Market Extent and Integration for Bioenergy Products: An Analysis Using Weekly Price Data

1. BACKGROUND

Biofuels have generated a great deal of interest among developed and developing countries as a way to simultaneously reduce imports of petroleum and reduce air pollution caused by the combustion of fossil fuels. Heightened concerns about global climate change, expanding demand and increasing oil prices, and instability in oil-exporting countries have led to considerable efforts in many nations to promote biofuels as an alternative to fossil fuels. In the U.S., attention has focused principally on ethanol derived from corn feed-stocks (Saitone et al, 2007), although in recent years, "2nd-generation biofuels" produced from crop and forest residues and from non-food energy crops are gaining importance. In the U.S., federal and state energy policies have also contributed considerably to the expansion of the biofuels industry.

Ethanol is also derived from renewable feed-stocks and used mostly in the Midwestern states- the Corn Belt and California. One disadvantage however of ethanol is that it cannot easily be transferred through petroleum pipelines and therefore, must be splash-blended at service stations. Although most ethanol consumption is in conventional gasoline engines (as an oxygenate, octane booster and gasoline extender), which are limited to a 10-percent ethanol blend (E10), there is growing demand for ethanol blended in higher concentrations, such as E85 (85 percent ethanol, 15 percent gasoline). Interestingly, the value of ethanol varies depending on how it is blended with gasoline (Dipardo, 2000). In purer forms, it can also be used as an alternative to gasoline in automobiles designed for its use.

For the reason that it is generally less expensive to produce ethanol close to the feedstock supply, it is not surprising that the top five corn-producing states in the U.S. are also among the top ethanol-producers. Most ethanol use is in the metropolitan centers of the Midwest, where it is produced. When ethanol is used in other regions, shipping costs tend to be high, since ethanol-blended gasoline cannot travel through petroleum pipelines. This geographic concentration is an obstacle to the use of ethanol on the East and West Coasts (Yacobucci et al, 2000).

As defined by Alfred Marshall (1961), “a market for a good is the area within which the price of a good tends to uniformity, allowance being made for transportation.” This definition is related to the economic market where differences in prices of the same commodity observed at different places are due to transaction costs. Therefore, according to the definition of a market, in the same geographic region, it is almost impossible that prices of the same commodity display a greater difference than the transaction costs over a long period of time. If a single price exists over several spatially separate markets, it implies that these markets are integrated as a single market (Yang, Bessler and Leathan, 2000). In other words, assuming market integration, prices of a commodity observed in different locations simultaneously will differ by the amount up to the transaction costs as maintained by the law of one price, the basis for defining spatial price relations and market extent. (Dawson and Dey 2002; FAO 2004). The reason being that arbitrage will always occur when the price differences in different geographic regions exceed the transaction costs (Egbendewe-Mondzozo, 2009).

The concept of market integration has been used in defining market boundaries in antitrust cases (see, for example, Horowitz 1981; Slade 1986; Spiller and Huang 1986; Kleit 2001) and international trade conflicts (e.g., Asche et al 1999). It has been suggested, for example, that a greater degree of integration leads to more transmission of price signals, which, in turn, encourages producers to specialize according to comparative advantage (Baulch 1997). The goal of analysis of market extent and the degree of market integration have led to the development and use of a myriad of methods. According to Fackler and Tastan (2008), economists commonly study measures of integration market by analyzing co-movement of prices (given that prices are often the only available data); which entails significant difficulty in estimating structural models capable of isolating the effect of regional demand shocks. The lack of complete data and the consequent presence of latent variables they add further compound this difficulty. The arbitrage cost approach has been shown in literature to have many advantages over the correlation approach, in particular, its ability to generate a precise number for arbitrage cost between markets, and how those arbitrage costs can change with changes in exogenous factors (Spiller and Huang, 1986).

The afore mentioned factors affecting the energy sector such as unstable energy prices, environmental concerns, pressures for oil independence and federal energy policies, thus create a strong market for renewable energy implying that conducting analyses in this evolving market is relevant. In addition, since increased demand and supply of ethanol is expected to have a significant effect on its pricing, in addition to the cost of transportation from areas of production to locations of need, among other factors, studies on spatial price determination and discovery is pertinent for the ethanol market as it emerges.

This study therefore aims to measure market integration and establish the extent of the market for ethanol with regards to cities in major bioenergy producing states as well as major consuming states (Iowa, Nebraska, Illinois, Minnesota, Indiana, New Mexico, California, Colorado and Texas). Maximum Likelihood Estimation (MLE) following Spiller and Huang (1996) earlier referred to as the arbitrage cost approach will be used. As a supplement, time series techniques namely the error correction model (ECM), directed acyclical graph (DAG), and impulse response analysis, will also be carried out to measure market integration and thereby substantiate or refute results obtained from the Spiller and Huang model.

