Journal of Agribusiness 30, 1 (Spring 2012)

Uncovering Dominant-Satellite Relationships in the U.S. Soybean Basis: Temporal Causation and Spatial Dependence

Daniel A. Lewis, Todd H. Kuethe, Mark R. Manfredo, and Dwight R. Sanders

This study examines the degree to which market information is shared in discovering the local basis. The analysis draws from tests of temporal causation and spatial dependence of weekly county-level soybean basis values across 13 markets. Time series analysis shows that local soybean basis levels have some tendency to follow basis levels at export locations (Toledo and U.S. Gulf). Processing centers tend to show the most independence in basis discovery. Spatial statistics suggest a similar phenomenon in which basis values at interior locations are highly correlated with neighboring locations. The patterns of spatial correlation appear consistent throughout the growing season.

Key words: basis, causality, soybeans, spatial dependence

Basis values, the difference between cash and futures price, play an important role in guiding commodities through the supply chain (Tilley and Campbell, 1988; and Tomek and Robinson, 1990). For storable commodities, namely grains and oilseeds, the difference between local cash and futures prices reflects the market determined price of storage for a particular market location, encompassing physical storage costs, quality differentials, and transportation costs from the local cash market to a par delivery point. While conventional wisdom suggests that basis is discovered and determined at the local market level, grain elevators may take cues from other locations, such as terminal or export locations, and make adjustments for transportation differentials when determining and quoting the basis for their particular market location. That is, the basis may not be entirely local, as some locations provide a source of market information used in

---

Daniel A. Lewis is a Ph.D. student in the Morrison School of Agribusiness and Resource Management, Arizona State University. Todd H. Kuethe is an economist with the USDA-Economic Research Service. Mark R. Manfredo is a professor in the Morrison School of Agribusiness and Resource Management, Arizona State University, and Dwight R. Sanders is a professor in the Department of Agribusiness Economics at Southern Illinois University. The authors would like to thank the participants of the 2010 NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management for their helpful comments and discussion on this paper. The authors also gratefully acknowledge the funding by the Marley Foundation Faculty/Student Research Grant at Arizona State University which partially supported this research. The views expressed are those of the authors and should not be attributed to ERS or USDA.

determining the basis at other locations. If this is indeed the case, these grain marketing locations may play an important role in discovering and determining the basis for other markets. This is often referred to as a dominant-satellite relationship (Koontz, Garcia, and Hudson, 1990). Moreover, if several local elevators within in a particular geographical location determine their basis in a similar fashion, then there is likely to be correlation in the basis across these neighboring elevators as well.

Some evidence exists to support the idea that the grain basis may not be an entirely local concept. In examining how soybean basis levels respond to barge rate shocks and other supply and demand disruptions, McKenzie (2005) found that, for Arkansas soybean markets, Gulf basis shocks cause simultaneous movements in the Memphis basis. McKenzie (2005) also found that Memphis basis shocks influence the basis for Arkansas Delta locations (Little Rock). However, basis shocks at Little Rock did not transmit to Memphis, suggesting a dominant price discovery role for Memphis. Similarly, Haigh and Bessler (2004) found that Illinois grain prices are highly influenced by both barge freight markets and Gulf export markets. Manfredo and Sanders (2006) also provide evidence of dominant-satellite relationships in the corn basis. Their findings show that the basis at export locations and terminal market locations tend to lead the basis at interior locations. In particular, they identified the importance of Toledo, Ohio, in that the basis at Toledo tended to lead the basis at various river terminal and interior locations. While each of these studies suggest that certain market locations may be dominant markets in terms of determining the basis for other locations with respect to timing of information, these studies do not account for the potential correlation in basis among locations at a given time.

Focusing on the basis for soybeans, this study attempts to provide further insight into the discovery of the basis by examining potential dominant-satellite relationships, as well as the degree to which markets simultaneously share pricing information across space. Following the previous work of Manfredo and Sanders (2006), McKenzie (2005), and Koontz, Garcia, and Hudson (1990), we test for causal relationships through time for a number of market definitions, including export terminals, interior river locations, processing centers, and interior markets. The analysis provides relevant information on market leaders or dominant points of basis discovery. We also address the degree to which markets simultaneously share pricing information across space through the use of spatial statistics. The analysis indicates the degree to which basis levels may be correlated over space. That is, the analysis provides a measure of whether basis exhibits a systematic pattern across locations or is generated by a random process. The spatial analysis therefore provides an alternative measure of the degree to which price information is shared among neighboring locations.

