Causality and Price Discovery:
An Application of
Directed Acyclic Graphs

by

Michael S. Haigh and David A. Bessler

WP 02-10

Department of Agricultural and Resource Economics

The University of Maryland, College Park

| **Current Title:** | Causality and Price Discovery: An Application of Directed Acyclic Graphs |

| **Authors:** | Michael S. Haigh[a], University of Maryland and David A. Bessler[b], Texas A& M University |

| **Abstract:** | Directed Acyclic Graphs (DAG's) and Error Correction Models (ECM's) are employed to analyze questions of price discovery between spatially separated commodity markets and the transportation market linking them together. Results from our analysis suggest these markets are highly interconnected but it is the inland commodity market that is strongly influenced by both the transportation and commodity export markets. However, the commodity markets affect the volatility of the transportation market over longer horizons. Our results suggest that transportation rates are critical in the price discovery process lending support for the recent development of exchange traded barge rate futures contracts. |

| **Key Words:** | Barge Rate Futures, Directed Acyclic Graphs, Causation, Integration. |

| **JEL Classification:** | C32, G14 |

| **First Draft:** | July 2001 |
| **Second Draft:** | May 2002 |

Michael S. Haigh[a] (corresponding author)
Assistant Professor
Dept. of Ag & Resource Economics
University of Maryland
College Park, MD 20742
(301) 405 – 7205 (Tel)
(301) 314 – 9091 (Fax)
mhaigh@arec.umd.edu

David A. Bessler[b]
Professor
Department of Ag Economics
Texas A&M University
College Station, TX 77840
(979) 845 – 3096 (Tel)
(979) 845 – 9769 (Fax)
d-bessler@tamu.edu

Causality and Price Discovery: An Application of Directed Acyclic Graphs

Directed Acyclic Graphs (DAG's) and Error Correction Models (ECM's) are employed to analyze questions of price discovery between spatially separated commodity markets and the transportation market linking them together. Results from our analysis suggest these markets are highly interconnected but it is the inland commodity market that is strongly influenced by both the transportation and commodity export markets. However, the commodity markets affect the volatility of the transportation market over longer horizons. Our results suggest that transportation rates are critical in the price discovery process lending support for the recent development of exchange traded barge rate futures contracts.

I   Introduction

To date, a large amount of research has been undertaken to evaluate the extent to which spatially separated markets are integrated. The popularity of the subject matter is driven in part by the fact that finding continual deviations from the equilibrium level of integration might imply that riskless arbitrage opportunities exist. However, despite the fact that freight rates are notoriously volatile, and the fact that over 5.5 billion tonnes, or 98% of annual world trade is carried by sea, the role of the transportation market in testing for integration within the marketing channel has been largely ignored in the literature.[1] A few notable exceptions do exist. For instance, important research undertaken by Geraci and Prewo (1977) confirmed that it is vital to include transportation costs in the study of integration among spatially separated markets. Goodwin, Grennes and Wohlegant (1990) conclude that failing to account for volatile freight rates can lead to erroneous conclusions in empirical trade research. They carefully demonstrate this point by finding stronger support for the Law of One Price only after they accounted for shipment costs.

Only a handful of studies have directly isolated the effect that volatile freight prices might have on the price discovery process. These include Haigh and Holt (2000), Hauser and Neff (1990) and Haigh and Bryant (2001). While the first contribution emphasized the importance of ocean freight volatility within the marketing channel, it was the latter contributions that isolated the extent to which domestic freight volatility (specifically volatile barge rates) contributed to the level uncertainty. However, both studies failed however to discuss in any detail exactly how the prices were linked and did not assess in any detail issues relating to causality among the markets.

Because of the importance of transportation rates in the price discovery process, there has always been considerable amount of interest in developing a forward/futures market for transportation services (Hauser and Buck, 1989). Indeed, in May of 1985 the BIFFEX freight futures contract was launched at the London International Financial Futures Exchange (LIFFE). The contract, based off an index of shipping prices compiled by the Baltic Exchange was designed to hedge freight price risk in the dry-bulk sector of the ocean shipping markets. Indeed, because of its uniqueness (it was the only futures contract on a service) and because of its potential importance, several researchers have investigated its use from a hedging standpoint. Examples include Thuong and Visscher (1990), Kavussanos and Nomikos (1999, 2000) and Haigh and Holt (2000). These studies invariably conclude that the BIFFEX market is not a particularly effective hedging instrument and does not provide the price risk protection evidenced in other futures markets. Each concludes that its weak performance as a hedging instrument is due to the fact that the contract was based on an index of shipping routes making the

hedge less appealing and hence the trading volume lower. As anticipated, in June 2001

LIFFE announced that trading in the BIFFEX contract would cease in April 2002 because

of low trading volumes.[2] It seems therefore that there is no way to predict with any

degree of accuracy whether or not a new futures contract will be successful, but it may be

possible however to provide some quantitative indicators of how important that market is

likely to be. This could be especially true if that market is influential or highly

influenced in the price discovery process within a marketing channel and causes a

significant source of risk for traders in that market.[3]

The current study makes contributions to the issue of evaluating the potential

success of a futures contract from several angles. Using recent high frequency price data

we adopt a new framework to analyze the relationship between inland grain prices in

Illinois, export grain prices at the U.S. Gulf and the barge market that links them

together. In particular, we employ Directed Acyclic Graph (DAG) theory which, to date,

has been surprisingly underutilized in both the economics and finance fields.[4] The

unique methodology allows us to examine the causal pattern of contemporary

relationships among the innovations in the three markets, based off of the familiar Error

Correction Model (ECM). Critically, our DAG analysis allows us to address the

construction of the data-determined othoganization on contemporaneous innovation

covariance, critical in providing sound inference in innovation accounting (Swanson and

Granger, 1997) and is used here to assess the dynamic relationships between these

markets. From a practical standpoint, this assessment of both the degree of

interconnectivity and direction of causation within the marketing channel will be of

practical use to physical traders in this marketplace and be of use to exchange administrators interested in assessing the potential role and success of a new derivative contract.

Therefore it is also the objective of this study to focus on the importance of the barge market and explain in detail its role in the price discovery mechanism of the export marketing channel. Indeed, data provided by the United States Federal Grain Inspection Service over the period May 6th 1999 – May 3rd 2001 (the same time period analyzed in this study) suggest a priori, the relative importance of the barge market. For instance, the total amount of grain exported out of the U.S. within this time period was 258.84 million tonnes of grain on 16586 different vessels to a total of 131 different countries. Of that total number, 134.26 million tonnes (51.9% of the total) was shipped out of the U.S. Gulf (the vast majority of which originated via barges along the Mississippi River) on 7187 different vessels to 101 different countries. [5] Isolating the importance of the barge market is of particular interest here simply because trading in barge futures contracts for the particular stretch of river analyzed in this study began at the St. Louis Merchants exchange in December of 2000. [6]

The rest of the paper is organized as follows. Section II provides an overview of the econometric methodologies employed in the paper. Section III describes the data, and Section IV presents the empirical results. The last section, Section V, concludes.

II   Econometric Methods

A considerable amount of research has attempted to evaluate the degree of interconnectivity between markets employing time-series techniques appropriate for non-

stationary and cointegrated data. In particular, much work on applied cointegration analysis has relied on Johansen's multivariate approach (Johansen, 1988, 1991; Johansen and Juselius, 1990) and readers are directed to these papers for a more complete explanation of the methodology. Examples of papers employing such techniques include Chowdhury (1991) and Goodwin and Piggott (2001).

First, assume an n-dimensional vector of nonstationary time series, Y_t, ($n = 3$ here). Assuming all variables are non-stationary and the existence of cointegration, then as shown by Johansen and Juselius (1990) the data generation process can be written as error-correction representation as follows:

$$\Delta Y_t = \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \Pi Y_{t-1} + e_t \, . \tag{1}$$

Equation (1) is nothing more than Vector Autoregression (VAR) (in first differences), with an inclusion of the lagged-level component, which is known as the Error Correction Term. The combination is simply known as an Error Correction Model (ECM). Π is a (3x3) coefficient matrix that may be factored as ab' if cointegration is present. The b matrix is a matrix of cointegrating parameters and the matrix a is a matrix of weights (also known as the speed of adjustment parameters) with which each cointegrating vector enters the n equations. Hypothesis tests can also be conducted on these matrices. However, as is the case of standard VAR's, the individual coefficients associated with the ECM can be somewhat difficult to interpret, particularly those associated with the short-run dynamics captured within Γ_i.[7] Consequently, innovation accounting techniques may be the best way to describe the structure and

interdependencies among the prices within the export marketing channel (Swanson and Granger, 1997). Therefore, given the ECM, impulse response analysis can be undertaken (based on an equivalent levels VAR) to summarize the dynamic interrelationships among the prices. Undertaking the impulse response analysis in this way addresses the necessity of imposing the cointegrating relationships into the system, which has very recently been proven to be crucial in yielding consistent impulse responses and forecast error decompositions (Philips, 1998).