This paper is unique because although several papers have been written on testing the extent of different markets, it employs more recent data and ventures into a market for which no similar study has been found. Organization of the paper is as follows: Section 2 introduces the models and methodologies to be employed; Section 3 presents the data and preliminary data analysis. The estimation results and discussion are contained in Section 4. Section 5 summarizes and concludes the paper.

2. MODELS AND METHODOLOGIES

MAXIMUM LIKELIHOOD ESTIMATION

The first empirical models to directly examine arbitrage outcomes assess whether two locations are in the same economic market by estimating the probability that their prices differ by the transaction cost, which is stochastic (Spiller and Huang, 1986; Spiller and Wood, 1988). Prices in the two locations either differ by the transaction cost (successful arbitrage), or by less than the transaction cost (autarky). This paper follows identically the method used to estimate extent of the markets by Spiller and Huang (1986). It is assumed that all regions of a state are within the same market and so one city in each state is used as a representative. The methodology implies the estimation of a switching regimes model. One regime is characterized by the prices between the two products differing by the arbitrage (or transaction) costs. In the other, when there is no (explicit or implicit) arbitrage between the two products, their prices differ by less than the transaction costs. This regime is statistically identified by a truncation in its error structure, similar to the stochastic frontier models estimated elsewhere in the literature (Spiller and Haung, 1986).

Now to the model itself:

Assume that the autarky prices for two markets in a given period, P_t^{1A} and P_t^{2A}, can be defined by the following reduced form equations:

(1a) $P^1{}_t = \pi^1 + \varepsilon_t^1$

(1b) $P^2{}_t = \pi^2 + \varepsilon_t^2$

Where π^1 and π^2 are nonstochastic elements of prices determined by supply and demand conditions in local markets, and ε_t^1 and ε_t^2 are zero mean stochastic disturbances (shocks) in each region. Next, define a transaction cost T_t, of moving the commodity from location A to B. In the absence of legal trading barriers but with finite transaction costs, the observed prices P_t^1 and P_t^2 may diverge from the autarky prices. Arbitrage opportunities arise if the autarky prices differ by more than T_t and do not arise if the reverse is the case. For simplicity, it is assumed that $P_t^{1A} < P_t^{2A}$. Then if

(2) $0 < P_t^{2A} - P_t^{1A} < T_t$ where $P_t^1 = P_t^{1A}$ and $P_t^2 = P_t^{2A}$. This implies that

(3) $0 < P_t^2 - P_t^1 < T_t$

Where arbitrage arises, the observed equilibrium prices in the two regions differ only by T_t therefore implying that a shock in one region translates to the other (as long as the autarky prices do not fall below T_t. Thus if

(4) $0 < T_t < P_t^{2A} - P_t^{1A}$ then

(5) $0 < P_t^2 - P_t^1 = T_t$

Now suppose that the transaction costs T_t are distributed geometrically with mean

$T_t = Te^{vt}$

Where V_t is normally distributed with zero mean and constant variance σ_v^2**.** The probability of no arbitrage opportunities and hence the probability of observing (3), is a constant λ

(6) Prob $[0< P_t^2 - P_t^1 < T_t]$ = Prob Prob $[0< P_t^{2A} - P_t^{1A} < Te^{vt}]$

$= \text{Prob } \{\log[(\pi^2 - \pi^1) + (\varepsilon_t^2 - \varepsilon_t^1) - V_t < \log T) = \lambda$

The probability of arbitrage and hence the probability of observing (5) is (1-λ). This probability measures how integrated the two areas are. If (1-λ) is very close to one (zero), then the two areas are almost always (never) in the same economic market. The value l is in other words the probability that prices in region B do not constrain prices in A. Thus, 1 - l is the probability that the two regions are directly "connected," that is that prices in region B act to constrain prices in region A.

Define a positive random variable U_t, and B = log T. It can be seen that the observed price equations in (3) and (5) are in fact a switching regressions system, where

(7) $\log (P_t^2 - P_t^1) = B + V_t - U_t$ with probability λ and

(8) $\log (P_t^2 - P_t^1) = B + V_t$ with probability (1-λ)

Equation (7) corresponds to the regime of no arbitrage opportunities or the autarky state, and (8) corresponds to the arbitrage state. Equation (7) is in fact a composite error regression with a positive component U_t. While the parameter λ measures the probability of being in autarky, the positive error U_t is a conditional measure of propensity to trade. The smaller the positive value of U_t, the higher is the propensity to trade.

U_t is assumed to be distributed independently of V_t, with a one-sided half-normal distribution, i.e., the distribution is derived from a normal distribution $N(0, \sigma_u^2)$ truncated from below at zero. Denote $\theta = (B, \sigma_u^2, \sigma_v^2, \lambda)$ as the parameter vector for the regressions (7) and (8); then the likelihood function for the n observations is given by:

(9) $$L = \prod_{t=1}^{n}[\lambda f_t^1 + (1-\lambda) f_t^2]$$

where f_t^1 and f_t^2 are the density functions of (7) and (8) respectively.