While a sizeable body of research exists examining the factors impacting the basis for grains (Garcia and Good, 1983; and Naik and Leuthold, 1991) and basis forecasting

(Sanders and Manfredo, 2006; and Jiang and Hayenga, 1997), little is known about the relationships between bases realized at various market locations throughout the grain marketing system. While the basis is typically thought to reflect local supply and demand conditions, the notion that the local basis follows a dominant basis location and/or markets simultaneously share pricing information across space has considerable ramifications for how economists and market participants should approach modeling and forecasting the basis. Indeed, understanding these relationships is important as accurate basis information is critical in developing successful risk management and marketing strategies (Tomek and Peterson, 2001). It has also been suggested that economists move away from explicitly predicting prices, and focus attention on forecasting basis (Brorsen and Irwin, 1996). That is, price forecasts can be formulated using prevailing futures prices and the expected basis (Kastens, Jones, and Schroeder, 1998). Most importantly, however, this research will add to the general body of literature examining basis behavior—an important and needed avenue of inquiry for agricultural economists (Tomek and Peterson, 2001).

Methods

This research takes a two-stage approach to analyzing how information from other market locations impacts the discovery of the local basis. First, potential dominant-satellite relationships among alternative soybean market locations are examined using time series techniques, namely Granger Causality tests (Manfredo and Sanders, 2006; Koontz, Garcia, and Hudson, 1990). This analysis will provide insight into whether one or more market locations play a dominant role in discovering and determining the local basis. If the basis in one market is found to lead the market in another, this would suggest that the lagging market takes cues from the leading market in terms of how pricing information is used through time. Second, spatial statistics are used to examine the degree to which neighboring market locations simultaneously share this pricing information across space. Thus, if basis patterns are affected by dominant-satellite relationships as demonstrated in time series-based tests, these relationships should also be observed in cross-sections over space. In other words, local markets may take cues from a dominant market location when discovering and determining the basis while simultaneously sharing information with neighboring markets across space.

Time Series Analysis—Granger Causality

Granger Causality provides one approach to identify whether markets share information. In a Granger Causality framework, market X is said to Granger cause market Y if market

X provides valuable information when forecasting market Y. The method has been used to test corn basis relationships among major export markets and interior locations (Manfredo and Sanders, 2006), as well as the relationship between spot and futures prices for live cattle (Koontz and Hudson, 1990; Ollerman and Farris, 1985). The causality test is based on the equation:

$$y_t = \alpha + \sum_{i=1}^{m} \lambda_i y_{t-i} + \sum_{j=1}^{n} \theta_j x_{t-j} + \omega \tag{1}$$

where y_t is the basis value at time t in market y, x_t is the basis value at time t in market x, and m and n are the optimal lag lengths for y_t and x_t, respectively. The optimal lag lengths m and n are determined using the method proposed by Beveridge and Oickle (1994), where equation (1) is estimated for all lag values of i=1 to 12 and j=1 to 12 with the model which minimizes Akaike's Information Criteria (AIC) ultimately used in the causality test. The null hypothesis that X does not Granger cause Y is examined using a Wald test on the restriction $\theta j=0 \ \forall j$ (Hamilton, 1994). If the null hypothesis is rejected, this suggests that market X dominates market Y, or more simply, that market X plays a role in the discovery of the basis at market Y. White's test is used to test equation (1) for heteroskedasticy, and White's heteroskedastic consistent covariance estimator is used to correct the covariance matrix, if necessary.

Spatial Analysis

The concept of serial correlation is widely recognized within the empirical analysis of commodity price behavior. The phenomenon occurs as the result of information being shared across time periods, as prices in one time period influence prices in the following time periods. A similar, although increasingly complex, phenomenon can also be observed across locations – called spatial autocorrelation. In the case of the local soybean basis, when grain buyers in one location establish price bids based in part on information from buyers at other locations, prices can exhibit a systematic pattern across locations. The spatial autocorrelation can arise as a result of buyers directly sharing information, but it may also be the result of similar geographic properties at each location, such as access to transportation or natural geographic features (McNew, 1996).

Spatial autocorrelation differs from the time series equivalent in two important ways. First, spatial autocorrelation is multi-dimensional because spatial locations exist in a multi-dimensional context (of two or three dimensions). Second, spatial autocorrelation is multi-directional. With time series analysis, neighboring locations are clearly defined as the time periods which precede the current time period. However, spatial relationships may exist in an infinite set of directions. For example, a spatial observation may have neighbors directly north or directly south or in a limitless number of surrounding angles. An additional component of multi-directionality is that locations may share information in two directions. This follows the common phrase "I am my neighbor's neighbor." As a result, the identification and measurement of spatial autocorrelation is increasingly complex.

The classic measure of spatial autocorrelation is Moran's I (Moran, 1950). Moran's I measures the degree of spatially weighted deviations from the global mean. The measure examines whether neighboring locations deviate from the global mean in a systematic way. Moran's I is expressed:

$$(2) \qquad I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2}$$

where N is the number of observations, w_{ij} is a neighbor definition which is strictly positive when observations i and j share a meaningful relationship in space, and $\bar{X}$ is global mean value of the random variable X.

