Potential Impacts of Food Borne Ill Incidence on Market Movements and Prices of Fresh Produce in the US

Marco Palma[1], Luis Ribera[1], David Bessler[2], Mechel Paggi[3], and Ronald Knutson[2]

[1]Assistant Professor and Extension Economist, [2]Professor, Department of Agricultural Economics, Texas A&M University, 2124 TAMU, College Station, Texas, 77845; [3]Professor, California State University, 2910 E. Barstow Ave, Fresno, California, 93740.

Selected Paper prepared for presentation at the Southern Agricultural Economics Association Annual Meeting, Atlanta, Georgia, January 31-February 3, 2009

Introduction

Historically the United States was perceived to have the safest food supply in the world. While, in fact, this may still be true, a number of incidents have led to questions regarding the safety of the U.S. food supply. Aside from the safety of the products they produce, fresh fruit and vegetable growers face many challenges. These include water availability for irrigation, increased energy and chemical costs, pest control, increased competition from globally sourced products, and the availability and cost of labor. With these many challenges, questions arise as to how much producers can afford to spend to assure the safety of their product? Put differently, what is the cost of not effectively controlling product safety? The following three case examples provide insight into the answers to these questions.

Consumers react to the news of a food safety alert by immediately changing their buying patterns and reducing consumption of the affected product. Since the initial reports of an outbreak may be indecisive as to the scope and origin of the problem, consumption/product demand may be affected nationally and even internationally. This shorter–term impact may actually shut down market movement until the source of the outbreak becomes clear by product, by the specific pathogen, by the source of the pathogen, and even by the handler and farm on which the product was produced. This may take several days or weeks. The reduction in sales depends on the severity of the outbreak, in terms of the number of people affected, number of deaths, regional scope, and the type of products and its origin.

Even after the source is identified there are potential longer-term impacts on consumption and the entire supply chain including issues such as legal liability from the incident, which may occur over a period of several months or years after the outbreak. This paper will study both, the contemporaneous and lagged effects of food borne illness in the fresh produce industry, and the length of time required to return to normal levels and the associated producer costs of the outbreaks.

Three case studies will be used to assess the potential impacts of outbreaks on product shipments and prices. Specifically, we analyzed the spinach outbreak of September, 2006; the cantaloupe outbreaks of the period 2000-2002 (April-June 2000, April-May 2001, and March-May 2002); and the tomato outbreak of July-September 2006. The data used in this study are monthly shipments, and average prices for domestic production and imports of spinach, cantaloupes, and tomatoes from the Agricultural Marketing Service (AMS), the U.S. Department of Agriculture. The prices are average monthly prices for all shipments, including national production and imports. Prices are expressed in dollars per hundredweight for tomatoes and cantaloupes, and in dollars per carton of bunched spinach.

Some of the most recent outbreaks occurred on spinach, cantaloupe and tomatoes. To illustrate, on September 13, 2006 the Food and Drug Administration (FDA) issued a warning of a multi-state *Escherichia coli* (E. Coli) O157:H7 outbreak associated with the consumption of bagged spinach. The first reports were confirmed by several states on bagged spinach having a “best if used by” date of August 30, 2006. By the time the outbreak was contained 227 people had become ill across the United States, 104 had been hospitalized, 31 had developed serious complications from hemolytic-uremic syndrome and 3 had died. An all-clear lifting of the warning alert was issued by FDA, although by about November 1, 2006, the sources of the contamination had been clearly identified and measures were being taken to assure that the incident was under control.

According to the CDC, more than 76 million people are affected and 5,000 die as a result of food poisoning every year. The most common food-borne illnesses are campylobacter, salmonella and E. Coli. Over the past 12 years, 22 leafy green E. Coli O157:H7 outbreaks have been identified. All 22 indicated a California source of the leafy greens. Since the mid-1990's foodborne illness outbreaks have occurred that were linked to raspberries, green onions, peppers, sprouts, and strawberries. In part as a reaction to these events increased efforts to enhance food safety have been undertaken by the government and associated industries groups. Efforts have focused on increased scrutiny of imported products and the improvement in domestic standards.