However, the basic problem of the orthoganalization of residuals from the ECM remains somewhat unresolved. Most studies employing ECM or VAR's have yet to fully address the problem associated with the contemporaneous relationships among variables. Despite this, innovation accounting techniques require that a causal assumption about contemporaneous correlation be made. Early work in this area employed the Choleski factorization, with more recent applications concentrating on a 'structural' factorization suggested by Bernanke (1986) and Sims (1986) simply because researchers may not view the world as being recursive (Cooley and Leroy (1985)). However, the problem with both the Bernanke (1986) and Sims (1986) approach is that it is assumed that the researcher has knowledge of the correct structural model (which is unlikely to be the case). Directed graphs offer a solution to this problem.

Directed Acyclic Graphs

The literature on Directed Acyclic Graphs is extensive. We refer the reader to Pearl (2000) and Spirtes, Glymour and Scheines (2000) for treatments. Briefly, a directed graph is a picture representing causal flow between or among a set of variables.

We use letters $(V_1, V_2, \ldots, V_n)$ to represent variables and lines (these lines are also called edges) with arrowheads at one end representing causal flows (e.g. $V_1 \rightarrow V_2$ is used to indicate V_1 causes V_2). We focus on graphs with directed edges (the edge in the representation $V_1 \rightarrow V_2$ is called a directed edge). More elaborate graphs involving undirected edges and bi-directed edges, while not resulting in directed graphs, are studied in the references cited above. Further, we will be concerned with paths (a path is a series of variables connected by line segments) containing no cycles; where a cyclic path begins with a variable, say V_1 and eventually returns to V_1 (e.g., $V_1 \rightarrow V_2 \rightarrow V_1$). Graphs with no cycles are said to be acyclic. Here we are interested in independence and conditional independence constraints put on variables under alternative causal flows between variables. Directed graphs provide the visual representation of that flow; the set of independence or conditional independence conditions which are implied by that graph are not (necessarily) obvious.

Pearl (1986) proposes d-separation as a graphical characterization of the conditional independence relations entailed by a graph. Verma and Pearl (1988) give a proof of this proposition. D-separation is understood using three subsets of variables A, B and C in a graph, *G*. For any path, say p between a variable in set A and a variable in set C, the set B is said to block the path p if there is a variable V_i on the path satisfying one of the following: (i) variable V_i has converging arrow along the path and neither V_i or any of its descendants are in B; or (ii) V_i does not have converging arrows along path p and V_i is a member of set B. Set B is said to d-separate A from C on graph *G* if and only if B blocks every path from a variable in set A to a variable in set C.

The conditional independence restrictions can thus be read off of the graph by merely asking whether two variables are d-separated in a particular graph. So, for example if we have a causal chain on four variables, $X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4$, and if we use correlation (denoted by $\mathbf{D}$) to measure dependence; d-separation tells us that the following conditions hold[8]:

$\mathbf{D}(X_i , X_j) \ldots 0; i,j = 1,2,3,4$

$D(X_1 , X_2 | X_j) \ldots 0 , j = 3,4 ; \quad D(X_j , X_3 | X_4) \ldots 0 , j = 1,2 ; \quad D(X_2 , X_j | X_1) \ldots 0 , j = 3,4;$

$D(X_3 , X_4 | X_j) \ldots 0 , j = 1,2 ; \quad D(X_1 , X_2 | X_3 , X_4) \ldots 0$

$D(X_1 , X_3 | X_2) = 0 ; D(X_1 , X_4 | X_3) = 0 ; D(X_1 , X_4 | X_2) = 0 ; D(X_2 , X_4 | X_3) = 0.$

Recently Swanson and Granger (1997) examined the contemporaneous relationships among innovations (residuals) from a vector autoregression. Under a maintained hypothesis that causal flow on innovations follows a causal chain, they consider the constraints on partial correlation (on observed innovations) and the use of ordinary least squares (ols) regression for testing these constraints. In particular, their Theorem 3.1 (page 361) makes three important points. First, they establish the relationship between vanishing ols estimates and vanishing partial correlation between observed innovations from a first stage VAR, when the original innovations are generated as a (causal chain) recursive ordering. That is to say, if we use ols regression of observed innovations $\hat{u}_{it}$ on observed innovations $\hat{u}_{jt}$ and $\hat{u}_{kt}$, where a recursive ordering generates the "true" innovations, $u_{kt} \rightarrow u_{jt} \rightarrow u_{it}$, the coefficient associated with $\hat{u}_{kt}$, call this $\bullet_k{}^*$, will be zero if and only if the partial correlation between the true u_{it} and u_{kt} given u_{jt} equals zero. Second they show the asymptotic equivalence of ols estimates based on observed innovations and ols estimates based on true innovations for this recursive

ordering. So a regression of observed innovations from series i ($\hat{u}_{it}$) on observed innovations from series j ($\hat{u}_{jt}$) and observed innovations from series k ($\hat{u}_{kt}$) will yield asymptotically valid inference on the relationship between true innovations u_{it}, u_{jt}, and u_{kt}. Finally, they point-out the problem with using ols regression to make inferences on other partial correlation constraints, for the recursive flow ($u_{kt} \rightarrow u_{jt} \rightarrow u_{it} \rightarrow u_{ht}$) due to the inconsistency of ols. That is to say, this recursive causal flow suggests that the partial correlation between $\hat{u}_{kt}$ and $\hat{u}_{jt}$ given $\hat{u}_{ht}$ should be non-zero (see results given above based on d-separation) . However an ols regression of $\hat{u}_{kt}$ on $\hat{u}_{jt}$ and $\hat{u}_{ht}$ will be inconsistent, since innovations on h are themselves a function of innovations on k. Below we use the partial correlation (based on Spirtes, Glymour and Scheines (1993)) and not ols regression to test for causal relations between observed innovations in series.

Spirtes, Glymour and Scheines (1993) have applied d-separation in an algorithm (PC Algorithm) for building directed graphs. PC algorithm is 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." The algorithm is described in detail in Spirtes, Glymour, and Scheines (1993, p.117).[9]

Briefly, the algorithm (we will summarize only the generic aspects of PC algorithm) begins with a complete undirected graph G on the vertex set **X**. The complete, undirected, graph shows an undirected line between every variable of the system (every variable in **X**). Lines between variables are removed sequentially based on zero correlation or partial correlation (conditional correlation). *The conditioning variable(s)*

on removed lines between two variables is called the sepset of the variables whose line

has been removed (for vanishing zero order conditioning information the sepset is the

empty set). Edges are directed by considering triples X — Y — Z, such that X and Y are

adjacent as are Y and Z, but X and Z are not adjacent. Direct lines between triples: X

— Y — Z as X → Y ← Z if Y is not in the sepset of X and Z. If X → Y, Y and Z are

adjacent, X and Z are not adjacent, and there is no arrowhead at Y, then orient Y — Z as

Y → Z. If there is a directed path from X to Y, and a line between X and Y, then direct

(X — Y) as: X → Y.

In applications, Fisher's z is used to test whether conditional correlations are

significantly different from zero. Fisher's z can be applied to test for significance from

zero; where:

$$z(\boldsymbol{r}(i,j\mid k),n) = \left[\frac{1}{2}\sqrt{n-\mid k\mid -3}\right]\ln\left\{\frac{\mid 1+\boldsymbol{r}(i,j\mid k)\mid}{1-\boldsymbol{r}(i,j\mid k\mid)}\right\}, \tag{6}$$

and *n* is the number of observations used to estimate the correlations, $\boldsymbol{r}(i,j\mid k)$ is the

population correlation between series *i* and *j* conditional on series *k* (removing the

influence of series k on each *i* and *j*), and |*k*| is the number of variables in *k* (that we

condition on). If *i,j* and *k* are normally distributed and *r(i,j|k)* is the sample conditional

correlation of *i* and *j* given *k*, the distribution of $z(\boldsymbol{r}(i,j\mid k),n) - z(r(i,j\mid k),n)$ is

standard normal. PC algorithm and its more refined extensions are marketed as the

software TETRAD II (Scheines, et al 1994).