Let $Y_t = \log(P_t^1 - P_t^2)$ then the density functions are

(10) $$f_t^1 = \left(\frac{2}{\sqrt{(\sigma_u^2+\sigma_v^2)}}\right)\phi\left(\frac{Y_t - B}{\sqrt{(\sigma_u^2+\sigma_v^2)}}\right)\left[1 - \Phi\left(\frac{(Y_t - B)\frac{\sigma_u}{\sigma_v}}{\sqrt{(\sigma_u^2+\sigma_v^2)}}\right)\right]$$

(11) $$f_t^2 = \frac{1}{\sigma_v}\phi\left(\frac{Y_t - B}{\sigma_v}\right)$$

Where ϕ and Φ are the standard normal density and distribution functions respectively. In this context, the goal of the maximum likelihood estimation is to maximize the value of L in (9) over the parameters θ.

TIME SERIES TECHNIQUES

As time series techniques will be used to supplement the arbitrage cost approach, the associated theoretical framework will not be discussed in depth in this paper.

The Error Correction Model (ECM)

The expectation of observing cointegration in the 9 ethanol price series lends to the data generating process of P_t (price at time t) being appropriately modeled in an error correction model (ECM) with k-1 lags following from Stockton et al. (2010) . The ECM for the 9 markets is:

$$\text{(12)} \qquad \Delta P_t = \Pi P_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta P_{t-i} + \mu + e_t \quad \text{where } t = 1,\ldots,T\text{ ; } e_t \sim \text{Niid}\,(0, \Sigma)$$

where Δ is the difference operator ($\Delta P_t = P_t - P_{t-1}$), P_t is a (9 x 1) vector of weekly prices at time t = 1,…, T ; Γ_i is a (9 x 9) matrix of parameters to be estimated reflecting the short-run relationships between past differences in prices and current differences in prices (price changes lagged i period to current changes in prices); $\Pi = \alpha\beta'$ is (9 x 9) matrix of parameters reflecting the relationship between lagged levels of prices to current changes in prices (or 9 x 10 if a constant is in the co-integration space); μ is a constant and ε_t is a (n x 1) vector of white noise innovations. The matrix β' reflects the long-run relationships between levels of price series and α is a matrix of adjustment parameters summarizing how each series adjusts to perturbations in each of the long-run relationships summarized in β' (Stockton et al., 2010).

Directed Acyclic Graph

Following Stockton et al. (2010), contemporaneous information flows are studied in a DAG structure using estimated innovations and their estimated covariances via the matrix, using PC algorithm. The principal idea of DAGs is to determine the causal relationship or flow among a set of variables then portray it using an arrow graph or picture (Vitale and Bessler, 2006). In Spirtes et al. (2000), the PC algorithm, one of the search algorithms associated with DAGs and employed in this study, is described as a sequential set of commands that begin with an unrestricted graph where every variable is connected with every other variable and proceeds step-wise to remove lines between variables and to direct "causal flow.” Although shortcomings of the PC algorithm have been documented in literature, its advantages and extensive usage have also been emphasized (see Spirtes et al. 2000 and Demeralp and Hoover, 2003).

In the graphs, given two variables X and Y, there are five possibilities between the variables: (1) no causal relationship when edges are removed, (2) Y causes X ($Y \rightarrow X$), (3) X causes Y ($X \rightarrow Y$), (4) Y and X simultaneously cause each other ($Y \leftrightarrow X$), and (5) the causal flow cannot be directed by the information contained in the sample($Y — X$).

3. DESCRIPTION OF DATA AND PRELIMINARY DATA ANALYSIS

The data employed consists of weekly prices of ethanol per gallon from Oil Price Information Service (OPIS) as reported in Hart's Oxy Fuel News. The assumption is that all regions of a state are within the same market and so one city in each state is used as a representative. The data spans 20 years (1989-2008 with 1036 observations) for Los Angeles (PLA); Denver (PDV); Cedar Rapids (PCR); Chicago (PCH); Indianapolis (PIN); Minneapolis (PMN); Albuquerque (PAL); Houston (PHO); Seattle (PSE). The selection of cities was based on production capacity and also utilization (Midwest), consumption (Los Angeles, Houston), and distance from major hubs (Seattle and Albuquerque), amongst other possible justifications.

Table 1. Summary statistics on ethanol prices in nine U.S cities, 1989-2008

City	**Mean**	**SD**	**CV**
Los Angeles	1.479631	0.503409	34.02261
Denver	1.469731	0.472928	32.17784
Cedar Rapids	1.415084	0.452329	31.96481
Chicago	1.421886	0.463686	32.61061
Indianapolis	1.408526	0.471286	33.45953
Minneapolis	1.439699	0.461545	32.05845
Albuquerque	1.489171	0.483881	32.49332
Houston	1.441413	0.529681	36.74736
Seattle	1.502976	0.491658	32.7123

Table 1 showcases descriptive statistics: the mean, standard deviation (SD), and coefficient of variation (CV) for six ethanol markets in the U.S. from January 1989 to February 2008. Seattle has the average highest price followed by Albuquerque. This agrees with the expectation that mean prices are most likely higher in consuming regions or regions far from production hotspots like the Midwest. A market such as Houston with a greater price CV indicates high variability (volatility) or possible susceptibility to shocks (receiver of price signals) from other markets, however, this supposition would be tested later in the study.

