FED CATTLE PRICE DISCOVERY: DIRECTED ACYCLIC GRAPH AND TIME SERIES MODELING

Andrew C. Lee
Economist
Environmental Protection Agency
Washington, DC 20460
Email: lee.andrew@epa.gov
Phone: 703-308-7226

Man-Keun Kim
Postdoctoral Fellow
University of Maryland
College Park, MD 20740
Email: man-keun.kim@pnl.gov
Phone: 301-314-6743

Corresponding author: Andrew C. Lee
Environmental Protection Agency
Office of Pesticide Program
1200 Pennsylvania Ave., NW
Washington, DC 20460
Mail Code: 7503P
Email: lee.andrew@epa.gov
Phone: 703-308-7226

Selected Paper prepared for presentation at the 2007 AAEA Annual Meeting, Portland, Oregon, July 29-August 1, 2007

This paper reflects work done by the authors at their prior jobs and does not represent the views of the U.S. Environmental Protection Agency or the Joint Global Change Research Institute of University of Maryland.

FED CATTLE PRICE DISCOVERY: DIRECTED ACYCLIC GRAPH AND TIME SERIES MODELING

Abstract

Locating fed cattle price discovery center is revisited using the framework that combines recent progress in causal research with time series analysis. The Bernanke factorization in innovation accounting is obtained by a graphical method called directed acyclic graph which uses data to identify the contemporaneous causal structure among the innovations. This may represent an improvement over the conventional methods which use human judgment and/or theory to supply such information. Results indicate that Kansas market is a dominant price leader where new information is discovered. Contrary to the previous studies, Nebraska market does not appear to be a price discovery location.

Key words: fed cattle, price discovery, directed acyclic graph, causality

Introduction

Price dynamics among regional fed cattle markets can provide useful information for delineating a relevant fed cattle procurement market. The dominant markets where information is discovered first may play the role of price leader providing significant market information to other markets (price followers) which may have insufficient activity to generate much new information (Schroeder 1996). Locating price discovery centers and estimating price interactions among the regional fed cattle markets can help define a relevant fed cattle procurement market.

Several studies have investigated price leadership and delineation of the relevant geographic market for fed cattle. Koontz, Garcia, and Hudson (1990), using weekly prices from 1973 through 1984 and the Granger causality analysis, found that the Nebraska direct market responded fastest to new information. Through the multivariate vector autoregression analysis (VAR) of weekly regional fed cattle prices from 1976 through 1987, Schroeder and Goodwin (1990) found that Iowa/Southern Minnesota and Eastern Nebraska tended to be the leading price discovery regions, with the western Kansas becoming more dominant over the time period. Schroeder (1996) used plant-level transaction prices (from March 23, 1992 to April 3, 1993) and VAR models and Granger causality analysis and found that slaughter plants in Kansas and Nebraska tended to be price discovery leaders. Plants in other states reacted most quickly to price changes at the Nebraska plants.

The objective of this study is two-fold. First, the literature on fed cattle price leadership is rather old. The existing research covers the various parts of the period from the early 1970's to early 1990's. However, in recent times, dramatic changes in the beef value chain have occurred. These include an increasing role of contract market in fed cattle procurement; implementation of mandatory price reporting; mad cow disease scare with its adverse impact on beef export; and changes in consumer tastes/preferences that may impact beef demand (e.g., due to potential health implications of red meat consumption). This study uses more current data in an attempt to reflect some of these changes. Secondly, a new method for identifying causal relationship among variables is employed in the investigation of the dynamics of the regional fed cattle prices. The framework that combines recent progress in causal research with time series analysis is used. The Bernanke factorization in the innovation accounting is obtained by a graphical method called directed acyclic graph which uses data to identify the contemporaneous causal structure among the innovations. This may represent an improvement over the conventional methods which use human judgment and/or economic theory to supply such information.