The main objective of this paper is to study the contemporaneous and lagged effects of food borne ill incidence on market movements and prices of fresh produce in the US. Due to data restrictions the case studies that will be used to assess the potential impacts of outbreaks on shipments and prices are: the spinach outbreak of September, 2006; the cantaloupe outbreaks of the period 2000-2002 (April-June 2000, April-May 2001, and March-May 2002); and the tomato outbreak of July-September 2006. The length of time necessary to recover to normal consumption levels will also be calculated.

Methodology

The working hypothesis is tested empirically using a time series econometric model. Specifically, the model explores how information is communicated across the three variables, price, imports and shipments for each vegetable product in a neighborhood of the aforementioned food events. The empirical analysis is based on a vector autoregression (VAR) model in which directed acyclic graphs are used to sort-out causal flows of price information in contemporaneous time. Let X_t denote a vector that includes the monthly price, imports and shipments of each vegetable product:

$$X_t = \begin{pmatrix} P_t \\ I_t \\ S_t \end{pmatrix}$$

where t is an index of time observed. Under fairly general conditions the dynamic correlation structure between these variables can be summarized as a structural vector autoregression. The structural VAR representing a $N \times 1$ vector of variables X_t can be written as:

$$\Phi_0 X_t - \sum_{k=1}^{K} \Phi_i X_{t-k} = \varepsilon_t \qquad (1)$$

Here contemporaneous and lagged values of observational measures on X at periods t-k, k = 0, 1, …, K are mapped into the white noise innovation term ε_t , where $Cov(\varepsilon_t) = \Omega$ and M_i , i=0, 1, …, K are square autoregressive matrices of order 3. The innovations ε_t are structural as they represent new information arising in each element of the X vector at time t. Under general conditions permitting matrix inversion an equivalent form exists as:

$$X_t - \Phi_0^{-1}\Phi_1 X_{t-1} - \cdots - \Phi_0^{-1}\Phi_k X_{t-k} = \Phi_0^{-1}\varepsilon_t .$$

The reduced form (non-structural) VAR is written in similar form as:

$$X_t - \Pi_1 X_{t-1} + \cdots + \Pi_k X_{t-k} = u_t ; \qquad (2)$$

where $\Pi_h = \Phi_0^{-1}\Phi_h$ for k = 1, …, K and $u_t = \Phi_0^{-1}\varepsilon_t$. The reduced form innovations (u_t) are "mongrel" or combinations of the structural innovations ε_t. It follows thus that $Cov(u_t) = \Sigma = \Phi_0^{-1}\Omega\left(\Phi_0^{-1}\right)$.

While the reduced form VAR has been championed as "atheoretic", the key to modeling structural VARs is proper identification of the matrix A_0. Bernanke (1986) and Sims (1986) used prior theory to achieve such identification. More recent work follows that of Swanson and Granger (1997) to use the causal pattern exhibited by observed innovations $\hat{u}_t$ to identify M_0. In this paper we use the machine learning algorithms of Spirtes, Glymour and Scheines (2000) as applied earlier in Bessler and Akleman (1998) and Hoover (2005) to achieve structural identification.

The dynamic response patterns summarized by a VAR are difficult to interpret (Sims, 1980; Swanson and Granger, 1997). The dynamic price relationships can be best summarized through the moving average representation. Given the estimated form of equation (1), we can algebraically re-express equation (2) as a levels VAR. We can then solve for its moving average representation, where the vector X_t is written as a function of the infinite sum of past innovations:

$$X_t = \sum_{i=0}^{\infty} \Theta_i u_{t-i} \qquad (3)$$

where G_i is a 3x3 matrix of moving average parameters, which map historical innovations at lag *i* into the current position of the vector *X*.[1] Notice 1_0 is not zero here as we use directed graph structures on the observed innovations from the reduced form VAR to translate these nonstructural innovations to structural innovations (as suggested first by Swanson and Granger (1997)).