Moran's I is weakly bounded by -1 and $+1$, where -1 indicates perfect negative spatial dependence and $+1$ indicates perfect positive spatial dependence. When observations are perfectly negatively (positively) spatially autocorrelated, a reduction in the observed value of X at any location leads to negative (positive) changes in the value of its neighbors. Thus, under perfect negative spatial autocorrelation, as basis weakens at one location, the basis values at its neighbors will fall by the same degree. The measure takes the value of 0 for spatially independent variables. There is not a distributional assumption for the Moran's I statistic, and as a result, hypothesis testing is conducted using a bootstrap procedure based on random draws with replacement. That is, the basis values are randomly reassigned with replacement, and the distribution of the Moran's I measure for each arrangement is compared against that of the observed data.

The Moran's I statistic presented in equation (2) provides a "global" measure of spatial autocorrelation because it yields only a single value for the entire study area. The homogeneity assumed by Moran's I can, however, mask clusters of local spatial autocorrelation. In other words, a small subset of observations may be spatially autocorrelated. For example, elevators in similar locations, such as country elevators in the Midwest, may share information to a greater degree than the population at large. A variation of the Moran's I statistic, called the local Moran's I, can be estimated at each location in an effort to detect clusters of spatial autocorrelation (Anselin, 1995). The local Moran's I is calculated in a manner similar to "moving" or "rolling" estimates in time series analysis. The statistic measures the correlation between one location and its neighbors, and the local Moran's I can be calculated for each observation in the study area. Using the same notation as above, the Local Moran's I for each location *i* can be expressed:

$$I_i = \frac{N(X_i - \bar{X})}{\sum_i (X_i - \bar{X})^2} \sum_j w_{ij}(X_j - \bar{X}) \qquad (3)$$

The local Moran's I provides potentially useful information in understanding spatial patterns in basis values as it signals the degree to which certain locations may be influenced by changes in basis levels at neighboring locations. It therefore informs concerns previously addressed by time series methods, yet it adds a new interpretation. That is, a particular market or group of markets may take cues from a dominant market, yet share this information among its neighbors. Although correlation does not necessarily imply causation in the same way time series methods may, the measure indicates the potential for what Fortin, Dale, and ver Hoef (2006) call "true spatial autocorrelation" which arises from casual space-time processes. True spatial autocorrelation can be interpreted as "a clue or 'signature' left by the past action of space-time processes" (Fortin, Dale, and ver Hoef, 2006, p. 1). Thus, if basis patterns are affected by dominant-satellite relationships, as demonstrated in time series-based tests, they should also be observed in cross-sections over space.

Data

To keep the Granger Causality analysis tractable, the data are selected for 13 markets shown in Figure 1. The sample locations include the major export terminals of Louisiana Gulf and Toledo, Ohio (Lucas County); interior river locations of St. Louis, Missouri (St. Louis County) and Peoria, Illinois (Peoria County); a major soybean processing facility in Bellevue, Ohio; and several interior locations including Omaha, Nebraska; Raleigh, North Carolina; Central Illinois (Champaign County); Northwestern Iowa (Buena Vista County); Central Iowa (Hamilton County); Eastern Iowa (Black Hawk County); Central Kansas (Pawnee County); and Southeastern Indiana (Decatur County). The data are