As an alternative to PC algorithm we also investigate direct scoring on alternative

causal models on observed innovations. In our problem under study we have three

variables: innovations on barge rates, soybean prices in Illinois and soybean prices at the Gulf of Mexico. This gives us a manageable number (25) of alternative acyclic graphs, each of which can be scored using one of several statistical loss func tions. It is of interest to see if the DAG selected by statistical scoring results in the same model as PC algorithm.

III   **Description of the Data**

The data for this study cover a two-year time period, from May 6th, 1999 to May 3rd 2001, totaling 520 daily observations for each of the time-series. The mid point of the original daily closing Illinois and Gulf soybean bid prices were provided by the Illinois Department of Agriculture and the USDA Agricultural Marketing Service respectively. Grain barge rate data covering the same period were also collected for the stretch of river beginning south of Peoria. Specifically, first, weekly barge rate information was collected from the USDA, Agricultural Marketing Service, Transportation and Marketing Division. This weekly rate (Wednesday quote) reflects the current rate as a percent of the historic benchmark tariff rate (southbound barge freight call session basis trading benchmark (July 1979)). From this figure the dollar per ton rate was obtained by multiplying the quoted rate (a percentage of the benchmark rate) by the historic benchmark rate associated with the south of Peoria region. Such a data series was used by Haigh and Bryant (2001). Daily rate data were also collected from a large grain trading company that transports grain on a daily basis along this stretch of the river. The data cannot be shared for confidentiality reasons, but to ensure its reliability the Wednesday daily price from the grain trading company was compared with the USDA

price. Because both series were found to be almost identical the daily grain and daily barge rates were used throughout.[10]

Summary statistics on all the prices are presented in Table 1. As one might expect, the average (mean) grain price at the Gulf is greater than that at Illinois, with the difference being slightly greater than the barge rate linking the two together. Indeed, Figure 1 Panel A plots the daily grain prices at Illinois and the Gulf, and the spread (Gulf – Illinois) and the barge rate are presented in Panel B. This graph also illustrates the strong degree of interconnectivity between these price series. As can be seen in Table 1 the degree of volatility varies among the price series with the grain price series exhibiting identical levels of uncertainty (as measured by the Coefficient of Variation). Interestingly, barge rate volatility is several times as great as the grain price volatility. Haigh and Byrant (2001) also found the excess volatility found in this market (relative to other markets). The discussion above indicates that the barge prices and the grain prices are linked together. However, it does not provide detailed evidence on the dynamics of these linkages as well as on the existence of causation among them. It is those issues that we now turn to.

IV   Empirical Application

Estimation of the Error Correction Model

In order to implement our ECM we first check the order of integration of each of the price series within the marketing channel. As can be seen from Table 2, each series is integrated of the first order confirming that the analyses will be conducted on the differenced price series. The ECM was then estimated using the maximum likelihood

technique outlined by Johansen and Juselius (1990). The lag length order was selected based on the Schwarz-loss criterion, (as shown in Table 3).[11] Table 3 (top panel) presents the decision rule based on the trace tests for the number of cointegrating vectors. Using critical values provided by Osterwald-Lenum (1992) we first fail to reject the null hypotheses on r ≤ 1and so the ECM is modeled with one cointegrating vector. Residuals from this estimated ECM seem to be well behaved. Specifically, Lagrangian Multiplier tests for first and fourth order autocorrelation are calculated using an auxiliary regression as described by Godfrey (1988, Chapter 5). We reject both first and fourth orders of autocorrelation comfortably with p-values of 0.74 and 0.20 respectively. We do however detect some evidence of remaining ARCH effects and reject the assumption of multivariate normality (at the 1 percent level of significance). While ARCH or GARCH effects and deviations from multivariate normality do not seem to adversely affect inference on cointegration analysis (Gonzalo (1994) and Lee and Tse (1996)), it is unclear of their influences on DAG analysis.[12]

We also test the individual elements of b' against zero in the factorization $ab' = \Pi$ and investigate the possibility of weak exogeneity of each of the series (testing whether each element of the a vector is equal to zero) and the results are presented in the lower portion of Table 3. In particular, the middle panel explores the possibility that one of the three series is not in the cointegrating space. We firmly reject the null that price i is not present in the cointegrating space for each series. With respect to the short-run adjustment toward the long run relationships, a, we also test for weak exogeneity on each market. For each market we test for whether or not it responds to perturbations in

the cointegrating space. Inspection of the lower panel of Table 3 suggests that both the Illinois and Gulf markets are weakly exogenous and the barge market does all the adjusting to the long-run equilibrium.

Perturbations in our equilibrium relationship was found to be represented as $z_t =$ Illinois - .066(Barge)-1.061(Gulf), where z_t represents stationary deviations in the long-run equilibrium between the two sets of prices. The t – statistic associated with the barge market suggests that the transportation market does respond to the export marketing channel equilibrium. Put simply, if the price of the Illinois market is high relative to its long-run equilibrium, the barge market responds downwards in period $t + 1$. This is an especially intuitive result given that one would expect the demand for barges to decrease (and hence prices fall) if the price of grain in Illinois increased. The other part of the ECM framework that isolates the short run dynamics is through the Γ matrix (see equation (1)). Casual inspection of the reported t – statistics associated with this matrix suggests that the dominant market is the barge market (all parameter estimates and associated t - statistics are excluded to conserve space but are available upon request). The coefficient associate with the lagged differences from the barge market is significant on itself and the Illinois market. Interestingly changes in the Illinois and Gulf markets in period t - 1 enter no market in period t with a statistically significant coefficient.

Innovation Accounting

As previously mentioned, the patterns of response and strengths of the relationships among the prices that make up the export marketing channel are quite difficult to decipher by focusing on the individual coefficients separately, either through

the speed of adjustment parameters, a_i or through the short run dynamics matrix, Γ. A more suitable way to summarize the dynamic relationships between these markets is through well-known innovation accounting techniques, applied to the ECM outlined in equation (7). However, as previously mentioned, crucial to such analysis is the method used to treat contemporaneous innovation. In this study we adopt the factorization known as the "Bernanke ordering". Write the innovation (residual) vector (n_t) from the ECM as $An_t = e_t$, where A is a 3 x 3 matrix and e_t is a 3 x 1 vector of orthogonal shocks. As illustrated by Doan (1992, 8 – 10), a factorization is identified if there is no combination of i and j ($i \neq j$) for matrix A where both $\{a_{ij}\}$ and $\{a_{ji}\}$ are non-zero where $\{a_{ij}\}$ is an element i,j of matrix A in this instance. Here we employ the algorithm presented in Spirtes et al. (1993) in order to place zeros in the A matrix.

Innovations from our ECM give us the contemporaneous innovation correlation matrix, Σ (representing the innovations as n_i). The equation below (7) presents the lower triangular elements of the correlation matrix on innovations ($\hat{n}$) from the ECM where the entries are presented in the order, Illinois, Barge and Gulf:

$$\Sigma(\hat{n}_t) = \begin{bmatrix} 1.00 & & \\ -.084 & 1.00 & \\ .919 & .049 & 1.00 \end{bmatrix}. \tag{7}$$

DAG theory points out that the off-diagonal elements of the scaled inverse of the $\Sigma(\hat{n}_t)$ matrix are in fact the negatives of the partial correlation coefficients between the corresponding pair of variables given the remaining variable(s) in the matrix (Whittaker

1990, p.4). The off-diagonal elements of the scaled inverse of the $\Sigma(\hat{\boldsymbol{n}}_t)$ matrix, denoted

by $\Sigma^*(\hat{\boldsymbol{n}}_t)$, where the * indicates that we have scaled the inverse matrix:

$$\Sigma^*(\hat{\boldsymbol{n}}_t) = \begin{bmatrix} 1.00 & & \\ -.327 & 1.00 & \\ .927 & .321 & 1.00 \end{bmatrix}. \tag{8}$$

For example, the partial correlation between innovations in prices in the Illinois

market and the barge market, given innovations in the Gulf market is -.327. Under the

assumption of multivariate normality, Fishers z statistic can be applied to test for

significance from zero (see Equation (6)). In this case, the correlation between Illinois

and the barge market (-.327) is significantly different from zero at all conventional

significance levels (with an associated p - value = .000). Interestingly, in this case all

conditional partial correlations are significantly different from zero. That is, the partial

correlation between the Illinois market and the Gulf market given innovations in the

barge market is .927 (p - value of .000) and the partial correlation between the barge

market and the Gulf market given innovations in the Illinois market is .321 (again a p-

value of .000). Curiously, the partial correlations between the Illinois and the barge

market and the Gulf and the barge market are of the intuitively correct sign. That is, one

would expect an increase in Illinois prices to cause a decrease in barge prices (less

demand for barges given the higher price of grain for export), a result found previously

when we standardized the cointegrating vector on the Illinois price. Moreover we find

here that an increase in Gulf prices tends to cause an increase in barge prices; a result

consistent with the notion that increase in demand for barges would drive these prices

upwards given the higher export prices at the Gulf.