A directed graph is a picture representing the causal flow among a set of variables. Lines with arrowheads are used to represent flows. For instance, A→ B

[1] While one can actually derive the first *n* terms of equation (2) analytically, we almost always allow the computer to do this following the zero-one simulation as described in Sims (1980).

indicates that variable A causes variable B. A line connecting two variables, C – D, indicates that C and D are connected by information flow but it's not certain whether C causes D or vice versa. Observed innovations from an estimated form of equation (2) are modeled as a directed acyclic graph for each vegetable commodity. The fundamental idea that enables detection of the direction of causal flow to a set of (observational) variables is the screening-off phenomena and its more formal representation as d-separation (Pearl, 2000). For three variables A, B and C, if we have variable A as a common cause of B and C so that B←A→C, then the unconditional association between B and C will be non-zero, as both have a common cause in A (this diagram is labeled a causal fork (Pearl 2000)). If we measure association (linear association by correlation) then B and C will have a non-zero correlation. However, if we condition on A, the partial correlation between B and C (given knowledge of A) will be zero. Knowledge of the common cause (A) "screens-off" association between its effects (B and C).

On the other hand, say we have variables D, E, and F such that D→E←F. Here we have E is a common effect of D and F (this diagram is labeled a causal inverted fork (Pearl 2000)). D and F will have no association (zero correlation if we constrain ourselves to linear association); however, if we condition on E, the association between D and F is non-zero (the partial correlation between D and F, given knowledge of E is non-zero). We say (in the vernacular) knowledge of the common effect does not "screen-off" association between its causes.

And if we have variables A, B and C forming a causal chain, A→B→ C, the unconditional association (correlation) between A and C will be non-zero, but the conditional correlation between A and C, given knowledge of B will be zero.

Spirtes, Glymour and Scheines (2000) and Pearl (2000) present algorithms with similar structures and outputs for inference on directed acyclic graphs from observational data. The former is labeled PC algorithm, embedded in the software TETRAD II and III (see the offering at http://www.phil.cmu.edu/projects/tetrad/ and Scheines *et al*., 1996) and described in Spirtes, Glymour and Scheines (2000); the latter is IC algorithm presented in Pearl (2000, pp.50-51). PC algorithm has been studied extensively in Monte Carlo simulations in Spirtes, Glymour and Scheines (2000) and Demiralp and Hoover (2003). The algorithm may make mistakes of two types: edge inclusion or exclusion and edge direction (orientation); the latter appears to be more likely than the former. Spirtes, Glymour and Scheines write: "In order of the methods to converge to correct decisions with probability 1, the significance level used in making decisions should decrease as the sample size increases and the use of higher significance levels (e.g., .2 at sample sizes less than 100, and .1 at sample sizes between 100 and 300) may improve performance at small sample sizes." (Spirtes, Glymour and Scheines, 2000, page 116). Nevertheless, the orientation (edge direction) decision is less reliable than the edge inclusion decision in PC algorithm; results presented below should be viewed with caution and/or interpreted with other relevant information.

Once the price innovations from the ECM estimation are orthoganized, the historical decomposition of the vector X at particular time t=T+k can be divided into two parts:

$$X_{T+k} = \sum_{s=k}^{\infty} \Theta_s u_{T+k-s} + \sum_{s=0}^{k-1} \Theta_s u_{T+k-s} \,. \qquad (4)$$

The first term in the right-hand side of equation (8), called the "base projection", utilizes information available up to time period T. The second term contains information available from time period T + 1 until T + k, including the animal disease outbreaks. The difference between the actual price $\left(X_{T+k}\right)$ and the base price projection $\left(\sum_{s=k}^{\infty}\Theta_s\varepsilon_{T+k-s}\right)$ is thus written as a linear function of innovations (new information) arising in teach series between the period T and period $T + k$ $\left(\sum_{s=0}^{k-1}\Theta_s\varepsilon_{T+k-s}\right)$. Through the partition, historical decomposition allows us to examine the behavior of each price series in the neighborhood of important historical events (animal disease outbreaks in our cases) and to infer how much each innovation contributes to the unexpected variation of X_{T+k}.