Method

In this study, the framework used by Bessler (e.g., Bessler, Yang, and Wongcharupan 2002; Haigh and Bessler 2004) that combines the directed acyclic graph (DAG) and the multivariate time series modeling is employed. A vector autoregression (VAR) model of the five regional fed cattle prices is estimated and the innovations are obtained. The DAG analysis then identifies the contemporaneous causal relationships among the innovations to supply the Bernanke factorization in the innovation accounting.

Weekly slaughter steer prices from February 25, 2002 to July 3, 2006 (United States Department of Agriculture, Livestock Market News) were used as the cash or spot market fed cattle prices. Descriptive statistics are shown in table 1. First, the data series were tested for non-stationarity using the Dickey Fuller (DF) and Augmented Dickey Fuller (ADF) tests (table 2). Results from both tests reveal that all five price series are stationary and thus VAR modeling is appropriate. The optimal lag of VAR can be determined based on the Schwarz Loss metric and Hanna and Quinn's Φ (table 3). The minimum values of these metrics, indicated by asterisks in table 3, determine the optimal lag length of VAR to be one. Let $\mathbf{X}_t$ denote the vector of variables under consideration, $\mathbf{X}'_t = [x_{1t}, x_{2t}, x_{3t}, x_{4t}, x_{5t}]$, where the subscript 1 represents Texas/Oklahoma fed cattle price, 2 represents Colorado fed cattle price, 3 represents Kansas fed cattle price, 4 represents Nebraska fed cattle price, and 5 represents Iowa/Southern Minnesota fed cattle price. The process can be modeled in the VAR of 1 as follows,

(1) $$\mathbf{X}_t = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1\mathbf{X}_{t-1} + \mathbf{e}_t \qquad (t = 2,...,T)$$

where, $\boldsymbol{\alpha}_0$ is 5 x 1 intercept vector and $\boldsymbol{\alpha}_1$ is the 5 x 5 coefficients matrix and $\mathbf{e}_t$ is the 5 x 1 innovation vector.

The impulse response functions and the forecast error variance decompositions generated from the estimated coefficients are used to describe the dynamics of the variables. In innovation accounting, when the contemporaneous innovations are correlated, one has to assign causal directions among them. Conventional approaches to this problem are the Choleski, Bernanke, or Sim's decomposition which uses human judgment and/or theory in assigning causal flows among the innovations. In this study, the DAG identifies the contemporaneous causal relationship among the innovations based on the causal information contained in the data. This causal structure is then used to obtain the Bernanke factorization for innovation accounting (Spirtes, Glymour, and

Scheines 2000; Pearl 1995, 2000; Swanson and Granger 1997; Bessler and Yang 2003; Bessler, Yang, and Wongcharupan 2002). Clearly, the proper determination of the contemporaneous causal structure is important because it governs all the subsequent dynamics of the variables in the system. The Bernanke factorization supplied by the DAG causal structure may represent an improvement over the conventional methods because the DAG approach is based on data whereas conventional methods must resort to human judgment, especially when no economic theory is available to guide the causal determination among the innovations. A brief description of the directed acyclic graph approach follows. Readers are referred to Spirtes, Glymour, and Scheines (2000) and Pearl (2000) for more detailed treatment of the subject.

Directed Acyclic Graph

A directed acyclic graph is an illustration using arrows and vertices (variables) to represent the causal flow among a set of variables (Pearl, 2000). An example of a DAG is shown in figure 1 where the causal flow among the regional fed cattle prices is indicated by the arrows.

Directed acyclic graphs represent a conditional independence relationship as given by the recursive decomposition:

(2) $$\Pr(x_1, x_2, x_3, \cdots, x_n) = \prod_{i=1}^{n} \Pr(x_i \mid pa_i)$$

where Pr (.) is the joint probability of variables $x_1, x_2, x_3, \cdots, x_n$ and pa_i represents parents of x_i which is the minimal set of x_i's predecessors that makes x_i independent of all its other predecessors (Pearl, 2000, p.14-15). There is a one-to-one correspondence between the graphical expression of variables in a directed acyclic graph and the set of conditional independencies among variables implied by (2) (Geiger, Verma, and Pearl, 1990).