The theory of DAG's as given in Spirtes et. al (1993) provides an algorithm for removing edges between different markets but also directs causal flow of information between markets. The algorithm starts with a complete undirected graph (like the one shown in the top panel of Figure 2) where innovations in every market are connected with innovations in every market. The algorithm then starts to remove edges based on simple correlations. Indeed, in this analysis, it was found that the sample correlation between the Gulf market and the barge market was not significantly different from zero ($r_{B,G}$ = .0486 with a p-value of .2681), so we could remove the edge between the Gulf market and the barge market. However, the edges between the Gulf market and the Illinois market and the correlation between the barge market and the Illinois market could not be removed. As such, only the edges connecting the barge market to Illinois and the Gulf market and Illinois remain. The next step of removing edges is based on the partial correlations. Here, correlations between the Gulf market and the Illinois market conditional on the barge market and between the barge market and the Illinois market conditional on the Gulf market are found to be non-zero. Accordingly, we can not remove the edges Illinois — barge and Illinois — Gulf.

Edge removal, based on correlations and partial correlation results in the triple: Gulf ¾ Illinois — barge, using the notation from Figure 2. Since the edge between Gulf and barge was removed using the unconditional correlation test (recall $r_{B,G}$ = .0486 with a p - value of .2681), we can direct this remaining triple as: Gulf $\rightarrow$ Illinois $\leftarrow$ barge, as we show in Panel B of Figure 2. Here, Illinois is a collider – receiving information from both the Gulf market and the barge market. As such (as a collider) it opens up the

information flow between the Gulf and barge markets. Recall from Equation 8, the conditional correlation between the Gulf market and the barge market is .321 and has a p – value of .000.

As further evidence of the causal path shown in Panel B of Figure 2, we consider scoring all possible acyclic representations on information flow in contemporaneous time. As we have three variables scoring all 25 alternative graphs is feasible. These are labeled, 1, 2, …, 25 in Figure 3. We apply a modified Schwarz-loss metric to each acyclic graph.[13] The score for each model is given beneath each graph in italics.

$$SL^* = \log(|\Sigma^*|) + k\log(T)/T.$$

Here Σ^* represents a diagonal matrix consisting of the diagonals elements of the variance covariance matrix associated with a linear representation of the disturbance terms from an acyclic graph fit to innovations from the ECM.

Notice that the minimum SL* metric results from model 13: B $\rightarrow$ I $\leftarrow$ G. The SL* statistic associated with this model is -.484. This is the same model suggested by TETRAD II; although we'll see below that at least one other model is not ruled out by our SL* scoring.   The edge I,G is clearly the most important edge in the system. Notice the score associated with models 4 and 5. Here the I,G edge gives us the lowest single edge score (-.304). Compared to SL* measures for models 2, 3, 8, and 9, where the I,G edge is missing. These latter models give SL* measures that are greater than +1.87. A qualitatively similar result is found from all two-edge models. Within the set of two edge models (these are models 11,12,13,14,16,17,18,19,21,22,23 and 24), those that eliminate the I,G edge (models 18, 19, 22, and 23) have larger SL* metrics (in a

neighborhood of 1.7 to 1.8), relative to models that include the I,G edge. The latter models result in SL* metrics ranging over -.296 to -.484.

Choice between model 13 (B → I ← G) and model 21 (I → G ← B) is not clear-cut, as both have the dominate I,G edge and both result in similar SL* metrics (-.484 for model 13 and -.482 for model 21). Recall from our use of TETRAD II, that the B,G edge has an unconditional correlation of .049 and associated p-value of 0.26. While the I,B has an unconditional correlation of -.084 with a p-value of 0.05. So even the TETRAD II is choosing between these two candidate models based on relatively slight differences in p-values.

Forecast error decompositions and impulse responses (one standard deviation shocks from the ECM's) based on the DAG's are provided in Table 4 and Figure 3 respectively. The forecast error decompositions allows us to consider which prices within the export marketing channel are statistically exogenous or endogenous relative to each other at differing forecast horizons. As can be seen from Table 4 we analyze a forecast horizon up to 14 days – more than enough time for a barge to travel from this part of the Illinois River (South of Peoria) to the U.S. Gulf. The impulse responses, which allow us to evaluate the dynamic paths of adjustment of each of the prices to shocks in the data series, are illustrated in Figure 3. They too allow a 14-day window.

The first column in the output for the forecast error decompositions is the standard error of the forecast for each particular price series. The remaining columns illustrate the error decompositions. As usual, each row should add up to 100 (but may not here due to rounding). As can be seen, the Illinois market is heavily influenced by

the Gulf market. The Gulf explains 84.78% of the variation in the Illinois market after just one-day. Recall, our results from the DAG analysis suggest that the Gulf market 'causes' the Illinois market in contemporaneous time, and apparently continues to do so in the short run (1 day) out to the longer term of 14 days, where it still explains over 78%. The barge market has some influence on the Illinois market, although its effect is not as large as the Gulf's. Indeed, the barge market explains about 1.6% of the variation after 1 day and finishes at about 3.9% after 14 days. Once again, this result is consistent with the DAG analysis. There, the barge market 'caused' the Illinois market in contemporaneous time. The remaining portion of the variation is attributed to the Illinois market itself (13.6% after 1 day and about 17.7% after 14 days).

Perhaps the most interesting finding is associated with the forecast error decompositions associated with the barge market. Consistent with the DAG graphs analysis, the barge market is not influenced by either the Illinois or the Gulf markets in the very short run (1 day). Indeed, after 1 day the barge market is exogenous, as it explains 100% of its own variation. Over time, however, a different pattern emerges. While some of the variation can be explained by the Gulf market at time passes, the vast majority of the variation of the barge rate can be attributed to the Illinois market. Indeed, after the 14 days have passed about 58% of the variation can be attributed to the Illinois market. Clearly, over time, the barge market is susceptible to large volatility shocks arising from the very market that it serves.

The Gulf market is also 100% exogenous in the short run (1 day) a result consistent with the direction of causality in the DAG analysis. Indeed, as time passes,

while not being completely exogenous, very little of the variation in the market is being explained by the domestic influences of the Illinois market and the barge market that connects the two together. It seems to be a plausible hypothesis therefore that the Gulf market is being influenced by other global factors, but it in turn affects the Illinois market which then influences the barge market as time passes. Put another way, the Gulf market does not seem to influence the barge market directly, but rather its informational effect is transmitted through the Illinois market and then onto the barge market shortly thereafter.

Focusing our attention on the impulse responses in Figure 3 we see an identical pattern emerge. For instance, the left-hand panel of the chart illustrates the response of each market to a shock in the Illinois market. While the Illinois and Gulf markets are somewhat affected by a shock from the Illinois market, it is the barge market that is most heavily influenced, a finding consistent with the error decompositions. Indeed, it is only after about the 14 days that the barge market stabilizes, yet still remains affected. Clearly a shock from the Illinois market creates considerable volatility in the barge industry, which could, if unhedged, be extremely detrimental to physical traders in this industry. Interestingly, the sign of the shock is as one might expect (negative), a result consistent with the finding of a negative conditional correlation between the markets. That is, an increase in Illinois prices should correspond with a decrease in barge rates (as explained previously). Note also that while the barge rate is affected by the Illinois price over time, it starts out at zero, a finding consistent with the DAG analysis whereby the Illinois market does not affect the barge market in contemporaneous time. This can also be said about the affect of the Illinois market on the Gulf market. An innovation in the barge

20

market has almost no affect the Gulf market (bottom graph of the middle panel), just like the innovation in the Illinois market had no affect. Once again, the Gulf market can be deemed to be exogenous to the other domestic linkages. However, as shown by the top graph in the middle panel, the Illinois market is somewhat affected by the barge market, and the sign of the response (negative) is, once again, consistent with earlier intuition.

The last panel of the impulse response graph illustrates the response of the inland markets to a shock in the Gulf market. As can be seen by the top graph, the Illinois market is immediately and strongly affected by a shock originating out of the Gulf. This is a result found previously in both the DAG framework and the forecast error decompositions. A shock to the Gulf market also has an affect on the barge market that feeds it. However, consistent with the DAG analysis, it does not have an immediate affect. However, as time passes, the barge market reacts positively, an intuitively pleasing result.