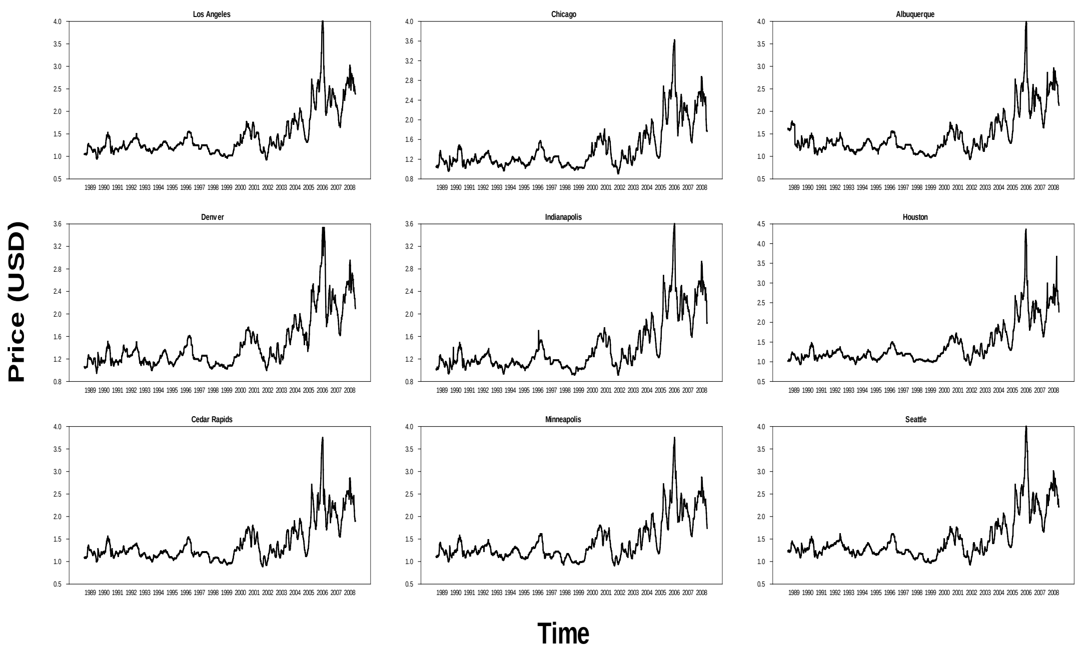

Figure 1. Graphs of price series in levels of ethanol in nine U.S. cities, 1989-2008

The prices of ethanol in the 9 markets (in figure 1) seem to be in “perfect” co-movement (co-integration). These prices show almost the same variation in terms of magnitude and amplitude (flat, peak) during the course of the period of study. This is an indication of good price signal transmission between ethanol markets and subsequent market integration. Prices begin trending upward from late 2005 and peaks observed in mid-late 2006 onwards could be attributed to the economic and food price crisis that was triggered in part by the increase in corn demand for ethanol fuel production. A quote from the New York Times in January, 2006 reads “High oil prices are dragging corn prices up with them, as the value of ethanol is pushed up by the value of the fuel it replaces” (Wald, 2006). During that period, there was also reported rising demand for animal feed in China, which helped push global grain prices to levels higher than had been observed in at least a decade.

4. RESULTS AND DISCUSSION

As mentioned earlier, prices tend to be lower in the Midwest region where most of the ethanol is produced and tends to increase in father regions. That being said, it appears that the price range across all markets in rather small. Considering only the major producing states in the Midwest for instance, it is anticipated that all will be in the same market, however for major consumers like California, Texas, Illinois and Iowa, based on geographic distances and the range of socio-economic characteristics of these states, interesting results with regards to market integration are expected. Obviously one would expect transportation costs to be a major determinant of how integrated two markets may be. It is important to recall that the ethanol market is still growing and thus may still be in a price discovery state especially with increasing federal subsidies and mandates encouraging production.

Looking at the correlation matrix below, as expected, in their levels, all prices are highly correlated. In first differences, however, price correlations are not as high. The lowest correlations are between Denver and Houston, Denver and Albuquerque and Los Angeles and Houston. Spiller and Huang in their paper suggest however that price correlations are not the proper statistic to infer whether two regions are usually in the same market. The results anticipated by using their model should therefore provide substantially different implications than those of simple correlation coefficients.