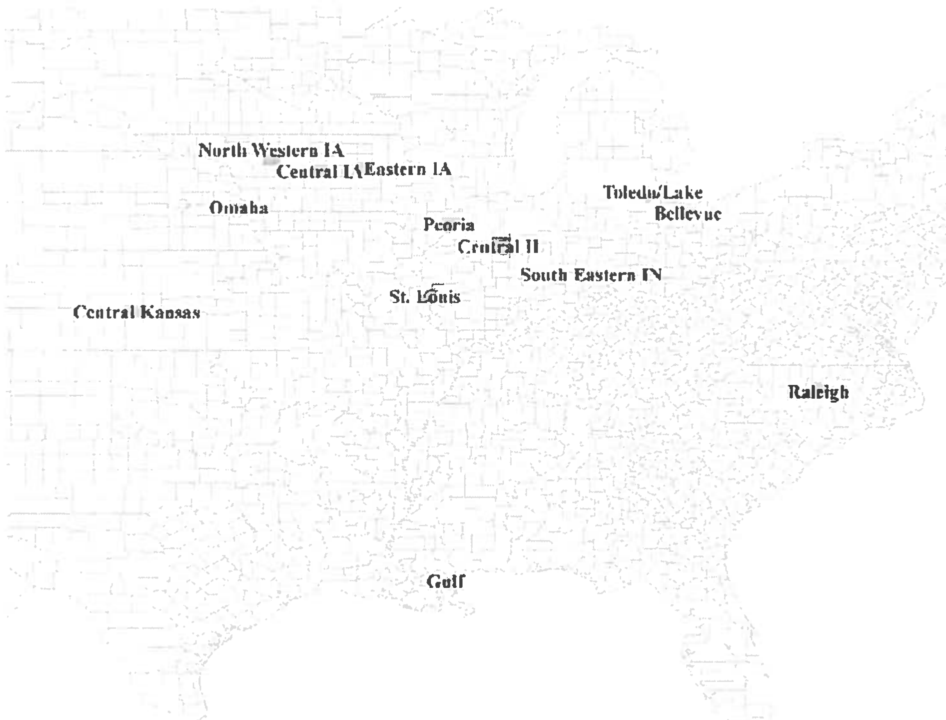

Figure 1. Market Definitions and Locations

comprised of weekly (Wednesday) nearby basis values obtained from cashgrainbids.com.[1] For the Louisiana Gulf, Omaha, and Raleigh locations, the basis data provided by cashgrainbids.com are USDA-AMS data. For all other locations except for Bellevue, the nearby basis used is an average of the basis reported at individual elevators within the county noted, with anywhere from two to eight elevators from each county composing the average. [2] Each time series spans January 2003-November 2009, providing 357 weekly observations of the basis for each location.

Given that the spatial statistics are cross-sectional in nature, the data are further aggregated for the purposes of spatial analysis. The county-level basis values are aggregated over four periods in the crop year over the years 2003 to 2009. Thus, each observation is the mean monthly basis value in each period. The four periods are defined

---

[1] In the event of missing observations for Wednesday, we selected from nearby data with the given priority: Tuesday, Thursday, Monday, or Friday.

[2] For St. Louis, Missouri, the basis data are drawn from an individual elevator in St. Louis County.

as spring (April-June), summer (July-August), fall (September-November), and winter (November-March). Thus, the degree of spatial correlation is calculated for 27 cross-sections (seasonal periods) ranging from spring 2003 to fall 2009, with each cross section containing 13 locations.

Results

Time Series Analysis

Augmented Dickey-Fuller tests (ADF) were first conducted to ensure that each of the basis series were indeed stationary in levels prior to conducting the Granger Causality tests. Following the procedure of Beveridge and Oickle (1994) the optimal lag length for the Granger Causality tests were found by estimating equation (1) for all lag combinations i=1,...,12 and j=1,....,12 and using the lag structure that minimizes Akaike Information Criterion (Akaike, 1974). In addition, we test for heteroskedasticity using White's test and apply White's consistent covariance estimator where necessary.

The Granger Causality results are reported in Table 1. Table 1 shows the information flow from the row to the column and vice-versa. For example, considering Toledo (row) and Omaha (column), the $\rightarrow$ symbol signifies that Toledo (x) leads Omaha (y), with the rejection of the null hypothesis that $\theta_j = 0 \; \forall \; j$ at the 1% level of confidence (equation 1). However, when the relationship is reversed, Omaha (x) does not lead Toledo (y), since there is a failure to reject the null hypothesis of $\theta_j = 0 \; \forall \; j$ at the 1% level. Therefore, it can be said that the direction of causality is from Toledo to Omaha. Considering Omaha, row, and Central Illinois (C IL), column, the $\leftarrow$ symbol signifies that Omaha (x) does not lead C IL (a failure to reject the null hypothesis at the 1% level), but C IL (x) does lead Omaha (y) since the null hypothesis of $\theta_j = 0 \; \forall \; j$ is rejected at the 1% level. Hence the direction of causality is from C IL to Omaha. The $\leftrightarrow$ symbol signifies two-way or simultaneous causality significant at the 1% level. For example, in the case of Peoria (row) and Raleigh (column), the null hypothesis that Peoria (x) does not cause Raleigh (y) is rejected at the 1% level. When the relationship is reversed the null hypothesis that Raleigh (x) does not cause Peoria (y) is also rejected, thus suggesting two-way or simultaneous causality. A zero (0) in any of the row/column combinations suggests that there is a failure to reject the null hypothesis in both directions, hence neither market leads the other.