Results and Discussion

We consider 1999 – 2007 monthly observations on price, imports and shipments of US cantaloupes, spinach and tomatoes. The data are plotted in Figure 1. Plotted with each series is a vertical line indicating the period of food scares. Our plots are offered to give the readers a sense of the seasonal pattern and response in a neighborhood of each food event. A VAR was fit with 2 lags of levels data a constant and eleven monthly dummy variables where Schwarz loss was used to select lag length (results are not reported here but are available from the third author).

Figure 2 shows the graph structure on innovation from each separate VAR. Cantaloupe innovations are modeled as in inverted fork, with price innovations being caused by innovations in shipments and imports. In contemporaneous time imports and shipments appear to be unrelated. Spinach innovations are contemporaneously independent. And contemporaneous innovations in tomatoes are an inverted fork with shipments serving as a collider. Based on the contemporaneous structures in Figure 2 and the estimated VARs for each series we offer historical decomposition of each price series below.

Historical decomposition of each price series following equation (4) are offered in tables 1, 2 and 3. In Table 1 we decompose cantaloupe price before and after each of three events: (April 2000, April 2001 and March 2002). We begin the decomposition one month prior to the event and run it for several months after the event to observe how information arising in each series, price, shipments and imports affected price at each monthly observation. For event 1, actual price was observed to be below the forecasted price in the months following the April event. Information arising in the shipments and price series were actually responsible for this downward movement in actual price; while activity in the import markets (series) was bringing forth a positive influence on price. Overall the negative price and shipments information dominated the positive imports pressure, so actual price was below its pre-event forecast. As one can see from the panel under event 1 in table 1, information arising in the price series (column 5) dominates other new information (column 3 is shipments and column 4 is imports). Relatively small decreases in supply (shipments and imports) resulted in large decreases in price. The depth of event 1 was at May 2000 with the dominate pressure for the $3.95 drop being accounted for from the price innovations itself. Recovery in the event 1 panel is noted clearly after August 2000 with information arising in the shipments series and imports showing a positive influence on price. The following month price also offers positive innovations on price.

Events 2 and 3 for cantaloupes appear to be fundamentally different than event 1. Whereas for event 1, pressure after the event was for a downward spiral in price, events 2 and 3 show price increases. The difference in response is perhaps due to differing consumer behavior toward the events. Whereas the strong negative pressure on prices following event one is perhaps consistent with a negative impression by consumers (reduced demand), there is no similar negative price response following events 2 and 3. Consumers appear not to have been reluctant to consume following these latter two events.

Table 2 summarizes a similar price, shipment and import innovation responses following the September 2006 food event in spinach. Here there is an overall negative response in price following the event. Most of this negative information arises in the price market, suggesting that a drop in consumer demand may be behind the fall off in prices. Innovations in shipments actually show a mildly positive influence on price and imports show a negative affect on price after October 2006. The short-lived effects of the 2006 spinach event are manifest by the return to positive difference between actual spinach price and forecasted spinach price of $3.47/ unit by January 2007 (of course our labeling four months preceding January 2007 as short lived is easy for us and may have not been easy for industry participants).

Table 3 offers price decompositions for tomatoes just before and following the September 2006 event in the tomatoes. Similar to what we observe in the cantaloupes event 2 and 3 the pressure on price after the event was actually positive, indicating no strong evidence of consumer withdrawal from the market. In fact, the large negative innovation arising in price one month before the event is the only negative evidence supporting a strong consumer reluctance to embrace tomatoes just before or after the July 2006 announcement.