The PC algorithm (Scheines et al. 1994) available as a free software TETRAD IV is used for identifying the causal flow among the variables. The causal determination begins with a completely undirected graph which shows an undirected edge (correlation) between every pair of variables in the model. Then, the undirected edges between variables are removed sequentially based on zero-order correlation, first-order correlation, and higher-order correlations. Finally, the PC algorithm determines the causal flow on the remaining edges using conditional independence relationships (Bessler, Yang, and Wongcharupan 2002).

Results and Discussion

The directed acyclic graph in figure 1 shows the contemporaneous causal relationships among the regional fed cattle prices. It is noted that the PC algorithm could not determine the causal direction between Kansas and Texas prices, leaving an undirected edge between them. The causal flow was determined by scoring based on the Schwarz loss metric and Hanna and Quinn's Φ. Two alternative systems (corresponding to the causal flows KS $\rightarrow$ TX and KS $\leftarrow$ TX) of the VAR innovations were estimated and the Schwarz loss metric and Hanna and Quinn's Φ were obtained. The system resulting from the causal flow KS $\rightarrow$ TX had a lower Schwarz loss metric and Hanna and Quinn's Φ, and thus KS $\rightarrow$ TX was selected as the causal flow between the two prices.

Results from the contemporaneous causal structure (figure 1), impulse response functions (figure 2), and forecast error variance decompositions (table 4) all point to the

Kansas regional market as the dominant price discovery center or a source of information. As shown in figure 2, in the contemporaneous and subsequent periods, Kansas prices have a major influence on Texas, Colorado, and Nebraska prices. This is also shown in table 4 where the variations in Texas, Colorado, and Nebraska fed cattle prices are mostly explained by Kansas fed cattle prices. Although Kansas prices do not influence Iowa prices in the contemporaneous period, they have large effects on Iowa prices in the subsequent weeks as shown in figure 2 and table 4.

Iowa/Southern Minnesota market appears to be a price discovery center for the area consisting of Iowa/Southern Minnesota, Nebraska, and Colorado as shown in figure 1, but the strength and the extent of the influence of this market is much smaller than the Kansas market. Table 4 shows Nebraska prices being significantly influenced by Iowa/Southern Minnesota prices although the effect fades over time as the influence of Kansas prices becomes dominant. It is interesting to note that Schroeder and Goodwin (1990) also found Iowa/Southern Minnesota (along with Eastern Nebraska) to be a leading price discovery region, with the western Kansas becoming more dominant over the time period (They used weekly data from 1976 through 1987). Thus, clearly, Kansas fed cattle market is a dominant price leader and a source of information. Contrary to the previous studies (including Koontz, Garcia, and Hudson 1990; Schroeder and Goodwin 1990; and Schroeder 1996), Nebraska prices do not influence other regional prices as seen in figures 1 and 2 and table 4. Rather, the Nebraska market appears to be a "sink" of information flow.

Conclusion

Using the framework that combines recent progress in causal research and time series modeling this research investigated the price dynamics among the regional fed cattle markets. This study also used more current data to reflect some of the structural changes in the cattle industry that have occurred in recent times. These may include changes in consumer tastes/preferences, industry technology, international trade, and government regulations such as the mandatory price reporting. Results indicate that the Kansas regional market is a dominant price leader where new information is discovered. Contrary to the previous studies, the Nebraska regional market does not appear to be a price discovery location. Although the application of a DAG to causal studies in economics is on the rise, more research is needed to establish its robustness. Authors hope this work contributes to that effort.

References

Bessler, D.A., and J. Yang. "The Structure of Interdependence of International Stock Markets." *Journal of International Money and Finance* 22(2003):261-287.

Bessler, D.A., J. Yang, and M. Wongcharupan. "Price Dynamics in the International Wheat Market: Modeling with Error Correction and Directed Acyclic Graphs." *Journal of Regional Science* 42(2002):793-825.

Geiger, D., T.S. Verma, and J. Pearl. "Identifying Independencies in Bayesian Networks." *Networks* 20(1990):507-34.