	Denver	Cedar Rapids	Chicago	Indianapolis	Minneapolis	Albuquerque	Houston	Seattle
Los Angeles	0.98/0.58	0.98/0.63	0.98/0.68	0.98/0.71	0.98/0.61	0.98/0.65	0.99/0.46	0.99/0.74
Denver		0.96/0.50	0.97/0.54	0.97/0.54	0.97/0.61	0.96/0.44	0.97/0.35	0.98/0.52
Cedar Rapids			0.99/0.85	0.99/0.78	0.99/0.53	0.98/0.64	0.98/0.51	0.98/0.59
Chicago				0.99/0.79	0.98/0.68	0.97/0.64	0.98/0.52	0.98/0.60
Indianapolis					0.99/0.73	0.97/0.62	0.99/0.50	0.98/0.63
Minneapolis						0.97/0.73	0.98/0.62	0.99/0.75
Albuquerque							0.98/0.71	0.98/0.77
Houston								0.98/0.63

Figure 2. Correlation matrix - levels/first difference of ethanol prices in nine U.S. cities

Results of the Augmented Dickey-Fuller (ADF) tests (testing the null hypothesis that each series is nonstationary) on levels and first difference are presented in table 2. All series were found to be nonstationary in levels and stationary in first differences, making each class series integrated of order one (denoted as I(1)). Also shown in table 2 are results of the Ljung-Box Q-test using the residuals or innovations from the ADF test. The p-values on Ljung-Box Q statistic show that the residuals are auto-correlated. The null hypothesis of non-auto-correlated residuals is therefore rejected. This is an obvious drawback for the veracity and interpretation of results, but resolving this issue goes beyond the scope of this study (I WILL CONSULT DR. BESSLER).

Table 2. Unit root test on prices of ethanol in nine U.S. cities, 1989-2008

Augmented Dickey-Fuller (levels)				Augmented Dickey-Fuller (1^{st} diff)		
Market	t-test	k	Q (p-value)	t-test	k	Q (p-value)
Los Angeles	-1.1091	2	586.07(0.00)	-6.2621	2	592.71 (0.00)
Denver	-1.1437	2	500.07 (0.00)	-6.2465	3	508.48 (0.00)
Cedar Rapids	-1.4061	1	579.86 (0.00)	-7.3843	2	589.71 (0.00)
Chicago	-1.47704	1	536.89 (0.00)	-8.0506	2	544.42 (0.00)
Indianapolis	-1.56612	2	412.41 (0.00)	-6.7885	2	418.05 (0.00)
Minnesota	-1.5302	2	487.09 (0.00)	-7.1065	2	501.25 (0.00)
Albuquerque	-1.32504	2	541.21 (0.00)	-6.10331	2	537.91 (0.00)
Houston	-0.98951	3	783.13 (0.00)	-4.47482	2	751.04 (0.00)
Seattle	-1.52972	3	567.28 (0.00)	-6.84231	2	579.85 (0.00)

The critical value (t-stat) to reject the null hypothesis (at 5% significance level) of non-stationarity is -2.89. The column named "k" indicates the number of lags of the dependent variable used to produce "white noise" residuals. The value of k results from the minimization of the Schwarz loss metric on values of k ranging from 1 to 3. The column labeled "Q (p-value)" refers to the Ljung-Box statistic (Portmanteau test) test of white noise residuals from ADF regression.

Table 3. Loss metrics (SL and HQ) on lag length from VARs in nine U.S. ethanol markets, 1989-2008

Lag length	**SL**	**HQ**
7	-59.8292	-60.7749
6	-59.7008	-60.513
5	**-59.7319**	-60.4106
4	-59.725	-60.2705
3	-59.5831	-59.9957
2	-59.3943	-59.6741
1	-58.9758	-59.1229

Metrics considered are Schwarz-loss (SL) and Hannan and Quinn's (HQ) measure on lag length (k) of a levels VAR: $SL = \log(|\Sigma|) + (6k)(\log T)/T$; $HQ = \log(|\Sigma|) + (2.00)(6k)\log(\log T))/T$
where Σ is the error covariance matrix and T is the total number of observations on each series. The symbol "| |" denotes the determinant operator and log is the natural logarithm. The single asterisk "*" indicates minimum of the Schwarz Loss metric and HQ measure.

The lag length for the ECM is established from the specification derived from an unrestricted VAR. Schwartz Loss (SL) and Hannan Quinn (HL) tests were performed (see table 3) to determine the maximum number of lag for the model, and the SL metric which is implemented subsequently shows a minimum of five lags for the VAR model. Table 4 displays results on the number of cointegrating vectors using the trace test. Failure to reject is at r=8 which indicates that there are 8 co-integrating vectors with a constant in the co-integrating space, implying that the series are highly co-integrated.

Table 4. Trace test (lag = 5) on ethanol prices from nine U.S. cities, 1989-2008

r	Trace	P-Value	D	Trace*	P-Value*	D*
=0	757.299	0	R	734.679	0	R
≤1	504.508	0	R	491.81	0	R
≤2	366.628	0	R	358.179	0	R
≤3	248.03	0	R	242.962	0	R
≤4	180.135	0	R	176.39	0	R
≤5	119.443	0	R	117.25	0	R
≤6	68.4	0	R	67.273	0	R
≤7	31.661	0.001	R	31.026	0.001	R
≤8	4.447	0.361	F	4.329	0.377	F

The test statistic (T) is the trace test corresponding to the number of co-integrating vectors (r) presented in the far left-hand column and a p-value. Entries associated with an asterisk have a constant within the co-integrating vectors. Entries without an asterisk have no constant in the co-integrating vector but instead have the constant outside the co-integrating vector. The column labeled "D" indicates the decision to reject (R) or fail to reject (F) at a 5% percent level of significance the null hypothesis Ho that the number of co-integrating vectors r=0, r≤1, …, r≤8.