V   Concluding Remarks

In recent years there has been a plethora of research looking at the level of interconnectivity between different yet related markets, but to date, no study has analyzed the degree of interconnectivity within a marketing channel in a truly dynamic manner. In this study, we apply Directed Acyclic Graphs (DAG's) to make causal statements amongst three related markets in contemporaneous time. We apply these results to the heretofore well-understand Error Correction Model to address issues surrounding dynamic patterns of price discovery using both forecast error decompositions and impulse responses.

Our results illustrate that information from the Gulf market is critical in the price discovery process over all time horizons. While the globally influenced Gulf market does not heavily influence the barge market that connects it to its inland grain source at Illinois in contemporaneous time, it is somewhat affected as time passes. However, it is the Illinois market that is immediately influenced by the Gulf. This affect seems to ripple through to the transportation market as time passes, reversing the direction of causation from the barge market influencing the Illinois market in contemporaneous time, to the Illinois market heavily influencing the barge market out into the longer term. It seems therefore, that over the longer term both the domestic and international markets heavily influence the barge market and shocks to these markets can greatly influence rates, negatively, or positively depending upon where the shock originates. These shocks, whether they originate from the Gulf or inland cause excess volatility in the barge market, which could be detrimental to unhedged physical traders in this marketing channel.

This paper has therefore, not only shed light on the degree of interconnectivity between several important markets using unique econometric methods but also sheds some light on the importance of the barge market critical in linking markets together. Our results seem to support the existence of the newly developed barge rate futures contract, but like so many other futures contracts that are designed, time can only tell whether the market will be successful.

References

Berg-Andreassen, J.A. 1997. Efficiency and Interconnectivity in International Shipping Market. *International Journal of Transport Economics* XXIV 2: 241 – 257.

Bernanke, B. 1986. Alternative Explanations of the Money-Income Correlation. *Carnegie-Rochester Conference Series on Public Policy* 25: 49 – 100.

Bessler, D.A. and D.G. Akleman. 1998. Farm Prices, Retail Prices, and Directed Graphs: Results for Pork and Beef. *American Journal of Agricultural Economics* 80 (5): 1144 – 1149.

Bessler, D.A. and S.W. Fuller. 2000. Railroad Wheat Transportation in the Central Plains: Modeling with Error Correction and Directed Graphs. *Transportation Research Part E* 36: 21 – 39.

Carlton, D.W. 1984. Futures Markets: Their Purpose, Their History, Their Growth, Their Successes and Failures. *The Journal of Futures Markets* 4 (3): 237 – 271.

Chowdhury, A.R. 1991. Futures Market Efficiency: Evidence from Cointegration Tests. *The Journal of Futures Markets* 11: 577 – 589.

Cooley, T.F. and S.F. LeRoy. 1985. Atheoretical Macroeconometrics: A Critique. *Journal of Monetary Economics* 16 (3): 283 – 308.

Dickey, D.A., and W.A. Fuller. 1981. The Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association* 74: 427 – 431.

Doan, T. 1992. RATS: User's Manual, Version 4.0 Evanston, Ill. ESTIMA.

Geiger, D., Verma, T., and Pearl, J. 1990. Identifying Independencies in Bayesian Networks, *Networks* 20: 507 – 534.

Geraci, V.J. and W.Prewo. 1977. Bilateral Trade Flows and Transport Costs. *Review of Economics and Statistics* 59: 67 –74.

Godfrey, L.G. 1988. Misspecification Tests in Econometrics. The Lagrange Multiplier Principle and Other Approaches. Cambridge University Press.

Gonzalo, J. 1994. Five Alternative Methods of Estimating Long-Run Equilibrium Relationships. *Journal of Econometrics* 60: 99 – 126.

Goodwin, B.K., T. Grennes, and M.K. Wohlgenant. 1990. Testing the Law of One Price when Trade Takes Time. *Journal of International Money and Finance* 9: 21 – 40.

Goodwin, B.K., and N.E. Piggott. 2001. Spatial Market Integration in the Presence of Threshold Effects. *American Journal of Agricultural Economics* 83(2): 302 – 317.

Glymour, C. 1997. A Review of Recent Work on the Foundation of Causal Inference. In McKim and Turner, editors, Causality in Crisis, Notre Dame: Notre Dame Press, pp. 201 – 248.

Haigh, M.S. and M.T. Holt. 2000. Hedging Multiple Uncertainty in International Grain Trade. *American Journal of Agricultural Economics* 82(4): 881 – 896.

Haigh, M.S. and H. Bryant. 2001. The Effect of Barge and Ocean Freight Price Volatility in International Grain Markets. *Agricultural Economics* 21: 41 – 58.

Hauser R.J. and J. Buck. 1989. The Feasibility of a Futures Market for Barge Grain Freight. *Review of Futures Markets* 8: 1 – 15.

Hauser R.J. and D. Neff. 1993. Export/Import Risks at Alternative Stages of U.S. Grain Export Trade. *The Journal of Futures Markets* 13 (6): 579 – 595.

Hausman, D.M. 1998. *Causal Asymmetries*, New York: Cambridge University Press.

Holland, P. 1986. Statistics and Causal Inference. *Journal of the American Statistical Association* 81: 945 – 960.

Johansen, S. 1988. Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control* 12: 231 – 254.

Johansen, S., Juselius, K. 1990. Maximum likelihood Estimation and Inference on Cointegration – with Application to the Demand for Money. *Oxford Bulletin of Economics and Statistics* 52: 169 – 210.

Johansen, S. 1991. Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models. *Econometrica* 59: 1551 – 1580.

Johansen, S. 1992. Determination of Cointegration Rank in the Presence of Linear Trend. *Oxford Bulletin of Economics and Statistics* 54: 383 – 397.

Johansen, S. 1995. Identifying Restrictions of Linear Equations with Applications to Simultaneous Equations and Cointegration. *Journal of Econometrics* 69: 111 – 132.

Juselius, K. 1995. Do Purchasing Power Parity and Uncovered Interest Rate Parity Hold in the Long Run? An Example of Likelihood Inference in Multivariate Time-Series Models. *Journal of Econometrics* 69 (1): 211 – 240.

Kavussanos, M and N. Nomikos. 1999. Futures Hedging when the Structure of the Underlying Asset Changes; the case of the BIFFEX contract. *The Journal of Futures Markets* 20(8): 775 – 801.

Kavussanos, M and N. Nomikos. 2000. Constant vs. Time-Varying Hedge Ratios and Hedging Efficiency in the BIFFEX Market. *Transportation Research: Part E: Logistics and Transportation Review* 36 (4): 229 – 248.

Lee, T and Y. Tse. 1996. Cointegration Tests with Conditional Heteroscedasticity. *Journal of Econometrics* 73: 401 – 410.

Orcutt, G. 1952. Toward a Partial Redirection of Econometrics. *Review of Economics and Statistics* 34:195 – 213.

Osterwald-Lenum, M. 1992. A Note with Fractiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank test Statistics: Four Cases. *Oxford Bulletin of Economics and Statistics* 54: 461 – 472.

Papineau, D. 1985. Causal Asymmetry. *British Journal of the Philosophy of Science* 36: 273 – 289.

Pearl, J. 1986. Fusion, Propogation, and Structuring in Belief Networks. *Artificial Intelligence* 29: 241 – 288.

Pearl, J. 1995. Causal Diagrams for Empirical Research. *Biometrica* 82: 669 – 710.

Pearl, J. 2000. *Causality.* Cambridge, Cambridge University Press.

Philips, P. 1994. Some exact Distribution Theory for Maximum Likelihood Estimators of Cointegrating Coefficients in Error Correction Models. *Econometrica* 62: 73 – 94.

Philips, P. 1998. Impulse Response and Forecast Error Variance Asymptotics in Nonstationary VARs. *Journal of Econometrics* 83: 21 – 56.

Reichenbach, H. 1956. *The Direction of Time.* Berkeley: University of California Press.

Rubin, D. 1978. Bayesian Inference for Causal Effect. *Annals of Statistics* 6: 34 – 58.

Scheines, R.P. Spirtes, C. Glymour, C. Meek. 1994. *Tetrad II: User's Manual and Software*, New Jersey: Lawrence Erlbaum Associates, Inc.

Schwarz, G. 1978. Estimating the Dimension of a Model. *Annals of Statistics* 6: 461 – 464.

Simon, H.A. 1953. Causal Ordering and Identifiability. In W.C. Hood and T.C. Koopmans (Eds) *Studies in Econometric Method*, 49 – 74. New York: Wiley.

Sims, C.A. 1986. Are Forecasting Models Usable for Policy Analysis? *Federal Reserve Bank of Minneapolis Quarterly Review*, Winter.