Table 1. Granger Causality Results[a,b]

Gulf	Toledo	St Louis	Peoria	Bellevue	Omaha	Raleigh	C IL	NW IA	C IA	E IA	C KS	SE IN
Gulf	↔	0	↔	→	→	→	↔	↔	↔	↔	↔	→
	Toledo	↔	↔	→	→	↔	→	→	↔	↔	←	→
		St Louis	0	↔	↔	→	→	↔	→	→	→	→
			Peoria	↔	→	↔	→	↔	↔	↔	↔	↔
				Bellevue	←	↔	0	0	↔	↔	0	←
					Omaha	↔	←	0	←	←	→	↔
						Raleigh	↔	↔	↔	↔	←	↔
							C IL	0	→	↔	0	↔
								NW IA	↔	↔	→	↔
									C IA	↔	←	→
										E IA	↔	→
											C KS	0
												SE IN

[a] *Results are interpreted from row to column. For example, for Peoria (row) and Raleigh (column), there is a simultaneous causality relationship in that there is a rejection of the null hypothesis that Peoria does not lead Raleigh, and a rejection of the null that Raleigh does not lead Peoria, both at the 1% level («). Similarly, for Peoria (row) and Omaha (column), Peoria leads Omaha (®) as there is a rejection of the null that Peoria does not lead Omaha, but a failure to reject the null that Omaha does not lead Peoria. For Omaha (row) and C IA (column), C IA is found to lead Omaha (¬) at the 1% level (rejection of the null at the 1% level), but Omaha does not lead C IA (failure to reject null). A zero (0) suggests that there is a failure to reject the null in each direction (no causality).*

[b] *C IL is Central Illinois, NW IA is Northwest Iowa, C IA is Central Iowa, E IA is Eastern Iowa, C KS is Central Kansas, and SE IN is Southeast Indiana.*

Table 2 summarizes the Granger Causality results that are presented in Table 1 and provides an indication of the connectivity of the individual markets. St. Louis, Toledo, and the Gulf have the largest number of instances where the basis leads that of other markets (6, 5, and 4 respectively). For example, the export terminal of Toledo is found to lead Bellevue, Omaha, C IL, Northwest Iowa (NW IA), and Southeast Indiana (SE IN), and the export terminal of Gulf is found to lead Bellevue, Omaha, Raleigh, and SE IN. In terms of lagging markets, the Omaha market (6) exhibits the largest degree of lagging information, followed by SE IN and Bellevue respectively at 5 and 4. In addition, Raleigh and Eastern Iowa (E IA) demonstrate the greatest number of two-way information flows (↔) with nine each, followed by Peoria (8) and Gulf, NW IA, and C IA each with 7. The Bellevue market (site of a major soybean processing facility), NW IA, C IL, and Central Kansas (C KS) appear to show the least amount of connectivity as indicated by the largest number of "0's" at 3 each.

Table 2: Summary of Granger Causality Results

	Simultaneous Causality	Lead	Lag	No Causality
Gulf	7	4	0	1
Toledo	6	5	1	0
St. Louis	4	6	0	2
Peoria	8	2	0	1
Bellevue	5	0	4	3
Omaha	4	2	6	1
Raleigh	9	0	3	0
C IL[a]	4	2	3	3
NW IA	7	1	1	3
C IA	7	2	3	0
E IA	9	2	1	0
C KS	3	3	3	3
SE IN	5	1	5	1

[a] C IL is Central Illinois, NW IA is Northwest Iowa, C IA is Central Iowa, E IA is Eastern Iowa,C KS is Central Kansas, and SE IN is Southeast Indiana.

One can also observe market dynamics across the four market categories (major export terminal, interior river, processing facility, and interior locations) by examining relationships based on these average linkages. For example, each of the major interior soybean production areas of C IL, NW IA, C IA, E IA, and SE IN are either led by or share a simultaneous feedback relationship with the major export (Gulf and Toledo) and river terminal markets (Peoria and St. Louis). The export markets appear to have the greatest amount of influence on average. That is, the export markets (Gulf and Toledo) have the greatest amount of forward linkages considering both leading and simultaneous relationships. Further, Omaha exhibits a lagging relationship, as demonstrated by the greatest number of lagging relationships, as well as independent relationships. Finally, the interior river (St. Louis and Peoria) and interior markets exhibit the greatest amount of combined forward and backward price transmissions. Overall, these results suggest that export markets tend to display dominate relationships with other markets displaying satellite behavior in terms of how basis is ultimately discovered and determined. This suggests that export bids maybe systematically bid back through the marketing channel.

Spatial Analysis

The measures of spatial autocorrelation are constructed using an inverse distance neighbor definition such that all observations are tied, yet the relative influence of each observation declines as the distance between two points increases. Distance is measured by the arc distance between county centroids (the geographical center of each county). The Moran's I test results are reported in Table 3. The first column (Global) reports the global Moran's I outlined in equation (2). The test statistics are statistically significant and positive across all periods except spring and fall of 2004. This suggests that basis values at each location are positively influenced by basis values at other locations. For example, when the basis strengthens, or for that matter weakens, it is doing so systematically across all of the market locations but to a varying degree.