Exploring how each series responds to innovations in every other series and the relative importance of each series in explaining (accounting for) the variation in the other series helps provide additional insight into the dynamic structure of ethanol prices in these nine cities. Contemporaneous correlations between price innovations were analyzed using the DAG and the results of this analysis (based on ECM and PC algorithm) are presented below in figure 3. The arrows and edges in illustrate the flow of information, or as stated by Stockton et al. (2000), show the causal structure of the contemporaneous innovations. The number of edges indicates a great deal of flow of information and interaction between the markets. While Houston, Seattle and Indianapolis show up to be price sinks, price signals originate majorly from Los Angeles, a chief consumer (demand pull). Chicago also provides price signals, and Illinois may well be the key production state in the Midwest (supply push).

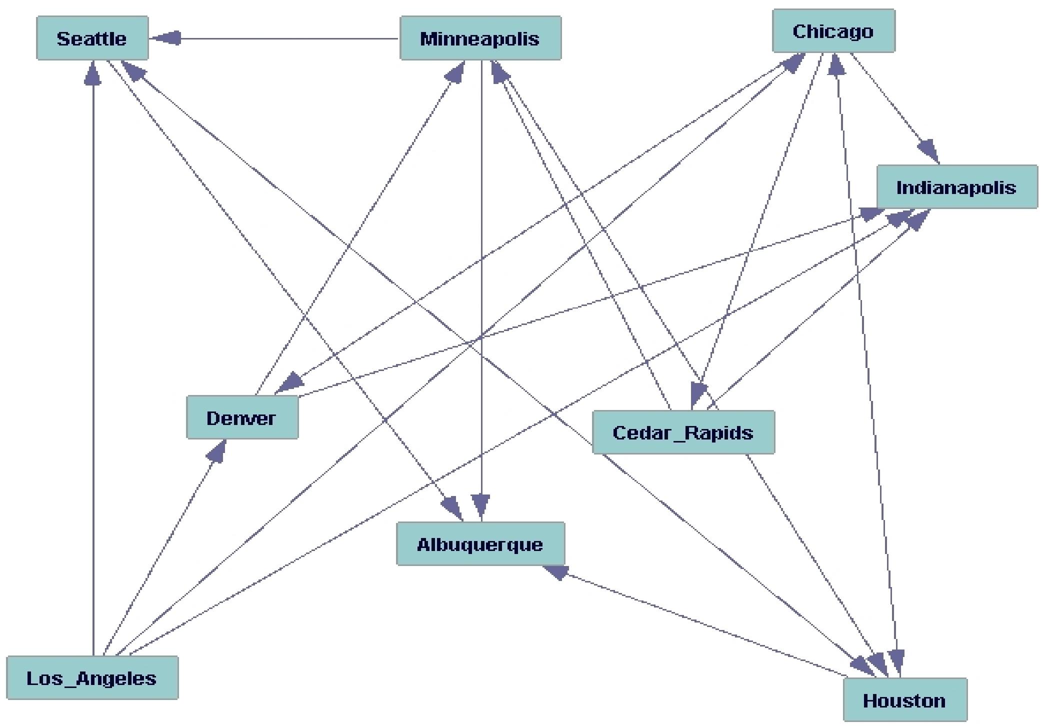

Figure 3. Causal flows found with PC algorithm at 5% significance level, on innovations from an ECM on ethanol prices from nine U.S. markets, 1989-2008

The impulse response function (figure 4) shows how different markets (listed at the beginning of each row) respond over a certain period of time (8 weeks) to a one-time-only shock or innovation from other markets (listed at the heading of each column). The impulse response function evaluates the dynamic responses to adjustment of each price to shocks in series. If the figure is read vertically, it shows how the innovation or shock (new information) from each market (listed at the heading of each column) affects prices in every market listed at the beginning of each row. Price innovations from Los Angeles are transmitted to all other markets over a period of 8 weeks (see spikes). From the results obtained and preceding discussion, it is extremely likely that new information is being transmitted within these markets rapidly.