Summary and Conclusions

Historically the United States was perceived to have the safest food supply in the world. While, in fact, this may still be true, a number of incidents have led to questions regarding the safety of the U.S. food supply. Three case studies were analyzed to assess the potential impacts of outbreaks on shipments and prices, the cantaloupe outbreaks of the period 2000 to 2002, and the spinach and tomato outbreaks of 2006. For cantaloupe, the model showed that price changes are caused by changes in shipments and imports. Changes in prices, shipments, and imports are independent for spinach, while changes in shipments are caused by changes in prices and imports.

On the historical price decomposition analyzes for cantaloupe, event 1 showed a lower actual price than the forecasted, while the opposite occurred for events 2 and 3, actual price was higher that forecasted price. For spinach there is an overall negative response in price following the event. Most of this negative information arises in the price market, suggesting that a drop in consumer demand may be behind the fall off in prices. Finally, for tomatoes the pressure on price after the event was actually positive, indicating no strong evidence of consumer withdrawal from the market.

Table 1. Historical Decomposition of Cantaloupe Price in a Neighborhood of the Three Food Scare Events.

(1) *Date*	**(2) Difference = Actual Price Minus Forecasted Price**	**(3) Due to Information Arising from Shipments**	**(4) Due to Information Arising from Imports**	**(5) Due to Information Arising from Price**
Event 1				
March 2000	-2.04	-.05	-.10	-1.89
April2000	-.58	-.08	-.03	-.47
May 2000	-3.95	-.31	.23	-3.87
June 2000	-.56	-.29	.30	-.57
July 2000	-.24	.58	.46	-1.28
August 2000	1.31	1.25	.22	-.17
September 2000	5.88	.85	-.02	5.05
Event 2				
March 2001	1.52	-.06	.25	1.33
April 2001	.52	-.18	.58	.12
May 2001	.87	.02	-.25	1.10
June 2001	1.41	.24	-.60	1.77
July 2001	5.74	.51	-.01	5.24
August 2001	1.98	1.15	.21	.62
September 2001	-1.04	.64	.13	-1.81
Event 3				
February 2002	.05	-.08	.17	-.04
March 2002	1.55	-.01	.03	1.53
April 2002	3.64	.02	-.02	3.64
May 2002	.32	-.08	-.17	.58
June 2002	-.52	-.14	.52	-.90
July 2002	-.81	-.20	.47	-1.08
August 2002	.43	1.41	.10	-1.08
September 2002	.73	.99	-.16	-.10

Note: This table decomposes the difference between the Actual Price and the Forecasted Price at each date, over a period just before and several periods after each of three food events. That difference at each date can be attributed to information arising in the shipments variable, the imports variable and the price variable. Accordingly, the column labeled (2) is decomposed at each date into the sum of columns (3), (4) and (5).

Table 2. Historical Decomposition of Spinach Price in a Neighborhood of the September 2006 Event.

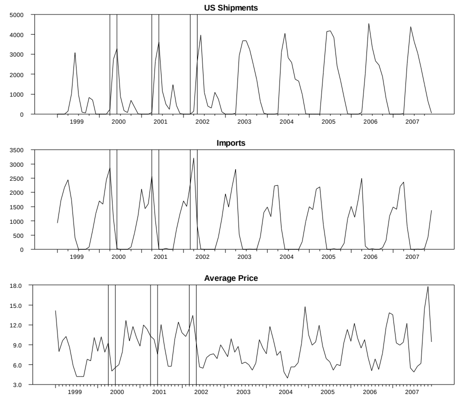
(1) *Date*	**(2) Difference = Actual Price Minus Forecasted Price**	**(3) Due to Information Arising from Shipments**	**(4) Due to Information Arising from Imports**	**(5) Due to Information Arising from Price**
August 2006	4.27	.40	1.11	2.76
September 2006	-.67	.16	.74	-1.58
October 2006	-1.90	.05	-.07	-1.89
November 2006	-1.17	.32	-.22	-1.28
December 2006	.63	.34	-.24	.52
January 2007	3.47	-.47	.66	3.28
February 2007	3.56	-1.53	.19	4.90
March 2007	-1.30	-1.10	.44	-.64
April 2007	-2.29	-.61	.42	-2.10
May 2007	-.48	-.64	.03	.13