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

Koontz, S.R. "Arbitrage Costs between Regional Fed Cattle Markets: Estimates Using Public Price Data." Hayenga, M.L., S.R. Koontz, and T.C. Schroeder, ed. *Definition of Regional Cattle Procurement Markets*, GIPSA-RR 96-1(1996):1-37.

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

Pearl, J. "Causal Diagrams for Empirical Research." *Biometrika* 82(1995):669-710.

Pearl, J. *Causality: Models, Reasoning, and Inference.* Cambridge: Cambridge University Press, 2000.

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

Schroeder, T.C. and B.K. Goodwin. “Regional Fed Cattle Price Dynamics.” *Western Journal of Agricultural Economics* 15(1990):111-122.

Schroeder, T.C. “Spatial Fed Cattle Transaction Price Relationships.” Hayenga, M.L., S.R. Koontz, and T.C. Schroeder, ed. *Definition of Regional Cattle Procurement Markets*, GIPSA-RR 96-1(1996):69-105.

Spirtes, P., C. Glymour, and R. Scheines. *Causation,Prediction, and Search.* Cambridge, MA: MIT press, 2000.

Swanson, N.R., and C.W.J. Granger. “Impulse Response Functions Based on a Causal Approach to Residual Orthogonalization in Vector Autoregressions.” *Journal of the American Statistical Association* 92(1997):357-367

United States Department of Agriculture, Livestock Market News. St. Joseph, MO.

Table 1. Descriptive Statistics of Data, February 25, 2002 – July 3, 2006 (Weekly)

Variables	Unit	Mean	Std Dev	Coef. of Variation	Min	Max
TX	$/cwt	81.70	9.27	0.11	61.96	107.21
CO	$/cwt	81.70	9.54	0.12	61.65	111.78
KS	$/cwt	81.67	9.36	0.11	61.80	106.68
NE	$/cwt	81.65	9.65	0.12	61.87	113.51
IA	$/cwt	81.37	9.56	0.12	61.74	112.09

Table 2. Non-Stationarity Tests[a]

	Dickey-Fuller			Augmented Dickey-Fuller				Final
Variables	t-value	SL	Decision	t-value	Lags	SL	Decision	Decision[b]
TX	-2.4242	1.927	S	-2.1094	2	1.939	S	S
CO	-2.4088	2.027	S	-2.1258	2	2.038	S	S
KS	-2.3571	1.889	S	-2.1737	2	1.898	S	S
NE	-2.3802	2.040	S	-2.0485	2	2.030	S	S
IA	-2.2509	1.922	S	-2.0011	2	1.941	S	S

[a] For the both tests the null hypothesis is that the variables are non-stationary and the 5% critical value is approximately –2.89. The null hypothesis is rejected if the observed t-value is less than this critical value. For Augmented Dickey-Fuller test, the optimal lag length for the augmented terms is determined by minimizing Schwarz-loss metric (SL in the table). NS indicates non-stationarity and S indicates stationarity.

[b] Where the decisions of Dickey-Fuller and Augmented Dickey-Fuller tests conflict the decision with the smaller SL is selected as the final decision and reported in the last column.

Table 3. Loss Metrics on the Order of Lags (k) in a Level Vector Autoregression[a]

Lag = k	Schwarz Loss	Φ
0	4.947	4.903
1	1.449*	1.180*
2	1.692	1.199
3	2.042	1.324
4	2.502	1.560
5	2.916	1.749

[a] The minimum values of Schwarz Loss and Hanna and Quinn's Φ, indicated by * determine the optimal lag length of VAR.