The remaining columns report the local Moran's I test results (equation 3) for the thirteen individual market definitions. As previously stated, the local Moran's I statistic is used to detect clusters of high spatial autocorrelation. For example, the first row of Table 3 reports the local Moran's I estimate at each location for Spring 2003. The results detect five markets with statistically significant levels of local spatial autocorrelation: Raleigh, C KS, and all three regions of Iowa (NW IA, C IA, and E IA). Local spatial autocorrelation appears to be especially strong among the three Iowa regions in terms of the magnitude of the Moran's I, which range anywhere from approximately 0.30 to 0.69 across the time periods.

Looking across each time period, the results consistently suggest a high degree of spatial clustering in interior locations, and the patterns appear consistent across the seasonal time periods defined. Again, the markets which exhibit the highest degree of local spatial autocorrelation include Raleigh, C KS; NW IA, C IA, and E IA. For example, the local Moran's I for C IA was statistically significant in 20 of the 27 time periods, and is statistically significant for 19 of the 27 time periods for NW IA, E IA, and C KS respectively. Overall, these results suggest that there is a considerable amount of basis information being shared among these neighboring interior locations. As suggested by Fortin, Dale, and ver Hoef (2006), the spatial autocorrelation tests also support the temporal causality tests. Indeed, the markets with the highest degree of local spatial autocorrelation also exhibit a large degree of temporal causality, as reported in Table 2. This suggests that there is a high degree of basis information being shared by neighboring interior market locations, and, at the same time, these locations may also be led by, or maintain a simultaneous feedback relationship with, the dominant markets such as Gulf and Toledo.

Table 3. Spatial Autocorrelation Test Results

		Market												
	Global	Gulf	Toledo	St. Louis	Peoria	Bellevue	Omaha	Raleigh	C IL	NW IA	C IA	E IA	C KS	SE IN
2003														
Spring	0.164***	-0.153	0.089	-0.124	0.035	0.142	0.050	0.225**	-0.019	0.511**	0.572***	0.510***	0.309**	-0.023
Summer	0.196***	-0.292	0.089	-0.158	0.075	0.172	0.219	0.251**	-0.078	0.571***	0.685***	0.599***	0.342**	0.067
Fall	0.111***	-0.565	-0.007	-0.171	0.122	-0.054	0.332	0.176	-0.016	0.489**	0.501**	0.461**	0.273**	-0.100
Winter	0.047*	-0.305	0.018	-0.066	0.008	0.043	0.027	0.101	0.017	0.254	0.270	0.228	0.005	0.004
2004														
Spring	-0.021	-0.389	0.010	-0.018	0.017	0.068	0.018	0.148	-0.017	0.005	0.030	0.076	-0.226	0.010
Summer	0.049*	-0.311	-0.154	-0.022	0.077	-0.072	0.249	0.362***	-0.030	-0.068	0.010	0.160	0.449***	-0.014
Fall	0.017	-0.700	0.007	-0.103	0.130	0.025	0.191	0.147	0.022	0.038	0.118	0.234	0.224*	-0.116
Winter	0.057**	-0.333	-0.008	-0.101	-0.015	0.004	0.148	0.093	0.001	0.285	0.311*	0.267*	0.092	0.001
2005														
Spring	0.079**	-0.242	0.019	-0.101	0.031	0.049	-0.003	0.193*	0.001	0.266	0.338*	0.324*	0.117	0.031
Summer	0.098**	-0.339	0.045	-0.064	0.129	0.073	-0.369	0.160*	-0.050	0.375*	0.529**	0.539***	0.226*	0.019
Fall	0.102***	-0.657	0.025	0.044	0.262	0.043	0.052	0.097	-0.031	0.340*	0.445**	0.502***	0.227*	-0.019
Winter	0.099***	-0.363	-0.002	-0.132	0.032	0.025	0.221	0.150	-0.003	0.399**	0.393**	0.354**	0.220*	-0.005
2006														
Spring	0.146***	-0.332	0.031	-0.079	0.065	0.088	0.216	0.233**	-0.012	0.433**	0.446	0.444**	0.327**	0.033
Summer	0.174***	-0.425	0.040	-0.112	0.110	0.098	0.327*	0.202*	-0.009	0.548**	0.586***	0.576***	0.306**	0.013
Fall	0.140***	-0.353	0.007	-0.174	0.050	0.038	0.252	0.123	-0.007	0.540**	0.594***	0.557***	0.223*	-0.034
Winter	0.125***	-0.282	0.000	-0.176	0.033	0.023	0.145	0.134	-0.001	0.496**	0.548***	0.482**	0.217*	0.013
2007														
Spring	0.196***	-0.147	0.076	-0.094	0.025	0.132	0.254	0.296**	0.003	0.500**	0.540	0.499***	0.376***	0.094
Summer	0.207***	-0.288	0.037	-0.034	0.093	0.091	0.334	0.175*	-0.021	0.662***	0.690***	0.633***	0.298**	0.025
Fall	0.176***	-0.210	0.000	-0.098	0.019	0.013	0.398	0.157	-0.027	0.653***	0.609***	0.476**	0.276**	0.016
Winter	0.143***	-0.096	0.000	-0.098	-0.016	-0.047	0.330	0.169	0.017	0.486	0.463	0.352	0.251	0.041
2008														
Spring	0.195***	-0.216	0.119	-0.025	-0.005	0.199	0.306	0.331	0.016	0.463**	0.457**	0.394**	0.399	0.095
Summer	0.188***	-0.195	0.134	-0.022	-0.004	0.201	0.077	0.295**	0.003	0.476**	0.554*	0.500*	0.362***	0.070
Fall	0.044*	-0.561	-0.056	0.017	0.027	-0.046	-0.091	0.084	0.010	0.309*	0.353*	0.337*	0.196	-0.001
Winter	0.051**	-0.322	0.007	-0.017	-0.025	0.037	-0.121	0.077	0.012	0.298	0.375*	0.306*	0.011	0.023
2009														
Spring	0.167***	-0.143	0.071	-0.052	-0.017	0.123	-0.023	0.212*	0.018	0.575***	0.632***	0.443	0.265**	0.064
Summer	0.083**	-0.235	-0.035	-0.049	-0.144	-0.008	0.126	0.067	-0.021	0.451**	0.425**	0.283	0.212*	0.006
Fall	0.064**	-0.302	-0.070	-0.019	-0.104	-0.033	0.109	0.120	0.000	0.365*	0.309*	0.241	0.265**	-0.055