Spirtes, P., C. Glymour, and R. Scheines. 1993. Causation, Prediction and Search, New York: Springer-Verlag.

Spirtes, P., C. Glymour, R. Scheines, C. Meek, S. Fienberg, and E. Slate. 1999. "Prediction and Experimental Design with Graphical Model," in Clark Glymour and Gregory F. Cooper editors: *Computation, Causation and Discovery*, Cambridge, Massachusetts: MIT Press, pp. 65-93.

Suppes, P. 1970. "*A Probabilistic Theory of Causality.*" Amsterdam: North Holland.

Swanson, N.R., Granger, C.W.J. 1997. Impulse Response Functions based on a Causal Approach to Residual Orthogonalization in VAR. *Journal of the American Statistical Association* 92: 357 – 367.

Thoung, L.T., and S.V. Visscher. 1990. The Hedging Effectiveness of Dry-Bulk Freight Rate Futures. *Transportation Journal* 29: 58 – 65.

Toda, H.Y and P. Phillips. 1994. Vector Autoregression and Causality: A Theoretical Overview and Simulation Study. *Econometric Reviews* 13: 259 – 285.

Verma, T. and J. Pearl. 1988. Causal Networks: Semantics and Expressiveness. In Proceedings of the 4th Workshop on Uncertainty in Artificial Intelligence (Mountain View California) pp. 352 – 359. Reprinted in R. Shachter, T.S. Levitt and L.N. Kanal (Editors), Uncertainty in Artificial Intelligence, (4) pp. 69 – 76, Amsterdam, Elsevier.

Verma, T. and J. Pearl. 1990. Equivalence and Synthesis of Causal Models. In *Proceedings of the 6th conference on Uncertainty in Artificial Intelligence* (Cambridge, Massachusetts), pp. 220 – 27. Reprinted in P. Bonissone, M.Henrion, L.Kanal, and J. Lemmer (Editors), *Uncertainty in Artificial Intelligence*, Vol. 6, pp. 255-68. Amsterdam:Elsevier.

Whittaker, J. 1990. Graphical Models in Applied Multivariate Statistics. Wiley, Chichester, UK.

Table 1. Descriptive statistics and correlation analysis on daily prices

	Illinois	Barge	Gulf
Mean	172.98	8.915	186.34
Median	170.15	8.584	184.28
Standard deviation	10.806	2.228	10.369
CV	0.056	0.250	0.056
m_3	-0.181	-0.531	-0.503
m_4	0.377	0.505	0.186
Min	147.17	5.336	161.11
Max	202.95	16.008	213.41

Unconditional Correlations

	Illinois	Barge	Gulf
Illinois	1		
Barge	-0.3268	1	
Gulf	0.9700	-0.1003	1

Summary statistics are presented for daily grain and barge prices for the period 6th May 1999 – 3rd May 2001. CV represents the Coefficient of Variation and m_3 and m_4 represent sample skewness and kurtosis respectively.

Table 2. Augmented Dickey-Fuller (ADF) tests for order of integration on prices

Test is on the estimated coefficient θ_1 from the following prototype model:

$$\Delta X_t = q_0 + q_1 X_{t\text{-}1} + \sum_{k=1}^{K} b_k \Delta X_{t\text{-}k}$$

Price	K	HO: $I(1)$ vs. HA: $I(0)$ ADF	HO: $I(2)$ vs. HA: $I(1)$ ADF
Illinois	0	-2.354	-24.190
Barge	1	-3.158	-18.960
Gulf	0	-2.341	-23.728

Critical values are taken from Fuller (1976). They are –2.57 (10%), -2.88* (5%) and –3.46 (1%). Therefore, based on these results are series are I(1). The optimal lag length (K) was based on the Schwarz Bayesian Criterion (1978).

Table 3. Cointegration analysis of prices

	Trace tests on order of cointegration[a]	
λ_{trace} *test statistic*	*H_O:*	*critical value*
103.96	r = 0	29.68
14.21	r ≤ 1	15.41
4.93	r ≤ 2	3.76

	Tests for exclusion from the cointegrating vector[b]	
	H_O:	*$c^2_{(1)}$ value*
Illinois	$\boldsymbol{b}_I = 0$	79.25
Barge	$\boldsymbol{b}_B = 0$	79.17
Gulf	$\boldsymbol{b}_G = 0$	78.85
	Tests for weak exogeneity[b]	
	H_O:	*$c^2_{(1)}$ value*
Illinois	$\boldsymbol{a}_I = 0$	1.11
Barge	$\boldsymbol{a}_B = 0$	7.57
Gulf	$\boldsymbol{a}_G = 0$	1.85

[a]Tests are on eigenvalues with the Π matrix. The λ_{trace} statistic is - $N(\sum_{i=r+1}^{2} \ln(1-\boldsymbol{I}_i))$, where λ_i are ordered (largest to smallest) eigenvalues on Π. Critical values for the λ_{trace} statistics (at the 10% level) are from Osterwald-Lenum (1992). The optimal lag length (k) is based on the Schwarz Bayesian Criterion (1978). The sample size (N) is equal to 520.

[b]Tests are based on the following: T = $N(\ln(1-\lambda_R) - \ln(1-\lambda_U)$, where λ_R is the eigenvalue calculated with the restriction and λ_U the eigenvalue calculated without the restriction. With one cointegrating vector the critical $c^2_{(1)}$ value is 3.84. Based upon these results all prices in the marketing channel appear to be a part of the cointegrating relationship, and both the Gulf and the Illinois prices are weakly exogenous.

Table 4. Forecast error decompositions

Steps ahead (days)	Std. Error	Illinois	Barge	Gulf
(Illinois)				
1	0.013	13.622	1.598	84.781
2	0.018	11.586	3.689	84.725
3	0.022	12.101	4.335	83.564
7	0.033	14.465	4.545	82.138
14	0.047	17.655	3.933	78.412
(Barge)				
1	0.056	0.000	100.00	0.000
2	0.086	1.990	97.786	0.224
3	0.109	6.195	93.263	0.542
7	0.171	31.405	66.670	1.924
14	0.245	58.334	38.510	3.156
(Gulf)				
1	0.012	0.000	0.000	100.00
2	0.017	0.432	0.004	99.563
3	0.020	0.356	0.032	99.611
7	0.031	0.386	0.051	99.560
14	0.043	0.415	0.057	99.527

The decompositions for each step ahead are given for a Bernanke factorization of contemporaneous covariances, which treats each price series as exogenous in contemporaneous time. The justification for this is based on the DAG on observed innovations from the error correction model shown in equation (7). The decompositions may not sum to one hundred in each row due to rounding.

Figure 1. Daily price data. The sample period is 6th May 1999 – 3rd May 2001 (520 observations):