Note: This table decomposes the difference between the Actual Price and the Forecasted Price at each date, between August 2006 and May 2007. That difference at each date can be attributed to information arising in the shipments variable, the imports variable and the price variable. Accordingly, the column labeled (2) is decomposed at each date into the sum of columns (3), (4) and (5).

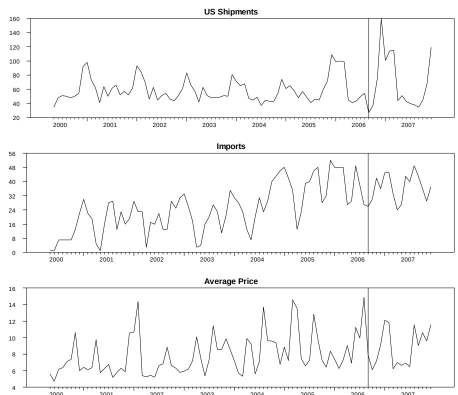
Table 3. Historical Decomposition of Tomato Price in a Neighborhood of the July – September 2006 Event.

(1) *Date*	**(2) Difference = Actual Price Minus Forecasted Price**	**(3) Due to Information Arising from Shipments**	**(4) Due to Information Arising from Imports**	**(5) Due to Information Arising from Price**
June 2006	-.40	2.50	1.31	-4.21
July 2006	.22	.68	.90	-1.37
August 2006	2.82	-1.43	.93	3.32
September 2006	13.72	.41	.59	12.71
October 2006	2.75	.64	.64	1.46
November 2006	-5.60	.13	1.36	-7.09
December 2007	-5.25	-.20	.32	-5.38
January 2007	-1.95	.39	.41	-2.76
February 2007	-.24	.10	2.25	-2.60
March 2007	-2.62	-.65	1.39	-3.35

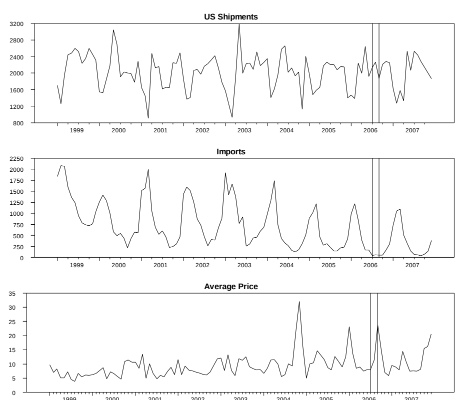
Note: This table decomposes the difference between the Actual Price and the Forecasted Price at each date, between July 2006 and March 2007. That difference at each date can be attributed to information arising in the shipments variable, the imports variable and the price variable. Accordingly, the column labeled (2) is decomposed at each date into the sum of columns (3), (4) and (5).



Cantaloupes



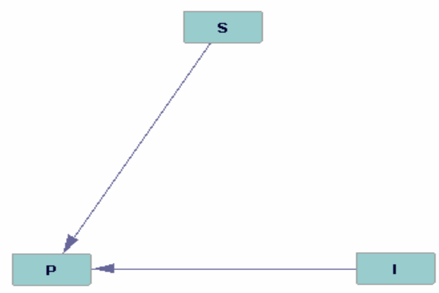
Spinach



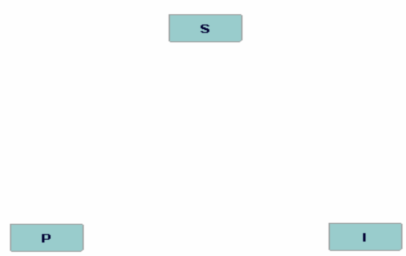
Tomatoes

Figure 1. Time Series Plots of Shipments, Imports, and Prices of Cantaloupes - Monthly Data, 1999 – 2007.