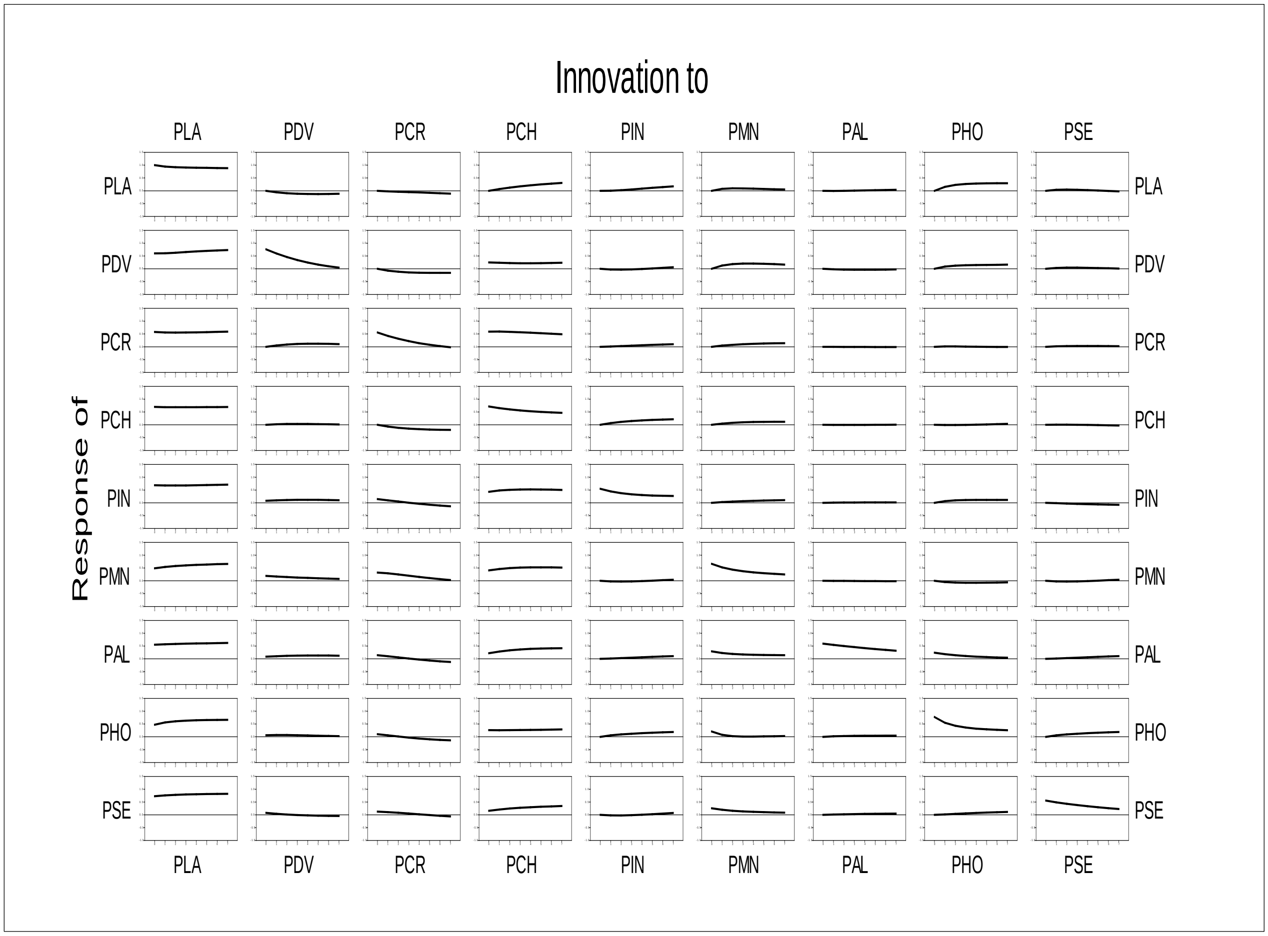

Figure 4. Response of each market to a one-time-only shock (innovation) in each series

Table 5. Maximum likelihood estimates (Log of T, λ and Log L) for selected pairs of cities

	PLA-PDV	PLA-PCR	PMN-PCH	PLA-PCH	PLA-PHO	PHO-PSE	PLA-PAL	PCH-PAL	PHO-PCR
Log of T	-0.59	-2.97	-2.75	-2.57	-2.53	-2.79	-3.56	-3.52	-3.56
λ	0.01	0.82	0.43	0.49	0.74	0.01	0.01	0.01	0.01
Log L	1670.03	214.28	720.63	978.73	955.35	351.07	660.59	201.25	660.59

Log of T represents log of transactions costs, λ is the probability of no binding arbitrage and Log L is the value of the likelihood function. The cutoff point chosen (similar to that employed by Spiller and Huang) would be such that any λ less than approximately 0.30 would be indicative of a high probability of being in the same market (1- λ).

Conjectures based on the MLE (results in table 5) would be that: city pairs with high λ are most likely not in the same market- Los Angeles is shown not to be in the same market with Cedar Rapids, Chicago and Houston. Results obtained do not necessarily reinforce the intuition that nearby states are more integrated than those further-away (take Minneapolis and Chicago for instance), however, market-pairs with higher transactions costs appear to be less integrated than those with smaller transactions costs with the exception of the Los Angeles-Denver pairing. It is therefore inconclusive that all nine cities belong to the same economic market.