Panel A. U.S. Gulf and Illinois Soybean Prices.
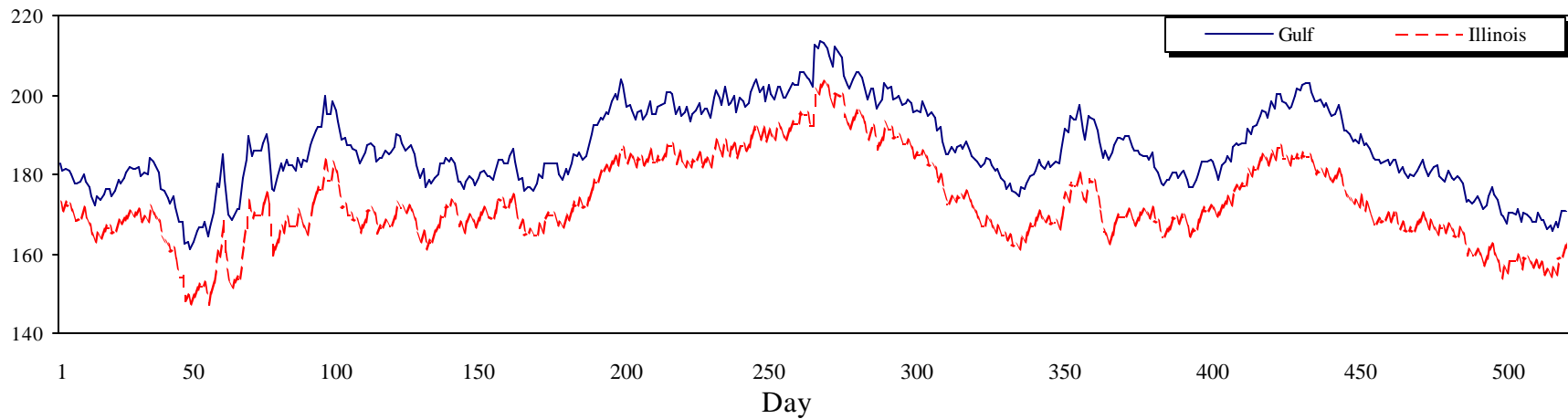
Panel B. Gulf-Illinois Soybean Price Spread and Mississippi Barge Rate.
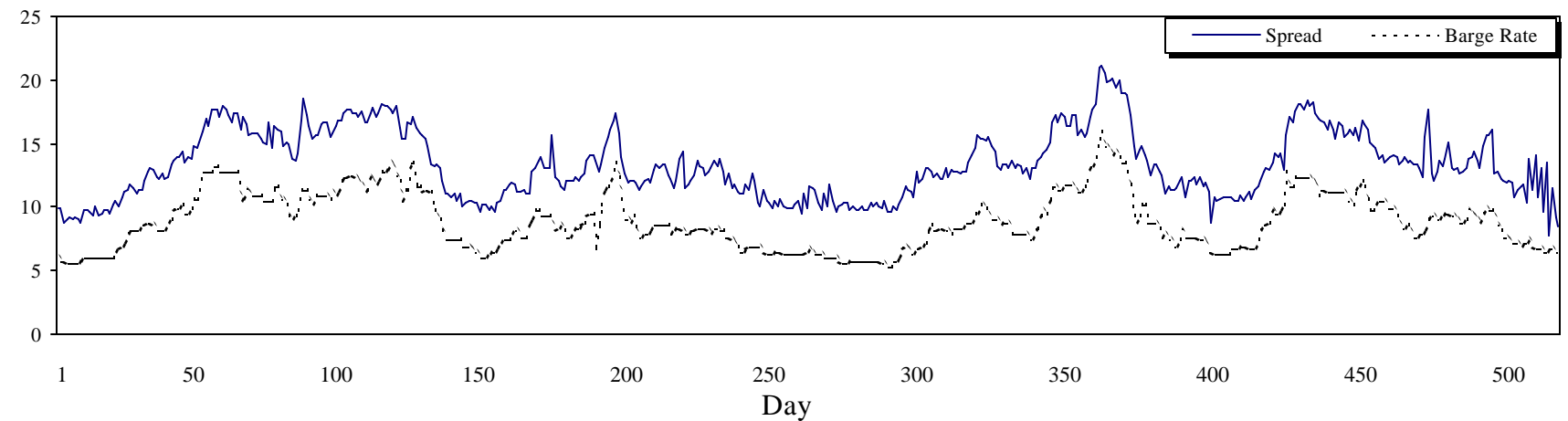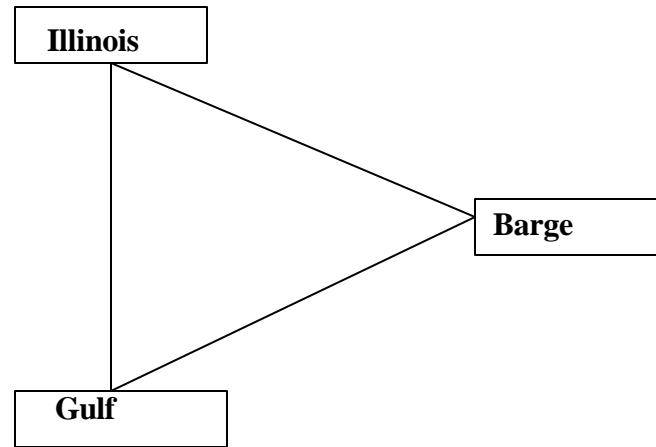
Figure 2. Undirected and Directed Acyclic Graphs

Panel A. Complete undirected graph.

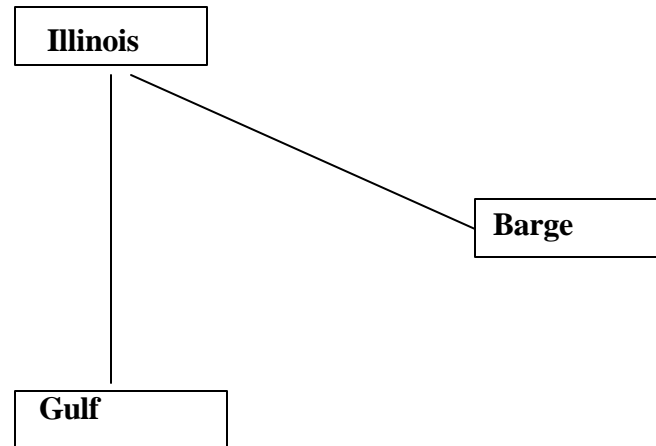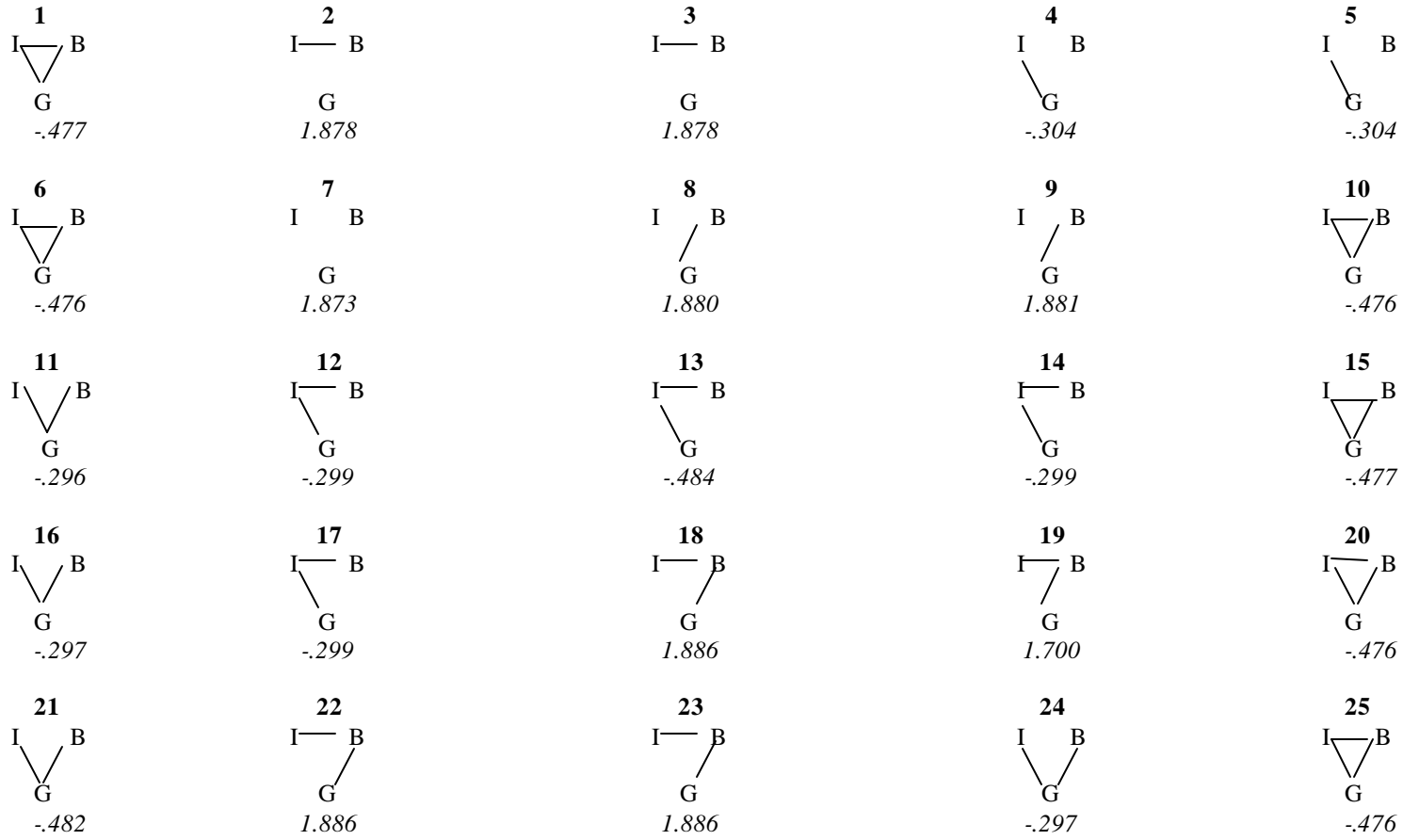Panel B. Directed graph (lines are significant at the 10% level).

Figure 3. Alternative Directed Acyclic Graphs on Innovations on Illinois Grain Prices (I), Gulf Grain Prices (G) and Barge Rates (B) with Schwarz-Loss Statistics (in italics) on Error Correction Residual Covariance Matrix.