a. *** represents statistical significance at the 1% level, ** 5% level, and * 10% level respectively.

b. C IL is Central Illinois, NW IA is Northwest Iowa, C IA is Central Iowa, E IA is Eastern Iowa, C KS is Central Kansas, and SE IN is Southeast Indiana.

Conclusions

Our analysis demonstrates the degree to which soybean markets share information in determining basis levels over both time and space. We define four market categories: major export terminals, interior river locations, soybean processing facilities, and interior locations. The evaluation is conducted through time series analysis and spatial econometrics. The time series analysis consists of Granger causality tests of weekly basis values in each market over the period January 2003 – November 2009. The results suggest that export markets may play a dominant role in basis discovery at other locations.

The spatial analysis, on the other hand, examines the degree of simultaneous spatial spillovers in observed basis over the period January 2006 – December 2009. The spatial analysis addressed four periods in the soybean crop year: spring, summer, fall, and winter. The results suggest that the basis values are globally spatially dependent as a result of positive spatial autocorrelation. The local measures of spatial autocorrelation also suggest that interior locations exhibit the greatest degree of spatial association.

Collectively, the results suggest a dominant-satellite relationship where the export markets are the dominant markets in terms of discovering the basis. This information is then transmitted to the satellite markets, predominantly those located at interior or origination points in the marketing channel. Moreover, the fact that neighboring interior locations are responding to changes in the same dominant basis creates local spatial correlation for neighboring elevators. This is evidenced by the relatively high local (spatial) correlations within, for example, the interior Iowa locations. Then, those interior locations have causal (temporal) relationships with export locations such as the Gulf or Toledo.

This research has both important academic and practical business implications. Namely, the findings suggest that basis modeling efforts need to include the potential for dominant-satellite effects. Efforts to model and forecast the interior Illinois basis for soybeans, for example, should include lagged values of the basis at major export locations (Gulf or Toledo) or river terminal (St. Louis or Peoria). Ignoring the information flow from these locations could result in a misspecified model. Basis modeling efforts should also include ways to account for the simultaneous sharing of basis information among neighboring regions. Improved basis modeling is not only an academic issue. Indeed, incorporating the findings from this study into current basis forecasting efforts has the potential to improve basis forecasts, especially as the behavior of the basis, namely the convergence of cash and futures, has become more difficult to predict in recent years (Irwin et al., 2004). Ultimately, improved basis forecasts help shape the marketing and risk management strategies implemented by farmers, elevators, processors, and others along the soybean supply chain.

Recognizing the dominant-satellite relationships and the spatial connectivity among markets should also provide managers of grain merchandising firms additional information in terms of how they form their individual cash bids. This may be particularly relevant for medium to small operations that do not get price information from a centralized office. Basis information from dominant and/or neighboring locations can be used as additional information in determining their bids, thus creating a more efficient bid in terms of information content that reflects both local supply and demand factors, as well as more general market information that is being transmitted among neighboring locations and/or through dominant market locations such as export markets. The findings from this research should also be of use to soybean farmers as it provides them additional insight into how basis is formed at their local elevators, allowing farmers to potentially make more informed marketing decisions. Insights into how soybean basis information is transmitted to local markets through dominant/satellite relationships, as well as how the information is shared among neighboring locations, should also be of particular interest to commodity exchange officials as the recent concerns with basis non-convergence threatens to reduce the hedging effectiveness of the soybean contract, as well as other futures contracts for storable commodities. In summary, this research helps to broaden the literature and our general understanding of basis behavior – a critical area of inquiry in agribusiness (Tomek, 1993; and Tomek and Peterson, 2001).