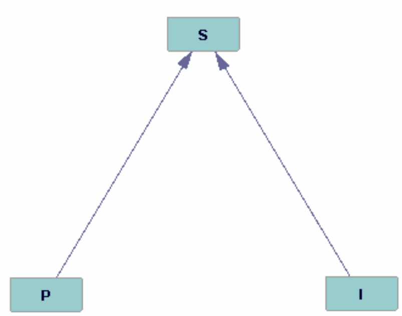
Note: Vertical Lines are Placed at Dates of Interest: beginning date of food scare and ending date on food scare) for each commodity.



Cantaloupes



Spinach



Tomatoes

Figure 2. Causal Pattern on Innovations from a Vector Autoregressions Models Fit to Monthly Observations on Shipments (S), Imports (I), and Prices (P) for Cantaloupes, Spinach and Tomatoes.

References

Bernanke, B.S. 1986. "Alternative Explanations of the Money-Income Correlation." Carnegie-Rochester Conference Series on Public Policy 25: 49-99.

Bessler, D.A. and Akleman, D. G.*1998. "Farm Prices, Retail Prices, and Directed Graphs: Results for Pork and Beef" *American Journal of Agricultural Economics* 80(5): 1145-1150.

Dermilap, S. and Hoover, Kevin D. (2003): "Searching for the Causal Structure of a Vector Autoregression," *Oxford Bulletin of Economics and Statistics*, 65:745-767.

Food and Drug Administration. (September 14, 2008) FDA Warning on Serious Foodborne E. Coli O157:H7 Outbreak. Available at: http://www.fda.gov/bbs/topics/NEWS/2006/NEW01450.html

Food and Drug Administration. (October 6, 2006). FDA Statement on Foodborne E. coli O157:H7 Outbreak. Available at: http://www.fda.gov/bbs/topics/NEWS/2006/NEW01486.html

Food and Drug Administration. (March 21, 2008). Archived Updates on the FDA Investigation in Honduras. Available at: http://www.fda.gov/oc/opacom/hottopics/cantaloupearchive.html

Food and Drug Administration. (March 22, 2008). FDA Warns of Salmonella Risk with Cantaloupes from Agropecuaria Montelibano. Available at: http://www.fda.gov/bbs/topics/NEWS/2008/NEW01808.html

Food and Drug Administration. (June 3, 2008). FDA Warns consumers in New Mexico and Texas Not to Eat Certain Types of Raw Red Tomatoes. Available at: http://www.fda.gov/bbs/topics/NEWS/2008/NEW01843.html

Food and Drug Administration. (July 17, 2008). FDA Lifts Warning About Eating Certain Types of Tomatoes. Available at: http://www.fda.gov/bbs/topics/NEWS/2008/NEW01862.html

Food and Drug Administration. (August 28, 2008). Salmonella Saintpaul Outbreak. Available at: http://www.fda.gov/oc/opacom/hottopics/tomatoes.html

Hoover, Kevin D. (2005): "Automatic Inference of the Contemporaneous Causal Order of a System of Equations," *Econometric Theory*, 21:69-77.

Pearl, J. 2000. Causality, New York: Cambridge University Press.

Sims, C. 1980."Macroeconomics and Reality, Econometrica, 48:1 – 48.

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

Swanson NR, Granger C. 1997. Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association* **92**: 357-367.

United States Department of Agriculture. Agricultural Marketing Service. *Fruit and Vegetable Portal News.* Price and Shipment Information. 2000-2008