Table 4. Forecast Error Variance Decompositions

	Weeks	TX	CO	KS	NE	IA
TX	0	11.18	0.00	88.83	0.00	0.00
	1	5.85	0.10	93.79	0.27	0.00
	2	4.38	0.16	95.27	0.17	0.02
	3	3.74	0.17	95.91	0.14	0.04
	6	3.01	0.17	96.64	0.12	0.06
	12	2.64	0.16	97.02	0.12	0.06
CO	0	2.77	9.95	77.18	0.00	10.09
	1	2.65	4.99	85.27	0.29	6.80
	2	2.84	3.24	88.24	0.54	5.15
	3	2.91	2.42	89.87	0.61	4.19
	6	2.86	1.47	92.37	0.55	2.75
	12	2.65	0.97	94.19	0.43	1.76
KS	0	0.00	0.00	100.00	0.00	0.00
	1	0.85	0.04	99.10	0.00	0.00
	2	1.19	0.07	98.70	0.03	0.01
	3	1.36	0.09	98.49	0.05	0.02
	6	1.58	0.10	98.21	0.08	0.02
	12	1.71	0.11	98.06	0.10	0.03
NE	0	0.98	3.51	37.60	4.92	53.00
	1	3.51	1.85	50.73	3.12	40.78
	2	4.27	1.21	59.53	2.66	32.32
	3	4.45	0.91	65.75	2.33	26.56
	6	4.19	0.56	76.62	1.66	16.97
	12	3.56	0.38	84.63	1.11	10.32
IA	0	0.00	0.00	0.00	0.00	100.00
	1	2.95	0.11	6.78	0.11	90.05
	2	4.25	0.17	16.38	0.40	78.81
	3	4.78	0.18	25.96	0.57	68.50
	6	4.83	0.18	47.71	0.67	46.62
	12	4.08	0.15	66.49	0.55	28.73

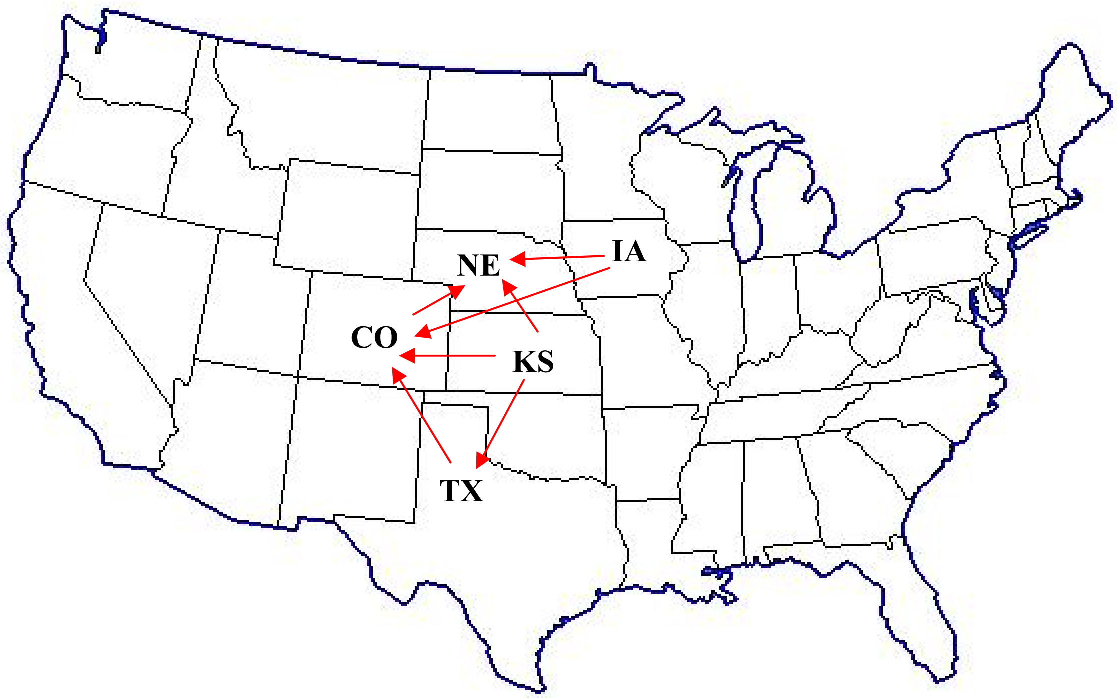

Figure 1. Directed acyclic graph -- Contemporaneous causal structure among regional fed cattle prices