1	2	3	4	5
I—B G *-.477*	I—B G *1.878*	I—B G *1.878*	I B G *-.304*	I B G *-.304*

6	7	8	9	10
I B G *-.476*	I B G *1.873*	I B G *1.880*	I B G *1.881*	I—B G *-.476*

11	12	13	14	15
I B G *-.296*	I—B G *-.299*	I—B G *-.484*	I—B G *-.299*	I—B G *-.477*

16	17	18	19	20
I B G *-.297*	I—B G *-.299*	I—B G *1.886*	I—B G *1.700*	I—B G *-.476*

21	22	23	24	25
I B G *-.482*	I—B G *1.886*	I—B G *1.886*	I B G *-.297*	I—B G *-.476*

Each graph is scored (in italics) using a modified Schwarz loss metric: $SL^* = \log(|\Sigma|) + k\log(T)/T$; where Σ is the matrix of diagonal elements from the residual error covariance matrix, T is the number of observations (518) and k is the number of regressors in the three variable system.

Figure 4. Impulse responses over 14 days from one standard deviation shocks.

Innovation to:
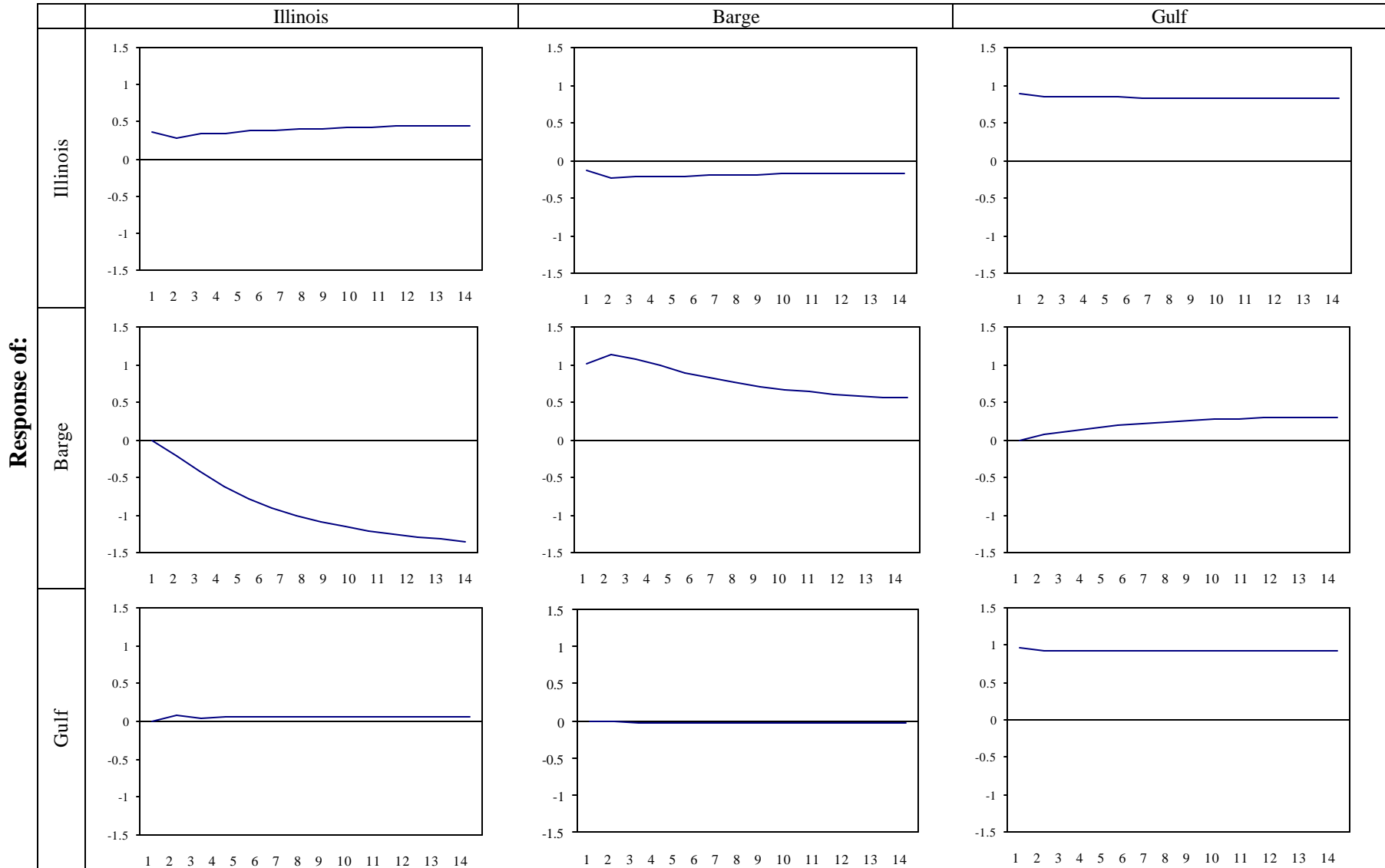
Endnotes

[1] Source: Baltic Exchange, London, UK.

[2] Over its life, the BIFFEX contract has generated a varying degree of trading interest. For instance, at is peak in 1988 the volume reached 97335 contracts (or about 383 contracts per day). However, since November 1999, when the underlying index was changed for the last time (it has been changed a total of 13 times to try and generate trading interest) trading volume only reached an average of 17 contracts a day. Indeed, according to Carlton (1984), 31% of futures contracts introduced in the United States between 1921 and 1983 died within their first two years of trading.

[3] See Carlton (1984) for a description of the important features that a commodity traded on a futures exchange should possess in order to be successful.

[4] Only a handful of papers have employed DAG analysis in economics. Examples include: Bessler and Akleman (1998) and Bessler and Fuller (2000).

[5] For soybeans in particular, which are analyzed in this study, the total tonnage exported out of the U.S. from all ports in this time period was 57.93 million tonnes on a total of 3864 vessels to 44 different countries. However, 40.07 million of those tonnes (or 69.2% of the total) left the U.S. Gulf at New Orleans from the Mississippi River on a total of 1686 vessels to 39 different countries.

[6] Unlike the BIFFEX futures contract, the barge futures contract is not based on an index of prices. Full details on the barge rate futures contract can be found at the Exchange web site: www.merchants-exchange.com. To date, only one paper has attempted to analyze the feasibility of a futures market for barge freight (Hauser and Buck, 1989). That research, except for some static regression techniques, analyzed the potential role of

the market in a largely qualitative manner. The research did recommend that a barge futures contract be developed.

[7] Papers by Phillips (1994) and Toda and Philips (1994) illustrate that simply observing the loading vectors associated with the ECM may not be sufficient to properly define/distinguish between short and long run causation. We recognize this limitation but point out that it is common for researchers to use the terms 'short' and 'long' run in applied cointegration research. We use these terms here but are aware of the inability to properly differentiate between these effects.

[8] Spirtes, Glymour and Scheines (1993) give the conditions, faithfulness and a causal Markov assumption, which allow us to use patterns in correlation structure to identify a graphical structure from data.

[9] It is worth noting that the PC algorithm assumes that there are no latent common causes of the three measured variables (Illinois, barge and Gulf prices). In contrast, the related FCI algorithm does not make such an assumption and while the conclusions that it draws from this data are weaker than the conclusions that follow the PC algorithm they do not rely on the somewhat dubious assumption of no latent common causes. The more complex FCI algorithm is documented in Spirtes, Glymour and Scheines (1993).

[10] These data (like all data used by the authors) are available upon request. A small number of price quotes were missing in each of these markets. On these days, the missing observations were replaced with the most recent price, thus construc ting a price series consistent with a random walk.

[11] We excluded the constant from inside the Π matrix due to its statistical insignificance.

[12] Scheines et al. (1994) point out that statistical tests that are based on the assumption of normality should also be useful for distributions that depart from normality, although the authors provide no formal evidence on how sensitive Tetrad II is to deviations from normality. The authors did apply Tetrad II to a set of variables (that were provided from a study on causes of publishing productivity). These variables were shown to violate the joint normality condition, but the authors discovered that Tetrad II provided reasonable suggestions (similar to the models that were provided by the original authors). The authors did not discuss the presence/affect of (G)ARCH effects on the underlying data set.

[13] We apply a modification of Schwarz loss because scoring alternative models results in the same log determinant for every alternative acyclic model. Alternative models merely move correlation from the diagonal elements to the off-diagonal elements in the residual covariance matrix from alternative models. Accordingly we score alternatives using the determinant of the main diagonal of the residual covariance matrix. See the note associated with Figure 3 for more details.