References

Anselin, L. (1995). "Local Indicators of Spatial Association – LISA." *Geographical Analysis* 27, 93-115.

Beveridge, S. and C. Oickle. (1994). "A Comparison of Box-Jenkins and Objective Methods for Determining the Order of a Non-Seasonal ARMA Model." *Journal of Forecasting* 13, 419-434.

Brorsen, B.W. and S.H. Irwin. (1996). "Improving the Relevance of Research on Price Forecasting and Marketing Strategies." *Agriculture and Resource Economics Review* 25, 68-75.

Fortin, M., M.R.T. Dale, and J. ver Hoef. (2006). "Spatial Analysis in Ecology" in *Encyclopedia of Environmetrics*, Wiley.

Garcia, P. and D.L. Good. (1983). "An Analysis of the Factors Influencing the Illinois Corn Basis, 1971-1981." Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Des Moines, Iowa. Online. Available at http://www.farmdoc.uiuc.edu/nccc134.

Haigh, M.S. and D.A. Bessler. (2004). "Causality and Price Discovery: An Application of Directed Acyclic Graphs." *Journal of Business* 77, 1099-1121.

Hamilton, J.D. (1994). *Time Series Analysis*. New Jersey; Princeton University Press.

Irwin, S.H., P. Garcia, D.L. Good, and E.L. Kunda. (2009, March). "Poor Convergence Performance of CBOT Corn, Soybean and Wheat Futures Contracts: Causes and Solutions." Marketing and Outlook Research Report 2009-02, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.

Jiang, B., and M. Hayenga, M. (1997). "Corn and Soybean Basis Behavior and Forecasting: Fundamental and Alternative Approaches." Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, Illinois. Online. Available at http://www.farmdoc.uiuc.edu/nccc134.

Kastens, T.L., R. Jones, and T.C. Schroeder. (1998). "Futures-Based Price Forecasts for Agricultural Producers and Businesses." *Journal of Agricultural and Resource Economics* 23, 294-307.

Koontz, S.R., P. Garcia, and M.A. Hudson. (1990). "Dominant-Satellite Relationships Between Live Cattle and Cash Futures Markets." *The Journal of Futures Markets* 10, 123-136.

Manfredo, M. R., and Sanders D. R. (2006). "Is the Local Basis Really Local?" Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, Missouri. Online. Available at http://www.farmdoc.uiuc.edu/nccc134.

McKenzie, A. M. (2005). "The Effects of Barge Shocks on Soybean Basis Levels in Arkansas: A Study of Market Integration." *Agribusiness: An International Journal* 21, 37-52.

McNew, K. (1996). "Spatial Market Integration: Definition, Theory and Evidence." *Agricultural and Resource Economics Review* 25, 1–11.

Moran, P. (1950). "Notes on Continuous Stochastic Phenomena." *Biometrika* 37, 17–33.

Naik, G. and R.M. Leuthold. (1991). "A Note on the Factors Affecting Corn Basis Relationships." *Southern Journal of Agricultural Economics* 23, 147-153.

Ollerman, C.M. and P.L. Farris. (1985). "Futures or Cash: Which Market Leads Beef Cattle Prices?" *The Journal of Futures Markets* 5, 529-538.

Sanders, D.R. and M.R. Manfredo. (2006). "Forecasting Basis Levels in the Soybean Complex: A Comparison of Time Series Methods." *Journal of Agricultural and Applied Economics* 38, 513-523.

Tilley, D.S. and S.K. Campbell. (1988). "Performance of the Weekly Gulf-Kansas City Hard-Red Winter Wheat Basis." *American Journal of Agricultural Economics* 70, 929-935.

Tomek, W.G. (1993). "Dynamics of Price Changes: Implications for Agricultural Futures Markets," Research Frontiers in Futures and Options Markets: An Exchange of Ideas, Proceedings from the Symposium in Recognition of Thomas A. Hieronymus, Office for Futures and Options Research, University of Illinois at Urbana-Champaign, pp. 45-55.

Tomek, W.G. and H.H. Peterson. (2001). "Risk Management in Agricultural Markets: A Review." *Journal of Futures Markets* 21, 953-985.

Tomek, W.G. and K.L. Robinson. (1990). *Agricultural Product Prices*, 3rd ed. N.Y.: Cornell University Press.