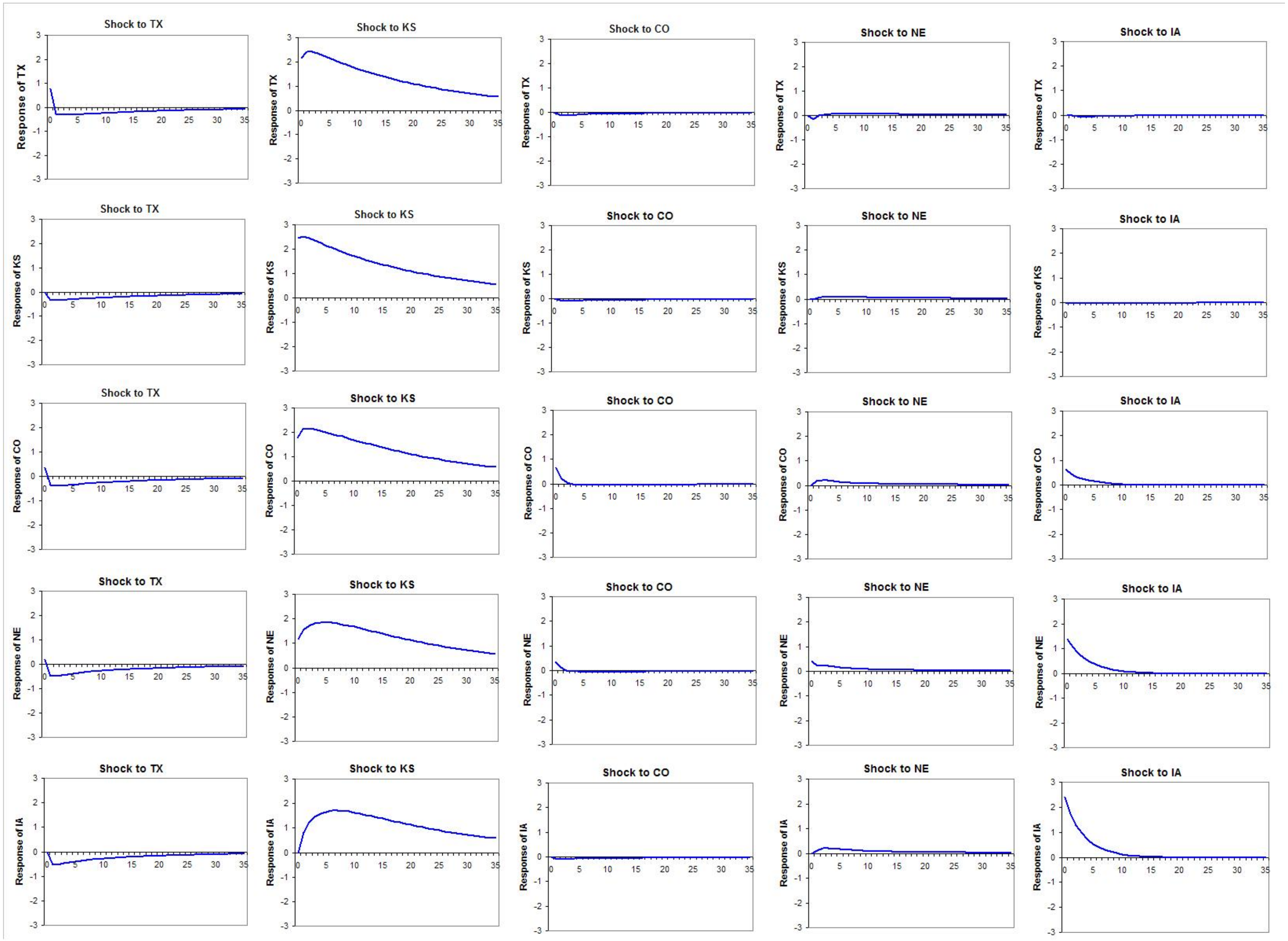

Figure 2. Impulse response functions