5. SUMMARY AND CONCLUSION

This study aimed to measure market integration and establish the extent of the market for ethanol with regards to cities in major bioenergy producing states as well as major consuming states (Iowa, Nebraska, Illinois, Minnesota, Indiana, New Mexico, California, Colorado and Texas). Maximum Likelihood Estimation (MLE) following Spiller and Huang (1996) to estimate the transaction costs required to arbitrage in the burgeoning ethanol market was applied to weekly ethanol price data spanning 20 years (1989-2008 with 1036 observations). In addition, time series techniques (the ECM, DAG and impulse response analysis) were carried out to measure market integration. All three time series techniques show that the markets under study are co-integrated and strongly related with the observable high levels of interaction between all nine cities. The arbitrage cost approach on the other hand showed Los Angeles to likely not be in the same market with Cedar Rapids, Chicago and Houston, giving room for arbitrage opportunities. These results also do not necessarily reinforce the intuition that nearby states are more integrated than those further-away; however, market-pairs with higher transactions costs appeared to be less integrated than those with smaller transactions costs. It is therefore impractical to conclude that all nine cities belong to the same economic market (I WILL REDO SOME ESTIMATIONS WHEN DR. WU GETS BACK).

Further work would include carrying out exclusion and weak exogeneity tests to determine which markets are not parts of the co-integrating space and which ones do not respond to shocks, respectively. Using the same or a modified MLE framework, more cities-pairs could also be investigated. Ethanol is said not to compete with regular gasoline but with gasoline additives. It becomes even more interesting to investigate whether ethanol and its alternatives (such as the now largely banned MTBE) are in the same market or whether ethanol-blended gasoline and regular/ premium gasoline are in the same market.

REFERENCES

Asche, F., H. Bremnes, and C.R. Wessells. 1999. "Product Aggregation, Market Integration, and Relationships Between Prices: An Application to World Salmon Markets." American Journal of Agricultural Economics 81:568-81.

Baulch, B., 1997. "Transfer Costs, Spatial Arbitrage, and Testing for Food Market Integration." AJAE. 79(May 1997):477-87.

Egbendewe-Mondzozo A., 2009. "On the extent of the market: a Monte Carlo study and an Application to the United States egg market", Selected Paper prepared for presentation at the Southern Agricultural Economics Association Annual Meeting, Atlanta, Georgia, January 31-February 3, 2009

DiPardo, J., 2000. Outlook for Biomass Ethanol Production and Demand. http://www.drivingethanol.org/research/Misc/Biomass.pdf

Horowitz, I. 1981. "Market Definition in Antitrust Analysis: A Regression Based Approach." Southern Economic Journal 48:1-16.

Kleit, A.N. 2001. "Are Regional Oil Markets Growing Closer Together?: An Arbitrage Cost Approach." Energy Journal 22:1-15.

Marshall, A., 1961, "Principles of Economics", Variorum edition.

Saitone, T. L, Sexton R. J., and Sexton S. E. "Effects of Market Power on the Size and Distribution of Subsidy Benefits: The Case of Ethanol Promotion". Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Portland, OR, July 29-August 1, 2007

Slade, M.E. 1986. "Exogeneity Tests of Market Boundaries Applied to Petroleum Products." Journal of Industrial Economics 34:291-303.

Spiller, P.T. and Huang, C. J. 1986, "On the extent of the Market: Wholesale gasoline in the northern United State". Journal of industrial economics, Vol. XXXV, No.2. (December).

Spiller, PT., and R.O. Wood. 1988. "The Estimation of Transaction Costs in Arbitrage Models." J. Econometrics 39(November 1988): 309-26.

Yacobucci, B. D. and Womach, J., 2000. "Fuel Ethanol: Background and Public Policy Issues", http://ncseonline.org/NLE/CRSreports/energy/eng-59.cfm

NY Times 2007 http://www.nytimes.com/2007/09/19/opinion/19wed1.html

MATTHEW L. WALD Published: January 16, 2006
http://www.nytimes.com/2006/01/16/national/16ethanol.html

Appendix

Residual	**S.E.**	**and**	**Cross-Correlations**						
	DPLA	DPDV	DPCR	DPCH	DPIN	DPMN	DPAL	DPHO	DPSE
Residuals	0.047445	0.051643	0.051834	0.05004529	0.048554	0.047681	0.052161	0.062675	0.046841
DPLA	1								
DPDV	0.604	1							
DPCR	0.656	0.55	1						
DPCH	0.696	0.602	0.826	1					
DPIN	0.708	0.596	0.765	0.786	1				
DPMN	0.598	0.57	0.729	0.665	0.661	1			
DPAL	0.616	0.47	0.605	0.602	0.585	0.658	1		
DPHO	0.444	0.395	0.524	0.518	0.51	0.557	0.671	1	
DPSE	0.77	0.548	0.62	0.61	0.63	0.716	0.748	0.595	1

In blue = correlation (levels);
in pink = residual S.E