T. B. (1997). Threshold co-integration. *International Economic Review* 38: 627-645.

Booth, G. G., Brockman, P., & Tse, Y. (1998). The relationship between US and Canadian wheat futures. *Applied Financial Economics* 8: 73-80.

Bracmort, Kelsi. 2015. “The Renewable Fuel Standard (RFS): Cellulosic Biofuels.” *Congressional Research Service, R41106. January 14, 2015. Available at http://nationalaglawcenter.org/wp-content/uploads/assets/crs/R41106.pdf*

Carter, C.K., and R. Kohn. 1994. “On Gibbs sampling for state space model.” *Biometrika* 81: 541–553.

Chow, G. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28: 591-603.

Christofoletti, M. A., Silva, R., & Mattos, F. (2012). “The increasing participation of China in the world soybean market and its impact on price linkages in futures markets.” *In proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO.*

Chu, Q. C., Hsieh, W. G., & Tse, Y. (1999). “Price discovery on the S&P 500 index markets: An analysis of spot index, index futures, and SPDRs.” *International Review of Financial Analysis* 8: 21-34.

Darr, Matt. 2013. “Harvest, Storage, and Transportation Systems for Large Square Bale Feedstock Supply Chains.” *Iowa State University.*

Dumortier, J. 2014. “Impact of Different Bioenergy Crop Yield Estimates on the Cellulosic Ethanol Feedstock Mix.” *Paper presented at AAEA annual meeting, Minneapolis, MN, July 27–29, 2014.*

Durbin, J., and S.J. Koopman. 2002. “A simple and efficient simulation smother for state space time series analysis.” *Biometrika* 2002: 603–615.

Engle, R.F., and C.W.J. Granger. 1987. “Co-integration and error correction: Representation, estimation, and testing.” *Econometrica* 55: 251–276.

French, B.C. 1960. “Some Considerations in Estimating Assembly Cost Functions for Agricultural Processing Operations.” *Journal of Farm Economics* 42: 767–778.

Frühwirth-Schnatter, S. 1994. “Data augmentation and dynamic linear models.” *Journal of Time Series Analysis* 15: 183-202.

Fung, H.G., Y. Tse, J. Yau, and L. Zhao. 2013. “A leader of the world commodity futures markets in the making? The case of China’s commodity futures.” *International Review of Financial Analysis* 27: 103–114.

Galbraith, C. 2010. “An examination of factors influencing fertilizer price adjustments.” *2010 AAEA Annual Meeting, July 25–27, 2010, Denver, CO.*

Ghosh, A. (1993). Co-integration and error correction models: Intertemporal causality between index and futures prices. *The Journal of Futures Markets* 13: 193-198.

Graham, R. L., R. Nelson, J. Sheehan, R.D. Perlack, and L.L. Wright. 2007. “Current and Potential U.S. Corn Stover Supplies.” *Agronomy Journal* 99: 1–11.

Granger, C.W.J. 1981. “Some properties of time series data and their use in econometric model specification.” *Journal of Econometrics* 16: 121–130.

Greenhut, M.L., Norman, G., and C-S. Hung. 1987. “The Economics of Imperfect Competition: A Spatial Approach.” *Cambridge University Press*.

Han, L., Liang, R., & Tang, K. (2013). “Cross-market soybean futures price discovery: does the Dalian Commodity Exchange affect the Chicago Board of Trade?” *Quantitative Finance* 13: 613-636.

Hansen, B. E. (2001). “The new econometrics of structural change: dating breaks in U.S. labor productivity.” *Journal of Economic Perspectives* 15: 117-128.

Hansen, B. E., & Seo, B. (2002). “Testing for two-regime threshold co-integration in vector error-correction models.” *Journal of Econometrics* 110: 293-318.

Huang, W. 2007. “Impact of rising natural gas prices on U.S. ammonia supply.” *Outlook No. WRS-0702.*

Humber, J. 2004. “A time series approach to retrospective merger analysis: evidence in the U.S. nitrogen fertilizer industry.” *2014 AAEA Annual Meeting, July 27–29, 2010, Minneapolis, MN.*

Johansen S. (1988). “Statistical analysis of co-integration vectors.” *Journal of Economic Dynamics and Control* 12: 231-254.

Johansen, S. 1991. “Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models.” Econometrica 59: 1551–1580.

Jong, P.D., and N. Shephard. 1995. “The simulation smoother for time series models.” *Biometrika* 82: 339–350.

Koop, G., R. Leon-Gonzalez, and R.W. Strachan. 2001. “Bayesian inference in a time varying cointegration model.” *Journal of Econometrics* 165: 210–220.

Liu, B. J., Wang, Y., Wang, J., Wu, X., & Zhang, S. (2015). “Is China the price taker in soybean futures?” *China Agricultural Economic Review* 7, 389-404.

Martens, M., Kofman, P., & Vorst, T. C. F. (1998). “A threshold error-correction model for intraday futures and index returns.” *Journal of Applied Econometrics* 13, 245-263.

Merener, N. (2015). “Globally distributed production and the pricing of CME commodity futures.” *Journal of Futures Markets* 35, 1-30.

Natanelov, V., Alam, M. J., McKenzie, A. M., & Huylenbroeck, G. V. (2011). “Is there co-movement of agricultural commodities futures prices and crude oil?” *Energy Policy* 39, 4971-4984.

Ogden, J.M., and L. Anderson, eds. 2011. “Sustainable Transportation Energy Pathways: A Research Summary for Decision Makers.” *UC Davis, California: Institute of Transportation Studies.*

Park. S.Y., and G. Zhao. 2010. “An estimation of U.S. gasoline demand: A smooth time-varying cointegration approach.” *Energy Economics* 32: 110–120.

Peri, M., & Baldi, L. (2010). Vegetable oil market and biofuel policy: An asymmetric co-integration approach. *Energy Economics*, 32, 687-693.

Petris, G., S. Petrone, and P. Campagnoli. 2009. "Dynamic linear model with R." *pp. 31-206. New York: Springer.*

Pieper, John. 2015. "Executive in charge of stover purchasing for Du Pont Cellulosic Ethanol Plant Nevada, Iowa." *Personal communication with author. March 14th 2015.*

Seo, M. (2006). "Bootstrap testing for the null of no co-integration in a threshold vecor error correction model." *Journal of Econometrics* 134: 129-150.

Swoboda, Rod. 2014. "Fuel of the Future: Project Liberty Cellulosic Ethanol Plant Opens."

US Department of Energy (US DoE) 2011. "US Billion-ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry." *R.D. Perlack and B.J. Stokes (leads), ORNL/TM-2011/221. Oak Ridge National Laboratory, Oak Ridge, TN.*

US Department of Agriculture National Agriculture Statistics Service (USDA NASS). 2015. "2015 State Agriculture Overview: Iowa."

Varian, H.R. 2009. "Intermediate Microeconomics: A Modern Approach." *New York: W.W. Norton & Co.*

Wahab, M., & Lashgari, M. (1993). "Price dynamics and error correction in stock index and stock index futures markets: A co-integration approach." *The Journal of Futures Markets* 13: 711-742.